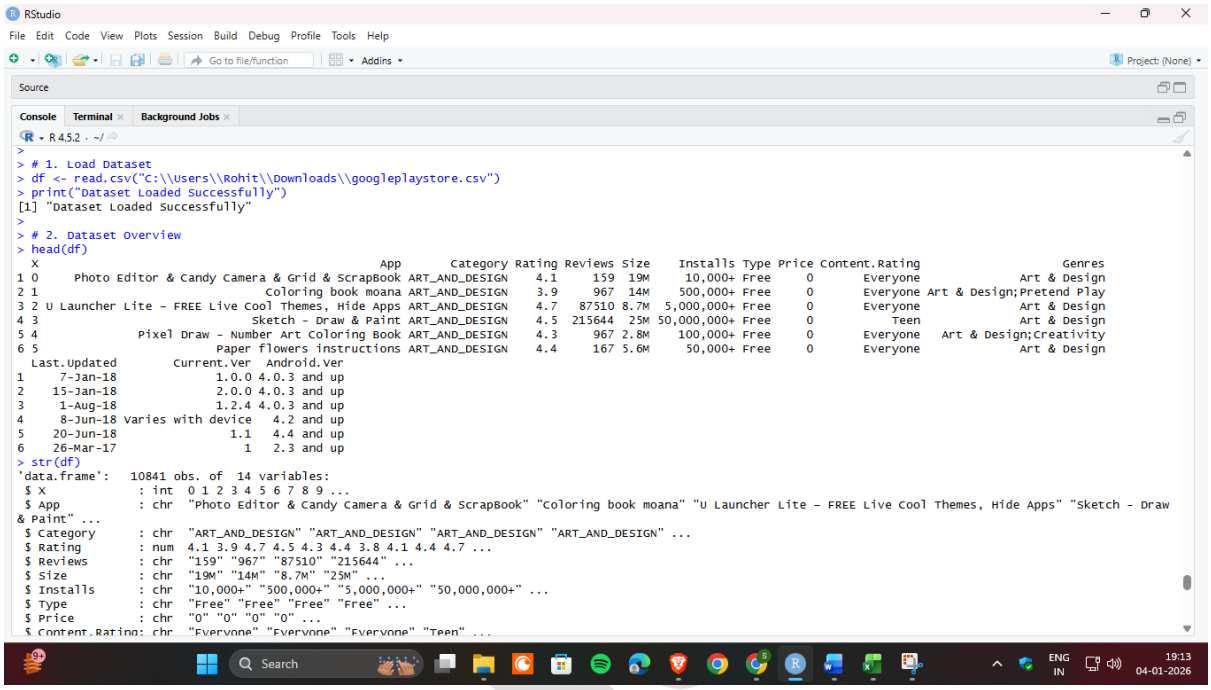
SHETH L.U.J AND SIR M.V. COLLEGE

SUBJECT :- R PROGRAMMING

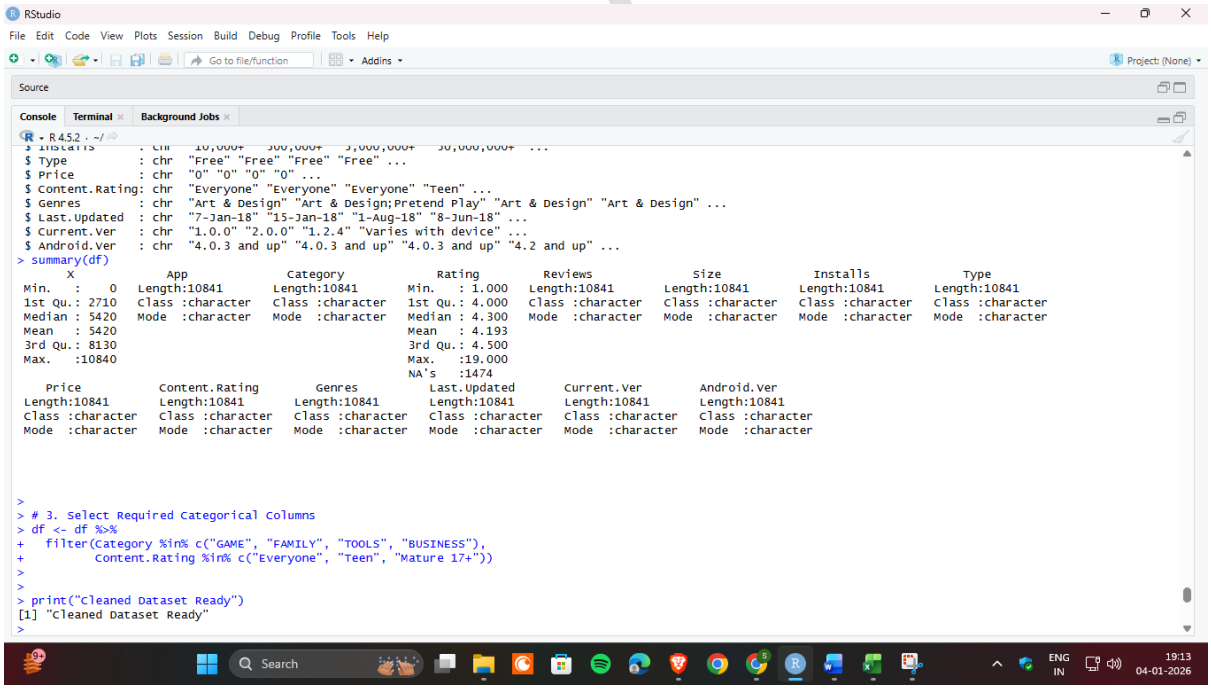
MODULE 2 – PRACTICAL - 9

AIM: Conducting Chi-square tests using chisq.test() (R)

OUTPUT:-



```
> # 1. Load Dataset
> df <- read.csv("C:\\Users\\Rohit\\Downloads\\googleplaystore.csv")
> print("Dataset Loaded Successfully")
[1] "Dataset Loaded Successfully"
>
> # 2. Dataset Overview
> head(df)
  X                                                 App        Category Rating Reviews Size     Installs Type Price Content.Rating                     Genres
1 0     Photo Editor & Candy Camera & Grid & ScrapBook ART_AND_DESIGN    4.1     159  19M      10,000+ Free     0       Everyone               Art & Design
2 1                                Coloring book moana ART_AND_DESIGN    3.9     967  14M     500,000+ Free     0       Everyone Art & Design;Pretend Play
3 2 U Launcher Lite – FREE Live Cool Themes, Hide Apps ART_AND_DESIGN    4.7   87510 8.7M   5,000,000+ Free     0       Everyone               Art & Design
4 3                              Sketch - Draw & Paint ART_AND_DESIGN    4.5  215644  25M 50,000,000+ Free     0           Teen               Art & Design
5 4               Pixel Draw - Number Art Coloring Book ART_AND_DESIGN    4.3     967 2.8M     100,000+ Free     0       Everyone   Art & Design;Creativity
