Project on EXAM_SCORE

Hide

```
# Here we install and load all required packages
install.packages("tidyverse")
library(tidyverse)
library(dplyr)
library(ggplot2)
```

Hide

```
# Here we will read the file using read.csv() function
score_df <- read_csv("Exam_Score.csv")
```

Hide

```
View(score_df)
```

Hide

```
str(score_df)
```

```
spc_tbl_ [25 × 13] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
 $ Roll_Number  : chr [1:25] "A_25" "A_19" "A_8" "A_17" ...
 $ First_Nmae   : chr [1:25] "Aniket" "Nayan" "Parthib" "Soumyadeep" ...
 $ Last_Nmae    : chr [1:25] "Chakraborty" "Kumar" "Pal" "Sarkar" ...
 $ Age          : num [1:25] 21 20 22 21 20 19 21 22 20 19 ...
 $ Gender       : chr [1:25] "Male" "Male" "Male" "Male" ...
 $ Maths_Score  : num [1:25] 98 97 95 85 78 79 96 58 48 78 ...
 $ Science_Score: num [1:25] 58 78 89 78 85 96 89 99 98 97 ...
 $ English_Score: num [1:25] 98 98 58 78 78 98 58 85 58 88 ...
 $ Bengali_Score: num [1:25] 99 98 95 96 96 98 52 78 85 52 ...
 $ Total_Marks  : num [1:25] 353 371 337 337 337 371 295 320 289 315 ...
 $ Out_Of       : num [1:25] 400 400 400 400 400 400 400 400 400 400 ...
 $ Percentage   : num [1:25] 88.2 92.8 84.2 84.2 84.2 ...
 $ Remarks      : chr [1:25] "Outsanding" "Excellent" "Outsanding" "Outsandin
g" ...
 - attr(*, "spec")=
  .. cols(
  ..   Roll_Number = col_character(),
  ..   First_Nmae = col_character(),
```

```
  ..   Last_Nmae = col_character(),
  ..   Age = col_double(),
  ..   Gender = col_character(),
  ..   Maths_Score = col_double(),
  ..   Science_Score = col_double(),
  ..   English_Score = col_double(),
  ..   Bengali_Score = col_double(),
  ..   Total_Marks = col_double(),
  ..   Out_Of = col_double(),
  ..   Percentage = col_double(),
  ..   Remarks = col_character()
  .. )
 - attr(*, "problems")=<externalptr>
```

Hide

```
colnames(score_df)
```

```
 [1] "Roll_Number"   "First_Nmae"    "Last_Nmae"     "Age"           "Gender"
 [6] "Maths_Score"   "Science_Score" "English_Score" "Bengali_Score" "Total_M
arks"
[11] "Out_Of"        "Percentage"    "Remarks"
```

Hide

```
summary(score_df)
```

```
 Roll_Number         First_Nmae         Last_Nmae              Age
 Length:25          Length:25          Length:25          Min.   :18.00
 Class :character   Class :character   Class :character   1st Qu.:19.00
 Mode  :character   Mode  :character   Mode  :character   Median :20.00
                                                          Mean   :20.28
                                                          3rd Qu.:21.00
                                                          Max.   :22.00
    Gender           Maths_Score    Science_Score   English_Score   Bengali_S
core
 Length:25          Min.   :45.00   Min.   :48.00   Min.   :48.00   Min.   :5
2.00
 Class :character   1st Qu.:78.00   1st Qu.:78.00   1st Qu.:78.00   1st Qu.:7
8.00
 Mode  :character   Median :88.00   Median :89.00   Median :87.00   Median :8
5.00
```

```
                  Mean   :82.12   Mean   :85.92   Mean   :83.08   Mean   :8
1.36
                  3rd Qu.:95.00   3rd Qu.:97.00   3rd Qu.:92.00   3rd Qu.:9
6.00
                  Max.   :99.00   Max.   :99.00   Max.   :98.00   Max.   :9
9.00
  Total_Marks        Out_Of      Percentage      Remarks
 Min.   :289.0   Min.   :400   Min.   :72.25   Length:25
 1st Qu.:316.0   1st Qu.:400   1st Qu.:79.00   Class :character
 Median :332.0   Median :400   Median :83.00   Mode  :character
 Mean   :332.5   Mean   :400   Mean   :83.12
 3rd Qu.:348.0   3rd Qu.:400   3rd Qu.:87.00
 Max.   :371.0   Max.   :400   Max.   :92.75
```

