

IBM COGNOS TO ANALYZE AND VISUALIZE NEW YORK CITY BIKE RIDE

Mini Project documentation submitted to

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY, HYDERABAD

In partial fulfillment of the requirements for the award of the degree

BACHELOR OF TECHNOLOGY

In

COMPUTER SCIENCE AND ENGINEERING

Submitted By

MEKALA PRAVEEN

19UK5A0508

KAVATI ANANTHESHWAR

19UK5A0515

BHUKYA VASANTH NAIK RATHOD

17UK1A05F8

Under the guidance of

MS. T.SUSHMA

(Associate Professor)



DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

VAAGDEVI ENGINEERING COLLEGE

Affiliated to JNTUH, HYDERABAD

BOLLIKUNTA, WARANGAL (T.S) –506005

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

VAAGDEVI ENGINEERING COLLEGE

WARANGAL



CERTIFICATE

This is to Certify that the Project entitled “**IBM COGONS TO ANALYSE AND VISUALIZE NEW YORK CITYBIKE RIDE SHARE DATA**” is submitted by **MEKALA PRAVEEN(19UK5A0508),KAVATI ANANTHESHW(19UK5A0515), BHUKYA VASANTH NAIK RATHOD(17UK1A05F8)** In partial fulfilment of the requirements the award of the Degree of Bachelor of Technology in Computer science & Engineering to Jawaharlal Nehru Technological University Hyderabad during the Academic year 2021-2022

Project Guide

HOD

MS.T.SUSHMA

DR.R. NAVEEN KUMAR

EXTERNAL EXAMINER

ACKNOWLEDGEMENT

We wish to take this opportunity to express our sincere gratitude and deep sense of respect to our beloved **DR. P. PRASAD RAO**, Principal, Vaagdevi Engineering College for making us available all the required assistance and for his support and inspiration to carry out this mini project in the institute.

we extend our heartfelt thanks to **Dr. R. NAVEEN KUMAR**, Head, Department of CSE, Vaagdevi Engineering College for providing us necessary infrastructure and thereby giving us freedom to carry out the mini project.

We express heartfelt thanks to the guide, **MRS.SUSHMA** Assistant Professor, Department of CSE for her constant support and giving necessary guidance for completion of this mini project.

We express our sincere thanks and gratitude to **THE SMART BRIDGE**, for providing internship.

Finally, we express our sincere thanks and gratitude to my family members, friends for their encouragement and outpouring their knowledge and experiencing throughout these.

ABSTRACT

Many bike share systems make available their trip data for those who want to understand how their systems are used. The bike share system in New York City, Citi Bike, is one of them. The trip data files contain one record for each ride, around two million records per month, depending on the season. It's a traditional bike share system with fixed stations where a user picks up a bike at one dock, using a key fob or a code, and returns it at another. The station and time when the ride started and stopped is recorded for each ride. Some limited information about the rider is also recorded: their gender and year of birth.

The goal of this analysis is to create an operating report of Citi Bike. Let us create data visualizations to understand. Total Number of Trip What is Customer and subscriber with gender and find the top bike used with respect to trip duration and calculating the number of bikes used by respective age groups and Top 10 Start Station Names with respect to Customer age group. Trip Duration (seconds), Start Time and Date, Stop Time and Date, Start Station Name End Station Name Station ID, Station Lat/Long, Bike ID, User Type (Customer = 24-hour pass or single ride user; Subscriber = Annual Member), Gender (Zero=unknown; 1=male; 2=female) year of Birth, the kinds of questions we wanted to answer included ones like these: What's the most common ride duration? What times of the day does the system get the most usage? How much does ridership vary over the course of a month? What are the most used stations? How old are the riders?

While the answers to these questions can be found in the trip data files, the data needs to be augmented to provide easy answers. For example the trip duration in seconds is too granular; minutes would be more useful.

