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

Perform the below operations:

a. Create a visual for representing missing values in the dataset.

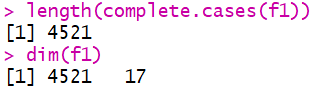
ANS.

Data…

**unzip("C:/Users/Ashwin/Desktop/AcadGildassignments/assignment\_11.1/data/bank.zip")**

**f1 <- read.csv2("C:/Users/Ashwin/Documents/bank\_data/bank.csv")**

There are no missing values in dataset. This is being tested using summary() command and also by complete.cases() method.



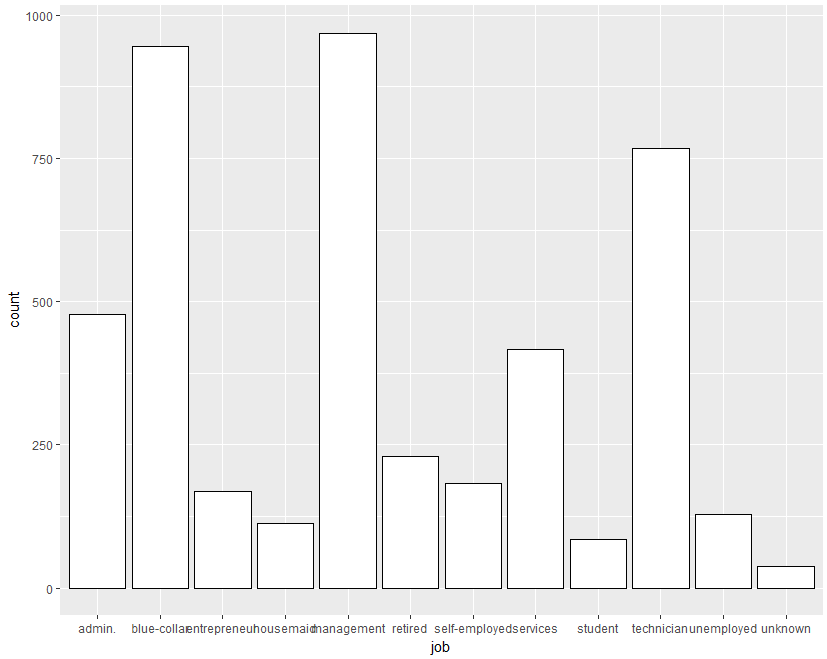
Complete.cases() returns a logical vector indicating absence or presence of missing values. ‘True’ if no missing value. ‘False’ if there is a missing value.

b. Show a distribution of clients based on a Job.

ANS.

library(ggplot2)

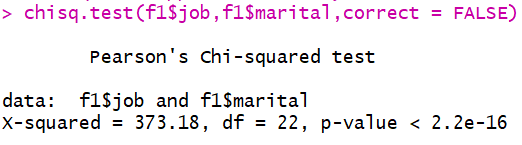
ggplot(data=f1,aes(x=job)) + geom\_bar(color="black",fill="white")



c. Check whether is there any relation between Job and Marital Status?

ANS.

Both variables Job and Marital Status are categorical. Whenever we want to test relation or association between two such variables we use CHI-SQUARE TEST.

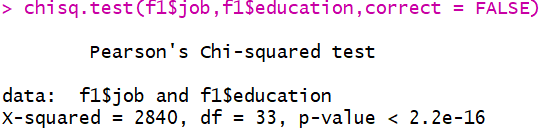


Since p-value is smaller than 0.05 we reject the hull hypothesis that they are not related. Thus the two variables do have a relation.

d. Check whether is there any association between Job and Education?

ANS.

Both variables Job and Education are categorical. Whenever we want to test relation or association between two such variables we use CHI-SQUARE TEST.



Since p-value is smaller than 0.05 we reject the hull hypothesis that they are not related. Thus the two variables do have a relation.