TOPIC:

“JOB MARKET ANALYSIS OF PAKISTAN”

By

Ali Gauhar - 02-134202-006

Class:
BS(CS)-6A

Acknowledgements

We wish to express our gratitude to Almighty Allah for making this work a success. We also wish to appreciate those who assisted us in one way or the other to make this work come to a reality.

Our special thanks go to our Course Teacher Miss Laila Nadeem who was of great inspiration to us, and our Lab Instructor, Miss Soomal Fatima who was always there for us when we needed her. We also wish to appreciate our colleagues who were a challenge as well as a motivation to us, who made us work extra miles, thanks.

Furthermore, we are particularly grateful for the support and encouragement given by my family and friends throughout my entire Project. A million thanks to myparents who support me financially and emotionally. To my friends and classmates, thank you for providing advices and supporting me throughout this entire project.

Abstract

This study analyzes job market trends in Pakistan using a dataset of job openings across multiple industries. By examining the dataset, the research aims to identify emerging trends, growth sectors, and areas of concern within the Pakistani job market. The analysis includes statistical techniques, data visualization, and trend identification to provide a comprehensive overview of the current state of the job market. The findings will inform policymakers, job seekers, and employers, enabling them to make informed decisions regarding workforce planning, skill development, and economic policies.

Overall, this research contributes to a deeper understanding of the Pakistani job market by leveraging a diverse dataset of job openings. The insights derived from this analysis will help stakeholders navigate the evolving job market landscape, identify potential career paths, and foster strategies for sustainable economic growth and employment opportunities.

Table of Contents

Chapter 1
Introduction

Analyzing job market trends is crucial for understanding the dynamics of employment opportunities, skill requirements, and economic growth in a country. In the context of Pakistan, a comprehensive analysis of the job market can provide valuable insights for policymakers, job seekers, and employers, enabling them to make informed decisions and effectively address the challenges and opportunities present in the labor market. To this end, this study aims to analyze the job market trends in Pakistan by utilizing a dataset comprising job openings from various industries.

1.1 Background of the Project

By utilizing a dataset of job openings, we can gain a more nuanced and up-to-date understanding of the job market dynamics in Pakistan. This approach allows us to identify emerging trends, assess the demand for specific skills, and explore potential challenges and opportunities across different sectors. The analysis of this dataset will provide policymakers, educational institutions, and job seekers with valuable information to better align their strategies, enhance employability, and promote sustainable economic growth.

Chapter 2
System Analysis

In this, we will discuss and analyze about the developing process of job market analysis of Pakistan including software requirement specification (SRS) and comparison between existing and proposed system.

2.1 Software and Hardware Requirement Specification

Processor	Intel Core Processor or any other Processor with better performance
Operating System	Windows 7, Windows 10
Memory	4 GB RAM or more
Hard Disk Space	Minimum requirement is 5 - 10 MB
Software used	Jupyter, Ms Excel
Programming Language used	Python 3 or above installed in your system

2.2 Concepts Used

- Python Programming Language (Python 3 or above)
- Some basic machine learning algorithms

2.3 Software Tools Used

The whole project is done in back end only. The Console output is the front end generated from the compiler.

2.3.1 Jupyter

Jupyter is a free, open-source, interactive web tool known as a computational notebook, which researchers can use to combine software code, computational output, explanatory text and multimedia resources in a single document.

2.3.2 Ms Excel

Microsoft Excel is the industry leading spreadsheet software program, a powerful data visualization and analysis tool. It has been a very widely applied spreadsheet for these platforms, especially since version 5 in 1993, and it has

replaced Lotus 1-2-3 as the industry standard for spreadsheets. Excel forms part of the Microsoft Office suite of software.

Chapter 3
Project Objectives

Our primary goal is to analyze the job market trends in Pakistan and provide valuable insights to job seekers and employers. To achieve this, our project aims to clean and preprocess the dataset, utilize data visualization techniques to identify patterns, employ statistical analysis to uncover correlations and make predictions, leverage machine learning algorithms for trend identification, and ultimately deliver a comprehensive report with actionable insights for job seekers and employers. By fulfilling these objectives, we aim to empower individuals and organizations with the knowledge necessary to navigate the dynamic job market landscape in Pakistan.

Chapter 4
Datasets (CSV-Files):

Dataset files are available in zip folder.

