

# Glass Dataset Exploration

*Naren Vijayakumar*

*November 19, 2016*

```
library(mlbench)
library(dplyr)
library(e1071)
library(ggplot2)
```

## Glass Dataset

```
data(Glass)
str(Glass)
```

```
## 'data.frame': 214 obs. of 10 variables:
## $ RI : num 1.52 1.52 1.52 1.52 1.52 ...
## $ Na : num 13.6 13.9 13.5 13.2 13.3 ...
## $ Mg : num 4.49 3.6 3.55 3.69 3.62 3.61 3.6 3.61 3.58 3.6 ...
## $ Al : num 1.1 1.36 1.54 1.29 1.24 1.62 1.14 1.05 1.37 1.36 ...
## $ Si : num 71.8 72.7 73 72.6 73.1 ...
## $ K : num 0.06 0.48 0.39 0.57 0.55 0.64 0.58 0.57 0.56 0.57 ...
## $ Ca : num 8.75 7.83 7.78 8.22 8.07 8.07 8.17 8.24 8.3 8.4 ...
## $ Ba : num 0 0 0 0 0 0 0 0 0 0 ...
## $ Fe : num 0 0 0 0 0 0.26 0 0 0 0.11 ...
## $ Type: Factor w/ 6 levels "1","2","3","5",...: 1 1 1 1 1 1 1 1 1 1 ...
```

There are 9 predictors and 1 outcome. Separating the predictor and outcome variables, we get.

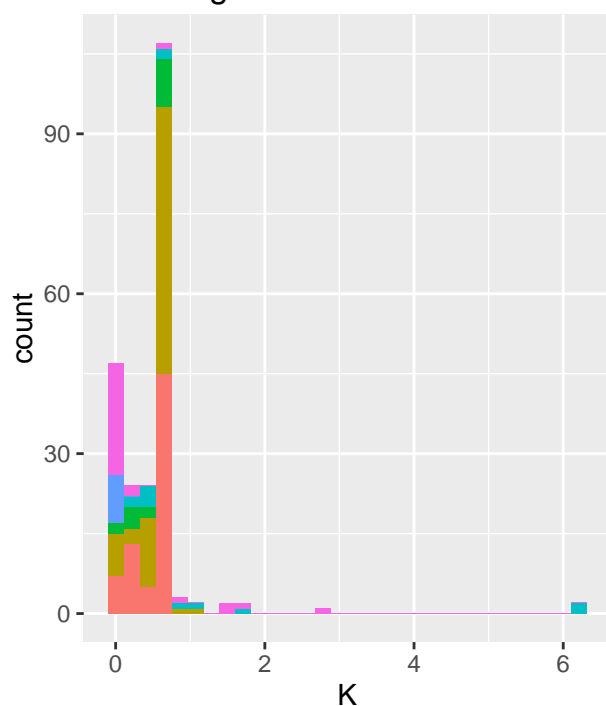
```
glassPredictors <- Glass %>% select(-Type)
glassTypes <- Glass %>% select(Type)
```

## Skewness

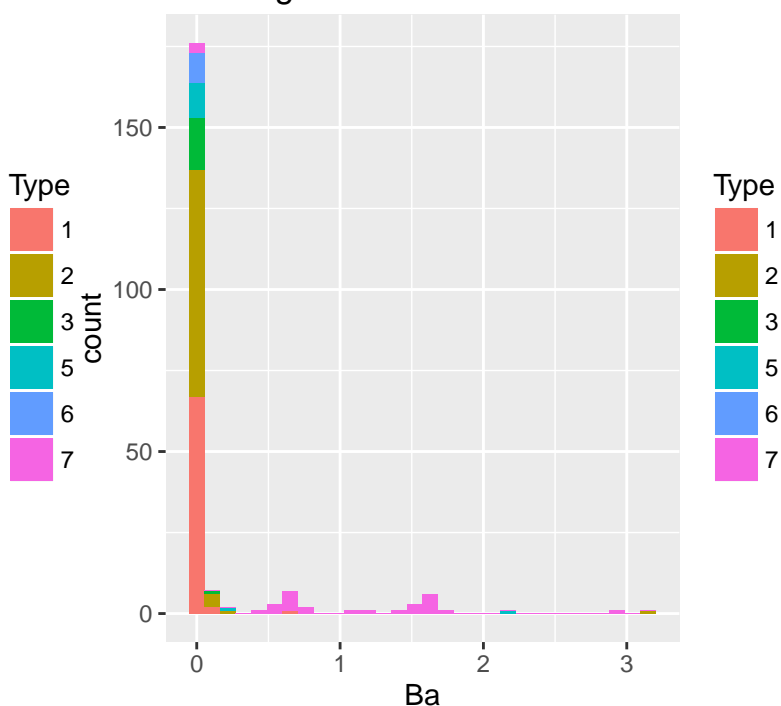
```
skewNess <- apply(glassPredictors,2,skewness)
skewNess[order(-skewNess)]
```

```
##           K           Ba           Ca           Fe           RI           Al
## 6.4600889  3.3686800  2.0184463  1.7298107  1.6027151  0.8946104
##           Na           Si           Mg
## 0.4478343 -0.7202392 -1.1364523
```

