INTERNSHIP REPORT

ON

JOB PORTAL DATA ANALYSIS

AT

NULLCLASS ED TECH PVT LTD

BY DEEPALI B A

E:deepali02300@gmail.com

M:9741349017

TABLE OF CONTENTS

FIGURE TABLE

01.INTRODUCTION

This report consists about my internship experience of Job Portal Data Analysis carried out at Nullclass Edu Pvt Ltd Company, who are involved in providing hands on experience of real world projects and acquire the skills by working with this projects**.**

02.BACKGROUND

The internship was centered on analyzing a job listings dataset from Kaggle to understand employment trends, in-demand skills, and job market dynamics across industries. As a data analyst intern, I was tasked with transforming raw data into actionable insights through cleaning, exploration, and visualization using notebook and Tableau.

03.LEARNING OBJECTIVES

- To apply data cleaning and preprocessing techniques on real-world datasets.
- To develop visualizations and dashboards using tools like Python and Tableau.
- To understand and carry out Data Analysis on datasets.

04.ACTIVITIES AND TASKS

In this Internship Data Analysis with Job Analytics Portal using Tableau was a course of making graphs using provided filters**.**

4.1.Training Program

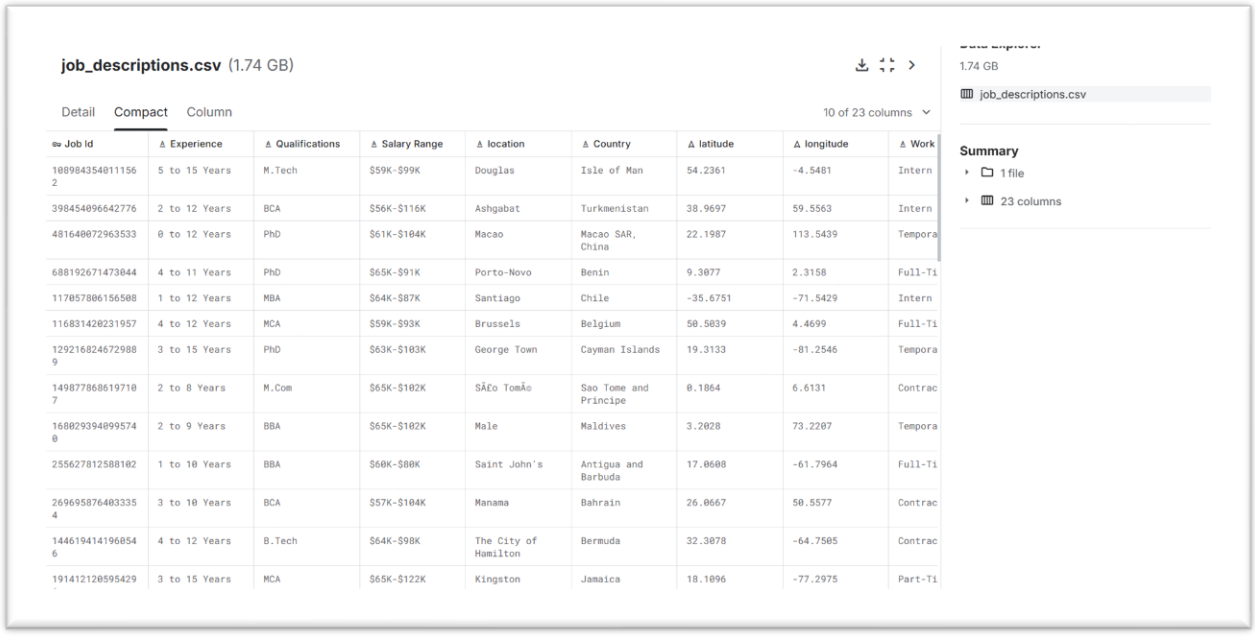
This internship is a platform where I was assigned with the specific task. In the initial days of the internship, I was trained on the following:

➢ Python Programming (Pandas and NumPy libraries)

➢ Tableau

4.2.DATA SET

This section describes, in brief, the data that has been used for the research. Data from 'Job Dataset' was used in this project, the data was extracted from public website Kaggle (Kaggle.com),this dataset offers a diverse set of job listings across various industries and job types. Below table shows the different types of columns present in the data-set.



job_descriptions.csv (1.74 GB)

Detail | Compact | Column — 10 of 23 columns

Job Id	Experience	Qualifications	Salary Range	location	Country	latitude	longitude	Work
1089843540111562	5 to 15 Years	M.Tech	$59K-$99K	Douglas	Isle of Man	54.2361	-4.5481	Intern
398454096642776	2 to 12 Years	BCA	$56K-$116K	Ashgabat	Turkmenistan	38.9697	59.5563	Intern
481640072963533	0 to 12 Years	PhD	$61K-$104K	Macao	Macao SAR, China	22.1987	113.5439	Tempora
688192671473044	4 to 11 Years	PhD	$65K-$91K	Porto-Novo	Benin	9.3077	2.3158	Full-Ti
1170578061565088	1 to 12 Years	MBA	$64K-$87K	Santiago	Chile	-35.6751	-71.5429	Intern
1168314202319577	4 to 12 Years	MCA	$59K-$93K	Brussels	Belgium	50.5039	4.4699	Full-Ti
1292168246729889	3 to 15 Years	PhD	$63K-$103K	George Town	Cayman Islands	19.3133	-81.2546	Tempora
1498778686197107	2 to 8 Years	M.Com	$65K-$102K	SÃ£o TomÃ©	Sao Tome and Principe	0.1864	6.6131	Contrac
1680293940995740	2 to 9 Years	BBA	$65K-$102K	Male	Maldives	3.2028	73.2207	Tempora
2556278125881020	1 to 10 Years	BBA	$60K-$80K	Saint John's	Antigua and Barbuda	17.0608	-61.7964	Full-Ti
2696958764033354	3 to 10 Years	BCA	$57K-$104K	Manama	Bahrain	26.0667	50.5577	Contrac
1446194141960546	4 to 12 Years	B.Tech	$64K-$98K	The City of Hamilton	Bermuda	32.3078	-64.7505	Contrac
191412120595429	3 to 15 Years	MCA	$65K-$122K	Kingston	Jamaica	18.1096	-77.2975	Part-Ti

1.74 GB
job_descriptions.csv

Summary
1 file
23 columns

Source: https://www.kaggle.com/datasets/ravindrasinghrana/job-description-dataset

Fig:4.1

4.3.DATASET TRANSFORMATION AND EXTRACTION.

- The data reading was first done by importing Pandas and NumPy libraries in Kaggle notebook.

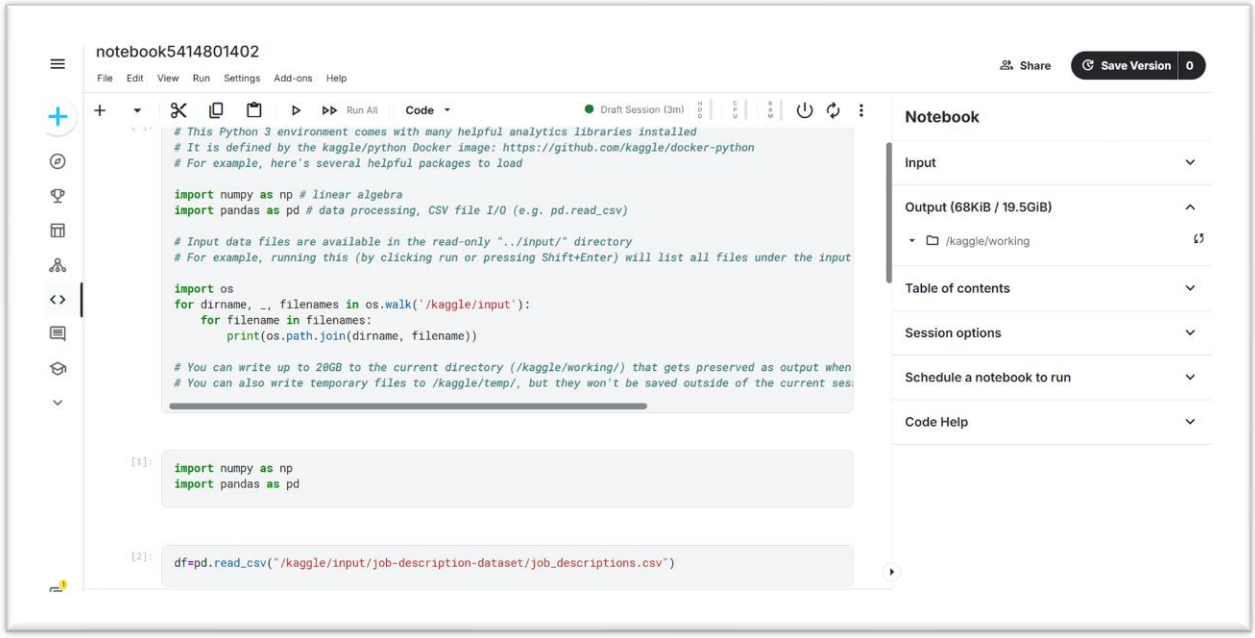


Fig:4.2

- It was followed by renaming the columns and data cleaning,by identifying null and duplicate values and removing it.

