

Airport Authority Data Analysis



POWERBI PROJECT

BY ELMA FORTUNATE PHIRI

Introduction

- This project aims to analyze the total number of flights (incoming and departing flights), the total number of flight delays (arrivals and departures), ground processing time
- This analysis could assist airport management authorities in making timely data-driven choices

Data Source

US Airline delays and cancellation dataset 2009-2018



Objectives

- Analyze the total number of incoming and departing flights to provide a complete picture of airport traffic.
- Examine the total number of flight delays, both for arrivals and departures, to understand the extent of disruptions.
- Assess ground processing time to identify potential bottlenecks and areas for optimization.



Overview

- Data Loading
- Data Transformation
- Time Series Analysis
- Trend Analysis
- Dashboard
- Insights
- Recommendation
- Reference and Acknowledgements



DATA LOADING

Get Data

All

Excel Workbook

Text/CSV

XML

JSON

Folder

PDF

Parquet

SharePoint folder

SQL Server database

Access database

SQL Server Analysis Services database

Oracle database

IBM Db2 database

IBM Informix database (Beta)

IBM Netezza

MySQL database

Certified Connectors

Template Apps

Connect

Cancel

2009.csv

File Origin

1252: Western European (Windows)

Delimiter

Comma

Data Type Detection

Based on first 200 rows

FL_DATE	OP_CARRIER	OP_CARRIER_FL_NUM	ORIGIN	DEST	CRS_DEP_TIME	DEP_TIME	DEP_DELAY	TAXI_OUT	WHEELS_O
01-01-2009	XE	1204	DCA	EWR	1100	1058	-2	18	
01-01-2009	XE	1206	EWR	IAD	1510	1509	-1	28	
01-01-2009	XE	1207	EWR	DCA	1100	1059	-1	20	
01-01-2009	XE	1208	DCA	EWR	1240	1249	9	10	
01-01-2009	XE	1209	IAD	EWR	1715	1705	-10	24	
01-01-2009	XE	1212	ATL	EWR	1915	1913	-2	19	
01-01-2009	XE	1212	CLE	ATL	1645	1637	-8	12	
01-01-2009	XE	1214	DCA	EWR	1915	1908	-7	9	
01-01-2009	XE	1215	EWR	DCA	1715	1710	-5	28	
01-01-2009	XE	1217	EWR	DCA	1300	1255	-5	15	
01-01-2009	XE	1218	DCA	EWR	1500	1457	-3	14	
01-01-2009	XE	1219	EWR	DCA	2135	2131	-4	21	
01-01-2009	XE	1220	CLE	DCA	1905	1855	-10	10	
01-01-2009	XE	1220	DCA	EWR	2100	2049	-11	10	
01-01-2009	XE	1232	ORD	EWR	905	900	-5	16	
01-01-2009	XE	1233	EWR	ORD	1000	1035	35	14	
01-01-2009	XE	1234	ORD	EWR	1230	1234	4	8	
01-01-2009	XE	1235	EWR	ORD	1343	1406	23	13	
01-01-2009	XE	1236	ORD	EWR	1630	1619	-11	19	
01-01-2009	XE	1237	EWR	ORD	1930	1927	-3	16	

