

# **SHINE JOB SCRAPING & TREND ANALYSIS**

**BY NISHANT**

# AGENDA

1. IMPORTS LIBRARIES
2. JOB DETAILS SCRAPING FROM SHINE
3. DATA IMPORT & EXPORT
4. DATA CLEANING & WRANGLING
5. DATA ANALYSIS



# THE POWER OF WEB-SCRAPING

BEAUTIFUL SOUP IS VERY POWERFUL TOOL TO SCRAP DATA FROM ANY WEB SO I HAVE USED IT IN THIS PROJECT BY SCRAPING DETAILS FROM SHINE.

## SCRAPED DATA FROM SHINE:

1. TITLE
2. FIRM NAME
3. LOCATION
4. EXPERIENCE
5. REQUIREMENT
6. POSITION
7. SALARY RANGE
8. JOB TYPE
9. INDUSTRY
10. DEPARTMENT
11. RECRUITER
12. JOB TAGS
13. JOB POSTING DAY

Premium0 days ago

### Customer Support Representative

Tech Mahindra Ltd.

📍 Hyderabad 📅 1 to 4 Yrs 💰 < Rs 50,000 - 3.0 Lakh/Yr

• Regular • 30 Positions

[Apply](#) [Share](#) [Similar Jobs](#)

#### Other Details

Department	Technical Support / HelpdeskCustomer Service (Domestic)
Industry	BPO / Call Center
Recruiter Details	Tech Mahindra Ltd.
Job Tags	customer service
Job Type	Full time

#### Key Skills

reading

customer service

hindi

speaking



Great careers start here

# IMPORT LIBRARIES

```
import requests
from bs4 import BeautifulSoup
import pandas as pd
import numpy as np
from skimpy import skim
import seaborn as sns
import matplotlib.pyplot as plt
import warnings
warnings.filterwarnings("ignore")
headers = {"User-Agent": "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/91.0.4472.124 Safari/537.36"}
import re
```

A large, expressive red brushstroke that sweeps across the upper half of the image, creating a dynamic, hand-painted effect. The stroke is thick and textured, with visible bristles and varying shades of red.

# **SHINE** **WEB-SCRAPING**





Great careers start here

## SCRAP JOB POST LINKS FROM 20 PAGES

```
job_links = []
for page_num in range(1, 20+1):
    link = f"https://www.shine.com/job-search/jobs-{page_num}?job_type=2&top_companies_boost=true&sort=1"
    page = requests.get(link, headers=headers)
    soup = BeautifulSoup(page.content, 'html.parser')
    # store links in job_links list
    all_jobs = soup.find_all("meta", {"itemprop": "url"})
    for i in all_jobs:
        job_links.append(i.get("content"))
        print(i.get("content"))
    print("Page", page_num, "links scraped and stored in job_links")
```

```
https://www.shine.com/jobs/customer-support-representative/tech-mahindra-ltd/16130789
https://www.shine.com/jobs/customer-service-manager/grampro-business-services-private-limited/16130741
```

```
len(job_links)
```

```
400
```



# CREATE MULTIPLE FUNCTIONS TO SCRAPE DETAILS FROM JOB POST PAGE

```
def get_job_title(job_soup):  
    try:  
        get_job_title=job_soup.find("h1",{ "class":"font-size-24"}).text.strip()  
    except:  
        get_job_title= "-"  
  
    return get_job_title
```

```
get_job_title(job_soup)
```

```
'Customer Support Representative'
```

```
def get_job_type(job_soup):  
    try:  
        # Finding all span elements  
        type_elements = job_soup.find_all("span")  
  
        # Iterate to find 'Job Type' and get the next element's text  
        for i, element in enumerate(type_elements):  
            if "Job Type" in element.text:  
                return type_elements[i + 1].text.strip() # Return the following element's text  
  
        # If 'Job Type' is not found  
        return "-"  
  
    except (AttributeError, IndexError):  
        return "-"
```

```
get_job_type(job_soup)
```

```
'Full time'
```



# SCRAPE DETAILS FROM JOB POST & STORE IN DICTIONARY

```
shine={"title":[],"firm":[],"location":[],"exp":[],"requirement":[],"position":[],"salary_range":[],"post_day":[],"job_type":[],
      "job_industries":[],"job_department":[],"job_recruiter":[],"job_tags":[]}

for i in job_links:
    product_page = requests.get(i, headers=headers)
    p_soup=BeautifulSoup(product_page.content,"html.parser")

    shine["title"].append(get_job_title(p_soup))
    shine["firm"].append(get_job_firm(p_soup))
    shine["location"].append(get_job_location(p_soup))
    shine["exp"].append(get_job_exp(p_soup))
    shine["requirement"].append(get_job_req(p_soup))
    shine["position"].append(get_job_position(p_soup))
    shine["salary_range"].append(get_job_salary(p_soup))
    shine["job_type"].append(get_job_type(p_soup))
    shine["job_industries"].append(get_job_industry(p_soup))
    shine["job_department"].append(get_job_Department(p_soup))
    shine["job_recruiter"].append(get_job_Recruiter_Details(p_soup))
    shine["job_tags"].append(get_job_Job_Tags(p_soup))
    shine["post_day"].append(job_post_day(p_soup))

print("job Links Data Extracted")
```

job Links Data Extracted





Great careers start here

## TOTAL EXTRACTED DATA

```
print("title:", len(shine["title"]))
print("firm:", len(shine["firm"]))
print("location:", len(shine["location"]))
print("exp:", len(shine["exp"]))
print("requirement:", len(shine["requirement"]))
print("position:", len(shine["position"]))
print("post_day:", len(shine["post_day"]))
print("salary_range:", len(shine["salary_range"]))
print("job_type:", len(shine["job_type"]))
print("job_industries:", len(shine["job_industries"]))
print("job_department:", len(shine["job_department"]))
print("job_recruiter:", len(shine["job_recruiter"]))
print("job_tags:", len(shine["job_tags"]))
```

```
title: 400
firm: 400
location: 400
exp: 400
```

# MERGE JOB LINK WITH DICT. AND CONVERT IT INTO DATA FRAME

add link in dictionary

```
shine["job_links"]=job_links
```

dictionary to dataframe

```
df=pd.DataFrame.from_dict(shine)
df
```

	title	firm	location	exp	requirement	position	salary_range	post_day	job_type	job_industries	job_department	job_recruiter
0	Customer Support Representative	Tech Mahindra Ltd.	Hyderabad	1 to 4 Yrs	[reading, customer service, hindi, speaking]	Regular30 Positions	< Rs 50,000 - 3.0 Lakh/Yr	[Premium0 days ago]	Full time	BPO / Call Center	Technical Support / HelpdeskCustomer Service (...)	Mahindra
1	Customer Service Manager	GRAMPRO BUSINESS SERVICES PRIVATE LIMITED	Bilaspur+1Raipur	4 to 9 Yrs	[loan sales, team handling, branch handling]	Regular	Rs 2.5 - 4.5 Lakh/Yr	[Premium0 days ago]	Full time	MFI ( Micro Finance )	Investment Banking / M&A	GRAMPRO BUSINESS SERVICES PRIVATE LIMITED
		ADITYA BIRLA		1	[cross-sell							ADITYA BIRLA



Great careers start here

## IMPORT & EXPORT DATA

export data into csv

```
df.to_csv("shine_data",header=True,index=False)
```

import file shine\_data

```
df=pd.read_csv("shine_data")
df
```

	title	firm	location	exp	requirement	position	salary_range	post_day	job_type	job_industries	job_department	job_location
0	Customer Support Representative	Tech Mahindra Ltd.	Hyderabad	1 to 4 Yrs	['reading', 'customer service', 'hindi', 'spea...	Regular30 Positions	< Rs 50,000 - 3.0 Lakh/Yr	['Premium0 days ago']	Full time	BPO / Call Center	Technical Support / HelpdeskCustomer Service (...)	Mah...
	Customer	GRAMPRO BUSINESS		4	['loan sales', 'team							G...

A large, vibrant red brushstroke graphic that sweeps across the upper half of the slide, framing the title text. The brushstroke has a dynamic, hand-painted appearance with varying thickness and feathered edges.

# **DATA CLEANING & WRANGLING**



Great careers start here

## DATA SUMMARY

```
skim(df)
```

skimpy summary

Data Summary

Data Types

dataframe	Values
Number of rows	400
Number of columns	14

Column Type	Count
string	14

string

column_name	NA	NA %	words per row	total words
title	0	0	5.7	2261
firm	0	0	4.5	1805
location	0	0	1.9	775
exp	0	0	3.9	1554
requirement	0	0	10	4020
position	0	0	1.9	753
salary_range	0	0	5	2019
post_day	0	0	3	1200
job_type	16	4	1.9	749
job_industries	1	0.25	3.4	1349
job_department	1	0.25	3.7	1477
job_recruiter	0	0	4.5	1805
job_tags	0	0	1.7	699
job_links	0	0	1	400

End





Great careers start here

## CHECK MISSING VALUES & MISSING PERCENTAGE

```
missing_values=df.isnull().sum().reset_index().sort_values(by=0,ascending=False).rename(columns={0:"Total_Missing_Data"})  
missing_values["missing_percentage"]=missing_values["Total_Missing_Data"]/len(df)*100  
missing_values
```

	index	Total_Missing_Data	missing_percentage
8	job_type	16	4.00
9	job_industries	1	0.25
10	job_department	1	0.25
0	title	0	0.00
1	firm	0	0.00
2	location	0	0.00
3	exp	0	0.00
4	requirement	0	0.00
5	position	0	0.00

```
# finding all columns that have missing values
```

## FIND ALL COLUMNS NAMES WHICH HAVE MISSING VALUES

```
# finding all columns that have missing values
missing_data_columns=missing_values[missing_values["Total_Missing_Data"]>0].set_index("index")
missing_data_columns
```

	Total_Missing_Data	missing_percentage
index		
job_type	16	4.00
job_industries	1	0.25
job_department	1	0.25



Great careers start here

## HANDLE MISSING VALUES ALL 3 COLUMNS

```
df[df["job_industries"].isnull()]
```

	title	firm	location	exp	requirement	position	salary_range	post_day	job_type	job_industries	job_department	job_recruiter	job_tags
67	Onspot Offer    Reliance Nippon Life Insurance...	RELIANCE NIPPON LIFE INSURANCE COMPANY LIMITED	Navi Mumbai+2Thane, Mumbai City	1 to 6 Yrs	['sales', 'agency', 'insurance sales', 'relati...	Regular50 Positions	Rs 1.5 - 4.0 Lakh/Yr	['Premium1 month ago']	Full time	NaN	NaN	RELIANCE NIPPON LIFE INSURANCE COMPANY LIMITED	-

```
# check job type by manually by link of missing values row number
```

```
df["job_links"].loc[67]
```

```
'https://www.shine.com/jobs/onspot-offer-reliance-nippon-life-insurance-relationship-manager/reliance-nippon-life-insurance-company-limited/15784164'
```

```
# need to replace nan value in job_department with Sales / BD
```

```
df["job_department"].replace(np.NaN,"Sales / BD",inplace=True)
```

```
# need to replace nan value in job_industries with Insurance
```

```
df["job_industries"].replace(np.NaN,"Insurance",inplace=True)
```



