Orlando Theme Park Statistics

Pete Way

Spring 2019

Introduction

The entertainment sector provided by the Disney and Universal companies and their properties is an especially unique implementation of Data Science research, while a variety of vocations working tirelessly to treat guests of the attractions to an exciting visit each and every time. There are countless applications that Data Science visualizations could assist with in this sector, both in preventing losses or damages as well as boosting returns and revenue. Hundreds if not thousands of data scientists are already employed by these two entertainment giants, with countless more voluntarily dedicating their time to peek into any available data source (the author of this report included). However, the author expressed regret that due to the brief space available, this report details only a few examples of these applications and available data sources.

Please be aware that all visualization and coding completed within this report was achieved via the use of R Studio, and the R programming language. Native commands such as *plot()*, *barplot()*, *hist()*, and *png()* were utilized, as well as a variety of sorting, refining, and pre-processing techniques. Specific libraries that required installation included:

- Rio - Provided ability for reading in multi-sheet Excel files
- GGPlot2 - Provided more streamline and cohesive visualizations

- GGMap - Extension of GGPlot2, specifically for retrieving Google/Stamen Map bases.
- Scatter3D - Provides 3-dimensional plotting abilities. (Used here to provide color as Z dimension)

Any questions regarding the work included in this report may be sent to KXC378@mocs.utc.edu, and the author would gladly respond accordingly.

1 Exploration of Employed Data Sets

This paper includes data retrieved or provided via several sources. Any work in this report is indebted to those individuals and teams who took time out of their schedules to provide the data explored from this point one. Each of these data sets will now be briefly introduced and explained, before their use later on in their independent sections.

Data regarding the spatial layout of the two Orlando based theme park brands was graciously received from The Park Database [2]. However, this data set is so much more robust than only location. The entire set of variables used from The Park Database is shown in Table 1, with sections for each sheet of the excel file received.

The Park Database data set includes attendance counts for the various parks for the time frame from 1990 to 2017 in the 'Attendance' sheet of the provided excel file. Ticket prices were provided for the Universal and Disney parks in the 'Tickets' sheet, although they were originally incomplete. The author of this report used another well known Disney resource (All Ears) to fill in any blanks in the price data for Disney properties [3]. Unfortunately, a similar resource could not be found for Universal properties, and as such those values remained as they were. The 'Parks' sheet provided area information, as well as cost, construction year, and other interesting fields for each of the unique park properties.

Due to the extensive nature of this data set, this data set will be explored in Section 2 in regards to the parks locations, as well as Sections 3 and 4 for Financial and Attendance

statistical analysis, respectively.

Additionally information regarding individual rides was provided by Ms. Lynne Passanisi, of ECCO USA [7]. Ms. Passanisi originally posted this data set to the Data World in December of 2018, and described the set as a result from research for her Masters Capstone Project. This data set originally included those variables presented in Table 2. Any variable presented with an asterisk was provided in binary fashion.

This data set was expanded upon by the author with the addition of latitude and longitude columns for the provided rides, as well as re-hauling the 'Park Location' column into a string format displaying the entire name of the park instead of the initials, and as such was renamed 'Park Name'. This data set will be combined with the previously mentioned Park Database set in Section 2.

Data regarding reported injuries within Orlando area parks was originally available for download in PDF form from a Florida government agency web page [5]. However, Dr. Jegar Pitchforth personally retrieved the data set via website scraping and clever use of code, and as such will be considered the source of said injury data for the purposes of this report. This data set covered the time frame from 2003 to 2017, and originally featured more properties other than Disney and Universal owned. The variables presented by Dr. Pitchforth's data set are presented in Table 3. This data set will be explored further in Section 5.

The final data set used by this report will be information retrieved from a study performed by Go Banking Rates [6]. This is a simpler data set, only featuring three different variables as shown in Table 4, these being 'Year' (the year the park opened), 'Cost' (the original advertised ticket price for the given year), and 'Adjusted Cost' (original cost adjusted to the inflation rates of 2017). All three variables are presented as numeric. This data set will mostly be used for support in analysis across the entirety of the report, but is of particular use to Section 3.

Table 1: Park Database Data Set Overview

Rides		Parks		Tickets		Attendance	
Variable	**Type**	**Variable**	**Type**	**Variable**	**Type**	**Variable**	**Type**
Park Code	numeric	Park Code	numeric	Park Code	numeric	Park Code	numeric
Ride Code	numeric	Park Name	text	Ticket Code	text	1990	numeric
Ride Name	text	Size	numeric	Segment	text	..	
Ride Type	text	Cost	numeric	2009	numeric	2017	numeric
Ride Sub Type	text	Type	text	...			
Comments	website URL	Brand	text	2019	numeric		
Cost	numeric	Year Built	numeric				
Hourly Capacity	numeric	Latitude	numeric				
Interval Time	numeric	Longitude	numeric				
Total Seats	numeric						
Year	numeric						

Table 2: Passanisi Data Set Variable Overview

Ride Name (text)	Ride Type Thrill *
Park Location (text)	Ride Type Spinning *
Park Area (text)	Ride Type Slow *
Fast Pass *	Ride Type Small Drops *
Open Date (date)	Ride Type Big Drops *
Height Requirement (inches)	Ride Type Dark *
Ride Duration (Minutes)	Ride Type Scary *
Age of Ride (years)	Ride Type Water *
Age Interest PreSchooler	Age Interest Kids *
Age Interest Tweens *	Age Interest Teens *
Age Interest Adults	Classic *

2 Location

Before delving into the deeper aspects of the introduced data sets, one must have an understanding of where these parks are, and how their locations relate to one another. Therefore, the first aspect of the data sets provided to be explored will be location. As shown in Figure 1, both Universal and Disney properties exist in rather close proximity to one another. Note that this map includes the now defunct 'Disney Quest' arcade styled attraction in addition to the properties more traditionally considered as theme parks.

A deeper look into the area of the Disney Theme parks produced Figure 2a. This particular map demonstrates the locations of the rides within, through the use of the Passanisi data set's GPS coordinates. A more concrete count of the number of rides per property is shown in Figure 2b. As shown in the map and bar graph, Magic Kingdom is furthest away from the other parks, as well as the most heavy populated with rides.

