

Foundations of Data Science (UCS548)

Dashboard Submission

CAMPUS PLACEMENT ANALYSIS



Submitted By:

Jigisha Chopra
102003650
3COE26

Submitted To:

Dr. Sharad Saxena

July 2022 – December 2022

About the dataset

The dataset consists of 215 rows representing students that completed an MBA course at an Indian University. In this report we analyse what factors influence the chances of finding good jobs after graduating from an MBA at an Indian university. The columns are :

SerialNumber: An increasing number which represents the row id.

Gender: The gender of the student (M / F).

LowerSecondarySchollGrade%: The average of grades in middle school (in %).

LowerSecondarySchollBoard: Schools in India are separated in “Boards”, depending on their region (Central / Others).

HigherSecondarySchollGrade%: The average of grades in high school (in %).

HigherSecondarySchollBoard: Same as middle school board, but for high school (Central / Others).

HigherSecondarySchollSpecialization: The area of specialization during high school (Commerce / Science / Arts).

DegreeGrade%: The average of grades in college (in %).

DegreeSpecialization: The area of specialization in college (Sci&Tech / Comm&Mgmt / Others).

WorkExperience: Whether the student had previous work experience (Yes / No).

EmployabilityTestGrade%: A test that was made by the university to see if a student will find a job or not.

MBA Specialization: The area of specialization in the MBA (Mkt&HR / Mkt&Fin).

MBA Grade%: The average of grades in the MBA (in %).

PlacementStatus: Whether the student finds a job or not after completing the MBA (Placed / Not Placed).

Salary: The student salary after getting a job (only available if PlacementStatus == “Placed”. NaN otherwise).

This dataset is divided into 2 tables:

Table 1:

	A	B	C	D	E	F	G	H	I	J
1	sl_no	gender	ssc_p	ssc_b	hsc_p	hsc_b	hsc_s	degree_p	degree_t	
2	1	M	67	Others	91	Others	Commerce	58	Sci&Tech	
3	2	M	79.33	Central	78.33	Others	Science	77.48	Sci&Tech	
4	3	M	65	Central		Central	Arts		Comm&Mgmt	
5	4	M	56	Central	52	Central	Science	52	Sci&Tech	
6	5	M		Central	73.6	Central	Commerce	73.3	Comm&Mgmt	
7	6	M	55	Others	49.8	Others	Science	67.25	Sci&Tech	
8	7	F	46	Others	49.2	Others	Commerce	79	Comm&Mgmt	
9	8	M	82	Central		Central	Science	66	Sci&Tech	
10	9	M	73	Central	79	Central	Commerce	72	Comm&Mgmt	
11	10	M	58	Central	70	Central	Commerce	61	Comm&Mgmt	
12	11	M	58	Central	61	Central	Commerce	60	Comm&Mgmt	
13	12	M	69.6	Central	68.4	Central	Commerce	78.3	Comm&Mgmt	
14	13	F	47	Central	55	Others	Science	65	Comm&Mgmt	
15	14	F	77	Central	87	Central	Commerce	59	Comm&Mgmt	
16	15	M	62	Central	47	Central	Commerce	50	Comm&Mgmt	
17	16	F	65	Central	75	Central	Commerce		Comm&Mgmt	
18	17	M	63	Central		Central	Commerce	65.6	Comm&Mgmt	
19	18	F	55	Central	67	Central	Commerce	64	Comm&Mgmt	
20	19	F	63	Central	66	Central	Commerce	64	Comm&Mgmt	
21	20	M	60	Others	67	Others	Arts	70	Comm&Mgmt	
22	21	M	62	Others		Others	Commerce	66	Comm&Mgmt	
23	22	F	79	Others	76	Others	Commerce	85	Comm&Mgmt	
24	23	F	69.8	Others	60.8	Others	Science	72.23	Sci&Tech	
25	24	F		Others	60	Others	Science	64.74	Sci&Tech	
26	25	M	76.5	Others	97.7	Others	Science	78.86	Sci&Tech	
27	26	F	52.58	Others	54.6	Central	Commerce	50.2	Comm&Mgmt	

Table 2:

	A	B	C	D	E	F	G	
1	sl_no	workex	etest_p	specialisat	mba_p	status	salary	
2	1	No	55	Mkt&HR	58.8	Placed	270000	
3	2	Yes	86.5	Mkt&Fin	66.28	Placed	200000	
4	3	No	75	Mkt&Fin	57.8	Placed	250000	
5	4	No		Mkt&HR		Not Placed		
6	5	No	96.8	Mkt&Fin	55.5	Placed	425000	
7	6	Yes	55	Mkt&Fin	51.58	Not Placed		
8	7	No	74.28	Mkt&Fin	53.29	Not Placed		
9	8	Yes	67	Mkt&Fin	62.14	Placed	252000	
0	9	No	91.34	Mkt&Fin	61.29	Placed	231000	
1	10	No	54	Mkt&Fin	52.21	Not Placed		
2	11	Yes	62	Mkt&HR	60.85	Placed	260000	
3	12	Yes	60	Mkt&Fin	63.7	Placed	250000	
4	13	No	62	Mkt&HR	65.04	Not Placed		
5	14	No	68	Mkt&Fin	68.63	Placed	218000	
6	15	No	76	Mkt&HR	54.96	Not Placed		
7	16	Yes	72	Mkt&Fin	64.66	Placed	200000	
8	17	Yes	60	Mkt&Fin	62.54	Placed	300000	
9	18	No	60	Mkt&Fin	67.28	Not Placed		
0	19	No	68	Mkt&HR	64.08	Not Placed		
1	20	Yes	50.48	Mkt&Fin	77.89	Placed	236000	
2	21	No	50	Mkt&HR	56.7	Placed	265000	
3	22	No	95	Mkt&Fin	69.06	Placed	393000	
4	23	No	55.53	Mkt&HR	68.81	Placed	360000	
5	24	Yes	92	Mkt&Fin	63.62	Placed	300000	
6	25	No	97.4	Mkt&Fin	74.01	Placed	360000	
7	26	Yes	76	Mkt&Fin	65.33	Not Placed		

Now, running queries in R to merge these tables into a single dataset.

