FAU (MADE)

Analysis of Domestic Flights to USA Cities

Presented by Muhammad Talha

Main Question?

Introduction

Analyze 2020 flight data between major USA cities to understand the relationship between city attributes and incoming flights. Explore top destinations, routes, and conduct network analysis. Utilize ETL, Data Modeling, and Data Warehousing with two datasets: a Parquet of flight info and a CSV file with city details from 2020.

Motivation

Goals

01
Identify Trends

02
City Type Analysis

03
Top Cities and Routes

04
Temporal Analysis

05
Network Analysis

06
Data Visualization

Datasources

Air Flight Dataset

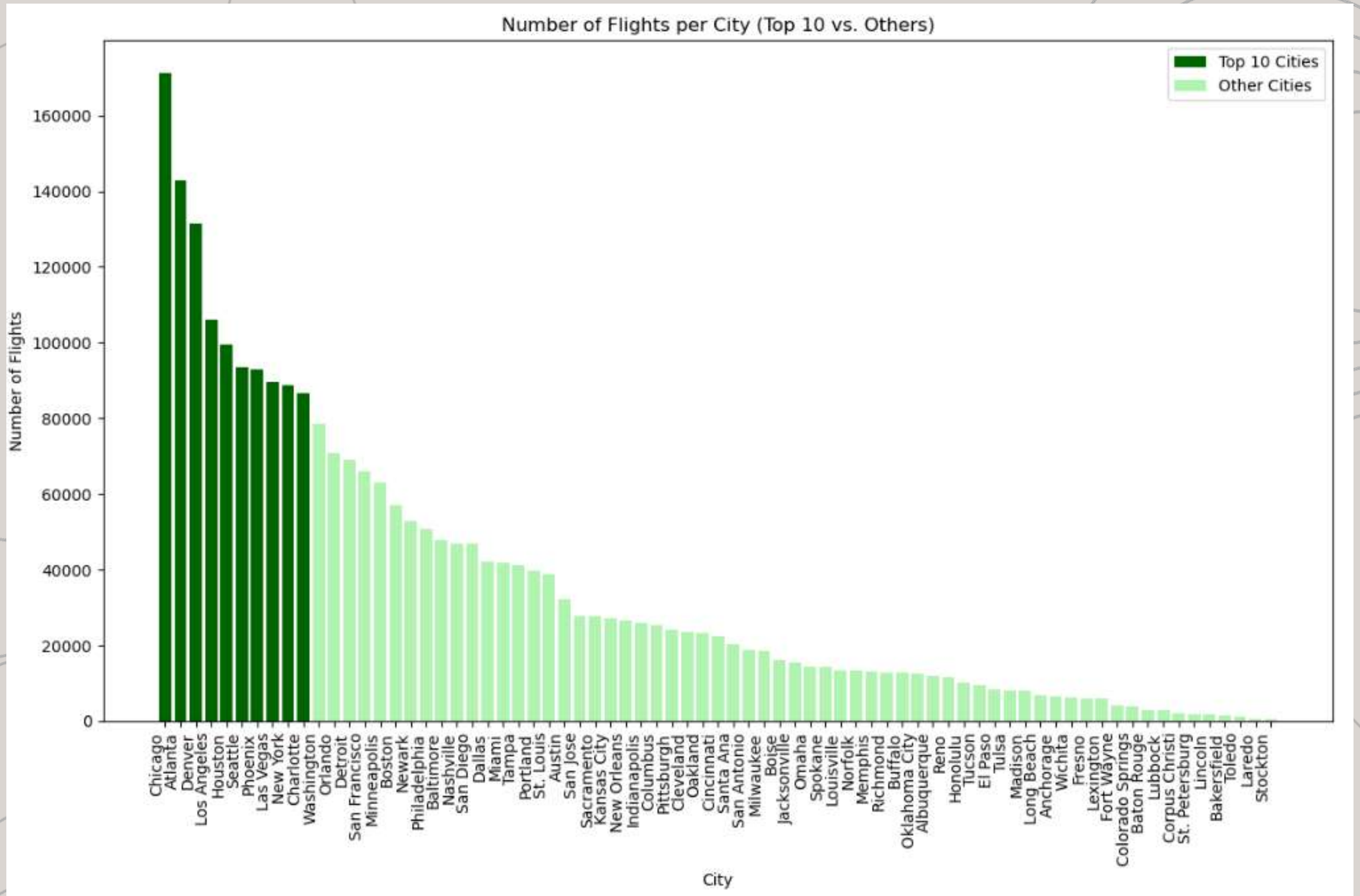
- **Data Type:** Parquet
- **License:** CC0: Public Domain
- **Description:** This dataset encompasses all flight information, including details on cancellations and delays by airline. The data spans from January 2018, extracted from the "On-Time" database of the TranStats data library.

