Worcester Polytechnic Institute
Digital WPI

Major Qualifying Projects (All Years) | Major Qualifying Projects

April 2008

Insurance Product Development

Alexander Tolivaisa
Worcester Polytechnic Institute

Jessica Lee Efstathiou
Worcester Polytechnic Institute

Yurong Y. Mao
Worcester Polytechnic Institute

Insurance Product Development

A Major Qualifying Project Report for
A requirement of the Degree of Bachelor of Science from
WORCESTER POLYTECHNIC INSTITUTE

Written by:
Jessica Copp

____

Yurong Mao

____

Alexander Tolivaisa

____

Project Center: Worcester, Massachusetts

Term: A07-D08

Sponsoring Agency: N/A

Project Advisor: Professor Jon Abraham

Abstract

The project's objective was to develop and price a new insurance product. Rollercoaster insurance was chosen for its uniqueness. Data for injury numbers, number of riders and deaths were necessary for pricing, but primarily private. Risk rate and ridership were modeled based on limited parks' data. A portfolio was developed using the modeled data. The purpose of portfolios was to share the risk and reduce premiums. After adding in expense and profit, a final premium was determined for each rollercoaster.

Executive Summary

This entire project was based around one objective: developing and pricing a new insurance product.

After discussing and researching some unexplored insurance areas (personal watercraft, truckers, etc.), roller coaster insurance was chosen. Rollercoaster insurance would provide financial coverage to amusement parks for injuries that occurred on their roller coasters. Roller coasters were chosen for the following reasons:

- There exists a real market to sell this product to. Roller coasters operate in 42 different states; each state government requires the ride's ownership to carry liability insurance.
- There was good data available to build the project on. At the start of the project, national government safety data was easily available. The reports provided extensive national injury totals and trends.
- This type of product is unique.

To start the project, data was needed to understand the severity and likelihood of roller coaster injuries. The government data provided a good foundation for this process. Once all of the government data were sorted out, obtaining specific park injury data was the next step, which became a great challenge. Park data was needed to accurately represent the claims amount. However, individual amusement parks choose to disclose as little information as possible with regards to injuries sustained on their rides (similar to tobacco companies with regards to the side effects of their products). One website, saferparks.org, did provide published reports of injuries occurring on roller coasters. For instance, a complete history of every injury that happened in Disneyland during 2001 was displayed.

However, there was great inconsistency with the saferparks.org reports. In 2001, there were 58 reported injuries in Disneyland, the next year there were only two. Obviously, there is some bias in the data, because they only collect data from hospitals; some riders that suffer small injuries will forgo a hospital trip. Other amusement parks make no injury numbers available at all, for instance many Six Flags parks are proficient at keeping their injury data under wraps. In fact, Six Flags does not even publish their park-by-park attendance numbers. With information from saferparks.org being the closest source of real data, another direction of acquiring data was necessary, modeling the data based on the samples that were found.

Creating injury data gives an output of expected claims amount for a particular coaster at a given year.

The first task was coming up with a list of factors which determine the safety of a roller coaster. Three factors were used, obviously there are more factors that determine the safety of a roller coaster; however, size, safety/location and material were considered. These three factors were assigned a certain weight to an average risk-rate, which is the ratio of injured riders to riders in general. Size was given the largest weight, safety/located was the middle weight while material was the smallest weight. From here, a certain number of variables were assigned to each factor, which represented a certain degree of the factor. Size had three variables: small, medium and large. Safety/location had four variables: level one, two, three and four, with level one being the safest. Lastly, material had two variables: wood and steel, with wood being the safer of the two. This produced 24 different categories of roller coasters (3 sizes * 4 safety levels * 2 sizes equals 24).

One additional factor, the age of the coaster is an important factor influencing the risk rate. To account for this situation, the risk-rate of a coaster at a certain age would be created. The coaster's age would range from creation (age zero) to age 50. This meant 1,200 different types of roller coasters would need risk-rate data. Age is separated from the other three factors because the previous factors determine the underlying risk-rate, while age changes and adjusts the risk-rate.

The data was modeled following a trend that would incorporate the following logic (Figure 1). The ride would have a relatively normal risk-rate then steadily increase until age 12, when the park would make safety improvements on the ride. The downward slope would continue until age 26, when old age starts to catch up to the coaster. The risk rate would slowly increase from that point on until it reaches age 50, when the observation period ends. Many other risk patterns are possible, but in the absence of solid data, this approach has been used to illustrate the process.

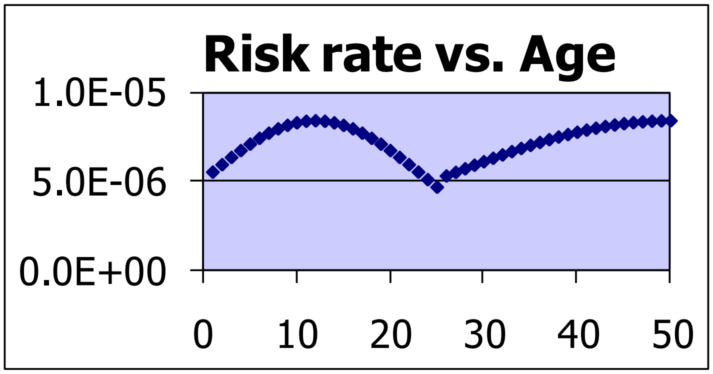


Figure 1 - Ideal Risk Rate Graph

For the risk rates to follow this trend, they needed to follow two normal distributions, one from age 0 to 25 and other from age 26 to 50. The risk rate that was calculated from using the characteristics of the coaster would be the maximum risk rates, at age 12 and 50.

To simulate real data, some variation was added (Figure 2). The variation was factored in by creating an interval that had its endpoints at a certain distance away from the original risk-rate. A random uniform interval created was, where any value within the interval could be selected as a sample risk-rate (illustrated in figure 2, with arbitrary risk rates).

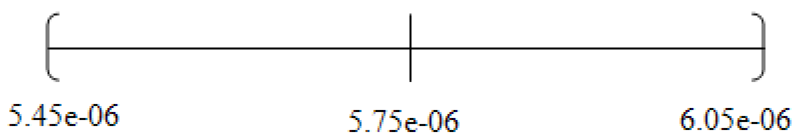


Figure 2 - Sample interval

Thirty sample risk rates were chosen for each age, the averages of these samples would be the "real" risk rates which are graphed below.

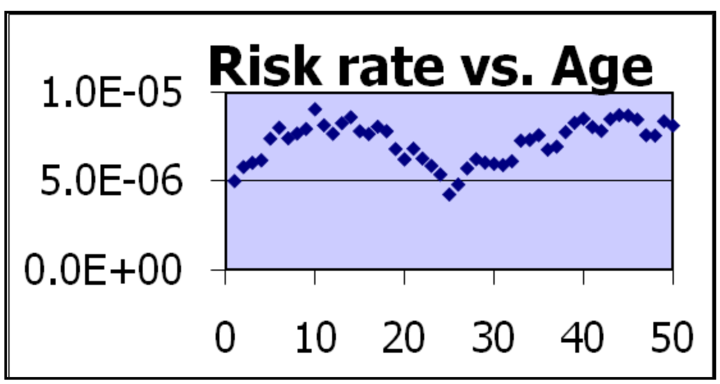


Figure 3 - Risk rate graph with variation

The formula to calculate the number of injuries on a particular coaster for any age was: risk-rate * number of riders. The safety factors that were used to create the risk-rates determined the number of riders of a coaster. When the total injuries were calculated, they were divided into small and large, with 50-80% of the injuries being small, the rest being large. The exact percentage was determined by a random number generator.

The formula for a coaster's expected total claims amount was: (number of small injuries * expected small injury claim) + (number of large injuries * expected large injury claim). The expected small and large injury claims ranged from $50-$1,000 and $2,000-$6,000, respectively. The exact claim amount was determined by a random number generator. Following the methodology of the risk-rate calculations, 30 samples of claim amounts for each coaster's age were taken to determine the average expected total claims amount and its standard deviation.

Before determining the pricing method, a few goals were established to ensure that the pricing method devised was sufficient.

- There would be a 95% chance that the premiums collected would cover all claims.
- The coaster's premium would be competitive.
- Having a fair way of allocating savings to different coasters, if possible.
- The method would be easily explainable to potential customers.

Keeping in mind these goals, there were a few options to calculate a coaster's premium.

The first option was taking the coaster's average total claim amount and adding with its standard deviation multiplied by 1.645, which would produce a 95% chance that the coaster's yearly claim amount would be covered. This option would make the premiums quite large, forcing each individual coaster to "stand alone" and there would be a little element of risk sharing. Adding that figure to an already expensive price would not make good business sense.

The second option was forming groups or "clusters" of coasters with close expected claim amounts and assigning a premium to that group. This option would take advantage of risk pooling to lower a coaster's premium; however, the method of forming the clusters was not completely subjective.

