Airline Performance Analysis

Drexel University LeBow College of Business

Summer 2023

Team-5

Dhruv Kandhari

Kavya Kishore Mudumbai

Manoj Suchendra Reddy

Mufaddal Zulfiquar

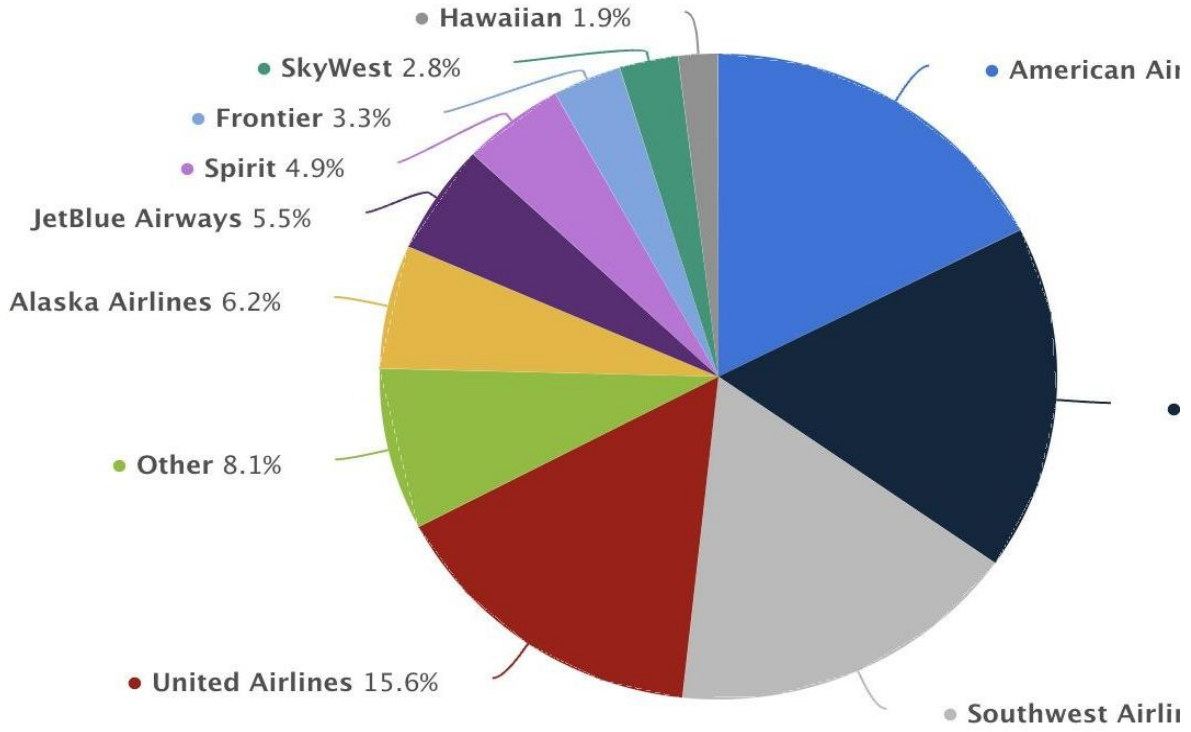
DREXEL UNIVERSITY
LeBow
College of Business

Filtered Data

~11 Million rows total

Filtered out Airlines:

- American Airlines
- Southwest Airlines
- United Airlines
- Alaska Airlines
- JetBlue Airlines



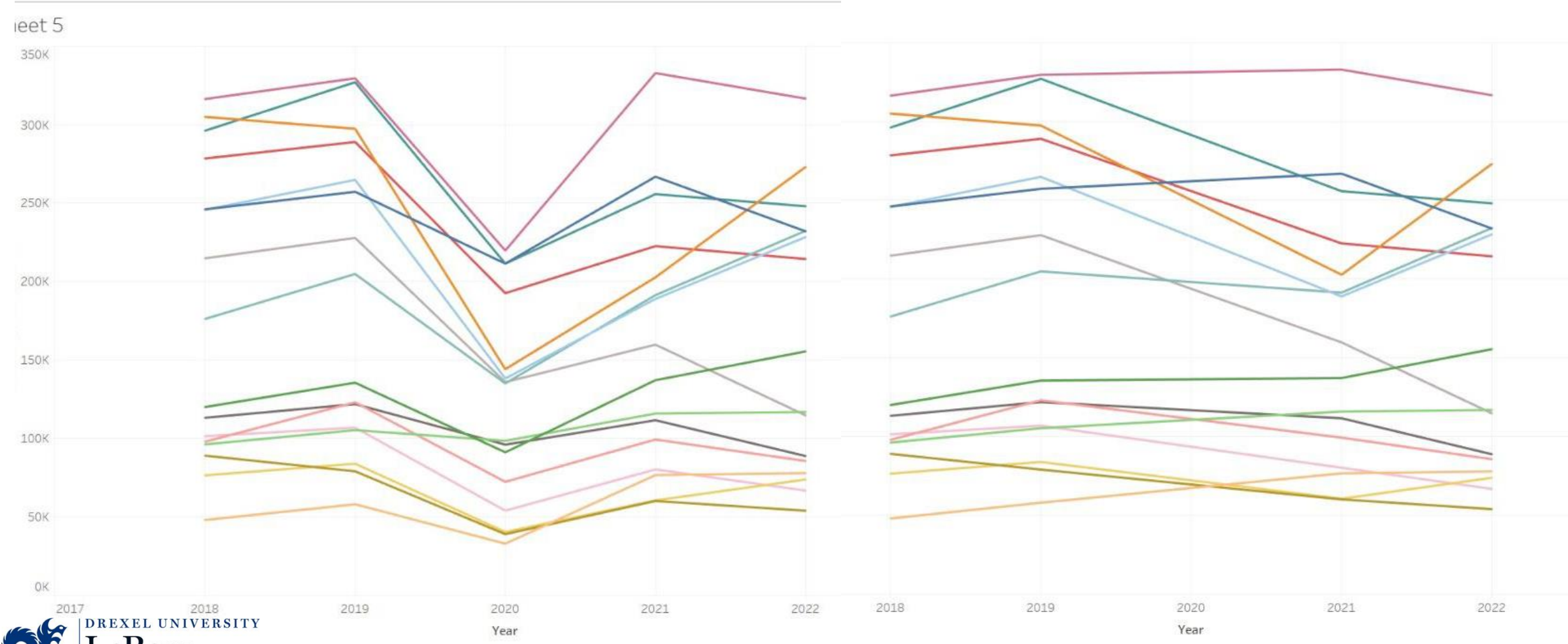
DREXEL UNIVERSITY
LeBow College of Business

Filtered Airlines

- **9E – Endeavor Air Inc.**
- **AS – Alaska Airlines Inc.**
- **B6 – JetBlue Airways**
- **C5 – CommuteAir LLC**
- **F9 – Frontier Airlines Inc.**
- **G4 – Allegiant Air**
- **G7 – GoJet**
- **YX – Republic Airlines**
- **MQ – Envoy Air**
- **NK – Spirit Air Lines**
- **OH – PSA Airlines Inc.**
- **PT – Piedmont Airlines**
- **QX – Horizon Air**
- **YV – Mesa Airlines Inc.**
- **ZW – Air Wisconsin Airlines Corp**
- **HA – Hawaiian Airlines**

