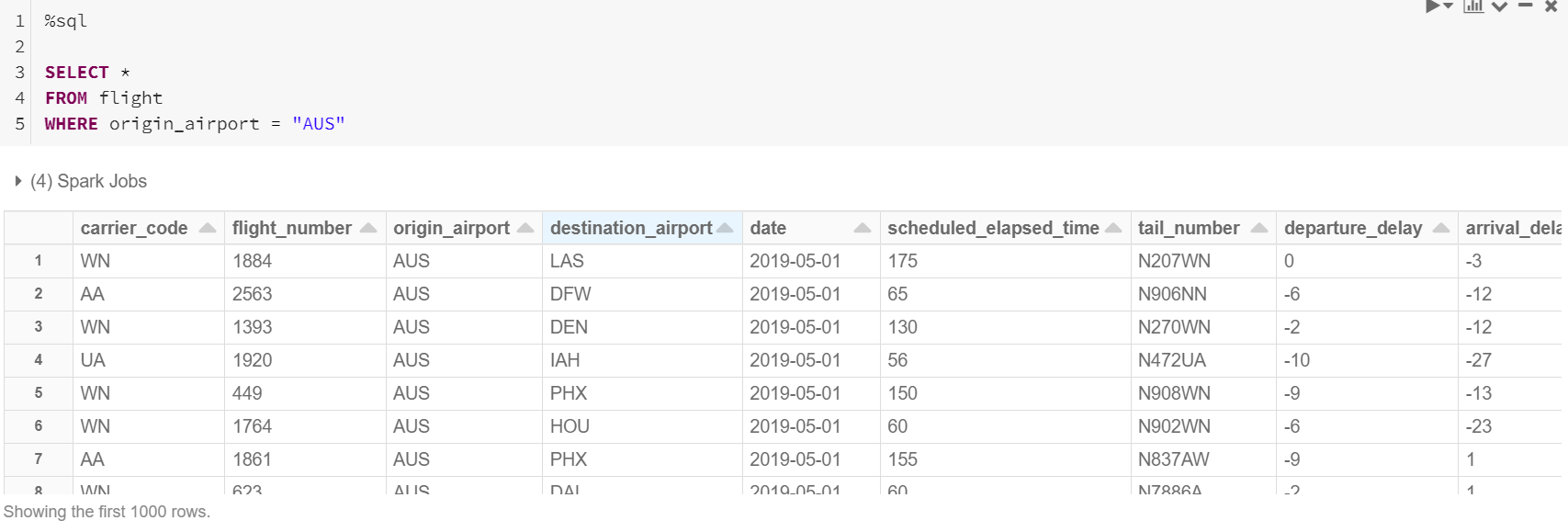
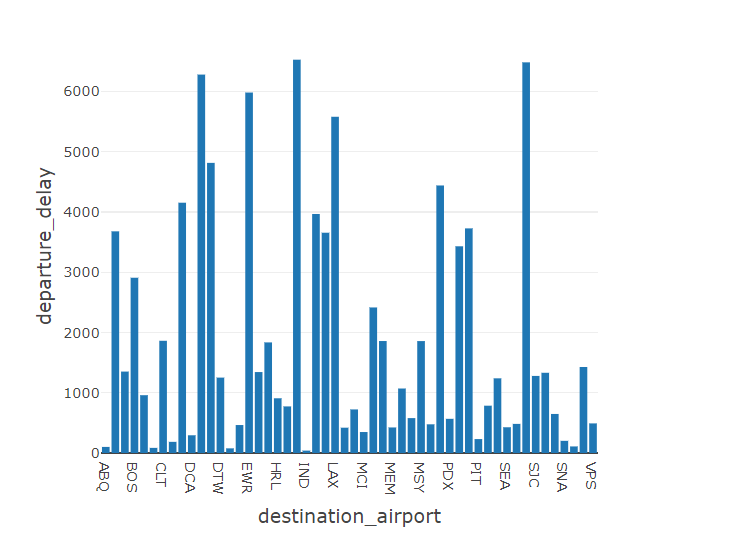
Out of my curiosity to know about the behavior of domestic flight delays, I decided to use the dataset on Kaggle for this assignment. Data is taken of the flights departing from Austin to different US airports for the month of May 2019.This CSV was read into Spark and analyzed further to generate meaningful information.

