March 2012

Roof Loss Control Analysis for The Hanover Insurance Group

Dennis Paul Griffin
Worcester Polytechnic Institute

Nathan Alexander Rivard
Worcester Polytechnic Institute

Nicholas William Bean
Worcester Polytechnic Institute

Scott Edward Brady
Worcester Polytechnic Institute

Ted William Fitts
Worcester Polytechnic Institute

Follow this and additional works at: https://digitalcommons.wpi.edu/mqp-all

Repository Citation

This Unrestricted is brought to you for free and open access by the Major Qualifying Projects at Digital WPI. It has been accepted for inclusion in Major Qualifying Projects (All Years) by an authorized administrator of Digital WPI. For more information, please contact digitalwpi@wpi.edu.
Roof Loss Control Analysis for The Hanover Insurance Group

A Major Qualifying Project

Submitted to

The Hanover Insurance Group

And to the Faculty of

Worcester Polytechnic Institute

March 2\textsuperscript{nd}, 2012

Authors:

Nick Bean
Scott Brady
Ted Fitts
Dennis Griffin
Nathan Rivard

Advisors:

Professor Fabienne Miller
Professor Guillermo Salazar

Project Liaisons:

Chris Beckman
Joan Wooley
Abstract

The Hanover Insurance Group has identified roof claims as an area to investigate for trends, which could be used to help minimize business loss. This report analyzes different roof and building qualities through data mining, statistical testing, and an interview. Through this research, the team was able to gather information regarding roof claims and conduct statistical analysis on this data. From these findings, it is recommended that The Hanover research further into roof loss claims.
The Hanover Insurance Group – Roof Loss Control Analysis

Acknowledgements

Our group would like to take the time to thank many individuals who, without their cooperation and support, the success of this Major Qualifying Project would not have been possible:

Our project liaisons, Chris Beckman and Joan Wooley, for providing guidance to the team as well as making the team feel welcome throughout the duration of our project.

Our project advisors, Professors Fabienne Miller and Guillermo Salazar, for their time, constructive feedback, and support throughout this project.

Jim Ducey, for taking the time out of his schedule to interview with us and providing valuable information to the team.

The Hanover Insurance Group, for providing a welcoming and outstanding work environment.
Authorship

Nick Bean, Scott Brady, Ted Fitts, Dennis Griffin and Nathan Rivard all contributed to the data gathering and research for this project.

As for the composition of this project, the Introduction was written by Ted and Nate. The Background, Literature Review, and Methodology were all written equally by each group member. Nick, Scott and Nathan all contributed to the Findings section. The Recommendations were constructed by Nick, Dennis and Nathan. The Conclusions were written by Nick, Dennis and Nathan. The Executive Summary was written by Nick, Dennis and Nathan. The Authorship was collaboratively written by the team as was the abstract. After the compilation of the report, the entirety of the paper was edited by all members of the group.
Executive Summary

Introduction

The Hanover Insurance Group, with its headquarters in Worcester, MA, is a nation-wide insurance company that offers a wide range of insurance products to a variety of customers. The Hanover utilizes its Loss Control employees to assess the risk of insuring buildings when determine whether to insure and at what rate. It uses this information going forward to identify trends in building/location data to optimize a protection plan that insures the building and protects The Hanover from unnecessary loss.

Goal

The goal of this project was to gather and analyze historical data of roof loss claims for large property losses insured by The Hanover, as well as gather data on these claims from outside sources, to identify additional trends in roof failures. The trends and information gathered throughout the course of this project could help The Hanover to identify common causes of roof loss claims, as well as provide recommendations for reducing potential roof losses. However, before this was to occur, a comprehensive review of literary sources were analyzed and used to make predictions and hypotheses about roof related claims. From these hypotheses, there were multiple steps taken to gather and analyze the data. The steps taken to achieve this goal include:

1. Retrieving, sorting and analyzing historical claims data on large property losses provided by The Hanover,

2. Researching the claims data further through the use of external sources,
The Hanover Insurance Group – Roof Loss Control Analysis

3. Using statistical analysis to identify and compare trends in the data that could potentially help The Hanover predict future losses,

4. Provide recommendations to The Hanover to help them protect themselves and the insured against unwanted losses.

Methodology

There were four steps used to collect and analyze this data. At the onset of the project it was hoped that a fifth step, interviews, would be included to provide the real-world experience they have regarding the insurance process. However, due to timing and schedule constraints, this was not possible.

The first step in collecting the group’s data was to analyze the PDF of large property loss claims, which totaled over $100,000, dating back to 2006 that was provided to the group by The Hanover. The group sorted this data by whether there was a roof claim or not. When this was completed, the information was transferred to a Loss Data Excel file where further research and modifications were now possible.

The next step was to take all of the claim information in Excel and, through use of The Hanover’s databases such as ARIES, CAAMs, CSS and HCS, and collect data on these claims for as many variables as possible. The variables to be collected by the group for this project include Addition, Ages of Building, Age of Roof Cover, Catastrophe Code, Claim Number, Date of Loss, Dollar Amount of Loss, Elevation Difference, ISO Building Codes, Loss Type, Name of Insured, Number of Stories, Occupancy, Pitch, Square Footage, Street Addresses, and Roof Cover Material. After all of The Hanover’s databases were exhausted, the group found that the
information gathered was not sufficient to conduct the analysis that had been planned. This step also involved thoroughly searching databases outside of The Hanover to gather the still required information on the buildings involved in the claims. This was achieved by identifying and searching websites and tools to gather this information, some of the data gathered includes square footage and elevation differences. Some of these outside databases were websites such as propertyshark.com and Google and were instrumental in finding as much relevant data as possible. All claims still requiring data were researched in this manner until all databases were exhausted and all possible data points were filled.

The third step was the descriptive and statistical analysis of the data, which was conducted using regression and ANOVA testing to identify potential trends and correlations in the data. With the results of these analyses in hand, the group was able to move on to the final step of the project, which involved making recommendations to The Hanover including some recommendations for future research. With this in mind, the purpose of the project was to identify potential correlations in relevant roof loss data that The Hanover would be able to use to assist in predicting future loss. However, the analysis did not yield the desired results.

Results and Recommendations

The results of the analysis performed on the data returned several recommendations. The analysis conducted by the group of the relationship between different variables was largely inconclusive. Small samples sizes and the lack of a control group greatly affected our analyses. For instance, the lack of a control group containing information on buildings without roof loss claims made it difficult for the group to evaluate things such as the frequency of claims in
different states and regions of the country. If a control group had been present the different loss types could have been compared to the total population of insured buildings. Instead the different loss types had to be compared against one another and to building characteristics in order to try and draw conclusions. The results of these comparisons proved to be statistically insignificant and therefore, the variables were not accurate predictors of roof loss. Although the team could not obtain control groups due to privacy laws and confidentiality policies of The Hanover, future research conducted within The Hanover would allow access to these control groups and could, accordingly, test our predictions further. Along with the recommendation to obtain a control group, other recommendations provided to The Hanover include a consolidation of its databases and collection of additional data, implementation of a roof maintenance benefit system, the gathering of information on building contractors, and the draft of an interior and exterior roof checklist.

With a consolidation of databases, all information can be centrally located and easily accessed, rather than having data from some years in one database and the other years in another. Even if this recommendation is disregarded and the practice of multiple databases is continued, The Hanover might attempt to streamline the process of recording and displaying information so that there are no discrepancies between the two. A second recommendation would be to make sure records of the assessments of the property are maintained and stored somewhere that can be easily accessed. While it is known the required research is being conducted by The Hanover when assessing a building to insure, there are a limited amount of these records available.
Another recommendation the group would like to extend to The Hanover, based on its research, involves the possibility of implementing a trial incentives program regarding roof maintenance, similar to the one in place for sprinkler systems. This recommendation stems from extensive research of literature that concludes that the best way to avoid a claim on a roof is regular inspection and maintenance. If this could be carried out by businesses that The Hanover insures, then at regular intervals determined by The Hanover, and with documentation of execution provided, The Hanover might consider offering some type of incentive in return to encourage the continuation of this practice. Hopefully, this incentives program could lead to a reduction in roof loss claims.

A third recommendation for The Hanover, involves beginning to collect information on building contractors. This database could provide a track record for each contractor of buildings built and insured by The Hanover, including and any buildings that have had roof loss claims attributed to that contractor. With this database, policies can be avoided for buildings that are built by contractors with a poor track record.

In addition to this, a standardized checklist containing traits and characteristics on the exterior and interior of buildings could be provided to loss control employees evaluating buildings. This list would contain common traits deemed as high risks in buildings and the presence, or lack of presence, of these traits would help to determine whether a policy should be granted.
Conclusions

While the analysis of the data did not yield results that proved to be statistically significant and did not enable the team to test the predictors of roof loss we identified in the literature review, it did provide areas of future research that might be beneficial for The Hanover to explore further. While the team’s analyses were limited by small sample sizes and lack of control groups, if The Hanover could obtain a control group of data and increase these sample sizes, different conclusions could be drawn.
Table of Contents

Abstract........................................................................................................................... ii
Acknowledgements........................................................................................................ iii
Authorship ........................................................................................................................ iv
Executive Summary......................................................................................................... v
  Introduction .................................................................................................................... v
  Goal ................................................................................................................................. v
  Methodology .................................................................................................................. vi
  Results and Recommendations ....................................................................................... vii
  Conclusions .................................................................................................................... x

Table of Contents ............................................................................................................. 1
Table of Figures ................................................................................................................ 4
Table of Tables .................................................................................................................. 5
1.0 Introduction ................................................................................................................ 6
2.0 Background ................................................................................................................ 8
  2.1 The Hanover Insurance Group Background ............................................................ 8
  2.2 Project Background ................................................................................................... 9
3.0 Literature Review ...................................................................................................... 11
  3.1 General Roof Information ....................................................................................... 11
    3.1.1 Common Roof Types ......................................................................................... 11
    3.1.2 Roof Materials ................................................................................................. 11
    3.1.3 Service Life ....................................................................................................... 16
    3.1.4 Construction Methods ..................................................................................... 17
    3.1.5 OSHA ............................................................................................................... 18
  3.2 Roof Failure ............................................................................................................ 19
    3.2.1 Causes of Roof Failure .................................................................................... 19
    3.2.2 Common Deficiencies of Roofs ....................................................................... 20
    3.2.3 Causes of Failure Specific to this Project ......................................................... 21
  3.3 Roof Failure Prevention .......................................................................................... 24
    3.3.1 Choosing a Contractor .................................................................................... 25
    3.3.2 Maintenance Program ..................................................................................... 25
    3.3.3 Roof Inspection ............................................................................................... 26
  3.4 Insurance ................................................................................................................ 28
    3.4.1 Property ............................................................................................................ 29
    3.4.2 Additional Structures ....................................................................................... 29
    3.4.3 Loss Control ..................................................................................................... 30
    3.4.4 Evaluating Loss ............................................................................................... 30
6.1 Recommendations ........................................................................................................................................94
6.2 Future Research ..........................................................................................................................................98
   6.2.1 Presence of Control Groups and Scarcity of Data.............................................................................99
   6.2.2 Hypothesis 2: Age of Roof Cover .......................................................... ..................................100
   6.2.3 Hypothesis 4: Claims with Additions...............................................................................................100
   6.2.4 Hypothesis 6: Maintenance and Inspection ..................................................................................101
7.0 Conclusion ...................................................................................................................................................102
References .........................................................................................................................................................103
Appendix A: Background ................................................................................................................................106
   Insurance Terms (Merriam Webster 2011) .................................................................................................106
   Roofing Terms (Common Roof Types 2011) ...............................................................................................106
Appendix B: Literature Review ........................................................................................................................108
   ISO Codes ...............................................................................................................................................108
Appendix C: Methodology .................................................................................................................................111
   Step by Step Instructions for Hanover Databases .....................................................................................111
      HCS .............................................................................................................................................111
      CSS ..............................................................................................................................................111
      CAAMS ........................................................................................................................................112
      ARIES ...........................................................................................................................................114
   Interview Contact List .................................................................................................................................114
   Interview Questions .................................................................................................................................115
Appendix D: Findings .......................................................................................................................................116
   Statistics ..............................................................................................................................................116
      Linear Regressions and ANOVA assumptions ......................................................................................116
# Table of Figures

<table>
<thead>
<tr>
<th>Figure Number</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Methodology Flow Chart</td>
<td>35</td>
</tr>
<tr>
<td>2</td>
<td>Claims with Catastrophe Information</td>
<td>63</td>
</tr>
<tr>
<td>3</td>
<td>Number of Claims per ISO Code</td>
<td>64</td>
</tr>
<tr>
<td>4</td>
<td>Number of Claims per Type of Roof Cover</td>
<td>65</td>
</tr>
<tr>
<td>5</td>
<td>Number of Claims per Building Age Category</td>
<td>70</td>
</tr>
<tr>
<td>6</td>
<td>All Weather Related Data Minus Outliers: Breakdown by Loss</td>
<td>78</td>
</tr>
<tr>
<td>7</td>
<td>All Weather Related Data Minus Outliers: Average Dollar Amount of Loss Type</td>
<td>79</td>
</tr>
<tr>
<td>8</td>
<td>All Weather Related Data Minus Outliers: Total Dollar Amount of Loss by Loss Type</td>
<td>80</td>
</tr>
<tr>
<td>9</td>
<td>Count of Claims by State by Loss Type</td>
<td>84</td>
</tr>
<tr>
<td>10</td>
<td>Sum of Dollar Amount of Loss by State by Loss Type</td>
<td>85</td>
</tr>
<tr>
<td>11</td>
<td>Number of Claims per Sq. Footage Category with Outliers Removed</td>
<td>87</td>
</tr>
<tr>
<td>12</td>
<td>Percentage of Occupancy Categories for All Roof Loss Claims Data</td>
<td>90</td>
</tr>
<tr>
<td>13</td>
<td>Average Dollar Amount of Loss per Occupancy Category</td>
<td>91</td>
</tr>
<tr>
<td>14</td>
<td>Total Dollar Amount of Loss per Occupancy Category</td>
<td>92</td>
</tr>
</tbody>
</table>
Table of Tables

Table 1. Variable Definitions........................................................................................................................................34
Table 2. Which Database to Search for Each Variable..............................................................................................49
Table 3. Summary of Observations per Variable ........................................................................................................66
Table 4. Age of Building Code & Log Transformation of Sq. Footage by Dollar Amount of Loss .....................71
Table 5. Number of Claims with Evidence of Elevation Difference and Loss Types of Snow or Collapse ....72
Table 6. Claim Frequency by Naturally Occurring Weather Related Data Minus Outliers ....................................73
Table 7. Average Dollar Amount of Loss for Naturally Occurring Weather Related Data Minus Outliers ..........73
Table 8. ANOVA Table for Naturally Occurring Weather Related Claims by Dollar Amount of Loss ............74
Table 9. Natural Disaster Claims Data Minus Outliers ...............................................................................................74
Table 10. Natural Disaster Claims Data Minus Outliers .............................................................................................75
Table 11. Natural Disaster Claims by Dollar Amount of Loss ....................................................................................75
Table 12. All Weather Related Data Combined Minus Outliers .............................................................................76
Table 13. All Weather Related Data Combined Minus Outliers .............................................................................76
Table 14. All Weather Related Loss Claims by Dollar Amount of Loss .................................................................77
Table 15. All Naturally Occurring Weather Related Data & Log Transformation of Sq. Footage by Dollar Amount of Loss ........................................................................................................81
Table 16. Natural Disaster loss & Log Transformation of Sq. Footage by Dollar Amount of Loss ..................82
Table 17. All Weather Related Loss Types & Log Transformation of Sq. Footage by Dollar Amount of Loss ......................................................................................................................83
Table 18. Data Source Code & Log Transformation of Sq. Footage by Dollar Amount of Loss ......................88
Table 19. Number of Claims due to Snow and Collapse versus Roof Pitch ........................................................89
1.0 Introduction

Insurance companies, throughout the United States, dedicate numerous resources and funds every year to loss control. Loss control is an organized and continuous effort to help decrease the possibility of unforeseen losses as well as the impact of those that do occur. Loss control can be applied to all kinds of losses such as those caused by wind, hurricane, snow, burglary, or anything that causes unexpected damages.

One particular area that has seen significant unexpected losses within the last two decades is roof claims. The Hanover Insurance Group (The Hanover) aims to reduce losses and prevent future losses related to commercial roof loss claims within the United States. Recently, these roof loss claims have been large and regularly occurring. Because of this, The Hanover wants to research if there are actionable underwriting or loss control programs that can be used to reduce their roof loss exposure.

The research conducted within this project could be used to aid The Hanover in their current efforts to reduce roof loss exposure. Given a list of commercial property losses, this project plans to use The Hanover’s current claims databases, as well as numerous external databases, to compile the informational data that is the basis for the group’s descriptive and statistical analyses.

Overall, this project, sponsored by The Hanover Insurance Group, intends to make recommendations for ways in which to improve the current roof evaluation process, and to identify trends in roof loss claims by collecting and analyzing the roof loss data provided.
The Hanover Insurance Group – Roof Loss Control Analysis

This report includes a brief discussion on the background of The Hanover, a literature review that provides the reader with the knowledge to process the information presented in the later sections of the paper, a methodology describing the methods used to gather and interpret the data. It also includes a findings section that outlines the results of the group’s data mining and statistical analysis, a recommendations section that details the group’s recommendations and future research for The Hanover, and finally a conclusion that ties together all of the information gathered from this project.
2.0 Background

The background section provides the reader with information on The Hanover Insurance Group, as well as an overview of the project itself.

2.1 The Hanover Insurance Group Background

The Hanover Insurance Group consists of Hanover Insurance, Citizen’s Insurance Company of America, Chaucer Holdings PLC and their affiliates. The original company began in 1852, as a property and casualty company in New York City. In 1995, a merger between the aforementioned companies occurred, thus forming The Hanover Insurance Group. The Hanover’s headquarters are located in Worcester, Massachusetts and have been located there since 1969. In the same year, it became affiliated with State Mutual, a company which, after going public in 1995, would become known as Allmerica Financial Corporation. In 2005, the company made the decision to sell its life insurance affiliates and changed the company name to The Hanover Insurance Group. The Hanover Insurance Group is publicly traded on the New York Stock Exchange, under the abbreviation THG, and is consistently ranked among the top 25 property and casualty insurers in the country while maintaining excellent ratings from A.M. Best, Standard & Poor’s and Moody’s (About The Hanover Insurance Group 2011). The Hanover Insurance Group currently employs more than 4,000 employees nationwide and 2,000 agents around the world, with annual sales totaling approximately eight and a half billion dollars.

The Hanover Insurance Group provides products and services across the United States, in the small to mid-size personal and commercial insurance market. The personal line segment includes automobile and homeowners insurance, as well as other expensive personal items that
The Hanover Insurance Group – Roof Loss Control Analysis

need to be insured. The commercial market segment covers company cars, worker’s compensation and other general lines of commercial property (About The Hanover Insurance Group 2011). The Hanover also provides advice on investments and can provide asset management to unaffiliated organizations.