Queries for merging dataset:

Reading csv files and storing in two dataframes:

```
1 df1<-read.csv("../Placement_Data_Full_Class.csv")
2 df1[1:10,]
3 df2<-read.csv("../placement_data_2.csv")
4 df2[1:10,]
5
6 summary(df1)
7 summary(df2)
8
9 df= df1 %>% inner_join(df2,by="sl_no")
10 df[1:10,]
11 #remove NA values
12 df$ssc_p[is.na(df$ssc_p)] <- mean(df$ssc_p, na.rm = TRUE)
13 df[1:10,]
14
15 df$hsc_p[is.na(df$hsc_p)] <- mean(df$hsc_p, na.rm = TRUE)
16 df[1:10,]
17
18
19 df$degree_p[is.na(df$degree_p)] <- mean(df$degree_p, na.rm = TRUE)
20 df[1:10,]
```

4:11 (Top Level) R Script

Console Terminal Jobs

```
R 4.2.0 ~ /ds project/
> df1<-read.csv("../Placement_Data_Full_Class.csv")
> df1[1:10,]
  sl_no gender ssc_p ssc_b hsc_p hsc_b hsc_s degree_p degree_t
1      1      M  67.00 Others 91.00 Others Commerce  58.00  Sci&Tech
2      2      M  79.33 Central 78.33 Others Science  77.48  Sci&Tech
3      3      M  65.00 Central NA Central Arts NA Comm&Mgmt
4      4      M  56.00 Central 52.00 Central Science  52.00  Sci&Tech
5      5      M    NA Central 73.60 Central Commerce  73.30 Comm&Mgmt
6      6      M  55.00 Others 49.80 Others Science  67.25  Sci&Tech
7      7      F  46.00 Others 49.20 Others Commerce  79.00 Comm&Mgmt
8      8      M  82.00 Central NA Central Science  66.00  Sci&Tech
9      9      M  73.00 Central 79.00 Central Commerce  72.00 Comm&Mgmt
10     10      M  58.00 Central 70.00 Central Commerce  61.00 Comm&Mgmt
> df2<-read.csv("../placement_data_2.csv")
> df2[1:10,]
  sl_no workex etest_p specialisation mba_p status salary
1      1      No  55.00 Mkt&HR 58.80 Placed 270000
2      2      Yes  86.50 Mkt&Fin 66.28 Placed 200000
3      3      No  75.00 Mkt&Fin 57.80 Placed 250000
4      4      No    NA Mkt&HR NA Not Placed NA
5      5      No  96.80 Mkt&Fin 55.50 Placed 425000
6      6      Yes  55.00 Mkt&Fin 51.58 Not Placed NA
7      7      No  74.28 Mkt&Fin 53.29 Not Placed NA
8      8      Yes  67.00 Mkt&Fin 62.14 Placed 252000
9      9      No  91.34 Mkt&Fin 61.29 Placed 231000
10     10      No  54.00 Mkt&Fin 52.21 Not Placed NA
>
```

Merging all dataset into a single dataset by using inner join function by serial number

```

1 df1<-read.csv("./Placement_Data_Full_Class.csv")
2 df1[1:10,]
3 df2<-read.csv("./placement_data_2.csv")
4 df2[1:10,]
5
6 summary(df1)
7 summary(df2)
8
9 df= df1 %>% inner_join(df2,by="sl_no")
10 df[1:10,]
11 #remove NA values
12 df$ssc_p[is.na(df$ssc_p)] <- mean(df$ssc_p, na.rm = TRUE)
13 df[1:10,]
14
15 df$hsc_p[is.na(df$hsc_p)] <- mean(df$hsc_p, na.rm = TRUE)
16 df[1:10,]
17
18 df$degree_p[is.na(df$degree_p)] <- mean(df$degree_p, na.rm = TRUE)
19 df[1:10,]
20

```

4:11 (Top Level) R Script

Console Terminal Jobs

```

R 4.2.0 - ~/ds project/
> df1<-read.csv("./Placement_Data_Full_Class.csv")
> df1[1:10,]
  sl_no gender ssc_p ssc_b hsc_p hsc_b hsc_s degree_p degree_t
1      1     M  67.00 Others 91.00 Others Commerce  58.00  Sci&Tech
2      2     M  79.33 Central 78.33 Others Science  77.48  Sci&Tech
3      3     M  65.00 Central NA Central Arts NA Comm&Mgmt
4      4     M  56.00 Central 52.00 Central Science  52.00  Sci&Tech
5      5     M    NA Central 73.60 Central Commerce  73.30 Comm&Mgmt
6      6     M  55.00 Others 49.80 Others Science  67.25  Sci&Tech
7      7     F  46.00 Others 49.20 Others Commerce  79.00 Comm&Mgmt
8      8     M  82.00 Central NA Central Science  66.00  Sci&Tech
9      9     M  73.00 Central 79.00 Central Commerce  72.00 Comm&Mgmt
10     10     M  58.00 Central 70.00 Central Commerce  61.00 Comm&Mgmt
> df2<-read.csv("./placement_data_2.csv")
> df2[1:10,]
  sl_no workex etest_p specialisation mba_p status salary
1      1     No  55.00 Mkt&HR 58.80 Placed 270000
2      2     Yes  86.50 Mkt&Fin 66.28 Placed 200000
3      3     No  75.00 Mkt&Fin 57.80 Placed 250000
4      4     No    NA Mkt&HR NA Not Placed NA
5      5     No  96.80 Mkt&Fin 55.50 Placed 425000
6      6     Yes  55.00 Mkt&Fin 51.58 Not Placed NA
7      7     No  74.28 Mkt&Fin 53.29 Not Placed NA
8      8     Yes  67.00 Mkt&Fin 62.14 Placed 252000
9      9     No  91.34 Mkt&Fin 61.29 Placed 231000
10     10     No  54.00 Mkt&Fin 52.21 Not Placed NA
>

```

After merging the data into a single dataset, next step is to clean the dataset by analysing the dataset.

