chola-batch-2-application-details

October 26, 2023

Senchola Batch 2 Application Details

Import Data from excel file

```
[36]: import numpy as np
      import pandas as pd
      import seaborn as sns
      import matplotlib.pyplot as plt
      from wordcloud import WordCloud
```

```
[37]: df=pd.read_excel('/content/Batch-2 Senchola application details - Data cleaned.
      ↪xlsx')
```

```
[38]: df.head()
```

```
[38]:                 Timestamp Are you open to learn ?  \
      0 2023-05-30 18:51:45.391                      Yes
      1 2023-05-30 18:56:19.324                      Yes
      2 2023-05-30 18:57:32.362                      Yes
      3 2023-05-30 19:00:37.964                      Yes
      4 2023-05-30 19:11:38.697                      Yes

                    Why you want to join this program ? Do you have laptop   \
      0                         To upgrade my skills                    Yes
      1                      Eager to learn and work.                   Yes
      2  I want to need and explore the real-time exper…              Yes
      3  I am interested in joining this program becaus…              Yes
      4  To learn new technology and explore my knowled…             Yes

                  Name                                         Address  \
      0    Aishwarya G  1/104,Reddiyar strt, alathudaiyanpatty,Thuraiy…
      1         Vidhya            Mandabam street, rayagiri, thenkasi.
      2        Ashen A  212/2 Sinthannagar 1st Street,Krishnampalayam,…
      3      Azhagar M  4/498, South Street, vaippam, Ariyalur Distric…
      4  Keerthivaasan M  43/d Chinnaswamy naidu street kadathur,Dharmap…

           Qualification                                       Degree  \
      0     Engineering  B.E Electronics and communication engineering
```

```
1        Others                                           BCA
2  Arts & Science                                    B.A English
3  Arts & Science                                           BCA
4  Arts & Science                              M.sc.Data Analytics

  Pass-out Year                                College Name  \
0          2021  Dhanalakshmi srinivasan engineering college,pe…
1          2023                               The American College
2          2021                      Nandha Arts And Science College
3          2023  Thanthai Hans Roever college of arts and scien…
4          2023                              Bharathiar university.

          City        State What you wan to learn ? Technical Feeback by  \
0   Perambalur    Tamilnadu     Frontend Development             Vignesh
1      Madurai    Tamilnadu             UI/UX Design             Vignesh
2        Erode    Tamilnadu      Backend Development             Vignesh
3      Ariyalur    Tamilnadu     Backend Development             Vignesh
4  Dharmapuri   Tamil Nadu            Cybersecurity             Vignesh

                                 Technical Comments  \
0                                                  s
1                                      Not pic call
2                                      Not pic call
3  He is interested in Python Backend Developer, …
4  He is interested in Cyber security and did a c…

                                   HR Comments
0                     she wants to gain knowledge
1                                     no response
2                                     no response
3  8870956961 please add his  no.  on the whatsap…
4  9894233004 add his no. in  whatsapp group, he …
```

[39]: `df.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 387 entries, 0 to 386
Data columns (total 16 columns):
 #   Column                         Non-Null Count  Dtype
---  ------                         --------------  -----
 0   Timestamp                      387 non-null    datetime64[ns]
 1   Are you open to learn ?        387 non-null    object
 2   Why you want to join this program ?  387 non-null    object
 3   Do you have laptop             387 non-null    object
 4   Name                           387 non-null    object
 5   Address                        387 non-null    object
 6   Qualification                  387 non-null    object
```

```
7    Degree                              387 non-null    object
8    Pass-out Year                       387 non-null    object
9    College Name                        387 non-null    object
10   City                                387 non-null    object
11   State                               387 non-null    object
12   What you wan to learn ?             387 non-null    object
13   Technical Feeback by                322 non-null    object
14   Technical Comments                  322 non-null    object
15   HR Comments                         60 non-null     object
dtypes: datetime64[ns](1), object(15)
memory usage: 48.5+ KB
```

[40]: `df.isnull().sum()`

```
[40]: Timestamp                           0
      Are you open to learn ?             0
      Why you want to join this program ?   0
      Do you have laptop                  0
      Name                                0
      Address                             0
      Qualification                       0
      Degree                              0
      Pass-out Year                       0
      College Name                        0
      City                                0
      State                               0
      What you wan to learn ?             0
      Technical Feeback by               65
      Technical Comments                 65
      HR Comments                       327
      dtype: int64
```

[41]: `df.columns`

```
[41]: Index(['Timestamp', 'Are you open to learn ?',
             'Why you want to join this program ?', 'Do you have laptop ', 'Name ',
             'Address', 'Qualification', 'Degree', 'Pass-out Year', 'College Name',
             'City', 'State', 'What you wan to learn ?', 'Technical Feeback by',
             'Technical Comments', 'HR Comments'],
            dtype='object')
```

0.1 Insights

1. Total number of Applicants

[42]:
```python
total_applicants=df['Name '].count()
print("Total number of Applicants:" ,total_applicants)
```

Total number of Applicants: 387

2.Gender classification

