

Corporate Employee Attrition Analytics

A PROJECT REPORT

Submitted By

Team ID : PNT2022TMID32181
Team Leader : SANTHOSH S (732119104065)
Team Member : SHARMILA N (732119104068)
Team Member : SIVASUBRAMANI T (732119104070)
Team Member : SNEGA S (732119104071)

in partial fulfillment for the award of the degree

of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE AND ENGINEERING

NANDHA COLLEGE OF TECHNOLOGY, ERODE

ANNA UNIVERSITY::CHENNAI 600025

JUNE 2022

List Of Contents

- 1. INTRODUCTION**
 - 1.1 Project Overview
 - 1.2 Purpose
- 2. LITERATURE SURVEY**
 - 2.1 Existing problem
 - 2.2 References
 - 2.3 Problem Statement Definition
- 3. IDEATION & PROPOSED SOLUTION**
 - 3.1 Empathy Map Canvas
 - 3.2 Ideation & Brainstorming
 - 3.3 Proposed Solution
 - 3.4 Problem Solution fit
- 4. REQUIREMENT ANALYSIS**
 - 4.1 Functional requirement
 - 4.2 Non-Functional requirements
- 5. PROJECT DESIGN**
 - 5.1 Data Flow Diagrams
 - 5.2 Solution & Technical Architecture
 - 5.3 User Stories
- 6. PROJECT PLANNING & SCHEDULING**
 - 6.1 Sprint Planning & Estimation
 - 6.2 Sprint Delivery Schedule
 - 6.3 Reports from JIRA
- 7. CODING & SOLUTIONING**
 - 7.1 Feature 1
 - 7.2 Feature 2
 - 7.3 Database Schema
- 8. TESTING**
 - 8.1 Test Cases
 - 8.2 User Acceptance Testing
- 9. RESULT**
 - 9.1 Performance Metrics
- 10. ADVANTAGES & DISADVANTAGES**
- 11. CONCLUSION**
- 12. FUTURE SCOPE**
- 13. APPENDIX**
 - Source Code
 - GitHub & Project Demo Link

1. INTRODUCTION

1.1 Project overview

Employee attrition has become a vital problem across the world. It is one of the crucial issues faced by business leaders within companies where they lose the most talented employees. A good employee is always an asset to the organization and their resignation can lead to various problems like financial losses, overall performance, and loss of acquired knowledge. Furthermore, hiring new employees is far exorbitant, taxing, and time-consuming in comparison to recruiting the existing one. It is very time-consuming to recruit a new employee as it takes him months for training, adjusting to the culture, rules, and environment. Therefore, upcoming trends and technology using Machine Learning Algorithms must be exploited for the benefit of business organizations. Knowing the reason beforehand for the employee attrition, companies can mitigate this loss. This analysis provides a conclusive review of employee attrition from the dataset 'IBM HR Analytics Employee Attrition Performance

1.2 Purpose

Hardik P. K. (2016) , researched on “a study on employee attrition: with special reference to Kerala IT Industry”. His research examined the relationship between organizational factors and attrition of IT professionals. The result can conclude that the organizational factors played significant role in predicting the variance in turnover intention (attrition) of Kerala IT professionals. Therefore, the HR managers in IT organizations may take into consideration the problems with organizational factors of their workers to reduce the turnover intention of the skilled employees.

2.LITERATURE SURVEY

2.1 Existing Problem

The Existing system includes only few attributes for analysis and deals with qualitative observations and simple statistical analysis. The qualitative observations deal with data and can be observed through human senses. They do not involve measurements or number. Due to the increase in IOT and connected device, we now have access to so much of data and along with it an increase needs to manage and understand data

2.2 References

From Big Data to Deep Data to support people analytics for employee attrition prediction, Nesrine Ben Yahia, Hlel Jihen, Ricardo Colomo-Palacio(2021)

Machine Learning Approach for Employee Attrition Analysis.Dr. R.

S. Kamath | Dr. S. S. Jamsandekar | Dr. P. G. Naik ,Published in International Journal of Trend in Scientific Research and Development (ijtsrd), (March 2019)

Investigation of early career teacher attrition(ECT) and the impact of induction programs in Western Australia, Janine E.Wyatt, MichaelO'Neill (2021)

EMPLOYEE ATTRITION PREDICTION USING DEEP NEURAL

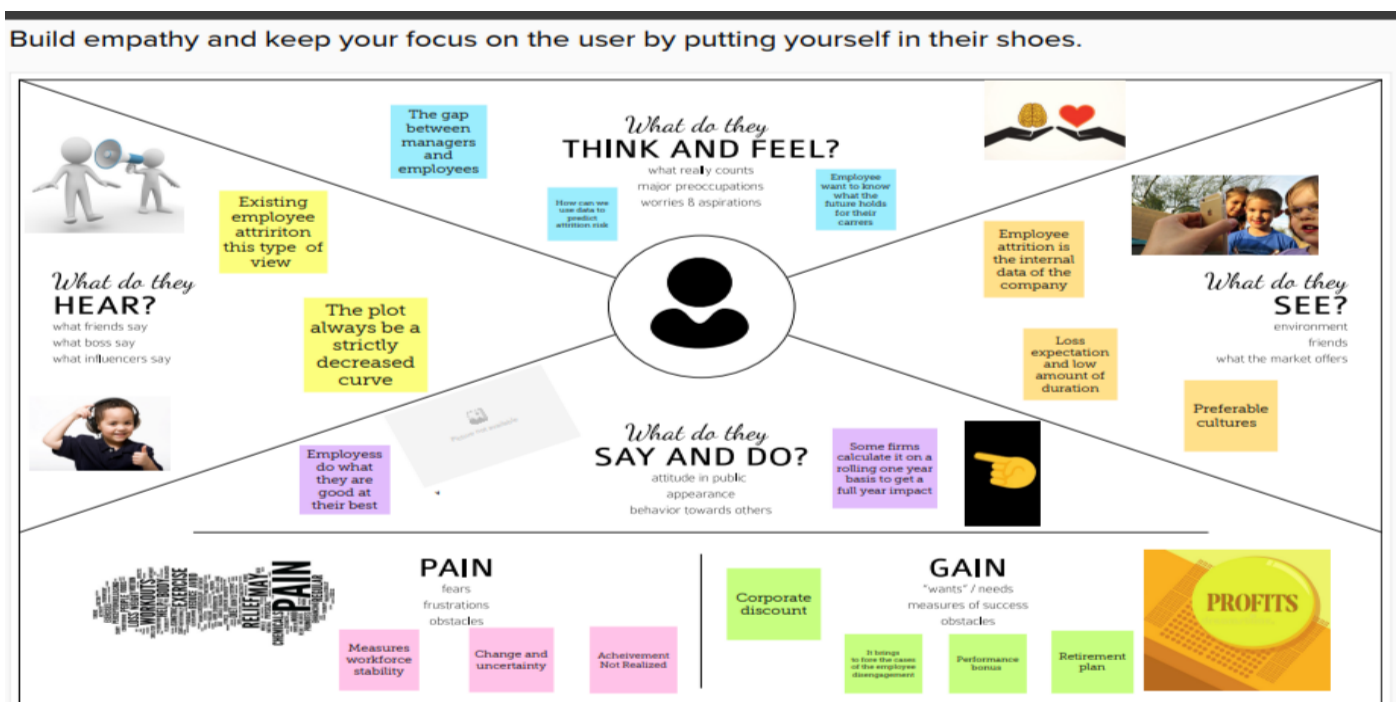
NETWORK, Salah Al-Darraj, Dhafer G. Honi, Francesca Fallucchi, Ayad I. Abdulsada, Romeo Giuliano and Husam A. Abdulmalik,(3 November 2021)

2.3 Problem Statement Definition

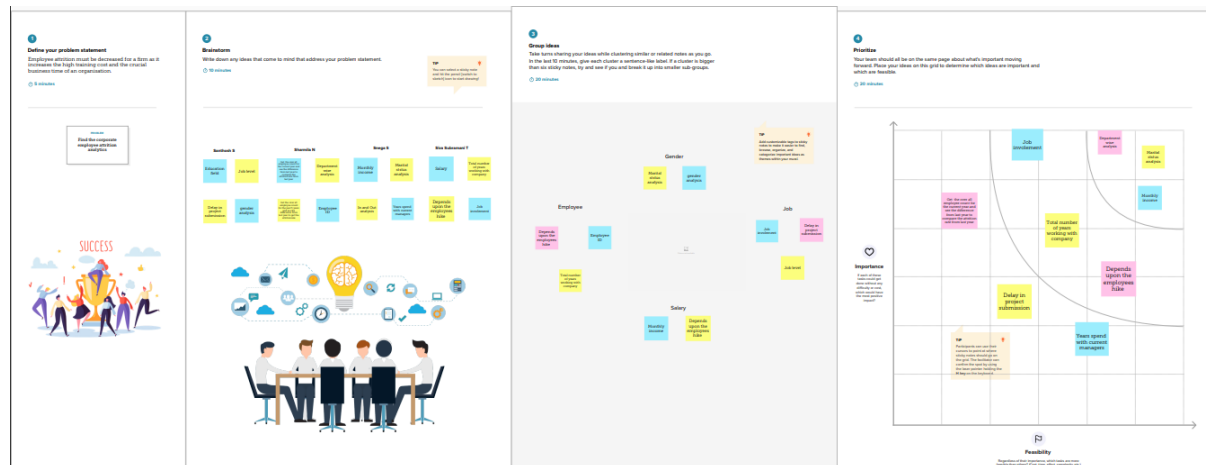
To create a dashboard and perform analysis of employee attrition in corporates using IBM Cognos analytics platform. To reduce the employee attrition rate through data analytics, data visualization by analyzing the major factors that causes attrition.