# CREATE NEW COLUMNS ON THE BASIS ON THE DATA

1. POSITION\_TYPE
2. START\_SALARY
3. LAST\_SALARY
4. CLASSIFY\_EXP
5. TOTAL\_COMPANY\_BRANCH
6. 1ST\_BRANCH\_NAME
7. 2ND\_BRANCH\_NAME

```
df["classify_exp"]=['Fresher' if '0' in i else 'Experienced' for i in df['exp']]
```

```
df.head(2)
```

	post_day	job_type	job_industries	job_department	job_recruiter	job_tags	job_links	position_type	Start_Salary	Last_Salary	classify_exp
0		Full time	BPO / Call Center	Technical Support / HelpdeskCustomer Service (...)	Tech Mahindra Ltd.	customer service	https://www.shine.com/jobs/customer-support-re...	Regular	50000.0	3.0	Experienced
0		Full time	MFI ( Micro Finance )	Investment Banking / M&A	GRAMPRO BUSINESS SERVICES PRIVATE LIMITED	customer service	https://www.shine.com/jobs/customer-service-ma...	Regular	2.5	4.5	Experienced



## COLUMNS INCORRECT VALUE CORRECTION IN BELOW COLUMNS

1. JOB\_TYPE
2. POSITION
3. JOB\_TAGS
4. POST\_DAY
5. START\_SALARY
6. LAST\_SALARY

### position data correction

```
df["position"].unique()

array(['Regular30 Positions', 'Regular', 'Regular15 Positions',
      'Regular5 Positions', 'Regular10 Positions', 'Regular99 Positions',
      'Regular2 Positions', 'Internship10 Positions',
      'Regular50 Positions', 'Regular20 Positions',
      'Regular40 Positions', 'Regular3 Positions', 'Regular12 Positions',
      'Regular51 Positions', 'Regular25 Positions',
      'Regular34 Positions', 'Regular21 Positions', 'Regular4 Positions',
      'Regular90 Positions', 'Regular49 Positions',
      'Internship99 Positions', 'Regular60 Positions',
      'Contractual40 Positions', 'Regular19 Positions',
      'Regular18 Positions', 'Regular17 Positions'], dtype=object)
```

```
# drop regular from column position
df["position"] = df["position"].str.extract(r'(\d+.*\s)')

df["position"].unique()

array(['30 Positions', 'Not Available', '15 Positions', '5 Positions',
      '10 Positions', '99 Positions', '2 Positions', '50 Positions',
      '20 Positions', '40 Positions', '3 Positions', '12 Positions',
      '51 Positions', '25 Positions', '34 Positions', '21 Positions',
      '4 Positions', '90 Positions', '49 Positions', '60 Positions',
      '19 Positions', '18 Positions', '17 Positions'], dtype=object)

df["position"] = df["position"].replace(r"Positions|Not Available", "", regex=True)

# fill Nan value with Not available
df["position"].replace('', 0, inplace=True)

df["position"] = df["position"].astype("int")

df["position"].unique()

array([30, 0, 15, 5, 10, 99, 2, 50, 20, 40, 3, 12, 51, 25, 34, 21, 4,
      90, 49, 60, 19, 18, 17])
```



## CHANGE DATA TYPE

```
df["Start_Salary"]=df["Start_Salary"].astype("float")
df["Last_Salary"]=df["Last_Salary"].astype("float")
```

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 400 entries, 0 to 399
Data columns (total 16 columns):
 #   Column                Non-Null Count  Dtype  
---  -
 0   title                 400 non-null   object 
 1   firm                  400 non-null   object 
 2   location              400 non-null   object 
 3   exp                   400 non-null   object 
 4   requirement           400 non-null   object 
 5   position              400 non-null   int32   
 6   post_day              400 non-null   int64   
 7   job_type              400 non-null   object 
 8   job_industries        400 non-null   object 
 9   job_department        400 non-null   object 
10   job_recruiter         400 non-null   object 
11   job_tags              400 non-null   object 
12   job_links             400 non-null   object 
13   position_type         400 non-null   object 
14   Start_Salary          400 non-null   float64 
15   Last_Salary           400 non-null   float64 
dtypes: float64(2), int32(1), int64(1), object(12)
memory usage: 48.6+ KB
```



## DROP COLUMNS

```
df.columns
```

```
Index(['title', 'firm', 'location', 'exp', 'requirement', 'position',  
      'salary_range', 'post_day', 'job_type', 'job_industries',  
      'job_department', 'job_recruiter', 'job_tags', 'job_links',  
      'position_type', 'Start_Salary', 'Last_Salary'],  
      dtype='object')
```

```
df=df.drop(columns=("salary_range"))
```



Great careers start here

## RENAME COLUMNS NAME

```
df.columns
```

```
Index(['title', 'firm', 'location', 'exp', 'requirement', 'position',  
      'post_day', 'job_type', 'job_industries', 'job_department',  
      'job_recruiter', 'job_tags', 'job_links', 'position_type',  
      'Start_Salary', 'Last_Salary'],  
      dtype='object')
```

```
df=df.rename(columns={"requirement":"Skills","position":"Total_Vacancy","job_industries":"Industries","job_department":"Department","job_recruiter":"Recr  
                    "job_tags":"Tags","job_links":"Apply_Link","position_type":"Position_Type","total_company_branch":"No_of_branch","1st_Branch_Name":"  
                    "2nd_Branch_Name":"2nd_Branch"  
                    })
```



Great careers start here

## DROPPING DUPLICATE VALUES FROM DATASET

```
# Dropping duplicate values and keep the first values
```

```
df=df.drop_duplicates(keep="first")
```

```
df
```

116	Laboratory Technician in Sweden and Europe	Flight to Success Immigration LLP	Canada+1Sweden	27 Yrs	['laboratory safety', 'microbiology', 'hematol...]	18	2	Full time	Medical / Healthcare		Doctor	Flight to Success Immigration LLP	Avail
117	Medical Representative in Portugal no ielts	Flight To Success Immigration LLP	Canada+1Hong Kong	510 Yrs	['medical equipment', 'medical billing', 'medi...]	17	2	Full time	Medical / Healthcare	Transcription	Medical Pharmacist / Medical Repr...	Flight To Success Immigration LLP	Avail
119	Insurance Sales Manager	WORKFREAKS BUSINESS SERVICES PRIVATE LIMITED	Chennai	37 Yrs	['agency channel', 'agency sales', 'agent recr...]	10	1	Full time	Insurance		Sales / BD	WORKFREAKS BUSINESS SERVICES PRIVATE LIMITED	s

117 rows × 20 columns



Great careers start here

## IMPORT & EXPORT CLEAN DATA

### export clean data

```
df.to_csv("shine_clean_data",header=True,index=False)
```

### Import clean data

```
pd.read_csv("shine_clean_data")
```

	title	firm	location	exp	Skills	Total_Vacancy	post_day	job_type	Industries	Department	Recruiter	
0	Customer Support Representative	Tech Mahindra Ltd.	Hyderabad	1 to 4 Yrs	['reading', 'customer service', 'hindi', 'spea...	30	0	Full time	BPO / Call Center	Technical Support / HelpdeskCustomer Service (...)	Tech Mahindra Ltd.	custo sel
1	Customer Service Manager	GRAMPRO BUSINESS SERVICES PRIVATE LIMITED	Bilaspur+1Raipur	4 to 9 Yrs	['loan sales', 'team handling', 'branch handli...	0	0	Full time	MFI ( Micro Finance )	Investment Banking / M&A	GRAMPRO BUSINESS SERVICES PRIVATE LIMITED	custo sel
2	Relationship Manager	ADITYA BIRLA SUN LIFE INSURANCE COMPANY	Delhi	1 to 6 Yrs	['cross-sell', 'banca sales', 'life insurance'...	15	0	Full time	Insurance	Field Sales ExecutiveSales / BD	ADITYA BIRLA SUN LIFE INSURANCE COMPANY	Avail








A large, vibrant red brushstroke graphic that serves as a background for the text. It has a rough, hand-painted texture with visible bristles and varying shades of red, creating a dynamic and energetic feel. The stroke is centered horizontally and occupies the middle portion of the frame.

# **EXPLORATORY DATA ANALYSIS (EDA)**

# DATA EXPLORATION

Data Summary		Data Types	
dataframe	Values	Column Type	Count
Number of rows	117	string	15
Number of columns	20	int32	3
		float64	2

number										
column_name	NA	NA %	mean	sd	p0	p25	p50	p75	p100	hist
Total_Vacancy	0	0	26.89	29.74	0	5	17	50	99	
post_day	0	0	24.28	22.8	0	4	14	30	60	
Start_Salary	0	0	8550	18900	0	1.5	2	6	50000	
Last_Salary	0	0	5.748	11.2	0	2	3	4.5	90	
No_of_branch	0	0	1.504	1.164	1	1	1	1	7	

string				
column_name	NA	NA %	words per row	total words
title	0	0	3.8	442
firm	0	0	4.8	561
location	0	0	1.7	203
exp	0	0	3.9	452
Skills	0	0	9.4	1097
job_type	0	0	2	234
Industries	0	0	2.9	337
Department	0	0	3.7	430
Recruiter	0	0	4.8	561
Tags	0	0	2.1	250
Apply_Link	0	0	1	117
Position_Type	0	0	1	117
classify_exp	0	0	1	117
1st_Branch	0	0	1.3	154
2nd_Branch	88	75.21	0.36	42

# ABOUT DATA

## About shine Data

The dataset has 117 entries and 20 columns, which cover various job details such as title, company, location, experience, skills, salary, job type, and recruiter information.

[8]: df.head()

	title	firm	location	exp	Skills	Total_Vacancy	post_day	job_type	Industries	Department	Recruiter	Tags
0	Customer Support Representative	Tech Mahindra Ltd.	Hyderabad	1 to 4 Yrs	['reading', 'customer service', 'hindi', 'spea...	30	0	Full time	BPO / Call Center	Technical Support / HelpdeskCustomer Service (...)	Tech Mahindra Ltd.	customer service https://
1	Customer Service Manager	GRAMPRO BUSINESS SERVICES PRIVATE LIMITED	Bilaspur+1Raipur	4 to 9 Yrs	['loan sales', 'team handling', 'branch handli...	0	0	Full time	MFI ( Micro Finance )	Investment Banking / M&A	GRAMPRO BUSINESS SERVICES PRIVATE LIMITED	customer service https://
2	Relationship Manager	ADITYA BIRLA SUN LIFE INSURANCE COMPANY LIMITED	Delhi	1 to 6 Yrs	['cross-sell', 'banca sales', 'life insurance'...	15	0	Full time	Insurance	Field Sales ExecutiveSales / BD	ADITYA BIRLA SUN LIFE INSURANCE COMPANY LIMITED	Not Available https://
3	Relationship Manager Walking Interview at Rajk...	RBL BANK LIMITED	Rajkot	2 to 7 Yrs	['sales', 'relationship marketing', 'insurance'...	5	0	Full time	Banking	Sales / BDPre-Sales	RBL BANK LIMITED	Not Available https://
4	Hr Recruiter	Shankar Hiring For Hr Recruiter	Pondicherry	0 to 3 Yrs	['hr', 'hr administration', 'hr operations']	10	0	Full time	Recruitment Services	RecruitmentHR	Shankar Hiring For Hr Recruiter	talent acquisition, recruitment



Great careers start here

## NUMERICAL & CATEGORICAL COLUMN SELECTION

column selection : numerical and categorical

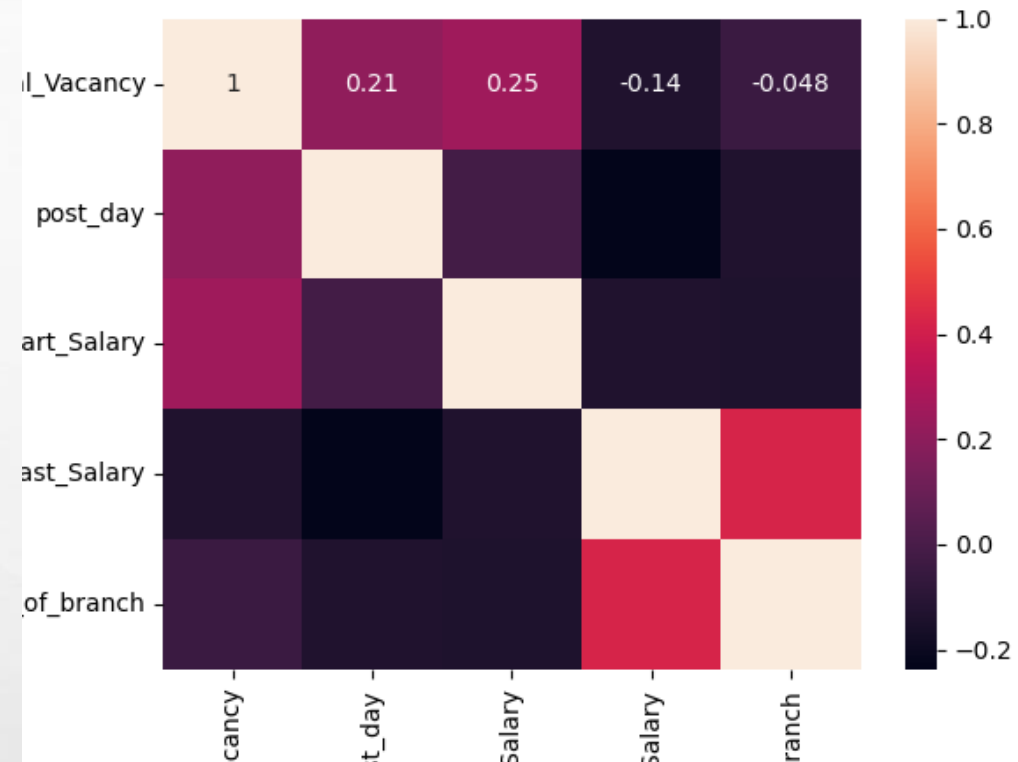
```
[12]: numerical=df.select_dtypes(include=['number']).columns.tolist()  
numerical
```