```
[43]: df1= pd.read_excel('/content/Batch-2 Senchola application details - Data␣
      ↪cleaned.xlsx', sheet_name='Name & Gender')
      gender_counts = df1['Gender'].value_counts()
      print(gender_counts)
```

```
Male      298
Female     89
Name: Gender, dtype: int64
```

3.Openness to learn

```
[44]: openness_to_learn = df['Are you open to learn ?'].value_counts()
      colors = ['green', 'red', 'yellow']
      # Create a pie chart
      plt.figure(figsize=(5,5))
      plt.pie(openness_to_learn, labels=openness_to_learn.index, autopct='%1.1f%%',␣
      ↪startangle=140, colors=colors)
      plt.title('Openness to Learn')
      plt.show()
```
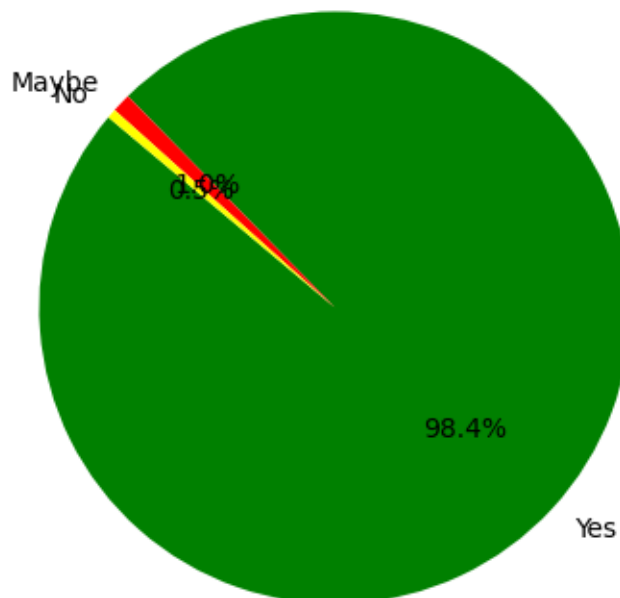
4.Motivation for Join

```
[45]: text_data = ' '.join(df['Why you want to join this program ?'])

      # Create a WordCloud object
      wordcloud = WordCloud(width=800, height=400, background_color='white').
        ↪generate(text_data)

      # Display the WordCloud image
      plt.figure(figsize=(5, 5))
      plt.imshow(wordcloud, interpolation='bilinear')
      plt.axis('off')
      plt.show()
```
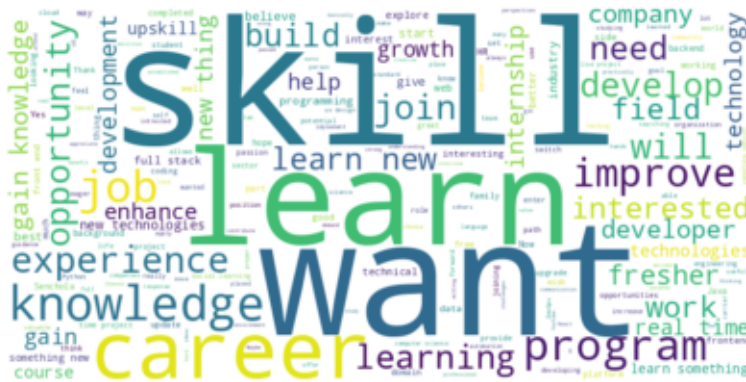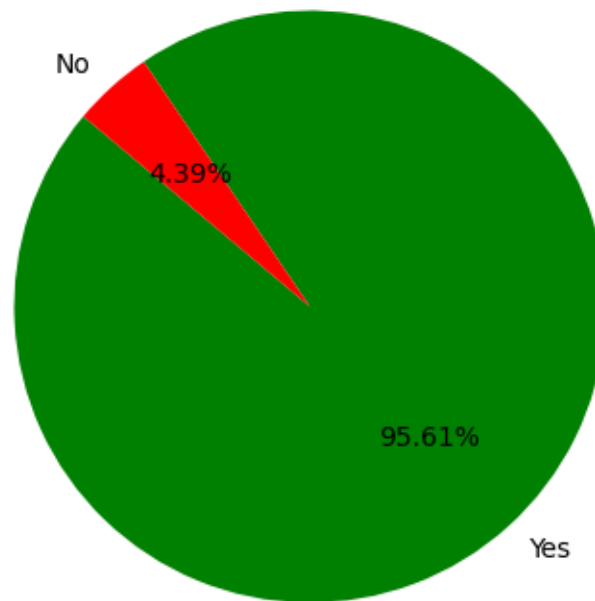

5.Laptop Availability