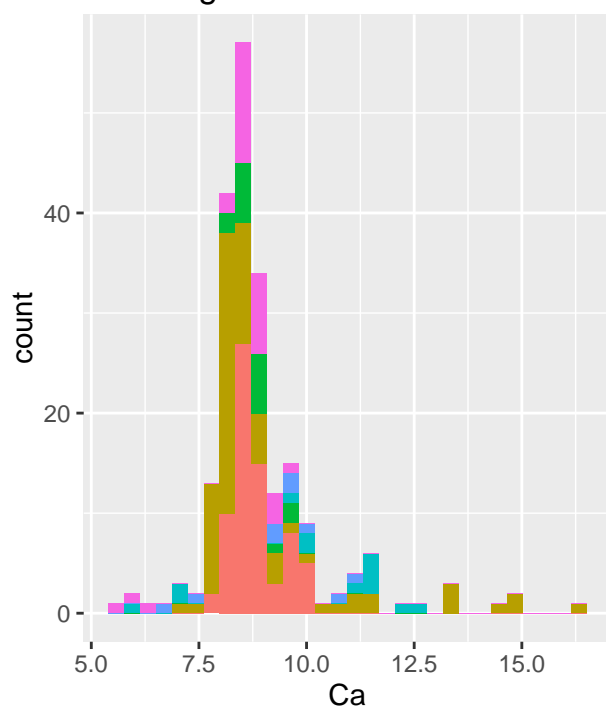
Histogram for the Element K



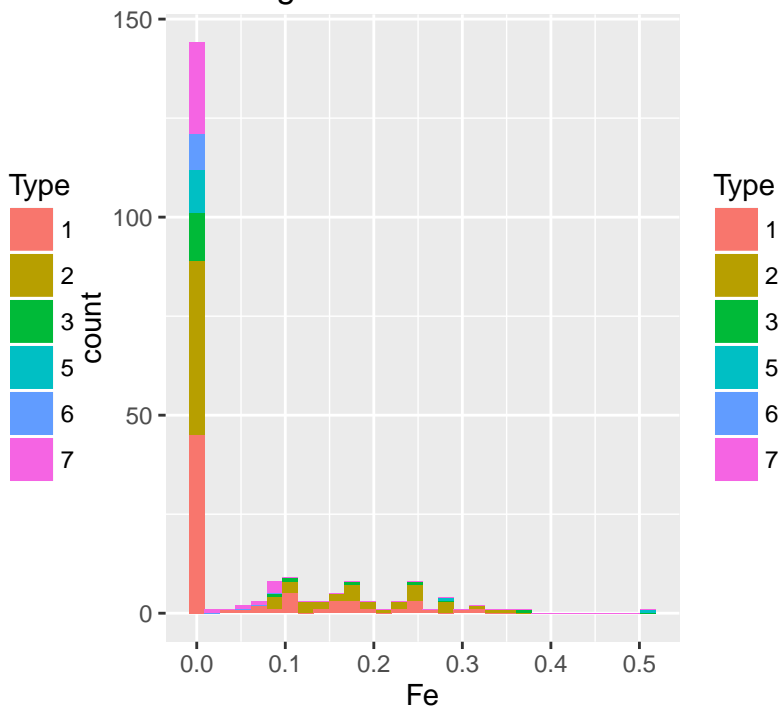
Histogram for the Element Ba



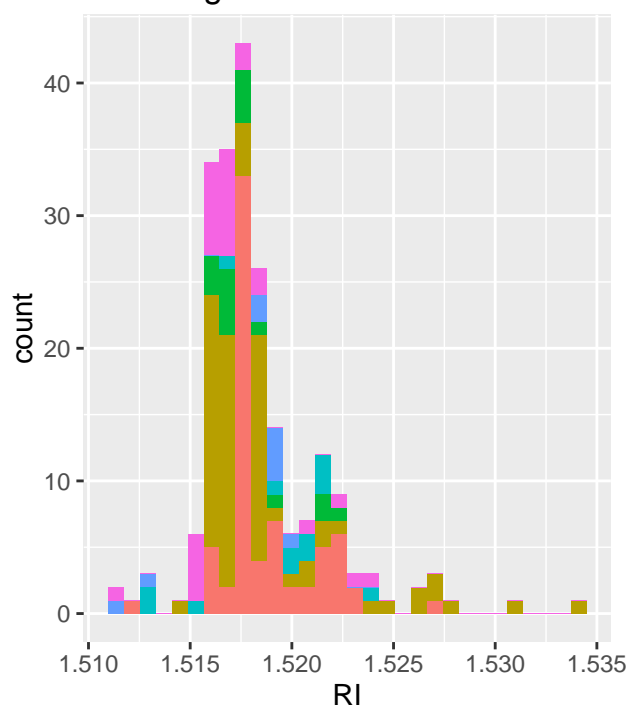
Histogram for the Element Ca



