Week 2 Statistical computing

MS 276

September, 2017

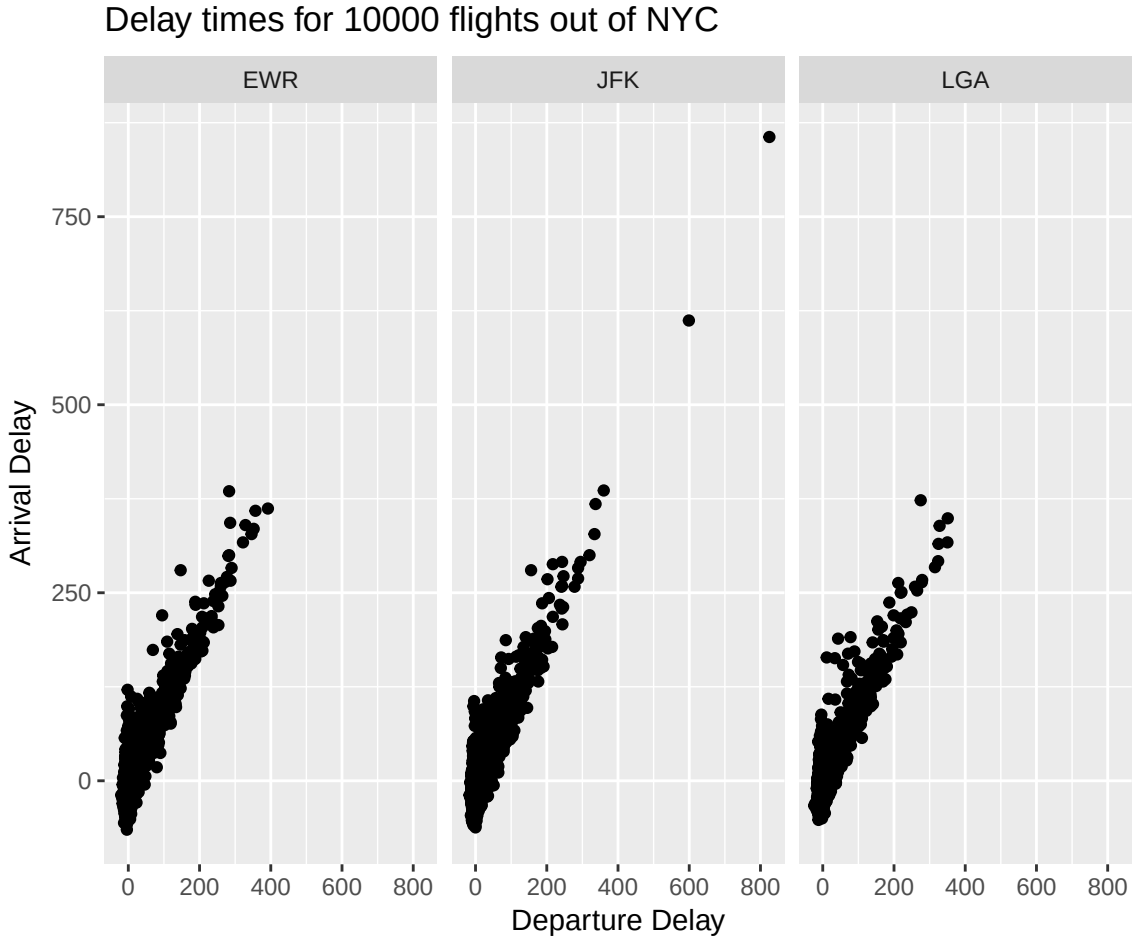
For each plot

1. Write the code that, intuitively, should get you most of the way towards a final plot
2. Identify any new aspects of visualization that you'll need to look up