Extract Table Using Examples

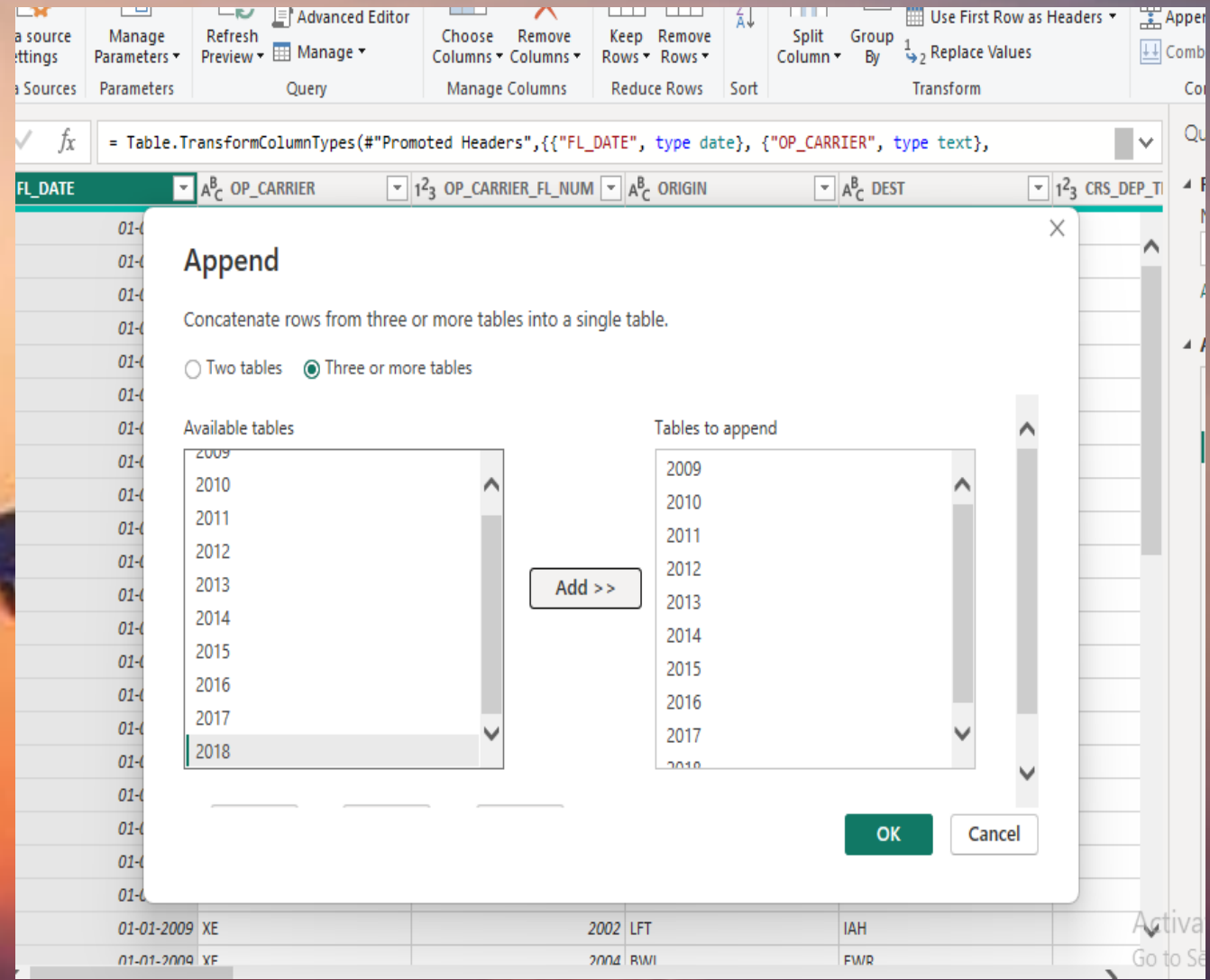
Load

Transform Data

Cancel

APPENDING DATASETS

The US Airline delays and cancellation dataset contains 10 different datasets grouped according to the years 2009-2018. We first had to append these datasets into one dataset to facilitate easy analysis of the data.



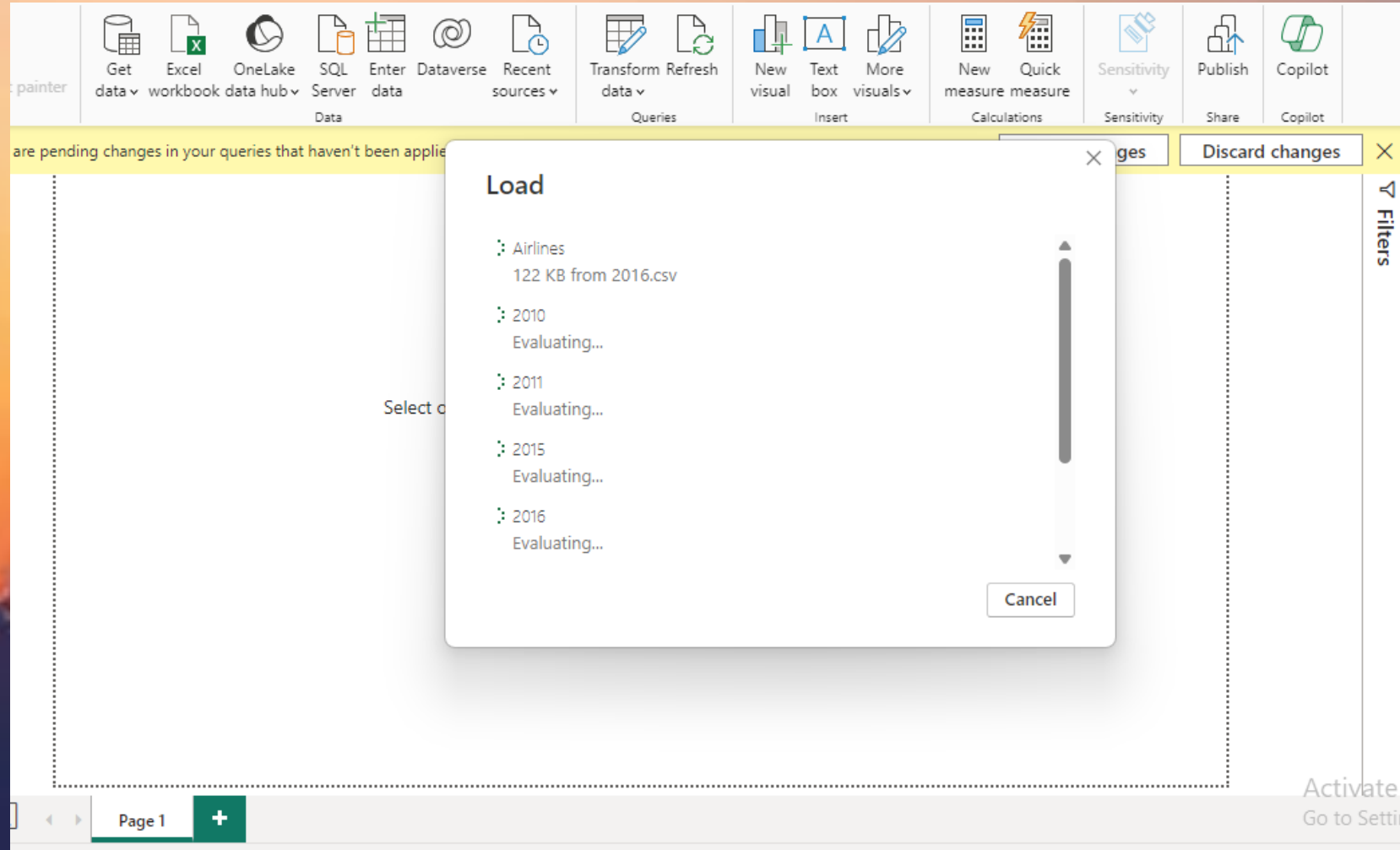
DATA TRANSFORMING

- Filtered Rows
- Removed unnecessary Columns
- Renamed Header Columns
- Replaced values

The screenshot displays the Microsoft Power Query Editor interface. The title bar indicates the data source is 'Airlines delays and cancellations'. The ribbon includes tabs for File, Home, Transform, Add Column, View, Tools, and Help. The Transform tab is active, showing various data manipulation options like 'Remove Columns', 'Keep Rows', 'Sort', 'Split Column', 'Group By', 'Replace Values', 'Merge Queries', 'Append Queries', and 'Combine Files'. The main area shows a table with columns: FL_DATE, Airline Identifier, Flight Number, ORIGIN, DEST, and Planned_Depature_Time. The table contains 21 rows of data. The right sidebar shows the 'Query Settings' for 'Airlines' and a list of 'APPLIED STEPS' including 'Filtered Rows', 'Changed Type', 'Filtered Rows1', 'Removed Columns', 'Renamed Columns', 'Replaced Value', 'Filtered Rows2', 'Replaced Value1', 'Filtered Rows3', 'Replaced Value2', 'Replaced Value3', 'Replaced Value4', 'Replaced Value5', 'Replaced Value6', 'Replaced Value7', and 'Replaced Value8'.

