

[← Go Back to Python Foundations](#)[Course Content](#)

Practice Quiz 3 - Python for Visualization

Type	:	Practice Quiz
Attempts	:	1/1
Questions	:	15
Time	:	2h
Your Marks	:	13/15

Instructions



Attempt History

Attempt #1

Nov 18, 8:30 PM

Marks: 13



Q No: 1

Correct Answer

Marks: 1/1

Load the 'Placement_Data.csv' into a Pandas dataframe.

What is the number of "Placed" candidates in the provided dataset?

148

You Selected

 67 140 48

There are various ways to solve this problem:

Method 1:

We can use the following code to check the shape of data whose status is equal to Placed.

```
df[df['status']=='Placed'].shape[0]
```

Method 2:

It can be counted by using the value_counts:

```
df['status'].value_counts()
```

Q No: 2

Incorrect Answer

Marks: 0/1

Print the statistical summary of the data using the pandas describe() function.

Based on the summary, which of the following statement is true?

- A. Student's MBA percentage ranges between 57.945 - 77.89%
- B. Students score 66.37% on average in their degree.

Only A Only B

Correct Option

 Neither A nor B Both A & B

You Selected

We can generate the statistical summary of the data using the describe function present in the pandas library.

```
df.describe()
```

Based on the generated summary we can observe:

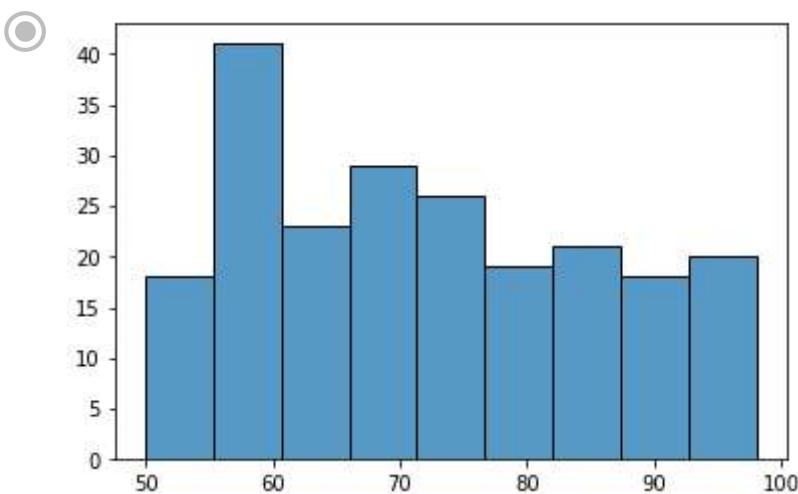
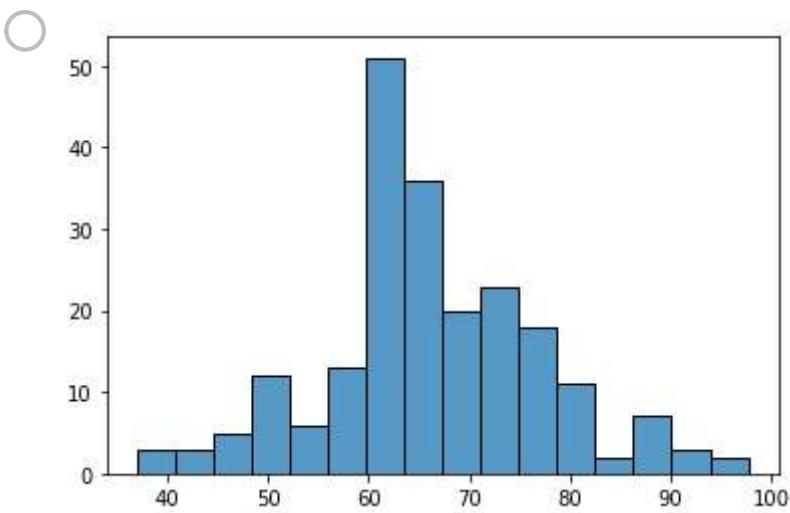
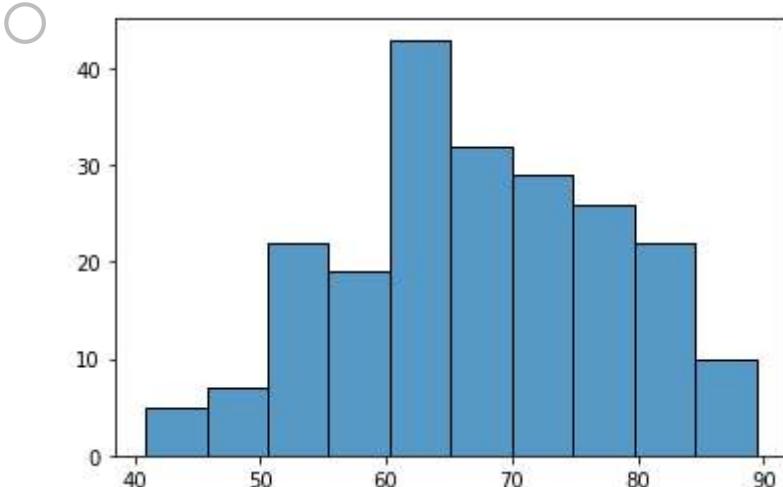
- MBA percentage('mba_p') of the students ranges between 51.21 - 77.89.
- Students score 66.37% on average in their degree.