6 5                         Paper flowers instructions ART_AND_DESIGN    4.4     167 5.6M      50,000+ Free     0       Everyone               Art & Design
  Last.Updated         Current.Ver  Android.Ver
1     7-Jan-18               1.0.0 4.0.3 and up
2    15-Jan-18               2.0.0 4.0.3 and up
3     1-Aug-18               1.2.4 4.0.3 and up
4     8-Jun-18 Varies with device   4.2 and up
5    20-Jun-18                 1.1   4.4 and up
6    26-Mar-17                   1   2.3 and up
> str(df)
'data.frame':   10841 obs. of  14 variables:
 $ X             : int  0 1 2 3 4 5 6 7 8 9 ...
 $ App           : chr  "Photo Editor & Candy Camera & Grid & ScrapBook" "Coloring book moana" "U Launcher Lite – FREE Live Cool Themes, Hide Apps" "Sketch - Draw
& Paint" ...
 $ Category      : chr  "ART_AND_DESIGN" "ART_AND_DESIGN" "ART_AND_DESIGN" "ART_AND_DESIGN" ...
 $ Rating        : num  4.1 3.9 4.7 4.5 4.3 4.4 3.8 4.1 4.4 4.7 ...
 $ Reviews       : chr  "159" "967" "87510" "215644" ...
 $ Size          : chr  "19M" "14M" "8.7M" "25M" ...
 $ Installs      : chr  "10,000+" "500,000+" "5,000,000+" "50,000,000+" ...
 $ Type          : chr  "Free" "Free" "Free" "Free" ...
 $ Price         : chr  "0" "0" "0" "0" ...
 $ Content.Rating: chr  "Everyone" "Everyone" "Everyone" "Everyone" "Teen" ...
```

