

Big Data

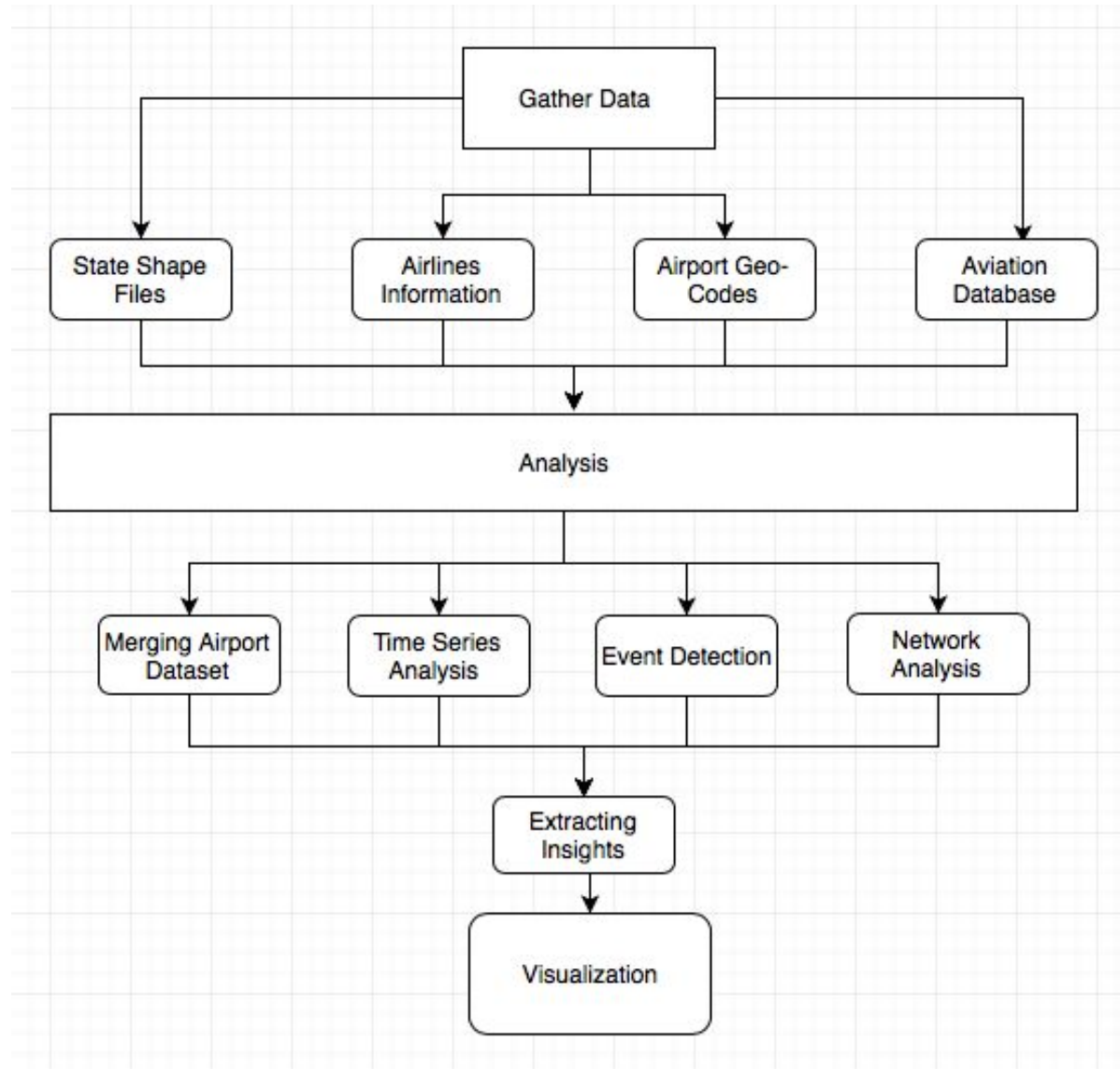
Descriptive Analysis for domestic flights in the United States

- Arnav Somani
- Hrafnkell Hjörleifsson
- Fernando Melchor
- Subhanshu Gupta

Insight of Data

- Flight records from 2014 - 2016
 - 17 million rows of data (flights)
 - 110 different attributes
- Geospatial information for airports
- Shapefiles for US states

