A New Hint to Transportation-Analysis of the NYC Bike Share System

Project Report

Submitted by

Team ID	**:**	**PNT2022TMID32188**
Team Leader	**:**	**Suganth J**
Team Member 1	**:**	**Boobesh K**
Team Member 2	**:**	**Dhatchayan G**
Team Member 3	**:**	**Nafreen Banu N**

in partial fulfillment for the award of the degree

of

BACHELOR OF ENGINEER

in

COMPUTER SCIENCE AND ENGNEERING

NANDHA COLLEGE OF TECHNOLOGY, ERODE

ANNA UNIVERSITY :: CHENNAI 600025

JUNE 2022

List of contents

1.Introduction:

1.1 Project Overview

The goal of this analysis is to create an operating report of Citi Bike .Let us create data visualizations to understand

1.Total Number of Trips

2.What is Customer and subscriber with gender

3.Find the top bike used with respect to trip duration?

4.Calculating the number of bikes used by respective age groups.

5.Top 10 Start Station Names with respect to Customer age group

1.2 Purpose

By the end of this project, you will:

- Know fundamental concepts and can work on IBM Cognos Analytics.
- Gain a broad understanding of plotting different graphs.
- Able to create meaningful dashboards

2.Literature Survey

2.1 Existing problem

Bike share programs have risen in popularity in recent years and have been promoted as a lower carbon alternative to other forms of transit. Interest in bicycle sharing has been growing exponentially over the past decade, resulting in a proliferation of bike share systems in 712 cities across the world, encompassing 806,000 bicycles and 37,500 stations (Shaheen et al., 2014). This can be largely attributed to the successful incorporation of information technology in docking stations and mobile devices as well as improved logistics such as bicycle re-balancing to ensure responsive supply management.

2.2 References

1. Shaheen, S., Martin, E., Chan, N., Cohen, A., & Pogodzinsk, M. (2014). Public Bike-sharing in North America During a Period of Rapid Expansion: Understanding Business Models, Industry Trends & User Impacts. All Mineta Transportation Institute Publications. Book retrieved from **https://scholarworks.sjsu.edu/**on September 18, 2015.

2. Analysis of the New York City bike share system done by the analyst Aditya Khandekar has the analysis and described as a report given for reference https://www.ocf.berkeley.edu/~adityakh/2015/12/14/analysis-of-the-new-york-city-bike-share-system/

2.3 Problem Statement Definition

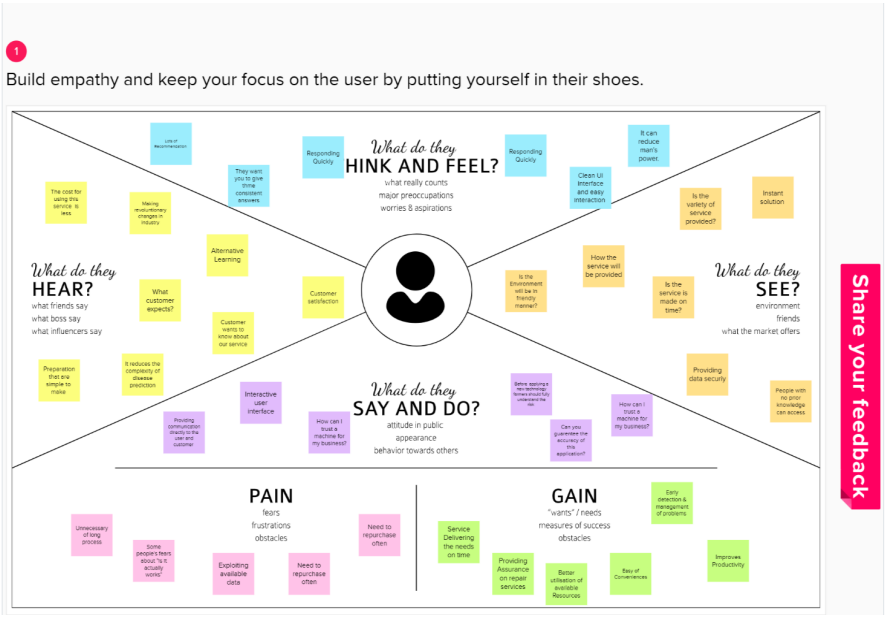
1. Citi Bike officials are pushing to make the program more robust and to broaden its reach. Financial viability increases with such larger bike-sharing programs.

2. The top bikes used with respect to trip duration could be found by this analysis so that more of these bikes can be produced and more users can be attracted.

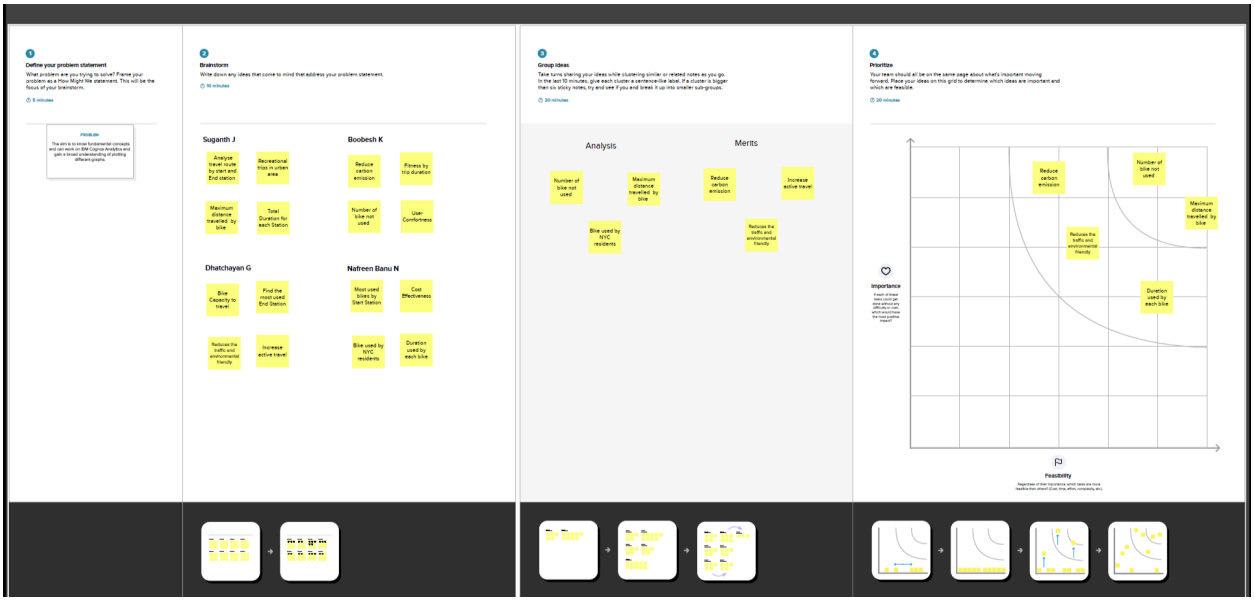
3. With the help of this analysis, top 10 Start station names with respect to customer age group could be found that the government can broaden the bike sharing system by increasing the number of bikes in those stations to make them readily available to all the potential users