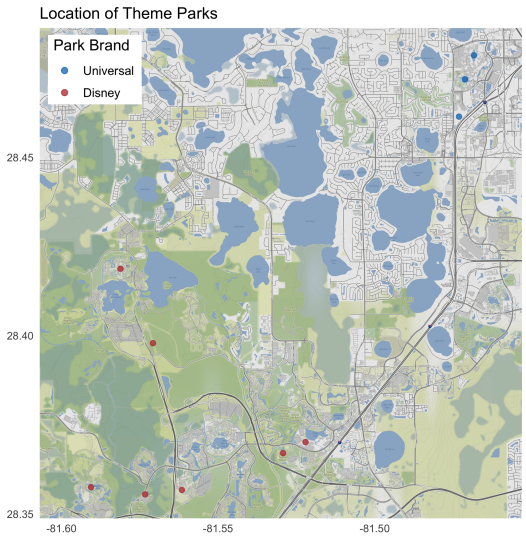
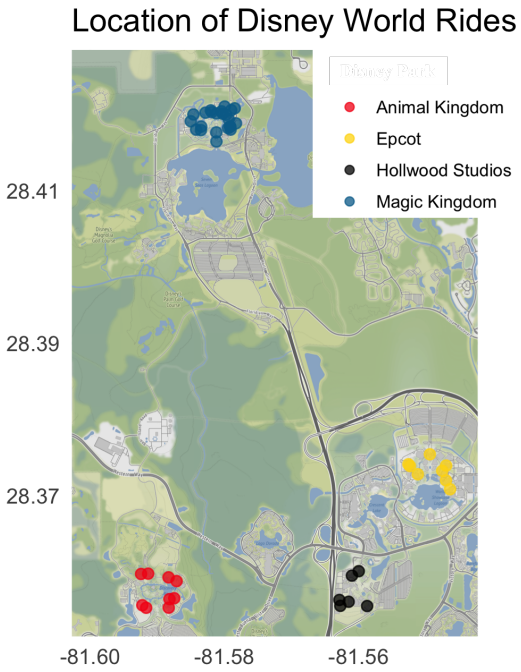
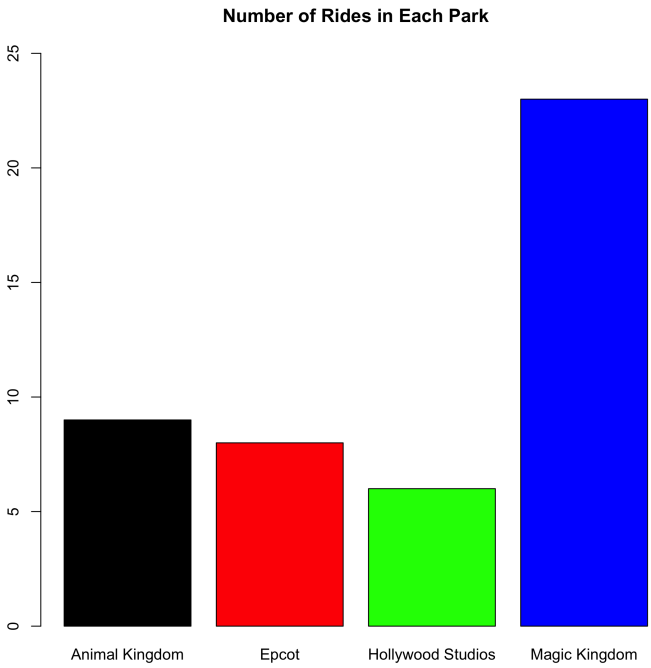


Figure 1: Elementary Map of the Orlando Area and Parks Found Within

Unfortunately, locational information for individual rides in the Universal Studios properties was unavailable. However, Figure 3a displays the aerial view of the Universal Parks while Figure 3b shows the ride count at each of the three locations, giving a vague idea of the entertainment options at each location.

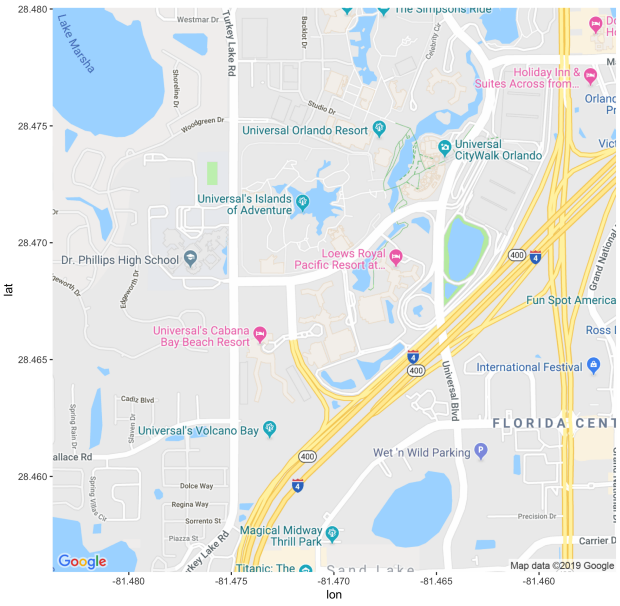


(a) Disney Park Map

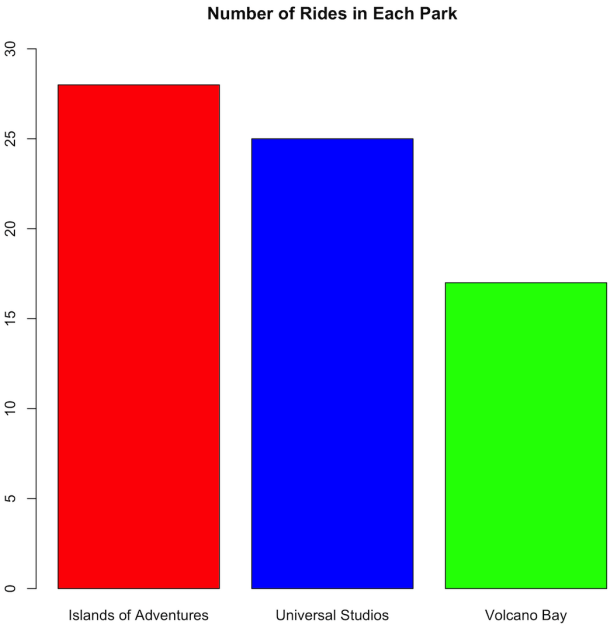


(b) Rides per Disney Park

Figure 2: Disney Ride Locations and Ride Counts



(a) Aerial Map of Universal Studios Properties



(b) Rides per Universal Park

Figure 3: Universal Parks Locations and Ride Counts

However, area information does exist for all of the properties in Orlando. As shown in Figure 4, Disney properties take the top three positions in terms of area. Animal Kingdom, the top ranking park in terms of area, covers an amazing 1,174,422 square meters, which equals to nearly a half mile. To further explore this area size it must be stated that according to The Park Database, the source for this data, this estimate only includes the park area accessible to guests. Therefore, it is to be stated that animal habitat areas are not included and the actual area covered by Animal Kingdom may well be tremendously larger than this estimate. The second largest area goes to Epcot, with 822,070 square meters. This equals to about a third of a mile. It is truly amazing to think about all of the entertainment options that Disney has managed to fit within these relatively compact spaces.

