Crowdsourced Data and Visualizing Theme Park Wait Times

Sarah Beals

Library and Information Science Program: Capstone Project

August 2020

Advisor: Dr. Peter Organisciak

Abstract

If good things come to those who wait, then attractions at Walt Disney World must be very good things indeed since some attractions involve three-to-five hour waits. To optimize maximum ride time at the parks, TouringPlans offers a crowd-sourced service that calculates actual wait times. TouringPlans users have asked to accommodate rope drop times that start a few minutes before a park opens. In this capstone, we will visualize what it would mean to offer this service to the TouringPlans community. This study merges three TouringPlans datasets, two of which are interchangeable depending on the ride and park desired. The final visualization splits the data according to daily crowd density and Extra Magic Hours (EMH) when the parks are exclusively open to resort guests. Results suggest that wait times become increasingly scattered throughout the first hour regardless of crowd level during general entry, but generally hold steady during EMH.

Crowdsourced Data and Visualizing Theme Park Wait Times

Introduction

If good things come to those who wait, then attractions at Walt Disney World must be very good things indeed. Disney's Magic Kingdom ranked fist as the most attended theme park in the world with over 20 million visitors in 2018, and 10 more Disney parks in the top 25 attended parks worldwide (Rubin, 2019). With tens of millions of people visiting, popular attractions can involve a three-to-five hour line during peak times (Tschinkel, 2018). TouringPlans is one way park guests can plan their trips around wait timess by collecting and distributing data as a group. Powered by community-sourced data, TouringPlans offers subscription-based services to access crowd calendars and phone apps that minimize waiting in the parks. TouringPlans currently does not offer predictions for wait times before a park officially opens, even though guests are often allowed in the park. As a measurable unit, wait times can be extracted and plotted alongside other data to illicit predictable patterns. This study will visually present the wait in the first hour of park opening with three, large datasets.

The importance of this study is rooted in community, data, and fun. Libraries are critical institutions within a community, as their mission is to assess and provide applicable resources to the community. This mission can be seen in the Libraries Transform campaign by the American Library Association. Libraries Transform aims to educate both the library patrons and the general public about the missions of a modern library (*Libraries Transform*, n.d.). The "libraries today are less about what they have for people and more about what they do for and with people" (*Libraries Transform*, n.d.). Investment in a local library is reciprocated by the library advancing and transforming the community with lifelong learning (*Libraries Transform*, n.d.). Lifelong learning, by nature, incorporates different topics and learning styles. For example, the campaign states, “adding Minecraft to curriculum may inspire more future engineers than mathematics alone” (*Libraries Transform*, n.d.). Similarly, organizations like Black Girls Code are teaching coding skills by helping kids build their own games (*Black Girls Code*, n.d.). This study takes community sourced information, turns it into numeric data, and ends with a more efficient vacation. This study demonstrates a tangible example of mathematical calculations with conceptual consequences at an accessible level for people beginning to learn code (i.e. a little, group math equals more attractions for everyone).

Librarians are ideal data scientists because they are experts in information management, organization, and communication, i.e. taking large amounts of information and transforming it into something their user understands (Stanton, 2012). Librarians are invested in users and their experiences, and therefore should study queuing theory and crowd distribution, as it pertains to many situations outside of theme parks. Libraries can transform when they examine a community working together to reduce wait times, and apply it locally.

Disney guests will usually wait long periods of time for high-demand attractions, but not always. This study looks at data from 2015-2019 to examine if the length of guest’s wait is indicated by when a guest waits. Factors that impact a wait include the crowd and the time of day. These topics are often discussed on the TouringPlans forum.

TouringPlans hosts a forum, where users discuss the crowd calendar, touring plans, dining, accommodations, attractions, deals and money, trip reports, and miscellaneous (TouringPlans Discussion Forums, n.d.). The forum can be difficult to interpret due to the abbreviated terms that stand for rides, parks, times, and events. For example, one user opened a thread with, “When I put it in TP and optimize it pushes SDD down to later in the day (5pm). I thought the best advice was to still RD…am I missing something” (Another SDD and Rope Drop Question, 2019)?

Translated, the question asks, “When I put it [the ride] in TouringPlans and optimize it (the plan), [TouringPlans] pushes Slinky Dog Dash down to later in the day (5pm). I thought the best advice was to still rope drop (arrive before park opening and follow a pace-setting rope to the ride)… am I missing something?” There is an updated thread with the abbreviations, and there are other resources that use the same or similar short hand, including the Disney Parks (Acronyms and Other Trivia, 2014; “Common Disney Abbreviations and Acronyms,” n.d.; Champion, 2018).

The question, however is one that is repeated throughout the forums (My Wish for Touring Plans - Rope Drop Option, 2019; Rope Drop to SWGE during EMH, 2019). The first few minutes of TouringPlans is not currently designed to accommodate rope dropping enthusiasts. Rope drop refers to the actual time Parks begin allowing guests through the gates, which is typically a few minutes before the scheduled opening. There is a way to force TouringPlans to accept the attraction, but it is another step and it will not optimize the rest of the day’s plan correctly (My Wish for Touring Plans - Rope Drop Option, 2019).

TouringPlans already has the data to estimate wait times for the first hour of park opening. TouringPlans users have also asked for a way to “pin” or “fix” the ride that they are determined to rope drop, and optimize the schedule around that item. In this capstone, we will visualize what it would mean to offer this service to the TouringPlans community.

The following paper includes a background, methodology, results, and a conclusion. The background explains theories in waiting, TouringPlans’ data collection, the selected rides in this study. The background includes sources from the TouringPlans community, such as message boards and blogs, and while they are not academic sources, they provide necessary insight. The methodology describes the merging of three datasets to build the visualization. The results section looks at the visualization outputs of the first hour at three rides and discusses how that information is useful to guests and the park. The conclusion offers suggestions for further research in the field. This paper was written with information sourced from TouringPlans.