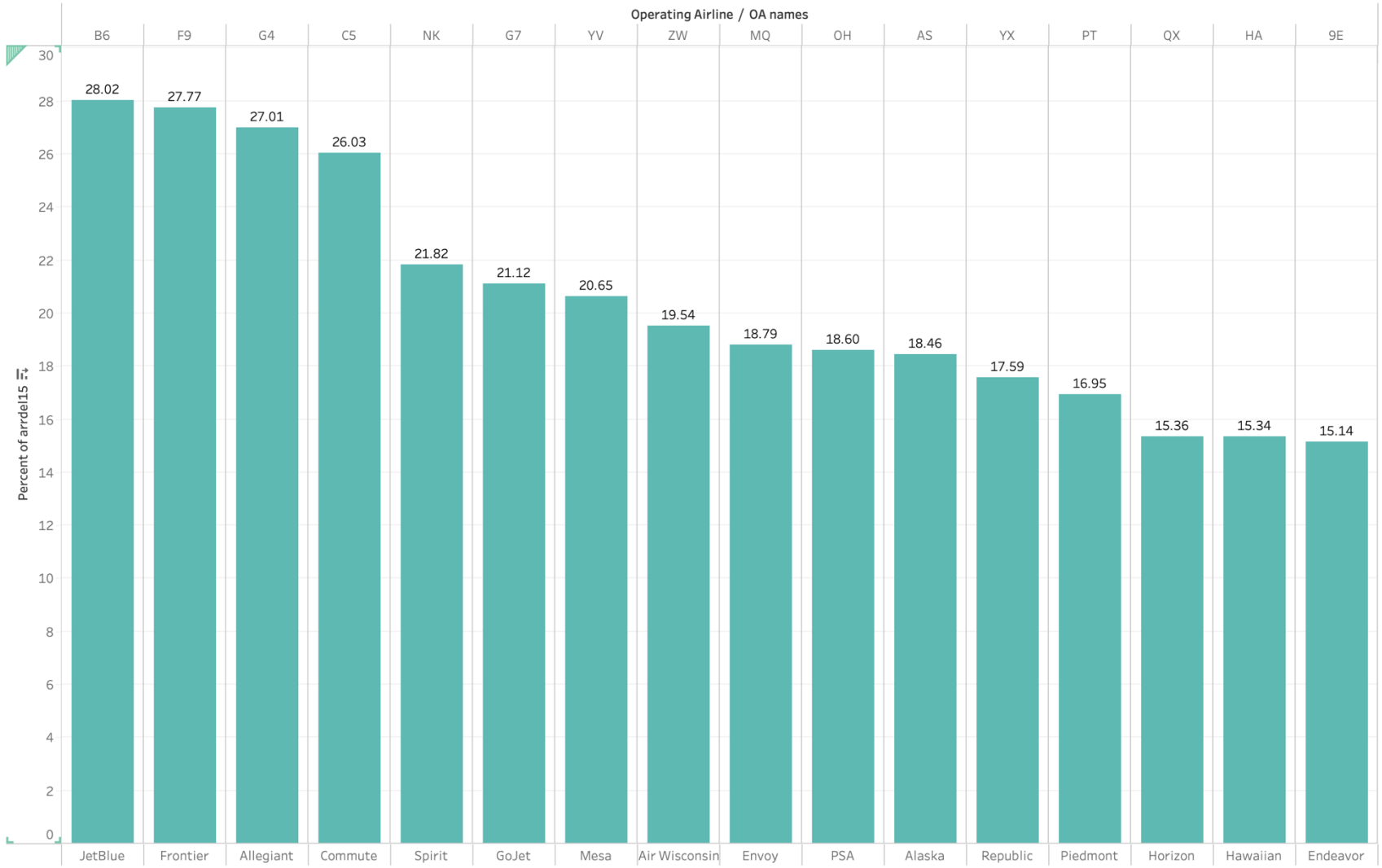
Insights & Key Findings

Excluding COVID Data



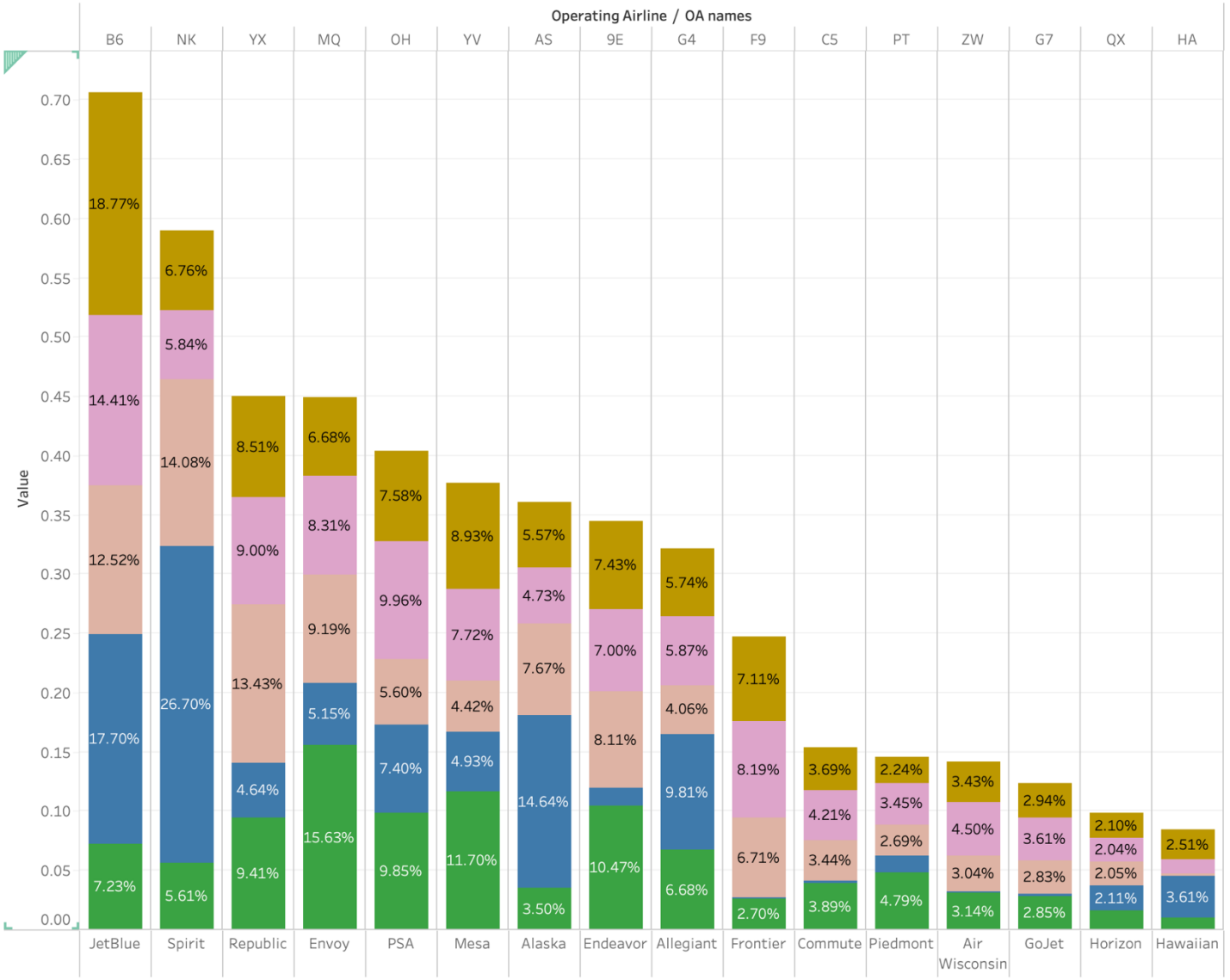
DREXEL UNIVERSITY
LeBow
College of Business

Delays by Airlines



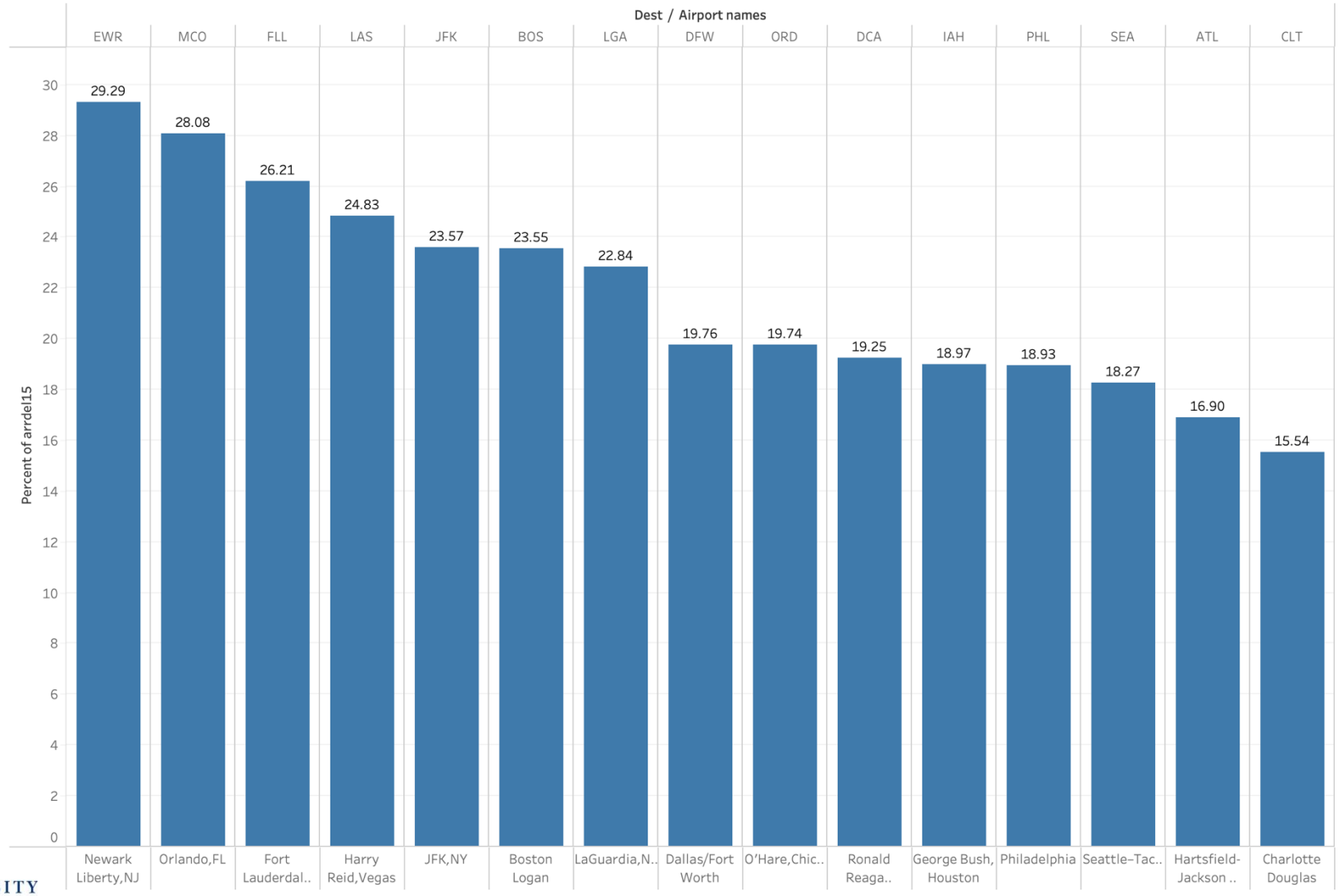
DREXEL UNIVERSITY
LeBow College of Business

