Sample Outputs

```
> names(df)
 [1] "sl_no"          "gender"        "ssc_p"         "ssc_b"         "hsc_p"         "hsc_b"
 [7] "hsc_s"          "degree_p"      "degree_t"      "workex"        "etest_p"       "specialisation"
[13] "mba_p"          "status"        "salary"
```

Figure 1: Column names in the Dataframe before renaming them.

```
> names(df)
 [1] "serial_number"                          "gender"
 [3] "secondary_education_percentage"         "secondary_education_board"
 [5] "higher_secondary_education_percentage" "higher_education_board"
 [7] "higher_secondary_specialization"        "degree_percentage"
 [9] "degree_field"                           "work_experience"
[11] "employability_test_percentage"          "mba_specialization"
[13] "mba_percentage"                         "placement_status"
[15] "salary"
```

Figure 2: Dataframe column names after renaming the columns.

```
> head(df)
  serial_number gender secondary_education_percentage secondary_education_board
1             1      M                         67.00                     Others
2             2      M                         79.33                    Central
3             3      M                         65.00                    Central
4             4      M                         56.00                    Central
5             5      M                         85.80                    Central
6             6      M                         55.00                     Others
  higher_secondary_education_percentage higher_education_board higher_secondary_specialization
1                                 91.00                 Others                        Commerce
2                                 78.33                 Others                         Science
3                                 68.00                Central                            Arts
4                                 52.00                Central                         Science
5                                 73.60                Central                        Commerce
6                                 49.80                 Others                         Science
  degree_percentage degree_field work_experience employability_test_percentage mba_specialization
1             58.00     Sci&Tech              No                          55.0             Mkt&HR
2             77.48     Sci&Tech             Yes                          86.5            Mkt&Fin
3             64.00    Comm&Mgmt              No                          75.0            Mkt&Fin
4             52.00     Sci&Tech              No                          66.0             Mkt&HR
5             73.30    Comm&Mgmt              No                          96.8            Mkt&Fin
6             67.25     Sci&Tech             Yes                          55.0            Mkt&Fin
  mba_percentage placement_status salary
1          58.80           Placed 270000
2          66.28           Placed 200000
3          57.80           Placed 250000
4          59.43       Not Placed     NA
5          55.50           Placed 425000
6          51.58       Not Placed     NA
```

Figure 3: The top 6 rows in the dataframe.

```
> glimpse(df)
Rows: 215
Columns: 15
$ serial_number                      <int> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18,…
$ gender                             <chr> "M", "M", "M", "M", "M", "M", "F", "M", "M", "M", "M", "M", "F…
$ secondary_education_percentage     <dbl> 67.00, 79.33, 65.00, 56.00, 85.80, 55.00, 46.00, 82.00, 73.00,…
$ secondary_education_board          <chr> "Others", "Central", "Central", "Central", "Central", "Others"…
$ higher_secondary_education_percentage <dbl> 91.00, 78.33, 68.00, 52.00, 73.60, 49.80, 49.20, 64.00, 79.00,…
$ higher_education_board             <chr> "Others", "Others", "Central", "Central", "Central", "Others",…
$ higher_secondary_specialization    <chr> "Commerce", "Science", "Arts", "Science", "Commerce", "Science…
$ degree_percentage                  <dbl> 58.00, 77.48, 64.00, 52.00, 73.30, 67.25, 79.00, 66.00, 72.00,…
$ degree_field                       <chr> "Sci&Tech", "Sci&Tech", "Comm&Mgmt", "Sci&Tech", "Comm&Mgmt", …
$ work_experience                    <chr> "No", "Yes", "No", "No", "No", "Yes", "No", "Yes", "No", "No",…
$ employability_test_percentage      <dbl> 55.00, 86.50, 75.00, 66.00, 96.80, 55.00, 74.28, 67.00, 91.34,…
$ mba_specialization                 <chr> "Mkt&HR", "Mkt&Fin", "Mkt&Fin", "Mkt&HR", "Mkt&Fin", "Mkt&Fin"…
$ mba_percentage                     <dbl> 58.80, 66.28, 57.80, 59.43, 55.50, 51.58, 53.29, 62.14, 61.29,…
$ placement_status                   <chr> "Placed", "Placed", "Placed", "Not Placed", "Placed", "Not Pla…
$ salary                             <int> 270000, 200000, 250000, NA, 425000, NA, NA, 252000, 231000, NA…
```