Background

Park Controlled Variables

The waiting experience has been the subject of many studies, especially in theme parks. Because theme parks wish to maintain the image of entertainment rather than waiting, they innovate new ways to shorten or justify wait times.

In the *Psychology of Waiting*, David Maister's first law of service is : satisfaction = perception – expectation (Maister, 1985). In order to keep clients, customers, and guests happy, they must experience something slightly better than their preconceptions. Maister also provided eight principles for managing customer satisfaction regarding wait times, and scholars have updated the 1985 guidelines (Katz et al., 1991; Norman, n.d.) For the purpose of this capstone, I will utilize Don Norman's eight design principles for waiting lines:

1. Emotions Dominate
2. Eliminate Confusion: Provide a Conceptual Model, Feedback and Explanation
3. The Wait Must Be Appropriate
4. Set Expectations, Then Meet or Exceed Them
5. Keep People Occupied: Filled Time Passes More Quickly Than Unfilled Time
6. Be Fair
7. End Strong, Start Strong
8. Memory of an Event Is More Important than the Experience (Norman, n.d.)

As a host, the Disney Parks are responsible for managing all eight design principles, but the first principle (Emotions Dominate) is the most important. In the great spectrum of emotions, theme parks are designed to emphasize joy, happiness, excitement, and wonder, and in in order for those emotions to be the memorable ones (i.e. principle number eight), the other design principles must work together. Rope drop combats dominating emotions by asking guests to respectfully walk behind a rope (similar to a parade banner), instead of rushing and trampling each other in anticipation. The emotional and memorable stakes are high at theme parks as lines disrupt the relaxing or magical environment guests paid to experience. Therefore, parks will spend time and money to create and maintain a suspended reality while guests are waiting in line (Daniels, 2016)

Queuing theory is a branch of applied mathematics that is “used in operational research for assessing waiting time, storage time, and average time taken to reach the front of the queue… the theory can be used to assess traffic flow, material distribution, and waiting time” (C. G. Gorse et al., 2020). The three pillars of queuing theory are utilization, variability, and pooling (Stanford, 2013). Utilization is the rate at which guests enter the line to when they get on the ride, or the wait time. Variability is when people decide to enter the line, which is noticeable when many people enter the line at once, or significantly fewer people are entering the line. Pooling is when attractions can hold large quantities of people in their queues and on the ride, thereby reducing wait times for other park attractions.

Queueing theory is found in hospital waiting rooms, customer support call centers, and grocery store checkouts to name a few. Theme parks are different because they are supposed to be fun, and waiting is not fun. Guests “typically spend 20% of their time experiencing attractions but over half their time waiting” and “as much as 80%... during peak seasons” (Zhang, Li, Su, et al., 2017). Interestingly, guests during peak season who waited longer for an attraction are satisfied with fewer experienced attractions than those who experienced more rides during non-peak times (Liang & Dong, 2011). This could be because expectations on both the park and the guest were set and met by each party.

Therefore, parks are anticipating waits and enhancing queue areas by:

1) Fostering engagement, like videos or games
2) Maintaining interest in the attraction by continuing the décor, design, and theme
3) Supporting positive environments by avoiding loud, repetitive music or sounds
4) Maintaining comfort with temperature, flooring, and places to lean or rest
5) Separating longer lines from expedited lines to avoid feelings of injustice
6) Facilitating interpersonal interaction by allowing space for group conversation
7) Presenting wait times, but not during the wait unless there is a significant delay
8) Providing a sense of progress with multiple, themed, pre-attraction areas (Ledbetter et al., 2013)

Parks are working to maintain guest satisfaction, support the park operations team, and increase revenue. Successful implementation requires math, design, and experience. Despite the park's preparation, there will be crowd surges and wait times will noticeably increase. To a point, the Disney Parks can control the number of guests in the park (*Phased Closures Due to Capacity Constraints | Walt Disney World*, n.d.), but it is more likely the guest who will decide if they will visit during historically, crowded time periods. This is because the number of guests is a number affecting the park and its physical tourism carrying capacity (TCC), but the imperceptible number does not directly affect the guest (Zhang, Li, Su, et al., 2017). The guests' satisfaction can be measured by "the lowest acceptable visited-attraction number and visitors' maximum tolerable waiting time" to determine the psychological carrying capacity (Zhang, Li, Su, et al., 2017). This is essentially Maister's equations for satisfaction = perception – expectation. Psychological carrying capacity is important for repeatability because if a guest was at maximum capacity, they were probably not enjoying the vacation, which highlights Don Norman's eighth principle: Memory of an Event Is More Important than the Experience.

To increase capacity and memories, Disney offers resort guests the option to come early or stay late during Extra Magic Hours (EMH). EMH is not a daily operation, and therefore requires previous knowledge and planning. In this study, it affects the first hour of park opening because on some days, the park is open to a smaller group.

Guest Controlled Variables

In 2020-2021, tickets to Walt Disney World started at $109/day/person for guests aged 10 years and older (*Disney World Theme Park Tickets in Orlando, Florida | Walt Disney World Resort*, n.d.). For a family of four to visit for five days, the cost is nearly $2000, not including food, lodging, or travel. Because these vacations are expensive, some families decided to plan efficient tours and maximize their time on the attractions. It is difficult to predict wait times because of the variables, and as shown in Norman's eight principles for waiting lines, unknowns can be uncomfortable for guests.

With a tool like TouringPlans, guests can "make more informed choices to suit their interests," and create more personalized and accurate expectations, increasing the possibility of higher satisfaction (Zhang, Li, & Su, 2017). TouringPlans is paired with good, spatial design because "well designed space could decrease waiting time and optimize a park's tourism carrying capacity" or TCC (Zhang, Li, & Su, 2017). Zhang, Li, and Su observed crowds at a popular theme park in 2010, and they offered these 10 propositions that distribute crowds while allowing guests to optimize their day at the park:

1) Ceteris paribus, attraction experience value affects visitor movement in a theme park.

