

1. Develop 5 different visuals using GGLOT with descriptions of the insights they convey.

ANS)

```
url<- "https://github.com/SavioSal/datasets/raw/master/Bank%20Churn_Modelling.csv"
```

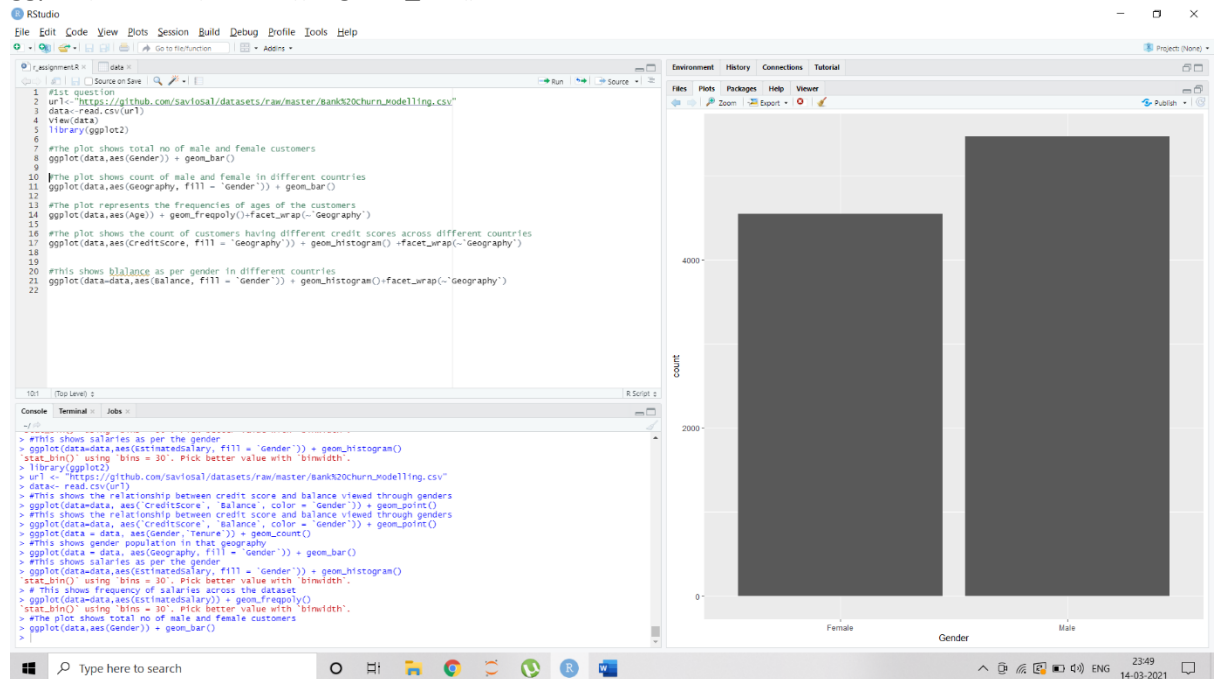
```
data<-read.csv(url)
```

```
View(data)
```

```
library(ggplot2)
```

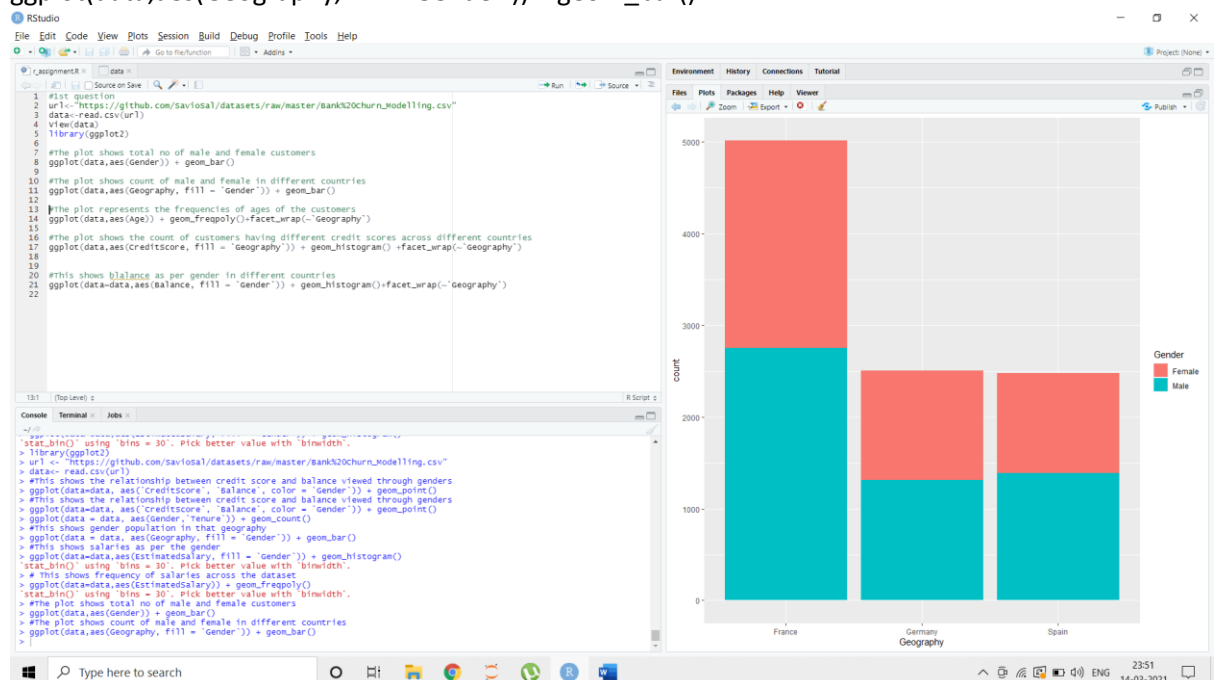
#The plot shows total no of male and female customers

```
ggplot(data,aes(Gender)) + geom_bar()
```