```
[51]: laptop_availability = df['Do you have laptop '].value_counts()

      # Define custom colors
      colors = ['green', 'red']

      # Create a pie chart
      plt.figure(figsize=(5,5))
      plt.pie(laptop_availability, labels=laptop_availability.index, autopct='%2.
        ↪2f%%', startangle=140, colors=colors)
      plt.title('Laptop Availability')
      plt.show()
```

Laptop Availability

6.Qualification Distribution

```
[52]: qualification_distribution = df['Qualification'].value_counts()

plt.figure(figsize=(5,5))
plt.bar(qualification_distribution.index, qualification_distribution.values,⊔
 ↪color='blue')
plt.title('Qualification Distribution')
plt.xlabel('Qualification Level')
plt.ylabel('Count')
plt.xticks(rotation=45, ha="right")

for i, value in enumerate(qualification_distribution.values):
    plt.text(i, value + 0.2, str(value), ha='center', va='bottom')

plt.tight_layout()
plt.show()
```

Qualification Distribution
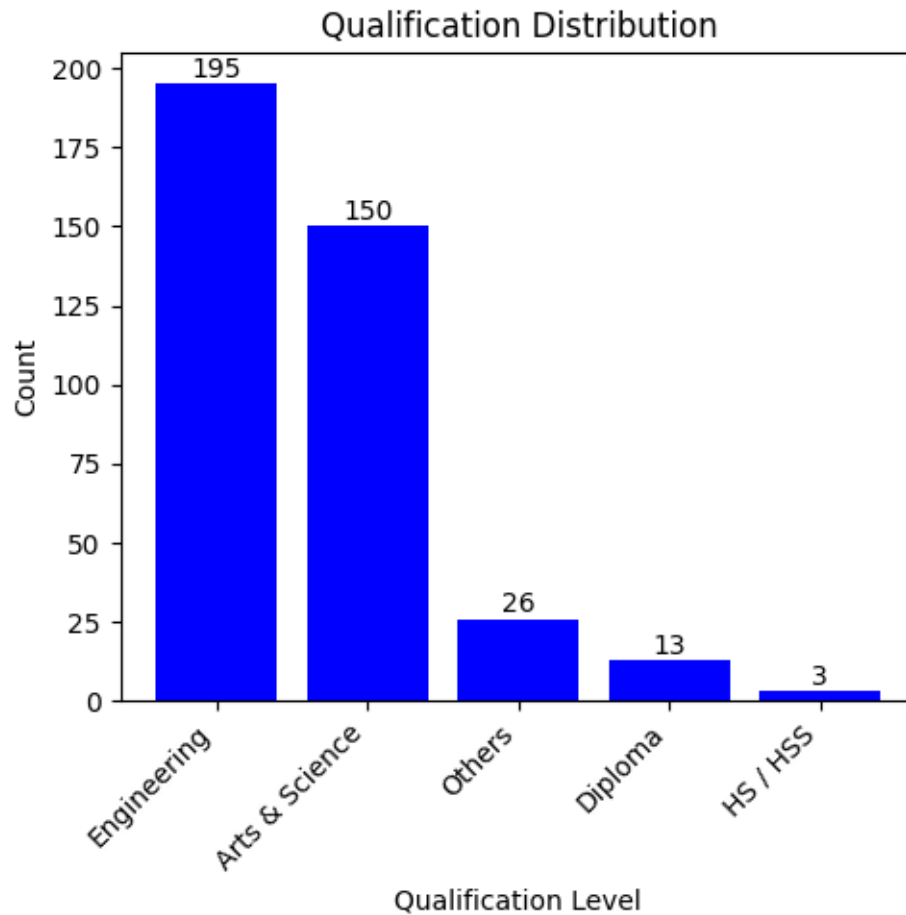
7.Popular Degree

```
[54]:  degree_distribution = df['Degree'].str.upper().value_counts()


       top_degrees = degree_distribution.head(10)


       plt.figure(figsize=(5,5))
       plt.bar(top_degrees.index, top_degrees.values, color='blue')
       plt.title('Popular Degrees')
       plt.xlabel('Degree')
       plt.ylabel('Count')
       plt.xticks(rotation=45, ha="right")


       for i, value in enumerate(top_degrees.values):
           plt.text(i, value + 0.2, str(value), ha='center', va='bottom')
```