2.2 Project Background

The Hanover Insurance Group has a desire to reduce the losses and safeguard against future losses related to commercial and residential roof loss claims across the country. Each year insurance companies cover losses on roofs which can cost them millions of dollars. As of late, these claims have become numerous and expensive. For example, The Hanover provided a list to the team of about 3,000 of The Hanover’s largest loss claims from the past 20 years. The Hanover has defined a large loss claim as one of $100,000 or greater. Of these claims, 405 were roof related claims. If every claim has a minimum of $100,000, then the minimum dollar amount paid by The Hanover, for only roof related claims, would be over $40 million. However, some individual claims can skyrocket to $5 million and above, making the total far larger. The recent string of natural disasters in the United States, such as the hurricanes and tornados that have wreaked havoc up and down the east coast, caused a spike in the number of roof related claims. This caused The Hanover to take a closer look at their evaluation process and the vast number of roof loss claims they have incurred.

From this project, the team’s goal is to provide research and analyze data about roofing claims to help The Hanover minimize the losses incurred. The group’s goal is to investigate the data from the roof loss claims, provided by The Hanover, and identify trends or patterns that
The Hanover Insurance Group – Roof Loss Control Analysis

The Hanover can use in its evaluation process to help protect against future preventable losses. If trends or specific causes for roof damage and failure can be identified by this group, it has the potential to save The Hanover large amounts of money on claims.

The following literature review provides material on general roof knowledge that the reader can use as background information for evaluating roofs. The literature review also provides knowledge of other roofing practices worldwide, causes of roof failures, and roof failure prevention. Information also included in the literature review is regarding the insurance industry and the ways in which losses are evaluated. The group’s hypotheses are explained at the conclusion of the literature review.
3.0 Literature Review

The main purpose of this project is to identify trends in roof loss claims; however it is important to first understand the characteristics of a roof and other general roof information. The rest of this literature review contains descriptions of the different kinds of common roof failures, roof failure prevention, and insurance as it relates to roof failures.

3.1 General Roof Information

Roofs have only one basic job, to protect the building and its contents from outside elements. It is also important to remember that no single roofing system universally outperforms all other types, but usually one or two are probably the best choice for each building scenario (Holzhauer 1997).

3.1.1 Common Roof Types

Most roofs can be classified into three categories, steep sloped, flat, or a combination of both. A steep sloped roof can be most commonly seen on residential buildings and homes, and are generally categorized by having 14 degrees of slope or more. Steep sloped roofs will have a variety of different shapes, including the gable roof, hip roof and A-frame roof. On the contrary, a flat or, sometimes called, low slope roof will generally have less than 14 degrees of slope. These types of roofs are mostly seen on commercial buildings and offices (Roof System Types 2011). Roofs can also be further classified based upon their materials.

3.1.2 Roof Materials

Roof materials are some of the most critical elements of the roofing system. The materials chosen in a roof system must be able to protect the building from the elements and
The Hanover Insurance Group – Roof Loss Control Analysis

withstand wear and tear from everyday use. With a clear understanding of all the various types of materials used in roofing systems, The Hanover can better prepare inexperienced underwriters and loss control agents to rate the quality of the roof they are insuring. There are roofing materials that are unique to both steep sloped roofs and flat or low slope roofs, the following sections describe these materials.

3.1.2.1 Roof Materials for Steep Sloped Roofs

With steep sloped and flat roofs, each classification has its own materials and techniques that are most common. Typically, steep sloped roofs consist of three parts: roof deck, underlayment and roof covering. Some of the most common types of steep sloped roof coverings are asphalt shingles, clay and concrete tile, slate, wood shakes/wood shingles, synthetic, and metal.

Asphalt shingles are made of a base material, asphalt and fillers, and surfacing material. The base material is either organic felt or glass-fiber mat, which provides support for the weather resistant components and gives shingles their strength. The surfacing material is generally in the form of mineral granules that provide protection from impact and UV degradation and also improve fire resistance (Asphalt Shingles 2011).

Clay and concrete tile are two very similar steep slope coverings and are often categorized together. Clay tile is produced by baking molded clay forms into tile. The density of the clay is determined by the length of time and temperature at which it is heated. Concrete tiles, on the other hand, are made of Portland cement, sand and water in varying proportions. The material is mixed and extruded on molds under high pressure and then cured to reach the required strength. For either tile, a glaze or surface texture treatments can be applied and there
are a wide variety of profiles, styles, finishes and colors (Clay Tile & Concrete Tile 2011).

_Slate_ used for roofing is a dense, durable, naturally occurring material that is essentially nonabsorbent. The surface texture of slate after being split for commercial use derives from the characteristics of the rock from which it was quarried. Some slate splits to a smooth, practically even surface, while other yields a surface that is rough and uneven (Slate 2011).

_Wood shakes and wood shingles_ are commonly manufactured from western red cedar, cypress, pine and redwood trees. The shakes are split from logs and reshaped by manufacturers for commercial use. They are thicker at the butt end than shingles and generally one or both surfaces are split to obtain a textured effect. Wood shingles are cut on both sides and have an even taper and uniform thickness (Wood Shakes & Wood Shingles 2011).

_Synthetic_, as it pertains to steep-slope roofing materials refers to manufactured products that replicate asphalt shingles, concrete tile, clay tile, metal panels, slate, wood shakes and wood shingles. Synthetic roof coverings contain recycled plastic and/or rubber as a key ingredient. Synthetic roof coverings are relatively new and there is no proven track record about their performance (Synthetic 2011).

_Metal_ typically includes three general categories of metal roof systems used for steep-slope roofing applications: architectural metal panel, structural metal panel and metal shingle/shingle panels. Generally, architectural metal panel roof systems are water-shedding and are intended for use on steep slope roofs. Most structural metal panel roof systems are designed to resist the passage of water at laps and other joints, as sealant or anti capillary designs can be used in the seams. Metal shingles and shingle panels are available in numerous varieties for use as steep-slope roof coverings. Most of the metal shingles are press-formed
during the manufacturing process to provide a variety of shapes. These products can take the shape of individual or multiple asphalt, tile, slate or wood shingle configurations (Metal Roof Systems 2011).

### 3.1.2.2 Roof Materials for Flat or Low Sloped Roofs

Like steep slope roofs, flat roofs have three principle components. For flat or low sloped roofs they are, weatherproofing, reinforcement, and surfacing, which together can be called a “roofing system”. There are multiple different types of flat or low sloped roof systems: built up roofs, modified bitumen, single ply membrane, metal, foam based membranes, and green roofs.

**Built up roofs** include asphalt or coal tar bitumen, which act as an adhesive, binding the membrane together and holding it to the substrate and reinforcing material, giving the roof its toughness. Asphalt impregnated fiber glass felts, available in several different types and strengths are the most common reinforcement material. Polyester fibers are used occasionally. Surfacing material, such as aggregate, is placed on top of the membrane for protection and resistance to UV radiation. Smooth surfaces with an aluminum pain or white coating, protected by some kind of emulsion coating are also available (Holzhauer 1997).

**Modified bitumen** closely resembles a built up roof in materials, equipment and expertise to manufacture and install the system. Bitumen includes elastomers or polymers mixed with basic asphalt to create a product with improved elasticity and toughness. Several modifiers are available, but two are most commonly used: styrene butadiene styrene (SBS) and atactic polypropylene (APP). Reinforcement is added to keep the product in place and add tensile strength (Holzhauer 1997).
Single ply roofs are much cheaper and easier to install than the previous two types of roofing since, as the name implies, only one layer is used. Since only one layer is used installation is critical to the overall success of the roof. There are three basic types of membrane: cured or vulcanized elastomers, non-cured elastomers, and thermoplastics. Cured elastomers possess great elasticity and are durable while uncured elastomers weather well and resist a broad range of contaminants. Thermoplastic elastomers are less resilient than the other versions (Holzhauer 1997).

Metal roofing, in contrast with single ply roofing, is rather expensive but is durable and requires low maintenance. Most metal roofs feature a standing seam design, where seams are located between the panels, elevated above the roof line and fasteners are concealed within the beams. Standing seam roof systems from metal panels are available in two designs: architectural (for steep slope roofs) and structural (for low slope roofs) (Holzhauer 1997).

Foam is a unique approach to roofing compared to the other four methods. The urethane foam material is sprayed onto the substrate to form an elastomer coating that insulates and waterproofs. The urethane and a sprayed-on coating protect the foam from UV radiation and other hazards. Foam roofs offer high thermal resistance and moisture protection, however are more susceptible to moisture damage (Holzhauer 1997).

Green roofs or sustainable roofs, along with those previously mentioned roof systems, are another type of roof that is gaining popularity. A green roof, or rooftop garden, is a vegetative layer grown on a rooftop. Green roofs provide shade and remove heat from the air through evapotranspiration, reducing temperatures of the roof surface and the surrounding air. Green roofs can be installed on a wide range of buildings, from industrial facilities to private
residences. They can be as simple as a 2-inch covering of groundcover or as complex as a fully accessible park, complete with trees (Green Roofs 2010).

With green roofs, as with all roofs, the service life of the roof can be an important factor in selecting the right roof for a building.

3.1.3 Service Life

Determining the service life of a roof is a difficult task to perform, especially when trying to evaluate the impact of the weather on the materials that make up the roof. Before a roof is installed on a building or home, manufacturers perform laboratory tests on sample materials to simulate weathering and its effect on various physical properties over time. Although these tests can be helpful in estimating the life of a roof, the only true test for a roofing system is actual time and observation of how each material reacts to different patterns of weather. Manufacturers, through controlled laboratory testing, may claim their roofing membrane material will pass through a simulation of 20, 30 or more years of exposure, but the only way to be certain is through examination of real life roofing systems (Benson 2006).

To determine the actual service life of a roofing system one should take into account certain factors. One of the most important factors is type of surface protection, meaning the topmost layer of the roof and how it protects the rest of the roof. Other factors used in determining service life include: deck type, type of system, building structure, drainage, proper design, workmanship, flashing details, environment, regular inspection, maintenance, and construction methods (Benson 2006). Another critical factor in determining the service life of a roof is the construction method used.
3.1.4 Construction Methods

Usually, steep sloped roofs are composed of components that are installed in shingle fashion. Almost all steep slope roof systems have five basic components: roof covering, sheathing, roof structure, flashing, and drainage. Roof coverings, as discussed in the previous section, are the top layer that protects the sheathing from weather. Sheathing consists of boards or sheet material that is fastened to roof rafters to cover a building. The roof structure is made up of the rafters and trusses that are constructed to support the sheathing. Flashing is the sheet metal or other material that is installed into a roof systems various joints and valleys to prevent water seepage. Lastly, drainage is a roof systems design features, such as shape, slope and layout that affect its ability to shed water (Consumer Info 2011).

With flat or low slope roofs, there are a variety of different construction methods. For example, built up roofs are usually applied hot in the field by a roofing crew that lay bitumen on the substrate with a mop or cart. Felts are then rolled into the hot bitumen shingle style in order to obtain the desired number of plies. A different technique used to install a built up roof is a cold applied installation where the asphalt is warmed and applied with a broom or squeegee. The cold approach allows the contractor a longer time to apply the felts and improves adhesion between the plies. For foam based applications, the material is sprayed directly onto the roof. With modified bitumen, the membranes can be mopped or applied with cold adhesives, or it can also be torched or put down with hot air. Single ply membranes can be applied using a variety of methods: loose-laid and ballasted, fully adhered, partially adhered or mechanically fastened (Holzhauer 1997).
With each different contractor or roofing company there are a variety of different techniques and methods of installing each of the roof material types, but each technique should contain all of the five basic components as mentioned previously, roof covering, sheathing, roof structure, flashing, and drainage. Although techniques may sometimes vary between contractors, there is a standardized system for construction methods, called Insurance Services Office (ISO) codes, which will be discussed in the Appendix B. Builders and contractors must also be fully aware of the conditions in which they are working and the safety precautions that they are taking. By keeping a job site safe and within regulations, a contractor can ensure that a roof is installed without compromise, thus resulting in a roof with structural integrity that can resist some major causes of failure and can also last to its fullest potential service life.

One way in which safety is maintained and regulated is by the organization known as the Occupational Safety and Health Administration.

3.1.5 OSHA

The Occupational Safety and Health Administration (OSHA) is an organization designed to ensure a safe and proper working condition for working men and women by setting and enforcing standards by providing training, outreach, education and assistance. These standards are a way in which safety and quality can be maintained during construction jobs, which contributes to a reduction in common roof deficiencies and failures (At a Glance – OSHA). The OSHA was formed as a part of the United States Department of Labor, as a result of the Occupational Safety and Health Act of 1970, to prevent workers from being killed or seriously harmed at work. With the OSHA, there is a certain set of extensive standards that exist to protect the employees that must be enforced by employers at all times. Construction sites are
The Hanover Insurance Group – Roof Loss Control Analysis

suspect to on-site inspections without notice and failure to comply with the OSHA standards can result in penalties. Some examples of the OSHA standards include requirements to provide fall protection, ensure the safety of workers who enter confined spaces, prevent exposure to harmful substances and chemicals, provide respirators and other protective safety equipment, and provision of training for certain dangerous jobs (At a Glance – OSHA). The next section discusses the specific ways in which roofs are inclined to failure.

3.2 Roof Failure

Roof claims and the negative effects of failing roofs are an area that needs further explanation. The previous section describes the different styles of roofs, the materials used to build them, as well as some common construction methods. This section provides some insight into the shortcomings of these current methods along with addressing the specific failure causes called into question by The Hanover.

3.2.1 Causes of Roof Failure

There are multiple causes of roof failure; however a definition of roof failure is first required. A failing roof is a roof which does not protect the items below it. There are many causes of a roof failure such as weather and construction methods.

Prevalent causes of roof failure such as wind and natural disaster are very straightforward categories and are easily understood, therefore, these causes do not require explanation. For example, when a tornado strikes, the wind force can detach part of a roof. A later section will focus on specific causes that are applicable to this report such as snow loading, ponding and hail damage.
Another main cause of roofing failures is improper construction methods (Warseck 2003). The incorrect use or placement of flashing can lead to leakage and eventually, the collapse of the roof. Furthermore, the improper use of venting systems and overhangs in roof construction can result in a roof being blown away by high winds. The following section describes, in further detail, some of the common deficiencies that can lead to roof failures.

3.2.2 Common Deficiencies of Roofs

When a roof fails, it is usually because of one of two reasons, a large weather storm such as a natural disaster or the gradual deterioration of the roof’s integrity. As mentioned before, a roof failure due to a natural disaster is rather self-explanatory. However, there are many different factors that can lead to the gradual deterioration of a roof. Some of these include design flaws, improper installation, and the prevalence of foot traffic.

The most common reason for failure of roofing systems is design flaws. For example, the inclusion of details and corners in a design that cannot be created on the actual job site could lead to improper coverage of the roof, and a compromise of the structural integrity of the building. Another flaw that may lead to a compromise of the roof’s integrity is flashing deterioration due to weather. Many roofs leak due to a compromise of the flashing because without good tight flashings around chimneys, vents, skylights and wall/roof junctions, water can enter a building and cause damage to the interior (Consumer Info 2011).

Inadequate seam sealing is another very common reason for failure of roofing systems. Improperly sealed laps lead to seams that do not stick, or seams that appear to stick, but do not. Another common error is improper installation of a second roof over an existing roof without examining the moisture conditions underneath the original roofing system. Installing a
roofing system over a wet deck or existing roof can create or hide structural damage (Windle 2005).

An additional common problem seen on roofs is the prevalence of foot traffic. Even with a durable multiple ply built up roof, punctures can occur from dropped equipment and tools, sheet metal covers and access doors splitting the membrane, and tented flashings that are accidentally kicked while being stepped around. To protect against these damages, some roofs may have traffic mats, pads or boards. However, these traffic protectors are not strictly regulated and could actually cause more damage to the roof by imposing localized stresses, concentrating solar heat, or trapping water beneath them (Fricklaus 2004)

The next coming sections will describe in further detail the specific ways in which roofs are likely to fail and the maintenance and inspection techniques that could be implemented to ensure that every roof lasts to its full, potential service life.

3.2.3 Causes of Failure Specific to this Project

In the last few sections, there has been much discussion of the roofing system and the many problems that can occur on the roof. These failures can be difficult to pinpoint during an insurance appraisal and many times are overlooked because of the presence of a larger cause of roof failure. Most of these larger causes are related to specific weather occurrences and need further examination. For the purposes of this project, the group will inspect and discuss the cause of losses due to snow loading, ponding and hail damage.

3.2.3.1 Snow Loads

One major cause of a failing roof is a disproportionate weight load due to snow (Taylor 1984). The Hanover is headquartered out of Worcester, Massachusetts and a large portion of
their business is conducted in this region. Elevation differences between sections of a roof can cause snow drifts and potentially roof failure. Snow drifts that form on roofs where there is an elevation difference present can be detrimental to the structural integrity of the building (Taylor 1984). The snow loads study evaluated was conducted in Canada and is more than 20 years old, but the problem being discussed is still relevant today due to the number of claims related to snow damage on roofs and provides much useful information for our group. These drifts are dangerous because the density of the snow on the roof is significantly greater than if it were collected on the ground. The gradual melting and freezing of snow results in a compacting effect with an overload of weight in a condensed area. The magnitude of the elevation does have an effect on the severity of snow drifting on roofs. Snow can cause a roof to collapse because of such reasons as snow accumulation, roof contour, roof obstructions and energy conservation efforts that have damaged the integrity of the roof (Hoover 1996).

Preventative measures, which will be explained in a later section, are the best ways to avoid a roof failure. This includes regular inspections and other maintenance to keep the roof in a satisfactory condition. When snow loads melt, the remaining water can cause another type of failure known as ponding.

3.2.3.2 Ponding

Another phenomenon that requires further development is ponding. Ponding is the event that occurs when water gathers on a roof due to gutter back up from lack of maintenance or a compromised pitch in the roof. Ponding is most commonly seen on flat or low sloped roofs but can ultimately affect all roof types because when the water begins to back up, it begins to compromise the integrity of the roof structure. Once the integrity is compromised, leakage
occurs and this can cause a lot of damage to the contents of the building and further damage to the integrity of the structure as well. If left unchecked, the ultimate consequence of ponding and the subsequent water leakage can be roof failure, which will cause a massive amount of damage to the contents of the building and the building itself. In most cases of ponding, the damage that has occurred can become exponentially worse due to neglect or poor construction methods and therefore cause the insurance company to incur unnecessary losses.

One case study, performed in the Netherlands, attributes roof failures to the usage of low-quality materials, the failure to follow building codes as promised and lackluster inspections (Vammersky 2006). Another major cause is that the materials used on the flat roofs have become more lightweight and less reliable for cost cutting reasons and therefore the load capacity has decreased and failures, like those due to ponding, are more prevalent. A case study conducted by Luis Estenssoro discusses in great detail the civil engineering facts and nuances of roof failure due to ponding (Estenssoro 1989). The article provides an engineer’s perspective into the numbers behind ponding while relating it to how building codes and regular inspections can prevent such from occurring. Ponding can also be the result of melting hail, which is discussed further in the next section.

3.2.3.3 Hail Damage

Hail stones are comprised of frozen water that ranges from a few millimeters in diameter to a few hundred millimeters and could lead to roof failures. Hail storms can cause severe widespread damage causing out of state contractors to chase hail storms in search of business (Munoz 2000). In New England, buildings and home owners who have insurance are often covered for hail damage. Since The Hanover is responsible for the repairs, and damages
seen from hail are most often hard to notice, it is important to take extra care in evaluating a roof. Shingle granules will deteriorate from the repeated pounding of the ice. The roof will be left vulnerable to the elements, and roof failure will show more quickly. Loads of hail can pose as a serious threat to flat roofs. Damages from the hail loads will be comparable to the snow load damages. When melting occurs, ponding may be left. Even if there is no visible damage from the hail on the roof, it is important to still examine for damages closely after a storm. Hail damage usually does not lead to immediate leaks, but damage to the roofing materials can eventually make leakage most likely (Munoz 2000).

