EDS Assignment 5

Name :- Nandini Thorat

Div:- C ***Batch:-*** C4

Roll no:- 384

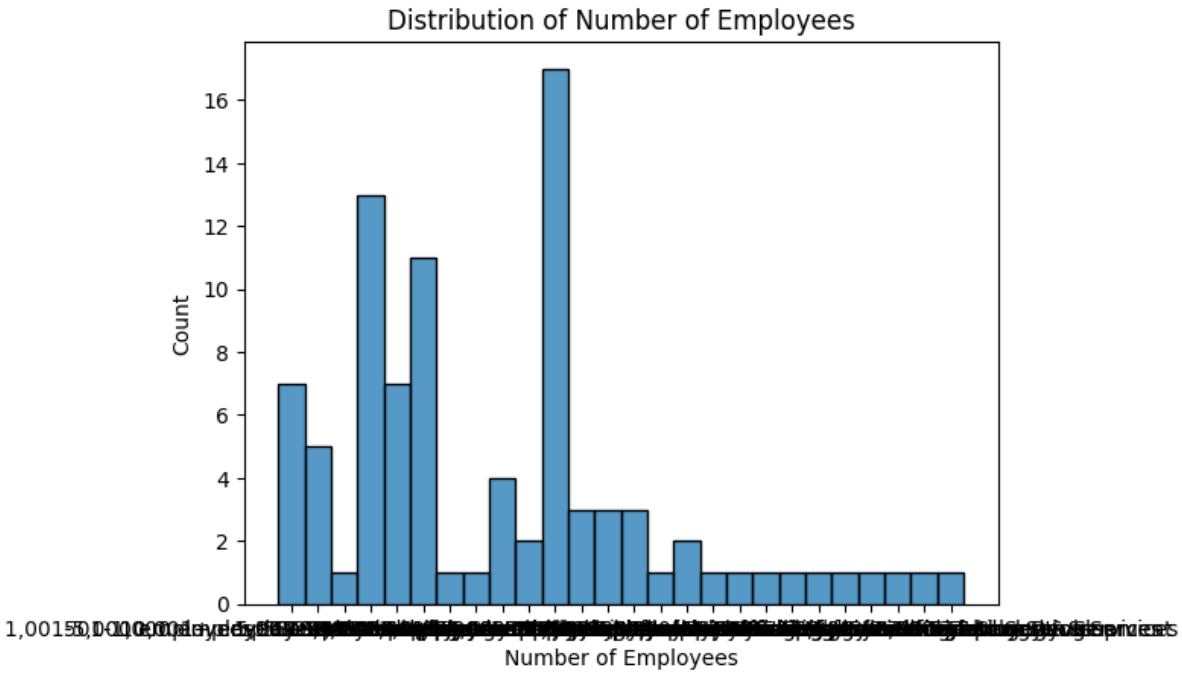
Code:-

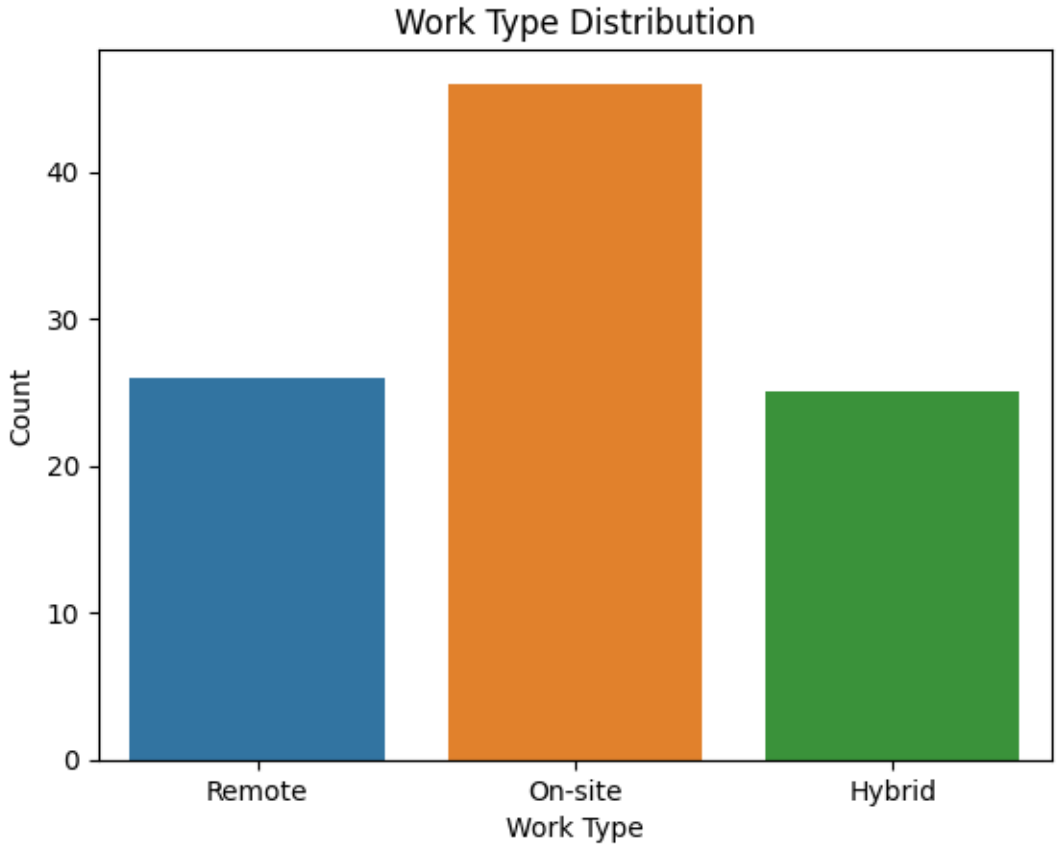
```
import seaborn as sns
import matplotlib.pyplot as plt
import pandas as pd

# Load the CSV file into a DataFrame
df = pd.read_csv("/content/CSV Dataset.csv")

# Distribution of the number of employees
sns.histplot(data=df, x='no_of_employ')
plt.xlabel('Number of Employees')
plt.ylabel('Count')
plt.title('Distribution of Number of Employees')
plt.show()

# Bar plot of work types
sns.countplot(data=df, x='work_type')
plt.xlabel('Work Type')
plt.ylabel('Count')
plt.title('Work Type Distribution')
plt.show()

# Bar plot of full-time remote jobs
sns.countplot(data=df, x='full_time_remote')
plt.xlabel('Full-Time Remote')
plt.ylabel('Count')
plt.title('Full-Time Remote Jobs Distribution')
plt.show()
```