2) Ceteris paribus, attraction facility capacity affects visitor movement in a theme park.
3) Ceteris paribus, attraction floor area affects visitor movement in a theme park.
4) Ceteris paribus, attraction types affect visitor movement in a theme park.
5) Ceteris paribus, attractions' indoor feature (i.e., being indoors or not) affects visitor movement in a theme park.
6) Ceteris paribus, distance between attractions affects visitor movement in a theme park.
7) Ceteris paribus, the path network affect visitor movement in a theme park.
8) Ceteris paribus, entrance location affects visitor movement in a theme park.
9) Ceteris paribus, subareas' attraction density affects visitor movement in a theme park.
10) Ceteris paribus, attraction spatial configuration affects visitor movement in a theme park. (Zhang, Li, & Su, 2017)

These ten propositions are integral to theme park design, meaning, they cannot be implemented at the last minute. Crowd distribution will be an important factor when placing walkways and attractions. The 2010 Zhang, Li, and Su study also recommends the radial path network seen in Disneyland and Walt Disney World's Magic Kingdom (see figure 1) for optimal capacity (Zhang, Li, & Su, 2017). The spoke wheel design places major attractions on the wheel's outer rim. The spoke wheel is efficient for crowd distribution, but not for the guest who is trying to plan the most direct route to each attraction.

The time dependent traveling salesman problem (TDTSP) is "a variant of traveling salesman problem (TSP) where the amount of time it takes to travel from one city to another varies depending on the time of day" (Testa et al., 1999) TDTSP is applicable to theme parks because "high experience value attractions' locations are closely tied to visitor movement," and "these attractions draw many visitors" (Zhang, Li, & Su, 2017). This means that popular, high value attractions are not typically located near one another, to help with pooling. In order for a guest to experience the high value attractions, they have to walk around the entire park. Luckily, computers can calculate TDTSP, and they can include an evolutionary algorithm to save meal times and walking speeds (Testa et al., 1999).

*Figure 1: Magic Kingdom Map, 2020. Spoke wheel outline added by author (*Magic Kingdom Park*, 2020)*

TouringPlans.com began when Len Testa expanded his business from publishing *Unofficial Guides* to Disney Parks, to planning subscribers' vacations online (Weinberger, 2019). Testa began the algorithms as his graduate thesis in 1997, and launched the site in 2012 (Weinberger, 2019). Their homepage explains, "Our trip planning tools show you the least crowded park to visit every day, customized touring plans for visiting Disney's best rides, honest restaurant reviews, how to save on Disney tickets, the best hotel rooms to ask for, and so much more" (*Touring Plans*, n.d.). By working together for eight years, 300-500 families contribute data daily, in real time, to build a large database of information (Weinberger, 2019). TouringPlans.com began as an unpaid project with Testa walking 18 miles a day physically recording wait times, but as of June 2019, has 140,000 paid subscribers, seven full-time employees, and 12 part-time employees (Lee, 2004; Weinberger, 2019).

Paid subscribers have access to the app, which updates the user's plan in real time when there are surges, drops, or closures affecting the plan. The app also features a chatroom, wait time estimates, menus and prices, and badges for users that submit wait times (*Disney World App | IPhone, Android, Mobile Web*, n.d.). Disney announced their own version of an app in 2019, but the late-2020 release will likely be delayed due to coronavirus related closures, phased re-openings, and a re-closure at Disney Parks (Barnes, 2020; Barton, 2020; Healey, n.d.; The DIS, 2019).

Example Attractions

This study examines data from three datasets, three attractions, and two parks at the Walt Disney World Resort. The three attractions are all dark rides which operate indoors, with varying levels of popularity and locations in the parks (Zika, 2018). These rides were chosen from TouringPlans website, where select attractions' data are available for students.

Rock 'n' Roller Coaster Starring Aerosmith is located in Hollywood Studios and is a thrill ride with long waits. The ride is loud, in the dark, with high speed, drops, and inverted loops (*Rock "n" Roller Coaster*, n.d.). Guests enter a queue design to look like a recording studio, and they find Aerosmith as the band realizes they are late for a concert. Aerosmith invites guests to the concert in a stretch limo through Los Angeles, and the ride begins. Rock 'n' Roller Coaster opened in 1999, and remains a popular attraction that does not close due to weather (Sanders, 2019b).

Toy Story Mania! is a 3D, dark ride in Hollywood Studios with long waits. The ride's spinning trams are on a track that guiding guests from scene to scene (Sanders, 2019a; *Toy Story*

Mania, n.d.). The queue is built to scale guests down to toy height, as they find themselves in Andy's room from *Toy Story* (1995). Andy created the Midway Games from toys, crayons, holiday lights, and board games. Guests board the tram and are propelled into a 3D carnival game with their party. The games change and count points within a tram, thus creating a repeatable attraction. Toy Story Mania! opened in 2007, maintains popularity, and does not close due to weather (Sanders, 2019a)

Spaceship Earth in a slow-moving, dark ride in Epcot located directly behind the main gate and is the first ride most guests see upon entering (Epcot has a second, smaller entrance for nearby resorts). The ride continuously loads, runs, and unloads at a slow pace (Sanders, 2020). Guests witness the history of human communication as narrated by Academy Award winner Judi Dench (Sanders, 2020; *Spaceship Earth*, n.d.). The ride is 16 minutes long, includes an interactive touch screen in each car. Spaceship Earth opened in 1982 during Epcot's debut, and maintains popularity but not as high as other thrill or game rides (Sanders, 2020). Spaceship Earth was scheduled to close in 2020 a reopen as a new attraction, but the COVID-19 closures postponed the updated attraction (Sanders, 2020).

Methodology

In order to visualize wait times during the first hour of Disney attractions, precise tools are required. The goal is to merge three datasets together and use specific values to visualize the wait within the first hour of a park based on crowd level and Extra Magic Hours. This section will review the data and tools within this project. The community supplied data comes from TouringPlans. The tools are the Python language, the *pandas* library, and Google Colab. TouringPlans is available for a $15.95 yearly subscription, and the other tools are freely available with resources on how to use them. Python, *pandas*, and Colab work together in this project.

