Stage 1 – Raw Data

1) Download the monthly air traffic data files from the [Bureau of Transportation Statistics](https://www.transtats.bts.gov/databases.asp?Z1qr_VQ=E&Z1qr_Qr5p=N8vn6v10&f7owrp6_VQF=D) website. Instructions are in the file Download Path and Schedule for RawData files

2) Download the list EAS from the [Department of Transportation website](https://www.transportation.gov/office-policy/aviation-policy/subsidized-eas-report-communities-outside-alaska-september-2020-0). It’s dated to 2020, but the EAS subsidies are in effect for several years. Hence, the airports listed in the CSV have been EAS participants between 2019 and 2020. Many of the contracts run until the 2023. You will use this data later to designate EAS airports in your dataset of 354 airports. It can done with a join in Pandas.

3) Download [the FAA National Plan of Integrated Airport Systems](https://www.faa.gov/sites/faa.gov/files/2022-10/ARP-NPIAS-2023-Appendix-A.pdf). It contains the list of airports as FAA’s classification: L – Large Hub, M – Medium Hub, S – Small Hub, N – Nonhub

I attempted various ways to turn the pdf’s table into a CSV (or equivalent) but failed. So I simply went through the list state by state and entered the proper data point for each airport in the list of 354 airports.