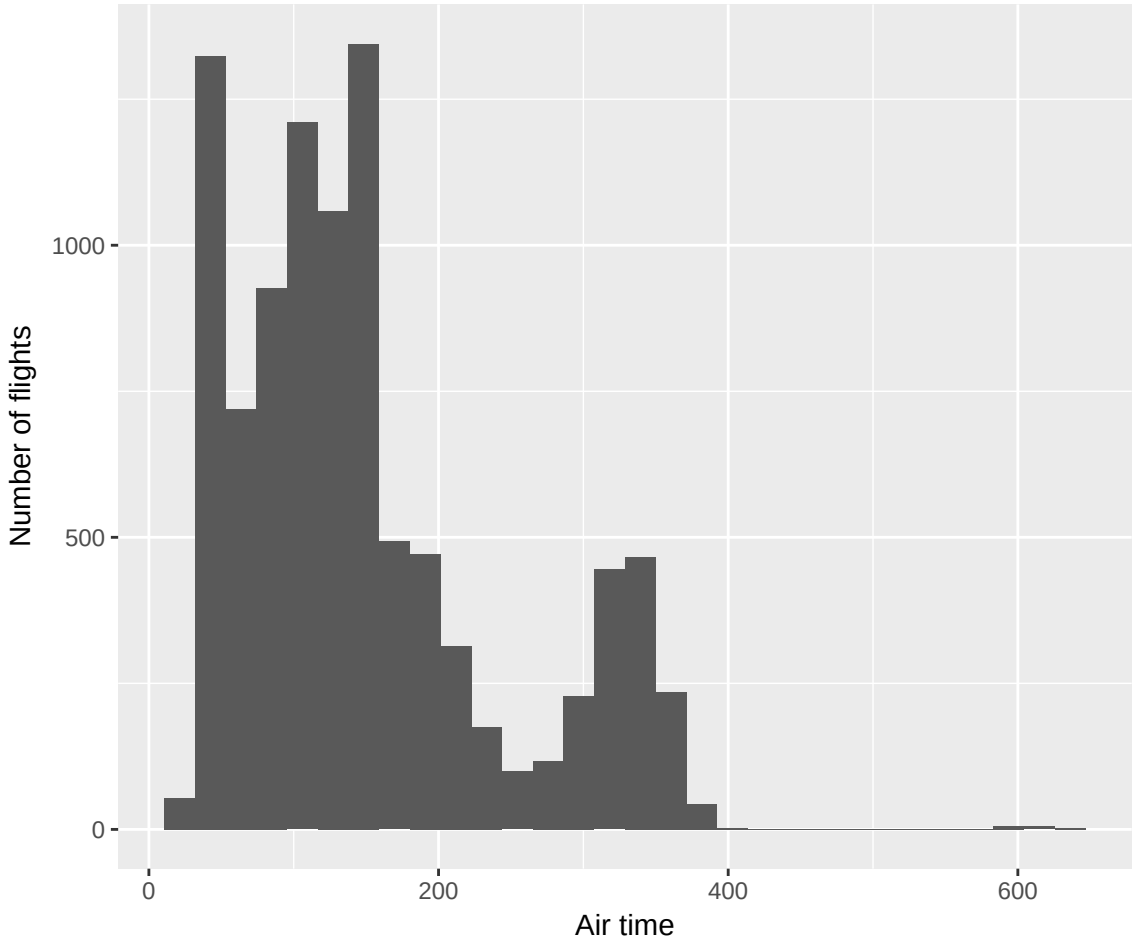
Use the following code to get started

```
set.seed(0)
library(nycflights13)
library(tidyverse)
flights <- flights %>% sample_n(10000)
```

