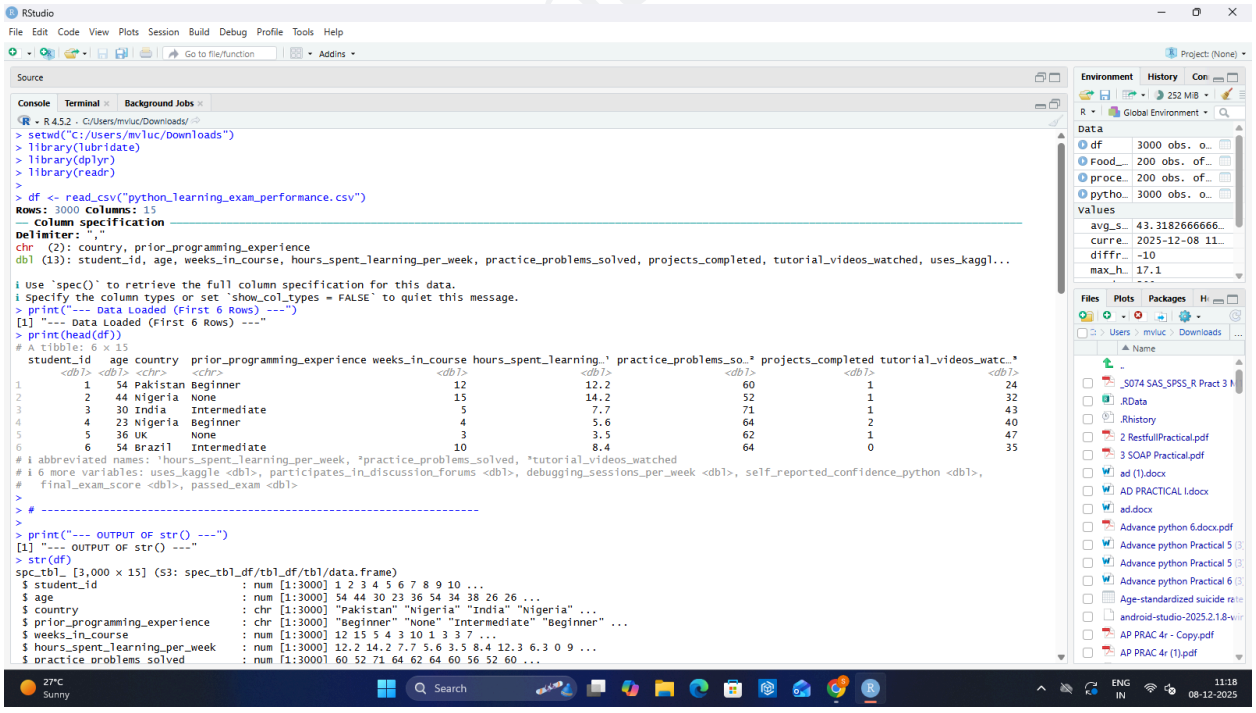
SHETH L.U.J AND SIR M.V. COLLEGE
SUBJECT NAME: DATA ANALYSIS WITH SAS/SPSS/R