3. IDEATION AND PROPOSED SOLUTION

3.1 Empathy Map Canvas



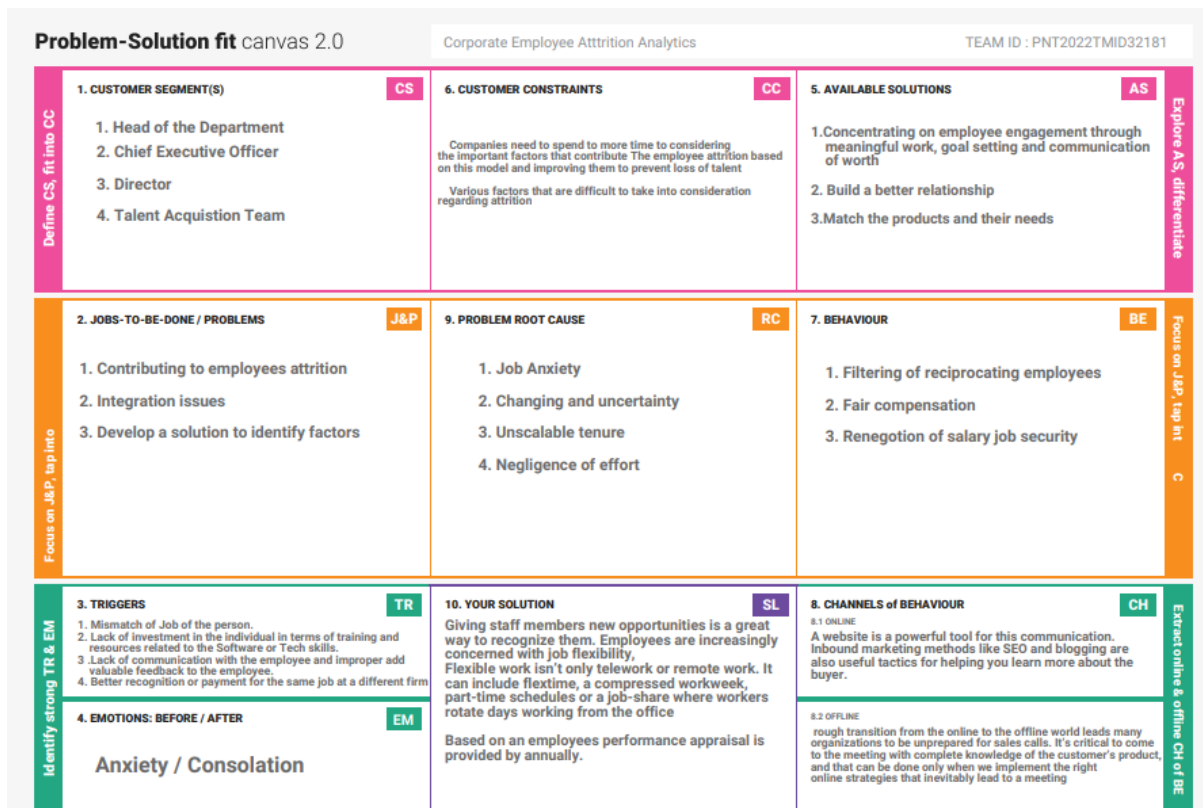
3.2 Ideation & Brainstorming



3.3 Proposed Solution

1.	Problem Statement (Problem to be solved)	<ul style="list-style-type: none"> ➤ It is difficult to identify all the potential employees through visual monitoring by HR. Potential employees leaving company may tend to decrease the company's profit. Layoffs, which occur due to a change or decrease of the business reduce employee morale and make it difficult to hire for other positions. When employees leave by choice, the company can decrease costs while moving forward with other hiring decisions. ➤ For an organisation to be successful it is important to attract and retain top talents. In order to do that, an organisation must determine the factors and the cause of an employee to leave or stay.
2.	Idea / Solution description	<ul style="list-style-type: none"> ➤ By collecting the data set of the corporate employee and using the methods of data analysis like visualizing, we can get some useful insights about the performance of the employees. ➤ Using algorithms to find the factors to analyze and understand the underlying pattern to improve on the factors leading to retention of employees.
3.	Novelty / Uniqueness	<ul style="list-style-type: none"> ➤ It is useful for the HR to identify their employees potential in order to retain them. This also helps the company to gain high profit by retaining good talents. ➤ Right before the employee joins, the analytical system would use the factors like gender, age, work duration in previous companies and etc to categorize whether the employee would stay for long term or short term hence controlling the attrition.
4.	Social Impact / Customer Satisfaction	<ul style="list-style-type: none"> ➤ Employees are of different types. Some employees after joining in the organization tend to be inactive and receive same salary as the one to work sincerely. ➤ This data analysis will be a warning sign to the employees too to work in the respective manner in order to achieve the organization goal.
5.	Business Model (Revenue Model)	<ul style="list-style-type: none"> ➤ The analysis can be performed with minimum cost and at the same time it can provide high performance.
6.	Scalability of the Solution	<ul style="list-style-type: none"> ➤ Predicting the exact result through data analysis helps us to increase the scalability of the solution.

3.4 Problem Solution fit



4.REQUIREMENT ANALYSIS

4.1 Functional requirement

Functional Requirements (Epic)	Sub Requirements (Story & Sub Task)
collect Dataset	Data from different sources are collected in order to get optimized result
Data cleaning	When combining data from multiple sources there are duplicated data and hence we clean the data 1 st
Data modelling	Identify the relationship between Various parameters.
Prediction and analysis	The length of stay is predicted with the Machine learning algorithm

4.2 Non-Functional requirements

Non-Functional Requirement	Description
Usability	User can view and visualize the data through the interactive dashboard and predict the length of stay of patients with machine learning algorithm
Security	IBM Cognos provides better security. The dataset uploaded to the dashboard cannot be downloaded or accessed by external sources
Reliability	The dashboard and the prediction is very reliable and provide prediction with more accuracy
Performance	The length of stay of patients is predicted with more accuracy
Availability	The predicted length of stay and the visualization will be available in Cognos analysis
Scalability	The software is scalable and extendable. Because it allow multiple user to handle the data at the same time