This section includes coding examples demonstrating how Python, *pandas*, and Colab work together.

TouringPlans

TouringPlans compiles attraction wait times from theme parks and volunteer guests into individual data sets for each attraction. TouringPlans also maintains an ongoing collection of metadata related to the park such as when the park opened that day, what season or events took place, and even what percentage of American schools are in session according to region. This metadata is key to the wait time data because it helps to break apart the variables and reveal how they affect each other. TouringPlans created a simplified indicator of the wait times and metadata into system of crowd levels. Crowds are defined as "the average posted wait time for the key attractions between 10:00 am and 5:00 pm" because it is a fixed time slot, with objective, collectable data (*Touring Plans*, n.d.). The crowd level system is numeric 1-10, where 1 indicates extremely low crowds and 10 indicates extremely high crowds. Together, the wait times, metadata, and crowd levels work together to visualize wait times.

Tools

Python is a programing language. The language has syntax rules like any other spoken language, but it is versatile enough to be applied in many coding projects (*General Python FAQ*, n.d.). Python is also designed to be logical as far as methodology and readability (Peters, 2004). Guido van Rossum release Python in 1991 as a successor to the ABC language, which he liked but found too limiting (*General Python FAQ*, n.d.). Rossum named Python after *Monty Python's Flying Circus* (1969-1974). Python is a small, core language that utilizes a large library (like *pandas*) in order to run code. This method allowed companies (like visual effects giant Industrial Light & Magic) to take their existing code and streamline it into fewer pieces of hardware and faster coding practices (Fortenberry, n.d.). There are many libraries associated with Python, and *pandas* is useful for the data frames in this project.

pandas is a Python library or "a collection of pre-combined codes that can be used iteratively to reduce the time required to code" (Advani, 2020). Libraries can store mathematical shortcuts in the commands, thereby saving the programmer from manually calculating most of the numerical problems. *pandas* is open source as of 2009, with the mission of providing "practical, real world data analysis" by prioritizing accessibility, flexibility, and power (*pandas - Python Data Analysis Library*, n.d.). In regards to this project, pandas is well suited for missing data (or NaN), inserting and deleting columns from the data frame, merging and joining datasets, and manipulating time (*pandas - Python Data Analysis Library*, n.d.). Other libraries in this project include seaborn, matplotlib, and NumPy.

Google Colaboratory (Colab) is a free option for writing and executing Python code in a browser, sharing code with other programmers, and synching with your Google account (*Google Colaboratory*, n.d.-a). Colab automatically saves changes and tracks revision histories within the code (*Google Colaboratory*, n.d.-b). Colab can save to an .ipynb file, the open source file for Jupyter Notebook's Python documents. One dataset for this project was stored on a Google Drive rather than the TouringPlans website, and Google required authentication before accessing the stored data. Colab holds Python and pandas together, and all three tools built the merged wait times.

Code

To visualize wait times during the first hour of park opening, during different crowd levels, three files must merge and show only the relevant data. Two files, a specific rides wait times and the Disney World metadata, are retrieved from the TouringPlans site and merged on date. Rides can be swapped out in the same code and line up with the metadata, essentially making this Colab file a template. The dates begin on January 1, 2015, and end on December 31, 2019. Columns with HS or EP in the title are park specific, and can be adjusted to match the park associated with a particular ride.

pandas then finds rows with an actual wait time recorded (as opposed to rows with only posted wait times) by looking for NaN entries and dismissing them from the new, merged dataset. Columns from the merged dataset are reduced from 194 to 8: date, datetime, SACTMIN, DATE, DAYOFWEEK, HS/EPEMHMORN, HS/EPOPEN, and HS/EPEMHOPEN. TouringPlans' data dictionary defines the variables as:

- date (ride): Park Day (not actual date stamp of the wait time, since some are after midnight). Format, MM/DD/YYYY
- datetime (ride): date-time stamp of wait time. Format: YYYY-MM-DD HH:MM:SS
- SACTMIN (ride): Actual Wait Time (in minutes). Format: numeric
- DATE (metadata): Park Day (not actual date stamp of the wait time, since some are after midnight). Format: MM/DD/YYYY
- DAYOFWEEK (metadata): Day of Week. Format: numeric
- HS/EPEMHMORN (metadata): Hollywood Studios/Epcot Extra Magic Hour Morning. Format: Boolean
- HS/EPOPEN (metadata): Opening Hour for Disney Hollywood Studios. Format: HH:MM
- HS/EPEMHOPEN (metadata): Hollywood Studios/Epcot Extra Magic Hour Opening Time. Format: HH:MM

Every date or time-based column had to be converted to a datetime format for *pandas* to read. Datetime format allowed for joining and time alignment. When converting into the datetime formats, three new columns were created: QueueStartTime, HSEMHOpenTime, and DateOnly.

Once joined and converted, a new column is created called TimeSinceOpen. TimeSinceOpen is the difference between the time a guest entered the ride queue and when the park opened. HSEMHOPEN has EMH opening times in the data, but on days without EMH, it still shows the general opening time. Therefore, TimeSinceOpen shows both EMH and general results. For this project, the only utilized times are within the first hour of opening, in addition to rope drop. TimeSinceOpen is then directed to not show results longer than one hour, and nothing less than -10 minutes.

TimeSinceOpen is in datetime format, and for the matplotlib graph, the time is converted to seconds in a numeric format. The new column Seconds acts as the x-axis representing the first hour (or 3600 seconds) of park opening, and actual, recorded wait times are the y-axis, and the graph is created with the matplotlib library (figure 2, in Results). Up to this point, the newly created columns are:

- QueueStartTime (ride): Time that guest entered the ride queue. Format: datetime
- HS/EPEMHOpenTime (metadata): Park opening time. Format: datetime
- DateOnly (ride): Park Day, from date. Format: datetime
- TimeSinceOpen (metadata and ride): When a guest entered the queue in relation to park opening. Format: datetime
- Seconds (metadata and ride): TimeSinceOpen in integers. Format: numeric.