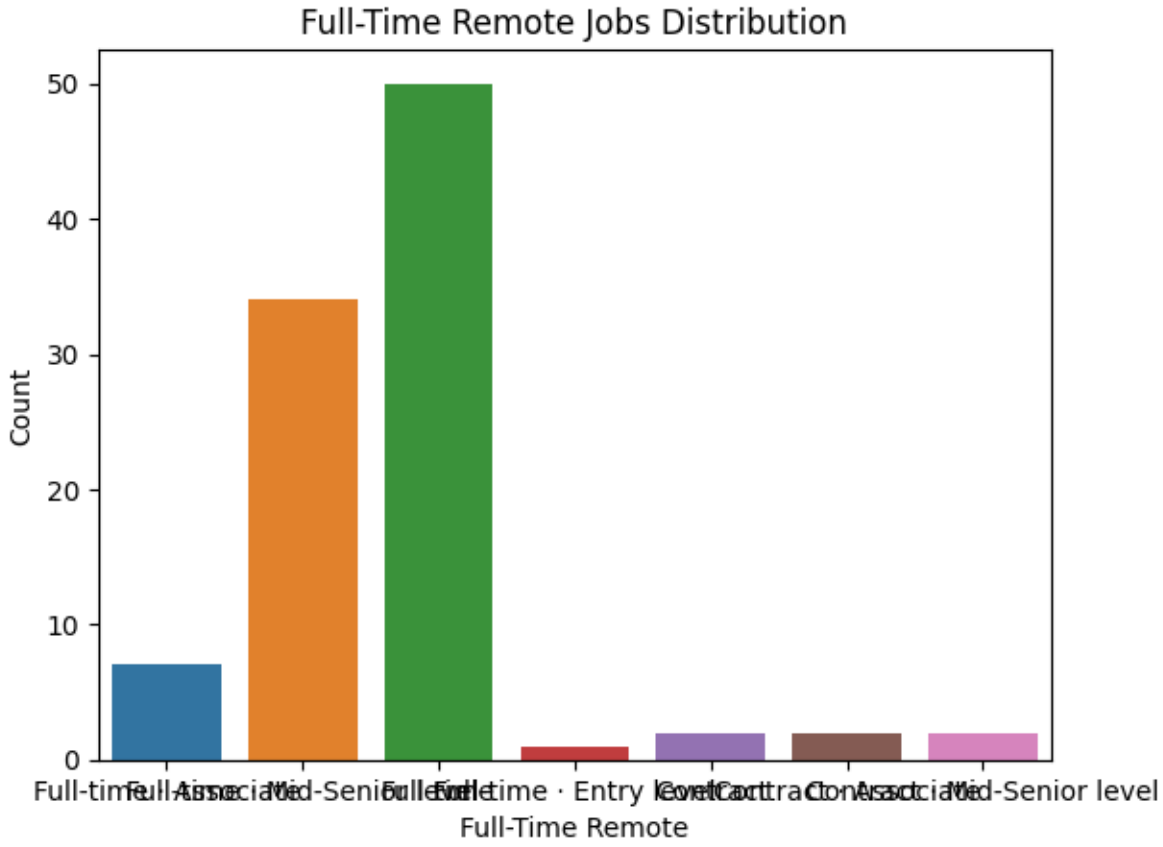
```
# Scatter plot of number of applications vs. number of employees
sns.scatterplot(data=df, x='no_of_employ', y='no_of_application')
plt.xlabel('Number of Employees')
plt.ylabel('Number of Applications')
plt.title('Number of Applications vs. Number of Employees')
plt.show()

# Bar plot of alumni
sns.countplot(data=df, x='alumni')
plt.xlabel('Alumni')
plt.ylabel('Count')
plt.title('Alumni Distribution')
plt.show()

# Bar plot of work locations
sns.countplot(data=df, x='location')
plt.xlabel('Location')
plt.ylabel('Count')
plt.title('Work Location Distribution')
plt.xticks(rotation=45)
plt.show()

# Histogram of LinkedIn followers
sns.histplot(data=df, x='linkedin_followers')
plt.xlabel('LinkedIn Followers')
plt.ylabel('Count')
plt.title('Distribution of LinkedIn Followers')
plt.show()

# Histogram of job details
sns.histplot(data=df, x='job_details')
plt.xlabel('Job Details')
plt.ylabel('Count')
plt.title('Distribution of Job Details')
plt.show()

# Bar plot of companies
sns.countplot(data=df, x='company_name')
plt.xlabel('Company Name')
plt.ylabel('Count')
plt.title('Company Distribution')
plt.xticks(rotation=45)
plt.show()

# Bar plot of job IDs
sns.countplot(data=df, x='job_ID')
plt.xlabel('Job ID')
plt.ylabel('Count')
```