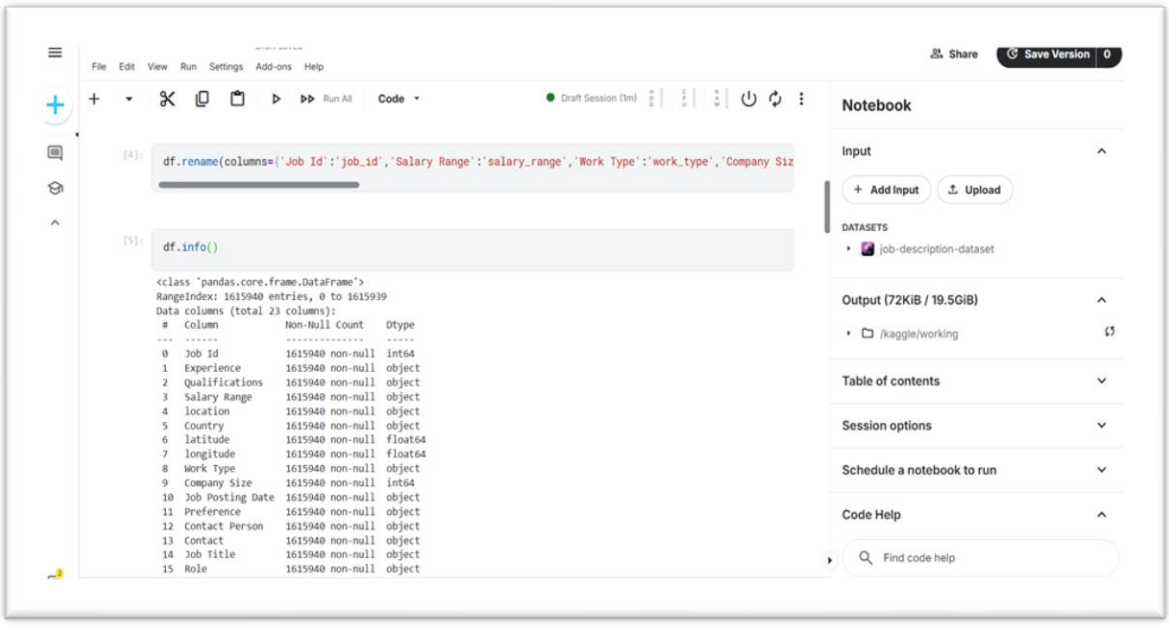


Fig:4.3

- The cleaned data was collected from Kaggle in .csv format.

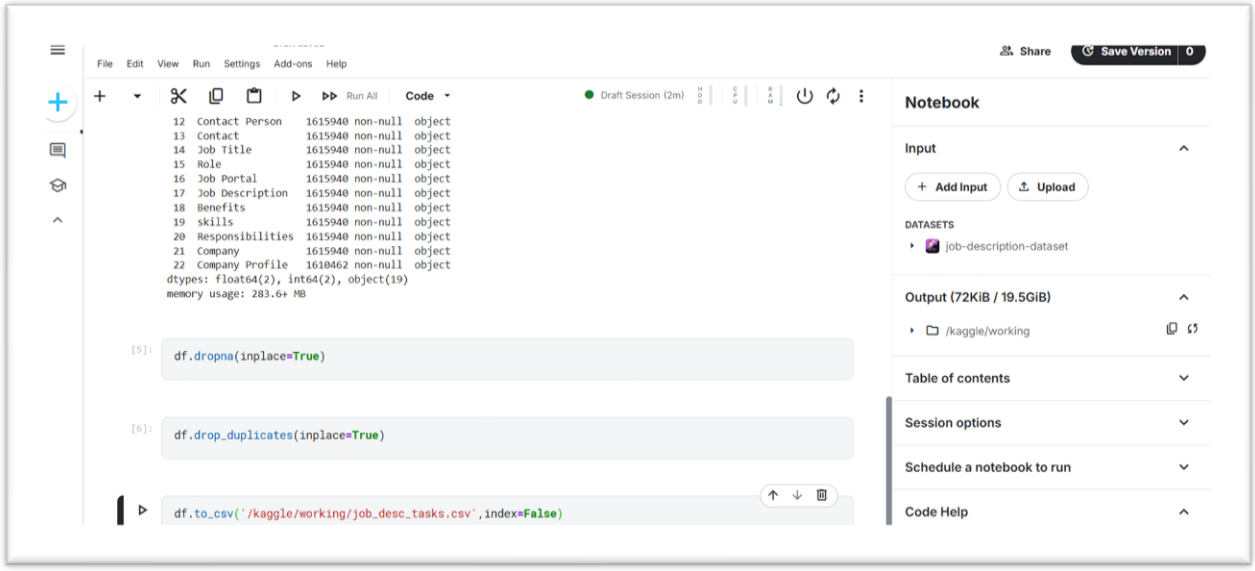


Fig:4.4

4.4.LOADING OF DATA INTO TABLEAU

The downloaded data was loaded into Tableau and two dashboards were created ,one related to Training tasks and other for Internship tasks.

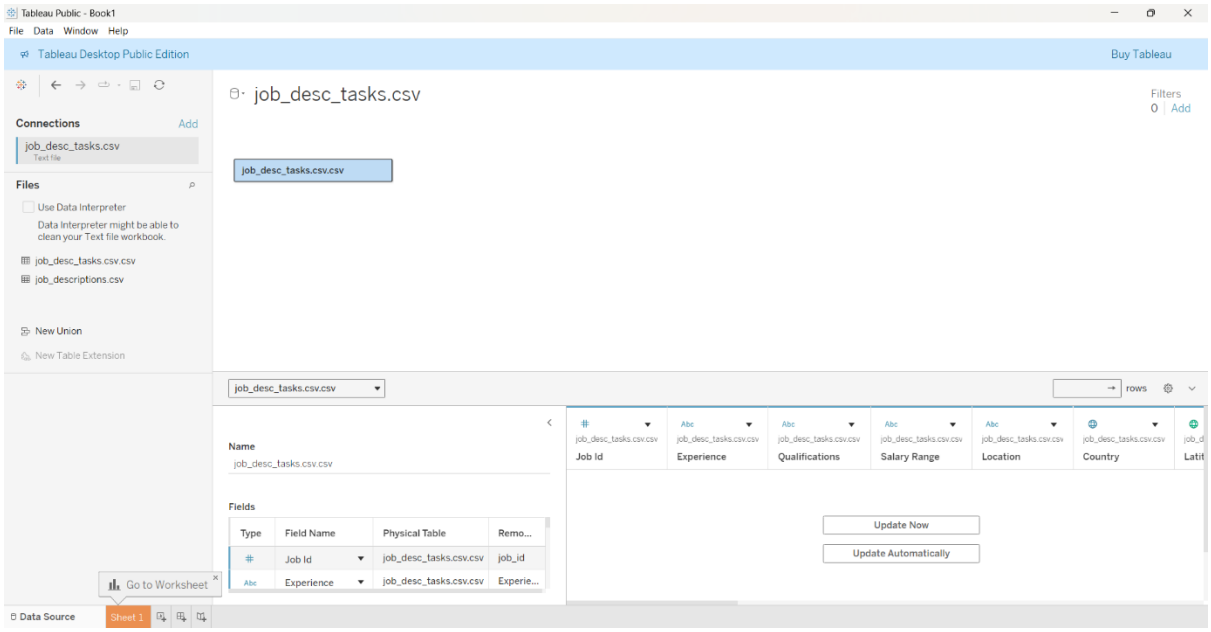


Fig:4.5

4.4.1.Training tasks:

The training tasks consisted of performing eight different graphs in each worksheets.