Q No: 3

Correct Answer

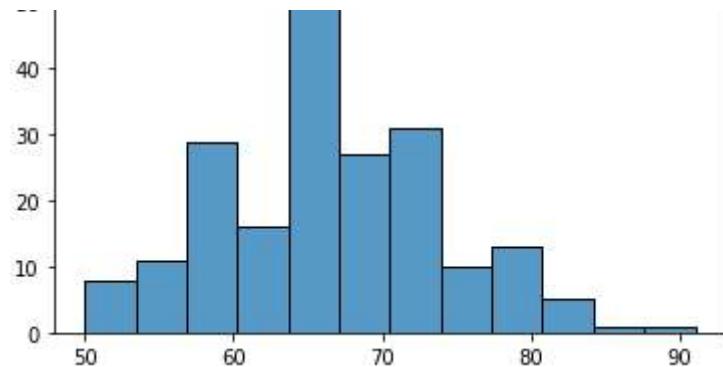
Marks: 1/1

Which one of the following graphs is the distribution plot of the Employability test?



You Selected





The above graph is a histogram, to generate a histogram we can use the [histplot function](#) [Links to an external site.](#) of the Seaborn library.

histplot() helps us to visualize the distribution of the variable

```
sns.histplot(x = 'etest_p', data = df)
```

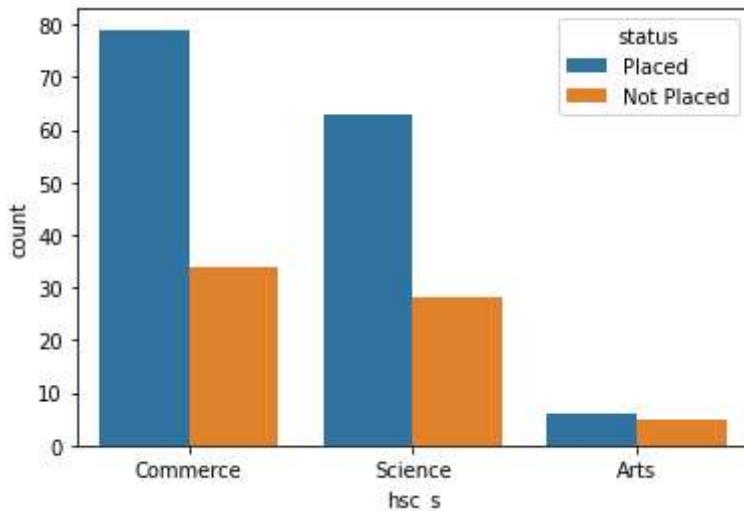
Q No: 4

Correct Answer

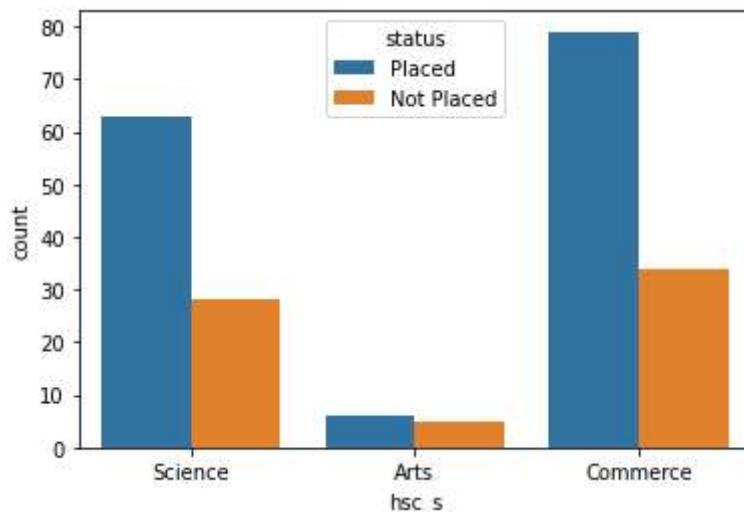
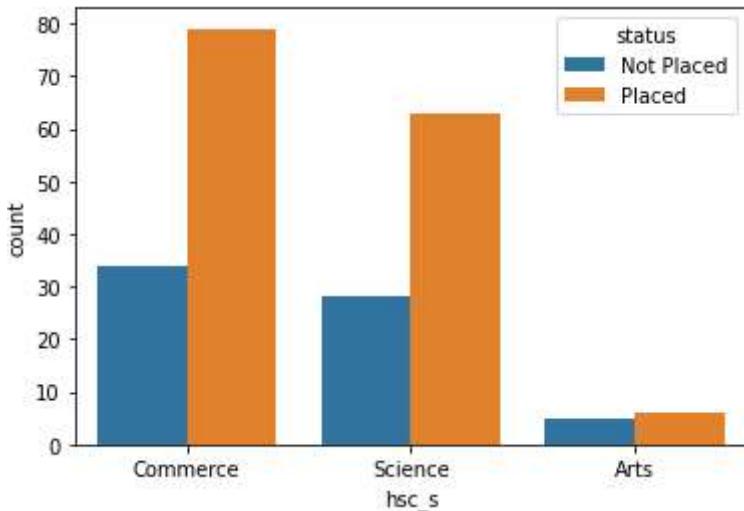
Marks: 1/1

What will be the output of the following code?