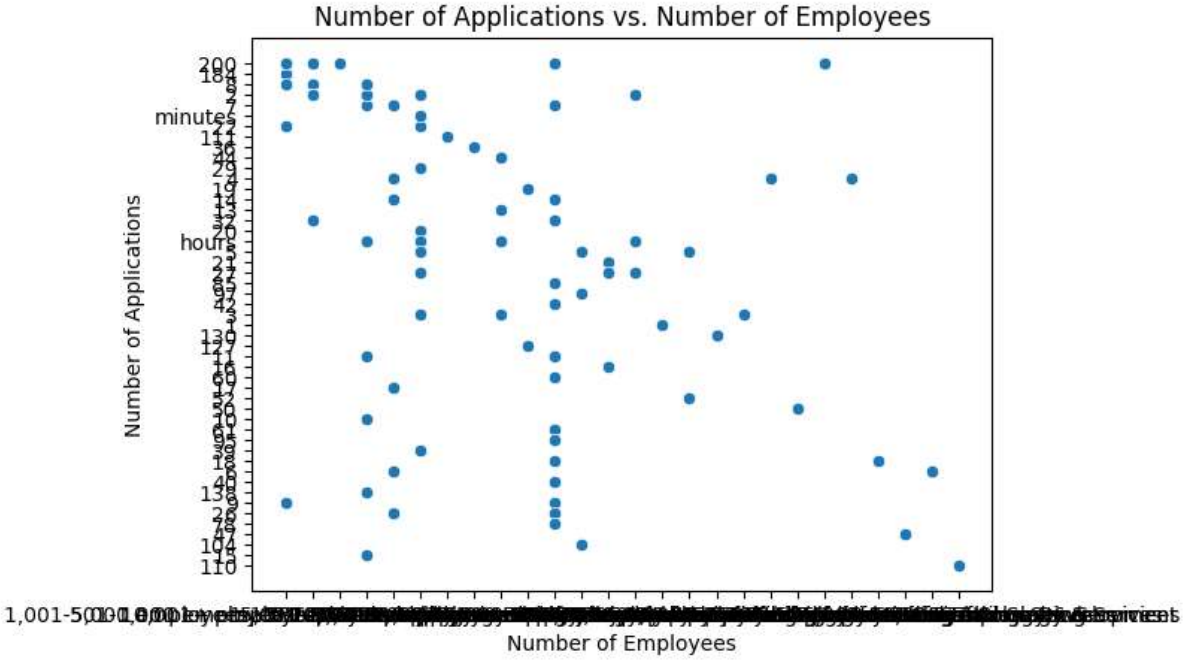
```
plt.title('Job ID Distribution')
plt.xticks(rotation=45)
plt.show()

# Bar plot of job locations
sns.countplot(data=df, x='location')
plt.xlabel('Location')
plt.ylabel('Count')
plt.title('Job Location Distribution')
plt.xticks(rotation=45)
plt.show()

# Bar plot of remote work types
sns.countplot(data=df, x='work_type')
plt.xlabel('Work Type')
plt.ylabel('Count')
plt.title('Remote Work Type Distribution')
plt.show()
```