CrowdHistory, the third dataset, is located on a personal Google Drive account. Colab is compatible with Drive, and only links the Drive files to matching Colab accounts. CrowdHistory's date is converted to datetime format (CL_Date2) and joined with DateOnly's matching column. The joined datasets are named RideWaitCrowd, and are reduced from 207 columns to 15. CL_Date2 is not revealed in the final data set, but the new columns are:

- CL_Date (CrowdHistory): Park Day. Format: M/D/YYYY
- CL_HS/EP (CrowdHistory): Park Crowd levels on a scale from 1-10 as calculated by TouringPlans. Format: numeric

The same SACTMIN/Seconds graph is plotted with the seaborn library, creating the same visual with a slightly different design (figure 3, in Results). The same variables are placed into a Lorentzian Model, so as to find the best fit for the overall data (figures 6-8, in Results).

The last visual splits RideWaitCrowd into 20 graphs (figure 9-11, in Results). Ten are for each crowd level, and then whether the day is EMH or not. This allows for side-by-side comparison of EMH vs general opening, and the ability to view how increasing crowds affect wait times.

Results

In visualizing wait times, there are four examples per ride. The matplotlib and seaborn graphs are essentially identical, differing primarily in aesthetics. The Lorentzian Model adds a line of best fit across the increasingly scattered times. The FacetGrid displays two variables side-by-side for easy comparison of those two variables. Unless otherwise stated, the examples in this section will also be from Rock 'n' Roller Coaster.

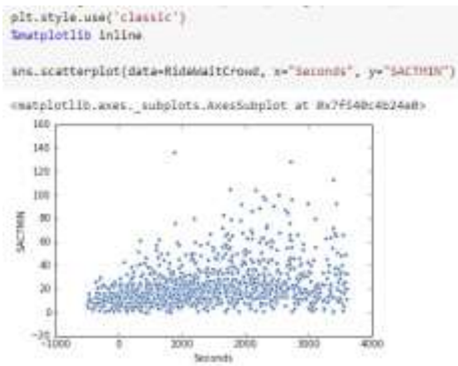

Figure 2

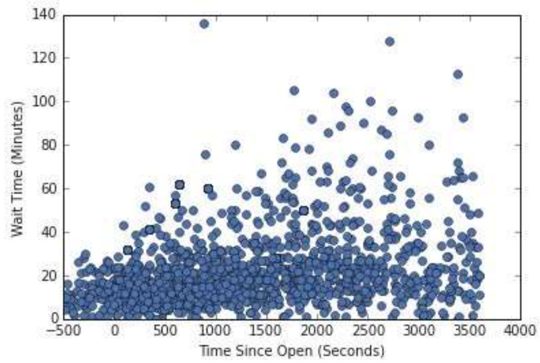

Figure 3

Matplotlib and seaborn are different libraries, though seaborn is based on matplotlib (Waskom, 2012). Both graphs reaveal the same data for Rock ‘n’ Roller Coaster, in that wait times are densly together closer to park opening, and fan out throughout the hour (figures 7 and 8). The graphs also show that guests arrive early to get in line before the park officially opens.

When comparing Rock ‘n’ Roller Coaster and Toy Story Mania, Toy Story Mania is a consistent, rectangular shape (figure 4), and Spaceship Earth shows a steep incline (figure 5). The rides are at different scales, with Toy Story Mania including a wait over 160 minutes and Spaceship Earth carrying noticeably fewer data points over the same four-year period.

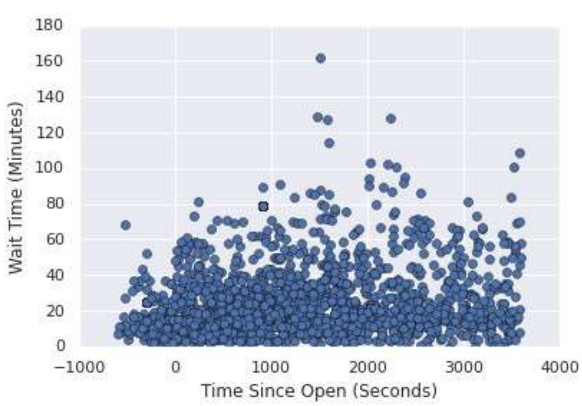

Figure 4: Toy Story Mania

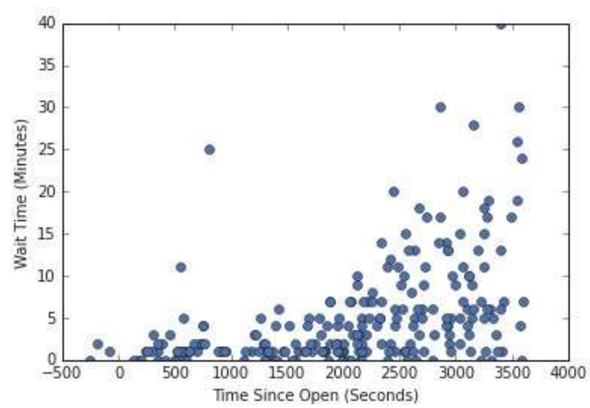

Figure 5: Spaceship Earth

While the Lorentzian Model did not offer the most flexible line to represent wait times within the first hour of opening, it did verify the difference between thrill rides located within the park, and slow rides at a park’s entrance. Both Rock ‘n’ Roller Coaster and Toy Story Mania maintain a relatively flat curve during the first hour (figures 6 and 7), but Spaceship Earth's incline is enough to invert the curve (figure 8).