3.IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas



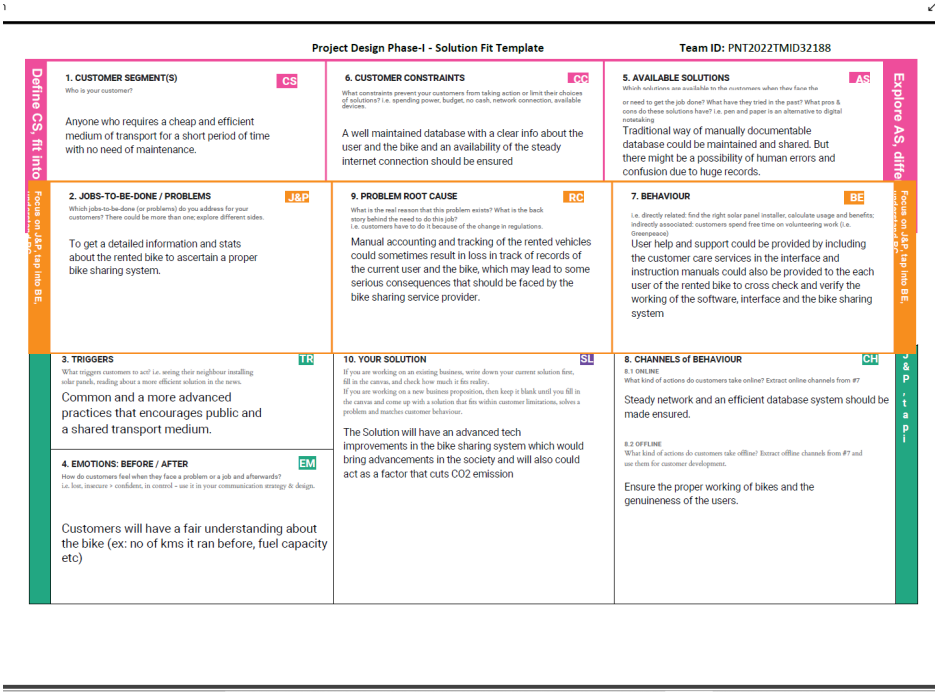
3.2 Ideation & Brainstorming



3.3 Proposed Solution

S.No.	Parameter	Description
1.	Problem Statement (problem to be solved)	**1.**Bike share system have risen popularity and have been promoted as low carbon alternative. **2.**The premise of sharing is short-term bike rental system, vary on timed memberships. **3.**It is an alternative for walking and using private bicycles.
2.	Idea / Solution Description	**1.**It is an open door policy, Encouraging inputs early and often the citizens of New York city **2.**Sites must ensure visibility and access
3.	Novelty / Uniqueness	**1.**Transport flexibility **2.**Health benefits **3.**Reduced congestion and fuel consumption
4.	Social Impact / Customer satisfaction	**1.**Transport **2.** Low cost **3.**Recreationn of cycling
5.	Business Model (Revenue Model)	**1.**Battery powered cycles **2.**Zero death since opened **3.**Lack of public subsidies
6.	Scalability of the solution	**1.**Economical development **2.**Reduce customer acquition cost **3.**Improved customer engagement

3.4 Problem Solution fit



4. REQUIREMENT ANALYSIS

4.1 Functional requirement

Functional Requirements:

Following are the functional requirements of the proposed solution.

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	user data collection:	The information is available on the official Lyft Citi Bike website to aid in research, development, visualisation, etc. These released files are where the data is gathered.
FR-2	Analysis of the user data	This data is used as input for various types of visualisations, as well as for analysis and the creation of a dashboard..
FR-3	Display of data	This dashboard is used to find the customer and subscriber with gender, to find the total number of trips, and to calculate the number of bikes used by respective age groups. It also displays the top bike used with respect to trip duration the top 10 Start Station and Names with respect to customer age group.

4.2 Non-Functional requirements

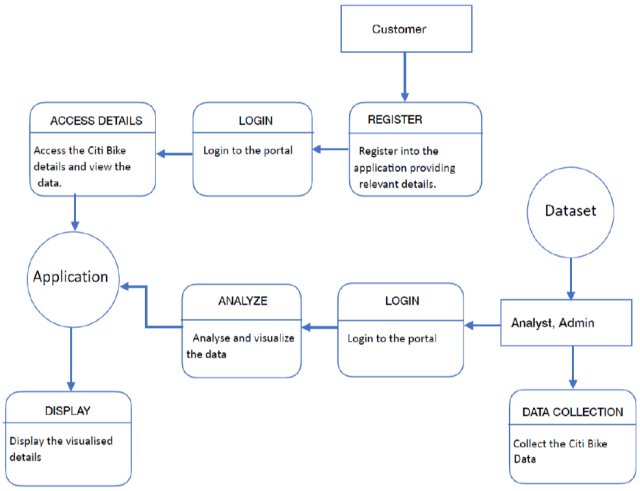
Non-functional Requirements:

Following are the non-functional requirements of the proposed solution.

FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	This dashboard offers an easily comprehensible report that helps many locals and visitors who use bicycles to get work done and have fun. It has several advantages, including measuring distance and assisting with duties like route planning, expanding the bicycle sharing system, producing required bikes, etc. Systems for sharing bicycles may result in lower car emissions, lower energy use, better health outcomes, personal financial savings, and lower traffic and fuel use.
NFR-2	**Security**	The need to secure data is very important as its crucial in a data analytics domain,hence access can be restricted using levels of authorization.
NFR-3	**Reliability**	It's always essential to have backup of data in case of corruption of server or complete wipe out In unfortunate situations
NFR-4	**Performance**	Operational effectiveness and spatial effectiveness are the two criteria used to measure bike sharing system performance. The operational effectiveness of the bike sharing program strives to comprehend the traits of public bike riders and assess the state of the bike lanes from their point of view. The dashboard for the bike sharing system's effectiveness looks at the characteristics of bike stations and the distances between them and other amenities. The public bicycle sharing program can be improved using the evaluation's findings.
NFR-5	**Availability**	A bicycle-sharing system is a type of shared transportation service in which bicycles are made available to individuals for short-term use at a low or no cost. The programs themselves include both docking and dockless systems, with docking systems allowing users to borrow a bike from a dock and return it to another node or dock within the system — and dockless systems providing a node-free system based on smart technology. Systems in either format may use smartphone web mapping to locate available bikes and docks.
NFR-6	**Scalability**	This analysis provides evidence of bike sharing systems' potential contribution to a more resilient transportation system, as they can quickly provide alternative transportation options to urban residents. As more data becomes available, particularly in other areas with similarly comprehensive bike sharing systems, this analysis will be able to provide a clearer picture of the role of this mode of transportation in these emergency situations and provide more accurate results. With increasing scale there shouldn't be any compromises

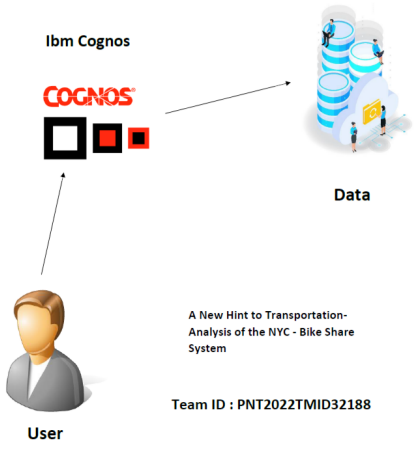
5. PROJECT DESIGN

5.1 Data Flow Diagrams

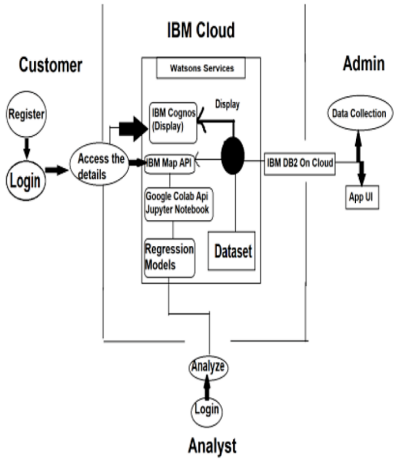


5.2 Solution & Technical Architecture

5.2.1 Solution Architecture



5.2.2 Technical Architecture



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
	User Data collection	USN-1	The information is available on the official Lyft Citi Bike website to aid in research, development, visualisation, etc.These released files are where the data is gathered.	Access the data from Lyft City Bike website	High	Sprint-1
	Login to the platform	USN-2	Login to the application to upload the analysed data.	Successful login	High	Sprint-2
	Analysis of the user data	USN-3	This data is used as input for various types of visualisations, as well as for analysis and the creation of a dashboard.	Ability to view the data analysed using the given tools	High	Sprint-3
	Display of data	USN-4	This dashboard is used to find the customer and subscriber with gender, to find the total number of trips, and to calculate the number of bikes used by respective age groups. It also displays the top bike used with respect to trip duration the top 10 Start Station and Names with respect to customer age group.	Access the dashboard and make changes	High	Sprint-3
Customer	Register	USN-5	Register with the application giving all the relevant details	Ability to access the account	High	Sprint-2
	Login to the platform	USN-6	Login to the application using the given username/email-id and password	Ability to access the account	Medium	Sprint-2
	Access the details	USN-7	Access the bike details and see the no.of bikes available	View the bike details	High	Sprint-3

