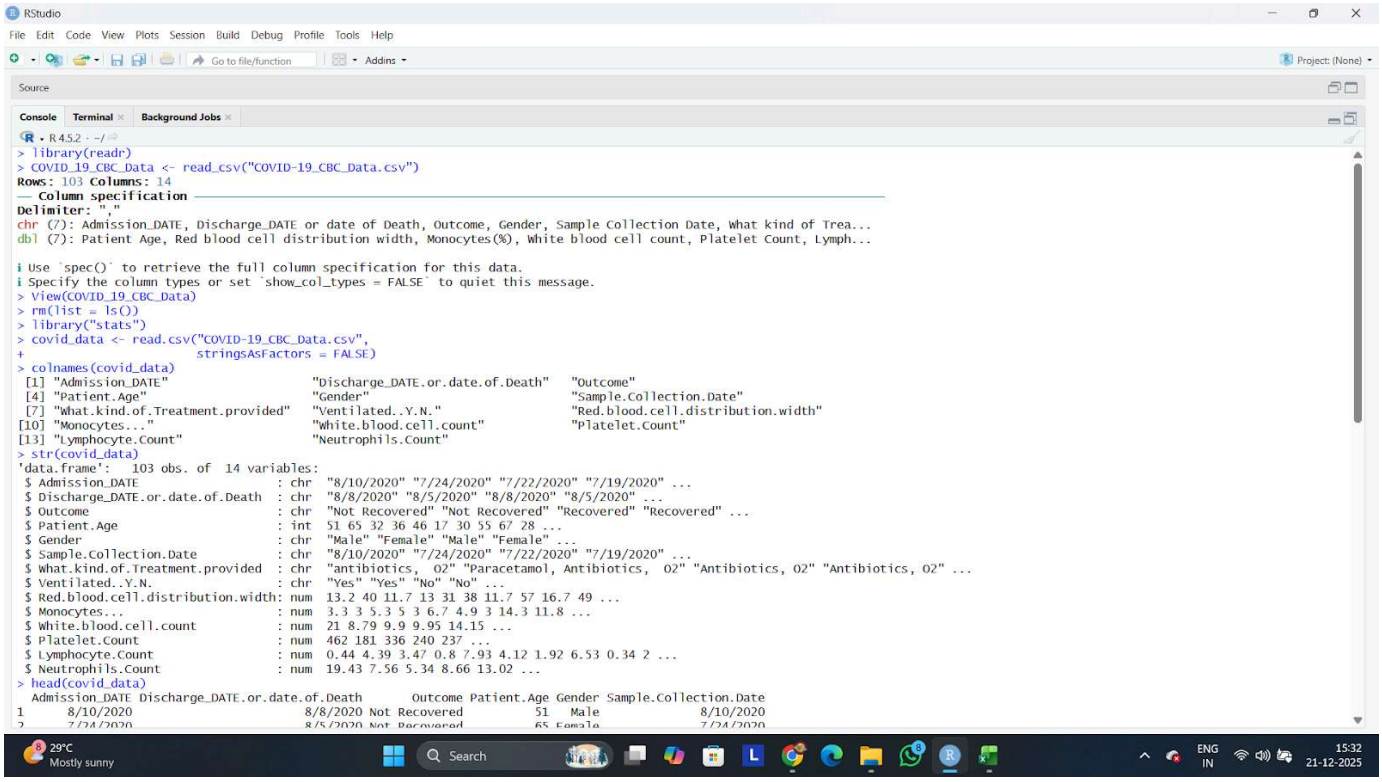
Sheth L.U.J & Sir M.V College Of Science
Subject :- Data Analysis With SAS/SPSS/R
Module 2 Practical No 9