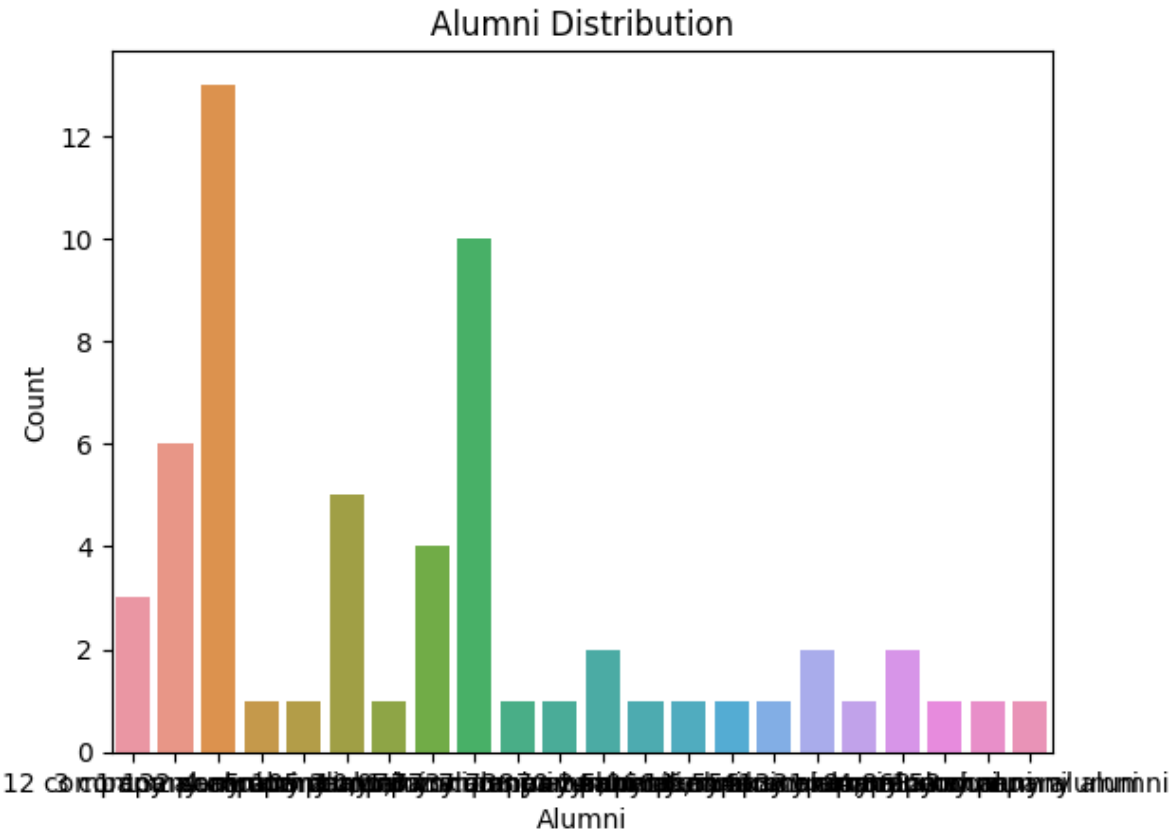
Output:-

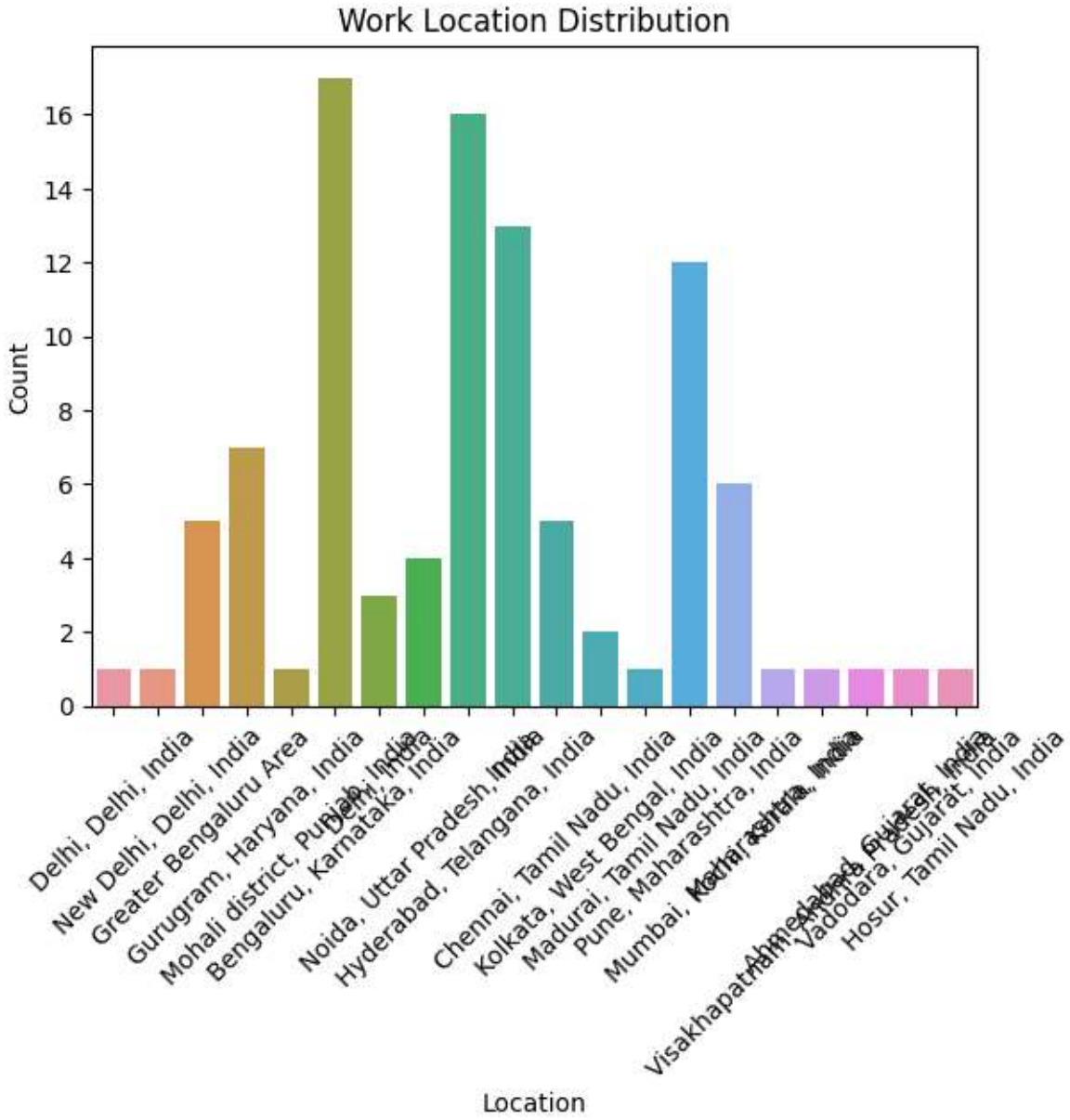

Distribution of Number of Employees
Count
0
2
4
6
8
10
12
14
16
Number of Employees