6. PROJECT PLANNING & SCHEDULING

6.1 Sprint Planning & Estimation

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Data Preparation	USN-1	As an analyst, I filter and extract the Citi-bike data for the year 2018 from the given bucket of datasets	4	Medium	Suganth, Boobesh
Sprint-1	Data Preparation	USN-2	As an analyst, I upload the filtered dataset to IBM Cognos.	1	Medium	Dhatchayan, Nafreen Banu
Sprint-2	Data Preparation	USN-3	As an analyst, I can prepare the data for analysis by handling missing values and outliers	7	Medium	Suganth, Nafreen Banu
Sprint-2	Analysis	USN-4	As an analyst, I perform Exploratory Data Analysis on the filtered dataset to identify patterns and relationships between various features present.	8	High	Boobesh, Dhatchayan
Sprint-3	Visualization	USN-5	As an analyst, I create various visualizations using IBM Cognos based on the knowledge obtained at the end of the EDA process	10	High	Suganth, Dhatchayan
Sprint-3	Visualization	USN-6	As an analyst, I create a dashboard with the created visualizations to supplement business insights during the decision-making process at Citi.	10	High	Boobesh, Nafreen Banu
Sprint-4	Visualization	USN-7	As an analyst, I apply predictive analytics and additional features to enhance visualizations	5	Medium	Suganth, Boobesh

---

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-4	Registration	USN-8	As a user, I can register for the application by entering my email and password, and confirming my password.	5	Low	Dhatchayan, Nafreen Banu

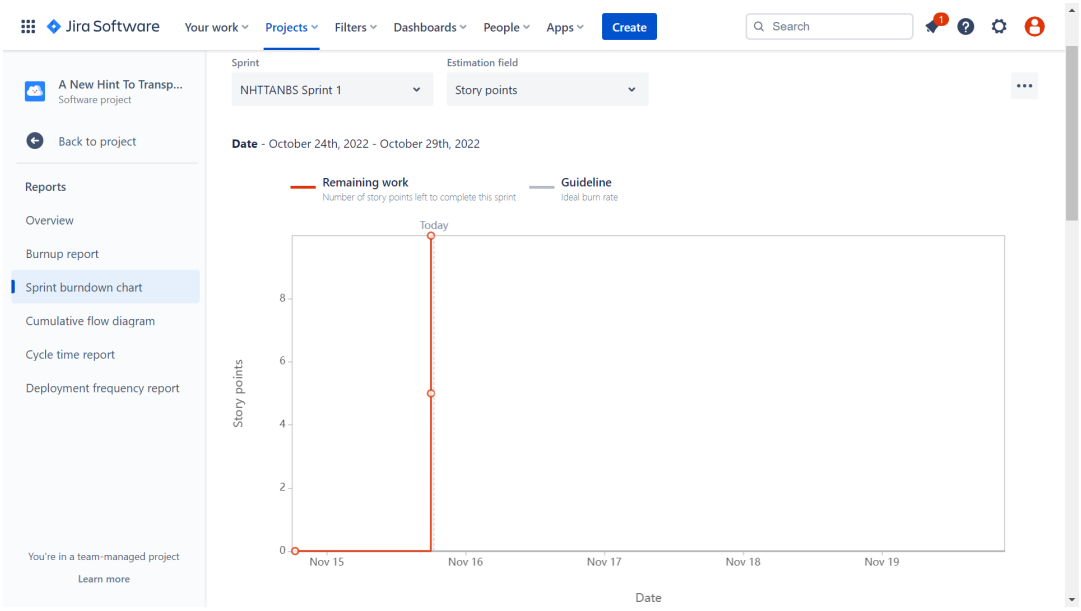
6.2 Sprint Delivery Schedule

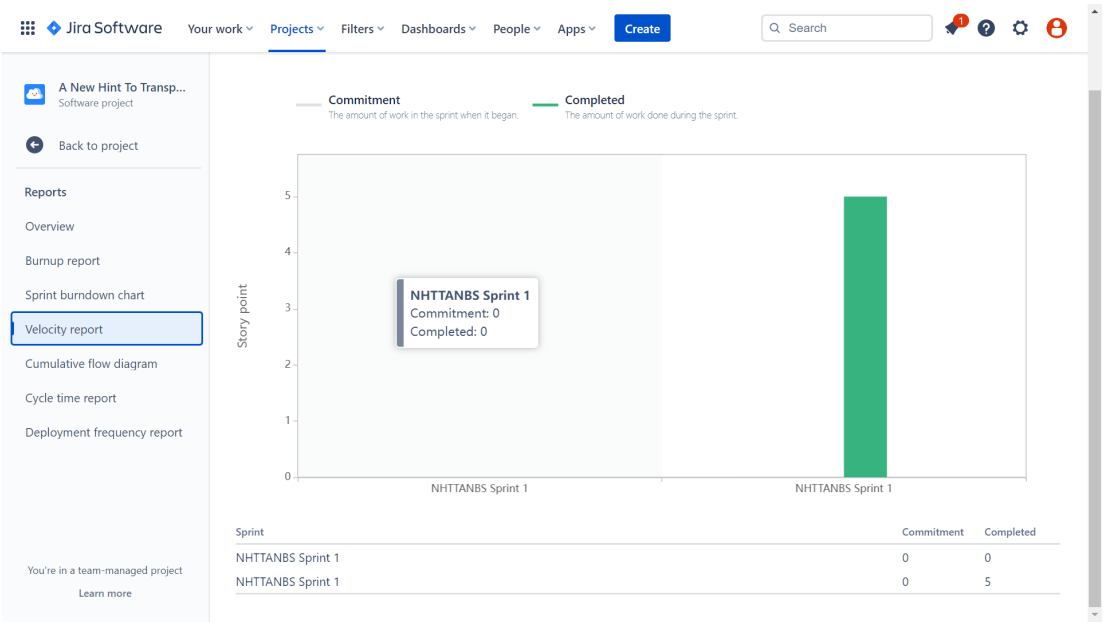
Project Tracker, Velocity & Burndown Chart: (4 Marks)

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	20	6 Days	24 Oct 2022	29 Oct 2022	5	29 Oct 2022
Sprint-2	20	6 Days	31 Oct 2022	05 Nov 2022	5	05 Nov 2022
Sprint-3	20	6 Days	07 Nov 2022	12 Nov 2022	5	12 Nov 2022
Sprint-4	20	6 Days	14 Nov 2022	19 Nov 2022	5	15 Nov 2022