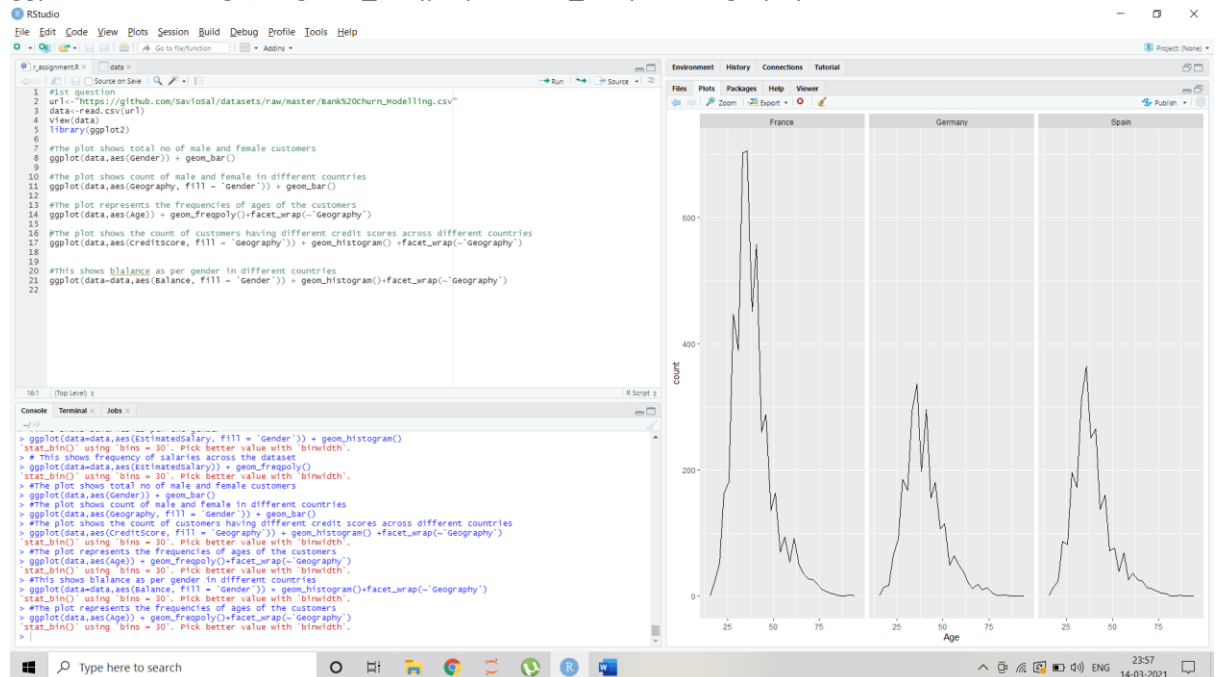
#The plot shows count of male and female in different countries

```
ggplot(data,aes(Geography, fill = 'Gender')) + geom_bar()
```