Delay Categories by Airline



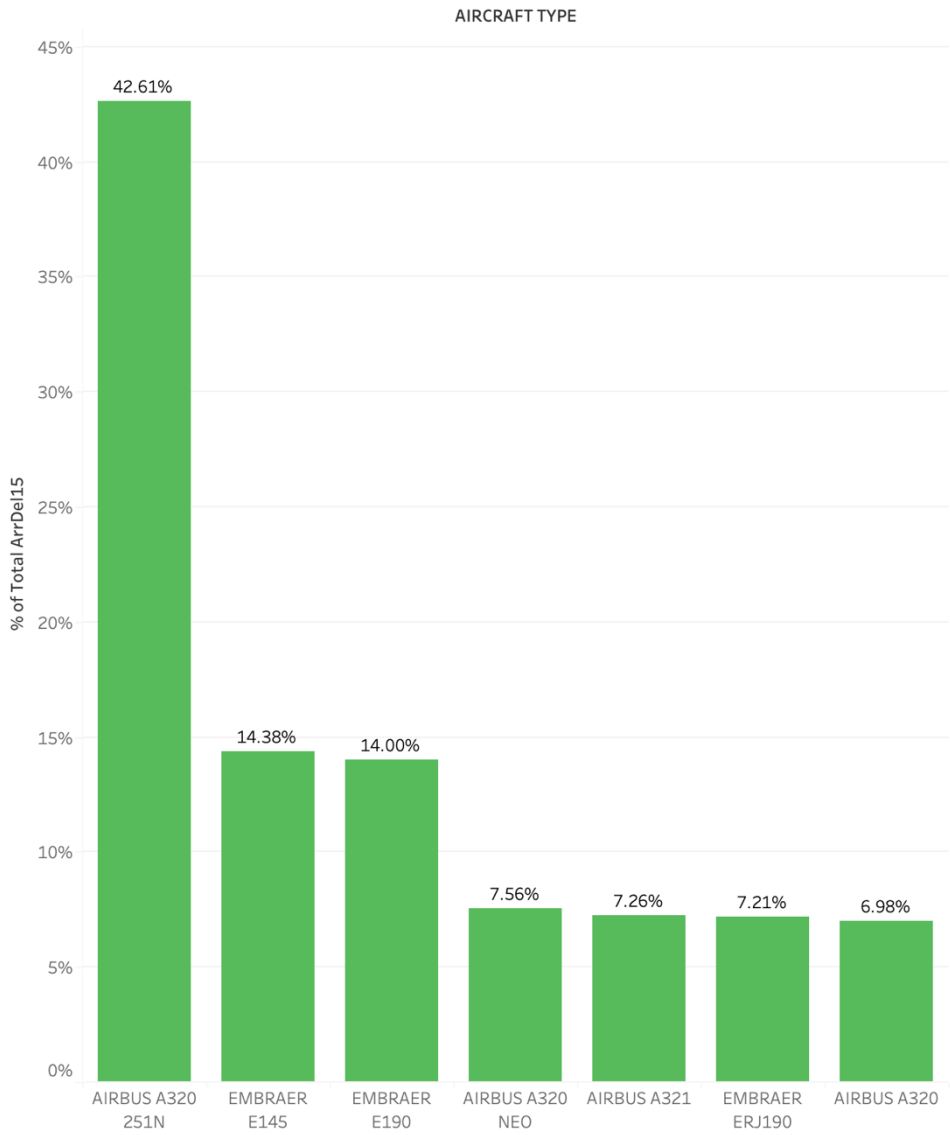
DREXEL UNIVERSITY
LeBow College of Business

Delays By Airport



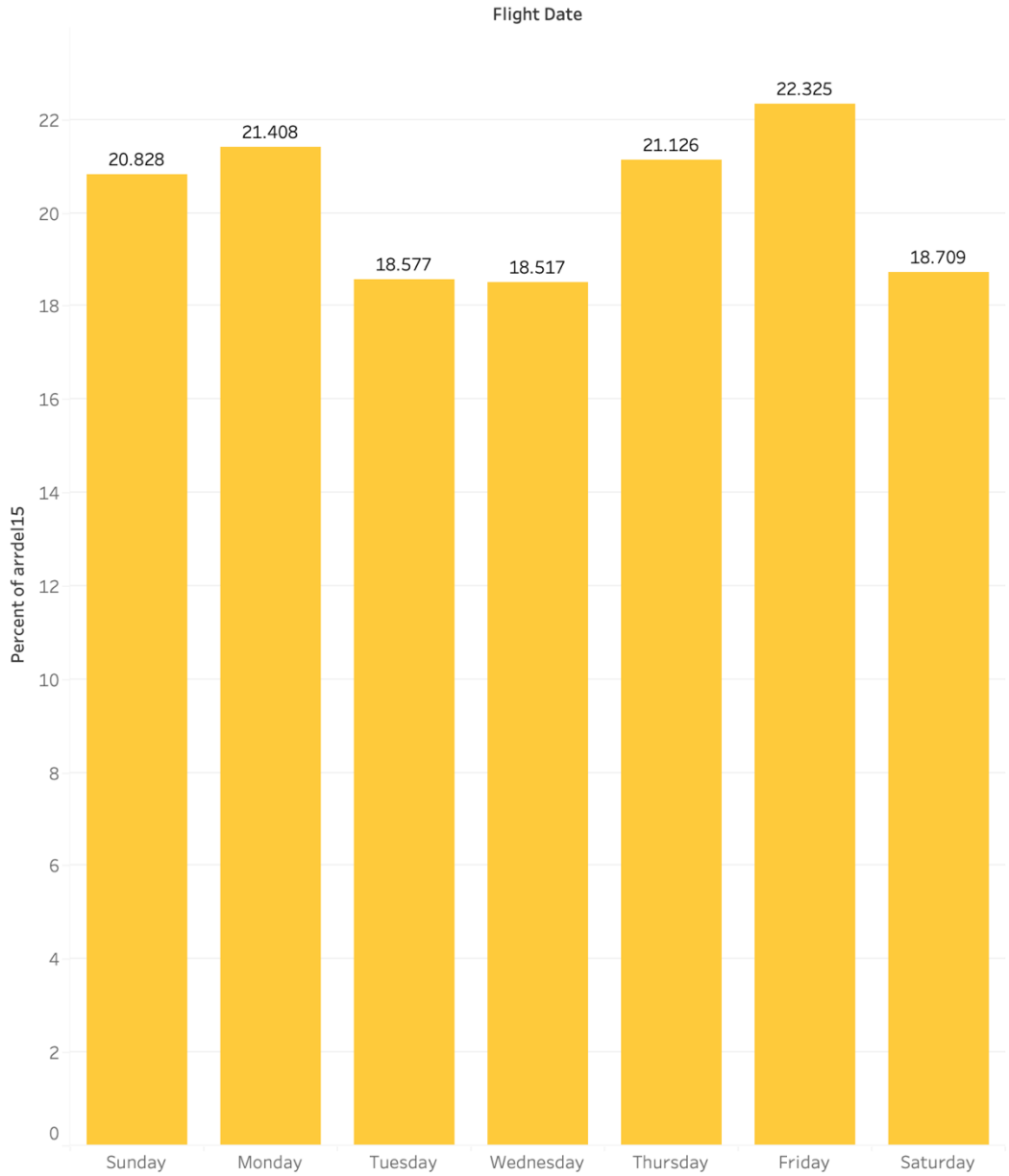
DREXEL UNIVERSITY
LeBow College of Business