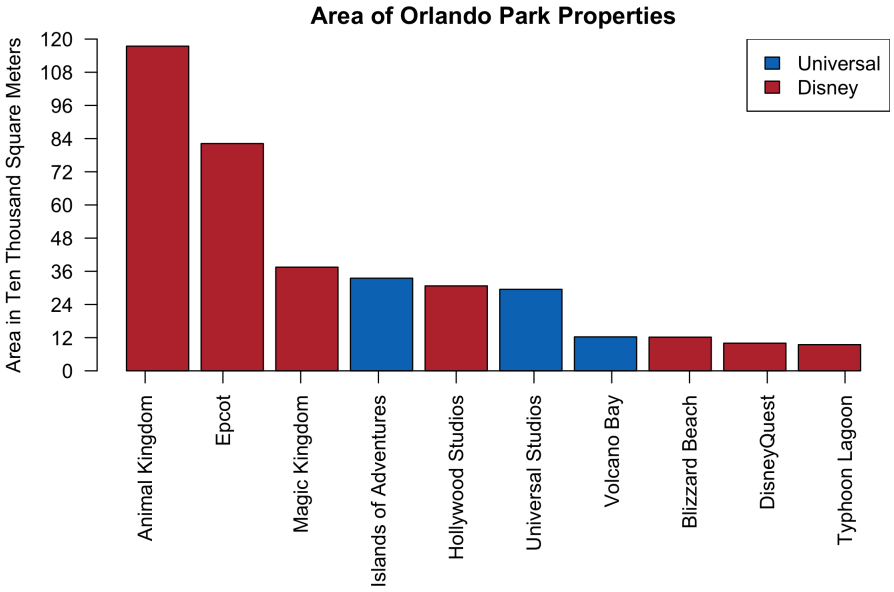


Figure 4: Area Comparison of Orlando Parks

3 Construction and Finances

Now that an elementary understanding of the area the properties in discussion inhabit exists, we can begin discussion of the financial questions one might have from these massive projects. To begin the exploration through Disney and Universal park financials, we must examine the construction costs. As displayed in Figure 5, these properties involved massive amounts

of investment. Please note that construction cost estimates were not able to be reliably sourced for both Typhoon Lagoon or Blizzard Beach (Disney's two water park properties in Orlando). To coincide with Figure 5, Table 5 shows the year that each property was premiered. Note the relatively large gap between the construction of Magic Kingdom and Epcot. This period of 11 years is the largest gap between any two construction dates aside from Islands of Adventure and Volcano Bay (18 years).

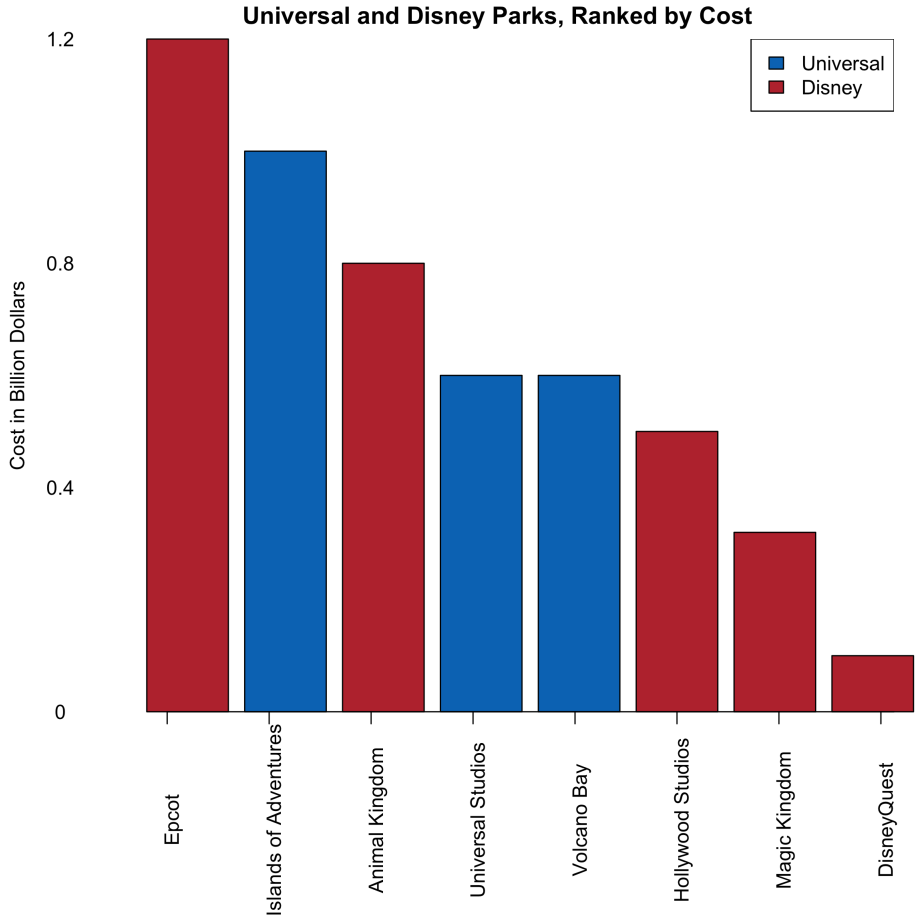


Figure 5: Construction Costs of Orlando area Disney and Universal Brand Properties

In addition to the construction costs, one must analyze when rides were premiered in the parks. As shown in 6, the highest number of rides premiered at Magic Kingdom at its inception. A a matter of fact, Magic Kingdom was the only property in the Orlando area until 1982, when the most expensive property Disney built in the Orlando area opened. Although Epcot is the highest area property, its main draw is not theme park styled rides.

Table 3: Pitchforth Injury Data Set Overview

Variable	Type	Description
Date	month/day/year	Date the injury occurred
Age	numeric	Age of the injured guest
Gender	binary	Gender of the injured guest (f/m)
Ride	text	Name of the ride guest was injured on
Condition	binary	Did the guest have a pre-existing condition?
Year	numeric	Year the injury occurred

Table 4: Go Banking Rates Data Set Variables

Variable	Description
Year	Year tickets sold
Cost	Cost of tickets for that year, as advertised
Adjusted Cost	Adjusted cost based off 2017 inflation

Table 5: Opening Years of Orlando Properties

Park	Opening Year
Magic Kingdom	1971
Epcot	1982
Typhoon Lagoon	1989
Hollywood Studios	1989
Universal Studios	1990
Blizzard Beach	1995
Animal Kingdom	1998
DisneyQuest	1998
Islands of Adventures	1999
Volcano Bay	2017

Therefore, no major increase of rides occurred in the 1980s. This trend of low number of premiers persisted until the later half of the 1990s, with the opening of both Blizzard Beach and Animal Kingdom, the second highest area property Disney owns in Orlando.

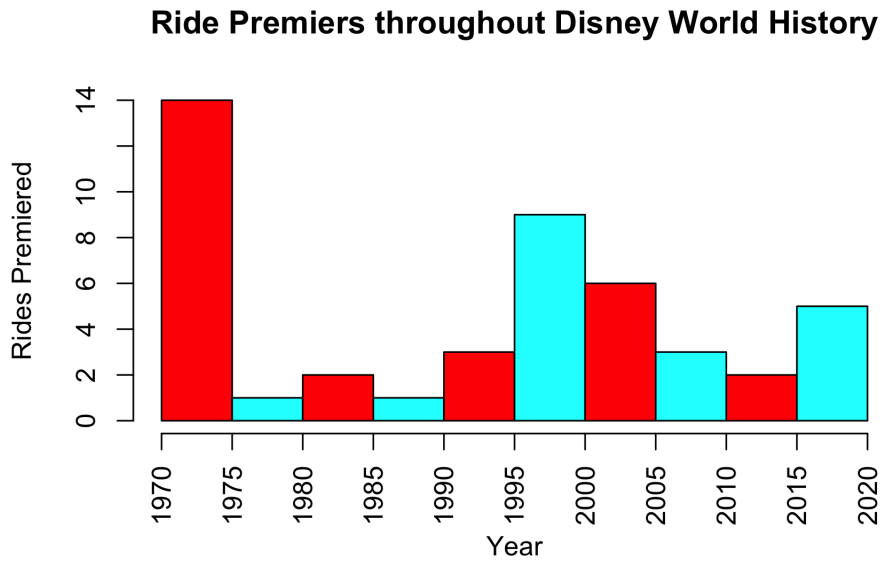


Figure 6: Ride Premiers in Disney World

The oldest and newest rides of Disney properties in Orlando are presented in Figures 7 and 8, respectively. Figure 7 reflects to twenty oldest rides in the parks, with the first 10 all opening together on Magic Kingdom's rope-drop day. Due to this large number of rides all opening at once, the normal 'Top Ten' format was restructured for this instance. As a matter of fact, only four of the oldest twenty rides premier inside other parks. One can even gauge the opening day of Epcot from this list, with the two original rides being listed at 16th and 17th oldest.

