IBM Cognos to Analyze and Visualize New York City Bike Ride Share Data

Developed by : Team Hex Clans

Team Members : Anusha R V, Chandumani S, Fareena S, Gagandeep B, Zaiba Muskan

Smart Bridge Remote Internship Program

1.INTRODUCTION

1.1 Overview

Data visualiza on is the graphical representa on of informa on and data. By using visual elements like charts, graphs and maps data visualiza on tools provide an accessible way to see and understand trends, outliers and pa erns in data. Data visualiza on helps us to tell stories by cura ng data into a form easier to understand, highligh ng the trends and outliers.

IBM Cognos Analy cs provides analy c insights that helps us to detect and validate important rela onships and meaningful differences based on the data that is presented by the visualiza on. IBM Cognos Analy cs provides a number of recommended visualiza ons based on the data that we are working with. IBM Cognos Analy cs integrates repor ng, modelling, analysis, explora on, dashboards, stories and event management so you can understand your organiza ons data and make effec ve business decisions.

1.2 Purpose

Our aim from the project is to make use of IBM Cognos Analy cs to analyze, visualize and create an opera ng report on New York City Bike Ride Share data for the year 2018. And in the end we will be crea ng a dashboard which displays the result of analysis.

2.LITERATURE SURVEY

2.1 Existing Problem

1. The oversimplification of data:One of the biggest draws of visualization is its ability to take big swaths of data and simplify them to more basic, understandable terms. However, it's easy to go too far with this; trying to take millions of data points and confine their conclusions to a handful of pictorial representations could lead to unfounded conclusions, or completely neglect certain significant modifiers that could completely change the assumptions you walk away with.

2. The human limita ons of algorithms: This is the biggest potential problem, and also the most complicated. Any algorithm used to reduce data to visual illustrations is based on human inputs, and human inputs can be fundamentally flawed.

3. Overreliance on visuals: This is more of a problem with consumers than it is with

developers, but it undermines the potential impact of visualization in general. When users start relying on visuals to interpret data, which they can use at-a-glance, they could easily start over-relying on this mode of input.

4. The inevitability of visualiza on: Already, there are dozens of tools available to help us understand complex data sets with visual diagrams, charts, and illustrations, and data visualization is too popular to ever go away. We're on a fast course to visualization taking over in multiple areas, and there's no real going back at this point. To some, this may not seem like a problem, but consider some of the effects—companies racing to develop visualization products, and consumers only seeking products that offer visualization. These effects may feed into user overreliance on visuals, and compound the limitations of human errors in algorithm development.

2.2 Proposed Solution

By using the IBM Cognos Analy cs we can overcome the exis ng problems and some of the benefits include:

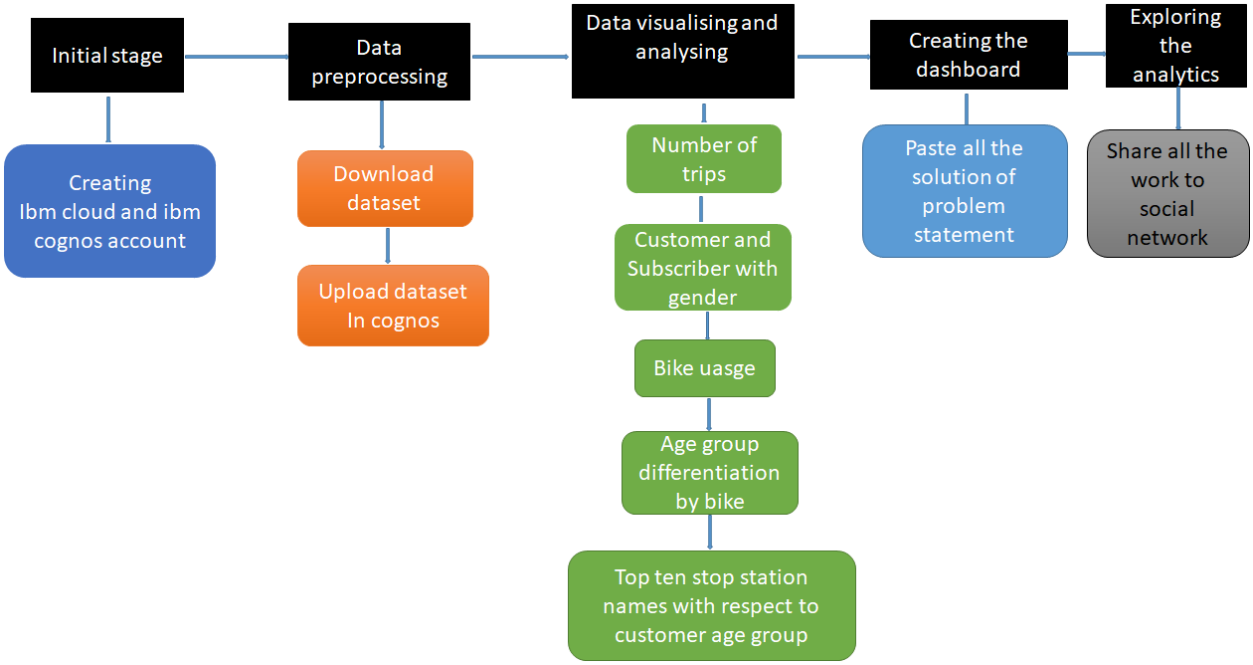
1. Use of varied data source.
2. An intuitive and straightforward interface.
3. Interactive content available online and offline.
4. Personalized experience.
5. Scheduling and alerts.
6. Smart search works in context.

These benefits create a strong case for using IBM Cognos over other solu ons.

IBM Cognos Analy cs provides a quick overview of rela onships among pairs of fields that focus on a single field of interest. Visualiza on comprises mul ple tabs, each for a different field of interest. This informa on is very useful in orien ng you regarding a mul tude of relevant rela onships available in data to be explored further as needed.

3.THEORETICAL ANALYSIS

3.1 Block Diagram



3.2 Hardware/Software Designing

So ware requirements

1. Operating System : Microsoft Windows 10
2. Analytics and Visualization tool : IBM Cognos Analytics

Hardware requirements

3. Main Processor : Intel core i3, Intel core i5
4. RAM Size : 4.00 GB
5. Processor Speed : 2.60 GHz

4.EXPERIMENTAL INVESTIGATIONS

In this project we made analysis on IBM Cognos Analy cs. The explosion of data, mainly unstructured data, over the past few years led to the developed of new type of computer system known as cogni ve system. Unlike the programmable computer that preceeded it.

1. Here we learn about visualisation of graphs
2. Working on fundamental concepts and on IBM Cognos Analytics
3. How to plot the graphs
4. Creating the wide dashboard

IBM Cognos Analy cs provides a number of recommended visualiza on based on the data that you are working with. Highligh ng condi onally forma ed data with colour. Use colour in your table or crosstab visualiza on to see distributed of your data and highlight excep onal data points.