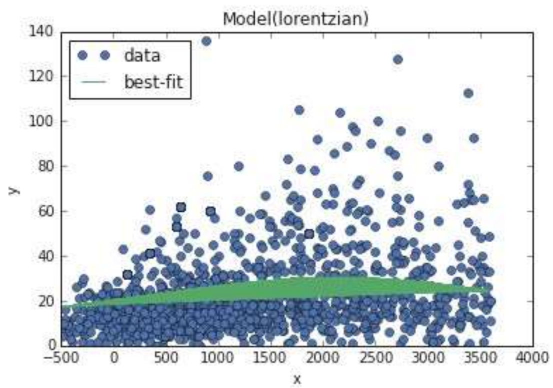

Figure 6: Rock 'n' Rollercoaster

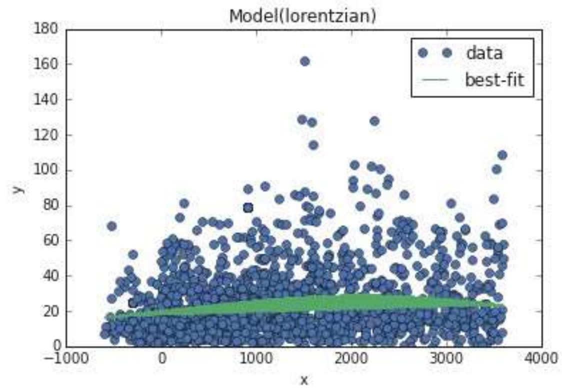

Figure 7: Toy Story Mania

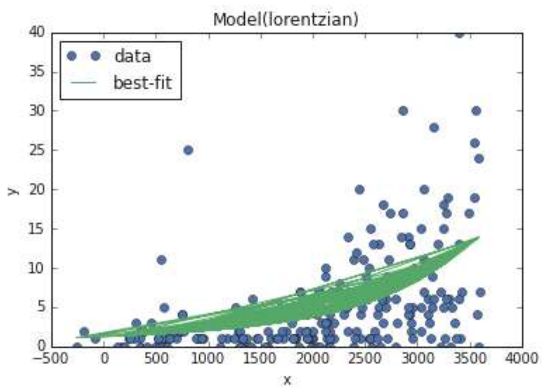

Figure 8: Spaceship Earth

In the FacetGrid, it is easier to see the impact of crowds and EMH on a given day. In the three grids, general opening is the top row with EMH making up the bottom row (figures 9-11). Each level is in numeric order, though, due to TouringPlans' numbering method, level 10 is between level 1 and 2. Each grid is therefore numbered as 1, 10, 2, 3, 4, 5, 6, 7, 8, 9. Crowd levels are defined by wait times, so a crowd level 1 will have the shortest wait times, and a crowd level 10 will have the highest. The crowd levels in these graphs are specific to the park with either the CL_HS/EP column implemented for the adjacent ride. By breaking each level and entrance time into a separate layer, it is easier to see how those variables affect wait times.