Figure 8 portrays the opposite end of the spectrum. This figure displays the ten youngest rides of the Orlando Disney properties. Note that in this figure, all four theme parks are represented. The two youngest rides shown are 'Alien Swirling Saucers' and 'Slinky Dog Dash', the two opening day rides for the 'Toy Story Land' addition to Hollywood Studios that rope-dropped April 26, 2018. This means the youngest rides in Disney World just became one year old (as of this report's completion). The next duo also reflect a park addition, this

time 'Pandora', the Avatar themed land from Animal Kingdom. This addition rope-dropped on May 27, 2017. The effects of these additions on theme park attendance will be examined further into this report, inside of Section 4.

Additionally, one must consider the age of the rides contrary to their duration. Do older rides tend to have a shorter or longer duration? This question led to the creation of Figure 9. As shown in the figure, many of the rides in Magic Kingdom tend to be much older, with many having a duration of 15 minutes or less, with two outliers having a duration of above 20 minutes. The two youngest rides overall, as discussed above, reside within Hollywood Studios, and have very short duration, only a few minutes each. Overall, much of the rides from Animal Kingdom are between 10 to 20 years old, with duration up to 5 minutes aside from two outliers. The oldest rides in Epcot, however, have the longest duration of the park at 14-15 minutes, with the younger rides mostly below 10 minutes duration.

Another point to consider when examining financial records is the day to day income generated by tickets for entry to each property. The past ten years of ticket prices for Disney and Universal properties in Orlando are shown in Figure 10. As the graph displays, the price of Disney Theme Park tickets increased 112% for adults in the 2009 to 2019 time frame, while the Water Park ticket price has increased by only twenty dollars for adults in the same time frame, an increase of roughly 44.44%.

To further explore Disney ticket prices, one must also look into the pricing of annual passes, that is, a ticket that allows the guest to visit many times throughout the year by paying a single set price. This price too has risen throughout the time frame, from $414 and $469 for children and adults in 2009 to $894 for both in 2019. This was a total increase of 115.94% for children, and 90% for adults. Figure **??** demonstrates this trend, with a notable occurrence in 2012. This occurrence was the discarding of the children specification, with either option set at $574.

Unfortunately, the prices of historical Universal tickets were much harder to procure, and as such, are not providing total coverage for the time frame shown by Figure 10.

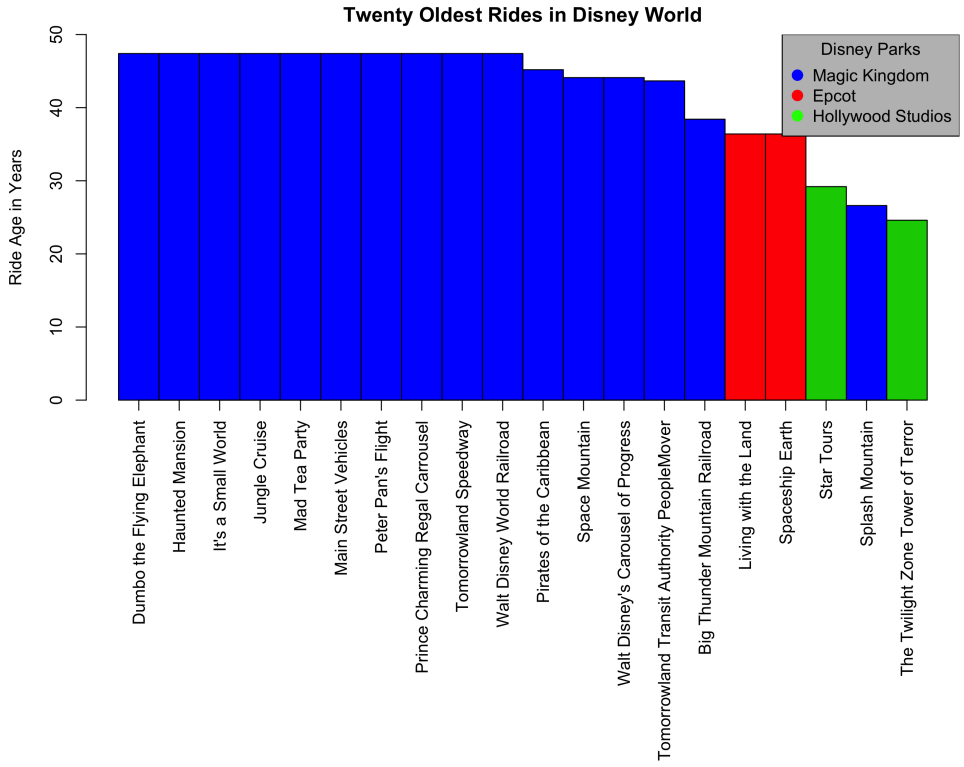


Figure 7: Twenty Oldest Rides

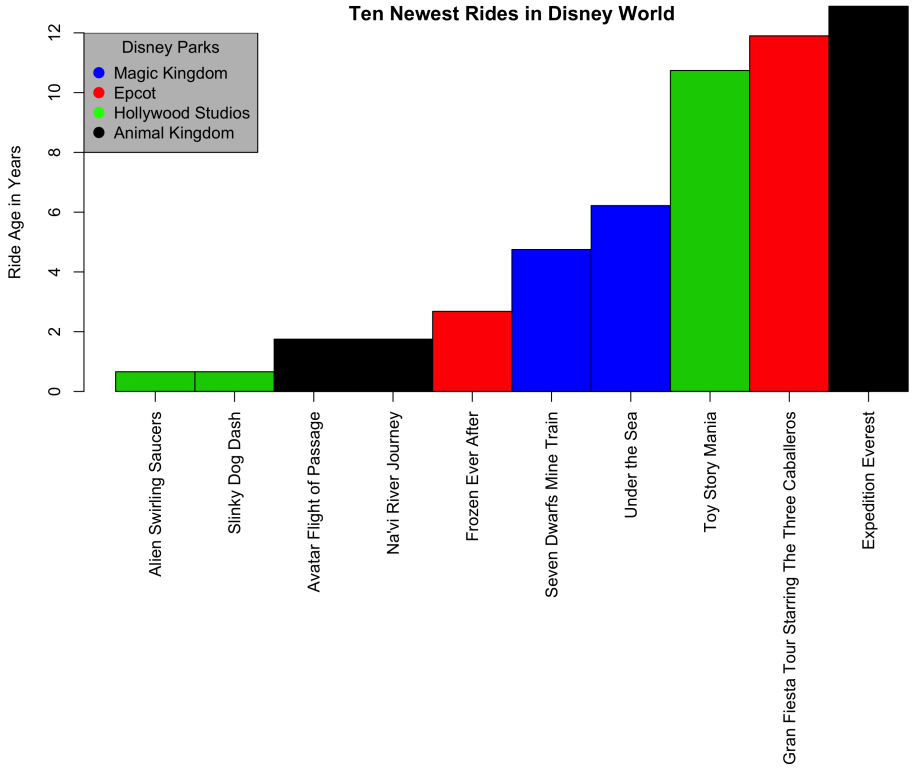


Figure 8: Ten Youngest Rides

However, if one observes the sparse data available, one can still see the increase from the adult pricing at $85 in 2011, to this years price at $170. This is a complete doubling of the price, so Universal had similar price increases to Disney. However, Universal did retain separation of children and adult pricing, leaving a bit of a price break for larger families.