	FL_DATE	Airline Identifier	Flight Number	ORIGIN	DEST	Planned_Depature_Time
1	01-01-2009	JSX	1233	EWB	ORD	1000
2	01-01-2009	JSX	1235	EWB	ORD	1343
3	01-01-2009	JSX	2010	IAH	IAH	1415
4	01-01-2009	JSX	2042	IAH	BHM	1905
5	01-01-2009	JSX	2044	IAH	CLT	1735
6	01-01-2009	JSX	2060	IAH	CVG	1855
7	01-01-2009	JSX	2081	COS	IAH	605
8	01-01-2009	JSX	2088	MAF	IAH	1030
9	01-01-2009	JSX	2092	EWB	DTW	1400
10	01-01-2009	JSX	2118	CLE	ALB	1855
11	01-01-2009	JSX	2128	CLE	LGA	1546
12	01-01-2009	JSX	2145	OMA	IAH	1420
13	01-01-2009	JSX	2201	CLE	RIC	1859
14	01-01-2009	JSX	2212	IAH	LIT	1545
15	01-01-2009	JSX	2217	IAH	SAV	1348
16	01-01-2009	JSX	2236	MSP	EWB	1410
17	01-01-2009	JSX	2237	IAH	OKC	1940
18	01-01-2009	JSX	2249	IAH	MEM	1112
19	01-01-2009	JSX	2261	IAH	AEX	2045
20	01-01-2009	JSX	2281	IAH	MCI	1735
21	01-01-2009	JSX	2287	EWB	MEM	1520

- Data Loading after transforming it



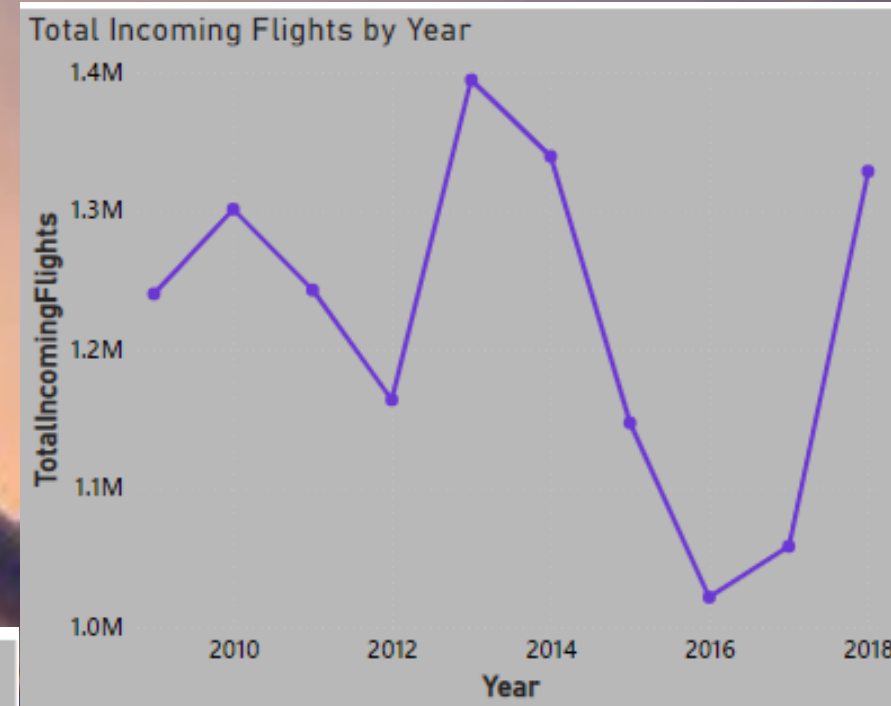
TIME SERIES ANALYSIS

- Analyzed the number of flights over time

Chart Type: Line Chart

Conclusion:

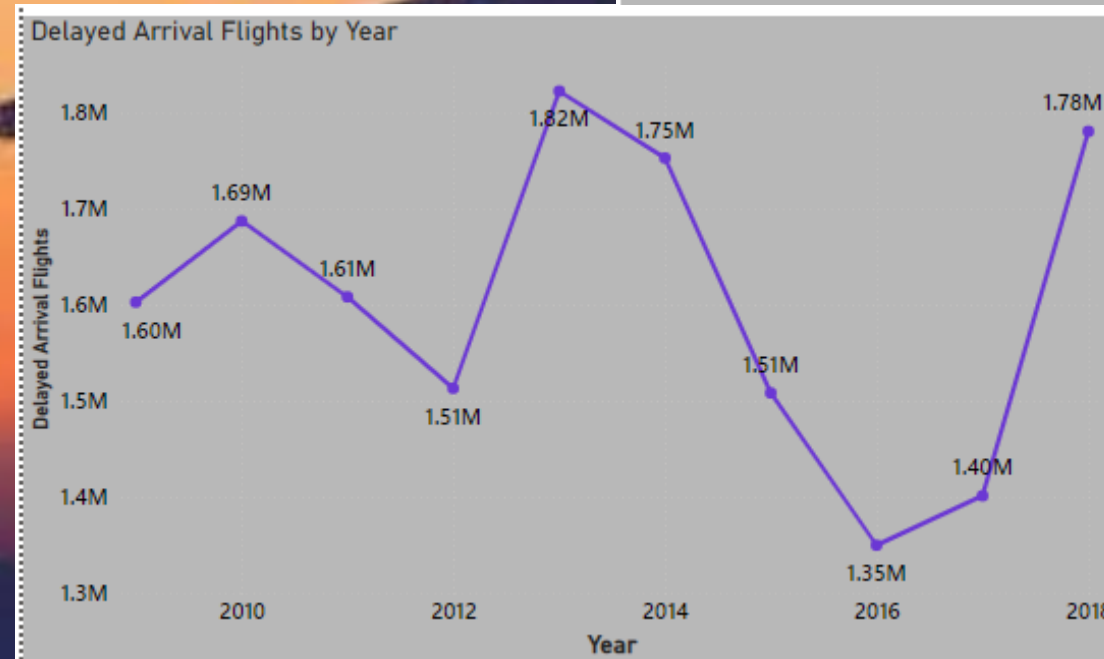
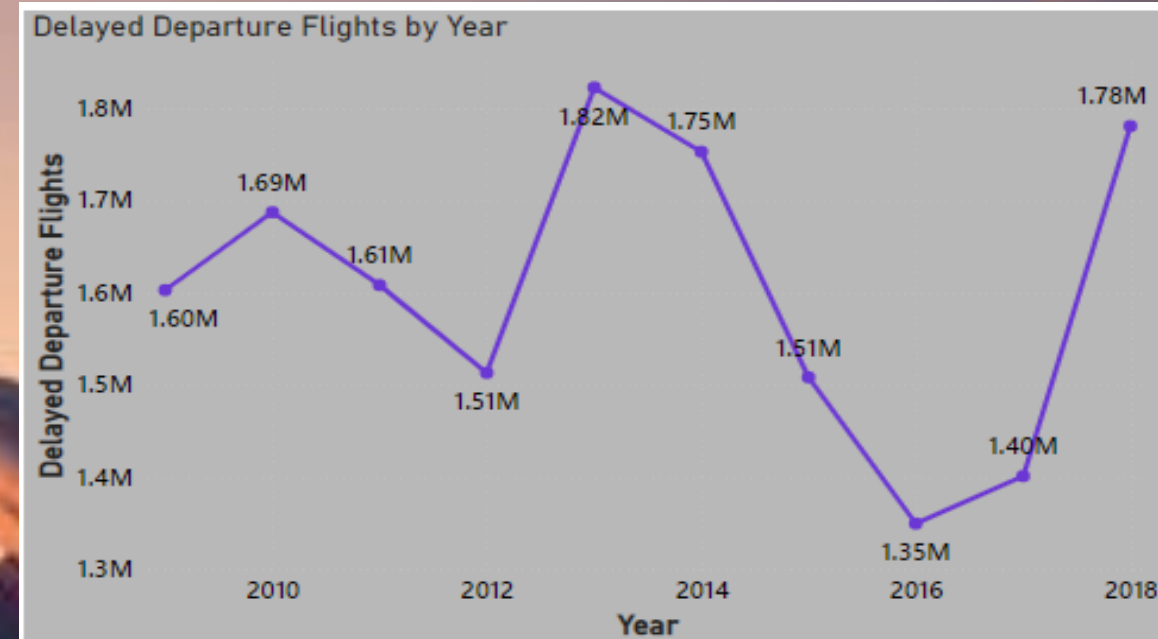
The highest number of incoming flights is 1 394 259 and departing flights is 826 690 both in the year 2013



