

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 visualization is the graphical representation of information and data. By using visual elements like charts, graphs and maps data visualization tools provide an accessible way to see and understand trends, outliers and patterns in data. Data visualization helps us to tell stories by curating data into a form easier to understand, highlighting the trends and outliers.

IBM Cognos Analytics provides analytic insights that helps us to detect and validate important relationships and meaningful differences based on the data that is presented by the visualization. IBM Cognos Analytics provides a number of recommended visualizations based on the data that we are working with. IBM Cognos Analytics integrates reporting, modelling, analysis, exploration, dashboards, stories and event management so you can understand your organizations data and make effective business decisions.

1.2 Purpose

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

2.LITERATURE SURVEY

2.1 Existing Problem

- **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.
- **The human limitations 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.
- **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.

- The inevitability of visualization: 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 Analytics we can overcome the existing problems and some of the benefits include:

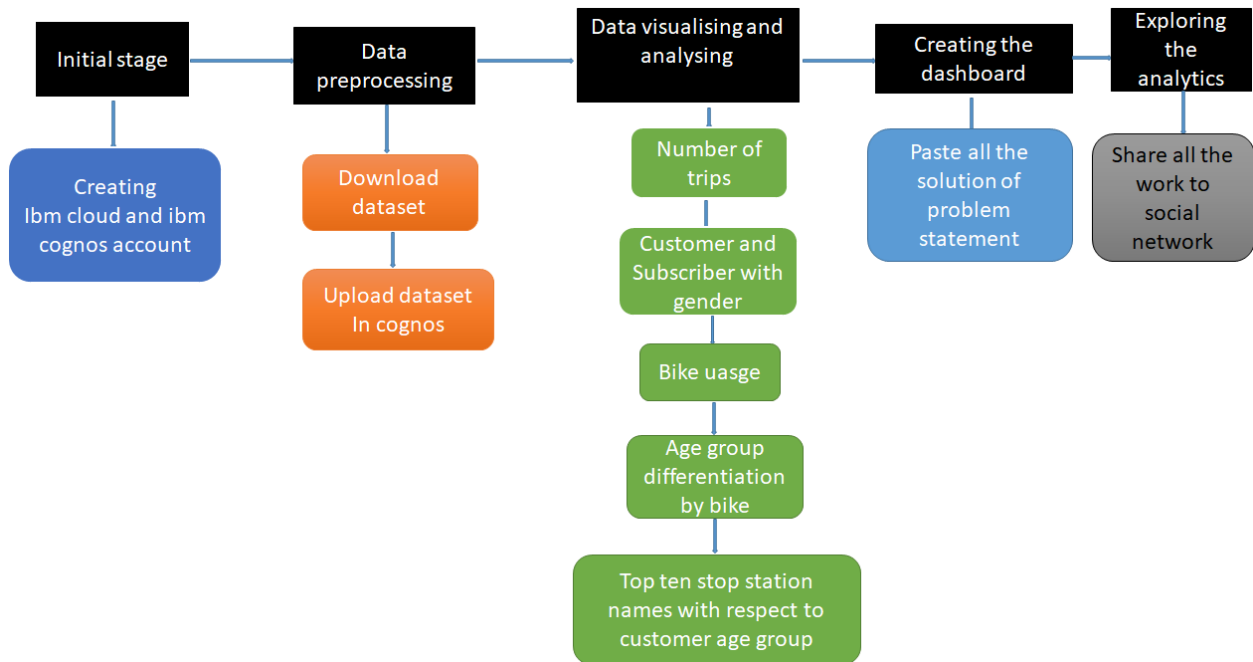
- Use of varied data source.
- An intuitive and straightforward interface.
- Interactive content available online and offline.
- Personalized experience.
- Scheduling and alerts.
- Smart search works in context.

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

IBM Cognos Analytics provides a quick overview of relationships among pairs of fields that focus on a single field of interest. Visualization comprises multiple tabs, each for a different field of interest. This information is very useful in orienting you regarding a multitude of relevant relationships available in data to be explored further as needed.

