In [1]:

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

In [2]:

```
file_name="C:\\Program Files\\Python311\\train_ctrUa4K (1).csv"
Loan_dataset=pd.read_csv(file_name)
```

In [3]:

```
Loan_dataset.head()
```

Out[3]:

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	Coa
0	LP001002	Male	No	0	Graduate	No	5849	
1	LP001003	Male	Yes	1	Graduate	No	4583	
2	LP001005	Male	Yes	0	Graduate	Yes	3000	
3	LP001006	Male	Yes	0	Not Graduate	No	2583	
4	LP001008	Male	No	0	Graduate	No	6000	

In [4]:

```
Loan_dataset['Education']
```

Out[4]:

```
0          Graduate
1          Graduate
2          Graduate
3      Not Graduate
4          Graduate
           ...
609        Graduate
610        Graduate
611        Graduate
612        Graduate
613        Graduate
Name: Education, Length: 614, dtype: object
```

In [6]:

```
Loan_dataset['Education'].unique()
```

Out[6]:

```
array(['Graduate', 'Not Graduate'], dtype=object)
```

In [7]:

```
Loan_dataset['Education'].nunique()
```

Out[7]:

```
2
```

In [8]:

```
len(Loan_dataset['Education'].unique())
```

Out[8]:

```
2
```

In [9]:

```
Education_values=Loan_dataset['Education'].value_counts()
```

In [10]:

```
Education_values
```

Out[10]:

```
Graduate        480
Not Graduate    134
Name: Education, dtype: int64
```

In [12]:

```
Education_values_dict=dict(Education_values )  #series type converting into dictionary ty
Education_values_dict        #keep all keys in one list and all values in another list
```

Out[12]:

```
{'Graduate': 480, 'Not Graduate': 134}
```

In [13]:

```
Education_values_dict.keys()
```

Out[13]:

```
dict_keys(['Graduate', 'Not Graduate'])
```

In [14]:

```
Education_keys=list(Education_values_dict.keys()) # list type casting
Education_keys
```

Out[14]:

```
['Graduate', 'Not Graduate']
```

In [15]:

```
Education_values_dict.values()
```

Out[15]:

```
dict_values([480, 134])
```

In [16]:

```
Education_v=list(Education_values_dict.values()) # Type cast to list
Education_v
```

Out[16]:

```
[480, 134]
```

In [17]:

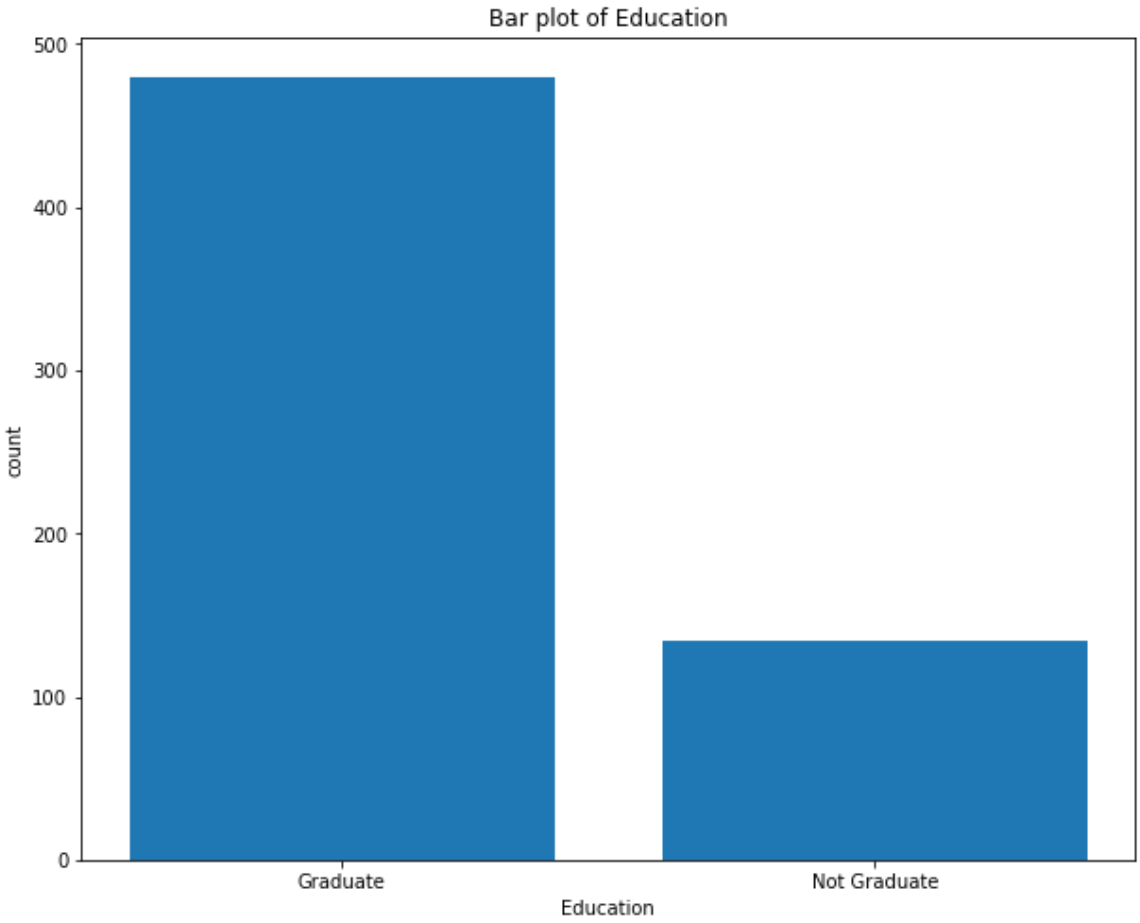
```
Education_df=pd.DataFrame(zip(Education_keys,Education_v),columns=['Education','count'])
Education_df
```

