Problem Statement

1. Use the link given below and locate the bank marketing dataset.
https://archive.ics.uci.edu/ml/machine-learning-databases/00222/
Perform the below operations:

a) Is there any association between job and default?

Answer :

```
> library(readr)
> bank <- read.csv("C:/Users/Vikram/Desktop/Acad/bank-
additional.csv",sep=";")
> View(bank)
> dim(bank)
[1] 4119 21
> str(bank)


> chisq.test(bank$job, bank$default)

        Pearson's Chi-squared test

data: bank$job and bank$default
X-squared = 224.29, df = 22, p-value < 2.2e-16



> with(bank,chisq.test( job, default))

      Pearson's Chi-squared test
data: job and default
X-squared = 224.29, df = 22, p-value < 2.2e-16

> with(bank, table( job, default) )


> with(bank, prop.table(table( job,default)))
```

There is NO association between Job and default.
Since the p-value is 2.2e-16 is less than the cut-off value of 0.05, we can reject the null hypothesis in favor of alternative hypothesis and conclude, that the variables, job & default are dependent to each other.

b) Is there any significant difference in duration of last call between? people having housing loan or not?

Answer :

```
> with(bank, chisq.test(duration,housing))
     Pearson's Chi-squared test

data: duration and housing
X-squared = 1616, df = 1654, p-value = 0.7433

> with(bank, table( duration,housing) )


[ reached getOption("max.print") -- omitted 495 rows ]
```

c) Is there any association between consumer price index and consumer?

Answer :

```
> chisq.test(bank$cons.price.idx,bank$cons.conf.idx)

          Pearson's Chi-squared test

data: bank$cons.price.idx and bank$cons.conf.idx
X-squared = 102980, df = 625, p-value < 2.2e-16

> #OR
>
> with(bank, chisq.test(cons.price.idx,cons.conf.idx))

Pearson's Chi-squared test

data: cons.price.idx and cons.conf.idx
X-squared = 102980, df = 625, p-value < 2.2e-16

> with(bank, table(cons.price.idx,cons.conf.idx))
```

There is NO association between Job and default.

d) Is the employment variation rate consistent across Job types?

Answer :

```
> chisq.test(bank$job,bank$emp.var.rate)

     Pearson's Chi-squared test

data: bank$job and bank$emp.var.rate
X-squared = 512.04, df = 99, p-value < 2.2e-16
```

```
> with(bank, chisq.test( job,emp.var.rate))

Pearson's Chi-squared test

data: job and emp.var.rate
X-squared = 512.04, df = 99, p-value < 2.2e-16

> with(bank, table( job,emp.var.rate) )
```

e) Is the employment variation rate same across Education?