TABLE OF CONTENT

	PAGE.NO
1. INTRODUCTION	
1.1 Introduction	6
1.2IBM services	6
2.LITERATURE SURVEY	
2.1introduction	7
2.2 history	7
2.3 Services	8
3. ANALYSIS	
3.1 Introduction	9
3.2Content diagram	10
4.DESIGN	
4.1 Introduction	11
4.2 Hardware/Software	12
5. IMPLEMENTATION AND RESULTS	
5.1 introduction	13
5.2 Result analysis	14
6. TESTING AND VALIDATION	
6.Number of trips	15
6.2 Customer & subscriber with Gender	16
7.CONCLUSION	17
8.FUTURE ENHANCEMENT	18
9.REFERENCES	19

1. INTRODUCTION

Users create multiple analysis graphs/charts. Using the analyzed chart creation of Dashboard is done. Saving and visualizing the final dashboard in the IBM Cognos Analytics. To accomplish this, we have to complete all the activities and tasks listed below, IBM Cloud Account, Login to Cognos Analytics, Working with the Dataset, Understand the Dataset, Loading the Dataset, Data visualization charts, Problem Statement Number of trips, Problem Statement, Percentage of Subscribers and Customers, Problem Statement, Bike Usage Problem Statement, Age Group differentiation by Bike, Problem Statement, Top 10 Start Station Names with Customer age group Dashboard Creation Export the Analytics and top 10 selected members like their bike cc and their gender and citys The bike share system in New York City, Citi Bike, is one of them. The trip data files contain one record for each ride, around two million records per month, depending on the season. It's a traditional bike share system with fixed stations where a user picks up a bike at one dock, using a key fob or a code, and returns it at another. The station and time when the ride started and stopped is recorded for each ride. Some limited information about the rider is also recorded: their gender and year of birth. The New York City Bike Share enables quick, easy, and affordable bike trips around the New York city boroughs. They make regular open data releases (this dataset is a transformed version of the data from.

,

1. IBM SERVICES

consulting, architecture, transformation on cloud, and management services to build a next generation data platform. Gain control of your data environment to drive actionable insights. IBM offers data strategy,

Accelerate innovation, while leveraging existing investments by transitioning traditional applications to the cloud in a cost-effective way that makes sense for your business.



2.0 LITERATURE SURVEY

2.1 INTRODUCTION:

IBM Services is the professional services arm of IBM, made up of business, technology and industry experts who apply advanced technology and help clients design, build and run businesses. It includes two divisions: IBM Global Business Services and IBM Global Technology Services.

IBM Cognos Business Intelligence is a web-based integrated business intelligence suite by IBM. It provides a toolset for reporting, analytics, scorecarding, and monitoring of events and metrics. The software consists of several components designed to meet the different information requirements in a company Import data from CSV files and spreadsheets. Connect to cloud or onpremises data sources, including SQL databases, Google Big Query, Amazon, Redshift, and more.

2.2 HISTORY:

IBM's early involvement in IT services can be traced back to the mid-seventies and its Data Processing Support Services (DPSS) division. At the time, the company was under regulatory pressure to unbundle its support services from the sale of hardware. Additionally, with the launch of its early teleprocessing and database products, such as CICS and IMS, many customers could only acquire the expertise needed to build their applications through hiring consulting services directly from IBM itself. Later, in 1989, Eastman Kodak Company and IBM completed an agreement by which IBM designed, built and managed a new state-of-the-art data center for under the brand name ISSC, Integrated Systems Solution Corporation.

Also in 1989, IBM introduced Business Recovery Services, an offering that enables a business to continue operations in the event of an unplanned outage or disaster. In 1992, the company formed IBM Consulting Group, as a new management consulting organization with service lines in Business Transformation and IT Strategy Consulting This initiative was led by Robert M. Howe, IBM vice-president and general manager of IBM .Consulting Group The group was rebranded to IBM Business Innovation Services in 2001.

2.3 SERVICES:

IBM SERVICES is the professional services arm of IBM, made up of business, technology and industry experts who apply advanced technology and help clients design, build and run businesses. It includes two divisions: IBM Global Business Services (GBS) and IBM Global Technology Services (GTS). Intelligent workflows built on data can transform the way work gets done in an enterprise. Create a unified view of enterprisewide data to build analytical and operational views for decision-making, and to automate end-to-end workflows to uncover new and better ways of working.