```
[12]: ['Total_Vacancy', 'post_day', 'Start_Salary', 'Last_Salary', 'No_of_branch']
```

```
[14]: categorical=df.select_dtypes(exclude=['number']).columns.tolist()  
categorical
```

```
[14]: ['title',  
      'firm',  
      'location',  
      'exp',  
      'Skills',  
      'job_type',  
      'Industries',  
      'Department',  
      'Recruiter',  
      'Tags',  
      'Apply_Link',  
      'Position_Type',  
      'classify_exp',  
      '1st_Branch',  
      '2nd_Branch']
```

## CORRELATION BETWEEN ALL NUMERICAL COLUMNS



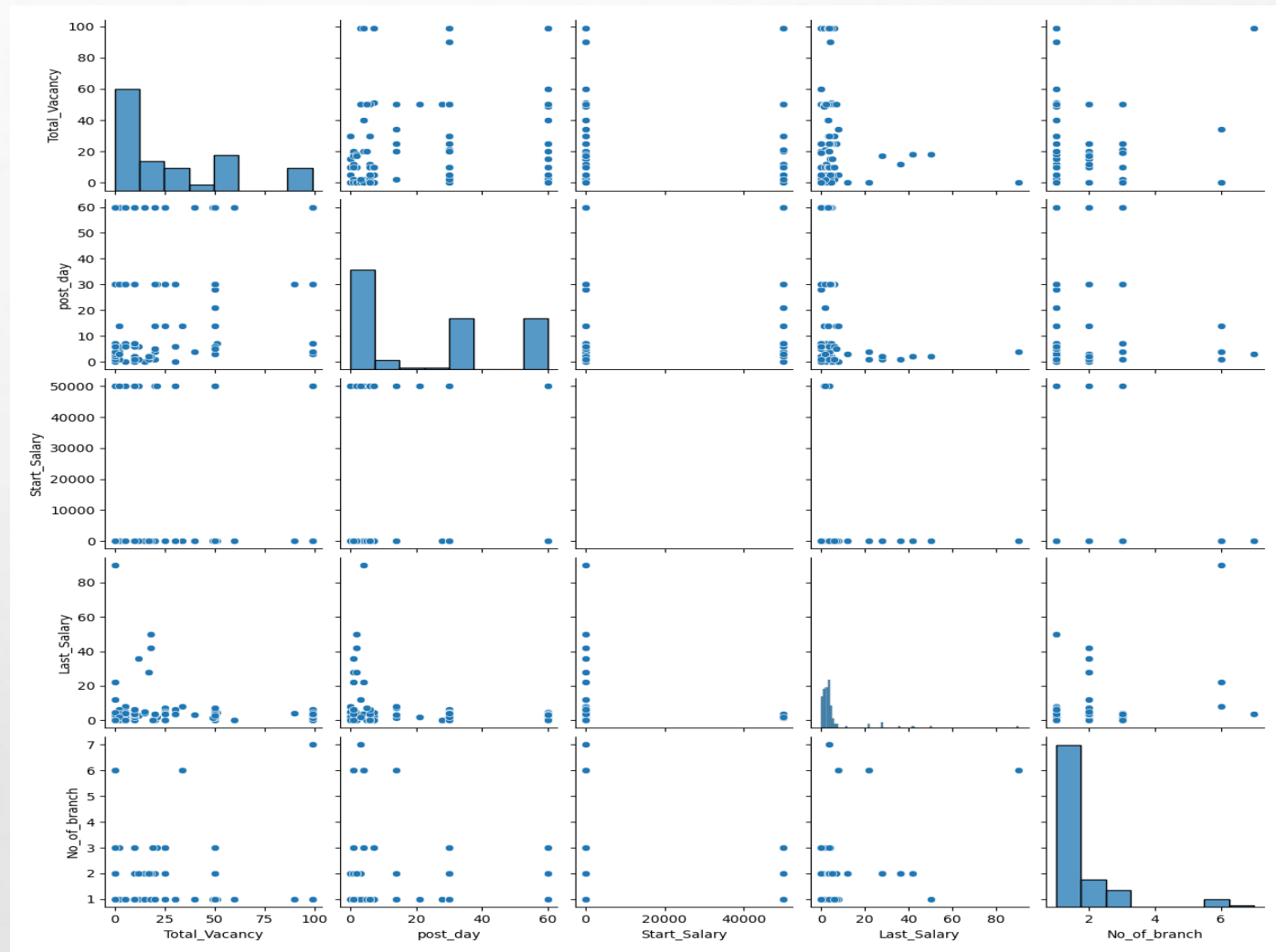
### Correlation between all numerical columns

