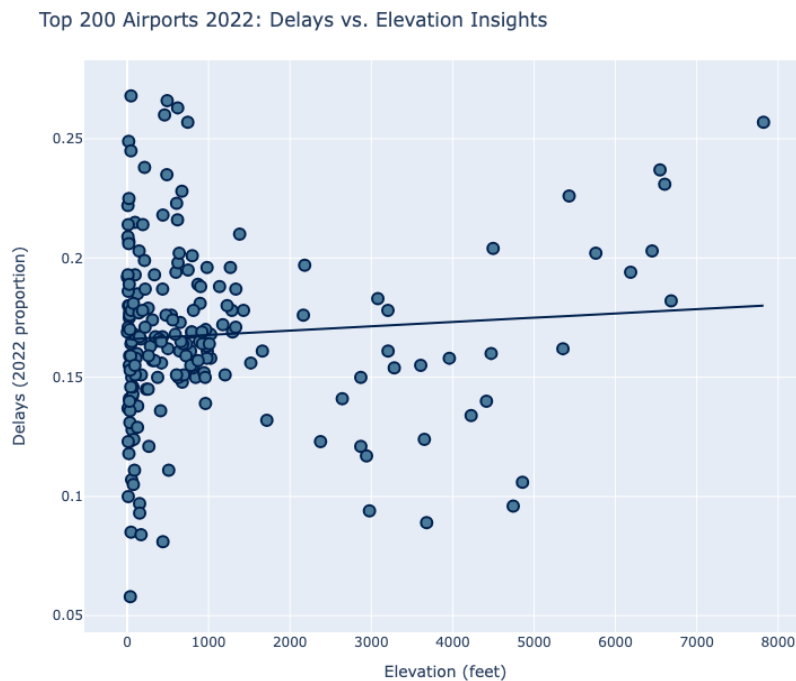


Part 1 Submission Details

- ☒ Submit a csv of your data through LearningSuite.
 - I attached my csv file, 2022Airports.csv, for submission.
- ☒ Link to the github repo with your code and data in the submission notes.
 - Github repo: <https://github.com/lucystorts/386Project>
- ☒ Submit a pdf, png, or jpeg or at least one figure you are using in your EDA.
 - Figure generated during EDA:



(I also attached this as a png)

Introduction

At the onset of the COVID pandemic in 2020, many opted to stop traveling and there was a noticeable decline in airport travel. By 2022, airports seemed to be back to pre-pandemic business. I want to investigate 2022 statistics for the most popular airports in the United States.

I created an extensive table for the top 200 ranked airports in the United States. Data for this ranked list was sourced from the Bureau of Transportation Statistics₁.

DataCleaning.ipynb is where the data was cleaned and organized.

- The ranked list file that was read into python needed to be cleaned.
- I cleaned it using a LOCID table₂ and a supplementary LOCID table₃.
- Next, I added originating₄ and enplaned₅ passenger data for each airport.

- Then, I used rapidAPI₆ to get ICAO codes for each row.
- Next, I used a second API₇ and the ICAO values to get long, lat, etc. columns.
- Last, I found a blog₈ that compiled an updated table of US airports for 2020 and interesting statistics. I accessed the table as a text file and merged it onto my pandas DataFrame.
- I exported my dataframe to 2022Airports.csv.

Using this new datatable, I will be able to do both Q→Q and C→Q analysis. I want to compare the 2022 *Delays* proportion to the *Elevation* metric using scatterplot and possibly regression. Next, I want to compare all airports by their *Avg. Delay (Mins)* in a bar chart, organized by descending order, not rank.

Resources

- <https://www.bts.gov/topics/airlines-and-airports/airport-rankings-2022>
- https://www.faa.gov/data_research
- Requested unknown LOCID values for the remaining airports from chatGPT (extraLOCID.txt)
- OriginatingPassengers - U.S. Airports ranked by 2022 Originating Domestic Passengers
Source: Bureau of Transportation Statistics, Origin & Destination Survey
DB1B Ticket, Based on 10 Percent Ticket Sample
O&D numbers are not comparable to T-100 Market Enplanement numbers
- [Bureau of Transportation Statistics DB](#)
- <https://rapidapi.com/proground/api/aviation-reference-data>
- <https://airportdb.io/#>
- <https://www.stratosjets.com/blog/us-airport-rankings/>