Sheet1:Company graph

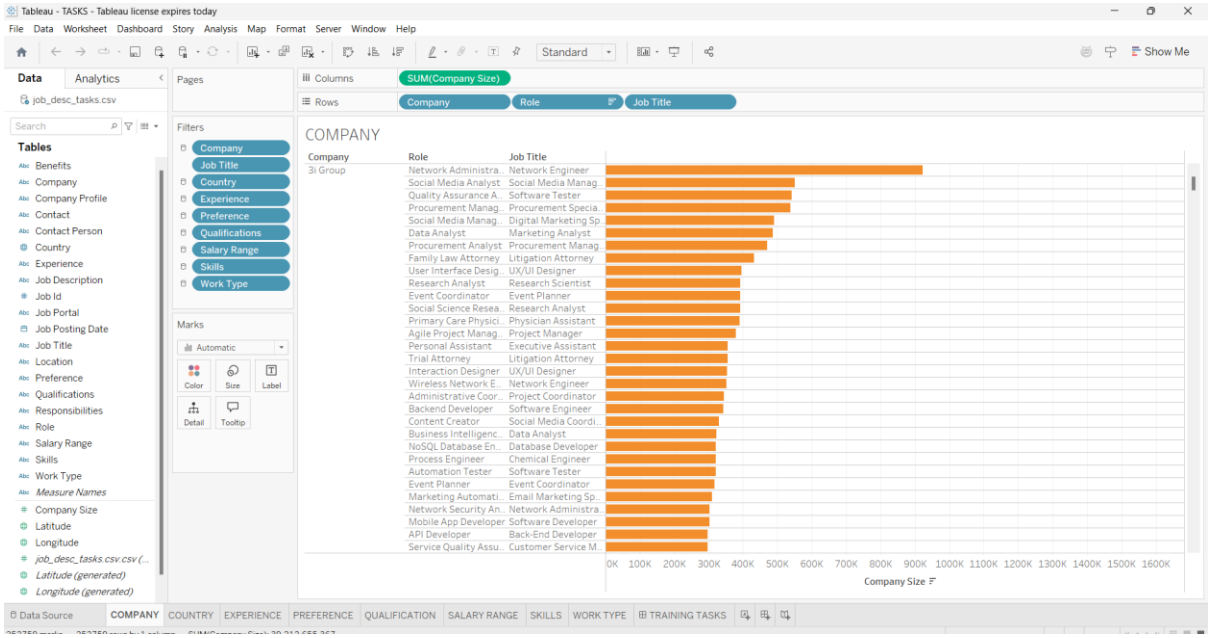


Fig:4.6

Sheet2:Country graph

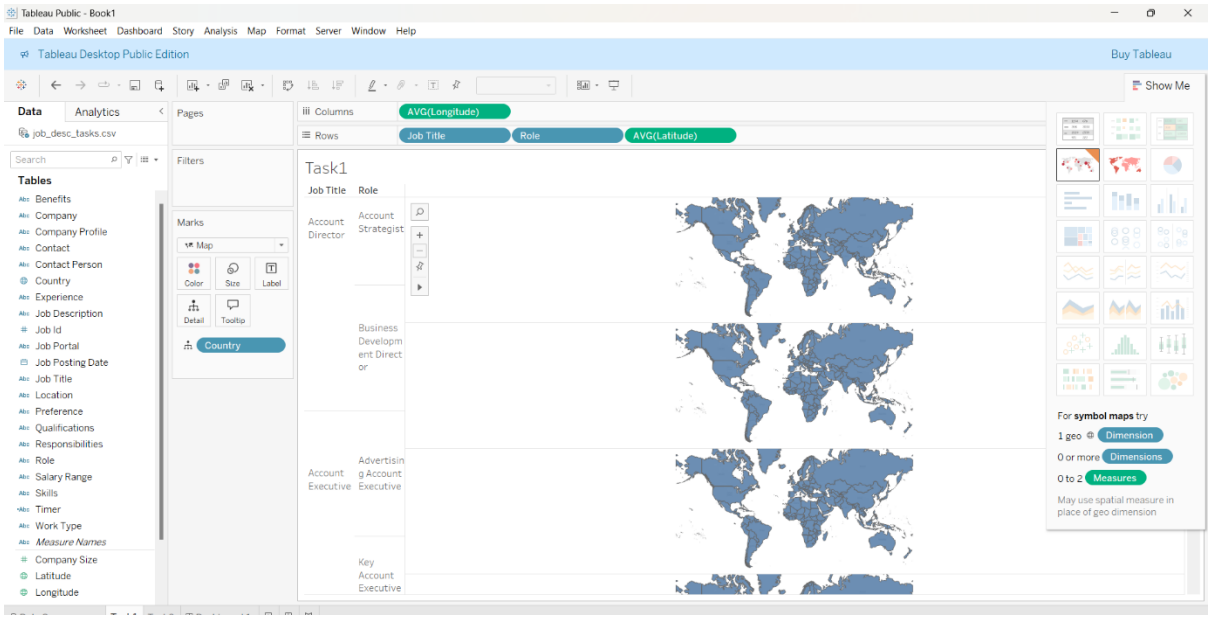


Fig:4.7

Sheet3:Experience graph

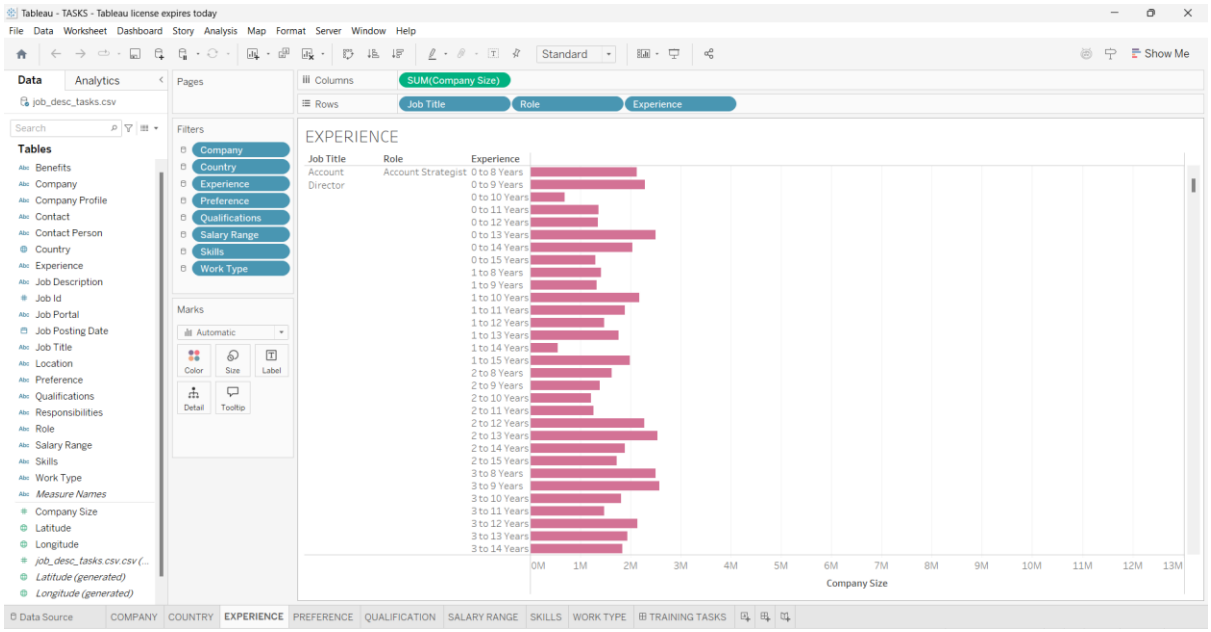


Fig:4.8

Sheet4:Preferance graph

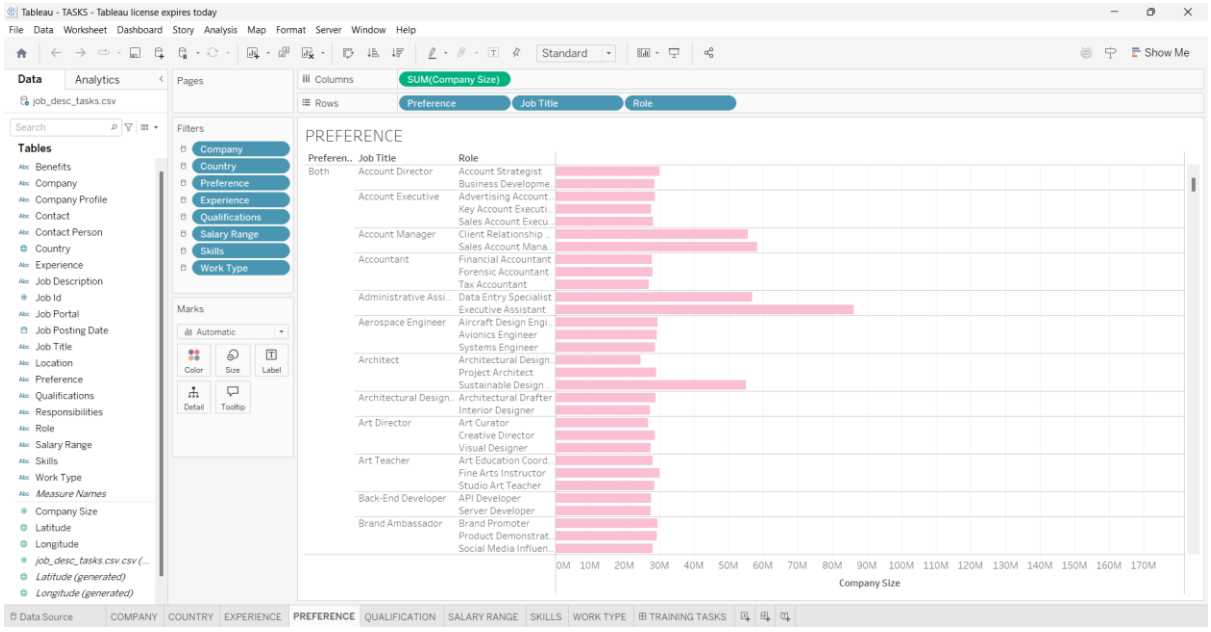


Fig:4.9

Sheet5:Qualification graph

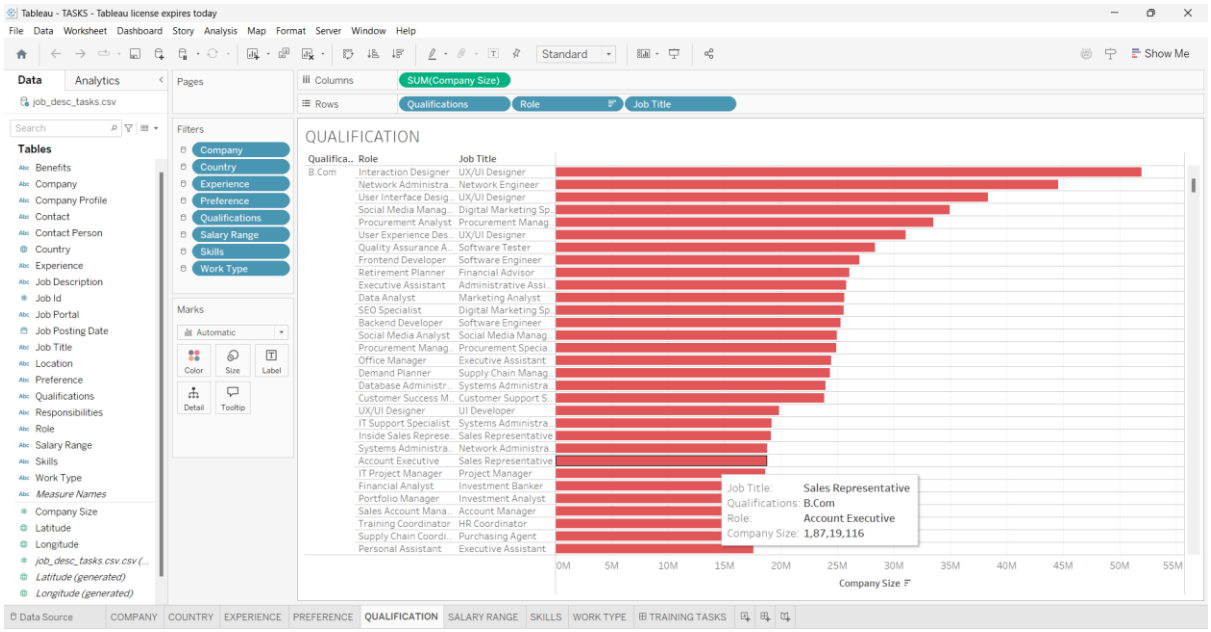


Fig:4.10

Sheet6:Salary Range graph