Data Cleaning

Checking missing and NA Values

```
> summary(df)
  sl_no      gender      ssc_p      ssc_b      hsc_p      hsc_b      hsc_s
Min.   : 1.0    Length:215    Min.   :40.89    Length:215    Min.   :37.00    Length:215
1st Qu.: 54.5    Class :character    1st Qu.:60.00    Class :character    1st Qu.:60.08    Class :character
Median :108.0    Mode  :character    Median :67.00    Mode  :character    Median :64.50    Mode  :character
Mean   :108.0                                Mean :66.75                                Mean :65.96
3rd Qu.:161.5                                3rd Qu.:74.00                                3rd Qu.:73.00
Max.   :215.0                                Max.   :89.40                                Max.   :97.70
NA's   :16                                    NA's   :25

  degree_p      degree_t      workex      etest_p      specialisation      mba_p      status
Min.   :50.00    Length:215    Length:215    Min.   :50.00    Length:215    Min.   :51.21    Length:215
1st Qu.:61.20    Class :character    Class :character    1st Qu.:60.00    Class :character    1st Qu.:57.69    Class :character
Median :66.00    Mode  :character    Mode  :character    Median :71.00                                Median :61.87    Mode  :character
Mean   :66.55                                Mean :72.10                                Mean :62.22
3rd Qu.:72.00                                3rd Qu.:83.75                                3rd Qu.:66.28
Max.   :91.00                                Max.   :98.00                                Max.   :77.89
NA's   :15                                    NA's   :18

  salary
Min.   :200000
1st Qu.:240000
Median :265000
Mean   :288655
3rd Qu.:300000
Max.   :940000
NA's   :67
```

```
> df[1:10,]
  sl_no gender ssc_p ssc_b hsc_p hsc_b hsc_s degree_p degree_t workex etest_p specialisation mba_p status
1     1     M  67.00 Others 91.00 Others Commerce 58.00 Sci&Tech No 55.00 Mkt&HR 58.80 Placed
2     2     M  79.33 Central 78.33 Others Science 77.48 Sci&Tech Yes 86.50 Mkt&Fin 66.28 Placed
3     3     M  65.00 Central NA Central Arts  NA Comm&Mgmt No 75.00 Mkt&Fin 57.80 Placed
4     4     M  56.00 Central 52.00 Central Science 52.00 Sci&Tech No NA Mkt&HR NA Not Placed
5     5     M  NA Central 73.60 Central Commerce 73.30 Comm&Mgmt No 96.80 Mkt&Fin 55.50 Placed
6     6     M  55.00 Others 49.80 Others Science 67.25 Sci&Tech Yes 55.00 Mkt&Fin 51.58 Not Placed
7     7     F  46.00 Others 49.20 Others Commerce 79.00 Comm&Mgmt No 74.28 Mkt&Fin 53.29 Not Placed
8     8     M  82.00 Central NA Central Science 66.00 Sci&Tech Yes 67.00 Mkt&Fin 62.14 Placed
9     9     M  73.00 Central 79.00 Central Commerce 72.00 Comm&Mgmt No 91.34 Mkt&Fin 61.29 Placed
10    10     M  58.00 Central 70.00 Central Commerce 61.00 Comm&Mgmt No 54.00 Mkt&Fin 52.21 Not Placed

  salary
1 270000
2 200000
3 250000
4 NA
5 425000
6 NA
7 NA
8 252000
9 231000
10 NA
```

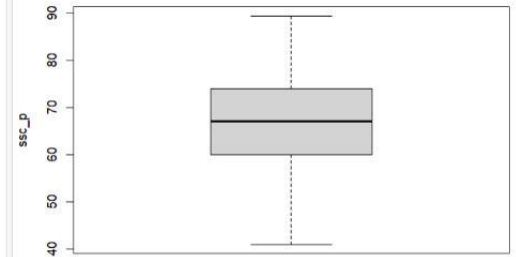
Removing NA values and replacing them with mean as its numerical data and there and doesn't have outliers.


```

> boxplot(df$ssc_p,
+         ylab = "ssc_p")
+ )
> df$ssc_p[is.na(df$ssc_p)] <- mean(df$ssc_p, na.rm = TRUE)
> df[1:10,]
  sl_no gender  ssc_p  ssc_b hsc_p  hsc_b  hsc_s degree_p degree_t workex etest_p specialisation mba_p status
1      1      M 67.0000 Others 91.00000 Others Commerce 58.00 Sci&Tech No 55.00 Mkt&HR 58.80 Placed
2      2      M 79.3300 Central 78.33000 Others Science 77.48 Sci&Tech Yes 86.50 Mkt&Fin 66.28 Placed
3      3      M 65.0000 Central 65.95795 Central Arts NA Comm&Mgmt No 75.00 Mkt&Fin 57.80 Placed
4      4      M 56.0000 Central 52.00000 Central Science 52.00 Sci&Tech No NA Mkt&HR NA Not Placed
5      5      M 66.7491 Central 73.60000 Central Commerce 73.30 Comm&Mgmt No 96.80 Mkt&Fin 55.50 Placed
6      6      M 55.0000 Others 49.80000 Others Science 67.25 Sci&Tech Yes 55.00 Mkt&Fin 51.58 Not Placed
7      7      F 46.0000 Others 49.20000 Others Commerce 79.00 Comm&Mgmt No 74.28 Mkt&Fin 53.29 Not Placed
8      8      M 82.0000 Central 65.95795 Central Science 66.00 Sci&Tech Yes 67.00 Mkt&Fin 62.14 Placed
9      9      M 73.0000 Central 79.00000 Central Commerce 72.00 Comm&Mgmt No 91.34 Mkt&Fin 61.29 Placed
10     10     M 58.0000 Central 70.00000 Central Commerce 61.00 Comm&Mgmt No 54.00 Mkt&Fin 52.21 Not Placed

  salary
1 270000
2 200000
3 250000
4 NA
5 425000
6 NA
7 NA
8 252000
9 231000
10 NA
> |

