Assignment - 2

Session 11 – Linear Models

1.Use the given link below and locate the bank marketing dataset. Data Set Link

Perform the below operations:

1. Is there any association between Job and default?

Ans:

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

Pearson's Chi-squared test

data: bank.additional.full$job and bank.additional.full$default

X-squared = 1927.7, df = 22, p-value < 2.2e-16

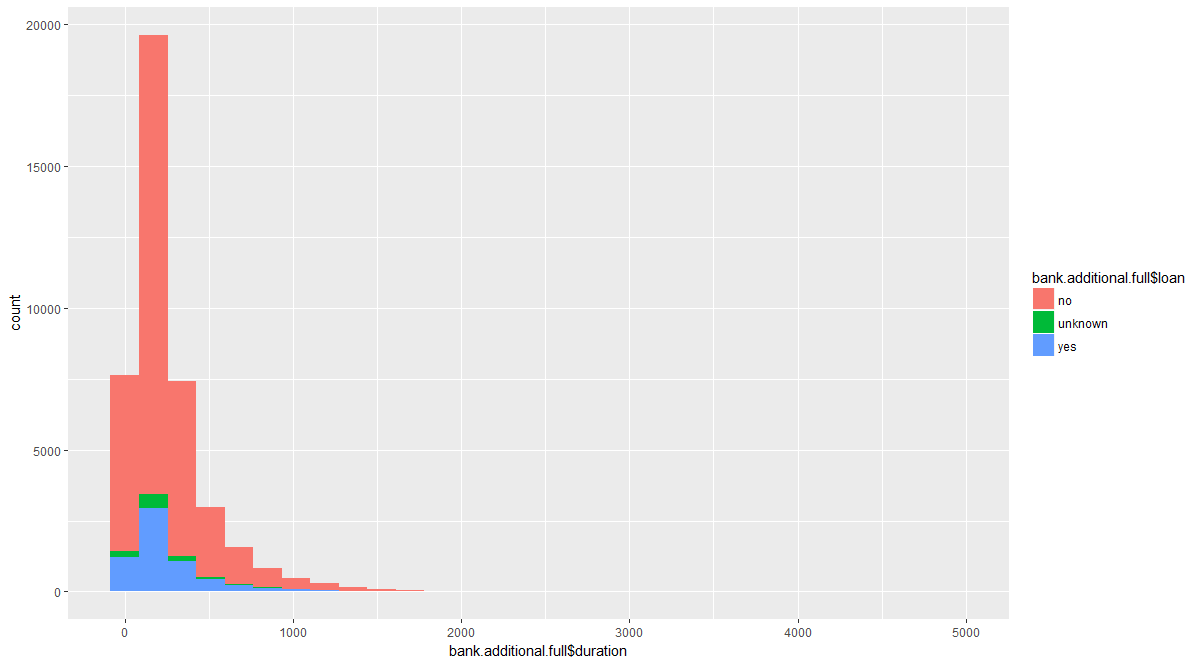
1. Is there any significant difference in duration of last call between people having housing loan or not?

Ans:

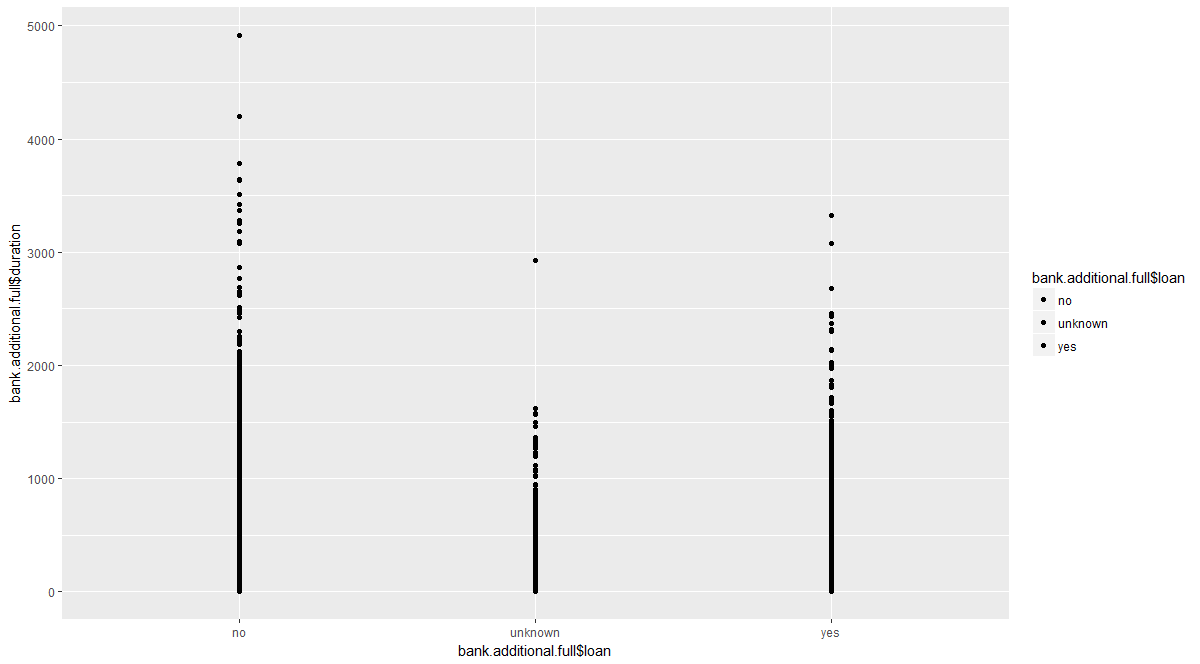
> library(ggplot2)

> bank.additional.full$duration<-as.numeric(bank.additional.full$duration)

> ggplot(bank.additional.full, aes(x=bank.additional.full$duration, fill=bank.additional.full$loan))+geom\_histogram()



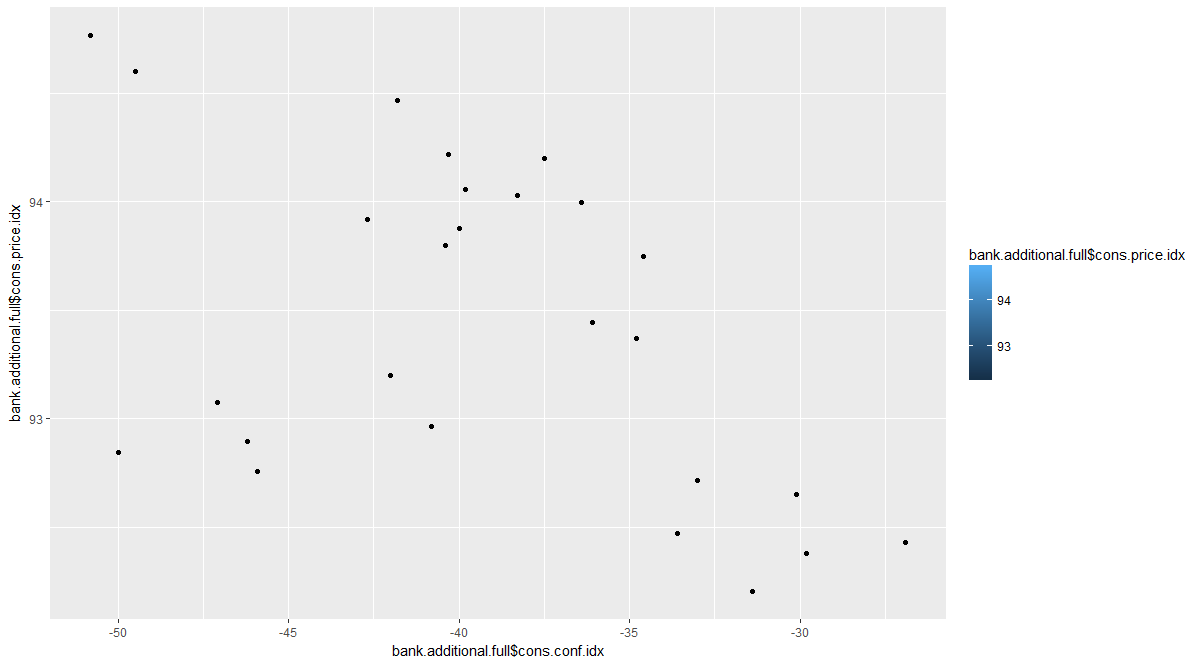
> ggplot(bank.additional.full, aes(x=bank.additional.full$loan,y=bank.additional.full$duration, fill=bank.additional.full$loan))+geom\_point()



