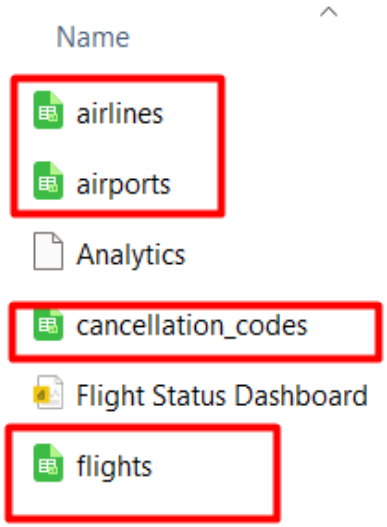
Generating intelligence and insight from Airline flight Data 2015

Introduction

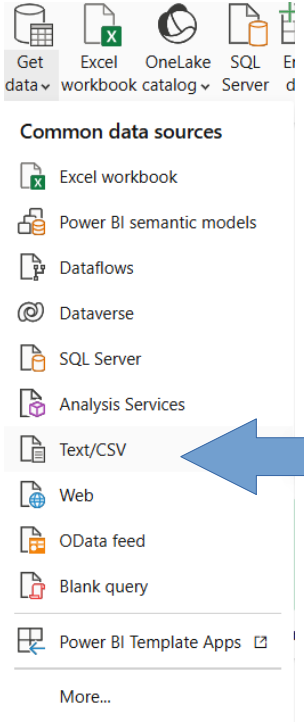


This analysis is based on a set of aviation data which includes the following files: -

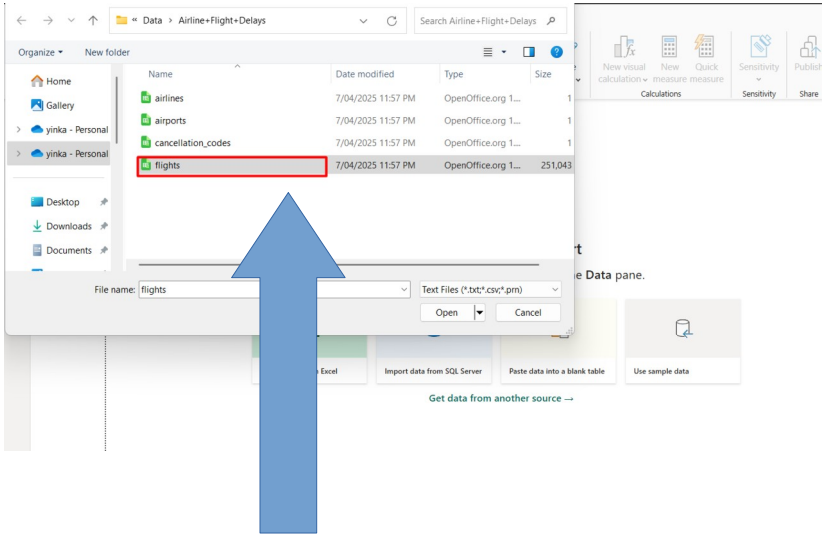
- "**Airlines.csv**" which consists detailed records of airline operations.
- "**Airport.csv**" which comprises of airport codes,
- "**flight.csv**" which comprises of flight information and,
- "**cancellation codes.csv**" which comprises of cancelation description and code

Our aviation dataset captures the operational dynamics between airlines and airports, with a particular focus on flight volumes and cancellation patterns. This dataset enables us to determine a structured assessment of airline performance, airport activity levels, and the underlying reasons for flight cancellations. This foundational overview sets the stage for deeper insights into operational efficiency, disruption causes, and potential areas for systemic improvement in air transport services.