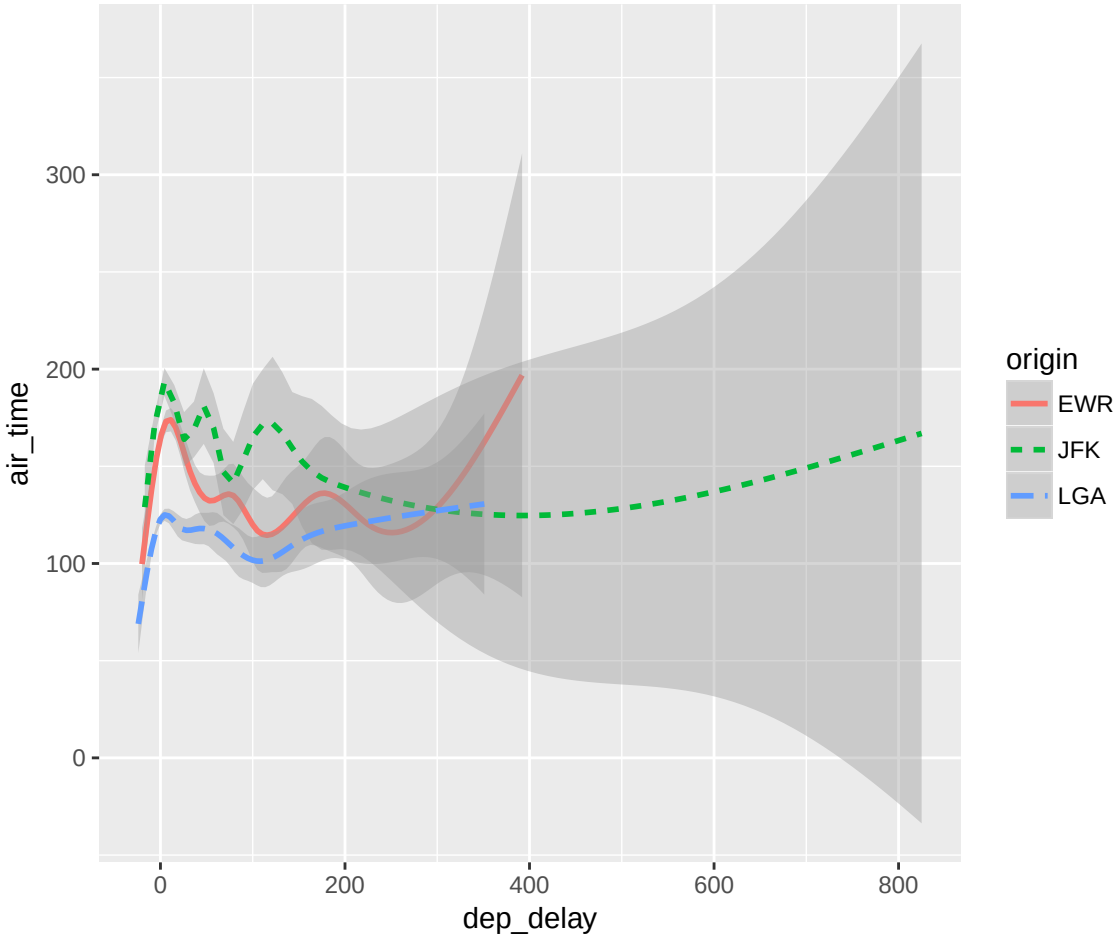
1.


Delay times for 10000 flights out of NYC
EWR
JFK
LGA
Arrival Delay
750
500
250
0
0
200
400
600
800
Departure Delay