Yet another source for Disney ticket prices was a study performed by Go Banking Rates [6], with ticket prices from 1971 to 2017 for the one day ticket for Magic Kingdom. This data set included the original ticket price, as well as the price adjusted for the inflation rates of 2017. Visualization performed with this data produced Figure 12. Note that the 1970 era prices would still only average out to around 22 modern USD. However, as previously shown in Figure 6, Disney World as a whole only featured 16 rides by the end of the 1970s.

4 Attendance

With the intense rise of pricing at the Orlando area Disney and Universal owned properties, one would assume that a lower number of guests would be visiting the parks.

However, as Figure 13 shows, as the prices at the parks rose there was not a drop in attendance whatsoever at Disney properties. Actually, since 2001, many of the properties have had slow but steady increases in guests, aside from Hollywood Studios who experienced a slight dip in attendance numbers in 2016-2017. However, it does appear that both Typhoon Lagoon and Blizzard Beach maintained lower attendance counts than all other Disney owned properties. However, this could also be attributed to the water parks' remarkably lower size as well (note that aside from Disney Quest, both water parks ranked lowest in total area as shown in Figure 4.

As mentioned above, new additions to the Disney parks tends to deliver a boost to the attendance of that given park, with occasionally other parks receiving a small decline. An excellent example of this would be 2017, when the 'Pandora' addition of Animal Kingdom was completed. As shown in Figure 13, Animal Kingdom outpaced Epcot and Hollywood Studios in that year, with even the Disney giant Magic Kingdom receiving a slight decline

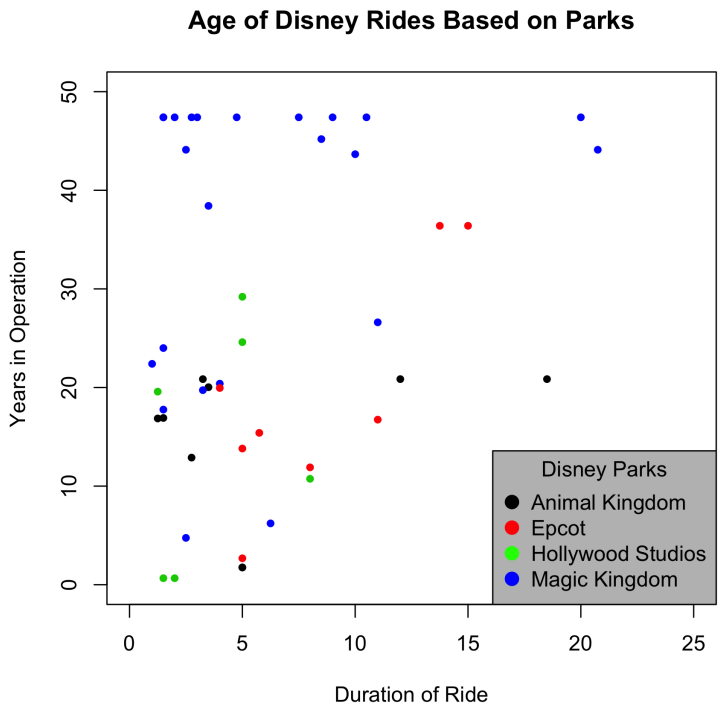


Figure 9: Ride Duration versus Age

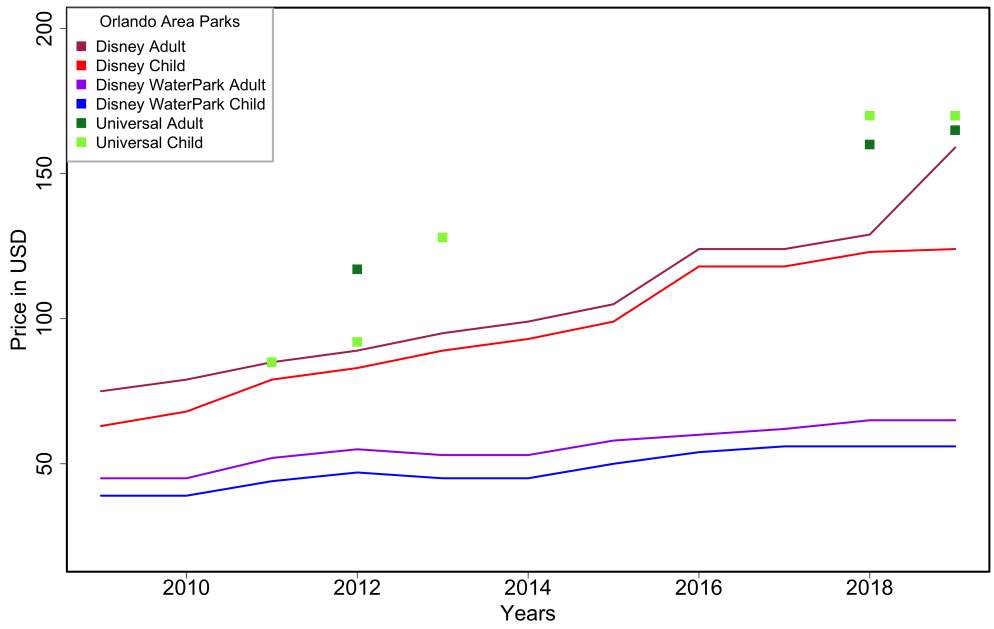


Figure 10: Ticket Prices of Orlando area Disney and Universal Brand Properties from 2009-2019

