Disney World Attractions
An Analysis of Wait Times and Sentiment

Kevin Lee
Computer Information Systems
Technology & Operations Management
Mentor: Dr. Rita Kumar
CBA RSCA Showcase

CAL POLY POMONA

CAL POLY POMONA CBA COLLEGE OF BUSINESS ADMINISTRATION

Abstract

As the availability of data from an expanding array of data streams grows exponentially, organizations and analysts are now increasingly reliant on data analytics techniques to generate insights for informed decision-making and to spearhead process improvements and innovations. The amusement park industry is no different: timely data analysis can be indicative of current conditions, influencing the daily operations and allocation of resources within Disney World and its selection of attractions and rides. Indeed, previous research applications of data analysis have led to mobile apps such as Lines, which utilize information such as historical wait times and live location data from app users to predict wait times with astonishing accuracy. The implications of such tools are extensive. To start, tourists can more easily plan trips by strategically minimizing wait times; Disney itself can utilize this data to manage demand and proactively plan services. In this study, we analyze a dataset of wait times and various other conditions from 2012-2019. Specifically, we employ data visualization and several analytic techniques to understand the overall trend and the various factors impacting Disney World's wait times. In addition, we combine these factors with time series forecasting techniques to predict wait times at a daily level.

Objective

- **Examine and analyze wait times for a selection of Disney World attractions**
  - **Determine the average Disney posted wait time, actual wait time, and the differences between the two**
  - **Understand the impact of certain conditions, such as time of year, holidays, events, etc. on wait time**
  - **Forecast wait times based on trend and seasonality**
  - **Extract the sentiment (feelings or attitudes towards something) on selection of rides**
  - **Summarize and relay data and findings in an interactive dashboard**

Method

- **Utilized a dataset provided by touringplans.com from Jan 2012 – Apr 2019 on 14 Disney World attractions**
- **Preprocessed and transformed the data; aggregated data from the original hourly level to the daily level using Python and MS Excel**
- **Forecasted wait times a year after the given timeframe with regression, taking into account general trend and seasonal components**
- **Extracted user-generated tweets and preprocessed Text using Python, then analyzed sentiment of text using TextBlob and VADER sentiment analyzers**
- **Visualized the results in Tableau and shared to the public via Tableau Public Server**

Excel

tableau SOFTWARE

python

Notable Data Fields

- **Date (Daily)**
- **Disney's Posted Wait Time**
- **Actual Wait Time**
- **Ride**
- **Ticket Type (Differential Seasonal Ticketing)**
- **Season**
- **Holidays**
- **Events**
- **Tweets**
- **Sentiment Scores**

Visualizations & Findings

Figure 1: Background Dashboard w/ Timeline of Rides, Background Information, and Links to Additional Dashboards | (Click for Link)

