Importing Necessory Library

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

Importing Dataset

```
data = pd.read_csv('/content/Applications_for_Machine_Learning_internship_edited.xlsx - Sheet1.csv')
data
```

	Name	Python (out of 3)	Machine Learning (out of 3)	Natural Language Processing (NLP) (out of 3)	Deep Learning (out of 3)	Other skills	Are you available for 3 months, starting immediately, for a full-time work from home internship?	Degree	Stream	Current Year Of Graduation	Performar
0	NaN	1	0	0	1	MS-Excel, MS-Word, Deep Learning, MySQL, Pytho...	Yes, I am available for 3 months starting imme...	Bachelor of Vocation (B.Voc.)	Software Engineering	2021	
1	NaN	2	0	0	0	Git, GitHub, Linux, Adobe After Effects, Adobe...	Yes, I am available for 3 months starting imme...	B.Tech	Computer Science & Engineering	2024	
2	NaN	2	2	0	0	Amazon Web Services (AWS), Docker, Hadoop, MS-...	Yes, I am available for 3 months starting imme...	Master of Science (M.S.)	Data Science And Analytics	2022	
3	NaN	3	2	2	0	Adobe XD, BIG DATA ANALYTICS, Canva, Data Anal...	Yes, I am available for 3 months starting imme...	Bachelor of Engineering (B.E)	NaN	2024	
4	NaN	2	2	0	0	C++ Programming, Data Science, Machine Learnin...	Yes, I am available for 3 months starting imme...	B.Tech	Computer Science	2023	
...	...	...	...	...	...	...	...	...	...	...	...
1131	NaN	2	2	0	2	Data Analytics, Amazon Web Services (AWS), Dat...	Yes, I am available for 3 months starting imme...	B.Tech	Mechanical Engineering	2021	
1132	NaN	3	3	2	3	Deep Learning, Docker, HTML, MS-Office, Machin...	Yes, I am available for 3 months starting imme...	B.Tech	Computer Science & Engineering	2024	
1133	NaN	3	1	3	3	Data Science, Deep Learning, English Proficien...	Yes, I am available for 3 months starting imme...	B.Tech	Electronics and Communication	2025	
1134	NaN	2	1	0	0	Python, Data Analytics, MS-Excel, Machine	Yes, I am available for 3 months starting	B.Tech	Computer Science	2024	

