senchola-task-4-batch-1-details

November 6, 2023

0.1 Sechola Batch 1 Applicant Details

Importing dataset from excel

[54]:
```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from wordcloud import WordCloud
```

[55]:
```
df_1=pd.read_excel('/content/batch 1 application.xlsx', sheet_name='Form␣
↪Responses 1')
df_2= pd.read_excel('/content/batch 1 application.xlsx', sheet_name='Selected')
```

[56]:
```
df_1.head()
```

[56]:

	Date	Time	Are you open to learn ?	\
0	2023-04-04	14:41:10	Yes	
1	2023-04-04	18:32:40	Yes	
2	2023-04-04	18:02:51	Yes	
3	2023-04-04	17:03:17	Yes	
4	2023-05-09	08:22:07	Yes	

	Why you want to join this program ?	Do you have laptop	\
0	As a non IT graduate,I want to start my career…	No	
1	Learn new technologies to upgrade my skills an…	Yes	
2	To enhance my knowledge	Yes	
3	To learn the technology with team	Yes	
4	Dream and Interesting	Yes	

	Name	Gender	Address	\
0	HARSHANI BALU	Female	No.9 collectrate Kanchipuram	
1	SUBHASHINI .S	Female	2/54 ,South Street , Sirangudi South	
2	JAYAPRAKASH V	Male	7/138-1 Second floor,Old arokya hospital,sanka…	
3	UGENDHAR .U	Male	Pudusampalli near vazga valamudan mettur dam s…	
4	KARTHIK .S	Male	NaN	

	Qualification	Degree	Branch	\

```
0     Engineering    B.E                          Civil Engineering
1     Engineering    B.E                Computer Science Engineering
2     Engineering    B.E  Electronics and Communication Engineering
3     Engineering    B.E                       Infromation Technology
4  Arts & Science    BCA            Bachelors of Computer Application

   Pass-out Year                                    College Name  \
0           2021                      A.V.C COLLEGE OF ENGINEERING
1           2022  AALIM MUHAMMED SALEGH COLLEGE OF ENGINEERING
2           2023     ADHI COLLEGE OF ENGINEERING AND TECHNOLOGY
3           2023               ADHIYAMAAN COLLEGE OF ENGINEERING
4           2023                             ALAGAPPA UNIVERSITY

             City      State What you wan to learn ?          Comments  \
0  MAYILADUTHURAI  Tamilnadu            Data Analyst          Not pick
1         CHENNAI  Tamilnadu            UI/UX Design        Not Pickup
2  TIRUVANNAMALAI  Tamilnadu  Full Stack Development  Data Not Entered
3          HOSUR  Tamilnadu            UI/UX Design  Data Not Entered
4     TIRUNELVELI  Tamilnadu  Full Stack Development  Data Not Entered

  Confidence Rating            Status Status checked
0  Data Not Entered  Data Not Entered   Not Selected
1  Data Not Entered  Data Not Entered   Not Selected
2  Data Not Entered  Data Not Entered   Not Selected
3  Data Not Entered  Data Not Entered   Not Selected
4  Data Not Entered  Data Not Entered   Not Selected
```