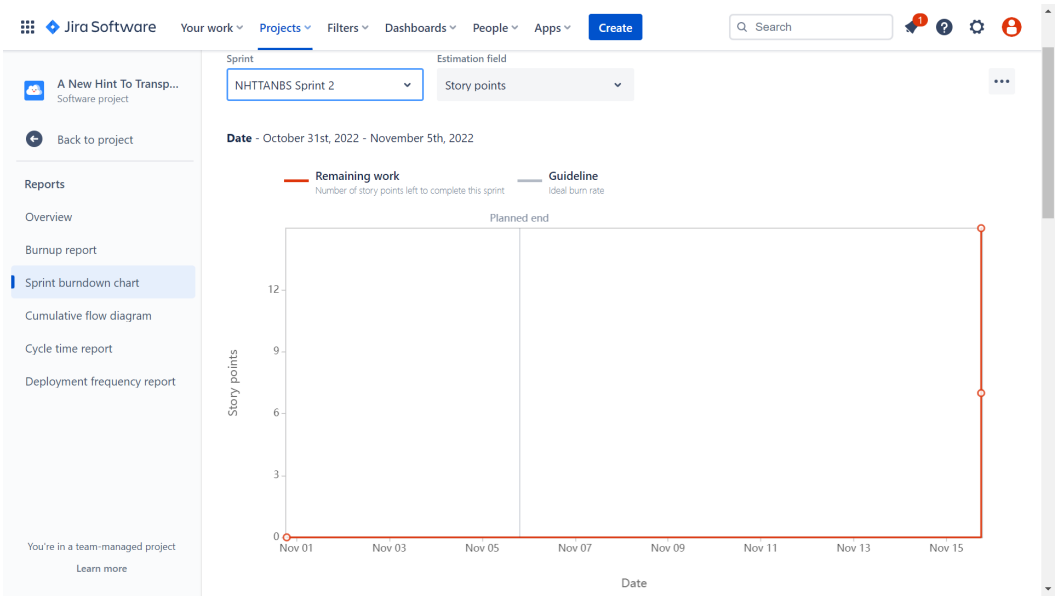
6.3 Reports from JIRA

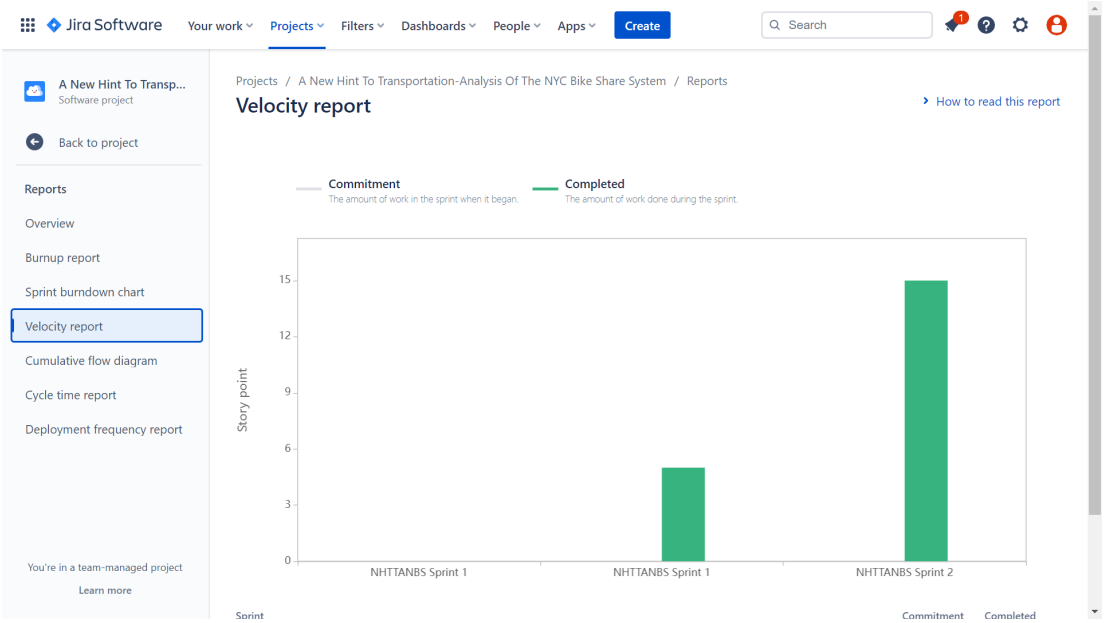
Sprint 1



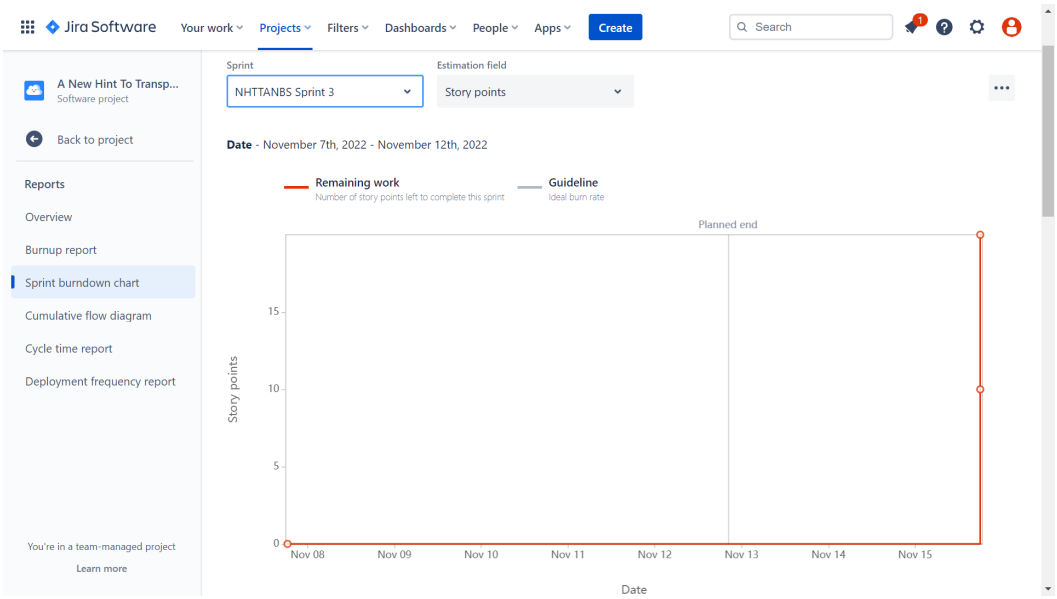


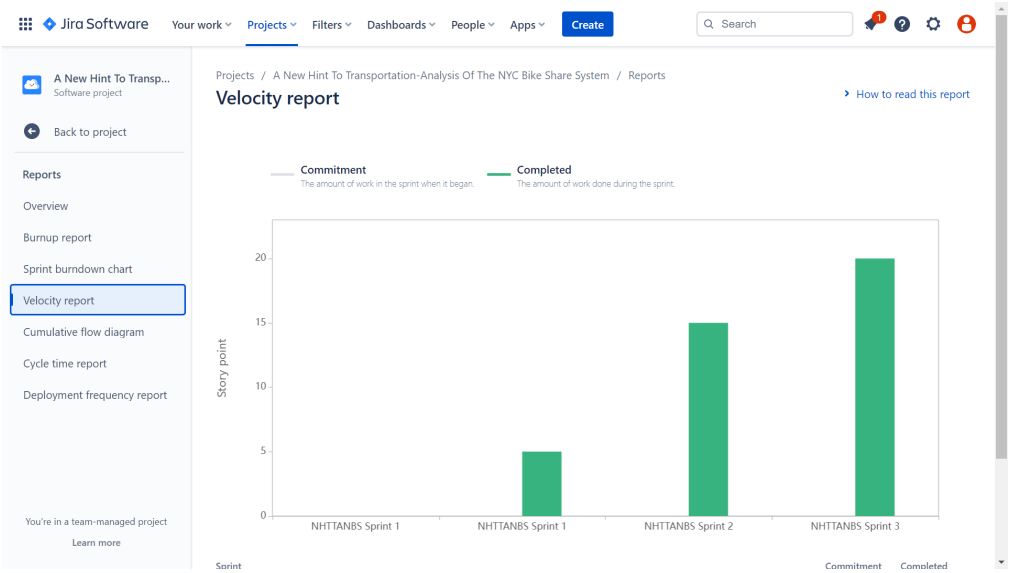
Sprint 2


Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
A New Hint To Transp...
Software project
Back to project
Reports
Overview
Burnup report
Sprint burndown chart
Cumulative flow diagram
Cycle time report
Deployment frequency report
You're in a team-managed project
Learn more
Sprint
NHTTANBS Sprint 2
Estimation field
Story points
Date - October 31st, 2022 - November 5th, 2022
Remaining work
Number of story points left to complete this sprint
Guideline
Ideal burn rate
Planned end
Story points
0
3
6
9
12
Nov 01
Nov 03
Nov 05
Nov 07
Nov 09
Nov 11
Nov 13
Nov 15
Date

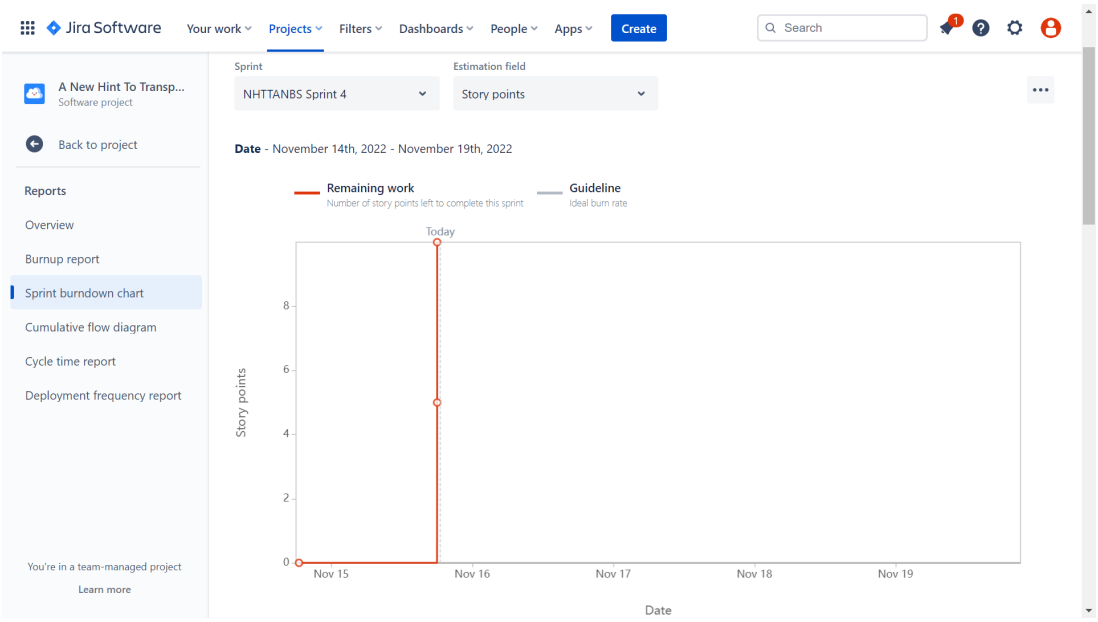

Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
A New Hint To Transp...
Software project
Back to project
Reports
Overview
Burnup report
Sprint burndown chart
Velocity report
Cumulative flow diagram
Cycle time report
Deployment frequency report
You're in a team-managed project
Learn more
Projects / A New Hint To Transportation-Analysis Of The NYC Bike Share System / Reports
Velocity report
How to read this report
Commitment
The amount of work in the sprint when it began.
Completed
The amount of work done during the sprint.
Story point
0
3
6
9
12
15
NHTTANBS Sprint 1
NHTTANBS Sprint 1
NHTTANBS Sprint 2
Sprint
Commitment
Completed