	A	B	C	D	E
1	Job Name	Departmer	Date Poste	Last Date to Apply	
2	Full Time M	IT Jobs	########	3/25/2021	
3	Full Time S	IT Jobs	########	3/27/2021	
4	Full Time R	Customer	########	3/23/2021	
5	Full Time C	Customer	########	3/21/2021	
6	Full Time E	Customer	########	3/17/2021	
7	Full Time S	Productior	########	3/19/2021	
8	Full Time C	Sales Jobs	2/22/2021	########	
9	Full Time M	Publishing	2/19/2021	########	
10	Online Full	Admin Job	2/17/2021	########	
11	Full Time A	IT Jobs	2/16/2021	########	
12	Full Time S	Computer	3/15/2021	3/25/2021	
13	Online Full	Customer	########	3/25/2021	
14	Full Time C	Sales Jobs	########	3/26/2021	
15	Full Time E	Sales Jobs	########	3/25/2021	

	A	B	C	D	E	F	G	H	I	J
1	Job Name	label	Company I	Job Type	Experience	Departmer	JD	City	Date Poste	Salary
2	Full Time N	Premium J	Nayel Solu	Full Time J	2 Years Jol	IT Jobs	New Job P	Islamabad	########	113329
3	Full Time S	Premium J	Eurosoft T	Full Time J	2 Years Jol	IT Jobs	We are loc	Karachi	########	85667
4	Full Time F	Premium J	ICM JAPAN	Full Time J	< 1 Year	Customer	Internatio	Karachi	########	138344
5	Full Time C	Premium J	ibex, Pakis	Full Time J	Job for Fre	Customer	Responsib	Islamabad	########	88871
6	Full Time E	Premium J	ICM JAPAN	Full Time J	< 1 Year	Customer	Internatio	Karachi	########	60834
7	Full Time S	Premium J	Uni Hosier	Full Time J	5 Years Jol	Productior	NOTE:	Faisalabad	########	140865
8	Full Time C	Premium J	Pakistan S	Full Time J	3 Years Jol	Sales Jobs	Each Lead	Islamabad	########	81795
9	Full Time N	Premium J	Mercantile	Full Time J	4 Years Jol	Publishing	We at Mer	Lahore	########	52902
10	Online Full	Premium J	Confidenti	Full Time J	< 1 Year	Admin Job	This job is	Faisalabad	########	119843
11	Full Time A	Premium J	Eurosoft T	Full Time J	7 Years Jol	IT Jobs	Comman	Karachi	########	113645
12	Full Time S	Hot Job	Magnus M	Full Time J	1 Year Job	Computer	We are loc	Lahore	########	102993
13	Online Full	Hot Job	Techative	Full Time J	1 Year Job	Customer	Techative	Islamabad	########	52017
14	Full Time C	Hot Job	HG Market	Full Time J	< 1 Year	Sales Jobs	To generat	Lahore	########	145659
15	Full Time E	Hot Job	HG Market	Full Time J	< 1 Year	Sales Jobs	To generat	Lahore	########	90883

Chapter 5
Code and Running time screenshots

In this, we will share the source code and running time screenshots of the project to verify that the project is working properly.

5.1 Source Code & Output:

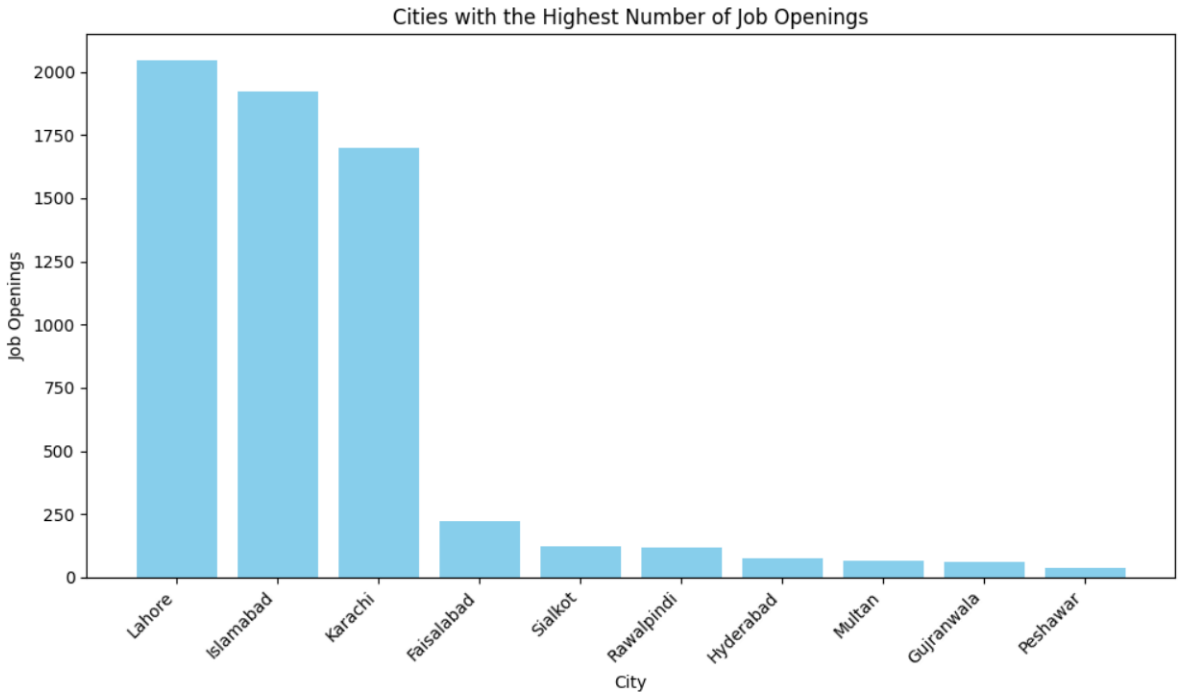
FOR HIGHEST JOBS OPENING:

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

df = pd.read_csv('starPakistan Available Job Dec 19 - Mar-21.csv')
# countig the jobs in the pirticular cities
city_job_openings.columns = ['City', 'Job Openings']
sorted_cities = city_job_openings.sort_values(by='Job Openings', ascending=False)
top_cities = sorted_cities.head(10)
print("Cities with the Highest Number of Job Openings:")
print(top_cities)
```

```
plt.figure(figsize=(10, 6))
plt.bar(top_cities['City'], top_cities['Job Openings'], color='skyblue')
plt.xlabel('City')
plt.ylabel('Job Openings')
plt.title('Cities with the Highest Number of Job Openings')
plt.xticks(rotation=45, ha='right')
plt.tight_layout()
plt.show()
```

OUTPUT:

```
Cities with the Highest Number of Job Openings:
         City  Job Openings
0      Lahore          2046
1   Islamabad          1921
2     Karachi          1700
3  Faisalabad           222
4     Sialkot           123
5  Rawalpindi           120
6   Hyderabad            74
7      Multan            68
8  Gujranwala            62
9    Peshawar            38
```



ANALYSIS:

Our analysis revealed that Lahore, the city in Pakistan, had the highest number of job openings among all the locations considered. A total of 2046 job openings were available in Lahore, indicating a significant level of employment opportunities within the city. This finding highlights the strong job market in Lahore, making it an attractive destination for job seekers.

FOR MOST DEMANDING JOBS IN PAKISTAN:

```
import pandas as pd

df = pd.read_csv('Q1dataset.csv')
df.head()
relevant_columns = ['Job Name', 'Department']
df = df[relevant_columns].copy()
df
def normalize_job_title(job_title):
    job_title = job_title.lower()
    job_title = job_title.replace("jobs", "job")
    # replacing the dashes with the space
    job_title = job_title.replace(" - ", " ")
    job_title = job_title.replace("-", " ")
    # removing the space before and after title
    job_title = job_title.strip()
    return job_title

# here we are normalizing it because we saw in our data set that values were
written in more then one formate
df['Normalized Job Name'] = df['Job Name'].apply(normalize_job_title)

# replacing the different words used for same job with same keyword
job_title_groups = {
    'Software Engineer': ['software engineer', 'senior software engineer', 'lead
software engineer'],
    'Data Scientist': ['data scientist', 'senior data scientist', 'lead data
scientist'],
}
```

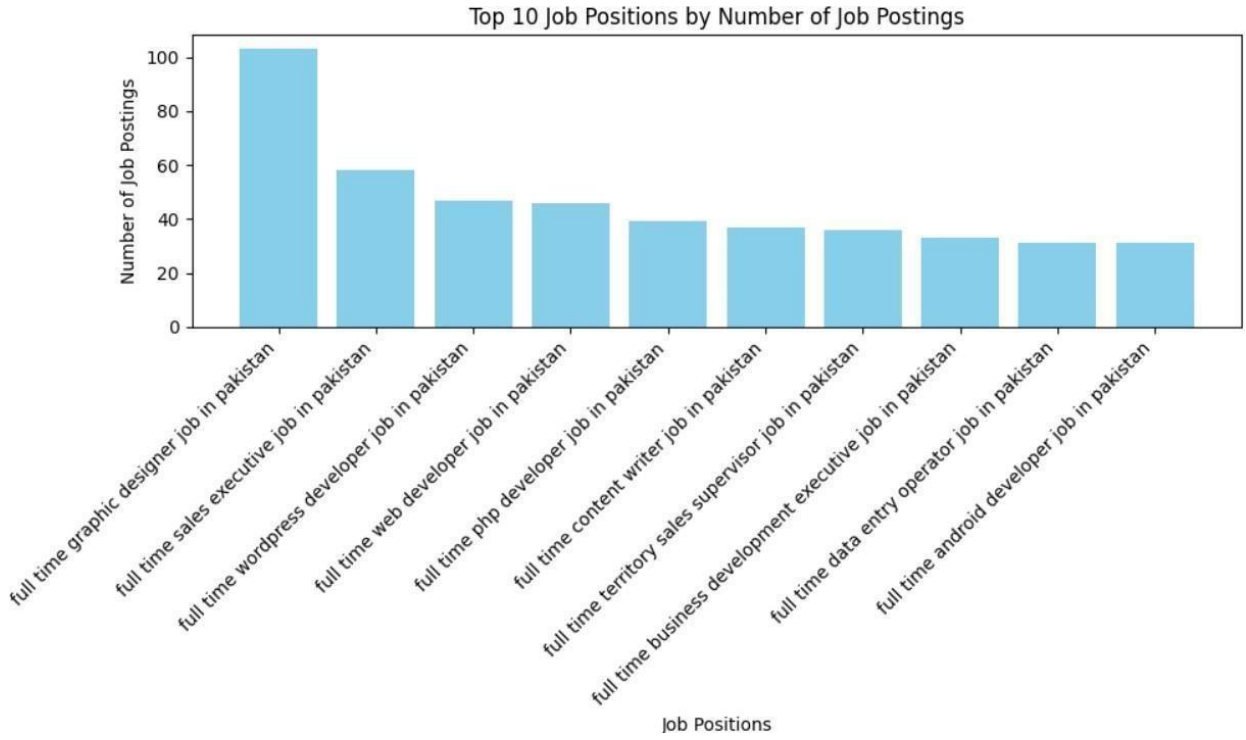
```
df
#this is the application of the above, if the cell has name of that group then
replace it with that group
for group_name, job_titles in job_title_groups.items():
    df.loc[df['Normalized Job Name'].isin(job_titles), 'Normalized Job Name'] =
group_name