Exploratory data analysis

```
data.shape
```

```
(1136, 14)
```

```
data.isnull().sum()
```

```
Name                                                                                                  1136
Python (out of 3)                                                                                        0
Machine Learning (out of 3)                                                                              0
Natural Language Processing (NLP) (out of 3)                                                             0
Deep Learning (out of 3)                                                                                 0
Other skills                                                                                            66
Are you available for 3 months, starting immediately, for a full-time work from home internship?         0
Degree                                                                                                  43
Stream                                                                                                 170
Current Year Of Graduation                                                                               0
Performance_PG                                                                                         952
Performance_UG                                                                                         654
Performance_12                                                                                         643
Performance_10                                                                                         709
dtype: int64
```

```
data.info
```

```
<bound method DataFrame.info of       Name  Python (out of 3)  Machine Learning (out of 3)  \
0      NaN                  1                           0
1      NaN                  2                           0
2      NaN                  2                           2
3      NaN                  3                           2
4      NaN                  2                           2
...    ...                ...                         ...
1131   NaN                  2                           2
1132   NaN                  3                           3
1133   NaN                  3                           1
1134   NaN                  2                           1
1135   NaN                  2                           0

      Natural Language Processing (NLP) (out of 3)  Deep Learning (out of 3)  \
0                                                0                         1
1                                                0                         0
2                                                0                         0
3                                                2                         0
4                                                0                         0
...                                            ...                       ...
1131                                             0                         2
1132                                             2                         3
1133                                             3                         3
1134                                             0                         0
1135                                             0                         0

                                         Other skills  \
0     MS-Excel, MS-Word, Deep Learning, MySQL, Pytho...
1     Git, GitHub, Linux, Adobe After Effects, Adobe...
2     Amazon Web Services (AWS), Docker, Hadoop, MS-...
3     Adobe XD, BIG DATA ANALYTICS, Canva, Data Anal...
4     C++ Programming, Data Science, Machine Learnin...
...                                                 ...
1131  Data Analytics, Amazon Web Services (AWS), Dat...
1132  Deep Learning, Docker, HTML, MS-Office, Machin...
1133  Data Science, Deep Learning, English Proficien...
1134  Python, Data Analytics, MS-Excel, Machine Lear...
1135  C++ Programming, Database Management System (D...

     Are you available for 3 months, starting immediately, for a full-time work from home internship?   \
0     Yes, I am available for 3 months starting imme...
1     Yes, I am available for 3 months starting imme...
2     Yes, I am available for 3 months starting imme...
3     Yes, I am available for 3 months starting imme...
4     Yes, I am available for 3 months starting imme...
...                                                 ...
1131  Yes, I am available for 3 months starting imme...
1132  Yes, I am available for 3 months starting imme...
1133  Yes, I am available for 3 months starting imme...
1134  Yes, I am available for 3 months starting imme...
1135  Yes, I am available for 3 months starting imme...

                            Degree                         Stream  \
0     Bachelor of Vocation (B.Voc.)          Software Engineering
1                           B.Tech  Computer Science & Engineering
2          Master of Science (M.S.)     Data Science And Analytics
3     Bachelor of Engineering (B.E)                            NaN
4                           B.Tech               Computer Science
```

```
data.describe
```

```
<bound method NDFrame.describe of       Name  Python (out of 3)  Machine Learning (out of 3)  \
0      NaN                  1                           0
1      NaN                  2                           0
2      NaN                  2                           2
3      NaN                  3                           2
4      NaN                  2                           2
...    ...                ...                         ...
1131   NaN                  2                           2
1132   NaN                  3                           3
1133   NaN                  3                           1
1134   NaN                  2                           1
1135   NaN                  2                           0

      Natural Language Processing (NLP) (out of 3)  Deep Learning (out of 3)  \
0                                                0                         1
1                                                0                         0
2                                                0                         0
3                                                2                         0
4                                                0                         0
...                                            ...                       ...
1131                                             0                         2
1132                                             2                         3
1133                                             3                         3
1134                                             0                         0
1135                                             0                         0

                                        Other skills  \
0     MS-Excel, MS-Word, Deep Learning, MySQL, Pytho...
1     Git, GitHub, Linux, Adobe After Effects, Adobe...
2     Amazon Web Services (AWS), Docker, Hadoop, MS-...
3     Adobe XD, BIG DATA ANALYTICS, Canva, Data Anal...
4     C++ Programming, Data Science, Machine Learnin...
...                                                 ...
1131  Data Analytics, Amazon Web Services (AWS), Dat...
1132  Deep Learning, Docker, HTML, MS-Office, Machin...
1133  Data Science, Deep Learning, English Proficien...
1134  Python, Data Analytics, MS-Excel, Machine Lear...
1135  C++ Programming, Database Management System (D...

     Are you available for 3 months, starting immediately, for a full-time work from home internship?   \
0     Yes, I am available for 3 months starting imme...
1     Yes, I am available for 3 months starting imme...
2     Yes, I am available for 3 months starting imme...
3     Yes, I am available for 3 months starting imme...
4     Yes, I am available for 3 months starting imme...
...                                                 ...
1131  Yes, I am available for 3 months starting imme...
1132  Yes, I am available for 3 months starting imme...
1133  Yes, I am available for 3 months starting imme...
1134  Yes, I am available for 3 months starting imme...
1135  Yes, I am available for 3 months starting imme...

                            Degree                          Stream  \
0     Bachelor of Vocation (B.Voc.)           Software Engineering
1                           B.Tech  Computer Science & Engineering
2         Master of Science (M.S.)      Data Science And Analytics
3     Bachelor of Engineering (B.E)                             NaN
4                           B.Tech               Computer Science
```

```
data.dtypes
```

```
Name                                                                                                  float64
Python (out of 3)                                                                                       int64
Machine Learning (out of 3)                                                                             int64
Natural Language Processing (NLP) (out of 3)                                                            int64
Deep Learning (out of 3)                                                                                int64
Other skills                                                                                           object
Are you available for 3 months, starting immediately, for a full-time work from home internship?       object
Degree                                                                                                 object
Stream                                                                                                 object
Current Year Of Graduation                                                                              int64
Performance_PG                                                                                         object
Performance_UG                                                                                         object
Performance_12                                                                                         object
Performance_10                                                                                         object
dtype: object
```

```
data['Stream'].unique()
```

```
array(['Software Engineering', 'Computer Science & Engineering',
       'Data Science And Analytics', nan, 'Computer Science',
       'Statistics',
       'Electronics and Telecommunication Engineering (ETE)',
       'Data Science And Machine Learning',
       'Information and Communication Technology', 'computer science',
       'Analytics And Finance', 'Information Technology', 'Commerce',
```

```
              'Instrumentation and Control Engineering',
              'Electronics and Communication',
              'Electrical and Electronics Engineering', 'Computer Applications',
              'Data Science And Business Analytics', 'Data Science', 'Finance',
              'Networks And Communication',
              'Statistics Computing( DATA Science )',
              'Computer  Science And Engineering', 'AIML',
              'Metallurgical And Materials Engineering',
              'Electronics Engineering', 'CSE Hons DS&AI', 'Physics',
              'Computer Sciences &Engineering', 'Computational Data Science',
              'Data Analytics', 'Computer Application', 'Python,sql',
              'Mechanical Engineering', 'Biomedical Engineering',
              'Computer Science And Engineering (Data Science)', 'Science',
              'Electrical Engineering',
              'Artificial Intelligence And Data Science', 'Mathematics',
              'Geography', 'Marketing', 'All Stream', 'Computer Science And IT',
              'Artificial Intelligence, Machine Learning And Robotics',
              'Civil Engineering', 'Food Technology',
              'Electronics and Instrumentation ',
              'Artificial Intelligence And Data Sciences', 'Cs(aiml)',
              'Artificial Intelligence',
              'CSE( Artificial Intelligence And Machine Learning )',
              'Computer Application(Data Science)', 'Nanotechnology',
              'Data Science And AI', 'AI & DATA SCIENCE',
              'Computer Science & Engineering(artificial Intelligence And Machine Learning)',
              'Arts', 'Mathemtaics With Computer Science',
              'Computer Science With Specialization In Data Science',
              'Data Science And Statistics',
              'Artificial Intelligence & Internet Of Things', 'Botany',
              'Business Analytics',
              'Artificial Intelligence And Machine Learning', 'IT(Data Science)',
              'Economics', 'COMPUTER', 'Metallurgical & Materials Engineering',
              'Computer Science & Technology', 'Business Development', 'cs',
              'B.E. Computer', 'Mechanical', 'Mathematics and Computing',
              'Computer Science And Technology', 'E&TC', 'Electronic Science',
              'Chemical Engineering', 'Statistics With Data Science', 'AIDS',
              'Computer Science And Engineering(data Science)',
              'communication engineering', 'AI/DA', 'Agriculture Management',
              'Ai&Ml', 'Information Technology (Data Science)',
              'Information science and engineering',
              'Metallurgical And Materials Science', 'Engineering Physics',
              'Statistics And Analytics', 'Electrical', 'BUSINESS ANALYST',
              'Mining Engineering',
              'Computer Science With Specialization In AI And ML',
              'Signal Processing Machine Learning',
              'Artificial Intelligence And Data Science Engineering',
              'Agricultural Science', 'Information Science',
              'Industrial Management', 'Machine Learning', 'Engineering',
              'Commerce With Maths', 'IT',
              'Computer & Communication Engineering',
              'Computer Science & Engineering(Artificial Intelligence)', 'AI',
```

```
data['Stream'].value_counts()
```

```
Computer Science & Engineering     220
Computer Science                    88
computer science                    68
Information Technology              60
Data Science                        56
                                  ...
E&TC                                 1
Computer Science And Technology      1
B.E. Computer                        1
Business Development                 1
Computational Mechanics              1
Name: Stream, Length: 186, dtype: int64
```

```
data['Other skills'].value_counts()
```

```
Machine Learning, Python
4
Python
4
Data Analytics, MS-Excel, Machine Learning, Python, SQL
3
Data Science, Machine Learning, Python
2
Data Science, Python, SQL
2

..
Data Science, Deep Learning, Machine Learning, Python, Java, ReactJS
1
MS-Excel, MS-Office, MySQL, Data Analytics, Data Science, Machine Learning, Neural Networks, Power BI, Python, SQL,
Deep Learning, Natural Language Processing (NLP)
1
English Proficiency (Spoken), English Proficiency (Written), Marathi Proficiency(Spoken), Adobe Premiere Pro, HTML, MS-
Excel, MS-PowerPoint, MS-Word, Report Writing, Statistical Modeling, Video Editing, CSS, MS-Office, MySQL, PHP, Python
```

```
1
Python, Algorithms, Artificial Intelligence, C++ Programming, Deep Learning, Machine Learning, Marketing Campaigns ,
Operations, SQL, Image Processing, Natural Language Processing (NLP)
1
C++ Programming, Database Management System (DBMS), Java, JavaScript, Python, C Programming, Data Science
1
Name: Other skills, Length: 1057, dtype: int64
```

```
data['Performance_12'].value_counts()
```

```
78.00/78.00    13
95.00/95.00    10
90.00/90.00     8
91.00/91.00     8
80.00/80.00     8
               ..
77.40/77.40     1
62.92/62.92     1
82.15/82.15     1
71.69/71.69     1
82.80/82.80     1
Name: Performance_12, Length: 266, dtype: int64
```