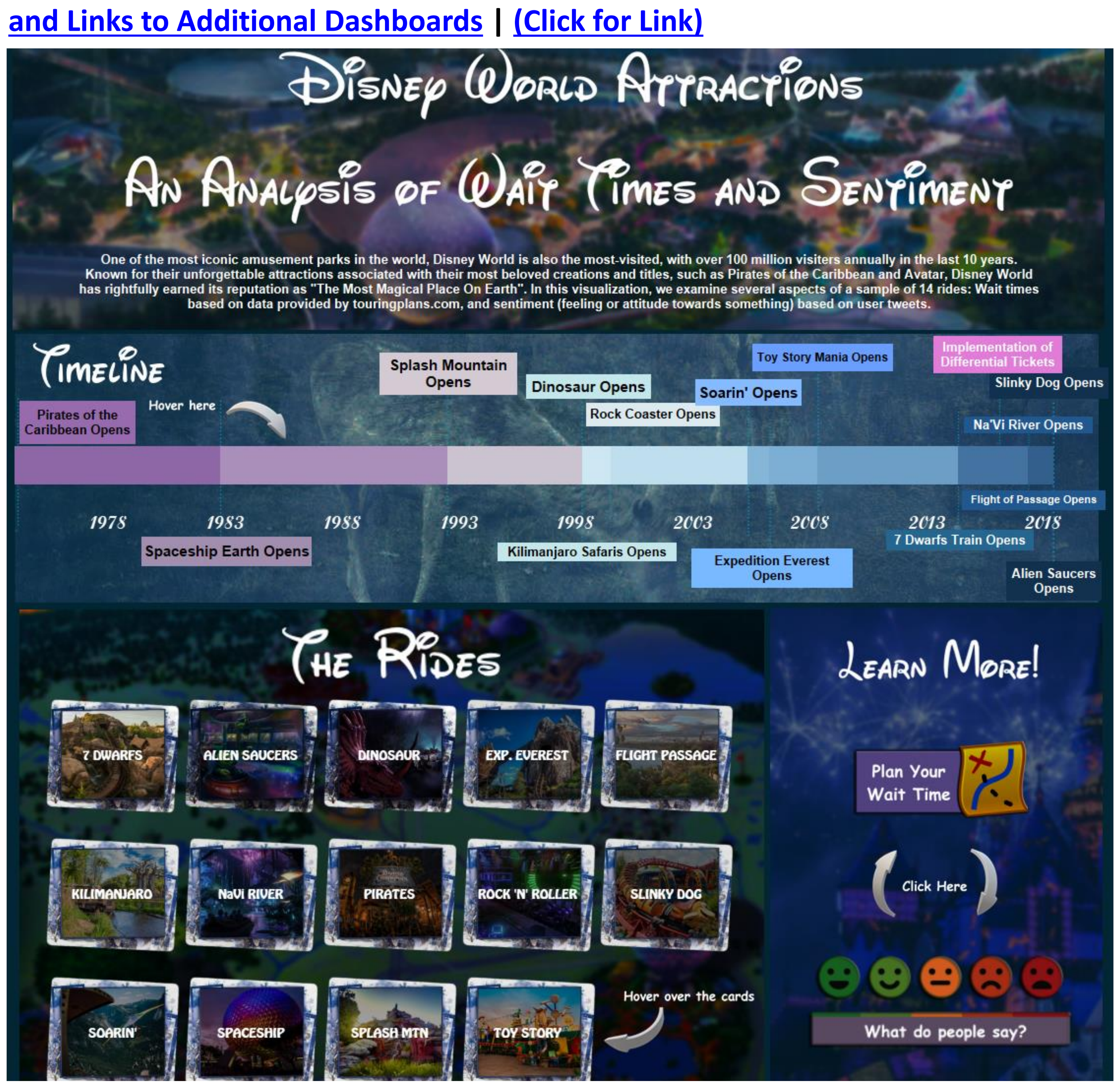


Figure 2: Average Posted and Actual Wait Times and Their % Difference, with selection of available filters (Season, Holiday, Events, etc.) | (Click for Link)