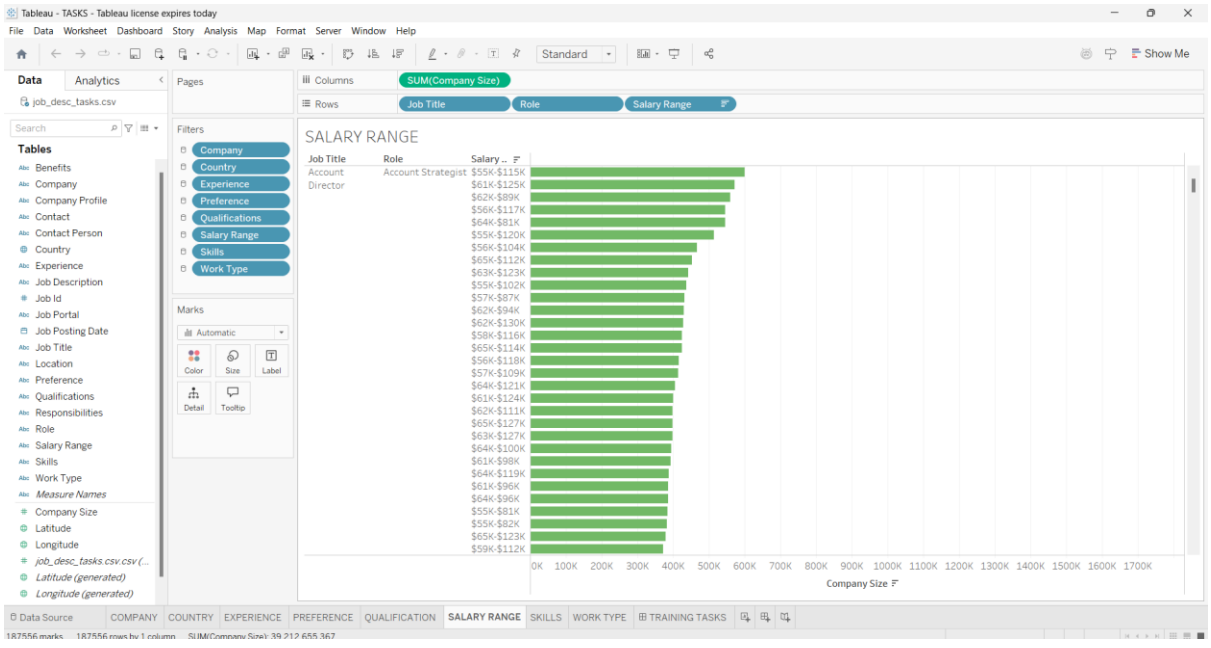


Fig:4.11

Sheet7:Skills graph

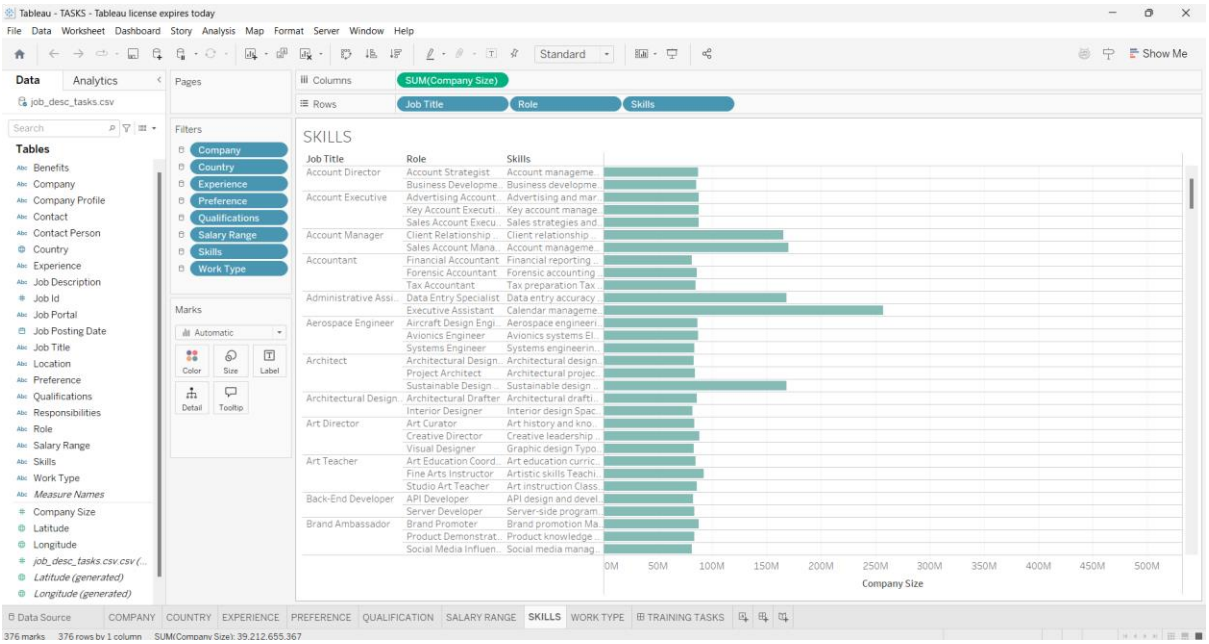


Fig:4.12

Sheet8:Work Type graph

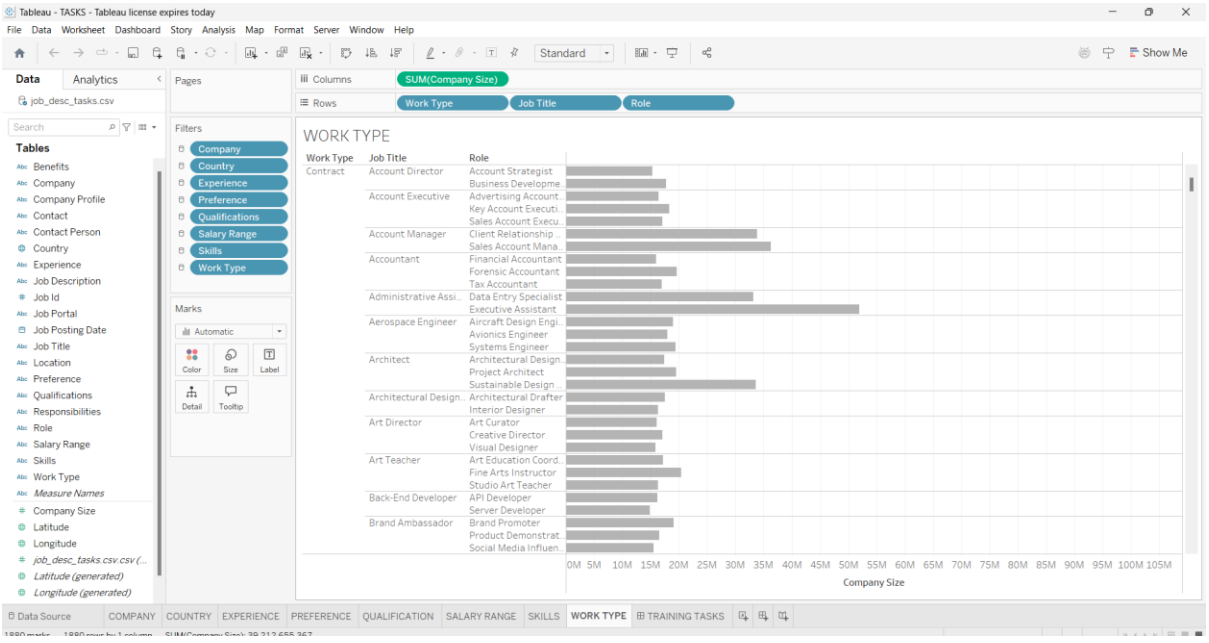
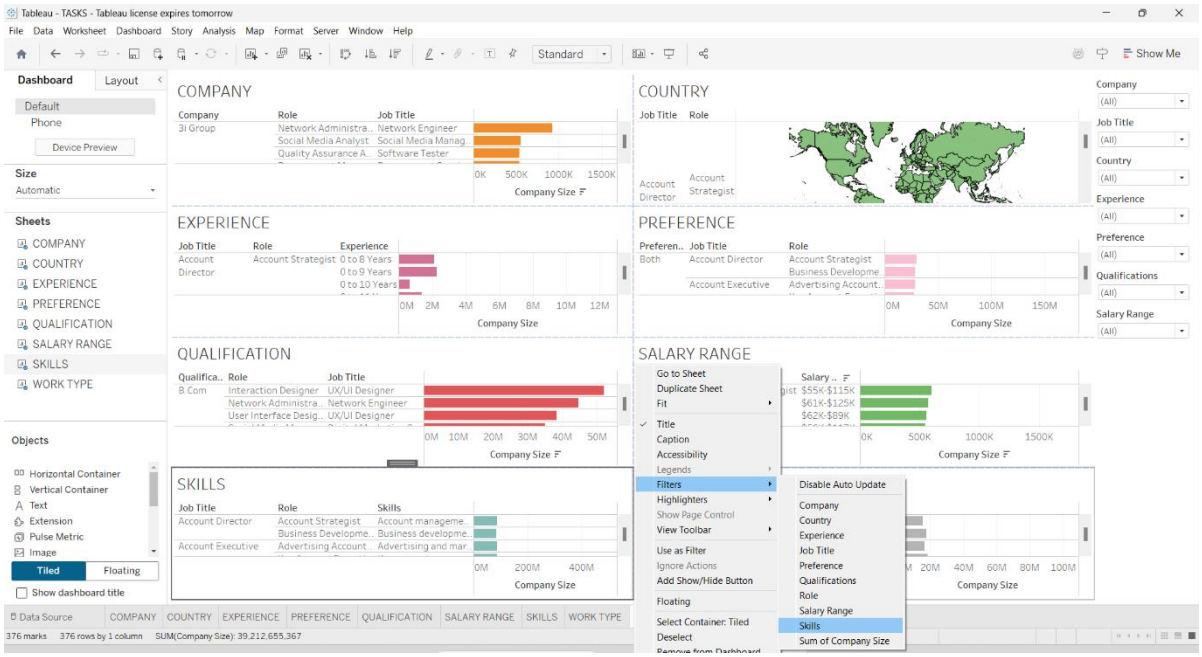
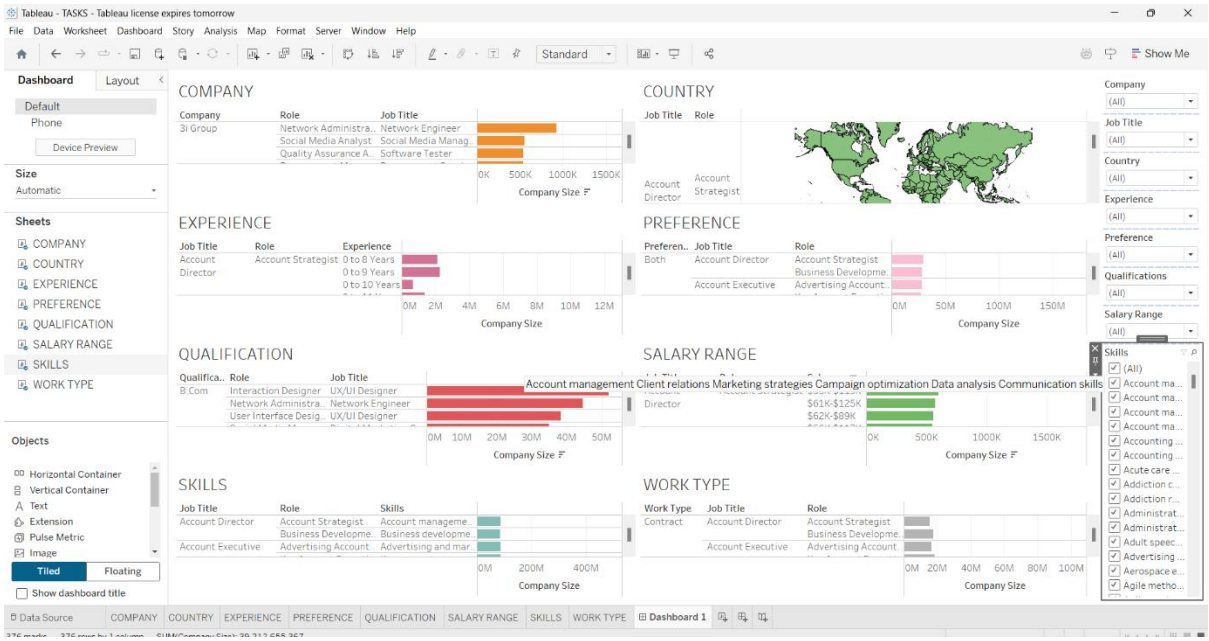


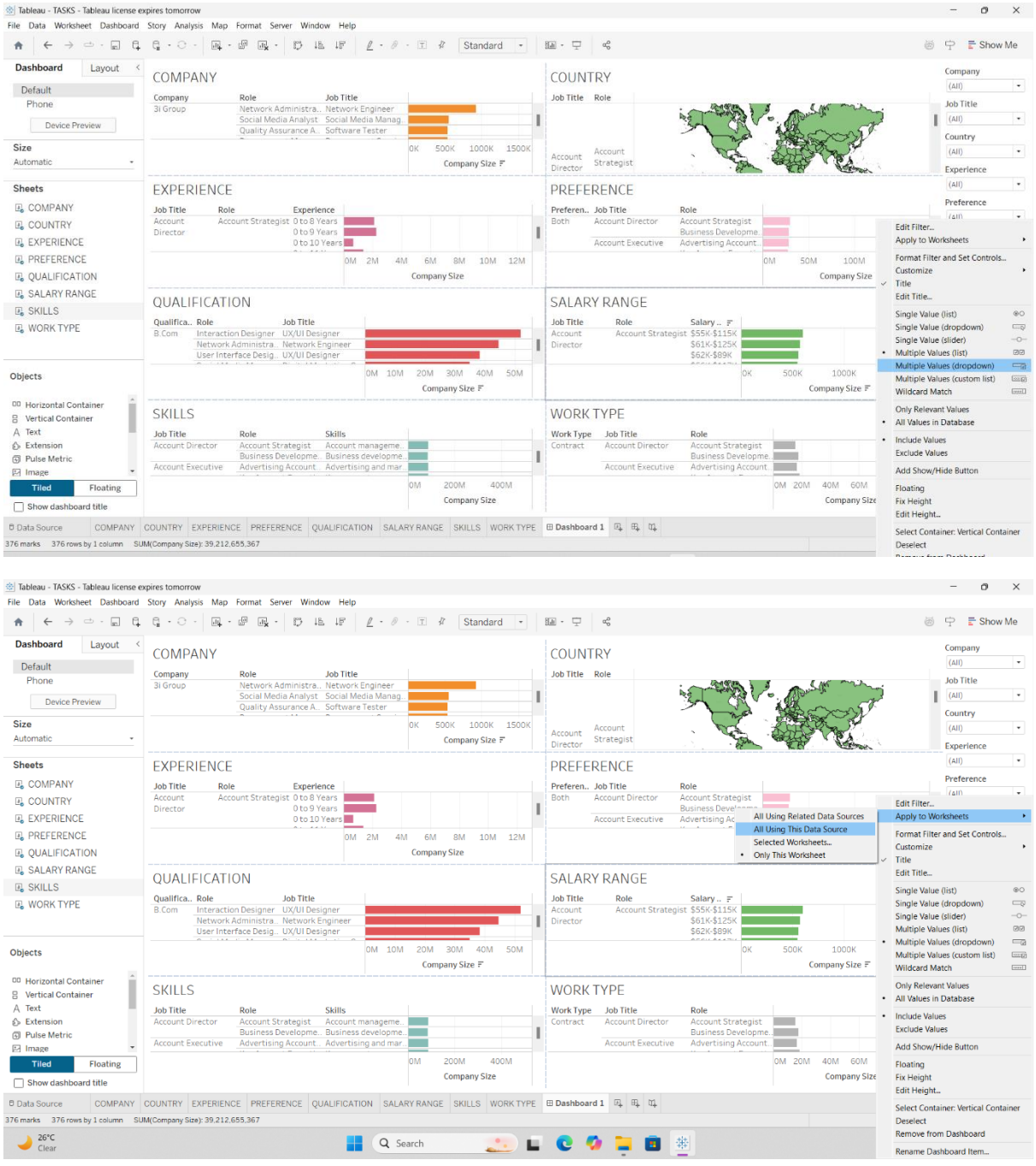
Fig:4.13

Later combining all this sheets ,a dashboard was created.