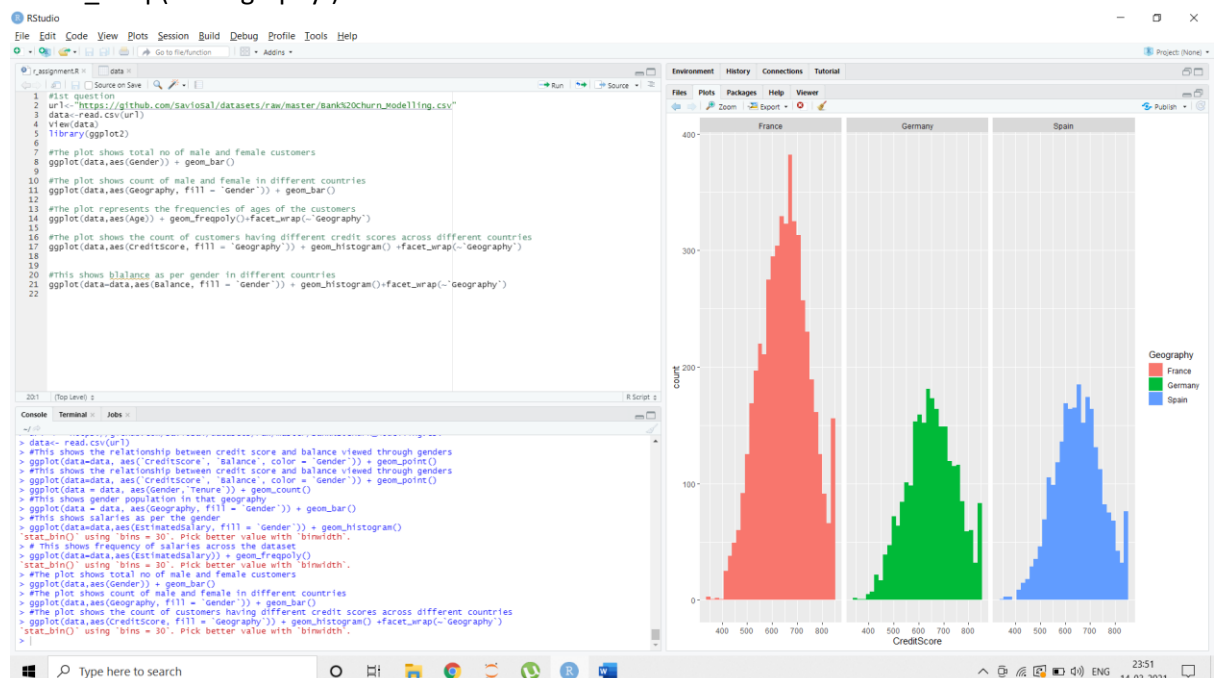
#The plot represents the frequencies of ages of the customers in different countries

`ggplot(data,aes(Age)) + geom_freqpoly()+facet_wrap(~`Geography`)`



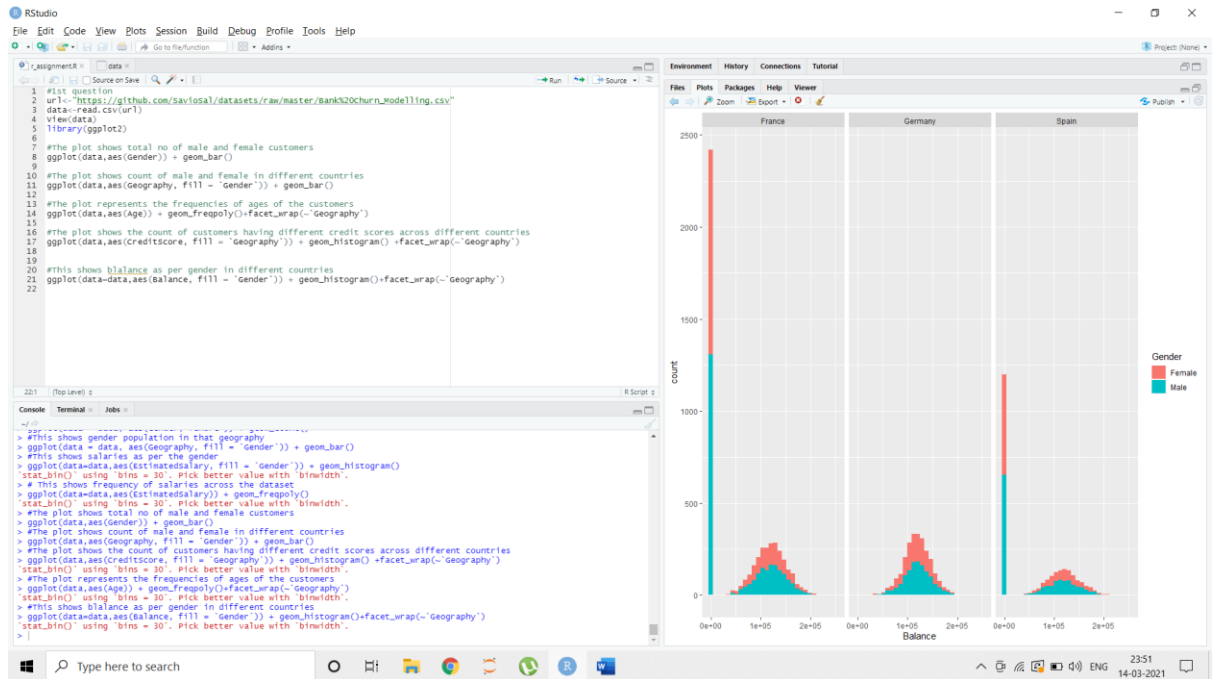
#The plot shows the count of customers having different credit scores across different countries

