

Project

1. Domain Description and Data Source:

Domain description:

Our domain for this project is an airline tracking system. The database will contain tables describing data for **airports**, **airlines**, **flights**, air **vehicles**, and arrival/departure **flight status** (with foreign key to flight, status, and date columns). There will be an additional relation connecting the Airport relation to Airline, because an airline can service many airports and an airport can have many airlines that visit it. Flight information will include the flight #, airline company, current destination, and gate #. Because frequent updates to this database are expected as planes arrive/depart, run late, change gates, etc., the database will be implemented with the goal of maximum simplicity and efficiency for tuple modifications either manually or automatically. For this reason, among others, we are going to make the **FlightStatus** relation the primary “dynamic” relation, where data is expected to be modified frequently; the others should only need to be updated occasionally. NOTE: this is not a flight-booking system, but rather a general-purpose flight-tracking system.

Questions:

1. What flights are departing today?
2. What flights are arriving in the next 3 days?
3. What is the distance to AMS from PDX?
4. What is the duration of the flight to St. Paul from San Francisco?
5. What flights were delayed today?
6. What flights are arriving from New York in next 2 hours?
7. What are all the different existing airline companies?
8. What flights are canceled this week?
9. What flights are departing before noon tomorrow?
10. What is the flight with the longest duration?
11. What airports have a flight scheduled for departure at 10:55 pm today?
12. What is the average travel distance for flights for flights departing from Columbus, OH?
13. What Alaska Airlines flights are departing this week?
14. What is the vehicle id of a particular flight?
15. What flights are departing for Chicago in the next hour?
16. What are the nearest and furthest departure/arrival times for all flights?
17. What flights are scheduled for bag pickup at carousel 3?
18. What helicopters are departing today?
19. How many flights visit PDX?
20. What is the capacity of a flight 788 from Frontier Airline that arrives in next 3 days?

Database:

As our data source, we will be using <http://openflights.org/data.html>, which has CSV files for more than 8K Airports, 20K Airlines lines, and 68K Flights. We will process CSV files to generate SQL queries (probably using python). While studying this data, we also decided that we wanted to add another relation AirlineToAirport, using the same code. We also noticed that the current dataset doesn't include data pertaining to statuses, departing times, or gates, so we decided to randomly generate data (probably also using python). We will then import the SQL file into our PostgreSQL (phpPgAdmin) account provided in class.