Filters were applied to each graphs in the dashboard.







Filtering of data in dashboard

Fig:4.14

4.4.2.Internship tasks:

TASK01:

Draw a chart to show relation between country, job title and role.

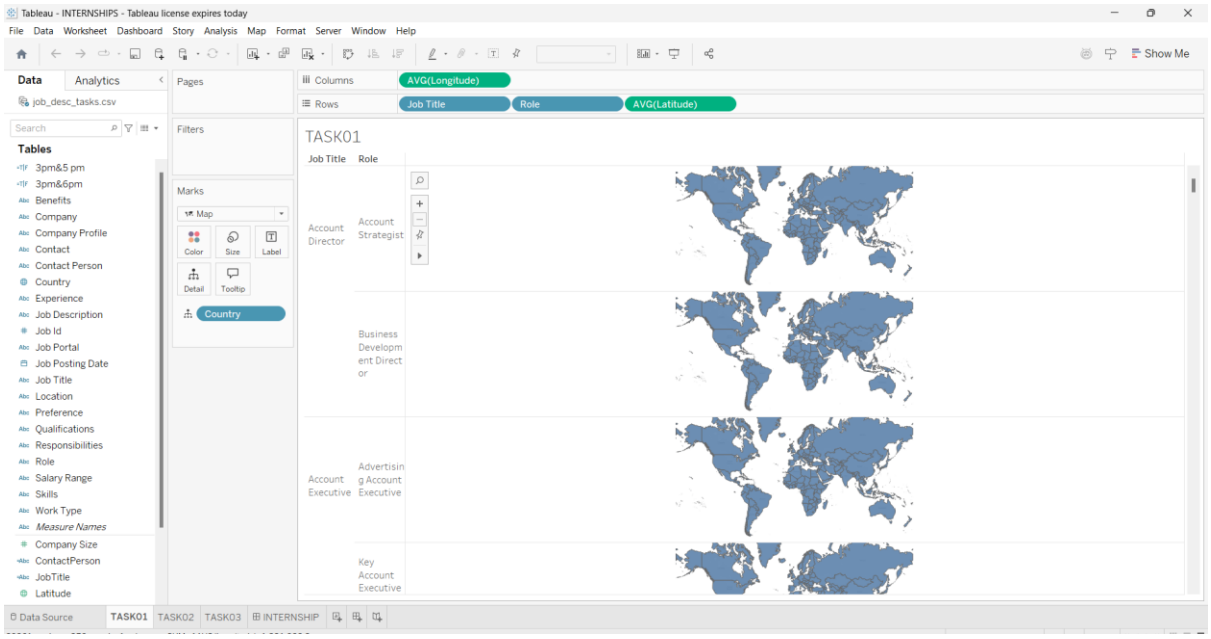


Fig:4.15

- In this graph,Job Title,Role and Country dimensions were placed in rows.
- Latitude from measures was placed in columns.
- The details are shown on country.