Delays by Aircraft Type



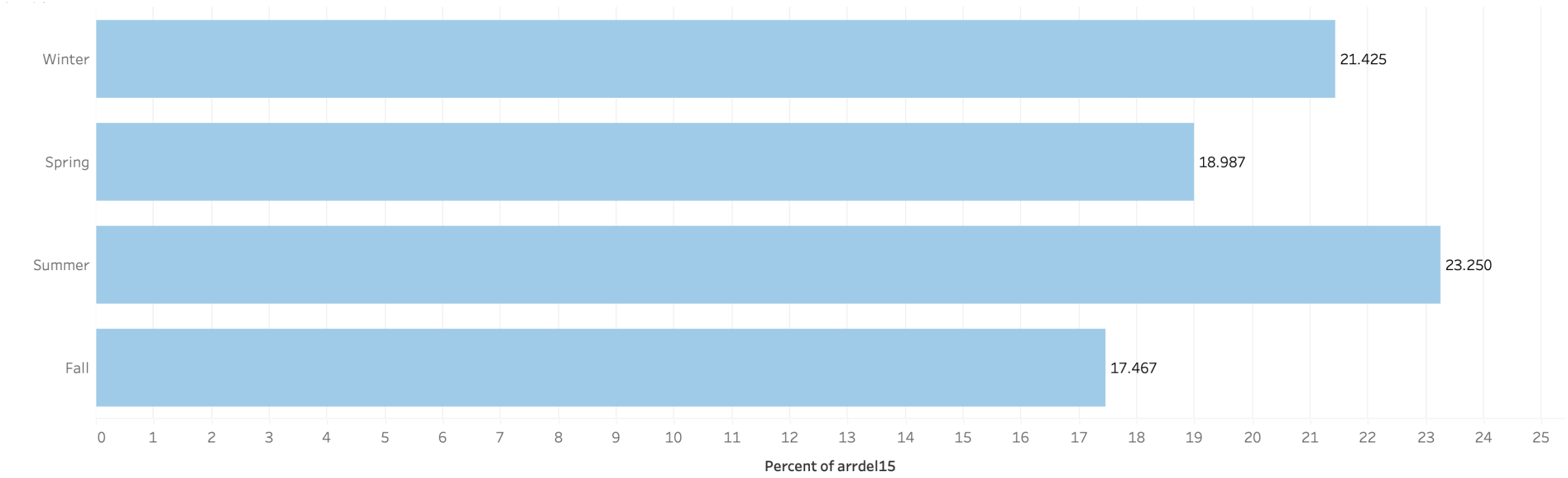
DREXEL UNIVERSITY
LeBow College of Business

Delays By Weekday



DREXEL UNIVERSITY
LeBow College of Business

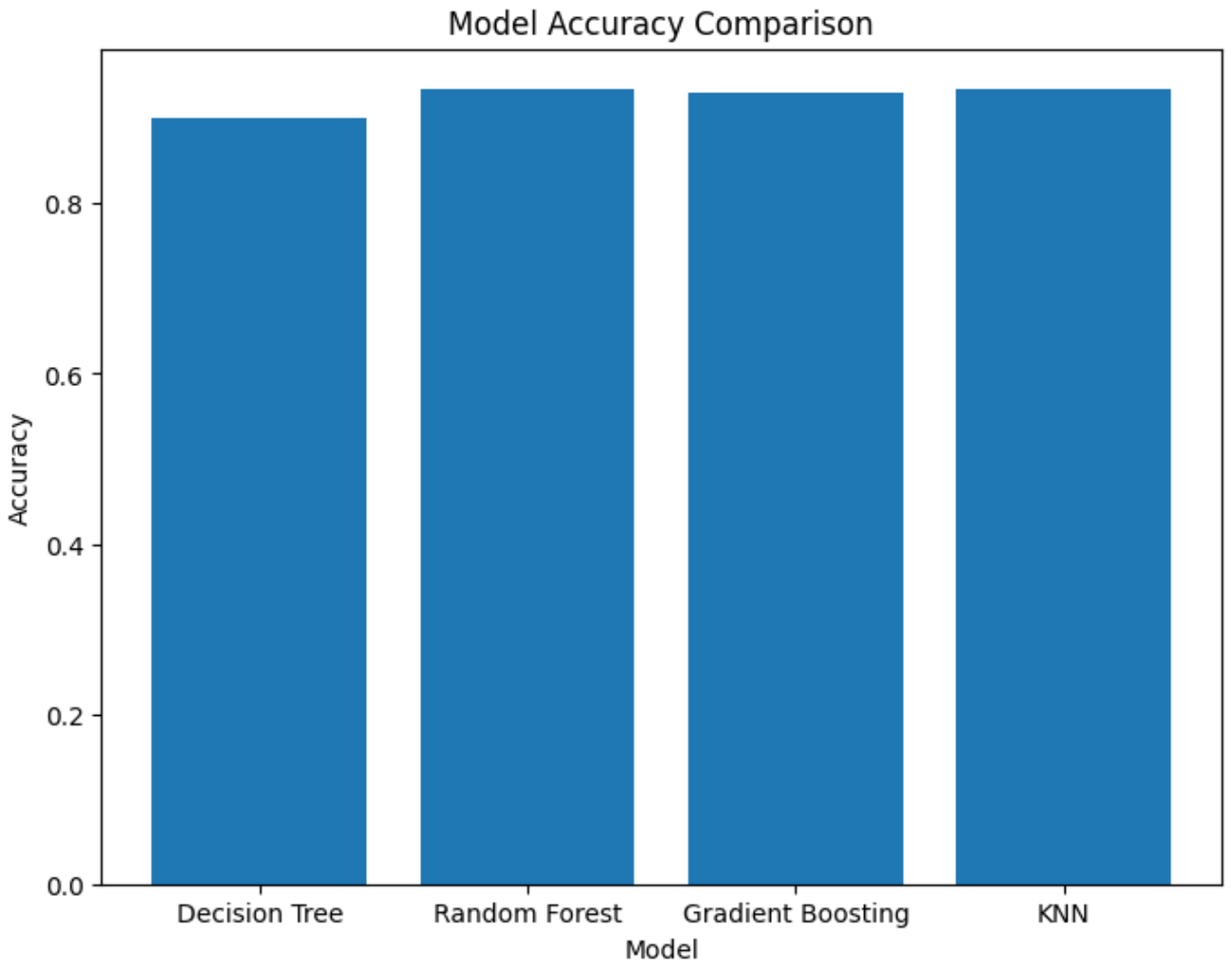
Seasonality



DREXEL UNIVERSITY
LeBow College of Business

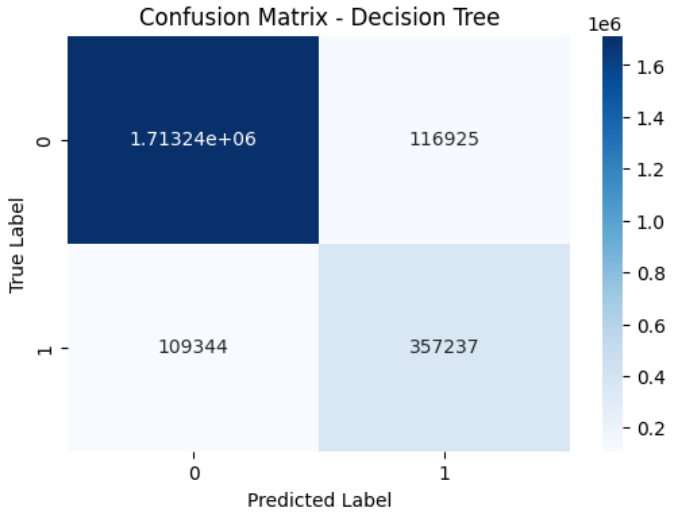
Models

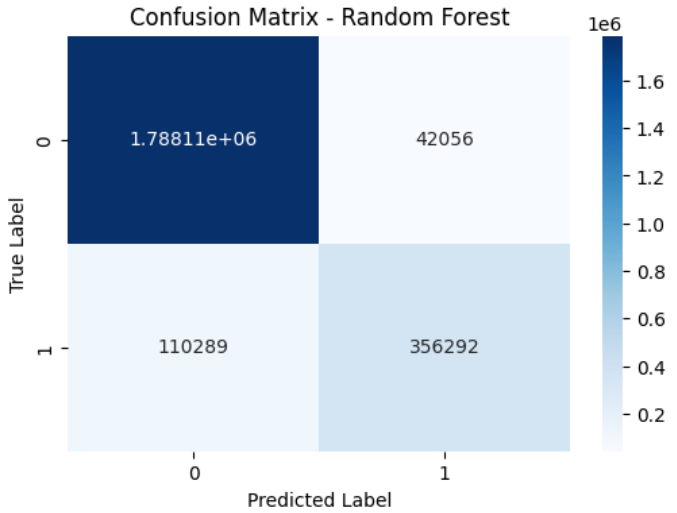
- **Accuracies:**
  - **Decision Tree-** 90.14%(~50s)
  - **Random Forest-** 93.36%(~23min)
  - **Gradient Boosting-** 93.05%(~15min)
  - **KNeighbors-** 93.53%(~2min)