```
plt.tight_layout()
plt.show()
```


Popular Degrees
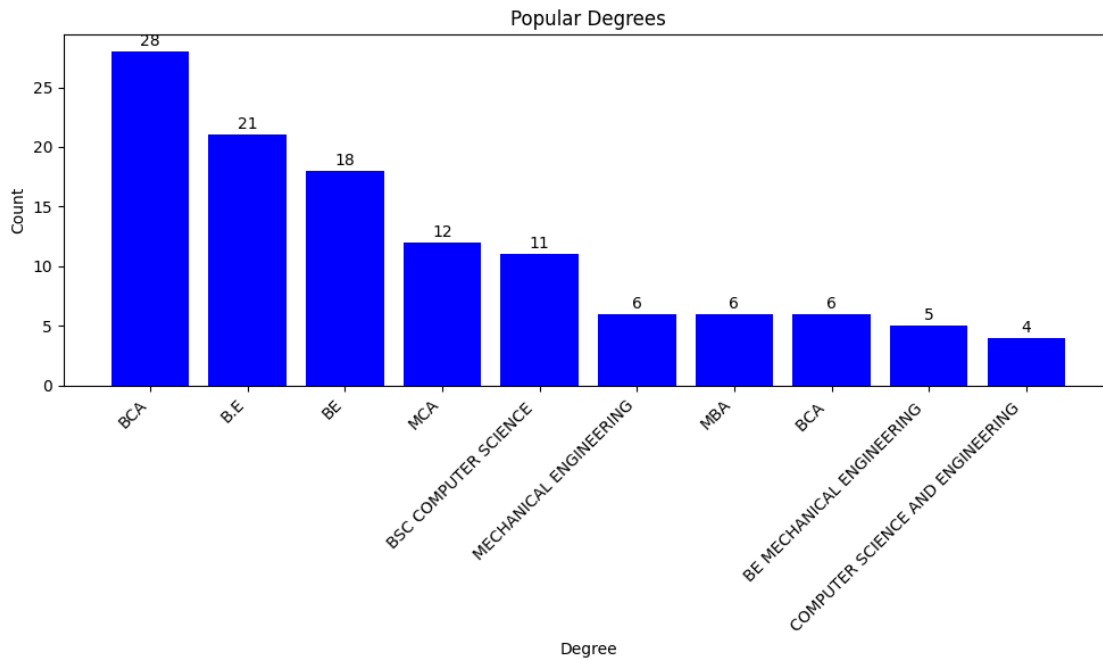
8.Gradution Year

```
[58]: df['Pass-out Year'] = pd.to_numeric(df['Pass-out Year'], errors='coerce')

      df = df.dropna(subset=['Pass-out Year'])

      pass_out_years = df['Pass-out Year'].value_counts().sort_index()

      plt.figure(figsize=(5,5))
      plt.plot(pass_out_years.index, pass_out_years.values, marker='o', color='blue',␣
       ↪linestyle='-')
      plt.title('Graduation Year Distribution')
      plt.xlabel('Year')
      plt.ylabel('Count')
      plt.xticks(rotation=45)
      for i, value in enumerate(pass_out_years.values):
          plt.text(pass_out_years.index[i], value + 1, str(value), ha='center',␣
       ↪va='bottom', fontsize=10, color='black')
      plt.tight_layout()
      plt.show()
```

## Graduation Year Distribution
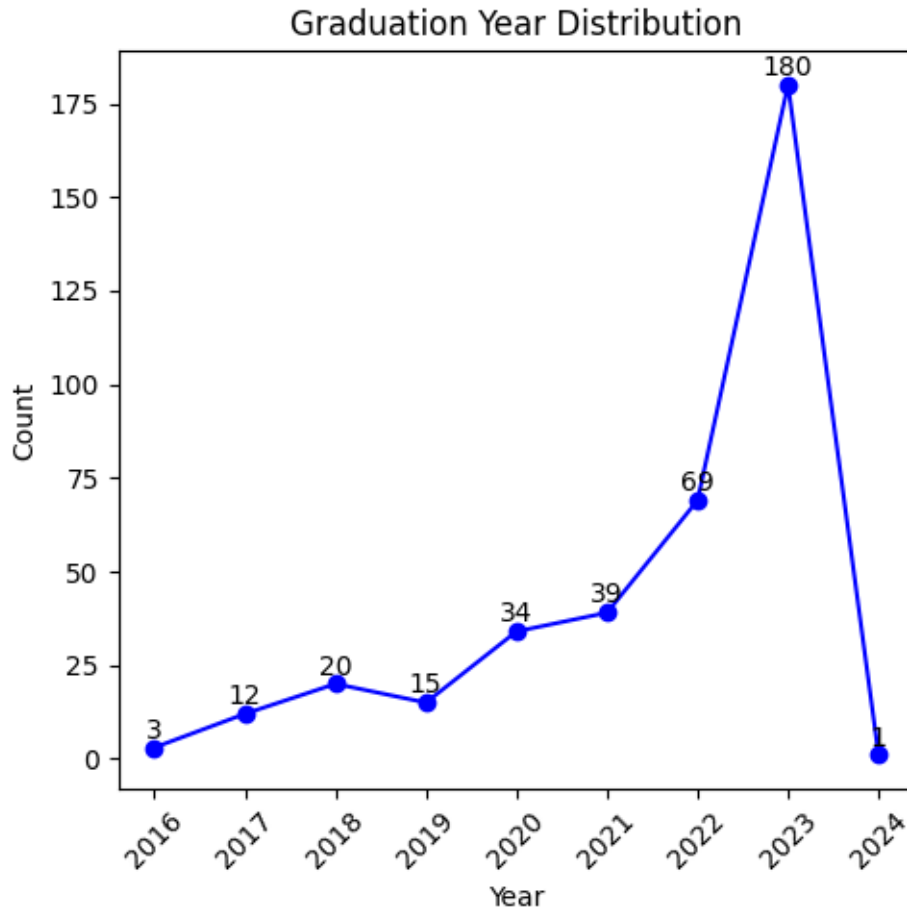
9.Top Colleges