[57]:
```
df_1.info()
df_2.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 97 entries, 0 to 96
Data columns (total 20 columns):
 #   Column                                Non-Null Count  Dtype
---  ------                                --------------  -----
 0   Date                                  97 non-null     datetime64[ns]
 1   Time                                  97 non-null     object
 2   Are you open to learn ?               97 non-null     object
 3   Why you want to join this program ?   97 non-null     object
 4   Do you have laptop                    97 non-null     object
 5   Name                                  97 non-null     object
 6   Gender                                97 non-null     object
 7   Address                               94 non-null     object
 8   Qualification                         97 non-null     object
 9   Degree                                97 non-null     object
 10  Branch                                97 non-null     object
 11  Pass-out Year                         97 non-null     int64
```

```
 12  College Name                          97 non-null     object
 13  City                                  97 non-null     object
 14  State                                 97 non-null     object
 15  What you wan to learn ?               97 non-null     object
 16  Comments                              97 non-null     object
 17  Confidence Rating                     97 non-null     object
 18  Status                                97 non-null     object
 19  Status checked                        97 non-null     object
dtypes: datetime64[ns](1), int64(1), object(18)
memory usage: 15.3+ KB
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 17 entries, 0 to 16
Data columns (total 20 columns):
 #   Column                                Non-Null Count  Dtype
---  ------                                --------------  -----
 0   Date                                  17 non-null     datetime64[ns]
 1   Time                                  17 non-null     object
 2   Are you open to learn ?               17 non-null     object
 3   Why you want to join this program ?   17 non-null     object
 4   Why you want to join this program ?.1 17 non-null     object
 5   Name                                  17 non-null     object
 6   Gender                                17 non-null     object
 7   Address                               17 non-null     object
 8   Qualification                         17 non-null     object
 9   Degree                                17 non-null     object
 10  Branch                                17 non-null     object
 11  Pass-out Year                         17 non-null     int64
 12  College Name                          17 non-null     object
 13  City                                  17 non-null     object
 14  State                                 17 non-null     object
 15  What you wan to learn ?               17 non-null     object
 16  Comments                              17 non-null     object
 17  Confidence Rating                     17 non-null     int64
 18  Confidence Rating out of              17 non-null     int64
 19  Status                                17 non-null     object
dtypes: datetime64[ns](1), int64(3), object(16)
memory usage: 2.8+ KB
```

[58]:
```
df_1.isnull().sum()
df_2.isnull().sum()
```

[58]:
```
Date                                  0
Time                                  0
Are you open to learn ?               0
Why you want to join this program ?   0
Why you want to join this program ?.1 0
Name                                  0
```

```
Gender                                0
Address                               0
Qualification                         0
Degree                                0
Branch                                0
Pass-out Year                         0
College Name                          0
City                                  0
State                                 0
What you wan to learn ?               0
Comments                              0
Confidence Rating                     0
Confidence Rating out of              0
Status                                0
dtype: int64
```

[59]:
```
df_1.columns
```

[59]:
```
Index(['Date', 'Time', 'Are you open to learn ?',
       'Why you want to join this program ?', 'Do you have laptop ', 'Name ',
       'Gender', 'Address', 'Qualification', 'Degree', 'Branch',
       'Pass-out Year', 'College Name', 'City', 'State',
       'What you wan to learn ?', 'Comments', 'Confidence Rating', 'Status',
       'Status checked'],
      dtype='object')
```

0.2 Insights

1)Total Applicants & Shortlisted

[60]:
```
Total_Applicants=df_1['Name '].value_counts().sum()
print('Total Applicants in Sechola Batch 1 Internship is', Total_Applicants)

Total_Shortlisted=df_2['Name '].value_counts().sum()
print('Total Shortlisted in Sechola Batch 1 Internship is', Total_Shortlisted)
```

```
Total Applicants in Sechola Batch 1 Internship is 97
Total Shortlisted in Sechola Batch 1 Internship is 17
```

2)Total Colleges for applicants & shortlisted

[61]:
```
Total_Colleges_1=df_1['College Name'].nunique()
print('Total Colleges for applicants in Sechola Batch 1 Internship is',␣
 ↪Total_Colleges_1)

Total_Colleges_2=df_2['College Name'].nunique()
print('Total Colleges for shortlisted in Sechola Batch 1 Internship is',␣
 ↪Total_Colleges_2)
```

```
Total Colleges for applicants in Sechola Batch 1 Internship is 76
Total Colleges for shortlisted in Sechola Batch 1 Internship is 17
```