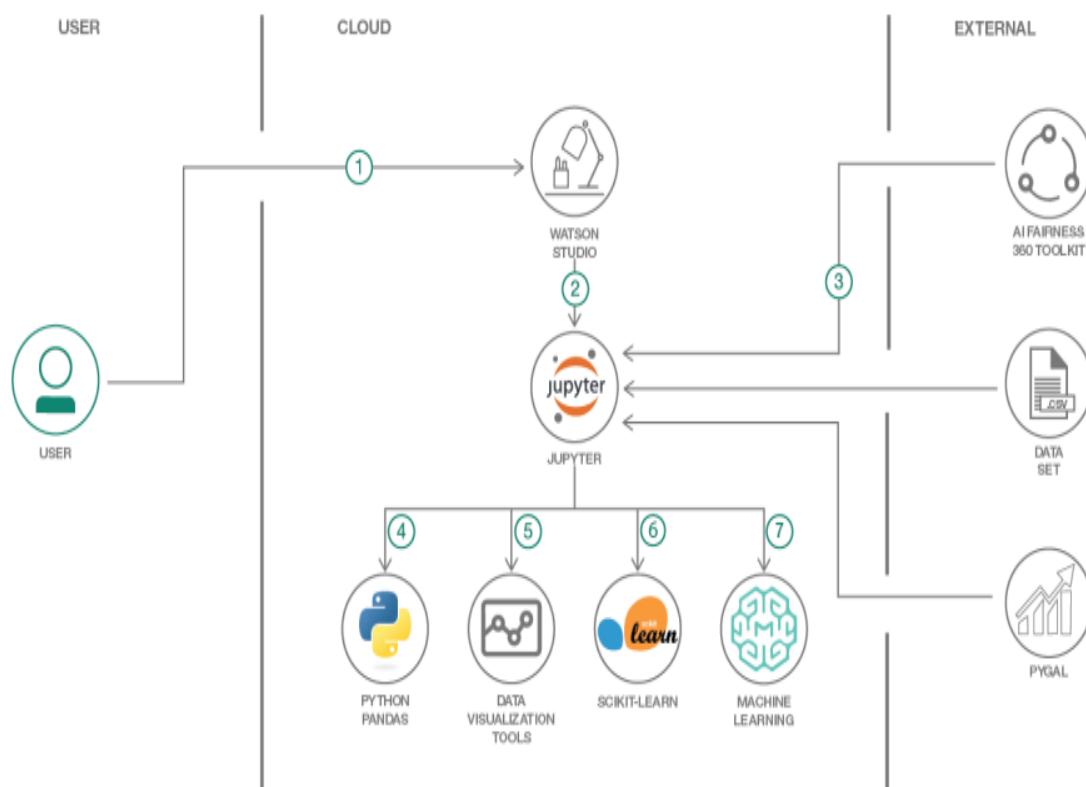
5. PROJECT DESIGN

5.1 Data Flow Diagrams

A two-dimensional diagram explains how data is processed and transferred in a system. The graphical depiction identifies each source of data and how it interacts with other data sources to reach a common output. Individuals seeking to draft a data flow diagram must identify external inputs and outputs, determine how the inputs and outputs relate to each other, and explain with graphics how these connections relate and what they result in. This type of diagram helps business development and design teams visualize how data is processed and identify or improve certain aspects.

Data Flow Diagrams:

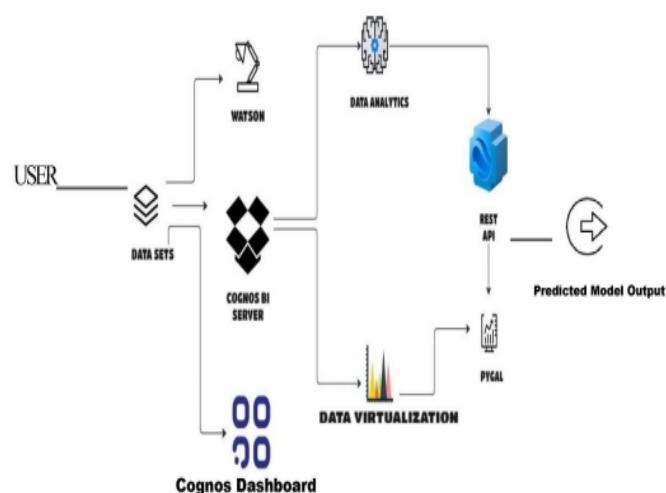
A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.



5.2 Solution & Technical Architecture

A system architecture or systems architecture is the conceptual model that defines the structure, behavior, and more views of a system. An architecture description is a formal description and representation of a system, organized in a way that supports reasoning about the structures and behaviors of the system. System architecture can comprise system components, the externally visible properties of those components, the relationships (e.g. the behavior) between them. It can provide a plan from which products can be procured, and systems developed, that will work together to implement the overall system. There have been efforts to formalize languages to describe system architecture, collectively these are called architecture description languages (ADLs).

Technical Architecture:



- As shortly after beginning their jobs, staffers register in the institution's database.
- After a specified period of time, the employees leaving from the organization in which uses machine learning model like Decision Tree, Random Forest model, K-Nearest Neighbour training accuracy, etc. to the training the dataset to get the accuracy by predicting the value.
- User interacts with the application using website UI HTML, CSS, JavaScript, React Js etc.
- This logic depends on the extracting the needed contents into the dataset using Python.

5.3 User Stories

User Stories

Use the below template to list all the user stories for the product.

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (Web user)	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	I can access my account / dashboard	High	Sprint-1
		USN-2	As a user, I will receive confirmation email once I have registered for the application	I can receive confirmation email & click confirm	High	Sprint-1
		USN-3	As a user, I can register for the application through Gmail	I can register & access the dashboard with Gmail Login	Medium	Sprint-1
	Login	USN-4	As a user, I can log into the application by entering email & password	I can access my account / dashboard	High	Sprint-1
	Dashboard	USN-5	Uploading the Dataset	I can be able to upload my dataset	High	Sprint 2
		USN-6	Working With Dataset	I can be able to access my dashboard	High	Sprint 2
		USN-7	Visualization	I can be able to view the visual attrition rate of my dataset	High	Sprint 3
		USN-8	Working with Dashboard	I can be able to view the various views of the attrition rate	High	Sprint 3
Customer Care Executive		USN-9	Asking Help / Feedback	I can be able to ask help if I can face any issues or problems while using the webpage	Medium	Sprint 4
Administrator		USN-10	Managing the Database	I can assure that my data is in secure state	High	Sprint 4
		USN-11	Managing the over all process	I can assure that my data and process is going good	High	Sprint 4

6 PROJECT PLANNING

6.1 Sprint Planning & Estimation

Product Backlog, Sprint Schedule, and Estimation (4 Marks)

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Datasets	USN-1	As a user, I can enter the details of the employees working in our organization for the attrition detail.	5	High	Sharmila N, Sneha S, Santhosh S
Sprint-1		USN-2	As an Analyst, I Prepare the data & Provide meaningful insights through EDA in Cognos Analytics	3	High	Sneha S, Sivasubramani T, Santhosh S
Sprint-2	Exploring data and creating model	USN-3	As a user, I want to find connections between various visualization that leads to attrition	2	Low	Sivasubramani T, Santhosh S
Sprint-2		USN-4	As an Analyst, I will create a prediction model for predicting the attrition.	3	Medium	Santhosh S
Sprint-3	Prediction	USN-5	As an Analyst, I will create different type of model to identify which give the correct prediction.	3	Medium	Sharmila N, Santhosh S, Sivasubramani T
Sprint-3		USN-6	As an Analyst, I will use Cognos Analytics to generate a report	3	Medium	Sneha S, Sharmila N, Santhosh S
Sprint-4	Dashboard	USN-7	As a user, I can only understand the Analysis in animated presentation of dataset	5	Medium	Sivasubramani T, Santhosh S
Sprint-4		USN-8	As an Analyst, I use Cognos Analytics to create an animated presentation (Story) of the dataset	3	High	Santhosh S