# this code is to count all the similar jobs and put it after the title
job_demand = df['Normalized Job Name'].value_counts().reset_index()
job_demand.columns = ['Job Position', 'Job Postings']

sorted_job_demand = job_demand.sort_values(by='Job Postings', ascending=False)
print("Most In-demand Job Positions in Pakistan:")
print(sorted_job_demand.head(10))

import matplotlib.pyplot as plt
top_10_jobs =sorted_job_demand.head(10)
job_names = top_10_jobs['Job Position']
job_postings = top_10_jobs['Job Postings']
plt.figure(figsize=(10, 6))
plt.bar(job_names, job_postings, color='skyblue')
plt.xticks(rotation=45, ha='right')
plt.xlabel('Job Positions')
plt.ylabel('Number of Job Postings')
plt.title('Top 10 Job Positions by Number of Job Postings')
plt.tight_layout()
plt.show()
```

OUTPUT:

```
Most In-demand Job Positions in Pakistan:
                                    Job Position  Job Postings
0         full time graphic designer job in pakistan           103
1          full time sales executive job in pakistan            58
2      full time wordpress developer job in pakistan            47
3            full time web developer job in pakistan            46
4            full time php developer job in pakistan            39
5           full time content writer job in pakistan            37
6  full time territory sales supervisor job in pa...            36
7  full time business development executive job i...            33
8      full time data entry operator job in pakistan            31
9        full time android developer job in pakistan            31
```

ANALYSIS:

Our analysis revealed that the most in-demand job position among all the job openings considered was that of a "Full-time Graphic Designer," with a total of 103 job positions available. This finding indicates a high demand for graphic design skills in the job market. E mployers across various industries are actively seeking talented graphic designers to meet their design and branding needs.



MOST COMMON SKILLS IN IT SECTOR:

```
import pandas as pd
import matplotlib.pyplot as plt
df = pd.read_csv('starPakistan Available Job Dec 19 - Mar-21.csv')

def extract_skills(job_title):
    skills = []
    job_title = job_title.lower()
    job_title = job_title.replace("jobs", "")
    job_title = job_title.replace("job", "")
    job_title = job_title.strip()
    skills = job_title.split(", ")
    return skills