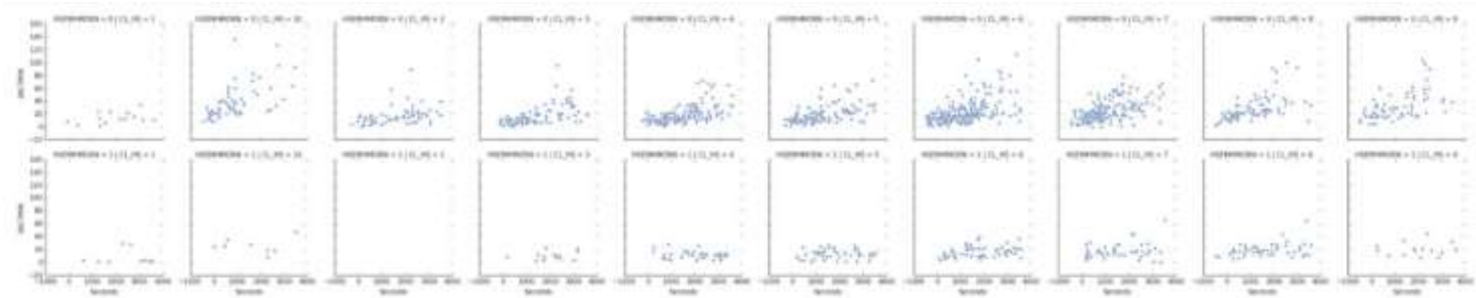

Figure 9: Rock 'n' Roller Coaster

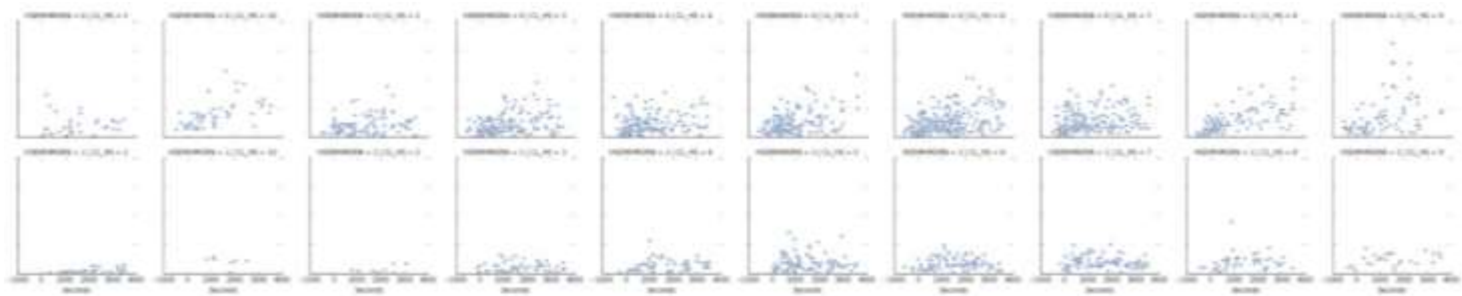

Figure 10: Toy Story Mania

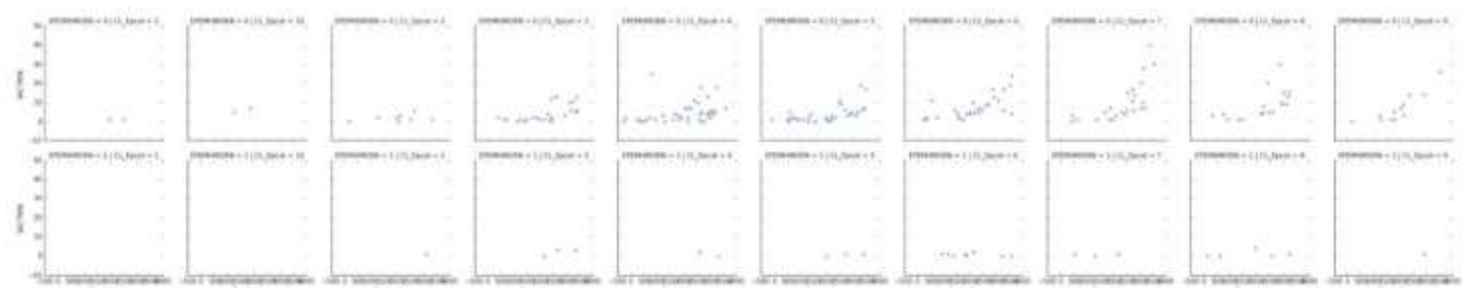

Figure 11: Spaceship Earth

Crowd data appears most dense at levels 4-7. This probably due to guest ability to travel and enter the ride's line. Lower crowd levels are typically associated with times that schools are in session and there are no holidays, and less traveling would likely generate less data. During peak season however, guests can spend hours in line for a single ride, thereby decreasing data points for all rides even if there were more guests present. The data also shows that wait times during EMH are steadier and lower than general opening days. This is a park controlled variable when they only let certain guests into the park early, significantly reducing the number of people in the park.

Overall, these data suggest that the first ride in a scheduled touring plan is important and that wait times become increasingly scattered throughout the first hour regardless of crowd level during general entry, but generally hold steady during EMH.

Conclusion

Wait times come with any Disney Parks vacation, but there is a community outside of Disney gathering information to better manage wait times. This information is voluntarily submitted from within the community and it is available online for students to practice learning data science. With the Python language and the *pandas* library, this project visualized mathematical results in Google Colaboratory for three rides and two parks. The scatterplot visuals revealed patterns with the rides, the crowds, and the hour exclusively reserved for guests at Disney resorts. This project demonstrates an applied angle of coding and mathematics with real benefits at a theme park. The project is low risk, and only applicable to people with the socio-economic freedom to visit Walt Disney World, but the project itself is meant to be accessible and fun.

Both the park and the guest are responsible for variables contributing to their wait times. Working with these variables requires familiarity with queuing theory, the time dependent traveling salesman problem, and David Maister's first law of service: satisfaction = perception – expectation. Together, the TouringPlans community work to build better, more efficient vacations for themselves.

As a community-oriented institution, libraries are prepared to meet people where they are—whether that means creating a computer science project, or helping plan a family vacation. Libraries are versatile enough to accommodate both. The Libraries Transform initiative also supports different cultures, including popular culture. Libraries are transforming "because superfans are welcome," and "fandoms thrive at the library" (*Libraries Transform*, n.d.). TouringPlans is essentially a group of superfans in the Disney fandom working together in an open-ended challenge, creating applicable results with creative problem solving.

Planning a trip is part of the wait, and the anticipation continues throughout the journey. Families and friends create memories while they are on vacation, and those memories can be little miracles upon return to daily life. Those miracles are a worthy investment within the TouringPlans community, and they prove true the words of Cinderella's Fairy Godmother, "Even miracles take a little time."

Bibliography

Advani, V. (2020, May 1). *34 Open-Source Python Libraries You Should Know About*. GreatLearning. https://www.mygreatlearning.com/blog/open-source-python-libraries/

ALA Libraries Transform | An initiative of the American Library Association. (n.d.). Retrieved August 23, 2020, from http://www.ilovelibraries.org/librariestransform/

Barnes, B. (2020, July 13). Hong Kong Disneyland to Close Again, Days After Disney World Reopens. *The New York Times*. https://www.nytimes.com/2020/07/13/business/hong-kong-disneyland-closing.html

Barton, R. (2020, June 9). New Disney Genie Could Be the App for Disney Park Reservations. *Inside the Magic*. https://insidethemagic.net/2020/06/disney-genie-new-reservation-system-rwb1/

Black Girls Code, BlackGirlsCode, Women of Color in Technology. (n.d.). BlackGirlsCode. Retrieved August 25, 2020, from https://www.blackgirlscode.com/

Daniels, E. C. (2016). *The perception of time waiting in theme park queue lines* [M.A., Michigan State University]. http://search.proquest.com/docview/1792069046/abstract/EFD4AD0E77364493PQ/1

Disney World App | iPhone, Android, Mobile Web. (n.d.). Retrieved July 20, 2020, from https://touringplans.com/disney-world-app

Disney World Theme Park Tickets in Orlando, Florida | Walt Disney World Resort. (n.d.). Retrieved July 16, 2020, from https://disneyworld.disney.go.com/admission/tickets/

Fortenberry, T. (n.d.). *Python Success Stories*. Python. Retrieved August 19, 2020, from https://www.python.org/about/success/ilm/

General Python FAQ. (n.d.). Python. Retrieved August 19, 2020, from https://docs.python.org/3/faq/general.html

Google Colaboratory. (n.d.-a). Retrieved August 20, 2020, from https://colab.research.google.com/notebooks/welcome.ipynb

Google Colaboratory. (n.d.-b). Retrieved August 20, 2020, from https://colab.research.google.com/notebooks/basic_features_overview.ipynb

Gorse, C. G., Johnston, D. J., & Pritchard, M. P. (2020). Queuing theory. In C. Gorse, D. Johnston, & M. Pritchard (Eds.), *A Dictionary of Construction, Surveying and Civil Engineering*. Oxford University Press. http://www.oxfordreference.com/view/10.1093/acref/9780198832485.001.0001/acref-9780198832485-e-5517

Healey, K. (n.d.). *What will your pleasure be? "Disney Genie" app will help plan your Orlando vacation*. WFTV. Retrieved July 20, 2020, from https://www.wftv.com/news/local/what-will-your-pleasure-be-disney-genie-app-will-help-plan-your-orlando-vacation/979736916/

Katz, K. L., Larson, B. M., & Larson, R. C. (1991). Prescription for the Waiting-In-Line Blues: Entertain, Enlighten, and Engage. *Sloan Management Review; Cambridge*, *32*(2), 44.