5.FLOW CHART

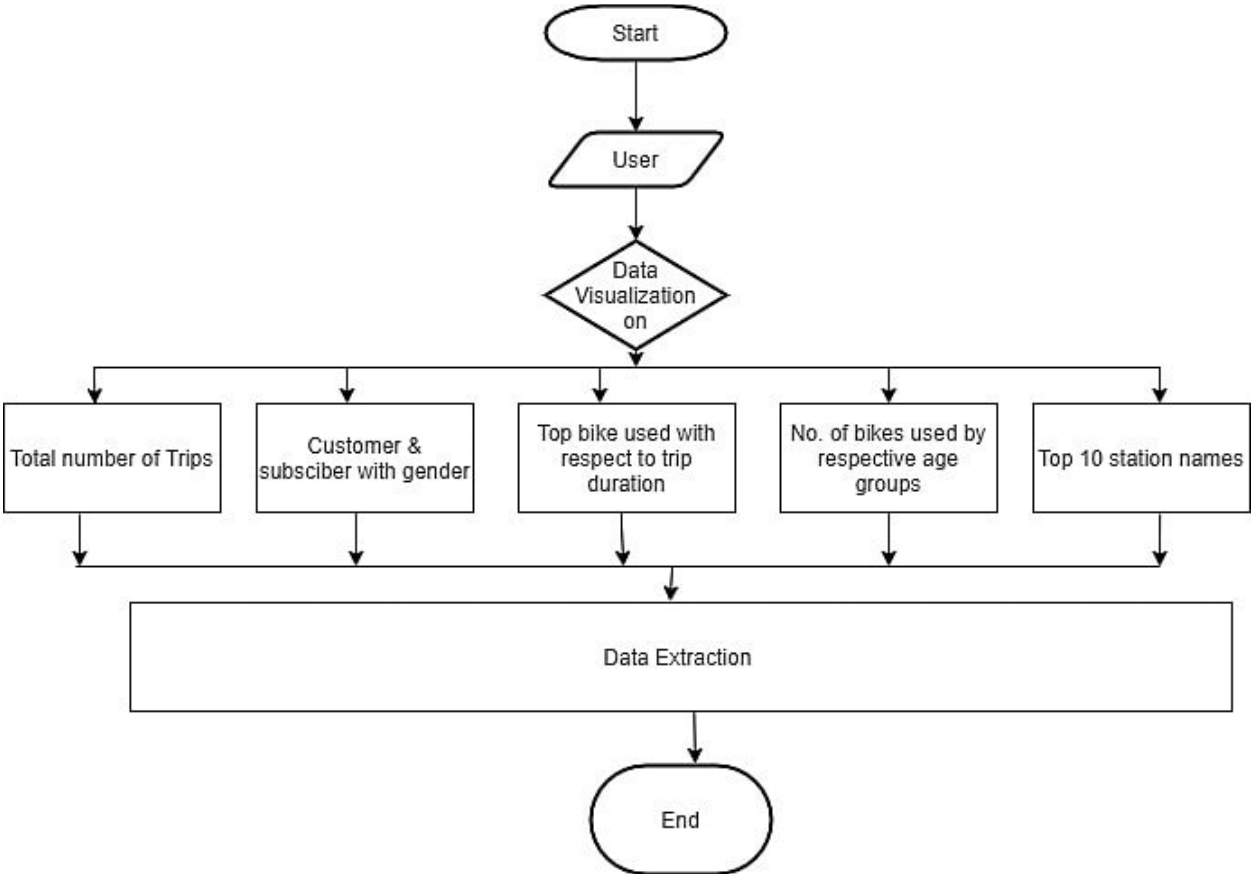


Fig: Diagram showing the control flow of the solu on.

6.RESULT

In this page, Data Visualiza on is used to visualize the data. Following are the snapshots of visualized data.

Snapshots:

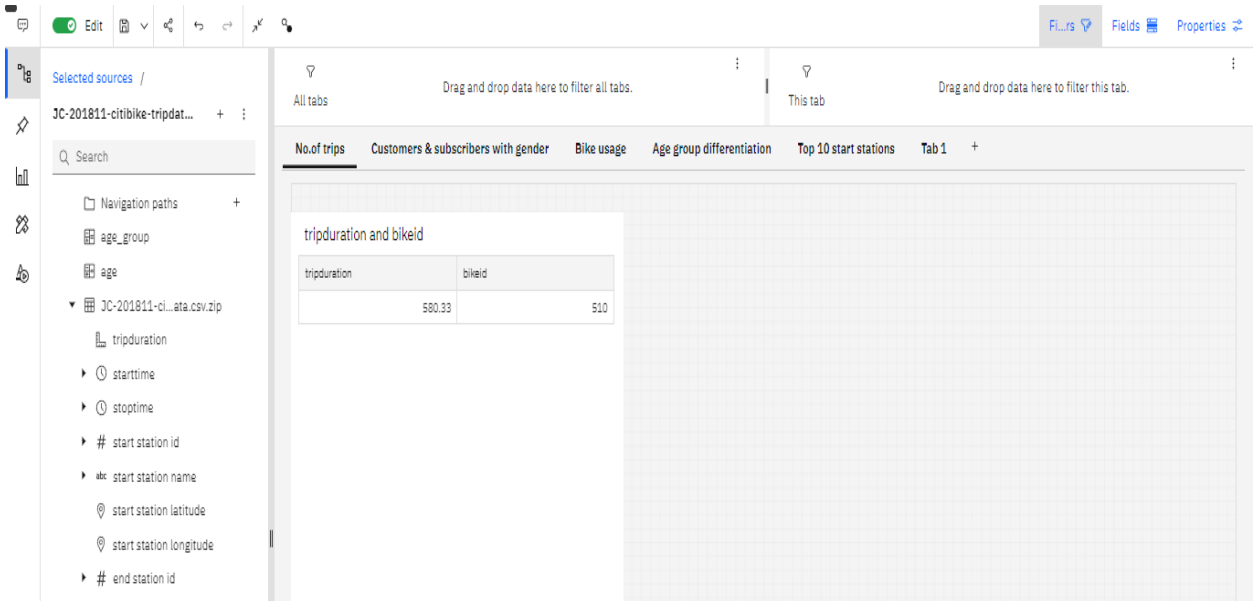


Fig: Snapshot of data visualization on total number of trips

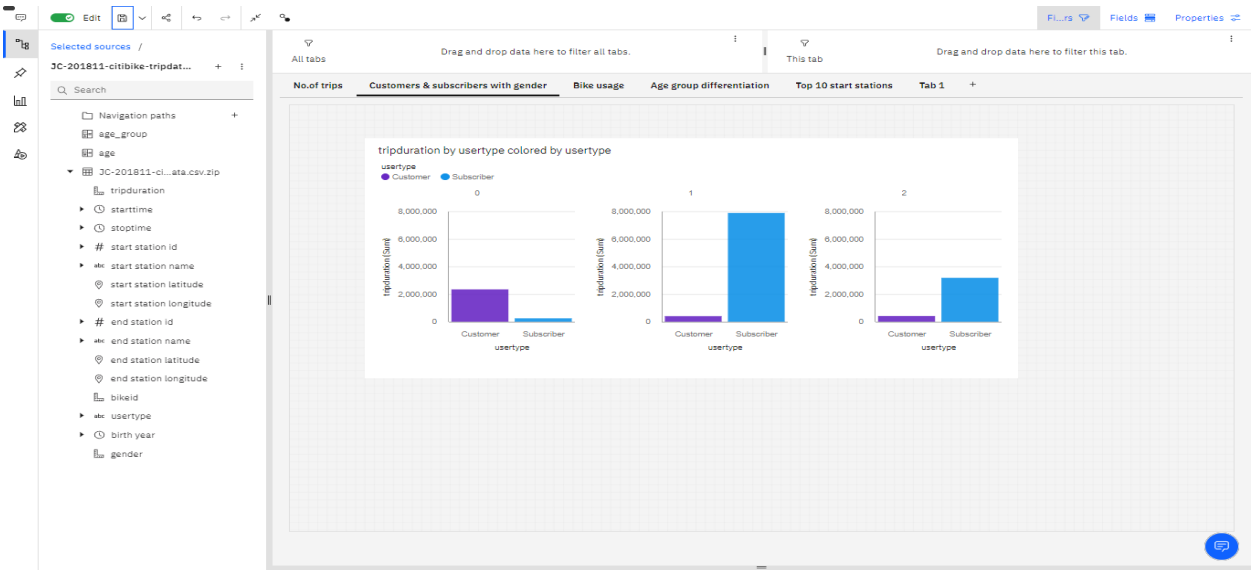