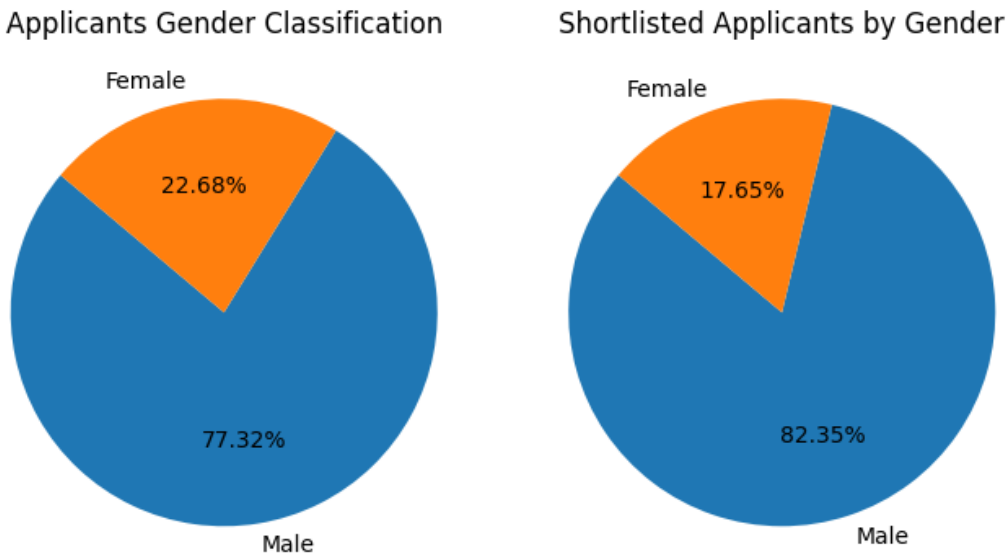
3)Total Cities for applicants & shortlisted

[62]:
```
Total_Cities_1=df_1['City'].nunique()
print('Total Cities for applicants in Sechola Batch 1 Internship is',␣
 ↪Total_Cities_1)

Total_Cities_2=df_2['City'].nunique()
print('Total Cities for shortlisted in Sechola Batch 1 Internship is',␣
 ↪Total_Cities_2)
```

```
Total Cities for applicants in Sechola Batch 1 Internship is 33
Total Cities for shortlisted in Sechola Batch 1 Internship is 8
```

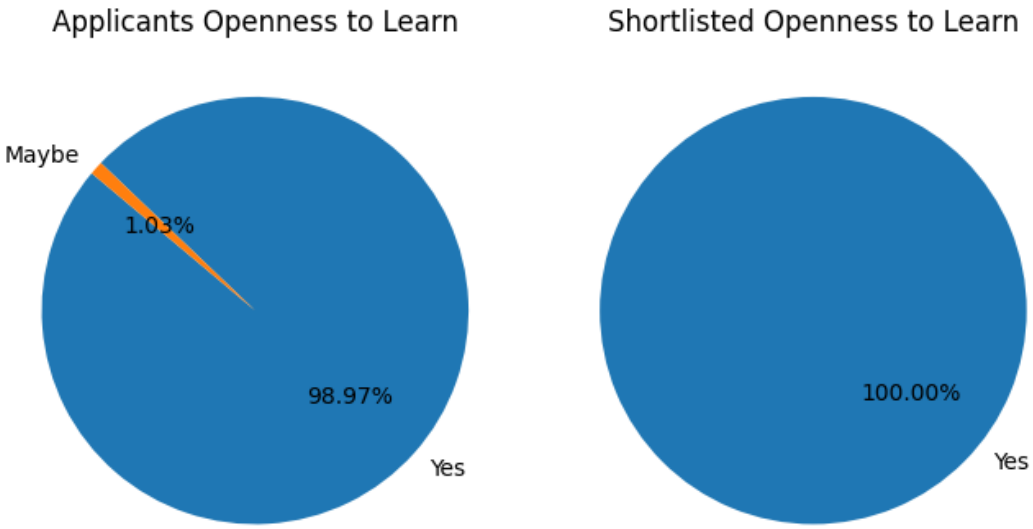
4)Applicants & Shortlisted Applicants by gender

[63]:
```
gender_counts_1 = df_1['Gender'].value_counts()
gender_counts_2 = df_2['Gender'].value_counts()

fig, (ax1, ax2) = plt.subplots(1, 2, figsize=(7, 5))
ax1.pie(gender_counts_1, labels=gender_counts_1.index, autopct='%1.2f%%',␣
 ↪startangle=140,)
ax1.set_title('Applicants Gender Classification')
ax2.pie(gender_counts_2, labels=gender_counts_2.index, autopct='%1.2f%%',␣
 ↪startangle=140, )
ax2.set_title('Shortlisted Applicants by Gender')
plt.tight_layout()
plt.show()
```