```



```

> df$hsc_p[is.na(df$hsc_p)] <- mean(df$hsc_p, na.rm = TRUE)
> df[1:10,]

```

```

  sl_no gender  ssc_p  ssc_b  hsc_p  hsc_b  hsc_s degree_p degree_t workex etest_p specialisation mba_p status
1      1      M 67.0000 Others 91.00000 Others Commerce 58.00 Sci&Tech No 55.00 Mkt&HR 58.80 Placed
2      2      M 79.3300 Central 78.33000 Others Science 77.48 Sci&Tech Yes 86.50 Mkt&Fin 66.28 Placed
3      3      M 65.0000 Central 65.95795 Central Arts NA Comm&Mgmt No 75.00 Mkt&Fin 57.80 Placed
4      4      M 56.0000 Central 52.00000 Central Science 52.00 Sci&Tech No NA Mkt&HR NA Not Placed
5      5      M 66.7491 Central 73.60000 Central Commerce 73.30 Comm&Mgmt No 96.80 Mkt&Fin 55.50 Placed
6      6      M 55.0000 Others 49.80000 Others Science 67.25 Sci&Tech Yes 55.00 Mkt&Fin 51.58 Not Placed
7      7      F 46.0000 Others 49.20000 Others Commerce 79.00 Comm&Mgmt No 74.28 Mkt&Fin 53.29 Not Placed
8      8      M 82.0000 Central 65.95795 Central Science 66.00 Sci&Tech Yes 67.00 Mkt&Fin 62.14 Placed
9      9      M 73.0000 Central 79.00000 Central Commerce 72.00 Comm&Mgmt No 91.34 Mkt&Fin 61.29 Placed
10     10     M 58.0000 Central 70.00000 Central Commerce 61.00 Comm&Mgmt No 54.00 Mkt&Fin 52.21 Not Placed

  salary
1 270000
2 200000
3 250000
4 NA
5 425000
6 NA
7 NA
8 252000
9 231000
10 NA
> |

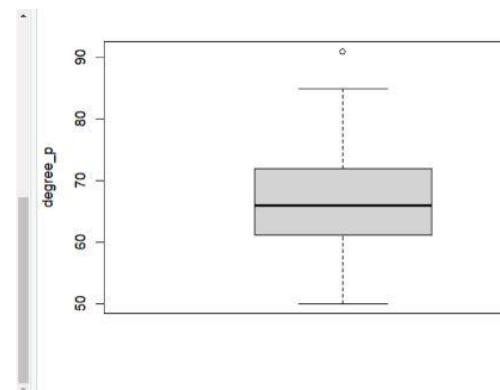
```

```

> boxplot(df$degree_p,
+         ylab = "degree_p")
+ )
> df$degree_p[is.na(df$degree_p)] <- mean(df$degree_p, na.rm = TRUE)
> df[1:10,]
  sl_no gender  ssc_p  ssc_b  hsc_p  hsc_b  hsc_s degree_p degree_t workex etest_p specialisation mba_p status
1      1      M 67.0000 Others 91.00000 Others Commerce 58.00000 Sci&Tech No 55.00 Mkt&HR 58.80 Placed
2      2      M 79.3300 Central 78.33000 Others Science 77.48000 Sci&Tech Yes 86.50 Mkt&Fin 66.28 Placed
3      3      M 65.0000 Central 65.95795 Central Arts 66.55055 Comm&Mgmt No 75.00 Mkt&Fin 57.80 Placed
4      4      M 56.0000 Central 52.00000 Central Science 52.00000 Sci&Tech No NA Mkt&HR NA Not Placed
5      5      M 66.7491 Central 73.60000 Central Commerce 73.30000 Comm&Mgmt No 96.80 Mkt&Fin 55.50 Placed
6      6      M 55.0000 Others 49.80000 Others Science 67.25000 Sci&Tech Yes 55.00 Mkt&Fin 51.58 Not Placed
7      7      F 46.0000 Others 49.20000 Others Commerce 79.00000 Comm&Mgmt No 74.28 Mkt&Fin 53.29 Not Placed
8      8      M 82.0000 Central 65.95795 Central Science 66.00000 Sci&Tech Yes 67.00 Mkt&Fin 62.14 Placed
9      9      M 73.0000 Central 79.00000 Central Commerce 72.00000 Comm&Mgmt No 91.34 Mkt&Fin 61.29 Placed
10     10     M 58.0000 Central 70.00000 Central Commerce 61.00000 Comm&Mgmt No 54.00 Mkt&Fin 52.21 Not Placed

  salary
1 270000
2 200000
3 250000
4 NA
5 425000
6 NA
7 NA
8 252000
9 231000
10 NA
> |

```




```

> boxplot(df$etest_p,
+         ylab = "etest_p"
+ )
> df$etest_p[is.na(df$etest_p)] <- mean(df$etest_p, na.rm = TRUE)
> df[1:10,]

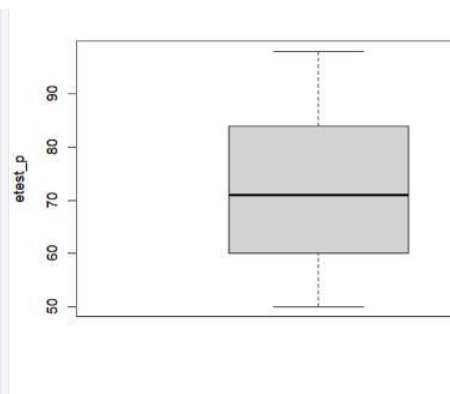
```