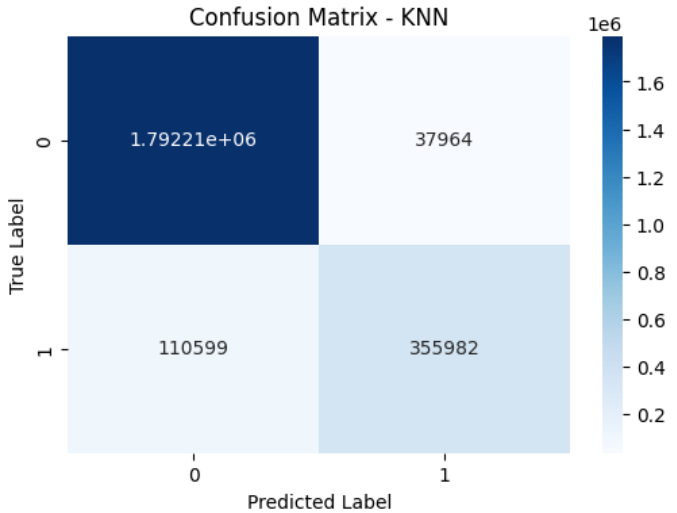


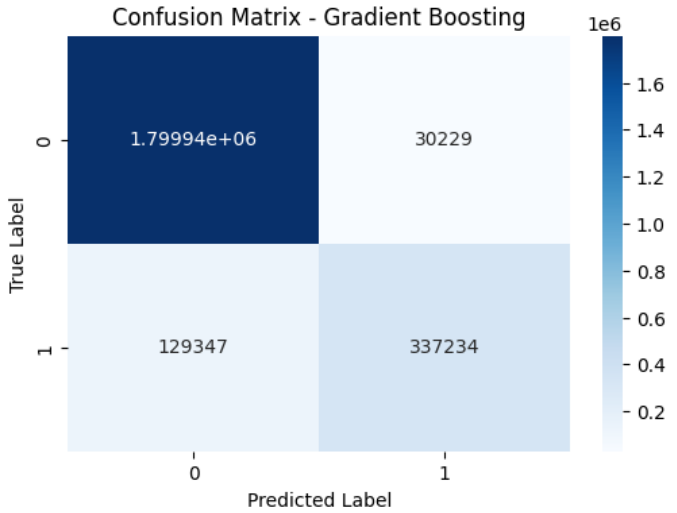
DREXEL UNIVERSITY LeBow College of Business

Confusion Matrices









DREXEL UNIVERSITY
LeBow
College of Business

Thank you

APPENDICES

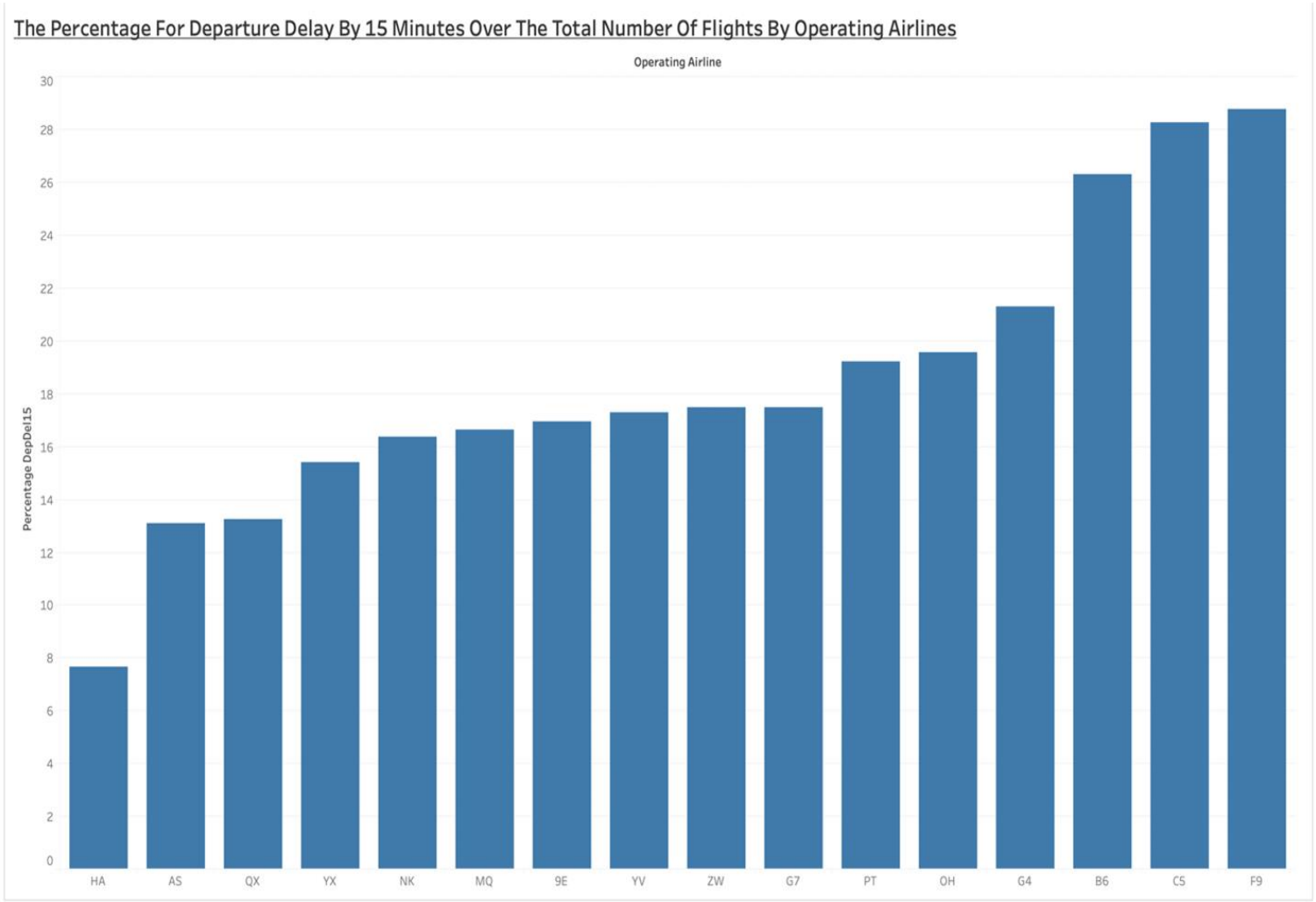
Background/Scope of Work

This project aims to identify the factors contributing to commercial flight delays across various airlines and establish a predictive model to forecast on-time performance of various airlines.

Utilize the machine learning algorithms to develop a predictive model for the on-time performance of an operating airline and train the model using historical data, and other identified features under different circumstances to ensure a higher customer experience.