3.THEORETICAL ANALYSIS

3.1 Block Diagram



3.2 Hardware/Software Designing

Software requirements

- Operating System : Microsoft Windows 10
- Analytics and Visualization tool : IBM Cognos Analytics

Hardware requirements

- Main Processor : Intel core i3, Intel core i5
- RAM Size : 4.00 GB
- Processor Speed : 2.60 GHz

4.EXPERIMENTAL INVESTIGATIONS

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

- Here we learn about visualisation of graphs
- Working on fundamental concepts and on IBM Cognos Analytics
- How to plot the graphs
- Creating the wide dashboard

IBM Cognos Analytics provides a number of recommended visualization based on the data that you are working with. Highlighting conditionally formatted data with colour. Use colour in your table or crosstab visualization to see distributed of your data and highlight exceptional data points.

5.FLOW CHART

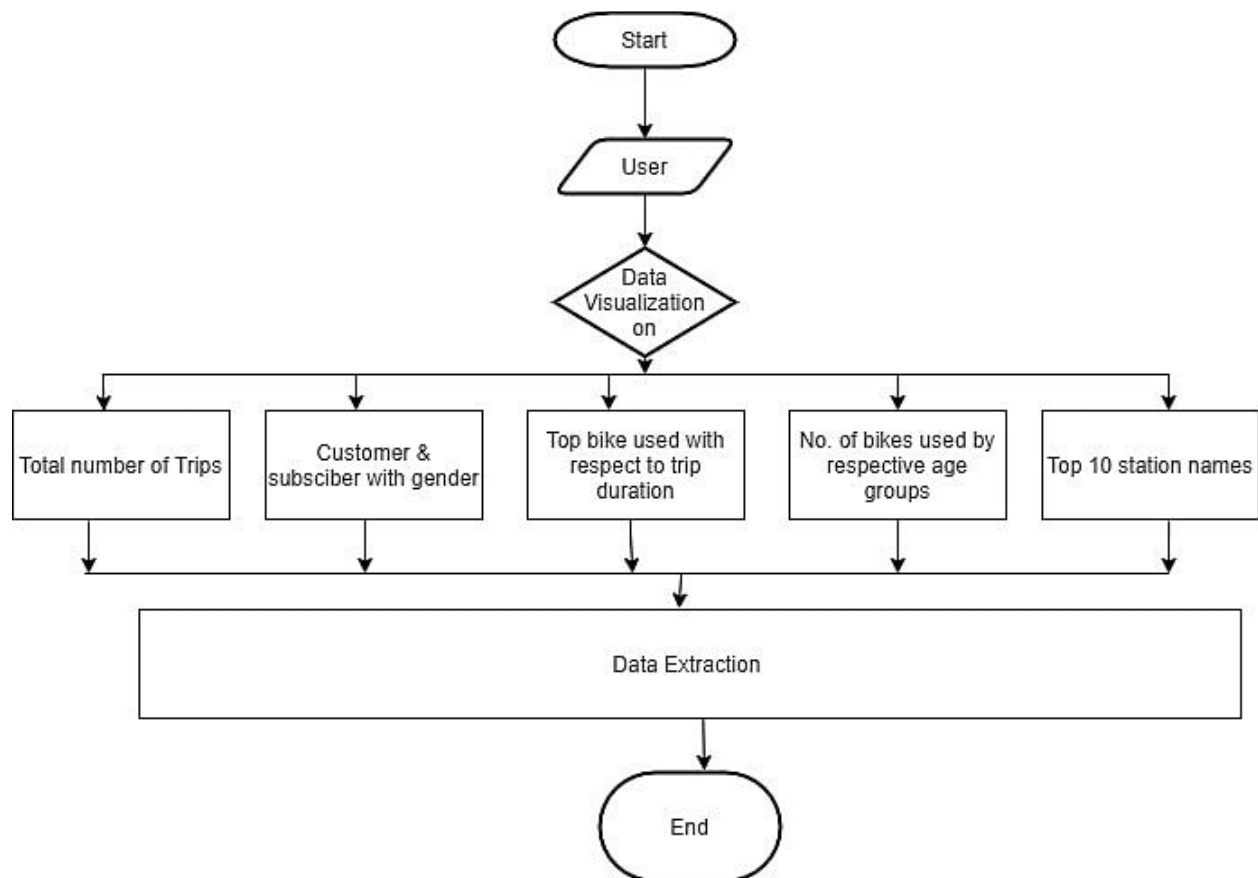


Fig: Diagram showing the control flow of the solution.

6.RESULT

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

Snapshots:

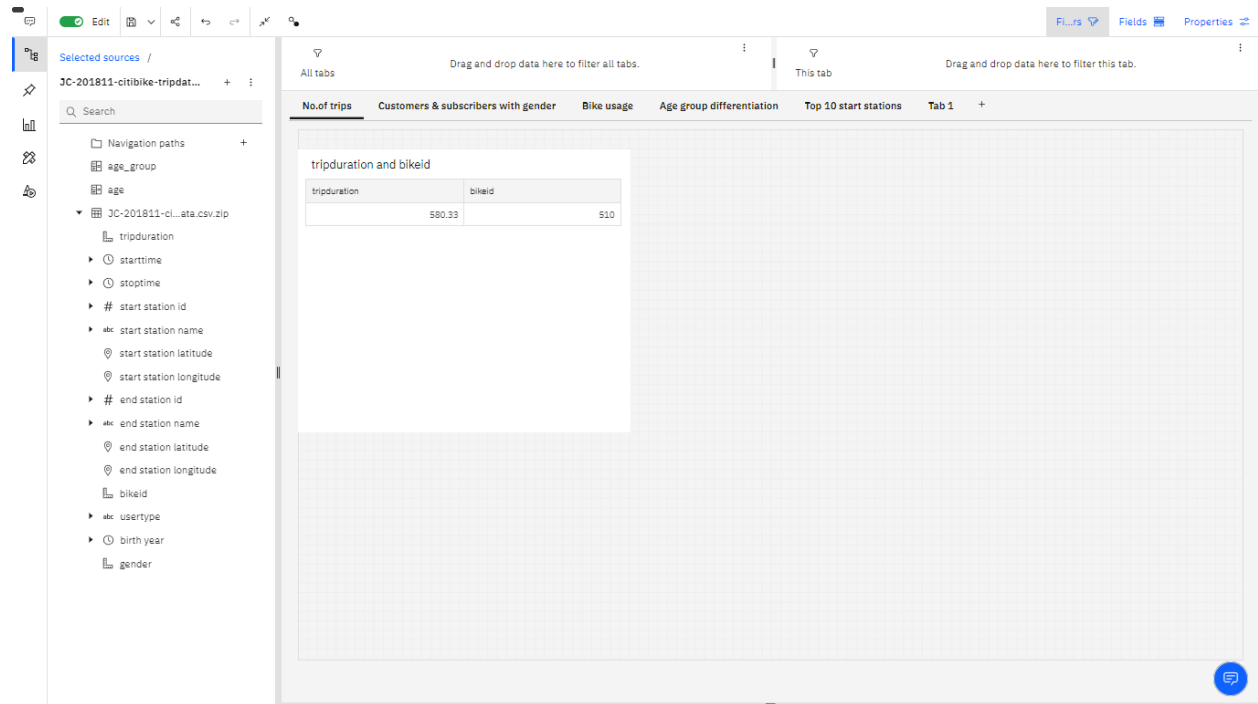


Fig: Snapshot of data visualization on total number of trips

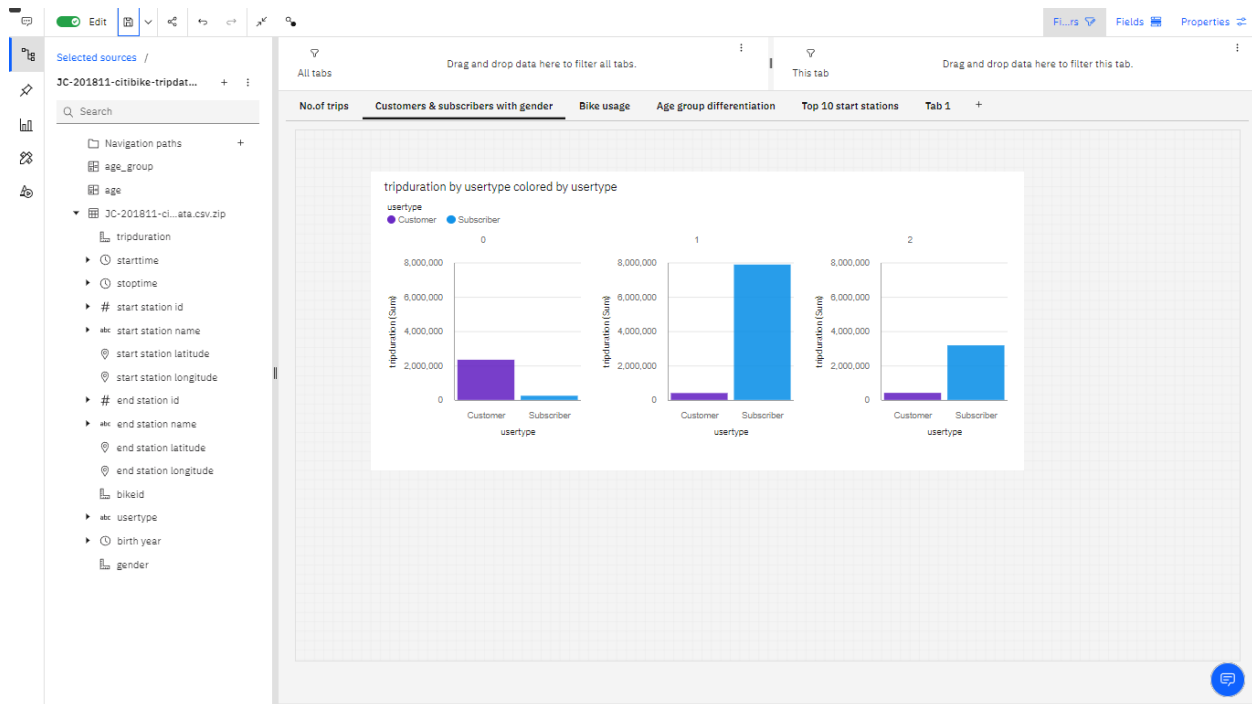


