Project File
On
"DATA ANALYTICS for CLIENT-010"
(PHASE-I of Internship)

Prepared by

Parth Patel-19DCS098

Under the guidance of

Mr. Ajay Shah

Team Lead (Team-003)

&

Prof. Dipak Ramoliya

Assistant Professor

A Report Submitted to

Charotar University of Science and Technology

for Partial Fulfillment of the Requirements for the

8th Semester **[CS453]** Software Project Major

Submitted at

COMPUTER SCIENCE & ENGINEERING
DEPSTAR

DECLARATION BY THE CANDIDATE

I declare that the project report entitled **"DATA ANALYTICS for CLIENT-010"(PHASE-I of Internship)"** submitted by me to Devang Patel Institute of Technology and Advanced Research, CHARUSAT, Changa and AVM COMPUTERS PVT LTD. in partial fulfilment of the requirement for the award of the degree of **B.Tech** in Computer Engineering, from Department of Computer Science and Engineering , DEPSTAR, is a record of Bonafide CS453 Software Project carried out by us under the guidance of **Prof. Dipak Ramoliya and Mr. Ajay Shah**. I further declare that the work carried out and documented in this project report has not been submitted anywhere else either in part or in full and it is the original work, for the award of any other degree or diploma in this institute or any other institute or university.

Parth Patel(19DCS098) (003-I)

Mr. Ajay Shah

Team Lead

Prof. Dipak Ramoliya

Assistant Professor

Computer Science & Engineering (CSE)

DEPSTAR, Changa, Gujarat

TABLE OF CONTENTS

1. OVERVIEW

The overall internship is divided into following phases:

1. Data Analytics and Dashboard development (first-half)
2. Data Science and Machine Learning Portfolio. (second-half)
3. Customer Relationship Management (Entire tenure)

The first-half of the project is to be commenced from December 19 2022 till February 24 2023. This half will be solely dedicated to team-003 i.e. the intern will be working as a part of team-003 for the project of client-010. The project mainly involves the data analytics and dashboard development. As this project is purely an analytics project, there is no requirement of UML diagrams, COCOCO Model and FP Value estimation. The details are described further.

The second-half of the project is to be commenced from March 1 2023 till April 19 2023. The intern will be assisting 2 teams i.e. Team-001 and Team-003. The major focus will be on Data Science and Machine Learning side of the projects. The details of the second phase are yet to be provided by the organization.

2. ABOUT THE TEAM:

- Team Name: Team-003
- Total Members: 6 (5 permanent & 1 intern)
- Team Lead: Mr. Ajay Shah
- Process model to be utilized: Agile
- Integration tool : NA
- Dashboard tool: Tableau
- Data handling tool : Excel & MySQL
- Data manipulation : Permitted with prior and post approval of client
- Mode of work: Offline (WFH if necessary)(need to take approval)
- Estimated Time : 2.5 months
- Project Deadline: March 1 2023
- CRM Update Frequency : twice in a week

ROLE OF TEAM MEMBERS:

- Team Lead:

Responsible for the overall management & Synchronization

- Team-member-002 & Team-member-003:

Data Gathering and pre-processing

- Team-member-004:

Dashboard Development

- Team-member-005:

Communication and Data Management and Security

- Team Member-003-I:

Assist all the team members (except for team-member-005) and handle CRM

3. REQUIREMENTS:

- Software Requirements:
    - Tableau (premium)
    - Active Internet Connections
- Hardware Requirements:
    - NA

- Functional Requirements:
    - Dashboard Interaction Pattern
    - Data Partitioning
    - System Responsiveness
- Non-Functional Requirements:
    - Scalability
    - Reliability
    - Maintainability

4. SRS:

A Software Requirements Specification (SRS) document is a formal document that describes the requirements for a software system. It typically includes detailed descriptions of the system's functions, features, and capabilities, as well as any constraints or dependencies that may impact the design or implementation of the system.

In some cases, an SRS document may not be necessary for a data analytics project, especially if the project is focused primarily on data analysis and visualization rather than software development. This is because the primary deliverables of the project are the insights and conclusions drawn from the data analysis, rather than a software system that implements those insights.

5. COCOMO MODEL & FP ESTIMATION:

COCOMO (Constructive Cost Model) and FP (Function Point) are software estimation models used to estimate the effort, time, and cost required to develop a software project. These models are typically used in traditional software development projects, where the software is being developed from scratch.

