```
from google.colab import drive
drive.mount('/content/drive')
```

```
Mounted at /content/drive
```

In [2]:

```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
import seaborn as sns
```

In [3]:

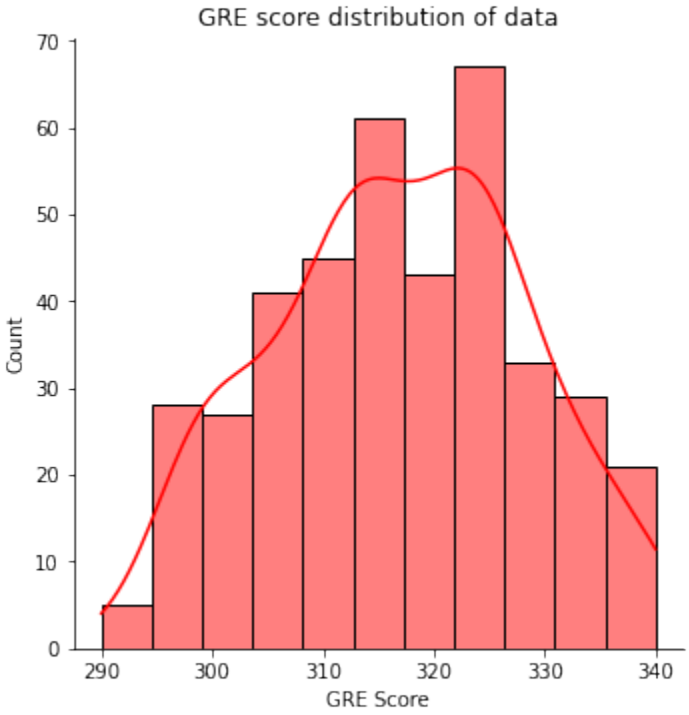
```
data=pd.read_csv('/content/drive/MyDrive/data/Admission_Predict.csv')
```

Data visualization

(i) Univariate Analysis

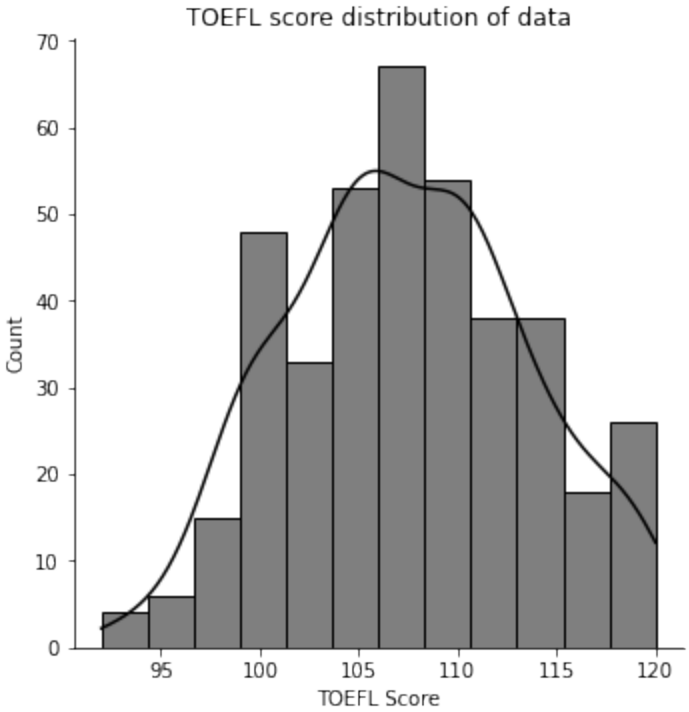
In [4]:

```
sns.displot(x=data["GRE Score"], kde=True, color='Red')
plt.title("GRE score distribution of data");
```