4)Applicants & Shortlisted Applicants by Openness to learn

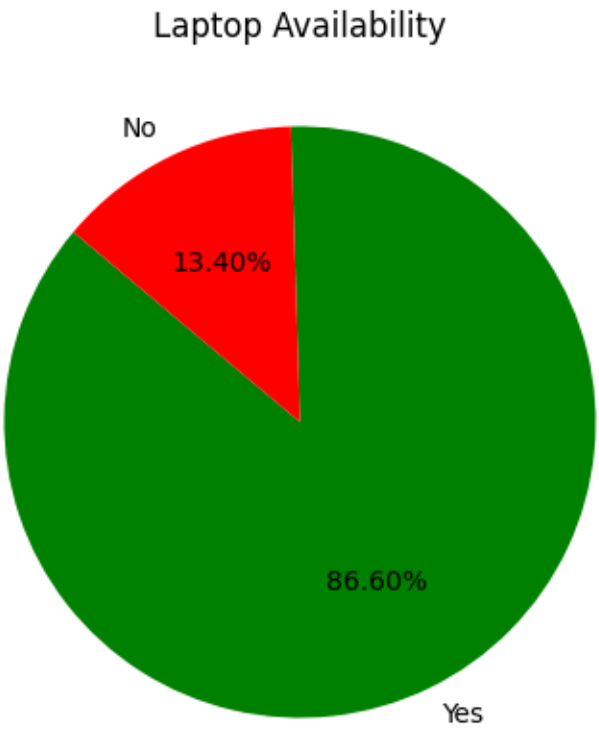
[64]:
```
openness_to_learn_1 = df_1['Are you open to learn ?'].value_counts()
openness_to_learn_2 = df_2['Are you open to learn ?'].value_counts()
colors = ['green', 'red', 'yellow']
fig, (ax1, ax2) = plt.subplots(1, 2, figsize=(7, 5))
ax1.pie(openness_to_learn_1, labels=openness_to_learn_1.index, autopct='%1.
 ↪2f%%', startangle=140,)
ax1.set_title('Applicants Openness to Learn')
ax2.pie(openness_to_learn_2, labels=openness_to_learn_2.index, autopct='%1.
 ↪2f%%', startangle=140, )
ax2.set_title('Shortlisted Openness to Learn')

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



5)Laptop Availability

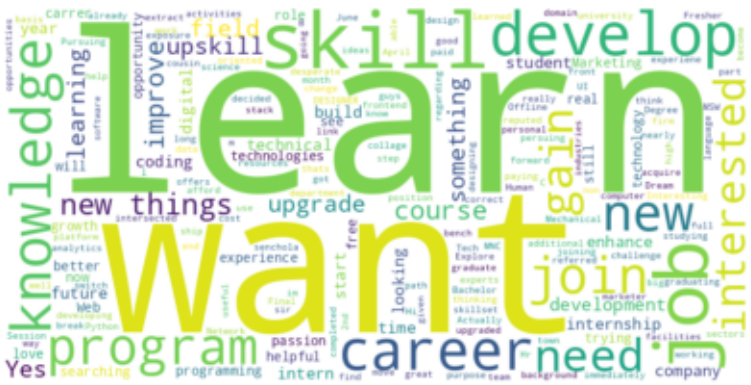
[65]:
```
laptop_availability = df_1['Do you have laptop '].value_counts()

colors = ['green', 'red']