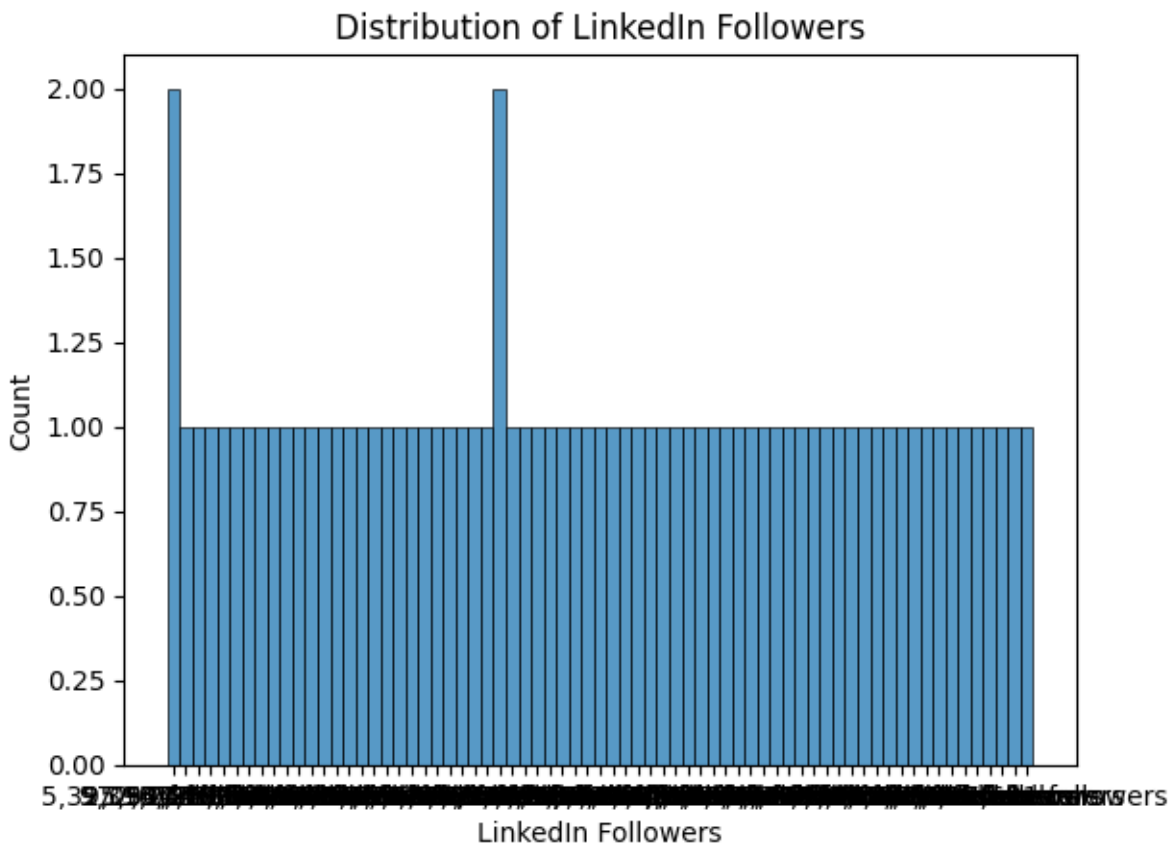

Work Type Distribution
Count
0
10
20
30
40
Remote
On-site
Hybrid
Work Type

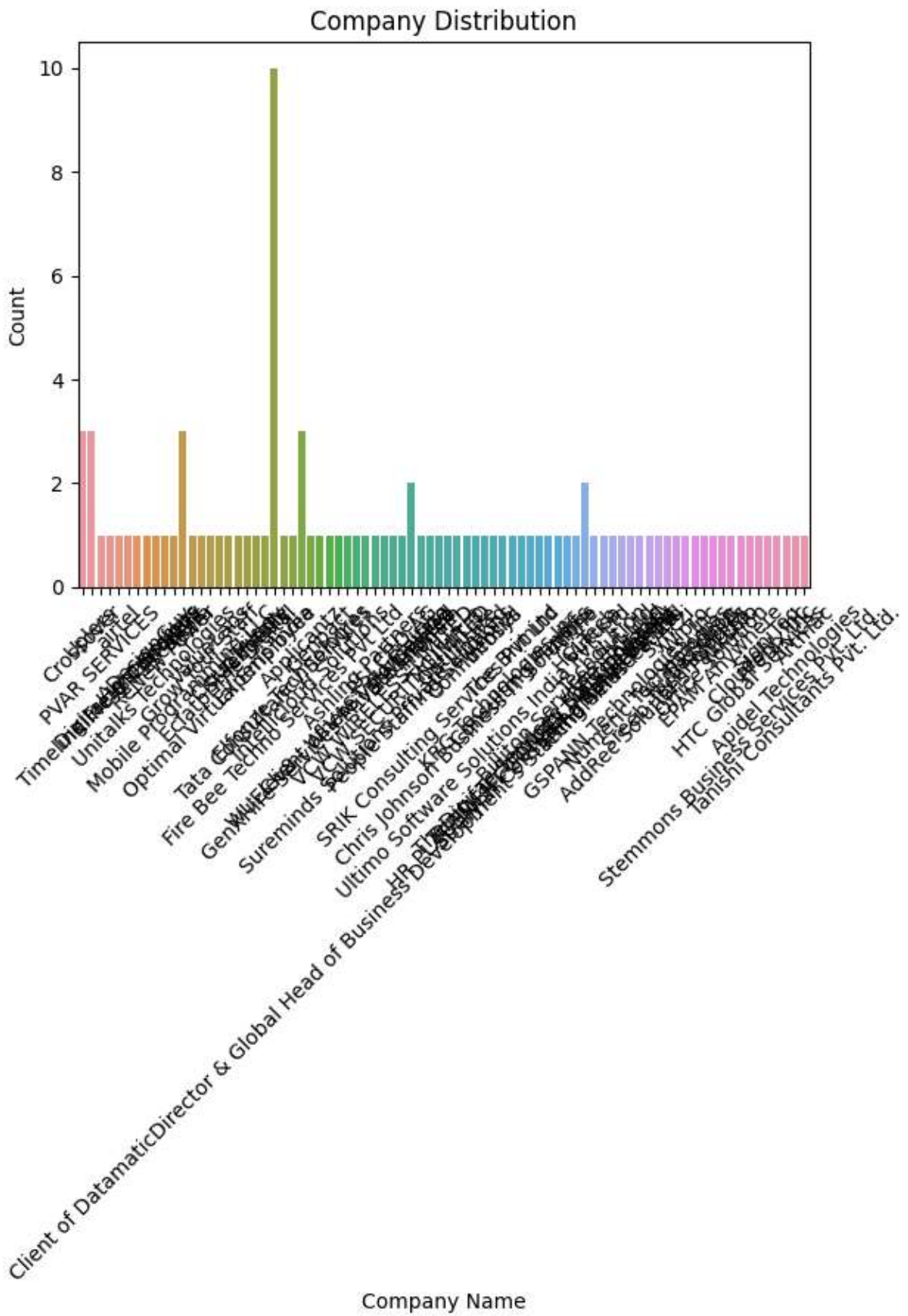

Full-Time Remote Jobs Distribution
Count
50
40
30
20
10
0
Full-time
Full-time · Entry level
Mid-Senior level
Full-Time Remote