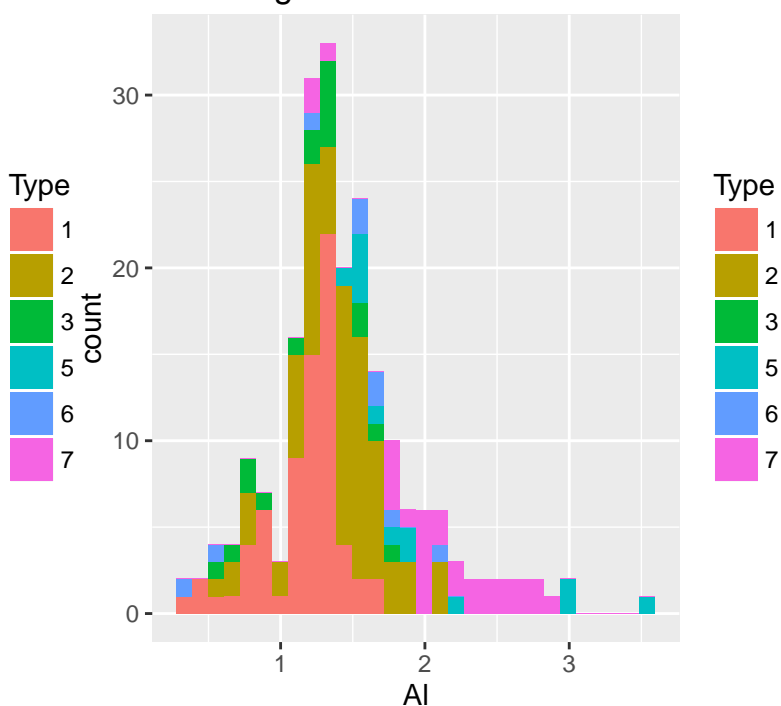
Histogram for the Element Fe



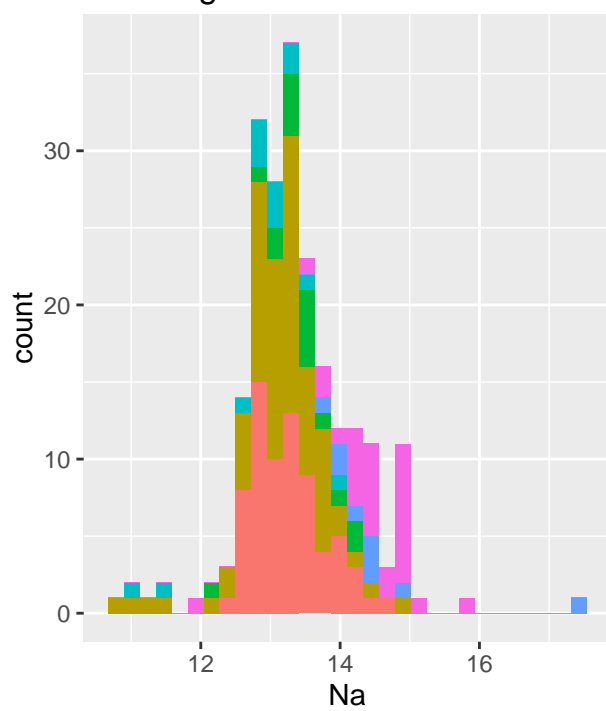
Histogram for the Element RI



Histogram for the Element Al



Histogram for the Element Na



Histogram for the Element Si