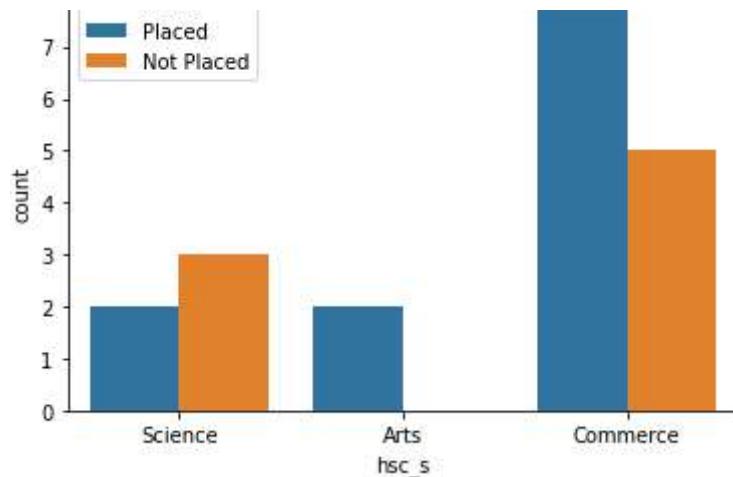
```
df = pd.read_csv('Placement_Data.csv') # Load the dataset  
sns.countplot(x = 'hsc_s', data = df, hue = 'status')
```



You Selected



8 status



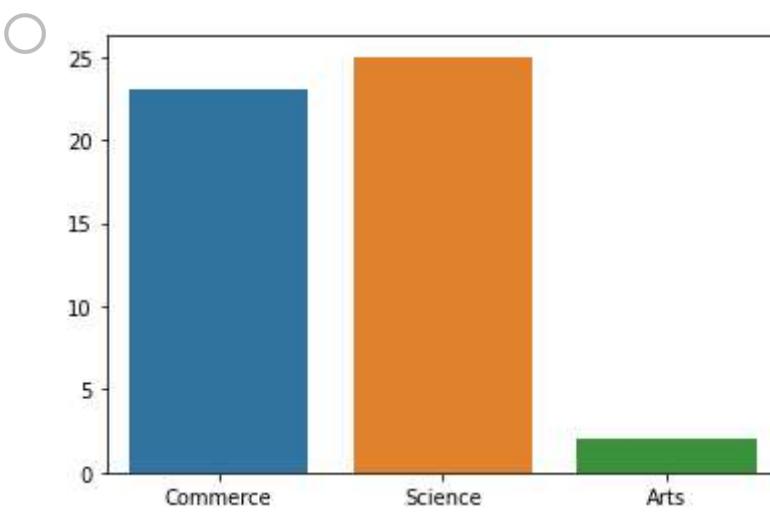
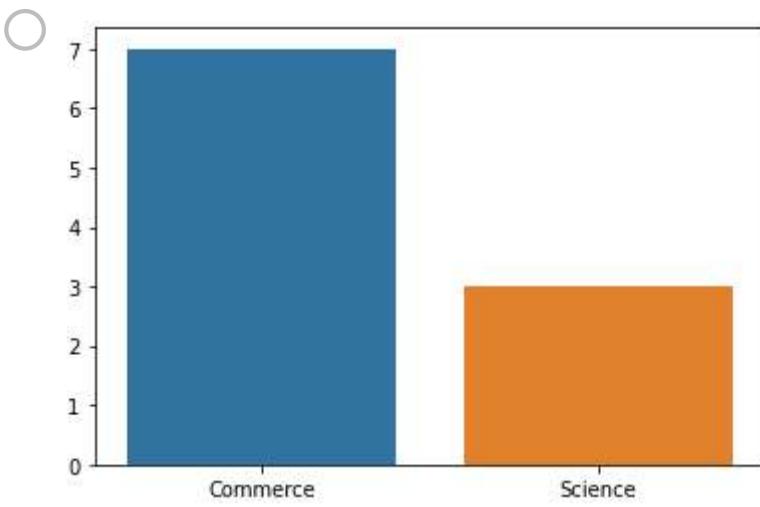
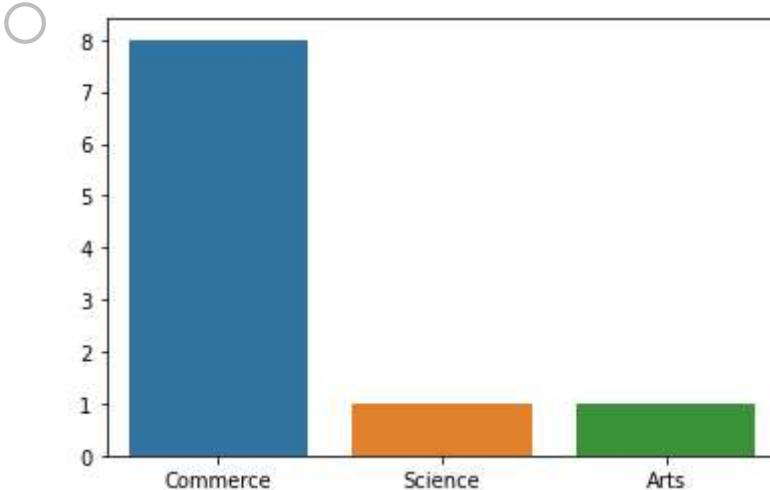
Here, the above code generates a count plot(bar graph) which categorizes the students on the x-axis based on their specialization in Higher Secondary Education(hsc_s). By using hue, we further categorize the plot and extract the student's employment status(hue = 'status').

