ECO395M Homework1: Evan Aldrich

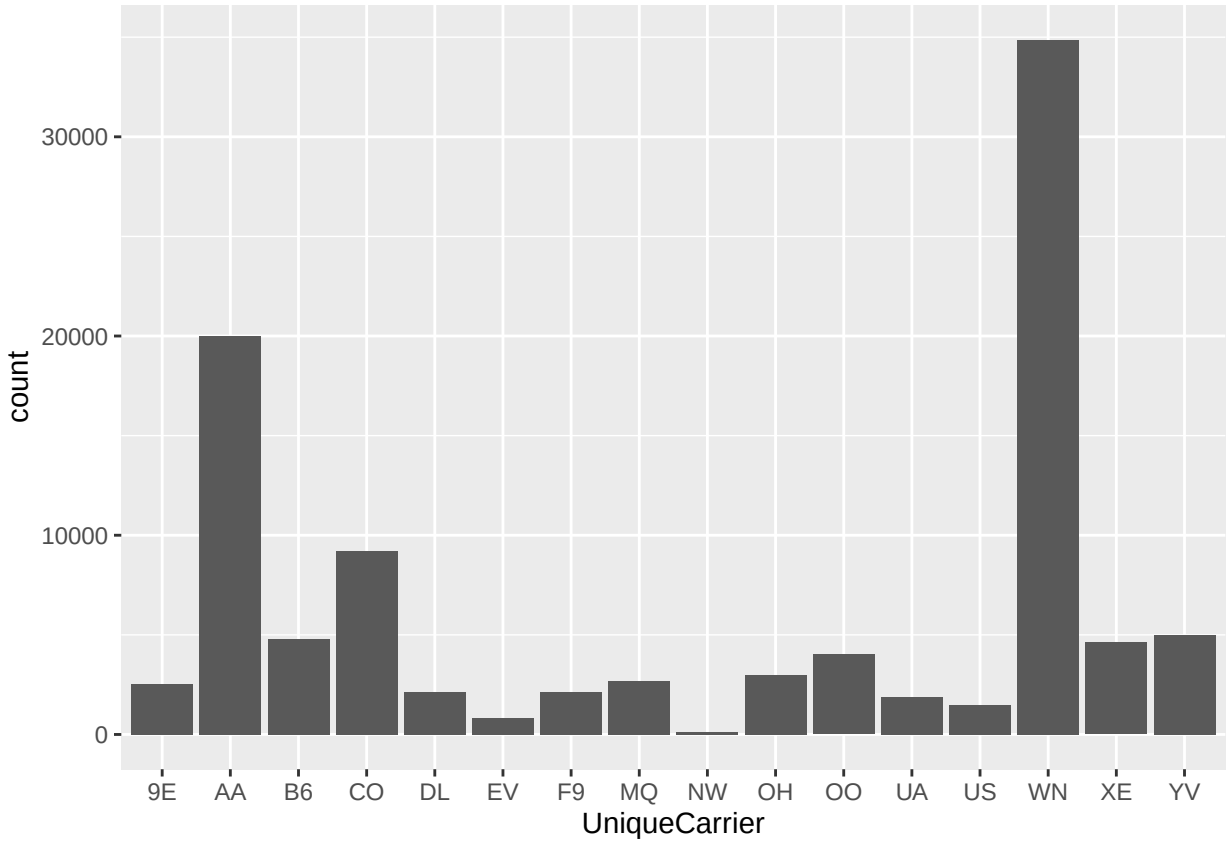
Evan Aldrich

2024-01-20

Question 1: Data visualization: flights at ABIA

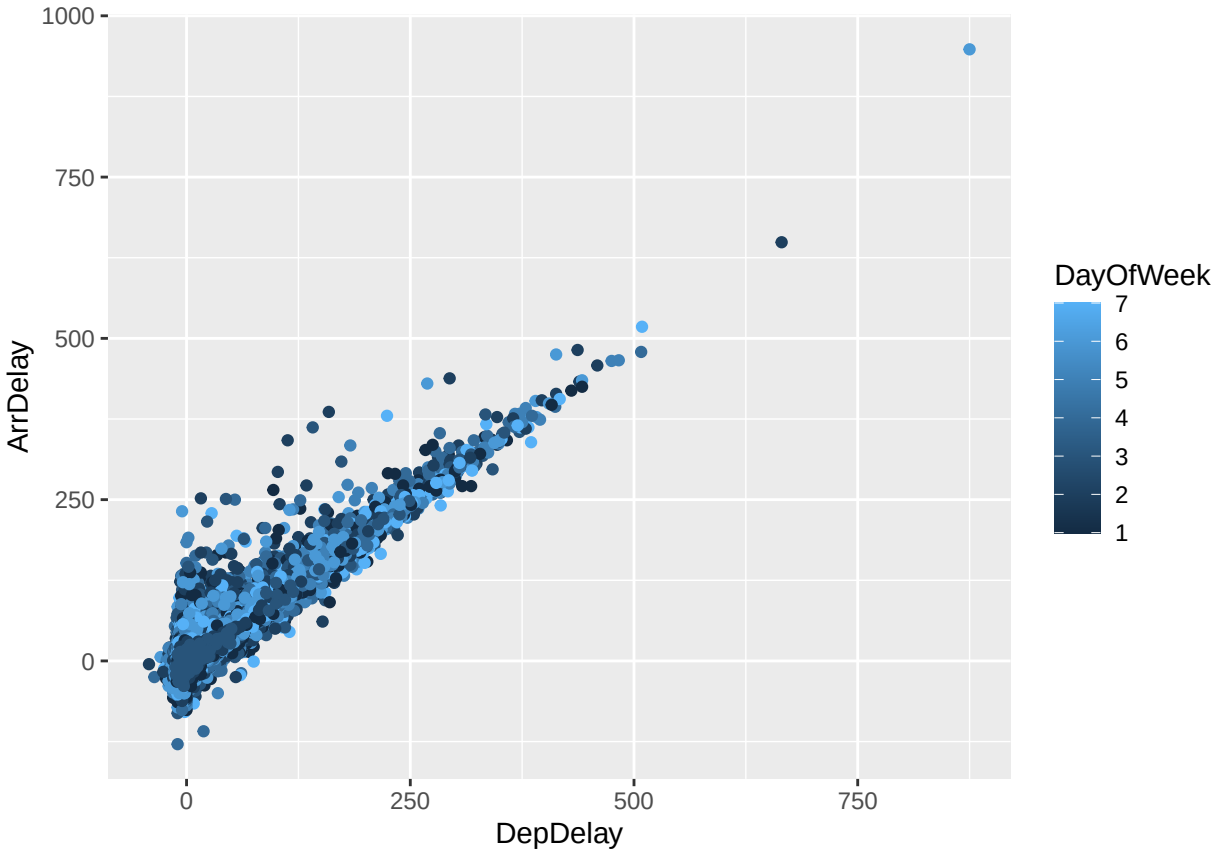
Part A Make a 2x2 table of conditional probabilities, conditional on the levels of the david.bowie variable. The bottom right entry in your 2x2 table should display P(plays Daft Punk | plays David Bowie).