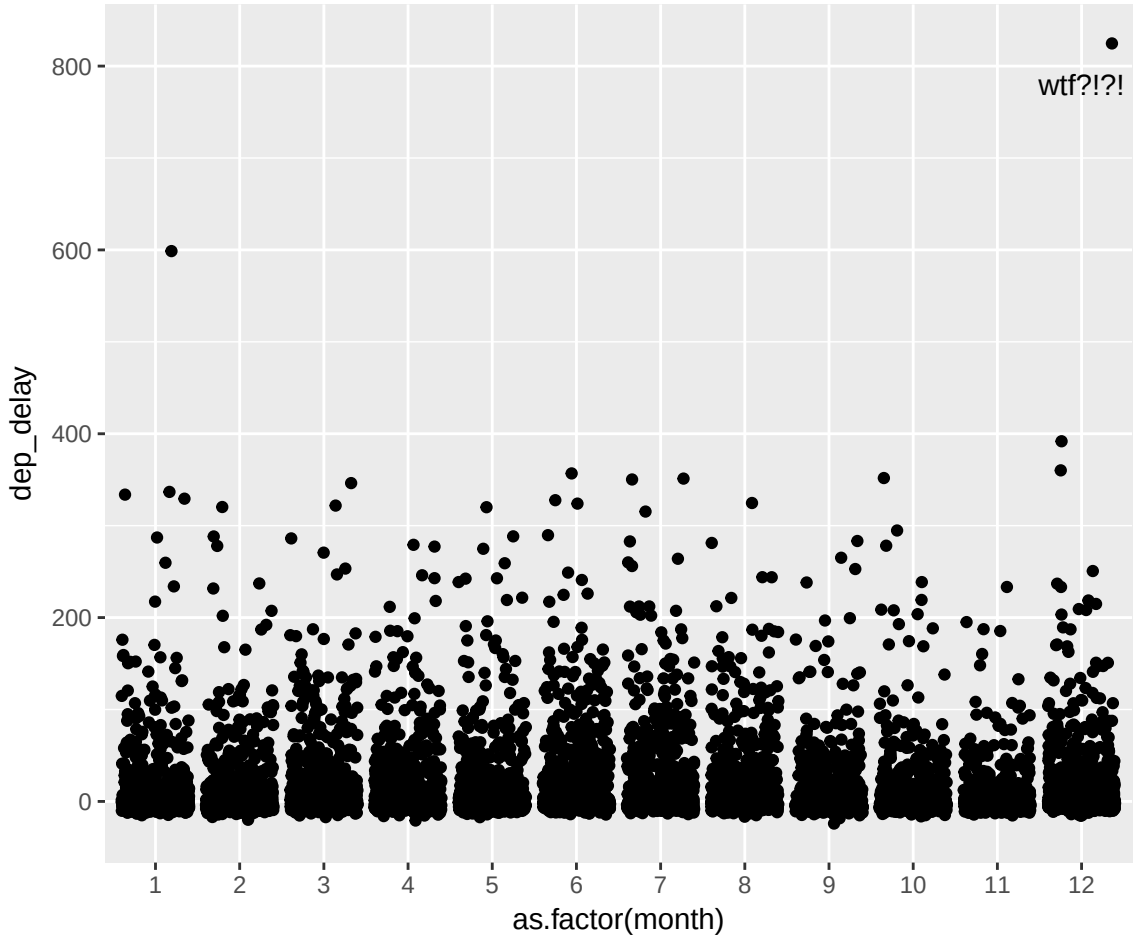
2.