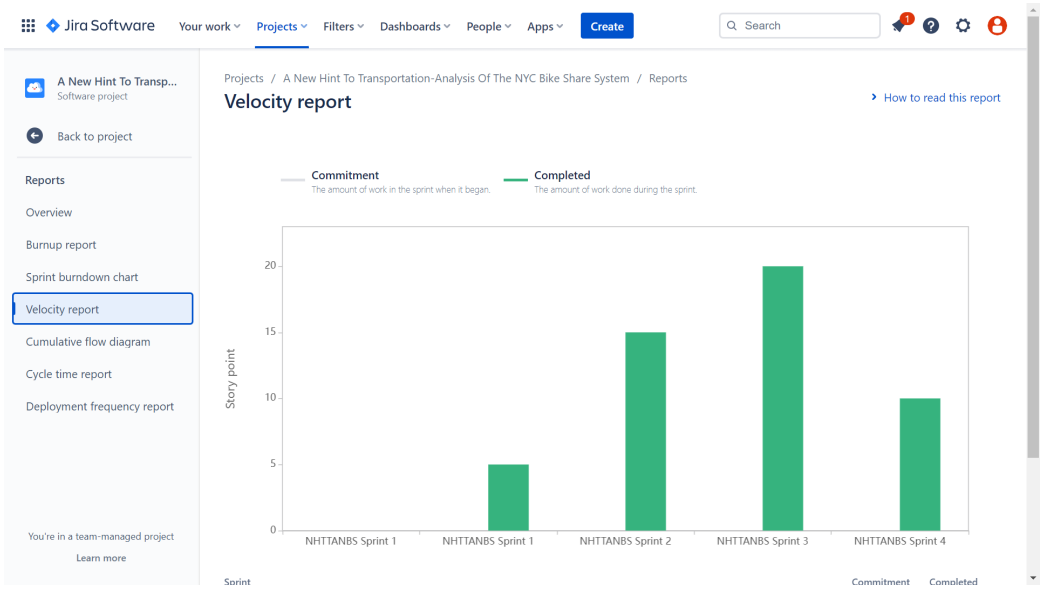
Sprint 3


Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
A New Hint To Transp...
Software project
Back to project
Reports
Overview
Burnup report
Sprint burndown chart
Cumulative flow diagram
Cycle time report
Deployment frequency report
You're in a team-managed project
Learn more
Sprint
NHTTANBS Sprint 3
Estimation field
Story points
Date - November 7th, 2022 - November 12th, 2022
Remaining work
Number of story points left to complete this sprint
Guideline
Ideal burn rate
Planned end
Story points
15
10
5
0
Nov 08
Nov 09
Nov 10
Nov 11
Nov 12
Nov 13
Nov 14
Nov 15
Date

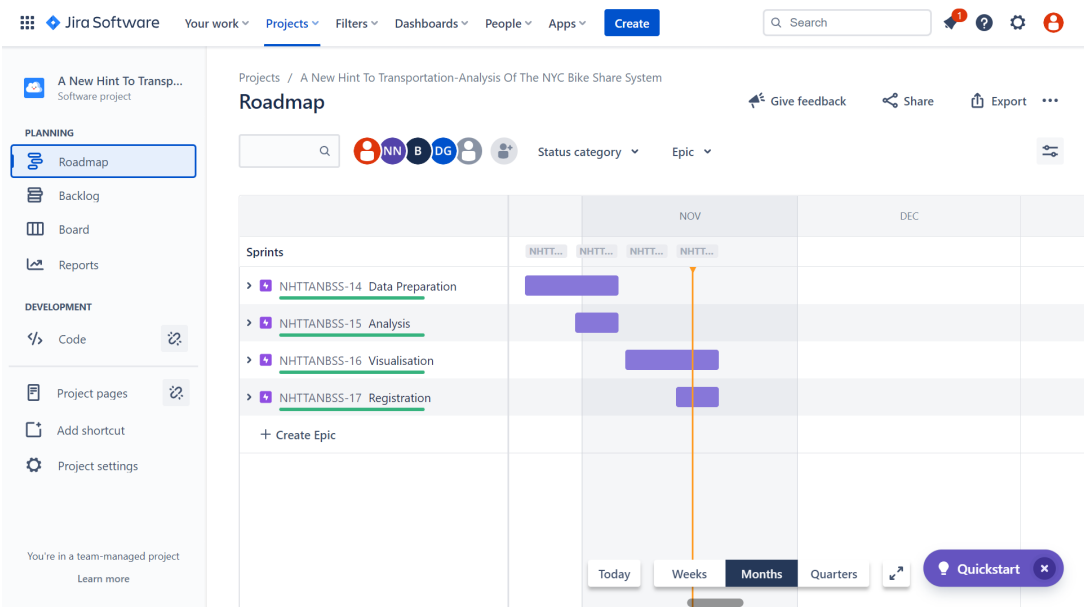

Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
A New Hint To Transp...
Software project
Back to project
Reports
Overview
Burnup report
Sprint burndown chart
Velocity report
Cumulative flow diagram
Cycle time report
Deployment frequency report
You're in a team-managed project
Learn more
Projects / A New Hint To Transportation-Analysis Of The NYC Bike Share System / Reports
Velocity report
How to read this report
Commitment
The amount of work in the sprint when it began.
Completed
The amount of work done during the sprint.
Story point
20
15
10
5
0
NHTTANBS Sprint 1
NHTTANBS Sprint 1
NHTTANBS Sprint 2
NHTTANBS Sprint 3
Sprint
Commitment
Completed

Sprint 4


Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
A New Hint To Transp...
Software project
Back to project
Reports
Overview
Burnup report
Sprint burndown chart
Cumulative flow diagram
Cycle time report
Deployment frequency report
You're in a team-managed project
Learn more
Sprint
NHTTANBS Sprint 4
Estimation field
Story points
Date - November 14th, 2022 - November 19th, 2022
Remaining work
Number of story points left to complete this sprint
Guideline
Ideal burn rate
Today
Story points
0
2
4
6
8
Nov 15
Nov 16
Nov 17
Nov 18
Nov 19
Date

