# Final Exam

Zihua Rong
ID:1505002
Econ 294A

Part a)

First of all, we consider the relationship between time delayed, whether the flight is canceled and the weather. I choose four important variables, temperature, humidity, wind speed and pressure. We use the code lm and glm to regress canceled and delay on these variables.

The results shows that temperature and wind speed have significant positive effect on time delayed.
Besides, humidity and pressure have significant negative effect on time delayed.
We can view the graphs which show the relationship between time delayed and the variables about weather.
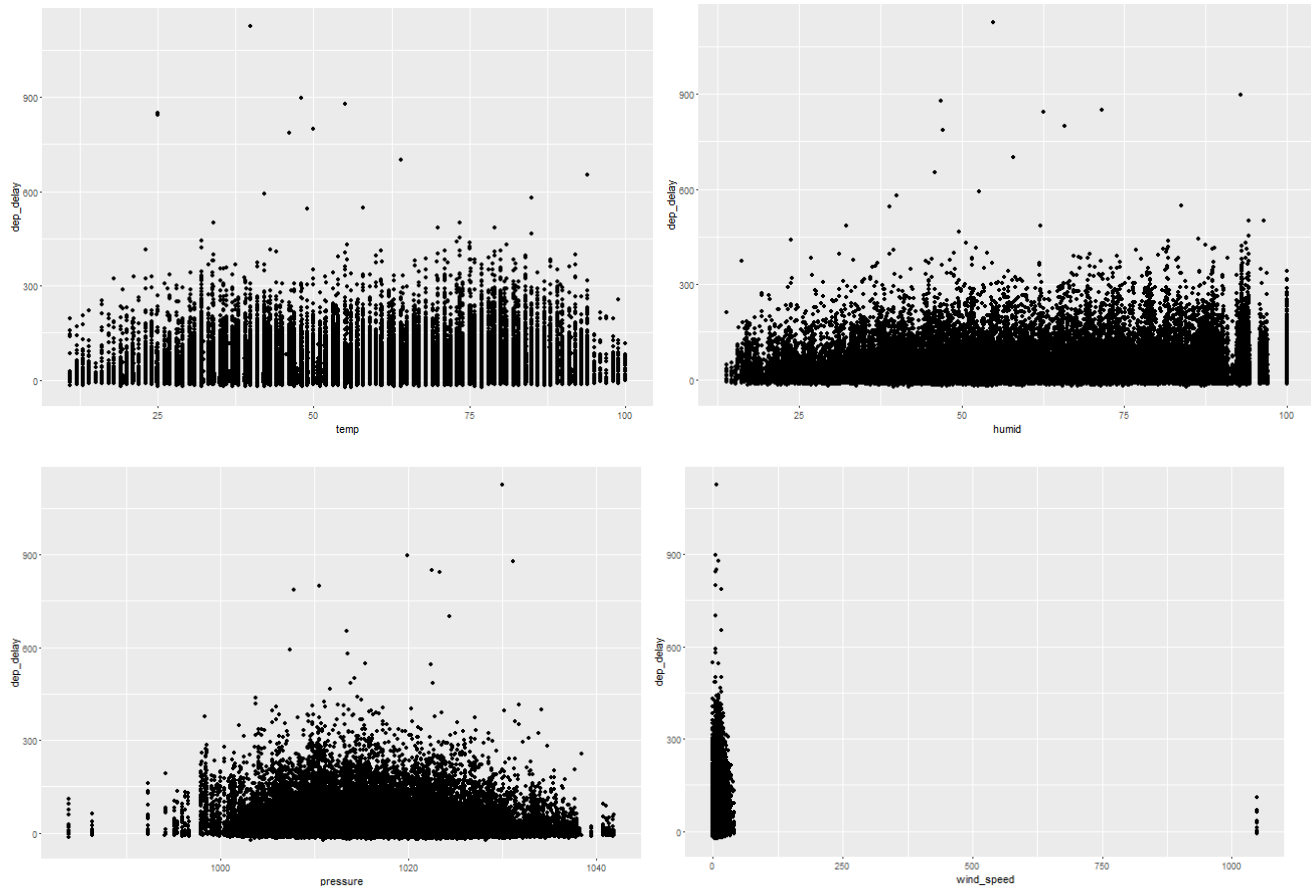
From the graphs, we may find that pressure is negative related to the time-delayed, which is same to the result of regression. However, we can't find strong relation between temperature and time delayed, neither between humidity and time delayed. So I think it's just significant in statistic but we can't find the relation in the graphs.