- Analyzed the number of delayed flights over time

Chart Type: Line Chart

Conclusion:
The highest number of delayed arrival and departing flights is 1823059 both in the year 2013



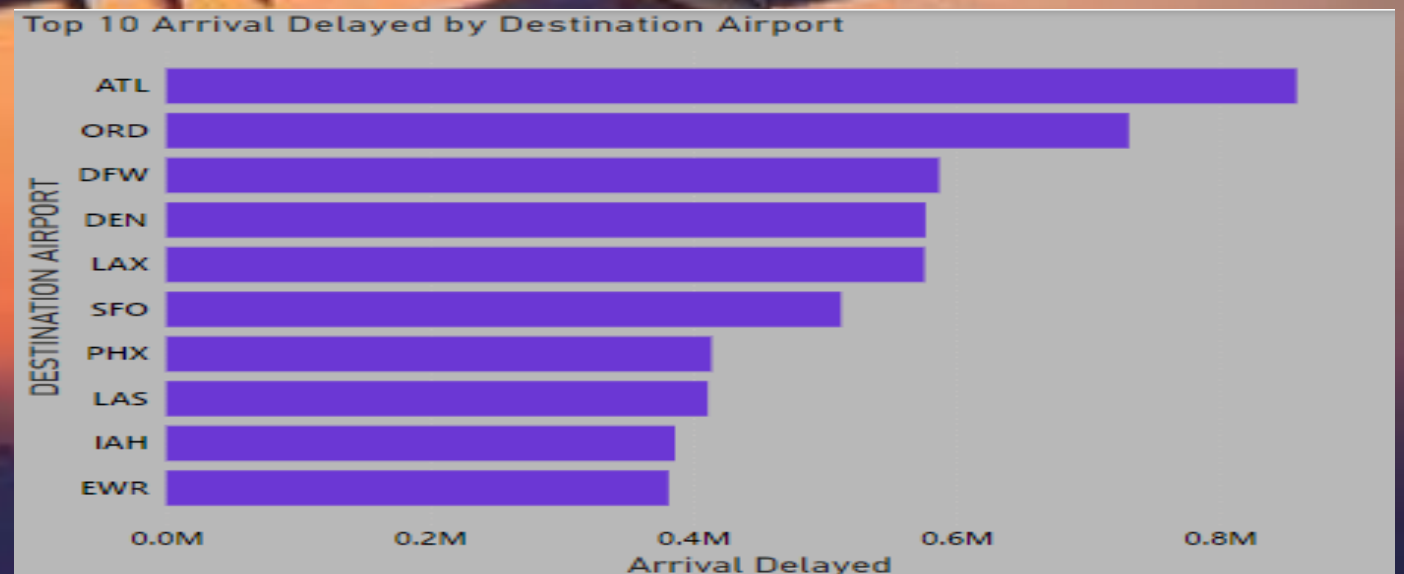
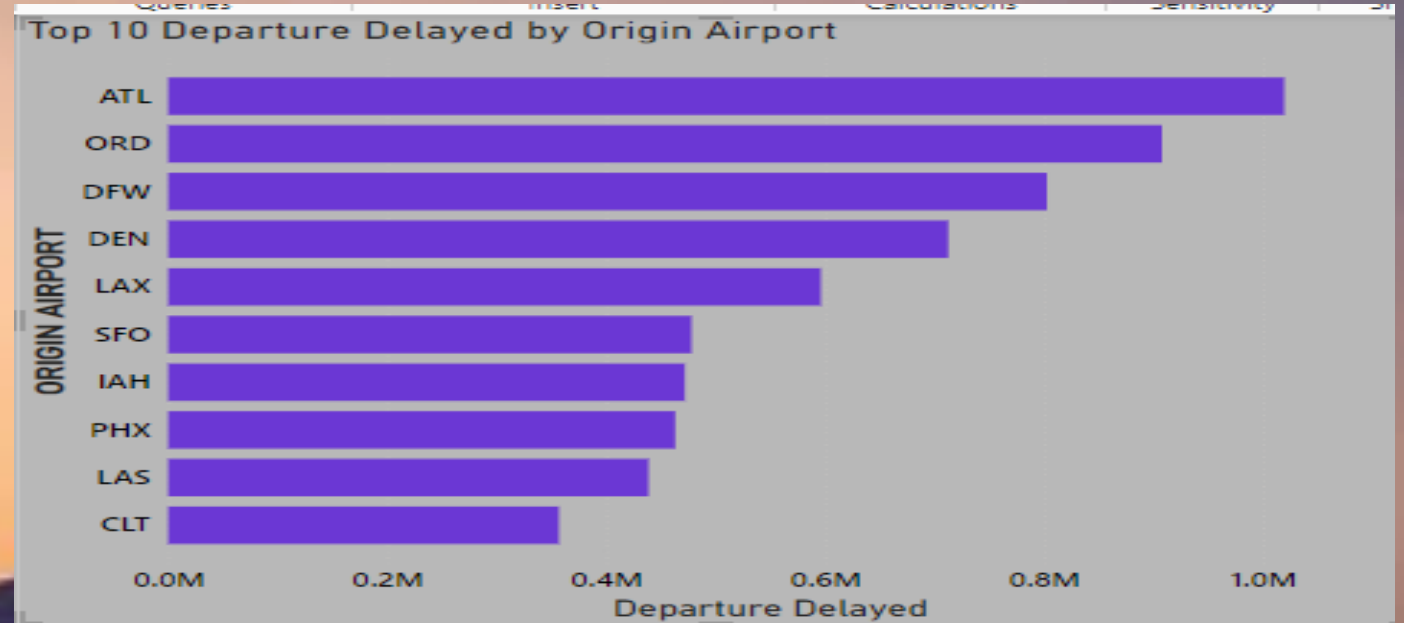
TREND ANALYSIS