Now that the team has discussed the main causes of roof failures, the following sections describe the ways in which roof failures can be prevented.

3.3 Roof Failure Prevention

Preventing a roof from failing and ensuring that a roof lives up to its fullest potential life expectancy is an important, but often overlooked factor by the people involved with the occupancy and insuring of the building. It takes a full effort from all parties to prevent or identify roof damages before serious damages occur. Good roof maintenance can be one of the best ways to prevent roof failure (Munoz 2000). A roof that is not checked regularly can be subject to minor problems. The little problems, such as a stick puncturing the membrane, can snow ball in to the rotting out of the roof framing and the eventual failure of the entire roofing system. When a roof collapses, an insurance company will cover this loss most of the time. However, in some cases where it can be proven, the liability will be placed on the party responsible, such as the contractor or homeowner, if neglect or poor maintenance is found to
be the cause of failure. One way in which this can be avoided is by choosing a contracting company that performs maintenance and can have a positive impact on the overall success of the roofing system.

3.3.1 Choosing a Contractor

Building a roof that suits its environment is essential to life of a roof. The most intricate maintenance plan is pointless unless the structure of the roof is sustainable. Choosing a quality contractor rather than one with a marked reputation is the first step towards maintaining a functional roof. A poor contractor can use cheap materials or cut corners during the building process, whereas high-quality contractors have experience and reputation to back up their work (Reming 2001). If possible, The Hanover Insurance Group should try and be aware of which contractors built or renovated any roof that they insure. If a certain contractor’s work is continuously seen as contractors associated with claims to be filled, then this contractor could be “red flagged”. A good rule of thumb is to make sure the contractor has been in business longer than the warranty offered.

Maintaining a poorly built roof can prove to be more costly and time consuming than maintaining a roof that was properly built (Reming 2001). The next section outlines the different ways in which maintenance can be implemented as a roof failure prevention measure.

3.3.2 Maintenance Program

In his 2001 article, A Lasting Relationship, Thomas Reming claims that even the highest performing and high quality of roof need to be inspected and maintained at least twice a year. Checking a roof after a harsh storm is also common roof protocol. Debris needs to be cleaned
away, pipes and gutters must be unclogged, and roof must be inspected for ponding or other visible signs of defects such as blisters. Another possible way to set up a quality maintenance program is to develop one with the people who are responsible for building or repairing the roof (Reming 2001). Also having a maintenance program that incorporates the material manufacturer can be beneficial. Doing so ensures that short cuts are not taken when repairing a roof. Patching up a leak is cheap, but patchwork is going to put the lifetime of a roof in jeopardy. Damages can be far more expensive in the long run unless the damage is correctly identified and fixed before patchwork is needed.

Conditions, limitations, and maintenance requirements are often embedded into an insurance policy. These requirements must be taken seriously. A policy holder is most often responsible for having a roof properly checked and maintained. It is suggested to walk on the roof periodically to inspect for early signs of roof failure (Reming 2001).

3.3.3 Roof Inspection

Roof inspections are the easiest way to prevent and protect a roof. Having a few inspectors checking a roof and the interior is very important. There are three areas that inspection can be broken into. These three areas are the simple maintenance issues, such as clearing drains and simple patchwork, the general repairs needed that has been identified through thorough inspection, and finally the methods needed to extend the service life of the roof such as a new reflective coating or flashings (Roof Inspection Checklist).

To ensure a complete inspection, checklists are used to make certain that maintenance programs are followed (Roof Inspection Checklist). Checklists should first incorporate the use of a database for the roof. This should include the entire history of the roof, such as roof plans,
pictures of repairs, previously finished inspection checklists, materials used, and also warranties and insurance policies.

3.3.3.1 Interior Checklist

Proper inspection can start by inspecting the interior under the roof. Items of interest would be characteristics such as cracks, paint peeling from the surface, watermarks, or even leaks (Roof Inspection Checklist). It is imperative to mark the roof plans with the location of each interior item of interest. This practice makes the roof inspection process as accurate and productive as possible (D’Annunzio 2008).

3.3.3.2 Exterior Checklist

Roof plans should first be checked for accuracy regarding all equipment and penetrations. Size and area of the roof should be confirmed as well. It is also important to make sure the information on the membranes and flashings are accurate (LaDuke 2008). General Roof conditions should be noted. The conditions of most interest would be the amount of debris, drainage, blatant physical damage, and obviously structural damages (Roof Inspection Checklist). Following that, inspectors should begin to look closely at the flashings for defects. The edges of the roof and also the penetrations should be examined for defects as well. Drains need to be cleared and pools of water must be noted in plans and vacuumed. The membrane is the next item to be inspected. The entire membrane, especially the areas subject to high foot traffic, need to be carefully assessed. Items of interest would include blistering, pooling, alligatoring, punctures, or fish-mouths. Also the condition of the coating should be noted. All the defects need to be marked with spray paint. Also, roof plans need to be updated with the location, type, and severity of the defect. Pictures of the defects should be taken. All minor
repairs that can be fixed by the inspectors can be performed next. The repairs also need to be
documented on the roof plans and photographed. More drastic repairs must be completed
using materials approved by the membrane manufacturer. The fixes must be carefully added to
the database because often repairs can lead to new defects (Roof Inspection Checklist).

The reality of the situation, however, is that roof failure does occur and when it does,
insurance companies provide the necessary coverage needed to overcome these roofing
failures. The next section details the insurance policies and coverage available from most
insurance companies as well as the loss control methods these companies use.

3.4 Insurance

When looking at roof loss, it is important to assess the category and type of insurance
policies these roof claims fall under. The two major categories of insurance policies analyzed
here are homeowner insurance and business insurance. Homeowner insurance typically offers
coverage in the following areas dwelling, additional structures, personal possessions, liability
claims, and medical payments (About The Hanover Insurance Group 2011). The second
category of insurance is business insurance. Business insurance, unlike homeowner insurance,
includes different insurance policies depending on the size and type of business. Typical
business insurance policies cover property, liability, commercial auto, health, life, and disability
(About The Hanover Insurance Group 2010). In this particular project, all roof loss claims are
commercial and therefore categorized under business insurance, and are covered by property
and additional structure policies. The next coming sections look more in detail at these
3.4.1 Property

Property insurance insures against loss or damage to the location of a company and its contents. The insured location can be owned, leased or rented. It also protects the property of other locations in the company’s control when loss occurs. Property insurance can be broad or specific. There are some business insurance policies that only insure against loss from specific perils such as fires. However, most insurance companies offer a packaged property policy or “business owners’ policy” (Insurance 2011). The premiums charged for business insurance vary widely from company to company. Insurance premiums are dependent on a number of factors, including location, age and type of building, use of building, local fire protection, choice of deductibles, application of discounts, and the amount of insurance policies one chooses to purchase. Conversely, there are some businesses that are not eligible for this package. Companies that have high associated risks, such as power plants, have to pay higher premiums for their property coverage. Lastly, businesses must work with insurance companies to keep their schedules up to date. Schedules are lists of the covered locations and property controlled by the business (Insurance 2011). They also must be updated anytime the business moves to another location or covered equipment changes.

3.4.2 Additional Structures

Business insurance also provides coverage on detached structures in addition to the main commercial property. Common structures include warehouses, offices, and garages. Often
additional structures are included within the “business owner’s policy”. The next section describes ways in which insurance companies attempt to reduce and control the amount of losses they incur.

3.4.3 Loss Control

Loss control is the ability to define areas of potential loss before they occur or to prevent further loss from accidents that have occurred in the past. To be fully effective, a loss control program must place great emphasis on known and suspected loss producing problems. To help target loss control efforts, insurers provide an in-depth analysis of past loss experience that includes statistical data on prevalent accident types, causes and uninsured cost factors (Kunreuther 1996). This enables insurers to see trends from loss causes and develop a cost effective method of control (Lee 2011). Overall, loss control is a way for insurers to control costs, effectively manage claims, and enhance profitability. One specific way in which losses can be controlled is through the evaluation of past losses using the Actual Cash Value of the loss.

3.4.4 Evaluating Loss

Actual Cash Value (ACV) is the most common calculation firms use to evaluate a loss (Actual Cash Value 2006). The ACV is simply the Replacement Value minus depreciation. Unlike Replacement Value, a loss or damage is valued at the value of the property loss (Actual Cash Value 2006). For example, if a roof claim is processed and the roof is fifteen years old its replacement value may be $100,000, but that is not the Actual Cash Value. After subtracting 20 years of depreciation ($4,000 per year), the roof may be actually valued at $20,000 at the time
The Hanover Insurance Group – Roof Loss Control Analysis

of loss. Therefore, the insurer will only pay for $20,000 worth of damages. Insurance firms using ACV typically have lower premiums due to their smaller loss payouts.

Now that a foundation of general roof and insurance knowledge has been established, the group has developed hypotheses using the knowledge gained from prior literature.

3.5 Hypotheses

Readings of the literature and discussions with staff from the Hanover lead to the following hypotheses:

1. As buildings increase in age, the likelihood of the building collapsing due to snow loads decreases.
   - Newer buildings are built with stricter budgets and are forced to cut costs, resulting possibly in the use of less expensive materials. This might cause newer buildings to be more prone to collapse due to snow loads.

2. As roof covers increase in age, the likelihood of a claim increases.
   - The older the materials of the roof, the longer they are exposed to weather, and the higher the likelihood of failure.

3. If the building has an elevation difference, then the roof has an increased likelihood of suffering a snow load or collapse claim.
   - When the roof of the building is not continuous and creates a meeting of a piece of roof and the wall of higher level, then this meeting point is suspect to snow drifts creating increased propensity for collapse.
4. If there is an addition added to the building, then the roof is more susceptible to failure.
   ▪ When there is an addition to a building, the entire building is not evaluated as one structure, rather as two separate structures. Therefore, the connection between the two structures is susceptible to increased loads causing failure.

5. The prevalent causes of roof claims are hurricanes, hail, and wind in order of significance.
   ▪ The most frequently occurring and costly claims are those that are due to hurricanes, hail, and wind, respectively. This is because these loss types can be the most destructive.

6. Roofs that are subjected to regular maintenance and inspection will be less likely to incur roof failures and claims.
   ▪ Regular maintenance and inspection can prevent most of the common causes of roof failure. Therefore, these roofs are less likely to have a claim.

7. Location of the building will have an effect on the susceptibility to certain types of damage.
   ▪ The geographical position of the building will determine the most likely weather related causes of failure. For example, buildings in the mid-west are more susceptible to tornado claims compared to buildings in the south-east which are more prone to hurricanes.

8. As the roof’s square footage increases in size, the claim amount increases as well.
The Hanover Insurance Group – Roof Loss Control Analysis

- The larger the area of the roof, the more damage can be sustained from the incident. Therefore, when there is a claim filed, the larger the area of the roof will lead to a larger dollar amount of loss.

9. If the roof has no pitch, then the roof is more susceptible to snow or collapse claims.
   - A roof with no pitch will not drain water as well as a roof with some degree of pitch. Therefore, when snow accumulates, these types of roofs will bear a larger load causing collapse.

The next section will describe what the team did to achieve the goals of this project. The various steps that were necessary to collect the data from the various databases will be explained as well as the statistical tests that were needed to test the hypotheses.
4.0 Methodology

The goal of this project is to produce a report that will provide information on current roof loss trends as well as recommendations to improve the current roof evaluation process. In order to organize the collected data, a list of the desired variables and their descriptions follows:

<table>
<thead>
<tr>
<th>Table 1. Variable Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Addition</strong></td>
</tr>
<tr>
<td><strong>Age of Building</strong></td>
</tr>
<tr>
<td><strong>Age of Roof Cover</strong></td>
</tr>
<tr>
<td><strong>Catastrophe Code</strong></td>
</tr>
<tr>
<td><strong>Claim Number</strong></td>
</tr>
<tr>
<td><strong>Claim Status</strong></td>
</tr>
<tr>
<td><strong>Date of Loss</strong></td>
</tr>
<tr>
<td><strong>Dollar Amount of Loss</strong></td>
</tr>
<tr>
<td><strong>Elevation Difference</strong></td>
</tr>
<tr>
<td><strong>ISO Building Construction Code</strong></td>
</tr>
<tr>
<td><strong>Loss Type</strong></td>
</tr>
<tr>
<td><strong>Name of Insured</strong></td>
</tr>
<tr>
<td><strong>Number of Stories</strong></td>
</tr>
<tr>
<td><strong>Occupancy</strong></td>
</tr>
<tr>
<td><strong>Pitch</strong></td>
</tr>
<tr>
<td><strong>Square Footage</strong></td>
</tr>
<tr>
<td><strong>Street Address</strong></td>
</tr>
<tr>
<td><strong>Type of Roof Cover</strong></td>
</tr>
</tbody>
</table>
Achieving the collection of this data requires an extensive analysis of databases (see Appendix C) both within The Hanover and other external sources. These variables will be collected through the process depicted below:

**Figure 1. Methodology Flow Chart**

This is a complex process requiring many types of expertise and data analysis. The following section describes, in detail, the databases that were used to extract as much data as possible. This section provides the steps to analyze the data.

### 4.1 Convert PDF to Loss Data Excel file

A PDF was given to the team by The Hanover and it contained about 3,000 of the largest (the largest claims were $100,000 and above) commercial claims that were filed over the past
The Hanover Insurance Group – Roof Loss Control Analysis

twenty years. The file contained important information about each claim, including, the claim number, the name of the insured, the amount of the claim, the date of the claim, and a description of the nature of the claim. These were the data points that were the most critical for the team to input into the Loss Data Excel file for further research to be executed.

Since there were various types of claims within the PDF, the information needed to be sorted by whether it was a roof claim or not and transferred to the Loss Data Excel file (See Appendix E) created by the team. The variables that were transferred from the PDF to the Loss Data Excel file were Insured Name, Date of Loss, Claim Number, Claim Status, Dollar Amount of Loss, and Loss Type. This was necessary because the group was unable to edit the PDF and all the claim data needed to be centralized for analysis. Transferring and organizing the data into the Loss Data Excel file made it easier to analyze data and identify trends. This was achieved by first utilizing the search function in Adobe Reader. The word “roof” was typed into the search box and a list of all claims pertaining to roofs was generated. The first set of claims that appeared supplied the group with the aforementioned information contained in the loss spreadsheet. This information was then analyzed and the team decided whether or not the claim was relevant. The claim’s relevancy was decided by determining if it was primarily a roof claim or whether it was a claim in which the roof was included as part of a total loss, such as in a fire. The claims involving fire were deemed irrelevant and therefore omitted from the search. The relevant information was then transferred to the Loss Data Excel file. This process was repeated until the entire document was searched for relevant claims, and all of the data was transferred. When this was completed, the Loss Data Excel file contained 405 roof claims. After the completion of the collection of the group’s data, a small sample of the data points was
selected to be validated for accuracy. This will ensure confidence in the group’s collection of data. The 405 claims needed to be researched further, in various databases, in order to obtain information to complete the Loss Data Excel file.

4.2 Hanover Database Information Retrieval

At this point, the PDF has provided the following variables: Insured Name, Claim Number, Claim Status, Amount of Loss, Loss Type, and Date of Loss. The team used four of The Hanover’s databases to gain Claims Addresses and if there was a Building Underwriting Report (BUR) available for the specific building, then all variables could be addressed for that particular claim. The Hanover Claims System (HCS) was one of the databases that provided the correct address of the loss of the claim as well as some specific information regarding the roof claim. The Hanover’s Claims Service System (CSS) was a similar database to HCS in that it provided addresses for the older claims on file at The Hanover. The CAAMS database produced information that was used to populate as many of the variables as possible. ARIES was a similar database that provided both addresses of the insured building as well as some preliminary information on the roof.

4.2.1 HCS

The claims system utilized by Hanover is HCS. This system proved valuable to the group because it provided a database to search all of the claims in the Loss Data Excel file from 2007 to the present. The search was based upon the claim number and the result of the search yielded the building’s address and whether or not there was a catastrophe code. The group was also able to gain information regarding the catastrophe code of each of these claims; however,
some claims numbers that began with 03 returned no results at all. See the HCS section in Appendix C for step by step instructions as to how the group gathered information from this database. After all the claims were searched in HCS, there were still many claims that did not have any address. The catastrophe codes were found for all the claims that started with a 14-xxxxxxx. The claims starting in a 03-xxxxxx or 56-xxxxxx needed to be found in an older database. Accordingly, the group used the CSS database.

4.2.2 CSS

As the team worked down the Loss Data Excel file, it became evident that claims beginning with the numbers 03 were older claims and were not accessible through the HCS database. Therefore, to gain the address and catastrophe code for the remaining claims, the group searched the remaining claim numbers in the CSS database. The address is needed by the group in order to be able to search for additional information in other databases such as CAAMS and ARIES, as these databases are not able to be searched using the claim number. Step by step instructions as to how the group conducted the search can be found in the CSS section of Appendix C. Once every claim has been searched for its address and catastrophe code, the team moved on to search the addresses in the CAAMS database for BURs and other reports that yielded data for all other variables.

4.2.3 CAAMS

CAAMS is an internal system for policy information at The Hanover and cannot be accessed outside the company. Documents within this system, such as Building Underwriting Reports, help describe the physical conditions at risk for the insurer. The CAAMS database
The Hanover Insurance Group – Roof Loss Control Analysis

houses all policy information, including any claim activity and documents as previously described.

After the address of the claim number was retrieved from HCS or CSS, this information was entered into the CAAMS database to find additional information on the claim. A detailed breakdown of the step by step instructions can be found in the CAAMS section of Appendix C. If the address produced any documents, then information such as square footage of roof, whether the roof is pitched, ISO codes, and the age of the building, was entered into the Loss Data Excel file. Since the group was unable to find information for every claim, the team needed to search the remaining addresses with no information in ARIES.

4.2.4 ARIES

The ARIES database is a web based application available via the internet; however one needs valid credentials from The Hanover in order to access this database from outside The Hanover’s building. ARIES is the loss control reporting system for The Hanover. It generates a comprehensive Survey Request that is first processed by the underwriters, and then electronically transmitted from the underwriter to the Loss Control Representatives. The Survey Request for each claim found in ARIES provided similar information to the BUR. These Survey Requests provided information on many variables that were still needed in the Loss Data Excel file. ARIES allows its users to search for reports by the name of the insured party. The ARIES section of Appendix C contains the step by step instructions for gathering the appropriate information from the database. The group searched through the reports contained within ARIES for information on all the remaining variables such as roof type, elevation difference, building age, etc.
After a thorough investigation into The Hanover’s databases the group had collected approximately 50% of the data for the variables. Therefore, there were still holes within the group’s data such as square footage, elevation difference, pitch, etc., and the team found it necessary to use outside resources. The Hanover suggested some online building information databases from specific regions of the country that would help fill in these gaps. The team also found similar databases for other geographical regions that were not covered by the databases suggested by The Hanover.