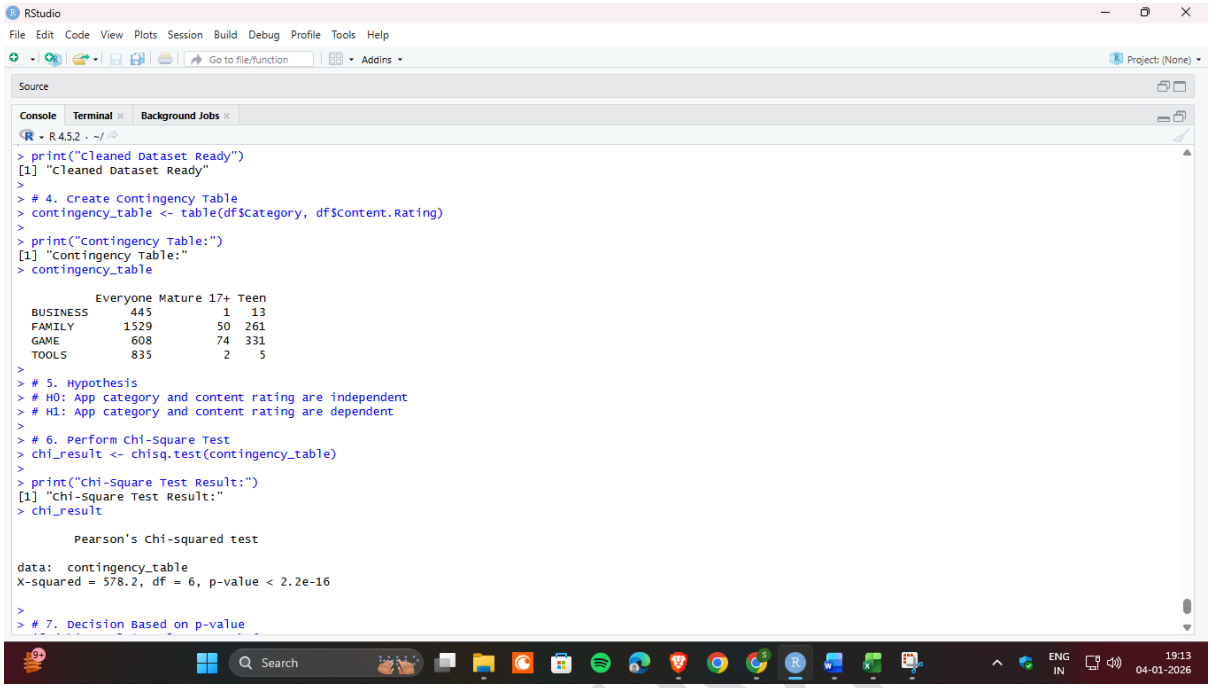


```
 $ Type          : chr  "Free" "Free" "Free" "Free" ...
 $ Price         : chr  "0" "0" "0" "0" ...
 $ Content.Rating: chr  "Everyone" "Everyone" "Everyone" "Teen" ...
 $ Genres        : chr  "Art & Design" "Art & Design;Pretend Play" "Art & Design" "Art & Design" ...
 $ Last.Updated  : chr  "7-Jan-18" "15-Jan-18" "1-Aug-18" "8-Jun-18" ...
 $ Current.Ver   : chr  "1.0.0" "2.0.0" "1.2.4" "Varies with device" ...
 $ Android.Ver   : chr  "4.0.3 and up" "4.0.3 and up" "4.0.3 and up" "4.2 and up" ...
> summary(df)
       X              App              Category            Rating         Reviews              Size              Installs           Type
 Min.   :    0   Length:10841       Length:10841       Min.   : 1.000   Length:10841       Length:10841       Length:10841       Length:10841
 1st Qu.: 2710   Class :character   Class :character   1st Qu.: 4.000   Class :character   Class :character   Class :character   Class :character
 Median : 5420   Mode  :character   Mode  :character   Median : 4.300   Mode  :character   Mode  :character   Mode  :character   Mode  :character
 Mean   : 5420                                         Mean   : 4.193
 3rd Qu.: 8130                                         3rd Qu.: 4.500
 Max.   :10840                                         Max.   :19.000
                                                       NA's   :1474
    Price           Content.Rating       Genres          Last.Updated       Current.Ver        Android.Ver
 Length:10841       Length:10841       Length:10841       Length:10841       Length:10841       Length:10841
 Class :character   Class :character   Class :character   Class :character   Class :character   Class :character
 Mode  :character   Mode  :character   Mode  :character   Mode  :character   Mode  :character   Mode  :character



>
> # 3. Select Required Categorical Columns
> df <- df %>%
+   filter(Category %in% c("GAME", "FAMILY", "TOOLS", "BUSINESS"),
+          Content.Rating %in% c("Everyone", "Teen", "Mature 17+"))
>
>
> print("Cleaned Dataset Ready")
[1] "Cleaned Dataset Ready"
>
```