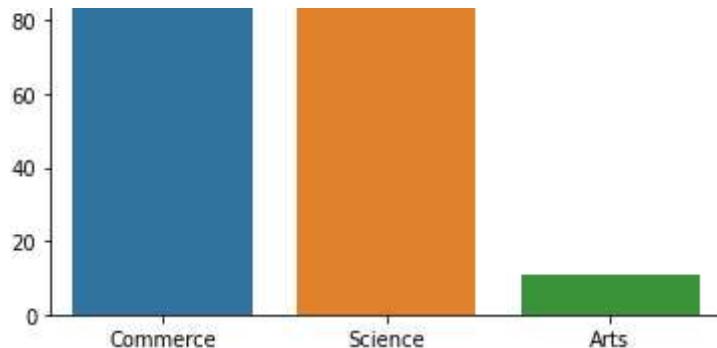
Q No: 5

Correct Answer

Marks: 1/1

Which of the following graph is of student's specialization in Higher Secondary Education?





The above plotted graphs are countplots. Countplot for student's specialization in Higher Secondary Education can be plotted using the following code:

```
sns.countplot(x = 'hsc_s', data = df)
```

We can also plot the graph by using catplot with kind = 'count'

```
sns.catplot(data = df, x = 'hsc_s', kind = 'count')
```

Q No: 6

Correct Answer

Marks: 1/1

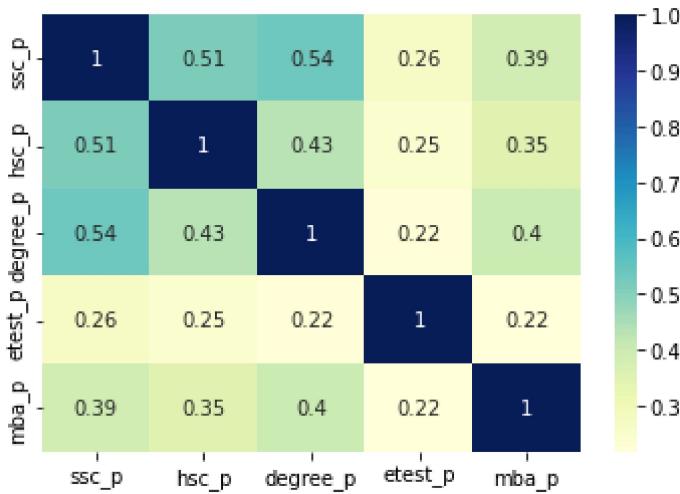
Which among the following columns have the highest correlation with each other?

- Employability test percentage and MBA percentage
- Secondary Education percentage and Degree percentage You Selected
- Higher Secondary Education percentage and Degree percentage
- Higher Secondary Education percentage and Secondary Education percentage

We can find the correlation of the numerical columns of the data using the following code:

```
sns.heatmap(df.corr(), annot = True, cmap='YlGnBu')
```

It generates a heatmap that will help plot the correlation among numerical data with color mapped index.

