

Imagine that you have been hired as a data analyst for a company that plans to disrupt the airline industry by building an underground high-speed passenger rail tunnel.

The company needs your help to decide which two major United States airports this tunnel should connect.

The distance between the airports must be within a specified range, and the airports must have a large volume of air travelers flying between them in both directions. The company believes that these air travelers can be persuaded to switch to high-speed rail because of frustratingly long flight delays.

You must write a SQL statement and analyze the result **to recommend which two airports this rail tunnel should connect**. Then you must create and upload a document describing the SQL statement you ran and the tunnel route you recommend.

The Task

Your job is to recommend which pair of United States airports should be connected with a high-speed passenger rail tunnel. The company you work for has given you the following strict requirements:

These two airports must:

- Be between **300** and **400** miles apart
- Average at least **5,000** (five thousand) flights per year between them, *in each direction*

Among the pairs of airports that meet these requirements, you must identify the one pair that has the **largest total number of seats on the planes that flew between them**.

The company is also interested to know the **average arrival delay** for flights between these two airports, because they believe that routes with a history of delayed arrivals will make it easier to persuade air travelers to switch to high-speed rail.

For the pair of airports you recommend, you must provide the following details:

- The **three-letter codes** identifying both airports
- The **average flight distance** in miles for flights between the airports, in each direction
- The **average number of flights per year between** the airports, in each direction
- The **average annual passenger capacity** (average yearly total number of seats on the planes) for flights between the airports, in each direction
- The **average arrival delay** for flights between the airports, in each direction

