# Project Description

## DS4A – Team 16

### Shantel Culver, Tosin Adebote, Ashley Drewry, Carolina Pacheco,

### Melissa Flores, Jiaxuan Ji, Sylvia Nanyangwe

1. **Overview of an industry, business or problem** 
   1. The airline industry provides air transportation for passengers and cargo by using aircraft. Flight delays and cancellation can significantly affect customer experience and cause financial loss to the business. The total cost of delays from 2016-2019 was over US$23 billion. The aim of this project is to assist customers to make better decisions when booking flights and also help businesses mitigate financial loss due to these delays.
2. **Define the specific problem that should be solved** 
   1. How does prediction of delayed flights improve overall efficiency of the airline industry in the US market?
3. **Why does this problem matter?**
   1. The airline industry is more prone to operational risks such as delayed flights, which take tolls on both the business and customers. In order to improve the overall efficiency of the airline industry, our team is aiming to develop statistical models such as machine learning and optimization to predict the likelihood of delayed flights. Smart algorithms on delayed flights can help airline businesses reduce delayed flights. By achieving this goal, airline companies will not only minimize operational costs but also improve customer experience that can potentially expand their customer base in the long term.
4. **What are the datasets that you will consider to solve this problem?** 
   1. Airline Delay and Cancellation Data, 2009 – 2018
      1. Flight info of US domestic flights
      2. All data files are downloaded from [OST website](https://www.transtats.bts.gov/DL_SelectFields.asp?Table_ID=236&DB_Short_Name=On-Time), which stores flights on-time performance from 1987 to present.
      3. <https://www.kaggle.com/yuanyuwendymu/airline-delay-and-cancellation-data-2009-2018>
   2. January Flight Delay Prediction
      1. US flight data for the month of Jan 2019 and Jan 2020
      2. This data is collected from the Bureau of Transportation Statistics, Govt. of the USA
      3. <https://www.kaggle.com/divyansh22/flight-delay-prediction>
   3. Feb 2020 US Flight Delay
      1. US Flight Data for the month of Feb 2020
      2. The data is collected by the Office of Airline Information, Bureau of Transportation Statistics (BTS).
      3. <https://www.kaggle.com/rowhitswami/feb-2020-us-flight-delay>