Module 1 Practical 15

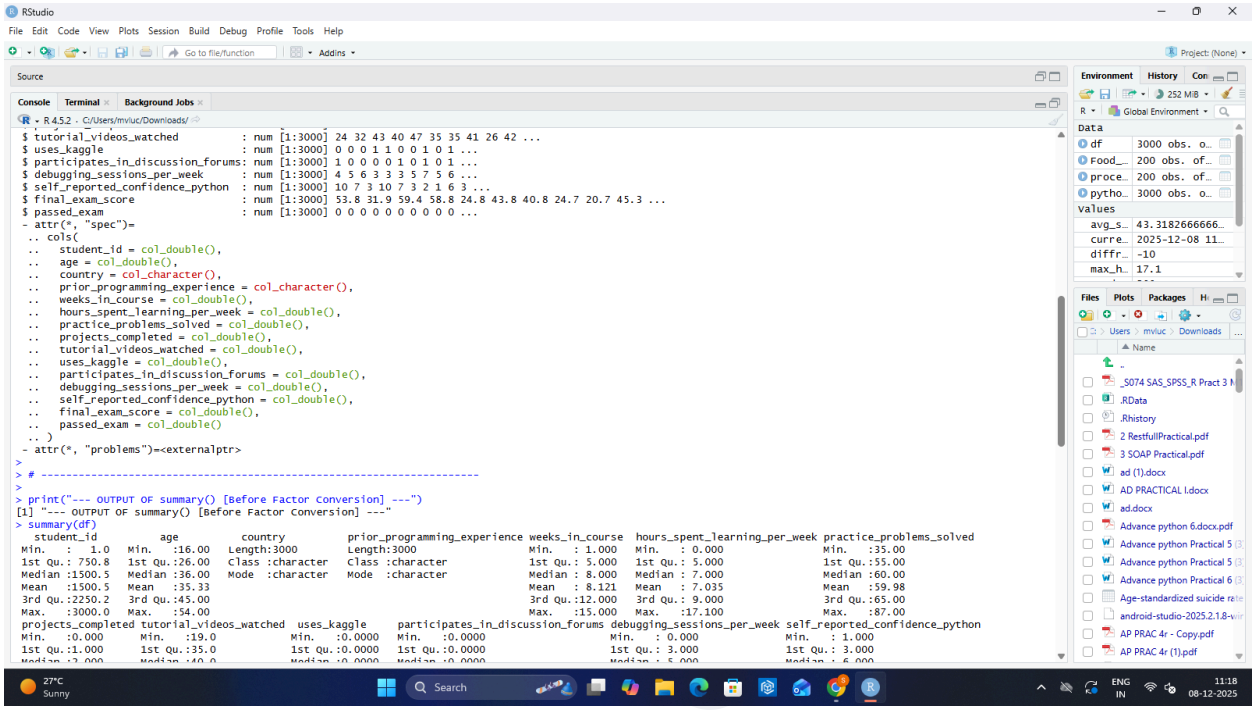
Aim: Generating basic summaries using str() or summary() (R).

OUTPUT:

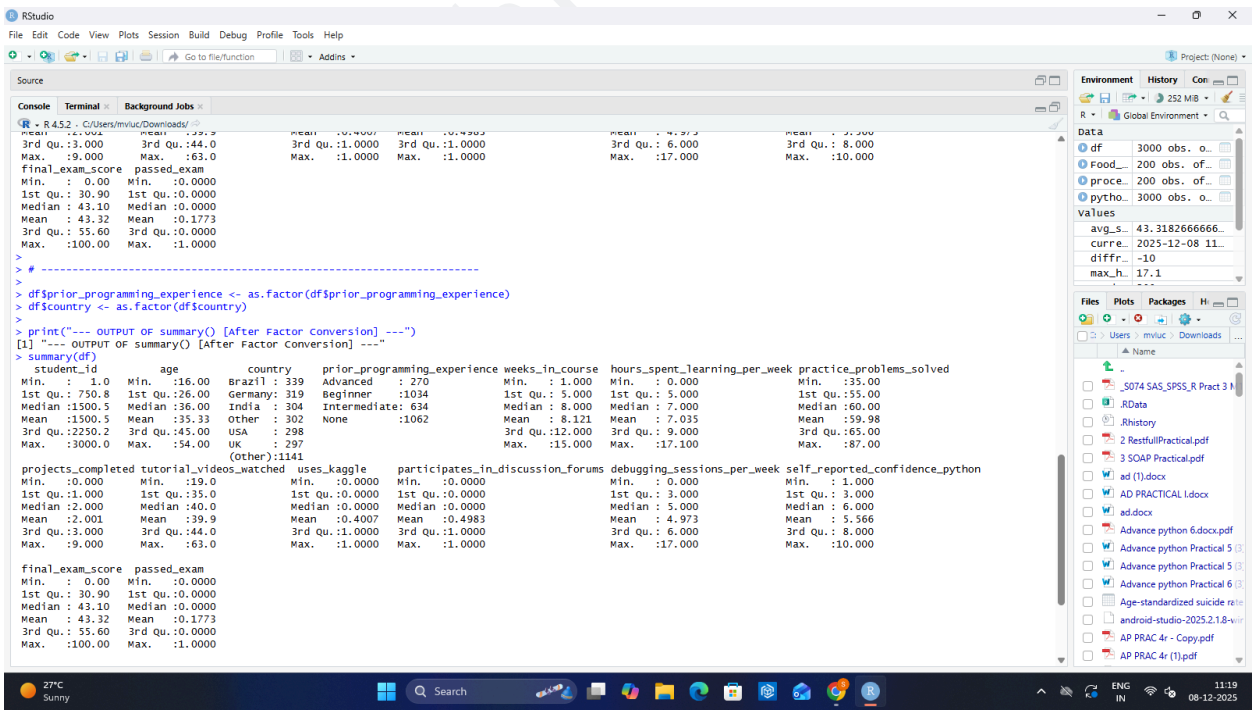


Krishna Nandkumar Chaudhary
S074

SHETH L.U.J AND SIR M.V. COLLEGE
SUBJECT NAME: DATA ANALYSIS WITH SAS/SPSS/R



```
 $ tutorial_videos_watched          : num [1:3000] 24 32 43 40 47 35 35 41 26 42 ...
 $ uses_kaggle                      : num [1:3000] 0 0 0 1 1 0 0 1 0 1 ...
 $ participates_in_discussion_forums: num [1:3000] 1 0 0 0 0 1 0 1 0 1 ...
 $ debugging_sessions_per_week      : num [1:3000] 4 5 6 3 3 3 5 7 5 6 ...
 $ self_reported_confidence_python  : num [1:3000] 10 7 3 10 7 3 2 1 6 3 ...
 $ final_exam_score                 : num [1:3000] 53.8 31.9 59.4 58.8 24.8 43.8 40.8 24.7 20.7 45.3 ...
 $ passed_exam                      : num [1:3000] 0 0 0 0 0 0 0 0 0 0 ...
 - attr(*, "spec")=
  .. cols(
  ..   student_id = col_double(),
  ..   age = col_double(),
  ..   country = col_character(),
  ..   prior_programming_experience = col_character(),
  ..   weeks_in_course = col_double(),
  ..   hours_spent_learning_per_week = col_double(),
  ..   practice_problems_solved = col_double(),
  ..   projects_completed = col_double(),
  ..   tutorial_videos_watched = col_double(),
  ..   uses_kaggle = col_double(),
  ..   participates_in_discussion_forums = col_double(),
  ..   debugging_sessions_per_week = col_double(),
  ..   self_reported_confidence_python = col_double(),
  ..   final_exam_score = col_double(),
  ..   passed_exam = col_double()
  .. )
 - attr(*, "problems")=<externalptr>
>
> # ------------------------------------------------------------------
>
> print("--- OUTPUT OF summary() [Before Factor Conversion] ---")
[1] "--- OUTPUT OF summary() [Before Factor Conversion] ---"
> summary(df)
   student_id          age          country          prior_programming_experience weeks_in_course  hours_spent_learning_per_week practice_problems_solved
 Min.   :   1.0   Min.   :16.00   Length:3000        Length:3000                   Min.   : 1.000   Min.   : 0.000                Min.   :35.00
 1st Qu.: 750.8   1st Qu.:26.00   Class :character   Class :character              1st Qu.: 5.000   1st Qu.: 5.000                1st Qu.:55.00
 Median :1500.5   Median :36.00   Mode  :character   Mode  :character              Median : 8.000   Median : 7.000                Median :60.00
 Mean   :1500.5   Mean   :35.33                                                   Mean   : 8.121   Mean   : 7.035                Mean   :59.98
 3rd Qu.:2250.2   3rd Qu.:45.00                                                   3rd Qu.:12.000   3rd Qu.: 9.000                3rd Qu.:65.00
 Max.   :3000.0   Max.   :54.00                                                   Max.   :15.000   Max.   :17.100                Max.   :87.00
 projects_completed tutorial_videos_watched  uses_kaggle     participates_in_discussion_forums debugging_sessions_per_week self_reported_confidence_python
 Min.   :0.000      Min.   :19.0            Min.   :0.0000   Min.   :0.0000                    Min.   : 0.000              Min.   : 1.000
 1st Qu.:1.000      1st Qu.:35.0            1st Qu.:0.0000   1st Qu.:0.0000                    1st Qu.: 3.000              1st Qu.: 3.000
```

