** TAMILNADU COLLEGE OF ENGINEERING **

**GLAOBAL SALES DATA ANALYTICS**

**TEAM ID**: PNT2022TMID43171

**A PROJECT REPORT**

**SUBMITTED BY**

**HARIHARAN.R (TEAM LEADER)**

**ANDAN BALASINGAM.S (TEAM MEMBER)**

**DENEESH.G (TEAM MEMBER)**

**SIVA HARIHARAN.V (TEAM MEMBER)**

in partial fulfillment for the award of degree of

**Bachelor of Engineering** (B.E.) in

**COMPUTER SCIENCE AND ENGINEERING**

**TABLE OF CONTENT**

1. **INTRODUCTION**
2. **LITERATURE SURVEY**
3. **IDEATION & PROPOSED SOLUTION**
   1. Empathy Map Canvas
   2. Ideation
   3. Proposed Solution
   4. Problem Solution fit
4. **REQUIREMENT ANALYSIS**
   1. Functional requirement
   2. Non-Functional requirements
5. **PROJECT DESIGN**
   1. Data Flow Diagrams
   2. Solution & Technical Architecture
6. **PROJECT PLANNING & SCHEDULING**
   1. Sprint Planning & Estimation
   2. Sprint Delivery Schedule
   3. Reports from JIRA
7. **CODING & SOLUTIONING (Explain the features added in the project along with code)**
   1. Feature 1
   2. Feature 2
8. **TESTING** 
   1. Test Cases
   2. User Acceptance Testing
9. **RESULTS**
   1. Performance Metrics
10. **CONCLUSION**
11. **ADVANTAGES & DISADVANTADES**
12. **FUTURE SCOPE**
13. **APPENDIX**

Source Code

GitHub& Project Demo Link

**1.INTRODUCTION**

Nowadays, many companies have started to exist and a few of them have grown to top positions. Some companies have very large amounts of data while some small companies have fewer amounts of data. This paper aims to research the company's growth in terms of the sales of the company’s products. The first aim of the paper is to make a web application that analyses a company’s sales data of varied products. This analysis is going to be effective if we use graphs and charts. This process is named Visualization. Manual Visualization of knowledge may be a time taking process. There is already some software that performs this task, but there are many disadvantages. Software like Jupyter Notebook can't be used on mobile phones whereas websites are often accessed using mobiles. Our paper takes company data of sales of products and data of a salesperson working in a particular company and depicts graphs between fields required.

**2.LITERATURE SURVEY**

[1] Analysis of sales data of a company or retailer has been becoming a widely discussed topic. The sales data can consist of many records, and filtering of sales data to find meaningful intuitions are common techniques in sales analysis. Tools like dashboards help managers and owners in visualizing aggregated data. Some tools usually show the items that are sold by different sales’ points. To understand data by visualization was used as early as 1137. In all fields, there has been vast Development in visualization techniques. Examining information and data visualization help to visualize and express ideas in architecture. With the coming of computer simulation visualization pertinence has been faster. Visualization of data is used to present design data with the aid of drawings and diagrams and data is usually conceptual or special, we require scientific visualization techniques like charts and graphs, etc. Visualization should have the power to present multidimensional data and it must be synergistic and permit efficacious communication. Some researchers focus on the tools used for data analysis. They highly focus on how easy it is to use the dashboard, create connections and store data in databases and ease of sharing information .

**ADVANDAGES:**

* The main aim of our research paper is to design and develop a web application which can help companies and retailers to analyses and visualize enterprise sales data in the form of graphs.
* Since we are developing a web application for this purpose, we do not need any storage space as once deployed the application can be used by any device with an internet connection. Using websites, we can see our output quickly

**DISADVANDAGES:**

* There is a need to install software’s to analyses and visualize the sales data, users need to shift between different applications and it takes more time to visualize the data. Since, these software’s cannot be used from a mobile phone, there is a need for a computer desktop and installing these software which in turn requires space.

[2] we analyzed the data sets of world’s largest retailers, Waymart Store to determine the business drivers and predict which departments are affected by the different scenarios (such as temperature, fuel price and holidays) and their impact on sales at stores’ of different locations. We have made use of Scala and Python API of the Spark framework to gain new insights into the consumer behaviors and comprehend Waymart’s marketing efforts and their data-driven strategies through visual representation of the analyzed data

**ADVANDAGES:**

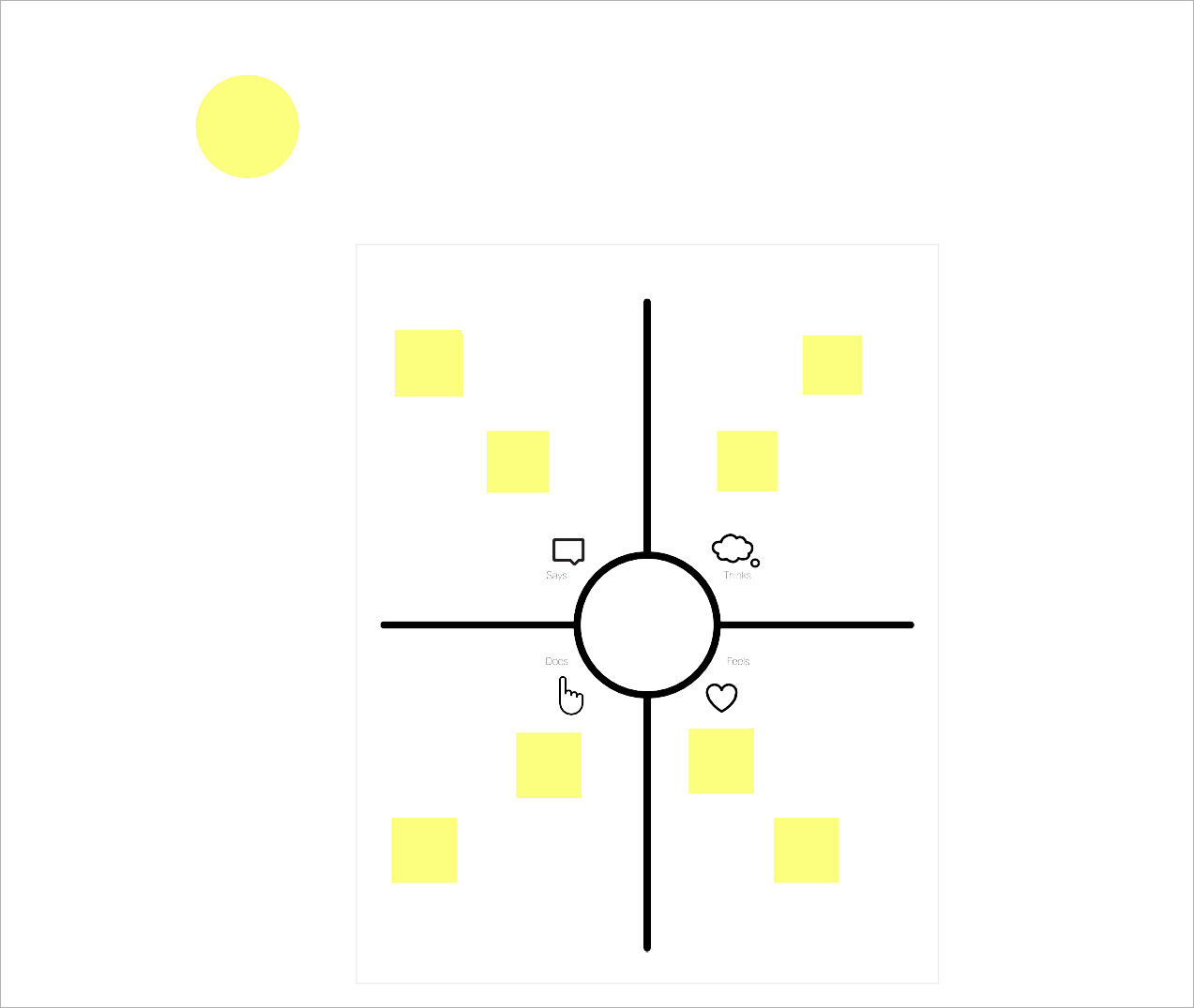
* After getting the analyzed data in key and value it is easier to graph and see relationships between values of the date and store location using GraphX library provided by Apache Spark using its python API.