In the case of data analytics projects, the focus is on analyzing and visualizing data, rather than developing software from scratch. As a result, the effort required for data analytics projects is primarily dependent on the size and complexity of the data set, as well as the specific analytical techniques used. The software tools or systems used for data analysis and visualization may also have an impact on the effort required, but this is typically not as significant as in traditional software development projects.

Therefore, COCOMO and FP models are generally not used for data analytics projects, as they are not well-suited to estimating the effort required for this type of project. Instead, project managers and stakeholders typically rely on other methods, such as expert judgment or historical data from similar projects, to estimate the effort required for data analytics projects.

6. UML DIAGRAMS:

UML (Unified Modeling Language) diagrams are a standardized way of representing software systems, including their structure, behavior, and interactions. They are commonly used in traditional software development projects to document software designs, communicate requirements, and aid in software development.

In the case of data analytics and visualization projects, the focus is on analyzing and visualizing data rather than developing software from scratch. While some software tools or systems may be used to support the data analysis and visualization, these tools are often off-the-shelf products that do not require significant customization or software development.

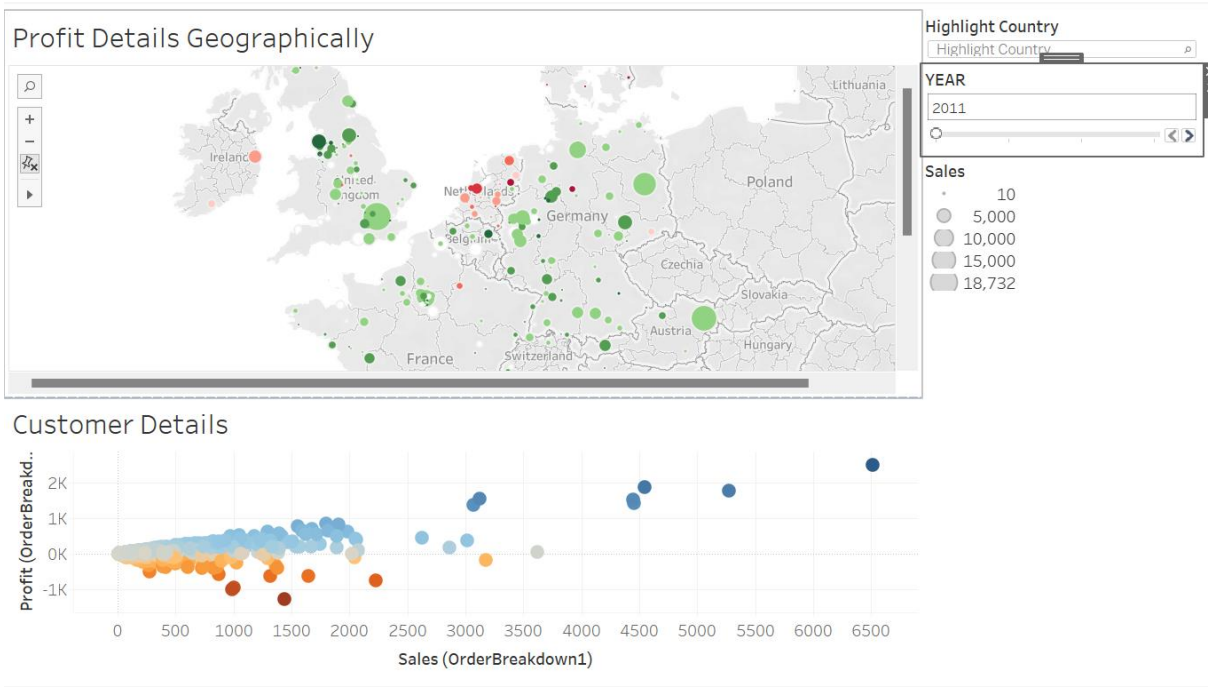
As a result, UML diagrams may not be as useful in data analytics and visualization projects, as the focus is on the data and the insights that can be gained from it, rather than the software systems that support the analysis and visualization

7. MINI-PROJECT-1:

Business Problem:

The client wants to keep an eye on their overall sales and profit across the Europe on yearly basis, along with minute details of the customer.

DASHBOARD:



Change in granularity: 1. Country wise and 2. Yearly basis.

Reason behind such design:

The customer wanted simple color combination.

The geographical map showcases yearly profits of the areas.

Towards red means loss, and profit for going towards green.

The lower chart depicts the sales and profit for every customer.