```
[17]: corr=df.select_dtypes(["int","float"]).corr()
corr
```

```
[17]:
```

	Total_Vacancy	post_day	Start_Salary	Last_Salary	No_of_branch
Total_Vacancy	1.000000	0.208494	0.250848	-0.135026	-0.047914
post_day	0.208494	1.000000	-0.018743	-0.238745	-0.133363
Start_Salary	0.250848	-0.018743	1.000000	-0.134917	-0.138617
Last_Salary	-0.135026	-0.238745	-0.134917	1.000000	0.423400
No_of_branch	-0.047914	-0.133363	-0.138617	0.423400	1.000000





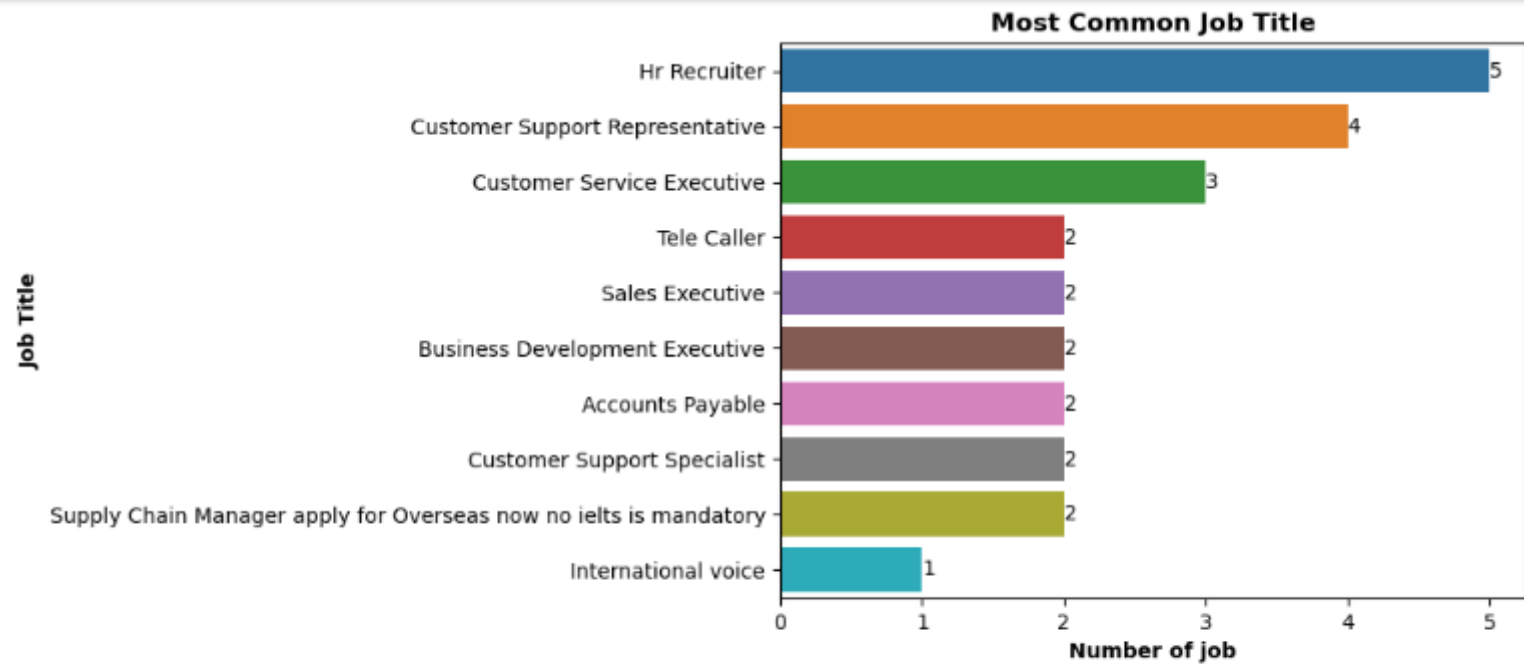
## Observation:

- **Total\_Vacancy and post\_day:** There's a weak negative correlation between these two variables, suggesting that as the number of days since the post increases, the total number of vacancies might decrease slightly.
- **Start\_Salary and Last\_Salary:** These two variables have a strong positive correlation, indicating that higher starting salaries tend to be associated with higher final salaries.
- **No\_of\_branch and Total\_Vacancy:** There's a weak positive correlation, suggesting that organizations with more branches might have a higher number of vacancies.



Great careers start here

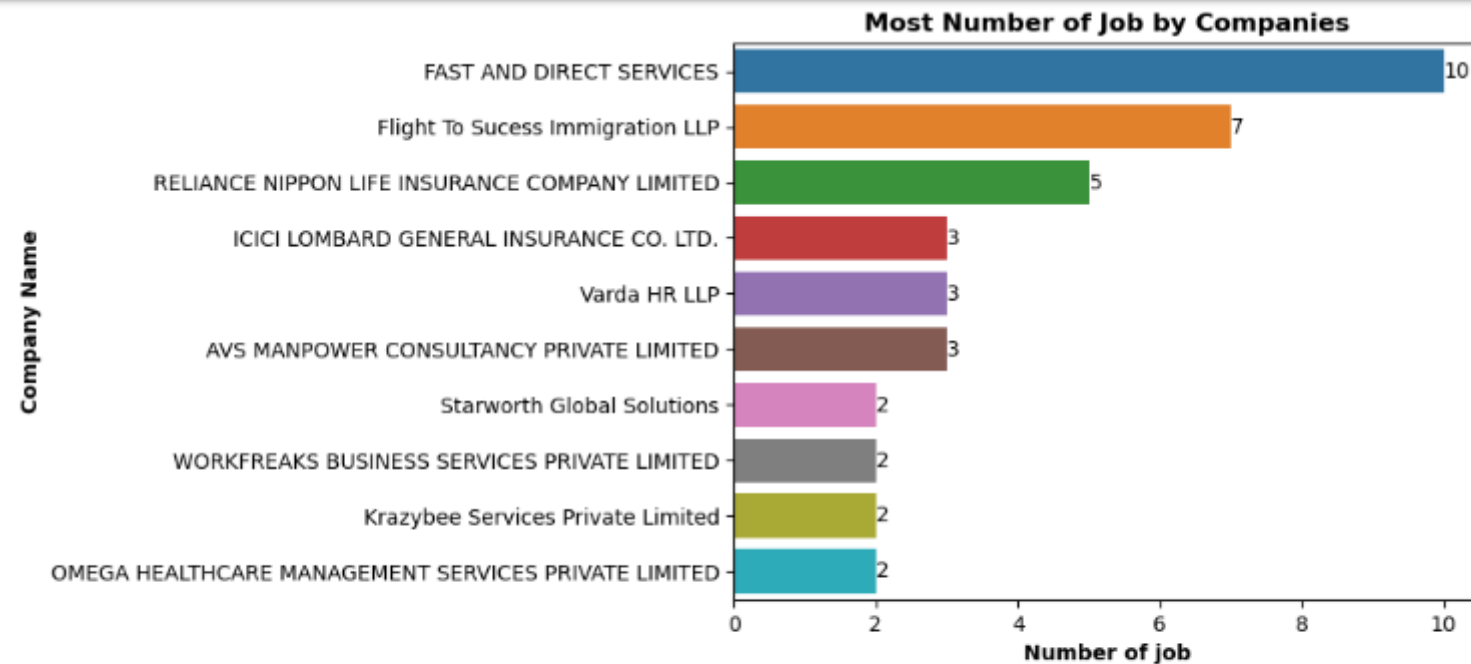
## WHAT ARE THE MOST COMMON JOB TITLES?



### Observation:

- Most common job title is HR recruiter
- after that Customer Support Representative

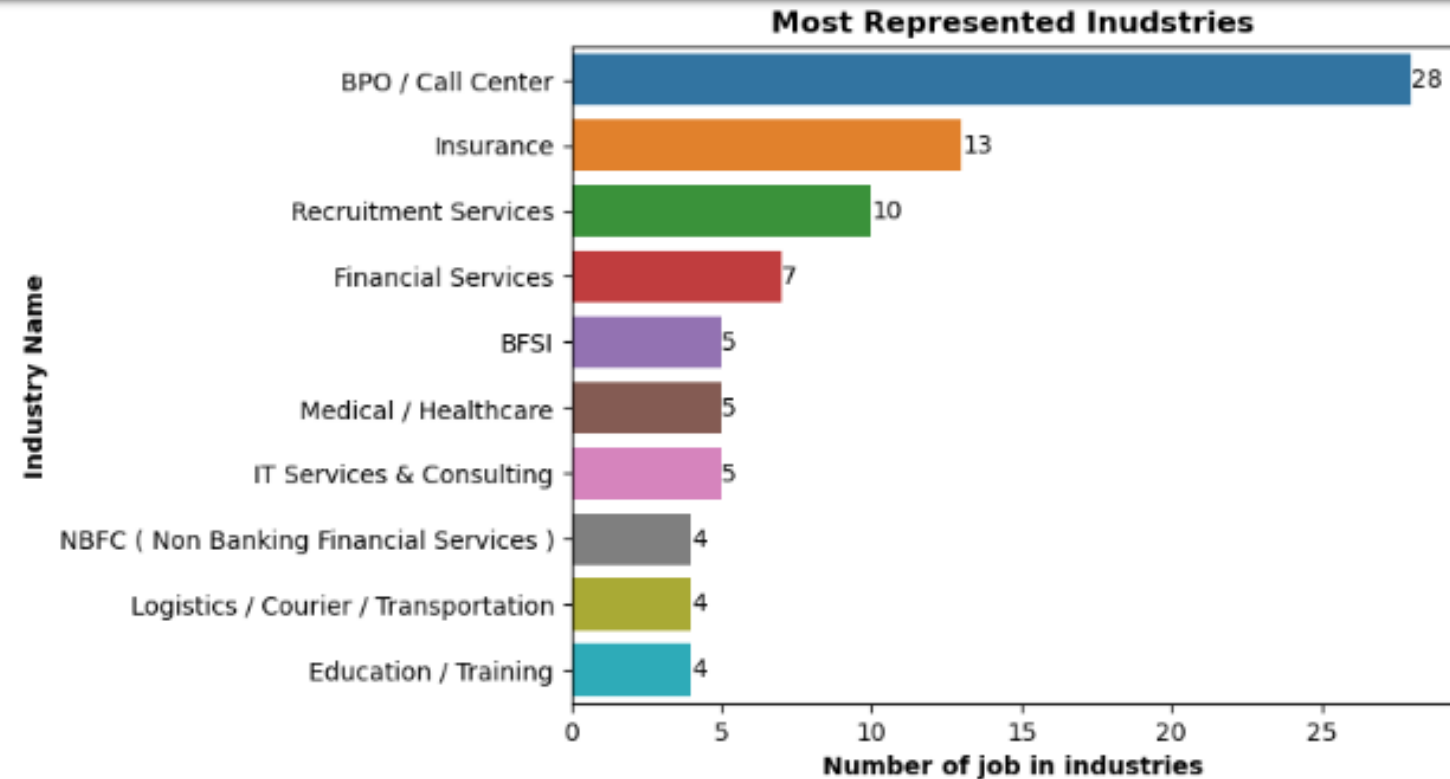
## WHICH COMPANIES HAVE THE MOST JOB OPENINGS?



### Obervation:

- Most Number of job opening are 10 from company "FAST AND DIRECT SERVICES".

## WHAT INDUSTRIES ARE MOST REPRESENTED IN THE DATASET?

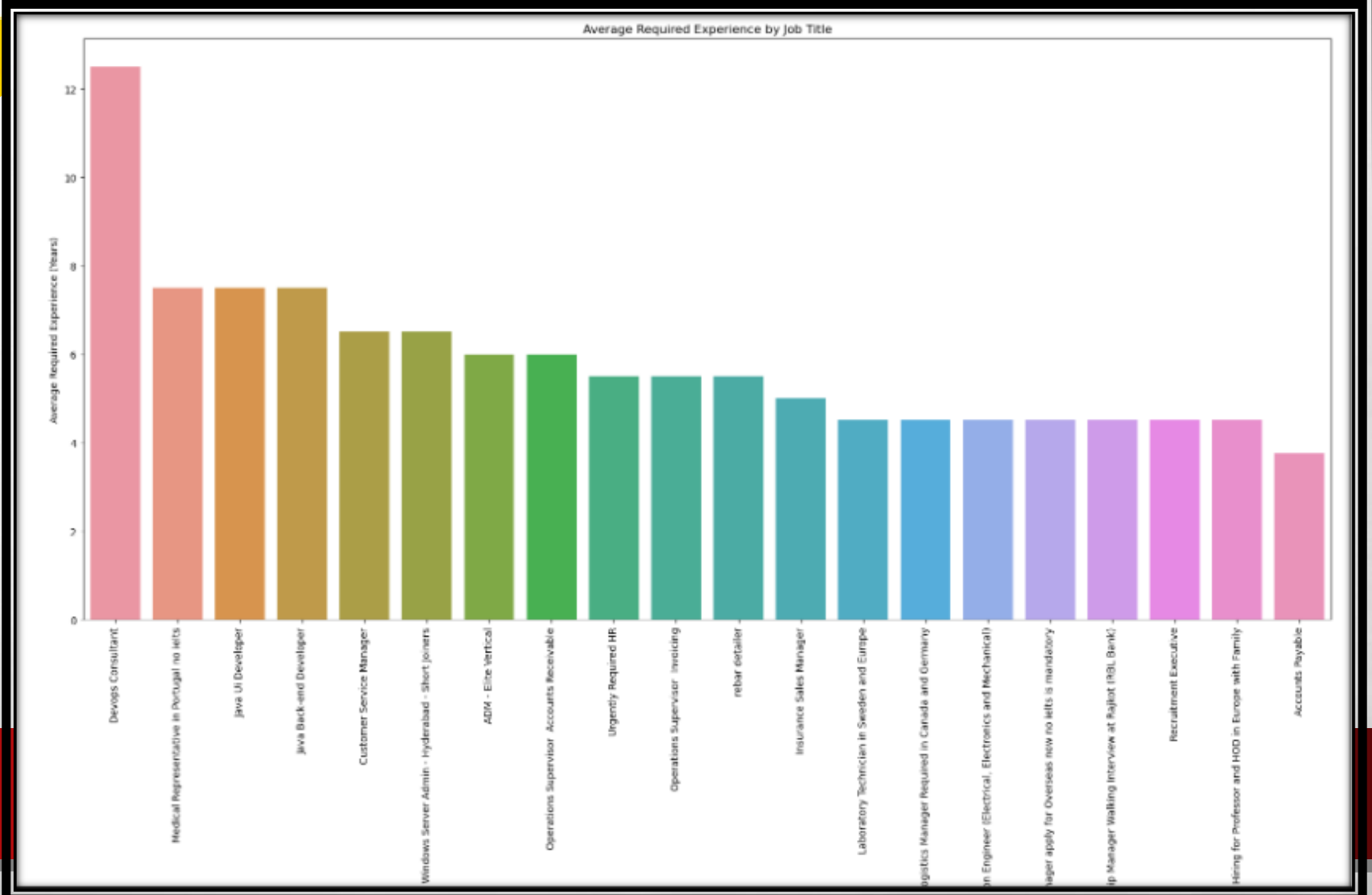


Observation:

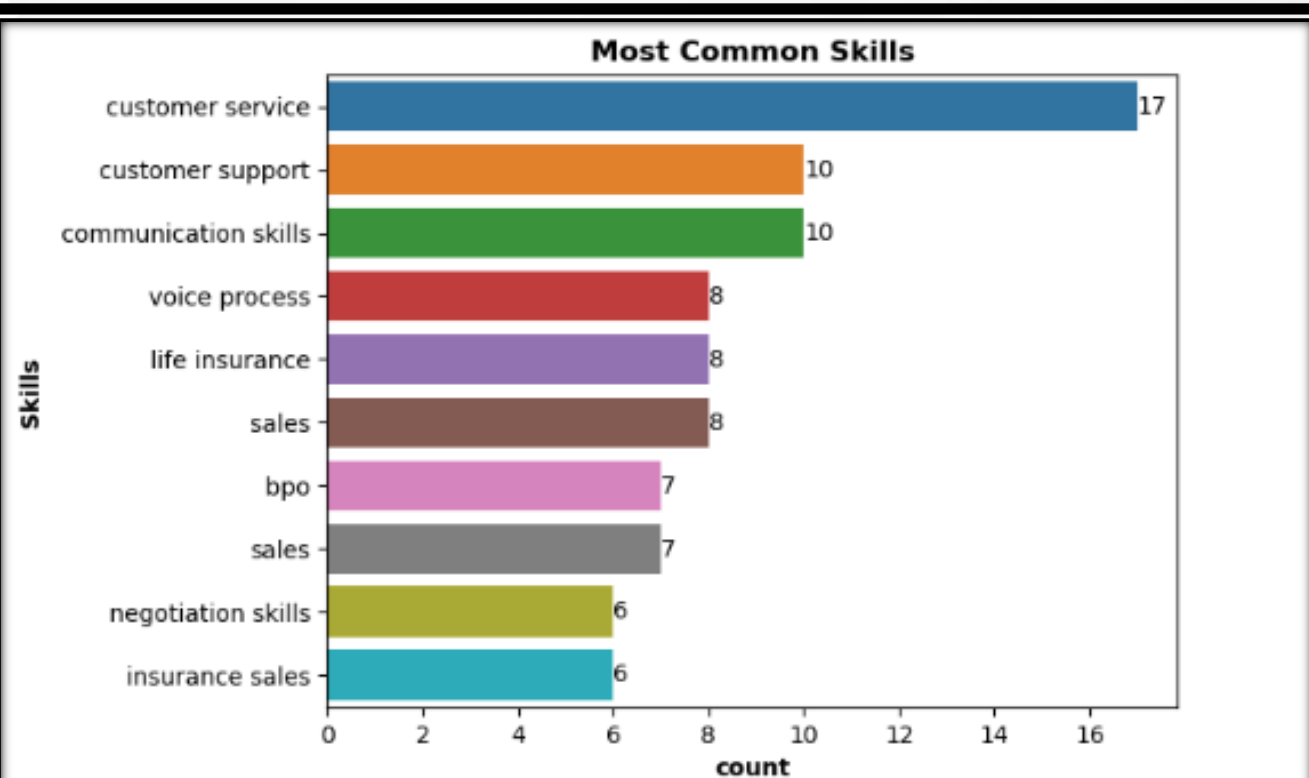
- BPO / Call Center has highest number of job opening with job number 28
- and following Insurance industries has 2nd highest number of opening with job number 13