**DISADVANDAGES:**

* Retailer’s ﬁrst priority is usually to understand their customers to be able to satisfy their needs so that these customers will return to the store for future needs, thus increasing the product demands and adding to the business value. These businesses want this information to plan where and when to invest proﬁtably.

**3.IDEATION& PROPOSEDSOLUTION**

**3.1 EMPATHY MAP CANVAS**



**EmpathyMap2**

maintaingcustomerloyalty

Employcybersecurityservice orengineers

Thehighteneddemand fordatasecurity

askingforcustomerfeedback

Givegoodqualityproduct

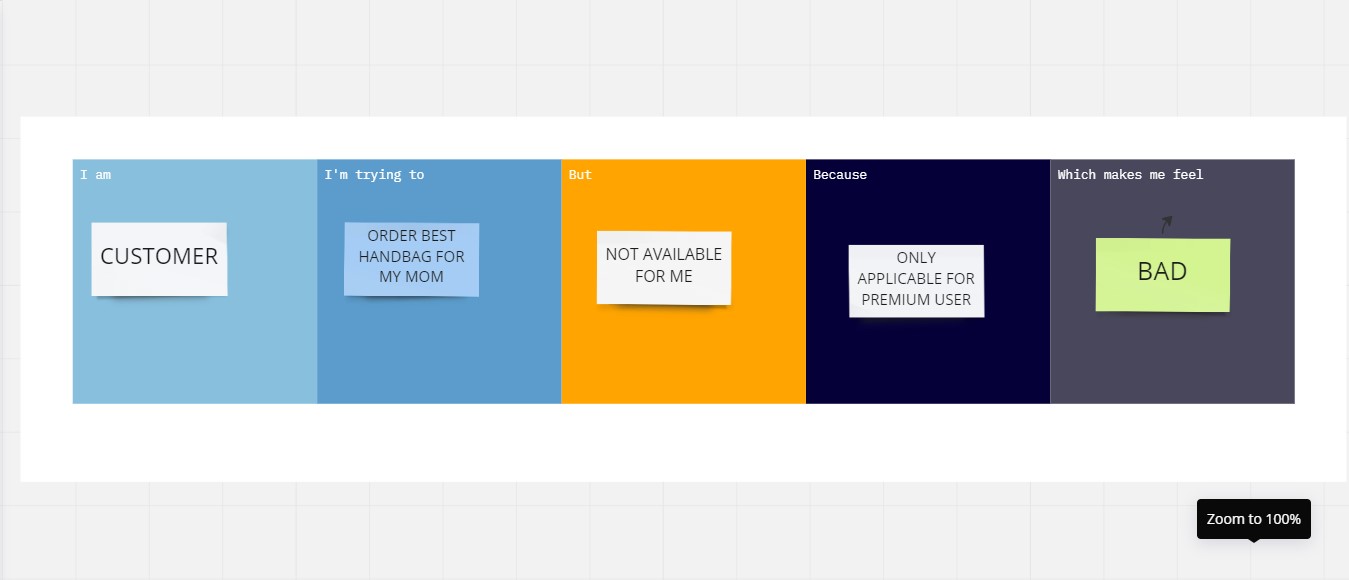
Beingresponsible

Attractive tothecustomer

Createawarenesstocustomer

3

3.2 IDEATION



3.3 PROPOSED SOLUTION

|  |  |  |
| --- | --- | --- |
| 1. | ProblemStatement(Problemtobesolved) | * CreateasimpleeasytounderstandanalyticsofSalesdata,by plottingdifferentvisualization. * Useof familiarmetricstoanalyze data. * EasytofindInsights of datawithclearand legiblecolorcoding. * Detailedinformationgathering * Localizationofareasofinterest,andcompleteanalysisonthem * Increasethe customerbuyingcapacity |
| 2. | Idea /Solutiondescription | * Identifythe customer’spriority * Creating an Interactive Dashboard. ResponsiveDesign for every screen sizes. Modular file basedanalytics.ManualInsights for eachinteraction |
| 3. | Novelty/Uniqueness | * UseArtificialIntelligencetogivesolution * Analytics are modular with the help of exporting andimporting files. Has ability to add manual insights forlaterviewing. |
| 4. | SocialImpact/ CustomerSatisfaction | * Customerwillidentifytheirneedseventhey don’t   know   * Customer gets instant analytical diagrams when theyinput the file to the software, as long as the file is inthe correct format. Customer can reuse the same anynumberof times. |
| 5. | BusinessModel(RevenueModel) | * AnyAI modelwithgoodaccuracyrate * One-time payment – for a user. Free Trial for 30days. |
| 6. | Scalabilityofthe Solution | * The solution scales well by default, as its file based.Any number of similarly formatted files can besubmitted and the analytics will be drawn for thatparticularfile. |

**3.4 PROBLEM SOLUTION FIT**

# CUSTOMERSEGMENT:

* + A businessownerwhowouldlike tounderstandmoreabouthisbusinessperformanceinglobalscale.

# JOBSTO BEDONE/PROBLEM:

* + Determine inputfilestructure.
  + Whatanalysistoperformtobeuseful?andhowto performthem?

# TRIGGERS:

* + haveyoueverfeltthatyouareunawareofhowyourbusinessisperforming?
  + haveyouever hadadecisionfatigue?Notknowingwhattonextinordertoprogress?

# EMOTIONSBEFORE /AFTER:

* + Before:Anxiety,decisionfatigue,laziness.
  + After:Clearmind,peacefulness.