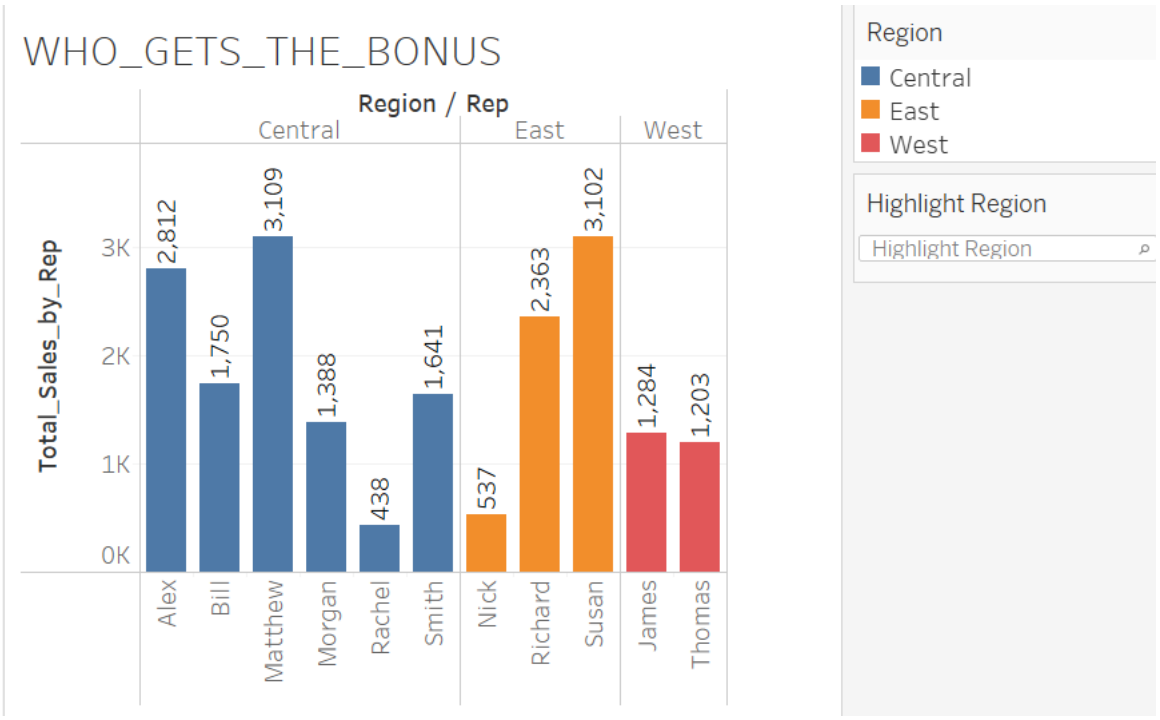
The charts are interactive and filter is applicable to both.

8. MINI-PROJECT-2:

BUSINESS PROBLEM:

The client wants to give yearly bonus to top-performing employee from each region. Help them in finding out the one who deserves to get the bonus.

DASHBOARD:



MY INSIGHTS:

Firstly, from the data provided, there were no direct indications for the sales and profits made by every employee. Thus, I created a calculated field which calculates the overall sales and profit made by all the employees.

There are 3 regions: Central, West and East.

Thus, 3 employees in total will get the bonus.

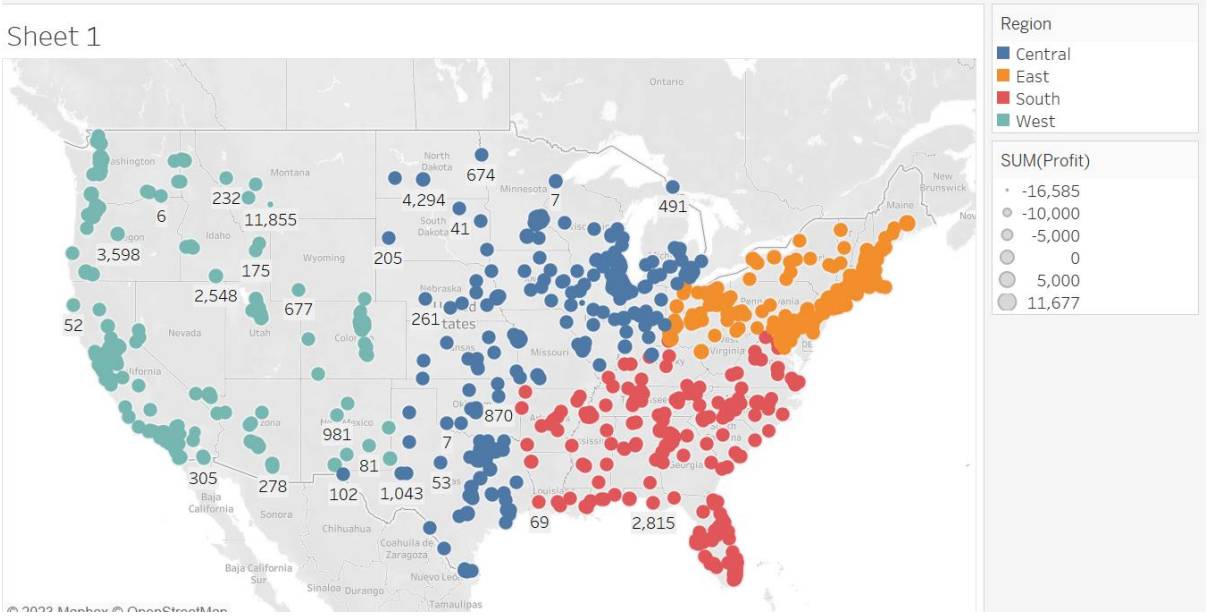
So, it is evident that Matthew from Central, Susan from East and James from West will get the bonus.