Fig: Snapshot of data visualization on customers & subscribers with gender

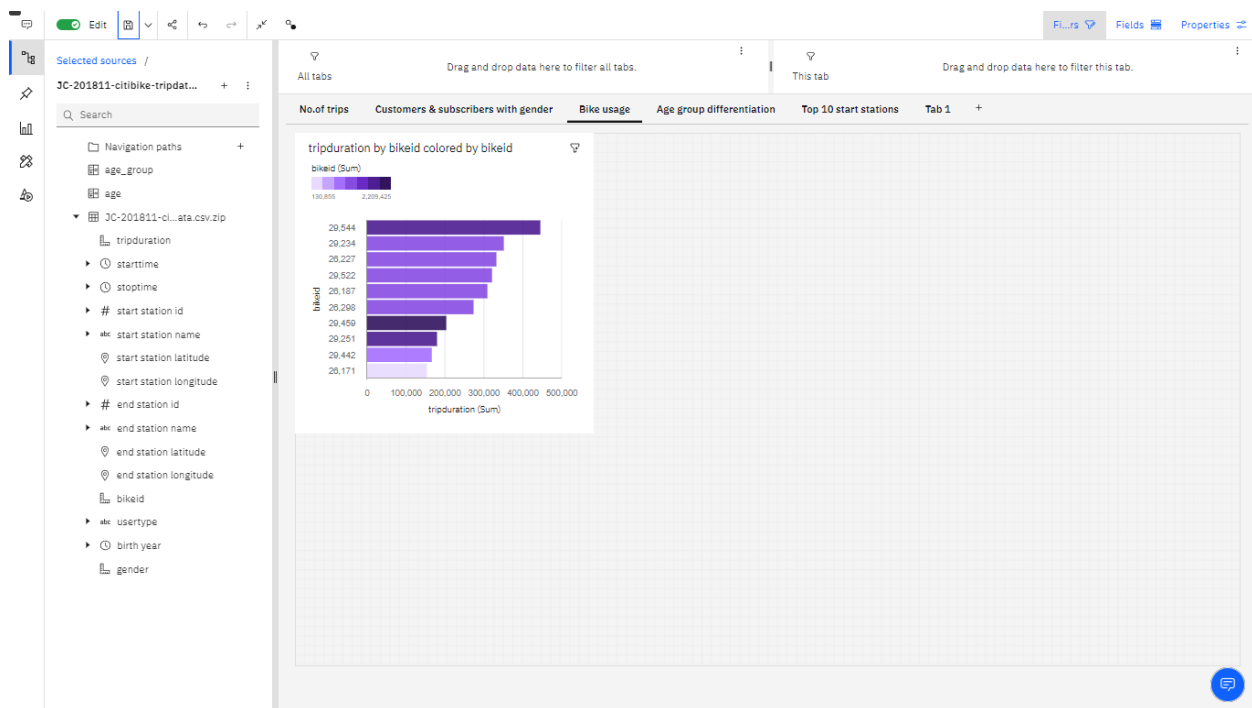


