BST 693/CPH 608 Assignment 2

1) Clone repository

2) Create a new branch with edits to the R file.

3) Using the flights dataset, look at the number of cancelled flights per day. Is there a pattern? Is the proportion of cancelled flights related to the average delay?
4) Which carrier has the worst delays? Challenge: can you disentangle the effects of bad airports versus bad carriers? Why/why not? (hint: think about

```
flights %>%
        group_by(carrier,dest) %>%
        summarize(n()). )
```

5) For each plane, count the number of flights before the first delay of greater than 1 hour.
6) Which plane (tailnum) has the worst on-time record?
7) What time of day should you fly if you want to avoid delays as much as possible?
8) For each destination, compute the total minutes of delay. For each flight, compute the proportion of the total delay for its destination.
9) Look at each destination. Can you find flights that are suspiciously fast? (That is, flights that represent a potential data entry error.) Compute the air time of a flight relative to the shortest flight to that destination. which flights were most delayed in the air?

10) Commit changes to your branch

11) When finished completing assignment, merge your branch into master

12) Push your changes