# AVAILABLESOLUTION:

* + Thecompetitionanalyticsanddisplaydashboardwithautogeneratedinsights.
  + Outproductprovidefacilitytoaddmanualinsightsto theanalysisperformed.

# CUSTOMERCONSTRAINS:

* + Noonlinepaymentsavailable.buydirectlyfromus.
  + Needtocheckinputfilestructurebeforeuploading.

# BEHAVIOUR:

* + Collectingsalesdataandusingofficesoftwaretoanalyzeit.
  + Un-intuitivewayofanalyzing dataandlot ofmanuallabor.

# 8.CHANNELSOFBEHAVIOUR:

* + Using third partly services with automatically insights and subscription based servicestoanalysesdata.
  + Usingofficesoftwaretoanalyzecomplexdatainun-intuitiveway.

# 9.PROBLEMROOTCASE:

* + IBM.
  + Annauniversity.
  + Businessmodel.
  + Society.

# 10.YOURSOLUTION:

* + Creatinganinteractivedashboard.
  + Responsivedesign foreveryscreensizes.
  + Manualinsights foreachinteraction.
  + Onetimepayment.

**4.REQUIREMENT ANALYSIS**

**4.1 FUNCTIONAL REQUIREMENT:**

|  |  |  |
| --- | --- | --- |
| **SI**  **No.** | **FunctionalRequirement(Epic)** | **SubRequirement(Story/Sub-Task)** |
| 1 | UserRegistration | Registration through FormRegistrationthroughGmail |
| 2 | UserConfirmation | Confirmation via EmailConfirmationvia OTP |
| 3 | UserLogin | LoginviaEmailandpassword |
| 4 | Useruploadingdata(administrative) | Tostorethe datasetthroughthecloud |
| 5 | Enduserbenefits | Gettinghigherstateofefficiencyand alsotoknow entire dataanalysis |

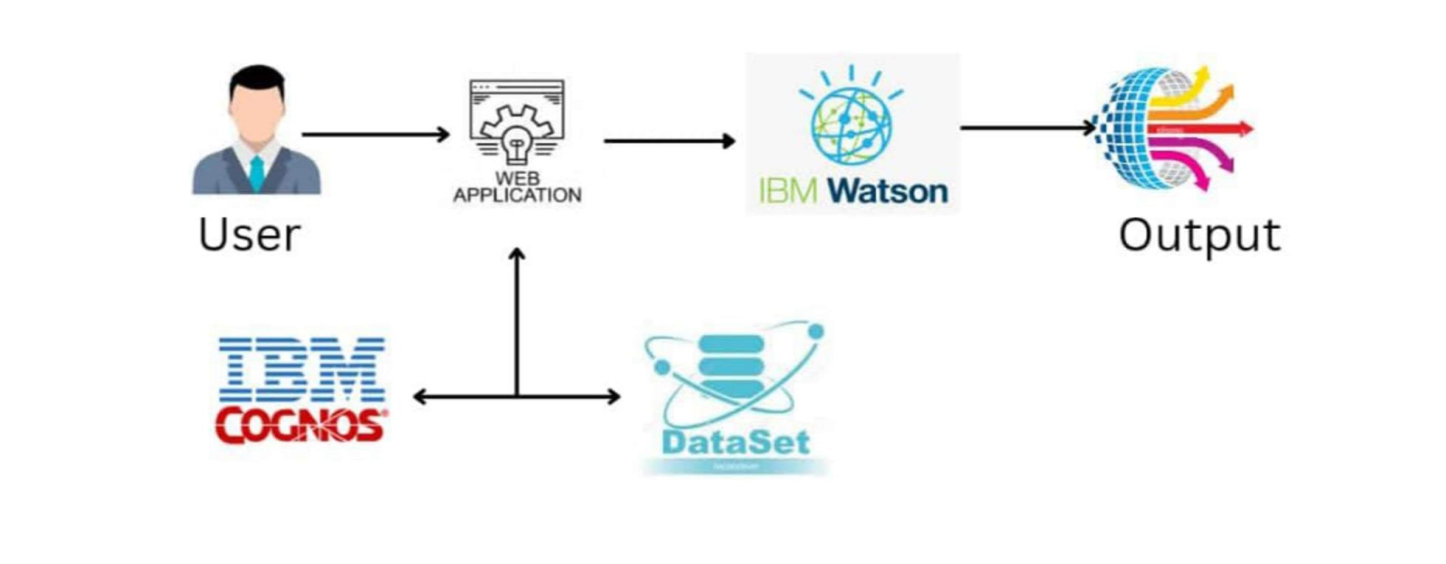
**4.2 NON FUNCTIONAL REQUIREMENTS:**

|  |  |  |
| --- | --- | --- |
| **SI**  **No.** | **Non-FunctionalRequirements** | **Description** |
| 1 | Usability | Optimizedresourcesanditcanbeusedbyeveryone |
| 2 | Security | Ithassecurablebecauseithasendtoendencryption |
| 3 | Reliability | It hashighreliabilitybasedondevelopment. |
| 4 | Performance | Ithashighstateofperformanceand efficiency. |

|  |  |  |
| --- | --- | --- |
| 5 | Availability | Ithasavailableinallplatformsandwebsites. |
| 6 | Scalability | The ability of a hardware and software parallelSystemtoexploitincreasingcomputingresourcesefficiencyintheanalysisofthe(very)largedatasets |

**5.PROJECT DESIGN**

**5.1 DATAFLOW DIAGRAM:**

A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clearDFDcandepicttherightamountofthesystemrequirementgraphically.Itshowshowdataentersandleavesthesystem,whatchangestheinformation,andwheredataisstored.

**5.2 SOLUTION & TECHNICAL ARCHITECTURE**

Solution architecture is a complex process – with many sub-processes – that bridges the gapbetween business problems and technology solutions. Its goals are to:

* Find the best tech solution to solve existing business problems.
* Describethestructure,characteristics,behavior,andotheraspectsofthesoftwaretoproject stakeholders.
* Define features, development phases, and solution requirements.
* Provide specifications according to which the solution is defined, managed, and delivered.