The third option was creating a portfolio of multiple coasters with a variety of characteristics. The option requires that at least 30 coasters be put in the same portfolio, regardless of its expected claim amount. This was the option that was finally settled upon.

Summing all the coasters' claims amounts and their variances was the first step. This led to taking the square root of the total variance, to calculate the standard deviation of the entire portfolio. The portfolio's standard deviation was multiplied by 1.645 then added to the sum of the claim amounts. The output of this process would be the "fund amount" or the needed total amount of premiums collected to ensure a 95% chance the premiums would cover the claims.

The next step was multiplying each coaster's standard deviation by 1.645 and adding it to its expected claims amount, which was the coaster's "pure premium." Next, was summing the pure premiums and subtracting the sum by the "fund amount." This would produce the portfolio's "total savings."

To fairly divide the savings among the roller coasters, the following process was used. Each coaster's standard deviation was divided by the sum of the coasters' standard deviation. The percentage was multiplied by the total savings, which produced the individual coaster's savings amount. The coaster's pure premium is subtracted by the coaster's savings amount, which produced the coaster's real premium or the amount they would have to pay for insurance before any consideration of expenses or profit.

A coaster's standard deviation is the determining factor for how much a coaster should save for the following reasons:

- It is easy to explain to amusement park operators.
- Coasters with a high standard deviation have two properties. 1. A high expected claims amount. 2. The likelihood of having lower claims than projected, which is favorable from a business's prospective. For instance, there are two coasters that both have $10,000 in expected claims and $1,000 and $4,000 as a standard deviation, respectively. The likely claim amount for the first coaster in a given year ranges from $9,000 to $11,000. The likely range for the second coaster is $6,000 to $14,000. Since there is a significant chance that the second coaster will have a favorable claim amount, compared to the first coaster, the second coaster will receive more savings.

After the coaster's premium was calculated, an expense ratio and profit are developed. The expense ratio is included for business expenses. The formula for the including the expense ratio and profit is: $\frac{\text{Premium}}{(1-\text{Expense Ratio})} * (1+\text{Profit Ratio})$. For this project, the expense ratio is 20% and the profit ratio is 15%, which means a coaster's actual premium would be its real premium multiplied by 1.4375 or 23/16.

The objective of this project was developing a pricing function for a new insurance product. To complete the objective, creative logic, unfamiliar spreadsheet tools and extensive mathematical projections were used. These resources would create a process that would make this insurance product a success.

Table of Contents

Table of Figures

Table of Tables

1. Introduction

Insurance companies throughout the world are constantly looking for new areas to sell insurance. For instance, many are now insuring specific types of deaths (i.e. cancer-related, Parkinson's, etc.). Even non-insurance companies practice this behavior. Cell phones companies offer "cell phone insurance" for customers who worry that their phone may become damaged or lost. Several things must be in place before launching the new insurance, and pricing is among the most important.

A primary objective for a company introducing a new product is to increase their after tax income. To achieve that objective, companies must charge the appropriate price for their product. This is one reason insurance companies employ actuaries, in order to create optimal prices for certain policies and readjust them as certain factors change over time.

The assigned goal of this project was to create a new insurance product and determine how to price the product. Developing rollercoaster insurance, which provided many challenges and obstacles, was chosen.

The product is designed to provide parks with financial security when injuries occur on their rides. In personal injury cases, plaintiffs often ask for incredible settlements which put defendants in a huge financial hole. Few companies can afford a multi-million dollar injury settlement, including amusement parks. A major accident has bleak short-term consequences for the park:

- Attendance drops because people are overly considered about their safety.
- Profits shrink considerably because of injury settlements and legal fees.
- The ride on which the accident occurred maybe shut down for an extensive amount of time due to repairs and constant inspections, which is costly.
- The media gives the park negative publicity.

The product would ease the financial consequences of an accident, so the amusement park operators can focus on other things (i.e. media relations). However, this protection is not being offered because of decency; it is to make a profit by selling policies as many as possible. Pricing is important as customers very sensitive to costs.

To price accurately, several factors were accounted for. Intuitively, it was assumed that roller coasters are different in some aspects, such as speeds and heights of a roller coaster; hence that was taken into account. It also became apparent that no single pricing method would make all customers happy, which meant the model had to make sense. A huge factor that was brought to our attention was the difference between insuring an injury and a death. It was estimated that a significant injury (i.e. a broken arm) would result in a $3,000 claim, whereas a death claim would be seven figures.

The rest of the report will show the specific types of problems encountered and how they were solved. In the end, a price function was created that should seem fair to all parties involved.

2. Background

2.1 Purpose

When assigned with this project, a decision on which insurance area that would be explored was needed. There were many areas of interest that were looked at, such as personal watercrafts (PWC), roller coasters, truckers, and umbrella policies. After some preliminary research, we examined our findings. PWC offered limited information, umbrella policies were simultaneously too complex and conventional and truckers insurance felt somewhat boring. There were two general requirements for the subject area:

1. The subject would be interesting.
2. There would be a good amount of information available on the subject.

We uncovered a substantial amount of government data on roller coasters. The government offered great, cumulative injury data that occurred in the United States. For instance, there was breakdown on what kind of ride (rollercoaster, waterslide, etc.) an injury occurred. Lastly, rollercoaster insurance had the added benefit of being pretty unique and uncharted. To illustrate this statement, only two foreign companies offered this type of insurance.

2.2 Terminology

Throughout this report there will be several terms used that are unfamiliar to many people. To make the report clearer, the following words will be defined:

- Risk Rate: The likelihood of an accident happening. In the case, the probability of an injury occurring on a rollercoaster.
- Raw Price: The initial, calculated price. In this case, the amount we will charge a customer before expenses and profit are added in.
- Centroid: Average or mean value of the objects contained in the cluster on each variable.[1]
- Account Credit: When you have more than one type of coverage for the same account (example, death and injury coverage).

[1]http://www.google.com/url?sa=X&start=11&oi=define&q=http://www.richmond.edu/~pli/psy538/Cluster%2520Analysis/Key%2520Concepts.html&usg=AFQjCNGIBnkAag1XkRz2M-kvvpzPhdI7Qg

- "Fuzzy Factor": How closely data points match an underlying trend line, with 100% being a perfect fit.

2.3 Safety Regulations

Roller coasters and amusement parks are not nationally regulated. Instead different states have different laws governing their safety. For instance, there are four different types of inspection programs used by states[2]:

1. Government inspections: State employees directly monitor the safety of all rides within their jurisdiction. There are 28 states in total that have this type of program in place.[3]
2. The Insurance/third party system: The park operators receive a list, from the state, of mandatory safety requirements for the rides. However, it is up to the owners' insurance company and/or an independent party to make sure that the park is satisfying the state requirements. There are 7 states that use this program.
3. Insurance requirements: Park operators are only required to carry liability insurance. The state does not require anyone to monitor the safety of the rides. There are only 9 states that use this "laissez-faire" approach.
4. Electrical inspections: Government agencies perform an electrical inspection of the rides, nothing else. The only state who has this policy is Idaho.

With regards to liability insurance, there are 42 states where rollercoaster operators are required to have some liability insurance; these states are the target market for our type of policies.

- In 41 states, amusement park owners must inspect their rides, on a regular basis, to make sure the ride is meeting the manufacturing requirements.
- In 37 states, parks must register new rides with state agencies and must have the registration renewed in order for the ride to function.
- In 37 states, parks must publicly report severe injuries sustained on their property, however only 20 states require parks to report any medical treatment received at their park.
- In 27 states, parks are required to allow easy access to their public safety records.

---

[2] http://saferparks.org/database/agencies/scoring_criteria.php
[3] http://saferparks.org/database/agencies/state_inventory.php?sort=scored

In addition, there are other laws that would help sell the product:

According to www.saferparks.org, Massachusetts and New Jersey have the best safety regulations out of all the states, while Nevada, Idaho and South Dakota have the lowest scores.

However, just because a state has an outstanding record with regards to regulations, accidents can still happen. On May 1st, 2004, an overweight man fell to his death after he slipped out of his seat because his seat's "T-bar" was not fully locked because of his girth. The accident happened at Six Flags New England located in Agawam, Massachusetts.

2.4 Waiver laws

As a mean to protect themselves from liability, parks use waivers. A waiver is "the act of intentionally relinquishing or abandoning a known right, claim or privilege," and also as "the legal instrument evidencing such an act." A waiver is useful for protection, but it does not guarantee to relieve the amusement park's responsibility or liability. Personal injury lawsuits are still brought to the amusement park, even if a waiver was signed by plaintiff, for recovering some medical expenses.