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()
```



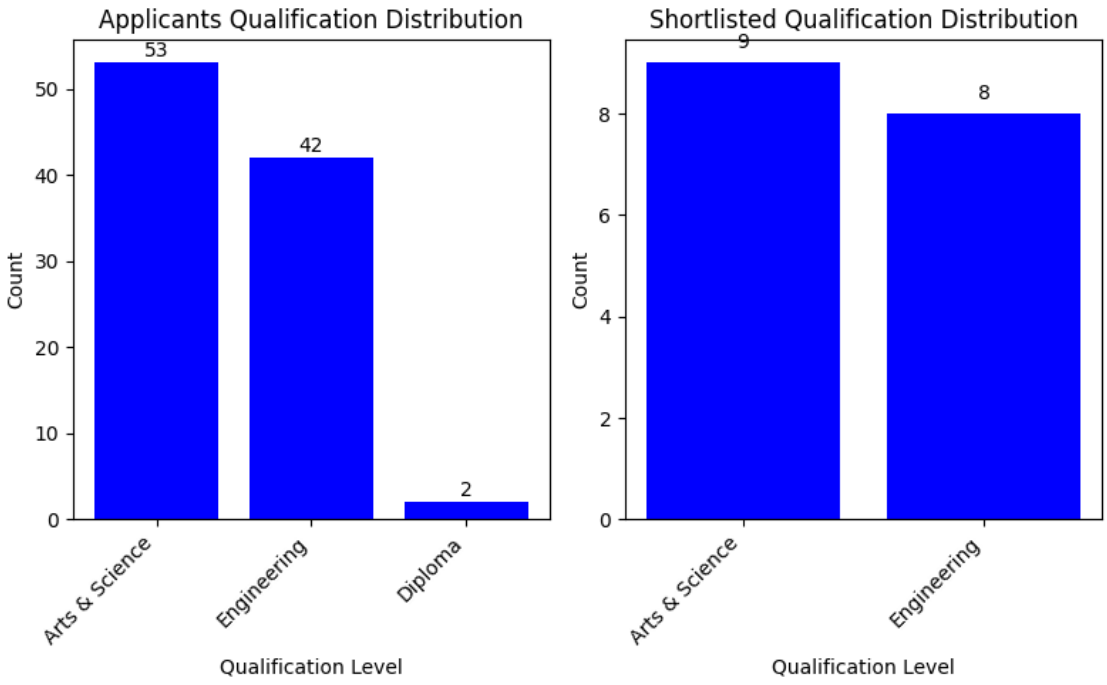
6)Motivation for joining

[66]:
```
text_data = ' '.join(df_1['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()
```



7)Applicants & Shortlisted Applicants by Qualification

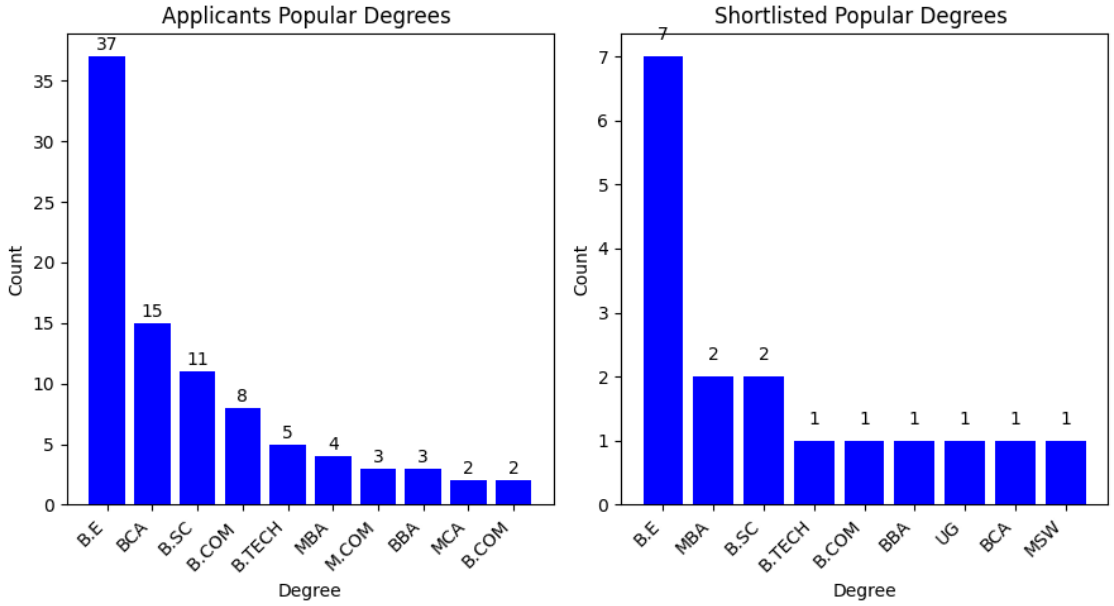
[67]:
```
qualification_distribution_1 = df_1['Qualification'].value_counts()
qualification_distribution_2 = df_2['Qualification'].value_counts()

fig, (ax1, ax2) = plt.subplots(1, 2, figsize=(8, 5))


ax1.bar(qualification_distribution_1.index, qualification_distribution_1.
 ↪values, color='blue')
ax1.set_title('Applicants Qualification Distribution')
ax1.set_xlabel('Qualification Level')
ax1.set_ylabel('Count')
ax1.set_xticks(qualification_distribution_1.index)
ax1.set_xticklabels(qualification_distribution_1.index, rotation=45, ha='right')