4.3 Online Building Information Databases

Since the databases provided by The Hanover did not provide data for each claim, the group searched independent databases available to the general public online. These databases were chosen after a recommendation by The Hanover and after the group’s own research into the sites. These databases provided data for numerous variables including addition, age of the building, age of roof, ISO building construction code, type of roof material, number of stories, elevation difference, square footage, pitch, and occupancy. The databases searched were mainly appraisal and property assessment websites. These databases included Vision Appraisal, Appraiser Central, Dallas Central Appraisal District, Property Shark, Tarrent Appraisal District and Property Assessment Directory.

4.3.1 Vision Appraisal

Vision Appraisal, http://www.visionappraisal.com/databases/index.htm, is an online database that provides users with tax and property information used for tax re-evaluation. The database unfortunately does not cover much more than the states within New England. Even
within the available states, many towns and cities were missing information. The information searched for in this and all subsequent databases included information from all necessary categories for the various claims. This information included, but was not limited to, categories such as square footage, elevation difference, pitch and others.

To find missing data points for the claims listed in the Loss Data Excel file, the team first needed to search the spreadsheet for claims within New England. After the group had located the claim, the team proceeded to click on the corresponding state in Vision Appraisal. After clicking the correct state, a list of available cities will pop up in alphabetical order. Next, the team checked whether or not the city is one of the options available. After the team found the correct city name, Vision Appraisal sometimes asked the group to sign in. This login just requires filling in a name and address. Once the team successfully “logged in”, the next page of the search required the team to enter the street name and number of the claim desired. A list of options appeared, hopefully, one of which included the desired claim. Once the claim was found, the team clicked on “View the Property”. This provided pictures of the property, if available, and some, if any, of the information needed to fill in the data in the Loss Data Excel file. Once all the information has been collected the group used the back button to return to the state selection page. The continued to search using this process until there were no longer any claims from the New England states remaining.

The group was able to gain five data points from this process. After thoroughly searching the databases at The Hanover and Vision Appraisal, the group still had approximately 25% of the claims from Massachusetts that needed to be researched further.
4.3.2 Appraiser Central

After finding that tax assessor data was helpful, a Google search for “Massachusetts Assessor Database” was performed and the link to Appraiser Central was returned as a result. Appraiser Central is another online database that is very helpful for providing tax assessor and zoning information for claims within Massachusetts.

It was necessary to sort through the Loss Data Excel file to identify claims with addresses from Massachusetts that still were lacking data. After locating all of the Massachusetts addresses, the group proceeded to open Appraiser Central (http://www.appraisercentral.com/st/massasses.htm). The website lists the cities and towns of Massachusetts in alphabetical order. After the team clicked on the town or city corresponding to the desired claim, a page specific to that town or city appeared. The team clicked on the link that leads to the online database of Assessor Records. The next page displayed is a search engine for the town database that has now been accessed. Again, using the information from the Loss Data Excel file for the corresponding claim, the street address was entered into the search engine in the corresponding boxes. After the correct information had been entered, the search results were produced on a new page with a list of possible addresses. When the correct claim address had been identified, the group clicked the Parcel ID # on the left side of the page to view the full information. Using the picture provided and the narrative description, the group was able to pull information on various characteristics of the building such as type of roof cover and year built. After extracting all the useful information, the tabs for the specific town can be closed. Repeat the process for all towns in Massachusetts.
The team found four observations from this source. After searching the New England area as much as possible, the group moved on to another large concentration of claims that were missing in The Hanover databases. That next area was the Dallas, Texas region and further research outside The Hanover databases was required.

4.3.3 Dallas Central Appraisal District

Dallas is a city that contains numerous claims within the Loss Data Excel file. These claims have missing data points which could be filled in by this database. The group was able to gain information on building age, number of stories, pitch, elevation difference and square footage from this database.

In order to gather this information, the team opened an internet window and input the address http://www.dallascad.org to bring the user to the Dallas Central Appraisal District (DCAD) homepage. After, the team clicked on the Search Appraisals link under the list of Navigation Links on the left side of the screen. This brings up a search engine that asks for the Owners Name. Since the Loss Data Excel file is somewhat vague in terms of the name of the claim, the team needed to access the search by address option. This was done by clicking the blue link above the search engine titled Street Address. By inputting the claim address and city into the search engine, the database can be accessed and the corresponding information for that address will be displayed. When there were multiple options offered, the group used the information in the Loss Data Excel file, such as name of insured, to confirm which option is the correct building. This database in particular proved helpful in providing pictures of the buildings, which allowed the group to gauge the pitch of the roof. It also contained property information such as square footage and the number of stories, which were also part of our
research. After completing the search, our team simply returned to the ‘find property by address’ search engine page and entered the next claim address. After, the team repeated this process as many times as necessary.

The team found 16 data points from this website. Staying in Texas, the group then moved on to the next concentration of claims in the state, Tarrent County.

4.3.4 Tarrent Appraisal District

This database is used to search for appraisal documents and data in Tarrent County, Texas. Since we have claims that pertain to that area, this database was suggested to us by The Hanover. Tarrent Appraisal District (TAD) is a public site (http://www.tad.org) that required no signing in or the entering of credentials. First, the team opened the TAD homepage and selected the ‘Property Data Search’ located on the left side box of the homepage. This opened a list of search options. Then the team clicked on the ‘Search Location by Address’ option. It was important to select the ‘Business/Personal Property’ option, not the ‘Real Estate’ option. The team had to use the drop box selection to pick the correct city, if available. The team entered the street name and number and executed the search. Using the Loss Data Excel file document, the group selected the correct search result and opened the appropriate file. After reviewing the document for any information pertaining to the Loss Data Excel file such as building specifications and the ISO construction method, the team proceeded to record the findings. For the group to repeat the search for other claims from this county, the team returned to the ‘Search Location by Address’ page. The team repeated this process as many times as necessary.

Since this website was specific to the district of Tarrent which provided limited data, the team still needed to research claims that were located in other locations of the state.
4.3.5 Property Shark

Property Shark (http://www.propertyshark.com/) is a tool that the group found to search for the remaining claim information such as square footage, number of stories, slope of roof and whether an elevation difference was present. This site was recommended to the team by The Hanover in order to further search for claim information. This site was specifically helpful for claims located in metropolitan areas.

This site required the group to create an account and sign in. After a few minutes an email was sent to confirm the account has been verified. Once the account was confirmed, the team was able to enter an address, along with a corresponding city and state. After entering the search criteria, the site returned a list of search results that could possibly be the building in question. After opening the file, data and pictures were usually available. When all the pertinent data has been extracted, the group initiated a new search by returning to the Property Shark home page. The team repeated this process as many times as necessary.

The team found two data points from this process. After searching the large amount of claims in Texas, a tool to research the rest of the claims scattered across the country that still required information was needed. The tool that satisfied that need was the Property Assessment Directory.

4.3.6 Property Assessment Directory

The Property Assessment Directory (http://www.propertyassessmentdirectory.com/) will provide assessor information, which is broken up by county. This website did not require login information.
First, the group conducted a Google search to find the claims county using the city name. The Property Assessment Directory directs the user to the assessor database. The team was able to pull up a list of counties after clicking on the correct state link. Sometimes, multiple counties were found in a region database, much like Vision Appraisal. If this was the case, the counties pertaining to that specific database are listed in the Property Assessment Directory. The group selected the correct county for the roof claim and the website redirected the group to a more specific, county database. The group used the address from the Loss Data Excel file to target the correct building. After the database produced the file corresponding to the correct building, information on the roof was extracted. If a picture or sketch is provided, it is easy to determine if the roof is pitched or has an elevation difference.

After all the possible information had been acquired, the team returned to the Property Assessment Directory homepage and began working on the next claim location. The team repeated this process as many times as necessary.

From this database, the group was able to gather information on 19 claims. In looking at the information gathered thus far, it was clear that the team still had many claims that did not have information in the Loss Data Excel file, and that Google and Google maps could help the team fill in the blanks within the Pitch, Elevation Difference and Square Footage variables.

**4.4 Google**

Another avenue for accessing useful information on the claims was to use Google searches. One can search for the building name provided on Google. From these searches, real estate websites would provide general information on the roof, such as the year built. A picture...
accompanies the real estate information as well. The pictures can be used to determine the presence of elevation differences, the number of stories, pitched roofs, and sometimes roof materials. The team could use this information gathered from the Google search once it was validated that the pictured building was the correct building from the claim.

4.4.1 Google Maps

Searching an address on Google will trigger Google Maps. This is another way to obtain a satellite image of the building. Most of the time, the claim address is the correct address of the location of the roof failure, however, sometimes the address on the Loss Data Excel file is not the actual address of the location where the loss occurred. The address may also be unclear on Google Maps as to which building the address is describing. Google Maps will provide insight into whether there is a pitched roof and the number of stories. The Google Earth plug-in can also be used to help determine if there is an elevation difference, and can approximate the square footage by utilizing the distance tool. Hanover employees have access to this tool; however, the team needed to use an outside resource to determine the square footage of the building.

4.4.2 Free Map Tools

For many of the claims, the group could not locate any information through the various databases used by The Hanover. Since the team wanted to come up with the most possible quantity of data points, Google Maps was used in conjunction with an area calculator map tool to determine the square footage, pitch and any possible elevation change of the roofs. The area calculator we chose to use was on the website freemaptools.com.
This tool can be accessed at the following web address, http://www.freemaptools.com/area-calculator.htm. Once on the website, one can enter the claim address into the search bar in the bottom left hand corner of the map on the screen. This will bring the user to the address of the roof claim desired. Zoom in on the map to better view the roof, and then, in a clockwise fashion, select the corners of the roof to create a polygon in the same shape as the roof. Below the map will display various metrics of the area of the polygon drawn on the map. For our purposes, the metric labeled in meters squared will be the most useful. Since the majority of the data had previously been in feet squared, the group then converted the area from meters squared to feet squared with an online conversion calculator located at the web address, http://www.metric-conversions.org/area/square-meters-to-square-feet.htm. This square footage was input into the Loss Data Excel file document. The process was repeated as many times as necessary. The team was able to gather additional information for eighty different claims.

The team exhausted all possible databases and there were still missing data points. The following table depicts which database or website provided information for each variable:
If a variable (the list along the left side of the table) has an ‘X’ in its’ row, then the column heading (the data source) is where the information could be collected. Data mining provided information about the property associated with the loss; however, an alternative source of collecting information about the process was utilized. Interviewing employees of The Hanover who were experts in their respective fields allowed the group to gain further insight into the process and formulate conclusions that would not be possible from data mining alone.
4.5 Interview Protocol

Due to the team’s lack of expertise within the insurance industry, the group decided it was necessary to interview employees in areas that needed more insight. The interviews were meant to be a secondary method by which the team gathered information. The information gathered from the interviews would supplement the information gathered in the Literature Review as well as the statistical analysis. Joan Wooley recommended the employees that were chosen to be interviewed. Many individuals were contacted in hopes of gaining an interview with each department; one within loss control, one within underwriting, as well as one within the claims department.

The interview was to be conducted in person and to follow a semi-structured interview protocol. Each interview had a specific set of questions that were tailored to the interviewee (see Appendix C). Flexibility was allowed for additional questions to be added by the interviewer to gain more information regarding the interviewee’s domain of expertise. One member was designated to speak with the interviewee for each interview; while the other member’s responsibility was to take detailed notes. These roles were not to be strictly enforced. The group wanted to ensure any questions that might arise, would be asked.

Initially, the team anticipated that the interviewees would be asked all the questions from the protocol and in the order listed. However, due to the different settings and structures of each interview, questions from the interview protocol were asked, as needed, in a more natural order. It was not necessary for all protocol questions to be asked during the interview, either
due to interviewee providing relevant answers on their own or the questions fell outside of the expertise of the interviewee.

The information gathered through this process provided an insight into how The Hanover process works. Conducting these interviews would enable the group to gain an understanding of how the loss control consultants access the building, how the premium is generated, and finally, how the loss control employees file the claim. With this information and the analysis of the data and statistics; the team was able to make recommendations to The Hanover.

4.6 Statistics

The following sections describe the statistical methods used to analyze and interpret the collected data in the Loss Data Excel file.

4.6.1 Statistics

Statistics is the process of collecting, organizing and analyzing data. The previous sections have covered the ways in which the data has been collected and the upcoming section will describe the different ways in which we organized and analyzed this data. Although the group was unable to obtain any control groups, the group attempted to test hypotheses with regressions and ANOVAs.

4.6.2 Determining Outliers

For some of the group’s frequency tests, it was observed that the data collected contained points considered to be outliers. In order to specifically determine and remove the outliers to obtain more representative data, the median was first calculated. After the median
was found, the upper and lower quartile ranges (UQR and LQR) were found by completing the formulas: $UQR = 0.75(n + 1)$ and $LQR = 0.25(n + 1)$, where $n$ equals the number of total claims.

Next, the interquartile range (IQR) was found by evaluating the formula: $IQR = UQR - LQR$. Once the IQR was known, the group could compute the formulae for the upper and lower outer fence (UOF and LOF). These formulae are: $UOF = UQR + 3(IQR)$ and $LOF = LQR - 3(IQR)$. Going forward, any points that were outside the limits of the upper outer fence and lower outer fence would be considered outliers.

4.6.3 Regression and ANOVA Tests

This upcoming section details the specific method in which the regression tests were performed by the group. The team conducted regression tests to identify if the independent variable(s) are a significant predictor of the dependent variable. To do this, there are three main outputs from the regression that need to be identified.

First, the adjusted R squared represents the variance explained by the regression model. Accordingly, it describes how well the model predicts the data. The ideal number for this variable is 1. However, a 1 indicates that 100% of the data is explained by the model and this is something that very rarely happens. Therefore, when the group was conducting regression tests, the goal was to have an adjusted R squared as close as possible to 1.

The second regression variable that needs to be examined is the $p$-value. This number identifies whether the associated independent variable is a significant predictor of the dependent variable. Ideally, the smaller the $P$-value, then the variable has stronger significance. However, due to the nature of the data, the group allowed $P$-values of 0.1 and even 0.2 to be of
acceptable significance. This was only to include control variables, not to necessarily determine the significance of the model.

The third and final regression output that needs to be examined is the coefficient of the independent variables. The sign of the coefficient is important to test whether or not the independent variable is associated with an increase or decrease of the dependent variable. However, if the coefficient is not significant, then the coefficient is not significantly different from zero.

For ANOVA tests, the mean differences are evaluated. This can be seen by looking at the corresponding value for the Significance of $F$. In order to prove significance, the ideal value for the Significance of $F$ is 0. The further the value gets from 0, the more insignificant the test becomes. The insignificance of an ANOVA test implies that the differences between the means are not great enough for the group to say that they are different. However, the significance of an ANOVA test implies that the means differ more than what would be expected by chance alone. The Significance of $F$ value is visible in the table produced by Excel when the test is conducted.

In Excel, when a regression test is run, the result is given in two tables. One of which is an ANOVA table and the other is a regression table. For each test, only one of the result tables is used, based upon the data tested. The results and analysis of the regression and ANOVA tests, as well as a detailed breakdown of the data collection process, will take place in the Findings section.
4.6.3.1 Age of Building per Number of Claims

The group wanted to evaluate hypothesis 1, that as buildings increase in age, the likelihood of the building collapsing due to snow loads decreases. In order to test hypothesis 1, it is necessary to test whether older buildings are more likely to incur roof collapse due to snow than newer buildings. The group could only obtain information for buildings that had incurred roof failures and claims. Thus, the team was unable to fully test hypothesis 1 because of the lack of a control group; however, the team wanted to investigate this prediction as far as possible. Therefore, the group ran a regression to evaluate a slightly different question. The group wanted to see if the age of the building helps predict the amount of the loss incurred.

In order to run this regression test, the group used Age of Building Code as the independent variable, Log Transformation of Sq. Footage as the control variable and Dollar Amount of Loss as the dependent variable. The building age categories used were from 1900 to 1919, from 1920 to 1939, from 1940 to 1959, from 1960 to 1979, from 1980 to 1999, and from 2000 years to the present year. Next, the group gave a code number of 1 to any building built after 1980 and 0 to any building prior to 1980. This code number was presented in a column labeled Age of Building Code. To normalize the data the Log Transformation of Sq. Footage was calculated by taking the natural log of each square footage. The group chose to control the square footage variable because it is possible that larger roofs are associated with larger costs of the roofs and accordingly larger claims. For the purpose of this test, the regression table results were used in the group’s analysis because the Log Transformation of Sq. Footage variable is continuous.
4.6.3.2 Age of Roof Cover per Number of Claims

The group wanted to test hypothesis 2, which states, as roof covers increase in age, the likelihood of a claim increases. In order to test hypothesis 2, it is necessary to record the age of roof covers for all the buildings insured by The Hanover that have not experienced roof loss claims. With this information, the group could then compare the ages of the roof covers for the buildings that do not have roof claims with those that have experienced roof loss claims. Since this information was unobtainable by the group, hypothesis 2 was unable to be fully tested at this time. However, further discussion of this hypothesis takes place in the Future Research section.

4.6.3.3 Number of Claims with Evidence of Elevation Difference and Loss Types of Snow or Collapse

The group wanted to evaluate hypothesis 3, which states that if there is evidence of elevation difference then the roof will be more prone to snow load and collapse damages. In order to test hypothesis 3, it is necessary to look at the variables Elevation Difference, Loss Type and Secondary Loss Type. These variables were first sorted by elevation difference (yes or no) and then again by loss type. Since the data for the variable Elevation Difference was limited, the group was unable to run an ANOVA test on this hypothesis. However from the data that was available, one pivot table was created using elevation difference as a row label and the sum of count as the values. A second pivot table was created using elevation difference as a row label, loss type as a column label and sum of count as the values. These tables and an analysis of the data available can be seen in the Findings section.
4.6.3.4 Number of Claims with Evidence of an Addition

The group wanted to evaluate hypothesis 4, which states that if there is an addition added to a building, then the roof is more susceptible to failure. In order to test hypothesis 4, it is necessary to record the number of buildings without a roof loss claim and that have an addition as well as those buildings that have an addition and also have a roof loss claim. With this information, the group could then compare these two numbers to evaluate whether or not the presence of an addition to a building makes the building more likely to experience a roof loss. Since the information on buildings without roof loss claims was unobtainable by the group, hypothesis 4 was unable to be fully tested at this time. However, further discussion of this hypothesis takes place in the Future Research section.

4.6.3.5 Dollar Amount of Loss Compared to Loss Type

The group wanted to evaluate hypothesis 5, which states that the prevalent cause of roof claims are hurricanes, hail and wind in order of significance. In order to test hypothesis 5, it is necessary to test the total amount of claims involving hurricanes, hail and wind versus the total amount of other weather related claims. To do this the group wanted to run an ANOVA that looked at the variables for Dollar Amount of Loss and Loss Type.

The group first separated the data into two categories, one for natural disasters and another for naturally occurring weather related losses. In this test, a natural disaster was classified as a hurricane or tornado and all other loss types were considered naturally occurring weather related losses. For the purposes of testing, all collapses were assumed to be due to a weather related event and were included in naturally occurring weather related losses. Losses that were categorized as Vandalism or Other were omitted from this frequency test. From this
data, the team was able to run multiple ANOVAs because the variable Loss Type was not a continuous variable.

The group ran three separate ANOVA tests, one for natural disaster loss data, one for naturally occurring weather related loss data, and one for all weather related loss data. Each ANOVA test used Loss Type as the independent variable and Dollar Amount of Loss as the dependent variable.