df['Skills'] = df['Job Name'].apply(extract_skills)
```

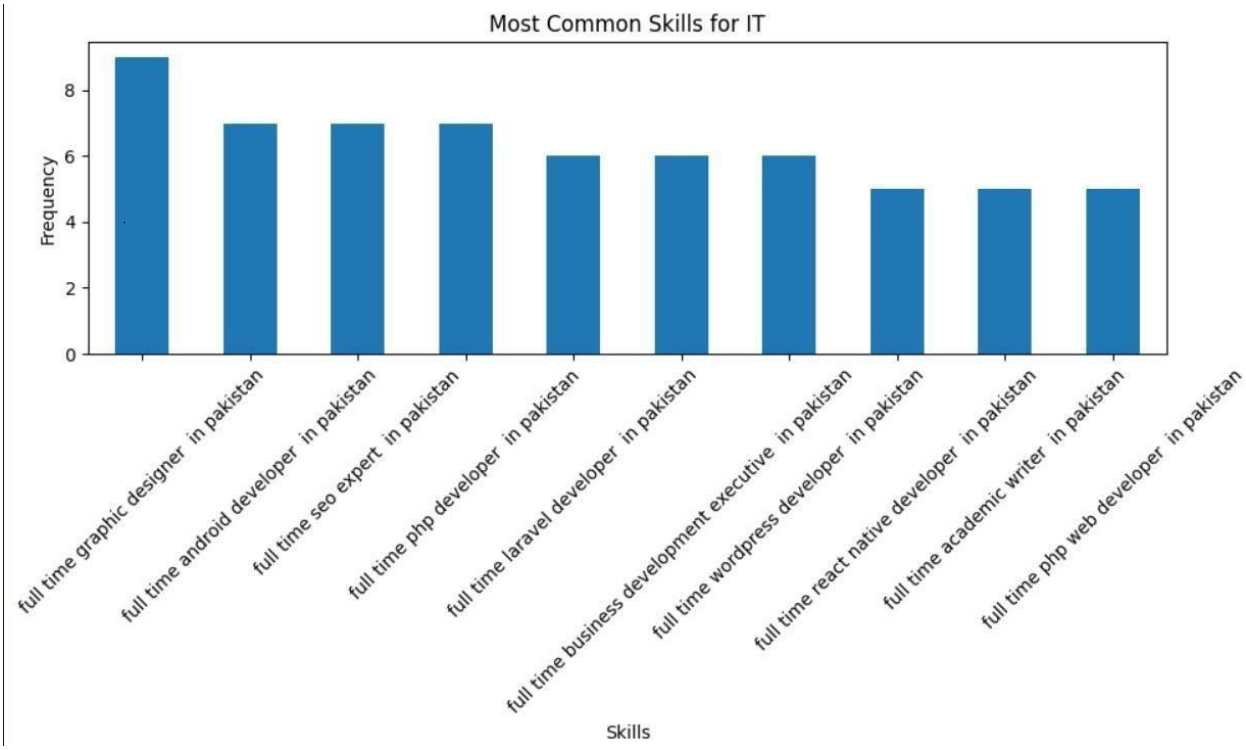
```
filtered_df = df[df['Department'].isin(['IT Jobs'])]
skill_counts = filtered_df['Skills'].explode().value_counts()
most_common_skills = skill_counts.head(10)
print("Most Common Skills for IT:")
print(most_common_skills)

plt.figure(figsize=(10, 6))
most_common_skills.plot(kind='bar')
plt.xlabel('Skills')
plt.ylabel('Frequency')
plt.title('Most Common Skills for IT')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```

OUTPUT:

```
Most Common Skills for IT:
Skills
full time graphic designer  in pakistan                 9
full time android developer  in pakistan                7
full time seo expert  in pakistan                       7
full time php developer  in pakistan                    7
full time laravel developer  in pakistan                6
full time business development executive  in pakistan   6
full time wordpress developer  in pakistan              6
full time react native developer  in pakistan           5
full time academic writer  in pakistan                  5
full time php web developer  in pakistan                5
Name: count, dtype: int64
```

ANALYSIS:

Our analysis of the IT job market revealed that the most common skill sought after by employers was that of a "Full-time Graphic Designer." This finding suggests a significant demand for graphic design expertise within the IT industry. Employers in the IT sector recognize the importance of visual communication and design in creating engaging user experiences and compelling digital products.



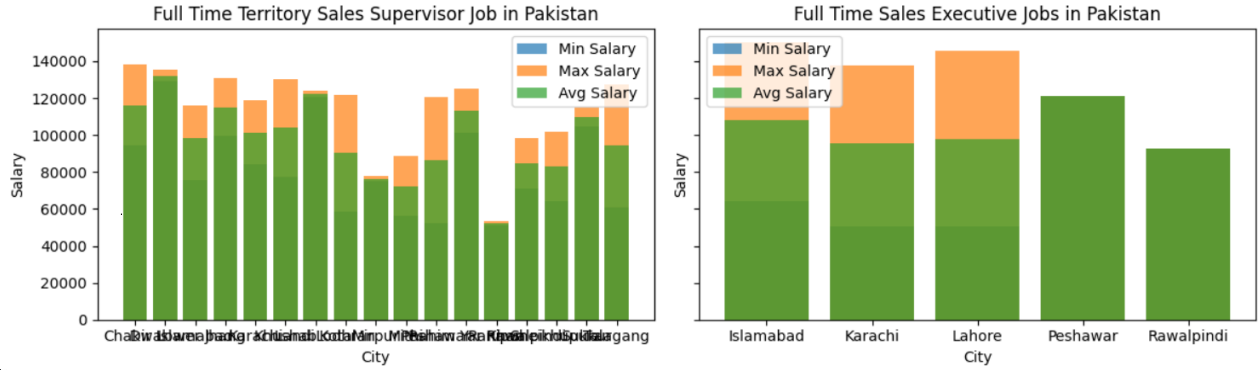
HIGHEST PAYING JOBS IN PAKISTAN:

```
import pandas as pd

df = pd.read_csv('starPakistan Available Job Dec 19 - Mar-21.csv')

required_columns = ['Job Name', 'City', 'Salary']

df_selected = df[required_columns]
df_selected.head()

# performing anaggregation function on the datase with respect to all the cities
and min, max and average
salary_ranges = df_selected.groupby(['Job Name', 'City']).agg(
    min_salary=('Salary', 'min'),
    max_salary=('Salary', 'max'),
    avg_salary=('Salary',  'mean')
)

top_20_records = salary_ranges.head(20)
print("Summary of Salary Ranges (Top 20 records):")
print(top_20_records.to_string())
```