There are a number of other occurrences in which courts may rule that waivers are invalid:

1. Parents have signed waivers and then did not remain on site when their children went on the ride.
2. It can be shown that the equipment was not properly set up. Once negligence is proven, the court usually negates the validity of the waiver.
3. In cases of rentals, the waiver was only signed by the owner of the site on which the equipment was set up, and not only by individual parents. In such cases, when the site owner has been added to the policy as an "Additional Insured," the court may negate the validity of the waiver.
4. The waiver was signed by people who were not related to the children and are therefore not authorized to sign, such as chaperones, teachers, group leaders etc.
5. It can be shown that the signature was requested in such a way that the signatories did not know exactly what they were signing.

There is a high probability for these occurrences to happen. If the injured person is a child, when he/she reaches legal age, there is a chance that he/she could sue the park, claiming that their parents were without any right to sign the waiver for them. All these cases demonstrated the role of insurance companies in this situation. Rollercoaster insurance product provides the kind of protection amusement parks need, even for big theme parks like Disneyland. They could end up with a number of payouts for injuries in a particular year when a new, state of the art rollercoaster is launched. It would be wiser with more profit in the income statement for them to buy insurance to cover the claims instead of paying on their own.

2.5 K-Means Clustering

One statistical method that was tried to price was K-Means Clustering. For this section, the process will be defined. The application of this method is in Section 4.2.

The simple purpose of K-Means Clustering is to divide a number of objects into "K" groups or clusters, where K is a pre-determined number.[4] Groups are determined by the following process[5]:

1. Place K points into the space represented by the objects that are being clustered. These points represent initial group centroids.
2. Assign each object to the group that has the closest centroid.
3. When all objects have been assigned, recalculate the positions of the K centroids.
4. Repeat Steps 2 and 3 until the centroids no longer move. This produces a separation of the objects into groups from which the metric to be minimized can be calculated.

Note that it is possible to end up with a different set of clusters if the process is repeated. K-Means Clustering is not guaranteed to result in a unique answer.

Since there were 1,200 different types of roller coasters that were under observation, it was believed that this method would save a significant amount of time.

---

[4] http://fconyx.ncifcrf.gov/lukeb/kmeans.html
[5] Ibid

2.6 Research Techniques

In order to obtain the necessary data for the models, several research outlets were used. Among the outlets were search engines, lobbyists' websites, accounting sheets and government reports.

The initial data was used to obtain a general idea of how often rollercoaster accidents happen. The reports were looked at to see if there was any specific data on particular theme parks; however, there was not. Searching company websites to see if they would post any information regarding injuries was another option, but the companies did not post anything directly useful for our purpose. Next, www.saferparks.org was visited, which ended up being the primary source for park injury data for this project. On the website, there were injury reports for each year from specific parks. They also had information on state laws and regulations and updates on safety modifications. However, the park injury reports were limited as they only published reports that were made public; it appears there are some gaps in the information.

The information on the www.saferparks.org was incomplete in other ways, as well. Settlement figures for major accidents were needed to come up with realistic figures for the premiums. Nearly all of these cases are settled out-of-court which means that the final settlement figures are kept private.

Researching the primary factors of rollercoaster accidents was next. After some basic factors were decided upon (e.g., size), the specifics of each factor was explored. For size, a report from Cedar Point about their different sizes of coasters was looked at and the size criteria were created afterwards.

It was difficult to put a number on safety by ride or even by park; however, www.saferparks.org already gave safety scores by state in their "Inventory of State Amusement Ride Regulations." The inventory was used to create four safety groups.

To represent a diverse range of roller coasters, a third category that was independent of the first two was needed. Looking at some of the government reports, it was discovered that is a significant difference between the injury rate between a wood (safer) and steel rollercoaster. Hence, material became the third category.

Once the model was ready for building, research some ways to build an efficient premium function was next. Clustering together certain coasters to save time and hassle was considered; hence K-means clustering was studied. Other than that, trial-and-error was mainly used to create an efficient method.

2.7 Marketing

Having a large enough portfolio to provide savings to the parks and covering financial risks is necessary for success in sales. One of the best ways to appeal to customers is to offer discounts, similar to the discounts offered for auto insurance. Some offers could be a multi-coaster discount and account credit on a per park basis. Multi-coaster discounts could be applied to a park's total payment (including all the roller coasters) if they insure multiple coasters. Account credit would apply to the park if they chose to purchase our death coverage as well as insure at least one coaster. Another offer would be a safe-coaster discount, if the parks have more stringent safety regulations than the state requires. Different payment methods would be available. The premiums would be based on a one-payment plan, but if the park would prefer to make installment payments, they could choose to do so.

A primary goal would be to maintain close, personal, and friendly relationship with customers. The initial months would be spent going to the parks to meet with the owners and explain our policies and the benefits of having our insurance. A copy of the product's brochure can be found in Appendix G.

3. Methodology

3.1 The Problem of Data

One of the biggest problems during this project was finding data. What amusement park would want to disclose information about how unsafe their main attraction is? After extensive research, it was concluded that actual, accurate data was virtually unattainable. However, some numbers about injuries were found, though they were not complete, but at least they were a place to start from.

A theory was created about how the injury rate graph should look like. As pictured in below, first the injury rate would be increasing, due to popularity of a new ride, aging and constant use. After a certain number of years, especially if there have been complications, the park operator would spend time and money on repairs, causing the risk to decrease slightly. Then, the risk would continue to increase over the remaining life of the coaster.

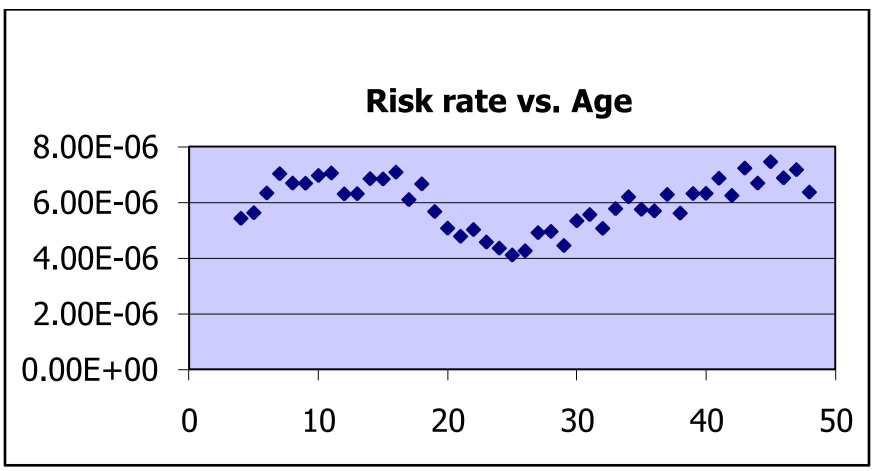


Figure 4 - A sample risk rate graph

This would not be a smooth curve due to occasional data discrepancies and "noise", but it would still follow a basic shape. Using this basic shape, an input sheet was developed and data was created to fit the curve.

3.2 Factors

$$Price = Average\ Accident\ Rate \times Risk\ Factor_1 \times Risk\ Factor_2 \times Risk\ Factor_3$$

To calculate the risk rate, it was necessary to decide what made a rollercoaster potentially dangerous. Had actual data been available, it would have been worked until correlations were found. However, that was not the case, so it was decided to list likely risk factors. To simplify the model, only four factors were chosen, but more could have easily been added. Each factor has a different effect on the risky rate of a given coaster. While there may be some correlation between the factors, the ones with the least correlation were chosen. The four chosen factors chosen, in order from of highest to lowest impact on the risk rate, were size, state safety score, material, and age.

Logical thinking was used to set numbers for the factors because of the lack of actual data. However, it was realized that these numbers are more of place holders. These numbers were determined by discussion of importance to risk and how it should impact the overall outcome. For example, it was believed that a large rollercoaster has more of an impact than any other factor. If there was actual data, mining could have been used to find correlations.

It was decided that size would be the largest risk. Anytime an accident or death from a rollercoaster incident is heard about through the media, it occurred on a large, fast coaster. The small ones that young children ride are low to the ground, travel slowly and usually only go in a small circle. Therefore, size is an important factor. Size was broken down into three categories; small, medium and large. The size is determined by speed, which is shown below. Speed was chosen because it had a definite, unquestionable, measure, but there is a direct correlation between the speed of the rollercoaster and its height.

Size	**Speed**
Small	Less than or equal to 30 mph
Medium	Between 30 and 60 mph
Large	Greater than or equal to 60 mph

Table 1 - Size categories

Since large coasters were more risky, it has a larger factor. It was believed that medium coasters were not right in the middle, but leaned more towards the smaller coasters than it did the larger ones. Medium roller coasters are skewed more to the side of the smaller roller coasters because they tend to be less popular, and while they can still reach large heights, they usually do not have loops or steep drops. These characteristics make the risk factor lean in favor of the smaller coasters. Below is the breakdown of the different risk factors associated with each size that was developed.