Figure 4: Output for glimpse(dataframe).

```
> summary(df)
 serial_number      gender          secondary_education_percentage secondary_education_board
 Min.   :  1.0   Length:215         Min.   :40.89                  Length:215
 1st Qu.: 54.5   Class :character   1st Qu.:60.60                  Class :character
 Median :108.0   Mode  :character   Median :67.00                  Mode  :character
 Mean   :108.0                      Mean   :67.30
 3rd Qu.:161.5                      3rd Qu.:75.70
 Max.   :215.0                      Max.   :89.40

 higher_secondary_education_percentage higher_education_board higher_secondary_specialization
 Min.   :37.00                         Length:215             Length:215
 1st Qu.:60.90                         Class :character       Class :character
 Median :65.00                         Mode  :character       Mode  :character
 Mean   :66.33
 3rd Qu.:73.00
 Max.   :97.70

 degree_percentage degree_field       work_experience    employability_test_percentage mba_specialization
 Min.   :50.00     Length:215         Length:215         Min.   :50.0                  Length:215
 1st Qu.:61.00     Class :character   Class :character   1st Qu.:60.0                  Class :character
 Median :66.00     Mode  :character   Mode  :character   Median :71.0                  Mode  :character
 Mean   :66.37                                           Mean   :72.1
 3rd Qu.:72.00                                           3rd Qu.:83.5
 Max.   :91.00                                           Max.   :98.0

 mba_percentage  placement_status       salary
 Min.   :51.21   Length:215         Min.   :200000
 1st Qu.:57.95   Class :character   1st Qu.:240000
 Median :62.00   Mode  :character   Median :265000
 Mean   :62.28                      Mean   :288655
 3rd Qu.:66.25                      3rd Qu.:300000
 Max.   :77.89                      Max.   :940000
                                    NA's   :67
```

Figure 5: Output for summary(dataframe).

```
> #convert binary values in placement_status column into 0 or 1
> dfCopy$placement_status <- ifelse(dfCopy$placement_status == "Placed", 1,0)
>
> #calculate percentage students who got a placement
> dfCopy %>% summarize(total = n(), percent_placed = mean(placement_status == 1))
  total percent_placed
1   215      0.6883721
>
> #calculate percentage students who got a placement based on MBA specialization
> dfCopy %>% group_by(mba_specialization) %>% summarize(total = n(), percent_placed = mean(placement_status == 1))
# A tibble: 2 × 3
  mba_specialization total percent_placed
  <chr>              <int>          <dbl>
1 Mkt&Fin              120          0.792
2 Mkt&HR                95          0.558
```

Figure 6: Output for percentages of students who secured a placement in the workforce and percentages grouped by MBA Specialization.

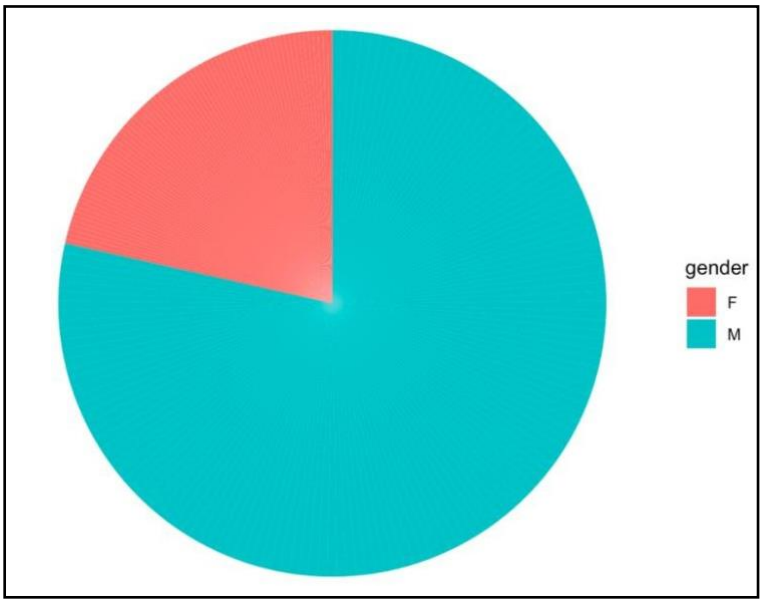

Figure 7: A pie chart to visualize the distribution of two genders in the dataset.

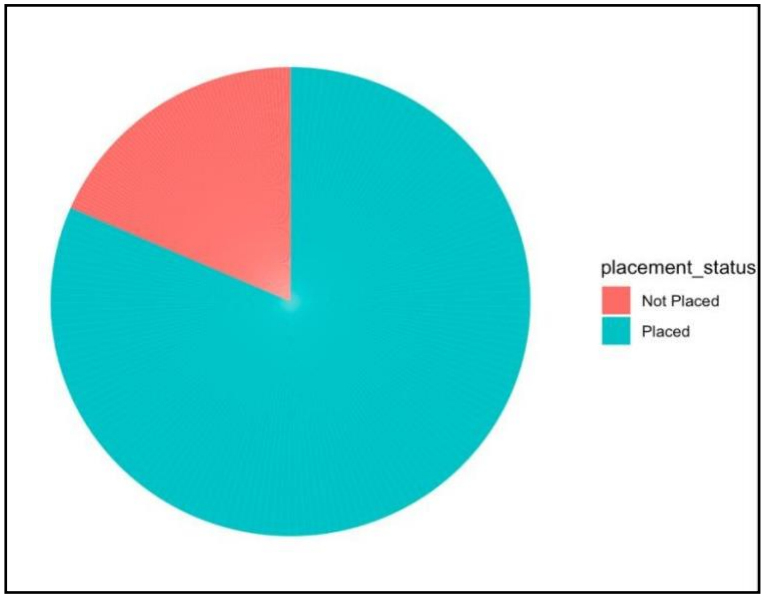

Figure 8: A pie chart to visualize the distribution of placement status of the students in the dataset.