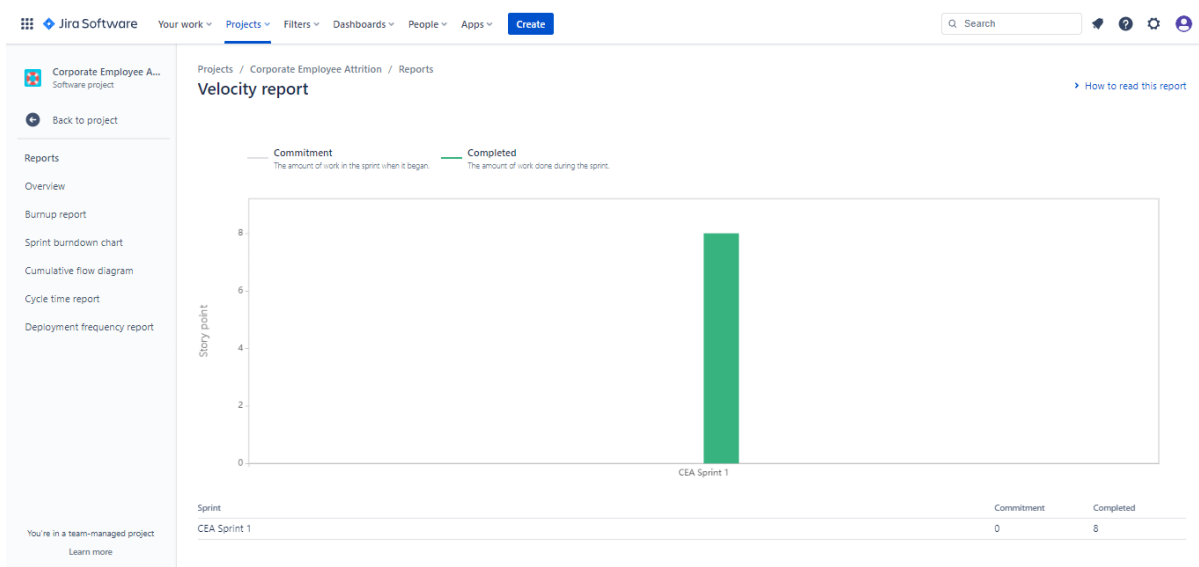
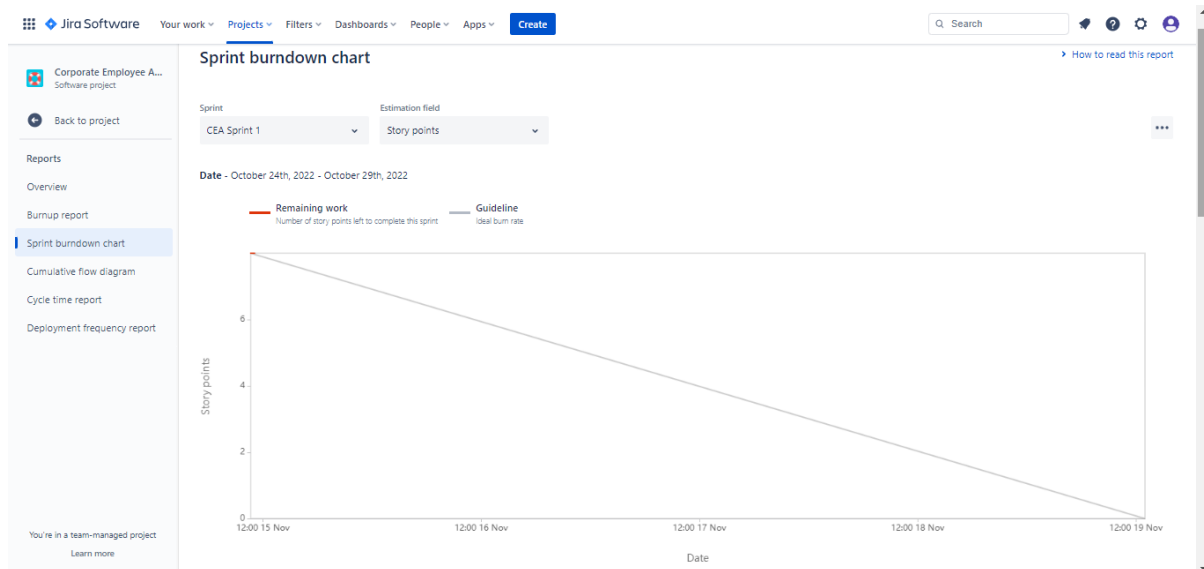
6.2 Sprint Delivery Schedule

Project Tracker, Velocity & Burndown Chart: (4 Marks)

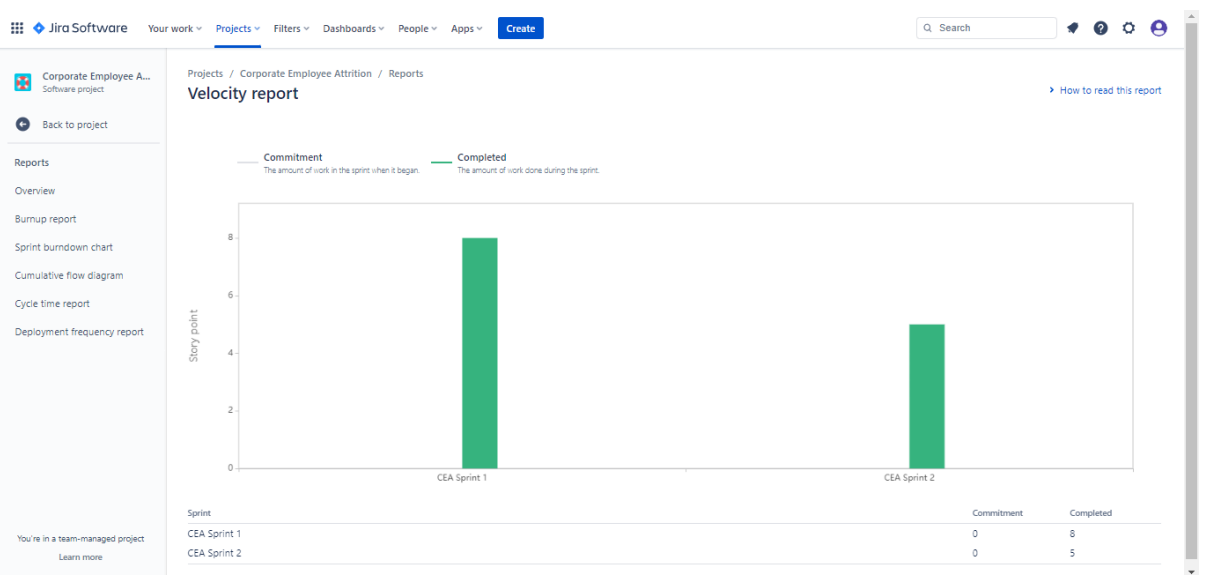
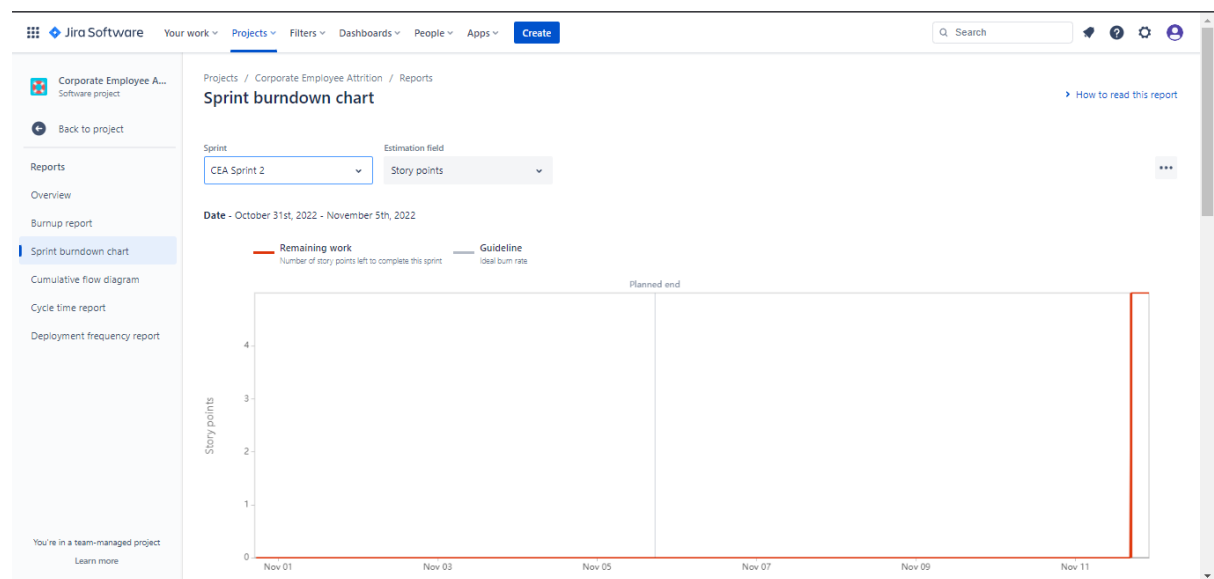
Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint-1	5	6 Days	24 Oct 2022	29 Oct 2022	5	29 Oct 2022
Sprint-2	5	6 Days	31 Oct 2022	05 Nov 2022	5	05 Nov 2022
Sprint-3	5	6 Days	07 Nov 2022	12 Nov 2022	5	12 Nov 2022
Sprint-4	5	6 Days	14 Nov 2022	19 Nov 2022	5	15 Nov 2022