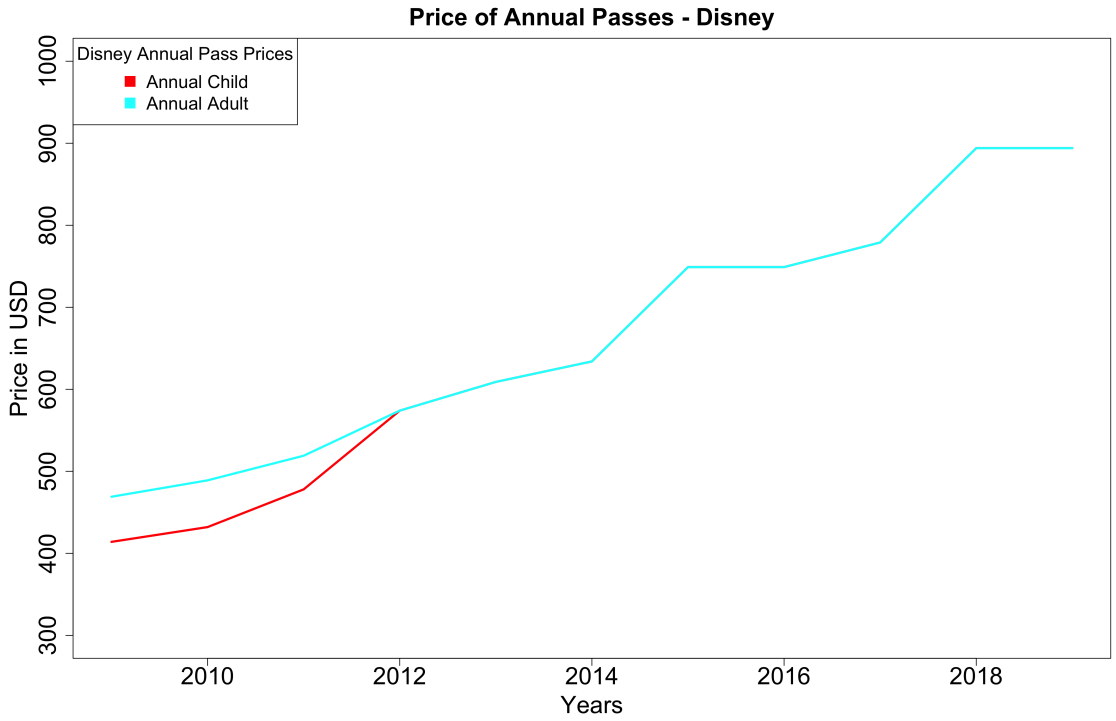


Figure 11: Price of Annual Passes from 2009-2019 Disney

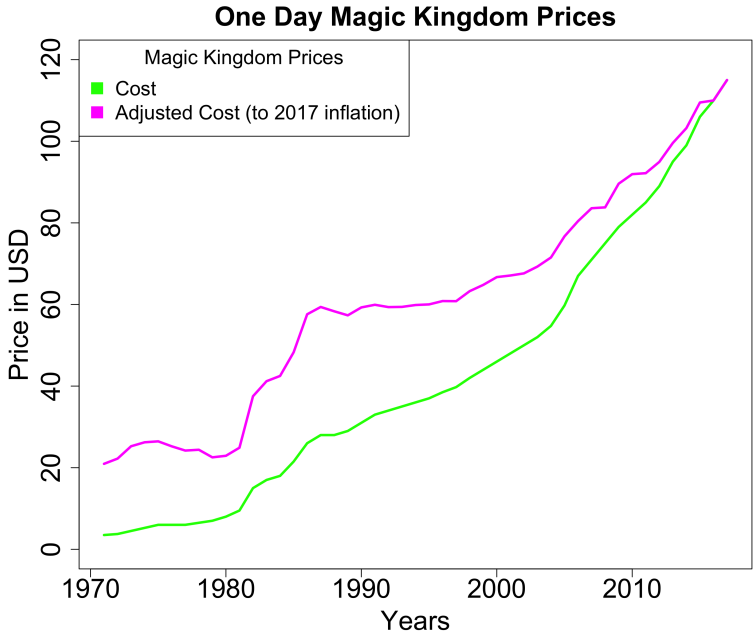


Figure 12: Price of Magic Kingdom Tickets, 1971-2017 with Adjusted Cost

on guest attendance. Unfortunately, the addition of 'Toy Story Land' to Hollywood Studios cannot yet be analyzed properly, due to the data set ending the year the addition premiered.

Now that overall trends have been pointed out, one can consider individual segments of the data. For example, park attendance for Magic Kingdom in 1993 took a drastic plummet, possibly caused in part by the early 1990s recession that rocked the United States economy[1]. The Magic Kingdom was able to rebound rather quickly, with a relative spike at 1997. One possible explanation for this excellent rebound was the release of Aladdin, Lion King, and Toy Story in the time of rebound. Also of note is the year 1996. It appears for 1996 that no attendance information was available to The Park Database team across all properties. It also appears that 2003 to 2006 was also a period of unavailable data for Disney Water Parks.