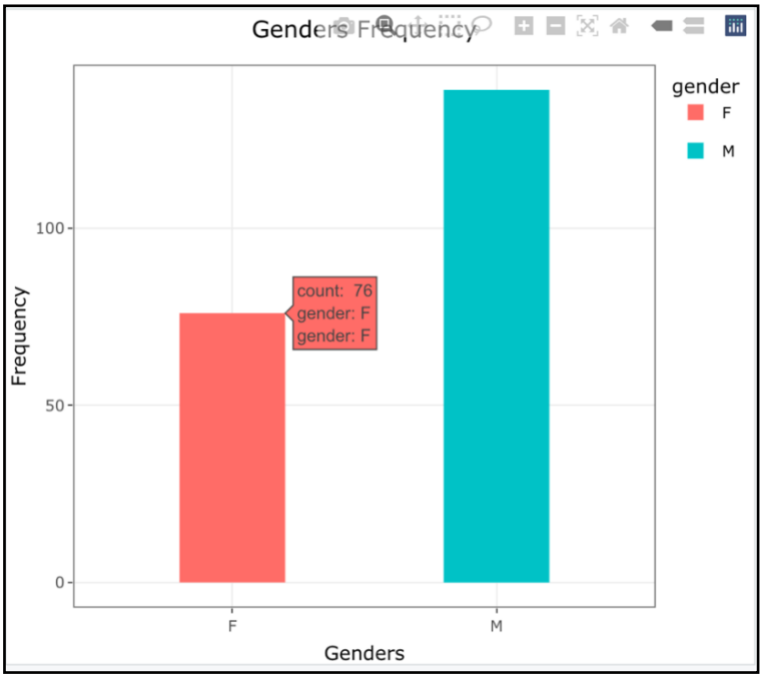

Figure 9: A bar plot to visualize the distribution of two genders in the dataset.

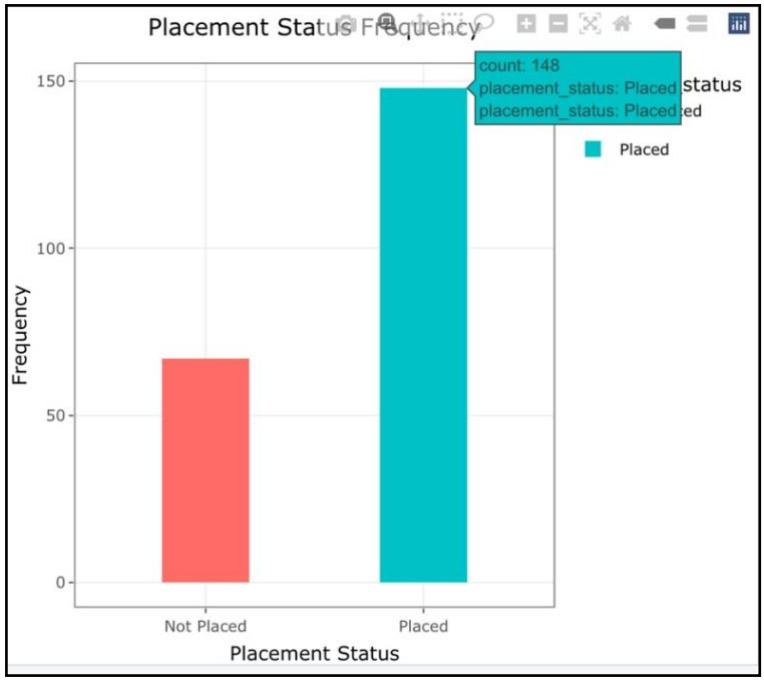

Figure 10: A bar plot to visualize the distribution of placement status of the students in the dataset.

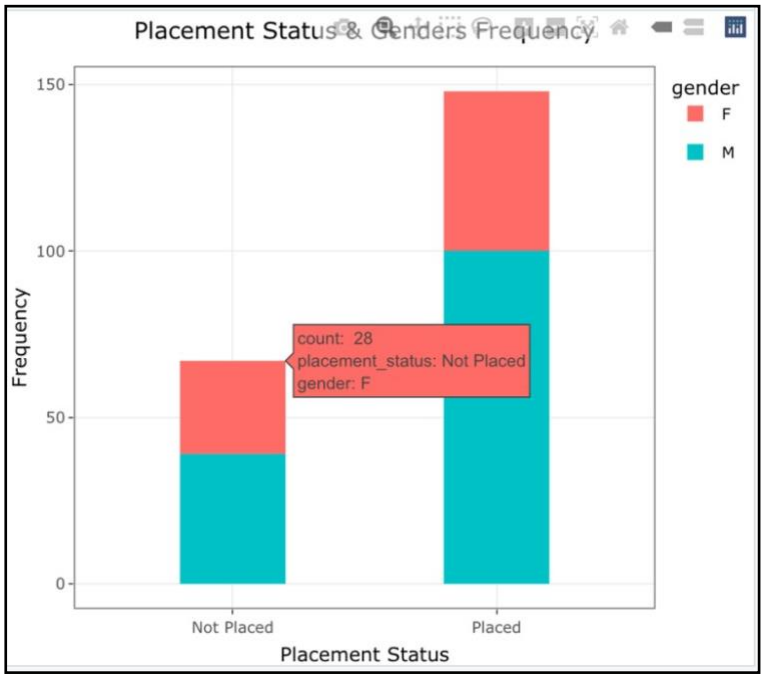

Figure 11: A stacked bar plot to visualize the distribution of two genders based on placement status.