Reading our Dataset

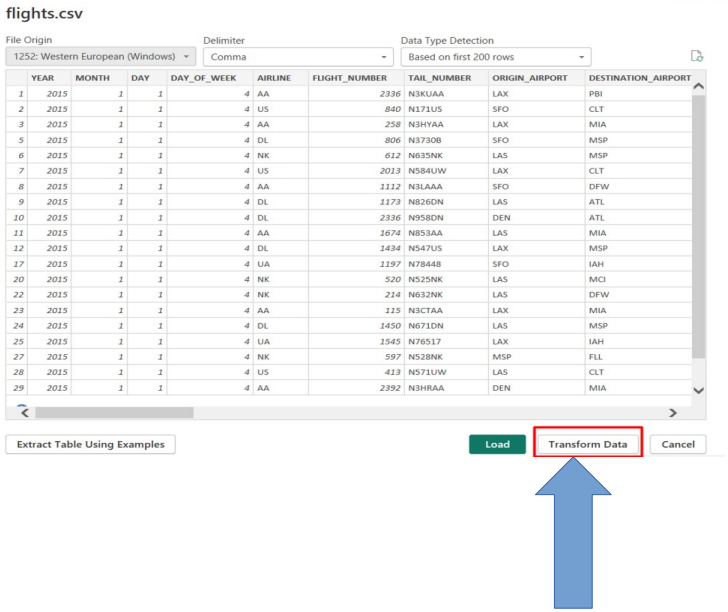


To analyse the data provided using PowerBI we read our CSV files “Text/CSV” option using the “Get Data”

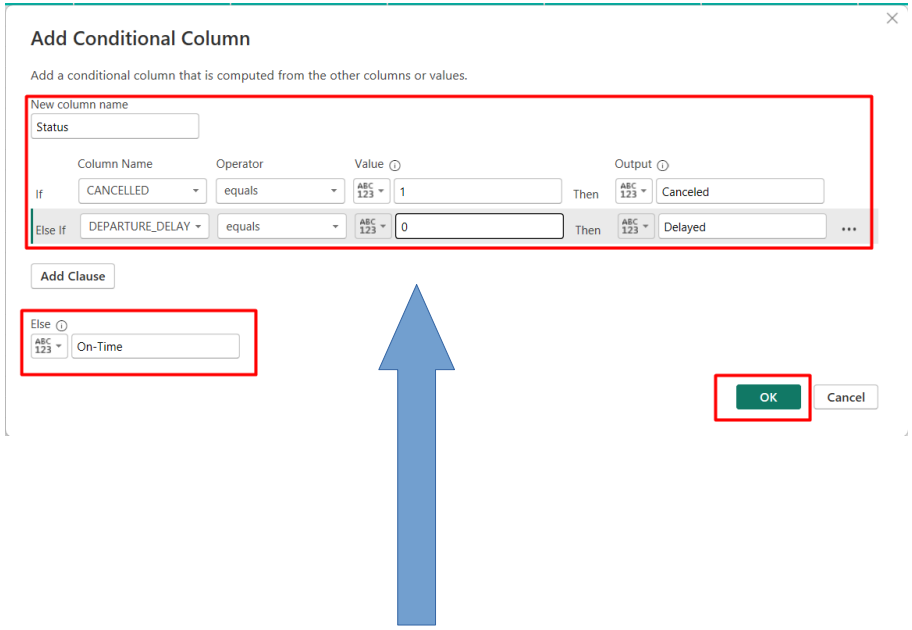


To analyse the data provided using PowerBI we read our CSV files “Text/CSV” option using the “Get Data”

Analyzing and transforming our Dataset

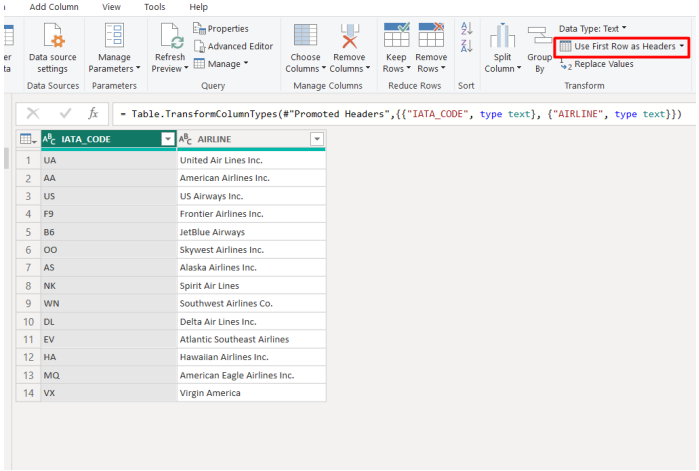


Reviewing our dataset allows us to clean our dataset or carry out operation such as turning our rows into headers. This activity can be carried out using our “Transform data” button