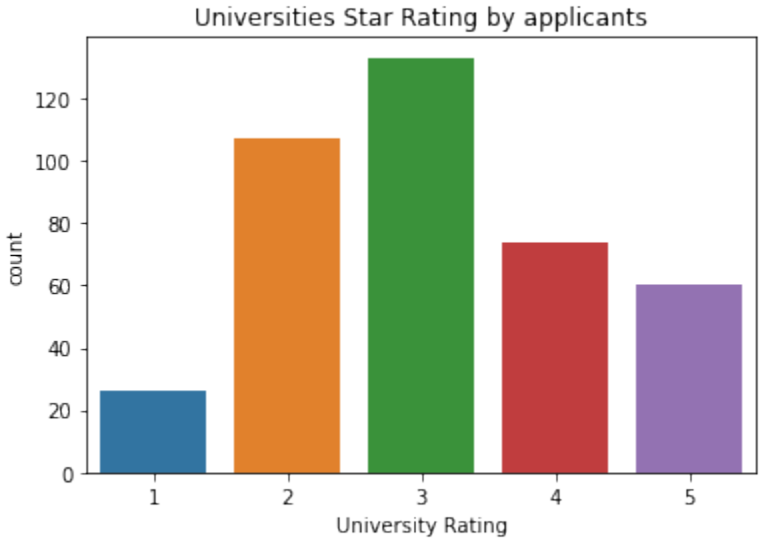
In [5]:

```
sns.displot(x=data["TOEFL Score"], kde=True, color='Black')
plt.title("TOEFL score distribution of data");
```



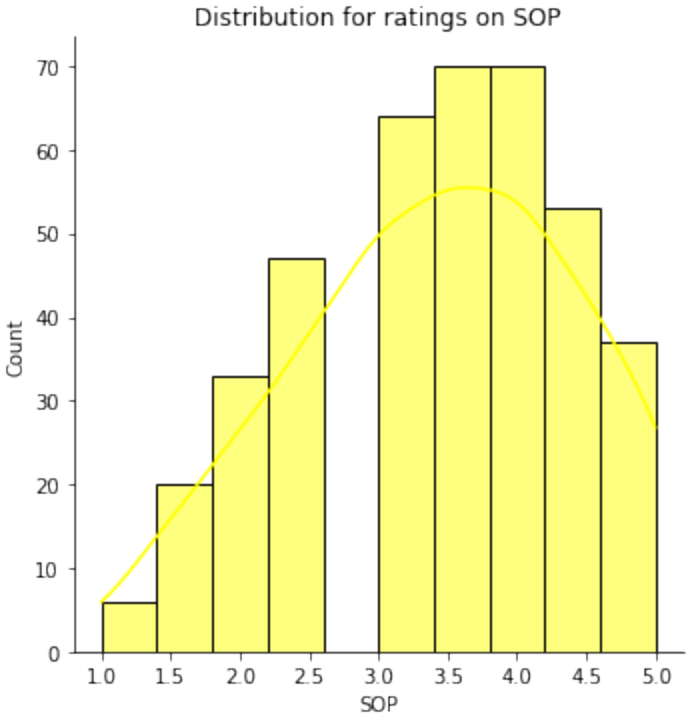
In [6]:

```
sns.countplot(x=data["University Rating"]);
plt.title("Universities Star Rating by applicants");
```



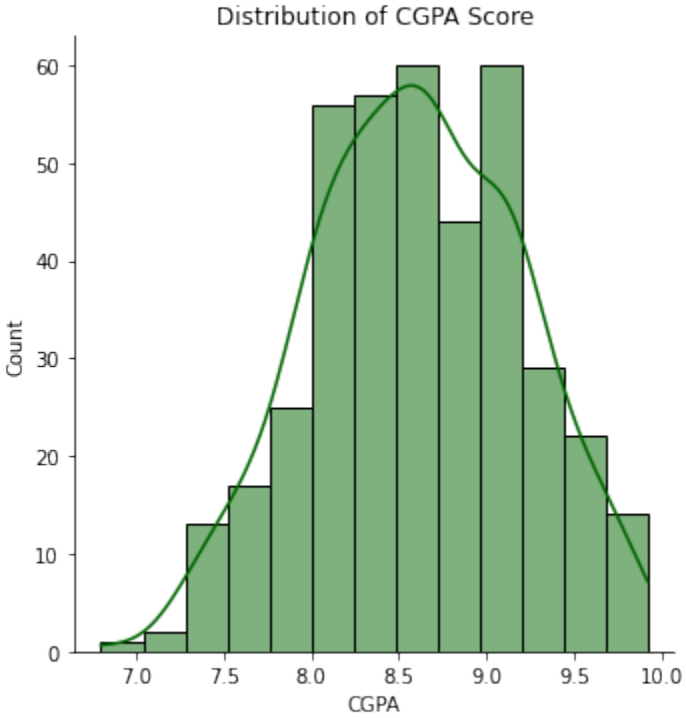
In [7]:

```
sns.displot(x=data["SOP"], kde=True, color='yellow');
plt.title("Distribution for ratings on SOP");
```