Top 100 US Cities Population

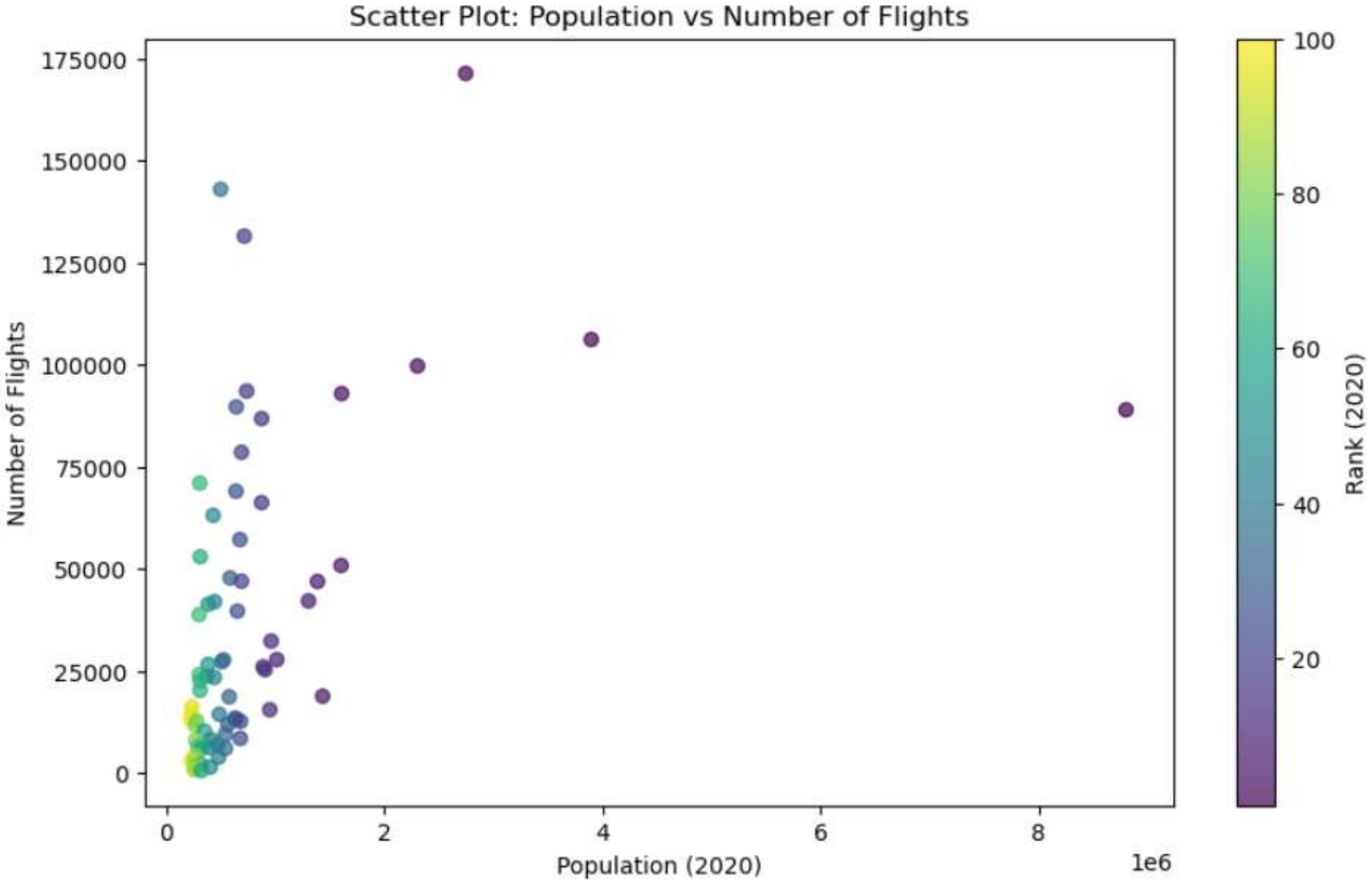
- **Data Type:** CSV
- **License:** CC0: Public Domain
- **Description:** The dataset includes information on the top 100 US cities by population, derived from Wikipedia. It lists each city, its rank based on 2020 population, area data, and populations in both 2020 and 2010.