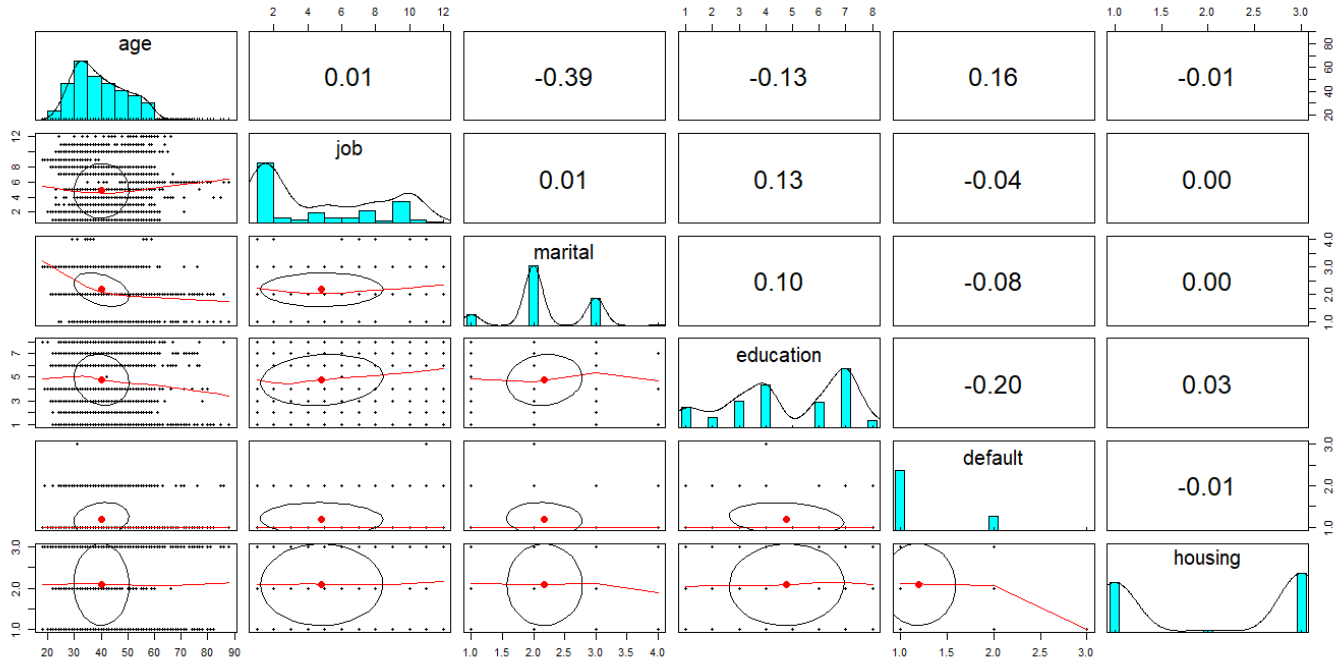
Answer :

```
> with(bank, chisq.test( education,emp.var.rate))

     Pearson's Chi-squared test

data: education and emp.var.rate
X-squared = 193.46, df = 63, p-value = 3.5e-15

> with(bank, table( education, emp.var.rate) )
```

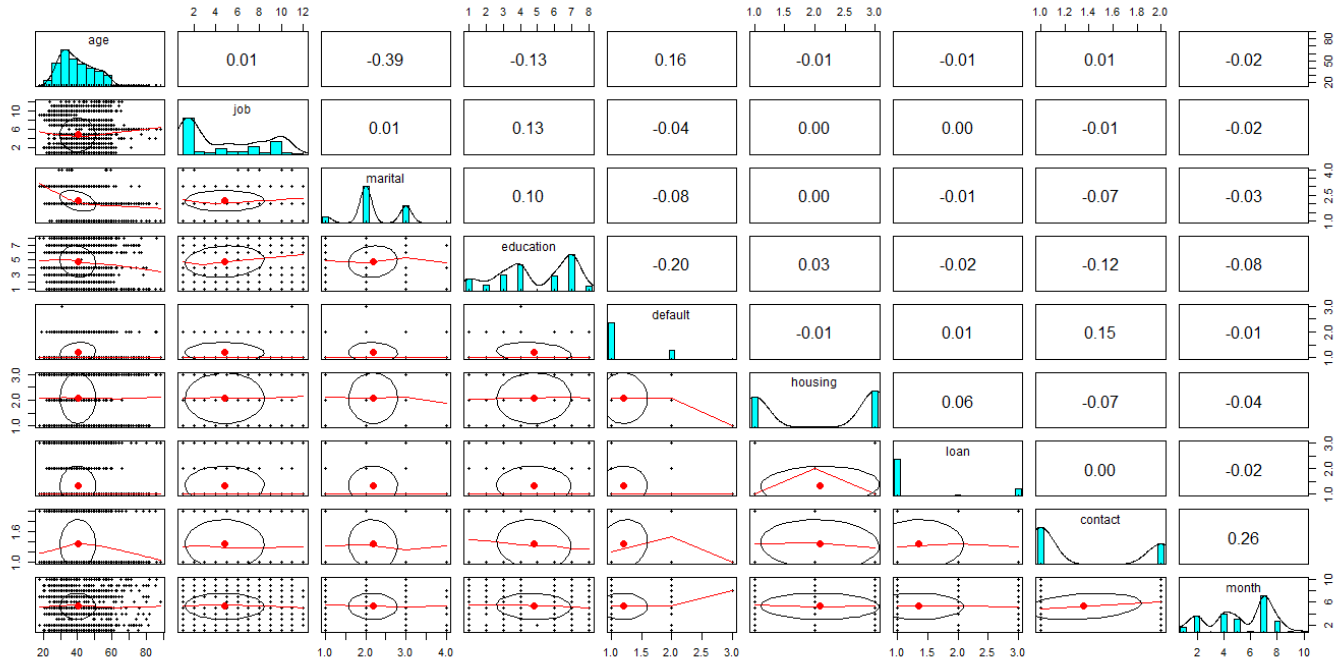
f) Which group is more confident?