# HOW DOES REQUIRED EXPERIENCE (EXP) VARY ACROSS JOB TITLES?



# WHAT ARE THE MOST COMMON SKILLS LISTED FOR JOBS?

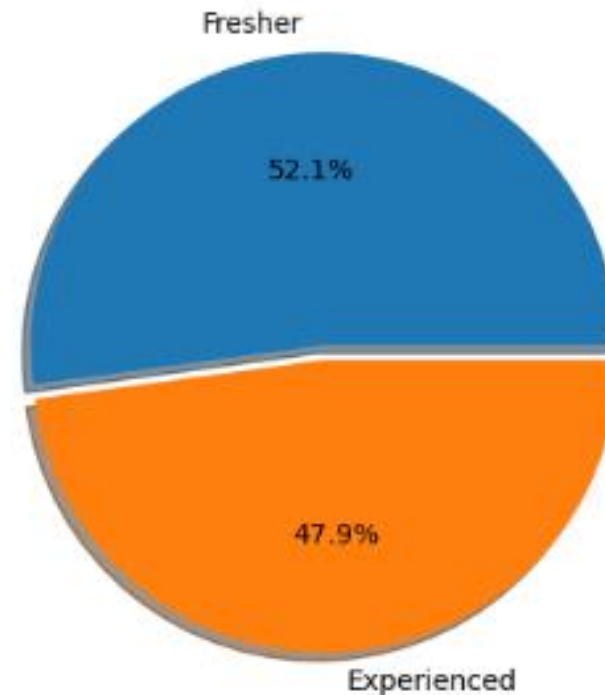


## Observation:

- most common skill are customer service to get the job
- after that customer support has second highest skill
- and following communication skills are at 3rd most common skill for job



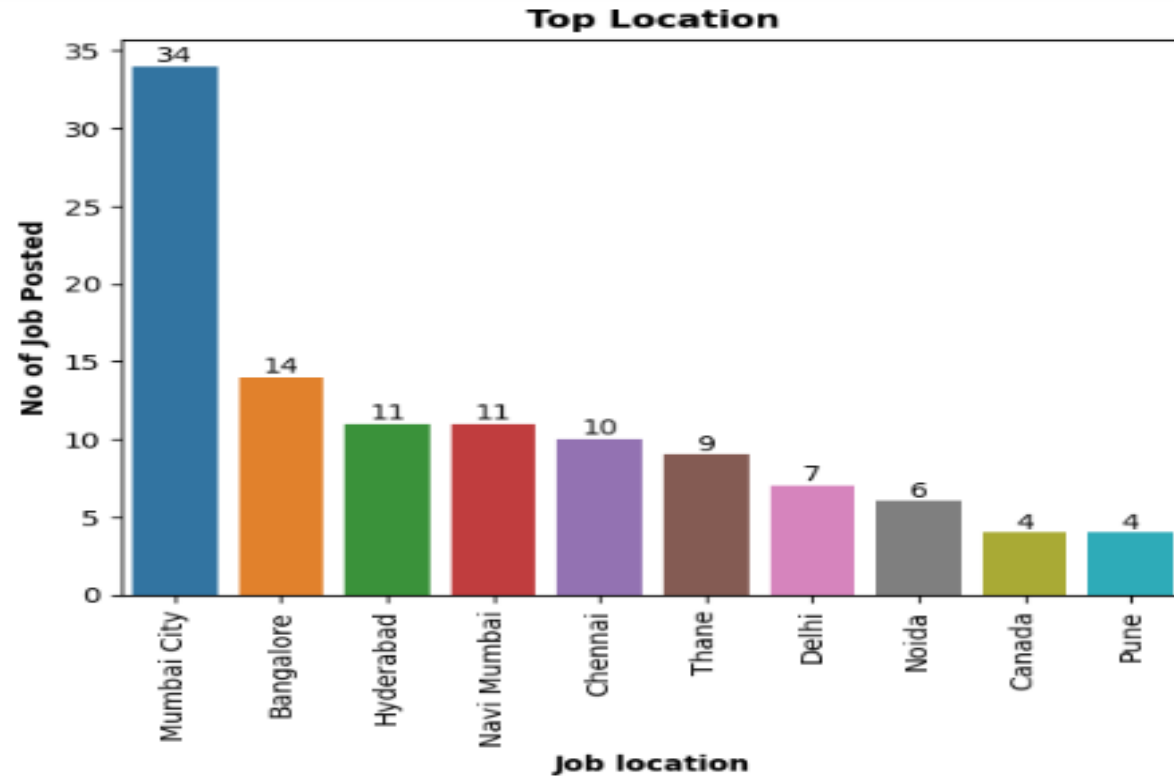
## HOW DOES THE DISTRIBUTION OF EXPERIENCE LEVELS (E.G., "FRESHER," "EXPERIENCED") LOOK ACROSS THE DATASET?



### Observation:

- The distribution is nearly equal, with a slight majority of job postings targeting freshers.

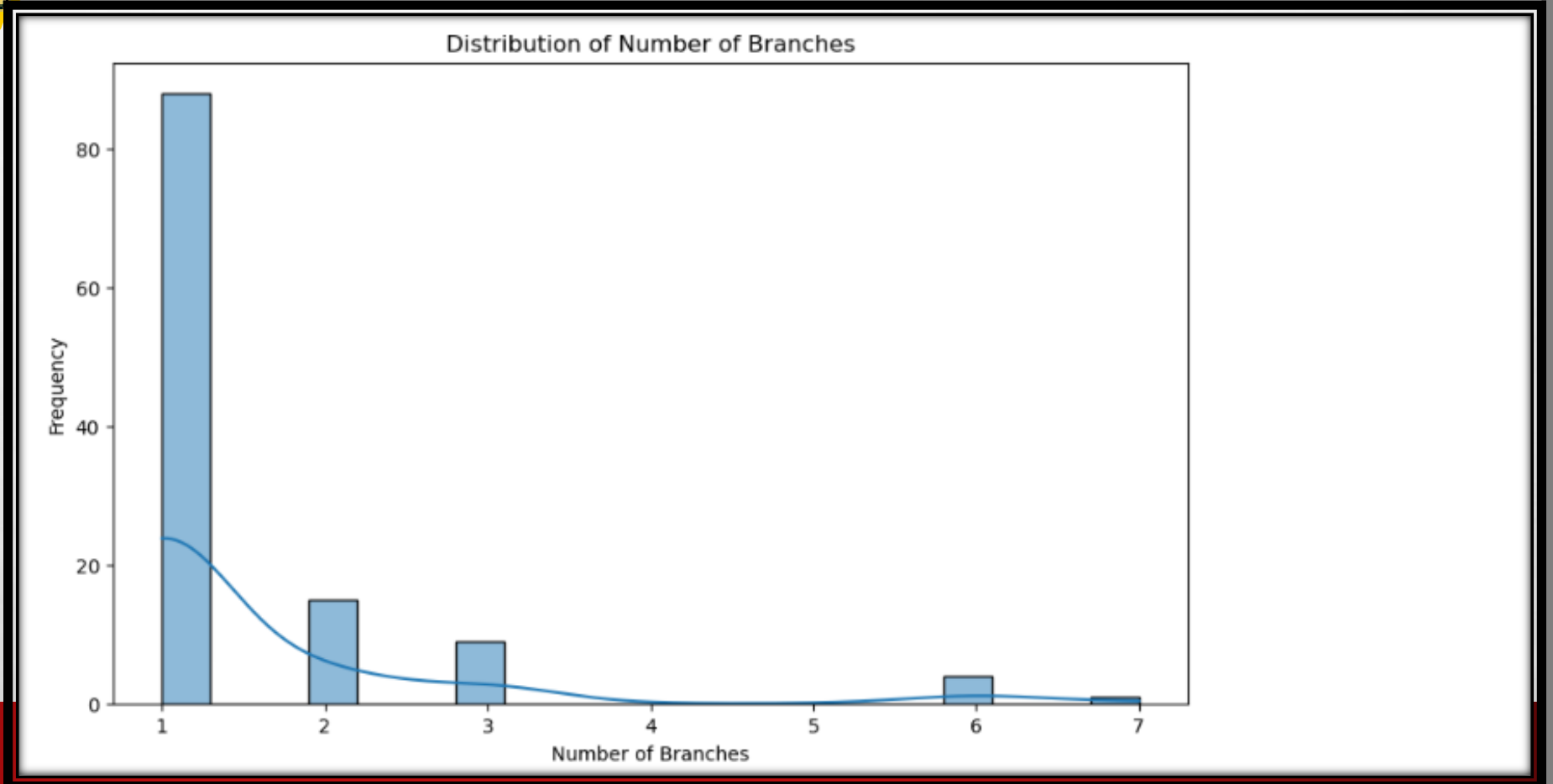
## WHAT ARE THE TOP LOCATIONS FOR THESE JOB LISTINGS

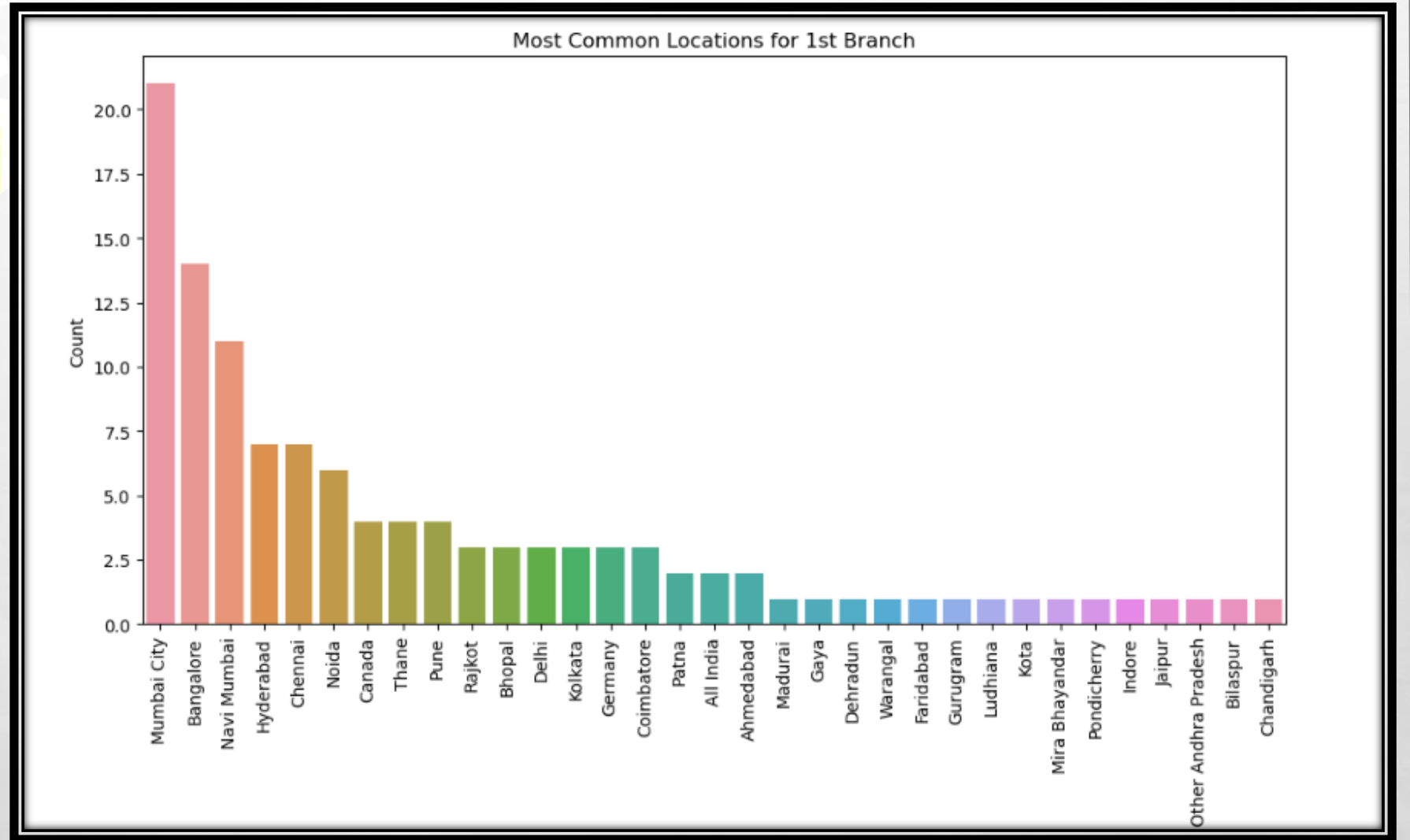


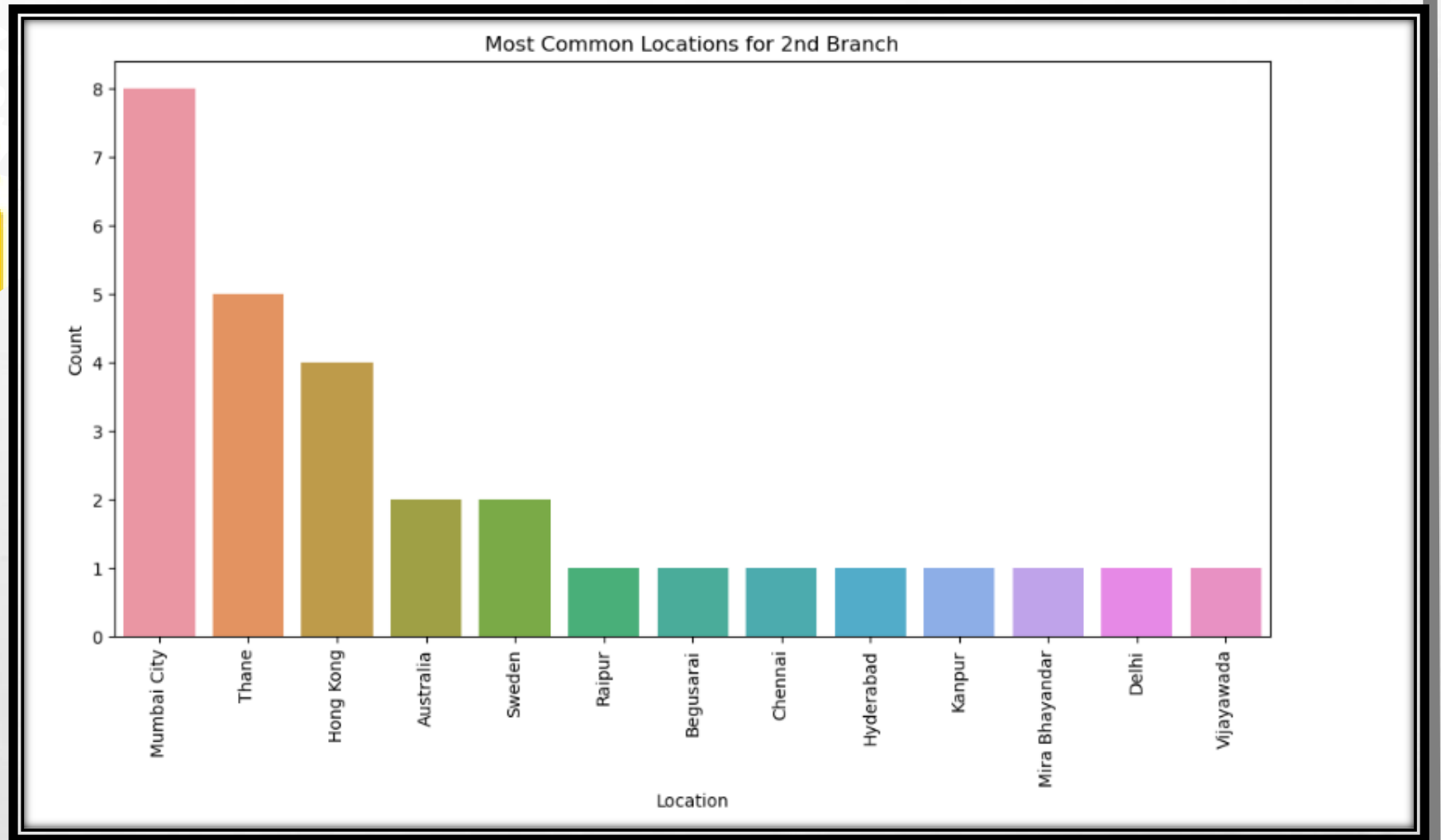
### Observation:

- mumbai city are top location with the highest number of job 34
- after that bangalore city 14 jobs and then hyderabad with 11 jobs.

## HOW MANY BRANCHES DO COMPANIES TYPICALLY HAVE, AND WHERE ARE THEY LOCATED?







### Observation:

- **Dominance of Mumbai City:** Mumbai City has the highest frequency, indicating a significant number of job postings or activities associated with this location.
- **Top Cities:** Bangalore, Hyderabad, and Navi Mumbai follow Mumbai City in terms of frequency, suggesting these cities are also major hubs for the activities represented in the dataset.
- **Long Tail:** A long tail of locations with lower frequencies is observed, indicating a diverse geographical spread of the data.



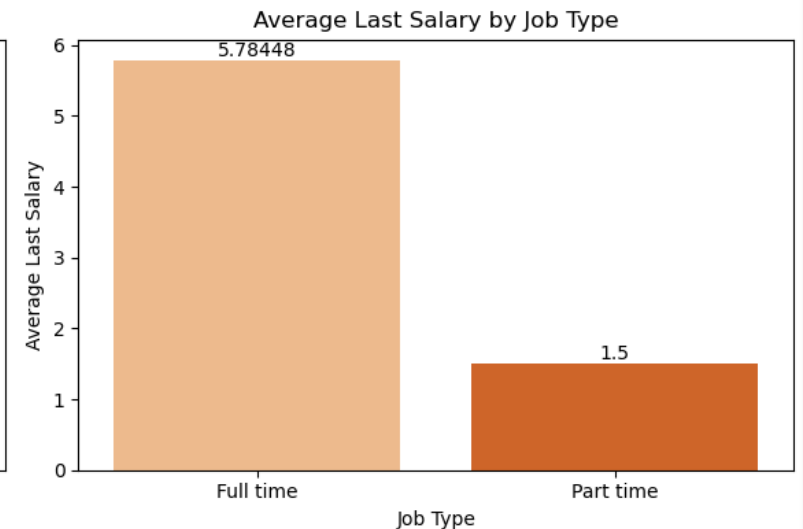
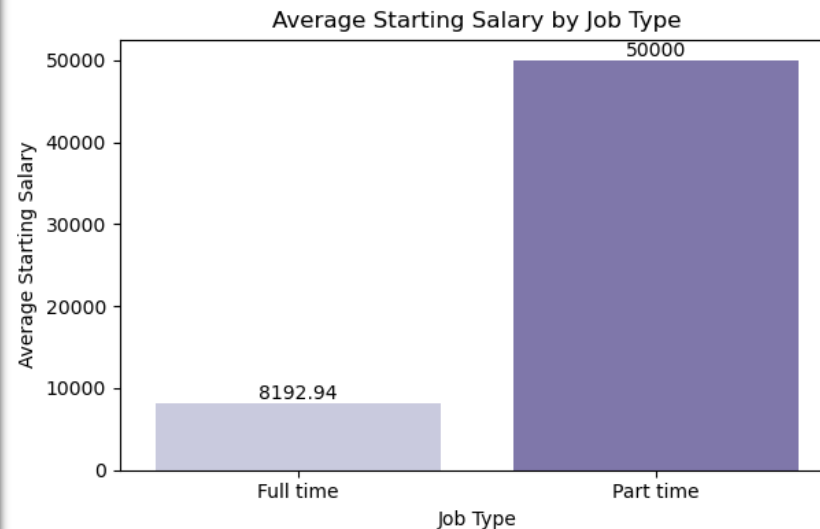
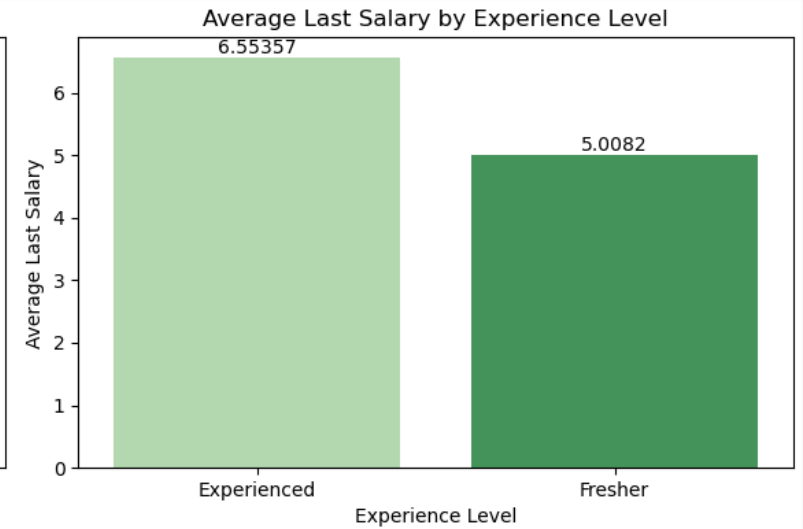
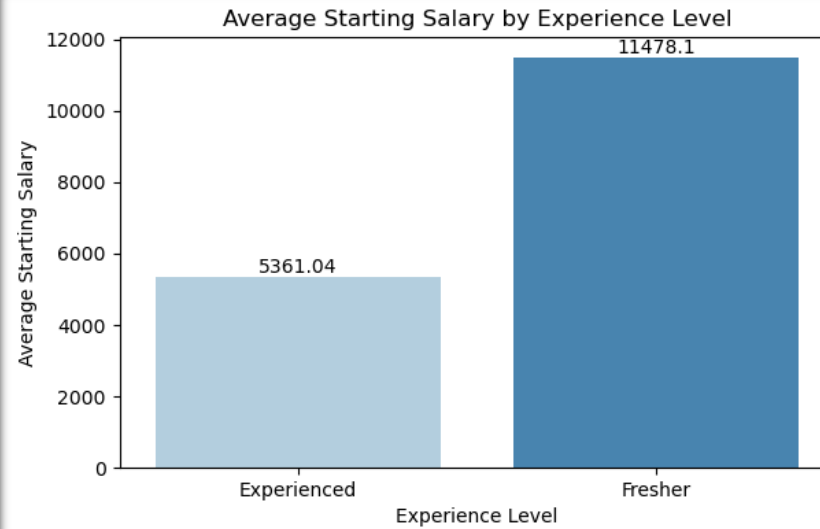
## WHAT IS THE RANGE OF STARTING AND LAST SALARIES FOR JOBS IN EACH INDUSTRY?

	Industries	start_salary_min	start_salary_max	last_salary_min	last_salary_max