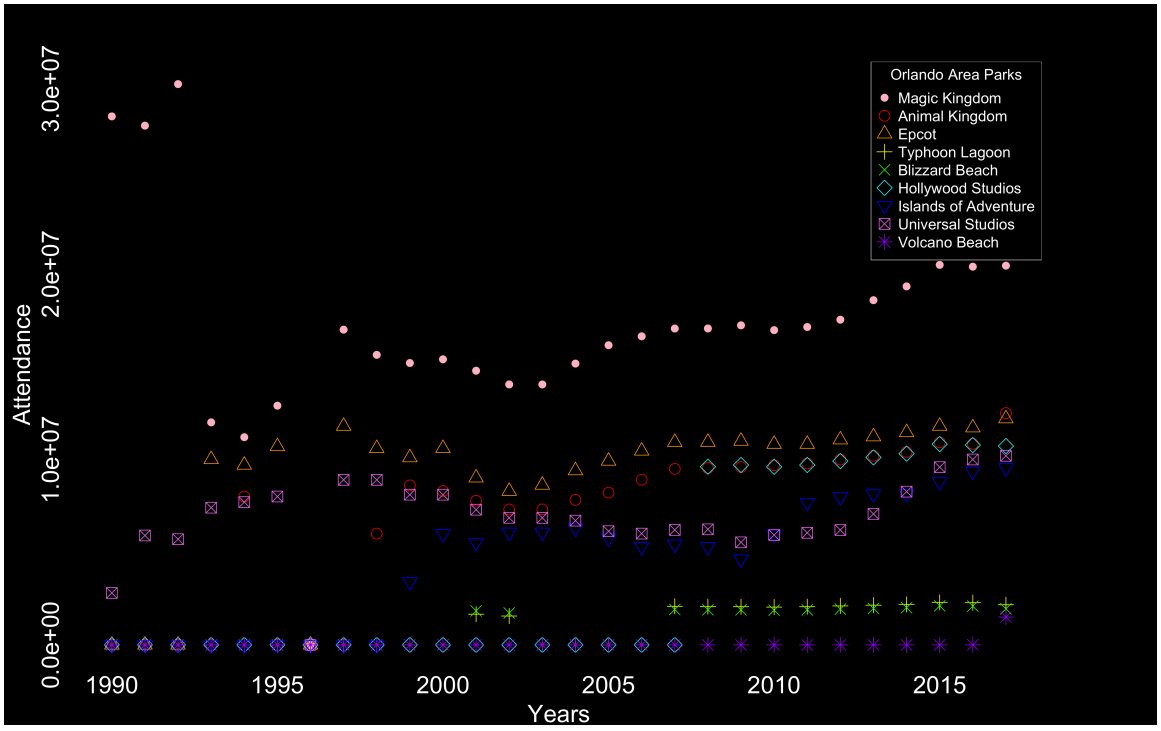


Figure 13: Disney and Universal Brand Property Attendance, 1990-2017

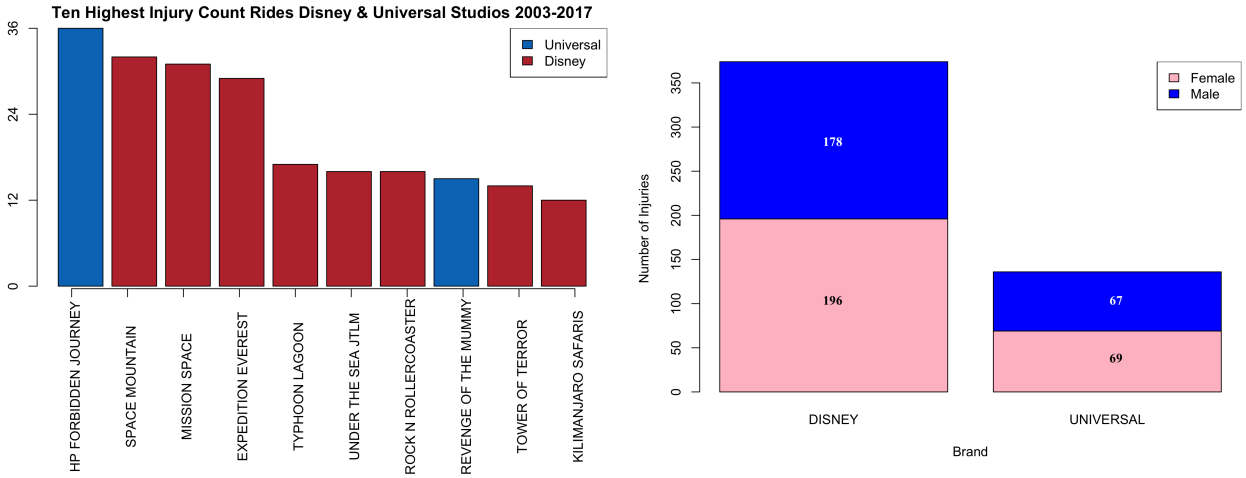
5 Damages

The final section of the Orlando theme park statistics to be analyzed is the history of injuries across Disney and Universal properties from 2003 to 2017. The total number of injuries across both brands can be observed in Figure 14a. The highest injury count was reported on a Universal ride, at 36 reported across the 14 year period. Unfortunately the Park Database data set does not feature the opening date of the Harry Potter and the Forbidden Journey. However, a quick google search produced the opening date of the Universal 'Wizarding World of Harry Potter' as June 15, 2010 [4]. This means that the ride acquired 36 reported injuries over a time frame of just over seven years. Meanwhile, Space Mountain (Disney's highest injury count at 32) premiered in 1975 making it one of the oldest rides in the parks. This means that the injury count accumulated by Space Mountain occurred over a time frame twice that of Forbidden Journey. Mission Space, the third highest collector at 31 injuries, opened in 2003. This means that similarly to Space Mountain, Mission Space also had the entire 14 year time frame to accumulate the injuries reported.

Alongside Figure 14a is Figure 14b, which illustrates the breakdown of injuries among visitors of female and male classification in the parks. Disney clearly shows a higher total injury count, with 178 for male guests and 196 for female guests, for a total of 374. This means the top ten rides shown in Figure 15a account for 50% (189) of the total injuries. Interestingly, the top three rides shown here (Space Mountain, Mission Space, and Expedition Everest are nearly tied, with 32, 31, and 29 respectively. It is worth mentioning that the total number of injuries reported at the Disney properties may seem quite high at first thought, but one must bear in mind that Disney's Orlando properties are reported by the Park Database as having 230 unique rides. This means that the total number of 374 rides is from 230 different rides.

Meanwhile, Universal is reported as having 67 injuries for male guests, with 69 reported for female guests. This sets the Universal total of injuries at 136. A top ten list was also produced for Universal (shown in Figure 15b) with a marked leader in Harry Potter and the Forbidden Journey. Surprisingly given the strong lead of Forbidden Journey, the top

ten rides in Universal account for roughly 80% (108) of the total. However, one must bear in mind the total numbers of rides found in Universal in drastically less than Disney's, with Universal having a total of 70 rides from the count from The Park Database. Additionally, many of Universal rides were opened rather recently. Harkening back to Table 6, Universal Studios opened in 1990, Islands of Adventure in 1999, and Volcano Bay in 2017. This should mean that many of the rides in Universal would be newer than their competition in Disney World. But once again this cannot be confirmed due to lack of concrete data sets regarding this.



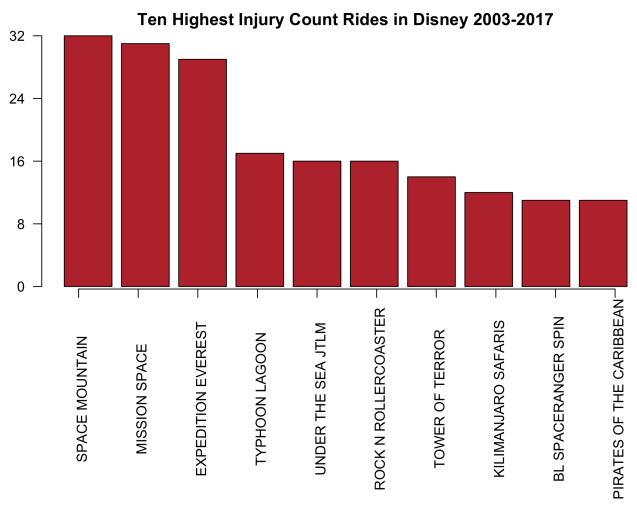
(a) Top Ten Injury Rides Reported at Disney and Universal Brand Properties 2003-2017

(b) Injuries Reported, by Gender and Brand

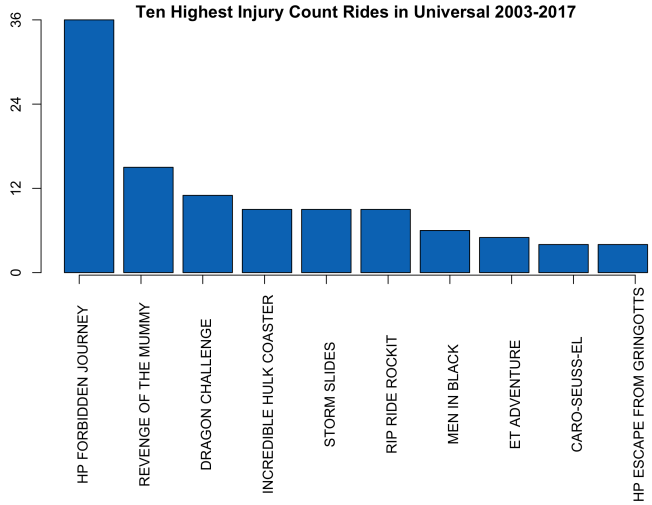
Figure 14: Injury Data Overview, Top Ten Rides Reports, Gender Distribution

While many would assume the majority of injuries sustained at a theme park property would be regarding young children, the data suggests otherwise. The distribution frequency of both Disney and Universal injuries of female and male guests are shown in Figures 16. Females injured in both parks tended to be middle-aged to elders, with the highest frequencies at 40-50 and 60-70 years old. Males injured in the parks had more of a normalized distribution, with the highest frequency occurring at 50-60 years old.