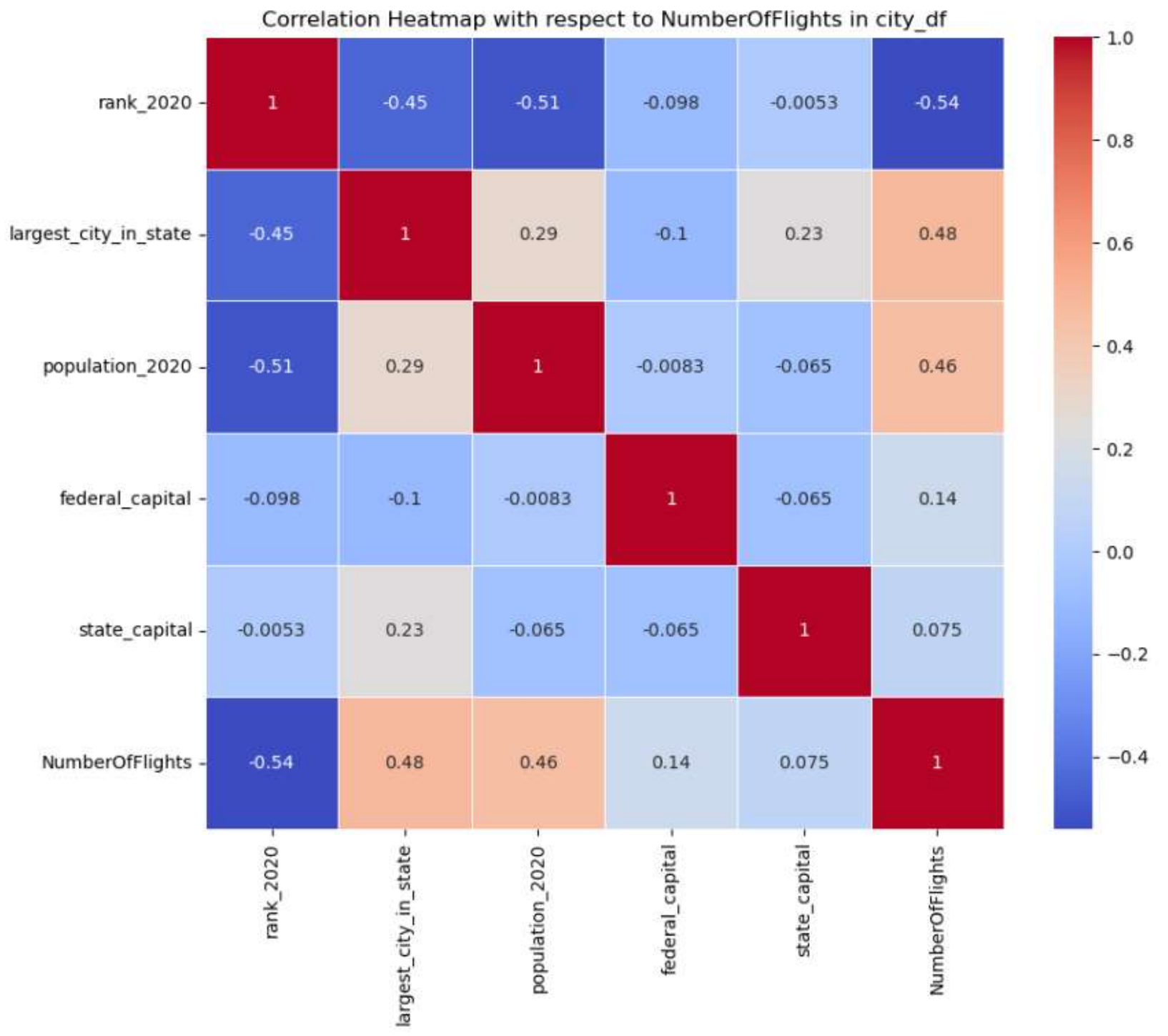
Analysis

Number of Flights per City


Number of Flights per City (Top 10 vs. Others)
Top 10 Cities
Other Cities
Number of Flights
0
20000
40000
60000
80000
100000
120000
140000
160000
City
Chicago
Atlanta
Denver
Los Angeles
Houston
Seattle
Phoenix
Las Vegas
New York
Charlotte
Washington
Orlando
Detroit
San Francisco
Minneapolis
Boston
Newark
Philadelphia
Baltimore
Nashville
San Diego
Dallas
Miami
Tampa
Portland
St. Louis
Austin
San Jose
Sacramento
Kansas City
New Orleans
Indianapolis
Columbus
Pittsburgh
Cleveland
Oakland
Cincinnati
Santa Ana
San Antonio
Milwaukee
Boise
Jacksonville
Omaha
Spokane
Louisville
Norfolk
Memphis
Richmond
Buffalo
Oklahoma City
Albuquerque
Reno
Honolulu
Tucson
El Paso
Tulsa
Madison
Long Beach
Anchorage
Wichita
Fresno
Lexington
Fort Wayne
Colorado Springs
Baton Rouge