```
[59]: top_colleges = df['College Name'].value_counts().head(10)

      plt.figure(figsize=(5,5))
      plt.bar(top_colleges.index, top_colleges.values, color='blue')
      plt.title('Top Colleges')
      plt.xlabel('College Name')
      plt.ylabel('Count')
      plt.xticks(rotation=45, ha="right")

      for i, value in enumerate(top_colleges.values):
          plt.text(i, value + 0.2, str(value), ha='center', va='bottom')

      plt.tight_layout()
      plt.show()
```

Top Colleges
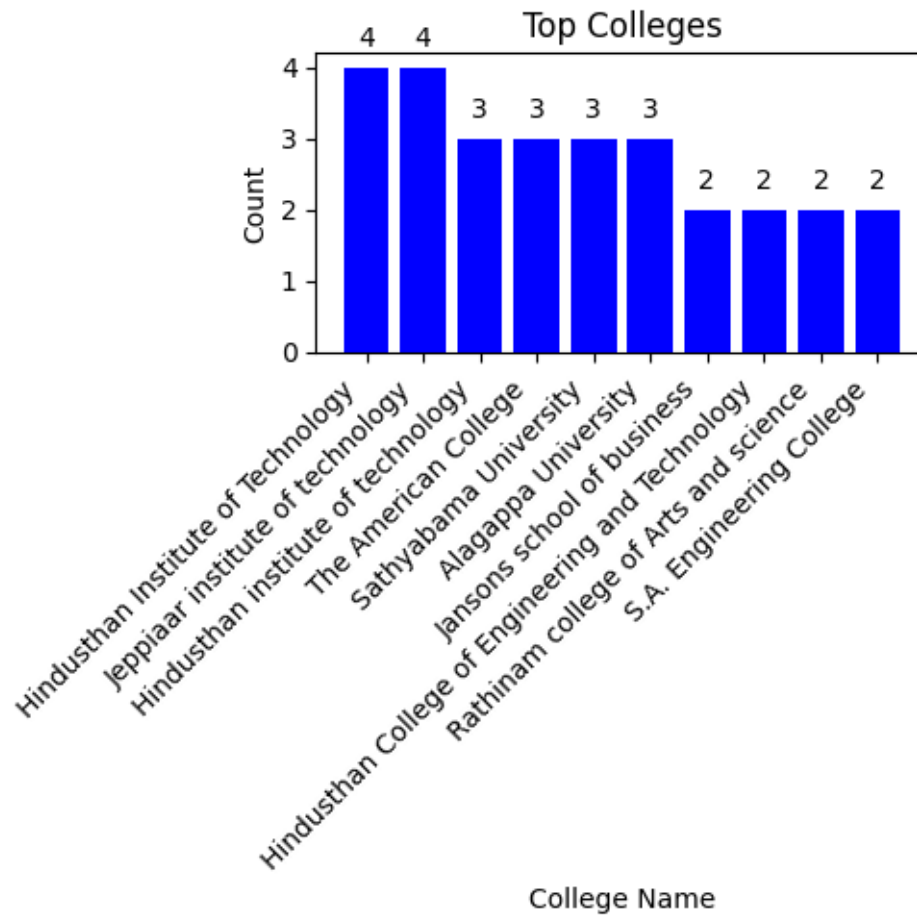
10.City Distribution