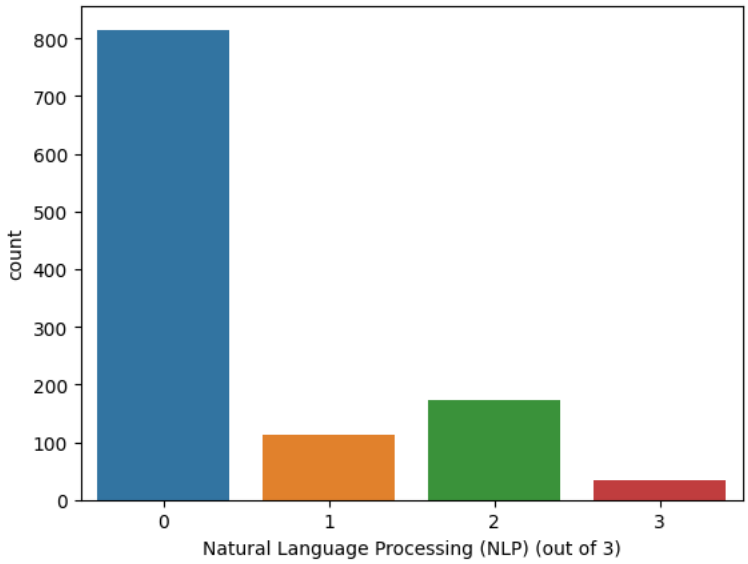
Data visulization

```
data['Natural Language Processing (NLP) (out of 3)'].value_counts()
```

```
0    815
2    174
1    112
3     35
Name: Natural Language Processing (NLP) (out of 3), dtype: int64
```

```
sns.countplot(x=data['Natural Language Processing (NLP) (out of 3)'])
```

```
<Axes: xlabel='Natural Language Processing (NLP) (out of 3)', ylabel='count'>
```




```
data['Python (out of 3)'].value_counts()
```

```
2    624
3    220
0    164
1    128
Name: Python (out of 3), dtype: int64
```

```
sns.countplot(x=data['Python (out of 3)'])
```