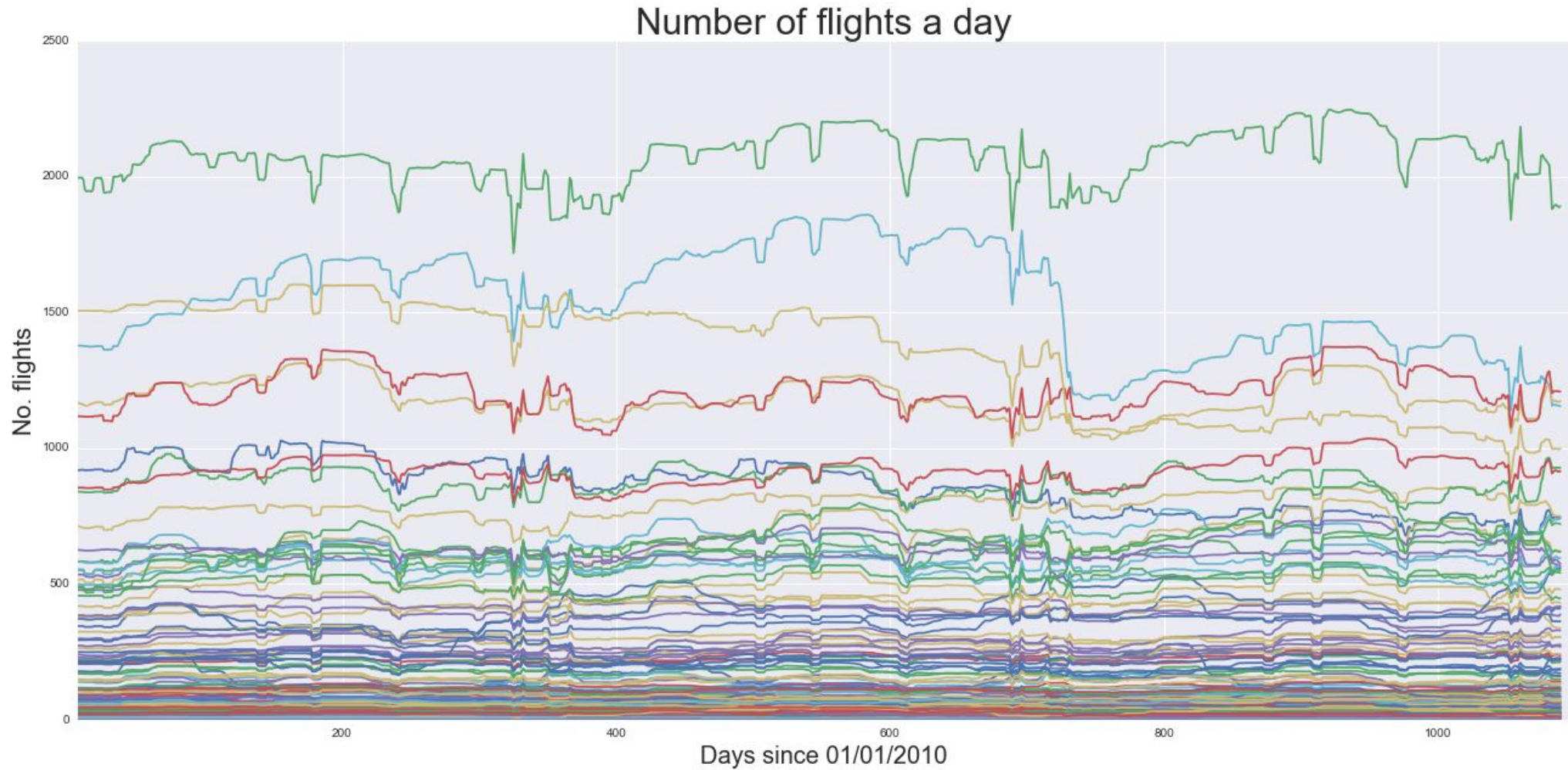
Project Flow



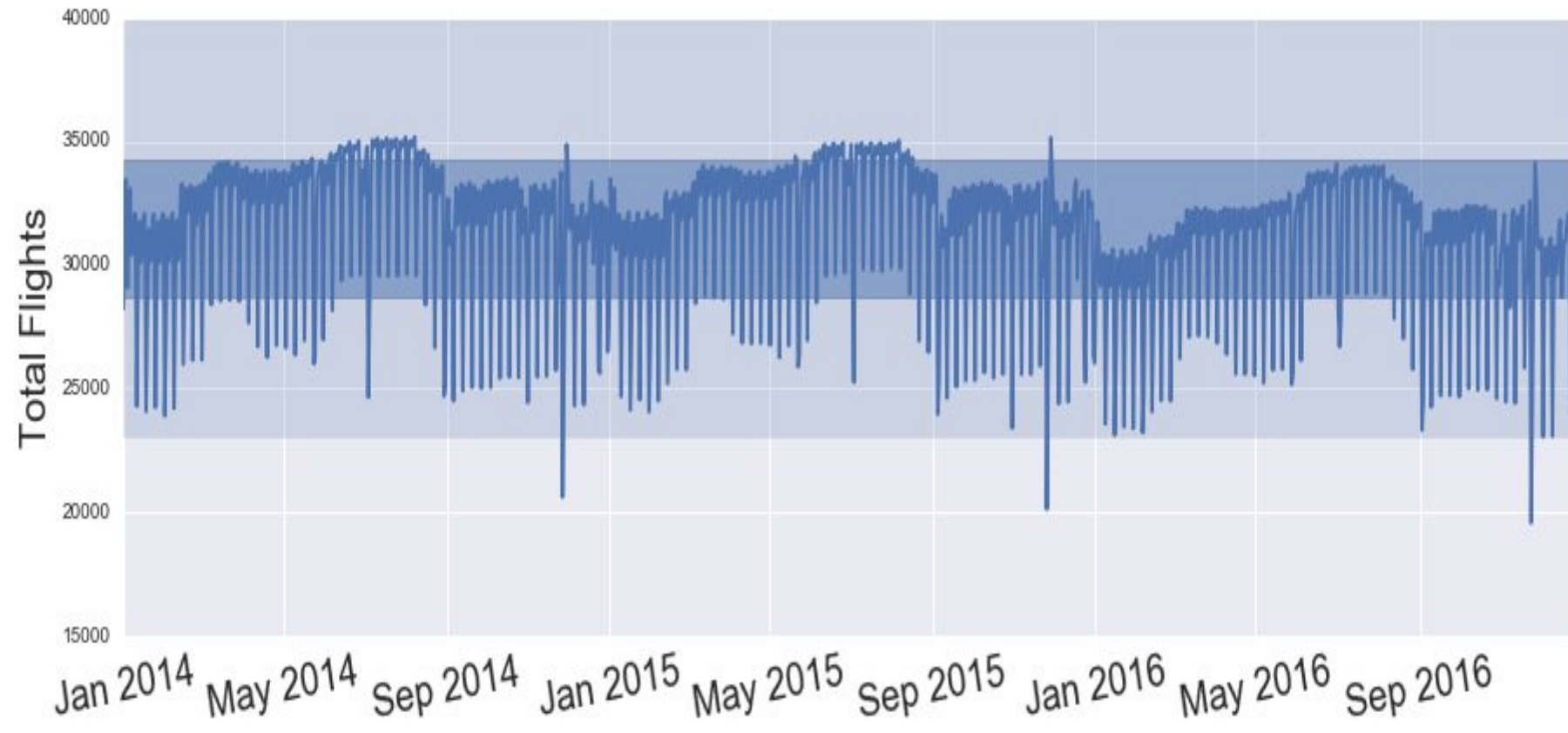
Methodology

- Extracting valuable information by running spark scripts on dumbo
 - turning 8 gb of data into easily manageable csv files (4 - 1,600 kb)
- Time Series Analysis
- Merging airports lat/lon information with airport identifiers
- Extracting facts about domestic aviation in the US
- Network analysis on statewide connectivity