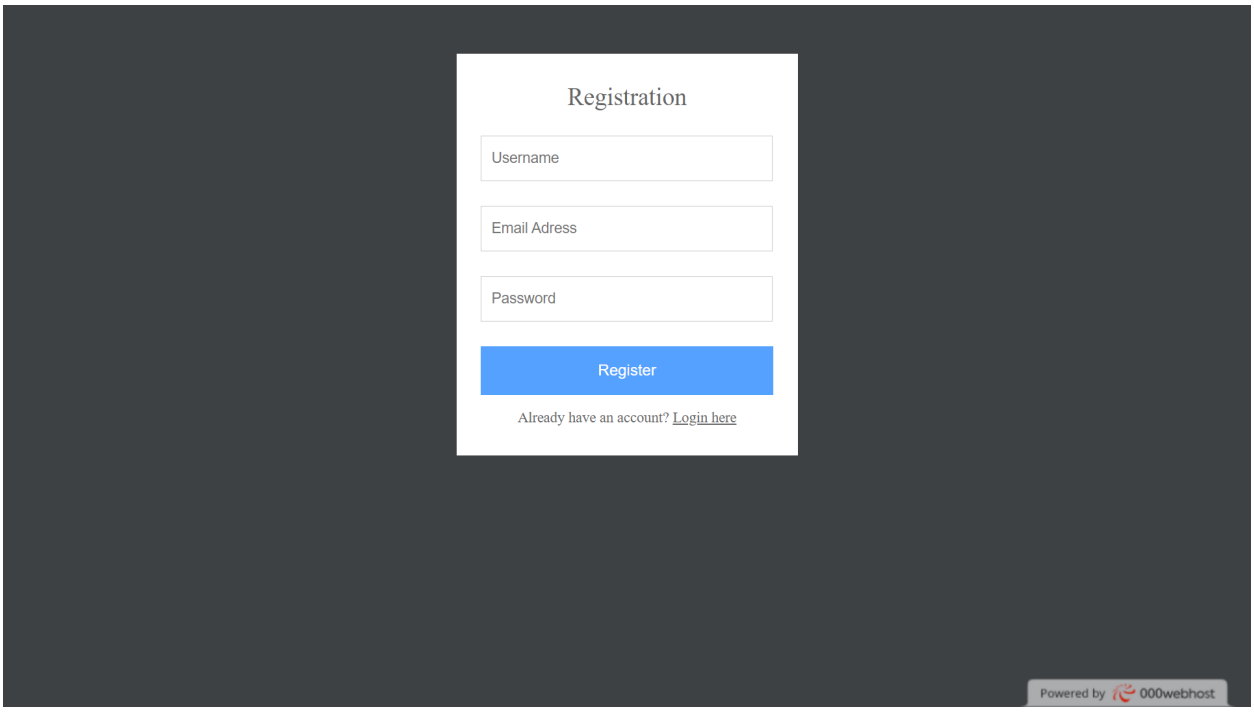

Jira Software
Your work
Projects
Filters
Dashboards
People
Apps
Create
Search
A New Hint To Transp...
Software project
Back to project
Reports
Overview
Burnup report
Sprint burndown chart
Velocity report
Cumulative flow diagram
Cycle time report
Deployment frequency report
You're in a team-managed project
Learn more
Projects / A New Hint To Transportation-Analysis Of The NYC Bike Share System / Reports
Velocity report
How to read this report
Commitment
The amount of work in the sprint when it began.
Completed
The amount of work done during the sprint.
Story point
0
5
10
15
20
NHTTANBS Sprint 1
NHTTANBS Sprint 1
NHTTANBS Sprint 2
NHTTANBS Sprint 3
NHTTANBS Sprint 4
Sprint
Commitment
Completed