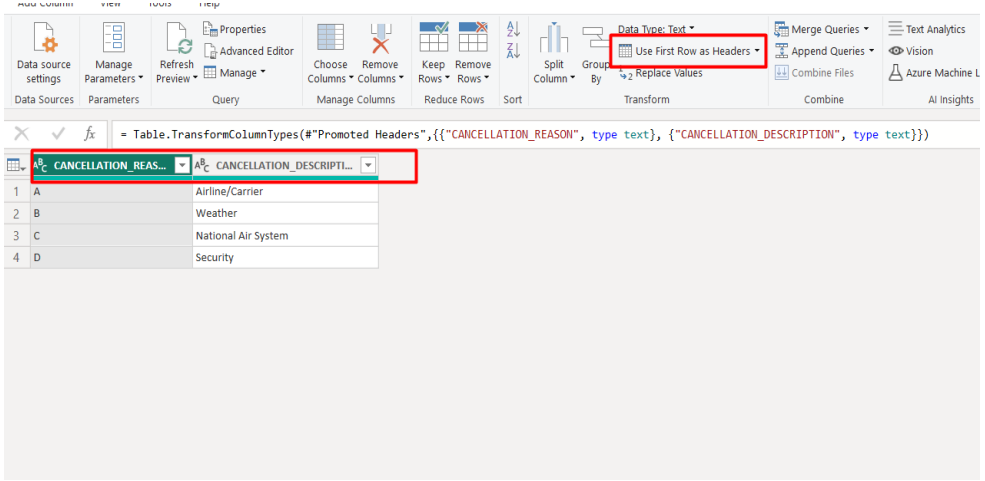


We can modify and add columns using conditional statements

Analyzing and transforming our Dataset

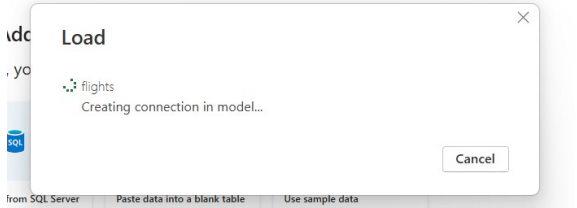


	IATA_CODE	AIRLINE
1	UA	United Air Lines Inc.
2	AA	American Airlines Inc.
3	US	US Airways Inc.
4	F9	Frontier Airlines Inc.
5	B6	JetBlue Airways
6	OO	Skywest Airlines Inc.
7	AS	Alaska Airlines Inc.
8	NK	Spirit Air Lines
9	WN	Southwest Airlines Co.
10	DL	Delta Air Lines Inc.
11	EV	Atlantic Southeast Airlines
12	HA	Hawaiian Airlines Inc.
13	MQ	American Eagle Airlines Inc.
14	VX	Virgin America

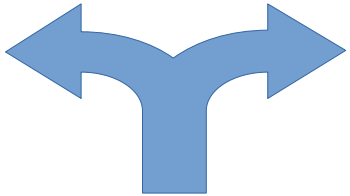


	CANCELLATION_REAS...	CANCELLATION_DESCRIPTI...
1	A	Airline/Carrier
2	B	Weather
3	C	National Air System
4	D	Security

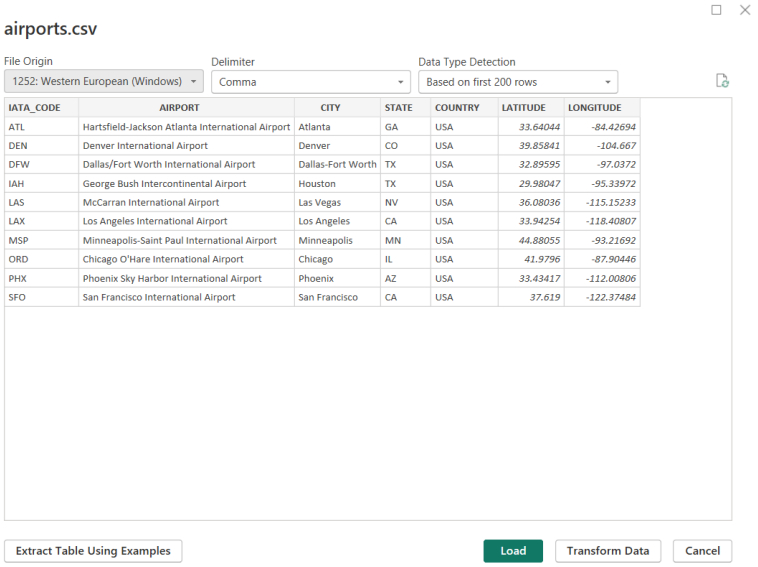
Reviweing and loading our Dataset



Column1	Column2
IATA_CODE	AIRLINE
UA	United Air Lines Inc.
AA	American Airlines Inc.
US	US Airways Inc.
F9	Frontier Airlines Inc.
B6	JetBlue Airways
OO	Skywest Airlines Inc.
AS	Alaska Airlines Inc.
NK	Spirit Air Lines
WN	Southwest Airlines Co.
DL	Delta Air Lines Inc.
EV	Atlantic Southeast Airlines
HA	Hawaiian Airlines Inc.
MQ	American Eagle Airlines Inc.
VX	Virgin America