s_l_no	gender	ssc_p	ssc_b	hsc_p	hsc_b	hsc_s	degree_p	degree_t	workex	etest_p	specialisation	mba_p	status	
1	1	M	67.0000	Others	91.00000	Others	Commerce	58.00000	Sci&Tech	No	55.00000	Mkt&HR	58.80	Placed
2	2	M	79.3300	Central	78.33000	Others	Science	77.48000	Sci&Tech	Yes	86.50000	Mkt&Fin	66.28	Placed
3	3	M	65.0000	Central	65.95795	Central	Arts	66.55055	Comm&Mgmt	No	75.00000	Mkt&Fin	57.80	Placed
4	4	M	56.0000	Central	52.00000	Central	Science	52.00000	Sci&Tech	No	72.09758	Mkt&HR	NA	Not Placed
5	5	M	66.7491	Central	73.60000	Central	Commerce	73.30000	Comm&Mgmt	No	96.80000	Mkt&Fin	55.50	Placed
6	6	M	55.0000	Others	49.80000	Others	Science	67.25000	Sci&Tech	Yes	55.00000	Mkt&Fin	51.58	Not Placed
7	7	F	46.0000	Others	49.20000	Others	Commerce	79.00000	Comm&Mgmt	No	74.28000	Mkt&Fin	53.29	Not Placed
8	8	M	82.0000	Central	65.95795	Central	Science	66.00000	Sci&Tech	Yes	67.00000	Mkt&Fin	62.14	Placed
9	9	M	73.0000	Central	79.00000	Central	Commerce	72.00000	Comm&Mgmt	No	91.34000	Mkt&Fin	61.29	Placed
10	10	M	58.0000	Central	70.00000	Central	Commerce	61.00000	Comm&Mgmt	No	54.00000	Mkt&Fin	52.21	Not Placed

```

1 salary
2 270000
3 200000
4 250000
5 NA
6 425000
7 NA
8 252000
9 231000
10 NA
>

```



```

> boxplot(df$mba_p,
+         ylab = "mba_p"
+ )
> df$mba_p[is.na(df$mba_p)] <- mean(df$mba_p, na.rm = TRUE)
> df[1:10,]

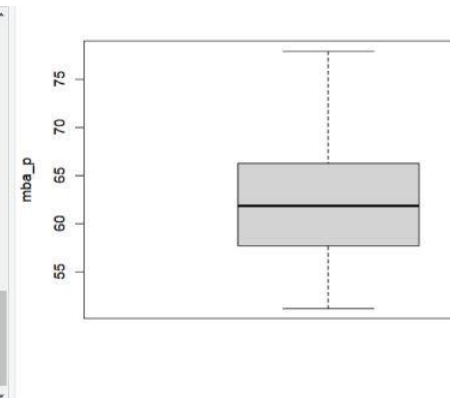
```

s_l_no	gender	ssc_p	ssc_b	hsc_p	hsc_b	hsc_s	degree_p	degree_t	workex	etest_p	specialisation	mba_p	
1	1	M	67.0000	Others	91.00000	Others	Commerce	58.00000	Sci&Tech	No	55.00000	Mkt&HR	58.80000
2	2	M	79.3300	Central	78.33000	Others	Science	77.48000	Sci&Tech	Yes	86.50000	Mkt&Fin	66.28000
3	3	M	65.0000	Central	65.95795	Central	Arts	66.55055	Comm&Mgmt	No	75.00000	Mkt&Fin	57.80000
4	4	M	56.0000	Central	52.00000	Central	Science	52.00000	Sci&Tech	No	72.09758	Mkt&HR	62.22294
5	5	M	66.7491	Central	73.60000	Central	Commerce	73.30000	Comm&Mgmt	No	96.80000	Mkt&Fin	55.50000
6	6	M	55.0000	Others	49.80000	Others	Science	67.25000	Sci&Tech	Yes	55.00000	Mkt&Fin	51.58000
7	7	F	46.0000	Others	49.20000	Others	Commerce	79.00000	Comm&Mgmt	No	74.28000	Mkt&Fin	53.29000
8	8	M	82.0000	Central	65.95795	Central	Science	66.00000	Sci&Tech	Yes	67.00000	Mkt&Fin	62.14000
9	9	M	73.0000	Central	79.00000	Central	Commerce	72.00000	Comm&Mgmt	No	91.34000	Mkt&Fin	61.29000
10	10	M	58.0000	Central	70.00000	Central	Commerce	61.00000	Comm&Mgmt	No	54.00000	Mkt&Fin	52.21000

```

1 status salary
2 Placed 270000
3 Placed 200000
4 Placed 250000
5 Not Placed NA
6 Placed 425000
7 Not Placed NA
8 Placed 252000
9 Placed 231000
10 Not Placed NA
>

```



For salary we have replaced the NA values with 0 as salary was given as NA for all those who weren't placed

```

> df$salary[is.na(df$salary)] <- 0
> df[1:10,]

```

s_l_no	gender	ssc_p	ssc_b	hsc_p	hsc_b	hsc_s	degree_p	degree_t	workex	etest_p	specialisation	mba_p	status	salary	
1	1	M	67.0000	Others	91.00000	Others	Commerce	58.00000	Sci&Tech	No	55.00000	Mkt&HR	58.80000	Placed	270000
2	2	M	79.3300	Central	78.33000	Others	Science	77.48000	Sci&Tech	Yes	86.50000	Mkt&Fin	66.28000	Placed	200000
3	3	M	65.0000	Central	65.95795	Central	Arts	66.55055	Comm&Mgmt	No	75.00000	Mkt&Fin	57.80000	Placed	250000
4	4	M	56.0000	Central	52.00000	Central	Science	52.00000	Sci&Tech	No	72.09758	Mkt&HR	62.22294	Not Placed	0
5	5	M	66.7491	Central	73.60000	Central	Commerce	73.30000	Comm&Mgmt	No	96.80000	Mkt&Fin	55.50000	Placed	425000
6	6	M	55.0000	Others	49.80000	Others	Science	67.25000	Sci&Tech	Yes	55.00000	Mkt&Fin	51.58000	Not Placed	0
7	7	F	46.0000	Others	49.20000	Others	Commerce	79.00000	Comm&Mgmt	No	74.28000	Mkt&Fin	53.29000	Not Placed	0
8	8	M	82.0000	Central	65.95795	Central	Science	66.00000	Sci&Tech	Yes	67.00000	Mkt&Fin	62.14000	Placed	252000
9	9	M	73.0000	Central	79.00000	Central	Commerce	72.00000	Comm&Mgmt	No	91.34000	Mkt&Fin	61.29000	Placed	231000
10	10	M	58.0000	Central	70.00000	Central	Commerce	61.00000	Comm&Mgmt	No	54.00000	Mkt&Fin	52.21000	Not Placed	0

```

>

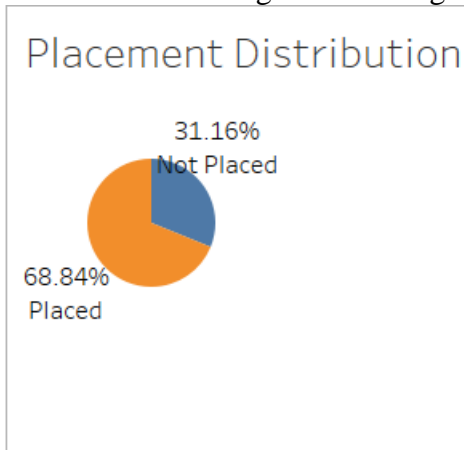
```