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



```
> library(readr)
> COVID_19_CBC_Data <- read_csv("COVID-19_CBC_Data.csv")
Rows: 103 Columns: 14
── Column specification ────────────────────────────────────────────────────────
Delimiter: ","
chr (7): Admission_DATE, Discharge_DATE or date of Death, Outcome, Gender, Sample Collection Date, What kind of Trea...
dbl (7): Patient Age, Red blood cell distribution width, Monocytes(%), White blood cell count, Platelet Count, Lymph...

ℹ Use `spec()` to retrieve the full column specification for this data.
ℹ Specify the column types or set `show_col_types = FALSE` to quiet this message.
> View(COVID_19_CBC_Data)
> rm(list = ls())
> library("stats")
> covid_data <- read.csv("COVID-19_CBC_Data.csv",
+                        stringsAsFactors = FALSE)
> colnames(covid_data)
 [1] "Admission_DATE"                  "Discharge_DATE.or.date.of.Death"   "Outcome"
 [4] "Patient.Age"                     "Gender"                            "Sample.Collection.Date"
 [7] "What.kind.of.Treatment.provided" "Ventilated..Y.N."                  "Red.blood.cell.distribution.width"
[10] "Monocytes..."                    "White.blood.cell.count"            "Platelet.Count"
[13] "Lymphocyte.Count"                "Neutrophils.Count"
> str(covid_data)
'data.frame':   103 obs. of  14 variables:
 $ Admission_DATE                  : chr  "8/10/2020" "7/24/2020" "7/22/2020" "7/19/2020" ...
 $ Discharge_DATE.or.date.of.Death : chr  "8/8/2020" "8/5/2020" "8/8/2020" "8/5/2020" ...
 $ Outcome                         : chr  "Not Recovered" "Not Recovered" "Recovered" "Recovered" ...
 $ Patient.Age                     : int  51 65 32 36 46 17 30 55 67 28 ...
 $ Gender                          : chr  "Male" "Female" "Male" "Female" ...
 $ Sample.Collection.Date          : chr  "8/10/2020" "7/24/2020" "7/22/2020" "7/19/2020" ...
 $ What.kind.of.Treatment.provided : chr  "antibiotics,  O2" "Paracetamol, Antibiotics,  O2" "Antibiotics, O2" "Antibiotics, O2" ...
 $ Ventilated..Y.N.                : chr  "Yes" "Yes" "No" "No" ...
 $ Red.blood.cell.distribution.width: num  13.2 40 11.7 13 31 38 11.7 57 16.7 49 ...
 $ Monocytes...                    : num  3.3 3 5.3 5 3 6.7 4.9 3 14.3 11.8 ...
 $ White.blood.cell.count          : num  21 8.79 9.9 9.95 14.15 ...
 $ Platelet.Count                  : num  462 181 336 240 237 ...
 $ Lymphocyte.Count                : num  0.44 4.39 3.47 0.8 7.93 4.12 1.92 6.53 0.34 2 ...
 $ Neutrophils.Count               : num  19.43 7.56 5.34 8.66 13.02 ...
> head(covid_data)
  Admission_DATE Discharge_DATE.or.date.of.Death      Outcome Patient.Age Gender Sample.Collection.Date
1      8/10/2020                      8/8/2020 Not Recovered          51   Male              8/10/2020
2      7/24/2020                      8/5/2020 Not Recovered          65 Female              7/24/2020
```



```
3      7/22/2020                      8/8/2020     Recovered          32   Male              7/22/2020
4      7/19/2020                      8/5/2020     Recovered          36 Female              7/19/2020
5      7/18/2020                      8/5/2020     Recovered          46   Male              7/18/2020
6      7/17/2020                      8/7/2020 Not Recovered          17   Male              7/17/2020
  What.kind.of.Treatment.provided Ventilated..Y.N. Red.blood.cell.distribution.width Monocytes...
1              antibiotics,  O2              Yes                            13.2          3.3
2   Paracetamol, Antibiotics,  O2              Yes                            40.0          3.0
3               Antibiotics, O2               No                            11.7          5.3
4               Antibiotics, O2               No                            13.0          5.0
5   Paracetamol, Antibiotics,  O2               No                            31.0          3.0
6              antibiotics,  O2              Yes                            38.0          6.7
  White.blood.cell.count Platelet.Count Lymphocyte.Count Neutrophils.Count
1                  21.00         462.00             0.44             19.43
2                   8.79         180.66             4.39              7.56
3                   9.90         336.00             3.47              5.34
4                   9.95         240.10             0.80              8.66
5                  14.15         236.58             7.93             13.02
6                  13.30         249.00             4.12              8.15
> covid_data$Gender  <- as.factor(covid_data$Gender)
> covid_data$Outcome <- as.factor(covid_data$Outcome)
> contingency_table <- table(covid_data$Gender,
+                            covid_data$Outcome)
> # Display contingency table
> contingency_table

        Not Recovered Recovered
  Female            15        33
  Male              27        28
> chi_square_test <- chisq.test(contingency_table)
> # Display test results
> chi_square_test

        Pearson's Chi-squared test with Yates' continuity correction

data:  contingency_table
X-squared = 2.6799, df = 1, p-value = 0.1016

> chi_square_test$expected

        Not Recovered Recovered
```

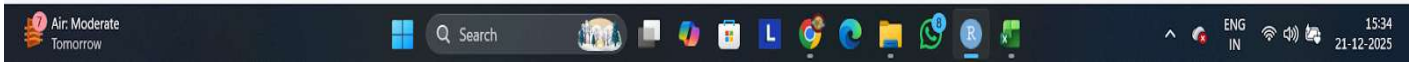
Name :- Vinod Mali

Roll No :- S092

Sheth L.U.J & Sir M.V College Of Science
Subject :- Data Analysis With SAS/SPSS/R
Module 2 Practical No 9

```
> chi_square_test$expected

        Not Recovered Recovered
Female      19.57282  28.42718
Male        22.42718  32.57282
>
>
```



Name :- Vinod Mali
Roll No :- S092