Fig: Snapshot of data visualiza on customers & subscribers with gender

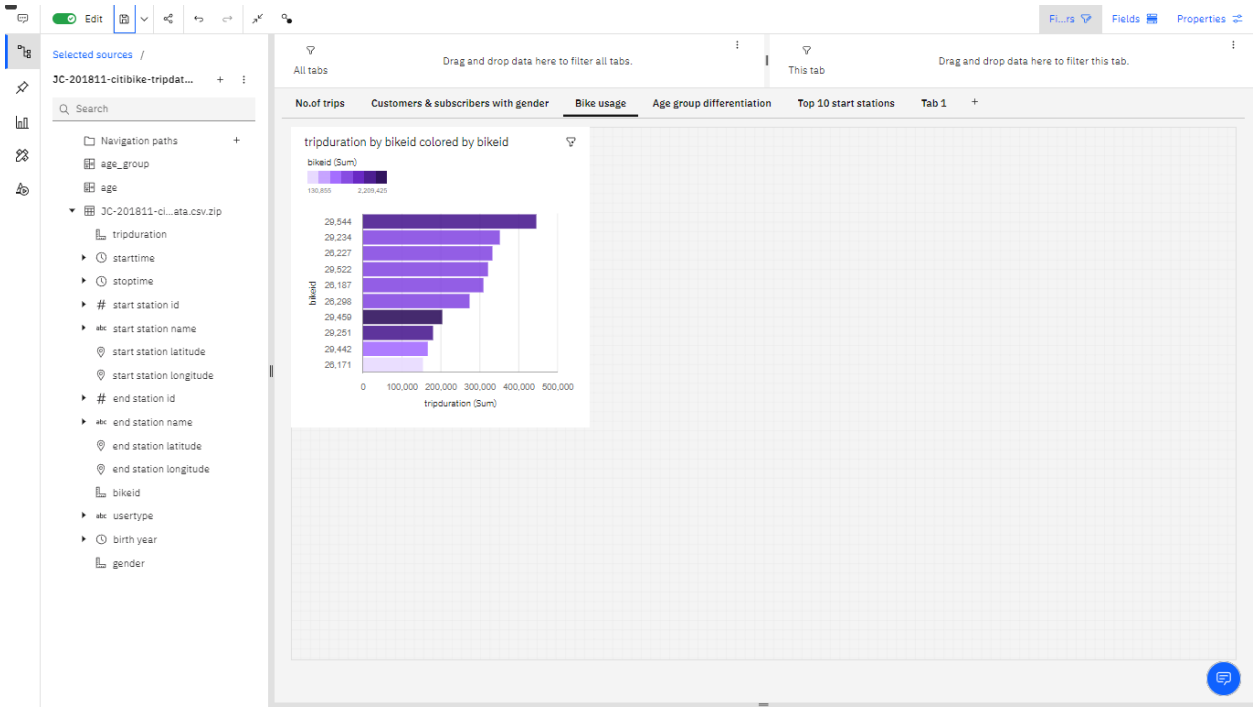


Fig: Snapshot of data visualiza on bike usage