Lubbock
Corpus Christi
St. Petersburg
Lincoln
Bakersfield
Toledo
Laredo
Stockton

City Attributes vs. Number of Flights


Scatter Plot: Population vs Number of Flights
Number of Flights
175000
150000
125000
100000
75000
50000
25000
0
0
2
4
6
8
1e6
Population (2020)
Rank (2020)
100
80
60
40
20


Correlation Heatmap with respect to NumberOfFlights in city_df
rank_2020
largest_city_in_state
population_2020
federal_capital
state_capital
NumberOfFlights
1
-0.45
-0.51
-0.098
-0.0053
-0.54
-0.45
1
0.29
-0.1
0.23
0.48
-0.51
0.29
1
-0.0083
-0.065
0.46
-0.098
-0.1
-0.0083
1
-0.065
0.14
-0.0053
0.23
-0.065
-0.065
1
0.075
-0.54
0.48
0.46
0.14
0.075
1
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4

Major Relations Analysis

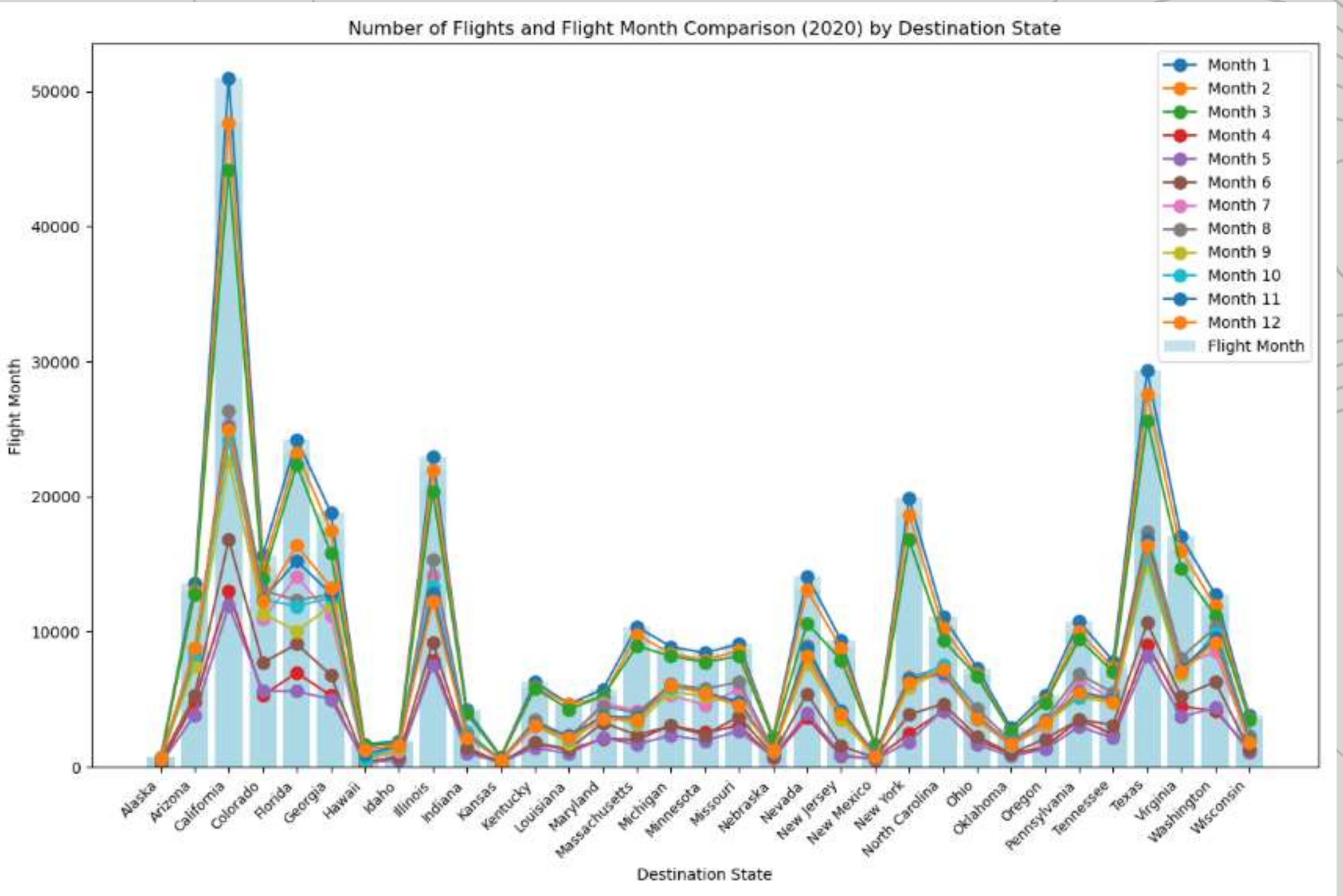
01
Rank vs. Population, Largest City in State, NumberOfFlights