Hide

```
nrow(score_df)
```

```
[1] 25
```

Hide

```
ncol(score_df)
```

```
[1] 13
```

Hide

```
# Selecting the Total_Marks and Percentage Column
score_df %>% select(Total_Marks,Percentage) %>% arrange(Total_Marks)
```

Total_Marks <dbl>	Percentage <dbl>
289	72.25
295	73.75
309	77.25
310	77.50
313	78.25
315	78.75
316	79.00
319	79.75
320	80.00

Total_Marks <dbl>	Percentage <dbl>
325	81.25

Next
123
Previous
1-10 of 25 rows

Hide

```
# Filtering rows having good remarks and maths_score greater than or equal to 90

score_df %>% filter(Remarks=='Good' & Maths_Score >=90)
```

Roll_Number <chr>	First_Nmae <chr>	Last_Nmae <chr>	Age <dbl>	Gender <chr>	Maths_Score <dbl>	Science_Score <dbl>	English_Score <dbl>	B
A_16	Aishi	Pramanik	21	Female	96	89	58	

1 row | 1-9 of 13 columns

Hide

```
NA
```

Hide

```
# Finding the mean and minimum maths_marks and maximum math_marks

score_df %>% summarize(maths_mean=mean(Maths_Score),math_min=min(Maths_Score)
,
                       math_max=max(Maths_Score))
```

maths_mean <dbl>	math_min <dbl>	math_max <dbl>
82.12	45	99

1 row

Hide

```
# Creating a new column called geography _marks and naming it as updated_scor
e_df

updated_score_df <- score_df %>% mutate(Geography_Score= (Maths_Score-10))

head(updated_score_df,5)
```

Roll_Number <chr>	First_Nmae <chr>	Last_Nmae <chr>	Age <dbl>	Gender <chr>	Maths_Score <dbl>	Science_Score <dbl>	English_Score <dbl>
A_25	Aniket	Chakraborty	21	Male	98	58	98
A_19	Nayan	Kumar	20	Male	97	78	98

Roll_Number <chr>	First_Nmae <chr>	Last_Nmae <chr>	Age <dbl>	Gender <chr>	Maths_Score <dbl>	Science_Score <dbl>	English_Score <dbl>
A_8	Parthib	Pal	22	Male	95	89	58
A_17	Soumyadeep	Sarkar	21	Male	85	78	78
A_10	Ujjal	Pattanayek	20	Male	78	85	78

5 rows | 1-8 of 14 columns

Hide

```
# The new variable name is Gepography_Score

# We obtain the values of the variable by substracting 10 from Maths_Score
```