TASK02:

Draw a chart where qualification=’B.tech,M.tech,PhD’ and work type=’Full time’. The country should be African continent and ignore other countries and this graph should work only between 3 PM IST to 6 PM IST apart from that time we should not show this graph in dashboard itself. Job title should be starting with letter ‘D’ and preference should be a Male.The company size should more than 80000. The contact person should be starting with letter ‘A’ and job portal should be indeed . We need to place a latitude and longitude and if we click on latitude and longitude we will open a map and show the exact location.

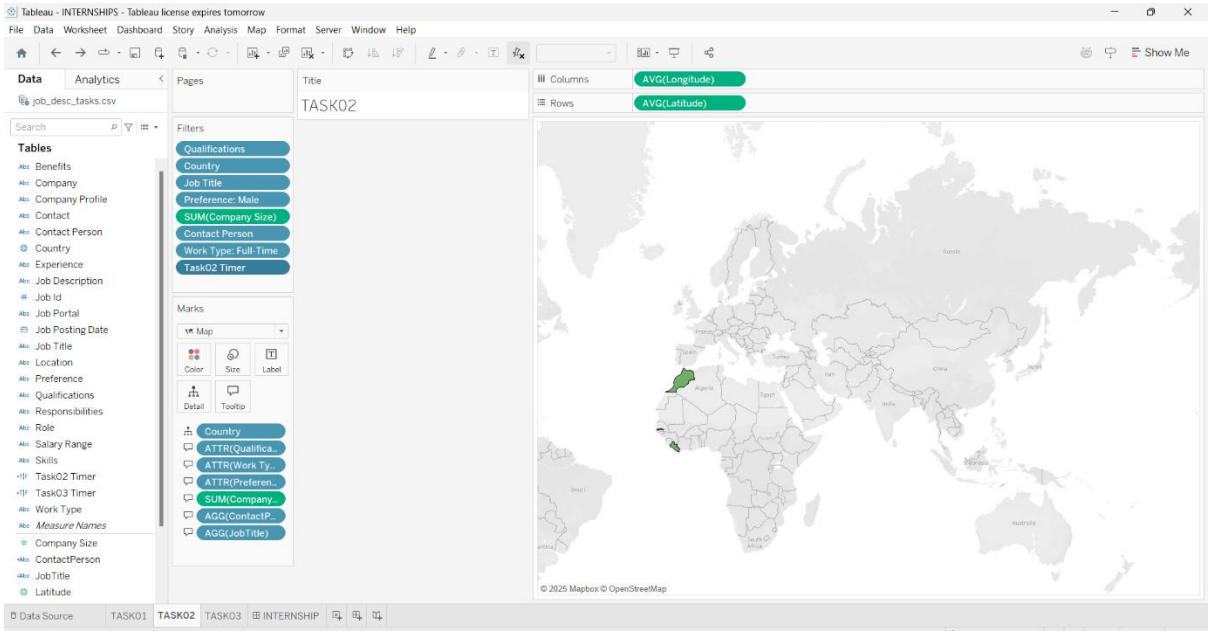


Fig:4.16

- The graph is based on longitude and latitude placed on columns and rows respectively.
- The filtered data are on displayed on graph by placing the columns from ‘Data Pane’ and placing it on ‘Tooltip’ from ‘Marks’ section.
- The data is filtered on
  (1) The Qualifications filter keep ’B.Tech','M.Tech’ and ‘PhD’
  (2) The Job Title filter keeps 8 of 147 members which starts with letter ‘D’ and is done using wildcard.

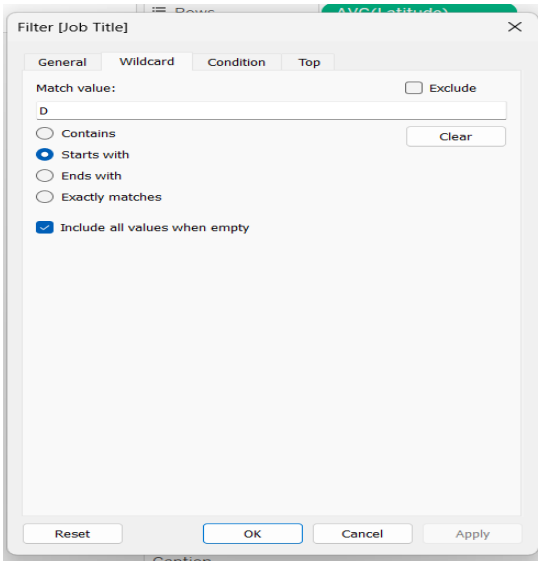
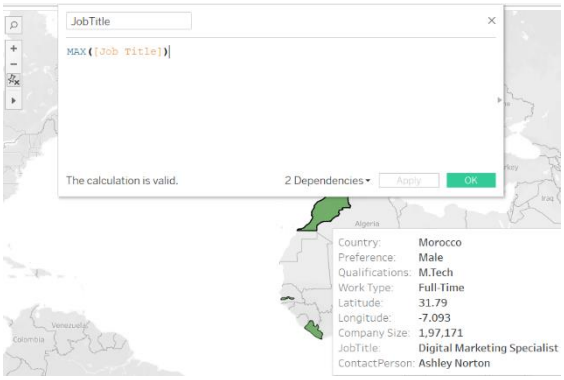


Fig:4.17

(3) The Preference filter keeps ‘Male’.

(4) The Contact Person filter keeps 19,540 of 2,13,692 members whose name start with ‘A’ and is done using wildcard.To return respective value of 19,540 for each filter aggregate function was used.



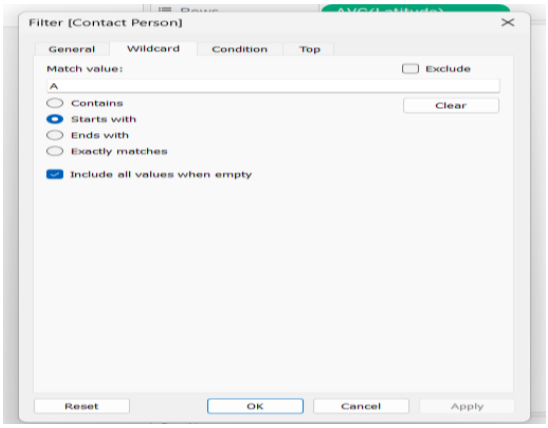


Fig:4.18

(5) The Work Type filter keeps ‘Full-Time’.

(6) The Country filter keeps 52 of 216 members which belong to African continent.

(7) The Company Size filter includes values greater than or equal to 80,001.

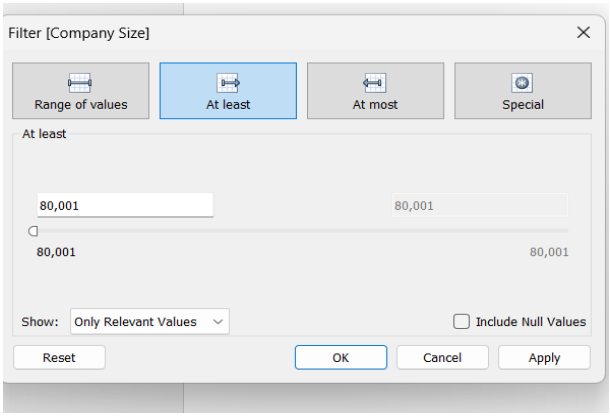


Fig:4.19

(8)**THE JOB PORTAL FILTER KEEPS 'INDEED' WHICH DOESN'T RETURN ANY DATA FOR THE GRAPH.**

(9)To show exact location on map Actions from Worksheet was selected and URL given.

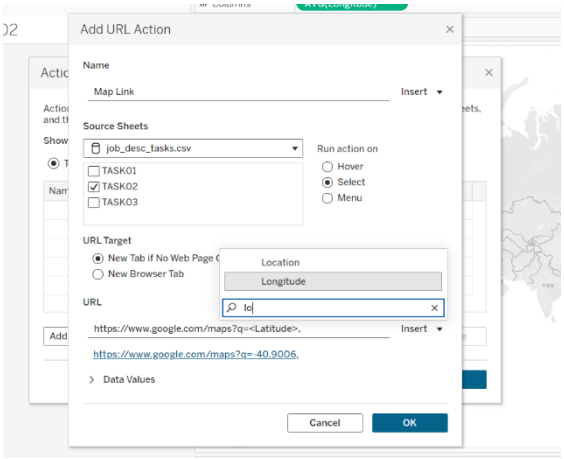


Fig:4.20

(10)To make the graph work between 3PM and 6PM ,Create Calculated Field was used named as '3PM&6PM' as shown below.This returns Boolean value based on the written condition.Here '+(5.5/24)' is done because Tableau works at UTC timings.