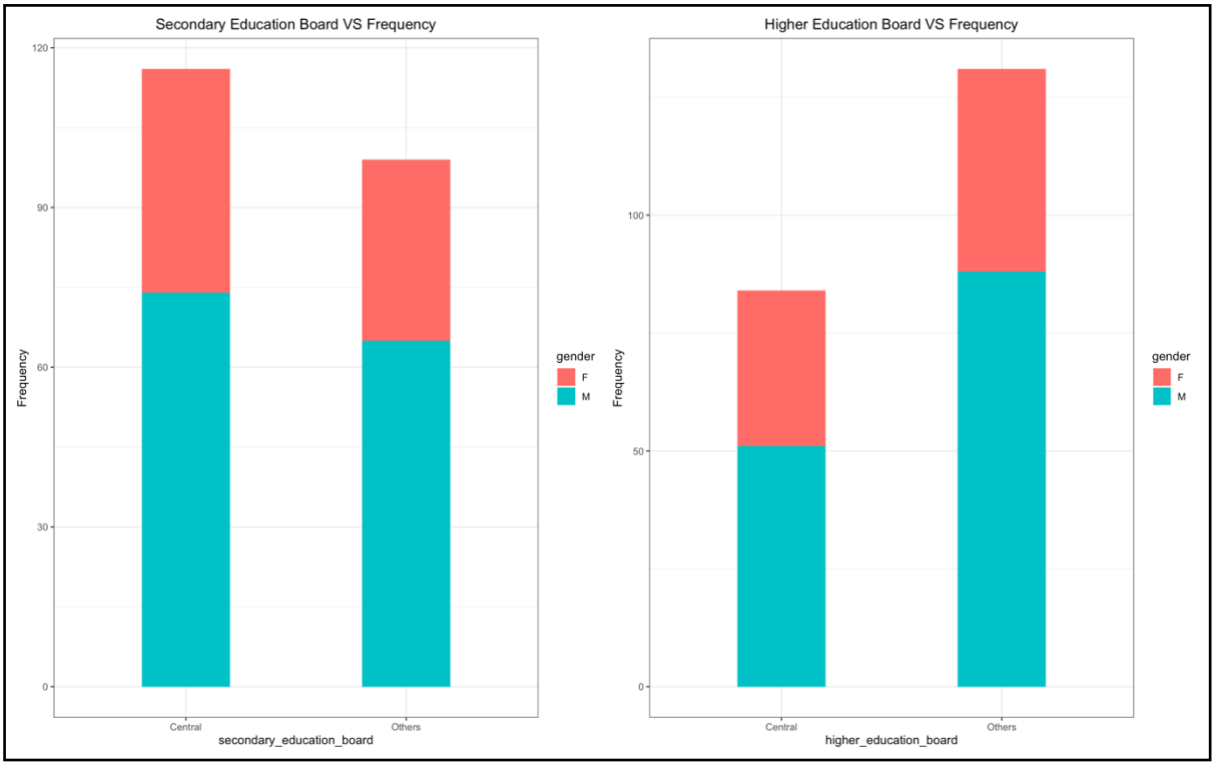

Figure 12: A stacked bar plot to visualize the distribution of two genders based on secondary and higher education board.

