**Airline Flight Delays**

Team members:

|  |  |  |
| --- | --- | --- |
| Cristina Iovu | Mwamba Mwape | Khuseyma Egaal |
| Daniela Shae-Bebeyi | Kirran Kayani | Chadi Ghosn |

Project outline:

The project will analyze the flights’ delays, cancellations and other statuses in the US for 322 different airlines. The dataset contains just above 5 million records with 31 columns. The findings will be presented to a potential airline company “Multico Air”, that want to enter the market, looking at flight delays in the industry to target a niche between the competitors. The presentation will target the GM, board of directors and the aviation team.

Research questions to answer:

1. How does the overall flight volume vary by month in 2015?
2. What percentage of flights experienced a departure delay in 2015? Among those flights, what was the average delay time, in minutes?
3. How does the delayed flights % vary throughout the year (monthly)?
4. What are the top 5 busiest origin airports in 2015? And what are the top 5 origin airports with the highest departure and arrival delays?
5. How many flights were cancelled every month for the top 5 origin airports?
6. How many flights were cancelled in 2015? What % of cancellations were due to weather? What % were due to the Airline/Carrier fault?
7. Which airlines seem to be most and least reliable, in terms of on-time departure in 2015?
8. Which airlines on average depart before the set schedule in 2015?
9. How does the distance of a trip have an influence (if any) on the arrival/departure delay of a flight?

Datasets to be used:

[Free Data Sets & Dataset Samples | Maven Analytics](https://www.mavenanalytics.io/data-playground?order=-fields.numberOfRecords) – Airline Flight Delays (2015 data)

Rough breakdown of tasks:

|  |  |  |
| --- | --- | --- |
| Cristina Iovu:   * Google graph with marker for the airports * Q4 * Q5 * G4 – Pie chart | Mwamba Mwape:   * Create power point presentation * Q7 * G7 – Bar chart | Khuseyma Egaal:   * Q1 * Q8 * G1 – Bar chart * G2 – Whisker box * Q9 + G9 |
| Daniela Shae-Bebeyi:   * Check for duplicates and null values * Q3 * G3 – Line chart | Kirran Kayani:   * Create word document explaining findings * Q2 * G8 – Scatter plot * Null statement testing | Chadi Ghosn:   * Ingest data, create data frame and common information * Power Bi report to visualize findings * Q6 * G6 – Line chart * G5 – Bar char |