6.3 Reports from JIRA

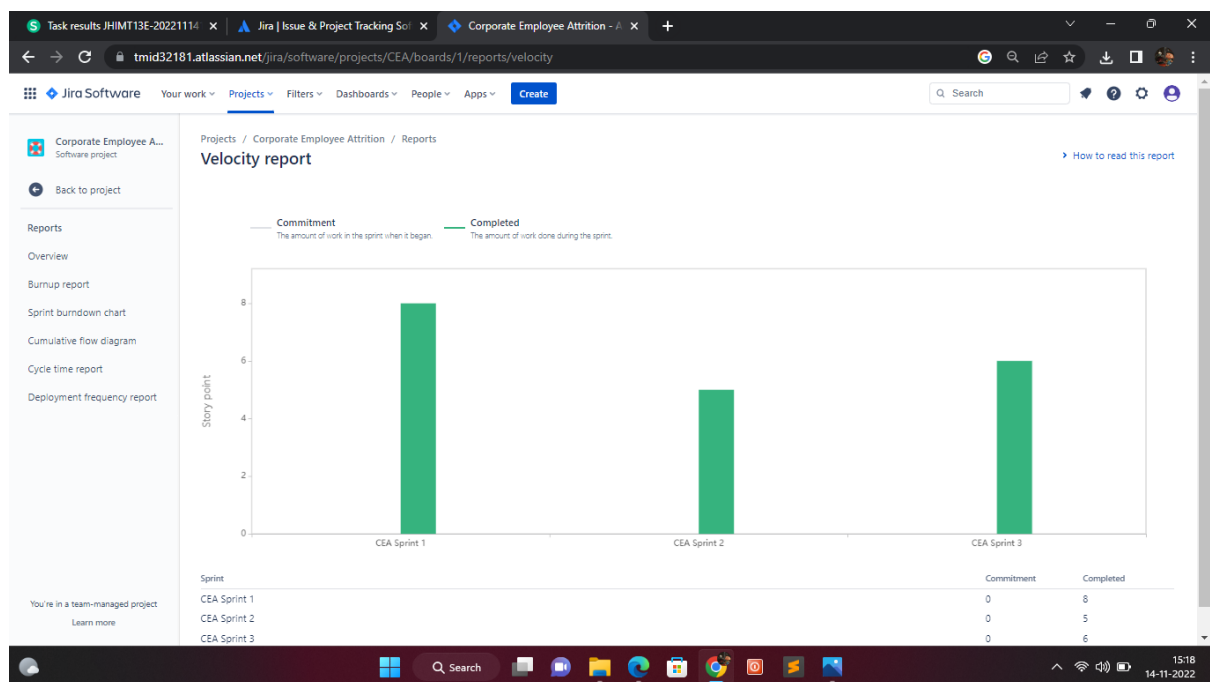
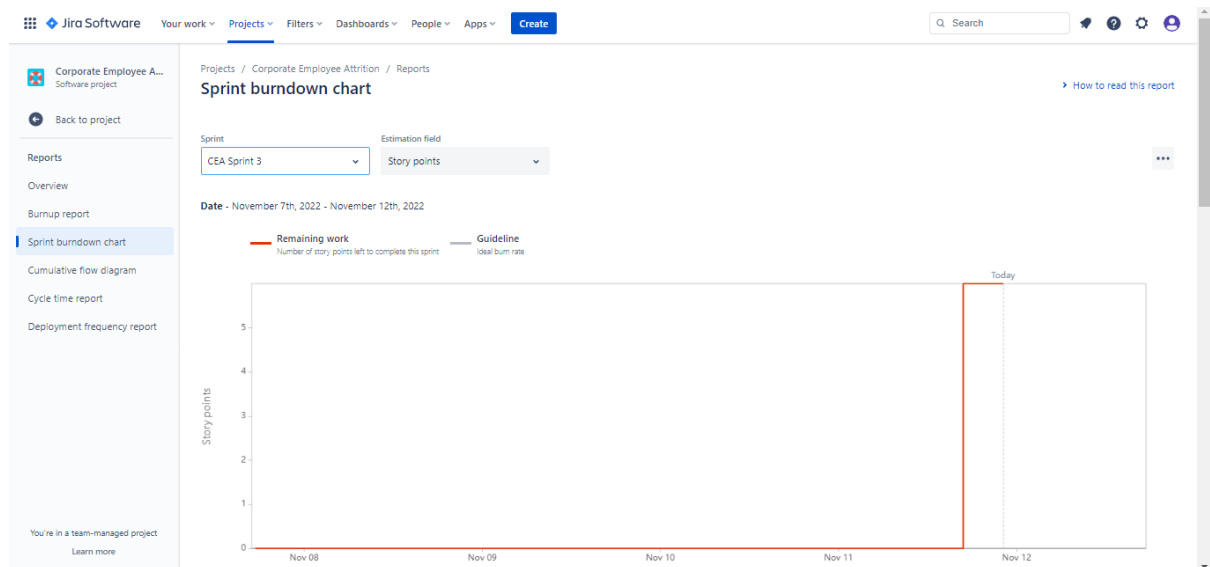
Sprint 1:



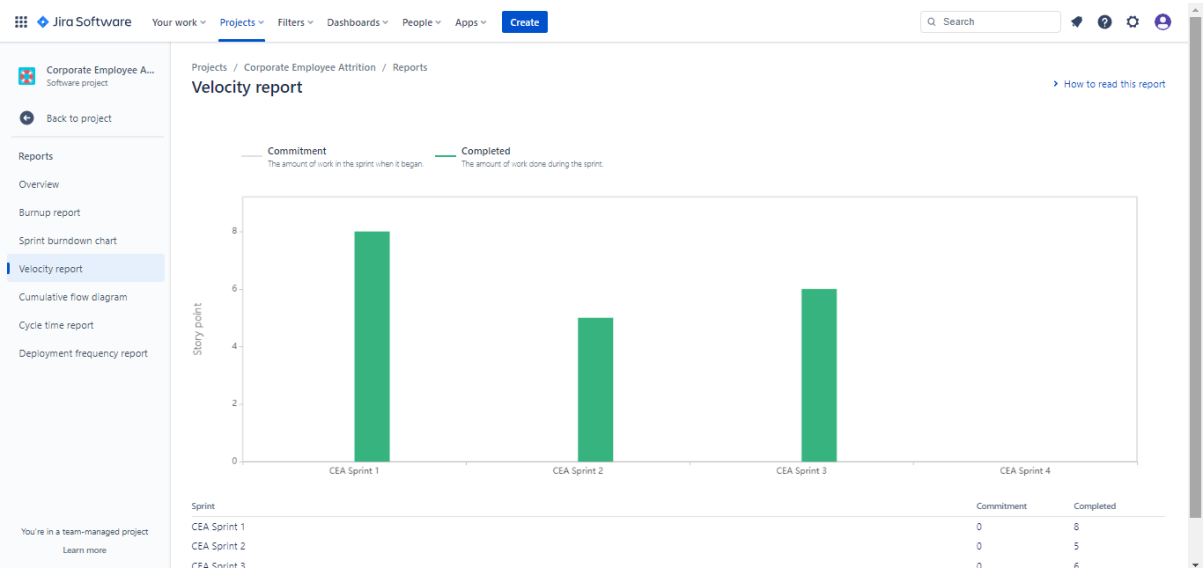
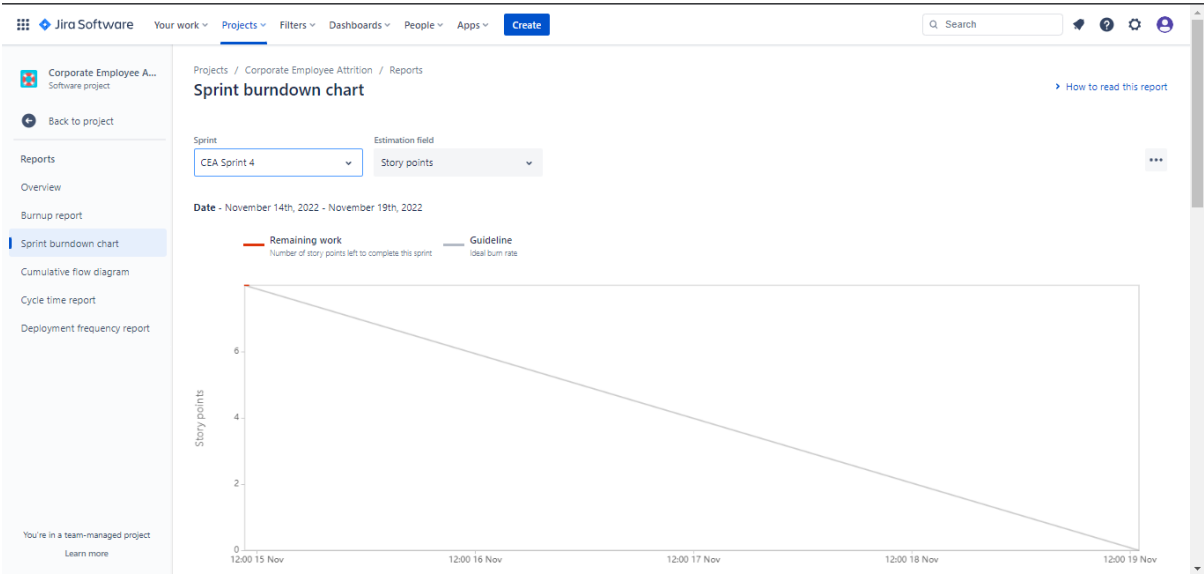
Sprint 2:














Sprint 3 :



Sprint 4:

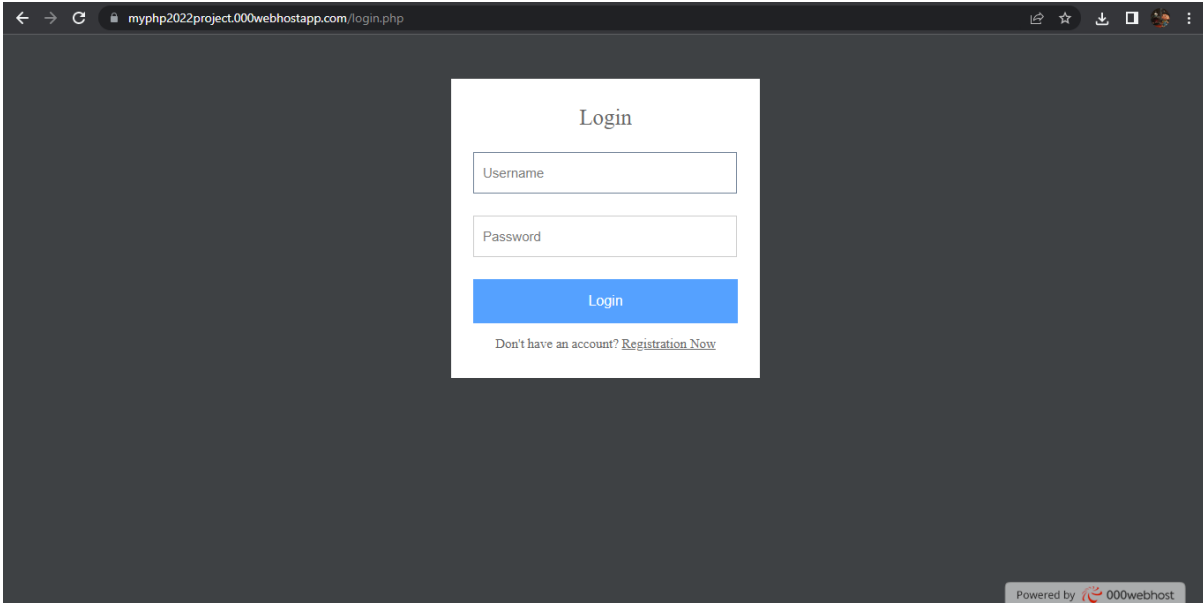


Road Map :

	OCT – DEC	JAN – MAR
Sprints		
>  <u>CEA-2 Datasets</u>	 	
>  <u>CEA-10 Exploring data and creating model</u>		
>  <u>CEA-11 Prediction</u>		
 CEA-12 Prediction		
>  <u>CEA-13 Dashboard</u>		

7. CODING & SOLUTIONING (Explain the features added in the project along with code)

7.1 Feature 1



← → ↻ myphp2022project.000webhostapp.com/login.php

Login

Login

Don't have an account? [Registration Now](#)

Powered by 000webhost

7.2 Feature 2

myphp2022project.000webhostapp.com/registration.php

Registration

Username

Email Address

Password

Register

Already have an account? [Login here](#)


Powered by 000webhost

myphp2022project.000webhostapp.com/index.html

Corporate Employee Attrition Analytics

DashBoard1 DashBoard2 DashBoard3

Corporate Employee Attrition Analytics

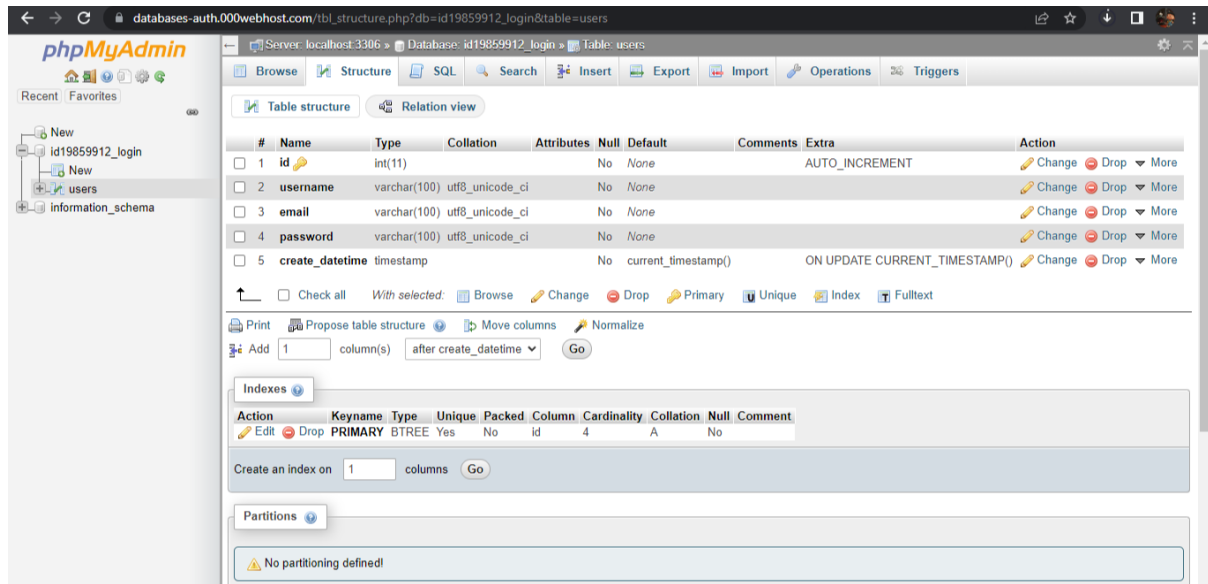
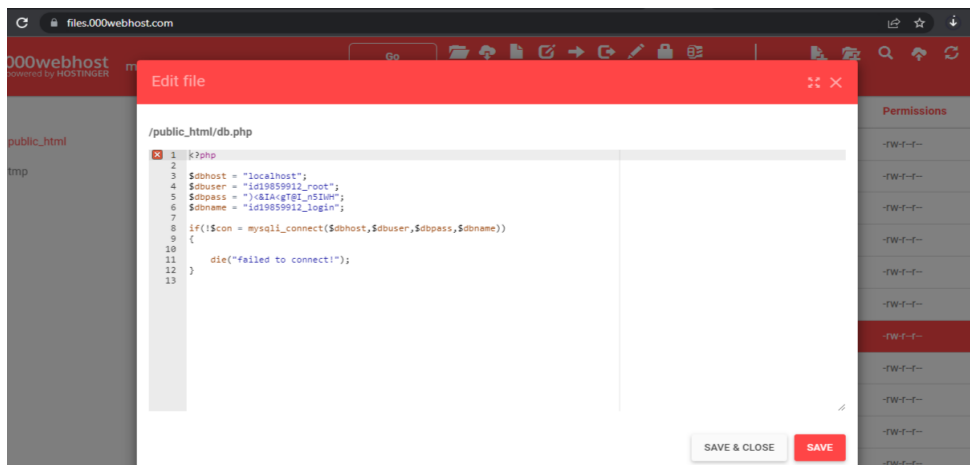


732119104065 - Santhosh S
732119104068 - Sharmila N
732119104070 - Sivasubramani T
732119104071 - Snega S

Team Id : PNT2022TMID32181

Powered by 000webhost

7.3 Database Schema



8. TESTING

8.1 Test Cases

A Test cases is a set of actions performed on a system to determine if it satisfies software requirements and functions correctly. The test cases are the following: usability testing, Functionality Testing, Compatibility testing, Database testing, Security testing, Performed Testing.