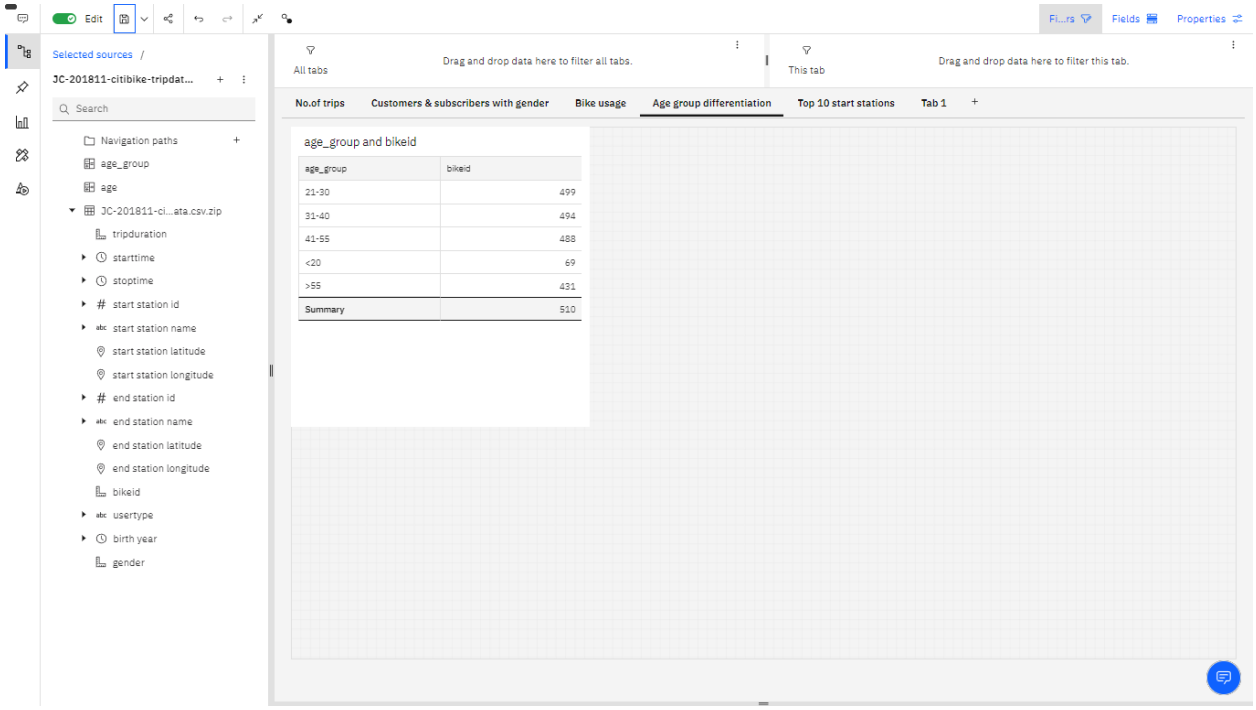


Fig: Snapshot of data visualiza on age group differenence

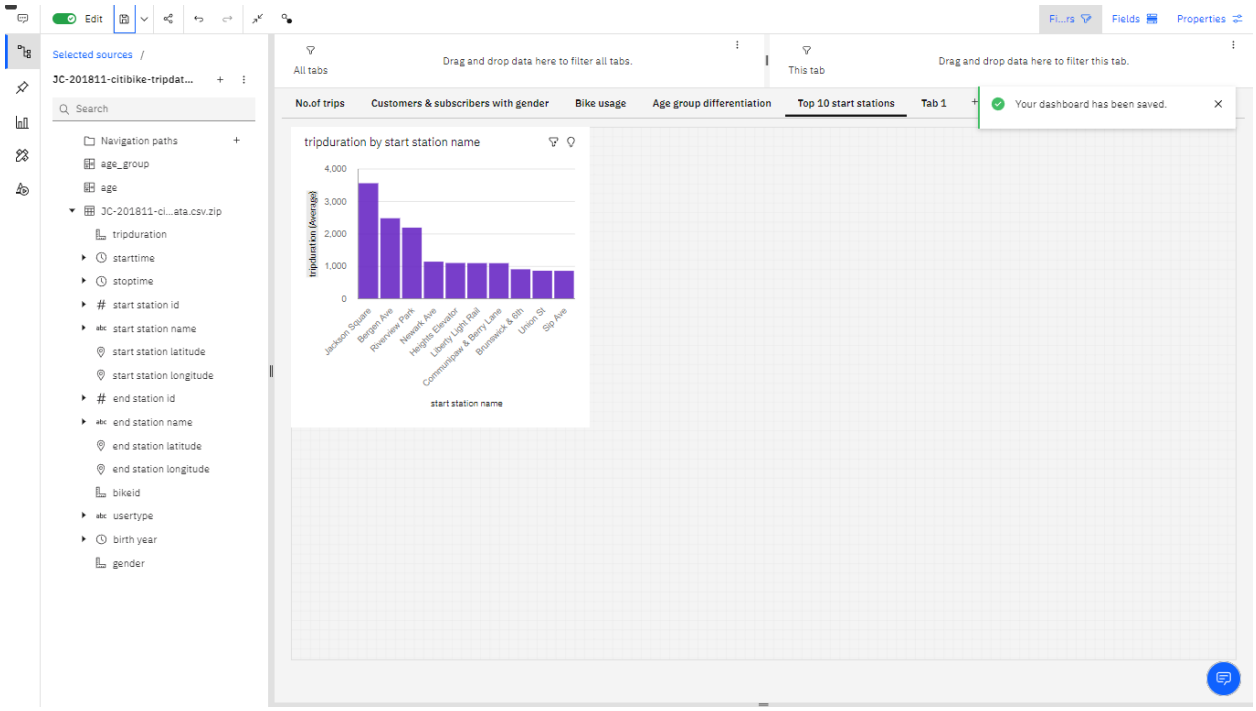


Fig: Snapshot of data visualiza