Out[17]:

	Education	count
0	Graduate	480
1	Not Graduate	134

In [19]:

```
plt.figure(figsize=(10,8))
plt.bar('Education','count',data=Education_df)
plt.xlabel('Education')
plt.ylabel('count')
plt.title("Bar plot of Education")
plt.savefig('Barplot_Education.png')  # x axis categorical column, y axis numerical colum
```



In [20]:

```
Education_df
```

Out[20]:

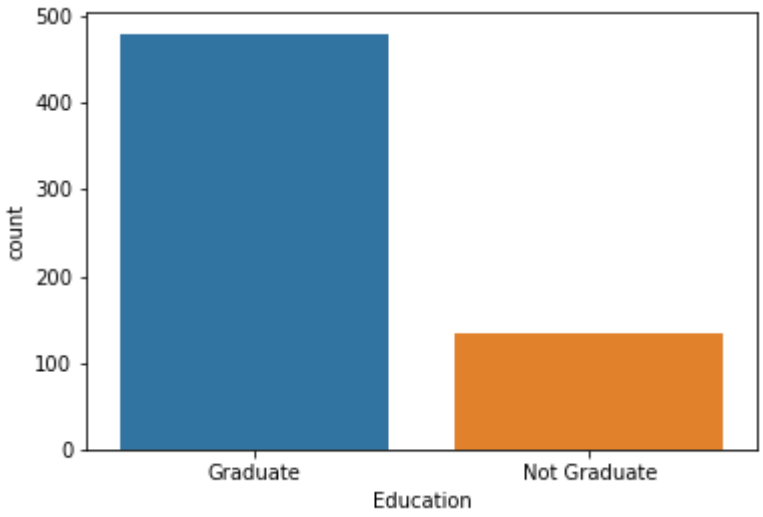
	Education	count
0	Graduate	480
1	Not Graduate	134

In [21]:

```
sns.countplot(data=Loan_dataset,x='Education')
```

Out[21]:

```
<AxesSubplot:xlabel='Education', ylabel='count'>
```



In [22]:

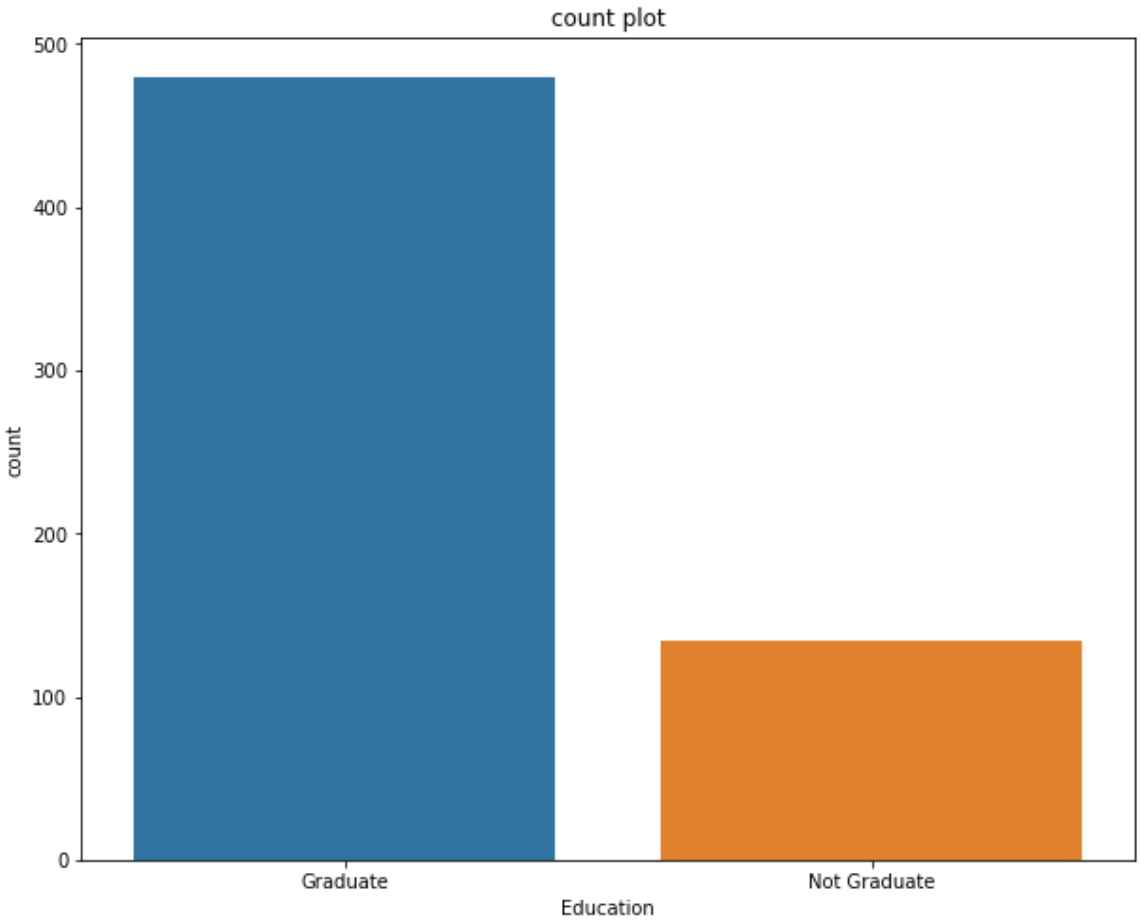
```
order_keys=Loan_dataset['Education'].value_counts().keys()
order_keys
```

Out[22]:

```
Index(['Graduate', 'Not Graduate'], dtype='object')
```

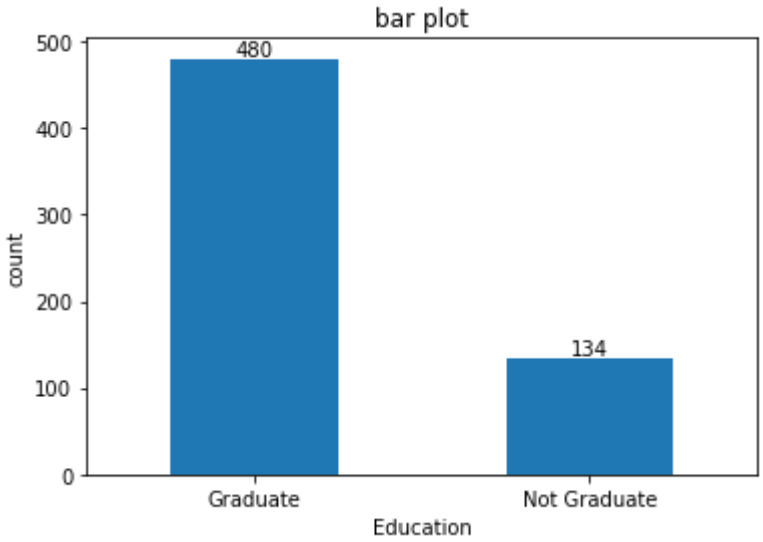
In [23]:

```
plt.figure(figsize=(10,8))
sns.countplot(data=Loan_dataset,x='Education',order=order_keys)
plt.xlabel('Education')
plt.ylabel('count')
plt.title('count plot')  # bar plot and count plot are different
plt.show()
```