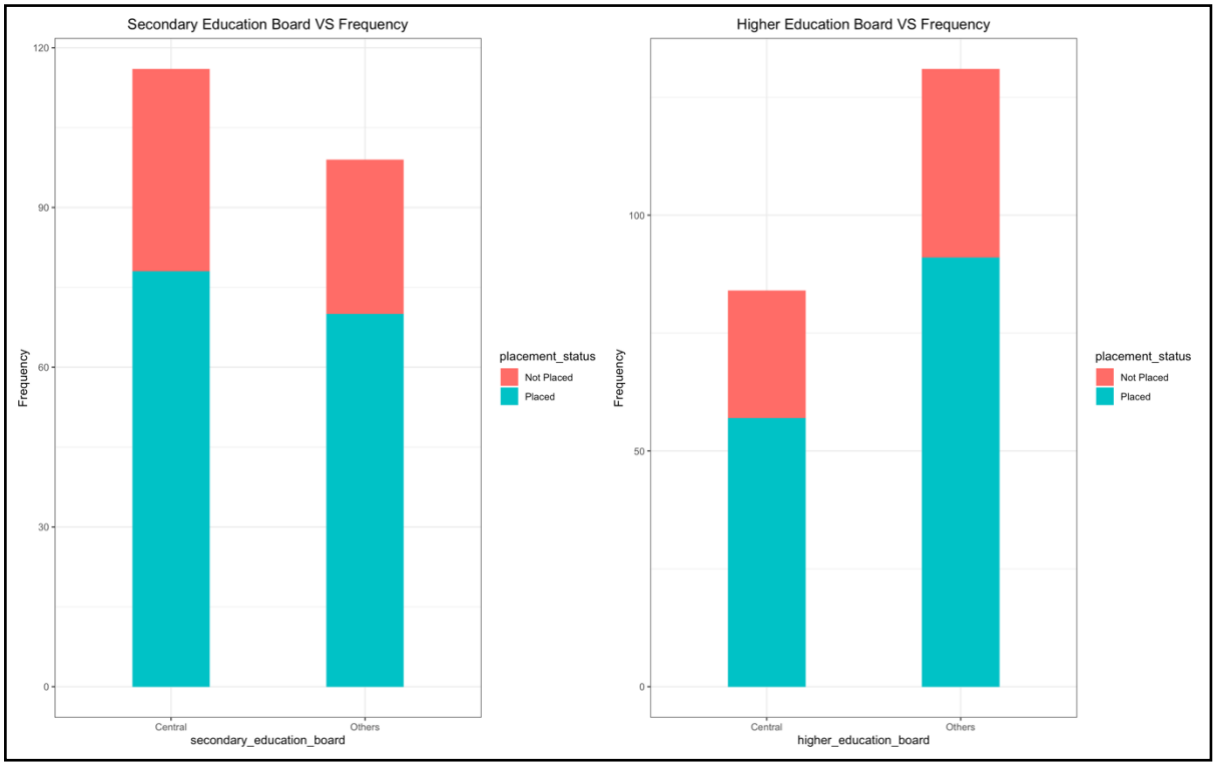

Figure 13: A stacked bar plot to visualize the distribution of placement status based on secondary and higher education board.

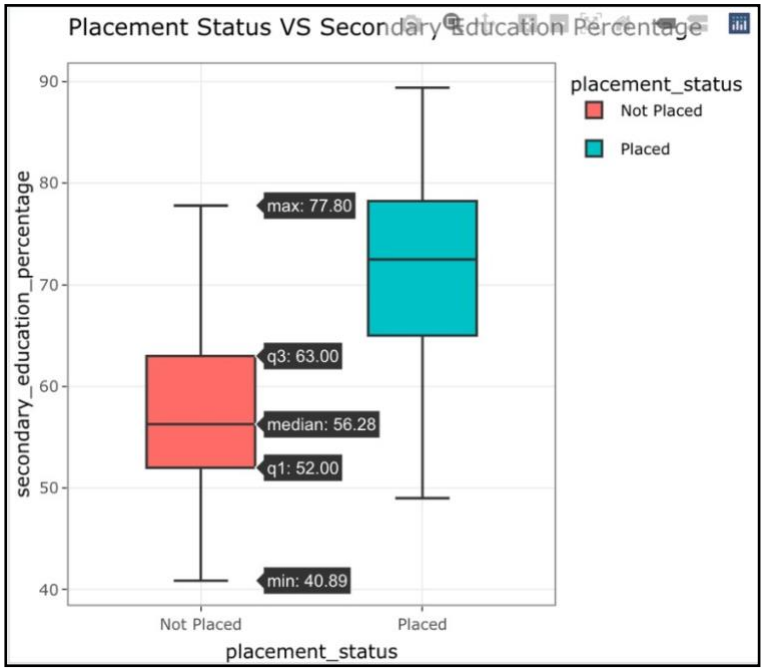

Figure 14: An interactive box plot to visualize the distribution of marks vs placement status.