9. MINI-PROJECT-3:

BUSINESS PROBLEM:

The client wants city-wise summarized details of how their stores are performing?

DASHBOARD:

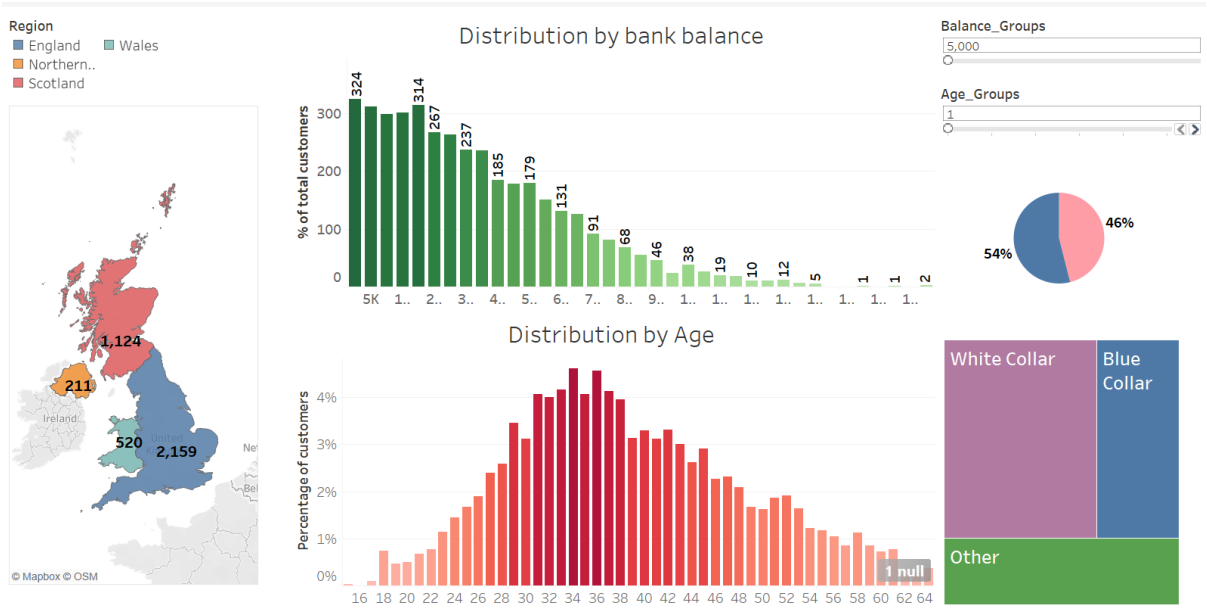


10. MINI-PROJECT-4:

BUSINESS PROBLEM:

There is a bank in United Kingdom. Now, the bank wants to retain their customers and to add new one. The board of directors have agreed to develop multiple schemes for different sections of people and regions. They want you to design a dashboard through which they can play with it to determine the new policies. Also, list out the important parameters that they should consider for development of new policies.

DASHBOARD:



Filter: Applicable to all

Individual Granularity: Limited to Age Distribution and Bank Balance.

MY INSIGHTS:

There are 4 regions in the UK where bank have presence. Out of the 4, the England have largest customer base. Thus, it will be a good idea to implement your new policies and incentives from England.

Secondly, most of the customers are having white collar jobs, and it is obvious as the economy of the UK is more of technology oriented where white collar jobs form a major workforce. Thus, new incentives should be focused more on benefitting white collar job holders.