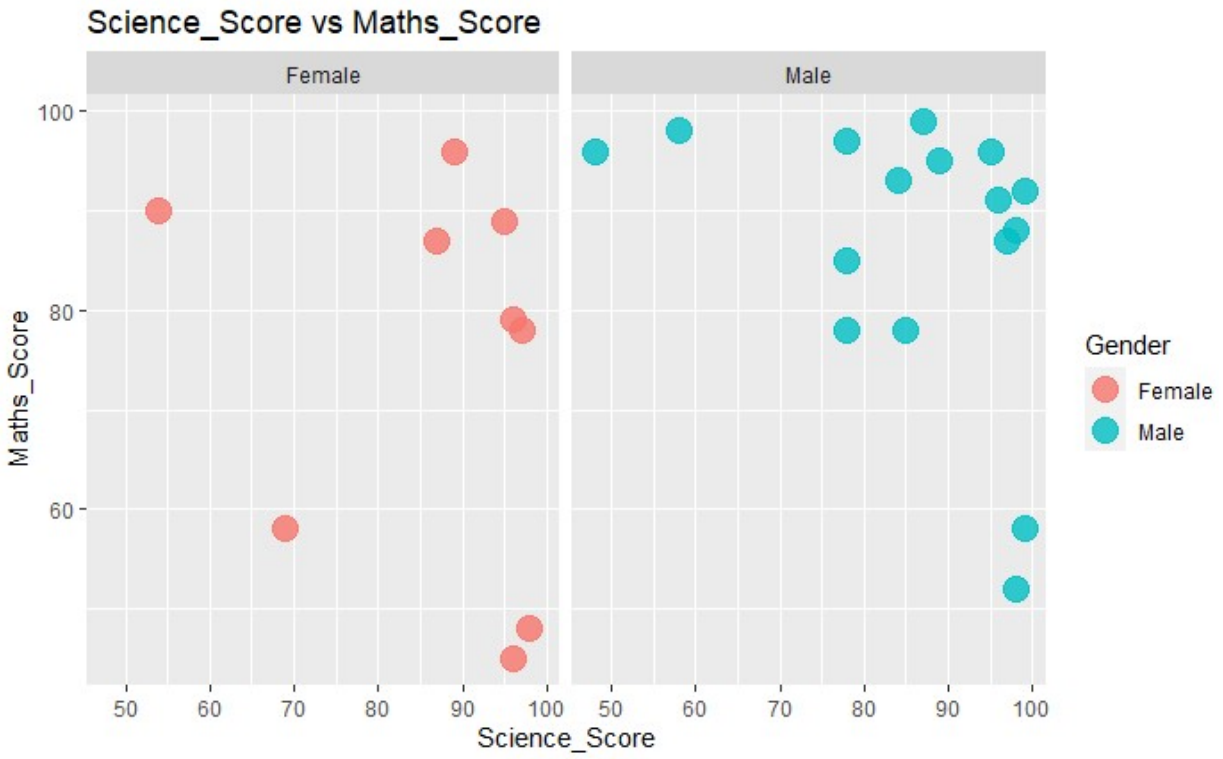
Hide

```
colnames(updated_score_df)
```

```
 [1] "Roll_Number"     "First_Nmae"      "Last_Nmae"       "Age"
 [5] "Gender"          "Maths_Score"     "Science_Score"   "English_Score"
 [9] "Bengali_Score"   "Total_Marks"     "Out_Of"          "Percentage"
[13] "Remarks"         "Geography_Score"
```

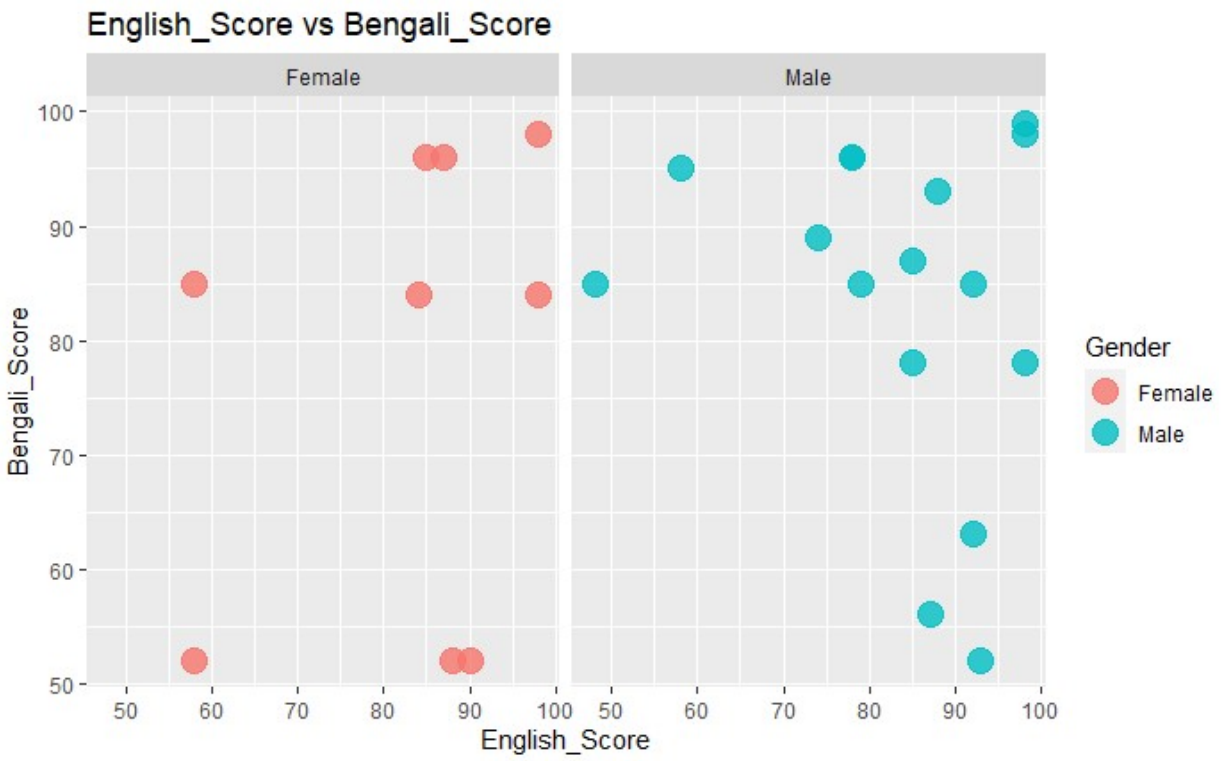
Hide

```
# Plotting scatter plot for Science_Marks vs Maths_Score and customizing it

p=ggplot(data=updated_score_df)+

geom_point(mapping=aes(x=Science_Score,y=Maths_Score,color=Gender),alpha=0.8,
size=5)+

facet_wrap(~Gender)+

labs(title="Science_Score vs Maths_Score ",x='Science_Score',y='Maths_Score')

p
```



