NULLCLASS DATA SCIENCE INTERNSHIP

TASKS COMPLETION REPORT

1. INTRODUCTION

After completing training in data science from Nullclass, I am well equipped for the internship tasks provided to me. Tasks are related to data science and deploying them to a website, a complete end-to-end project.

After completing them on Tableau software, deploying them to the website as different tasks.

2. BACKGROUND

For the task, background knowledge of Tableau, data analysis, data cleaning and filtering was required, some features were not present in the data, which were required for the task, where data filtration and extraction techniques comes in action.

Python knowledge to code such techniques was essential to accomplish the tasks to the fullest.

3. LEARNING OBJECTIVES

Learning objectives were:

- To understand how data extraction and feature engineering is done on a data using python.
- How tableau features can give us the desired results.
- How the deployment of different tasks is done on a single website.

4. ACTIVITIES AND TASKS

The tasks were:

- Draw a chart between company size and company name where company size<50000 and job title should be mechanical engineer and experience should not 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 chart should work on only between 3 Pm to 6 Pm.

- 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. The job title should be starting with letter 'D' and preference should be a Male.The company size should be 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.

- Draw a chart where country=’india and Germany ’ ,qualification=’B.tech’ and work type=’Full time’ where experience should be more than 2 years and job should be Data Science and salary range should be more than $100k . The india details should be on orange color and Germany details should be on green color . The job portal should be indeed and preference should be female . The job posting date should be below 08/01/2023 and jon portal should be indeed . This chart should swap color after 12 PM to 6 PM .

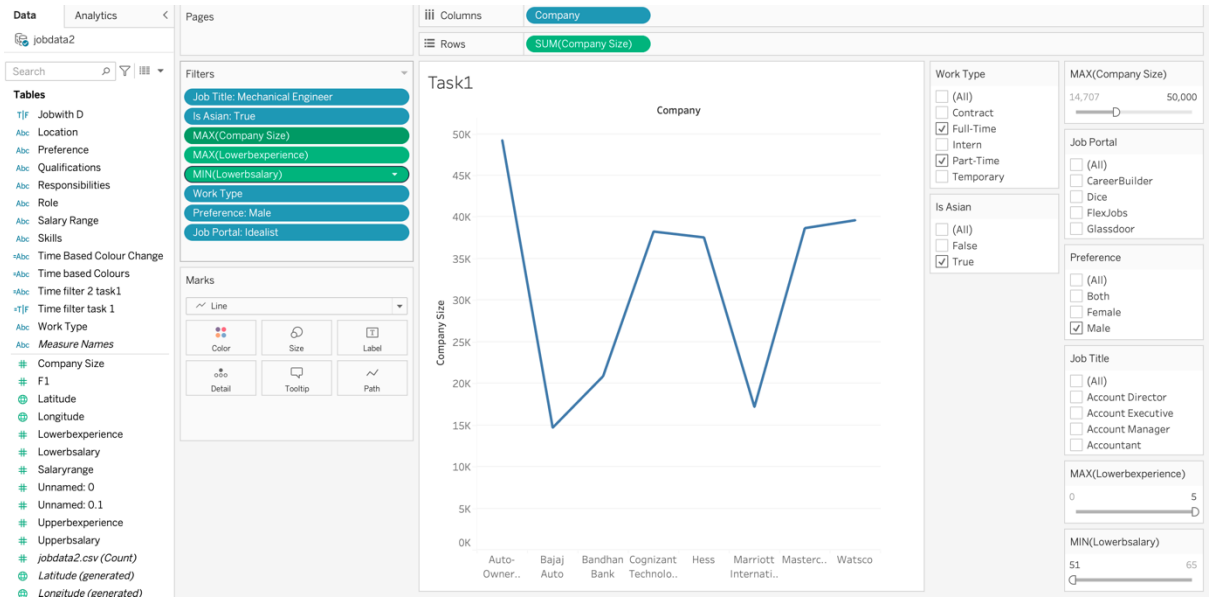
5. SKILLS AND COMPETENCIES

For the tasks to be completed, one should know the tableau basics, data engineering basics and deployment basics.