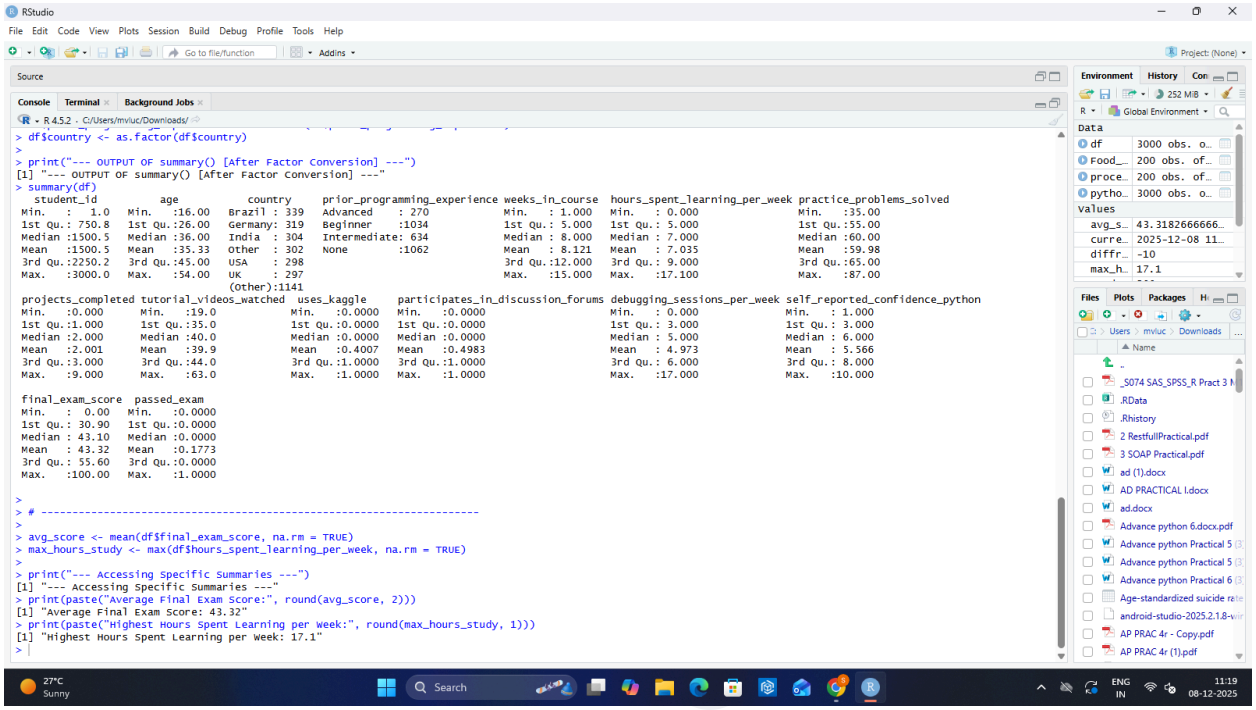


```
 3rd Qu.:3.000      3rd Qu.:44.0            3rd Qu.:1.0000   3rd Qu.:1.0000                    3rd Qu.: 6.000              3rd Qu.: 8.000
 Max.   :9.000      Max.   :63.0            Max.   :1.0000   Max.   :1.0000                    Max.   :17.000              Max.   :10.000
 final_exam_score  passed_exam
 Min.   :  0.00   Min.   :0.0000
 1st Qu.: 30.90   1st Qu.:0.0000
 Median : 43.10   Median :0.0000
 Mean   : 43.32   Mean   :0.1773
 3rd Qu.: 55.60   3rd Qu.:0.0000
 Max.   :100.00   Max.   :1.0000
>
> # ------------------------------------------------------------------
>
> df$prior_programming_experience <- as.factor(df$prior_programming_experience)
> df$country <- as.factor(df$country)
>
> print("--- OUTPUT OF summary() [After Factor Conversion] ---")
[1] "--- OUTPUT OF summary() [After Factor Conversion] ---"
> summary(df)
   student_id          age             country     prior_programming_experience weeks_in_course  hours_spent_learning_per_week practice_problems_solved
 Min.   :   1.0   Min.   :16.00   Brazil : 339   Advanced    : 270              Min.   : 1.000   Min.   : 0.000                Min.   :35.00
 1st Qu.: 750.8   1st Qu.:26.00   Germany: 319   Beginner    :1034              1st Qu.: 5.000   1st Qu.: 5.000                1st Qu.:55.00
 Median :1500.5   Median :36.00   India  : 304   Intermediate: 634              Median : 8.000   Median : 7.000                Median :60.00
 Mean   :1500.5   Mean   :35.33   Other  : 302   None        :1062              Mean   : 8.121   Mean   : 7.035                Mean   :59.98
