

## **Data Appendix**

### **Original Data**

Data Dictionary for all initial data in the DATA folder. This includes the following files:

- April.csv
- August.csv
- Dec.csv
- Feb.csv
- Jan.csv
- July.csv
- June.csv
- March.csv
- May.csv
- Nov.csv
- Oct.csv
- Sept.csv

<b>Column</b>	<b>Description</b>	<b>Potential Responses</b>
FL_DATE	Date that the flight is departing	1/1/2023
OP_UNIQUE	Unique carrier code of the airline departure	9E
ORIGIN	Origin Airport Code	DTW, ATL, ORD
DEP_DELAY_NEW	Difference in minutes between scheduled and actual departure time. Early departures set to 0.	0 or any other integer
CANCELED	Canceled Flight Indicator (1=Yes & 0=No)	0 or 1
CARRIER_DELAY	Time in minutes for delay caused by the carrier	0 or any other integer
WEATHER_DELAY	Time in minutes for delay caused by the weather	0 or any other integer

NAS_DELAY	Time in minutes for delay caused by the NAS	0 or any other integer
SECURITY_DELAY	Time in minutes for delay caused by the National Aviation System (NAS)	0 or any other integer
LATE_AIRCRAFT_DELAY	Time in minutes for delay caused by a late aircraft	0 or any other integer

### **Analyzed Data**

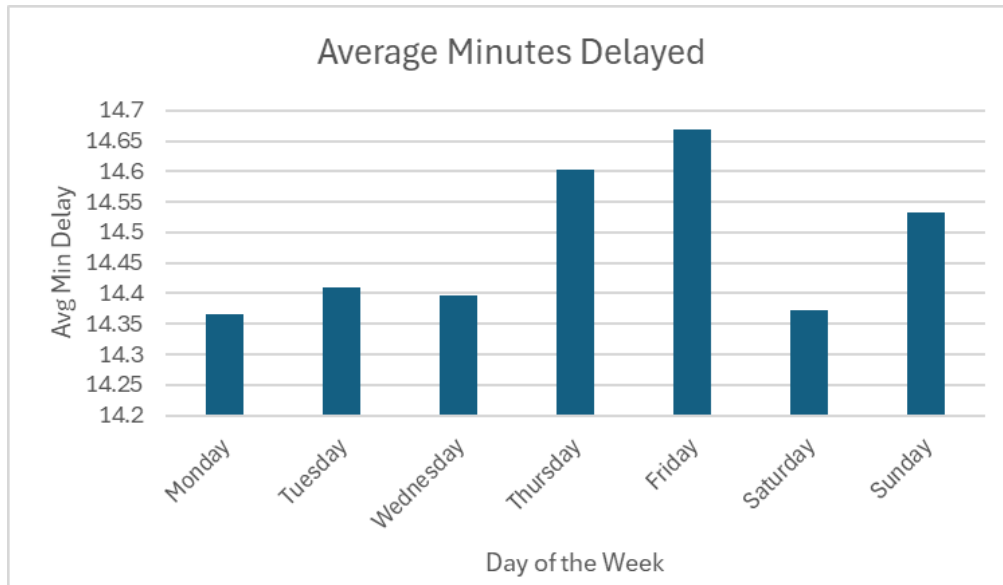
The analyzed dataset is a dataset consisting of all initial data combined into one dataset. It also includes the day\_of\_week column which indicates the day of the week a flight departed. To create this dataset, we combined all the data into one file through google sheets/excel, added a new column to determine the day of week, added column ORIGIN\_CODE to numerically categorize the airport, and filtered out airports not being studied.

Data Dictionary for **Airports2023.xlsx**.

Column	Description	Potential Responses
FL_DATE	Date that the flight is departing	1/1/2023
day_of_week	Day of the week the flight departed	Sunday
OP_UNIQUE	Unique carrier code of the airline departure	9E
ORIGIN	Origin Airport Code	DTW, ATL, ORD
ORIGIN_CODE	Origin Airport Code as a numeric value	1, 2, 3
DEP_DELAY_NEW	Difference in minutes between scheduled and actual departure time. Early departures set to 0.	0 or any other integer
CANCELED	Canceled Flight Indicator (1=Yes & 0=No)	0 or 1
CARRIER_DELAY	Time in minutes for delay caused by the carrier	0 or any other integer
WEATHER_DELAY	Time in minutes for delay caused by the weather	0 or any other integer
NAS_DELAY	Time in minutes for delay caused by the NAS	0 or any other integer
SECURITY_DELAY	Time in minutes for delay caused by the National Aviation System (NAS)	0 or any other integer
LATE_AIRCRAFT_DELAY	Time in minutes for delay caused by a late aircraft	0 or any other integer