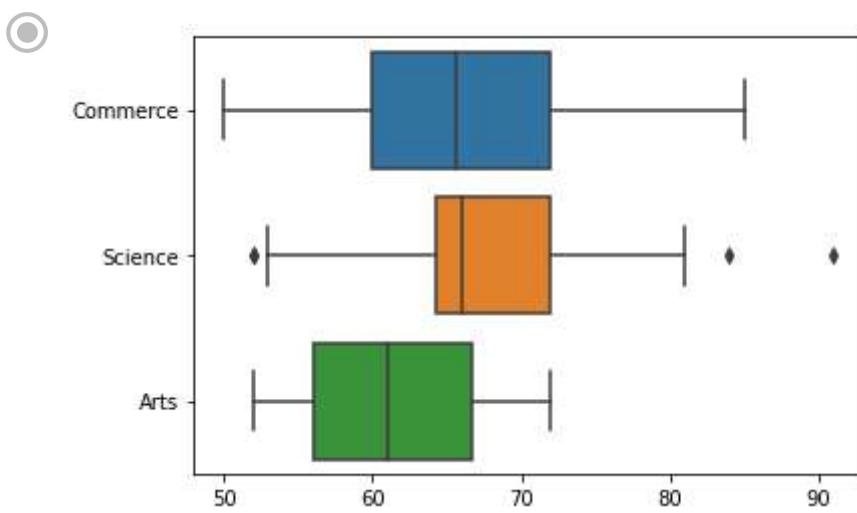
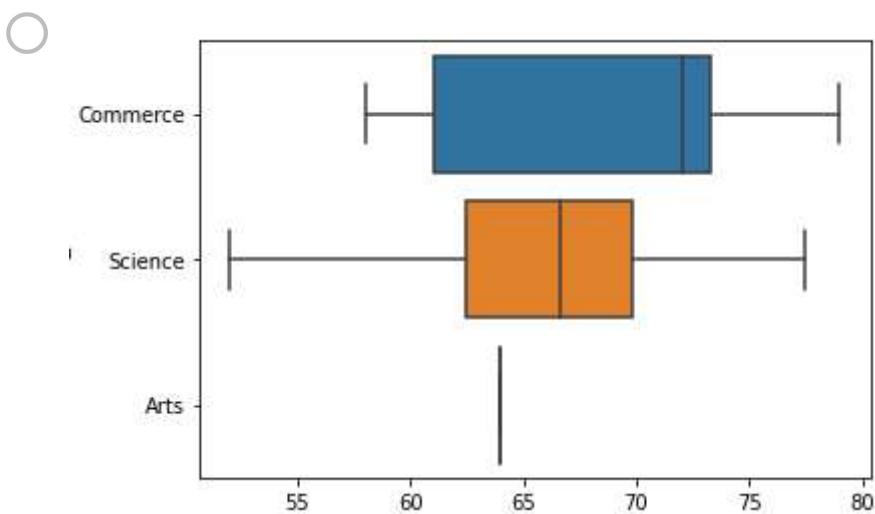
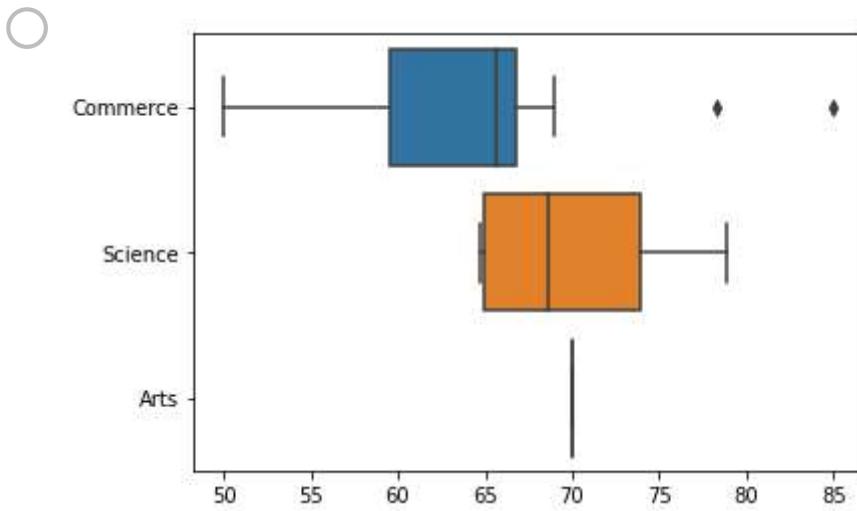


Q No: 7

Correct Answer

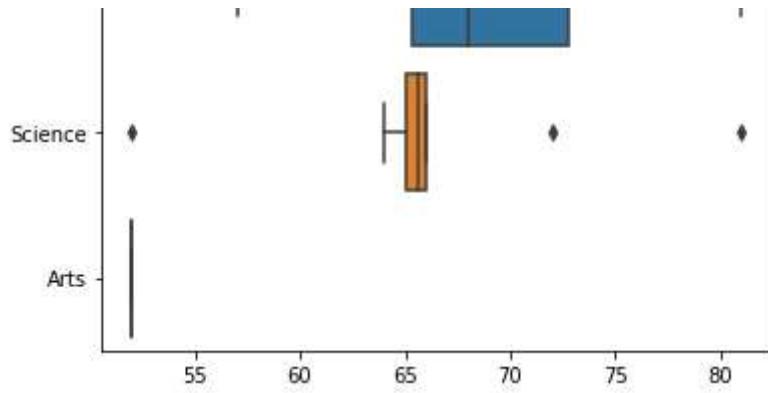
Marks: 1/1

Which one of the following is the boxplot visualization of students' degree percentage and their specialization in Higher Secondary Education?



You Selected





A box plot (or box-and-whisker plot) shows the distribution of quantitative data in a way that facilitates comparisons between variables or across levels of a categorical variable.

Boxplot can be created using boxplot function:

```
sns.boxplot(x = 'degree_p', y = 'hsc_s', data = df )
```

The same can be achieved using the catplot:

```
sns.catplot(x = 'degree_p', y = 'hsc_s', data = df, kind = 'box')
```

Q No: 8

Correct Answer

Marks: 1/1

Plot a graph between Secondary Education percentage and Higher Secondary Education percentage and categorize them based on the employment status. (hue = 'status')

Select True/False for the following statement:

Based on the graph plotted, the best fit line for Placed students is plotted higher than the best fit line for the not Placed students.