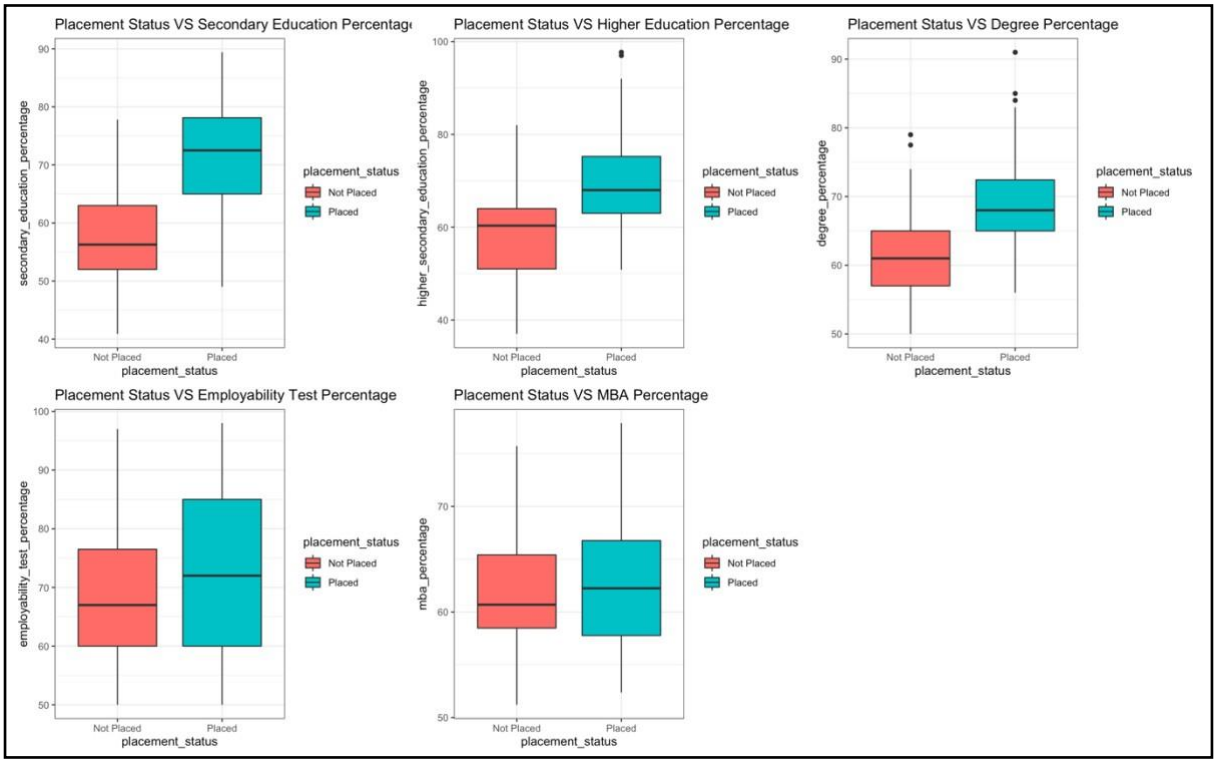

Figure 15: A box plot to visualize the distribution of marks vs placement status.

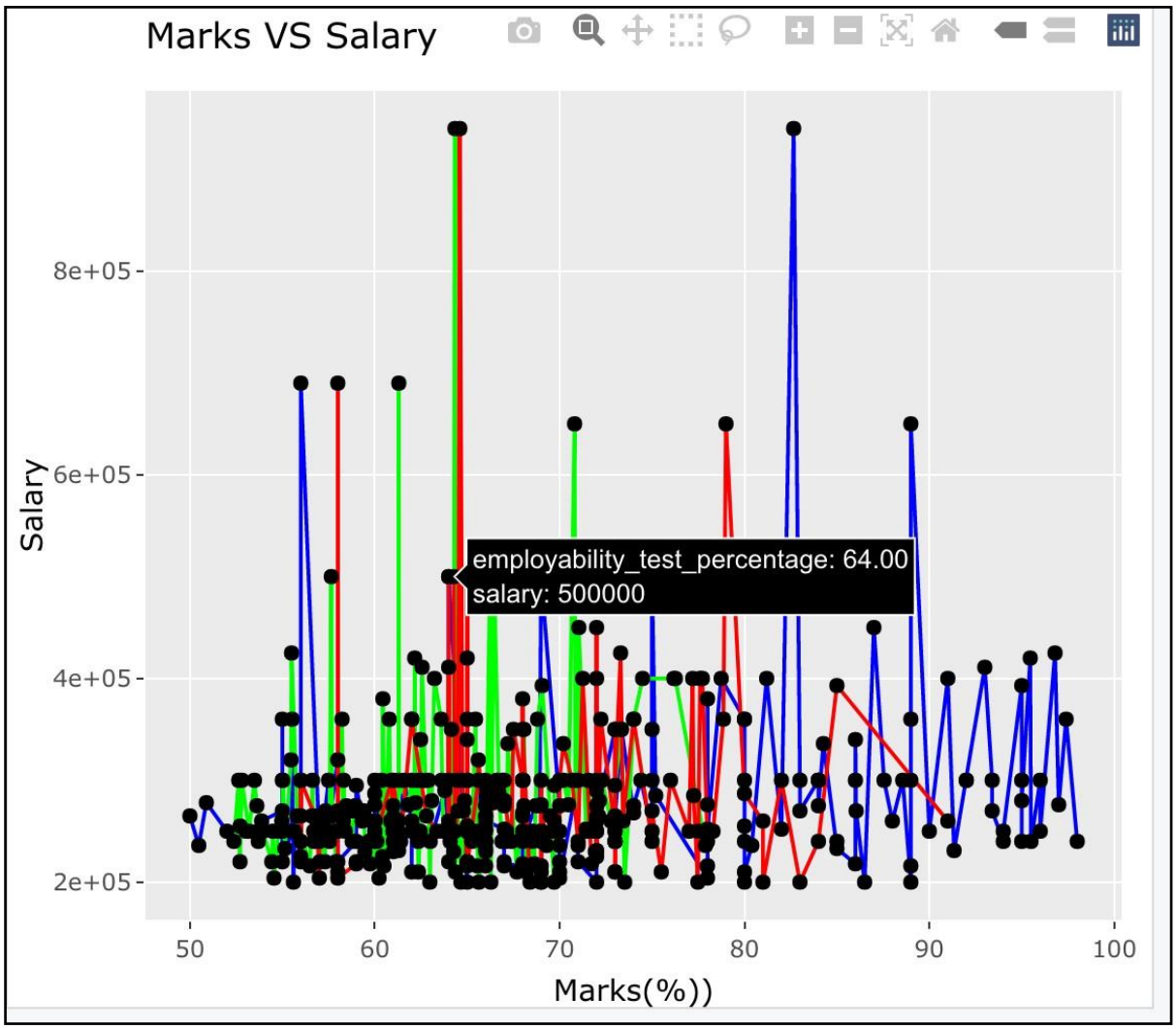

Figure 16: An interactive line graph to visualize marks vs salary.