```
[66]: city_distribution = df['City'].value_counts().head(10)

      plt.figure(figsize=(10,5))
      plt.subplot(121)
      plt.bar(city_distribution.index, city_distribution.values, color='blue')
      plt.title('City Distribution')
      plt.xlabel('City')
      plt.ylabel('Count')
      plt.xticks(rotation=45, ha="right")


      for i, value in enumerate(city_distribution.values):
          plt.text(i, value + 0.2, str(value), ha='center', va='bottom')
```

City Distribution
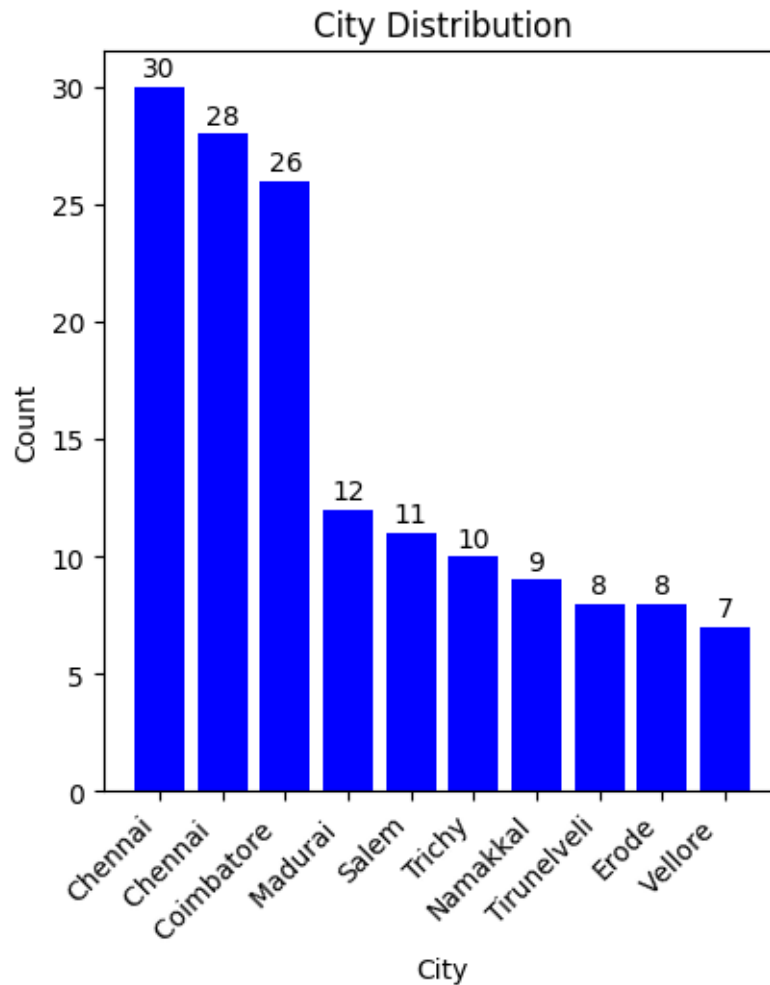
11. Areas of Interest

```
[82]: areas_of_interest = df['What you wan to learn ?'].value_counts().
      ↪sort_values(ascending=False).head(10)

      plt.figure(figsize=(8,8))
      plt.plot(areas_of_interest.index, areas_of_interest.values, marker='o',␣
       ↪color='blue', linestyle='-')
      plt.title('Top 10 Areas of Interest Distribution')
      plt.xlabel('Area of Interest')
      plt.ylabel('Count')
      plt.xticks(rotation=45)

      for x, y in zip(areas_of_interest.index, areas_of_interest.values):
          plt.text(x, y, str(y), ha='center', va='bottom', fontsize=10, color='black')
```