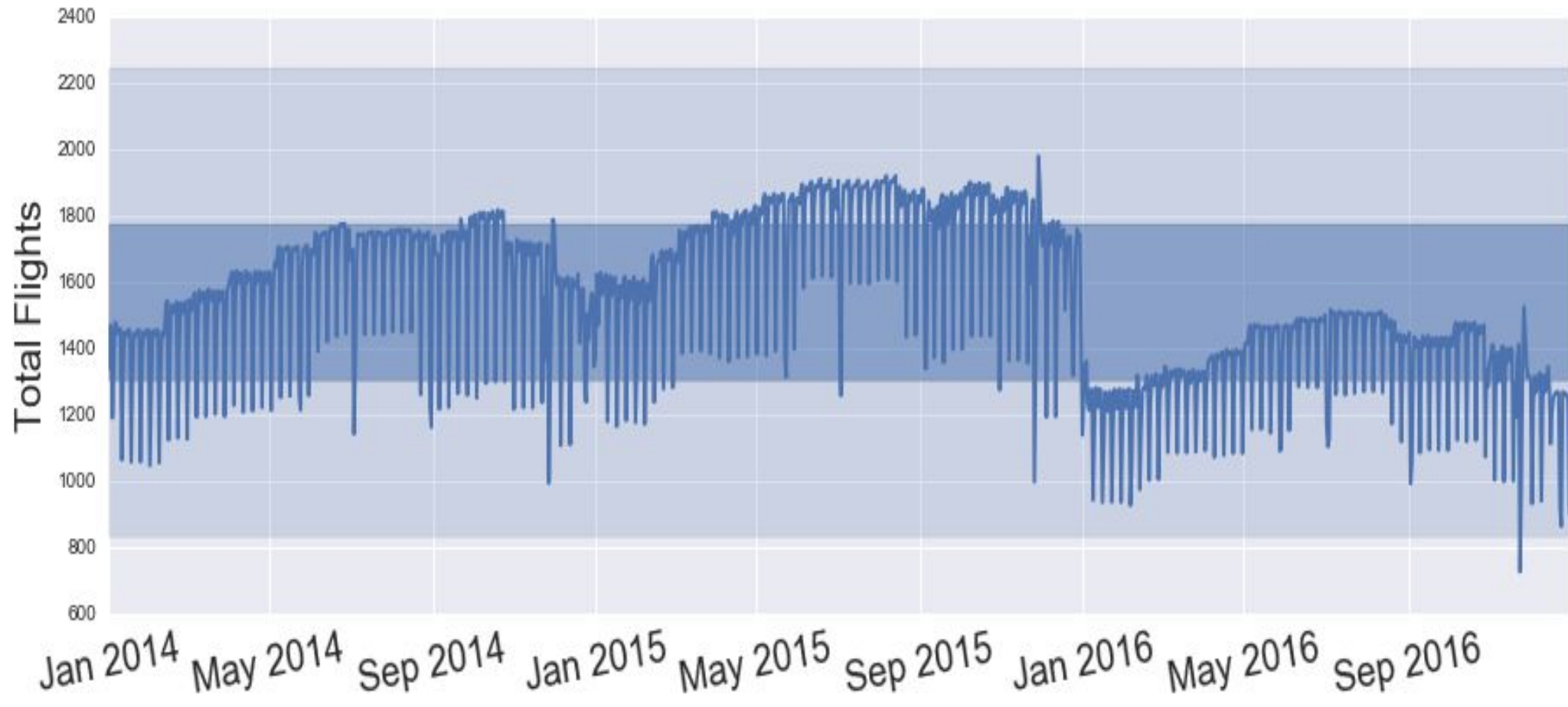
Total Flights per day grouped by Airport