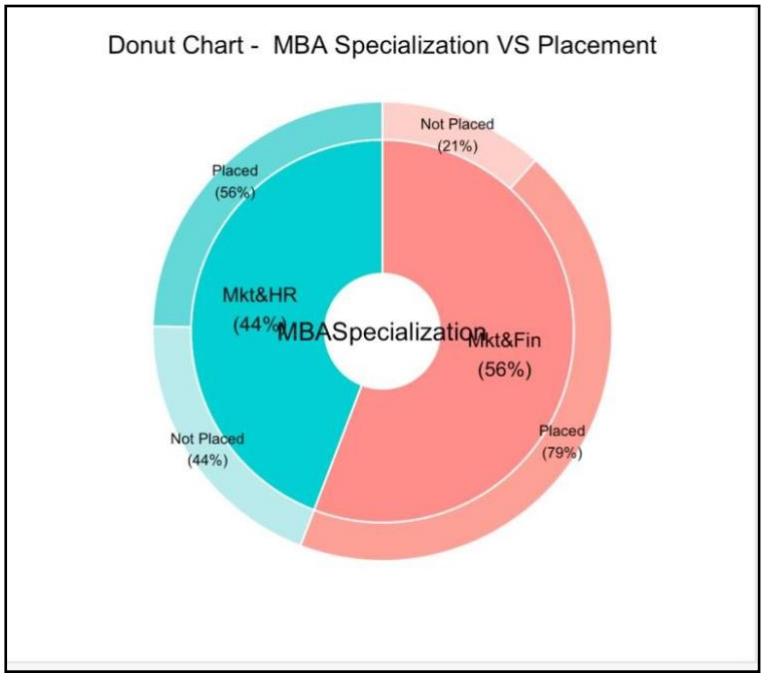

Figure 17: A donut chart to visualize MBA specialization vs placement status.

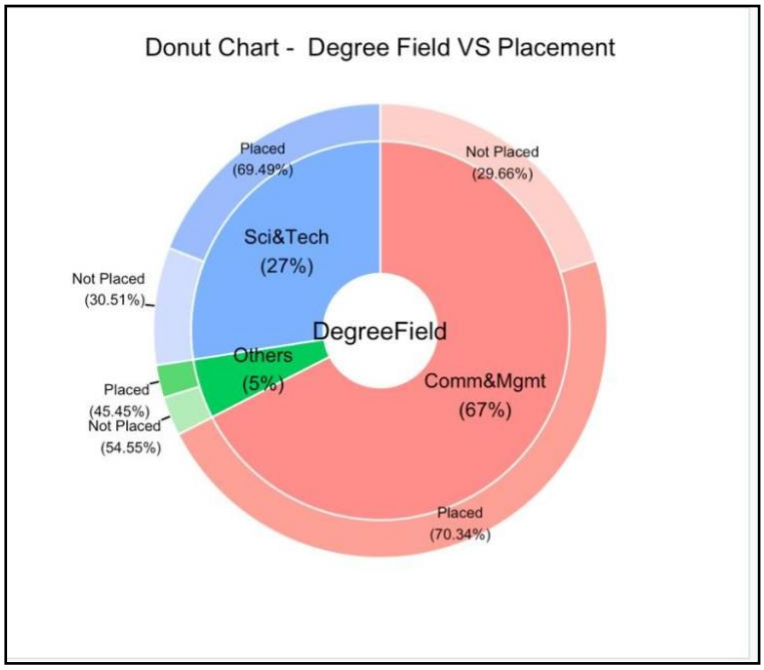

Figure 18: A donut chart to visualize degree field vs placement status.

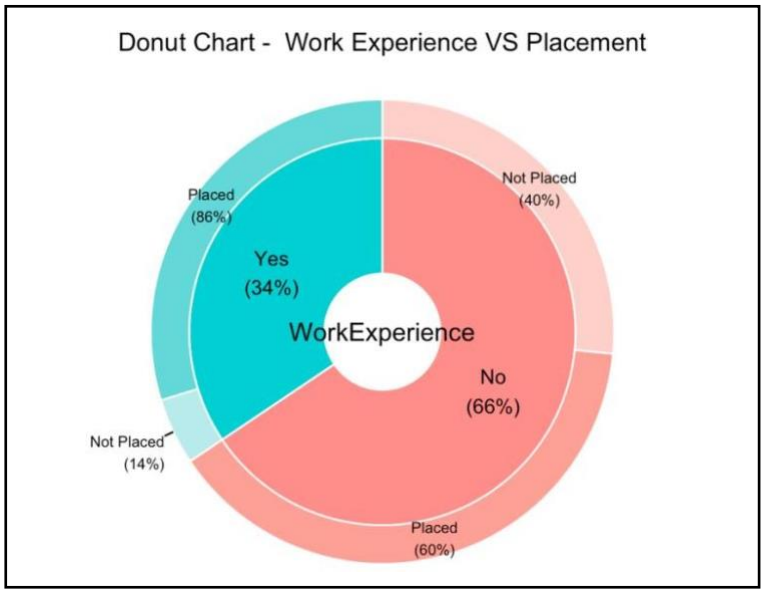

Figure 19: A donut chart to visualize work experience vs placement status.