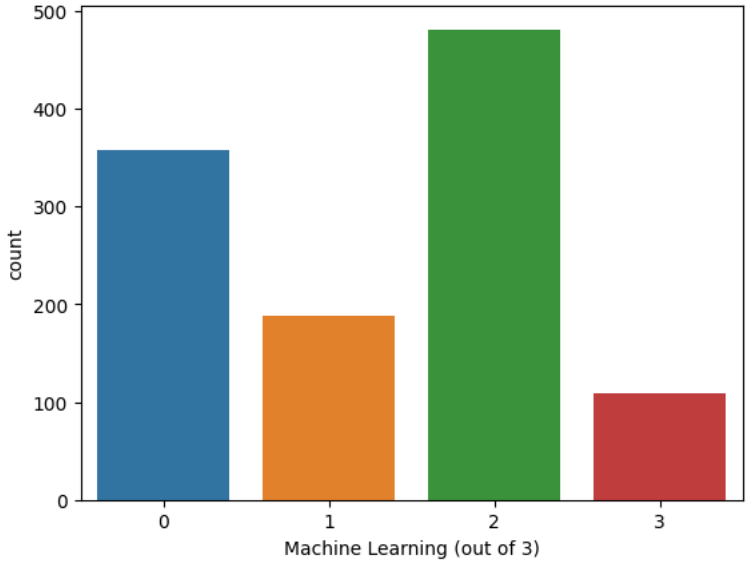
```
<Axes: xlabel='Python (out of 3)', ylabel='count'>
```


```
data['Machine Learning (out of 3)'].value_counts()
```

```
2    481
0    358
1    188
3    109
Name: Machine Learning (out of 3), dtype: int64
```

```
sns.countplot(x=data['Machine Learning (out of 3)'])
```

```
<Axes: xlabel='Machine Learning (out of 3)', ylabel='count'>
```




```
data['Natural Language Processing (NLP) (out of 3)'].value_counts()
```