IBM's data and analytics consulting services can help you bring more clarity needed to operationalize your enterprise data and build a data-driven organization to drive business value.

Gain control of your data environment to drive actionable insights. IBM offers data strategy, consulting, architecture, transformation on cloud, and management services to build a nextgeneration data platform.

Identify digital business transformation opportunities to quickly progress from ideation through to design, prototype and implementation. Build confidence in your analytics tools with ethical, trustworthy AI.

IBM analytics consulting helps you integrate and scale your business intelligence and automation efforts. Use data science, predictive analytics and data visualization to gain meaningful insights that transform your business.

Integrate data from multiple enterprise applications such as SAP, workday sales force as well as from documents to build intelligent workflows

3.ANALYSIS:

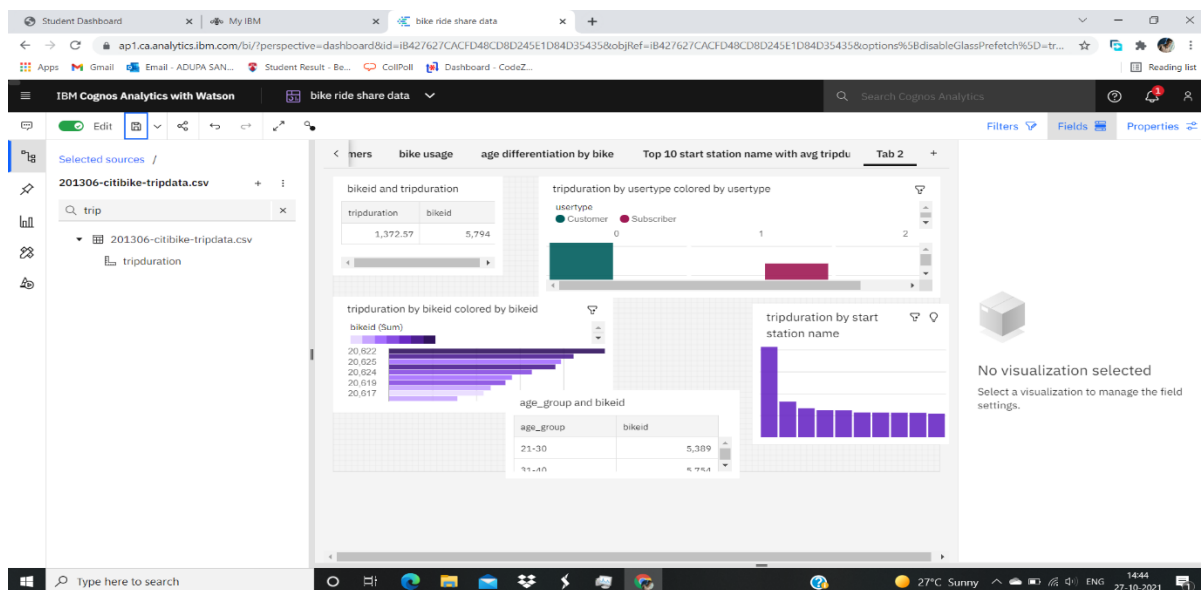
3.1 INTRODUCTION:

The objective of the Service Analysis Roadmap step is to use the Business Object Model (BOM) to define an enterprise view of the business concepts and capabilities that relate to the documented requirements, independent of implementation considerations.

Inputs to Service Analysis come from a variety of sources, depending on the needs to the Financial Institution. Examples include:

1. IBM Industry Model APM Service Candidate Tasks and Processes, identified during a Process analysis phase
2. IBM CBM (Component Business Model)

Business processes are analyzed to identify the service candidates that are mapped to from the BOM and new activities that represent new service candidates. The outputs of this phase are a new and customized capability operation for each identified service candidate, and customized data concepts that enable their interaction with other services.



3.2. DESIGN:

Process Analysis is used to identify and define the tasks and activities that take place during each phase of a business process. It is also used to describe the data inputs and outputs for each activity. Process Analysis using the APM Processes can help to understand how the process operates and to identify potential improvements in existing processes.