```
plt.tight_layout()
plt.show()
```

Top 10 Areas of Interest Distribution
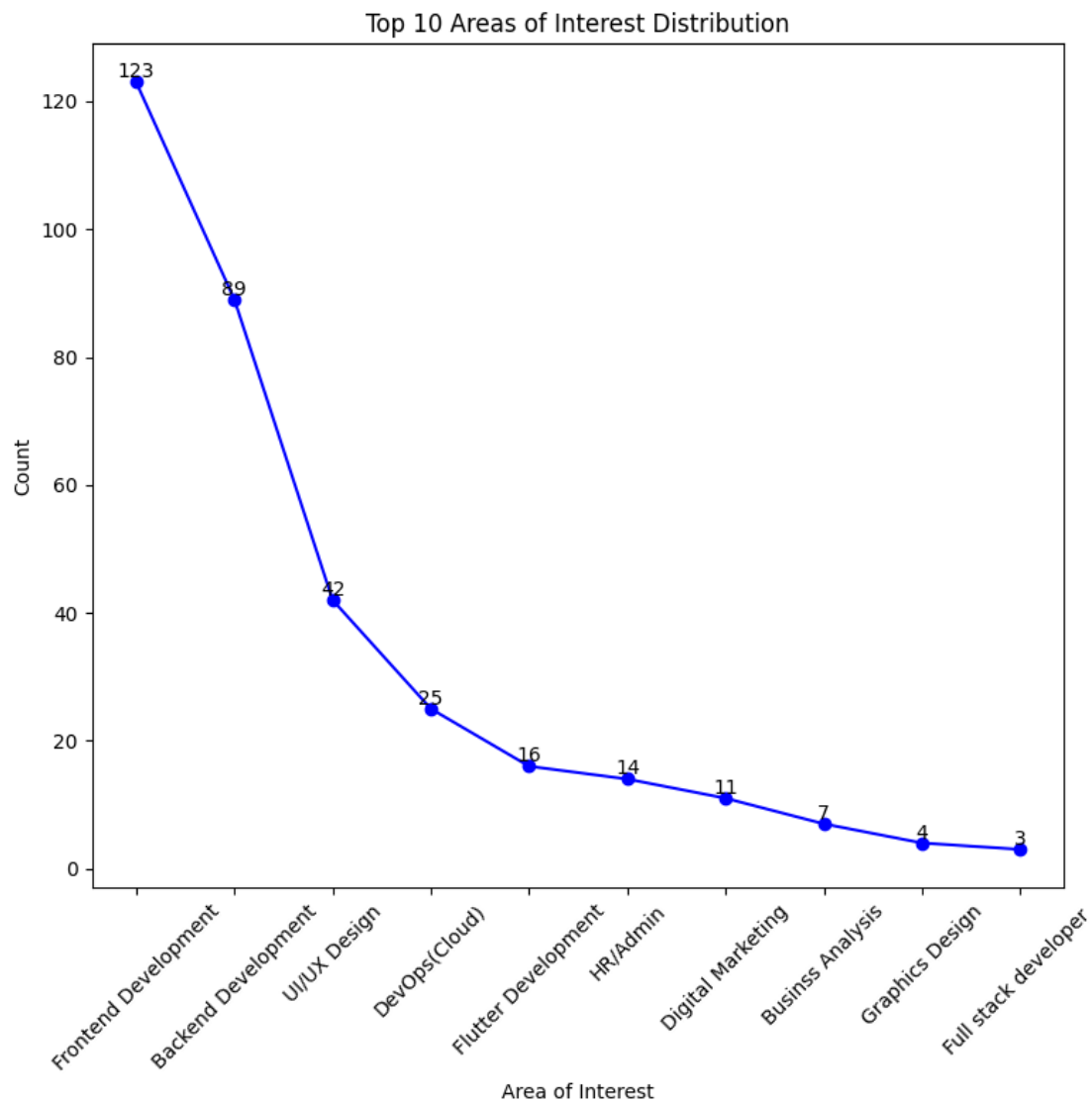
12.Tecnical Feedback by

```
[86]: feedback_counts = df['Technical Feeback by'].value_counts().head(10)

plt.figure(figsize=(5,5))
plt.bar(feedback_counts.index, feedback_counts.values, color='blue')
plt.title('Technical Feedback by')
plt.xlabel('Name')
plt.ylabel('Count')
plt.xticks(rotation=45, ha="right")
```
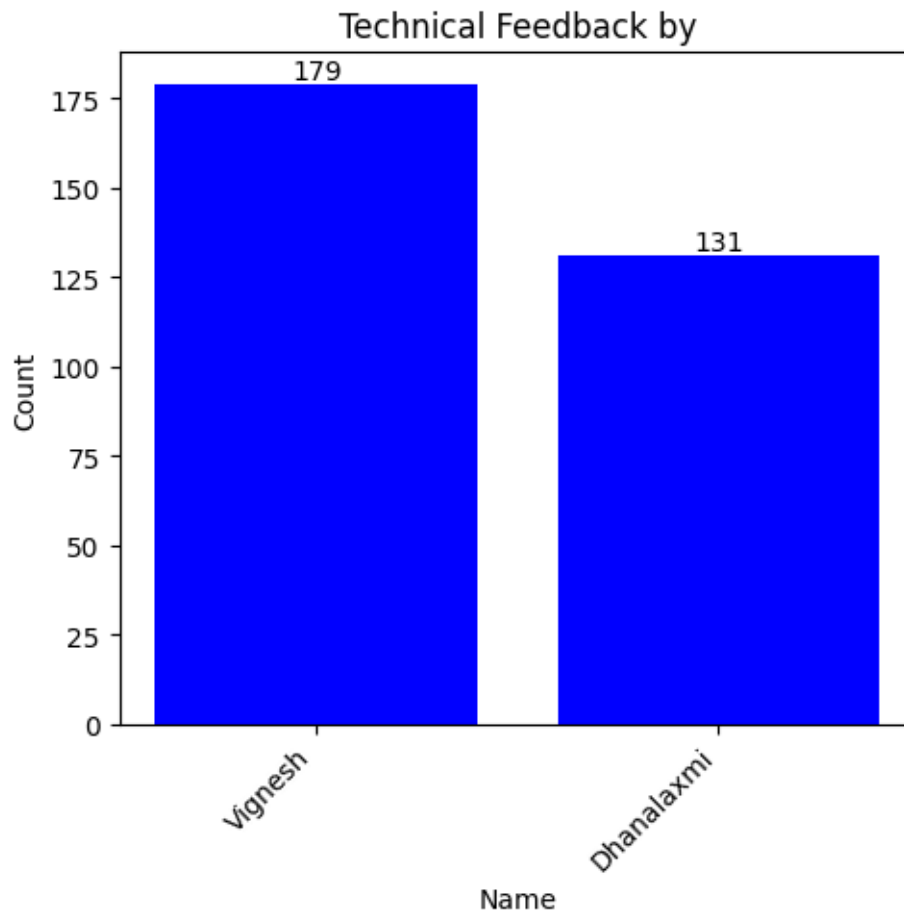
```
for i, value in enumerate(feedback_counts.values):
    plt.text(i, value + 0.2, str(value), ha='center', va='bottom')

