Team #G002

30th October 2025

Summary

Our algorithm to determine the Top 10 Roller Coaster List works through a three-step process. First, the criteria that we have deemed important are identified. Secondly, each criterion is assigned a value based on its importance. Each value of importance is then multiplied by how well each roller coaster can fulfill the said criteria. For example, a height of over 150 feet might be something that we find important. We could assign that criteria a value of 8 (out of 10). For example, one roller coaster in Russia, Velikolukskiy Myasokombinat, is 124 feet tall. While it isn't 150 feet tall, it gets close and thus could receive a 3 (out of 5). For this criterion, the roller coaster would receive a 24. Our algorithm performs this process for all the criteria we identify and sums up the values. The roller coaster with the highest sum will be considered the best, and likewise, the lowest sum the worst.

Top ten:

Ride Name	Rounded Point Value
Kingda Ka	41
Top Thrill Dragster	39
Steel Dragon 2000	38
Leviathan	36
Fury 325	36
Intimidator 305	36
Superman: Escape from Krypton	35
Millennium Force	35
Tower of Terror II	35
Coaster Through the Clouds	34

Team #G002

Introduction and Problem Statement

Rollercoasters are all very different and have unique, stunning qualities that set each one apart. However, some are significantly superior to others, and the users need to be aware before purchasing their ticket and travelling all the way. Can we build a mathematical model to objectively rank rollercoasters based on their recorded qualities, such as speed and height?

Assumptions and Justifications

This mathematical primarily assumes that certain qualities are more impactful than others, and that there is an overwhelming objective ranking to satisfy much of the population. This model also assumes that the objective status of a ride can be determined by quantitative factors, and that qualitative factors can be converted to a quantifiable form such as an index score or proportion.

Here's our personal ranking of the most important criteria:

1. Height- The maximum height the roller-coaster reaches relative to the ground
2. Speed- The maximum speed the rollercoaster reaches during its travel
3. Drop- The distance of a vertical drop, if applicable
4. G Force- How many g's experienced during the ride
5. # of Inversions- How many times the roller coaster takes you upside-down
6. Duration- How long the ride (usually) lasts
7. Vertical Angle- Highest vertical incline
8. Length- The length of the track (scalar, not displacement)
9. Type of Roller Coaster
10. Material (wood or steel)

This ranking was used to determine how valuable each criterion was. We looked at what current rankings believe is the best, and just general background knowledge about what makes a roller coaster great or memorable.

We also assumed that the further you get away from the ideal value, the worse your score will become, and that this process happens linearly. This decision was made sense coming up with non-linear (maybe exponential) equation for how much worse a roller coaster is based off distance from an ideal value would just add even more subjectivity and more assumptions to our already assumption heavy model.

Another assumption was that the highest value is the best one. For example, the tallest roller coaster would get the best score for height, and the fastest roller coaster would get the best score for speed. This was decided as we couldn't think of a reasonable scenario where an average roller coaster enjoyer would prefer a (for example) shorter roller coaster over a taller one, other than if the taller one reached up until space or some unrealistic scenario like that. For the roller coasters in the dataset, where no such thing occurred, we assumed the bigger value the better.

We also assume when there is no datapoint for a quality withing a roller-coaster, we can replace the point with the median value from those scores. This is justified because a roller-coaster should not be penalized for lacking a datapoint, as the roller-coaster cannot control who is recording it and what they are recording for.

Model and Solution

For an initial model, we had to make a way to evaluate a roller coaster's "amusement" given a wide selection of different forms of data. Aside from needing to be objective, and therefor take into account only the given information, more criteria for this initial model was that it needed to

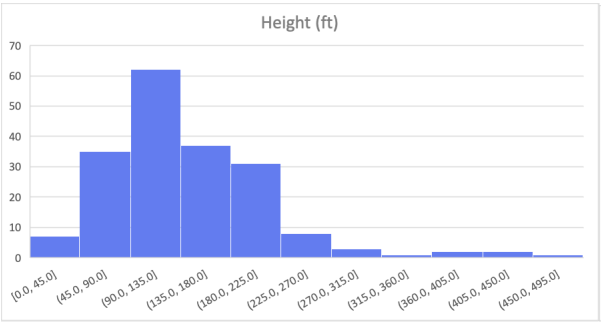
function with some missing data points, as many categories were only provided by some rollercoasters, and that it needed to not rely on any overly complicated strategies. This is because, in order to calculate a top ten list of roller coasters, the model must not rely on any human work to make any decisions, as it would make the process of evaluating every roller coaster take too long.