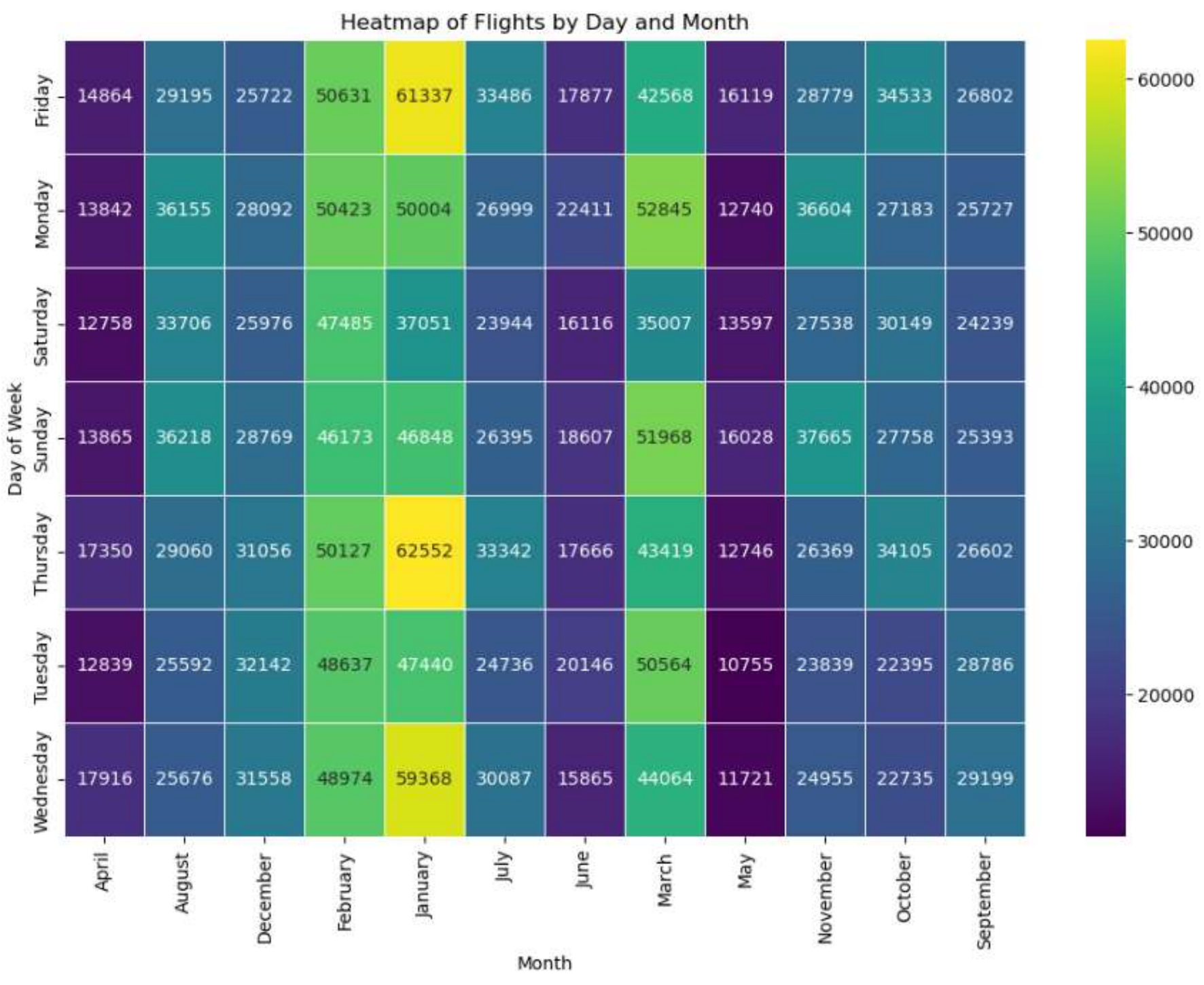
02
Largest City in State vs. Population, State Capital, NumberOfFlights

