**Report**

**Aviation Dataset Description:**

The dataset we have is containing 645299 Rows and 110 columns the dataset contains the missing values that have to be drop for the better visualization and to explore data better.

The Data Columns their Mean Median and mode Standard deviation Unique Values are as follows

* **Year**

|  |  |
| --- | --- |
| **Mean** | **2018** |
| **Median** | **2018** |
| **Min** | **2018** |
| **Max** | **2018** |
| **Standard Deviation** | **0** |
| **Unique Values** | **1** |
| **Missing Values** | **0** |

* **Quarter**

|  |  |
| --- | --- |
| **Mean** | **3** |
| **Median** | **3** |
| **Min** | **3** |
| **Max** | **3** |
| **Standard Deviation** | **0** |
| **Unique Values** | **1** |
| **Missing Values** | **0** |

* **Month**

|  |  |
| --- | --- |
| **Mean** | **7** |
| **Median** | **7** |
| **Min** | **7** |
| **Max** | **7** |
| **Standard Deviation** | **0** |
| **Unique Values** | **1** |
| **Missing Values** | **0** |

* **DayofMonth**

|  |  |
| --- | --- |
| **Mean** | **16.2818** |
| **Median** | **16** |
| **Min** | **1** |
| **Max** | **31** |
| **Standard Deviation** | **8.8428** |
| **Unique Values** | **31** |
| **Missing Values** | **0** |

* **DayOfWeek**

|  |  |
| --- | --- |
| **Mean** | **3.9036** |
| **Median** | **4** |
| **Min** | **1** |
| **Max** | **7** |
| **Standard Deviation** | **2.0795** |
| **Unique Values** | **7** |
| **Missing Values** | **0** |

* **FlightDate**

|  |  |
| --- | --- |
| **Unique Values** | **31** |
| **Missing Values** | **0** |

* **Reporting\_Airline**

|  |  |
| --- | --- |
| **Unique Values** | **17** |
| **Missing Values** | **0** |

* **DOT\_ID\_Reporting\_Airline**

|  |  |
| --- | --- |
| **Mean** | **19982.7207** |
| **Median** | **19977** |
| **Min** | **19393** |
| **Max** | **20452** |
| **Standard Deviation** | **374.223** |
| **Unique Values** | **17** |
| **Missing Values** | **0** |

* **IATA\_CODE\_Reporting\_Airline**

|  |  |
| --- | --- |
| **Unique Values** | **17** |
| **Missing Values** | **0** |