Part b)

This time, we try to search the relationship between time-delayed, cancel and time. I choose month, day, weekdays and time of day as dependent variables in the lm model and the glm model.

The result shows that Thursday has significant positive effect on the chance of canceled. On the other hand, the chance of cancel decreases by months, so less flights are canceled in the second half of the year.

We can check the results in the graphs:
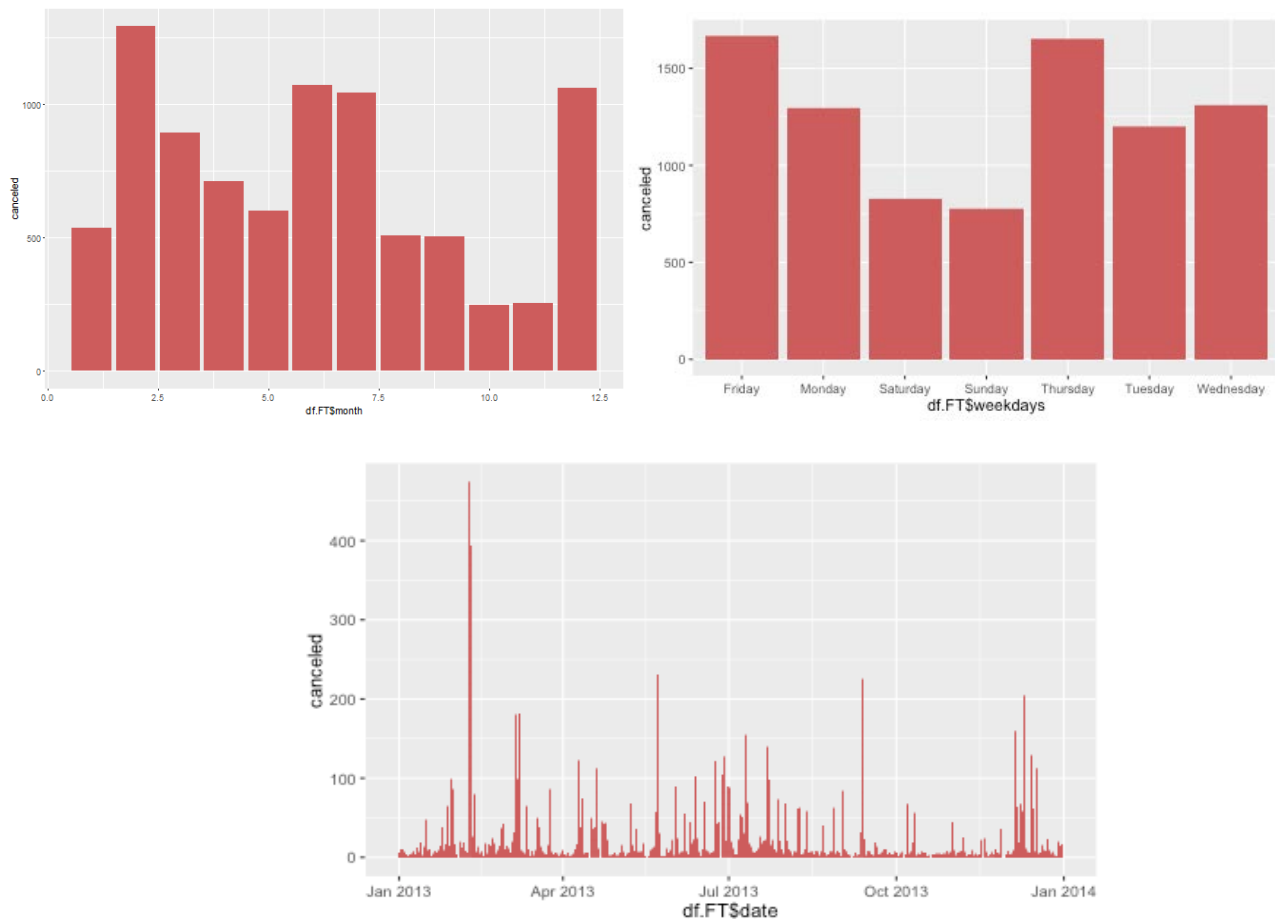
We find that Thursday and Friday have most canceled planes and Saturday and Sunday have least flights canceled. It's similar to the result of the regressions.