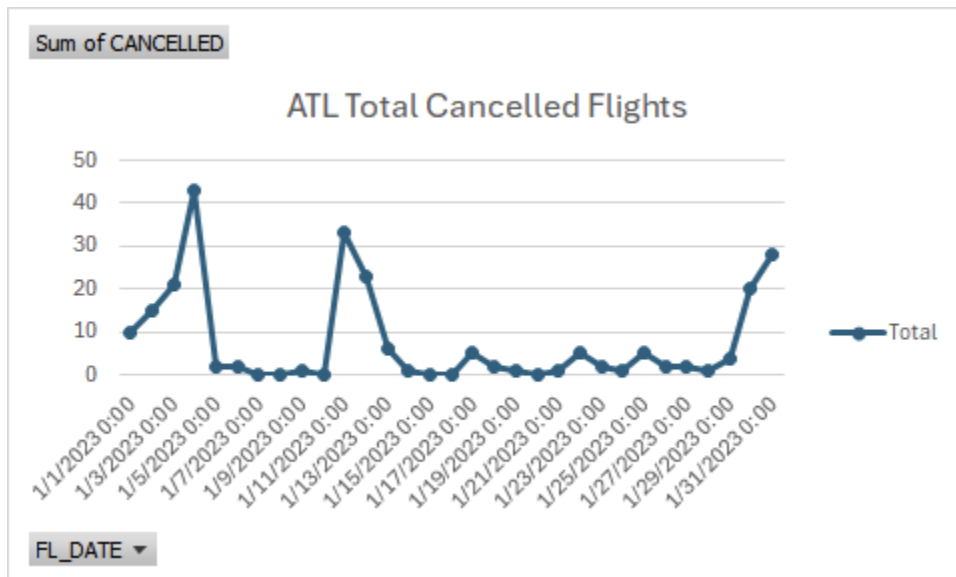
**Figure 1.**

A histogram displaying the average minutes delayed per flight across DFW, ATL, & ORD, per day.