for i, value in enumerate(qualification_distribution_1.values):
    ax1.text(i, value + 0.2, str(value), ha='center', va='bottom')


ax2.bar(qualification_distribution_2.index, qualification_distribution_2.
 ↪values, color='blue')
ax2.set_title('Shortlisted Qualification Distribution')
ax2.set_xlabel('Qualification Level')
ax2.set_ylabel('Count')
ax2.set_xticks(qualification_distribution_2.index)
ax2.set_xticklabels(qualification_distribution_2.index, rotation=45, ha='right')

for i, value in enumerate(qualification_distribution_2.values):
    ax2.text(i, value + 0.2, str(value), ha='center', va='bottom')

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



8)Applicants & Shortlisted Applicants by Degree

[68]:
```
degree_distribution_1 = df_1['Degree'].str.upper().value_counts()
top_degrees_1 = degree_distribution_1.head(10)
fig, (ax1, ax2) = plt.subplots(1, 2, figsize=(9, 5))
ax1.bar(top_degrees_1.index, top_degrees_1.values, color='blue')
ax1.set_title('Applicants Popular Degrees ')
ax1.set_xlabel('Degree')
ax1.set_ylabel('Count')
ax1.set_xticks(top_degrees_1.index)
ax1.set_xticklabels(top_degrees_1.index, rotation=45, ha='right')

for i, value in enumerate(top_degrees_1.values):
    ax1.text(i, value + 0.2, str(value), ha='center', va='bottom')


degree_distribution_2 = df_2['Degree'].str.upper().value_counts()
top_degrees_2 = degree_distribution_2.head(10)

ax2.bar(top_degrees_2.index, top_degrees_2.values, color='blue')
ax2.set_title('Shortlisted Popular Degrees ')
ax2.set_xlabel('Degree')
ax2.set_ylabel('Count')
ax2.set_xticks(top_degrees_2.index)
ax2.set_xticklabels(top_degrees_2.index, rotation=45, ha='right')
```

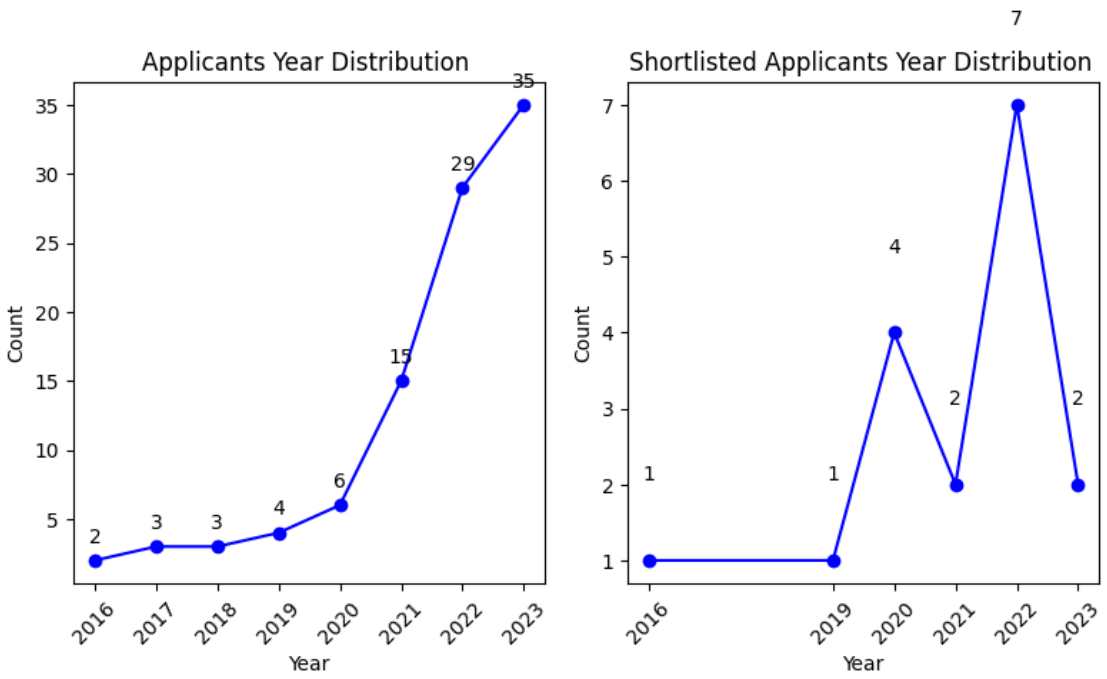
```
for i, value in enumerate(top_degrees_2.values):
    ax2.text(i, value + 0.2, str(value), ha='center', va='bottom')

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



9)Applicants & Shortlisted Applicants by Pass-Out year

[69]:
```
pass_out_years_1 = df_1['Pass-out Year'].value_counts().sort_index()
pass_out_years_2 = df_2['Pass-out Year'].value_counts().sort_index()