After the three original ANOVAs were completed, the group thought that adding the square footage to this test may prove to add further clarification to the data. Therefore, the group brought in the control variable Log Transformation of Sq. Footage and created three new groupings for natural disaster loss data, naturally occurring weather related loss data, and all weather related loss data, this time with the Log Transformation of Sq. Footage variable in addition to the Loss Type and Dollar Amount of Loss. The group then again calculated and removed the outliers from this data and performed three new regression tests. For these regression tests, the group used Loss Type as the independent variable, Log Transformation of Sq. Footage as control variable, and Dollar Amount of Loss as the dependent variable. The results and analysis of all six of these tests can be seen in the Findings section.

4.6.3.6 Number of Claims with Regular Maintenance

The group wanted to evaluate hypothesis 6, which states that roofs that are subjected to regular maintenance and inspection will be less likely to have claims filed against them. In order to test hypothesis 6, it is necessary to first evaluate the number of buildings that are subject to regular maintenance and do not have a roof loss claim. Since this data is not available
to the team, hypothesis 6 cannot be fully evaluated at this time. However, further discussion on this hypothesis takes place in the Future Research section.

4.6.3.7 Number of Claims per State by Loss Type

The group wanted to evaluate hypothesis 7, which states that the location of the building will have an effect on the susceptibility to certain types of damage. In order to test hypothesis 7, it is necessary to evaluate the variables for State, Loss Type and Dollar Amount of Loss. Ideally, the group would have liked to compare the number of claims in each state to the total number of policies for that state. However, the group did not have access to a control group of the total number of policies in each state to run the proper regression test. Since a regression test could not be run, the group thought the most efficient way to evaluate the variables would be using a frequency test. Therefore, the data was organized and from this data, two pivot tables were created. One used the state as a row label, the loss type as a column label and the sum of count as the values. The other table used the state as a row label, the loss type as a column label and the sum of the dollar amount of loss as the values. These tables, graphs and an analysis of the data available can be seen in the Findings section.

4.6.3.8 Square Footage versus Dollar Amount of Loss

The group wanted to evaluate hypothesis 8, that as the roof’s square footage increases in size, the claim amount increases as well. The group wanted to assess this hypothesis to try and draw the correlation between larger roofs and increased dollar amount of loss. The team felt that the best way in which to analyze this data is through the results of a regression test. In order to test hypothesis 8, it is necessary to evaluate the variables for Sq. Footage, Dollar
Amount of Loss and Data Source Code (Hanover or Outside). To normalize the data the Log Transformation of Sq. Footage was calculated by taking the natural log of each square footage. The group next ran the regression using Data Source Code as the control variable, Log Transformation of Sq. Footage as the independent variable, and Dollar Amount of Loss as the dependent variable. Since much of the square footage data was obtained from outside The Hanover, the data source is controlled in order to evaluate if the source of the data (from databases within The Hanover or outside) affects the team’s analysis. For the purpose of this test, the regression table results were used in the group’s analysis because the Log Transformation of Sq. Footage variable is continuous. The results and analysis of this test can be seen in the Findings section.

4.6.3.9 Number of Claims due to Snow and Collapse versus Roof Pitch

The group wanted to evaluate hypothesis 9, which states that if there is little or no slope, the roof will be more likely to be prone to snow or collapse damage. The hypothesis only includes snow and collapse failures because these are the claims that would be affected by whether a roof is flat or pitched. In order to test hypothesis 9, the group needed to evaluate the variables for ‘Pitch’, ‘Loss Type’ and ‘Secondary Loss Type’. The group wanted to run an ANOVA to evaluate hypothesis 9 and compare the number of claims due to snow or collapse on pitched roofs versus flat roofs. However, there was insufficient data and, instead, a frequency test was run. Two pivot tables were then created from this data. The first pivot table created used pitch as a row label and sum of count as values. The second pivot table used pitch again as
a row label, loss type as a column label and sum of count as values. These tables and an analysis of the data available can be seen in the Findings section.

4.6.3.10 Occupancy versus Dollar Amount of Loss

The group wanted to examine the possibility of a relationship between the dollar amount of loss and the occupancy. This was for the group to see if there were any patterns in the occupancy of buildings that were experiencing the most expensive roof loss claims. If the group had additionally had access to a control group of buildings that had not experienced claims, an ANOVA could have then been run to evaluate which occupancies were also the most prevalent. However, in order to test this assumption, the group needed to examine the variables for ‘Dollar Amount of Loss’ and ‘Occupancy’. The group felt that a frequency test could best analyze the data. Two pivot tables were then created to display the data. One of the two pivot tables created used the occupancy for the row labels and used a sum of the count for each occupancy category as the values. The other pivot table again used occupancy for the row labels, but used the average dollar amount of loss for the values. These tables and an analysis of the data available can be seen in the Findings section.
The Hanover Insurance Group – Roof Loss Control Analysis

5.0 Findings

This section discusses the results of the data gathering and also provides the descriptive statistics for the multiple databases that were searched. These results are followed by results from the tests used to test the group’s multiple hypotheses. Lastly, the third section provides information that was gathered during the interview with a loss control employee. The following information is used as to make recommendations to The Hanover.

5.1 Data

The results of the data mining process came from The Hanover’s databases as well as the external databases. From the claim numbers found in the Loss Data Excel file, the group was able to search for the address of each specific claim. By searching the databases within The Hanover, HCS and CSS, the group was able to locate the addresses of 197 claims from the 405 claim numbers of the PDF file. Therefore, the number of data points that come from searching various databases will depend upon the search criteria. If the data came from the PDF, then the group was able to test all 405 claims; however if the search criteria was the address, then the maximum number that could be returned was 197 data points. CSS was ideal for searching the older claims, and returned 48 addresses out of the 197 addresses found. Of the 405 total claims, 208, or 49%, of the claims were left empty because we could not find the address of the building in which the loss occurred. For the 197 address found, the group was able to conduct further searches for information pertaining to the roof’s failure.

The best source of information relating to the roofs materials and structure were the BURs, found in CAAMS, which provided the most useful information for the widest range of
variables within one document. From the 197 addresses in the Loss Data Excel file, CAAMS was able to return 25 BURs. Again, using the 197 addresses from the Loss Data Excel file, ARIES was able to produce information on 34 claims. CSS was able to provide information for 194 claims and HCS was able to provide information on 139 claims. After searching through The Hanover’s databases, it was clear that the group did not have all the data needed to analyze the roof claims in the Loss Data Excel file, and that it was necessary to use external internet sources for further data mining.

Since the addresses were already obtained, the team used appraisal and tax assessor websites to pull some useful data points. These appraisal websites Property Shark, Appraisal Central, and Vision Appraisal provided roof data on a combined 11 different claims. The database specifically for Texas, DCAD, was able to provide 16 successful searches. From Property Assessment, the team located 19 addresses and recorded the data. The final tool used to find information on the remaining claims was Google. Google’s feature, Google Earth, was able to provide information on another 81 addresses.

The combined efforts of The Hanover’s databases and the internet sources enabled the group to identify the following number of observations for each variable:

- The group found 27 data points with or without an Addition. Of these 27 data points, eight were found to have an addition while 19 were not.
- The group identified the age of 69 buildings out of 197 addresses and these ranged in age from 1900 to 2009. Further analysis of this variable is displayed later in this section.
The Hanover Insurance Group – Roof Loss Control Analysis

- 19 data points were identified for the Age of the Roof Cover variable, ranging in ages from being built or modified in 1972 to 2010.
- The group was able to collect 194 data points for the Catastrophe Code variable. Of these 194 data points, 127 had confirmation of a catastrophe code while 67 were not catastrophe related. The following chart displays the breakdown of catastrophe claim data:

Figure 2. Claims with Catastrophe Information

- Whether or not there was an Elevation Difference was confirmed for 153 buildings. Of the 153 buildings there were 55 without a difference in elevation and 98 with.
- 68 data points with an ISO Building Code were found. The results of which are displayed in the graph below.
The Number of Stories was reported for 119 buildings. There were 60 buildings with one story, 37 with two stories, 14 with three stories, six with four stories and one each with six and seven stories respectively.

The Occupancy was specified for 134 buildings. The results of the testing for this variable are displayed and discussed extensively later in this section.

Whether or not there was a pitched roof was identified for 153 buildings. Of these buildings 70 had evidence of a pitched roof while 83 did not.

The Square Footage was found or estimated by Google Earth, as well as many other references for 151 addresses. The results of the testing for this variable are displayed and discussed extensively later in this section.

Finally, there were 47 data points with descriptions of the roof cover material. The frequencies of the different types are displayed in the graph below.
These data points were not included in the analysis because no hypotheses were made regarding this information. The group felt that it would not be useful to analyze these roof characteristics at this time because it would be extremely difficult to draw conclusions using this data. These variables will be discussed further in the Future Research section of the Recommendations. Below is a table showing a breakdown of the total number of data points collected per variable.
Table 3. Summary of Observations per Variable

<table>
<thead>
<tr>
<th>Category of Data</th>
<th>Total Number of Data Points per Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Addition</td>
<td>27 Observations</td>
</tr>
<tr>
<td>Ages of Building</td>
<td>69 Observations</td>
</tr>
<tr>
<td>Age of Roof Cover</td>
<td>19 Observations</td>
</tr>
<tr>
<td>Catastrophe Code</td>
<td>194 Observations</td>
</tr>
<tr>
<td>Claim Numbers</td>
<td>405 Observations</td>
</tr>
<tr>
<td>Date of Loss</td>
<td>405 Observations</td>
</tr>
<tr>
<td>Dollar Amount of Loss</td>
<td>405 Observations</td>
</tr>
<tr>
<td>Elevation Difference</td>
<td>153 Observations</td>
</tr>
<tr>
<td>ISO Building Codes</td>
<td>68 Observations</td>
</tr>
<tr>
<td>Loss Type</td>
<td>405 Observations</td>
</tr>
<tr>
<td>Name of Insured</td>
<td>405 Observations</td>
</tr>
<tr>
<td>Number of Stories</td>
<td>119 Observations</td>
</tr>
<tr>
<td>Occupancy</td>
<td>134 Observations</td>
</tr>
<tr>
<td>Pitch</td>
<td>153 Observations</td>
</tr>
<tr>
<td>Square Footage</td>
<td>151 Observations</td>
</tr>
<tr>
<td>Street Addresses</td>
<td>197 Observations</td>
</tr>
<tr>
<td>Roof Claims taken from PDF</td>
<td>405 Observations</td>
</tr>
<tr>
<td>Roof Cover Material</td>
<td>47 Observations</td>
</tr>
</tbody>
</table>

5.2 Data Validation

It is imperative for our project to not only have data, but to have data that consistent and accurate. The team used a variety of data mining methods when filling in the data sheet. Since there were many sources of data providing information, the group decided it was important to perform a data validation. To perform a data validation required a return to the methodology and repeat the data mining steps for random data points. The team needs to not only confirm the data, but also confirm that through our data mining methods we were not able to locate empty data points. This required the group to return to The Hanover use random claim numbers found in the Loss Data Excel file.
With the claim numbers, the team can confirm the address found. After using this address and claim number, the team was able to move onto validating the other data points of interest using the claim databases at The Hanover. The team picked five random claim numbers without addresses and confirmed that these claim number do not lead to any useful data available in the databases. The next step was to check ten claims with data. The information that was found on these data points was located in The Hanover’s claim files, in online assessor and zoning databases, or in Google Earth. To verify that the data was valid, the group compared the data validation excel file to the Loss Data Excel file that was compiled throughout the project. The corresponding data matched up and it was clear that the group’s data mining had a high level of confidence was reasonably accurate. However, there may be some slight variation in the square footage information. Since an estimation tool was used, there could be some variance in the tracing of the shape of the roof, causing the square footage to be off by a couple hundred feet. For the testing of this project, because the variance is rather small, this does not affect the outcome of the results.

Armed with this information, the team conducted some of the statistical analyses to test the project’s hypotheses. Due to the lack of data in some categories or the lack of a control group, some tests were not able to be performed. Again, this will be discussed further in the Future Research section of the Recommendations. The following section will describe the information obtained from the team’s interview with an employee from The Hanover.
5.3 Interviews

To gain further insight into The Hanover, the team conducted an interview with veteran loss control consultant within the insurance industry (see Appendix C). Unfortunately, the loss control consultant was the only candidate that was able to make time for an interview and the remaining candidates, from underwriting and claims could not be interviewed. However, the information that the interviewee provided for the group proved useful for future recommendations.

5.3.1 Training

New loss control employees are required to fulfill a required number of hours in training, prior to their first property evaluation. They also must complete multiple evaluations with a more experienced loss control consultant before they are allowed to conduct their own on-site evaluation. The reasoning for this, according to the veteran loss control consultant, is that the younger field staff may not have an easy time identifying the difference between a high and low quality roof. Younger inexperienced loss control employees also may not have the same material knowledge as a veteran consultant.

5.3.2 Typical Building Assessment

A typical building assessment consists of the following procedure: (1) an on-site interview, (2) camera shots of the property or roof, (3) and finally, a walkthrough of the building.

During the on-site interview a loss control agent presents the property owner with a preset list of questions regarding the details of the property, as well as a review of any damages
or loss the property has previously incurred. Anything in question during an on-site interview is documented. Occasionally, prior to the interview, the employee already have reviewed images of the building, both loss and square footage, via Google Earth. Following this interview, and safety permitting, it is required to physically walk around a property or atop a roof and take pictures of the damages. The loss control employee does this in order to more accurately assess the extent of damage to the property or roof. However, roof access is not always possible. For example, in the case that the roof is covered in snow, OSHA regulations prohibit roof activity. Finally, to further evaluate the damages or loss, the loss control agent will walk through the inside of the property to gauge the structural integrity of the property as well as any interior damages.

5.3.3 Additional Information

The loss control agent, during the course of the interview, also provided the team with supplemental information that proved useful for our future recommendations. When evaluating a property loss, The Hanover provides discounts or a benefit system to buildings that have actively tested sprinkler systems. To maintain this discounted insurance premium, The Hanover requires insured companies to test their sprinkler systems every 36 months. The group then made the connection between this piece of information and the possibility of implementing the same sort of incentives program for roof systems. This, however, will be discussed in more detail in the future recommendations section. The following section will describe and analyze the findings from the statistical tests.
5.4 Statistics

The following section describes the hypothesis examined, the limitations of the tests conducted, the results of any regression or ANOVA testing, the analysis of the test results, and displays any graphs and tables for each of the different statistical and descriptive tests.

5.4.1 Age of Building per Number of Claims

The group tested hypothesis 1, which states, as buildings increase in age, the likelihood of the building collapsing due to snow loads decreases. The data for the age of the building can be represented by the following graph:

Figure 5. Number of Claims per Building Age Category

The graph shown above depicts the number of claims for each age category. As one can see, the age category with the most amounts of claims is from 1960 to 1979, with 29 claims. The age category with the least amount of claims is from 1920 to 1939, with zero claims. This data might lead one to believe that our hypothesis for this test set is confirmed and that as buildings increase in age, the likelihood of the building collapsing due to snow loads decreases. However,
since there is no control group of buildings that have not experienced a claim to compare this data to, the group cannot draw this conclusion.

From the building age data, the group ran a regression test with the Dollar Amount of Loss as the dependent variable and the Age of Building Code as the independent variable, and the Log Transformation of Square Footage as the control variable. Below is a table showing the relevant data from the regression.

Table 4. Age of Building Code & Log Transformation of Sq. Footage by Dollar Amount of Loss

<table>
<thead>
<tr>
<th>Regression Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adjusted R Square</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Coefficients</td>
</tr>
<tr>
<td>Intercept</td>
</tr>
<tr>
<td>Age of Building Code</td>
</tr>
<tr>
<td>Log Transformation of Sq. Footage</td>
</tr>
</tbody>
</table>

From these results shown above, it was concluded that the data does not fit the model. The Adjusted R Square value is nowhere close to the ideal value of 1 and the p-values infer that the coefficients are not significantly different from 0. The Age of Building Code and the Log Transformation of Square Footage are not predictors of the Dollar Amount of Loss. Therefore, there is no clear evidence of a relationship between the Age of Building Code, Log Transformation of Square Footage and Dollar Amount of Loss.

For this test, the group excluded the data from one claim because there is not a specific building for that claim. The group found that the variable, Age of Building, was a limiting factor.
The Hanover Insurance Group – Roof Loss Control Analysis

within this test because there were only 68 (69 minus the one data entry that was removed) observations out of 405 roof claims.

5.4.2 Number of Claims with Evidence of Elevation Difference and Loss Types of Snow or Collapse

The group analyzed hypothesis 3, which states if the building has an elevation difference, then the roof has an increased likelihood of suffering a snow load or collapse claim.

The data for this analysis can be represented by the following table:

Table 5. Number of Claims with Evidence of Elevation Difference and Loss Types of Snow or Collapse

<table>
<thead>
<tr>
<th>Row Labels</th>
<th>Collapse</th>
<th>Snow</th>
<th>Grand Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Elevation Difference</td>
<td>7</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td>Elevation Difference</td>
<td>4</td>
<td>9</td>
<td>13</td>
</tr>
<tr>
<td>Grand Total</td>
<td>11</td>
<td>15</td>
<td>26</td>
</tr>
</tbody>
</table>

The group was unable to test hypothesis 3 because of the absence of a control group that contains information on buildings that have not experienced a roof loss claim. The data was further limited by the number of data entries with information about evidence of elevation difference on the roof (153 data points) and then again by the number of data entries with a loss type of snow or collapse that also included data in the elevation difference column. These factors reduced our data for this test to 26 total data points. For additional information see Future Research section.

5.4.3 Dollar Amount of Loss Compared to Loss Type

The group tested hypothesis 5, which states that the prevalent causes of roof claims are hurricanes, hail, and wind in order of significance. The group created three loss type categories
for which to analyze the data, these categories were Naturally Occurring Weather Related Claims, Natural Disaster Claims and, finally, All Weather Related Claims. The following sections will describe the data as well as the testing done for each category.

### 5.4.3.1 Naturally Occurring Weather Related Claims

The data for Naturally Occurring Weather Related Claims can be represented by the two following tables below:

**Table 6. Claim Frequency by Naturally Occurring Weather Related Data Minus Outliers**

<table>
<thead>
<tr>
<th>Row Labels</th>
<th>Sum of Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collapse</td>
<td>21</td>
</tr>
<tr>
<td>Hail</td>
<td>59</td>
</tr>
<tr>
<td>Snow</td>
<td>21</td>
</tr>
<tr>
<td>Water</td>
<td>27</td>
</tr>
<tr>
<td>Wind</td>
<td>70</td>
</tr>
<tr>
<td><strong>Grand Total</strong></td>
<td><strong>198</strong></td>
</tr>
</tbody>
</table>

**Table 7. Average Dollar Amount of Loss for Naturally Occurring Weather Related Data Minus Outliers**

<table>
<thead>
<tr>
<th>Row Labels</th>
<th>Average of $ Amount of Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collapse</td>
<td>$ 286,186</td>
</tr>
<tr>
<td>Hail</td>
<td>$ 275,551</td>
</tr>
<tr>
<td>Snow</td>
<td>$ 318,627</td>
</tr>
<tr>
<td>Water</td>
<td>$ 222,076</td>
</tr>
<tr>
<td>Wind</td>
<td>$ 294,458</td>
</tr>
<tr>
<td><strong>Grand Total</strong></td>
<td><strong>$ 280,640</strong></td>
</tr>
</tbody>
</table>

For this category, there were 16 outliers. This test was not limited by any factors, as the data for loss type and dollar amount of loss were all collected from the PDF.
The Hanover Insurance Group – Roof Loss Control Analysis

The group was able to test the hypothesis using the Dollar Amount of Loss as the dependent variable and Naturally Occurring Weather Related Loss as the independent variable. With this test, the group evaluated the results of the ANOVA. Below is a table showing the relevant data from the ANOVA table.