**6.PROJECT PLANNING & SCHEDULING**

**6.1 SPRINT PLANNING & ESTIMATION:**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Sprint** | **FunctionalRequirement(Epic)** | **UserStoryNumber** | **UserStory/Task** | **StoryPoints** | **Priority** | **TeamMembers** |
| Sprint-1 | Registrationandlogin | USN-1 | Asauser,Icanregisterfortheapplicationbyentering my email, password, and confirmingmypassword. | 10 | High | Hari  Andan  Deneesh  siva |
| Sprint-1 |  | USN-2 | Asauser, IwillreceiveconfirmationemailonceIhaveregisteredfortheapplication | 10 | Medium | Hari  Andan  Deneesh  siva |
| Sprint-1 |  | USN-3 | Asauser,IcanregisterfortheapplicationthroughFacebook | 10 | High | Hari  Andan  Deneesh  siva |
| Sprint-1 |  | USN-4 | Asauser,IcanviewthedashboardcontainsappropriateSales Datasets | 10 | High | Hari  Andan  Deneesh  siva |
| Sprint-2 | AccessingDashboard | USN-5 | Asauser,IcanviewthedashboardcontainsappropriateSales Datasets | 10 | Medium | Hari  Andan  Deneesh  siva |
| Sprint-3 | Helpdesk | USN-6 | Asacustomercareexecutive,Icanprovideyoufreetrialtoaccesssome ofthedatasets | 05 | High | Hari  Andan  Deneesh  siva |
| Sprint-3 |  | USN-7 | Asacustomercareexecutive,Iwillcollectthefeedback and suggessions from thecustomer,afterusingthis application | 05 |  | Hari  Andan  Deneesh  siva |
| Sprint-4 | Admin | USN-8 | AsanAdmin,Icanauthenticatetheregistrationand logincredintialsofthe customer | 10 | Medium | Hari  Andan  Deneesh  siva |
| Sprint-4 |  | USN-9 | AsanAdmin,IensurethesecurityoftheCustomerdetails. | 05 | High | Hari  Andan  Deneesh  siva |

**6.2 SPRINT DELIVERY SCHEDULE:**

|  |  |  |
| --- | --- | --- |
| **TITLE** | **DESCRIPTION** | **DATE** |
| **Literature Survey &InformationGathering** | Literature survey on theselected project & gatheringinformation by referring thetechnical papers, researchpublicationsetc. | 28SEPTEMBER2022 |
| **PrepareEmpathyMap** | Prepare Empathy Map Canvasto capture the user Pains &Gains, Prepare list of problemstatements | 24SEPTEMBER2022 |
| **Ideation** | List the by organizing thebrainstormingsessionandprioritize the top 3 ideasbased on the feasibility &importance. | 25SEPTEMBER2022 |
| **ProposedSolution** | Preparetheproposedsolutiondocument, which includes thenovelty,feasibility ofidea,business model, social impact,scalabilityof solution,etc. | 23SEPTEMBER2022 |
| **ProblemSolutionFit** | Prepare problem - solution fitdocument. | 30SEPTEMBER2022 |
| **SolutionArchitecture** | Preparesolutionarchitecturedocument. | 28SEPTEMBER2022 |

|  |  |  |
| --- | --- | --- |
| **CustomerJourney** | Prepare the customer journeymaps to understand the userinteractions & experienceswith the application (entry toexit). | 20OCTOBER2022 |
| **FunctionalRequirement** | Prepare the functionalrequirementdocument. | 8OCTOBER2022 |
| **DataFlowDiagrams** | Draw the data flowdiagramsandsubmitthemforreview. | 9OCTOBER2022 |
| **TechnologyArchitecture** | Prepare thetechnology  architecturediagram. | 10OCTOBER2022 |
| **Prepare Milestone & ActivityList** | Preparethemilestones&activitylistoftheproject. | 10 NOVEMBER 2022 |
| **Project Development -DeliveryofSprint-1,2,3&4** | Develop & submit thedevelopedcodebytestingit. | INPROGRESS. |

**6.3 REPORT FROM JIR:**

**BURNDOWN CHART:**

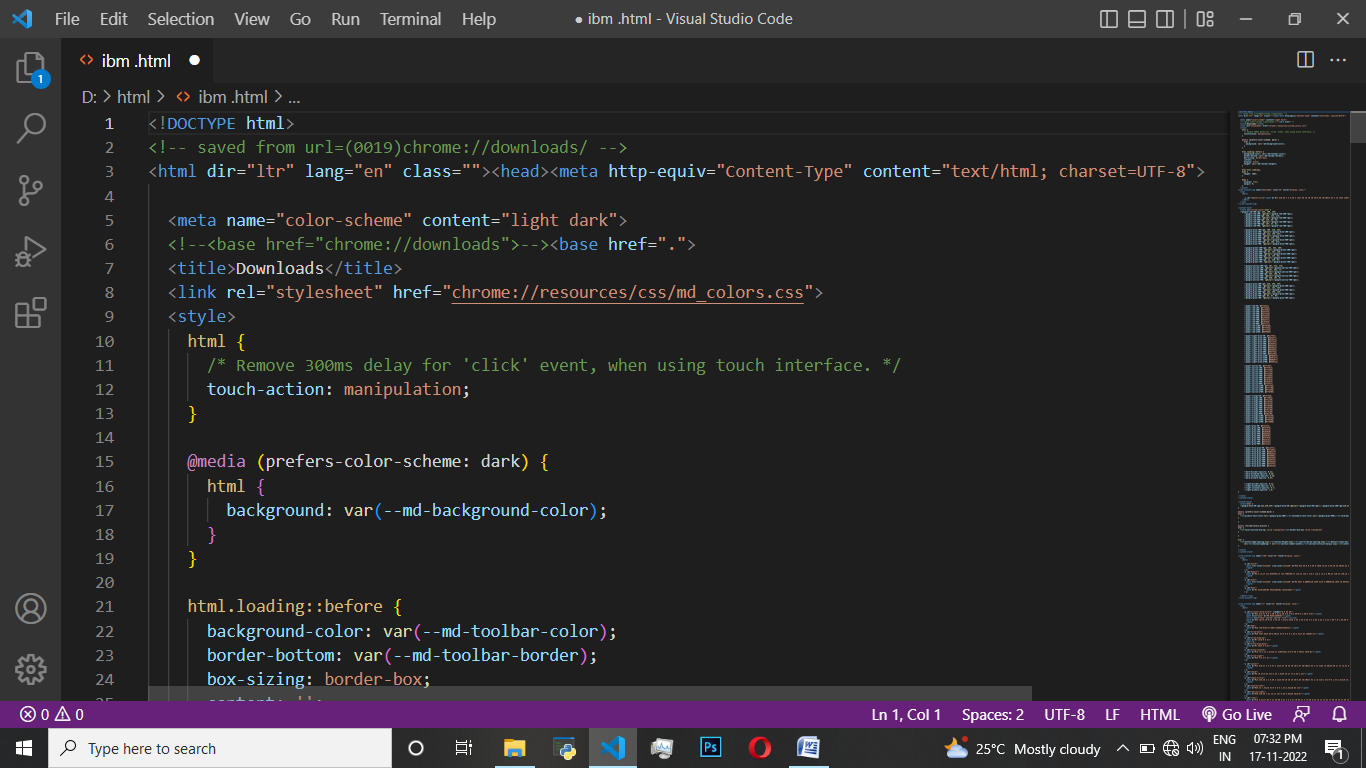
A burn down chart is a graphical representation of work left to do versus time. It is often used in agile [software development](https://www.visual-paradigm.com/scrum/what-is-agile-software-development/) methodologies suchas [Scrum](https://www.visual-paradigm.com/scrum/scrum-in-3-minutes/).However,burn down chartscanbe appliedto anyprojectcontaining measurableprogressovertime.