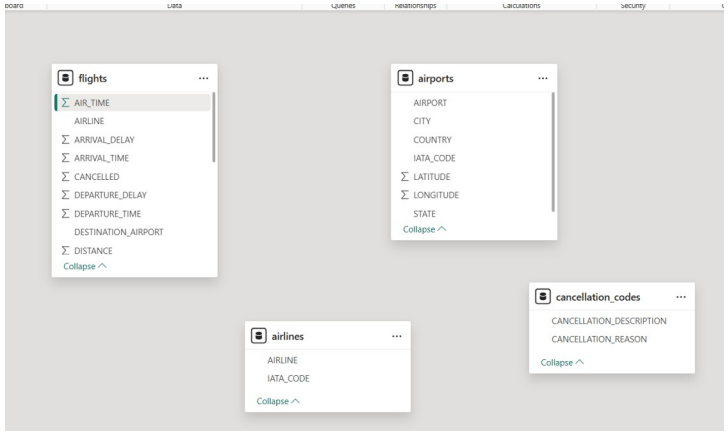


After reviewing and confirming the validity of our dataset, we can load our data for further analysis



IATA_CODE	AIRPORT	CITY	STATE	COUNTRY	LATITUDE	LONGITUDE
ATL	Hartsfield-Jackson Atlanta International Airport	Atlanta	GA	USA	33.64044	-84.42694
DEN	Denver International Airport	Denver	CO	USA	39.85841	-104.667
DFW	Dallas/Fort Worth International Airport	Dallas-Fort Worth	TX	USA	32.89595	-97.0372
IAH	George Bush Intercontinental Airport	Houston	TX	USA	29.98047	-95.33972
LAS	McCarran International Airport	Las Vegas	NV	USA	36.08036	-115.15233
LAX	Los Angeles International Airport	Los Angeles	CA	USA	33.94254	-118.40807
MSP	Minneapolis-Saint Paul International Airport	Minneapolis	MN	USA	44.88055	-93.21692
ORD	Chicago O'Hare International Airport	Chicago	IL	USA	41.9796	-87.90446
PHX	Phoenix Sky Harbor International Airport	Phoenix	AZ	USA	33.43417	-112.00806
SFO	San Francisco International Airport	San Francisco	CA	USA	37.619	-122.37484

Entity Relationship Mapping



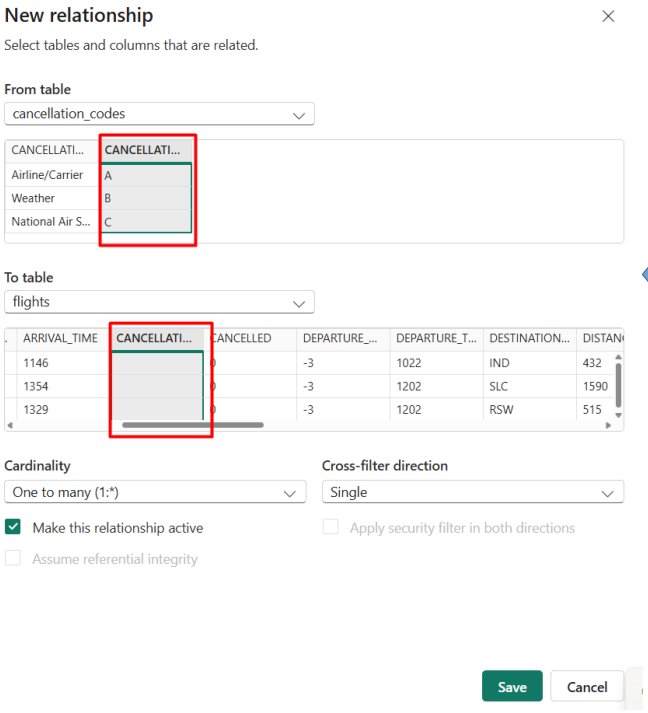
Our dataset is structured around four primary entities:

- Airlines,
- Airports,
- Flights, and
- Cancellation Codes.