S.No.	Parameter	Screenshot / Values
1.	Dashboard design	No of Visualizations /Graphs – 7 – 8 visualization /7 – 8 Graphs
2.	Data Responsiveness	Users and Analyst or Developers
3.	Amount Data to Rendered (DB2 Metrics)	Data(200MB)
4.	Utilization of Data Filters	Simple Mathematics and distribute the datasets
5.	Effective User Story	No of Scene Added – 7 stories
6.	Descriptive Reports	No of Visualizations / Graphs – 8 Graphs/ 3 - graph

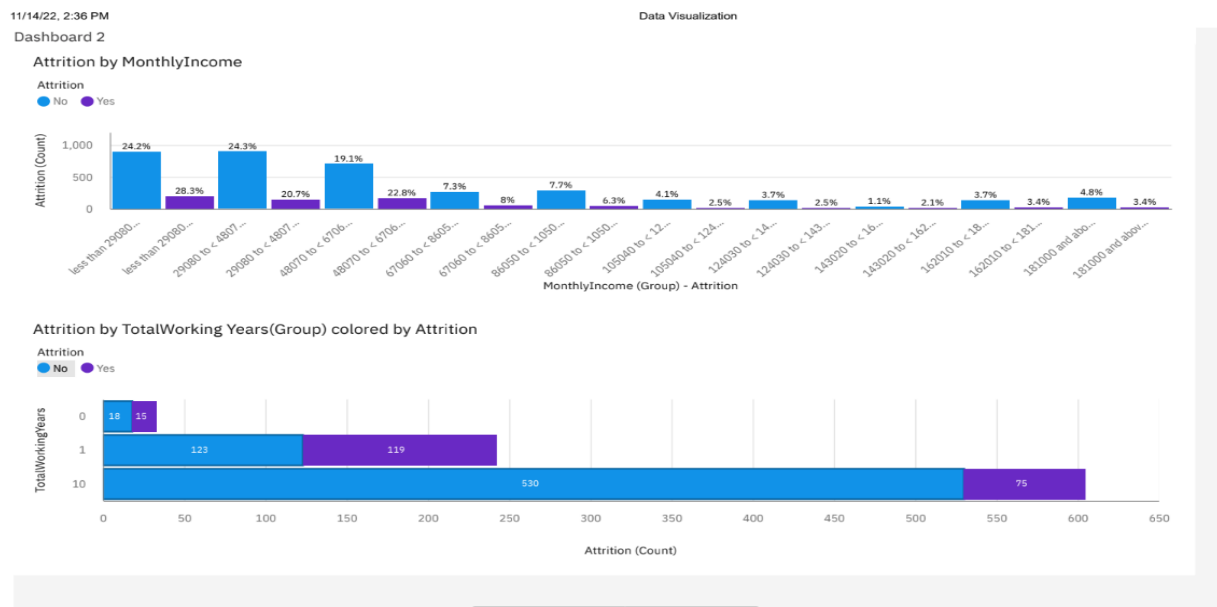
8.2 User Acceptance Testing

User Acceptance Testing (UAT), also called application testing or end-user-testing, is a phase of software development in which the software is tested in the real world by its intended audience. The purpose of this document is to briefly explain the test coverage and open issues of the [Corporate Employee Attrition Analytics] project at the time of the release to User Acceptance Testing (UAT).

Resolution	Severity 1	Severity 2	Severity 3	Severity 4	Subtotal
By Design	9	3	2	3	17
Duplicate	1	0	3	0	4
External	2	3	0	1	6
Fixed	10	2	4	20	36
Not Reproduced	0	0	1	0	1
Skipped	0	0	1	1	2
Won't Fix	0	5	2	1	8
Totals	22	13	13	26	74

9. RESULTS

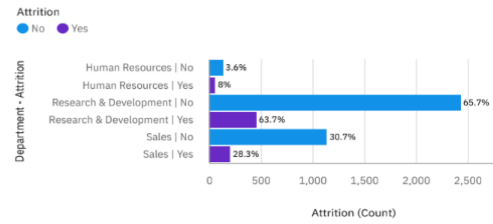
9.1 Performance Metrics



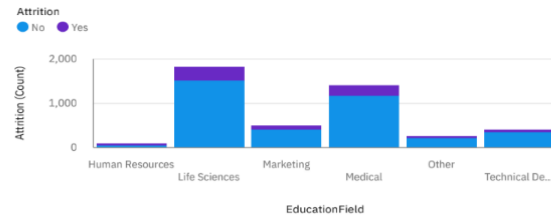
11/14/22, 2:36 PM

Dashboard 3

Attrition by Department

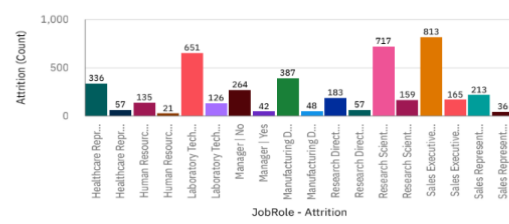


Attrition by Education Field

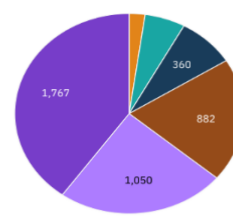


Data Visualization

Attrition by JobRole



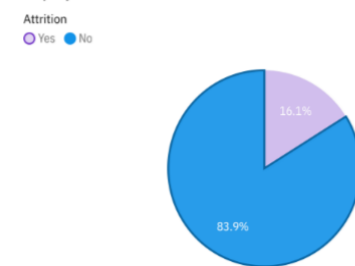
Attrition by Marital Status



11/14/22, 2:36 PM

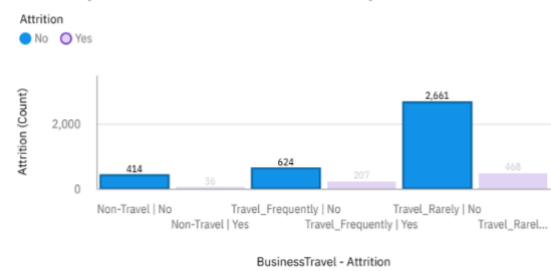
Dashboard 1

EmployeeAttrition



Data Visualization

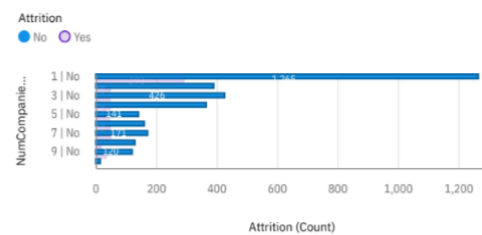
Attrition by BusinessTravel and Attrition colored by Attrition