Ledbetter, J. L., Mohamed-Ameen, A., Oglesby, J. M., & Boyce, M. W. (2013). Your Wait Time From This Point Will Be . . .: Practices for Designing Amusement Park Queues. *Ergonomics in Design*. https://doi.org/10.1177/1064804613477100

Lee, D. (2004, December 3). Disney by the Numbers. *The New York Times*. https://www.nytimes.com/2004/12/03/travel/escapes/disney-by-the-numbers.html

Liang, Z.-X., & Dong, G.-Z. (2011). A study on tourist psychological capacity of theme parks and its influencing factors: A case study of Happy valley in Shenzhen. *Human Geography*, *26*(2), 139–143.

Magic Kingdom Park. (n.d.). Walt Disney World. Retrieved July 20, 2020, from https://disneyworld.disney.go.com/destinations/magic-kingdom/

Maister, H. (1985). The Psychology of Waiting. *In the Service Encounter, Eds*.

Norman, D. (n.d.). *The Psychology of Waiting Lines*. Dropbox. Retrieved July 18, 2020, from https://www.dropbox.com/s/bhvj4xa4iqq2198/Norman%20The%20Psychology%20of%20Waiting%20Lines.pdf?dl=0

pandas—Python Data Analysis Library. (n.d.). Retrieved August 20, 2020, from https://pandas.pydata.org/about/

Peters, T. (2004, August 19). *PEP 20—The Zen of Python*. Python. https://www.python.org/dev/peps/pep-0020/

Phased Closures Due to Capacity Constraints | Walt Disney World. (n.d.). Retrieved July 18, 2020, from https://touringplans.com/walt-disney-world/capacity-closures

Rock "n" Roller Coaster® Starring Aerosmith | Hollywood Studios Attractions. (n.d.). Walt Disney World Resort. Retrieved August 19, 2020, from https://disneyworld.disney.go.com/attractions/hollywood-studios/rock-and-roller-coaster-starring-aerosmith/

Rubin, J. (Ed.). (2019). *TEA/AECOM 2018 Theme Index and Museum Index: The Global Attractions Attendance Report*. Themed Entertainment Association (TEA). https://www.teaconnect.org/images/files/TEA_328_381804_190528.pdf

Sanders, S. (2019a, October 13). Everything You Need to Know About Toy Story Mania! *TouringPlans Blog*. https://touringplans.com/blog/2019/10/13/everything-you-need-to-know-about-toy-story-mania/

Sanders, S. (2019b, December 24). Everything You Need to Know About Rock "n" Roller Coaster Starring Aerosmith. *TouringPlans Blog*. https://touringplans.com/blog/2019/12/24/everything-you-need-to-know-about-rock-n-roller-coaster-starring-aerosmith/

Sanders, S. (2020, June 28). Everything You Need to Know About Spaceship Earth. *TouringPlans Blog*. https://touringplans.com/blog/2020/06/28/everything-you-need-to-know-about-spaceship-earth/

Spaceship Earth | Epcot Attractions. (n.d.). Walt Disney World Resort. Retrieved August 19, 2020, from https://disneyworld.disney.go.com/attractions/epcot/spaceship-earth/

Stanford, D. (2013, December 13). *How Queueing Theory Can Improve Wait Times*. YouTube. https://www.youtube.com/watch?v=SRqI_Ntrcnc

Stanton, J. M. (2012, July 16). *Data Science: What's in it for the New Librarian?* ISchool | Syracuse University. https://ischool-dev.syr.edu/data-science-whats-in-it-for-the-new-librarian/

Testa, L., Esterline, A. C., & Dozier, G. V. (1999). Evolving efficient theme park tours. *Journal of Computing and Information Technology*, *7*(1), 77–92.

The DIS. (2019, August 25). *Disney Genie Announcement | D23 Expo 2019*. YouTube. https://www.youtube.com/watch?v=nQ1YD_TWvew&list=TLPQMjEwNzIwMjAR046IBwPHoA&index=1

Touring Plans. (n.d.). Retrieved July 20, 2020, from https://touringplans.com/

Toy Story Mania. (n.d.). Walt Disney World Resort. Retrieved August 19, 2020, from https://disneyworld.disney.go.com/attractions/hollywood-studios/toy-story-mania/

Tschinkel, A. (2018, August 3). *Disney rides with the longest wait*. Insider. https://www.insider.com/disney-rides-with-the-longest-wait-2018-8

Waskom, M. (2012). *An introduction to seaborn—Seaborn 0.10.1 documentation*. Seaborn. https://seaborn.pydata.org/introduction.html

Weinberger, M. (2019, June 2). *Meet the computer scientist using artificial intelligence to help 140,000 paying customers plan the perfect Disney vacation*. Business Insider. https://www.businessinsider.com/touringplans-disney-world-len-testa-interview-2019-5

Zhang, Y., Li, X. (Robert), & Su, Q. (2017). Does spatial layout matter to theme park tourism carrying capacity? *Tourism Management*, *61*, 82–95. https://doi.org/10.1016/j.tourman.2017.01.020

Zhang, Y., Li, X. (Robert), Su, Q., & Hu, X. (2017). Exploring a theme park's tourism carrying capacity: A demand-side analysis. *Tourism Management*, *59*, 564–578. https://doi.org/10.1016/j.tourman.2016.08.019

Zika, J. (2018). Dark rides and the evolution of immersive media. *Journal of Themed Experience and Attraction Studies*, *1*(1), 54–60.