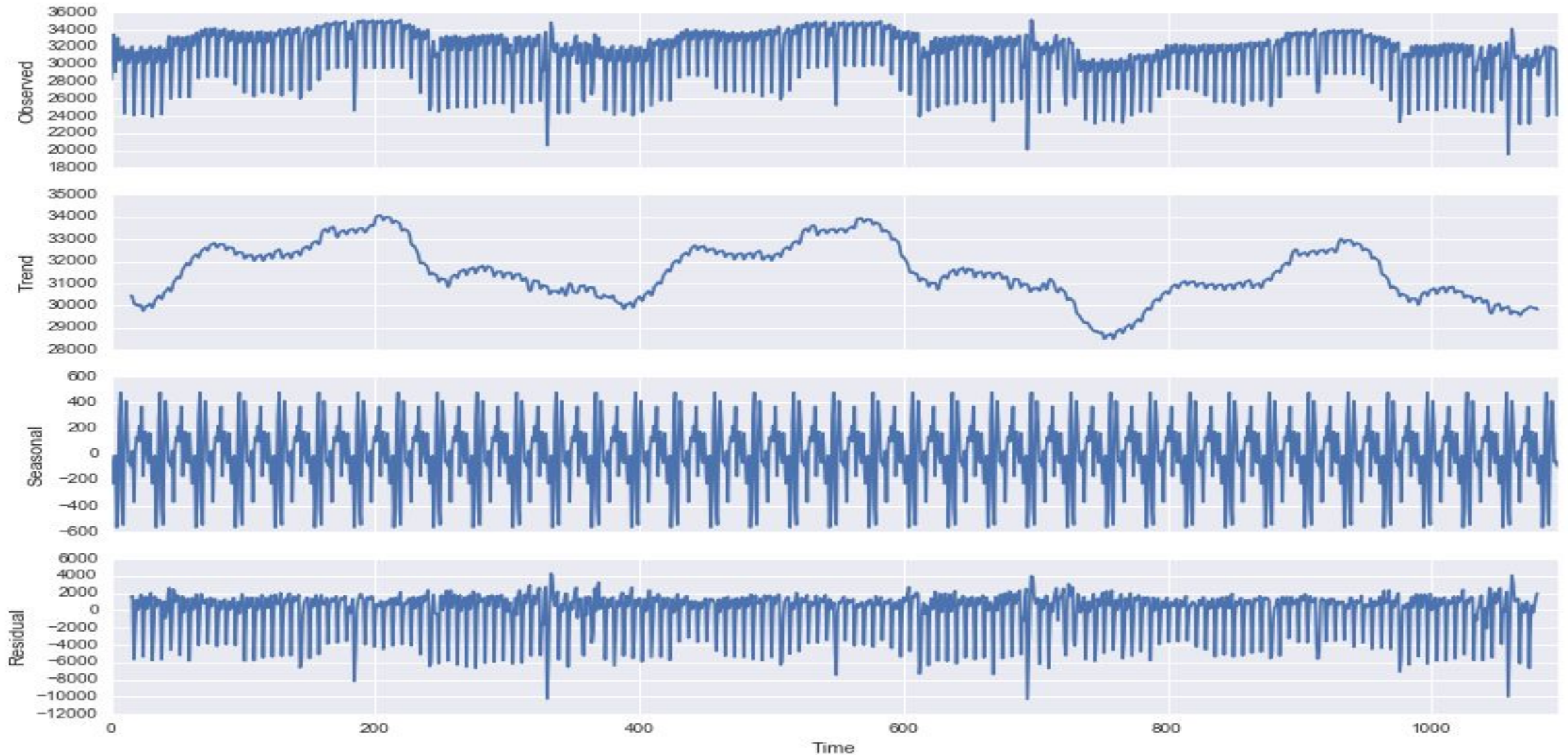
Total Flights - Searching for Event



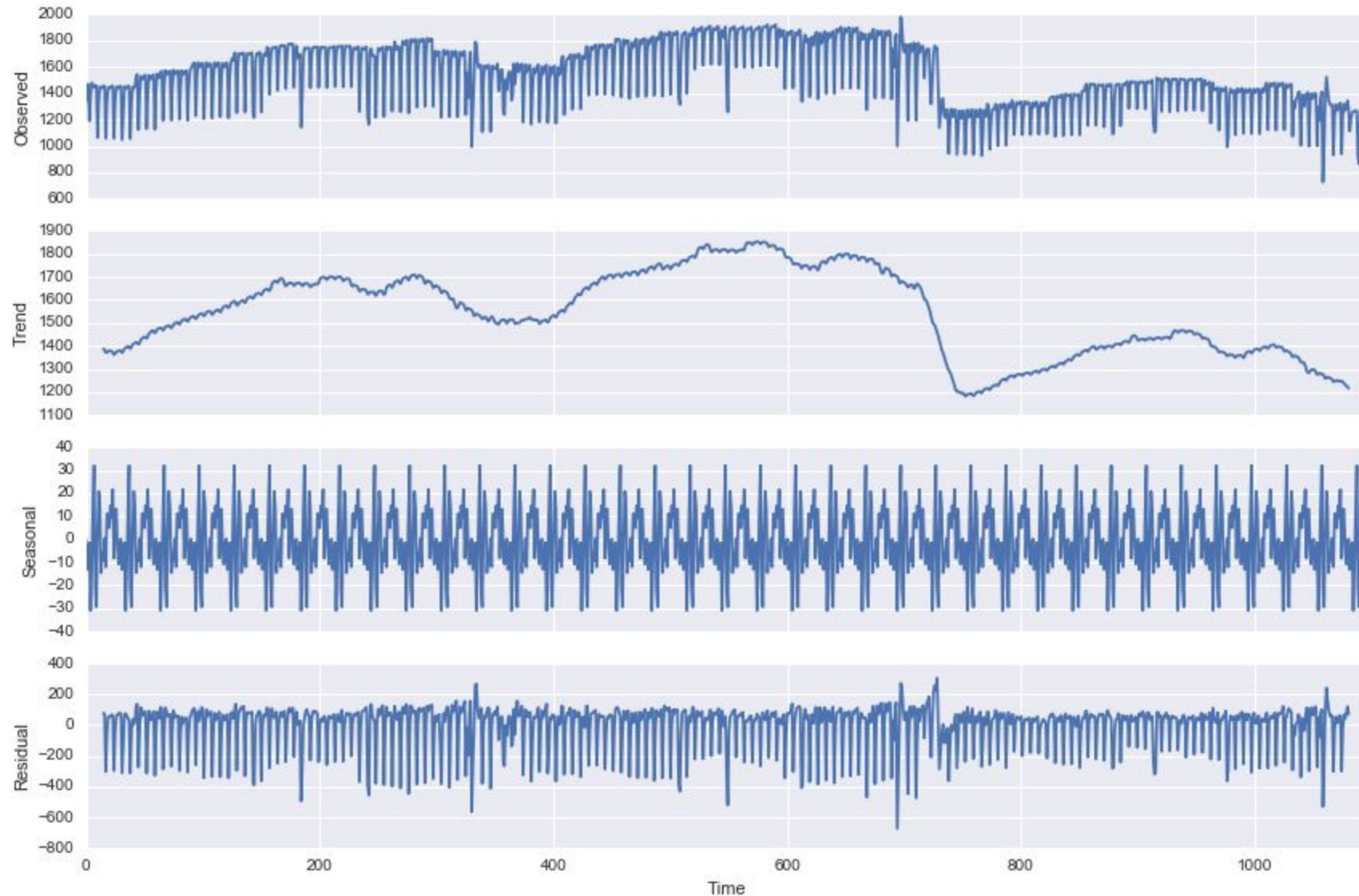