Besides, from the graphs we may find February has most flights canceled and October and November have least flights canceled. It also meets the result of the regression that less flights are canceled in the second half of the year.

Part c)

This time, we consider the effect of destination. The results tell us, flights to BHM, CAE, DSM, OKC, RIC, TUL and TYS have significant positive effects on time delayed.

Then we can see the results in the graphs, flights to several destinations have more flights canceled, I think the reason is that those airports are busier so there're more flights canceled.
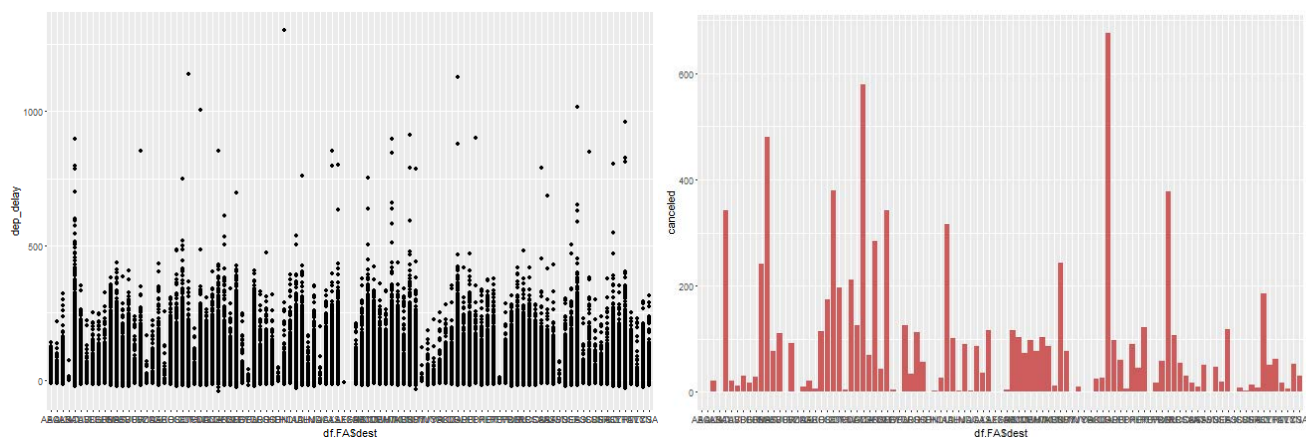
Part d)

This time we consider the effects of the characters of the planes. Among the characters of the planes, I think the number of seats and the number of engines are most important.

