

```
#Importing the Hepatitis c Dataset
```

```
HCV= read.csv("HCV-Egy-Data.csv")
```

```
HCV
```

```
#Summary
```

```
attach(HCV)
```

```
summary(HCV)
```

```
#Dimensions of the data set
```

```
NROW(HCV)
```

```
NCOL(HCV)
```

```
#Displaying the column names of the dataset
```

```
colnames(HCV)
```

```
#Another method for dimensions
```

```
dim(HCV)
```

```
#Preprocessing data was done but didn't find any discrepancies.
```

```
na= is.na(HCV)
```

```
na
```

```
any(is.na(HCV))
```

```
#Displaying the first six rows of the datasets
```

```
head(HCV)
```

```
tail(HCV)
```

```
#Differentiating data set based on gender
```

```
Gen_male = HCV[HCV$Gender== '1',]
```

```
Gen_female = HCV[HCV$Gender=='2',]
```

```
#Exploring symptoms
```

```
#####Male Data Exploration#####
```

```
#Fever
```

```
Fev_male = Gen_male[Gen_male$Fever == '2',]
```

```
Fev_male
```

```
summary(Fev_male)
```

```
#Vomiting and Nausea
```

```
Nau_male = Gen_male[Gen_male$Nausea.Vomting =='2',]
```

```
Nau_male
```

```
summary(Nau_male)
```

```
#Fatigue
```

```
Fat_male = Gen_male[Gen_male$Fatigue...generalized.bone.ache =='2',]
```

```
Fat_male
```

```
summary(Fat_male)
```

```
#Jaundice
```

```
Jau_male = Gen_male[Gen_male$Jaundice =='2',]
```

```
Jau_male
```

```
summary(Jau_male)
```

```
#Stomack pain
```

```
sto_male = Gen_female[Gen_female$Epigastric.pain == '2',]
```

```
sto_male
```

```
summary(sto_male)
```

```
#####Female Data Exploration#####
```

```
#Fever
```

```
Fev_female = Gen_female[Gen_female$Fever == '2',]
```

```
Fev_female
```

```
summary(Fev_female)
```

```
#Vomiting and Nausea
```

```
Nau_female = Gen_female[Gen_female$Nausea.Vomting == '2',]
```

```
Nau_female
```

```
summary(Nau_female)
```

```
#Fatigue
```

```
Fat_female = Gen_female[Gen_female$Fatigue...generalized.bone.ache == '2',]
```

```
Fat_female
```

```
summary(Fat_female)
```

```
#Jaundice
```

```
Jau_female = Gen_female[Gen_female$Jaundice == '2',]
```

```
Jau_female
```

```
summary(Jau_female)
```

```
#Stomack pain
```

```
sto_female = Gen_female[Gen_female$Epigastric.pain == '2',]
```

```
sto_female
```

```
summary(sto_female)
```

```
#Exploration of the data for high chances of HCV Infection
```

```
#Here RNA.base value if it is more than 700000 units then virus is detected in high quantity.
```

```
#Here ALT.1 if value is greater than 57 then it is not normal.
```

```
#we sorted the data on these two components.
```

```
library(dplyr)
```

```
HCV_male = HCV %>% filter(Gender == 1 & RNA.Base>= 700000 & ALT.1 >= 57)
```

```
HCV_male
```

```
HCV_female = HCV %>% filter(Gender == 2 & RNA.Base>= 700000 & ALT.1 >= 57)
```

```
HCV_female
```

```
#plotting, Are they in a straight line.
```

```
#Male Plotting of the dataset is done for five different attributes.
```

```
qqnorm(HCV_male[, "RNA.Base"], main = "RNA.Base"); qqline(HCV_male[, "RNA.Base"])
```

```
qqnorm(HCV_male[, "ALT.1"], main = "ALT.1"); qqline(HCV_male[, "ALT.1"])
```

```
qqnorm(HCV_male[, "WBC"], main = "WBC"); qqline(HCV_male[, "WBC"])
```

```
qqnorm(HCV_male[, "RBC"], main = "RBC"); qqline(HCV_male[, "RBC"])
```

```
qqnorm(HCV_male[, "AST.1"], main = "AST.1"); qqline(HCV_male[, "AST.1"])
```

```
#Female, Are they in a straight line.
```

```
#FeMale Plotting of the dataset is done for five different attributes.
```

```
qqnorm(HCV_female[, "RNA.Base"], main = "RNA.Base"); qqline(HCV_female[, "RNA.Base"])
```

```
qqnorm(HCV_female[, "ALT.1"], main = "ALT.1"); qqline(HCV_female[, "ALT.1"])
```

```
qqnorm(HCV_female[, "WBC"], main = "WBC"); qqline(HCV_female[, "WBC"])
```

```
qqnorm(HCV_female[, "RBC"], main = "RBC"); qqline(HCV_female[, "RBC"])
qqnorm(HCV_female[, "AST.1"], main = "AST.1"); qqline(HCV_female[, "AST.1"])
```

```
#Visualisatiom
```

```
#Chiplot
```

```
#For male data
```

```
with(HCV_male, plot(RNA.Base, ALT.1, xlab = mlab, ylab = plab, cex.lab = 0.9))
with(HCV_male, chiplot(RNA.Base, ALT.1))
```

```
#For Female Data
```

```
with(HCV_female, plot(RNA.Base, ALT.1, xlab = mlab, ylab = plab, cex.lab = 0.9))
with(HCV_female, chiplot(RNA.Base, ALT.1))
```

```
install.packages("scatterplot3d",
lib="/Library/Frameworks/R.framework/Versions/3.5/Resources/library")
library(scatterplot3d)
```

```
typeof(RNA.EOT)
```

```
hist(HCV, xlim=RNA.EOT, ylim=Age)
```

```
library(GGally)
```

```
ggpairs(HCV_male, columns=c("AST.1", "RNA.EOT", "WBC", "ALT.1", "RBC"), color="Survivorship")
ggpairs(HCV_female, columns=c("AST.1", "RNA.EOT", "WBC", "ALT.1", "RBC"), color="Survivorship")
summary(lm(data = HCV , RNA.EOT~Age))
summary(lm(data = HCV , RNA.EOT~Gender))
summary(lm(data = HCV , RNA.EOT~WBC))
```

```
summary(lm(data = HCV , RNA.EOT~ALT.1))
```

```
cor(HCV)
```