0	Accounting & Auditing	1.5	1.5	4.5	4.5
1	Automobile / Auto Ancillaries	1.5	1.5	2.5	2.5
2	BFSI	0.0	50000.0	0.0	4.5
3	BPO / Call Center	0.0	50000.0	0.0	6.0
4	Banking	2.5	3.5	3.5	8.0
5	Department	0.0	50000.0	0.0	5.0
6	Education / Training	1.0	50000.0	1.5	50.0
7	Engineering / Construction	1.0	3.0	2.0	4.0
8	FinTech	0.0	2.5	0.0	4.0
9	Financial Services	0.0	50000.0	0.0	4.0
10	Gifts / Toys / Stationary	2.0	2.0	2.5	2.5
11	Hotel / Restaurant	0.0	0.0	0.0	0.0
12	IT Services & Consulting	0.0	50000.0	0.0	90.0
13	Insurance	0.0	50000.0	0.0	8.0
14	KPO / Analytics	1.0	50000.0	1.5	1.5
15	Logistics / Courier / Transportation	2.0	30.0	5.0	36.0
16	MFI ( Micro Finance )	2.5	2.5	4.5	4.5
17	Management Consulting / Strategy	0.0	1.5	0.0	3.0
18	Manufacturing	1.5	2.0	3.0	4.5
19	Medical / Healthcare	2.0	30.0	3.0	42.0
20	Metal / Iron / Steel	1.0	1.0	1.5	1.5





# HOW DOES THE SALARY VARY WITH EXPERIENCE AND JOB TYPE?





salary\_by\_exp

	classify_exp	start_salary_min	start_salary_max	start_salary_avg	last_salary_min	last_salary_max	last_salary_avg
0	Experienced	0.0	50000.0	5361.044643	0.0	50.0	6.553571
1	Fresher	0.0	50000.0	11478.065574	0.0	90.0	5.008197

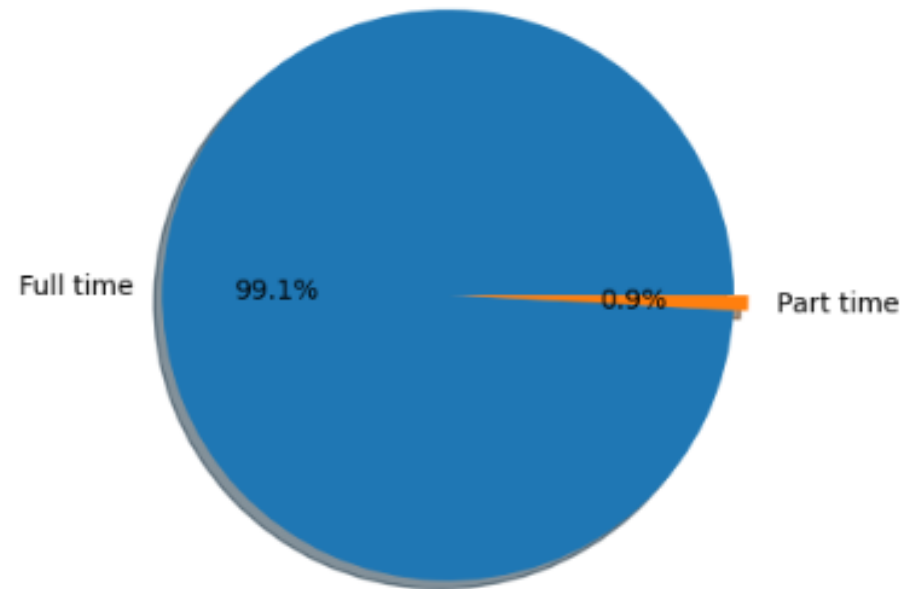
salary\_by\_job\_type

	job_type	start_salary_min	start_salary_max	start_salary_avg	last_salary_min	last_salary_max	last_salary_avg
0	Full time	0.0	50000.0	8192.935345	0.0	90.0	5.784483
1	Part time	50000.0	50000.0	50000.000000	1.5	1.5	1.500000

## Observation:

- **Higher Starting Salaries for Freshers:** Surprisingly, freshers seem to have a higher average starting salary compared to experienced candidates. This could be due to factors like higher demand for fresh talent in specific sectors or increased salary expectations for freshers.