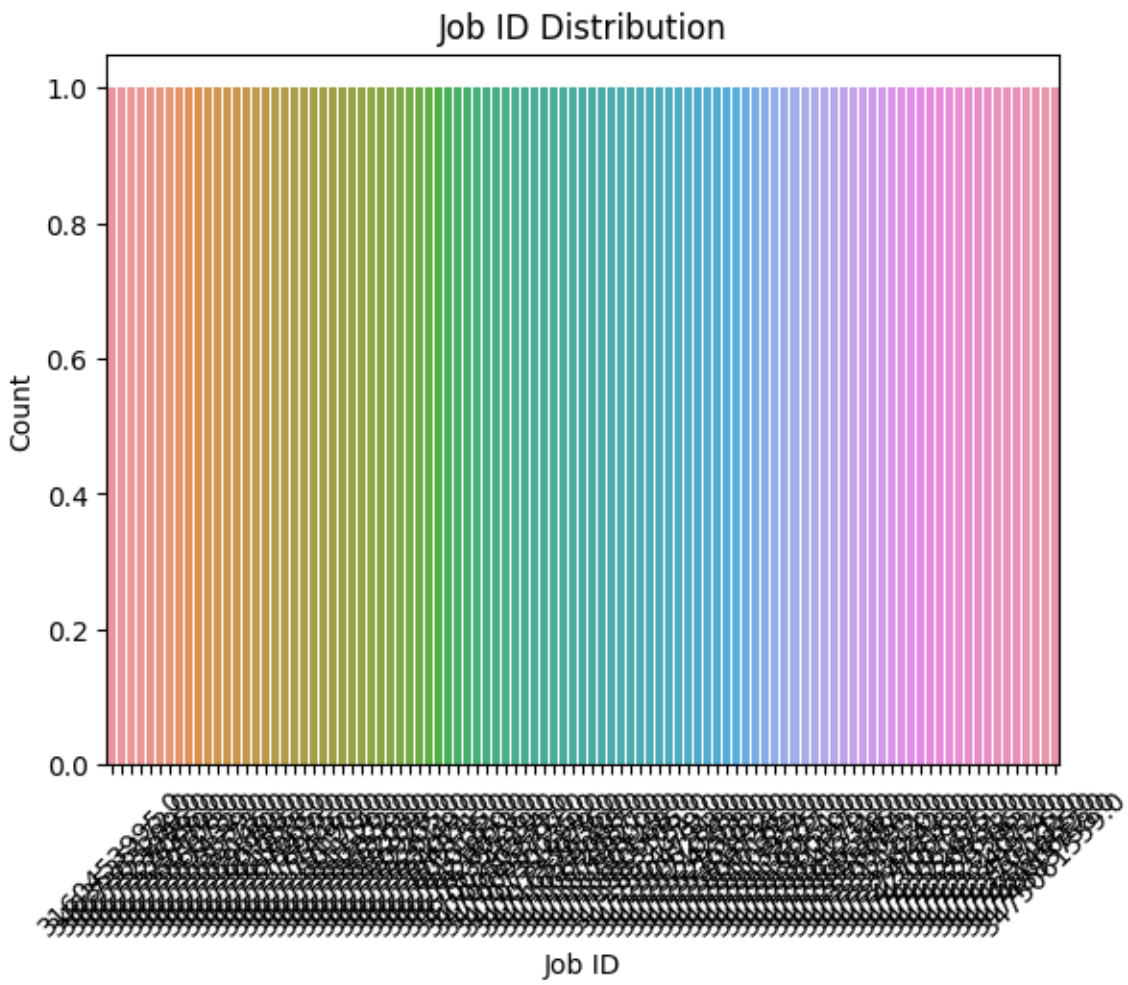

Number of Applications vs. Number of Employees
Number of Applications
minutes
hours
1,001-5,000
Number of Employees