Fig: Snapshot of data visualization on bike usage

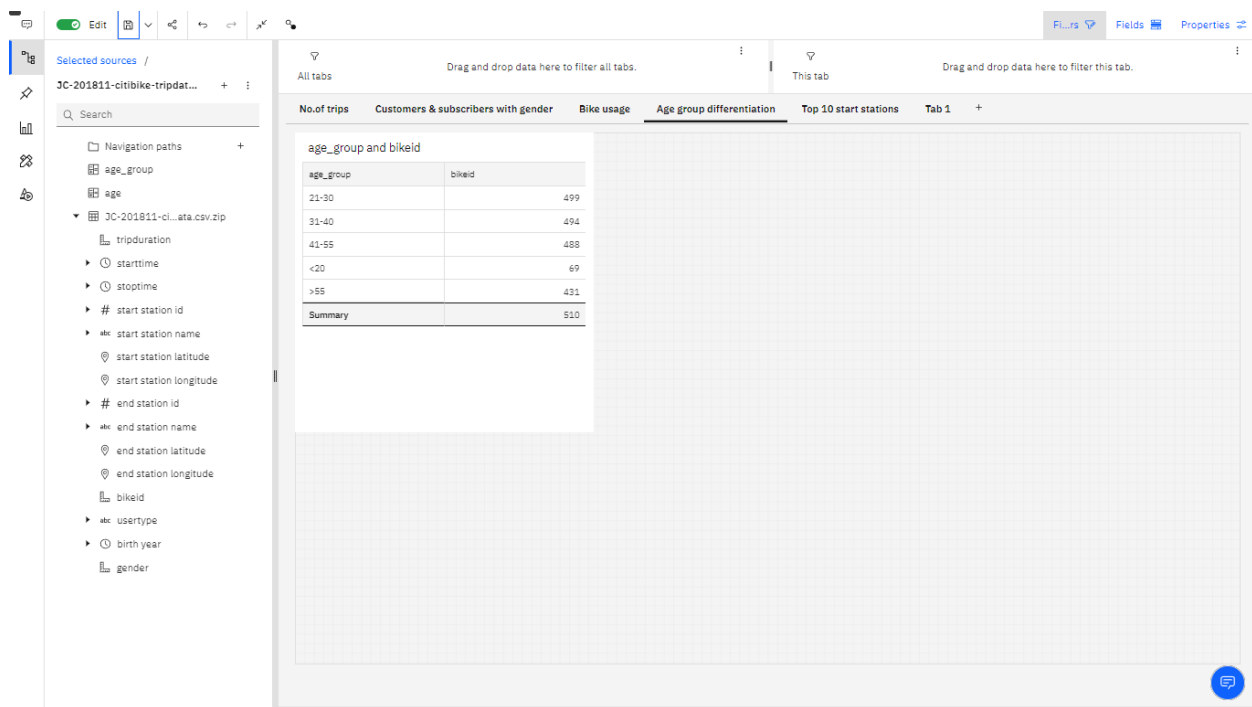


Fig: Snapshot of data visualization on