The objective of the Process Analysis phase is to:

- Apply the business process templates to understand the documented requirements
- Apply customizations as required to the business process templates based on the documented requirements
- Review the customized process to understand the implementation requirements for individual steps, determining whether manual or automated implementation is required

Inputs to Process Analysis come from a variety of sources, depending on the needs to the Financial Institution. Examples include:

1. IBM Industry Model Value Chains

For each of the sources, the inputs to the Process Analysis phase are a set of Processes identified by an earlier process scoping exercise performed on the source artefact.

The output of Process Analysis is a set of customised business processes which satisfy the business requirements for the development project.

When following a top down modelling approach that includes modelling Services, the Process Analysis phase may also include identification of Service Candidate activities.

4.DESIGN:

4.1 INTRODUCTION:

Many bike share systems make available their trip data for those who want to understand how their systems are used. The bike share system in New York City, Citi Bike, is one of them, but they don't provide much more than the data. I've got some experience in obtaining and preparing their data for visualization, so in this article I will show you how to get started with this rich data source. In the Before Times I commuted from suburban New Jersey to my job as a Product Manager in New York City at an office, now shuttered, above Penn Station. To get around in the City at lunch or after work I often relied on Citi Bike, New York's bike share system. I found I could get to destinations in midtown and even further afield faster than walking and cheaper than the bus or subway. When I discovered that Citi Bike made trip data publicly available I thought that it might provide an interesting use case for the data preparation product that I managed. Using real data turned out to be much more interesting than the sample files that we had been using because there were actual anomalies that needed to be cleaned up to make the data useful for analysis, and there were interesting stories to tell from the data. The trip data files contain one record for each ride, around two million records per month, depending on the season. It's a traditional bike share system with fixed stations where a user picks up a bike at one dock, using a key fob or a code, and returns it at another. The station and time when the ride started and stopped is recorded for each ride.

4.2 Hardware Software: -

"IBM COGNOS ANALYTICS"

- ✓ IBM Cognos is a business intelligence tool for web-based reporting and analytics.
- ✓ It is a Online based Bi Tool.
- ✓ BI(Business Intelligence) is a set of processes, architects, and technologies that convert raw data into meaningful information.
- ✓ The various features of IBM Cognos are Dashboard Creation, Reporting, Analysis, Data Integration.
- ✓ Use built-in AI to accelerate and improve blending data or finding the optimal tables for your model.
- ✓ Create powerful visualizations, tell the story of your data and share insights via email, Slack, or the mobile app.
- ✓ Governed self-service adapts to your requirements and protects data from misuse.
- ✓ AI helps uncover hidden trends and drivers so you can get the facts behind your data and deliver insights in real time.
- ✓ Bring together advanced analytics and data science into daily operations and seize new opportunities.