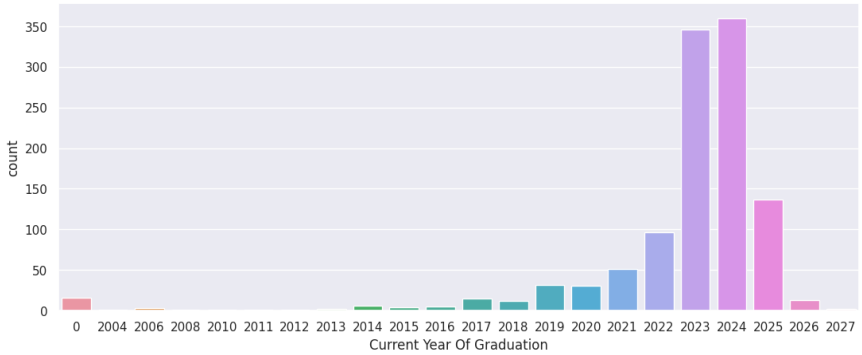
```
0    815
2    174
1    112
3     35
Name: Natural Language Processing (NLP) (out of 3), dtype: int64
```

```
sns.countplot(x=data['Natural Language Processing (NLP) (out of 3)'])
```

```
<Axes: xlabel='Natural Language Processing (NLP) (out of 3)', ylabel='count'>
```