- Analyzed the top 10 delayed flights by Airports

Chart Type: Clustered Bar Chart

Conclusion:

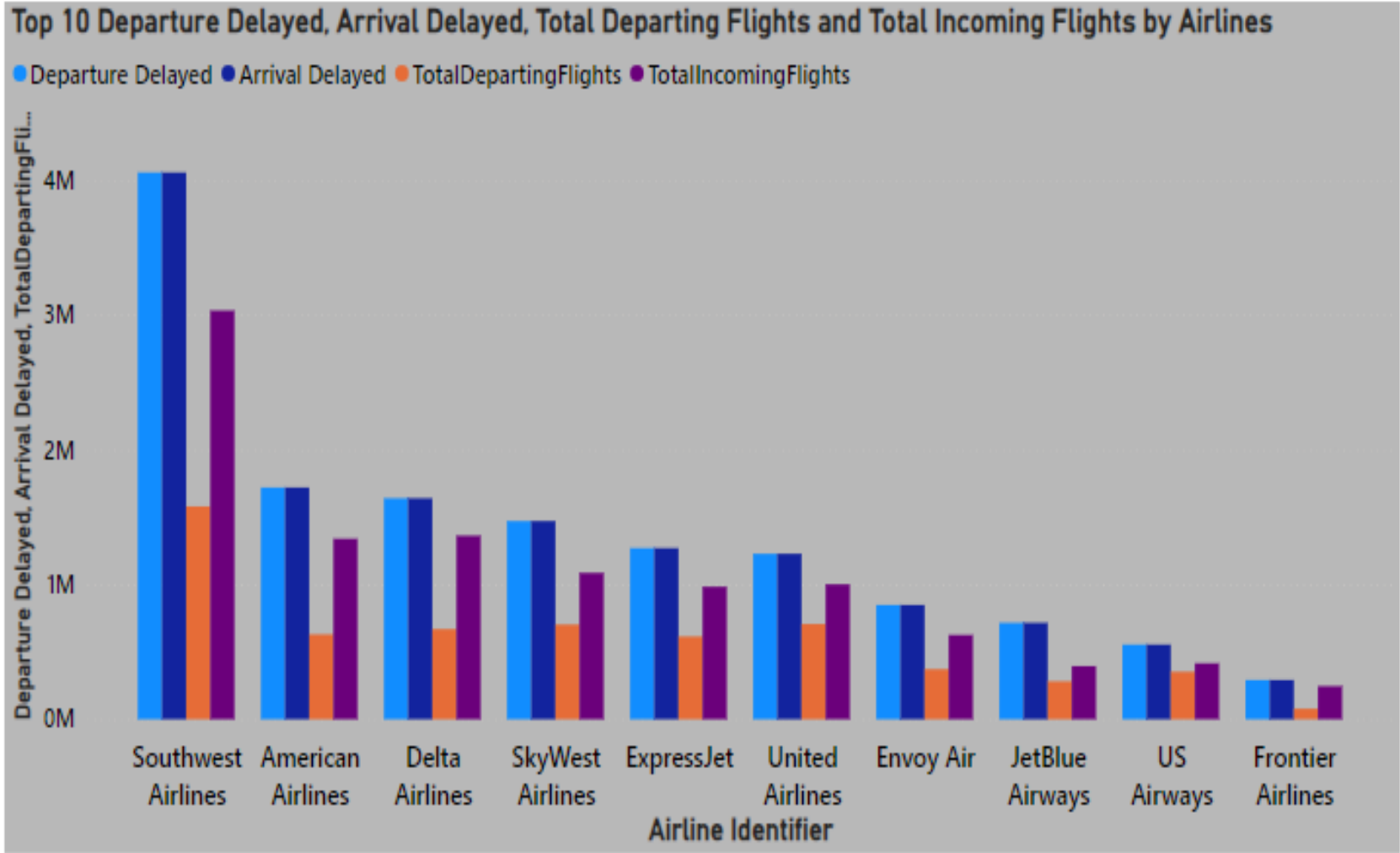
The most delayed flights were from the ATL airport followed by ORD airport



- Analyzed the number of delayed flights , total incoming and departing flights by Airlines

Chart Type: Clustered Column Chart

Conclusion:
The Airline with the most flight delays, most flights is Southwest Airlines