- **Significant Salary Growth for Experienced Candidates:** Experienced candidates, despite having a lower starting salary, experience a substantial increase in their last salary, indicating career progression and salary growth over time.
- **Higher Salaries for Part-Time Jobs:** Interestingly, part-time jobs have a higher average starting and last salary compared to full-time jobs. This could be attributed to specialized roles, higher pay rates for part-time work, or other factors.
- **Limited Data Points:** The lower average salary for full-time jobs might be influenced by a smaller sample size or a specific set of job roles.

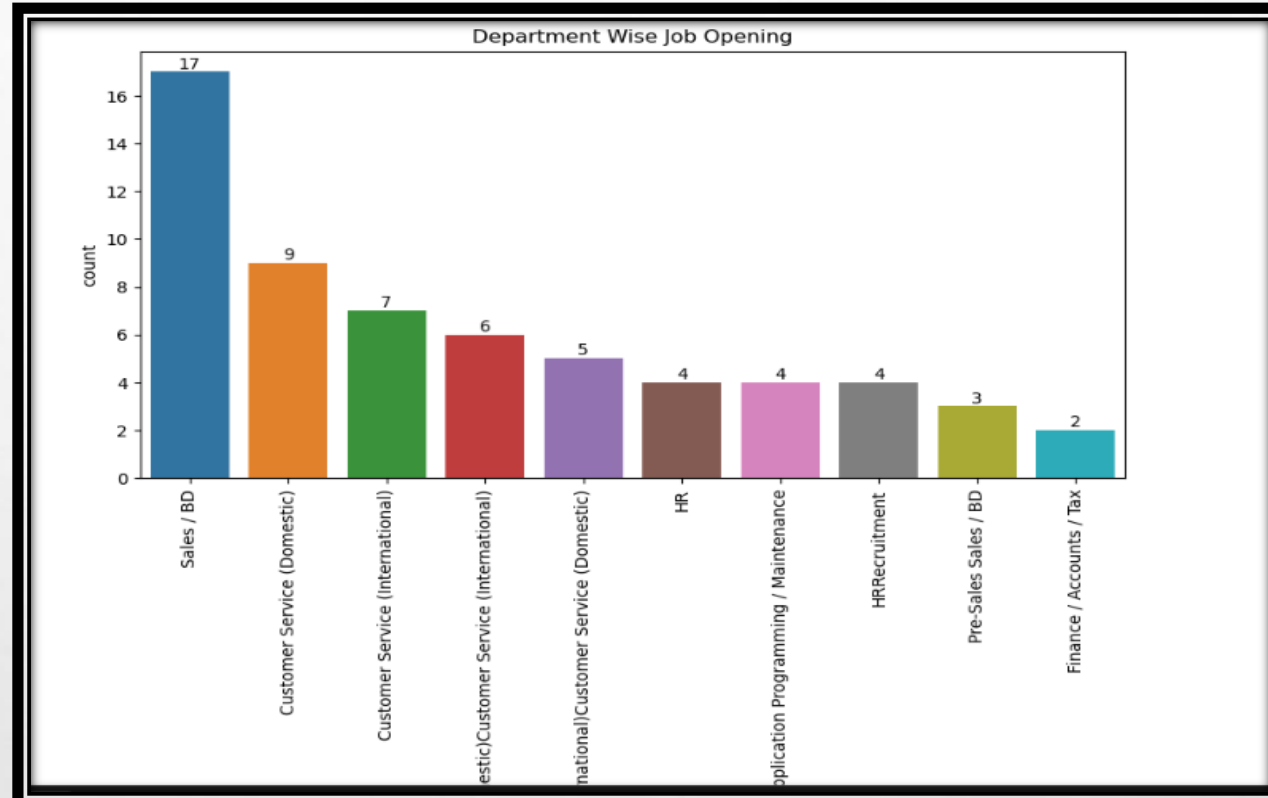
## WHAT PROPORTION OF JOBS ARE FULL-TIME VERSUS OTHER TYPES?



### Observation:

- Dominance of Full-Time Roles: A significant majority of the job postings (99.1%) are for full-time positions.
- Minority of Part-Time Roles: Only 0.9% of the job postings are for part-time positions.

# WHAT DEPARTMENTS HAVE THE MOST JOB OPENINGS?

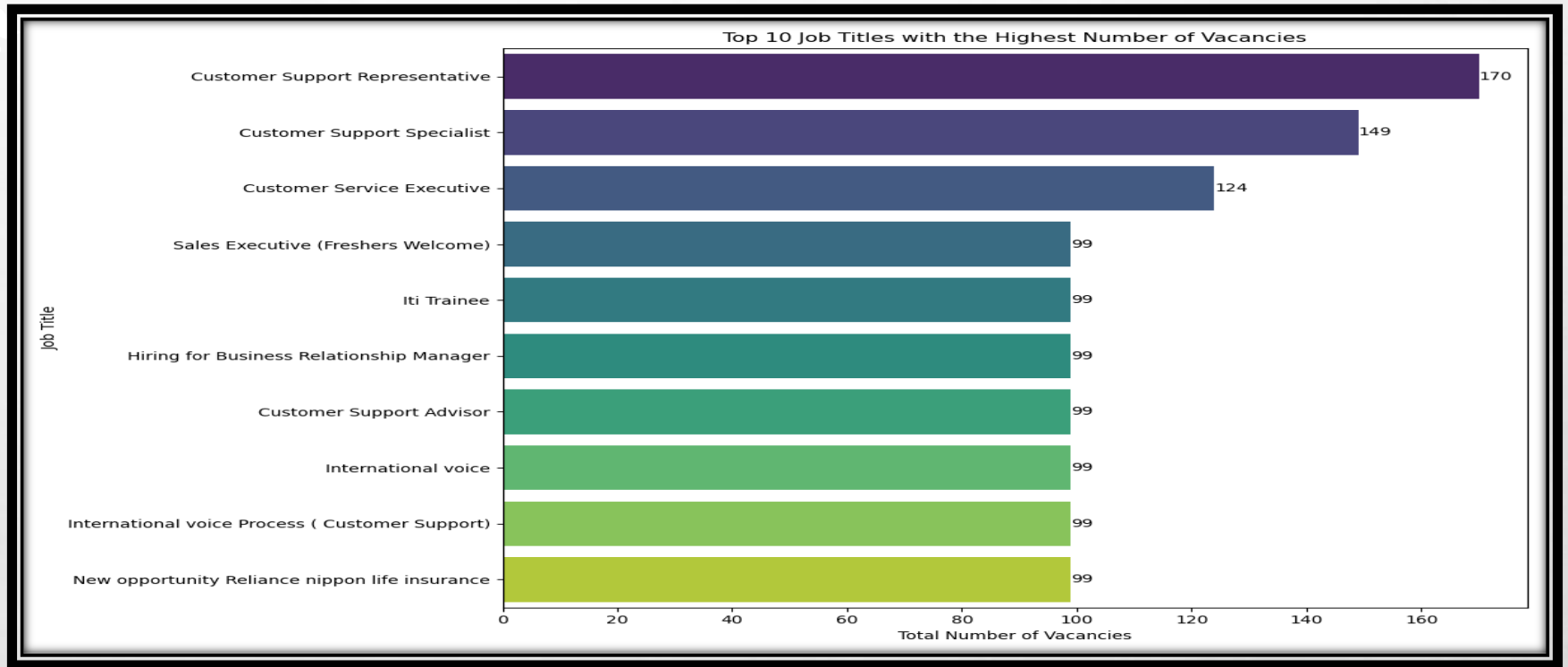


## Observation:

- Sales/BD: This department has the highest number of job openings, indicating a significant demand for sales and business development professionals.
- Customer Service: Both domestic and international customer service departments have a substantial number of openings, suggesting a need for customer support personnel.
- HR and HR Recruitment: The HR department and HR recruitment roles have a moderate number of openings, indicating ongoing hiring needs for human resources professionals.
- Technical Roles: Departments like Application Programming/Maintenance and Pre-Sales Sales/BD have a fewer number of openings, suggesting a lower demand for technical roles compared to sales and customer service.
- Finance and Accounts: The Finance/Accounts/Tax department has the lowest number of openings, indicating a relatively stable staffing requirement in this area.