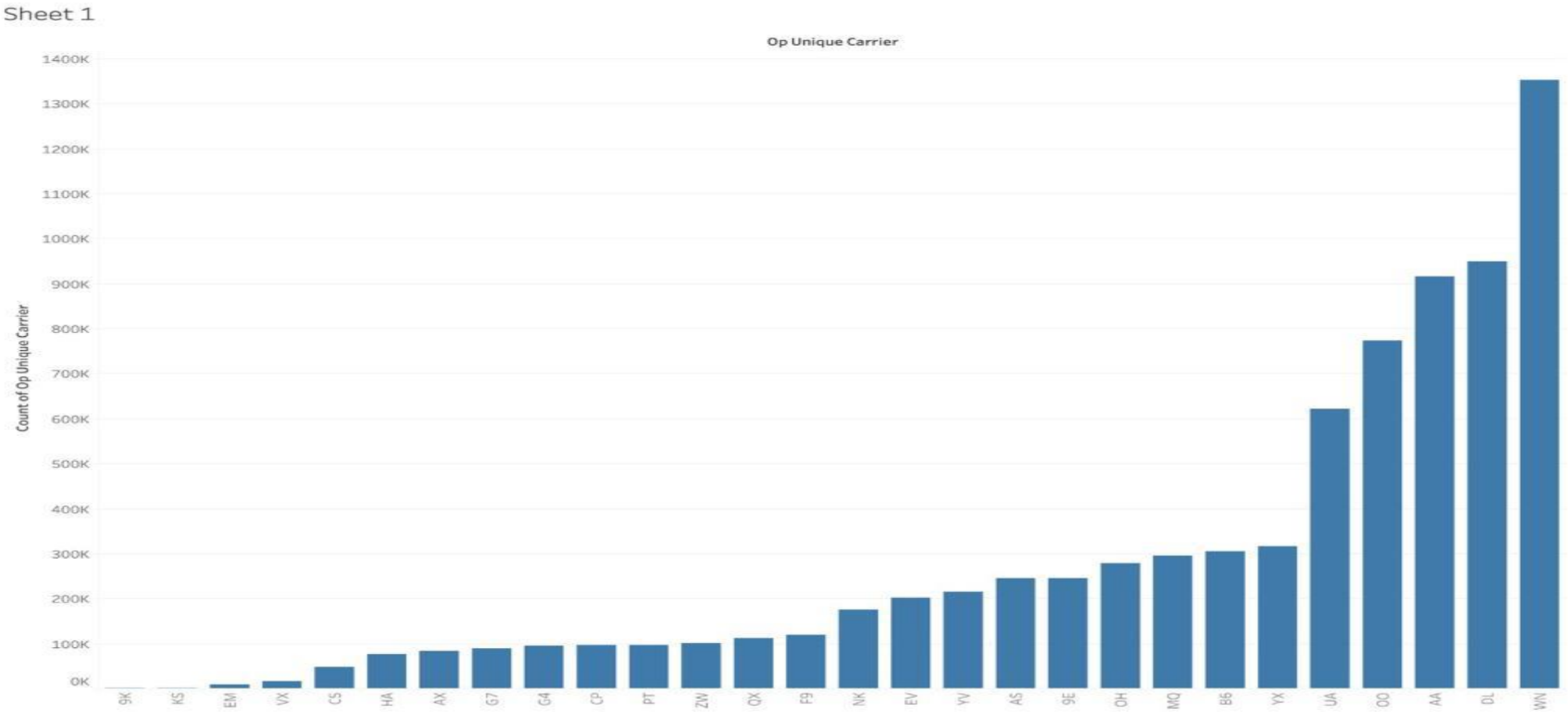
Departure Delay

- Frontier Airlines(F9) with the highest departure delay percentage of 28.7% and Hawaiian Airlines with the lowest percentage of 7.67%



DREXEL UNIVERSITY
LeBow College of Business

RANGE OF FLIGHTS(2018)



DREXEL UNIVERSITY
LeBow College of Business

Project Timeline

1. Data Collection and Preprocessing

2. Exploratory Data Analysis (EDA)

3. Identifying Delay Categories

4. Machine Learning Model Development

5. Insights & Key Results

DREXEL UNIVERSITY
LeBow College of Business

Dataset – Pre- Filter

~600K rows per dataset (per month)

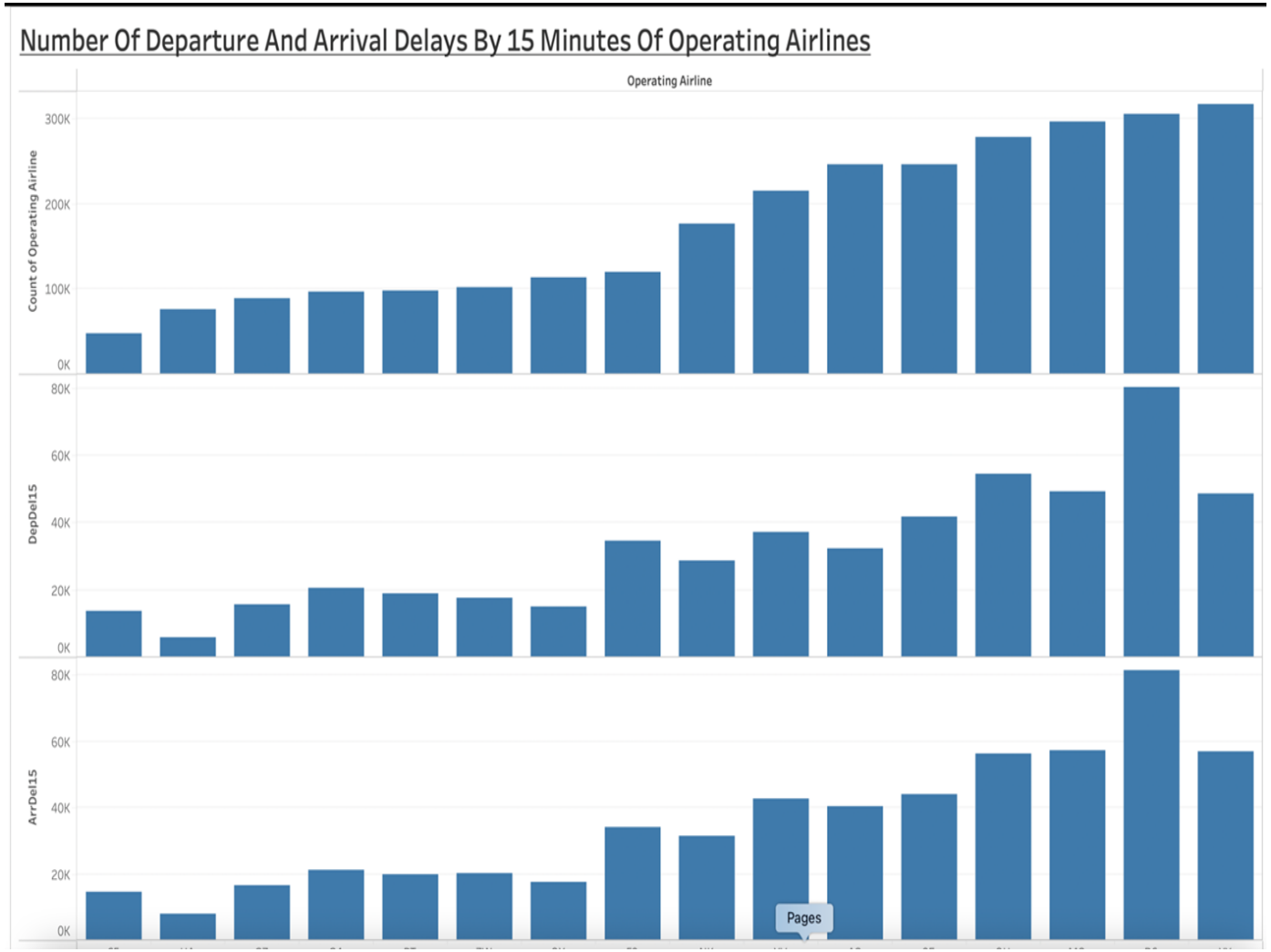
600K*12 months = 7.2 Million rows (per year)

7.2M*5 years = 36 million rows

DEPARTURE AND ARRIVAL DELAYS

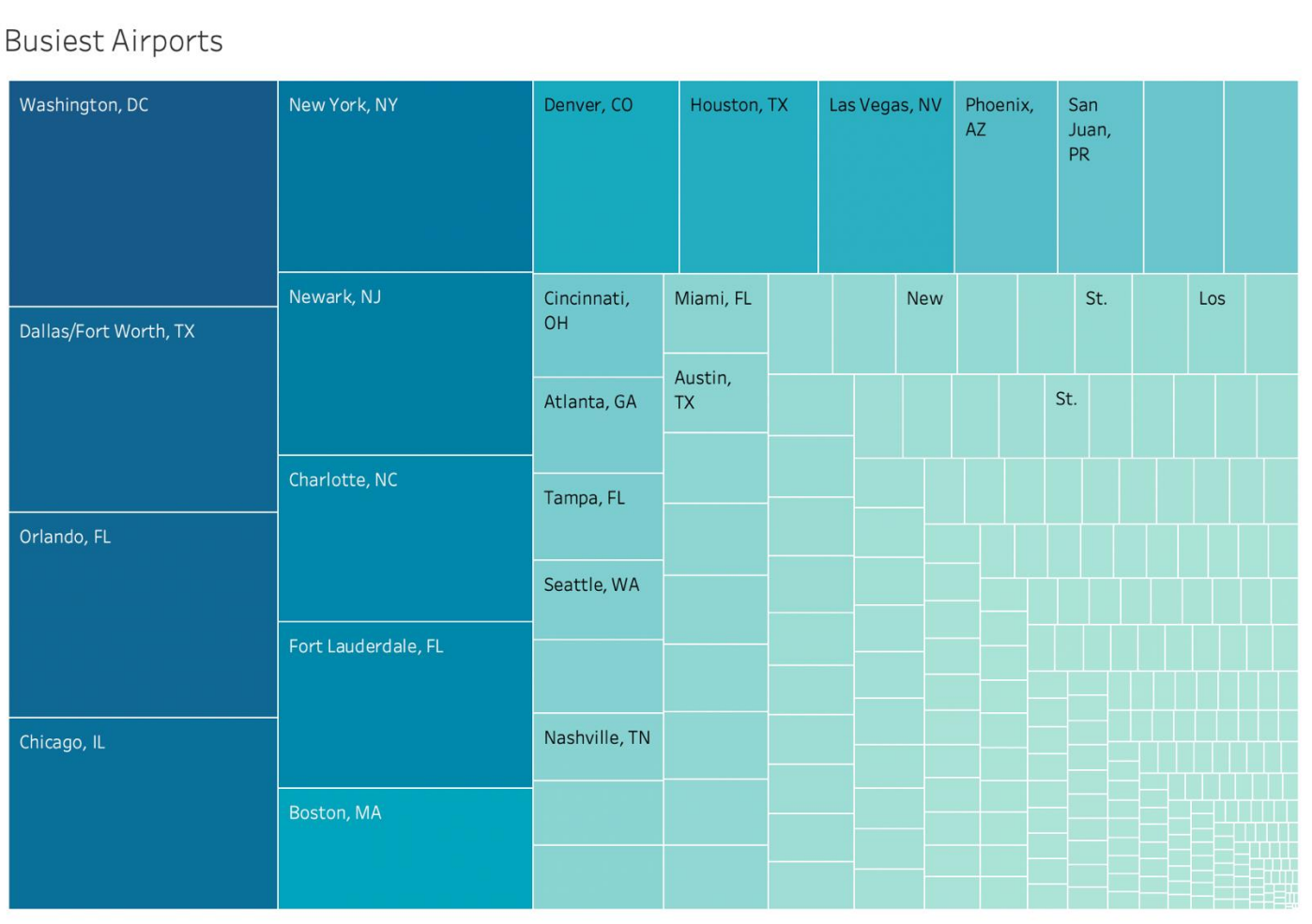
- Highest departure delay count of 80,223 Flights by JetBlue Airlines and the lowest Count of 5,865 flights by Hawaiian Airlines.