03
Population vs. NumberOfFlights

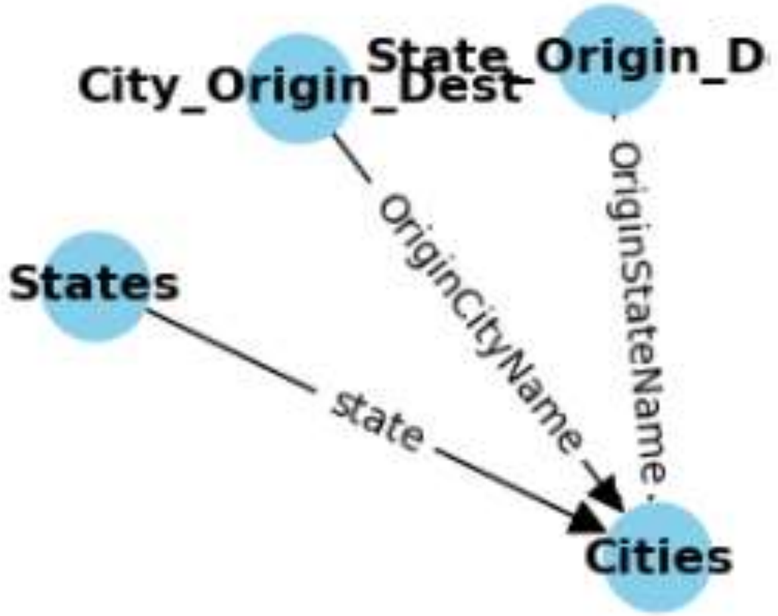
04
Federal Capital vs. NumberOfFlights

Time Analysis


Number of Flights and Flight Month Comparison (2020) by Destination State
Flight Month
Destination State
0
10000
20000
30000
40000
50000
Month 1
Month 2
Month 3
Month 4
Month 5
Month 6
Month 7
Month 8
Month 9
Month 10
Month 11
Month 12
Flight Month
Alaska
Arizona
California
Colorado
Florida
Georgia
Hawaii
Idaho
Illinois
Indiana
Kansas
Kentucky
Louisiana
Maryland
Massachusetts
Michigan
Minnesota
Missouri
Nebraska
Nevada
New Jersey
New Mexico
New York
North Carolina
Ohio
Oklahoma
Oregon
Pennsylvania
Tennessee
Texas
Virginia
Washington
Wisconsin


Heatmap of Flights by Day and Month
Day of Week
Friday
Monday
Saturday
Sunday
Thursday
Tuesday
Wednesday
14864 29195 25722 50631 61337 33486 17877 42568 16119 28779 34533 26802
13842 36155 28092 50423 50004 26999 22411 52845 12740 36604 27183 25727
12758 33706 25976 47485 37051 23944 16116 35007 13597 27538 30149 24239
13865 36218 28769 46173 46848 26395 18607 51968 16028 37665 27758 25393
17350 29060 31056 50127 62552 33342 17666 43419 12746 26369 34105 26602
12839 25592 32142 48637 47440 24736 20146 50564 10755 23839 22395 28786
17916 25676 31558 48974 59368 30087 15865 44064 11721 24955 22735 29199
April
August
December
February
January
July
June
March
May
November
October
September
Month
60000
50000
40000
30000
20000