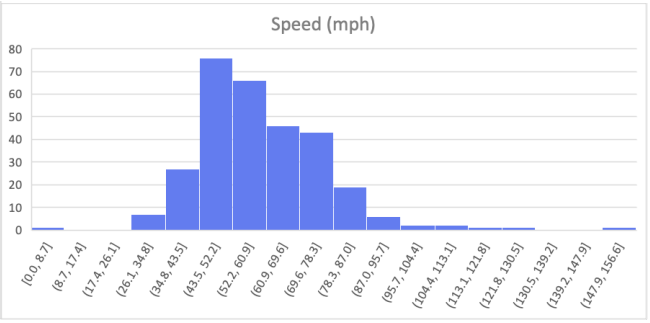
As a base of our model, we decided to use a strategy in which each variable would correspond to a different value and weight, where non-numerical data was substituted with determined value ratings while numerical data directly fed into the value. This was done because for most variables it was determined that having a greater value was always better than having a lesser value. After this step of inputting values, each value would then be multiplied by a number determining the importance of the value, and all of the resulting values would be summed to create a final point value for one roller coaster. Below is a table of the values decided for every category, depending on the type of data and the perceived importance of each category. Boxes are color-coded to better show how every value number links to a specific input:

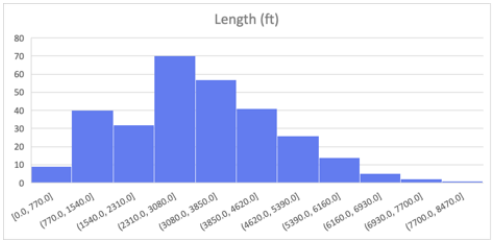
			Category:	Material	Type	Height	Speed	Length	Number of Inversions	Drop	Duration	G Force	Vertical Angle
			Weight of category:	0.1	0.4	1	0.9	0.2	0.7	0.9	0.7	0.8	0.5
Values (Including category weights):	Value Multiplier:	Flying point value:	Steel point value:	1	4	2/99	0.057471264	0.000236128	0.5	0.028125	0.019444444	1.538461538	0.041322314
		Stand up point value:	Wood point value:	0	2.4								
		Inverted point value:			2.8								
		Suspended point value:			1.2								
		Wing point value:			2.8								
		Sit down point value:			0								

This table is the most subjective aspect of the model, as much of the determination of point scale, category importance, and data ranking can only be evaluated through consensus.

To solve the problem of gaps in data, we decided to make the model assume that the chosen rollercoaster would have had the median data value of all roller coasters with data. This option was deemed most optimal because any other action such as assuming a zero or invalidating the entire roller coaster ends up either overly punishing the roller coaster for missing data or going against the purpose of the model by removing a large number of viable roller coasters from the dataset. Of the ways to calculate a general value for each data type, median was deemed best because all charts of the data showed that the data was skewed left. In this case, mean would have given higher values, which seemed unnecessary because there was no way of knowing the true value, so making a lower estimate is more reasonable. This makes more sense when considering how amusement parks may respond to the scoring system, in which it is better to choose an approach that more penalizes missing data in order to encourage the collection of data. Below are some examples of the data for various categories, in which all catagories seem to be skewed right (right skew indicates outliers, which are known to affect the mean):







These decisions were then made into equations in excel, building onto the already provided excel data. For the non-numerical data, IF statements had to be used to assign values to each possible data point, while the numerical data was simply multiplied to the value multipliers shown in the table shown previously. After this process was repeated across the span of the dataset, a pivot table was created to extract and order the roller coasters by score. Below are the top ten values from this pivot table, with scores rounded to the nearest whole number.

Ride Name	Sum of Point Value
Kingda Ka	41
Top Thrill Dragster	39
Steel Dragon 2000	38
Leviathan	36
Fury 325	36
Intimidator 305	36
Superman: Escape from Krypton	35
Millennium Force	35
Tower of Terror II	35
Coaster Through the Clouds	34

From this, it is shown that the very best roller coaster is the Kingda Ka, in the Six Flags Great Adventure amusement park in New Jersey, US.

App Design

User Input

Upon opening the application, the user will be prompted to select criteria for roller coasters which suit their own preferences.