fig, (ax1, ax2) = plt.subplots(1, 2, figsize=(8, 5))

ax1.plot(pass_out_years_1.index, pass_out_years_1.values, marker='o',␣
 ↪color='blue', linestyle='-')
ax1.set_title('Applicants Year Distribution ')
ax1.set_xlabel('Year')
ax1.set_ylabel('Count')
ax1.set_xticks(pass_out_years_1.index)
ax1.set_xticklabels(pass_out_years_1.index, rotation=45)

for i, value in enumerate(pass_out_years_1.values):
    ax1.text(pass_out_years_1.index[i], value + 1, str(value), ha='center',␣
 ↪va='bottom', fontsize=10, color='black')
```

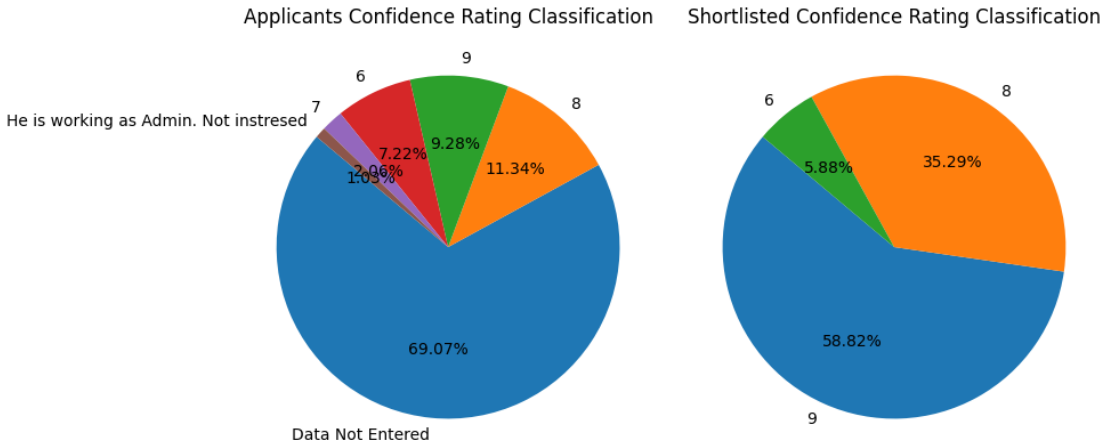
```
ax2.plot(pass_out_years_2.index, pass_out_years_2.values, marker='o',␣
 ↪color='blue', linestyle='-')
ax2.set_title('Shortlisted Applicants Year Distribution ')
ax2.set_xlabel('Year')
ax2.set_ylabel('Count')
ax2.set_xticks(pass_out_years_2.index)
ax2.set_xticklabels(pass_out_years_2.index, rotation=45)

for i, value in enumerate(pass_out_years_2.values):
    ax2.text(pass_out_years_2.index[i], value + 1, str(value), ha='center',␣
 ↪va='bottom', fontsize=10, color='black')