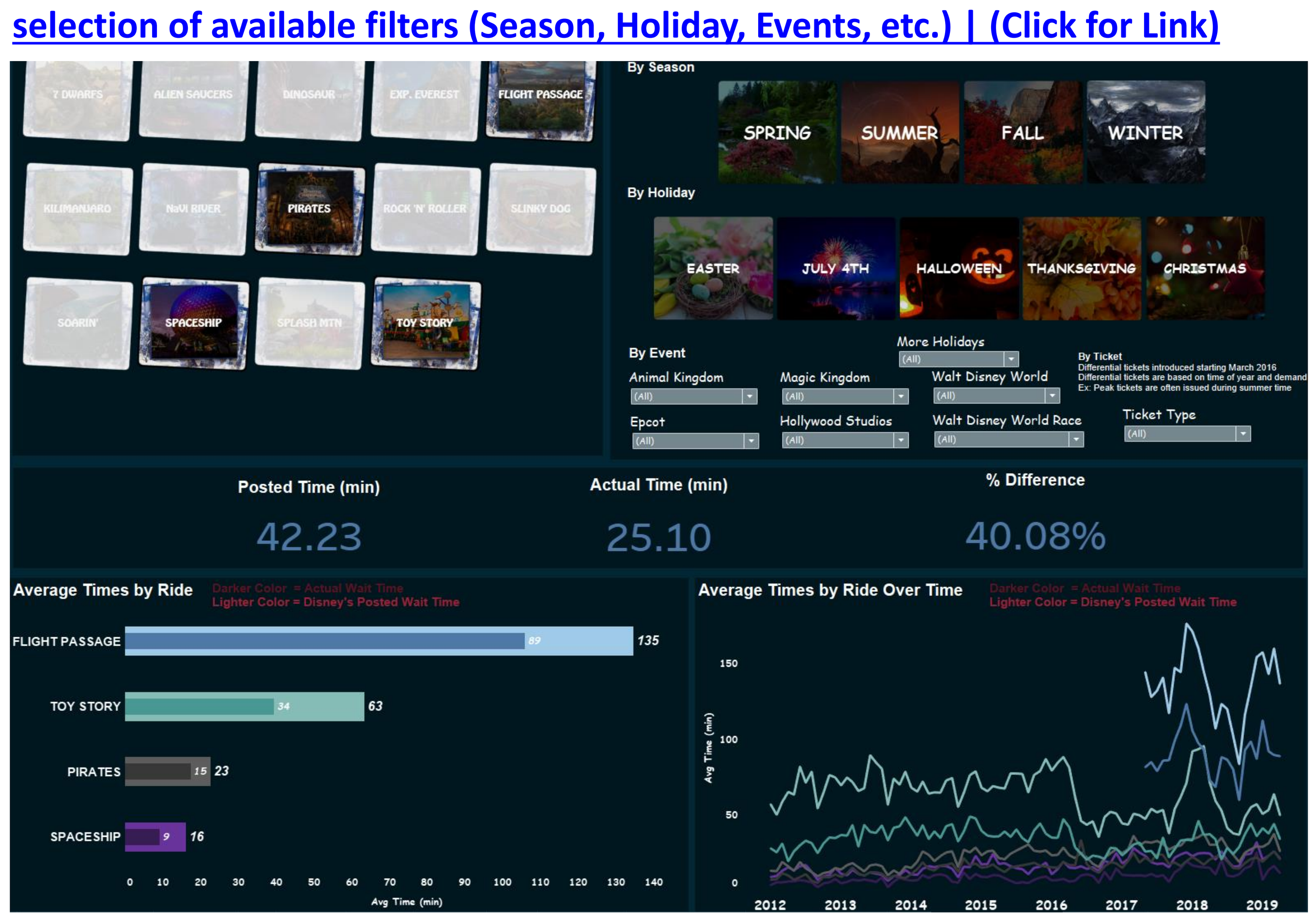


Figure 3: Average Posted and Actual Wait Times of All Rides in Dataset | (Click for Link)

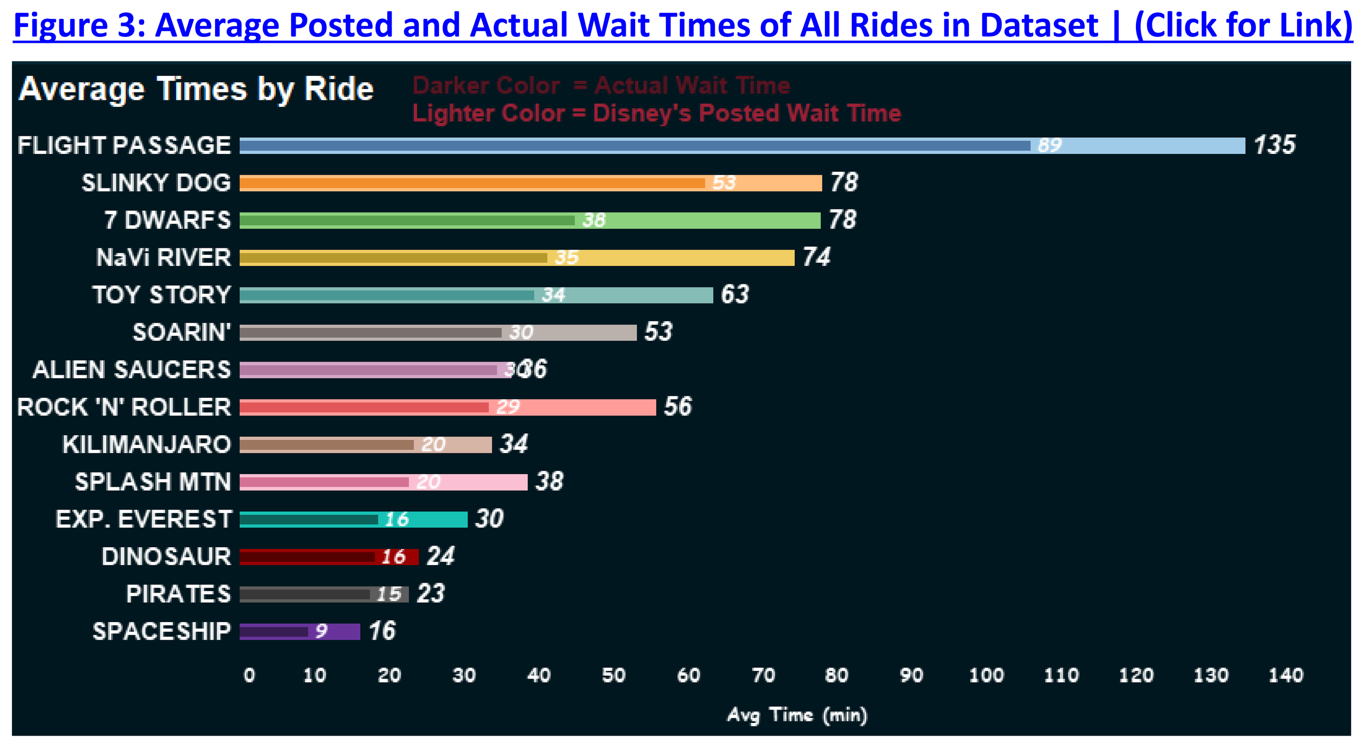
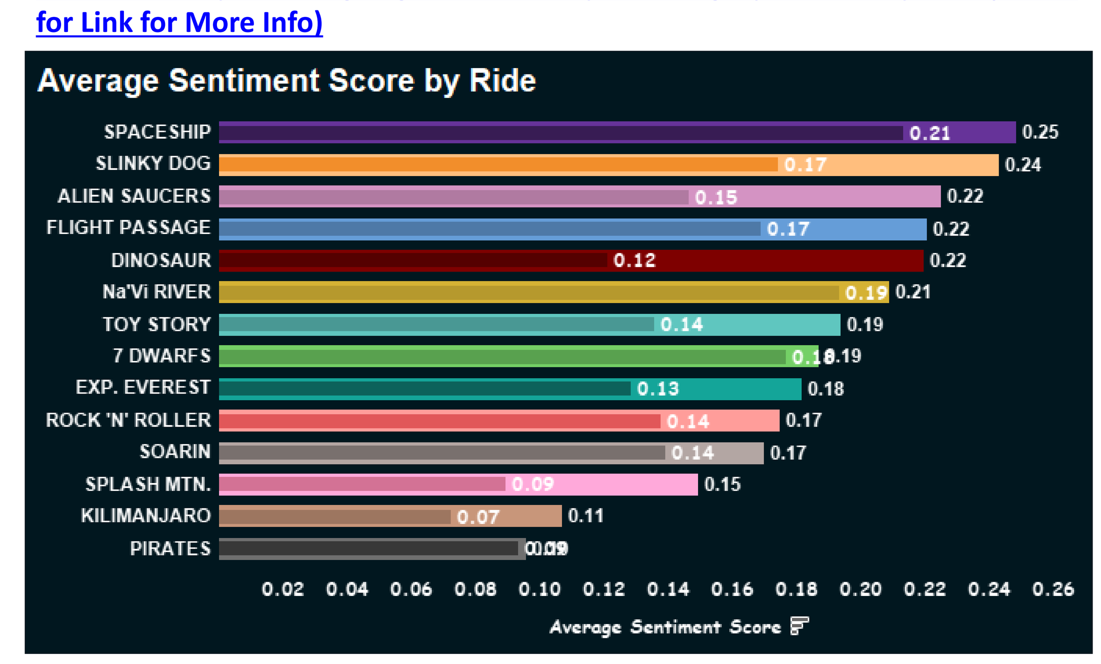