To finally use and analyze the data in tableau we have saved the data in RData format named as place.RData.

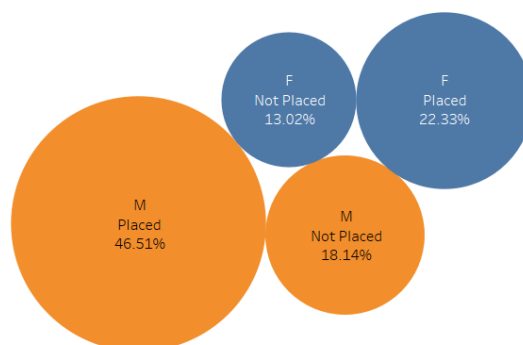
```
51 save(df,file = "C:/place.rdata")
51:33 (Top Level)
Console Terminal x Jobs x
R 4.2.0 ~/ds project/
save(df,file = "C:/place.rdata")
```

Tableau Queries

Shows more students find a job after completing MBA as compared to those who don't. Can be inferred that doing an MBA degree adds value to them

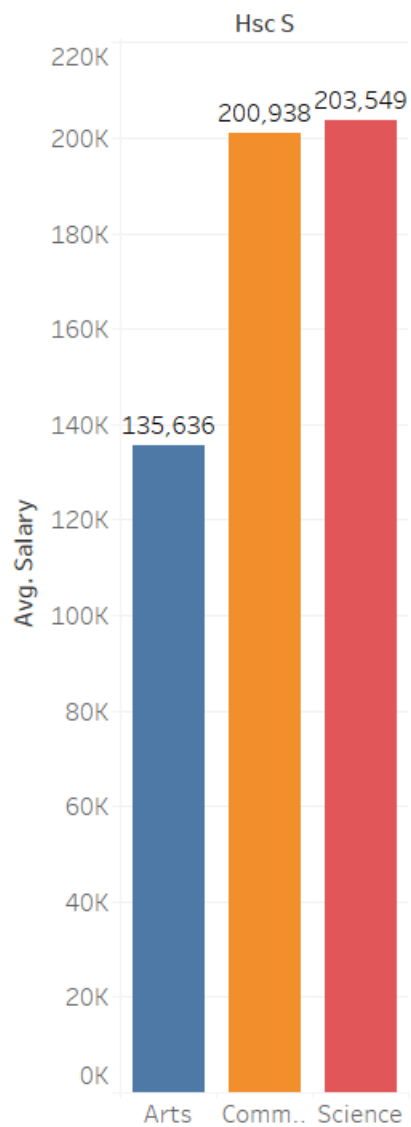


The number of male students placed are more than female students.
Gender Distribution



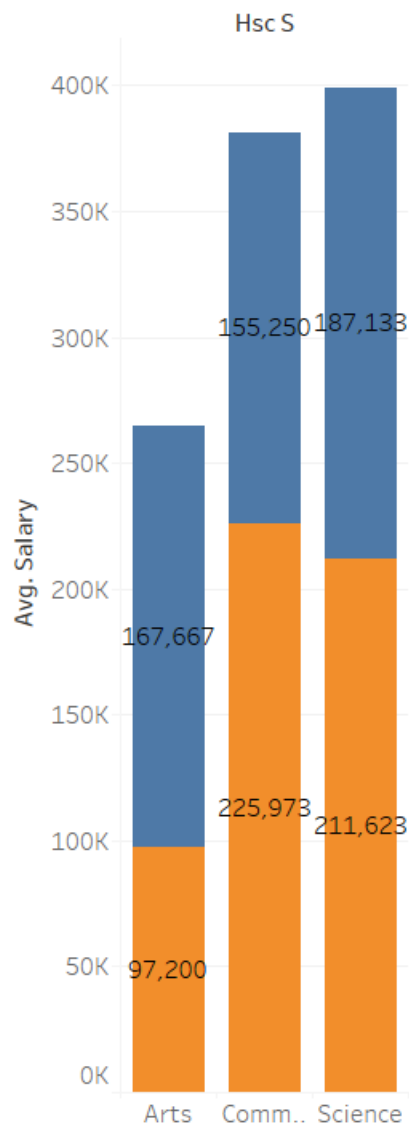
The specialization the person chose during High school plays a role. Science and Commerce again show a salary boost when compared to Arts as their salaries are higher.