In [8]:

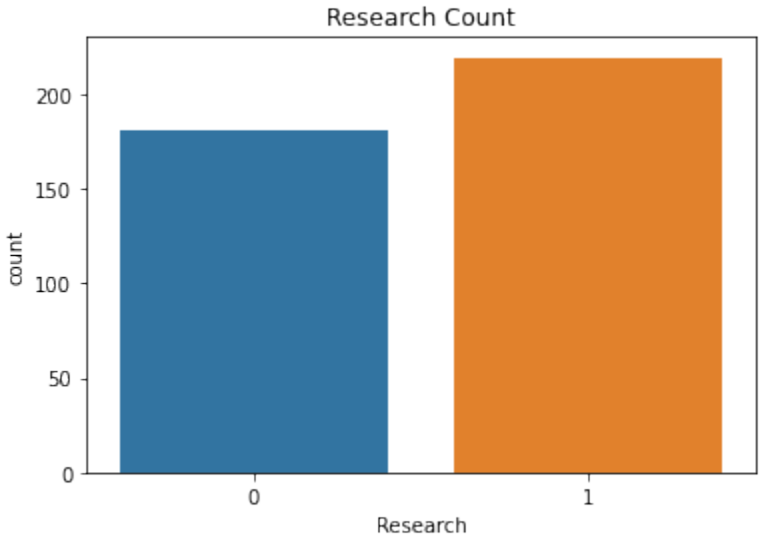
```
sns.displot(x=data["CGPA"], kde=True, color='Darkgreen');
plt.title("Distribution of CGPA Score");
```



In [9]:

```
sns.countplot(x=data["Research"]);
```

```
plt.title("Research Count");
```



(ii) Bivariate Analysis

In [10]:

```
cols = data.columns
features = [i for i in data.columns if i != 'Chance of Admit']
label = 'Chance of Admit'
features
```

Out[10]:

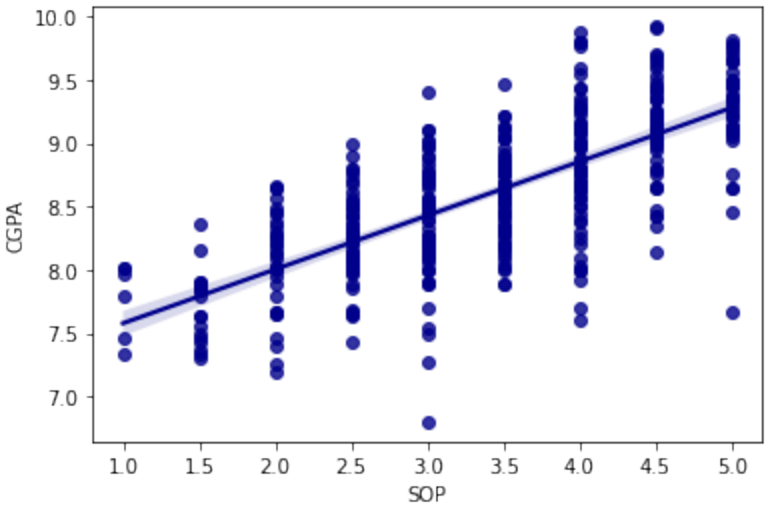
```
['Serial No.',
 'GRE Score',
 'TOEFL Score',
 'University Rating',
 'SOP',
 'LOR ',
 'CGPA',
 'Research',
 'Chance of Admit ']
```

