HW#1 Nan Deng

(a)

```
car_data <- read.csv("/Users/CandiceDeng\ 1/Desktop/STATS500/HW#1/vwjetta_gas.csv")
car_data$mileage[car_data$mileage == 999999] <- NA
car_data$vehage[car_data$vehage == 999] <- NA
car_data$mpg[car_data$mpg == 99] <- NA
car_data$congrade <- factor(car_data$congrade)
levels(car_data$congrade) <- c("rough","average","clean","exellent")
summary(car_data)
```

Missing mileage = 5
Missing vehage = 8
Missing mpg = 6

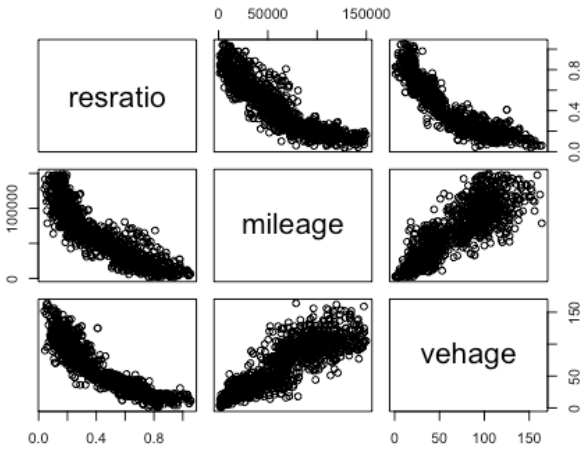
(b)

```
library(plyr)
count(car_data, 'congrade')
```

Number of rough = 578
Number of average = 378
Number of clean = 209
Number of exellent = 31

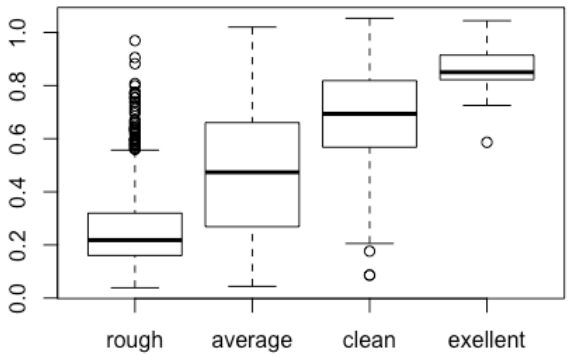
(c)

```
pairs(car_data[,c("resratio","mileage","vehage")])
```



(d)

```
plot(car_data$congrade,car_data$resratio,main="")
```

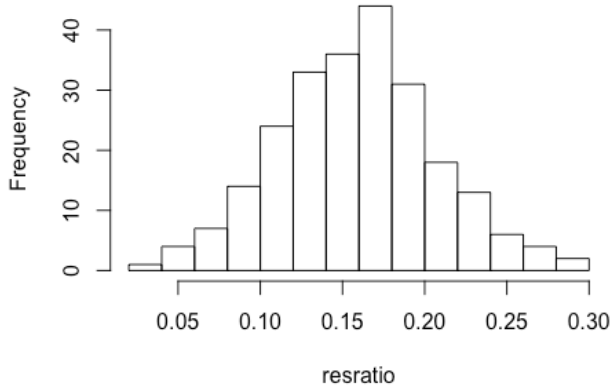



```
aggregate(car_data$resratio, list(car_data$congrade), mean)
```

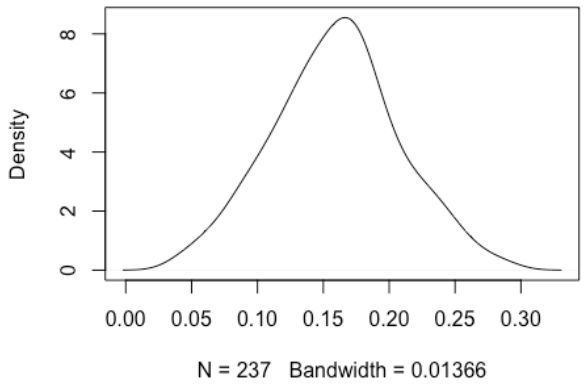
Mean resratio by rough = 0.2722137
Mean resratio by average = 0.4774244
Mean resratio by clean = 0.6732804
Mean resratio by exellent = 0.8596422

(e)

```
hist(car_data$resratio[car_data$mileage>100000],xlab="resratio",main="")
```




```
plot(density(car_data$resratio[car_data$mileage>100000],na.rm=T),main="")
```




```
mean(car_data$resratio[car_data$mileage>100000],na.rm=T)

sd(car_data$resratio[car_data$mileage>100000],na.rm=T)
```

The shape of this plot is a smoothing version of previous histogram, which represents the distribution of resratio for vehicles with mileage in each range.

Mean of resratio = 0.1593309

Standard Deviation of resratio = 0.04841774

(f)

```
car_data$mile15 <- car_data$mileage/15000
mean(car_data$resratio[car_data$congrade=="average"],na.rm=T)

sd(car_data$resratio[car_data$congrade=="average"],na.rm=T)

mean(car_data$mile15[car_data$congrade=="average"],na.rm=T)

sd(car_data$mile15[car_data$congrade=="average"],na.rm=T)

mean(car_data$vehage[car_data$congrade=="average"],na.rm=T)

sd(car_data$vehage[car_data$congrade=="average"],na.rm=T)
```

Mean of resratio = 0.4774244

Standard Deviation of resratio = 0.2266965

Mean of mile15 = 3.651413

Standard Deviation of mile15 = 2.220556

Mean of vehage = 58.17772

Standard Deviation of vehage = 34.79075