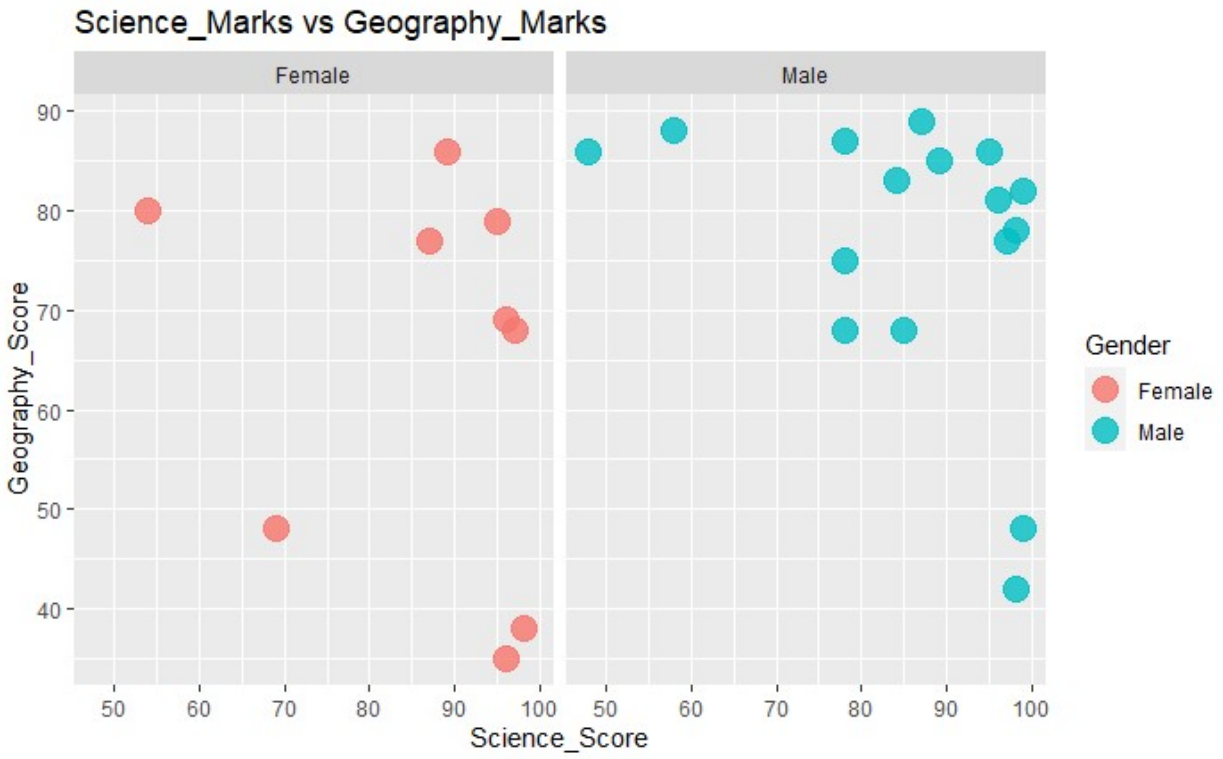
Hide

```
# Plotting scatter plot for English_Marks vs Bengali_Score and customizing it

p1=ggplot(data=updated_score_df)+

geom_point(mapping=aes(x=English_Score,y=Bengali_Score,color=Gender),alpha=0.
8,size=5)+

facet_wrap(~Gender)+

labs(title="English_Score vs Bengali_Score ",x='English_Score',y='Bengali_Sco
re')

p1
```