Figure 4: Linear Forecast of Navi River Ride: Decomposition of Trend (daily) and Seasonality (monthly), Comparison Between Posted and Actual Wait Time. This Color = Posted Time | This Color = Actual Time

Variable	R-Square	P-Value	Coefficient
Period (Daily: Increase by 1 for every day progressed in dataset)	0.105	1.52E-18	-0.025
Period (Daily: Increase by 1 for every day progressed in dataset)	0.017	0.001	-0.014

Month	Average (Posted Time)	Seasonal Factor (Posted Time)	Average (Actual Time)	Seasonal Factor (Actual Time)
1	76.67	1.02	39.63	1.13
2	76.05	1.01	33.44	0.95
3	80.76	1.08	38.38	1.09
4	74.00	0.99	35.90	1.02
5	69.51	0.93	35.55	1.01
6	70.43	0.94	32.89	0.93
7	63.56	0.85	30.69	0.87
8	63.07	0.84	33.35	0.95
9	61.11	0.81	32.21	0.91
10	78.87	1.05	39.69	1.13
11	88.70	1.18	35.98	1.02
12	94.50	1.26	35.05	0.99

Figure 5: Average Sentiment Score on User-generated tweets for each ride based on two sentiment analysis methods: TextBlob and VADER. Score ranges from [-1,1] with -1 representing negative and 1 representing a positive response | (Click for Link for More Info)



Implications

Actual and posted wait times have a high disparity for most rides, with average posted wait times consistently higher than actual wait times. Both time figures follow similar patterns over time: highly variable from a daily perspective, but relatively constant month-by-month, and notably impacted by external factors such as seasons and certain holidays. Timely analysis of current conditions such as number of people in line, equipment failures, etc. to accurately relay wait times greatly benefits the customer experience. Improving general forecast of demand and wait times using additional time-based data and events helps Disney plan ahead of time and allocate operations and resources accordingly. Extracting user-generated content to understand customer sentiment may help improve Disney's service offerings based on insight gained.

Data Sources:
https://touringplans.com/walt-disney-world/crowd-calendar#DataSets
Tableau Dashboard
https://public.tableau.com/profile/kevindlee98#!/vizhome/Disney_Rides_Background_WIP/BackgroundDashboard