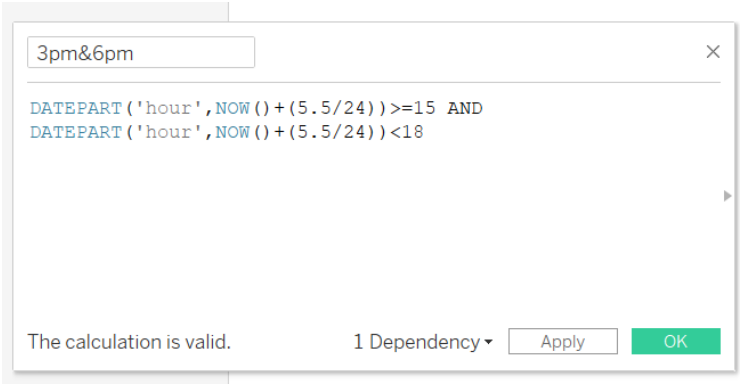


Fig:4.21

This filter was not applied on this worksheet as the tableau server works dynamically when published.So this filter was applied to this graph in dashborad using Containers and Layout option.

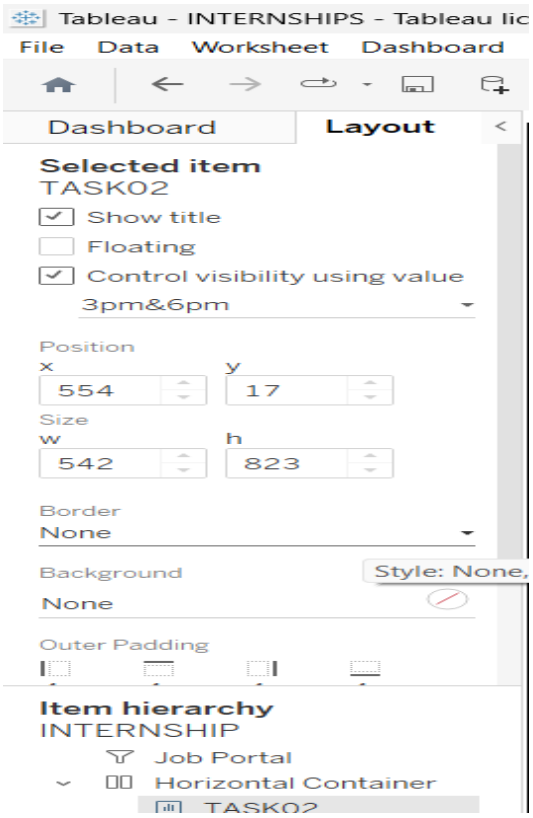


Fig:4.22

TASK03:

Draw a chart between company size and company name where company size<50000 and job title should be mechanical engineer and experience should be more than 5 years and country should be Asian as well as salary should be more than $50k and the work type should be both part time and full time and ignore all other work types and the preference should be male .we should filter candidates those who are applied on idealist and this graph should work only between 3PM IST to 5 PM IST apart from that time we should not show this graph in dashboard itself.

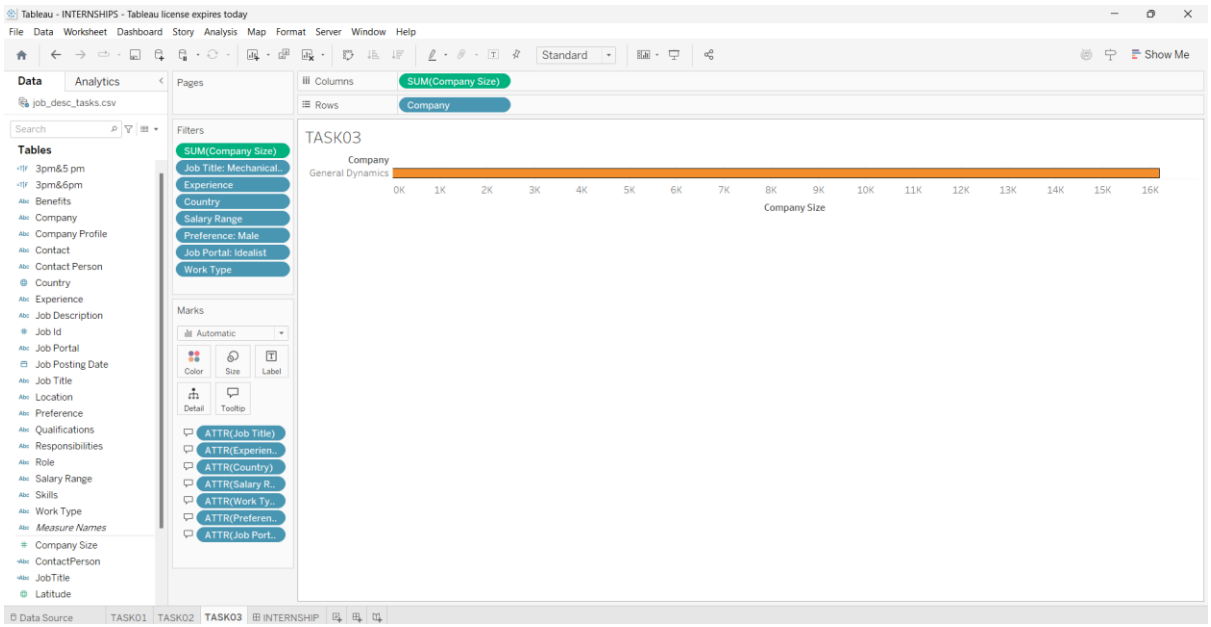


Fig:4.23

- The graph is based on aggregate of Company Size and Company placed on columns and rows respectively.
- The data is filtered on

  i) The Company Size is sum of each company size,which includes values lesser than 50,000.

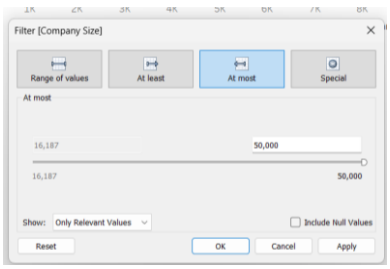


Fig:4.24

ii) The Job Title filter keeps ‘Mechanical Engineer’ which is 1 of 147 values.

iii) The Experience filter keeps 8 of 48 members.

iv) The Country filter keeps 46 of 216 which are part of Asian continent.

v) The Salary Range filter keeps 561 values whose salary is more than $50K.

vi) The Job Portal filter keeps ‘Idealist’.

vii) The Preference filter keeps ‘Male’.

viii) The Work Type filter keeps ‘Full-Time’ and ’Part-Time’.

ix) To make the graph work between 3PM and 5PM ,Create Calculated Field was used named as ‘3PM&5PM’ as shown below.This returns Boolean value based on the written condition.Here ‘+(5.5/24)’ is done because Tableau works at UTC timings.

Fig:4.25

This filter was not applied on this worksheet as the tableau server works dynamically when published.So this filter was applied to this graph in dashborad using Containers and Layout option.

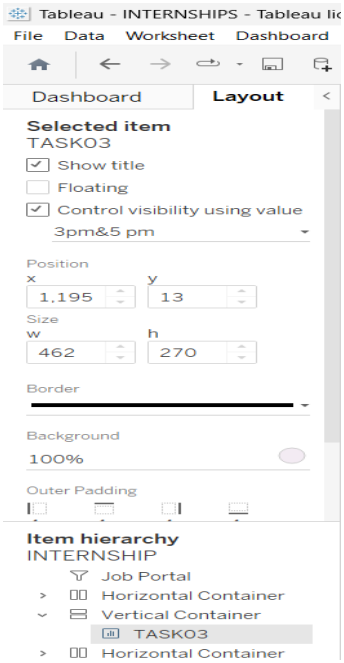


Fig:4.26

Later combining all this sheets in three different containers, a dashboard was created.And the graphs were modified as stated above with help of containers.

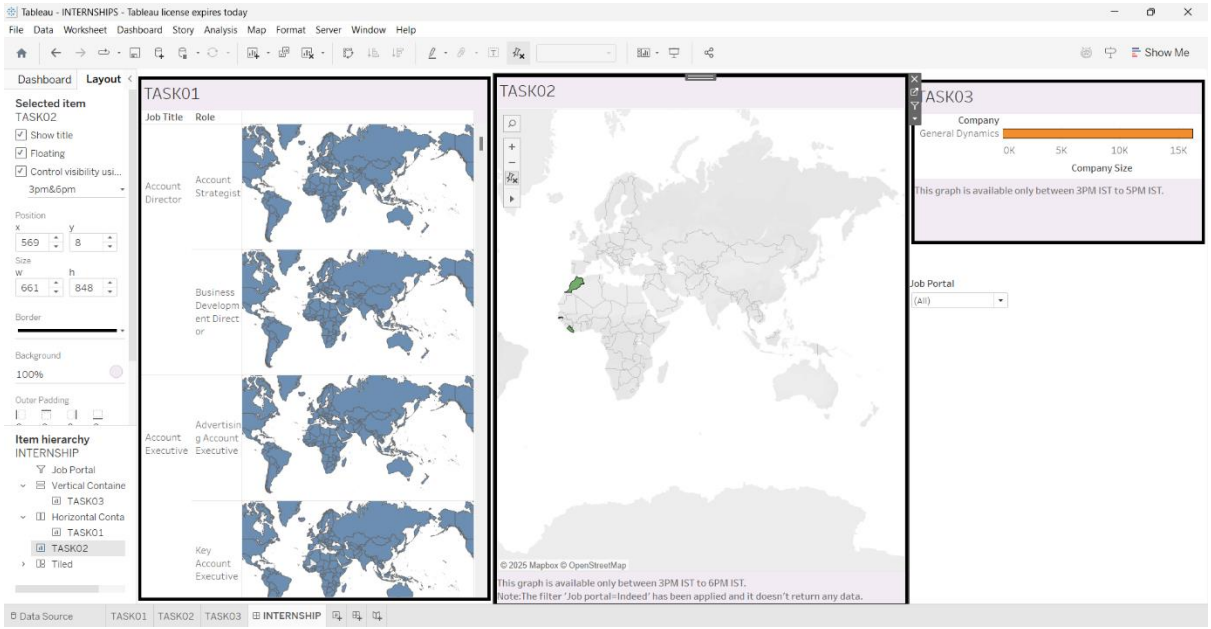


Fig:4.27

NOTE:Here ‘Job Portal’=”ALL”