Hide

```
# Plotting scatter plot for Science_Marks vs Geography_Marks and customizing
it

p2=ggplot(data=updated_score_df)+

geom_point(mapping=aes(x=Science_Score,y=Geography_Score,color=Gender),alpha=
0.8,size=5)+

facet_wrap(~Gender)+

labs(title='Science_Marks vs Geography_Marks',x='Science_Score',y='Geography_
Score')

p2
```



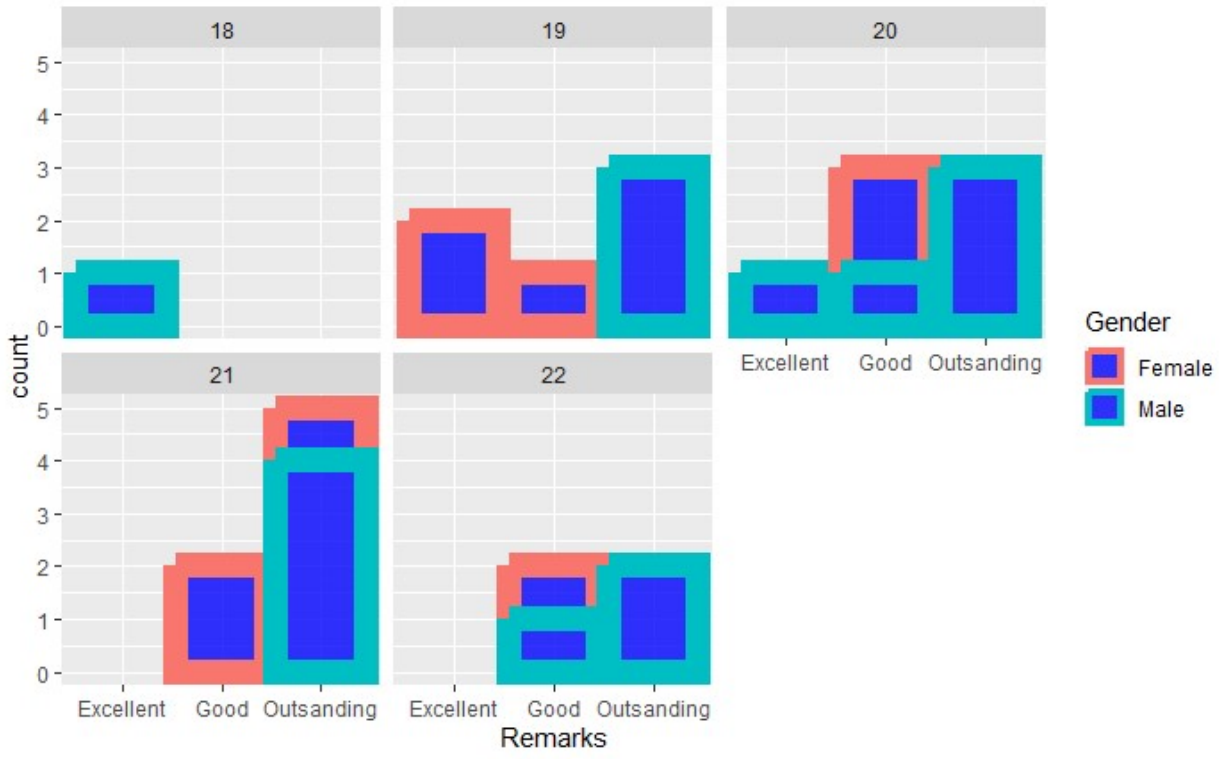
Hide

```
p3=ggplot(data=updated_score_df)+
geom_bar(mapping=aes(x=Remarks,color=Gender),alpha=0.8,size=5,fill='blue')+facet_wrap(~Age)
```

```
Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
Please use `linewidth` instead.
```

Hide

```
p3
```


18
19
20
21
22
count
0
1
2
3
4
5
Excellent
Good
Outsanding
Remarks
Gender
Female
Male