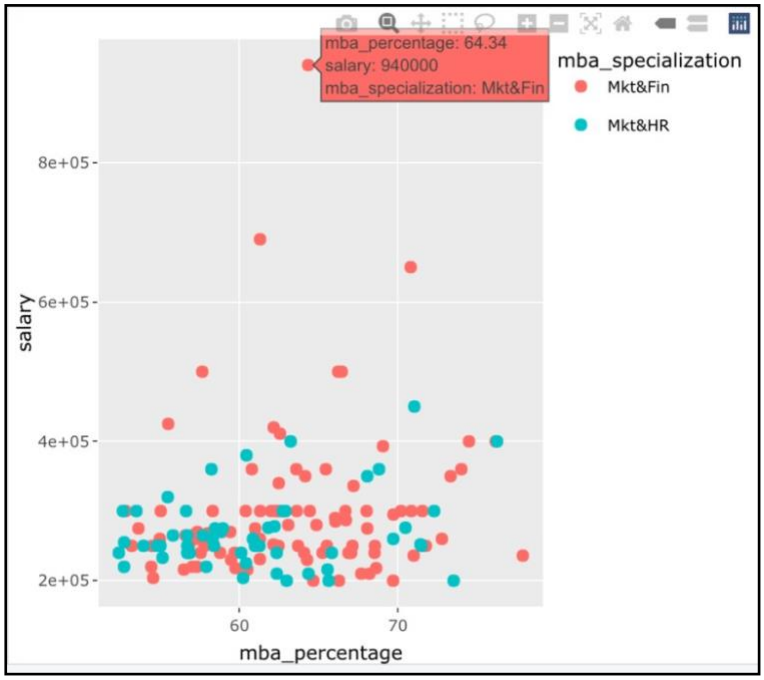

Figure 20: An interactive scatter plot to visualize MBA percentage vs salary according to its MBA specialization.

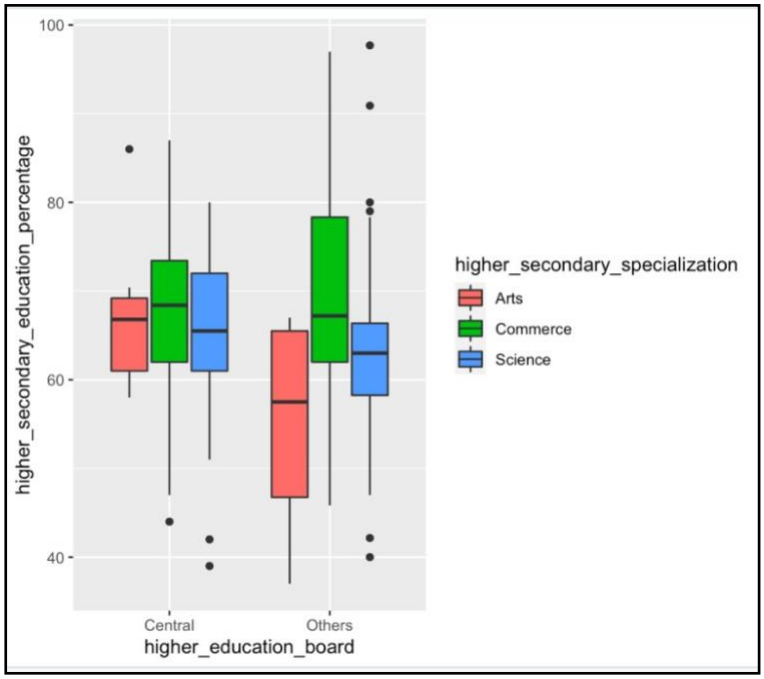

Figure 21: A box plot to visualize higher education board vs higher secondary education percentage according to the higher secondary specialization.

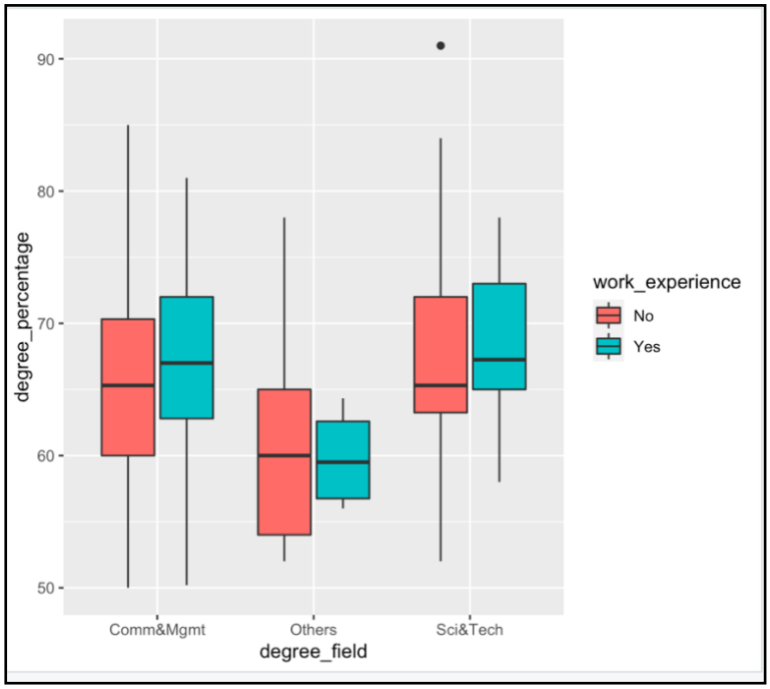

Figure 22: A box plot to visualize degree field vs degree percentage according to work experience.