5. IMPLEMENTATION AND RESULT

5.1 INTRODUCTION:

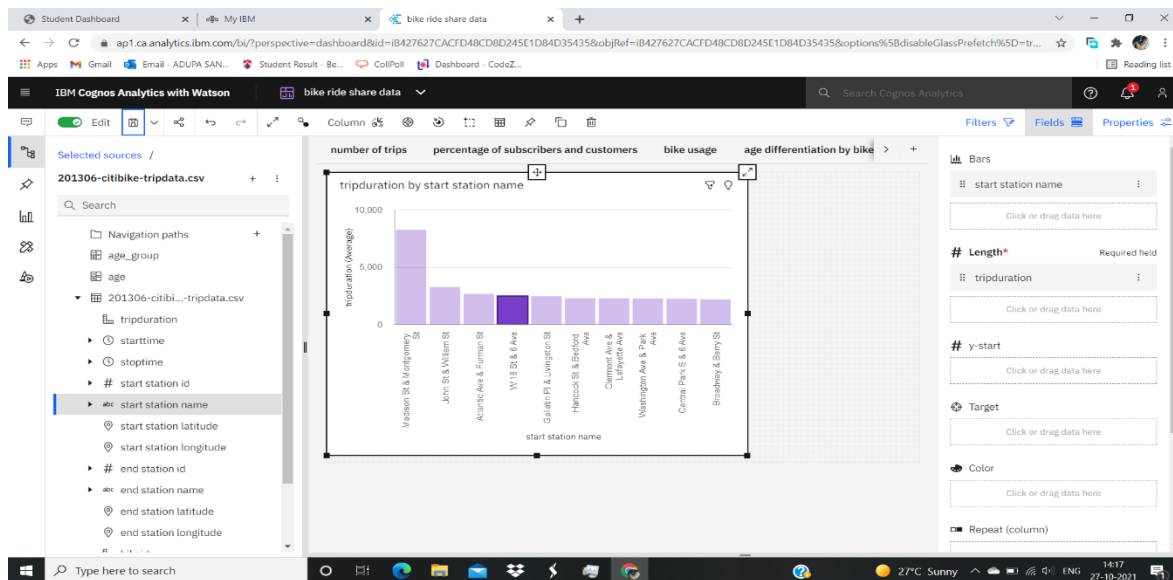
- In the implementation and result we have gone through the method of applying the different important concepts like:
- Trip Duration (seconds)
- Stop Time and Date
- Start Station Name
- End Station Name
- Station ID
- Station Lat/Long
- Bike ID
- User Type (Customer = 24-hour pass or single ride)
- Gender (Zero=unknown; 1=male; 2=female)
- Year of Birth

5.2 RESULT ANALYSIS:

The kinds of questions we wanted to answer included ones like these: What's the most common ride duration? What times of the day does the system get the most usage? How much does ridership vary over the course of a month? What are the most used stations? How old are the riders?

While the answers to these questions can be found in the trip data files, the data needs to be augmented to provide easy answers. For example the trip duration in seconds is too granular; minutes would be more useful.

Over the years I used this data for numerous presentations to customers and at user group meetings. And the cleansed data I created was used by the product managers for a visualization tool for their own presentations.



6.TESTING AND VALIDATION

6.1 INTRODUCTION:

Number Of Trips

Using the Citi-Bike Analysis dataset, we plan to create a dashboard showing the sales and profits for different segments and Sub-Category of products across all the states. To achieve this objective, the following are the steps.

What will be the number of trips in case of city bike dataset?

For visualizing the number of trips, we will require the following data.

- Bike id
- Trip duration
- Bike cc
- Gender
- City
- Age
- Manufacture Age
- Model

6.2 Customer And Subscriber With Gender:

What is the Customer and subscriber with gender?

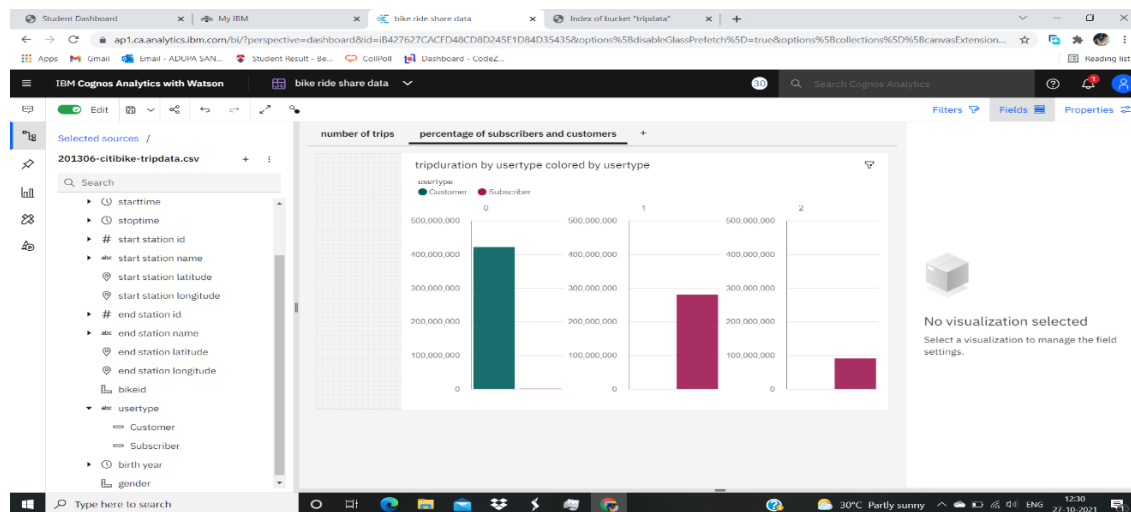
The column User type consists of Subscribers and Customers categories. Let us plot a stacked bar chart for showing the different gender count of the user type with the trip duration.