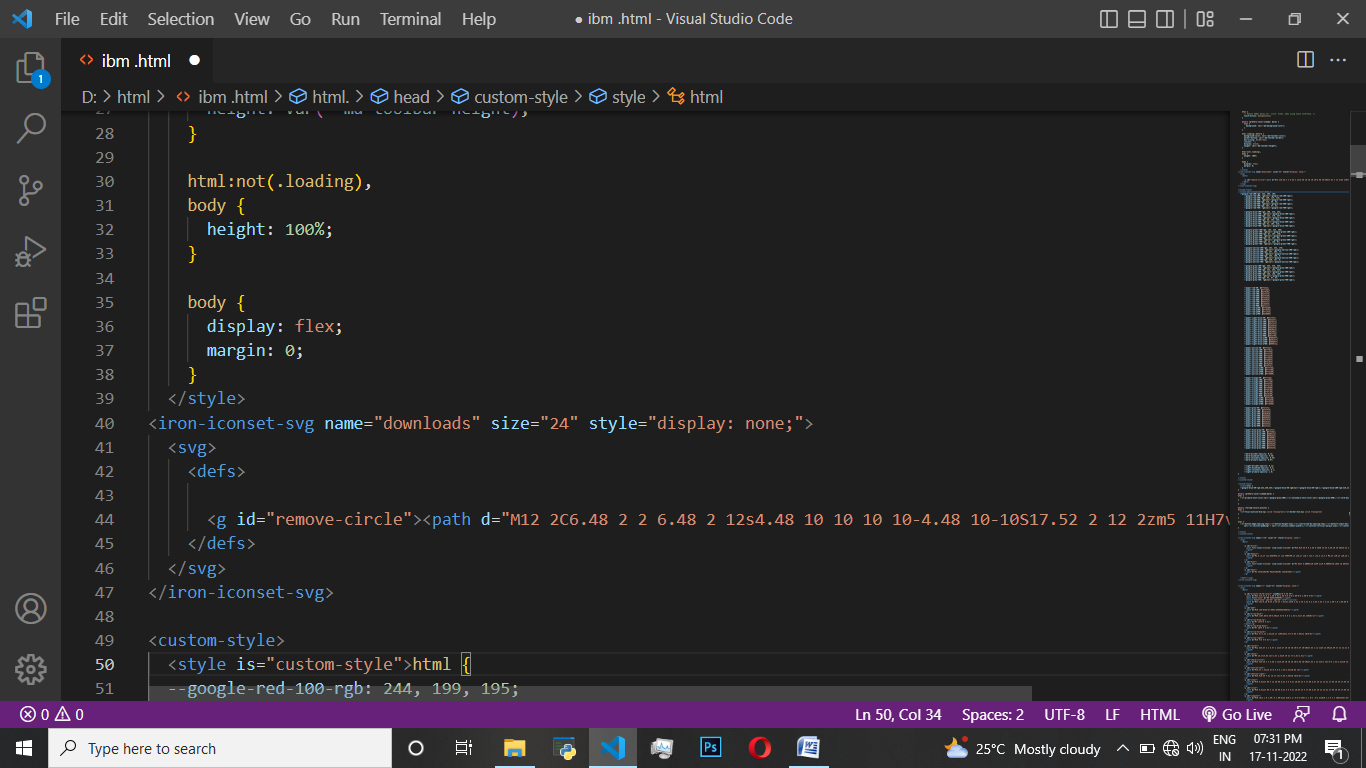
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Sprint** | **Total StoryPoints** | **Duration** | **SprintStartDate** | **SprintEndDate(Planned)** | **Story PointsCompleted (as onPlannedEndDate)** | **SprintReleaseDate(Actual)** |
| Sprint-1 | 20 | 6Days | 24Oct2022 | 29Oct2022 | 20 | 29Oct2022 |
| Sprint-2 | 20 | 6Days | 31Oct2022 | 05Nov2022 | 20 | 05Nov2022 |
| Sprint-3 | 20 | 6Days | 07Nov2022 | 12Nov2022 | 20 | 12Nov2022 |
| Sprint-4 | 20 | 6Days | 14Nov2022 | 19Nov2022 | 20 | 19Nov 2022 |