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



10) Confidence Rating

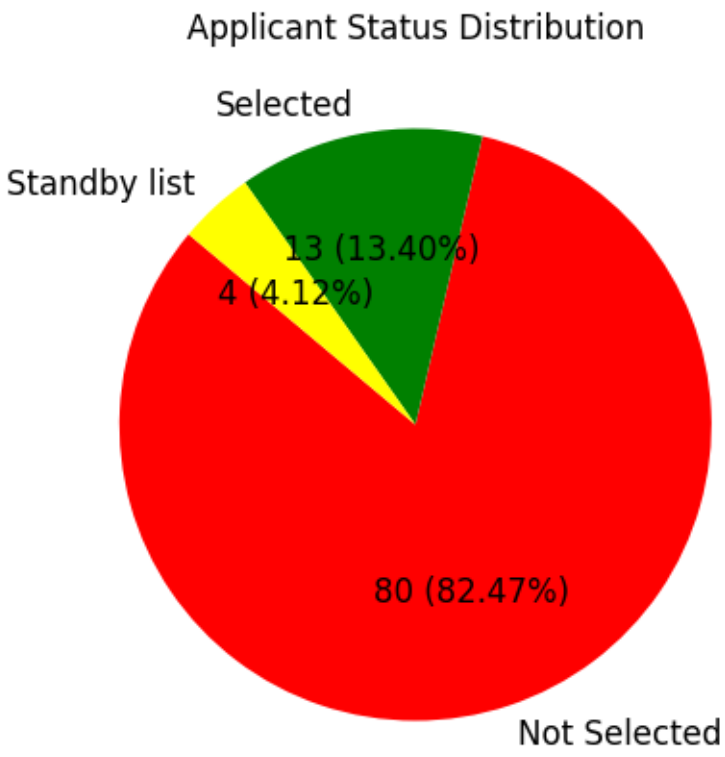
[73]:
```
confidence_rating_1 = df_1['Confidence Rating'].value_counts()
confidence_rating_2 = df_2['Confidence Rating'].value_counts()

fig, (ax1, ax2) = plt.subplots(1, 2, figsize=(10, 5))
ax1.pie(confidence_rating_1, labels=confidence_rating_1.index, autopct='%1.
 ↪2f%%', startangle=140,)
ax1.set_title('Applicants Confidence Rating Classification')
```

```
ax2.pie(confidence_rating_2, labels=confidence_rating_2.index, autopct='%1.
 ↪2f%%', startangle=140, )
ax2.set_title('Shortlisted Confidence Rating Classification')
plt.tight_layout()
plt.show()
```



11) Applicant status

[71]:
```
Applicant_status = df_1['Status checked'].value_counts()

colors = ['red', 'green', 'yellow']

plt.figure(figsize=(5, 5))
wedges, texts, autotexts = plt.pie(Applicant_status, labels=Applicant_status.
 ↪index, autopct='%1.1f%%', startangle=140, colors=colors,␣
 ↪textprops={'fontsize': 12})


total_count = Applicant_status.sum()
for i, count in enumerate(Applicant_status):
    percentage = count / total_count * 100
    autotexts[i].set_text(f'{count} ({percentage:.2f}%)')


plt.title('Applicant Status Distribution')
plt.show()
```



12)Shortlisted Applicants with comments

[72]:
```
selected_applicants = df_1[df_1['Status checked'] == 'Selected'][['Name ',␣
↪'Comments']]
standby_applicants = df_1[df_1['Status checked'] == 'Standby list'][['Name ',␣
↪'Comments']]

print('The selected applicants are:\n', selected_applicants)
print('\nThe Standby applicants are:\n', standby_applicants)
```

```
The selected applicants are:
               Name                                              Comments
6        VIDHIYA .A  Intrested in HR. One week itself she will arra…
17     KARTHIKRAJA   He is in chennai now. searching for job. He wi…
30     HARIHARAN .R          Did Course online. Prepare for gov exam.s
33  SARAVANAKUMAR.S                        Node JS, ecommerse website,
54    VIMAL SARATHY                         Worked in sales, we may try
58        SANJAI .B            he is intrested in full stack, he sill
63       VIGNESH .D  Reat, Boarstrap,worked in non IT field, have p…
69         SARAN .M  He worked in medical field. He is intrested to…
70     DHANALAKSHMI  Front end HTML, Intership in Payroll. Having p…
72        BHARATHAN                 Did frondend course, react js, java
```

76	SELVA KUMAR .V	Intrested,
78	VIGNESH	He is working in data entry. He is ready to le…
81	PANDIYARAJ	Sales Exective HDFC. Looking For Admin

The Standby applicants are:

	Name	Comments
38	GUNASEELAN .S	MEAN Stack, 1 year working in trainee, Looking…
39	BHUVANESHWARI .M	From Gov Eng college, did java and php course,…
47	MUGEETH	He is working in CTS. Looking to upgrade skill…
53	ANTONY .M	did english litrature. did intership. english …

THANK YOU