True

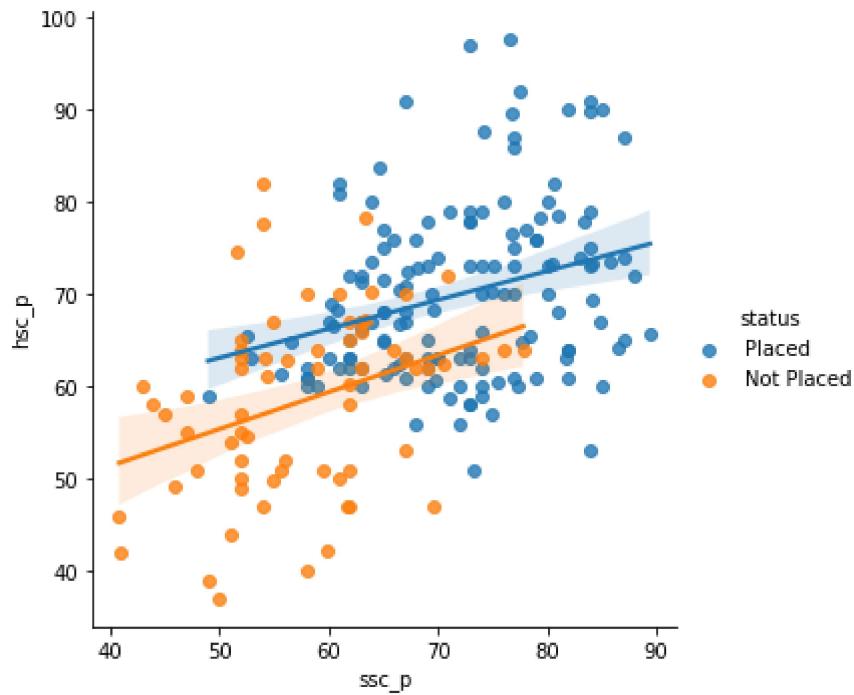
You Selected

False

Best fit lines can be generated in the [seaborn's lmplot function](#)[Links to an external site..](#) When we plot the graph using the following command:

```
sns.lmplot(x = 'ssc_p', y = 'hsc_p', hue = 'status', data = df)
```

We can observe that the best fit line for Placed students is plotted higher than the best fit line for the not Placed students.

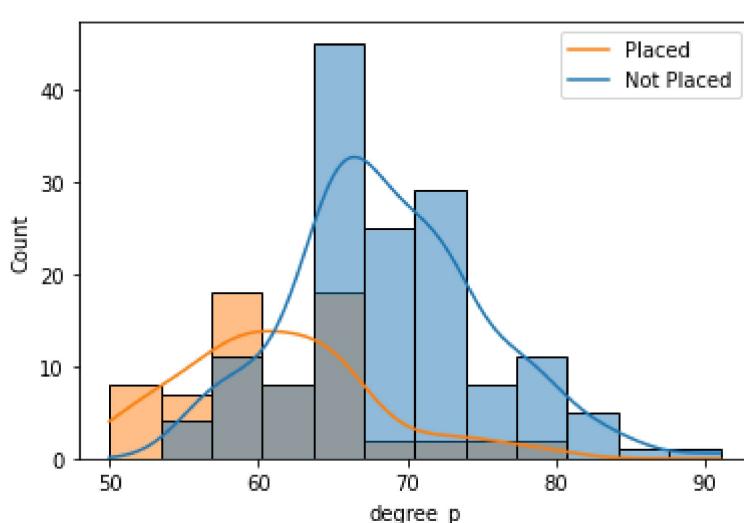
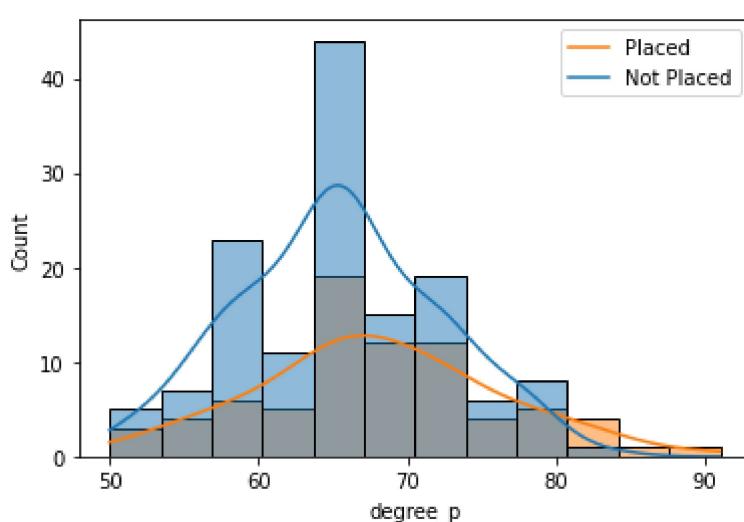
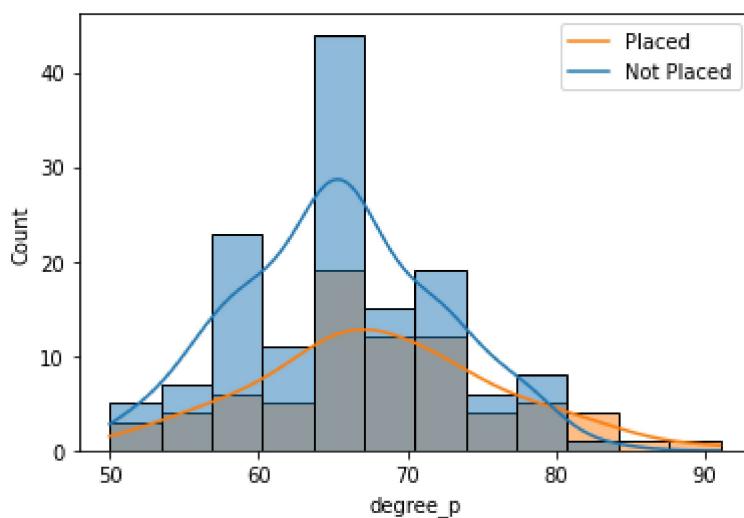