Size	Risk Factor
Small	0.5
Medium	1.7
Large	4.0

Table 2 - Size Risk Factors

The next factor was the state's safety score. The safety scores are based on the government's safety regulations for the state. This covered other areas that should be included, such as location, park regulations and their maintenance schedules. Even though it is state based, it still provides an overall idea of how safe each park should be within those states. The scores range from 0-30 with 30 being the highest. The scores were broken down into four levels, with level one being the highest scores. This was done by applying the Central Limit Theorem to the list of states and their respective scores and finding the 25th, 50th, and 75th percentiles. See Appendix A. The final grouping came out to the following decision.

Score	Level
23-30	1
18-22	2
14-17	3
0-13	4

Table 3 - Safety Score categories

As in size, each level was assigned a risk factor, with no effect on the states with the highest scores. Again, no actual data was available, so the created data was created by comparing the weight on the risk-rate to the risk factors that were assigned to the size of the roller coasters. The less regulated states looked as being very risky, and centered the two middle levels around 1.5 with level three taking a slightly harder hit than level two. This setup is intended to magnify the difference between the better regulated states and the "laissez-faire" ones.

Level	Risk Factor
1	1.0
2	1.3
3	1.8
4	3.0

Table 4 - Safety Score Risk Factors

Usually, large, fast roller coasters are made out of steel. While wooden roller coasters seem more unstable and shaky, it is the steel coasters are more likely to have an accident. Wooden coasters are built to feel unstable, which adds to the adrenaline rush, but they hardly ever malfunction. The steel coasters are designed for height, speed, and architectural grandeur. All these advances add to the thrill, but also add to the number of complications that can occur. The material is closely associated with the size of the coaster. This factor is still necessary because steel roller coasters spread across the spectrum of sizes, where wood roller coasters are not usually found in the large category.

Material	Risk Factor
Wood	1.0
Steel	1.2

Table 5 - Material Risk Factors

The age was also considered when calculating the risk of these coasters, though it is not applied the same way the other factors are. After the overall injury rate of the coaster, using the first three factors, is determined, a graph is plotted to show the risk rate of that particular type of coaster over its lifetime (ages 0-50). Once the graph is plotted, the risk rate at that age can be applied to the risk before calculating the expected injury numbers and payout amounts. The graph can be seen in Appendix B. The graph is not smooth because of a "fuzzy factor". The fuzzy factor adds variation to the data to make the graph more realistic. To create that data, the fuzzy factor was set to 90%, where 100% would be a line exactly following the curve.

Although it is not a direct input into the model for risk rate, the number of riders plays an important factor in the number of expected injuries for a given rollercoaster and therefore its expected payout. A larger rollercoaster is going to have more riders than a smaller coaster. The small roller coasters have lower risk rates as well as fewer people riding them, so their expected payout will not as high. On the other hand, a large, very popular rollercoaster has a higher risk rate, and with the increased number of riders, will have a much higher expected payout.

3.3 Modeling Data

Without data, pricing the roller coaster insurance would be impossible. Therefore, it was necessary to create the data before we used it to price our product. The process began by finding a range for the data using the little actual data available and created reasonable data to acquire the risk-rates found for different coasters by plugging its characteristics into an input sheet template which returns a coaster's risk-rate. See Appendix B. Then, keeping around the same risk, 30 different coasters were created (from their creation until 50 years out), all having the same characteristics, by taking a random distribution around the reasonable data. Creating a sample of thirty allowed for the application of the Central Limit Theorem to find an average risk rate for that type of coaster.

This method of creating thirty coasters and using the normal distribution was also used to find the average number of riders and number of injuries on a given rollercoaster type. There was some actual data in these areas. The actual ridership numbers seem correct; however the injury numbers seemed too low. The probable reason for the low injury rates were that only some major injuries were reported. However, there are also smaller incidences that require medical attention, whether just for precautions or simple procedures such as stitches. People over react in some situations and think a trip to the emergency room is necessary when it is really nothing. For example, a child may scrape their knee as they climbed out of the rollercoaster car, but a parent believes the child needs to go to the hospital when they just needed a Band-Aid. It is the non-serious injuries that are not reported and therefore keeping the injury numbers we found low.

Knowing that some payments would be for small cuts and bruises, while others would be for extensive injuries, such as loss of limb, the injuries were divided into two groups, small and large. Using the number of riders and injuries, the expected payout of the coaster was found by simply multiplying the number of small and large injuries by the average claim amounts for each type, respectively. After some research online to find out the cost for both small and large injuries, it was decided that a reasonable average claim amount for each would be $650 for small injuries and $3,000 for large injuries.

4. Pricing

4.1 Pricing a Single Rollercoaster

If there are customers buying rollercoaster insurance, an amount of money called the “premium” would be collected to offset losses. The premium is determined by the following: the expected payout of a claim, the standard deviation and the probability that a claim will occur. The methodology for pricing can be explained using the following example.

The probability of having a claim in one year for a 10-year-old rollercoaster, in the category of L/S/4 (large steel coaster with level four safety score) would be 5.4 accidents per million riders for every particular rider on the coaster. This number was calculated by modeling the injury rate of L/S/4 coasters following a curve for age 1 to 50 as discussed previously. This number also indicates that if there are 2,173,087 riders on this roller coaster, there would be an expectation of 12 injuries resulting from the 5.4 accidents per million times two million of riders. According to the actual data, 50-80% of the 12 injuries would be small, such as a headache or whiplash. The rest would be large injuries such as heart attacks and broken arms. From the injury severity of Disneyland in recent five years, it is estimated that the payout for a claim would on average be $650 per small injury and $3,000 per large injury. Hence, the expected payout would be $14,850 ($650*9 + $3000*3= $14,850) for this 10-year-old coaster in the category of L/S/4.

With 30 identical L/S/4 coasters, the mean and standard deviation of payouts for these 30 roller coasters are $18,221.70 and $1,246.18, respectively. The normal distribution was used to find the 95% confidence interval using our found mean and standard deviation. Therefore, the final premium amount would be $\sigma * 1.645 + \mu$, $\$1{,}246.18 * 1.645 + \$18{,}221.70 = \$20{,}271.67$. To turn this into a final price, the expense ratio and profit would be added onto the final premium amount.

Pricing a single rollercoaster would be the result of an unsuccessful sales policy. Selling a policy with such a high price is difficult and uncompetitive. This is why it is necessary to find some way to group the coasters together to share the risk and reduce the prices.

4.2 Clustering

According to the data collected by IAAPA (International Association of Amusement Parks and Attractions), there are approximately around 400 roller coasters totally in US.

Table 1. Survey Highlights

	2001-2002	2003*	2004	2005
Number of Facilities	459	403	403	398
Estimated Annual Attendance (millions)	302.9	300.4	300.0	300.4
Estimated Annual Ridership (billions)	---	1.95	1.81	1.82
Estimated Annual Number of Ride-Related Injuries	2,486	2,044	1,637	1,783
Injuries per Million Attendance	8.2	7.0	5.2	5.2

*Changes in the estimating method beginning with 2003 affect comparability with the 2001-2002 survey.

Figure 5 - IAAPA attendance figures

Furthermore, they are categorized by three types of size, two types of material and four types of safety scores, which create 24 possibilities for coaster type. When age of coaster is included, it brings the total possibilities up to 1,200 (24 categories * 50 ages=1200). It is impossible to group the exactly identical coasters to help share the risk and reduce the premium. However, clustering different types of coasters with various ages would help reduce the price by sharing risks.

The price of every rollercoaster policy should be cheap enough to attract customers. Using clustering techniques, coasters that have an expected payout at a reasonably close level are grouped together. The expected payouts of coasters were listed from 1 to 50 years old in each category in a spreadsheet where we could see where the clusters be grouped. We started by trying to cluster the wood medium sized coasters with safety score three and steel large safety score four coasters together. By ordering the two lists from smallest expected payout to largest, it becomes easy to identify how many of them and at what level would be clustered together.

Medium Wood 3		**Large Steel 4**	
Age	Expected Payout	Age	Expected Payout
25	$4,570.42		
26	$5,324.36		
●	●		
●	●		
●	●		
48	$9,800.39		
5	$9,903.54		
14	$9,938.25		
2	$9,943.24		
49	$10,093.60	*41*	*$10,092.80*
3	$10,174.09	25	$15,349.91
12	$10,590.37	29	●
●	●	24	●
●	●	31	●
●	●	26	$19,975.42

Table 6 - Ages and Expected Claims amount for certain MW3 and LS4 coasters

Suppose the 14-year old M/W/3 coaster would be the lowest expected payout in the portfolio and the range of expected payouts cannot exceed 1000 dollars. Based on these criteria, the resulting cluster is shown below.