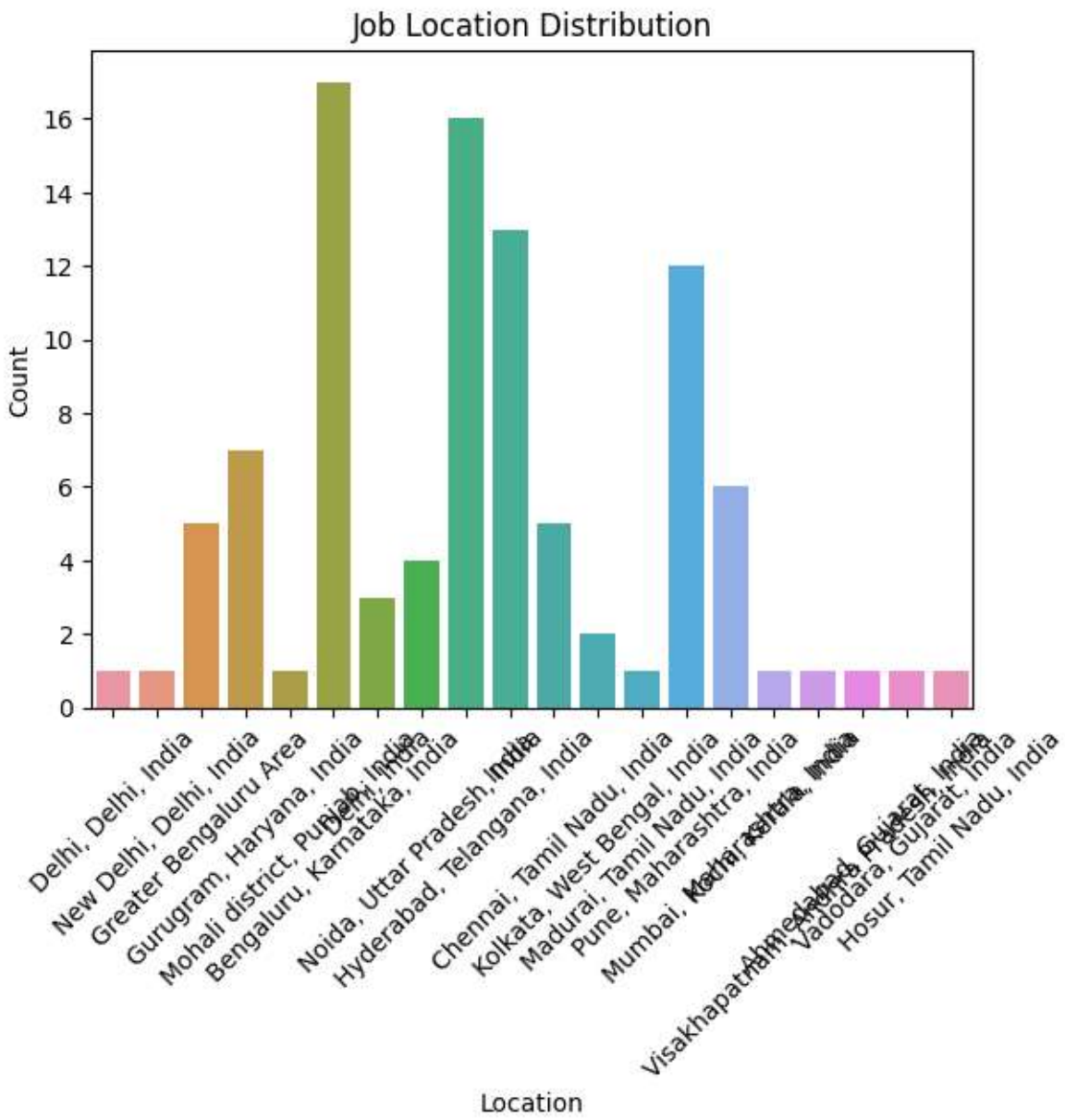

Alumni Distribution
Count
0
2
4
6
8
10
12
Alumni

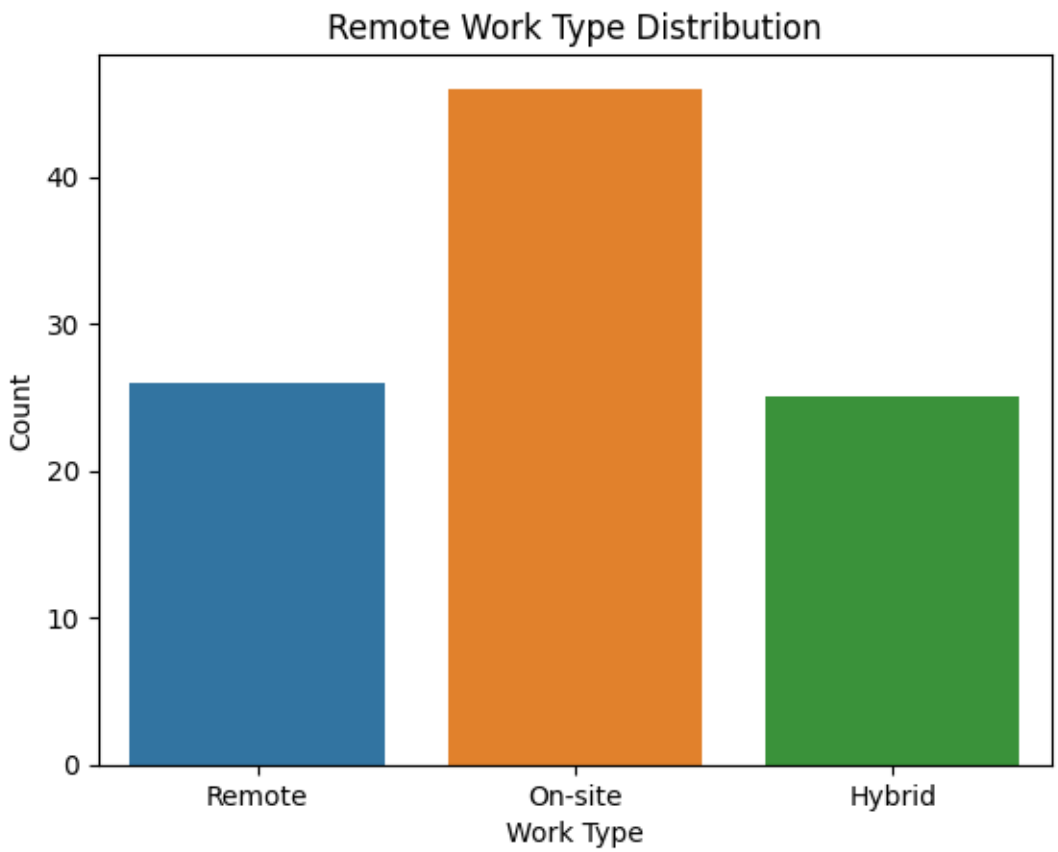

Work Location Distribution
Count
0
2
4
6
8
10
12
14
16
Delhi, Delhi, India
New Delhi, Delhi, India
Greater Bengaluru Area
Gurugram, Haryana, India
Mohali district, Punjab, India
Bengaluru, Karnataka, India
Noida, Uttar Pradesh, India
Hyderabad, Telangana, India
Chennai, Tamil Nadu, India
Kolkata, West Bengal, India
Madurai, Tamil Nadu, India
Pune, Maharashtra, India
Vadodara, Gujarat, India
Hosur, Tamil Nadu, India
Location


Distribution of LinkedIn Followers
Count
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
LinkedIn Followers


Company Distribution
Count
0
2
4
6
8
10
Company Name
Client of DatamaticDirector & Global Head of Business Development
Ultimo Software Solutions India Pvt Ltd
Chris Johnson Business Consulting
SRIK Consulting Services
Sureminds Solutions
Fire Bee Technologies
Tata Consultancy Services
Optimal Virtual Employee
Mobile Programming LLC
Unitalks Technologies
PVAR SERVICES
Cross Solar
Stemmons Business Services Pvt. Ltd.
Tanishi Consultants Pvt. Ltd.
HTC Global Services
Apidel Technologies
EPAM ANYCHECK LLC
GSPANN Technologies
AddRee Technologies


Job ID Distribution
1.0
0.8
0.6
0.4
0.2
0.0
Count
Job ID


Job Location Distribution
Count
0
2
4
6
8
10
12
14
16
Delhi, Delhi, India
New Delhi, Delhi, India
Greater Bengaluru Area
Gurugram, Haryana, India
Mohali district, Punjab, India
Bengaluru, Karnataka, India
Noida, Uttar Pradesh, India
Hyderabad, Telangana, India
Chennai, Tamil Nadu, India
Kolkata, West Bengal, India
Madurai, Tamil Nadu, India
Pune, Maharashtra, India
Mumbai, Maharashtra, India
Visakhapatnam, India
Vadodara, Gujarat, India
Hosur, Tamil Nadu, India
Location


Remote Work Type Distribution
Count
0
10
20
30
40
Remote
On-site
Hybrid
Work Type