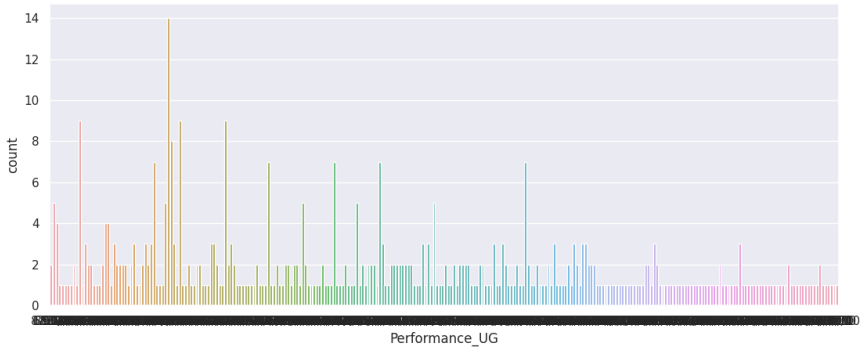

```
sns.set(rc={"figure.figsize":(13, 5)})
sns.countplot(x=data['Current Year Of Graduation'])
```

```
<Axes: xlabel='Current Year Of Graduation', ylabel='count'>
```




```
sns.countplot(x=data['Performance_UG'])
```

```
<Axes: xlabel='Performance_UG', ylabel='count'>
```



Take necessory Columns for Our model

```
df = data[['Python (out of 3)','Machine Learning (out of 3)','Natural Language Processing (NLP) (out of 3)','Deep Learning (
df
```

	Python (out of 3)	Machine Learning (out of 3)	Natural Language Processing (NLP) (out of 3)	Deep Learning (out of 3)	Current Year Of Graduation	Performance_UG	Perfo
0	1	0	0	1	2021	6.50/7	
1	2	0	0	0	2024	8.90/10	
2	2	2	0	0	2022	NaN	

Fill Null value Columns

```
df['Performance_UG'].fillna('0',inplace=True)
df['Performance_12'].fillna('0',inplace=True)
```

```
<ipython-input-30-5baa9644963c>:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-
  df['Performance_UG'].fillna('0',inplace=True)
<ipython-input-30-5baa9644963c>:2: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-
  df['Performance_12'].fillna('0',inplace=True)
```

```
df.isnull().sum()
```

```
Python (out of 3)                             0
Machine Learning (out of 3)                   0
Natural Language Processing (NLP) (out of 3)  0
Deep Learning (out of 3)                      0
Current Year Of Graduation                    0
Performance_UG                                0
Performance_12                                0
dtype: int64
```

```
df
```

	Python (out of 3)	Machine Learning (out of 3)	Natural Language Processing (NLP) (out of 3)	Deep Learning (out of 3)	Current Year Of Graduation	Performance_UG	Perfo
0	1	0	0	1	2021	6.50/7	
1	2	0	0	0	2024	8.90/10	
2	2	2	0	0	2022	0	
3	3	2	2	0	2024	0	
4	2	2	0	0	2023	8.10/10	
...	...	...	...	...	...	...	
1131	2	2	0	2	2021	0	
1132	3	3	2	3	2024	0	
1133	3	1	3	3	2025	8.77/10	
1134	2	1	0	0	2024	7.90/10	

```
df.dtypes
```

```
Python (out of 3)                              int64
Machine Learning (out of 3)                    int64
Natural Language Processing (NLP) (out of 3)   int64
Deep Learning (out of 3)                       int64
Current Year Of Graduation                     int64
Performance_UG                                object
Performance_12                                object
dtype: object
```