Age	Expected Payout
48	$9,800.39
5	$9,903.54
14	$9,938.25
2	$9,943.24
41	***$10,092.80***
49	$10,093.60
3	$10,174.09
12	$10,590.37

Table 7 - Cluster of MW3 and LS4

The only coaster that belongs to L/S/4 is italicized in Table 7 with an expected payout of $10,092.80. E(x) is the average payout for these eight coasters. If there are two coasters, the premium is equal to $2 * E(x) * a$, where “a” stands for the ratio of premiums by clustering two coasters over the total expected loss. The confidence level is set to be 99%; so that there is 1% that the amount of premiums will not cover the loss occurred. It helps to avoid bankruptcy.

E(X)=	$10,067.03	**premium p=**	a*E(X)*n	**Ratio**	
SD(X)=	$243.91				
if n=	2	**premium p=**	$56,151	**a=**	2.79
if n=	10	**premium p=**	$3,254	**a=**	0.03

Table 8 – Clustering formula

The table above showed that as more coaster policies are sold, the ratio decreases. The formula to calculate the premium is $E(\mathrm{a}) + \mathrm{inverse} - \mathrm{t} - \mathrm{distribution} * \mathrm{SD}(\mathrm{a})$; where $\mathrm{E}(\mathrm{a}) = \mathrm{n} * \mathrm{E}(\mathrm{x})$, n is the number of coasters, while E(x) is the average payout $10,067.03 for these eight clustered coasters; $\mathrm{SD}(\mathrm{a}) = \mathrm{SD}(\mathrm{x}) * \sqrt{\mathrm{n}}$ is the standard deviation $243.91 for the clustered coasters' payout. The ratio- "a" equals to the premium dividing the total expect payout of all the coasters. The formula notation to calculate is r $= {}^{\mathrm{p}}/_{(\mathrm{n} * \mathrm{E}(\mathrm{x}))}$.

Looking at the value of ratios makes the effect of clustering obvious. As more policies are sold and the portfolio becomes larger, both the ratio and every policyholder's price in the portfolio are decreasing. Therefore, the policyholder pays less with more coasters in the portfolio. This explains the benefits of clustering and the necessity for insurance companies to lower the charging price by clustering. An example of clustering can be found in Appendix C.

However, this method produced many mistakes involving in the calculation of the standard deviation and criteria selected to clustering. The standard deviation above is the difference between expected payout. It makes the premium decrease dramatically from $10,000 to $3,000 with only ten coasters clustered. Such a drastic decrease of payment, due to the wrong standard deviation applied, is impressive but impossible. However, the correct standard deviation should be different and unique for each coaster's payout, which shows the range of payout's domain. The expected payouts were plotted scatter plot in SAS by looking at in the 3D space. Further, for the rollercoaster payouts, the mean and standard deviation was calculated for each kind of coaster from 1 to 50 years old.

Initially, clustering seemed like it would work, but then some standard deviation problems arose. Clustering works only when there are clear differences between the importance levels. However, the importance of the standard deviation was ignored in this pricing process. There is no definitive way to know which coasters should be closer together if given something that looked like the following.

Type	Age	Expected Payout	Standard Deviation
MW3	13	$50,000	$16,000
MW3	17	$100,000	$16,000
LS4	4	$100,000	$20,000

Table 9 - Claim amounts of certain coasters

It would be a personal opinion of which ones should be grouped together and there is no park that will be supportive of paying a premium based on someone's opinion. Therefore, another method was needed to replace this clustering method to reduce premiums. It was discovered that creating different kinds of coaster portfolios with various types of coasters results in better savings received by the coasters' owners. It applies a proper method of spreading saving among the coasters. Many methods were practiced through the following section and finally the standard deviation ratio one is selected.

4.3 Independence and Pricing

To begin looking for new ways to group the roller coasters as a way to reduce price, we came across the idea of independence of the coasters and what that means for us. This brings up some questions, but look at the arguments individually. If the coasters are in different parks, or are different types of coasters, it may seem more obvious that they are independent. But what if they are both large, steel roller coasters within the same park? The safety of the coasters seems to be not independent because they are only a couple feet from each other. In fact, these have been factored in the location/safety score and ridership, and therefore this should not be considered anymore. This is the same with the number of people in the park and its correlation to the number of riders on any given coaster. By applying the number of riders, ridership on any coaster, the injury number for similar coasters in close distance would not be correlated with each other. Therefore, the coasters are independent of each other.

Independence is a valid assumption we can use to lower the cost of insurance for a particular rollercoaster. The book of business will be made up of as many roller coasters as we can sell to. By putting each coaster's information into the template we have set up, the coaster's expected payout as well as its standard deviation and variance will be computed by excel. Once this information is obtained, the total expected amount to pay out for the entire book of business and the portfolio's total variance can calculated. This is where independence comes into play because finding the totals only requires summing the expected values and the variances in the formula below.

$S=X_1+X_2+X_3+\ldots\ldots+X_n$ (X_i are independent from each other)

$E(S) = E(X_1) + E(X_2) + E(X_3) +\ldots\ldots+ E(X_n)$

$Var(S) = Var(X_1) + Var(X_2) +\ldots..+ Var(X_n)$

To cover the losses incurred for the entire block of business, premiums will be collected from the amusement park. The formula to calculate the total fund needed is: $\sqrt{\sigma^2_{Total}} * 1.645 + E(x)_{Total}$.

After the total premium needed is calculated, the prices can be fairly allocate to the different coasters. This process starts with each coaster's individual premium: $\sqrt{\sigma^2_{Individual}} * 1.645 + E(x)_{Individual}$. This value gets multiplied by their weighted average of their individual cost as compared to the total cost of all the coasters if calculated individually. The weighted average is multiplied by the total fund amount needed to get each coasters premium. This provides a discount to the coasters as compared to what they would have had to pay if everything was done individually and we had not grouped them together.

4.4 Developing a Portfolio

To simulate reality, portfolios of different coasters were created to determine the actual premium an individual rollercoaster would have to pay if they were insured. The main objectives of building a portfolio were:

- Determining how much money to collect so there can be a 95% confident interval of paying all claims.
- Allocating the savings from sharing risks and calculating a fair premium for each rollercoaster insured.

To create the pricing algorithm, multiple coasters (of all types) were grouped together to help bring down the total variance by sharing risks. Variation was kept variation until the premium amount was ready to get calculated. To achieve 30 samples of each coaster (over their entire life) were taken and put into the portfolio, and for each of these 30 coasters, the following was needed: the injury rate, the number of riders, small injuries, large injuries, the average amount paid out for each of the small and large injuries. The table below summarized the formulas involving in the calculation:

Frequency for injury accidents:

Injury number=	Injury Rate * Number of Riders
Small Injury No.= N_s	(70% to 80%) * Injury Number
Big Injury No. = N_b	Injury Number – Small Injury

Table 10 - Injury Frequencies

Severity for Injury accidents:

Small Injury Payout= P_s	300 to 700 dollars
Big Injury Payout= P_b	2500 to 8000 dollars
Total Payout=	$\mathbf{N_s * P_s + N_b * P_b}$

Table 11 - total claims formula

Once all of the necessary numbers were obtained, the average expected payout and standard deviation of the coaster were calculated at each age. See Appendix F. This is the information used in the portfolio to calculate the premium.

Thirty coasters of different ages and types are randomly chosen and grouped together. The expected payout of these 30 coasters is the sum of every individual one, while the total variance is just the sum of each one as they are independent. The total fund needed to cover the portfolio of coasters was calculated by multiplying the standard deviation by 1.645 and adding it to the sum of the means. In formula notation is

Fund $= \sqrt{\sigma_{Total}^2} * 1.645 + E(x)_{Total}$, while $E(x)_{Total} = \sum_{n=1}^{30} x_n$ $\sigma_{Total}^2 = \sum_{n=1}^{30} \sigma^2{}_{x_n}$

All the coaster's independent based premiums were added together to calculate the total amount that would be paid without grouping. The total fund needed was subtracted from the total amount that would be paid if done independently. The extra money was the total savings created by developing this portfolio.

Next was determining a way to fairly distribute savings among all coasters. Several methods of sharing the savings were tried and compared. Two of them are illustrated to demonstrate the strength of the one that was chosen.

The first method way to calculate the premium for coasters being pooled was based on the coaster's individual premium. The ratio- P of the individual premium to the sum of all the premiums was applied to the total fund to be collected. The resulting amount was the coaster's final premium by pooling in the portfolio. See Appendix D. Below is formulas to summarize this method.

Individual Premium for each coaster= $E(X) + 1.645 * SD(X)$

Sum of Individual Premium= $\sum_{i=1}^{n}[E(X_i) + 1.645 * SD(X_i)]$

P= $\frac{\text{Individual Premium}}{\text{Sum of Individual Premium}}$

(E(X) is the expected payout and SD(X) is the standard deviation)

However, distributing the saving this way was not fair, because some of the coasters, especially those small ones result to have a premium even less than the premium if they had been pooled with 30 same coasters. For example, the S/W/1 coaster at age of 22 ended up with a premium of $565.8, which is less than $698.24 (premium pooled with 30 same coasters) in Appendix D. Standard deviation should be also taken into consideration. A coaster with a high expected payment and low standard deviation should not get much benefit from pooling risks.