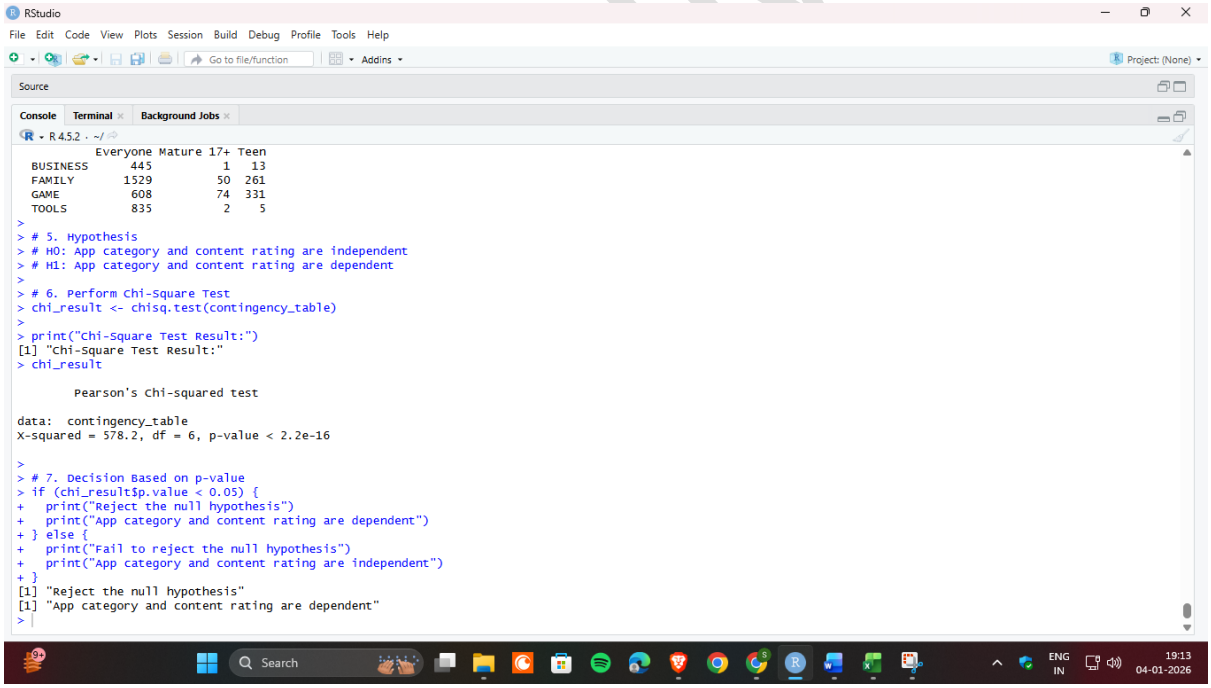
SUMEET YADAV
S124

SHETH L.U.J AND SIR M.V. COLLEGE

SUBJECT :- R PROGRAMMING



```
> print("Cleaned Dataset Ready")
[1] "Cleaned Dataset Ready"
>
> # 4. Create Contingency Table
> contingency_table <- table(df$Category, df$Content.Rating)
>
> print("Contingency Table:")
[1] "Contingency Table:"
> contingency_table

           Everyone Mature 17+ Teen
  BUSINESS      445          1   13
  FAMILY       1529         50  261
  GAME          608         74  331
  TOOLS         835          2    5
>
> # 5. Hypothesis
> # H0: App category and content rating are independent
> # H1: App category and content rating are dependent
>
> # 6. Perform Chi-Square Test
> chi_result <- chisq.test(contingency_table)
>
> print("Chi-Square Test Result:")
[1] "Chi-Square Test Result:"
> chi_result

        Pearson's Chi-squared test

data:  contingency_table
X-squared = 578.2, df = 6, p-value < 2.2e-16

>
> # 7. Decision Based on p-value
```



```
           Everyone Mature 17+ Teen
  BUSINESS      445          1   13
  FAMILY       1529         50  261
  GAME          608         74  331
  TOOLS         835          2    5
>
> # 5. Hypothesis
> # H0: App category and content rating are independent
> # H1: App category and content rating are dependent
>
> # 6. Perform Chi-Square Test
> chi_result <- chisq.test(contingency_table)
>
> print("Chi-Square Test Result:")
[1] "Chi-Square Test Result:"
> chi_result

        Pearson's Chi-squared test

data:  contingency_table
X-squared = 578.2, df = 6, p-value < 2.2e-16

>
> # 7. Decision Based on p-value
> if (chi_result$p.value < 0.05) {
+   print("Reject the null hypothesis")
+   print("App category and content rating are dependent")
+ } else {
+   print("Fail to reject the null hypothesis")
+   print("App category and content rating are independent")
+ }
[1] "Reject the null hypothesis"
[1] "App category and content rating are dependent"
>
```

SUMEET YADAV

S124