Q No: 9

Correct Answer

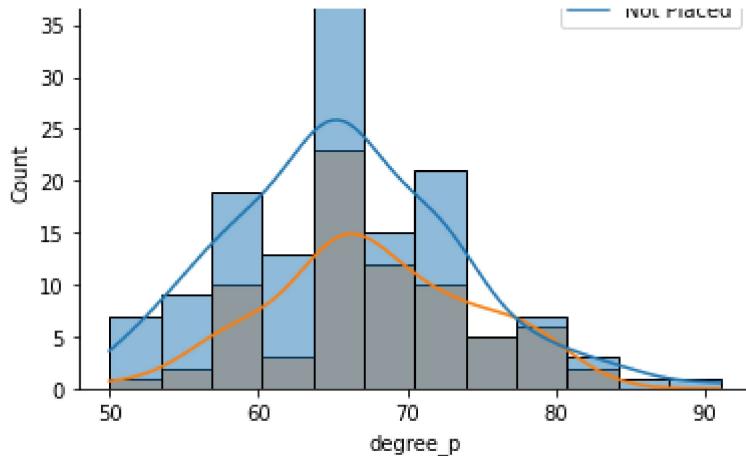
Marks: 1/1

Which plot among the following is of degree percentage with hue as status ?



You Selected





The correct graph can be generated using the following command:

```
sns.histplot(data = df,x="degree_p",kde= True,hue = 'status')  
plt.legend(["Placed", "Not Placed"])
```

Q No: 10

Correct Answer

Marks: 1/1

Select True/False for the following statement:

75% of not placed students have their Secondary Education percentage less than 65%.

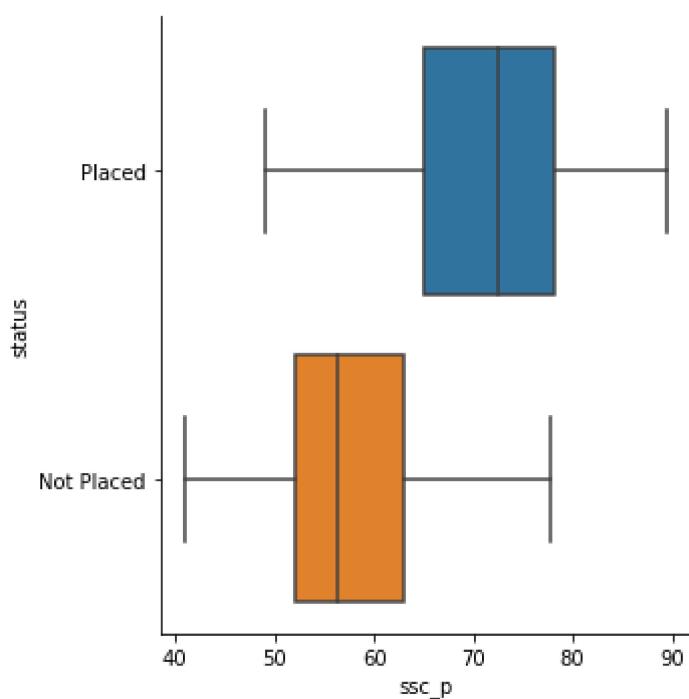
True

You Selected

False

To understand this relationship, we can plot a boxplot using the following code:

```
sns.boxplot(x = 'ssc_p', y = 'status', data = df)
```



As we can observe from the graph, 75% of not placed students have their Secondary Education percentage less than 65%.

Q No: 11

Correct Answer

Marks: 1/1

Which of the following is used to set the color mapped to the data values of heatmap?

corr

cmap

You Selected

map

color

Heatmaps can be generated using the following code:

```
sns.heatmap(df.corr(), annot = True, cmap = 'YlGnBu')
```

Here `annot` is used to print correlation on top of the colored boxes,

`cmap` : is used for mapping data values to color space. If not provided, the default will depend on whether center is set.