Chicago - Searching for Events



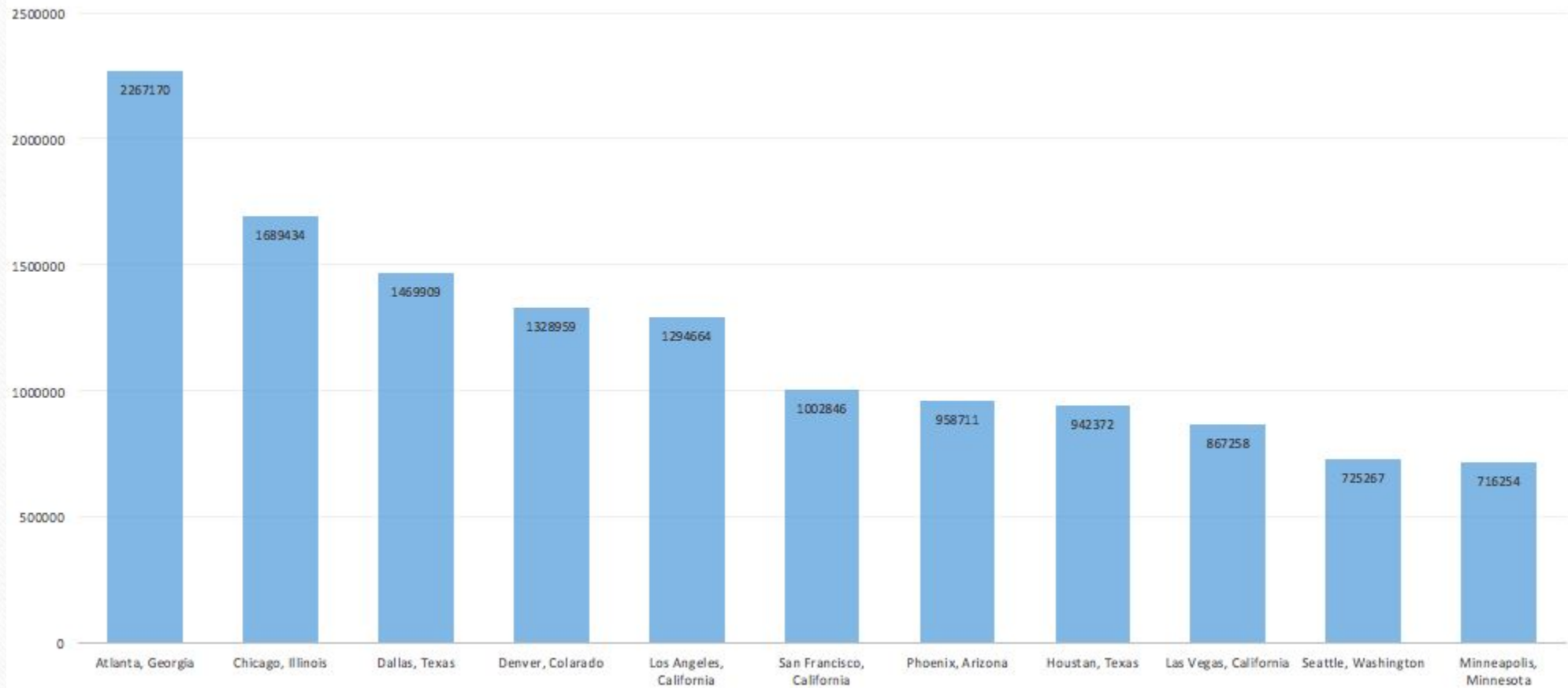
Seasonally Decomposing All Flights



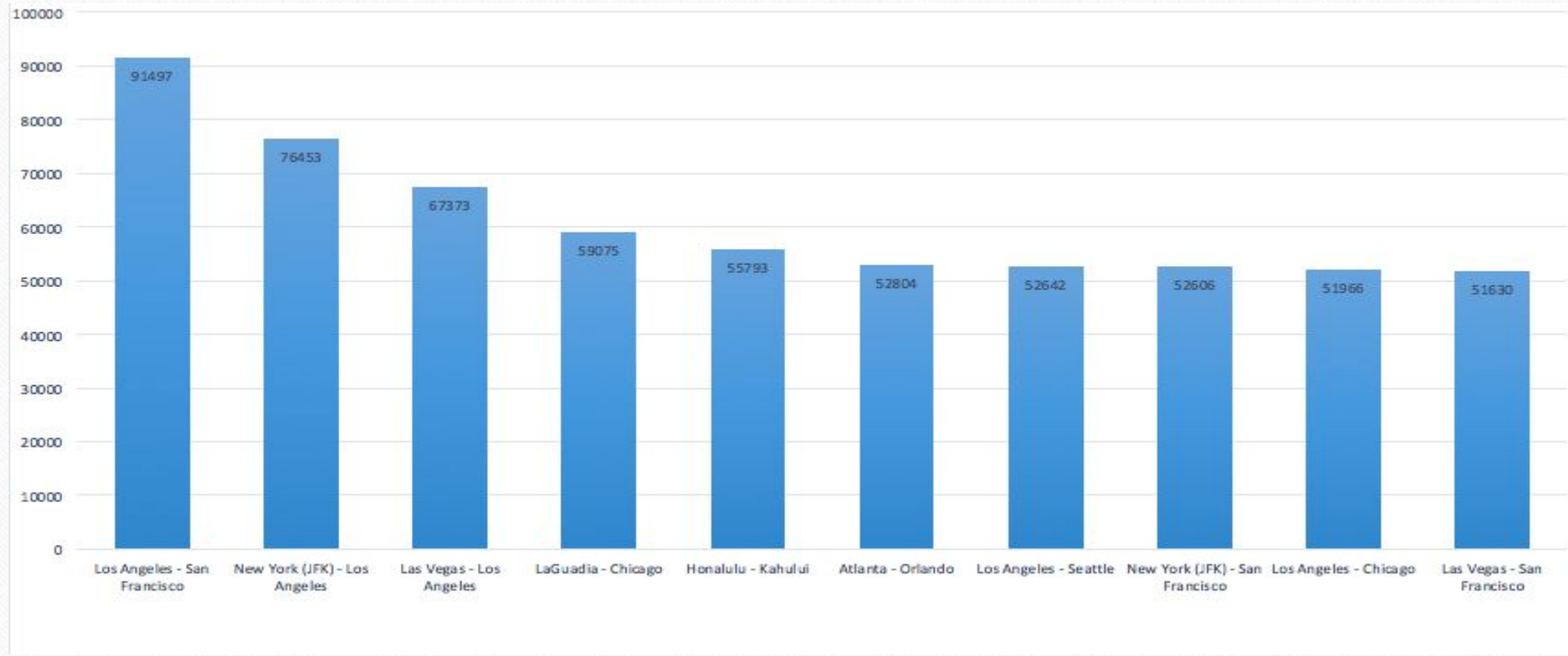