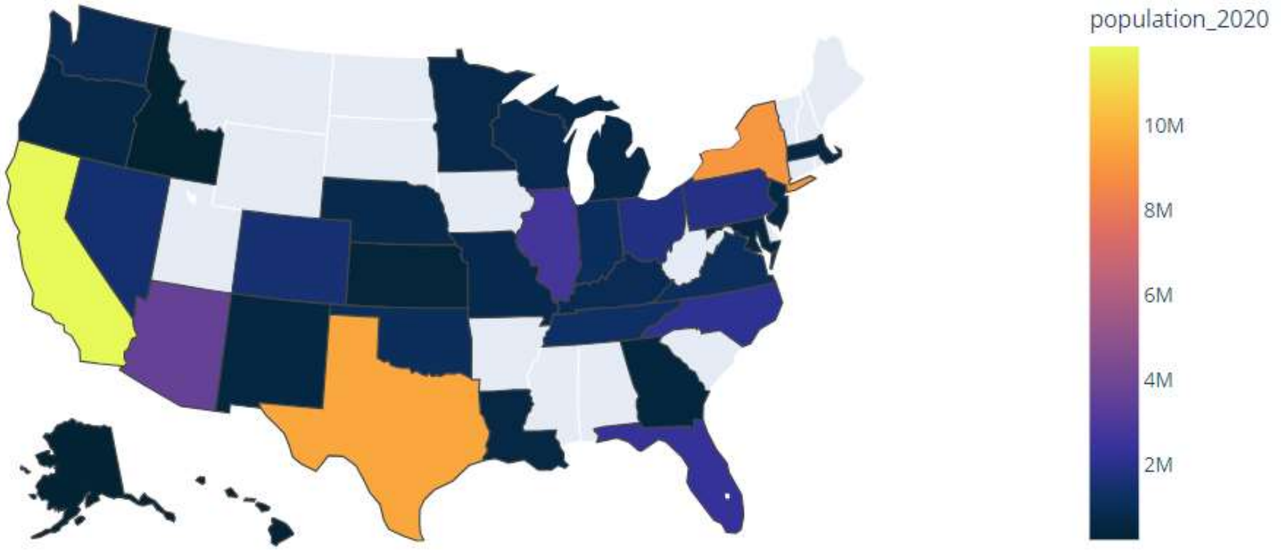
Data Engineering and ETL


City_Origin_Dest
State_Origin_D
OriginCityName
OriginStateName
States
state
Cities

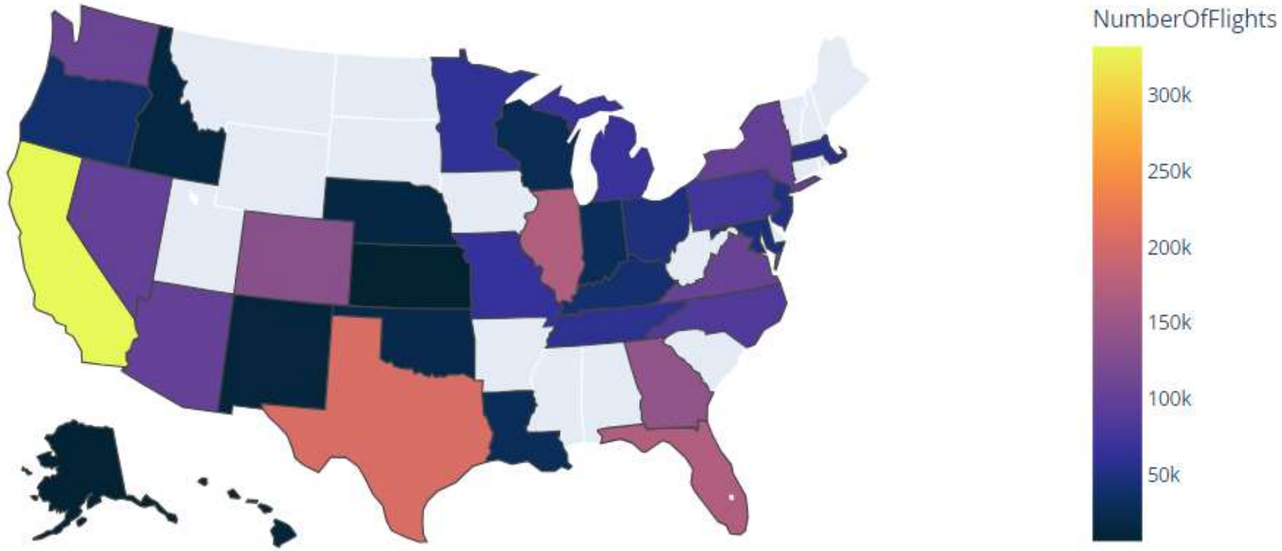
Flights

State Names and Number of Flights in comparison with State Names and their Population