RoadMap

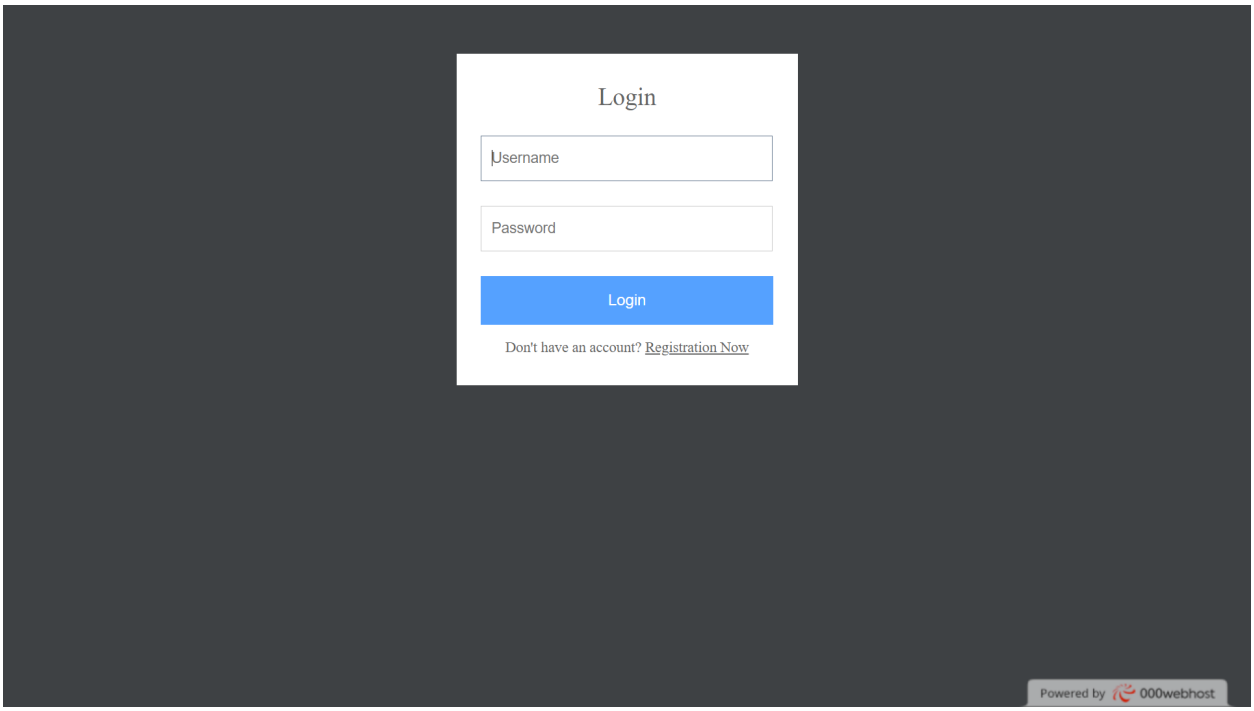


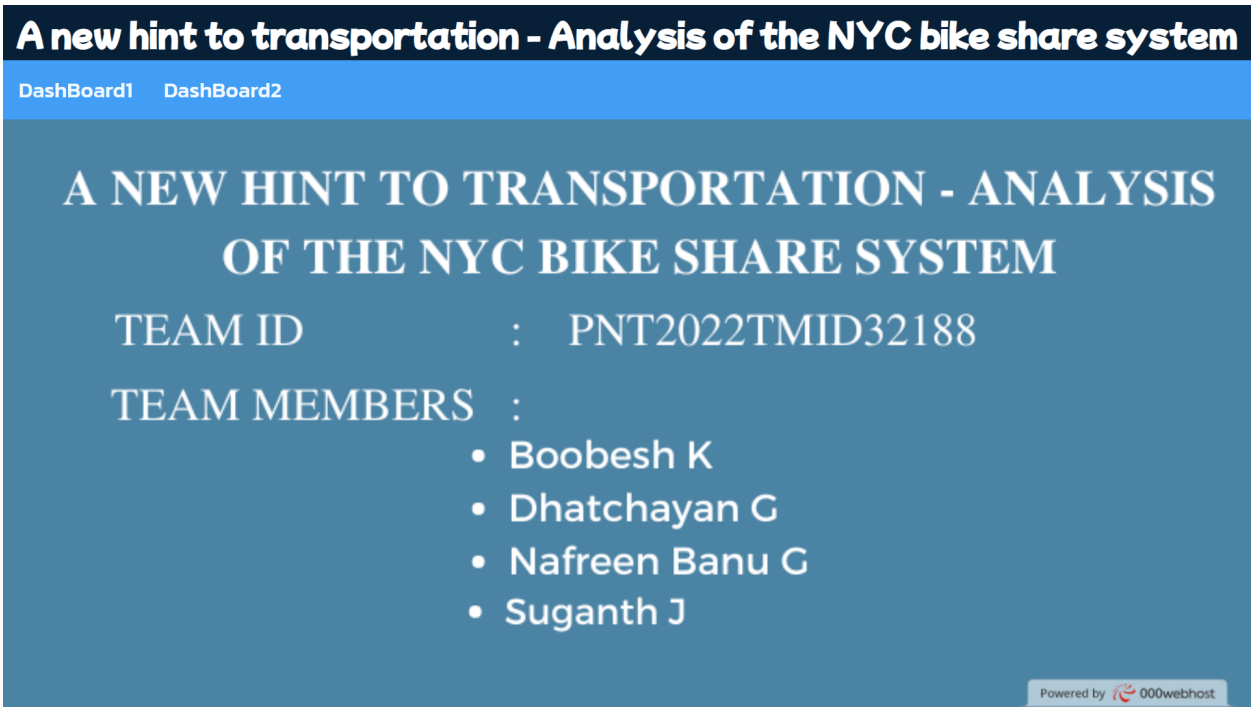
7. CODING & SOLUTIONING

7.1 Feature 1

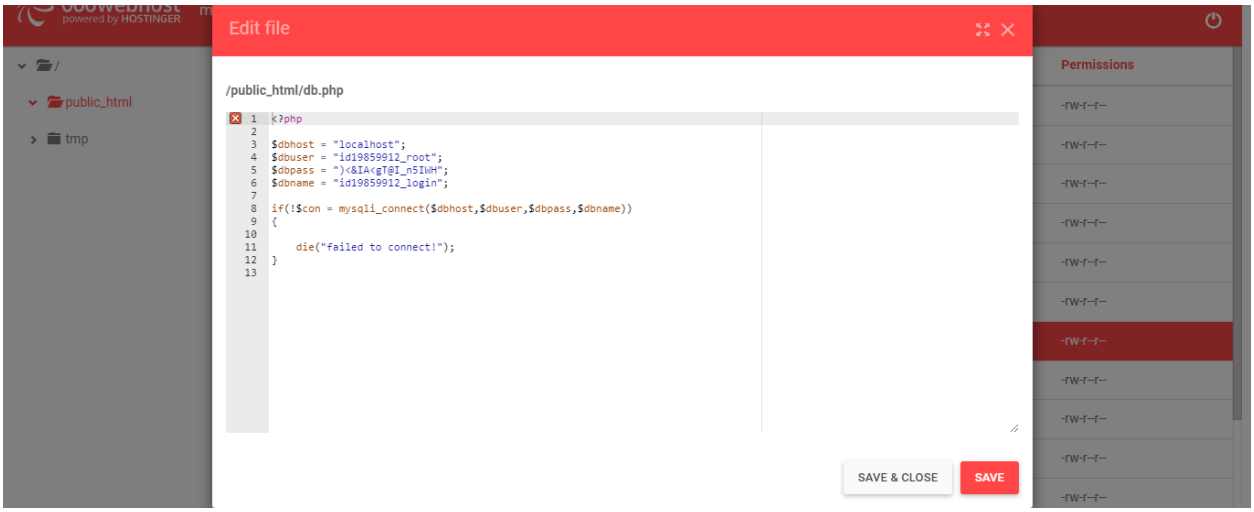


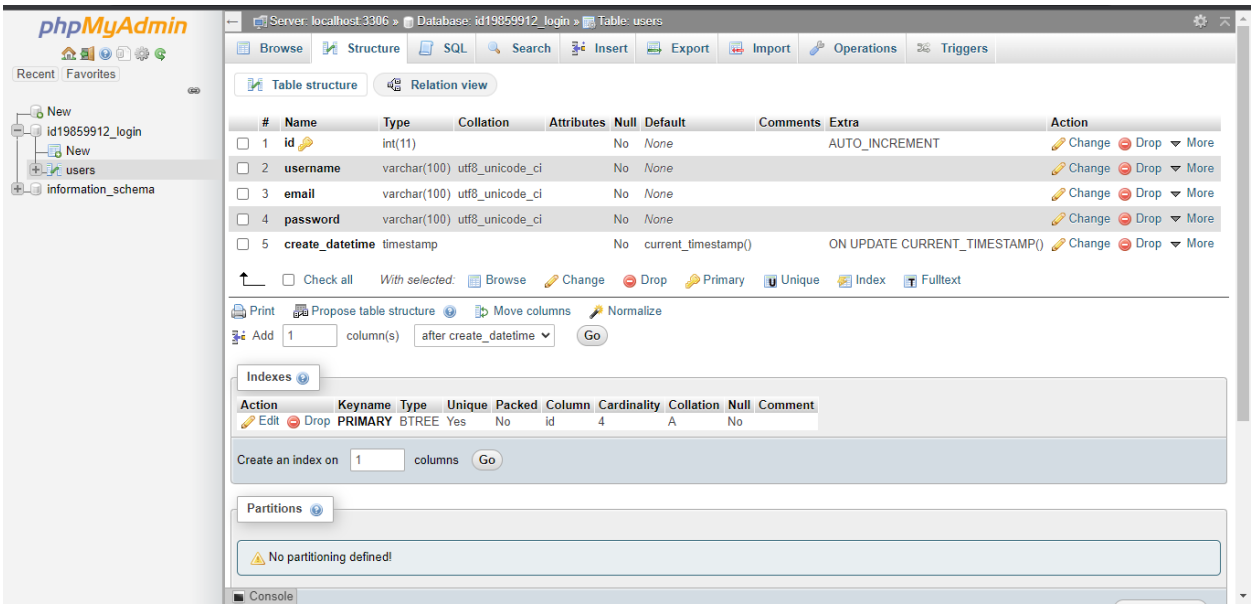
7.1 Feature 2


Login
Username
Password
Login
Don't have an account? Registration Now
Powered by 000webhost


A new hint to transportation - Analysis of the NYC bike share system
DashBoard1
DashBoard2
A NEW HINT TO TRANSPORTATION - ANALYSIS OF THE NYC BIKE SHARE SYSTEM
TEAM ID : PNT2022TMID32188
TEAM MEMBERS :
Boobesh K
Dhatchayan G
Nafreen Banu G
Suganth J
Powered by 000webhost

7.3 Database Schema





8. TESTING

8.1 Test Cases

A test case is a set of actions performed on a system to determine if it satisfies software requirements and functions correctly. The purpose of a test case is to determine if different features within a system are performing as expected and to confirm that the system satisfies all related standards, guidelines and customer requirements. The process of writing a test case can also help reveal errors or defects within the system.

S.No.	Parameter	Screenshot / Values
1.	Dashboard design	No of Visualizations / Graphs – 5-6 visulizations/4 – 5 graphs
2.	Data Responsiveness	Users and Analyst or Developer
3.	Amount Data to Rendered (DB2 Metrics)	Data (200 MB)
4.	Utilization of Data Filters	Remove null values and used needed data for visualization
5.	Effective User Story	No of Scene Added – 5 story
6.	Descriptive Reports	No of Visualizations / Graphs - 5