High school specialization and salary



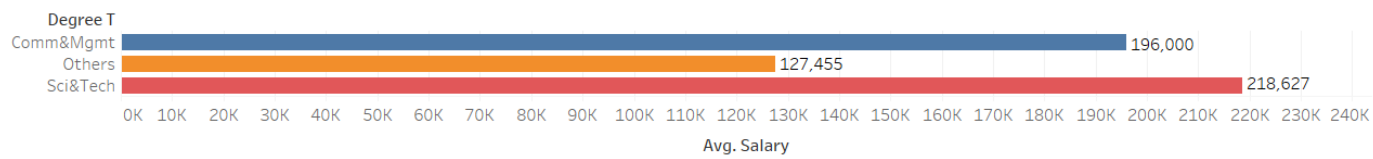
Analyzing the effects of gender on placement and salary, the conclusion was that women that came Science and Commerce tend to receive lower salaries and when they had a background different from these women receive higher salaries in this scenario.

Relation between high school specialization and salary



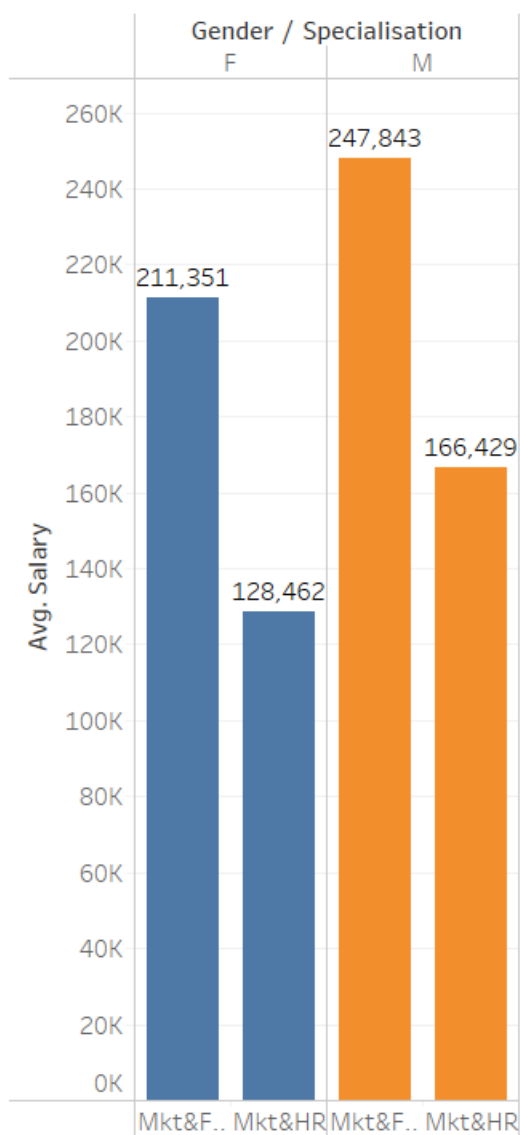
The person who chose science and tech as their college degree type had highest average salaries.

College degree area and salary

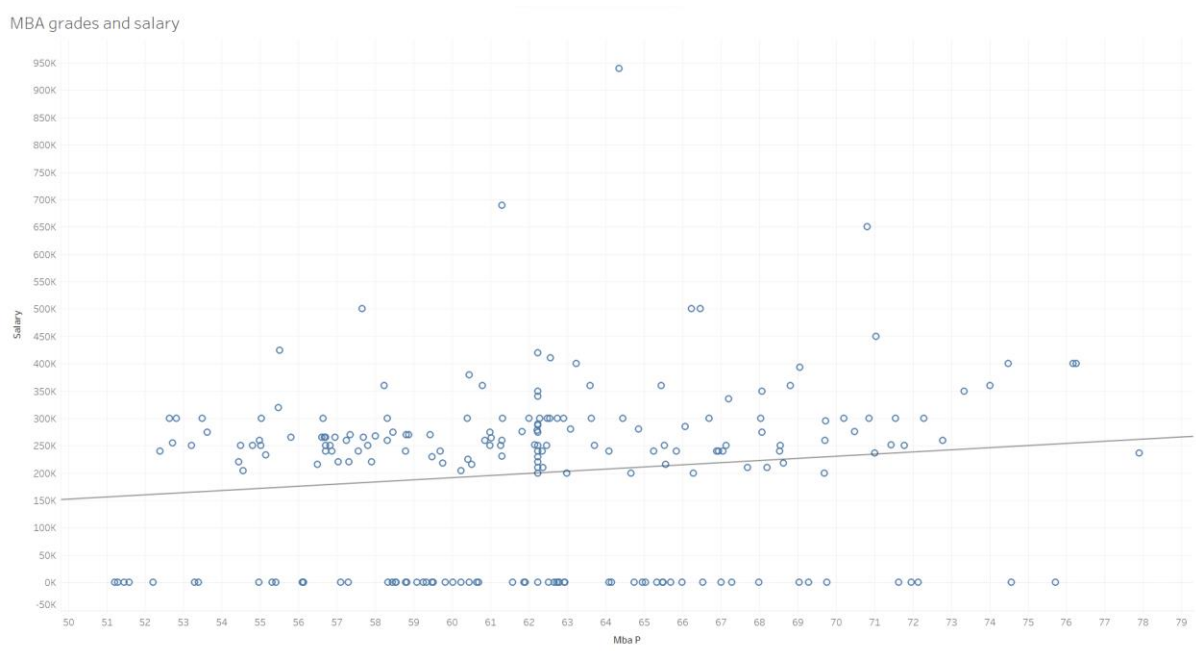


In both the MBA field choices the average salaries of men are more than female students.