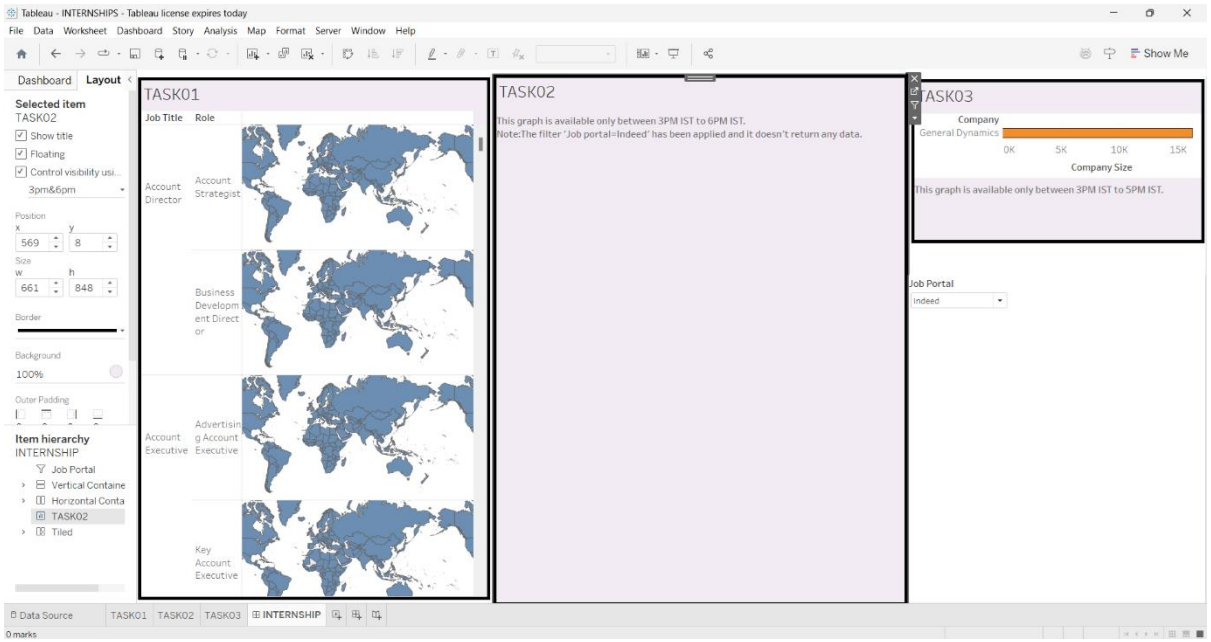


Fig:4.28

NOTE:Here ‘Job Portal’=”Indeed”

4.4.3.EXTRACTION OF LINK

For the above dashboards,the link was extracted by making it public through server.And an HTML code was written was uploaded to Netlify.

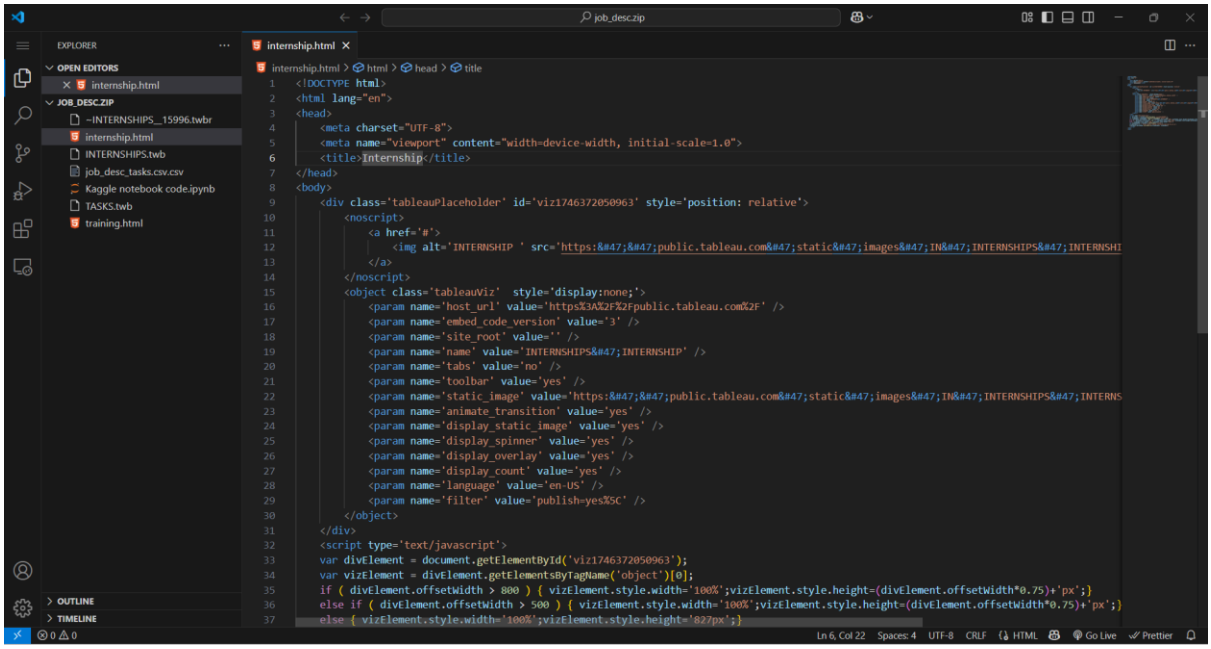


Fig:4.29

05.SKILLS AND COMPENTENCIES

Skills:

- Technical skills:Python(Pandas,NumPy),Tableau
- Analytical skills:Data cleaning,Data visualization

Compentencies:

- Ability to approach data related problems logically.
- Learning new tools

06.FEEDBACK AND EVIDENCES

- Gained real world scenario problem solving experience with perfect beginner friendly training and internship tasks.
- As for evidence,attaching screenshot related to participation in doubt clarification session.

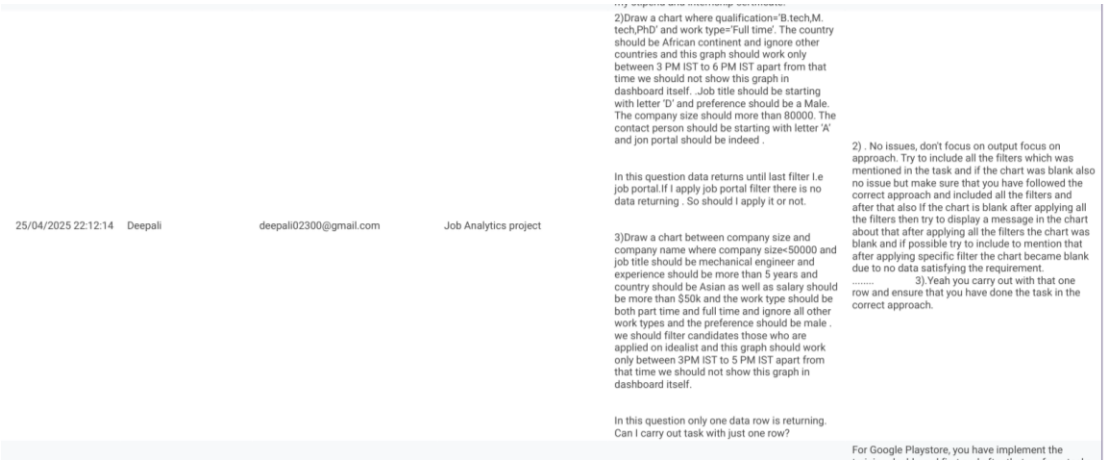


Fig:6.1

Training dashboard:
https://public.tableau.com/app/profile/deepali.b.a/viz/TASKStraining/TRAINING?publish=yes

Internship task dashboard:
https://public.tableau.com/views/INTERNSHIPS/INTERNSHIP?:language=en-US&publish=yes&:sid=&:redirect=auth&:display_count=n&:origin=viz_share_link

Github: deepa333777/Internship

07.CHALLENGES AND SOLUTIONS

7.1.Had trouble with uploading timer conditions to Tableau server.

Found out that tableau server works with UTC timing so added (5.5/24) to work at IST timings.

7.2.The timer filter applied on worksheet doesn’t work dynamically on server or autorefresh the filters.

Used containers from objects of dashboard and dragged graphs into it. And filtered out this containers in Layout by applying timer filters.

08.OUTCOMES AND IMPACT

During my internship, I analyzed a Job Dataset.I cleaned and standardized records and created interactive dashboards using Tableau.

The project demonstrated my ability to apply analytical tools to real-world data and generate impactful business insights efficiently and effectively.

09.CONCLUSION

The internship provided valuable experience and essential skills such as Pandas,NumPy and Tableau that are required in real world job scenarios.