**Table 8. ANOVA Table for Naturally Occurring Weather Related Claims by Dollar Amount of Loss**

<table>
<thead>
<tr>
<th>Naturally Occurring Weather Related Claims by Dollar Amount of Loss</th>
<th>ANOVA Table</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Adjusted R Square</strong></td>
<td>0.01</td>
</tr>
<tr>
<td><strong>Significance F</strong></td>
<td>0.106</td>
</tr>
</tbody>
</table>

From the table above, one can conclude that the data does not fit the model well; however, the model does explain about 1% of the variance. The significance of this ANOVA test implies that the means differ more than what would be expected by chance alone. This means that Naturally Occurring Weather Related Loss Types were not equal in Dollar Amount of Loss, however, this does not tell the group anything about what the loss is, just that there was loss.

**5.4.3.2 Natural Disaster Claims**

The data for Natural Disaster Claims can be represented by the two following tables below:

**Table 9. Natural Disaster Claims Data Minus Outliers**

<table>
<thead>
<tr>
<th>Natural Disaster Claims Data Minus Outliers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Row Labels</strong></td>
</tr>
<tr>
<td>Hurricane</td>
</tr>
<tr>
<td>Tornado</td>
</tr>
<tr>
<td><strong>Grand Total</strong></td>
</tr>
</tbody>
</table>
For this category, there were six outliers. This test was not limited by any factors, as the data for loss type and dollar amount of loss were all collected from the PDF.

The group used the Dollar Amount of Loss as the dependent variable and Natural Disaster Loss Type as the independent variable. With this test, the group evaluated the results of the ANOVA. Below is a table showing the relevant data from the ANOVA table.

Table 11. Natural Disaster Claims by Dollar Amount of Loss

<table>
<thead>
<tr>
<th>Natural Disaster Claims by Dollar Amount of Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANOVA Table</td>
</tr>
<tr>
<td>Regression</td>
</tr>
</tbody>
</table>

From the table above, one can conclude that the data does not fit the model. The insignificance of this ANOVA test implies that the differences between the means are not great enough for the group to say that they are different. No further interpretation can be attempted.

5.4.3.3 All Weather Related Claims

The data for All Weather Related Claims can be represented by the two following tables below:
Table 12. All Weather Related Data Combined Minus Outliers

<table>
<thead>
<tr>
<th>All Weather Related Data Combined Minus Outliers</th>
<th>Sum of Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collapse</td>
<td>23</td>
</tr>
<tr>
<td>Hail</td>
<td>60</td>
</tr>
<tr>
<td>Hurricane</td>
<td>148</td>
</tr>
<tr>
<td>Snow</td>
<td>21</td>
</tr>
<tr>
<td>Tornado</td>
<td>12</td>
</tr>
<tr>
<td>Water</td>
<td>27</td>
</tr>
<tr>
<td>Wind</td>
<td>71</td>
</tr>
<tr>
<td><strong>Grand Total</strong></td>
<td><strong>362</strong></td>
</tr>
</tbody>
</table>

Table 13. All Weather Related Data Combined Minus Outliers

<table>
<thead>
<tr>
<th>All Weather Related Data Combined Minus Outliers</th>
<th>Average of $ Amount of Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collapse</td>
<td>$370,368</td>
</tr>
<tr>
<td>Hail</td>
<td>$292,791</td>
</tr>
<tr>
<td>Hurricane</td>
<td>$346,672</td>
</tr>
<tr>
<td>Snow</td>
<td>$318,627</td>
</tr>
<tr>
<td>Tornado</td>
<td>$447,674</td>
</tr>
<tr>
<td>Water</td>
<td>$223,023</td>
</tr>
<tr>
<td>Wind</td>
<td>$306,463</td>
</tr>
<tr>
<td><strong>Grand Total</strong></td>
<td><strong>$323,860</strong></td>
</tr>
</tbody>
</table>

For this category, there were 23 outliers as a result of the method for determining outliers. This test was not limited by any factors, as the data for loss type and dollar amount of loss were all collected from the PDF.

The group used the Dollar Amount of Loss as the dependent variable and All Weather Related Loss as the independent variable. With this test, the group evaluated the results of the
The Hanover Insurance Group – Roof Loss Control Analysis

ANOVA because the variable, All Weather Related Loss, is not a continuous variable, it is categorical. Below is a table showing the relevant data from the ANOVA table.

**Table 14. All Weather Related Loss Claims by Dollar Amount of Loss**

<table>
<thead>
<tr>
<th>All Weather Related Loss Claims by Dollar Amount of Loss</th>
<th>ANOVA Table</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Adjusted R Square</td>
</tr>
<tr>
<td></td>
<td>Significance F</td>
</tr>
<tr>
<td></td>
<td>Regression</td>
</tr>
</tbody>
</table>

From the table above, one can conclude that the data does not fit the model well. The insignificance of this ANOVA test implies that the differences between the means are not great enough for the group to say that they are different. No further interpretation can be attempted. However, the three graphs shown below display more clearly the relationship between the different loss types.
Figure 6. All Weather Related Data Minus Outliers: Breakdown by Loss

All Weather Related Data (Minus Outliers): Breakdown of Loss Type

- Hurricane: 41%
- Hail: 17%
- Wind: 20%
- Water: 7%
- Tornado: 3%
- Snow: 6%
- Collapse: 6%

- Wind: 20%
- Hail: 17%
- Hurricane: 41%
- Water: 7%
- Tornado: 3%
- Snow: 6%
- Collapse: 6%
Figure 7. All Weather Related Data Minus Outliers: Average Dollar Amount of Loss by Loss Type

<table>
<thead>
<tr>
<th>Loss Type</th>
<th>Average Dollar Amount of Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wind</td>
<td>$306,463</td>
</tr>
<tr>
<td>Water</td>
<td>$223,023</td>
</tr>
<tr>
<td>Tornado</td>
<td>$447,674</td>
</tr>
<tr>
<td>Snow</td>
<td>$318,627</td>
</tr>
<tr>
<td>Hurricane</td>
<td>$346,672</td>
</tr>
<tr>
<td>Hail</td>
<td>$292,791</td>
</tr>
<tr>
<td>Collapse</td>
<td>$370,368</td>
</tr>
</tbody>
</table>
As one can see Figure 6, the claims caused by hurricanes make up 41% of all the loss claims. Hurricanes are the most costly in terms of total dollar amount of loss, but are not the most costly in terms of average dollar amount per loss. That distinction belongs to the tornado loss type with just under $450,000. Still, tornado losses make up only 3% of all the loss claims, the smallest percentage, and also only accounts for just over $5 million in total dollar amount of loss, again, the smallest of all loss types. It is also interesting to notice that wind, having the second highest percentage of all loss claims (20%), has the third lowest average dollar amount of loss.
5.4.3.4 Dollar Amount of Loss Compared to Loss Type: Regressions with Square Footage

As previously mentioned in the methodology, the group thought that adding the square footage as a control to the previous three ANOVA tests (which were for Naturally Occurring Weather Related Claims, Natural Disaster Claims, and All Weather Related Claims), may prove to strengthen the analysis.

The first regression test was of Naturally Occurring Weather Related Losses, which was limited by the number of square footage data points. To run this regression, the group used Naturally Occurring Weather Related Claims as an independent variable, the Log Transformation of the Square Footage as a control variable and the Dollar Amount of Loss as the dependent variable. For the purposes of this test, the group evaluated the regression table because the variable, Log Transformation of Sq. Footage, is continuous. The table below shows the relevant data that resulted from the regression test.

<table>
<thead>
<tr>
<th>All Naturally Occurring Weather Related Data &amp; Log Transformation of Sq. Footage by Dollar Amount of Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression Table</td>
</tr>
<tr>
<td>Adjusted R Square</td>
</tr>
<tr>
<td>Coefficients</td>
</tr>
<tr>
<td>Intercept</td>
</tr>
<tr>
<td>Naturally Occurring Weather Loss</td>
</tr>
<tr>
<td>Log Transformation of Sq. Footage</td>
</tr>
</tbody>
</table>

From the table above, one can conclude that the data is not a good predictor of the model. However, the variable Log Transformation of Square Footage is weakly significant. On the other hand, the variable Naturally Occurring Weather Related Loss is not a predictor of the Dollar
Amount of Loss. Therefore, there is no clear evidence of a relationship between the Dollar Amount of Loss and the Naturally Occurring Weather Related Loss, but there is some evidence of a positive relationship between the Dollar Amount of Loss and the Log Transformation of Square Footage.

The second regression test was of Natural Disaster Loss Claims, which was again limited by the number of square footage data points. To run this regression, the group used Natural Disaster Loss as the independent variable, the Log Transformation of the Square Footage as the control variable, and the Dollar Amount of Loss as the dependent variable. For the purposes of this test, the group evaluated the regression table because the variable, Log Transformation of Sq. Footage, is continuous. The table below shows the relevant data that resulted from the regression test.

**Table 16. Natural Disaster loss & Log Transformation of Sq. Footage by Dollar Amount of Loss**

<table>
<thead>
<tr>
<th>Regression Table</th>
<th>Coefficients</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adjusted R Square</td>
<td>-0.01</td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-665,332</td>
<td>0.41</td>
</tr>
<tr>
<td>Natural Disaster Loss Type</td>
<td>124773</td>
<td>0.33</td>
</tr>
<tr>
<td>Log Transformation of Sq. Footage</td>
<td>36,241</td>
<td>0.40</td>
</tr>
</tbody>
</table>

From the table above, one can conclude that the data does not fit the model. The variables Natural Disaster Loss Type and Log Transformation of Square Footage are not predictors of the Dollar Amount of Loss. Therefore, there is no clear evidence of a relationship between the Natural Disaster Loss Type, Log Transformation of Square Footage and the Dollar Amount of Loss.
The third regression test was of All Weather Related Loss Types, which was, once again, limited by the number of square footage data points. To run this regression, the group used All Weather Related Loss Types as the independent variable, the Log Transformation of the Square Footage as the control variable, and the Dollar Amount of Loss as the dependent variable. For the purposes of this test, the group evaluated the regression table because the variable, Log Transformation of Sq. Footage, is continuous. The table below shows the relevant data that resulted from the regression test.

**Table 17. All Weather Related Loss Types & Log Transformation of Sq. Footage by Dollar Amount of Loss**

<table>
<thead>
<tr>
<th>Regression Table</th>
<th>Coefficients</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adjusted R Square</td>
<td>-0.01</td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>179,299</td>
<td>0.09</td>
</tr>
<tr>
<td>All Weather Loss Types</td>
<td>6,099</td>
<td>0.33</td>
</tr>
<tr>
<td>Log Transformation of Sq. Footage</td>
<td>3,400</td>
<td>0.74</td>
</tr>
</tbody>
</table>

From the table above, one can conclude that the data does not fit the model well. The variables All Weather Related Loss and Log Transformation of Sq. Footage are not predictors of the Dollar Amount of Loss. Therefore, there is no clear evidence of a relationship between All Weather Related Loss, the Log Transformation of the Square Footage and the Dollar Amount of Loss.

**5.4.4 Number of Claims per State by Loss Type**

The group tested hypothesis 7, which states that the location of the building will have an effect on the susceptibility to certain types of damage. Outliers were kept for this test because The Hanover must pay the claim no matter the location. The data for this test can be represented by the tables below:
## Figure 9. Count of Claims by State by Loss Type

<table>
<thead>
<tr>
<th>State</th>
<th>Collapse</th>
<th>Hail</th>
<th>Hurricane</th>
<th>Other</th>
<th>Snow</th>
<th>Tornado</th>
<th>Vandalism</th>
<th>Water</th>
<th>Wind</th>
<th>Sum of Count</th>
<th>Grand Total</th>
<th>Percentages</th>
</tr>
</thead>
<tbody>
<tr>
<td>AL</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>0.5%</td>
</tr>
<tr>
<td>AR</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>1.0%</td>
</tr>
<tr>
<td>AZ</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>0.3%</td>
</tr>
<tr>
<td>CO</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>0.8%</td>
</tr>
<tr>
<td>CT</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0.5%</td>
</tr>
<tr>
<td>DC</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0.5%</td>
</tr>
<tr>
<td>FL</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>1.0%</td>
</tr>
<tr>
<td>GA</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>0.5%</td>
</tr>
<tr>
<td>IL</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>11</td>
<td>2</td>
<td>2.8%</td>
</tr>
<tr>
<td>IN</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>12</td>
<td>8</td>
<td>3.0%</td>
</tr>
<tr>
<td>KS</td>
<td></td>
<td>6</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>8</td>
<td>2</td>
<td>2.0%</td>
</tr>
<tr>
<td>KY</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>0.5%</td>
<td></td>
</tr>
<tr>
<td>LA</td>
<td></td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>10</td>
<td>3</td>
<td>2.5%</td>
</tr>
<tr>
<td>MA</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>13</td>
<td>4</td>
<td>3.3%</td>
</tr>
<tr>
<td>MD</td>
<td></td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>1.0%</td>
</tr>
<tr>
<td>ME</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>0.8%</td>
</tr>
<tr>
<td>MI</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>9</td>
<td>2</td>
<td>2.3%</td>
</tr>
<tr>
<td>MN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0.5%</td>
</tr>
<tr>
<td>MN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>0.5%</td>
<td></td>
</tr>
<tr>
<td>MO</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>1.0%</td>
</tr>
<tr>
<td>MT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>0.3%</td>
</tr>
<tr>
<td>NC</td>
<td></td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>1.3%</td>
</tr>
<tr>
<td>NH</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>1.0%</td>
</tr>
<tr>
<td>NJ</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>0.8%</td>
</tr>
<tr>
<td>NY</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>13</td>
<td>3</td>
<td>3.3%</td>
</tr>
<tr>
<td>OH</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>6</td>
<td>3</td>
<td>1.5%</td>
</tr>
<tr>
<td>OK</td>
<td>3</td>
<td>11</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>16</td>
<td>4</td>
<td>4.0%</td>
</tr>
<tr>
<td>SC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>0.3%</td>
</tr>
<tr>
<td>TN</td>
<td></td>
<td>5</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>7</td>
<td>1</td>
<td>1.8%</td>
</tr>
<tr>
<td>TX</td>
<td>3</td>
<td>12</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>43</td>
<td>16</td>
<td>10.8%</td>
</tr>
<tr>
<td>VA</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>0.8%</td>
</tr>
<tr>
<td>VT</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>0.3%</td>
</tr>
<tr>
<td>WI</td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>0.5%</td>
</tr>
<tr>
<td>N/A</td>
<td>4</td>
<td>15</td>
<td>123</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>6</td>
<td>36</td>
<td>134</td>
<td>399</td>
<td></td>
<td>48.6%</td>
</tr>
</tbody>
</table>

**Grand Total**: 28, 64, 159, 6, 21, 12, 8, 29, 72, 399

**Percentages**: 7.0%, 16.0%, 39.8%, 1.5%, 5.3%, 3.0%, 2.0%, 7.3%, 18.0%, 100.0%
## Figure 10. Sum of Dollar Amount of Loss by State by Loss Type