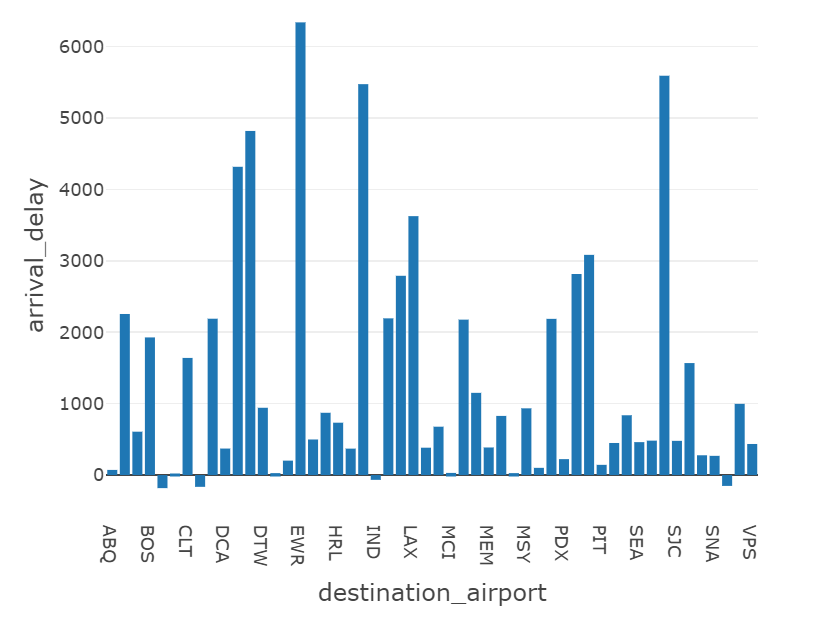
Firstly, this CSV was filtered to only show flights departing from AUS.



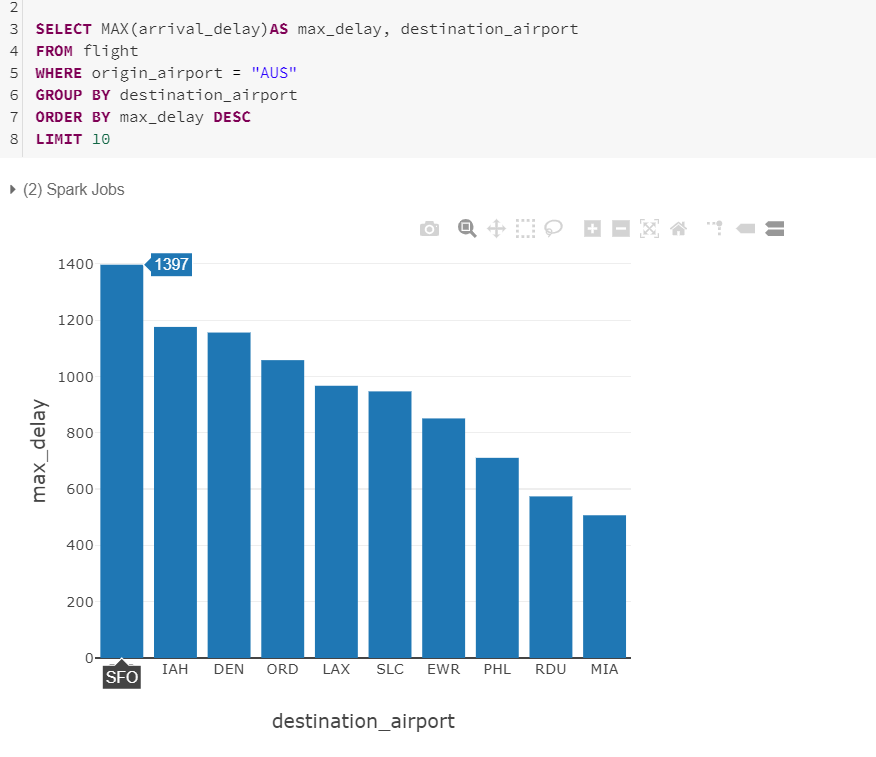
Next, the data was processed to show Departure Delay associated with different airports. Here we can see that flights going to DEN, EWR, IAH, LAX, ORD and SFO all experience high departure delays. Obvious reason is, since all these airports are big and major airports. So flights coming into these airports are sometimes held at their departure airports to ease the incoming flight traffic.