In [11]:

```
sns.regplot(data['SOP'],data['CGPA'],color='Darkblue')
```

```
/usr/local/lib/python3.7/dist-packages/seaborn/_decorators.py:43:
FutureWarning: Pass the following variables as keyword args: x, y. From
version 0.12, the only valid positional argument will be `data`, and
passing other arguments without an explicit keyword will result in an
error or misinterpretation.
  FutureWarning
```

Out[11]:

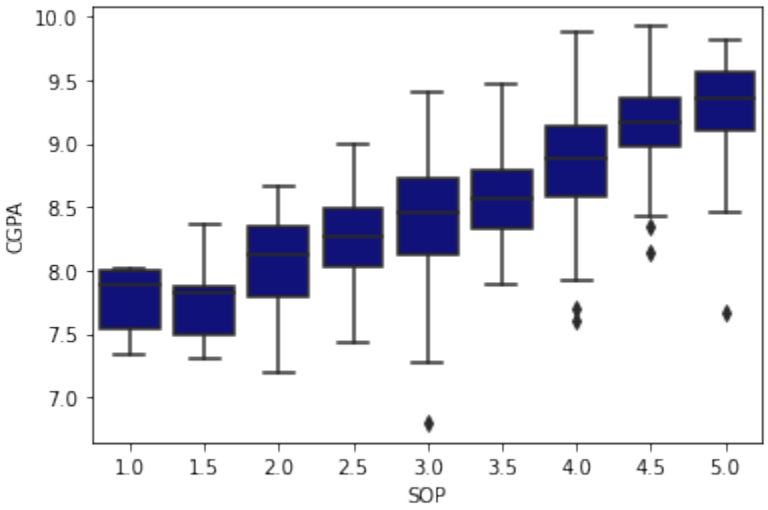


In [12]:

```
sns.boxplot(data['SOP'],data['CGPA'],color='Darkblue')
```

```
/usr/local/lib/python3.7/dist-packages/seaborn/_decorators.py:43:
FutureWarning: Pass the following variables as keyword args: x, y. From
version 0.12, the only valid positional argument will be `data`, and
passing other arguments without an explicit keyword will result in an
error or misinterpretation.
  FutureWarning
```

Out[12]:

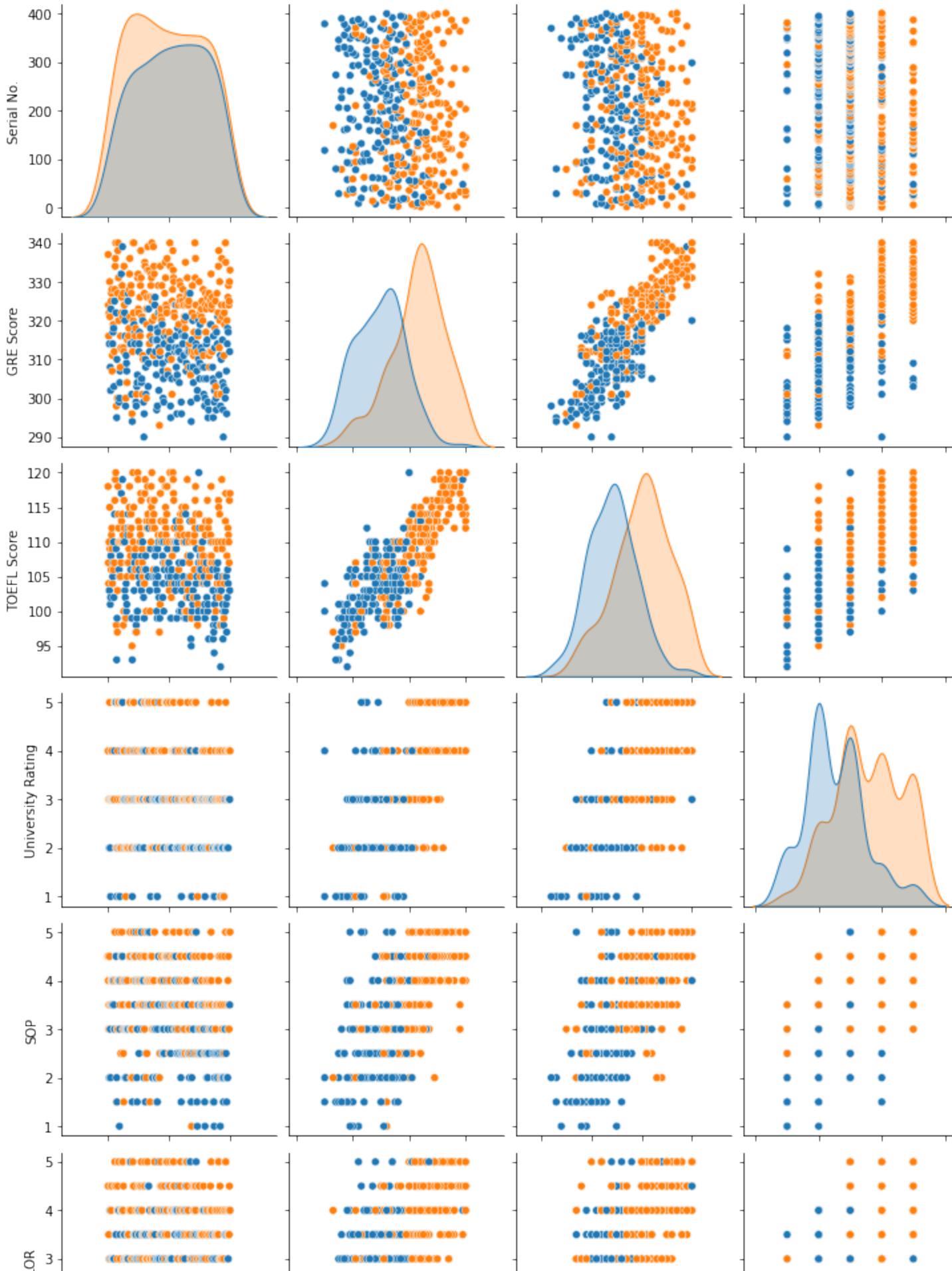


(iii) Multivariate Analysis

In [15]:

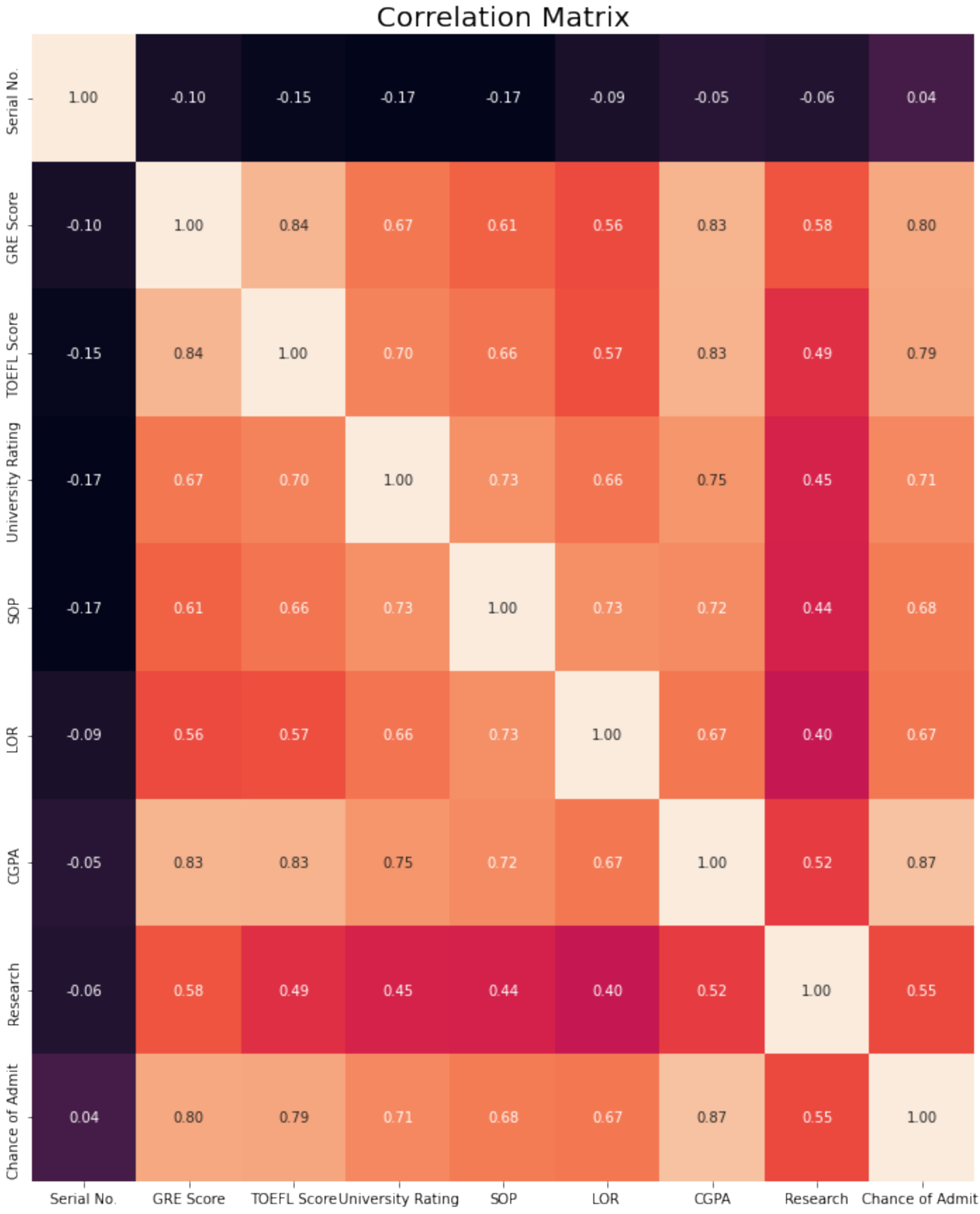
```
sns.pairplot(data,hue='Research')
```

Out[15]:


Serial No.
GRE Score
TOEFL Score
University Rating
SOP
LOR

In [13]:

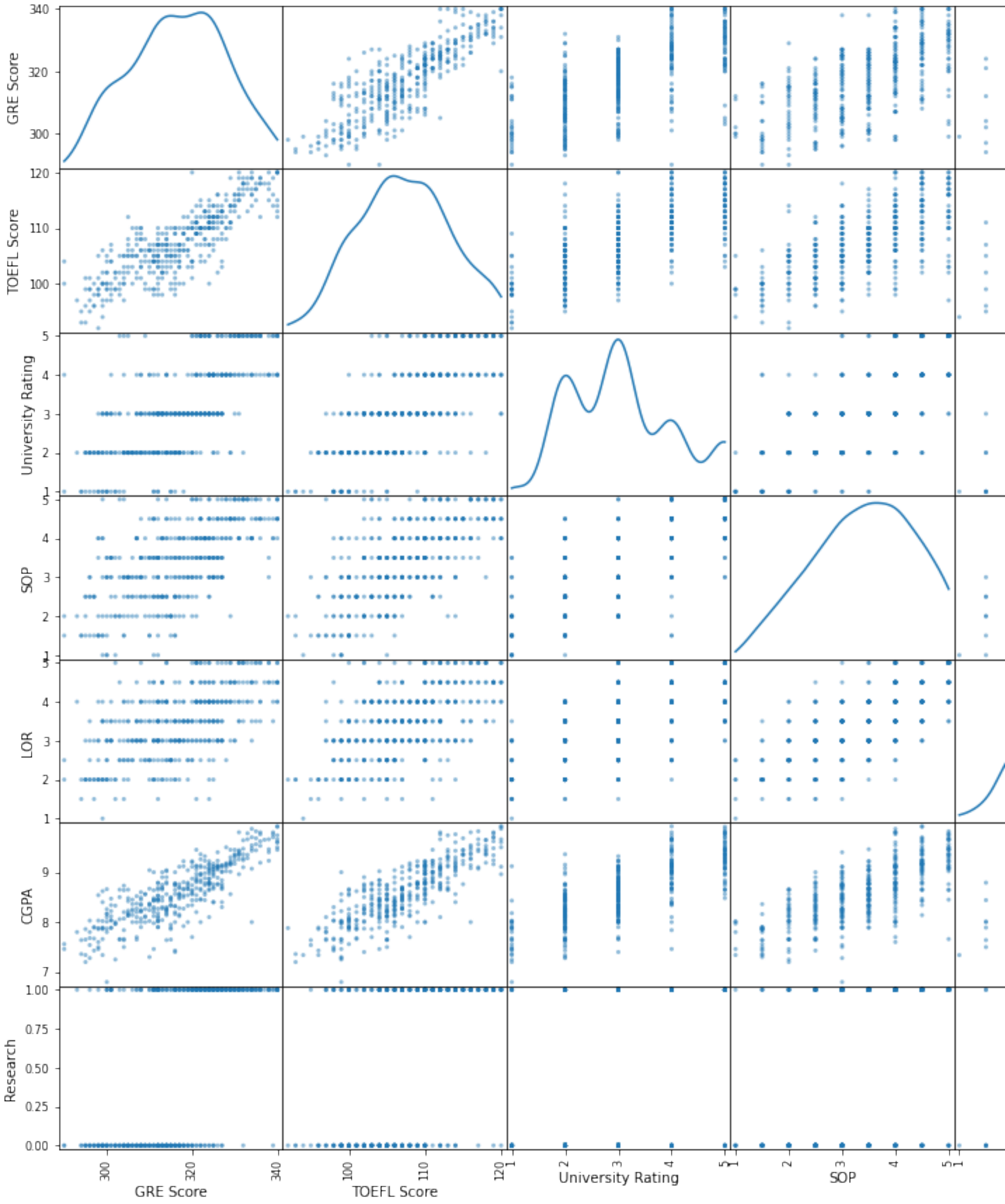
```
corr_matrix = data.corr()
plt.figure(figsize = (15, 15))
sns.heatmap(corr_matrix,annot=True,fmt='0.2f')
plt.title("Correlation Matrix", fontsize = 20)
plt.show()
```


Correlation Matrix
Serial No.
GRE Score
TOEFL Score
University Rating
SOP
LOR
CGPA
Research
Chance of Admit
1.00 -0.10 -0.15 -0.17 -0.17 -0.09 -0.05 -0.06 0.04
-0.10 1.00 0.84 0.67 0.61 0.56 0.83 0.58 0.80
-0.15 0.84 1.00 0.70 0.66 0.57 0.83 0.49 0.79
-0.17 0.67 0.70 1.00 0.73 0.66 0.75 0.45 0.71
-0.17 0.61 0.66 0.73 1.00 0.73 0.72 0.44 0.68
-0.09 0.56 0.57 0.66 0.73 1.00 0.67 0.40 0.67
-0.05 0.83 0.83 0.75 0.72 0.67 1.00 0.52 0.87
-0.06 0.58 0.49 0.45 0.44 0.40 0.52 1.00 0.55
0.04 0.80 0.79 0.71 0.68 0.67 0.87 0.55 1.00
Serial No.
GRE Score
TOEFL Score
University Rating
SOP
LOR
CGPA
Research
Chance of Admit

In [14]:

```
pd.plotting.scatter_matrix(data.loc[:,"GRE
Score":"Research"],diagonal="kde",figsize=(20,15))
plt.show
```

Out[14]:


GRE Score
TOEFL Score
University Rating
SOP
LOR
CGPA
Research
GRE Score
TOEFL Score
University Rating
SOP