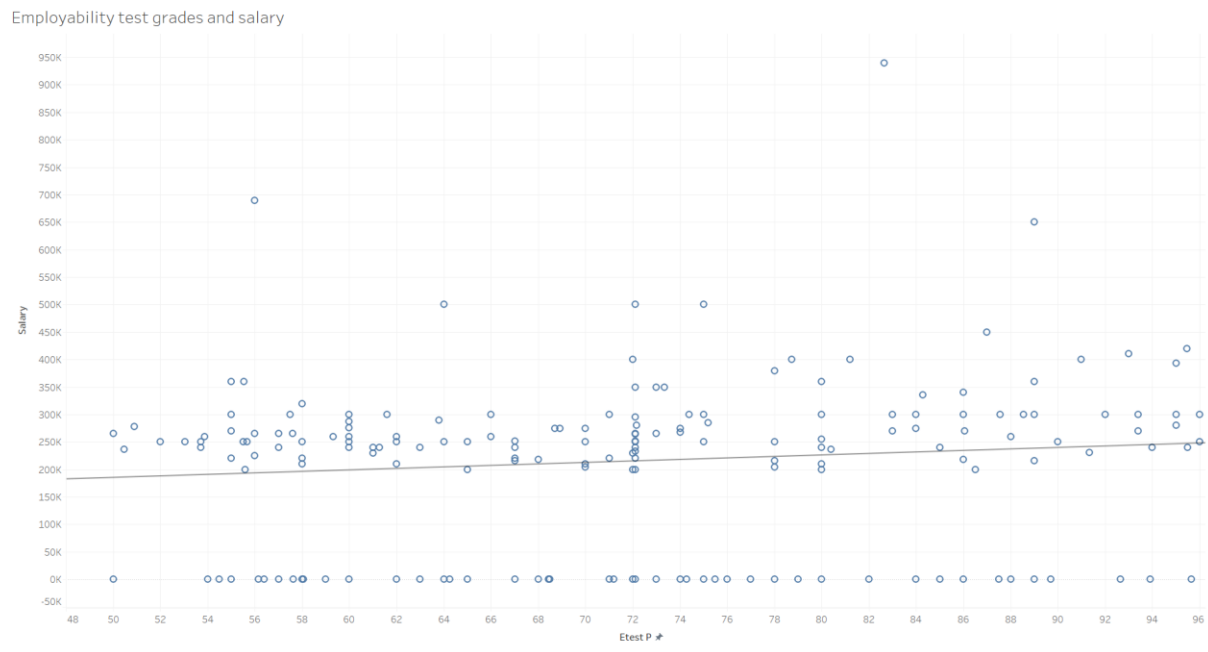
MBA field choice and salary



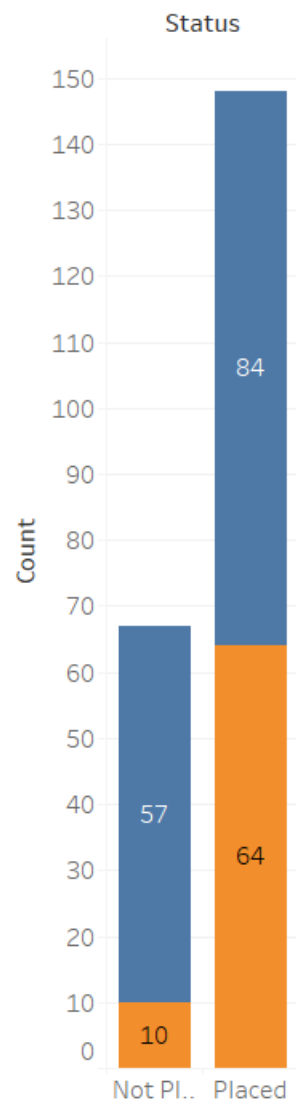
The data suggests that having good grades in MBA impacts positively on the salary a person will receive when employed.



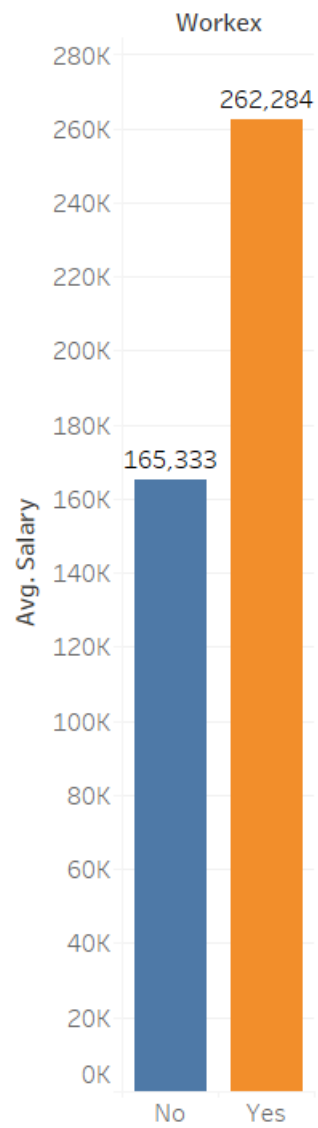
There is a positive relation between salaries and the employability test. Image shows an increase in salary for students that scored higher on the test.



Work experience and employability



Avg salary and work experience



Final Dashboard

CAMPUS RECRUITMENT ANALYSIS

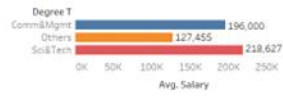
Placement Distribution



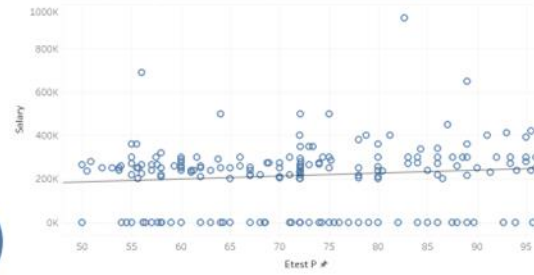
Gender Distribution



College degree area and salary



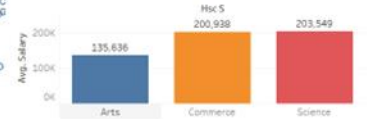
Employability test grades and salary



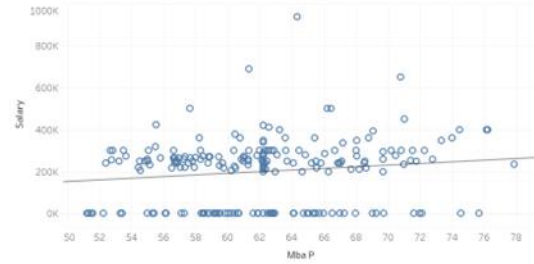
MBA field choice and salary



High school specialization and salary



MBA grades and salary



Work experience and employability