It was decided to base savings off of standard deviation, because it generated a fair way to break it down to distribute to the different coasters. If a coaster does not have a standard deviation, it should not share any of the savings. The premium was calculated by proportion P times the total savings of the portfolio. The formula to calculate the proportion for coaster x is P = $\frac{\text{Standard deviation for coaster x}}{\text{total standard deviation}}$.

To compare the two methods, the premium amounts were compared to the premium a coaster would pay if they were grouped in a portfolio of thirty coasters, all identical. The second method produces a premium that's less than the identical portfolio does. This shows that it is providing a greater amount of savings to the rollercoasters and is therefore a better method. The last step is to get each coaster's premium which is calculated by taking the coaster's individual premium and subtracting its independent savings. See Appendix E

To check that this method of pricing was efficient, the ideal premium per coaster in a portfolio of 30 identical coasters was calculated. The premium would be less than when coasters were grouped with different types. With variation in the portfolio, the risk would be spread throughout the portfolio, increasing their premium from the ideal premium. By increasing the number of coasters in the portfolio, some of the coasters would end up with a premium amount that was less than the ideal premium (this was still keeping the identical portfolio with 30 coasters).

4.5 Expense Ratio

Expense ratio is the measurement "used by insurance companies to relate income to administrative expenses." [6] The formula used to calculate the value: ${}^{\text{expenses}}/_{\text{earned premiums}}$. The expense excludes losses, loss adjusting expenses and policyholder dividends. In the assumption of rollercoaster insurance, expense ratio is constant and around 20% as seen in the following.

By looking at the accounting statement (e.g. balance sheet) of some developed insurance company, the expense ratio is calculated for Sun Life[7] and The Hanover Group Insurance[8] respectively:

Consolidated Statements of Operations

YEARS ENDED DECEMBER 31 (in millions of Canadian dollars, except for per share amounts)

	2007	2006 *	2005 *
REVENUE			
Premium income:			
Annuities	$ 3,530	$ 5,380	$ 4,556
Life insurance	6,010	6,168	5,683
Health insurance	3,584	3,061	2,701
	13,124	14,609	12,940
Net investment income *(Note 5)*:			
Changes in fair value of held-for-trading assets	(1,558)		
Income from derivative investments	86		
Realized gains on sales of available-for-sale assets	127		
Other net investment income	6,197	6,664	6,079
	4,852	6,664	6,079
Fee income	3,212	3,014	2,899
	21,188	24,287	21,918
Operating expenses *(Note 15)*	3,260	3,028	2,921
Premium taxes	240	205	190
Interest expense *(Notes 10, 11 and 12)*	349	323	273

Figure 5 - Sun Life's Accounting Sheet

[6] http://insurance.cch.com/rupps/expense-ratio.htm
[7] http://www.sunlife.com/static/worldwide/Investor%20Relations/Static%20files/pa_e_Q407_2007_FinStats.pdf
[8] http://www.hanover.com/thg/investors/anreport/06ar/pdf/07148.pdf

CONSOLIDATED STATEMENTS OF INCOME

FOR THE YEARS ENDED DECEMBER 31	2006	2005	2004
(In millions, except per share data)			
Revenues			
Premiums	**$2,254.6**	$2,198.2	$2,288.6
Fees and other income	**74.9**	80.9	83.1
Net investment income	**318.9**	321.4	329.3
Net realized investment (losses) gains	**(4.3)**	23.8	16.1
Total revenues	**2,644.1**	2,624.3	2,717.1
Benefits, Losses and Expenses			
Policy benefits, claims, losses and loss adjustment expenses	**1,471.8**	1,703.1	1,646.7
Policy acquisition expenses	**477.5**	465.2	477.0
Restructuring costs	**1.6**	2.1	8.5
Other operating expenses	**413.8**	382.6	440.4
Total benefits, losses and expenses	**2,364.7**	2,553.0	2,572.6

Figure 6 - Hanover's accounting sheet

Because the two companies represent the two great business sectors in the insurance industry: Life and Property & Casualty, their Consolidated Statement of Income consists of quite different expense categories.

The specific formula of calculating the expense ratio for Sun life is:

$$\left(\frac{\text{Premium Income} * (\text{Operational cost} + \text{Interest Cost})}{\text{Total Income}} + \text{Premium Tax}\right) \Big/ \text{Premium Income}$$

There are two types of business in Sun Life, which are investments and insurance. Income from stock investments should be not counted as premium income. Therefore, the operational and interest cost from the investment income are excluded. As premium income occupies portion of total income, the cost for premium income should be the same portion of operational and interest cost. In essence, we assume that operational and interest costs are from two sources: investment and premium income. After adjusting the more accurate operational and interest cost, premium tax as another cost is added in. This way, the total expense from insurance business is divided by premium income according to the expense ratio calculation formula.

For the Hanover's expense ratio, it is calculated by:

$$(\text{Restructuring costs} + \text{Other operating expenses}) \Big/ \text{Premiums}$$

The expense ratio calculation for both companies involves only insurance business and does not take into consideration of stock and acquisition income and losses. As a new insurance product, costs mostly result from operation and tax. The expense ratios are shown in the following table:

	Sun Life	Hanover
2007	18.86%	18.42%
2006	15.20%	17.50%
2005	16.04%	19.61%
Average	16.70%	18.51%
Expense Ratio		17.61%

Table 12 - Annual Expense Ratios from Sun Life and Hanover

In the section of Price Portfolio, 20% was used as the constant expense ratio. The raw price is without the profit and expense. Dividing the raw price by (1-expense ratio), the business expenses would be factored into the rollercoaster's premium. Raw price equals (1-expense ratio)*final price, using Hanover's equation, instead of using (1+expense ratio)*raw price. Theoretically, there would be a 95% chance that the raw price would be enough to cover the losses occurred.

However, these expense ratios were not actually calculated because creating actual accounting sheets would necessary. Instead, an estimated expense ratio of 20% and profit ratio of 15% was used for all price calculations. The assumption of constant expense ratio and profit ratio is to consider about the time constraints of the project. To adjust the ratios as in the business world, it could be another new project.

4.6 Pricing Deaths

Pricing deaths is more difficult than injuries. The number of occurrences is far less and the data is difficult to come across and decipher. One of the biggest questions would be whether or not deaths that are caused by something other than the rollercoaster. For example, what if someone has a heart attack on the ride, but it was a precondition, or someone choked on a piece of candy that they had in their mouth? Would the policy cover that death? The courts tend to side with the deceased, and reward retribution to the family, no matter what the actual is.

The extremely low risk, high pay out, and these questionable areas are why it would make more sense to price deaths for a park-wide, overall coverage. This way any death, no matter what the cause, as long as it happened within the boundaries of the amusement park, would be considered a loss and would result in a payment.