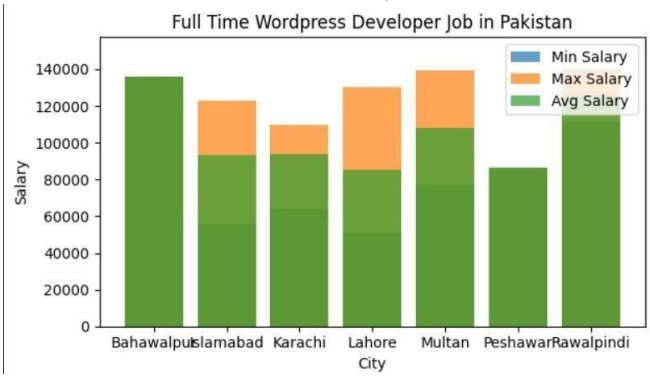
```
import matplotlib.pyplot as plt
# getting top 10 heighest paying jobs
top_10_jobs = df_selected['Job Name'].value_counts().nlargest(10).index
# generatig cub plots
fig, axes = plt.subplots(nrows=5, ncols=2, figsize=(12, 18), sharey=True)
# plotting for all jobs
for i, job_name in enumerate(top_10_jobs):
    ax = axes[i // 2, i % 2]
    job_data = salary_ranges.loc[job_name]
    cities = job_data.index
    min_salaries = job_data['min_salary']
    max_salaries = job_data['max_salary']
    avg_salaries = job_data['avg_salary']

    ax.bar(cities, min_salaries, label='Min Salary', alpha=0.7)
    ax.bar(cities, max_salaries, label='Max Salary', alpha=0.7)
    ax.bar(cities, avg_salaries, label='Avg Salary', alpha=0.7)

    ax.set_title(job_name)
    ax.set_xlabel('City')
    ax.set_ylabel('Salary')
    ax.legend()

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

OUTPUT:

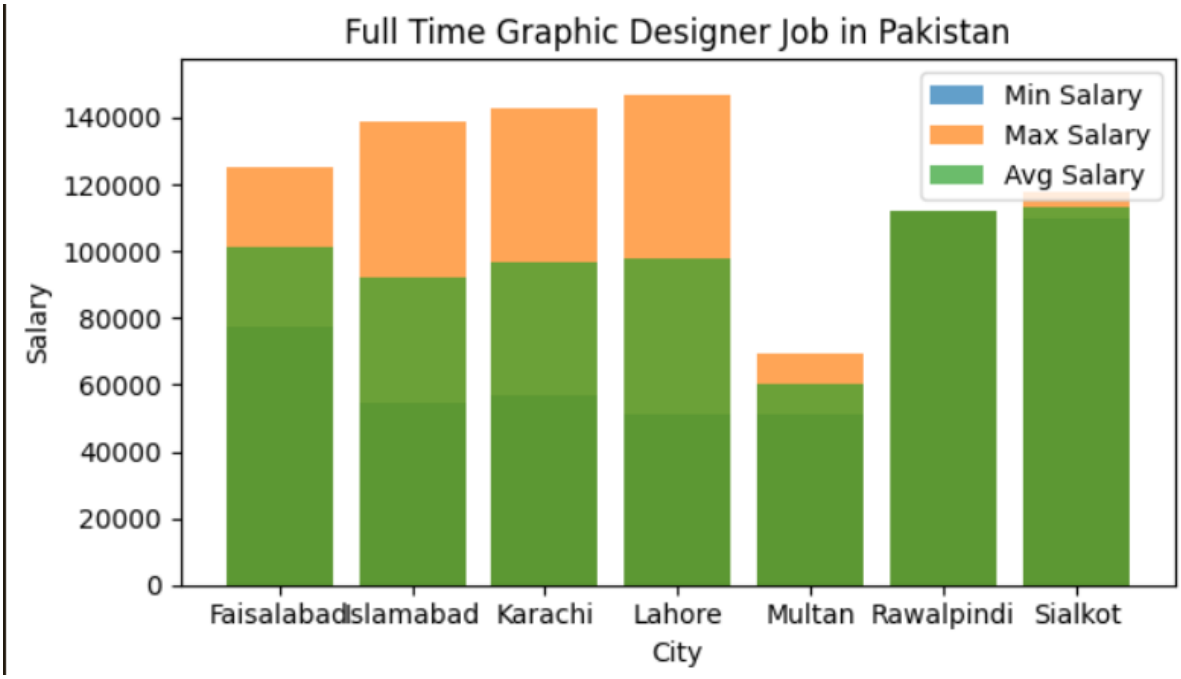
```
Summary of Salary Ranges (Top 20 records):
                                                                                                  min_salary  max_salary  avg_salary
Job Name                                                                                 City
Full Time  .NET Developer Or Application Developer Job in Pakistan                       Lahore        53181      128993     91087.0
Full Time  Adviser Financial Services / Assistant Branch Manager Jobs in Pakistan        Sialkot      140399      140399    140399.0
Full Time  Assistant Manager Job in Pakistan                                             Lahore        69641       69641     69641.0
Full Time  Associate Law / Legal Advisor / Lawyer  Job in Pakistan                       Lahore       136250      136250    136250.0
Full Time  Beauty Salon / Parlour Staff Jobs in Pakistan                                 Islamabad     73845       73845     73845.0
Full Time  Business Development Intern Job in Pakistan                                   Islamabad     92772       92772     92772.0
Full Time  Circle Technology Head - South Job in Pakistan                                Karachi      128113      134692    131402.5
Full Time  Customer Services Executive Jobs in Pakistan                                  Lahore        67277       86891     77084.0
Full Time  DAE + RAE Refrigeration And Air Conditioning Technician Required Jobs in Pakistan Islamabad  70011      114239     92125.0
Full Time  Data Entry Operator To Perform Daily Tasks. Job in Pakistan                   Karachi      118098      118098    118098.0