Population

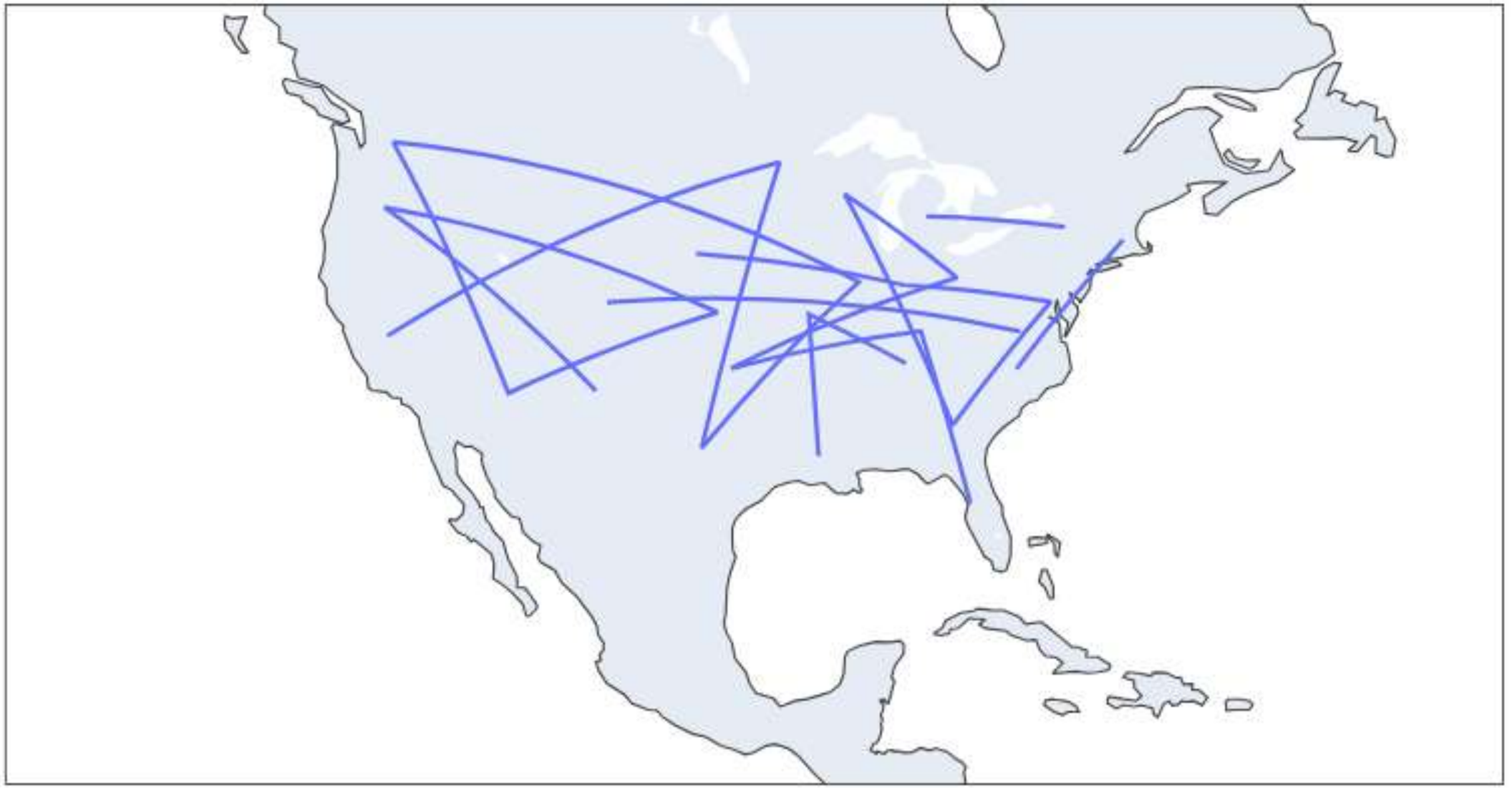


Flights Traffic



State Connections Map of Flights

State Connections Map



Flights Traffic Heatmap by City in 2020

Flights Traffic Heatmap by City in 2020

Conclusion

our exploration of flight data highlights the prominence of specific cities in air travel. The major relations analysis unveils correlations between city attributes and flight patterns, offering valuable insights for further investigation. Understanding these dynamics is crucial for efficient planning and decision-making in the realm of air travel.

Thank You