6. FEEDBACK AND EVIDENCE

For the task 1:

- Since the task 1 asks us to plot a chart between company size and company name with some filters like company size, country region, salary, etc. and the chart must be working between the time of 3pm to 6pm
  Here’s the chart looking like after applying the filters:

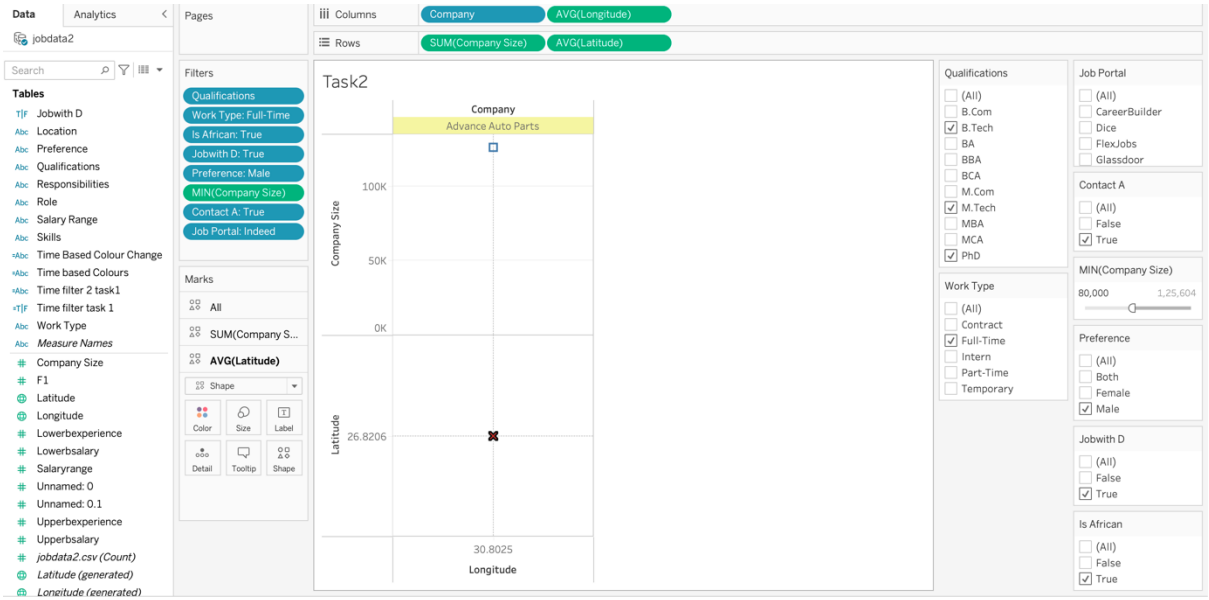


The chart clearly shows the applied filters on the data and the resulting graph.

For Task 2:

- We need to plot a chart again, but this time with some other filters, this time filters include qualification, work type, contact person, etc.
- And at last, we need to make it clickable for the longitude and latitude which opens the map for that location.

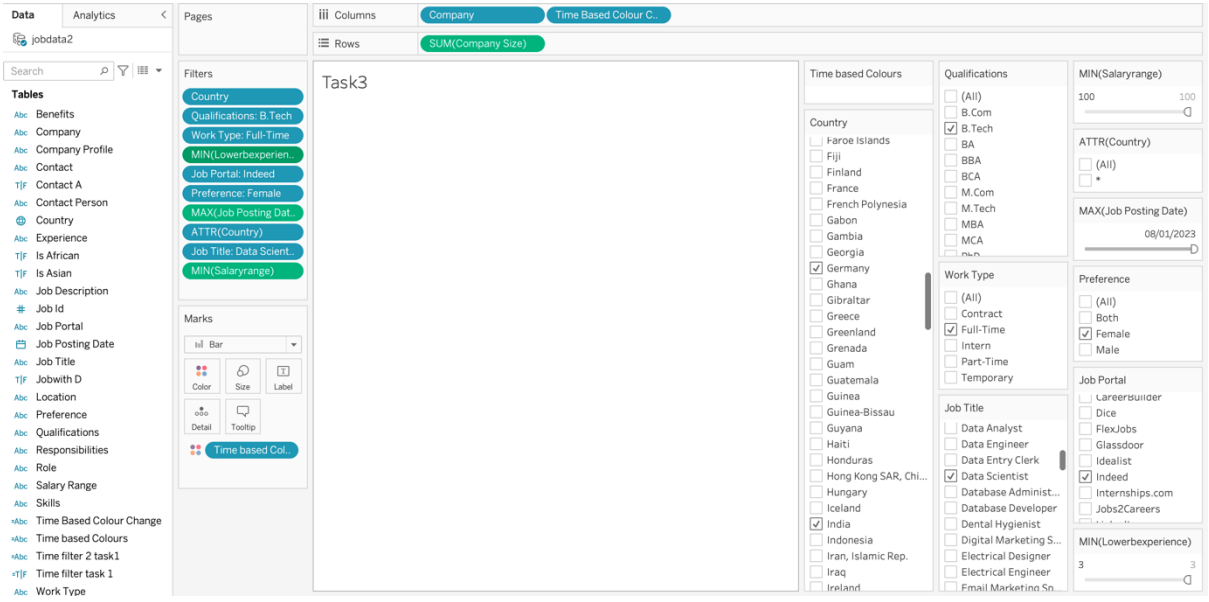
- Here’s the plot with all the filters:



- The X symbol represents the longitude and latitude of the location, after clicking on it we go to the map view.

For Task 3:

- There’s no data point in the original data set which follows all the needed requirement. As confirmed by your side, if there’s no data for that filter, then we are not showing anything.
  The plot look like this:



7. CHALLENGES AND SOLUTIONS

There were few challenges with all of the three tasks:

- Challenge 1:
  Original dataset didn't have the required columns needed for the task for example in task 1, they want the countries to be Asian.
  Solution:
  So we need to create a binary column for every row which indicates that the job posting is in asia or not.
  Similary we make many features list below:

  i) Lowerbexperience: the Lower bound of experience required
  ii) Upperbexperience: the Upper bound of experience required
  iii) Is_african: The job posting is in Africa or not
  iv) JobwithD: The job posting starting with letter D
  v) ContactA: The contact person whose name starts with A
  vi) Upperbsalary: the upper bound of salary
  vii) Is_asian: the job posting is in asia or not
  viii) lowerbsalary: the Lower bound of salary
  ix) salaryrange: Difference of upperbsalary and lowerbsalary .

- Challenge 2
  As asked in task1 salary should not be more than 5 years, but there's a range given in the dataset

  Solution:
  We interpreted and tacked the situation by creating a column as the lower bound of experience required and for the task, its value should not be more than 5 years.

  Same situation occurs for the salary filter, since we don't want the salary to be less than 50K. We created a lower bound of salary column and filter it with the value of 50K.

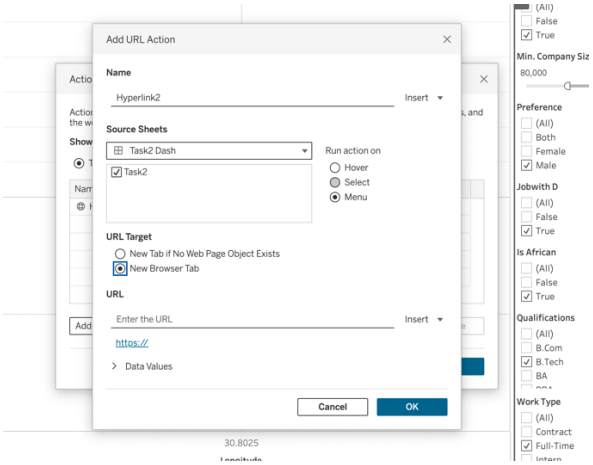
- Challenge 3
  Since we need to show chart between the time of 3pm and 6pm. We first used the code calculation technique of tableau where we write some code to be performed while execution. But that thing was really not responsive as we deploy the dashboard on the website.

  Solution:

We used JavaScript techniques to follow that order, where we used system's time to be used for the purpose. Detailed code is present in the source code given.

- Challenge 4:
  Since in task2, we were asked to create a system where click on the latitude and longitude and we got that location on the map, again here I used the tableau techniques to tackle this problem. But again the problem was when deployment was done.

  Solution:
  Since there's only one entry I manually added google map link of the location using the actions feature of dashboard as:



- Challenge 5:
  In task 3 we need to plot the charts with the required filters but in the original data we don't have any data fulfilling all the requirement of the filters.
  Then how to show color change with India and Germany.
  Solution:
  Since there's no data we show no chart hence, there is no need to do some coding for color change.

8. OUTCOMES AND IMPACT

The outcome is a website containing three dashboards for three different tasks of plotting job data.

Here's the link

Website : https://667ff847cd6edbf0962866e1--superb-vacherin-895050.netlify.app/

Github Link: https://github.com/zerobyzeroe/NullClass-Internship-Task

It creates an impact on our understanding of data visualization and Deployment techniques.

9. CONCLUSION

We created a Job analytics portal, which shows us different job trends on the basis of our filter settings and changes themselves based on time changes.