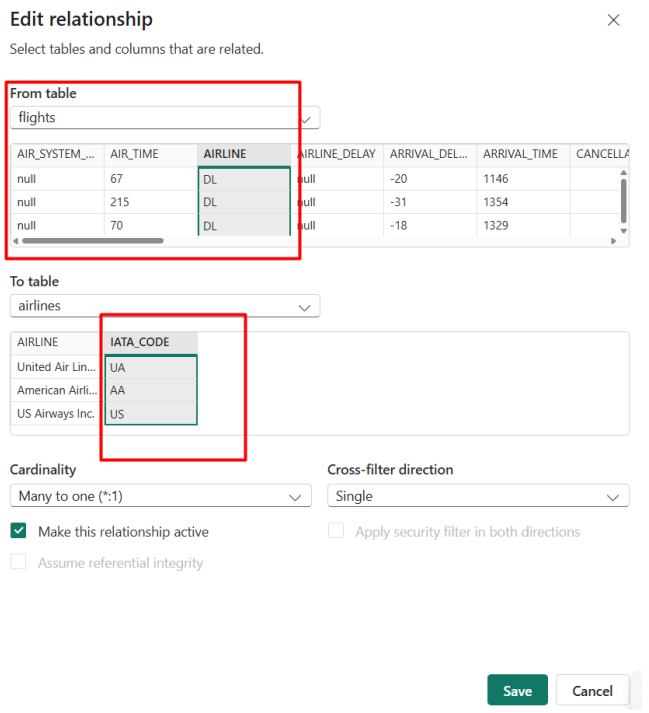
These entities are interconnected as follows:

- Flights serve as the central fact table, containing foreign keys referencing both Airlines and Airports.
- Each flight record links to a specific Airline (via airline code) and a Departure and Arrival Airport (via airport codes).
- Cancellation Codes are associated with canceled flights, providing categorical reasons for cancellation.
- Airlines and Airports function as dimension tables, offering descriptive metadata for reporting and analysis.

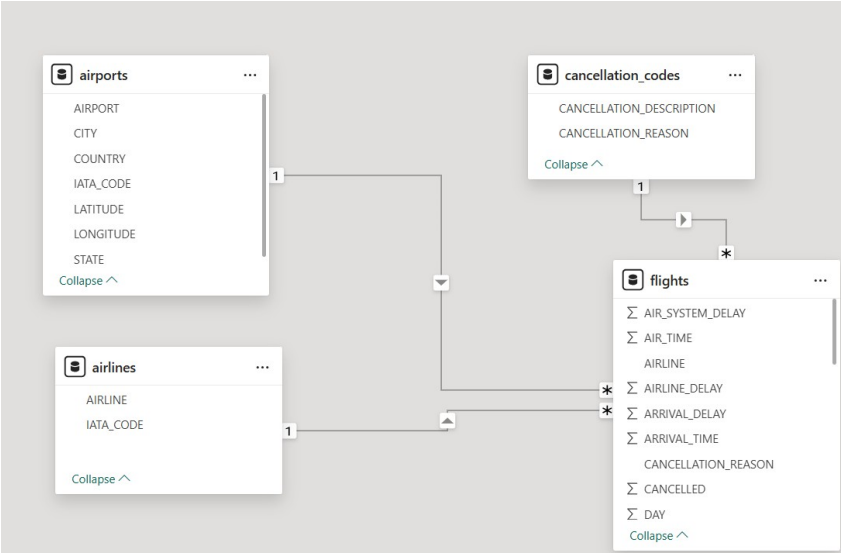
Entity Relationship Mapping...



Mapping and Editing Relationship across different tables allows us to map unique identifiers in one table to its corresponding foreign key in another table



Entity Relationship mapping...



This relational structure enables efficient querying and supports multidimensional analysis across time, carriers, locations, and disruption causes.

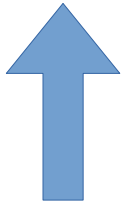
Using Dax Statements...

```
1 % Canceled Flights = FORMAT(DIVIDE([Canceled Flight], [Total Flight], "-") * 100, "0.00") & "%"
```

```
1 % Delayed Flights = FORMAT(DIVIDE([Delayed Flight], [Total Flight], "-") * 100, "0.00") & "%"

1 % On-Time Flights = FORMAT( DIVIDE([On-Time Flight], [Total Flight], "-") * 100, "0.00") & "%"

1 Canceled Flight = CALCULATE([Total Flight], flights[Status] = "Canceled")
```



DAX (Data Analysis Expressions) is used for calculations. It helps us to leverage the use of expressions to simulate or utilize our entity relationship in defining measures and calculated columns.