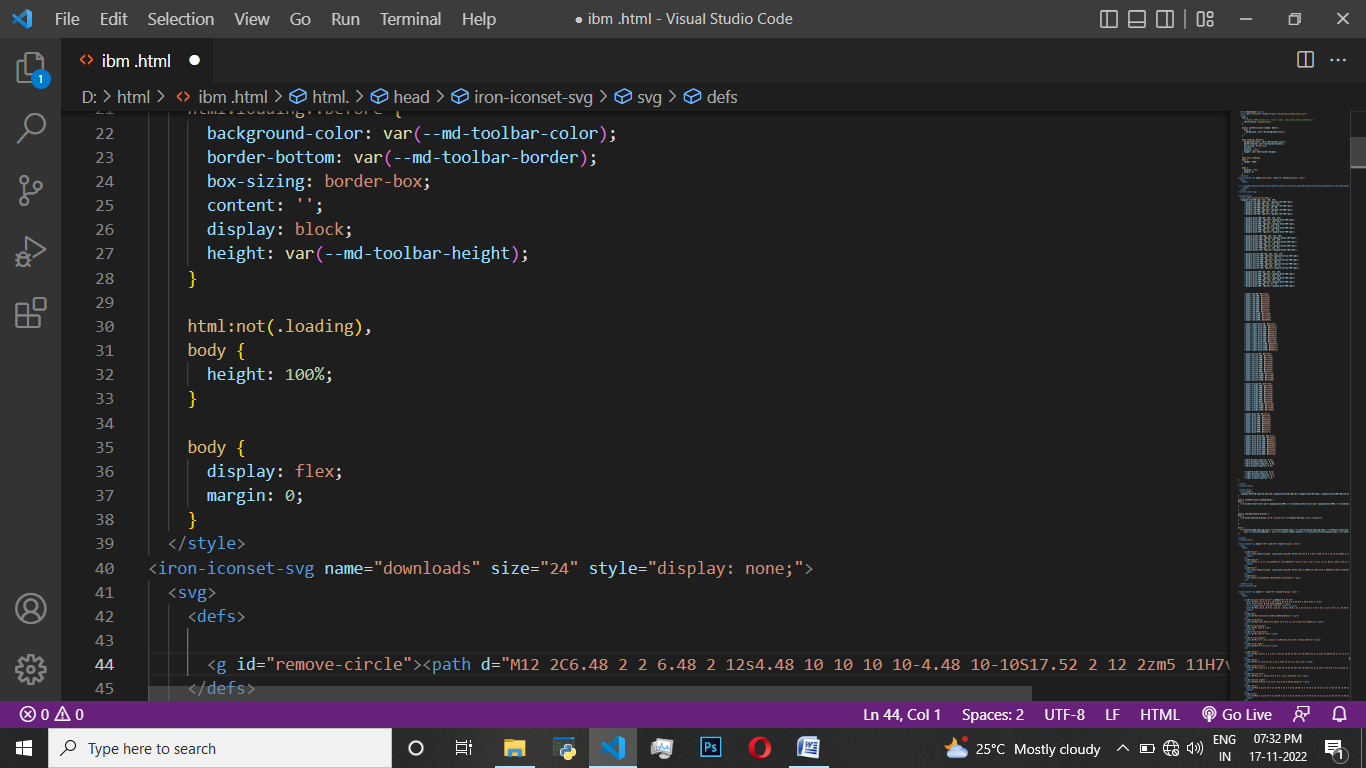


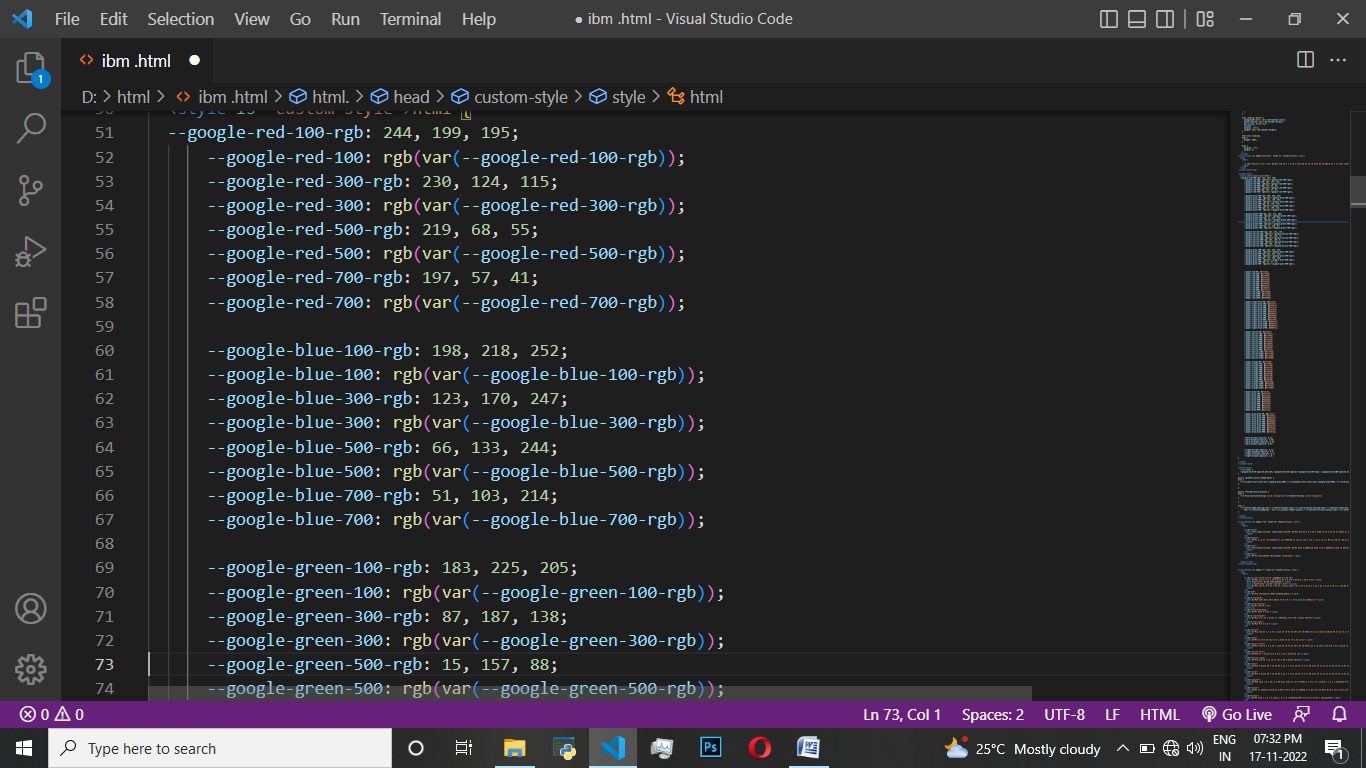
**7.CODING AND SOLUTION**:

**7.1 FEATURE 1:**

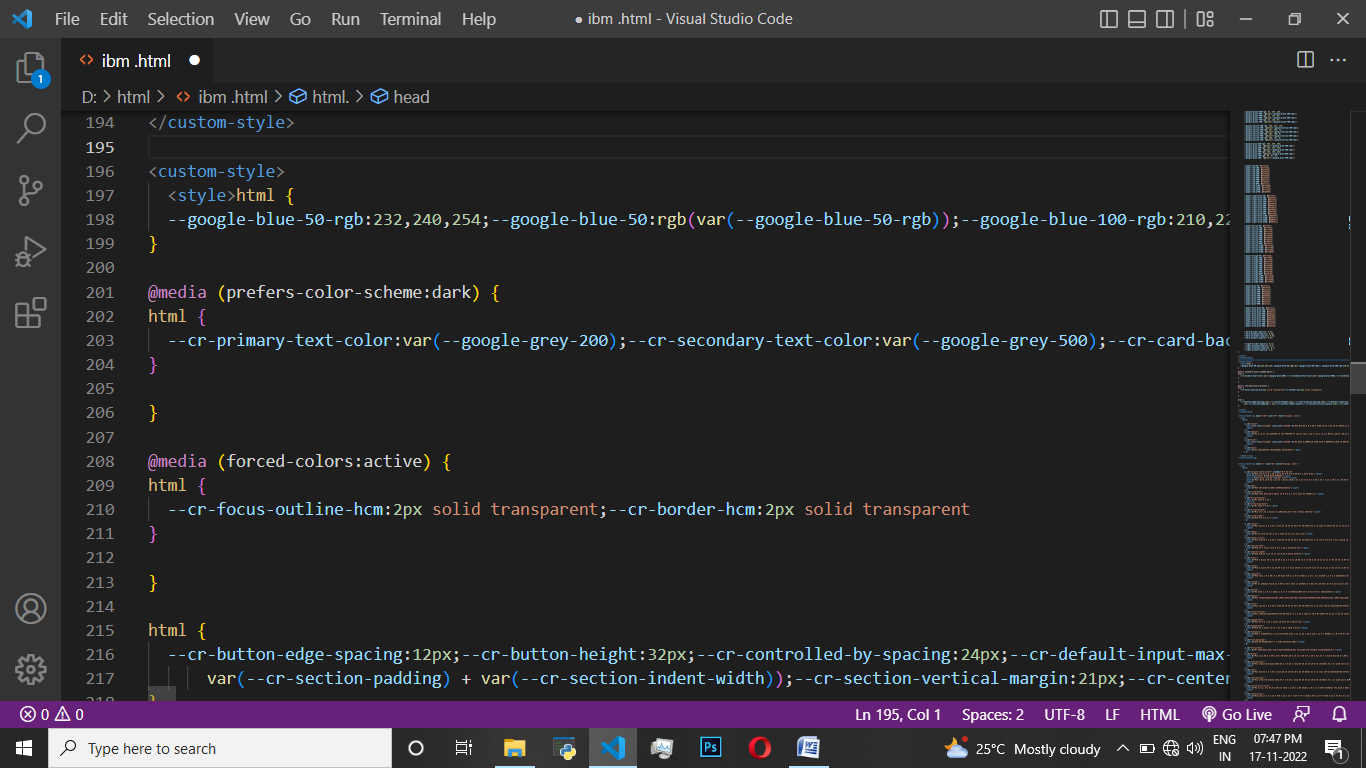


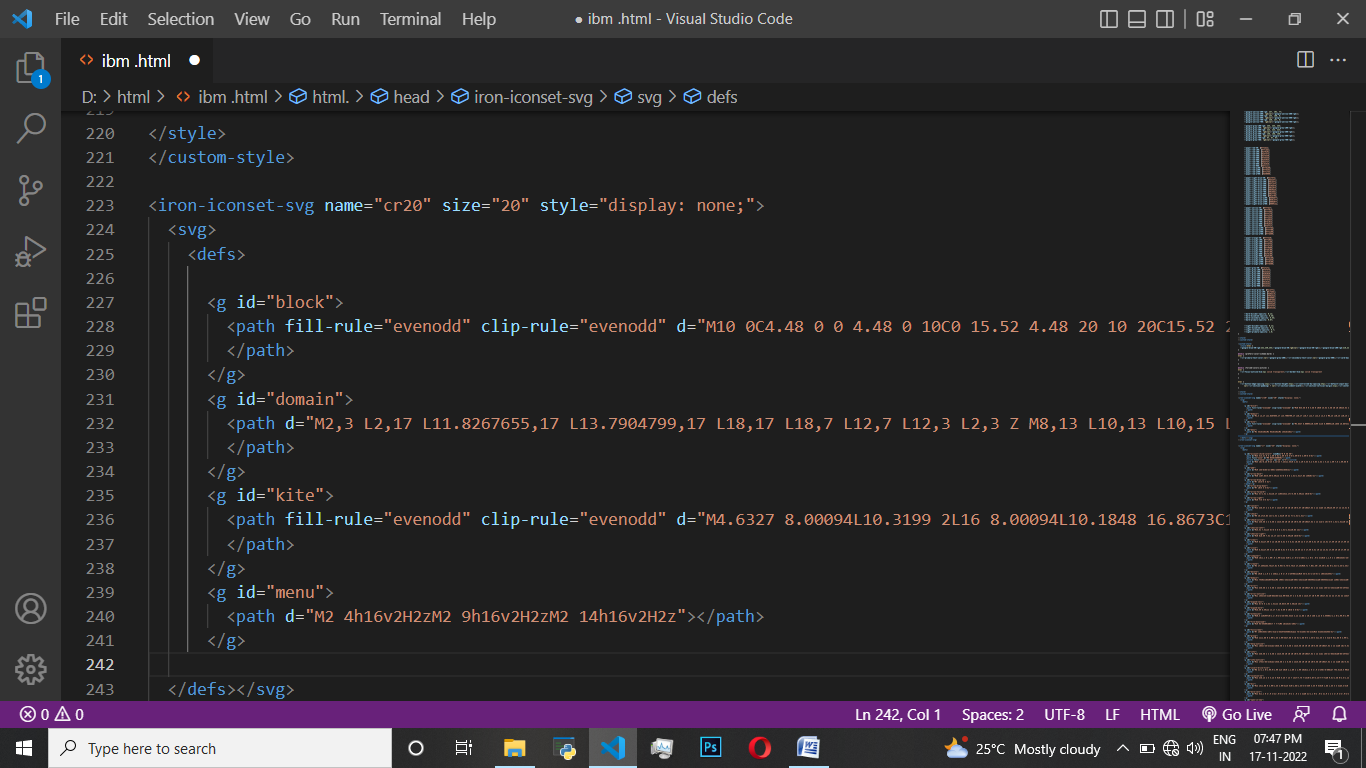


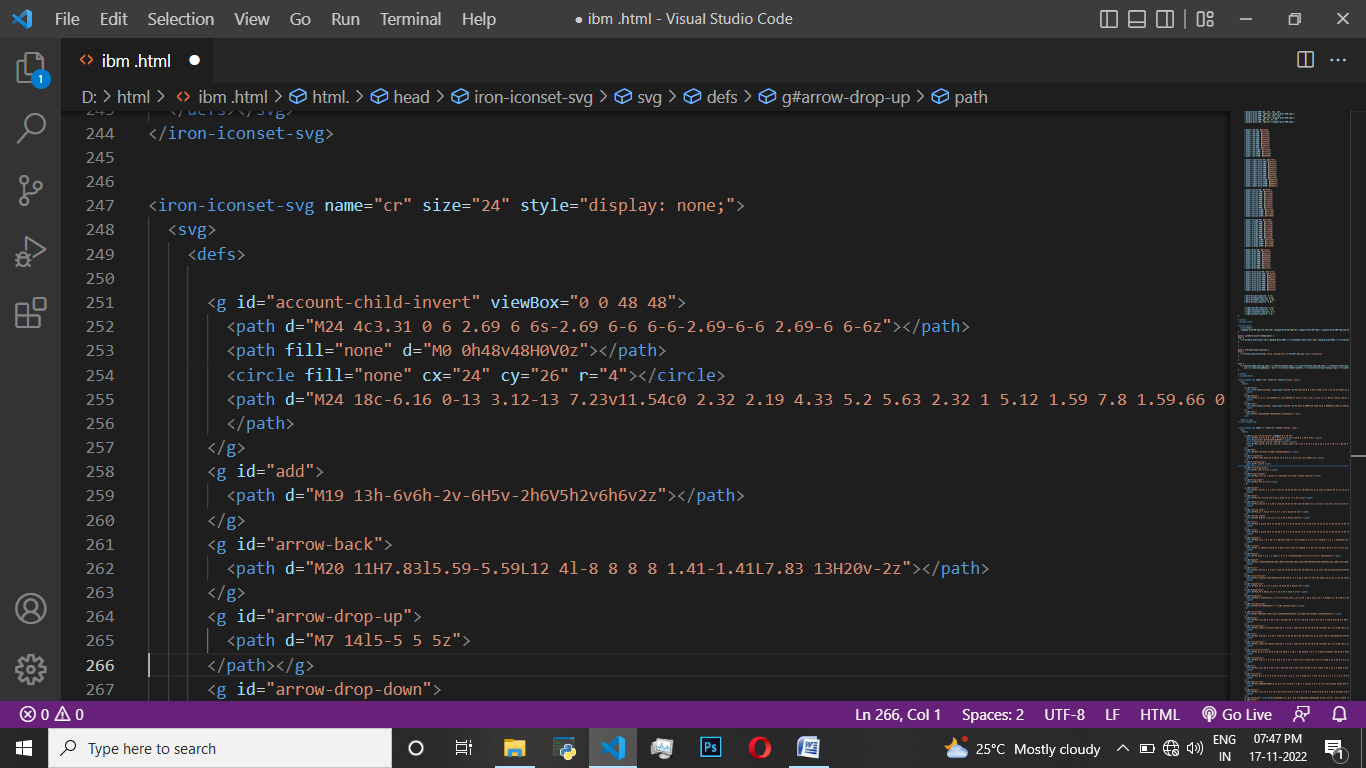


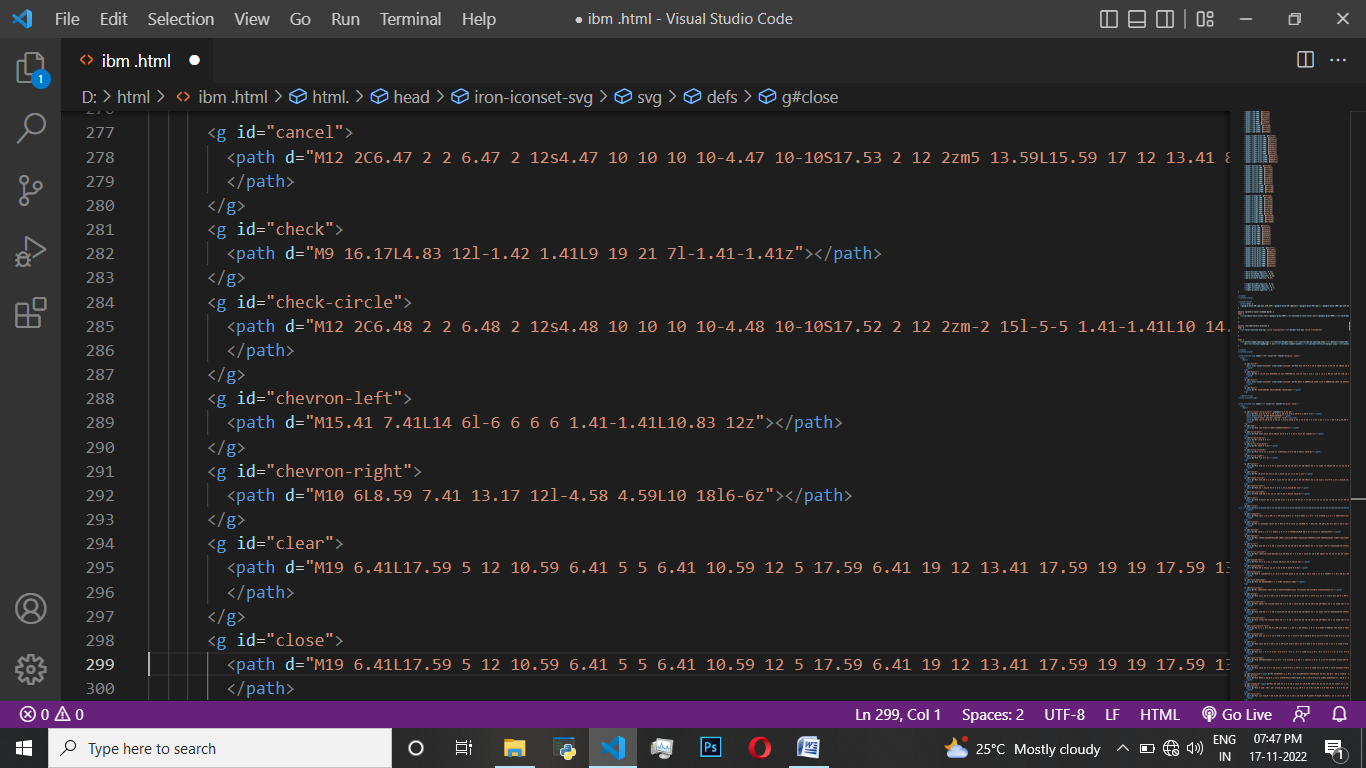


**7.2 FEATURE 2:**









1. **TESTING:**

Software testing is **the process of evaluating and verifying that a software product or application does what it is supposed to do**. The benefits of testing include preventing bugs, reducing development costs and improving performance. Test management plan. Types of software testing.

**8.1 TESTCASE :**

A test case includes information such as **test steps, expected results and data** while a test scenario only includes the functionality to be tested.

**8.2 USER ACCEPTANCE TESTING :**

User Acceptance Testing (UAT), which is performed on most UIT projects, sometimes called beta testing or end-user testing, is **a phase of software development in which the software is tested in the "real world" by the intended audience or business representative**.

**9.RESULT :**

**Performance Metrics**

This dashboard is created to understand a few things like, Customer Analysis and Product Analysis of the Global Super Store. This can be achieved by hearing out to the consumers and collecting their user preference data So that purchasing power will increase and beneficiary for both retailers and consumers.

1. **CONCLUSION:**

By implementing this analytics solution, the company brought their competitive and sales data reporting in-house, cut costs and increased the accuracy of their reporting and analysis. As the company moves forward with this new solution, their sales reporting costs will most likely be reduced by 50 to 70%. They are now able to analyze raw data themselves, respond more quickly to changes in market trends and perform root cause analysis to determine those shifts in the market. By securing quicker access to their data with the new solution, the company was also able to reduce the risk associated with delayed responses to changes in their markets.