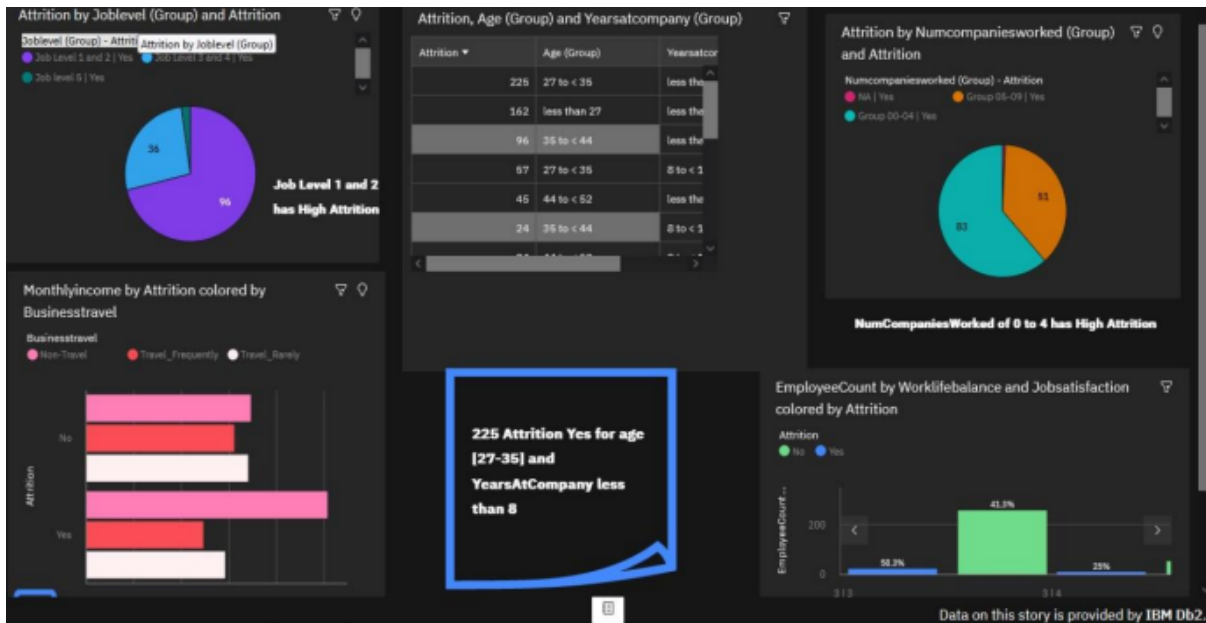


Based On Salary Hike Percent

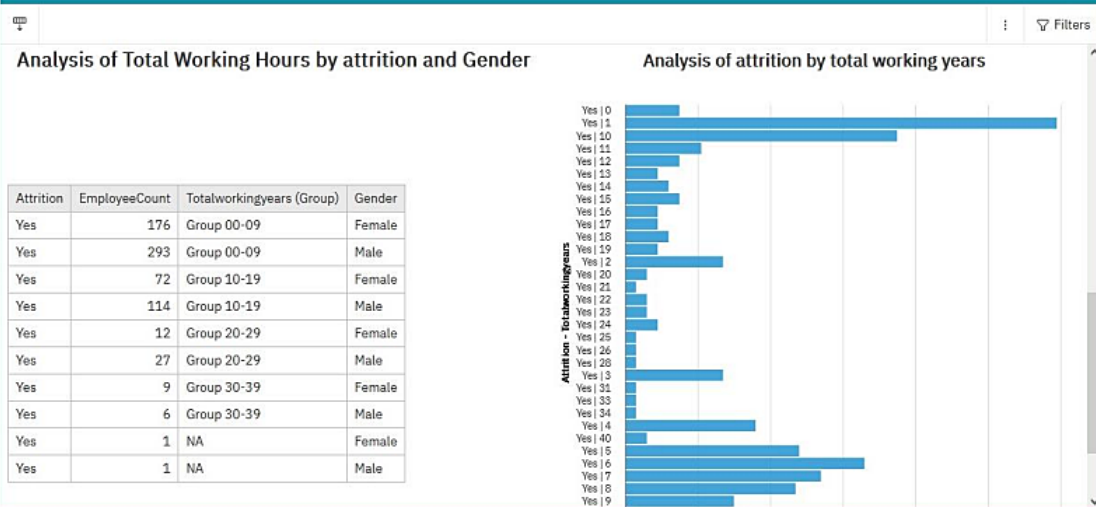


Attrition by NumCompaniesWorked and Attrition colored by Attrition





Report



10.ADVANTAGES & DISADVANTAGES

Advantages

1. The metrics offer accuracy in terms of the reasons given by employees themselves
2. It brings to fore the cause of employee disengagement.
3. Enables HR managers develop long term strategies to reduce attrition.
4. Develops & Shapes drills that benefit both the management and the employee.
5. Enhanced work culture.

Disadvantages

1. The might be radiation al reduction in the strength of the workforce.The employees might loss interest and sincerity.
2. The remaining employees might face an increase in the workload and feels stressed about it.

11.CONCLUSION

The may fail to perform at their job. This eventually leads to attrition. Organizations should state the requirements and expectations Employees as well as organizations must be clear with their expectations regarding unambiguously. This helps candidates decide upon to accept the job position or not. This eventually avoids further conflicts in the employment terms.job profile. Any sort of mismatch leads to discrepancy and employees.

.

12. FUTURE SCOPE

Research findings suggest that attrition reasons in IT organizations primarily revolve around professional growth and challenges in the organization. Although economic factors happen to the most influential factor, professionals may settle for second best criteria of their preference that is career growth and supportive work policies in the organization. On the other hand, candidates who aspire to have a better job than the one in hand are more interested in securing the next job. Young talent wants to work on latest technology and functional domain. professionals who are young career makers are less Research findings suggest that attrition reasons in IT organizations primarily revolve around professional growth and challenges in the organization. Although economic factors happen to the most influential factor, professionals may settle for second best criteria of their preference that is career growth and supportive work policies in the organization. On the other hand, candidates who aspire to have a better job than the one in hand are more interested in securing the next job. Young talent wants to work on latest technology and functional domain. IT professionals who are young career makers are less.

13. APPENDIX

Source Code :

index.php

```
1 <?php
2 header("Location: login.php");
3 ?>
```

login.php

```
1 <!DOCTYPE html>
2 <html>
3 <head>
4   <meta charset="utf-8"/>
5   <title>Login</title>
6   <link rel="stylesheet" href="style.css"/>
7 </head>
8 <body>
9 <?php
10  require('db.php');
11  session_start();
12  // When form submitted, check and create user session.
13  if (isset($_POST['username'])) {
14      $username = stripslashes($_REQUEST['username']); // removes backslashes
15      $username = mysqli_real_escape_string($con, $username);
16      $password = stripslashes($_REQUEST['password']);
17      $password = mysqli_real_escape_string($con, $password);
18      // Check user is exist in the database
19      $query  = "SELECT * FROM `users` WHERE username='$username'
20                AND password='" . md5($password) . "'";
21      $result = mysqli_query($con, $query) or die(mysqli_error());
22      $rows = mysqli_num_rows($result);
```

```

23     if ($rows == 1) {
24         $_SESSION['username'] = $username;
25         // Redirect to user dashboard page
26         header("Location: dashboard.php");
27     } else {
28         echo "<div class='form'>
29             <h3>Incorrect Username/password.</h3><br/>
30             <p class='link'>Click here to <a href='login.php'>Login</a> again.</p>
31             </div>";
32     }
33 } else {
34 ?>
35 <form class="form" method="post" name="login">
36     <h1 class="login-title">Login</h1>
37     <input type="text" class="login-input" name="username"
38         placeholder="Username" autofocus="true"/>
39     <input type="password" class="login-input" name="password"
40         placeholder="Password"/>
41     <input type="submit" value="Login" name="submit" class="login-button"/>
42     <p class="link">Don't have an account? <a href="registration.php">Registration
43         </a></p>
44 </form>
45 <?php
46 }
47 ?>
48 </body>
49 </html>

```

logout.php

```

1 <?php
2     session_start();
3     // Destroy session

```

```

4     if(session_destroy()) {
5         // Redirecting To Home Page
6         header("Location: login.php");
7     }
8     ?>

```

registration.php

```

1  <!DOCTYPE html>
2  <html>
3  <head>
4      <meta charset="utf-8"/>
5      <title>Registration</title>
6      <link rel="stylesheet" href="style.css"/>
7  </head>
8  <body>
9  <?php
10     require('db.php');
11     // When form submitted, insert values into the database.
12     if (isset($_REQUEST['username'])) {
13         // removes backslashes
14         $username = stripslashes($_REQUEST['username']);
15         //escapes special characters in a string
16         $username = mysqli_real_escape_string($con, $username);
17         $email    = stripslashes($_REQUEST['email']);
18         $email    = mysqli_real_escape_string($con, $email);
19         $password = stripslashes($_REQUEST['password']);
20         $password = mysqli_real_escape_string($con, $password);
21         $create_datetime = date("Y-m-d H:i:s");
22         $query    = "INSERT into `users` (username, password, email, create_datetime)
23             VALUES ('$username', '" . md5($password) . "', '$email',
24             '$create_datetime')";
25         $result    = mysqli_query($con, $query);
26         if ($result) {
27             echo "<div class='form'>
28                 <h3>You are registered successfully.</h3><br/>
29                 <p class='link'>Click here to <a href='login.php'>Login</a></p>
30             </div>";

```

```

30     } else {
31         echo "<div class='form'>
32             <h3>Required fields are missing.</h3><br/>
33             <p class='link'>Click here to <a href='registration.php'>registration</a>
again.</p>
34             </div>";
35     }
36 } else {
37 ?>
38 <form class="form" action="" method="post">
39     <h1 class="login-title">Registration</h1>
40     <input type="text" class="login-input" name="username"
placeholder="Username" required />
41     <input type="text" class="login-input" name="email" placeholder="Email
Address">
42     <input type="password" class="login-input" name="password"
placeholder="Password">
43     <input type="submit" name="submit" value="Register" class="login-button">
44     <p class="link">Already have an account? <a href="login.php">Login
here</a></p>
45 </form>
46 <?php
47     }
48 ?>
49 </body>
50 </html>

```

db.php

```

1 <?php
2 $dbhost = "localhost";
3 $dbuser = "root";
4 $dbpass = "";
5 $dbname = "loginsystem";
6 if(!$con = mysqli_connect($dbhost,$dbuser,$dbpass,$dbname))
7 {
8     d
9 }

```

DashBoard.php

```
1  <?php
2  //include auth_session.php file on all user panel pages
3  include("auth_session.php");
4  ?>
5  <!DOCTYPE html>
6  <html>
7  <head>
8      <meta charset="utf-8">
9      <title>Dashboard - Client area</title>
10     <link rel="stylesheet" href="style.css" />
11 </head>
12 <body>
13     <div class="form">
14         <p><b><i>Hey,</b></i> <?php echo $_SESSION['username']; ?>!</p>
15         <p><b><i>You are in user dashboard page.</b></i></p>
16         <form method="POST" action="indexx.php">
17             <center><input type="submit" value="DashBoard" name="submit"
18             /></center>
19         </form>
20         <p><a href="logout.php">Logout</a></p></i>
21     </div>
22 </body>
23 </html>
```

GitHub & Project Demo Link :

GitHub : <https://github.com/IBM-EPBL/IBM-Project-12544-1659453304>

Project Demo Link : <https://clipchamp.com/watch/ArJJqbx2qIf>