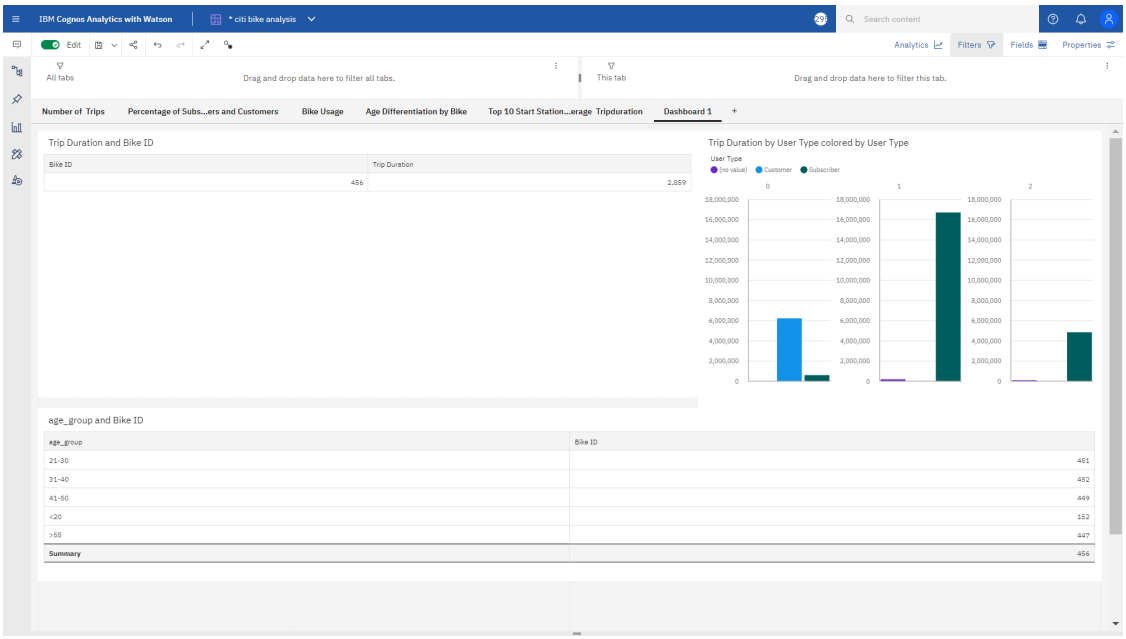
8.2 User Acceptance Testing

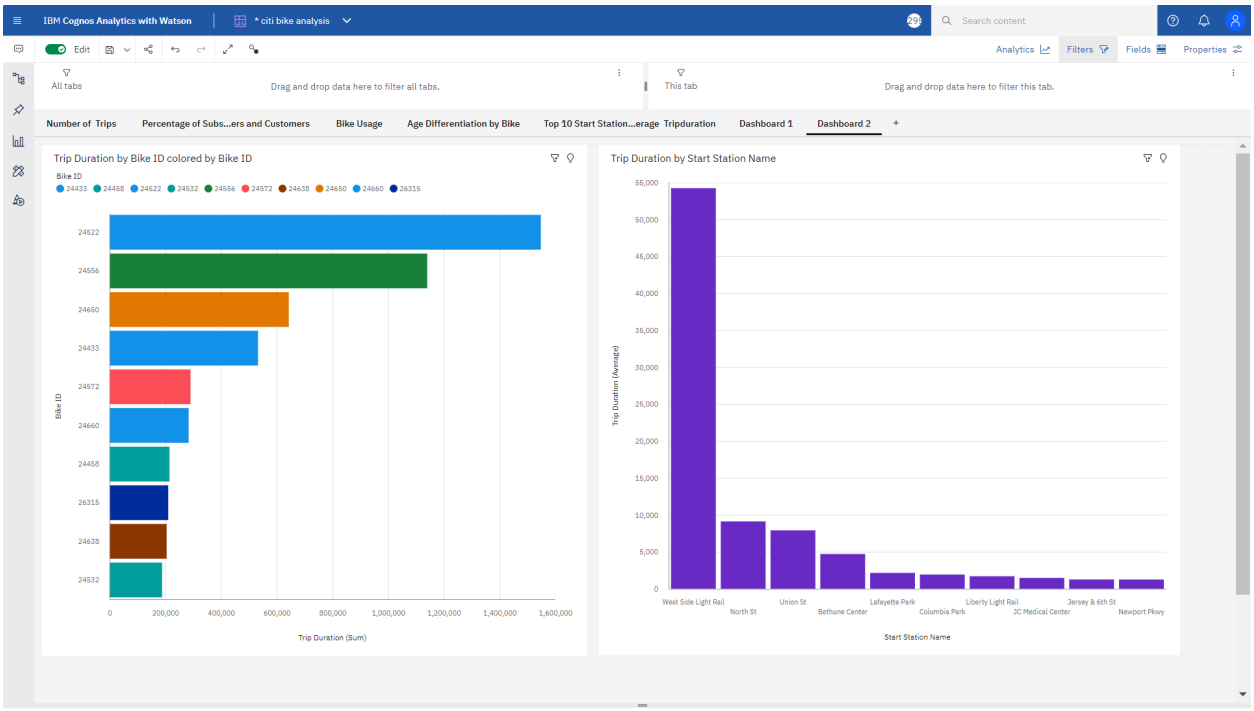
User Acceptance Testing (UAT) is a type of testing performed by the end user or the client to verify/accept the software system before moving the software application to the production environment. UAT is done in the final phase of testing after functional, integration and system testing is done.

Section	Total Cases	Not Tested	Fail	Pass
Login Page	1	0	0	1
Citi Bike details	1	0	0	1
Database	2	0	0	2
Dashboard	2	0	0	2
Visualize the data	8	0	0	8
Logistic Regression	4	0	0	4

9. RESULTS

9.1 Performance Metrics





10. ADVANTAGES & DISADVANTAGES

10.1 Advantages

- Riding a Citibike makes you really interesting for once**.**
- It’s the fastest way to get from A to B in certain situations.
- Docking is easy in the Citibike.

10.2 Disadvantages

- It’s extremely difficult to ride uphill.
- They are slow and bulky.
- There are other various technical difficulties.

11. CONCLUSION

Bike-share usage has an overall upward trend, yearly seasonality and weekly cycles; weather and system size are also important factors determining the number of trips. Better data could help make a better prediction. E.g. heavy rain in peak hours affects ridership dramatically and having detailed hourly weather data would help create a more accurate model.

Walking and cycling have become more important . The pattern of bike-share use has changed: more people rely on bike-sharing for recreation, shopping and other trips. Even without the daily work commutes overall ridership has increased compared . It will be interesting to see whether this pattern stays the same when people begin commuting again.

One can only welcome Citi Bike's expansion in 2020. But perhaps even more important is creating protected infrastructure like new footpaths and temporary and permanent bike lanes. New York City's Open Streets and similar programs in other cities around the world are crucial in order to accommodate this increased demand and create a safe environment for pedestrians and cyclists.

12. FUTURE SCOPE

One aspect of the data that I did not explore in great detail is the intra-day variation in usage of the system. This is also a key aspect that bike share system operators are very interested in because knowing the variation in demand on an hourly basis is another very useful metric for identifying the times of the day when the need for artificial re-balancing is maximum.

Additionally, this work will feed into a larger study calculating the life cycle environmental impacts of a bike-share system and its ability to substitute other modes of transit with the aim of reducing the overall Greenhouse gas (GHG) emissions due to transportation.

13. Appendix

Source code:

1.index.php

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

2.login.php

```php
<!DOCTYPE html>
<html>
<head>
    <meta charset="utf-8"/>
    <title>Login</title>
    <link rel="stylesheet" href="style.css"/>
</head>
<body>
<?php
    require('db.php');
    session_start();
    // When form submitted, check and create user session.
    if (isset($_POST['username'])) {
        $username = stripslashes($_REQUEST['username']);    //
removes backslashes
        $username = mysqli_real_escape_string($con, $username);
        $password = stripslashes($_REQUEST['password']);
        $password = mysqli_real_escape_string($con, $password);
        // Check user is exist in the database
        $query    = "SELECT * FROM `users` WHERE
username='$username'
```

```
                    AND password='" . md5($password) . "'";
        $result = mysqli_query($con, $query) or die(mysql_error());
        $rows = mysqli_num_rows($result);
        if ($rows == 1) {
            $_SESSION['username'] = $username;
            // Redirect to user dashboard page
            header("Location: dashboard.php");
        } else {
            echo "<div class='form'>
                  <h3>Incorrect Username/password.</h3><br/>
                  <p class='link'>Click here to <a href='login.php'>Login</a> again.</p>
                  </div>";
        }
    } else {
?>
    <form class="form" method="post" name="login">
        <h1 class="login-title">Login</h1>
        <input type="text" class="login-input" name="username" placeholder="Username" autofocus="true"/>
        <input type="password" class="login-input" name="password" placeholder="Password"/>
        <input type="submit" value="Login" name="submit" class="login-button"/>
        <p class="link">Don't have an account? <a href="registration.php">Registration Now</a></p>
  </form>
<?php
    }
?>
</body>
</html>
```

3.logout.php

```php
<?php
    session_start();
    // Destroy session
    if(session_destroy()) {
        // Redirecting To Home Page
        header("Location: login.php");
    }
?>
```

4.registration.php

```php
<!DOCTYPE html>
<html>
<head>
    <meta charset="utf-8"/>
    <title>Registration</title>
    <link rel="stylesheet" href="style.css"/>
</head>
<body>
<?php
    require('db.php');
    // When form submitted, insert values into the database.
    if (isset($_REQUEST['username'])) {
        // removes backslashes
        $username = stripslashes($_REQUEST['username']);
        //escapes special characters in a string
        $username = mysqli_real_escape_string($con, $username);
        $email    = stripslashes($_REQUEST['email']);
        $email    = mysqli_real_escape_string($con, $email);
        $password = stripslashes($_REQUEST['password']);
        $password = mysqli_real_escape_string($con, $password);
        $create_datetime = date("Y-m-d H:i:s");
        $query    = "INSERT into `users` (username, password, email,
create_datetime)
                    VALUES ('$username', '" . md5($password) . "',
```

```
    '$email', '$create_datetime')";
24          $result   = mysqli_query($con, $query);
25          if ($result) {
26              echo "<div class='form'>
27                    <h3>You are registered successfully.</h3><br/>
28                    <p class='link'>Click here to <a
    href='login.php'>Login</a></p>
29                    </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 Adress">
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>
```

5.db.php

```php
<?php
$dbhost = "localhost";
$dbuser = "root";
$dbpass = "";
$dbname = "loginsystem";
if(!$con = mysqli_connect($dbhost,$dbuser,$dbpass,$dbname))
{
    die("failed to connect!");
}
```

6.dashboard.php

```php
<?php
//include auth_session.php file on all user panel pages
include("auth_session.php");
?>
<!DOCTYPE html>
<html>
<head>
    <meta charset="utf-8">
    <title>Dashboard - Client area</title>
    <link rel="stylesheet" href="style.css" />
</head>
<body>
    <div class="form">
        <p><b><i>Hey,</b></i> <?php echo $_SESSION['username']; ?>!</p>
        <p><b><i>You are in user dashboard page.</b></p>
            <form method="POST" action="indexx.php">
                <center><input type="submit" value="DashBoard" name="submit" /></center>
            </form>
```

```
19          <p><a href="logout.php">Logout</a></p></i>
20      </div>
21 </body>
22 </html>
```

Github & Project Demo Link:

Github Link: https://github.com/IBM-EPBL/IBM-Project-12588-1659454470

Project Demo Link : https://clipchamp.com/watch/LcMs716Msxp