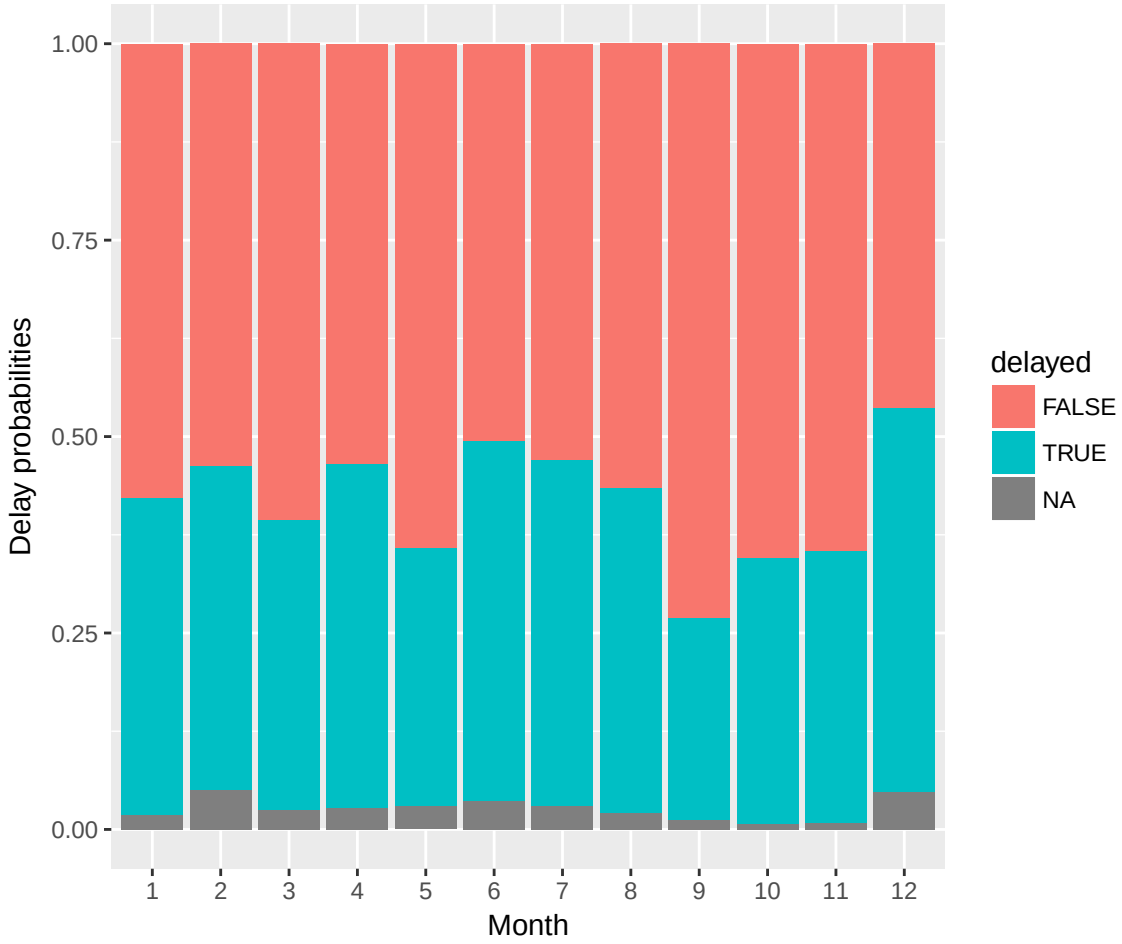
3.


air_time
300
200
100
0
0
200
400
600
800
dep_delay
origin
EWR
JFK
LGA

4.


wtf?!?!
800
600
400
200
0
dep_delay
1
2
3
4
5
6
7
8
9
10
11
12
as.factor(month)

5.


1.00
0.75
0.50
0.25
0.00
Delay probabilities
1
2
3
4
5
6
7
8
9
10
11
12
Month
delayed
FALSE
TRUE
NA

6. Using the `flights` data frame, construct code and a rough outline of a plot that you think would be interesting to look at.

```
head(flights) %>% print.data.frame()
```

```
##   year month day dep_time sched_dep_time dep_delay arr_time sched_arr_time
## 1 2013     8  24      547            545         2      816            815
## 2 2013    12   7     1447           1445         2     1842           1812
## 3 2013     2  17      809            810        -1     1024           1042
## 4 2013     4  30     1354           1359        -5     1447           1519
## 5 2013     8  28      905            910        -5     1203           1215
## 6 2013    11  14      555            600        -5      723            730
##   arr_delay carrier flight tailnum origin dest air_time distance hour
## 1         1      UA    390  N538UA    LGA  IAH      187     1416    5
## 2        30      UA   1554  N14731    EWR  SFO      396     2565   14
## 3       -18      DL    269  N302NB    JFK  ATL      104      760    8
## 4       -32      9E   3354  N931XJ    JFK  BOS       33      187   13
## 5       -12      AA    713  N3HFAA    LGA  DFW      178     1389    9
## 6        -7      AA    301  N555AA    LGA  ORD      121      733    6
##   minute           time_hour
## 1     45 2013-08-24 05:00:00
## 2     45 2013-12-07 14:00:00
## 3     10 2013-02-17 08:00:00
## 4     59 2013-04-30 13:00:00
## 5     10 2013-08-28 09:00:00
## 6      0 2013-11-14 06:00:00
```