Flight Performance Report

Total Flight

Total Flight

2M

Canceled Flight

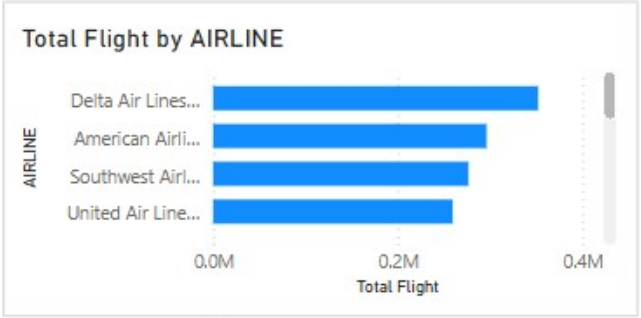
Canceled Flight

29K

Delayed Flight

Delayed Flight

790K

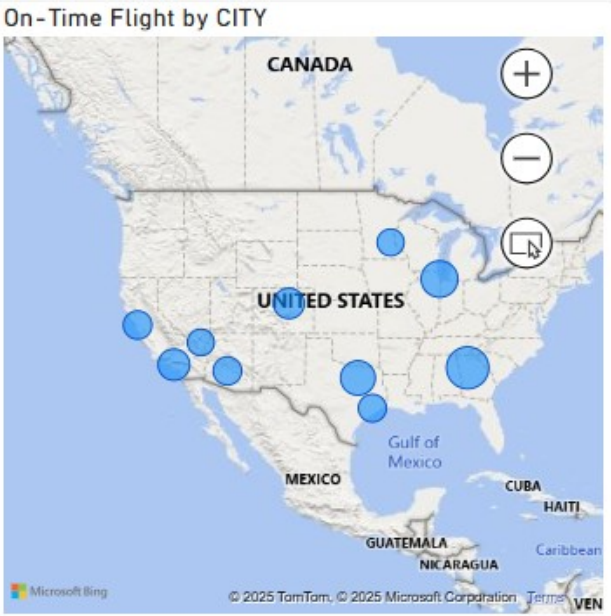


% Canceled Flights

1.47%

% Delayed Flights

40.54%



Conclusion...

From our dashboard, we can infer the following insights:

-

Total Flight
2M

We had about two million flights scheduled across all airports surveyed for the year 2015.

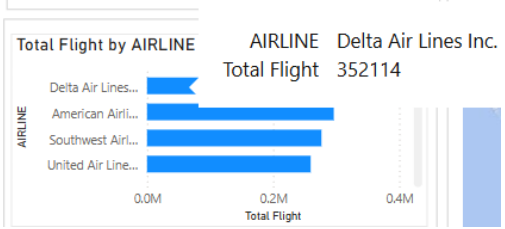
Canceled Flight
29K

We had about twenty nine thousand of those flights making 1.47% of the total were cancelled across all airports in 2015.

Delayed Flight
790K

We had over seven hundred and ninety thousand flights scheduled which were delayed making 40.54% of the total.

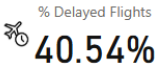
Chicago airport recorded the highest number of cancelled flight by city with 8,548



Delta Airline has the highest record of flights at over three hundred and fifty thousand for the year – 2015

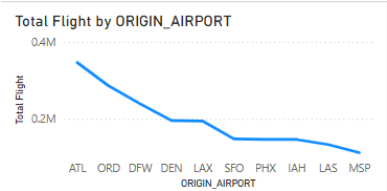
Conclusion...

From our dashboard, we can infer the following: -

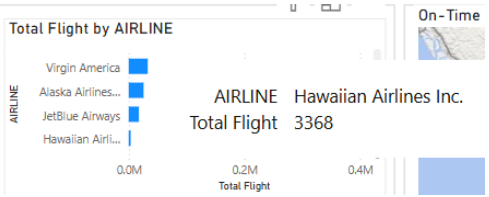


Out of the total flights scheduled, we had one million and a hundred flights were on time making 50.99%

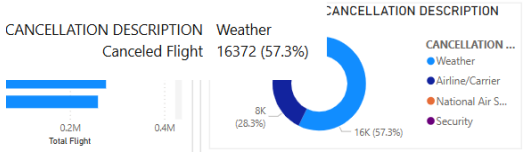
Out of the total flights scheduled, we had one million and a hundred flights were on time making 50.99% of the total number of flights across all airports



We can see a downward trend in total flights across Origin Airport



Hawaiian Airlines had the least number of flights with a total number of flights at 3,368



Cancellation due to weather reasons constitutes about 57.3% of total Cancellation