# ARE THERE PATTERNS IN THE NUMBER OF VACANCIES ACROSS DIFFERENT JOB TITLES?



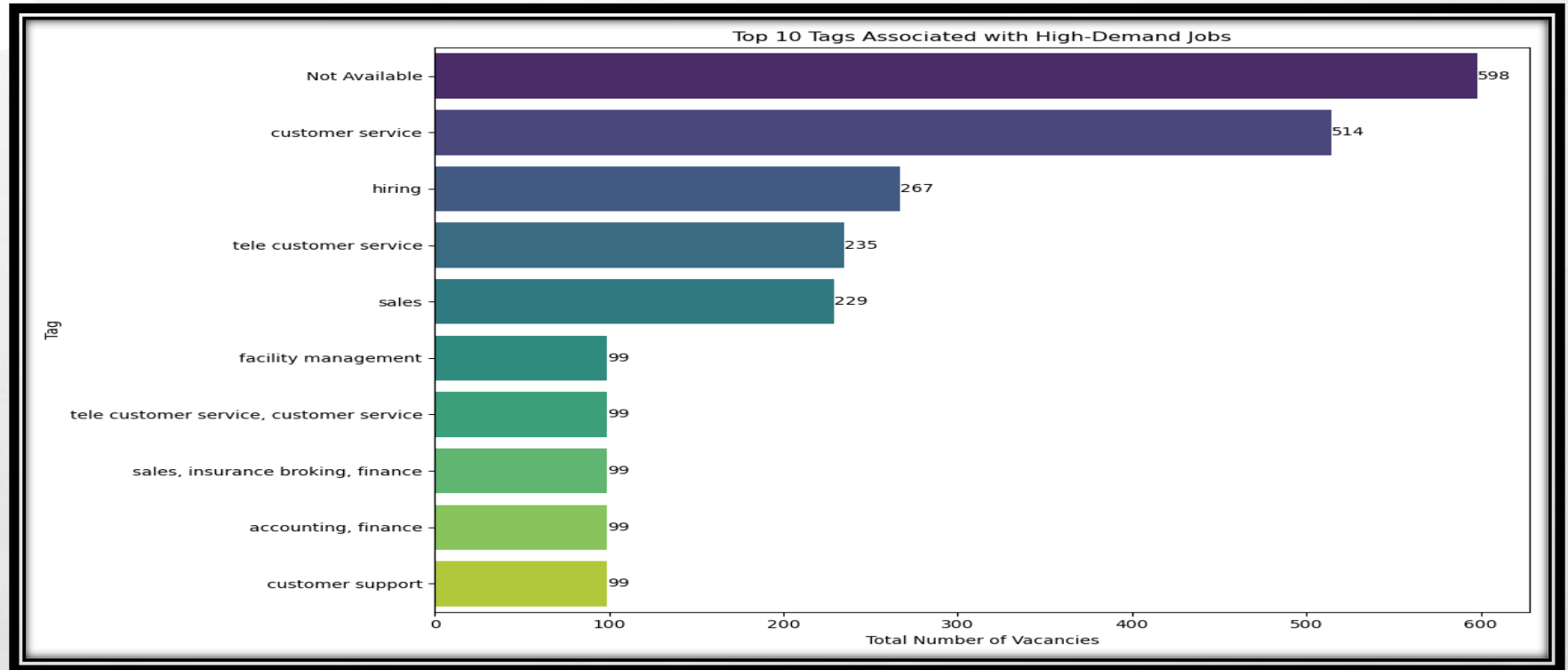
## Observation:

- **Customer Support Dominance:** Customer support roles, such as Customer Support Representative, Specialist, and Executive, are among the top job titles with the highest number of vacancies. This suggests a significant demand for customer support professionals.
- **Sales and Business Development:** Sales Executive (Freshers Welcome) and Hiring for Business Relationship Manager indicate a need for sales and business development talent.
- **IT and Technical Roles:** While not as prominent as customer support roles, there are vacancies for IT-related positions like IT Trainee and International Voice Process roles.





# WHAT TAGS ARE MOST ASSOCIATED WITH HIGH-DEMAND JOBS?



## Observation:

- **"Not Available" Tag Dominates:** The tag "Not Available" has the highest number of vacancies, indicating a significant number of job postings without specific tag information.
- **Customer Service and Related Tags:** Tags like "customer service," "tele customer service," and "customer support" are prominently featured, suggesting a high demand for customer service professionals.
- **Sales and Finance:** Tags related to sales, insurance broking, finance, and accounting also appear in the top 10, indicating a need for professionals in these domains.
- **Facility Management:** This tag, though less frequent, still represents a notable number of vacancies.



**THANK YOU**

**NISHANT**

**NISHANT575051@GMAIL.COM**

**LINKDIN : [WWW.LINKEDIN.COM/IN/NISHANT-KUMAR-DATA-ANALYST](http://WWW.LINKEDIN.COM/IN/NISHANT-KUMAR-DATA-ANALYST)**

**Github : <https://github.com/Nishant6020>**

**shine!**  
com

Great careers start here