- Highest arrival delay count of 81,308 flights by Jetblue Airlines and the Lowest of 8,090 flights by Hawaiian Airlines.



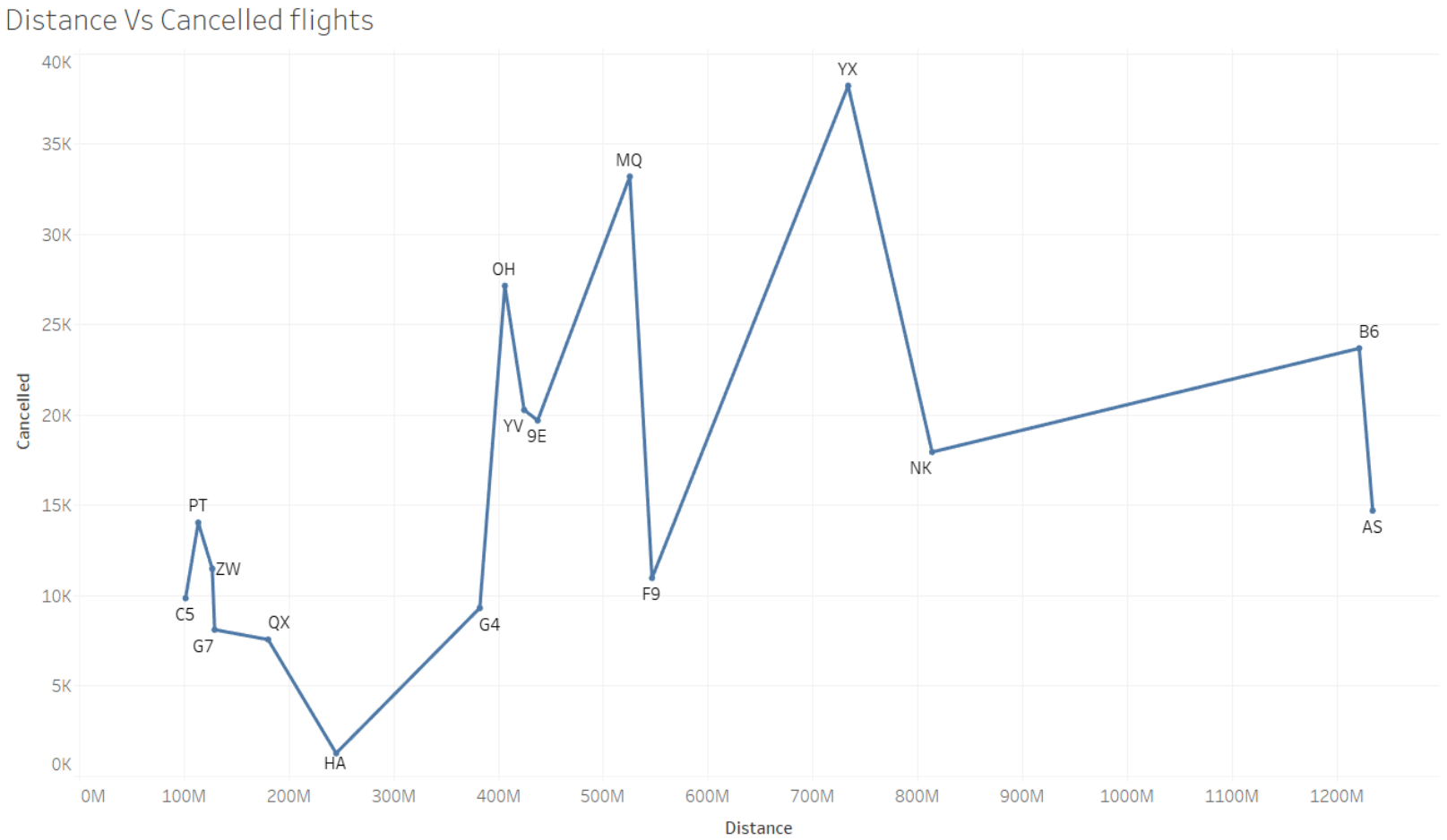
DREXEL UNIVERSITY
LeBow College of Business

Busiest Airports



DREXEL UNIVERSITY
LeBow College of Business

Distance vs Cancelled Flights

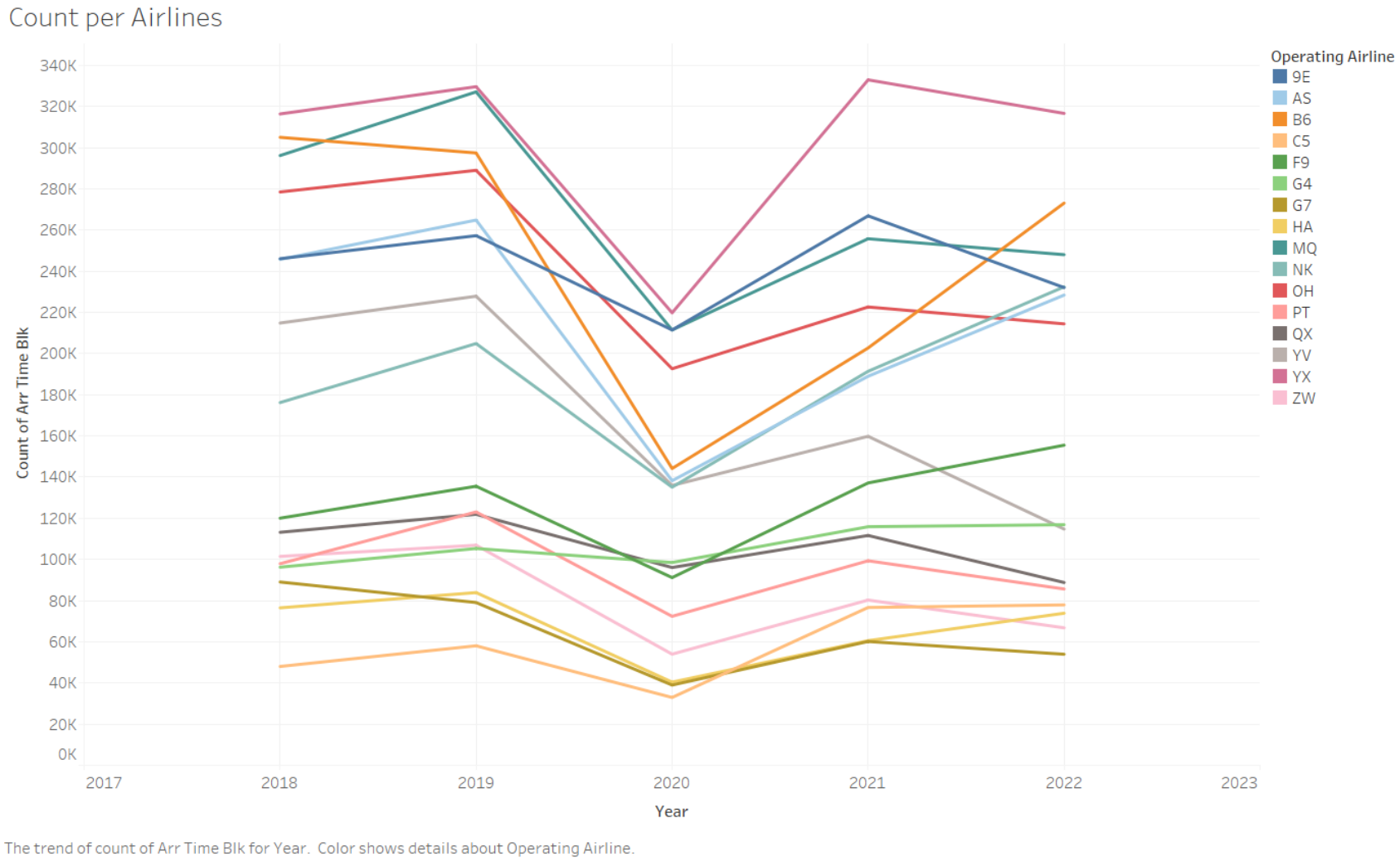


DREXEL UNIVERSITY
LeBow College of Business

FLIGHTS COUNTS YEARLY

Flights Count is where flights dropped for every airline in 2020.