on top 10 start stations

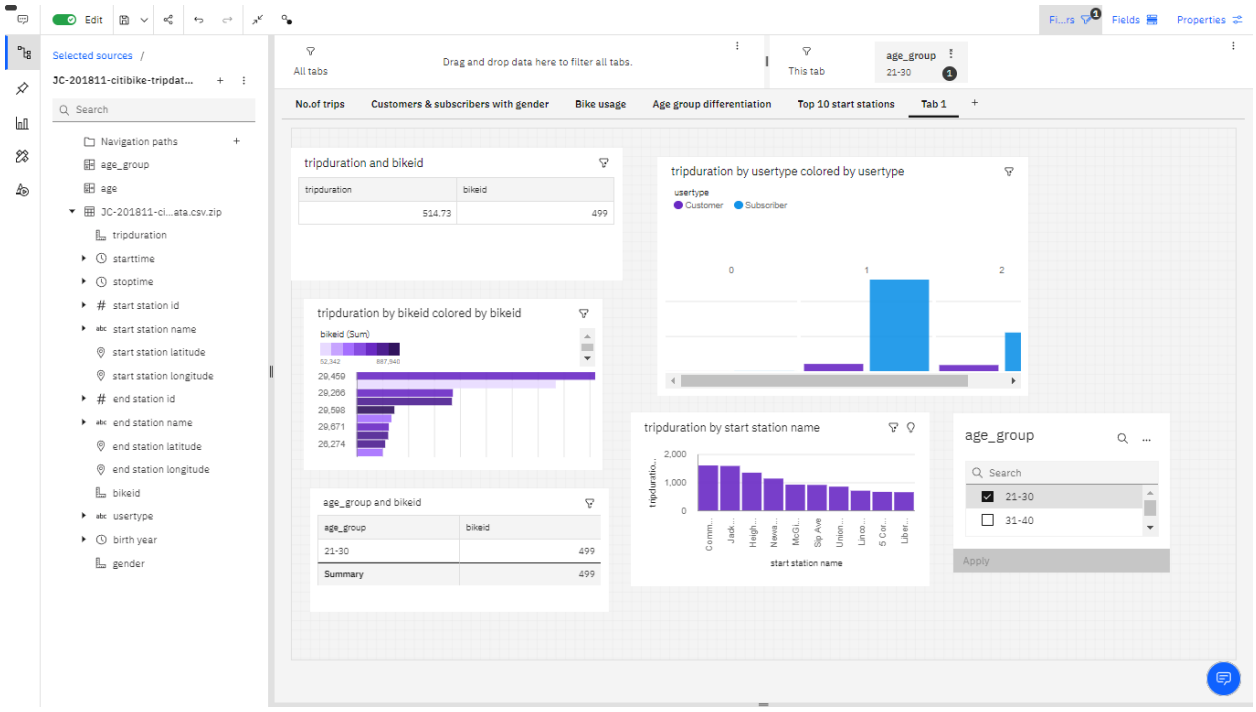


Fig: Snapshot of data visualiza on Dashboard

7.ADVANTAGES AND DISADVANTAGESAdvantages:

There are various advantages to integra ng with Cognos 10.2.2 for genera ng reports. Some of these are:

1. Lower costs—reduces maintenance due to complete report coverage and a zero-footprint environment.
2. Faster results—shortens repor ng me due to seamless integra on and adap ve authoring.
3. Improved decision making—reports and dashboards present data in easily-understood formats.
4. Adap ve authoring automa cally adjusts report layout when objects are added, moved, or removed.
5. Ability to work with data using familiar business terms.
6. Ability to use a variety of charts—crosstabs, bar or 3D bar, pie or doughnut, line, gauge, funnel, sca er, dot density, waterfall, and so forth.
7. Ability to create complex, mul -page layouts using different data sources.
8. High performance data access across all sources.
9. Complete connec vity regardless of environment.
10. Open architecture that leverages XML, SOAP, and WSDL.
11. Mul ple export formats—Excel, Portable Document Format (PDF), Extensible Markup Language (XML), Hypertext Markup Language (HTML), and Comma Separated Value (CSV).
12. Mul lingual capabili es automa cally deliver reports in the users' working language.
13. Ability to integrate seamlessly with the Selling and Fulfillment Founda on, without the user having to log in to the applica on again.

Disadvantages:

Along with the benefits of IBM Cognos Analy cs men oned above, there are a few drawbacks to know about, as well.

Some of the disadvantages are:

14. Total_Cost_o f_Ownership_(TCO) is more significant than other tools
15. Minimal forecast capabilites

16. Investment in Cognos R&D by IBM is declining
17. Won't work smoothly with large data sets having many parameters
18. Cross-browser compatability is often problematic

8.APPLICATIONS

a. The IBM Cognos Analytics applications tier contains one or more Cognos Analytics servers. The servers run requests, such as reports, analyses, and queries that are forwarded by the gateway, and renders the interfaces.

b. IBM Cognos Configuration is used to configure Cognos Analytics, and to start and stop its services.

c. Cognos Analytics portal provides a single access point to the corporate data available for its products.

d. Using the Reporting tool, report authors create, edit, and distribute a wide range of professional reports.

9.CONCLUSION

There does not seem to be a linear rela onship between the average ride and the age of the rider, however there seems to be a be er avenue of explora on along the line of the days for the ride and specifically between whether the ride occurred on a weekday or a weekend. The results of the analysis seem to indicate that on average ride on a weekend (or holidays) are longer than the ride on weekday for the sample of the popula on. The observa ons in our popula on of interest not being independent, we may not be able to infer similar conclusion. However, addi onal analysis based on day and me of rides may proved worth while to inves gate. Also, as future analysis, it would be interes ng to explore any geographical rela onships among the various bike docking sta ons.

10.FUTURE SCOPE

The future of Big Data Visualiza on drama cally increases efficiency and improves efficiency by delivering infographics that can be turned into valuable insights. Nothing is best than becoming able to communicate insights in real me using immersive visuals.

IBM Cognos provides a good vary of op ons and might be considered as an enterprise so ware package to produce versatile report setting and can be used for big and medium enterprise. It meets the requirements of power users, analysts, business managers and company executives.

11.BIBLIOGRAPHY

1. https://www.ibm.com/docs/en/cognos-analytics/11.1.0?topic=analytics-relationships
2. https://www.datasciencecentral.com/profiles/blogs/4-potential-problems-with-data visualization

APPENDIX

The link provided below gives the access to the dashboard and the data visualization charts we have created:

https://eu1.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.my_folders%2FNo%2Bof%2Btrips&action=view&mode=dashboard&subView=model0000017cd42159e0_00000000