<table>
<thead>
<tr>
<th>Row Labels</th>
<th>Collapse</th>
<th>Hall</th>
<th>Hurricane</th>
<th>Other</th>
<th>Snow</th>
<th>Tornado</th>
<th>Vandalism</th>
<th>Water</th>
<th>Wind</th>
<th>Grand Total</th>
<th>Averages</th>
<th>Percentages</th>
</tr>
</thead>
<tbody>
<tr>
<td>AL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$1,275,000</td>
<td>$195,465</td>
<td></td>
<td></td>
<td>$1,450,465</td>
<td>$154,232.50</td>
<td>0.7%</td>
</tr>
<tr>
<td>AR</td>
<td></td>
<td>$204,715</td>
<td></td>
<td></td>
<td></td>
<td>$1,001,178</td>
<td></td>
<td>$101,043</td>
<td></td>
<td>$1,752,877</td>
<td>$154,444.25</td>
<td>0.9%</td>
</tr>
<tr>
<td>AZ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$198,510</td>
<td>$158,310.00</td>
<td>0.1%</td>
</tr>
<tr>
<td>CO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$434,844</td>
<td>$235,000</td>
<td>0.3%</td>
</tr>
<tr>
<td>CT</td>
<td></td>
<td>$128,636</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$373,964</td>
<td>$189,532.00</td>
<td>0.2%</td>
</tr>
<tr>
<td>DC</td>
<td></td>
<td>$200,908</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$770,426</td>
<td>$158,214.00</td>
<td>0.2%</td>
</tr>
<tr>
<td>FL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$856,184</td>
<td>$600,800</td>
<td>0.2%</td>
</tr>
<tr>
<td>GA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$1,076,125</td>
<td>$538,097.50</td>
<td>0.5%</td>
</tr>
<tr>
<td>IL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$3,939,221</td>
<td>$816,178</td>
<td>$932,813</td>
<td></td>
<td>$1,272,215</td>
<td>$656,795</td>
<td>3.3%</td>
</tr>
<tr>
<td>IN</td>
<td></td>
<td>$1,006,570</td>
<td>$1,613,294</td>
<td></td>
<td></td>
<td>$330,672</td>
<td>$1,847,287</td>
<td></td>
<td></td>
<td>$1452,802</td>
<td>$1,244,316.88</td>
<td>7.5%</td>
</tr>
<tr>
<td>KS</td>
<td></td>
<td>$7,875,256</td>
<td>$372,487</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$8,558,338</td>
<td>4.2%</td>
</tr>
<tr>
<td>KY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$154,190</td>
<td>$159,892.50</td>
<td>0.2%</td>
</tr>
<tr>
<td>LA</td>
<td></td>
<td>$5,268,048</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$633,474</td>
<td>$5,901,522</td>
<td>3.0%</td>
</tr>
<tr>
<td>MA</td>
<td></td>
<td>$5,189,091</td>
<td>$217,255</td>
<td>$492,896</td>
<td></td>
<td>$925,000</td>
<td></td>
<td>$465,661</td>
<td></td>
<td>$1,644,470</td>
<td>$7,945,083</td>
<td>3.7%</td>
</tr>
<tr>
<td>MD</td>
<td></td>
<td>$235,376</td>
<td></td>
<td></td>
<td></td>
<td>$171,323</td>
<td></td>
<td>$494,972</td>
<td></td>
<td>$901,671</td>
<td>$225,417.75</td>
<td>0.5%</td>
</tr>
<tr>
<td>MI</td>
<td></td>
<td>$272,242</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$519,731</td>
<td></td>
<td>$794,973</td>
<td>$263,992.00</td>
<td>0.4%</td>
</tr>
<tr>
<td>MI</td>
<td></td>
<td>$185,158</td>
<td>$290,000</td>
<td></td>
<td></td>
<td>$176,612</td>
<td>$481,052</td>
<td>$1,246,006</td>
<td></td>
<td>$2,378,226</td>
<td>$1,244,257.50</td>
<td>1.2%</td>
</tr>
<tr>
<td>MN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$182,319</td>
<td></td>
<td>$376,985</td>
<td></td>
<td>$559,005</td>
<td>$279,502.50</td>
<td>0.3%</td>
</tr>
<tr>
<td>MN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$2,198,542</td>
<td>$1,14,386</td>
<td></td>
<td>$3,119,928</td>
<td>$1,158,464.00</td>
</tr>
<tr>
<td>MO</td>
<td></td>
<td>$552,711</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$202,492</td>
<td></td>
<td>$232,839</td>
<td>$232,495.75</td>
<td>0.5%</td>
</tr>
<tr>
<td>MT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$963,335</td>
<td></td>
<td>$693,553</td>
<td>$699,553.00</td>
<td>0.3%</td>
</tr>
<tr>
<td>NC</td>
<td></td>
<td>$750,000</td>
<td>$443,707</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$417,678</td>
<td></td>
<td></td>
<td>$1,272,270</td>
<td>0.8%</td>
</tr>
<tr>
<td>NH</td>
<td></td>
<td>$425,492</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$104,192</td>
<td></td>
<td>$104,896</td>
<td>$158,520.00</td>
<td>0.3%</td>
</tr>
<tr>
<td>NJ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$174,509</td>
<td></td>
<td>$254,747</td>
<td></td>
<td>$152,200</td>
<td>$891,456</td>
<td>0.3%</td>
</tr>
<tr>
<td>NY</td>
<td></td>
<td>$2,442,617</td>
<td>$350,000</td>
<td>$318,113</td>
<td></td>
<td>$1,284,121</td>
<td></td>
<td>$1,004,499</td>
<td></td>
<td>$6,199,390</td>
<td>$415,334.62</td>
<td>2.7%</td>
</tr>
<tr>
<td>OH</td>
<td></td>
<td>$361,385</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$125,798</td>
<td></td>
<td>$499,899</td>
<td>$991,282</td>
<td>0.5%</td>
</tr>
<tr>
<td>OK</td>
<td></td>
<td>$606,615</td>
<td>$5,879,334</td>
<td></td>
<td></td>
<td>$202,790</td>
<td></td>
<td>$104,603</td>
<td></td>
<td>$6,792,742</td>
<td>$240,546.38</td>
<td>1.4%</td>
</tr>
<tr>
<td>SC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$111,717</td>
<td></td>
<td>$111,717</td>
<td></td>
<td></td>
<td>$111,717.00</td>
<td>0.1%</td>
</tr>
<tr>
<td>TN</td>
<td></td>
<td>$1,247,269</td>
<td>$184,130</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$271,780</td>
<td></td>
<td>$1,073,179</td>
<td>$243,813.29</td>
<td>0.9%</td>
</tr>
<tr>
<td>TX</td>
<td></td>
<td>$616,830</td>
<td>$1,198,836</td>
<td>$7,004,518</td>
<td></td>
<td>$582,219</td>
<td>$2,016,970</td>
<td>$646,467</td>
<td></td>
<td>$14,065,640</td>
<td>$327,112.56</td>
<td>7.0%</td>
</tr>
<tr>
<td>VA</td>
<td></td>
<td>$254,915</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$134,068</td>
<td></td>
<td>$388,883</td>
<td>$129,661.00</td>
<td>0.2%</td>
</tr>
<tr>
<td>VT</td>
<td></td>
<td>$275,794</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$275,794</td>
<td></td>
<td></td>
<td>$275,794.00</td>
<td>0.1%</td>
</tr>
<tr>
<td>WI</td>
<td></td>
<td>$900,000</td>
<td>$600,140</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$1,300,140</td>
<td></td>
<td>$750,070</td>
<td></td>
<td>0.8%</td>
</tr>
<tr>
<td>N/A</td>
<td></td>
<td>$2,590,714</td>
<td>$5,241,965</td>
<td>$65,526,845</td>
<td>$356,256</td>
<td>$2,654,415</td>
<td>$4,425,524</td>
<td></td>
<td></td>
<td>$2,053,882</td>
<td>$2,515,228</td>
<td>3.3%</td>
</tr>
</tbody>
</table>
The Hanover Insurance Group – Roof Loss Control Analysis

The group was limited by the number of addresses that were available. The data points that are lacking an address are also lacking a state in which the roof loss occurred and are represented in this test by the notation, N/A. This test was also limited by the lack of a control group, this control group would give the number of buildings that The Hanover insured for each state.

It is interesting to note that of all the roof loss types (not including states represented by N/A), hurricane has, by far, the highest percentage of losses (39.8%) and is the most costly, with a loss grand total of over $85 million. This is a little over 42% of the entire grand total loss amount. The next highest percent of roof loss type is wind, comprising of 18% of all claims. As one would expect, wind is also the second most costly roof loss type at just slightly under $32 million. Hail and collapse, are the third and fourth highest in terms of grand total dollar amount of loss ($29.5 and $28.4 million respectively). However, hail makes up 16% of all roof loss claims, whereas collapse makes up only 7%.

As for roof losses by state, not including those which cannot be located (represented by N/A), Texas has the highest percentage of all roof loss claims at 10.8%. The next highest is Oklahoma, with only 4% of all roof loss claims. With Texas having the highest percentage of all roof loss claims, it has a low dollar amount of loss average with only $327,000. This is only the 14th highest, not including roof claims unable to be located, that are represented by N/A. Even though Indiana held only 3% of all the roof loss claims, it was the state with the highest dollar amount of loss average, with just under $1.25 million.
5.4.5 Square Footage versus Dollar Amount of Loss

The group tested hypothesis 8, which states that as the roof’s square footage increases in size, the claim amount increases as well. The data can be represented by the following graph shown below, depicting the number of claims in each square footage category with the outliers removed.

Figure 11. Number of Claims per Sq. Footage Category with Outliers Removed

Looking at this graph, one might be inclined to think that buildings that are smaller are more likely to have a roof loss claim. However, this data may be misleading. For example, The Hanover may insure five times as many smaller buildings versus larger buildings. Therefore, it would appear as though smaller buildings are more likely to have a roof loss claims even though that is not the case. The presence of a control group that includes buildings that have not experienced a roof loss claim for this test would be necessary to test the relationship between the area of the roof and the frequency of the claims.
The Hanover Insurance Group – Roof Loss Control Analysis

There were three outliers. For this test, the only limitations were that for each square footage range, the data was not evenly distributed.

The group was able to run a regression with the Dollar Amount of Loss as the dependent variable and the Data Source Code and Log Transformation of Square Footage as independent variables. The group evaluated the results of the regression because the variable, Log Transformation of Square Footage, is continuous. Below is a table showing the relevant data from the regression.

**Table 18. Data Source Code & Log Transformation of Sq. Footage by Dollar Amount of Loss**

<table>
<thead>
<tr>
<th>Data Source Code &amp; Log Transformation of Sq. Footage by Dollar Amount of Loss</th>
<th>Regression Table</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Adjusted R Square</td>
</tr>
<tr>
<td></td>
<td>Coefficients</td>
</tr>
<tr>
<td>Intercept</td>
<td>-185,598.62</td>
</tr>
<tr>
<td>Data Source Code</td>
<td>182,953.11</td>
</tr>
<tr>
<td>Log Transformation of Sq. Footage</td>
<td>41,956.61</td>
</tr>
</tbody>
</table>

From the table above, one can conclude that the data does not fit the model well; however, the model explains about 2% of the variance. The Data Source Code is a predictor of the Dollar Amount of Loss, but the Log Transformation of Square Footage is not a predictor of the Dollar Amount of Loss. Therefore, there is no clear evidence of a relationship between the Log Transformation of Square Footage and the Dollar Amount of Loss.

We can conclude that it does matter where the data is extracted from. The information that was gathered from outside resources was, on average, 60% more costly in Dollar Amount of Loss than information that was gathered from within The Hanover’s databases. This does not mean that the data collected from The Hanover or from outside sources is necessarily wrong,
just that there is a discrepancy between the two. It is inherently unreliable to use multiple databases since there can be bias introduced into the analysis and even large, and often, unrecognized errors in the data collected due to the increased likelihood of a user error. In conclusion, in order to reduce this discrepancy, one should minimize the use of outside sources.

5.4.6 Number of Claims due to Snow and Collapse versus Roof Pitch

The group analyzed hypothesis 9, which states that if the roof has no pitch, then the roof is more susceptible to snow or collapse claims. The data can be represented by the following table:

Table 19. Number of Claims due to Snow and Collapse versus Roof Pitch

<table>
<thead>
<tr>
<th>Row Labels</th>
<th>Collapse</th>
<th>Snow</th>
<th>Grand Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not Pitched</td>
<td>7</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td>Pitched</td>
<td>5</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td>Grand Total</td>
<td>12</td>
<td>14</td>
<td>26</td>
</tr>
</tbody>
</table>

The group was unable to run a regression due to a lack of data entries and an absence of a control group. This is because the data was limited by the number of data entries with information about the pitch of the roof (144 data points) and then again by the number of data entries with a loss type of snow or collapse that also included data in the pitch column. These factors reduced our data for this test to 26 total data points. For additional information see the Future Research in section 6.2.

5.4.7 Occupancy versus Dollar Amount of Loss

For this analysis, there was no hypothesis previously established. The group thought that it would be interesting to compare the different types of buildings in which roof claims
were filed, against the Dollar Amount of Loss. The data was limited by the number of roof claims which had a data entry for the Occupancy. Since there were a wide range of occupancies, the data was not normally distributed, which is expected since the variable is categorical.

After viewing the graphs below, the group was able to determine that the occupancies that contained the higher percentages of the data were churches, with 24%; warehouses, with 17%; schools, with 11%; and manufacturing, with 9%.

**Figure 12. Percentage of Occupancy Categories for All Roof Loss Claims Data**

![Pie chart showing percentage of occupancy categories for all roof loss claims data.](chart.png)
Figure 13. Average Dollar Amount of Loss per Occupancy Category

Average Dollar Amount of Loss per Occupancy Category

<table>
<thead>
<tr>
<th>Loss Type</th>
<th>Average Dollar Amount of Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>WELD SHOP</td>
<td>$154,190</td>
</tr>
<tr>
<td>WATER PARK</td>
<td>$2,196,542</td>
</tr>
<tr>
<td>WAREHOUSE</td>
<td>$448,458</td>
</tr>
<tr>
<td>THEATER</td>
<td>$450,683</td>
</tr>
<tr>
<td>SUPERMARKET</td>
<td>$2,979,217</td>
</tr>
<tr>
<td>STORAGE UNITS</td>
<td>$301,129</td>
</tr>
<tr>
<td>SCHOOL</td>
<td>$1,400,879</td>
</tr>
<tr>
<td>RETAIL</td>
<td>$838,068</td>
</tr>
<tr>
<td>RECREATIONAL FACILITY</td>
<td>$212,649</td>
</tr>
<tr>
<td>PUB</td>
<td>$494,972</td>
</tr>
<tr>
<td>OFFICE</td>
<td>$401,222</td>
</tr>
<tr>
<td>NURSING HOME</td>
<td>$187,825</td>
</tr>
<tr>
<td>METAL FAB &amp; WAREHOUSE</td>
<td>$1,317,963</td>
</tr>
<tr>
<td>MANUFACTURING</td>
<td>$299,110</td>
</tr>
<tr>
<td>LODGING</td>
<td>$330,508</td>
</tr>
<tr>
<td>GYM/FITNESS</td>
<td>$318,514</td>
</tr>
<tr>
<td>FOOD CHAIN</td>
<td>$180,000</td>
</tr>
<tr>
<td>EVENT VENUE</td>
<td>$391,393</td>
</tr>
<tr>
<td>DINER</td>
<td>$312,466</td>
</tr>
<tr>
<td>DELI</td>
<td>$265,000</td>
</tr>
<tr>
<td>CONDOS</td>
<td>$334,274</td>
</tr>
<tr>
<td>CHURCH</td>
<td>$269,115</td>
</tr>
<tr>
<td>AUTO</td>
<td>$220,000</td>
</tr>
<tr>
<td>APARTMENTS</td>
<td>$751,515</td>
</tr>
</tbody>
</table>

Average Dollar Amount of Loss
The group could also determine that the occupancies that had the highest average dollar amount of loss per category were supermarkets, water parks, schools, metal fab and warehouse, and retail. Whereas the occupancies with the highest total dollar amount of loss were school, warehouse, church, retail and supermarket.

The group found it interesting to note that churches comprised of 24% of the occupancy data but consisted of only $269,115 in average dollar amount of loss. This was the seventh lowest of all the occupancy categories and only just over $100,000 more than the lowest occupancy, which was weld shops at $154,190. It was also interesting to notice that there were only one instance of each, metal fab and warehouse roof loss, and water park roof loss.
The Hanover Insurance Group – Roof Loss Control Analysis

However, these losses were the second and fourth highest respectively, in terms of average dollar amount of loss per occupancy category. Also, buildings with occupancies of supermarket comprised of only 2% of the data and had only two instances of roof loss claims, but had the highest average dollar amount of loss, with just under $3 million.

5.5 Summary of Hypotheses Based on Findings

The group was able to discuss six of the nine hypotheses that were formed at the beginning of the project. The missing hypotheses, 2, 4, 6, will be discussed further in the future research section. The first hypothesis, as buildings increase in age, the likelihood of the building collapsing due to snow loads decreases, was unable to be fully tested because there was no control group to compare to the collected data. The third hypothesis, number of claims with elevation difference, was unable to be fully tested because there was not enough data to statistically test this hypothesis. The fifth hypothesis, dollar amount of loss compared to loss type, was tested with regressions. No independent variables, loss types, were deemed to be predictors of dollar amount of loss. The seventh hypothesis, location of loss has an effect on loss type, was not able to be fully tested because of the lack of a control group. However, it was determined that the most costly loss types were hurricanes and wind, combining for a claim total of approximately 117 million dollars. The eighth hypothesis, as roofs increase in size, so do the claim amounts, was inconclusive because of the lack of a control group. The ninth and final hypothesis, claims with or without pitch, was unable to be fully tested because of the lack of data and a control group. The next section will make some recommendations and suggestions for future based upon these hypotheses.
6.0 Recommendations and Future Research

Based upon the literature review, the results of the data analyses performed, and the interview conducted, the group would like to provide The Hanover with a few recommendations that could help minimize losses associated with roof loss claims.

6.1 Recommendations

1. Identification of Control Groups

The first recommendation that should be taken into consideration is to compile data for control groups against which different statistical tests can be run. With control groups, The Hanover will be able to compare the characteristics of the insured properties that incurred claims to those of the properties that did not incur claims. Absent such control groups, we can only describe trends observed in the data but cannot draw any inferences about causal relationships. There were many failure types and categories that would have benefitted from such a control group during the group's statistical testing. The control groups are required for comparison of the age of roofs, testing the effect of building additions, and of regular maintenance on roof failure.

2. Consolidation of Databases and Collection of Relevant Data

This recommendation stemmed from the actual process of retrieving the necessary data to conduct the study. During the data retrieval process the group became aware that the data was scattered across many different platforms and databases. This made the retrieval process long and arduous and increased the chance of human error. Through speaking with a loss control employee the team was made aware that loss control employees sometimes investigate
claims by using these same databases. Despite the collection of the data by loss control employees, there did not appear to be a single database where all this data was consolidated. It seemed as though the data was collected and then once the claim was resolved, it was discarded or stored in various databases. The group’s recommendation for The Hanover would be to consolidate the data, which is currently stored across multiple databases, by claim number. For example, CSS and HCS provided the same information to The Hanover employees yet HCS is the newer of the two and did not contain any of the older claims information.

In addition to this consolidation of databases, the group recommends that The Hanover begin to collect data types previously not considered. These data types include the presence of an elevation difference in the parts of a roof and whether an addition has been added to the building. Furthermore, if an addition has been added it should be noted whether or not the new addition was inspected. These new data types were collected by the group as they were recognized to be risk factors in potential roof loss claims by extant literature. The presence of an elevation difference in a roof allows snow to accumulate and water to pool in the crevice between the two heights. If allowed to remain, water and snow can cause severe roof damage and if continued to be left unchecked, can lead to collapse. The presence of an addition that has not been inspected can cause weak supports and joints to incur damage, form leaks and potentially cause a full or partial collapse if it was not discovered and remedied during an inspection.

This would allow for research similar to that conducted by this group to be conducted more easily. This would also provide accountability and increase the accuracy in the reporting of this data because it has been reported by The Hanover itself. The use of external databases, like
those required by this group to gather necessary data, would be reduced by the combination of this database and the newly collected data types. These measures would allow analyses to be taken much further.

3. Maintenance Benefit System

The third recommendation comes from the literature review, including articles written by experts in the field, as well as an interview conducted with a loss control employee. During the preparation stages of the project, an extensive amount of literature was reviewed and one recurring theme was that regular maintenance is the best way to prevent a roof claim. This theme was encountered again, during the interview, when the loss control consultant mentioned an incentives system for those insured by The Hanover regarding their sprinkler systems. If the sprinkler system can be proven to be regularly maintained, the policy holder then receives a discount on the policy. This led the group to the idea that a similar incentive program, for roof maintenance, could be implemented to help prevent roof loss claims. The Hanover would determine the maximum amount of time between maintenance to determine if the insured qualifies for the program. This could be further validated if the insured had their property inspected by a qualified source that assess whether the roof is in satisfactory condition or repairs are necessary. This program might offer incentives to the policy holder, such as discounts or compensation for the inspection. Both of these incentives might provide a mutually beneficial outcome, minimizing the risk of a roof loss claim.

The Hanover might want to further research the subject of the incentive based roof maintenance plan. If Hanover could collect data on whether inspection was conducted and compared claims to a control group of buildings not inspected to compare failure rates
between the two groups, then the Hanover would have preliminary support for the effectiveness of roof inspection as a prevention mechanism. The group would recommend to The Hanover that a study be conducted on whether a discount should be offered if regular maintenance occurred or whether a different course of action would be more appropriate. A survey could be carried out asking the policy holders of The Hanover whether or not they would be interested in such a program and what incentives interested them the most, i.e. discounts or free inspections, etc. This would help The Hanover gauge interest in the prospective program and could even be conducted at minimal cost by utilizing email, postal service, phone calls and posting a survey on the website itself so that visitors might take the survey upon browsing the site.

4. Create a Database for Building Contractors

The fourth recommendation builds off the research found in the literature review. One thing that would be useful in evaluating buildings to be insured by The Hanover is to begin recording and collecting information about the contractor responsible for the construction and maintenance of each building. Through research of literature and speaking with officials from The Hanover, it became clear that each building is thoroughly researched before being given a policy. This record keeping would allow The Hanover to quickly and easily search through its databases to see if a contractor has had multiple claims. If The Hanover finds this to be the case, then more risk could be associated with buildings that were built and maintained by the same contractor.
5. Draft an Interior and Exterior Roof Checklist

The team’s next recommendation is for The Hanover to develop a standard checklist of things to look for when inspecting a new property. This recommendation ties in to the inspection piece of the third recommendation and would also be used as an objective way to inspect buildings. The Hanover could develop a list of things that are typically bad traits and things that have been shown to be good traits for inspectors to look for. The presence or absence of these items could be used by the loss control employee to decide whether or not to grant a policy and how risky the building is. The items on this checklist could be developed through further study and statistical tests of the hypotheses we presented as well as consultation with civil engineers about structural deficiencies that could increase risk of roof failure.