The results of the regressions are, the number of seats have significant negative effects on both time delayed and the chance that the flights are canceled. Besides, the number of engines have significant positive effects on both time delayed and the chance that the flights are canceled. Then we can see the graphs.
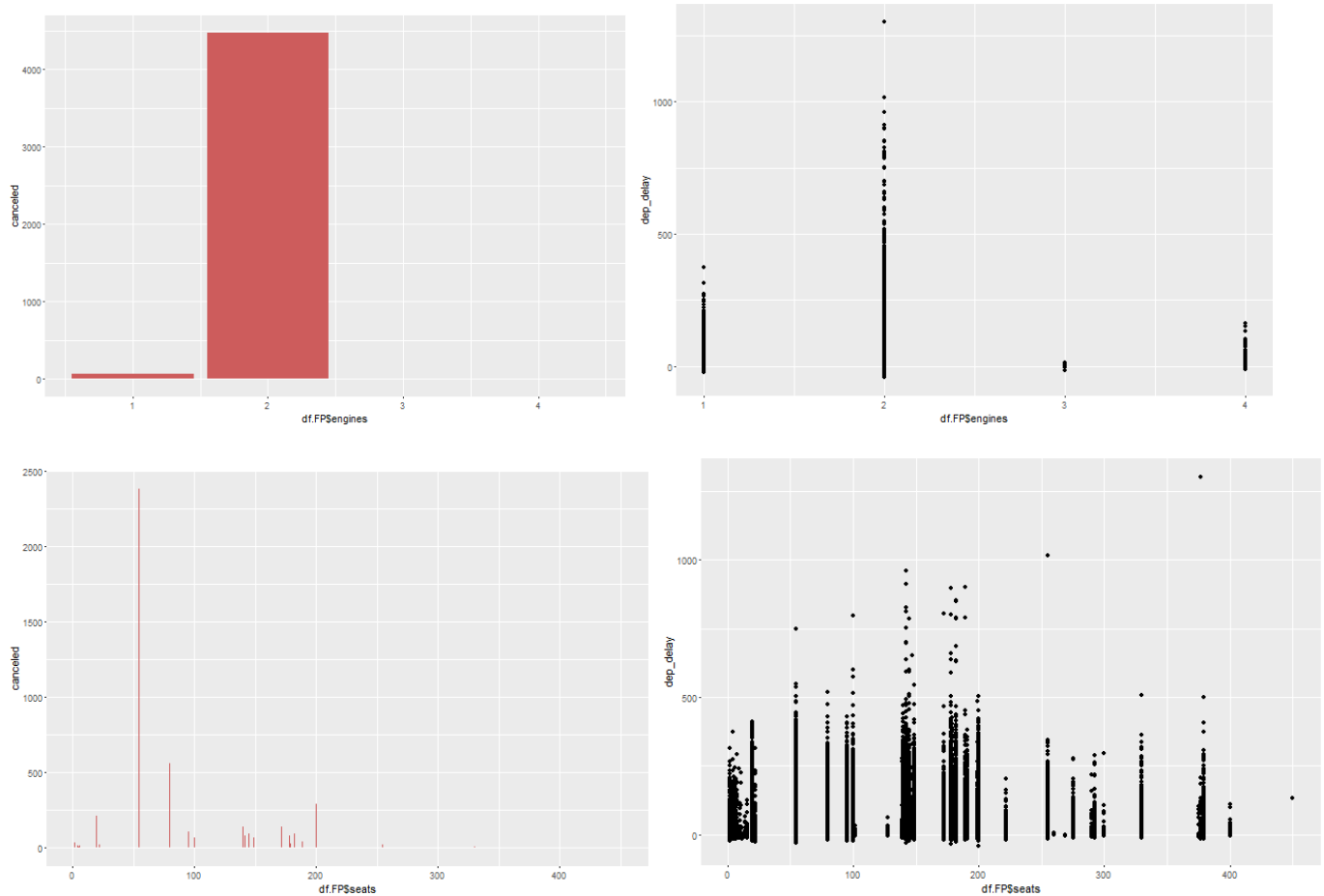
From the graphs above, we may find that when the number of seats increases, the delayed time will be shorter.

Besides, the planes with 2 engines have longer time delayed and more chance to delay than the planes with only one engine.

Part e)

In conclusion, temperature and wind speed have significant positive effect on time delayed, and humidity and pressure have significant negative effect on time delayed. Besides, Thursday and Friday have most canceled planes. Saturday and Sunday have least flights canceled. February has most flights canceled and October and November have least flights canceled. Next, flights to BHM, CAE, DSM, OKC, RIC, TUL and TYS have longest time delayed. Finally, when the number of seats decreases, the delayed time will be longer. And the planes with 2 engines may have the longest time delayed.