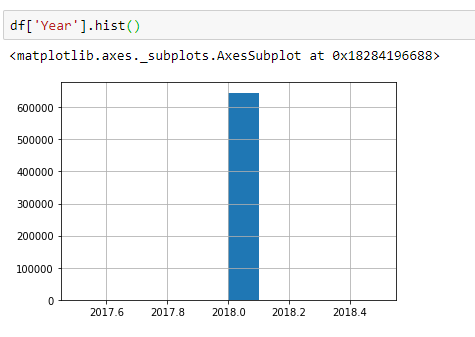
* **Tail\_Number**

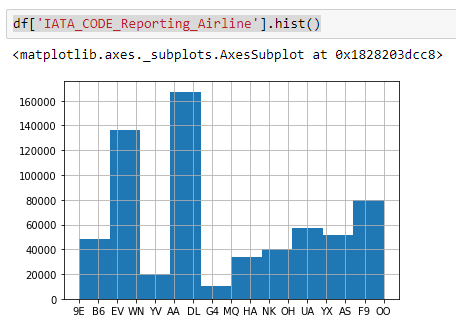
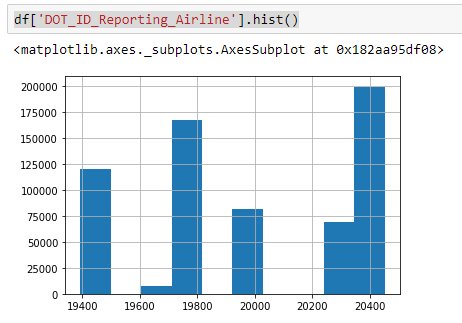
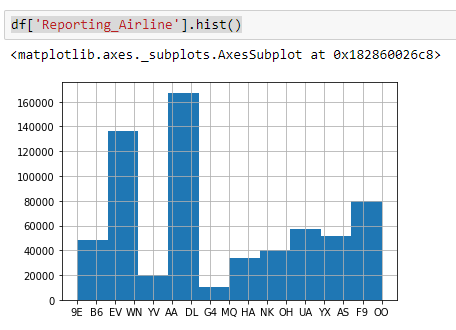
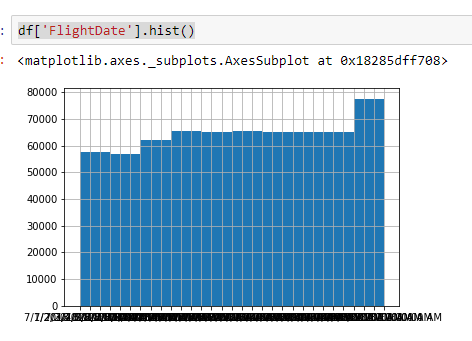
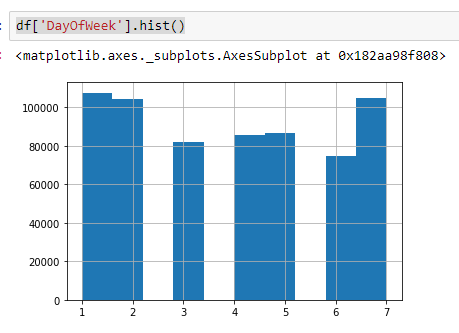
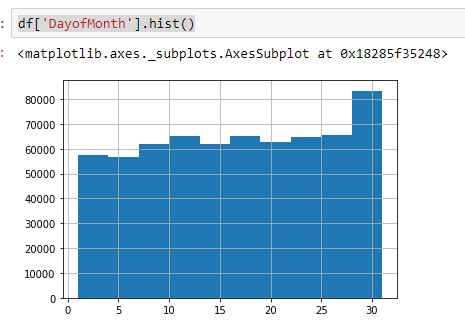
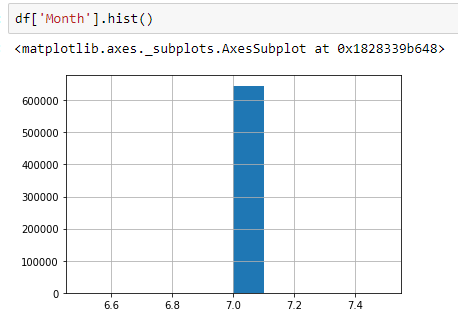
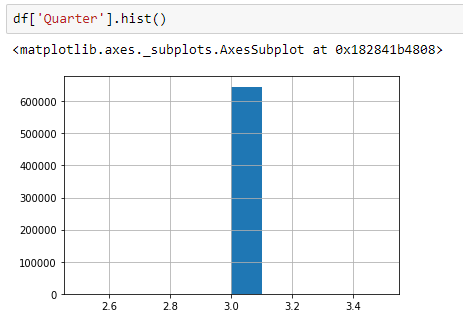
|  |  |
| --- | --- |
| **Unique Values** | **5388** |
| **Missing Values** | **0** |

* **Flight\_Number\_Reporting\_Airline**

|  |  |
| --- | --- |
| **Mean** | **2545.6994** |
| **Median** | **2122** |
| **Min** | **1** |
| **Max** | **7439** |
| **Standard Deviation** | **1818.7102** |
| **Unique Values** | **6495** |
| **Missing Values** | **0** |

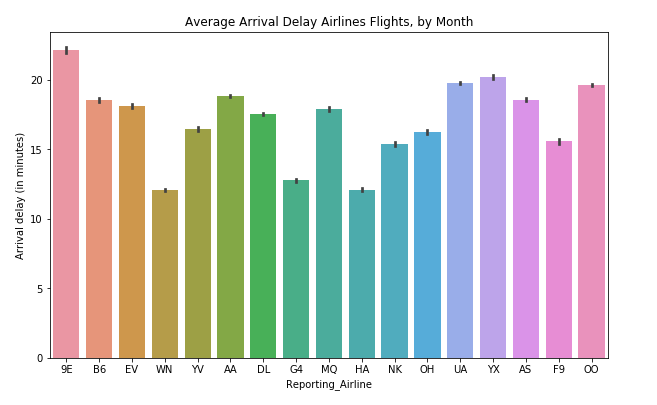
**The Histogram Visualization of the Data columns are as follows:**