Preferences include:

- Maximum speed of roller coaster
- Roller coaster height
- Roller coaster length
- Number of inversions
- Drop height
- Type
- Location

App Algorithm

The algorithm that our app uses is very similar to the original, with one key change. Each criterion is still assigned a value based on its importance. Each value of importance is then multiplied by how well each roller coaster can fulfill the said criteria. However, the key difference here is that the user gets to decide how important each criterion is.

The algorithm that our app uses is designed to assign a desirability score out of 5 marks to different ranges of magnitudes of a user's selected preference. The process can be defined as following:

1. Criteria evaluation: Each of our criteria will be presented to the user, and they will be able to input a value from 0-10 based on the importance that they assign that criterion with the help of a slider.
2. Range Segmentation: The full spectrum of the selected criterion (e.g. 0-100 mph, 0-100 ft) will then be divided into five ranges of equal intervals.
3. Scoring: The user will then be prompted to select a specific range from the five intervals to represent their most preferred value for the chosen criterion. This method is used instead of assuming that greater values are better because it allows for a greater subjectivity, adapting more easily to a user's known preferences.
4. The selected range is then assigned a maximum score of 5/5.
5. All other intervals are assigned a score based on their proximity to the desired range. The score decreases for each interval away from that range.
6. Evaluation: This score is then multiplied by the value assigned to the criterion by the user in step one, giving that rollercoaster a particular score for that criterion.
7. The score for each criterion is summed up for each coaster, and the roller coasters with the top ten sums will be returned to the user as the best fit for them.

Each selected criterion for the roller coasters can be evaluated this way. For multiple criteria, the separate scores can be summed or weighted to obtain a total preference score. This final score shall then be utilized to sort the Top 10 roller coasters that align with the users' preferences.

Output

Based on the user defined parameters and preferences, the app will input these values into our developed algorithm and display a top ten list of roller coasters that cater to the user's preferences. The user will be able to leave a review if/when they do choose to go to these roller coasters, which can contribute to later refinement of the algorithm.

Analysis

Our model can determine how good a roller coaster is in comparison to other coasters in a way that can be highly objective and consistently follows a set of criteria. It is formulated in a way such that if someone disagrees with our ranking, they simply disagree with the weight we have assigned each criterion, which is an issue that our app solves for. This model will change the way roller coasters are ranked and how people decide to go to roller coasters, shifting from decisions based on subjective, exaggerated reviews to a more objected, indisputable conclusion.

Strengths:

- The method of using pre-set weights and point values allows for a high level of objectivity across roller coasters.

- Easily understandable and modifiable
    - This is even more seen in the plan for the app, in which there will be a user interface that allows for easy manipulation of weights and preferences.
- It works for a wide variety of coasters, adapting to the potential absence of some data points.
- The method is highly adaptable and is able to accept the addition of more sources of data.

Limitations:

- Currently, any weightings are determined solely by us, possibly misrepresenting the general population.
- No research was done to determine if there are certain values that are most enjoyable.
    - Instead, higher value was evaluated to be better, which may not always be true.
- The point drop-off for not reaching a certain value is linear, which is not always realistic

Conclusion

An application that utilizes a two-step algorithm to rank/create a list of the Top 10 roller coasters was successfully developed, directly meeting the need to create an objective ranking system. The implementation of a weight-based system based on selected criteria successfully generated a list of the Top 10, with Kingda Ka placed in first. The success of this system is due to the application's algorithmic ability to take personalized criteria into account when determining this list, allowing a value of importance (0-10) to be assigned to preferred ranges of quantities of the

selected criteria. The calculation of point-value scores based on such user inputs allows for the fulfilment of a customized ranking system. As for future directions regarding this app, continuous feedback for this system could be integrated. Furthermore, the qualitative factors could be converted to numerical scores. The interval ranges for the weighted system could be user-defined instead of fixed intervals as well, to allow for further customizability.

News Release

Massachusetts high school students announce a novel real-time rollercoaster ranking application.

31st October 2025: Recently, some students have developed an app that uses a novel algorithm to objectively determine a list of the Top 10 best roller coasters in the world. But here is what is unique – this app compiles this list based on user input, meaning that the list is based completely on user preference. This is great for especially those who value customizability!

This new algorithm is comprised of two components: the user input and application algorithm. The user input section essentially collects data from the user regarding types of criteria that they would find relevant to making their decision. From this point forth, the app algorithm takes over.

References

- Microsoft Excel

AI Use Report

AI or any assistive technology was not used apart from Excel tabulation and simple functions.