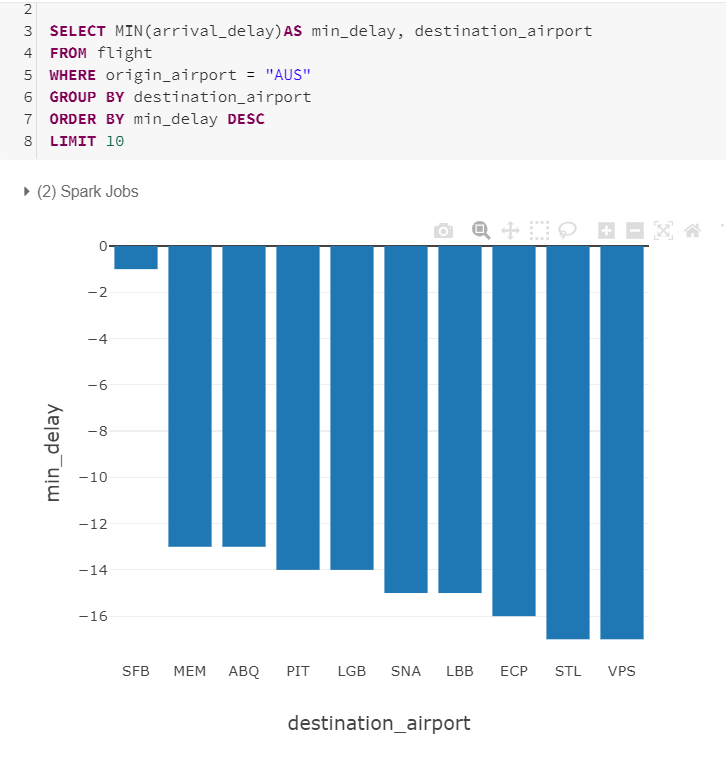


Then the data was further processed to show Arrival Delay associated with different airports. Here again we can see that airports like DEN, DFW, EWR and SFO all experience high arrival delays. Since aircrafts don’t normally get landing clearance, they get delayed landing on the airport since they keep going in rounds waiting to get their turn.



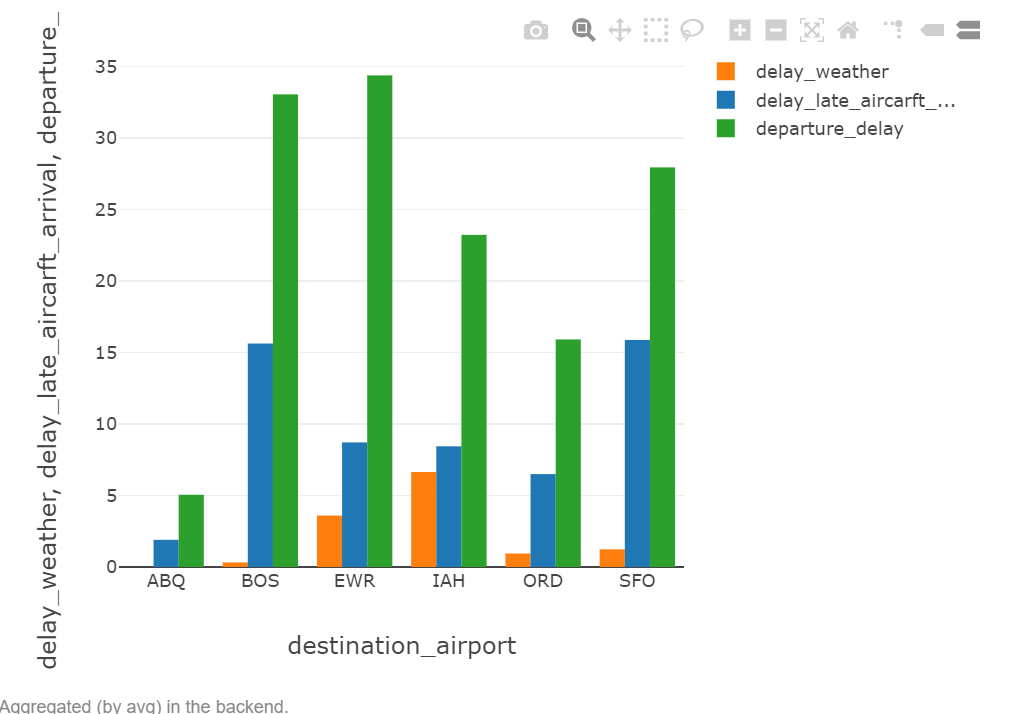
Next, I analyzed the data to show max and min delays associated with flights going out of AUS. It has been observed that max delay associated with the flight going to SFO. Which makes sense because airport has high incoming traffic and that is why their arrival is delayed a lot. It is also seen that the flights going to VPS experiences least delay and in fact it normally arrives before time compared to other airports.





I further broke down the delays caused between AUS and other airports. It is found that departure delays at BOS has been mostly because of late coming aircraft from its previous journey. And that is because flights are coming in late into BOS due to heavy traffic. Similar story is at SFO where the max departure delay is due to late incoming aircraft due to the busy nature of the airport.

As against, departure delays at IAH is mostly due to Weather delays and that is because of possible thunderstorms in the month of May in and around Houston.



Lastly, the arrivals into IAH are mostly delayed due to the stormy weather in the month of May around the area. Flights coming into ABQ were hardly delayed due to weather. In fact they experienced least arrival delay due to a very small airport