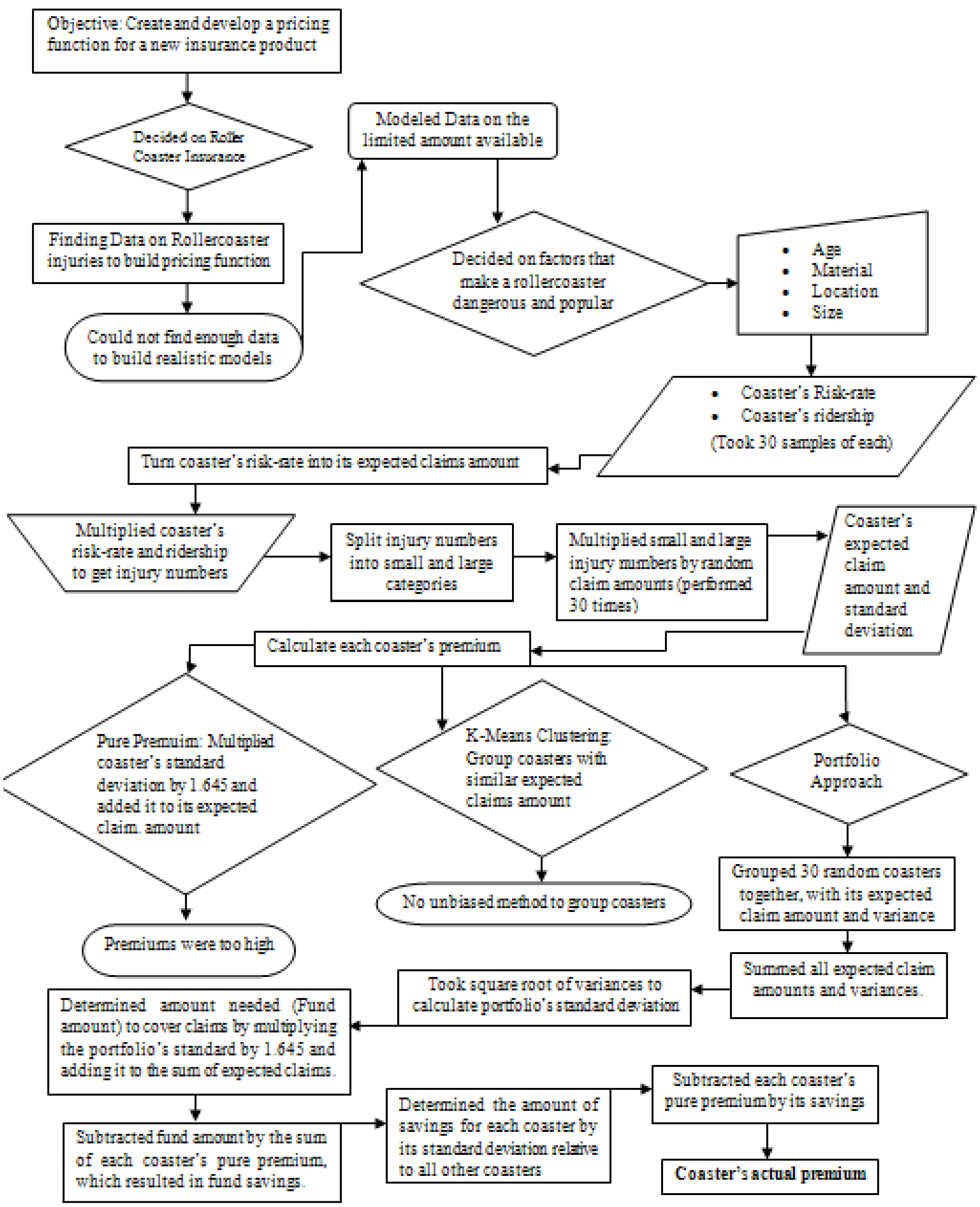
The court settlements for an unexpected death are usually in the millions. There may be some variation, but most claims are between one and two million dollars. There would be an upper limit on the policy of two million dollars.

The method to price a death coverage policy would be similar to the method used for injuries, but different factors would be used to determine the risk of a park. The safety score of the park and the number of visitors in a year would be the primary factors. This could be averaged out by the number of visitors to put the policy amount on a per visitor basis.

5. Conclusion

The following section will summarize, via flowchart, the process used to complete the project's objective and will conclude on how this project can be expanded upon for future projects.

5.1 Flow Chart of Project



5.2 Areas of future study

To make sure rollercoaster insurance successful, there are a few areas that would need to be further researched. Expense ratio and profit ratio were both held constant throughout this project, however, it would be necessary to actually go through and adjust these ratios as needed. Expense and profit ratio are directly linked to insurance company's cash flow and incomes. If the ratios are simply selected to be reasonable numbers, the company is not able to maximize profit or minimize losses.

Loss ratio is another area that would need to be looked into. Loss ratios observe the loss incurred versus premiums collected. The company wants to keep the loss ratio as low as possible without making the customers unhappy and satisfying the government regulation. However, the loss ratio cannot be too small; because the customers would believe that this insurance policy does not pay enough or over charge customers. The results would be that fewer and fewer customers buy this insurance product. On the other hand, it is crucial that the ratio is not too large that the company cannot afford the loss occurred. It would take time and research to ensure and adjust the loss ratio to optimize the sales revenue and minimize the payment.

To complete the rollercoaster coverage, death coverage needs to be refined. Even though basics of the coverage have been decided, specifying the death payment is essential, as well as clear coverage rules. This would allow for a reasonable price to be developed.

After rollercoaster insurance is developed as a product and has become a regular part of business, the company could focus on expanding coverage to new areas. While there are many potential clients to insure, it would be reasonable to stay within the areas during the early stage. Until the business grows into mature stage, exploring and expanding rollercoaster insurance would be the next stage.

Park-wide insurance is very similar to the death coverage already offered, and would therefore be a logical next step for the company to expand to. The policy would offer coverage to any injury that occurs within the park's borders.

Besides expanding coverage area, the parks that are currently covered are only stationary parks. A new product could reach out to the mobile carnivals. There are more risks involving when a ride has to be set up and taken back down frequently. It appears to be more error with the lack of stability. This coverage could be broken down by ride or by carnival-wide, depending on what the data showed the trend. The mobile amusement parks are quite different from fixed sites. They are getting more and more popular and the number of non-fixed amusement parks is becoming larger.

Water slides are also popular and could be a likely candidate to expand coverage. Combined with the rollercoaster insurance, the two most dangerous items are insured for the park. The water added a completely new element to the park and the types of rides that the park would have. This coverage would require the similar preliminary research as roller coasters, but would still be a large project to take on. It is likely that there would be a high demand for this coverage from the parks that are currently insured; especially most of them would also contain some water slides.

Expanding to these areas brings more profits to the insurance company and covers more entertaining and dangerous riders and slides for amusement parks. To offer policies on all these areas would make the company a convenient place for amusement clients and satisfy all their insurance needs.

Appenix A
State Safety Score Breakdown

State	Score
MA	30
NJ	29
PA	26
LA	25
NC	25
KY	23
MI	23
OH	22
OK	22
AR	22
CT	22
GA	22
MD	22
NH	21
WI	21
IL	21
RI	20
WV	20
AK	20
LA	19
WA	19
CA	19
ME	18
NY	17
VA	17
MO	17
NE	16
TX	15
FL	15
CO	15
IN	14
SC	14
HI	9
OR	9
DE	8
NM	8
ND	8
VT	8
TN	7
SD	6
ID	4
NV	4

Using Normal distribution

perctile	value	rounded value
0.25	12.63229507	13
0.5	17.19047619	17
0.75	21.74865731	22

Range	Color Code	Level	# of States in Level
(23,30)	BLUE	4	7
(18,22)	ORANGE	3	16
(14,17)	BROWN	2	9
(4,13)	YELLOW	1	10

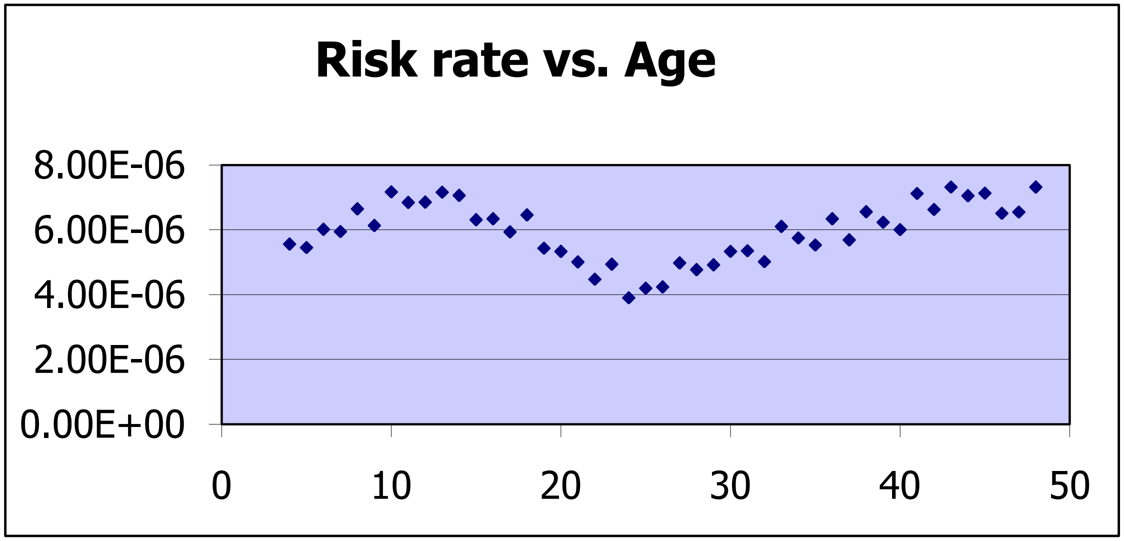
mean score	17.19048
stdev	6.757969

Appendix B
Input Sheet

INPUT ITEMS		
Lower limit for age at beginning of study	0	
Upper limit for age at end of study	50	
SAFETY RATING (Characteristic 1)	4 = 0-13	
SIZE (Characteristic 2)	3 = LARGE	
MATERIAL (Characteristic 3)	2 = STEEL	90.00% fuzzy factor

RISK FORMULA				
0.0000015	x	4	=	6.0E-6
0.00000025	x	3	=	750.0E-9
0.000000125	x	2	=	250.0E-9
inputs		SUM	v	**7.0E-6**

OUTPUT		
AGE	IDEAL RISK	FUZZY RISK
1	4.60E-06	4.84E-06
2	4.95E-06	4.47E-06
3	5.28E-06	5.18E-06
•	•	•
•	•	•
•	•	•
49	6.99E-06	7.29E-06
50	7.00E-06	7.17E-06