With the new solution, the company can now process sales reports faster than the outsourced solution, reducing turnaround time between 50% to 60%. The reporting needs of the company have been streamlined, consolidating over 10 reports into the centralized dashboard solution. The company’s competitive analysis group is also able to more quickly respond to internal data requests given they have the ability to pull the information themselves. With this quicker response, the company is better able to react to changes in the market and predict opportunities for its sales force. The business also experienced an increase in the overall understanding of their sales data throughout the organization. The company now has great flexibility in the presentation of their sales and competitive data, while also being able to integrate sales data with other key data points for the organization

.

**11.ADVANTAGES:**

* Lower marketing costs
* Consistency in brand image
* More revenue and more customer
* Optimum utilization of resources
* Growth and expansion opportunities

**DISADVANTAGS:**

* Differences in consumer needs, wants, usage patterns
* Lack of sales and marketing channel adaptation
* Chances of non-acceptance of product or services
* Non-specification of target markets
* Government restriction

**12.FUTURE SCOPE:**

* Contract manufacturing/ outsourcing
* Integration of economics
* Joint venture and collaboration
* Technical and managerial knowledge
* Foreign agent and distributor

13. APPENDIX

GitHub & Project Demo Link: <https://github.com/IBM-EPBL/IBM-Project-43982-1660720950>