Find the top bike used with respect to trip duration?

Let us plot a simple bar chart to show the top 10 most used Bike id along with its trip durations.

Calculating the number of bikes used by each age group.

We only have the birth year in our data. So, let us calculate the age of the person by subtracting the birth year from the current year to make visuals easier to interpret. Then let us show a text table for age group differentiation by bike.



7.CONCLUSION

Project conclusion:

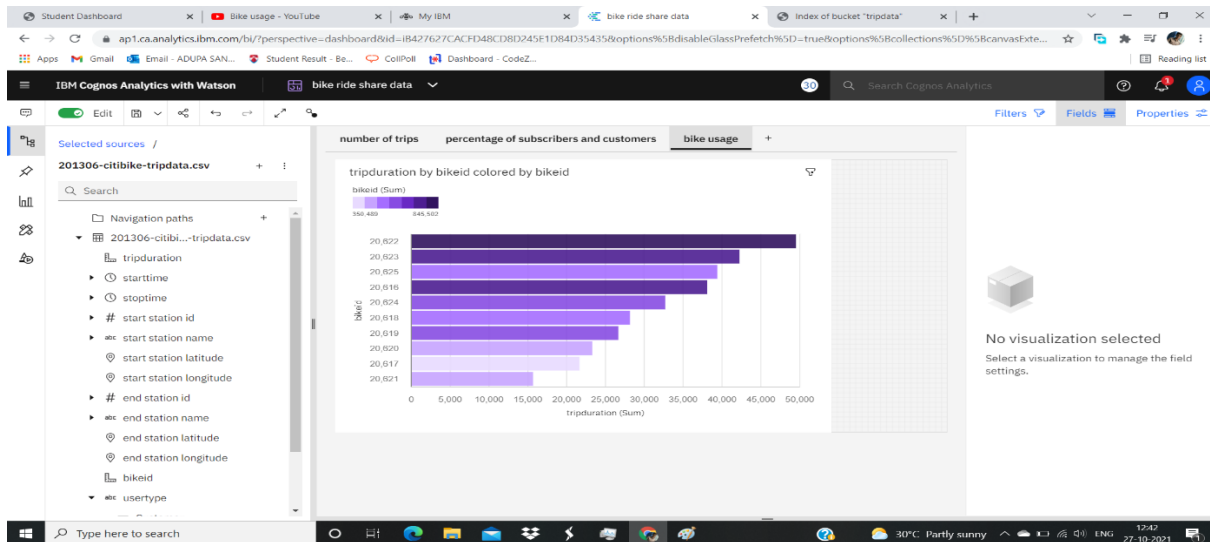
In this project we have to create a IBM Cloud Account and login to the IBM Cognos Analytics.

Here in this project Total number of trips, what is customer and subscriber with gender. Find the top bike used with respect to trip duration. calculating the number of bikes used by respective age groups top 10 Start Station Names with respect to Customer age group .know fundamental concepts and can work on IBM Cognos Analytics. Gain a broad understanding of plotting different graphs. Able to create meaningful dashboards

From the above report we can conclude that the overall performances at different departments can be displayed and can be understandable by seeing the Dashboard.

we can conclude that the overall performances at different departments can be displayed and can be understandable Customer age group .know fundamental concepts and can work on IBM Cognos Analytics. Gain a broad understanding

Different segments and Sub-Category of products across all the states. To achieve this objective



8. FUTURE ENHANCEMENT

In future we will get lot of bikes but present we don't what cc bikes are good, so here we know all the brand bikes which is gonna better cc and which is fast bike and let me know which gender is most participating in this race



REFERENCES

1. <https://s3.amazonaws.com/tripdata/index.html>