****

****

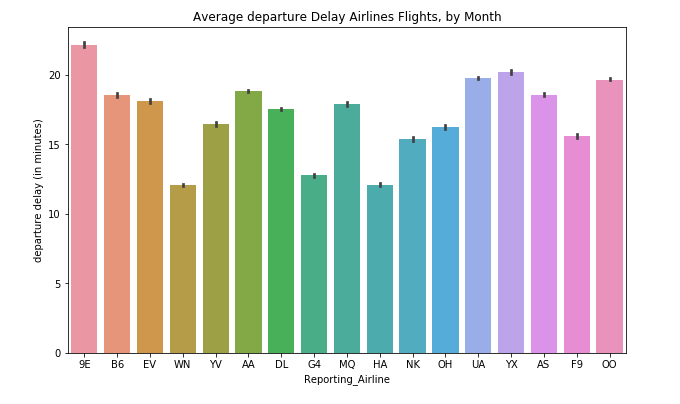
**Visualization of Data**

1. The first visualization to explore data is by comparing the average delay time of the flights with month



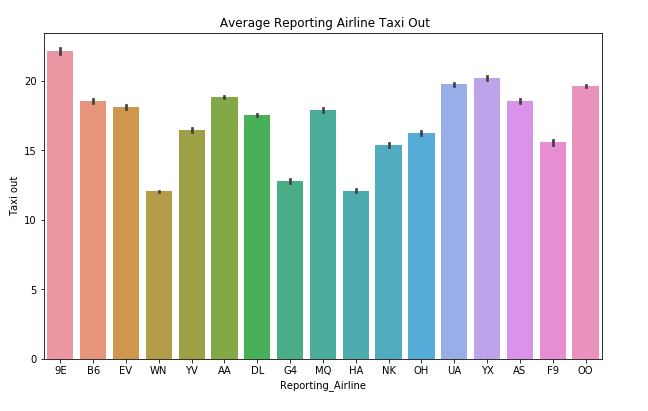
The insight we had is the on the average the highest delay time of the flight is 9e which is 27 minutes and the second most highest time of arrival delay is YX which is nearly 17 minutes

1. The second visualization is the average delay time of the departure of the flights as according to the month



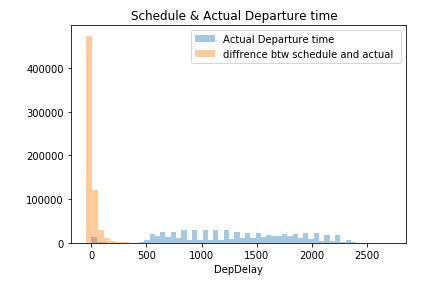
The highest average delay time of departure is of 9E that is almost 25 minutes and the second highest is YX which is nearly 17 minutes

1. Third visualization is of Average of the Reporting Airline taxi is as follows



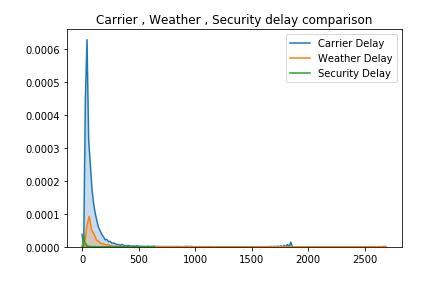
The highest taxi time is of 9E and the lowest is of HA

1. The fourth visualization to explore data is the difference between the schedule and the actual Departure time



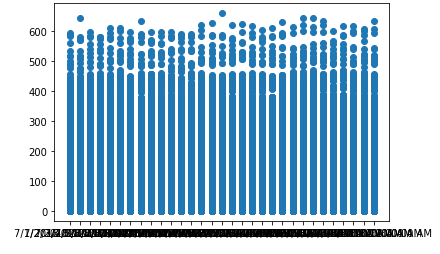
The difference from the visual is clear that the departure time is lying beneth whereas the schedule time is rising up

1. 5th visualization is of the comparison of the Carrier delay weather delay security delay



The highest going up curve is of the carrier delay and then weather delay and security delay

1. 6th visualization is of the time taken by the flights in air the visualization is as follows



Here we have date on the X axis and the time taken by the flight in air is on Y axis

1. 7th visualization is of the time taken by the flights