In [24]:

```
value_count=Loan_dataset['Education'].value_counts()
ax=value_count.plot(kind='bar')
ax.bar_label(ax.containers[0])
plt.xlabel('Education')
plt.xticks(rotation=0)
plt.ylabel('count')
plt.title('bar plot')
plt.show()
```



In [25]:

```
value_count
```

Out[25]:

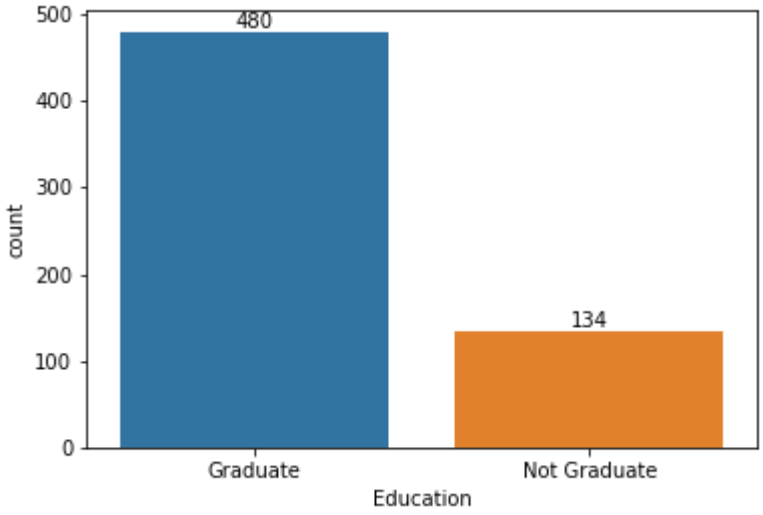
```
Graduate        480
Not Graduate    134
Name: Education, dtype: int64
```

In [26]:

```
ax=sns.countplot(data=Loan_dataset,x='Education')
ax.bar_label(ax.containers[0])
```

Out[26]:

```
[Text(0, 0, '480'), Text(0, 0, '134')]
```



In [27]:

```
Loan_dataset['Education'].value_counts(normalize=True) # provides data in percentage form
```

Out[27]:

```
Graduate        0.781759
Not Graduate    0.218241
Name: Education, dtype: float64
```

In [28]:

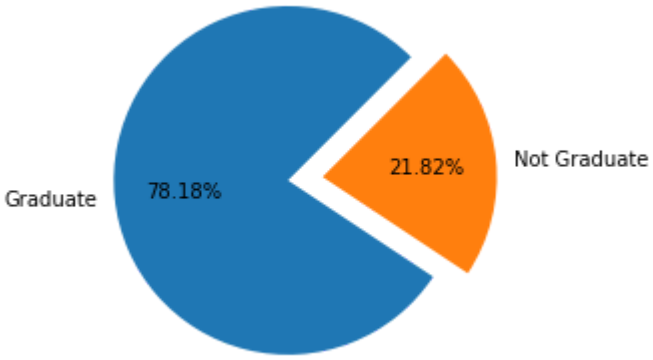
```
Loan_dataset['Education'].value_counts() # it is a series object
names=Loan_dataset['Education'].value_counts().keys()
values=Loan_dataset['Education'].value_counts().to_list()
names,values     # x and labels
```

Out[28]:

```
(Index(['Graduate', 'Not Graduate'], dtype='object'), [480, 134])
```

In [30]:

```
plt.pie(x=values,labels=names,autopct='%0.2f%%',explode=[0.1,0.1],startangle=45) # shows
plt.show()
```



In [ ]:

```

```