`ggplot(data,aes(CreditScore, fill = `Geography`)) + geom_histogram()+facet_wrap(~`Geography`)`



#This shows balance as per gender in different countries

`ggplot(data=data,aes(Balance, fill = `Gender`)) + geom_histogram()+facet_wrap(~`Geography`)`



2)

#a)What is the average credit score of females and males in France?

```
data%>% select(CreditScore, Gender, Geography) %>% filter(Geography == "France") %>%
  dplyr::group_by(Gender) %>%
  dplyr::summarise(Gender_Average = mean(CreditScore))
```

#b)What is the average credit score of people in the age brackets 20-30,31-40,41-50?

```
data %>% select(CreditScore, Age) %>% mutate(agegroup = case_when(Age >= 41 & Age <=
50 ~ '3', Age >= 31 & Age <= 40 ~ '2', Age >= 20 & Age <= 30 ~ '1')) %>%
  filter(agegroup == "1" | agegroup == "2" | agegroup == "3") %>%
  dplyr::group_by(agegroup) %>%
  dplyr::summarise(Age_Average = mean(CreditScore))
```

#c)What is the correlation between credit score and estimated salary?

```
data %>% select(CreditScore, EstimatedSalary) %>% cor()
model <- lm(CreditScore ~ Gender+Age+EstimatedSalary, data = data)
```

#d)printing the model

```
print(model)
summary(model)
```

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Project: Home

Source

```
1 #1st question
2 url<- "https://github.com/savioosal/datasets/raw/master/bank%20churn_modeling.csv"
3 data<-read.csv(url)
4 View(data)
5 library(ggplot2)
6
7 #the plot shows total no of male and female customers
8 ggplot(data,aes(Gender)) + geom_bar()
9
10 #The plot shows count of male and female in different countries
11 ggplot(data,aes(Geography, fill = "Gender")) + geom_bar()
12
13 #The plot represents the frequencies of ages of the customers
14 ggplot(data,aes(Age)) + geom_freqpoly()+facet_wrap(~"Geography")
15
16 #the plot shows the count of customers having different credit scores across different countries
17 ggplot(data,aes(Creditscore, fill = "Geography")) + geom_histogram()+facet_wrap(~"Geography")
18
19
20 #this shows blalance as per gender in different countries
21 ggplot(data+data,aes(balance, fill = "Gender")) + geom_histogram()+facet_wrap(~"Geography")
22
23
24
25
26
27 #2nd question
28 #a)what is the average credit score of females and males in France?
29 data%>% select(Creditscore, Gender, Geography) %>% filter(Geography == "France") %>%
30 dplyr::group_by(Gender) %>%
31 dplyr::summarise(Gender_Average = mean(Creditscore))
32
33
34 #b)what is the average credit score of people in the age brackets 20-30,31-40,41-50?
35 data %>% select(Creditscore, Age) %>% mutate(Agegroup = case_when(Age >= 41 & Age <= 50 ~ "3", Age >= 31 & Age <= 40 ~ "2", Age >= 20 & Age <= 30 ~ "1")) %>%
36 filter(Agegroup == "1" | Agegroup == "2" | Agegroup == "3") %>%
37 dplyr::group_by(Agegroup) %>%
38 dplyr::summarise(Age_Average = mean(Creditscore))
39
40
41 #c)what is the correlation between credit score and estimated salary?
42 data %>% select(Creditscore, EstimatedSalary) %>% cor()
43 model <- lm(Creditscore ~Gender+Age+EstimatedSalary, data = data)
44
45 #d)printing the model
46 print(model)
47 summary(model)
```

271 (Top Level) | R Script 1

Console

Type here to search

0008 15-03-2021