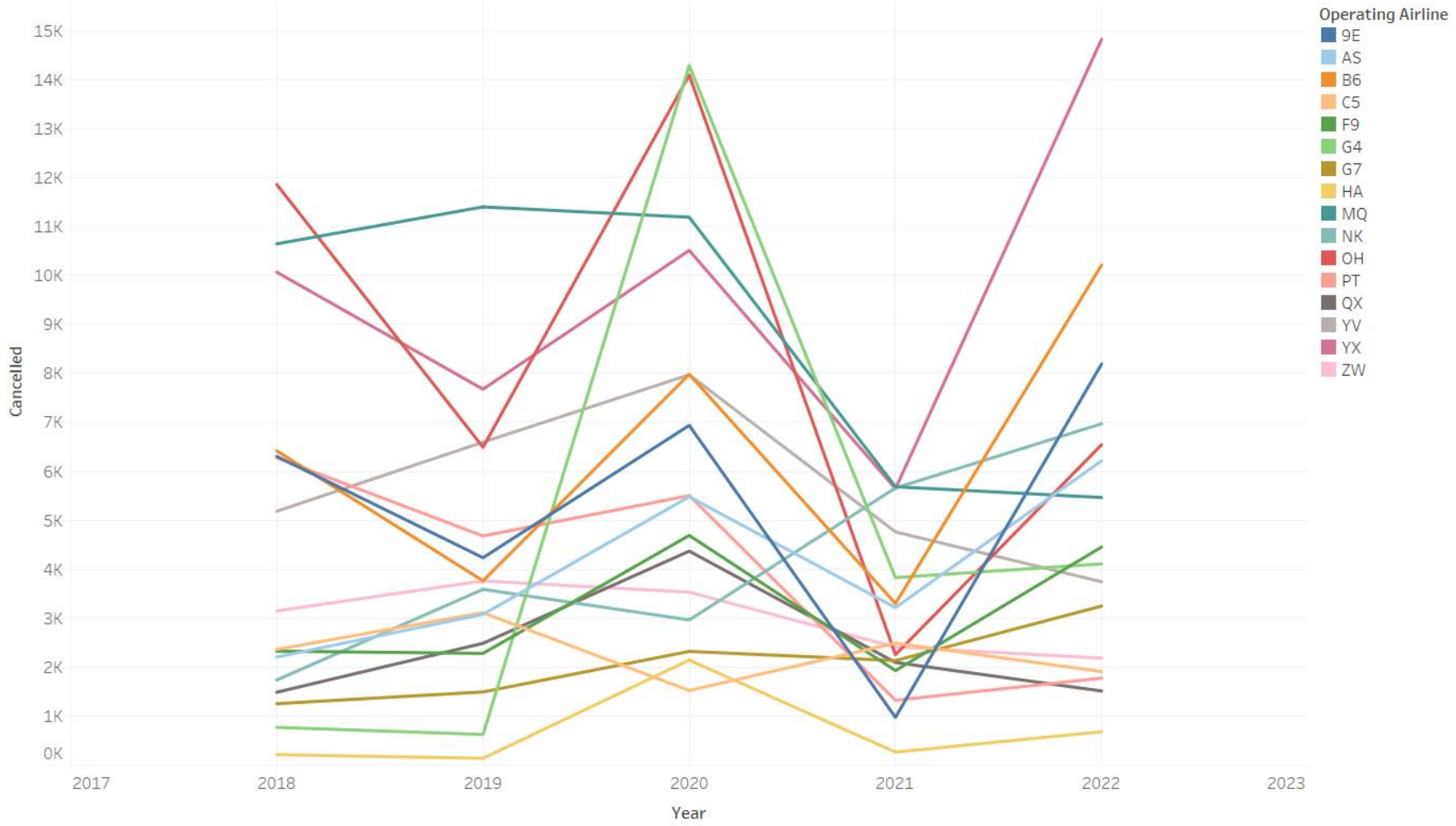
Taking YX for example
2018 - 316,636
2019 - 329,596
2020 - 219,772
2021 - 332,926
2022 - 316,622



DREXEL UNIVERSITY
LeBow College of Business

Cancellations per Year

Canceled flights where we see the most canceled flight were in 2020 by G4 and OH of around 14k/year.

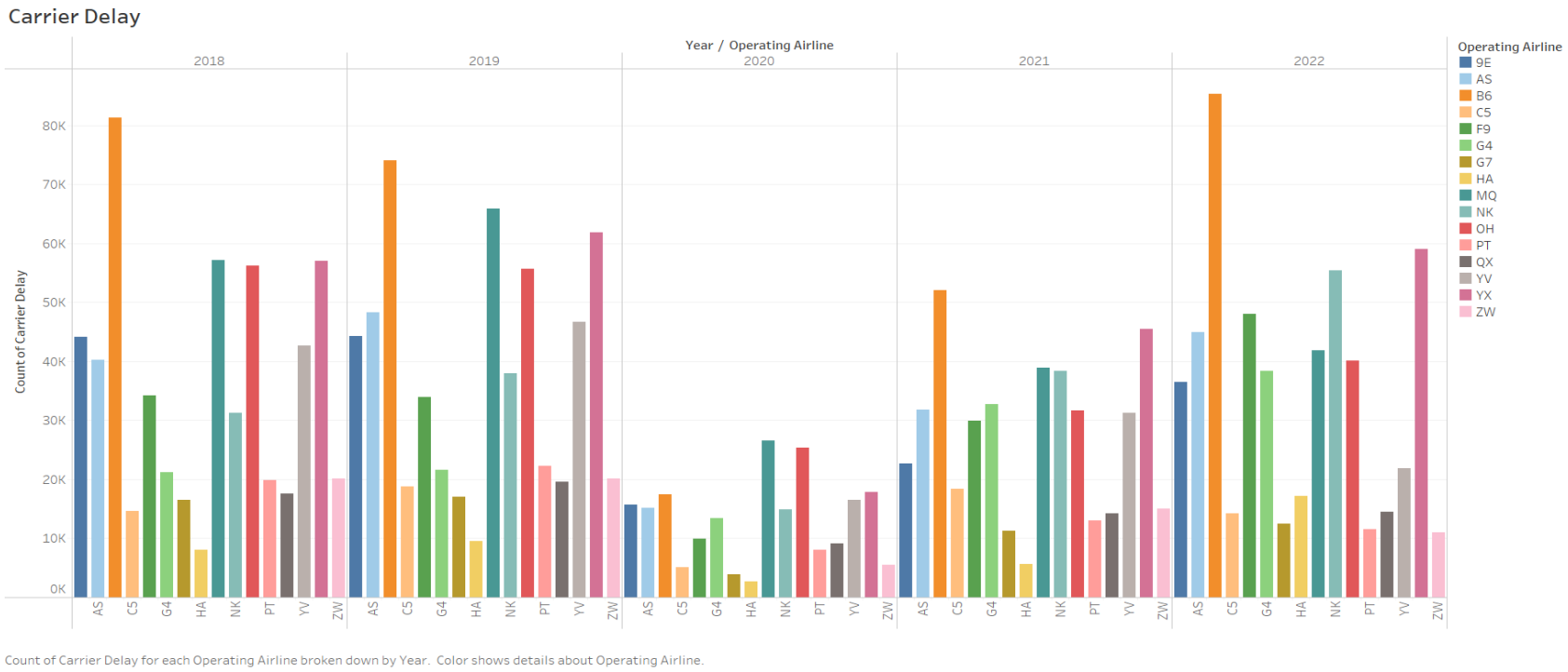


The trend of sum of Cancelled for Year. Color shows details about Operating Airline.

DREXEL UNIVERSITY LeBow College of Business

Carrier Delay

Carrier Delay by all the filtered airlines per year data, B6 leading the most carrier delays every year to HA having the most minor carrier delays per year.



Roadblocks

Kernel Failures
Not enough computing power
Lack of any real-time data of flights.