Modifiey String Columns

```
df['Performance_UG'] = df['Performance_UG'].str.replace('[\/]',"")
```

```
<ipython-input-34-da6dc128a1ba>:1: FutureWarning: The default value of regex will change from True to False in a future
  df['Performance_UG'] = df['Performance_UG'].str.replace('[\/]',"")
<ipython-input-34-da6dc128a1ba>:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-
  df['Performance_UG'] = df['Performance_UG'].str.replace('[\/]',"")
```

```
df['Performance_UG'] = pd.to_numeric(df['Performance_UG'])
```

```
<ipython-input-37-5a809f66e33b>:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-
  df['Performance_UG'] = pd.to_numeric(df['Performance_UG'])
```

```
df['Performance_12'] = df['Performance_12'].str.replace('[\/]',"")
```

```
<ipython-input-35-badd2e1e88a8>:1: FutureWarning: The default value of regex will change from True to False in a future
  df['Performance_12'] = df['Performance_12'].str.replace('[\/]',"")
<ipython-input-35-badd2e1e88a8>:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-
  df['Performance_12'] = df['Performance_12'].str.replace('[\/]',"")
```

```
df['Performance_12'] = df['Performance_12'].str.extract(r'(\d+\.\d+)', expand=False)
df['Performance_12'] = pd.to_numeric(df['Performance_12'])
```

```
<ipython-input-39-2ba183ec5526>:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-
  df['Performance_12'] = df['Performance_12'].str.extract(r'(\d+\.\d+)', expand=False)
<ipython-input-39-2ba183ec5526>:2: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-
  df['Performance_12'] = pd.to_numeric(df['Performance_12'])
```

```
df.dtypes
```

```
Python (out of 3)                                 int64
Machine Learning (out of 3)                       int64
Natural Language Processing (NLP) (out of 3)      int64
Deep Learning (out of 3)                          int64
Current Year Of Graduation                        int64
Performance_UG                                  float64
Performance_12                                  float64
dtype: object
```

```
df.isnull().sum()
```

```
Python (out of 3)                                   0
Machine Learning (out of 3)                         0
Natural Language Processing (NLP) (out of 3)        0
Deep Learning (out of 3)                            0
Current Year Of Graduation                          0
Performance_UG                                      0
Performance_12                                    643
dtype: int64
```

Convert DataFrame to Array

```
X = df.values
```

```
X
```

```
array([[1.0000e+00, 0.0000e+00, 0.0000e+00, ..., 2.0210e+03, 6.5070e+00,
                nan],
       [2.0000e+00, 0.0000e+00, 0.0000e+00, ..., 2.0240e+03, 8.9010e+00,
                nan],
       [2.0000e+00, 2.0000e+00, 0.0000e+00, ..., 2.0220e+03, 0.0000e+00,
```

```
        nan],
       ...,
       [3.0000e+00, 1.0000e+00, 3.0000e+00, ..., 2.0250e+03, 8.7710e+00,
        9.4090e+00],
       [2.0000e+00, 1.0000e+00, 0.0000e+00, ..., 2.0240e+03, 7.9010e+00,
        9.0009e+01],
       [2.0000e+00, 0.0000e+00, 0.0000e+00, ..., 2.0240e+03, 0.0000e+00,
        nan]])
```

Future Scaller

```
from sklearn.preprocessing import MinMaxScaler

ms = MinMaxScaler()

X = ms.fit_transform(X)
```

Find The Number Of CLuster Using Elbow Method

```
import matplotlib.pyplot as plt
from sklearn.cluster import KMeans
from sklearn.datasets import make_blobs

wcss = []
for k in range(1, 11):
    kmeans = KMeans(n_clusters=k, random_state=0)
    kmeans.fit(X)
    wcss.append(kmeans.inertia_)

plt.plot(range(1, 11), wcss)
plt.title('Elbow Method')
plt.xlabel('Number of Clusters (k)')
plt.ylabel('WCSS')
plt.show()
```

```
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870: Futur
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870: Futur
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870: Futur
```

Use K Means Cluster Model

```
  warnings.warn(
```

```
kmeans = KMeans(n_clusters=4, random_state=0)
kmeans.fit(X)
```

```
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870: Futur
  warnings.warn(
```

```
KMeans
KMeans(n_clusters=4, random_state=0)
```

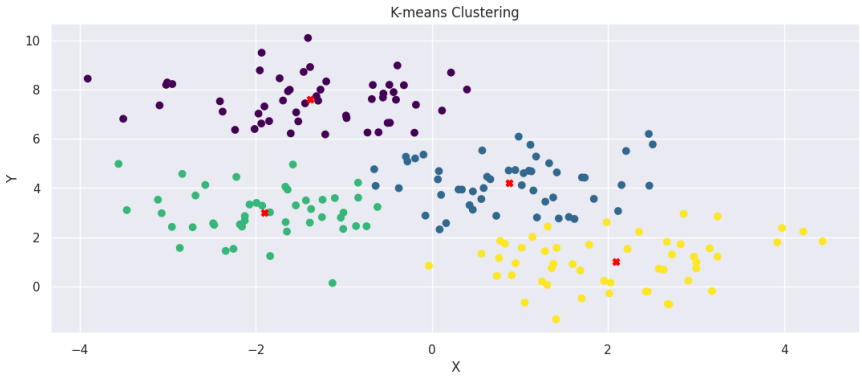
```
labels = kmeans.fit_predict(X)
```

```
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870: FutureWarning: The default value of `n_init` wil
  warnings.warn(
```

```
centroids = kmeans.cluster_centers_
```

