Team

McKenna Goffinet & Anyesha Ray

2/5/2020

1)

```r
dim(choco)
```

```
## [1] 1852    9
```

```r
head(choco)
```
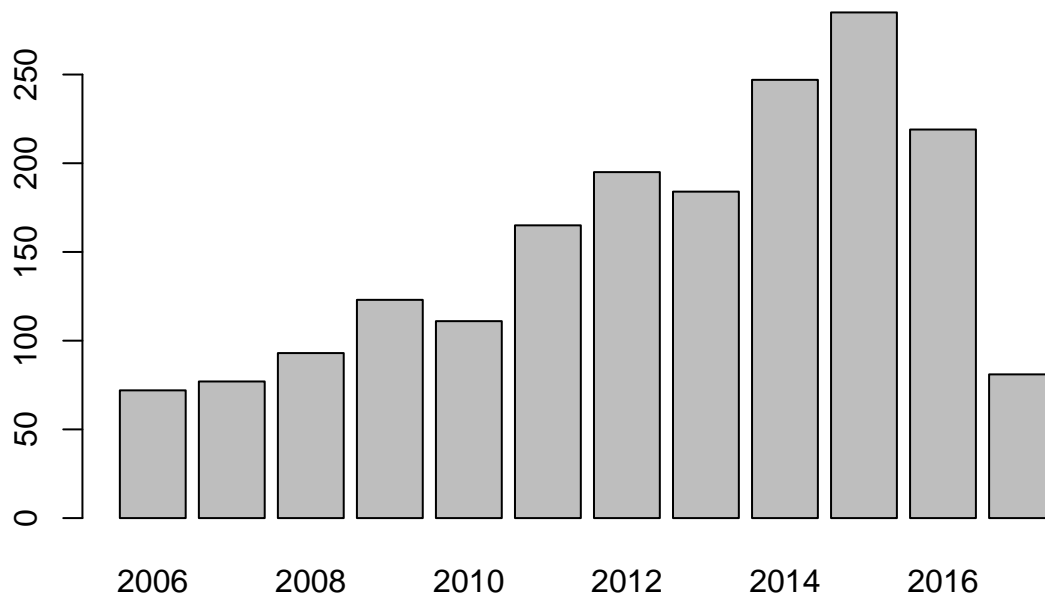
```
##      Company Specific.Bean.Origin  REF Review.Date Cocoa.Pct
## 1 A. Morin          Agua Grande 1876        2016        63
## 2 A. Morin                Kpime 1676        2015        70
## 3 A. Morin               Panama 1011        2013        70
## 4 A. Morin            Madagascar 1011        2013        70
## 5 A. Morin                Brazil 1011        2013        70
## 6 A. Morin             Equateur 1011        2013        70
##   Company.Location Rating Bean.Type Broad.Bean.Origin
## 1           France   3.75                    Sao Tome
## 2           France   2.75                        Togo
## 3           France   2.75                      Panama
## 4           France   3.00   Criollo         Madagascar
## 5           France   3.25                      Brazil
## 6           France   3.75                     Ecuador
```

1852 chocolate bars

2)

```r
barplot(table(choco$Review.Date))
```

The number of ratings per year has been generally increasing over the years, though inconsistently. The number of ratings in a single year peaked in 2015, followed by a sudden drop in 2017. The period of most increase occurred during the years of 2010 to 2015.
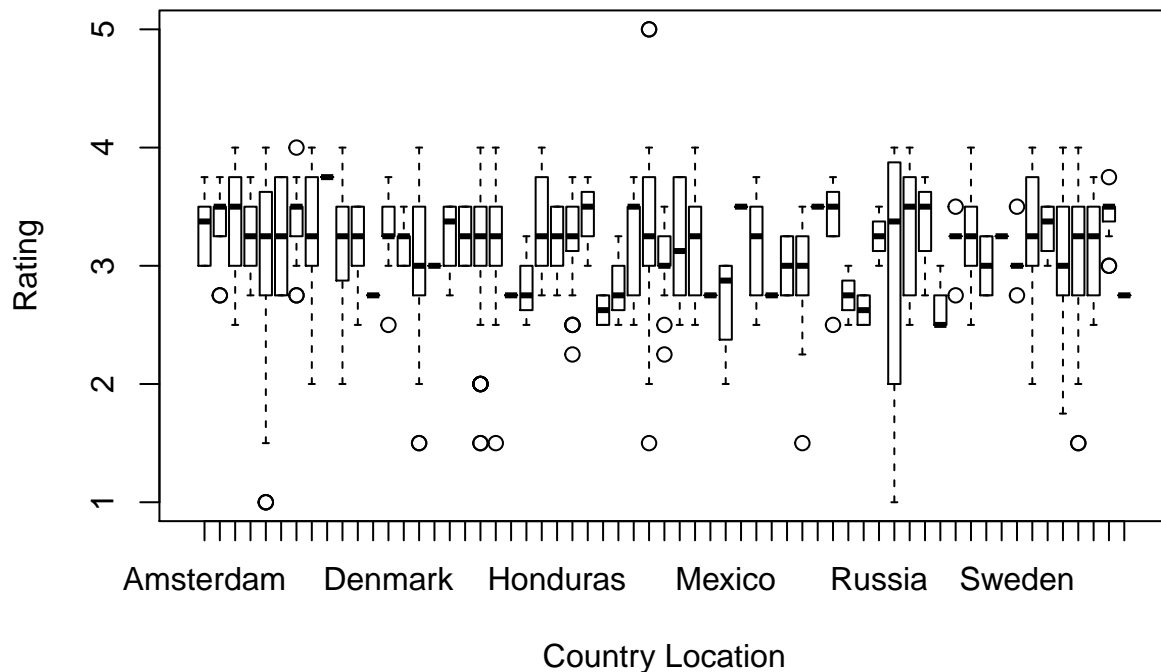
3)

```r
y <- count(choco$Company)
z <- y[order(-y$freq),]
z[1:3,]
```

```
##          x freq
## 370   Soma   47
## 44  Bonnat   27
## 161 Fresco   26
```

```r
boxplot(choco$Rating~choco$Company.Location, main="Country Location vs. Rating", xlab="Country Location"
```

Country Location vs. Rating

Location does not seem to have much of an impact on rating. They seem pretty consistedn in the 2 to 4 rating range with a few outliers.

4) Our Question: What are the most popular countries for producing beans? What are the least 3?

5) Most popular:

```
c <- count(choco$Broad.Bean.Origin)
d <- c[order(-c$freq),]
d[1:3,]
```

```
##            x freq
## 92 Venezuela  217
## 23   Ecuador  198
## 56      Peru  171
```

Least popular:

```
e <- count(choco$Broad.Bean.Origin)
f <- e[order(e$freq),]
f[1:10,]
```

```
##                           x freq
## 2  Africa, Carribean, C. Am.    1
## 7                      Burma    1
## 8                   Cameroon    1
## 10     Carribean(DR/Jam/Tri)    1
## 15          Cost Rica, Ven    1
## 18     Dom. Rep., Madagascar    1
```

```
## 20     Dominican Rep., Bali    1
## 22        DR, Ecuador, Peru     1
## 24     Ecuador, Costa Rica     1
## 25      Ecuador, Mad., PNG     1
```

6) Since we did this in lab, the workflow was smooth. We both worked on question 1 to 3 on Mckenna's laptop and then pushed it. We worked on the last 3 questions on my laptop. It was pretty balanced work load but it is slightly difficult to push and pull items from Github at least at this stage. Other wise work flow was smooth. We both did the individual part and then collaborated on the team questions on both laptops.