

# Capstone Project 1: Project Proposal

## **What is the problem you want to solve?**

In 2015 there were lots of flight delays in the United States. The U.S. Department of Transportation's (DOT) Bureau of Transportation Statistics tracks the status of the flights by providing a summary of the arrival delays, departure delays, scheduled departure, on-time arrival in their monthly report. The data is acquired from Kaggle:<https://www.kaggle.com/usdot/flight-delays>. The flight data set contains 5819079 rows with 31 columns and 14 different carriers.

## **Who is my client and why do they care about this problem? In other words, what will my client do or decide based on my analysis that they wouldn't have done otherwise?**

My clients are Airlines Passengers, vacationers, the Department of Transportation/National Aviation System.

No one enjoys spending extra time at the airport because flight delays have kept their plane grounded.

The analysis will help passengers to be mindful of the airlines that have the largest/lowest delays and cancellations; what days of the week are the best to fly. And what airports should they avoid? The analysis will also help The Department of

Transportation(DOT)/National Aviation System to implement regulations that will reduce delays and provide better protection for consumers.

Many U.S. airlines are now prohibited from allowing the domestic flight to remain on the Tarmac for more than three hours. Carriers are required to notify passengers of known delays and provide accommodations.

**What data are you using? How will you acquire the data?**

I am using this free dataset from Kaggle: 2015 Flight Delays and Cancellations: <https://www.kaggle.com/usdot/flight-delays>  
This data has 3 sources: airlines.csv, airports.csv, flights.csv.

**Briefly outline how you'll solve this problem.**

I am going to solve the problem by following these steps

- Process the data (data wrangling):
- Explore the data
- Perform in-depth analysis (machine learning, statistical models, algorithms):
- Communicate results of the analysis:

**What are your deliverables?** Code & Paper