Seasonally Decomposing Chicago



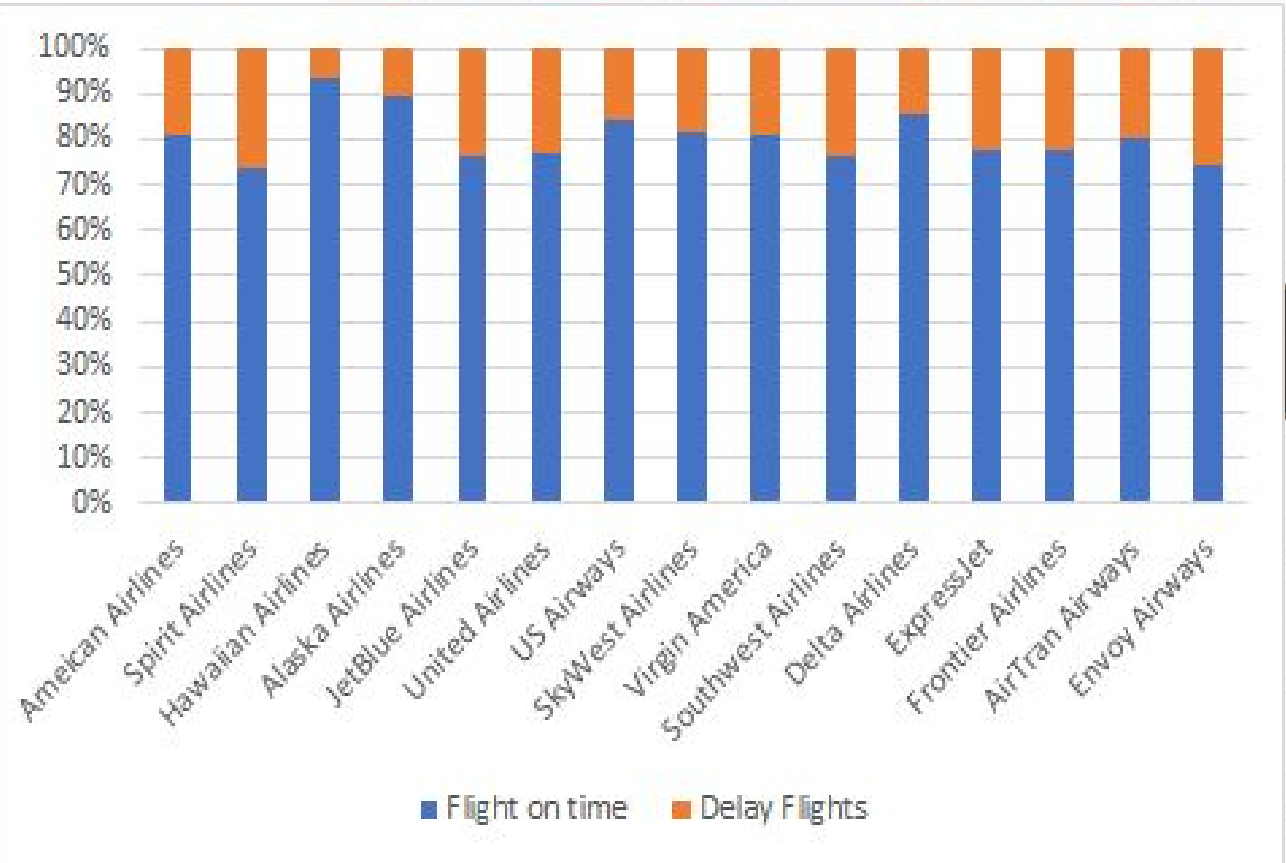
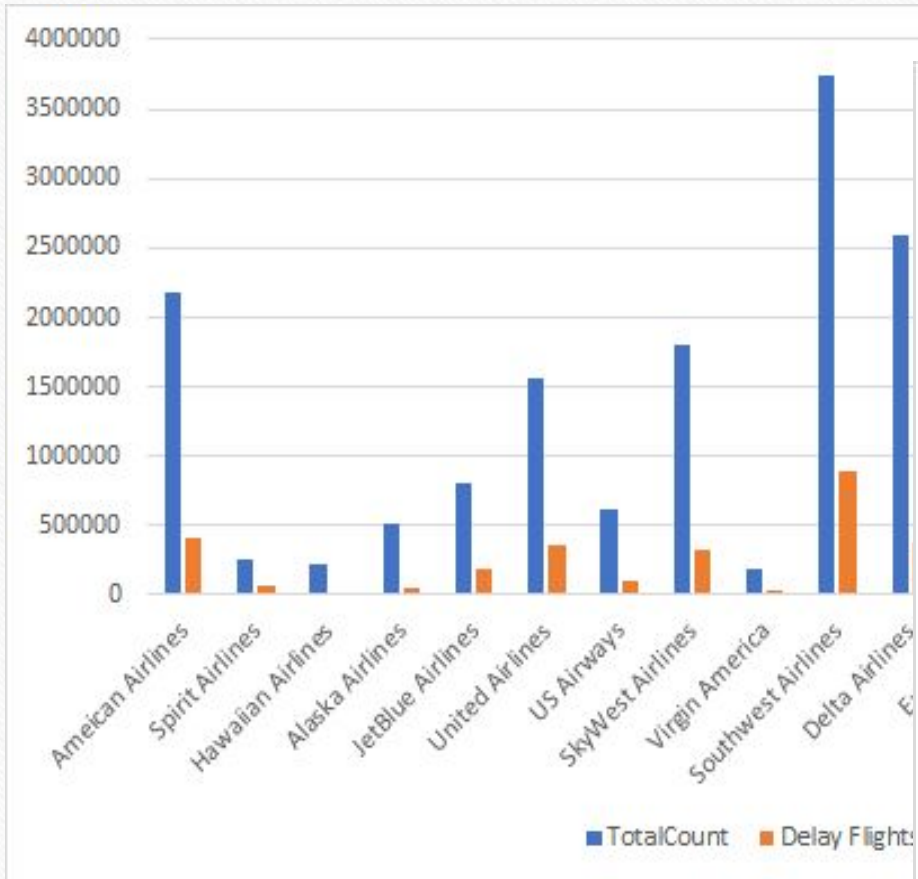
Busiest Airport