While these recommendations are steeped in research and can be supported by the literature and data, they are only recommendations and ideas that, in a perfect world, the group would have ideally been able to carry out and begin to set in motion. With the understanding that these recommendations cannot be implemented without further research and testing, the group wishes to also provide The Hanover with suggestions for future research that might make these recommendations more feasible and the results of the statistical studies more conclusive.

6.2 Future Research

The goal of this section is to provide The Hanover with suggestions for future research necessary to support the group’s previous recommendations and conclusions. The
The Hanover Insurance Group – Roof Loss Control Analysis

aforementioned recommendations will be expanded further and methods for testing or implementing them will be discussed.

6.2.1 Presence of Control Groups and Scarcity of Data

An area of future research the group would like to identify concerns the statistical analyses conducted by the group. As previously mentioned in the recommendations, the presence of control groups would have allowed further comparison of the different failure types and increased the ability of the group to test hypotheses. If these control groups containing all of the buildings insured by The Hanover were present, the group could have had a better idea of how the different failure types compared to the entire population. A few areas where control groups would have been particularly beneficial include data pertaining to roof material, age of roof cover, and the presence of an addition. However, it is important to note that these control groups would be useless if there is a lack of claims data in these categories.

This explains why the group recommends gathering not only the control groups, but beginning to record more of the data on these categories. The group recommends that future research efforts be directed towards comparing this data to see if there are any valid predictors of roof loss claims within the different categories. As it stands, the group was only able to identify many potential predictors of roof loss claim but could only test a limited number of variables.

With this project, the group only analyzed large property, commercial loss claims above $100,000 dollars. The group feels that this study could not only be conducted on this group of commercial claims, but also on all insured buildings.
6.2.2 Hypothesis 2: Age of Roof Cover

The second hypothesis, as roof covers increase in age, the likelihood of a claim increases, was unable to be tested because of the lack of data available pertaining to the variable, Age of Roof Cover. Additionally, the collection of a control group, a group of buildings without roof loss claims and within the same age demographic as the data gathered, would be necessary to test this hypothesis. The elapsed time since the roof was installed plays a role in the number of weather related events that the roof has endured, thus affecting the quality of the roof. For this reason, the team feels that The Hanover could conduct more research into this area. If this research occurs and the results are conclusive, then it would be beneficial for The Hanover to more regularly check the age of the roof cover before they insure the building.

6.2.3 Hypothesis 4: Claims with Additions

Hypothesis 4 is present in this section because it is a hypothesis that was provided by The Hanover to be tested; however a lack of data on whether an addition had been made to a building, as well as a control group to compare it to, limited the group’s effort. When an addition is made to a building, the building essentially needs to be re-evaluated as an entirely new structure. This is because the new addition changes the entire shape of the roof and can lead to increased loads along the seam between the new addition and the old structure. If this is not carefully examined, it could possibly lead to an increased chance of roof failure or collapse. It is for this reason the group feels that The Hanover could conduct future research on whether buildings insured by them include additions and whether they have been re-inspected. If The Hanover could gather more claims data on the presence of these additions and a control
group of buildings with additions but no claims, then they would be able to see whether the new additions do pose an increased risk of having a roof loss claim.

6.2.4 Hypothesis 6: Maintenance and Inspection

The sixth hypothesis, which states that roofs that are subjected to regular maintenance and inspection will be less likely to have claims filed against them could not be tested because of the lack of data available about the maintenance or inspection of roofs. If regular maintenance is made on the roof, then the roof will not be as vulnerable to roof loss claims. For this reason, the team feels that The Hanover could conduct more research into this area. If this research occurs and the results are conclusive, then it would be beneficial for The Hanover to record whether or not the building under examination has undergone maintenance.

If the collection of aforementioned data proves possible by The Hanover, the group feels it has the potential to be very beneficial. If time or financial restraints do not allow this data collection to be executed, then other avenues such as future WPI MQPs might just provide The Hanover an opportunity to conduct the research at a low cost. It is the hope and desire of this group that the research and subsequent conclusions and recommendations provided have proved beneficial to The Hanover or will be at some point in the future.
7.0 Conclusion

The initial goal of this group was to provide The Hanover with a complete analysis of the data collected and to examine which variables are predictors of roof loss claims. The extensive review of literature as well as consultation with officials from The Hanover provided the group with the list of data types to collect and analyze to satisfy this goal. This goal was modified however as the team realized that the complete analysis of this data would not be possible due to small sample sizes and lack of control groups. This resulted in the group identifying procedures and future data collection practices that could help identify these predictors in addition to analyzing data.

Based upon the information gathered, recommendations have been prepared that highlight and address the main findings from the research and statistical analyses conducted by the group. The team recognizes that the total cost of roof claims is high and recommends that the Hanover does take further steps in researching how to minimize this cost using the guidelines set forth by the group’s recommendations. Specifically, the Hanover could gather data related to specific control groups and variables that are outlined in the recommendations. For instance, this would provide valuable insight into which roof characteristics, loss types, or geographic regions are associated with the most roof claims. Armed with this information, The Hanover could make business decisions about which geographic areas should be targeted or avoided. The recommendations we have provided, as well as those for future research, will help the Hanover pursue these investigations further and provide The Hanover with ideas on how roof claims are could be minimized.
References


Estenssoro, Luis F. "Two roof Failures Due to Water Ponding and Related Code Requirements." ASCE. http://ascelibrary.org/cfo/resource/1/jpcfev/v3/i3/p184_s1


Roofing Contractors Association: http://www.nrca.net/consumer/roofsystems.aspx


Appendix A: Background

Insurance Terms (Merriam Webster 2011)

- Insurance is defined as the equitable transfer of the risk of a loss, from one entity to another, in exchange for payment.
- An insurer is a company selling the insurance; an insured, or policyholder, is the person or entity buying the insurance policy.
- The insurance rate is a factor used to determine the amount to be charged for a certain amount of insurance coverage, called the premium.
- A claim is a formal request to an insurance company asking for a payment based on the terms of the insurance policy. Insurance claims are reviewed by the company for their validity and then paid out to the insured or requesting party once approved.

Roofing Terms (Common Roof Types 2011)

- Since, this project will focus on evaluating roof loss claims, it is important to have a full understanding of the different shapes of roofs that are most common today. With the knowledge of common roof shapes established, the more specific materials and types associated with these roofs that will be examined in this project can be described in further detail.
- Gable roofs are very popular, easy to build, proficiently shed water, provide for ventilation and are applicable to a variety of different building designs. These roofs are commonly found on Capes and are shaped like an upside down V at an obtuse angle.
- A hip roof is slightly more difficult to build than a gable roof, but still are a popular choice. Ventilation is not that great compared with gable roofs and the chance for leakage to occur is increased due to hips and valleys in the design.
- An A-Frame roof provides not only a roof, but walls as well. Originally, A-frames were designed for cottages, but they have also been applied to churches, homes and other structures. They shed water exceptionally well and allow for decent ventilation. Maintenance can be difficult due to the steep slope. If damage occurs, it is not only the roof being compromised, but the walls being damaged as well. Usually, asphalt shingles are used...
in construction, which have a warranty of 5-10 years but realistically have a lifespan usually closer to 20 years.

- Flat roofs are the most economical roof to build out of the previous three mentioned. A flat roof requires a built up roof covering rather than conventional shingles. Built-up roofs consist of layers of roofing, felt and tar topped with gravel. Most so-called flat roofs are pitched in some degree, sometimes 1/8 to 1/2 inch per foot, to aid in drainage. Flat roofs are popular in warmer areas where the overhangs can provide shade and there is little or no snowfall. Generally, flat roofs are very poor at shedding water, and most often ponding, leakage, and snow will cause major problems for such roofs.
Appendix B: Literature Review

ISO Codes

The Insurance Services Office has developed a standardized class system for construction methods that help insurance companies to evaluate risk when assessing a property. The classes are meant to categorize the materials used in the entire structure of a building by the risk of a loss due to fire. However, for the purposes of this project, none of the roof loss data claims are fire related. There are numerous classes and this section will begin by describing the first six and most common of the classes and then progress through the lesser common classes.

Class 1 is known as the frame class. This consists of wooden buildings in which wood or other combustible materials have been used to construct the walls, floors and roof. The combination of noncombustible or slow burning exterior walls and combustible floors and roof is also considered “Frame” construction. This method is widely used in residential buildings as well as small commercial buildings.

Class 2 is joisted masonry. This class is similar to the frame class because the floors and the roof are usually constructed using wood or other combustible material. This material differs from Class 1 because it has a fire resistance rating of not less than one hour. The walls in a building built with joisted masonry are free standing and independent of the floors and roof. The walls act as load bearing walls and therefore, joisted masonry buildings usually, are not more than three stories high. This construction method is commonly found in older businesses, stores and more recently in some newer convenience stores. Brick or brick joisted can also be considered joisted masonry. Overall, the joisted masonry class usually provides a building with more structural support in the case of a fire.

Class 3 is noncombustible. In this construction method, the walls, floor, roof and materials supporting it are all constructed using noncombustible or slow burning materials. However, a fire resistance rating is not required for any part of this construction. A common example of a building constructed in the noncombustible method is an all steel building such as a warehouse with the insulation on the outer deck being the only combustible material.
allowed. With intense heat and highly combustible contents inside the building, the structure is susceptible to complete failure.

Class 4 is known as masonry noncombustible. This class covers a construction method that entails many similar characteristics to the noncombustible class. The difference between the two classes is that the walls are noncombustible and, have a fire resistance rating of not less than one hour, or, are made of masonry material that is not less than four inches thick. Interior materials must also be non-combustible.

Class 5 is modified fire resistive. This consists of masonry or noncombustible materials that must be used in the construction of the interior and exterior bearing walls or structural supports, floors and roof. In modified fire resistive structures the materials have a fire resistance rating of between one and two hours.

Class 6 is entitled fire resistive. This class is similar to the modified fire resistive class, except that the masonry materials required need to be thicker and the fire resistance rating needs to be not less than two hours. If hollow masonry materials are used in construction, they cannot be less than twelve inches thick and, if solid, the requirement is reduced to no less than four inches.

These six classes are the most common classifications and are the most widely used. While these are considered the general classifications, others do exist. These other classifications build off of the previous ones and are not fire related classifications, but weather related, and will be discussed in the following paragraphs. Also, in the case of mixed construction types, where a mix of materials is used, the final construction type to be designated must make up at least two thirds of the total bearing wall area and at least two thirds of the floor and roof areas. In the case that several types make up these areas, the one with the least fire resistive rating shall be designated.

The first of the alternative classifications is titled Class 7. This type of construction typically consists of heavy timber and joisted masonry construction that meets the requirements of the joisted masonry classification. The wood beams and girders supporting the roof cannot be less than six inches thick and have a deck that meets the same thickness requirements, depending on the material used.
The Hanover Insurance Group – Roof Loss Control Analysis

The second alternative, Class 8, is essentially the same as the noncombustible class, where the roof deck is a masonry material at least two inches thick and is on protected or unprotected supports. This method requires the use of twenty two gauge or heavier metal and/or has a documented wind uplift classification of ninety or equivalent.

The final class is Class 9. This class is almost the same as the masonry noncombustible class. However, the key differences are that the roof deck needs to be a masonry material of at least two inches thick and constructed of twenty two gauge or heavier metal and/or has a wind uplift classification of ninety or equivalent.

These ISO Codes are essential to the insurance industry and to builders and contractors, as they need to be cognizant of the type of construction they are performing. Builders and contractors must also be fully aware of the conditions in which they are working and the safety precautions that they are taking. One way in which safety is maintained and regulated is by the organization known as the Occupational Safety and Health Administration.
Appendix C: Methodology

Step by Step Instructions for Hanover Databases

HCS
- Log onto the Hanover system
- Open the Intranet and the homepage for Hanover employees is displayed.
- Select the department tab and when redirected to that page, and choose the claims option.
- After the new screen appears, click the HCS portal and a prompt to enter credentials appears. Enter user information.
- The only searchable criteria given in the PDF for this database is claim number. In the box for claim number, the corresponding claim number being searched is entered and the search button is clicked.
- The claim will then appear in a new window with a multitude of information and different side tabs. For our purposes, the initial screen that you are brought to is the correct one and the address of the claim can be retrieved, which will be useful when further searching for claim information.
- The catastrophe code will be found in the Loss Detail tab found on the left side toolbar. After clicking on this toolbar, the loss causes and descriptions can be found. The catastrophe code will be described here if it is assigned to the claim.
- After the information was retrieved and a new claim needed to be searched, the previous window that yielded the search results was returned to and the ‘new search’ link was selected. The process was repeated as many times as necessary.

CSS
- To access the CSS database, one must first log on to a computer that supports the Hanover System by entering the correct credentials into the Windows log-in.
- Open an Internet Explorer window, which directed the user to the homepage for The Hanover employees. On The Hanover’s homepage, the Departments tab is then selected, which brought up a dropdown menu. We selected the Claims link and then the Claims
Corner link in the upper right hand side of the screen. This link will prompt the user for a login name and password.

- Click on the CSS portal. A screen will then sometimes appear with a message asking if the user would like to close the previous window. Selecting yes will bring the user to the CSS homepage.
- Click search followed by ‘1-General Claim’ to access the screen which will allow the user to search the claims database.
- In the space designated for the claim number, enter the claim number from the Excel document in the format ‘##-#######’ (Note: If two zeros appear after the dash on the Excel document, omit the zeros).
- After the claim number is entered, it will either produce any links to documents related to the claim, or a message stating “The search result is empty.”
- If one or more documents appear, open the documents by clicking on the white square icons to the left side of the screen until the desired information is found. For our purposes, the forms that usually produced the information we sought were labeled ‘ACORD Form’, which was either a ‘Loss Report Form’ or a ‘Property Loss Notice’. Other information was sometimes included in documents containing correspondences, emails, bills, information on policies, rates, premiums, or litigation reports.
- The catastrophe codes can be inferred from details found in the ACORD forms. If the primary loss cause was due to a national disaster, you can assume there is a catastrophe code present. The ACORD Forms will have a category for the CAT #, but this section is always left empty regardless if a catastrophe is present or not.
- To go back, or to search another claim number, click ‘: Search 1-General Claim’ at the top of the page. This will bring the user back to the screen which will allow the user to search the claims database.
- The process was repeated as many times as necessary.

CAAMS

- This information was accessed by first logging onto The Hanover System.
Open the intranet and click the link on the left titled, CLi Bank.

A new screen will appear listing a large variety of different applications and databases. Under the systems heading, select CAAMS. At the next screen, under the heading ‘Daily Tools’, choose CAAMS application.

A prompt to enter credentials will appear and, after entered, a new screen with a few folders will appear.

Once in the CAAMS database, on the screen with multiple folders displayed, choose desktop, followed by selecting ‘Name and Address Search’.

Enter the appropriate name and address associated with the claim desired that was found using HCS. After clicking search, results were returned and the correct claim was found by matching the address of the excel sheet to the search results (if multiple addresses were returned).

The ‘Account Review Form’ tab was selected, at the top left of the screen, in the toolbar.

The main search window needs to be left open at this point.

Once the ‘Account Review Form’ tab is selected, it will either direct you to the available claim information or a message stating, "No review forms exist for this claim." will appear. If results did appear, the claim that was titled “in process” was selected, followed by choosing "edit/review" because they were the most current and held the most information.

When the account appeared, all information needed to be displayed by choosing "+all" near the top, to the right of the page. Available information and documents were reviewed, looking specifically for Building Underwriting Reports, Risk meters and Building Valuation Reports that are produced and which contain information desired by the group.

After the desired information was either obtained or not obtained due to lack of availability, and it was time to move on to the next claim, the ‘modify search’ button in the aforementioned window was selected and the process was repeated as many times as necessary.
ARIES

- To search through the ARIES database, begin by opening Internet Explorer and navigating to the address. Click the link that says, “Sign in ARIES Loss Control System”. Doing so will prompt the user to enter a user ID and password. When proper credentials are presented, the user will be brought to the ARIES homepage. From here, the user will click the link “Request Inventory” on the yellow banner at the top of the page. This link will bring the user to the page that will allow the user to search the ARIES database for loss control reports.

- At this page, next to the words “Type of Requests”, be sure to select the option “All”. Then begin to enter the desired claim name into the search bar directly above and hit search. It may take many search attempts to produce a claim that has a report within the database. Once a report is shown, click the link under the heading “Insured” of the report the user wishes to view. Doing so will bring the user to a new page displaying the report. The user should then click on the link to “View Entire Report”. This will again produce a new window, at the bottom of which will have a section for attachments. The most useful attachment will usually be a file that contains the word ‘Report’, which opens yet another window that should contain the relevant information sought.

- To begin another search, enter a different claim into the aforementioned search bar. Continue to repeat this process until all possible claims have been searched. After the completion of the search process for Hanover database there were still many missing data points from all categories. Due to the lack of information, alternative sources needed to be accessed in order to generate sufficient data and satisfy as much of the search criteria as possible.

Interview Contact List

<table>
<thead>
<tr>
<th>Name</th>
<th>Organization</th>
<th>Reason for Contact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jim Ducey</td>
<td>Hanover Insurance</td>
<td>Veteran Loss Control Consultant within the insurance industry</td>
</tr>
</tbody>
</table>
Interview Questions

Preliminary Questions:

- Will it be ok to identify you with you as a source?
- Are you averse to being quoted directly within the project?
- Do you have any questions or concerns for us before beginning?

Questions:

- Can you describe a typical building assessment?
- Are there any characteristics of a building that you look for during an evaluation, which provided either a discount on the premium or makes you increase the premium?
- Before an evaluation is performed is there a basic facts sheet about the building in question that would provide general building specifications/information?
  - Where are they located?
- What document do you fill out when evaluating a building, where are these located once complete?
  - Any databases that contain these documents/information?
- Is there any knowledge gained throughout your experiences that can only be acquired through experience?
  - Based on this knowledge, are you able to more accurately predict certain risk factors?
  - Do you feel you have identified any trends throughout your experiences?
- Could you briefly identify what you feel are some of the chief advantages and disadvantages of the current building evaluation system?
  - What is accounted for well and not so well?
- What do you feel can be done to address these shortcomings?
Appendix D: Findings

Statistics

Linear Regressions and ANOVA assumptions

There are assumptions that need to be made when running simple linear regression models. In fact there are four that have been documented and respected. Without these assumptions, simple linear regression would not be a valid approach to validate or assume (Testing the Assumptions of a Linear Regression 2005).

- The first assumption is linearity of the relationship between the dependent variable and the independent variable.
- The second assumption can be explained as independence of the error, otherwise known as no serial correlation.
- The third assumption also relates to the residuals, otherwise known as the error. The residuals have a normal distribution.
- The fourth and final assumption deals with the term homoscedasticity. Homoscedasticity means constant variance among the residuals. The residuals must vary versus time and independent variables. (Testing the Assumptions of Linear Regression 2005)

As in simple linear regression, there are also four assumptions that have to be made while using an ANOVA table. One assumption is that the expected values of the errors are zero. A second assumption is the residuals all have an equal variance. Thirdly, the residuals need to be independent of each other. Lastly, the errors need to be normally distributed on a histogram (ANOVA Test 2007)