**Dataset Description (Spring 2019 Version):**

This dataset contains n=194,833 observations of flight segment data collected by a major airline. Each row represents one flight segment by one airline customer. The dataset captures characteristics of the flight as well as characteristics of the customer. The dataset also contains a simple survey-based rating of each customer’s satisfaction with the flight segment.

**Attributes Name:**

1. **Satisfaction** – it is rated from 1 to 5, that how satisfied is the customer?
   1. 5 means higher satisfied, and 1 is lowest level of satisfaction.
2. **Airline Status** – each customer has a different type of airline status or package, which are platinum, gold, silver, and blue.
3. **Age** – the specific customer’s age. That is starting from 15 to 85 years old.
4. **Gender** – male or female.
5. **Price Sensitivity** – the grade to which the price affects to customers purchasing. The price sensitivity has a range from 0 to 5.
6. **Year of First Flight** – this attributes shows the first flight of each single customer. The range of year of the first flight for each customer has been started in 2003 until 2012.
7. **Flights Per Year** – The number of flights that each customer has taken in the most recent 12 months. The range starting from 0 to 100.
8. **Loyalty** – An index of loyalty ranging from -1 to 1 that reflects the proportion of flights taken on other airlines versus flights taken on this airline. A higher index means more loyalty.
9. **Type of Travel** – is provide three traveling purpose for each consumer, which are business travel, mileage tickets that based on loyalty card, and personal travel like to see the family or in vacation
10. **Total Frequent Flyer Accounts** – How many frequent flyer accounts the customer has.
11. **Shopping Amount at Airport** – The spending in dollars on non-food/drink goods and services at the airport(s) where the customer was before, between, or after flights.
12. **Eating and Drinking at Airport** – The spending in dollars on food/drink goods and services at the airport(s) where the customer was before, between, or after flights.
13. **Class** – it consisted of three different kinds of service level such as, business, and economy plus, economy. Moreover, customers have optional to choose their seat.
14. **Day of Month** – it means the traveling day of each costumer. In this attribute, shows total of 31 days of the month.
15. **Flight date** – all of these data are abbreviate the passenger’s flight date travel, which were since 2014 and only in January, February, and March.
16. **Partner Code** – This airline works with wholly- and partially-owned subsidiary companies to deliver regional flights. For example, AA, AS, B6, and DL.
17. **Partner Name** – These are the full names of the subsidiary airline companies. Pseudonyms have been substituted in place of the real names.
18. **Origin City** – refers to actual city that customers have departed from. For example, Yuma AZ, Waco TX, and Toledo HO.
19. **Origin State** – same thing as origin city such as, what state that customers have departed from? A good example, Texas, Ohio, Alaska, and Utah.
20. **Destination City** – the place to which passenger travels to. For example, Akron HO, Alpena MI, Austin TX, and Boston MA.
21. **Destination State** – also, it is the same thing as origin city, such as, to what state passenger travel to? Some example of destination states, Alaska, Kentucky, Iowa, and Florida.
22. **Scheduled Departure Hour** – the specific time at which passengers are scheduled to depart. In this data in scheduled departure hour is starting at 1 am until 23 pm.
23. **Departure Delay in Minutes** – which are minutes of departure delayed for each passenger, when compared to schedule. In this data the rage are starting from 0 until 1128 minutes.
24. **Arrival Delay in Minutes** – how many minutes of arrival delayed of each passenger. Rang of delayed minutes in this data are starting from 0 until 1115 minutes.
25. **Flight Cancelled** – occurs when the airline dose not operates the flight at all, and that is for a certain reason.
26. **Flight time in minutes** – indicate to period time to the destination.
27. **Flight Distance** – the extent of space between two places. Also, that means how many minutes are passenger traveling between two different places. Rang in this data starting from 31 until 4983 minutes.
28. **Arrival Delay greater 5 Minutes** – It means the delay of arrival airline time, which is more than 5 minutes per each passenger in the data.
29. **Long Duration Trip** – A Boolean variable that divides flight segments into two types: FALSE means a shorter duration segment (including average delays), TRUE means a longer duration segment.