# **Project Title:** Wait Less, Play More: Data-Driven Itinerary Planning

**Date:** 1/16/25

**Project Overview:** Automatically create the best itinerary for your Disney park of choice depending on wait times, popular times, your start time, and party size.

Idea: The goal of this project is to create a **Smart Park Planner** that helps visitors to Disney parks optimize their schedules based on predicted wait times, visitor flow patterns, and personalized preferences. By using machine learning, we aim to provide real-time itinerary recommendations that maximize enjoyment and minimize downtime.

We will focus on Disneyworld’s Epcot (Florida), Disneyland (California), Disney Paris, and Disney Hong Kong. If through data exploration and research, we are unable to find data for these specific locations, we may modify our project accordingly.

**Members and Initial Task Assignments:**

* Brian Hester
  + Weather data from Disney Park cities (Historical) to go alongside heatmap
* Israel Flores
  + Scraping
* Jenny Liu
  + Search for lat/longs of each ride/Disney park
  + Visualizations - leaflet/heat maps
  + HTML/CSS
* Jimmy Lee
  + Classification Model for showing High and Low Traffic Areas
  + Powerpoint
* Morgan Bee
  + Regression Model for Predicting Wait Times
* Kiley Karecki
  + Scraping

**Project Timeline:**

**1/22/25:** Datasets to use finalized and downloaded appropriately

**1/23/35:** API calls completed, data cleaned, building ML models and visualizations

**1/27/25:** All coding and visualizations completed by end of class

**1/29/25:** Finalize presentation and submit to BCS

**1/30/25:** Presentation

**Potential Datasets To Use:**

<https://www.thrill-data.com/waits/park/dlr/disneyland/>

<https://www.thrill-data.com/waits/park/cedar-fair/knotts-berry-farm/graphs/current/ds/high>

<https://www.waltsapartment.com/disneyland-ticket-prices-through-the-years/>

<https://www.thrill-data.com/waits/park/uh/universal-studios-hollywood/#typical-waits-area>

<https://queue-times.com/en-US>

**Accessing Historical Data:** If we are unable to find a suitable historic dataset in .csv format, we will scrape the data from the queue-times.com website listed above to collect at least a year of historical data in order to create, fit, and predict our model.

**Implementing the Model:** We will either use an API to pull real-time wait times from each park, or we will use our predictive model to show what our estimated wait time for any given ride in the park is.

**Two machine learning models for this project:**

Predicting Wait Times (Regression Model)

* **Problem:** Help park visitors predict ride wait times based on factors such as the time of day, day of the week, weather, season, or special events.
* **Approach:**
  + Use historical wait time data for various attractions. Train a regression model (e.g., Linear Regression, Random Forest Regressor) to predict wait times for specific rides.
  + Output: A recommended route based on predicted low wait time slots.

### Peak Time Prediction (Classification Model)

* **Problem**: Predict whether a ride or area of the park will experience high traffic at a given time.
* **Approach**:
  + Use historical data to train a classification model (e.g., Decision Trees, Logistic Regression) to label times as “high traffic” or “low traffic.”
  + Output: A heat map of peak and off-peak times for park areas.

The two models can be used simultaneously so that you can first narrow down the high or low-traffic areas of the park and then pick specific rides based on wait time when you arrive at the line. Overall, our models will help parkgoers to wait in lines for the least amount of time, maximizing their ride minutes and extending the value of their theme park tickets.

**Presentation of Data and Visualization Plans:**

1. **Park Maps:** If we are able to locate coordinates for each theme park’s rides, we will create an interactive map that shows users exactly where the model is suggesting the optimal next ride.
2. **Heatmap of Park Business:** Show High Traffic and Low Traffic rides based on user input for when they want to visit the park, and if applicable, the time of day they are considering.
3. **Wait Time Comparison Across Disney Parks:** Side-by-side comparison of each Disney Park’s wait times around the world, paired with weather data to demonstrate the best and worst times of year to visit any particular park.

**Final Data Sets:**

* Weather:
* Lat/Longs for each ride:
* Historical wait times:
  + Disneyland (CA) <https://queue-times.com/en-US/parks/16/stats> (Overall Average Wait Times Table)
  + Disneyworld Epcot (FL) <https://queue-times.com/en-US/parks/5/stats> (Overall Average Wait Times Table)
  + Disneyland Paris <https://queue-times.com/en-US/parks/4/stats> (Overall Average Wait Times Table)
  + Disneyland Hong Kong <https://queue-times.com/en-US/parks/31/stats> (Overall Average Wait Times Table)
* Live wait times: <https://queue-times.com/en-US/pages/api>

**Steps for Project:**

1. Finalize data sets for use
2. Scrape historical wait times if needed
3. Clean and store weather and historical wait times
4. Create models
   1. Predicting Wait Times (Regression Model)
   2. Peak Time Predictions (Classification Model)
5. Create visualizations for presentation
   1. Leaflet Park Map (Connected to API for live feed wait times)
   2. Heatmap of Park Peak Times
   3. Wait Time Comparisons Across all Disney Parks
6. Create presentation using Gamma App