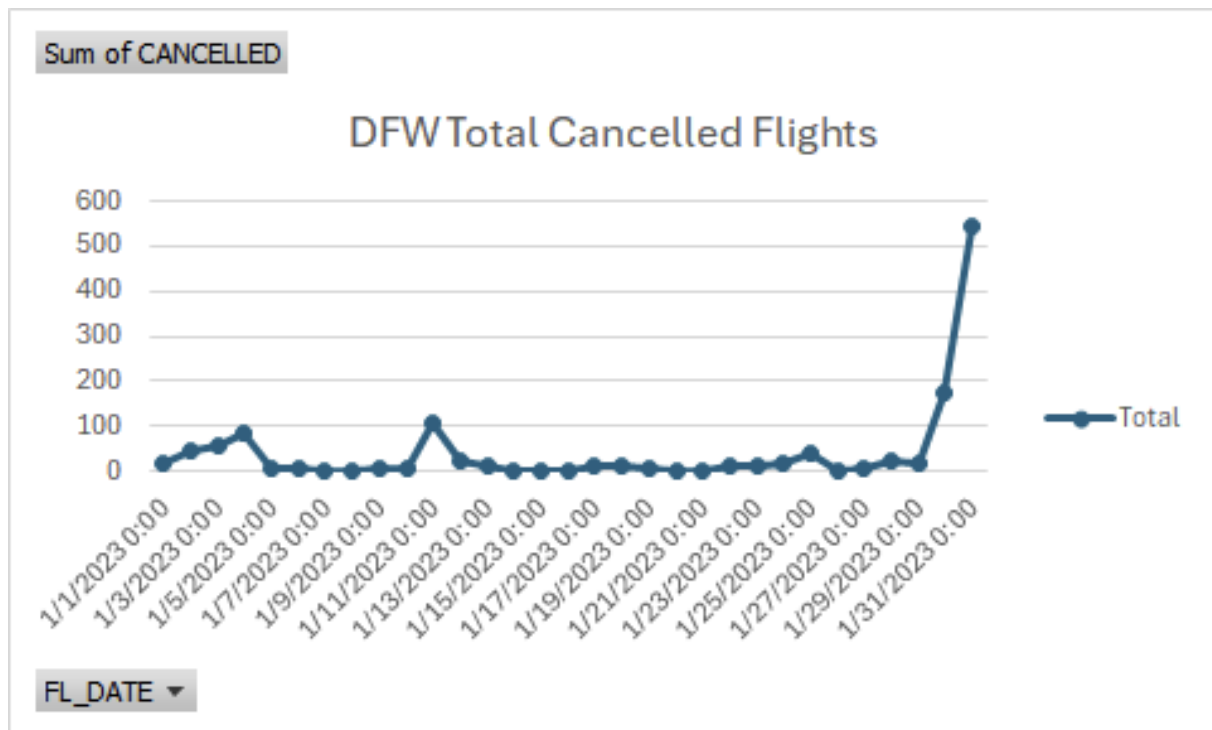
**Figure 2.**

A histogram displaying the number of total canceled flights per day in January for ATL.



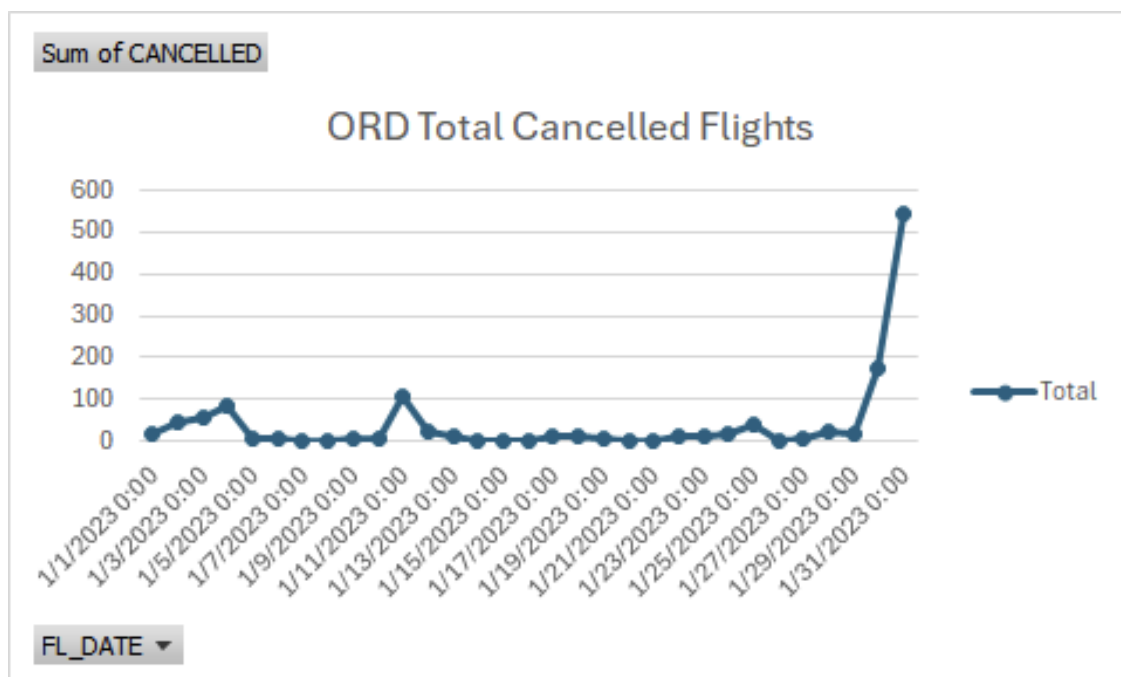
**Figure 3.**

A histogram displaying the number of total canceled flights per day in January for DFW.



**Figure 4.**

A histogram displaying the number of total canceled flights per day in January for ORD.



## References

- [1] U.S. Department of Transportation. (n.d.). OST\_R: BTS: Transtats. Available: [https://www.transtats.bts.gov/DL\\_SelectFields.aspx?gnoyr\\_VQ=FGJ&QO\\_fu146\\_anzr=b0-gvzr](https://www.transtats.bts.gov/DL_SelectFields.aspx?gnoyr_VQ=FGJ&QO_fu146_anzr=b0-gvzr) [Accessed Feb. 29, 2024].