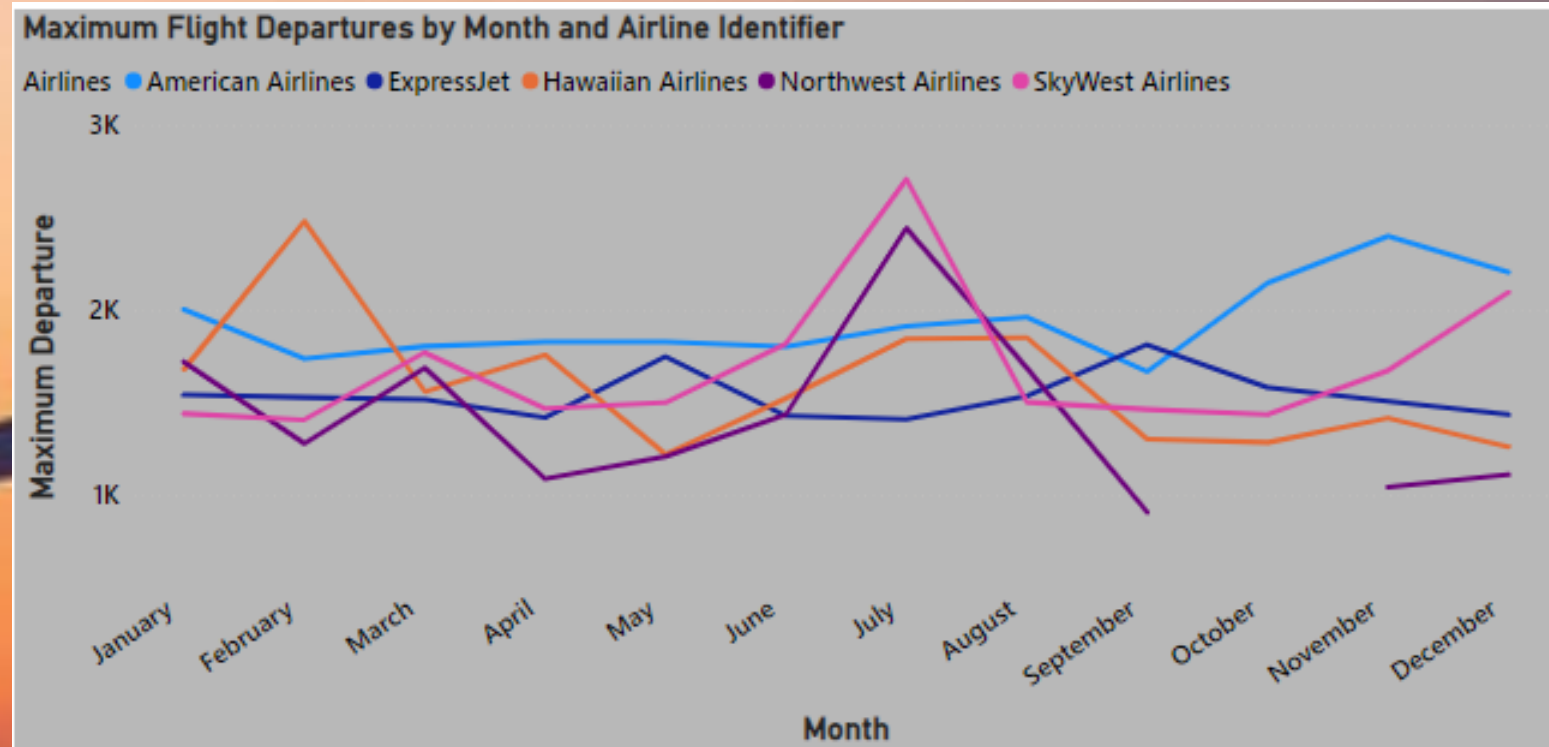
TREND ANALYSIS

- Analyzed the maximum flight departures by month and Airlines

Chart Type: Line Chart

Conclusion:

The maximum flight departures occurred in the month of July and SkyWest Airlines had the most flights at 2 710 followed by Northwest Airlines with 2 445

