**ADS-505 Team Project Form & Business Brief Templates Team Project Form**

Fill out this form and business brief and submit it by the end of Module 3 in Blackboard (2 pages max for each). Reference the file, “Final Project Business Brief Requirements.doc.”

Team Number: 4

Team Leader/Representative: Jose Luis Estrada

Full Names of Team Members:

1. Hanmaro Song

2. Eva Chow

3. Jose Luis Estrada

Title of Your Project: Airport Delays Analysis

Short Description of Your Project and Objectives:

There are a few things a company can utilize this dataset.

1. One can analyze the arrival delay time and see which dates and time there happens delay more than the other days. One way of interpreting this fact is there are too many planes in the airport that those waiting in the sky for their turns to land are wasting resources such as time and fuel among others. They could optimize the expense of operation by selecting an appropriate time of departure from other cities to NY City Airport.
2. Another way would be using departure delay time. A snack business, for example, can figure out what time of days are heavily delayed. With this finding, they can attract more customers by offering a some kind of promotion such as providing space and something to eat while they wait for their leave.
3. Can be merged with Air Quality of New York to see if that affects any delay
4. Can be merged with Taxi Fare to see any correlation between the two
5. Can be merged with Uber Pickup,Parking Tickets, and many more

Name of Your Selected Dataset and Programming Language: [New York City Airport Activity Dataset](https://www.kaggle.com/sveneschlbeck/new-york-city-airport-activity), Python

Description of Your Selected Dataset (source, number of variables, size of the dataset, etc.):

* The New York City Airport Activity Dataset was taken from Kaggle. It consists of airport activity data collected from the year 2013 from the airports in New York City for departing flights. The dataset consists of 16 variables and 32,735 observations. Total of 2.06 MB.
* Four object (string) dtypes within the dataset : carrier, tailnum, origin, and dest. Other features are in integer type
* Month and Day features are quite evenly distributed
* All columns have values without any missing or nan values.

Provide your team GitHub link here: <https://github.com/hsong1101/ADS-505-Project>

How many times have your team members met so far? Once

What was the agreed-upon method of communication? Are you using any teamwork project management software, such as [Deepnote](https://deepnote.com/), [Trello](https://trello.com/en-US), or [Asana](https://asana.com/)? If not, explain why?

* Slack will be the main form of communication. Will utilize Zoom meetings (for live communication) as well, if necessary.

Comments/ Roadblocks: Because the dataset contains quite a limited set of features, we may need to find third datasets to add more information.

**Team Project Business Brief**

**Purpose:**

The purpose of this two-part analysis is to identify delays on flights, and from the delayed flights, will there be a slight delay (mins/hours) or a long delay (hours/days). These outcomes will result in actionable items to recommend to different airlines and the NYC airport.

**Background:**

Delays result in a cost for the airlines and the airport. This analysis will help us identify which airlines have fewer delays and flight seasonality which can impact profit. Some of the expenses that delays can cause are paying for hotels for passengers and gate utilization for other flights. Additionally, airports have businesses such as souvenir stores and food stands that could increase revenue during these times if they are located in the right location.

**Current Situation:**

The three major airports from which data was collected are the John F. Kennedy, LaGuardia, and Newark Liberty International airports. From here, sixteen unique airline carriers serviced the three airports to 102 unique destinations. Data was collected from departure information for flights from the year 2013, with delays being noted for each unique flight. With delays extending to as long as a day, there is an opportunity for the airports and airlines to maximize potential revenue by strategically marketing additional services.

**Conclusion:**

If this problem can be solved by a model(s) to predict whether there will be a delay or not and if there is, for how long, a company can put up a business to attract those people in delay to get their products (drink, food, housing, etc) for profits. By having this predictive model, not only this can be used as-is but by merging with other external datasets, there are many other possibilities for companies.