Here, though male constitutes a whopping 56% customer base, the bank should focus on bringing more and more women customers for a simple reason that, the world is changing and so, more and more females are contributing in the economy. Hence, lucrative offers for women can be a master stroke. Remember, the bank need to grab the first mover advantage.

Considering the bank balance of the customers, majority of them falls in middle class category. Thus, the committee needs to have a dual strategy here. The one for adding more salaried middle class customer and second to lure upper class people or business tycoons.

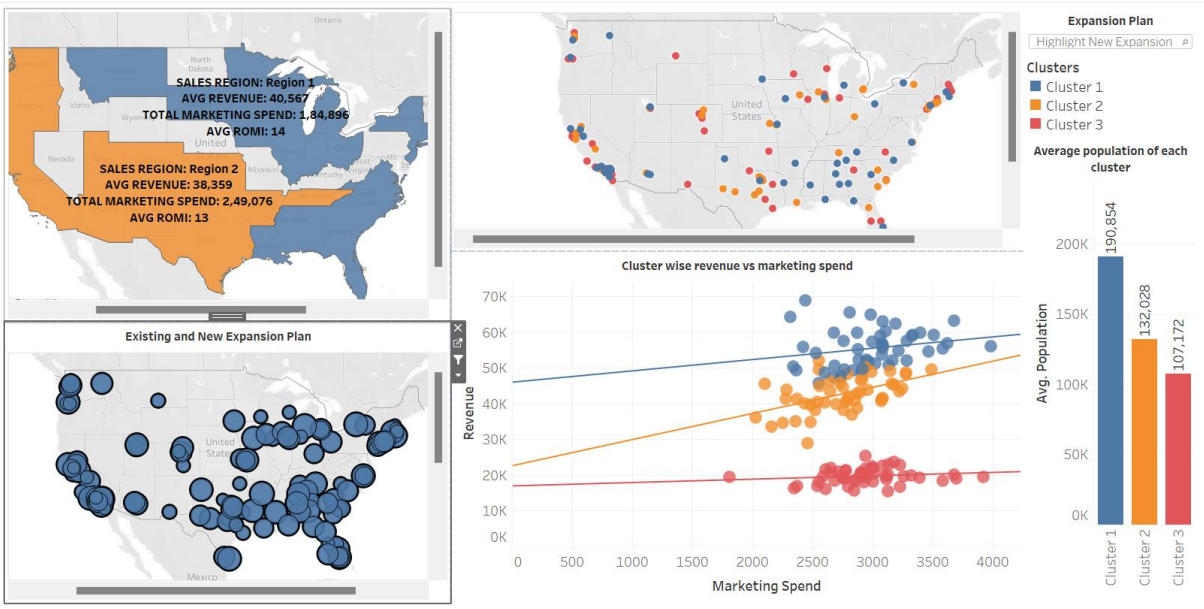
Considering the age distribution, it is quite opposite to the national trend. While UK has more aging population, however, the customer base of the bank is relatively younger. Thus, considering the pro-immigrant stance of the government, the bank should consider to target younger immigrants and citizens for sustainable growth.

11. MINI-PROJECT-5:

Business Problem:

A Laundry business owner wants to expand his business. Now, he is spending a fortune on advertisements and marketing as well. So, he wants to cut the marketing expenditure in the new expansion plan and wants to open more shops in the areas where he can maximise the profits while spending a little on marketing. Also, he wants to shut down his shops in loss making or less profit making areas.

DASHBOARD :



INSIGHTS:

NOTE: THE INSIGHTS REPORT HAS BEEN SUBMITTED TO THE TEAM LEAD. THE FEEDBACK IS PENDING TILL THE DAY OF CREATION OF THE REPORT. THUS, CANNOT INCLUDE IN THE REPORT.

NOTE:

THE INTERN HAS ALSO WORKED ON GROUP PROJECT ALONG WITH THE TEAM. HOWEVER, BECAUSE OF THE PRIVACY POLICY OF THE CLIENT AND ORGANIZATION, THE INTERN IS NOT ALLOWED TO DISCUSS AND SHARE ANY DETAILS OF THAT PROJECT.

AUTHORISED SIGNATURE:

12. PHASE-2:

The phase-2 of the internship will begin from March 1 2023.
The portfolio is updated from Data Analytics and Dashboard development to Data Science and Machine Learning Projects.

The details of the group project is yet to be shared by the organization.
Following are the details of the mini-project that the intern needs to complete:

1. Text Analytics and classification
2. Text Analyzer
3. Course Recommendation System
4. Car Acceptability Predictor
5. CRM Assistance