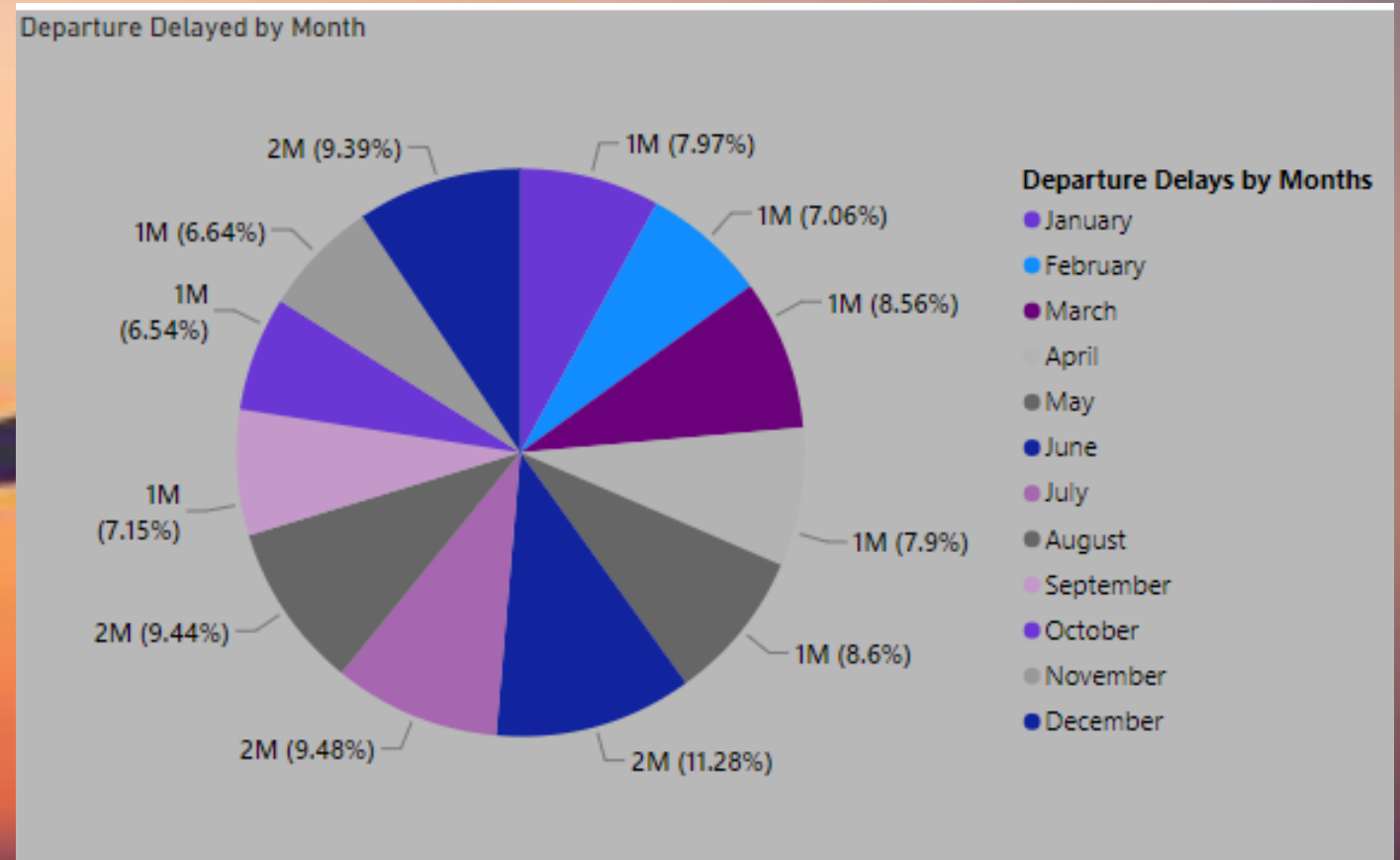


- Analyzed Departure delays by months

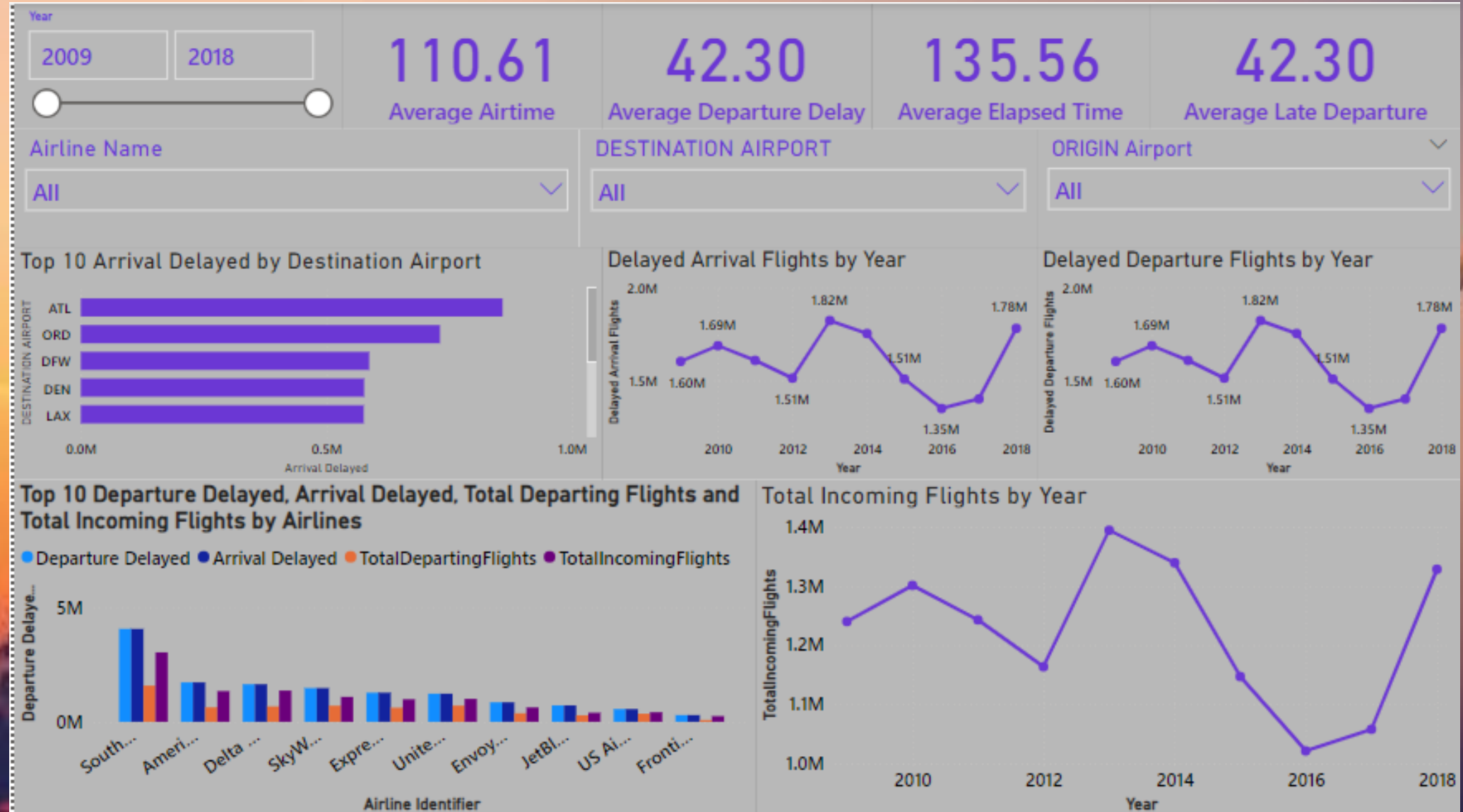
Chart Type: Pie Chart

Conclusion:

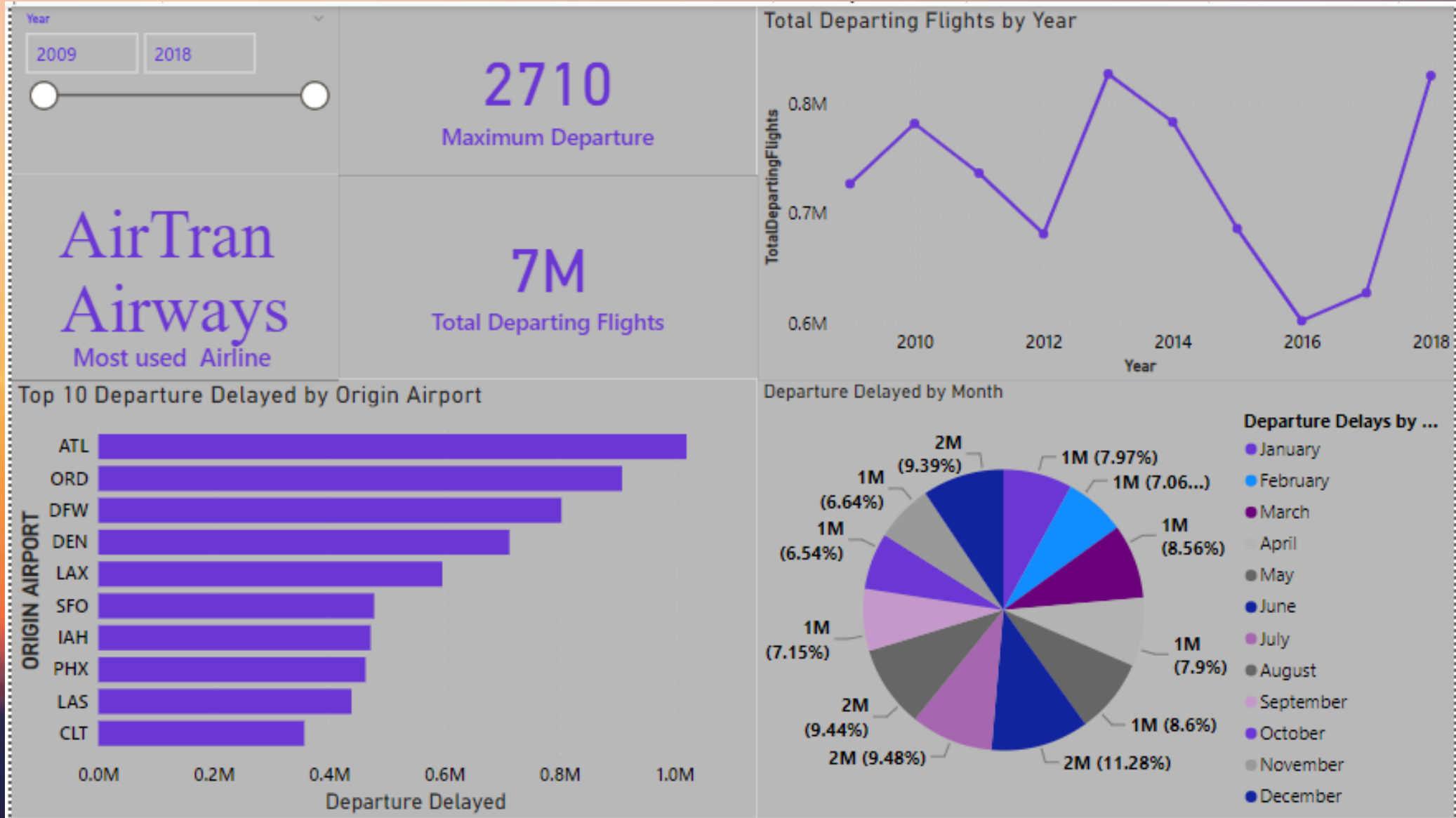
The most flight delays occurred in the month of June with 11.28% of the total proportion



DASHBOARD



DASHBOARD



Summary and Key Insights

US Airline delays and cancellation dataset 2009-2018 reveals the following key insights:

1. The highest number of incoming flights is 1 394 259 and departing flights is 826 690 both in the year 2013
2. The highest number of delayed arrival and departing flights is 1823 059 both in the year 2013
3. The most delayed flights were from the ATL airport followed by ORD airport
4. The Airline with the most flight delays, most flights is Southwest Airlines
5. The average Airtime is 110.61
6. The average departure delay is 42.30

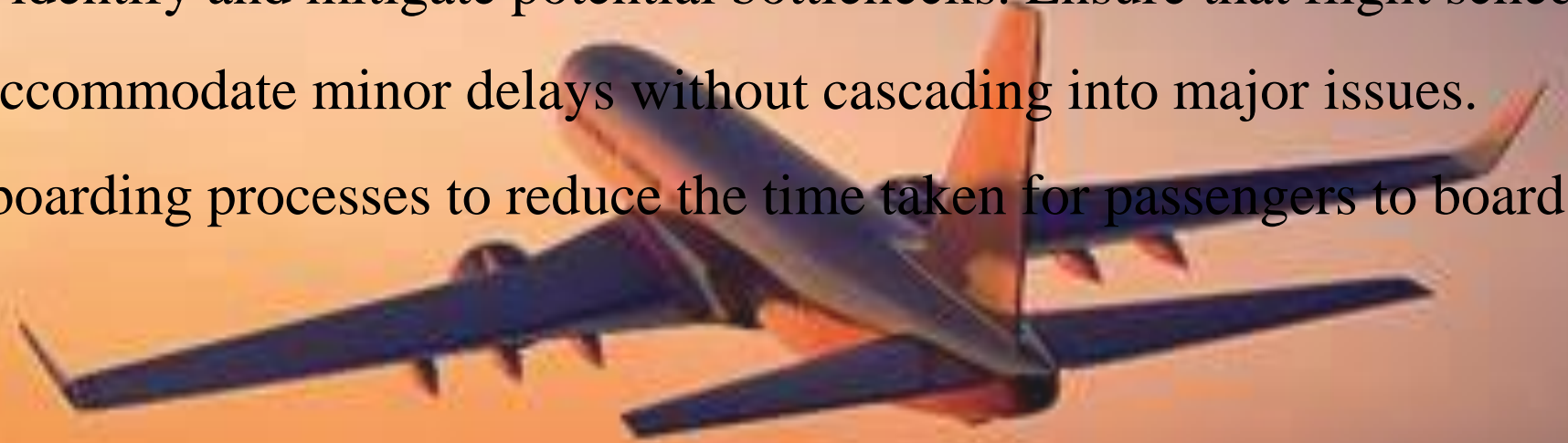


RECOMMENDATIONS

- **Actionable Recommendations:**

Use predictive analytics to identify and mitigate potential bottlenecks. Ensure that flight schedules have sufficient buffers to accommodate minor delays without cascading into major issues.

Implement more efficient boarding processes to reduce the time taken for passengers to board and disembark



REFERENCE AND ACKNOWLEDGEMENTS

•References:

US Airline delays and cancellation dataset 2009-2018

List of airlines of the United States Wikipedia

•Acknowledgements:

Many thanks to mentors, colleagues, and data providers.