Answer :

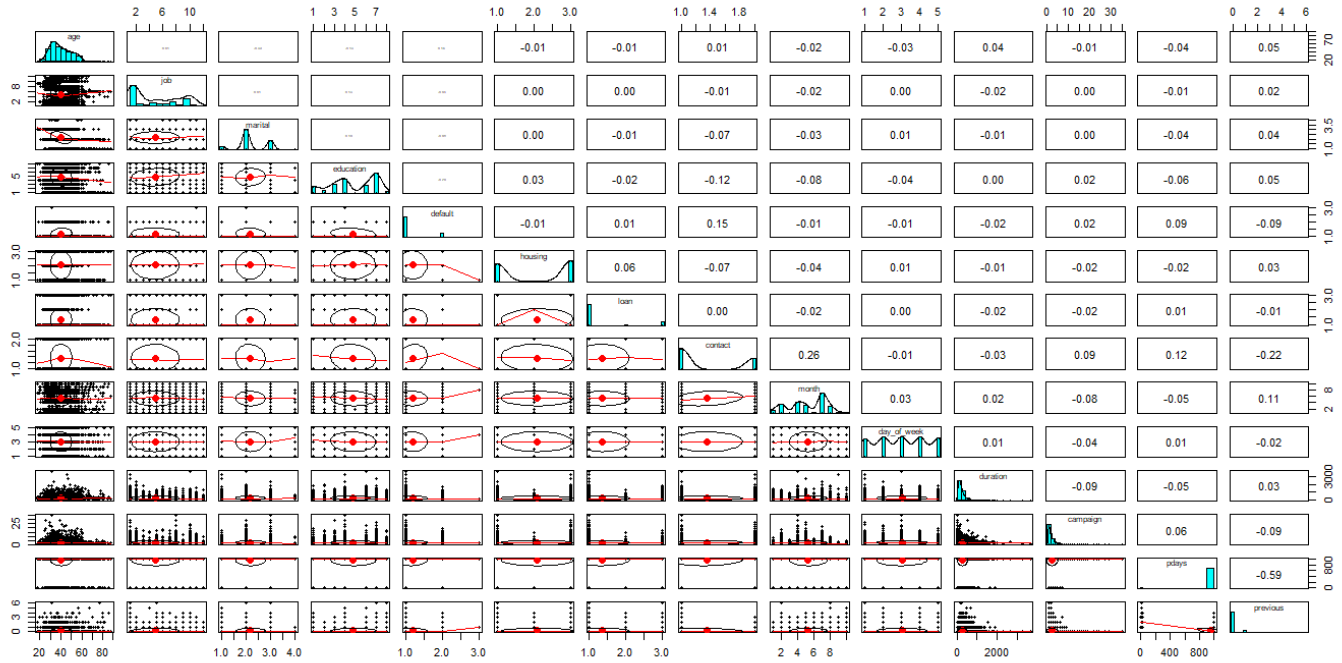
```
> library(psych)
> pairs.panels(bank[,1:6])
```




```
> pairs.panels(bank[,1:9])
```




```
> pairs.panels(bank[,1:14])
```




```
> summary(bank)
```