Full Time  Developer (.NET Core) Jobs in Pakistan                                        Islamabad     53593      138067     95830.0
Full Time  Fashion Designer Required For Brand 6FGR Clothing Line  Job in Pakistan       Islamabad    104243      104243    104243.0
Full Time  Finance Executive Jobs in Pakistan                                            Islamabad     76030       76030     76030.0
Full Time  Ionic Developer For Junior And Intern Positions Jobs in Pakistan              Lahore        63866       63866     63866.0
Full Time  Marketing Manager In Fabric  Job in Pakistan                                  Lahore        60698       60698     60698.0
Full Time  Naan / Paratha Maker Job in Pakistan                                          Islamabad     89845      102305     96075.0
Full Time  SEO Expert  Job in Pakistan                                                   Lahore       105438      105438    105438.0
Full Time  Sales & Marketing Executive Job in Pakistan                                   Quetta        54703       54703     54703.0
Full Time  Sales Executive Jobs in Pakistan                                              Islamabad     52752       81249     67000.5
Full Time  Search & Data Entry Officer Job in Pakistan                                   Lahore        97854       97854     97854.0
```

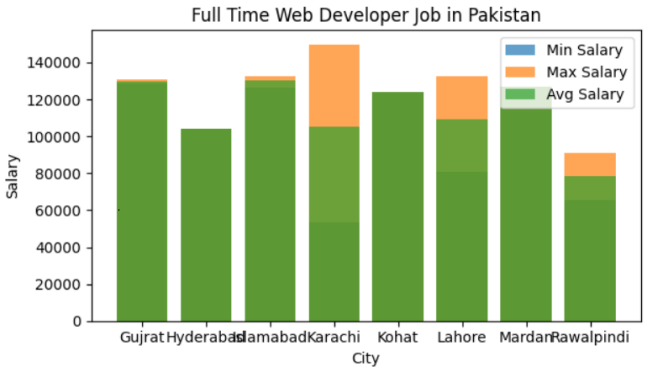

Full Time Territory Sales Supervisor Job in Pakistan
Min Salary
Max Salary
Avg Salary
Salary
City

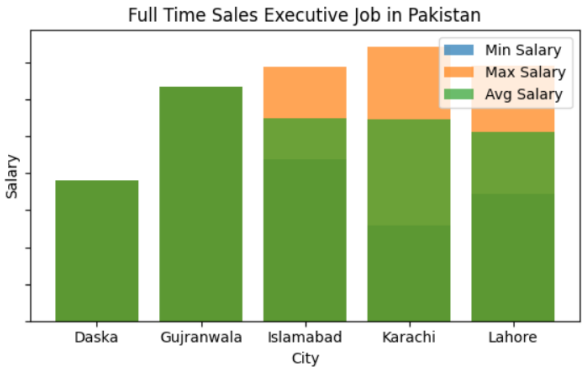
Full Time Sales Executive Jobs in Pakistan
Min Salary
Max Salary
Avg Salary
Salary
Islamabad
Karachi
Lahore
Peshawar
Rawalpindi
City

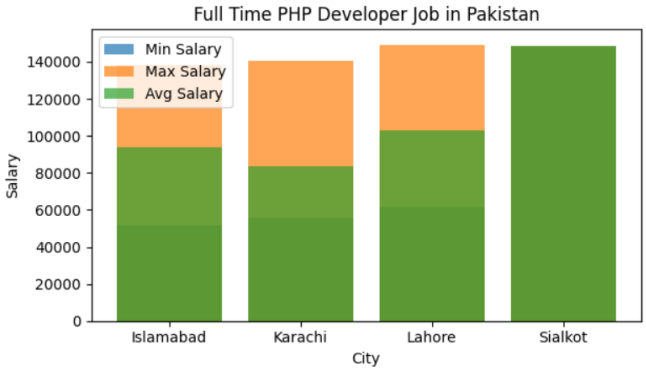

Full Time Wordpress Developer Job in Pakistan
Min Salary
Max Salary
Avg Salary
Salary
Bahawalpur
Islamabad
Karachi
Lahore
Multan
Peshawar
Rawalpindi
City

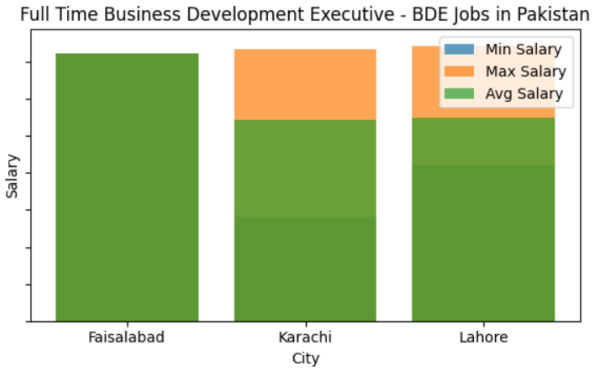

Full Time Content Writer Job in Pakistan
Min Salary
Max Salary
Avg Salary
Salary
Islamabad
Karachi
Lahore
Rawalpindi
City


Full Time Graphic Designer Job in Pakistan
Min Salary
Max Salary
Avg Salary
Salary
Faisalabad
Islamabad
Karachi
Lahore
Multan
Rawalpindi
Sialkot
City









ANALYSIS:

Our analysis of job salaries revealed that the highest paying job among all the positions considered was that of an "Advisor Financial Services." This finding highlights the potential for lucrative financial opportunities in the field of financial services. Employers in this sector recognize the value of experienced professionals who can provide expert financial advice to clients.

SALARY PREDICTION:

```
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.preprocessing import LabelEncoder

data = pd.read_csv('starPakistan Available Job Dec 19 - Mar-21.csv')
features = ['Job Name', 'Job Type', 'Experience Required', 'Department', 'City']
target = 'Salary'

le = LabelEncoder()
data_encoded = data[features].apply(le.fit_transform)

X = data_encoded
y = data[target]

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

```
model = LinearRegression()
model.fit(X_train, y_train)

job_name = input("Enter Job Name: ")
job_type = input("Enter Job Type: ")
experience = input("Enter Experience: ")
department = input("Enter Department: ")
city = input("Enter City: ")

user_input = pd.DataFrame([[job_name, job_type, experience, department, city]], columns=features)

for feature in features:
    le = LabelEncoder()
    le.fit(data[feature])
    user_input[feature] = user_input[feature].map(lambda s: '<unknown>' if s not in le.classes_ else s)
    le.classes_ = np.append(le.classes_, '<unknown>')

user_input_encoded = user_input.apply(le.transform)
prediction = model.predict(user_input_encoded)
print("Predicted Salary according to market condition is:", int(prediction), "PKR")
```

OUTPUT:

```
Enter Job Name:  PHP developer
Enter Job Type:  Full time
Enter Experience:  2
Enter Department:  IT
Enter City:  Karachi
Predicted Salary accoding to market condition is : 106061 PKR
```

Chapter 6
Conclusion and Future Work

In this, we will discuss about the Future Work of this system along with Executive Summary.

6.1 Future Work:

The idea we had in mind was to extend our analysis by visualizing the current year's job market trends for various skills, following a similar approach to how we examined job demand from previous year. we hypothesized that the most skilled demand was for a particular job, with 1000 job openings in that field, but only 200 positions were filled. Based on this information, we further speculated that the remaining 800 unfilled jobs from the previous year would likely be available in the current year. This supposition applies not only to that specific job but also to other skills in the dataset.

However, due to the unavailability of the dataset, we were unable to conduct this analysis. Nonetheless, this idea presents an intriguing avenue for future research. Obtaining the dataset would enable us to analyze the job market trends for various skills and identify which jobs from the previous year remained unfilled, thereby providing insights into potential job opportunities in the current year. This future work would contribute to a comprehensive understanding of the evolving job market dynamics and aid job seekers and employers in making informed decisions.

6.2 Conclusion

In conclusion, our analysis of the job market trends in Pakistan using a dataset of job openings from various industries has provided valuable insights for job seekers and employers. Through data cleaning, visualization, and statistical analysis, we have gained a deeper understanding of the employment landscape in the country. While limitations, such as the unavailability of the current year's dataset, hindered the full scope of our analysis, our project serves as a foundation for future research in this area.

By uncovering patterns and trends within the job market, our analysis offers guidance for job seekers to enhance their employability and make informed career decisions. Additionally, employers can benefit from the insights provided to identify emerging sectors and skills in high demand. Moving forward, it is crucial to continue collecting and analyzing up-to-date data, expanding the analysis to cover multiple years, and exploring advanced techniques such as machine learning to further enhance our understanding of the job market trends in Pakistan.