 3rd Qu.:2250.2   3rd Qu.:45.00   USA    : 298                                 3rd Qu.:12.000   3rd Qu.: 9.000                3rd Qu.:65.00
 Max.   :3000.0   Max.   :54.00   UK     : 297                                 Max.   :15.000   Max.   :17.100                Max.   :87.00
                                  (Other):1141
 projects_completed tutorial_videos_watched  uses_kaggle     participates_in_discussion_forums debugging_sessions_per_week self_reported_confidence_python
 Min.   :0.000      Min.   :19.0            Min.   :0.0000   Min.   :0.0000                    Min.   : 0.000              Min.   : 1.000
 1st Qu.:1.000      1st Qu.:35.0            1st Qu.:0.0000   1st Qu.:0.0000                    1st Qu.: 3.000              1st Qu.: 3.000
 Median :2.000      Median :40.0            Median :0.0000   Median :0.0000                    Median : 5.000              Median : 6.000
 Mean   :2.001      Mean   :39.9            Mean   :0.4007   Mean   :0.4983                    Mean   : 4.973              Mean   : 5.566
 3rd Qu.:3.000      3rd Qu.:44.0            3rd Qu.:1.0000   3rd Qu.:1.0000                    3rd Qu.: 6.000              3rd Qu.: 8.000
 Max.   :9.000      Max.   :63.0            Max.   :1.0000   Max.   :1.0000                    Max.   :17.000              Max.   :10.000

 final_exam_score  passed_exam
 Min.   :  0.00   Min.   :0.0000
 1st Qu.: 30.90   1st Qu.:0.0000
 Median : 43.10   Median :0.0000
 Mean   : 43.32   Mean   :0.1773
 3rd Qu.: 55.60   3rd Qu.:0.0000
 Max.   :100.00   Max.   :1.0000
```

Krishna Nandkumar Chaudhary
S074

SHETH L.U.J AND SIR M.V. COLLEGE
SUBJECT NAME: DATA ANALYSIS WITH SAS/SPSS/R



Krishna Nandkumar Chaudhary
S074