1. Is there any association between consumer price index and consumer?

Ans:

> ggplot(bank.additional.full, aes(x=bank.additional.full$cons.conf.idx, y= bank.additional.full$cons.price.idx, fill=bank.additional.full$cons.price.idx))+geom\_point()



1. Is the employment variation rate consistent across job types?

Ans:

> chisq.test(bank.additional.full$job ,bank.additional.full$education)

Pearson's Chi-squared test

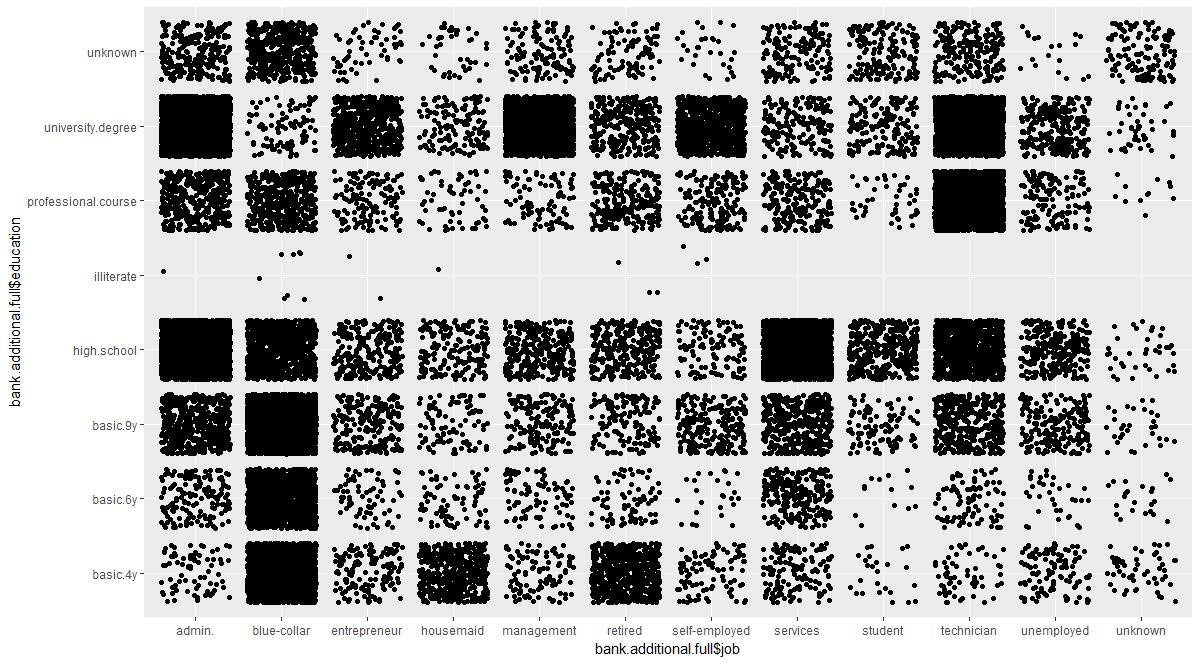
data: bank.additional.full$job and bank.additional.full$education

X-squared = 37338, df = 77, p-value < 2.2e-16

1. Is the employment variation rate same across education?

Ans:

> ggplot(bank.additional.full, aes(x=bank.additional.full$job, y=bank.additional.full$education))+geom\_jitter()



1. Which group is more confident?

Ans:

# Not covered in class