

MVLU COLLEGE R PRACTICAL 3 MODULE 2

Aim: -. Creating cross-tabulations and two-way tables using table() (R).

The screenshot displays the RStudio interface with the following components:

- Source Editor:** Contains R code for loading data, cleaning it, and creating cross-tabulations.
- Console:** Shows the output of the R code, including the data head and the final cross-tabulation results.
- Environment:** Lists the objects in the environment, including 'df', 'df\$runs', 'df\$bf', and 'df\$group'.
- Files:** Shows the file explorer with various files and folders.

```
# Load required library
library(dplyr)

# Read the csv file (user must select the file when running this)
df <- read.csv(file.choose(), stringsAsFactors = FALSE)

# Data cleaning fix (Robust Method using Column Index)
# This ensures '4s' (Column 5) and '6s' (Column 6) are numeric, preventing
# previous errors. It's safe to include this.
df[, 5] <- as.numeric(df[, 5]) # column '4s'
df[, 6] <- as.numeric(df[, 6]) # column '6s'

# View data
print("---- Data Head ----")
print(head(df))

# 1. Runs group: High / Low Runs (based on median Runs)
# We ensure the runs column is numeric before calculating the median
df$runs <- as.numeric(df$runs)
df$runs_group <- ifelse(df$runs > median(runs), 'High Runs', 'Low Runs')
# 2. Balls Faced group: High / Low BF (based on median BF)
# We ensure the bf column is numeric before calculating the median
df$bf <- as.numeric(df$bf)
df$bf_group <- ifelse(df$bf > median(bf), 'High BF', 'Low BF')
print("---- Cross-tabulation: Runs group vs. Balls Faced Group ----")
# Rows = Runs group, columns = BF group
print(table(df$runs_group, df$bf_group))

# Cross-tabulation with totals
print("---- Cross-tabulation: Runs group vs. Balls Faced Group ----")
print(margins(table(df$runs_group, df$bf_group)))
```

Output of the R code:

```
---- Data Head ----
  POS Player Runs BF four six Against Venue Match.date
1   1  KL Rahul  51  14   6   4   DC  IS Bindra Stadium 08-Apr-18
2   2  Ishan Kishan  62  17   5   6   KKR  Eden Gardens 09-May-18
3   3  Sunil Narine  50  17   4   5   RCB  Eden Gardens 08-Apr-18
4   4  Jos Buttler  67  18   4   7   DC  Arun Jaitley Stadium 02-May-18
5   5  Sam Billings  56  21   2   5   KKR  Chidambaram 10-Apr-18
6   6  KL Rahul  66  22   2   7   KKR  Holkar Cricket Stadium 12-May-18

# Cross-tabulation: Runs group vs. Balls Faced Group
# Rows = Runs group, columns = BF group
table(df$runs_group, df$bf_group)
      High BF Low BF
High Runs  23    30
Low Runs   27    53
Sum        50    83
```

Hariprasad Vishwakarma
S126
SYCS