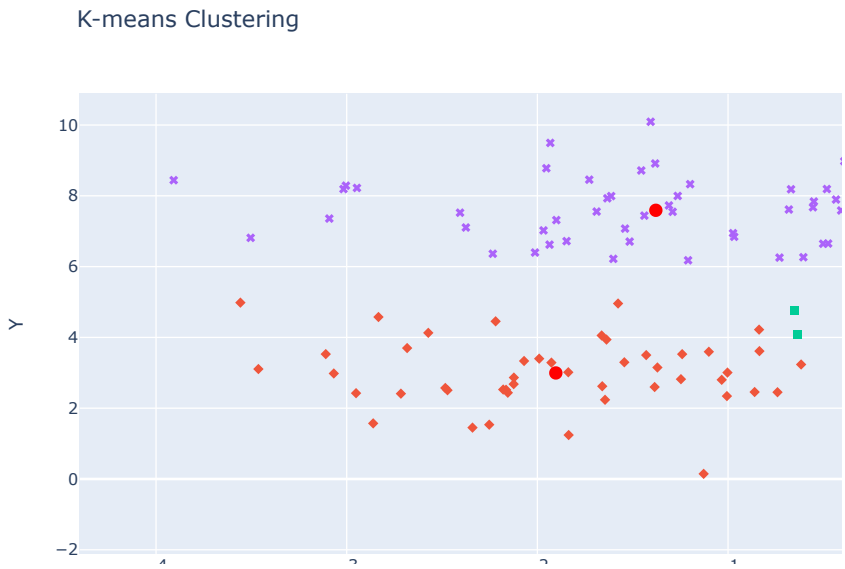
Visualize The Model

```
plt.scatter(X[:, 0], X[:, 1], c=labels, cmap='viridis')
plt.scatter(centroids[:, 0], centroids[:, 1], marker='X', c='red')
plt.title('K-means Clustering')
plt.xlabel('X')
plt.ylabel('Y')
plt.show()
```



Advanced Visulization Of the model

```
import plotly.express as px
df = pd.DataFrame(X, columns=['X', 'Y'])
df['Label'] = labels.astype(str)
fig = px.scatter(df, x='X', y='Y', color='Label', symbol='Label',title='K-means Clustering', hover_data=df.columns)
fig.add_scatter(x=centroids[:, 0], y=centroids[:, 1], mode='markers',marker=dict(size=10, color='red'), name='Centroids')
fig.show()
```



Check The Accouracy Of the model Using ARI and NMI

(To evaluate the clustering accuracy, we calculate the ARI and NMI scores using the adjusted_rand_score and normalized_mutual_info_score functions from scikit-learn. The ARI measures the similarity between the predicted clusters and the ground truth labels, while the NMI measures the mutual information between the two sets of labels, taking into account label permutations and the data distribution.)

```
from sklearn.metrics import adjusted_rand_score, normalized_mutual_info_score
X, y_true = make_blobs(n_samples=200, centers=4, random_state=0)
ari = adjusted_rand_score(y_true, labels)
nmi = normalized_mutual_info_score(y_true, labels)
print("Adjusted Rand Index (ARI):", ari)
print("Normalized Mutual Information (NMI):", nmi)
```

```
Adjusted Rand Index (ARI): 0.8348329935674467
Normalized Mutual Information (NMI): 0.8101793178662472
```

Overall, The Model gave 83% accuracy result that is pretty much good model. It may gave more good result if DateSet may incressed and also maybe gave good result to other model.