plt.tight_layout()
plt.show()
```


13. HR Comments

```
[93]: hr_comments_count = df.groupby('HR Comments').size().
    ↪sort_values(ascending=False).head(5)

plt.figure(figsize=(10,8))
plt.bar(hr_comments_count.index, hr_comments_count.values, color='blue')
plt.title('Top 5 HR Comments')
plt.xlabel('HR Comments')
plt.ylabel('Count')
```
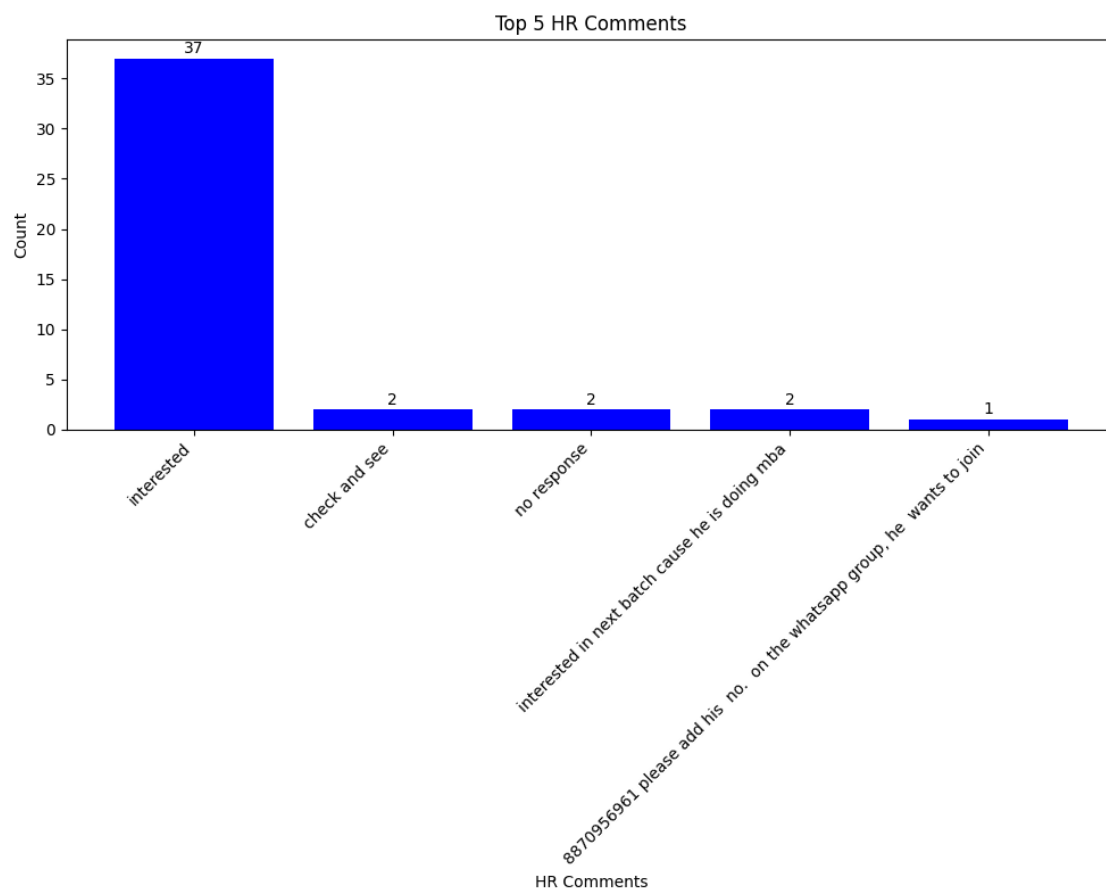
```python
plt.xticks(rotation=45, ha="right")

for i, value in enumerate(hr_comments_count.values):
    plt.text(i, value + 0.2, str(value), ha='center', va='bottom')

plt.tight_layout()
plt.show()
```


Top 5 HR Comments

Thank You!!