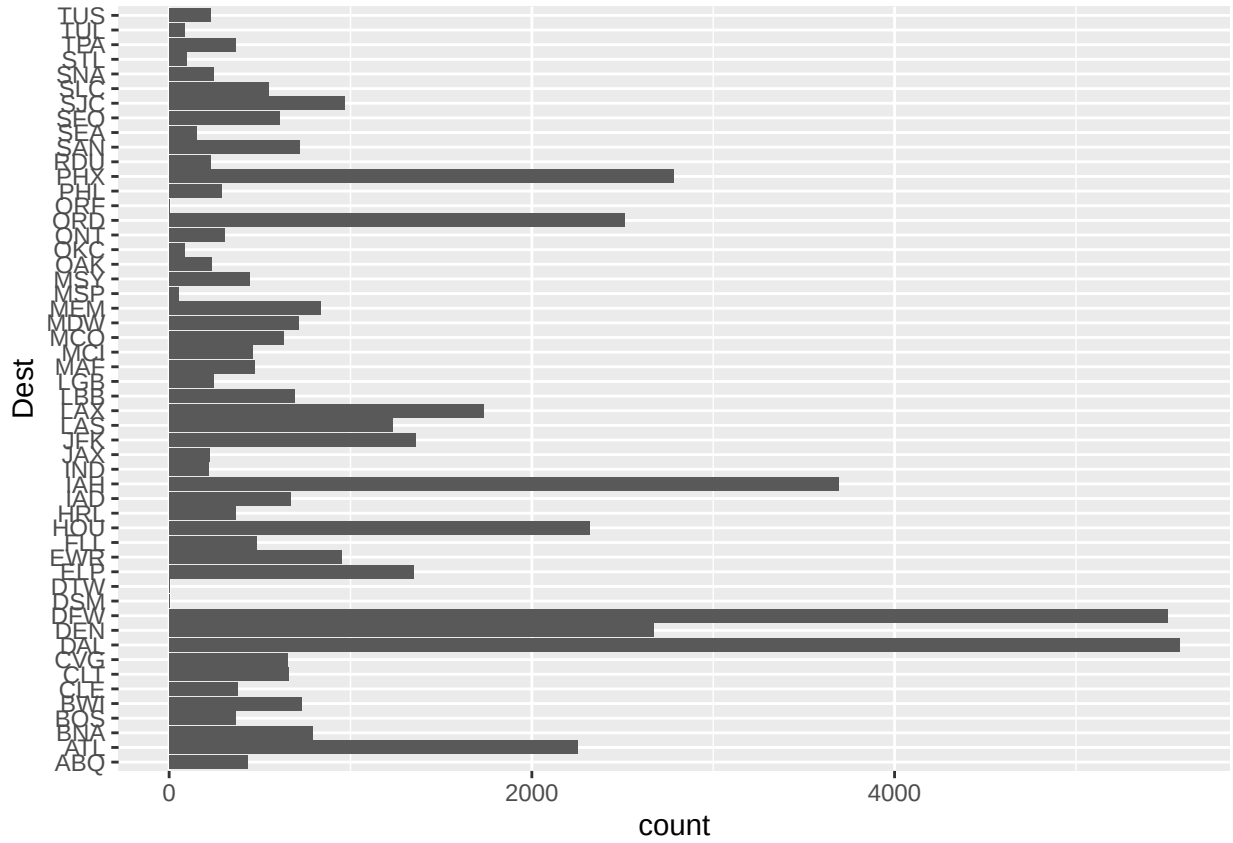
```
## Cancelled
##     0     1
## 97840  1420
```




```
## Warning in geom_point(aes(x = DepDelay, y = ArrDelay, color = DayOfWeek), :
## Ignoring unknown parameters: ‘width‘
```

```
## Warning: Removed 1601 rows containing missing values (‘geom_point()‘).
```


1000
750
500
250
0
ArrDelay
0
250
500
750
DepDelay
DayOfWeek
7
6
5
4
3
2
1



Question 2: Wrangling the Olympics

Part A What is the 95th percentile of heights for female competitors across all Athletics events (i.e., track and field)?

```
## 95%
## 183
```

Part B Which single women's event had the greatest variability in competitor's heights across the entire history of the Olympics, as measured by the standard deviation?

Part C How has the average age of Olympic swimmers changed over time? Does the trend look different for male swimmers relative to female swimmers? Create a data frame that can allow you to visualize these trends over time, then plot the data with a line graph with separate lines for male and female competitors. Give the plot an informative caption answering the two questions just posed.

Question 3: K-nearest neighbors: cars