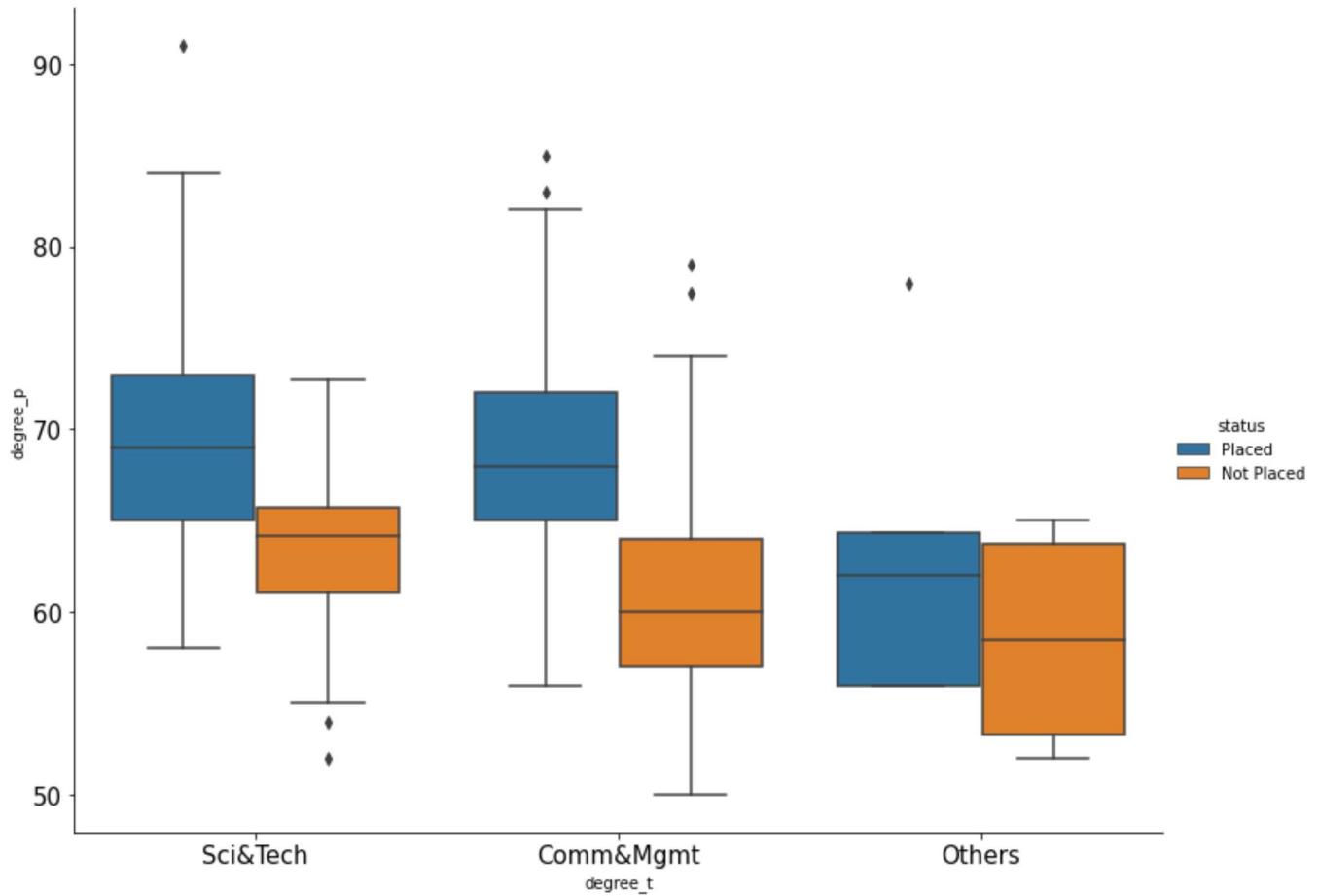
Some common `cmap`'s used for mapping are: '`YIGnBu`'[Stands for Yellow, Green and Blue], '`Blues`', '`Greens`' and '`Reds`'

Q No: 12

Correct Answer

Marks: 1/1

Following is a graph of the degree percentage of all the disciplines categorized for placed and not placed students.



Select True/False for the following statement:

Based on the graph, 50% of students who are placed have a higher degree percentage when compared with 50% of not-placed students.

True

You Selected

 False

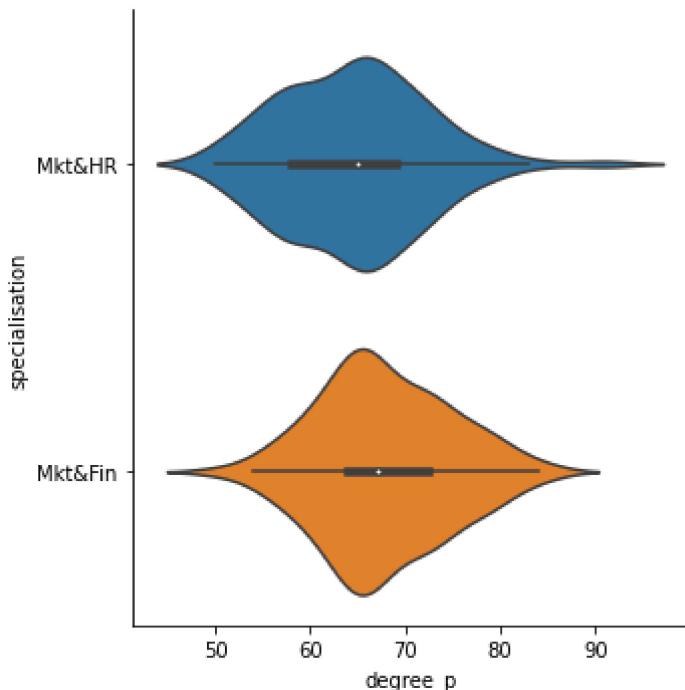
Based on the above graph, we can observe that all the median line for placed students is above the median line of the not placed students. This indicates that 50% of students that are placed have a higher degree percentage when compared with not placed students. (Since 50% of data lies above the median line.)

Q No: 13

Correct Answer

Marks: 1/1

Following is a violin plot of students with specialization and their degree percentage.



Select True/False for the following statement:

at least 75% of the students who opted for Marketing & HR and Marketing & Finance have their degree percentage below 80%

True

You Selected

 False

As we can observe from the violin plot above, the thick black line for Mkt&HR and Mkt&Fin ends before 80% on the degree percentage scale, meaning that 75% of students have their degree percentage less than 80%.

< Previous

Next >

Proprietary content. ©Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

© 2023 All rights reserved

[Privacy](#) [Terms of service](#) [Help](#)