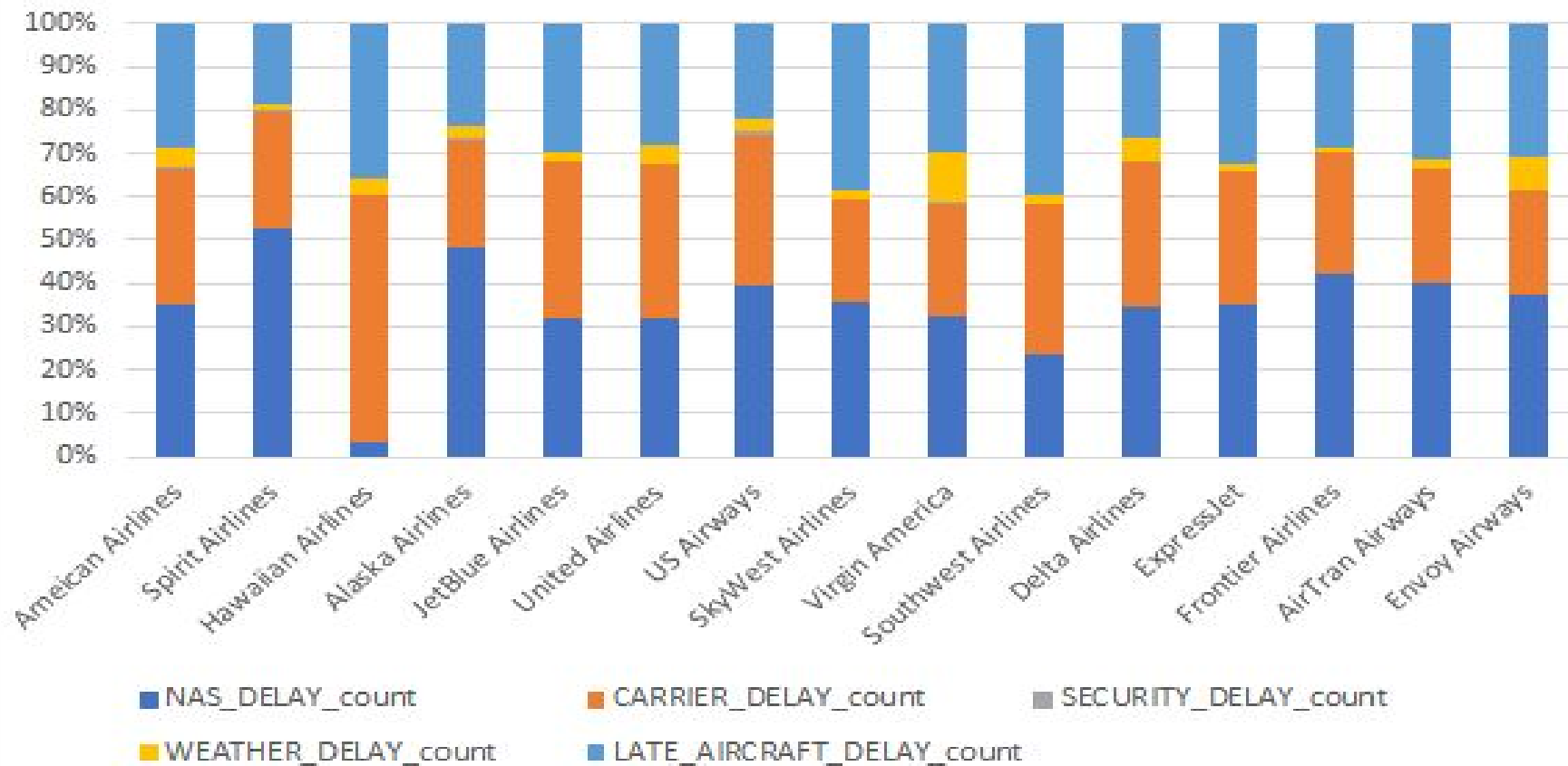
Busiest Route



Total Flights & Delayed Flights



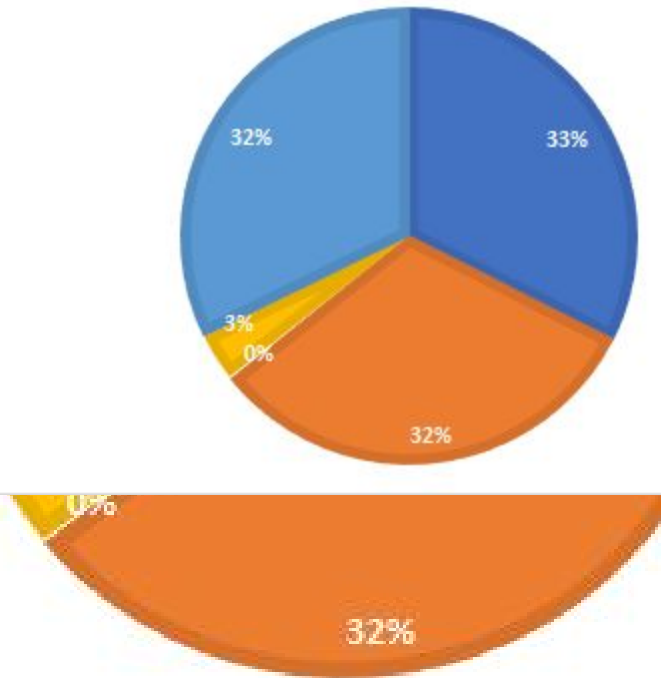
Reasons for Delay Airlines



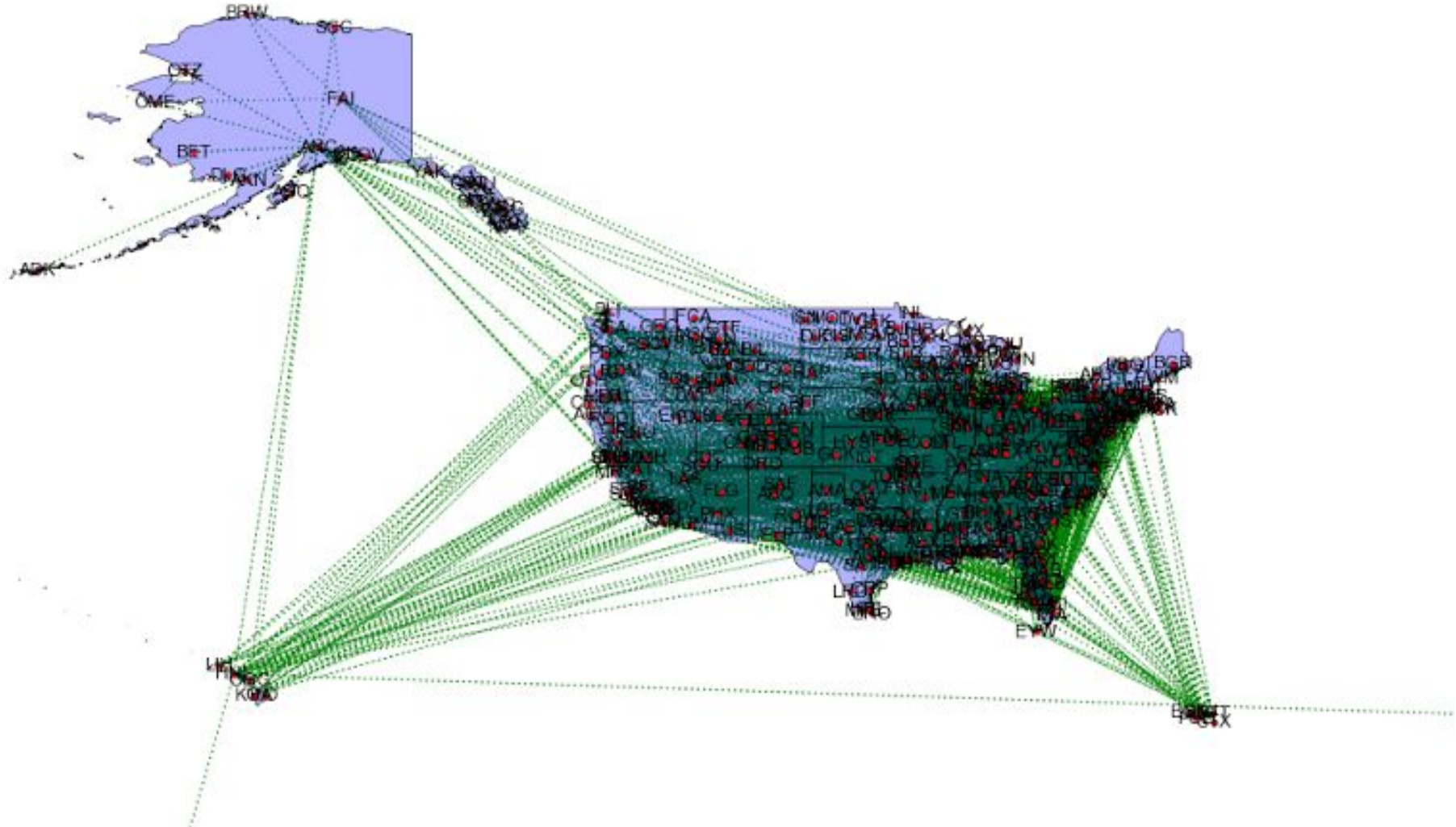
Delay Ratio

■ National Air System ■ CARRIER ■ SECURITY ■ WEATHER ■ LATE_AIRCRAFT

■ National Air System ■ CARRIER ■ SECURITY ■ WEATHER ■ LATE_AIRCRAFT

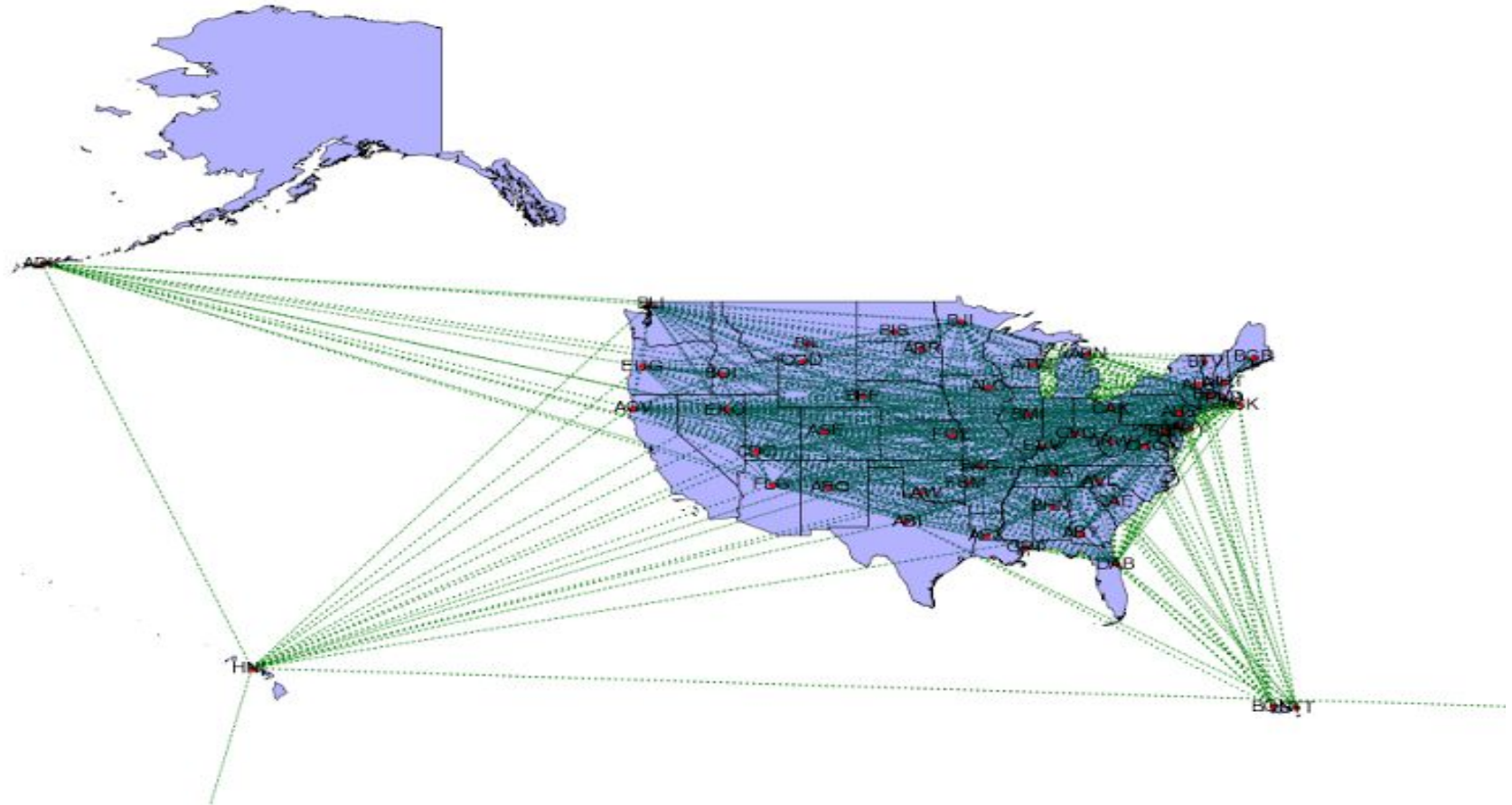


Network Analysis



Number of Nodes/Airports in the complete Network: 334

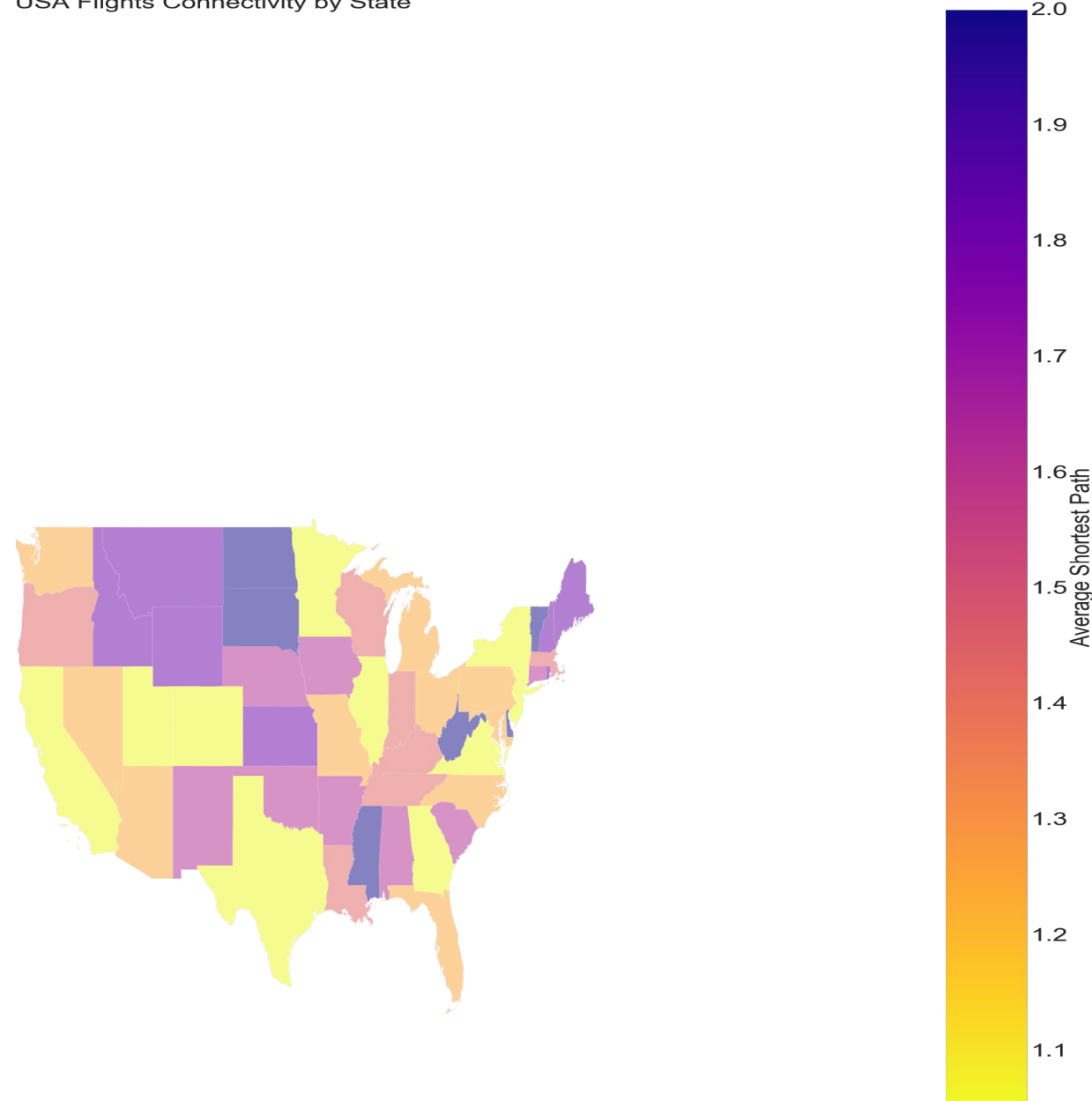
Reduced Network by State



Number of Nodes/Airports in the reduced Network: 55

Average Shortest Path - Connecting Flights

USA Flights Connectivity by State



Top 5

	NAME	Connectivity_Score
12	Illinois	1.054545
9	Georgia	1.072727
49	Texas	1.145455
5	Colorado	1.181818
17	Minnesota	1.200000

Last 10

	NAME	Connectivity_Score
54	Commonwealth of the Northern Mariana Islands	3.545455
32	American Samoa	2.618182
53	Guam	2.581818
18	Mississippi	1.945455
7	Delaware	1.945455
30	West Virginia	1.927273
23	North Dakota	1.890909
27	South Dakota	1.872727
29	Vermont	1.872727
15	Kansas	1.854545

Thank You