Figure 17 demonstrate the distribution of frequency of female and male guests to Universal properties. Note that the distributions here are almost identical to the combined data, just on a smaller scale. Once again, females are more likely to be injured at ages 40-50

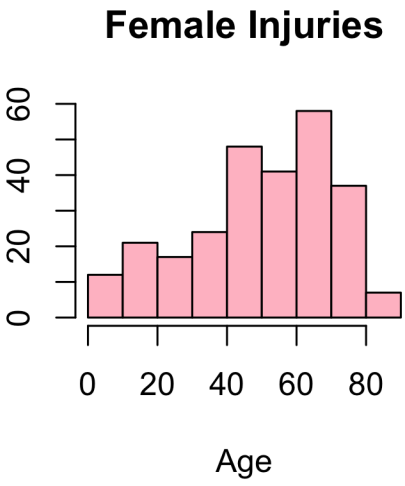


(a) Injuries on Disney Rides

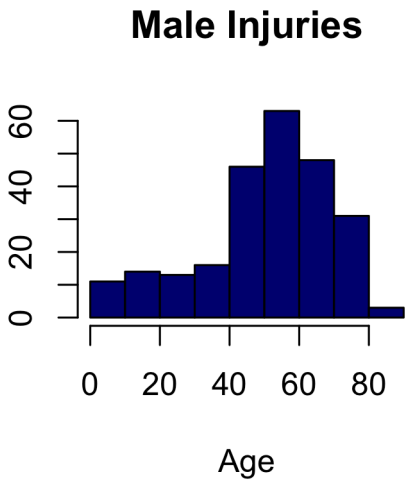


(b) Injuries on Universal Rides

Figure 15: Individual Top Ten Injury Counts per Brand



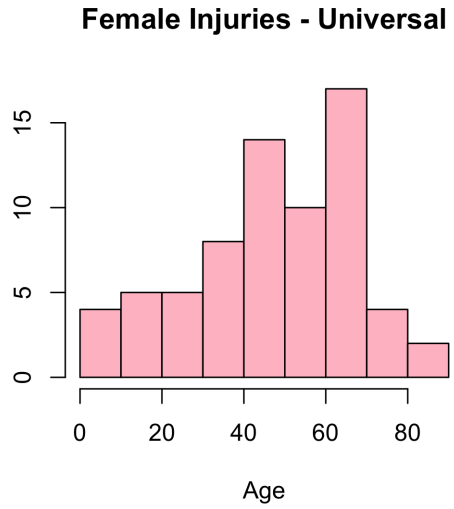
(a) Female Injuries Distribution



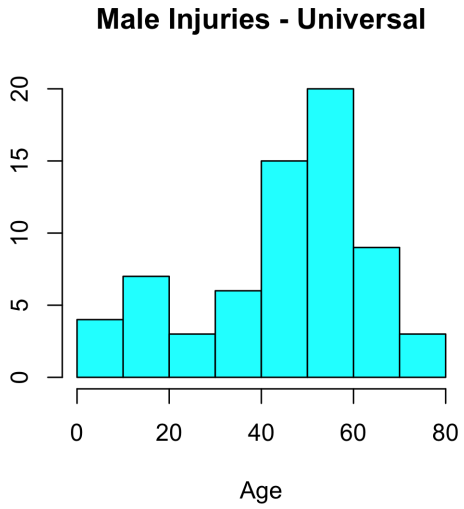
(b) Male Injuries Distribution

Figure 16: Injury Distribution by Gender

and 60-70. Males are more likely to be injured at ages between 40-60, but particularly from 50-60.



(a) Universal Female Injuries Distribution

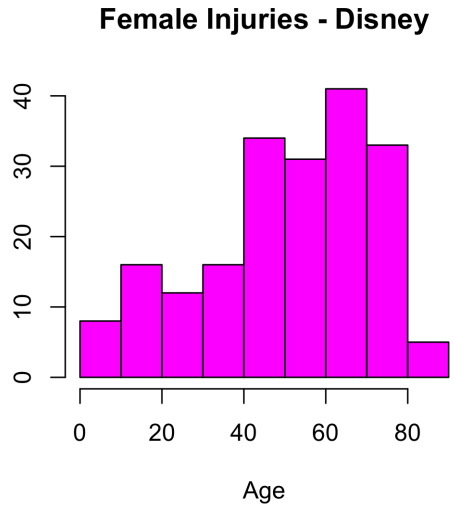


(b) Universal Male Injuries Distribution

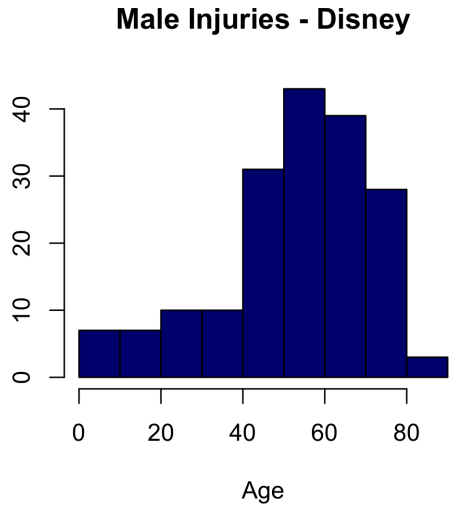
Figure 17: Universal Injury Distribution by Gender

Meanwhile, Disney property distribution of injuries looks a good bit more unique in the female graph. Injuries from female guests visiting Disney properties seem to follow more of a bell-shaped curve, with a steep increase at 40 years old which then peaks at the 60-70 age break. From 70-80 there is a downward trend, with a drastic reduction at 80. Male guest injuries from Disney properties closer follow the overall trend, with the beginning of the bell shape at 40-50, peaking at 50-60, and declining until the drastic drop once again at 80.

Overall, the study of the injuries reported at Orlando area theme parks reveals that children and young adults are not as much injured at the properties as their parents, aunts, uncles, or perhaps grandparents.



(a) Female Disney Injuries Distribution



(b) Male Disney Injuries Distribution

Figure 18: Disney Injury Distribution by Gender

Acknowledgements

The Park Database team were very generous and helpful in the completion of this report, providing an extensive amount of data for analysis. Particularly, the author would like to thank Aden for their fast response to the request for data. The data presented in this report is but a small sliver of the possibilities presented by the Park Database site, found at: http://www.theparkdb.com .

Additionally, Dr Jegar Pitchford generously provided personally retrieved data from Florida records for use in this report's injury analysis. Dr. Pitchford was also very receptive to the idea of allowing the author use of his data, and routinely provides insight to previously unexplored datasets on his site at https://themeparkanalysis.com .

Both data sites are phenomenal resources for the theme park research sector, and present the exact type of work the author wishes to engage with after graduation. Needless to say, the author will continue to check for any updated posts from both resources.

References

[1] 1990-92 early 1990s recession.

[2] A database of theme parks and attractions around the world.

[3] Wdw ticket history.

[4] Harry potter and the forbidden journey, Apr 2019.

[5] Dr. Pitchforth, J. Analysing florida theme park incidents: The road is long and full of regex.

[6] Nelson, D. This is how much it cost to go to disney world the year you were born, Oct 2017.

[7] Passanisi, L. Walt disney world ride data.