Appendix C
Clustering

Medium Wood 3		Large Steel 4	
Age	Expected Payout	Age	Expected Payout
25	$4,570.42		
26	$5,324.36		
●	●		
●	●		
●	●		
2	$9,943.24		
49	$10,093.60	41	$10,092.80
3	$10,174.09	25	$15,349.91
12	$10,590.37	29	$18,742.92
15	$10,697.70	24	$19,227.92
11	$10,971.42	31	$19,809.07
10	$11,967.59	26	$19,975.42
9	$13,304.00	22	$20,026.46
13	$13,376.73	32	$20,166.57
8	$14,171.37	33	$20,530.04
		46	$20,927.53
		30	$21,095.99
		●	●
		●	●
		●	●
		12	$37,715.89
		11	$40,342.10

Cluster	
48	$9,800.39
5	$9,903.54
14	$9,938.25
2	$9,943.24
41	$10,092.80
49	$10,093.60
3	$10,174.09
12	$10,590.37

Color Key
LS4
MW3

E(X)=	$10,067.03	expense ratio=	0.16	profit ratio=	0.15
SD(X)=	$243.91	premium p=	a*E(X)*n		
if n=	2	premium p=	$56,150.62	a=	2.79
if n=	10	premium p=	$3,253.77	a=	0.03

Appenix D

Portfolio Development (Prem.)

Coaster	Age	Mean	STD	Var	Ind. Premium	Ratio of Prem.	Savings	Premiums
LS1	34	$ 13,028.97	$ 4,419.37	$ 19,530,869.00	$ 20,298.841	0.044961571	$ 5,290.09	$ 15,008.76
LS4	24	$ 14,377.93	$ 4,523.19	$ 20,459,254.62	$ 21,818.581	0.048327768	$ 5,686.15	$ 16,132.44
LW1	14	$ 8,789.31	$ 2,763.78	$ 7,638,488.82	$ 13,335.735	0.029538415	$ 3,475.43	$ 9,860.31
MS2	43	$ 4,128.29	$ 1,525.16	$ 2,326,108.81	$ 6,637.171	0.014701216	$ 1,729.71	$ 4,907.46
MW3	3	$ 8,957.50	$ 2,867.74	$ 8,223,916.36	$ 13,674.931	0.030289729	$ 3,563.83	$ 10,111.10
●	●	●	●	●	●	●	●	●
●	●	●	●	●	●	●	●	●
●	●	●	●	●	●	●	●	●
SW1	31	$ 683.28	$ 49.82	$ 2,482.17	$ 765.235	0.001694982	$ 199.43	$ 565.81

Prem w/ 30 same	Prem. - 30 same
$ 14,356.26	$ 652.49
$ 15,736.40	$ 396.03
$ 9,619.37	$ 240.93
$ 4,586.34	$ 321.11
$ 9,818.78	$ 292.32
$ 698.24	$ (132.43)

Sum of Ind Means	$ 299,304.59
Sum of Ind STDs	$ 92,502.32
Sum of Ind Premiums	$ 451,470.907
Total Var.	$ 440,065,651.52
Total STD	$ 20,977.74
Fund total	$ 333,812.98
Total Savings	$ 117,657.928
Number of Coasters	30

Appenix E

Portfolio Development (STDs)

Coaster	Age	Mean	STD	Var	Ind. Premium	Ratio of STDs	Savings	Premium
LS1	34	$ 13,028.97	$ 4,419.37	$ 19,530,869.00	$ 20,298.841	0.035286544	$ 5,998.62	$ 14,300.219
LS4	24	$ 14,377.93	$ 4,523.19	$ 20,459,254.62	$ 21,818.581	0.036115468	$ 6,139.54	$ 15,679.044
LW1	14	$ 8,789.31	$ 2,763.78	$ 7,638,488.82	$ 13,335.735	0.022067446	$ 3,751.41	$ 9,584.326
MS2	43	$ 4,128.29	$ 1,525.16	$ 2,326,108.81	$ 6,637.171	0.012177646	$ 2,070.17	$ 4,567.003
MW3	33	$ 6,861.04	$ 2,133.56	$ 4,552,085.75	$ 10,370.752	0.017035448	$ 2,895.98	$ 7,474.769
●	●	●	●	●	●	●	●	●
●	●	●	●	●	●	●	●	●
●	●	●	●	●	●	●	●	●
SW1	6	$ 647.36	$ 201.71	$ 40,686.28	$ 979.166	0.001610543	$ 273.79	$ 705.378

Prem w/ 30 same	Prem. - 30 same
$ 14,356.261	$ (56.042)
$ 15,736.403	$ (57.359)
$ 9,619.374	$ (35.048)
$ 4,586.343	$ (19.341)
$ 7,501.825	$ (27.056)
$ 707.936	$ (2.558)

Number of Coasters	30
Sum of Ind Means	$ 299,304.59
Sum of Ind STDs	$ 92,502.32
Sum of Ind Premiums	$ 451,470.907
Total Var.	$ 440,065,651.52
Total STD	$ 20,977.74
Fund total	$ 333,812.98
Total Savings	$ 117,657.928

Appendix F
Calculating Mean and Standard Deviation

Large Steel with safety score level 4

Age	Coaster #	Risk Rate	Sm. Injuries	Lg. Injuries	Sm. Injuries $	Lg. Injuries $	# of Riders	Expected Payout		
1	1	5.11129E-06	0.62	0.38	441.0	3438	1,962,012	$ 15,843.49		
1	2	5.08358E-06	0.8	0.2	792.0	3484	2,087,856	$ 14,120.58		
1	3	4.30631E-06	0.74	0.26	256.0	5179	2,271,070	$ 15,021.77		
●	●	●	●	●	●	●	●	●		
●	●	●	●	●	●	●	●	●		
●	●	●	●	●	●	●	●	●		
1	29	4.90204E-06	0.8	0.2	645.0	3100	2,201,243	$ 12,258.09	Mean of Age 1	$ 18,775.24
1	30	5.19293E-06	0.62	0.38	92.0	4376	2,299,734	$ 20,539.91	Std of Age 1	$ 6,867.42
2	1	4.35935E-06	0.73	0.27	764.0	2180	2,247,232	$ 11,229.89		
2	2	5.02071E-06	0.69	0.31	779.0	4242	2,040,712	$ 18,980.70		
2	3	5.45463E-06	0.62	0.38	336.0	3270	2,045,685	$ 16,190.01		
●	●	●	●	●	●	●	●	●		
●	●	●	●	●	●	●	●	●		
●	●	●	●	●	●	●	●	●		
2	29	5.44416E-06	0.73	0.27	73.0	5909	2,060,662	$ 18,496.30	Mean of Age 2	$ 18,842.25
2	30	5.33468E-06	0.68	0.32	778.0	4258	2,204,595	$ 22,246.73	Std of Age 2	$ 5,558.04
3	1	4.91494E-06	0.71	0.29	367.0	5338	2,075,189	$ 18,446.59		
3	2	5.79832E-06	0.66	0.34	272.0	4133	2,230,225	$ 20,493.15		
3	3	5.8383E-06	0.57	0.43	968.0	3842	2,004,493	$ 25,790.95		
●	●	●	●	●	●	●	●	●		
●	●	●	●	●	●	●	●	●		
●	●	●	●	●	●	●	●	●		
3	29	5.43634E-06	0.59	0.41	316.0	3738	2,121,424	$ 19,825.10	Mean of Age 2	$ 20,071.52
3	30	5.27184E-06	0.74	0.26	710.0	3710	2,185,521	$ 17,167.37	Std of Age 2	$ 7,493.78
					●					
					●					
					●					
50	1	7.2646E-06	0.55	0.45	693.0	4588	1,501,808	$ 26,683.23		
50	2	7.89227E-06	0.79	0.21	103.0	5608	1,589,457	$ 15,794.06		
50	3	7.21051E-06	0.8	0.2	370.0	4322	1,490,296	$ 12,469.41		
●	●	●	●	●	●	●	●	●		
●	●	●	●	●	●	●	●	●		
●	●	●	●	●	●	●	●	●		
50	29	6.94423E-06	0.58	0.42	197.0	5329	1,598,925	$ 26,119.84	Mean of Age 50	$ 18,395.86
50	30	7.40768E-06	0.68	0.32	489.0	2926	1,600,368	$ 15,042.12	Std of Age 50	$ 5,434.50

Equations
Small Injuries=RANDBETWEEN(50,80)/100
Large Injuries=1-Small Injuries
Small Injury $=RANDBETWEEN(50,1000)
Large Injury $=RANDBETWEEN(2000,6000)
Expected Payout=Small Injury $*(Small Injuries*(Risk Rate*# of Riders))+Large Injury $*(Large Injuries*(Risk Rate*# of Riders))