age group differentiation

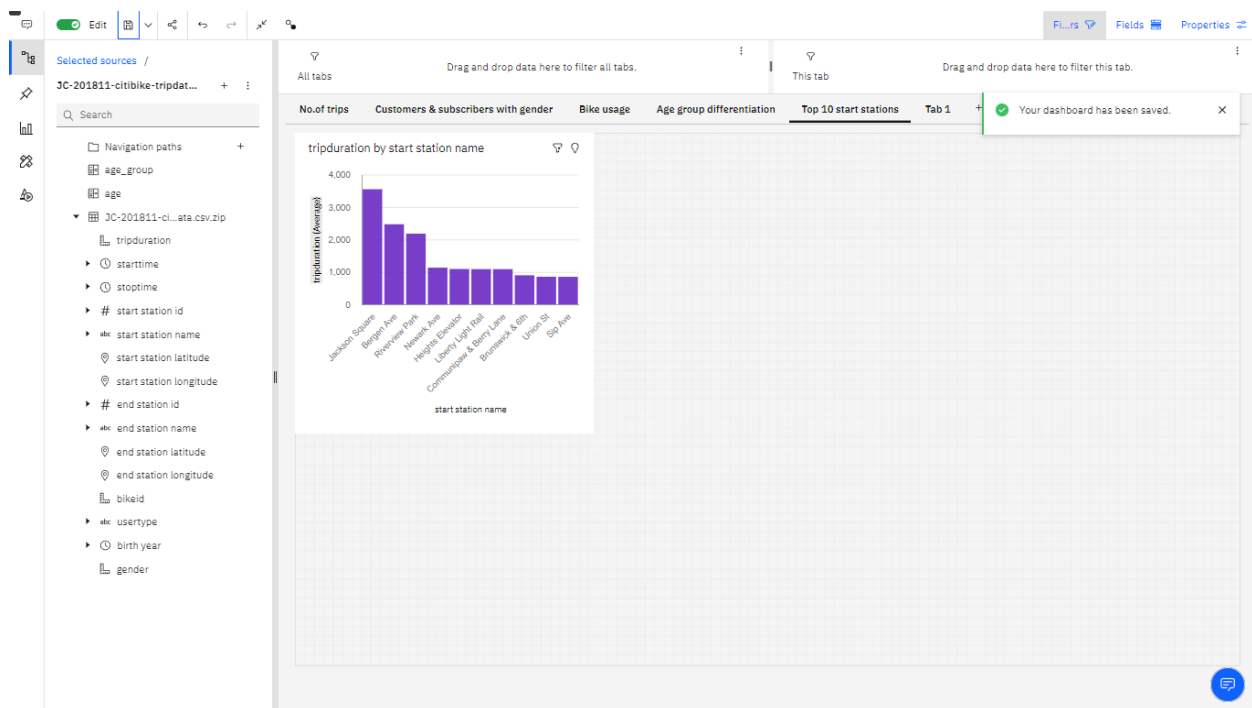


Fig: Snapshot of data visualization on top 10 start stations

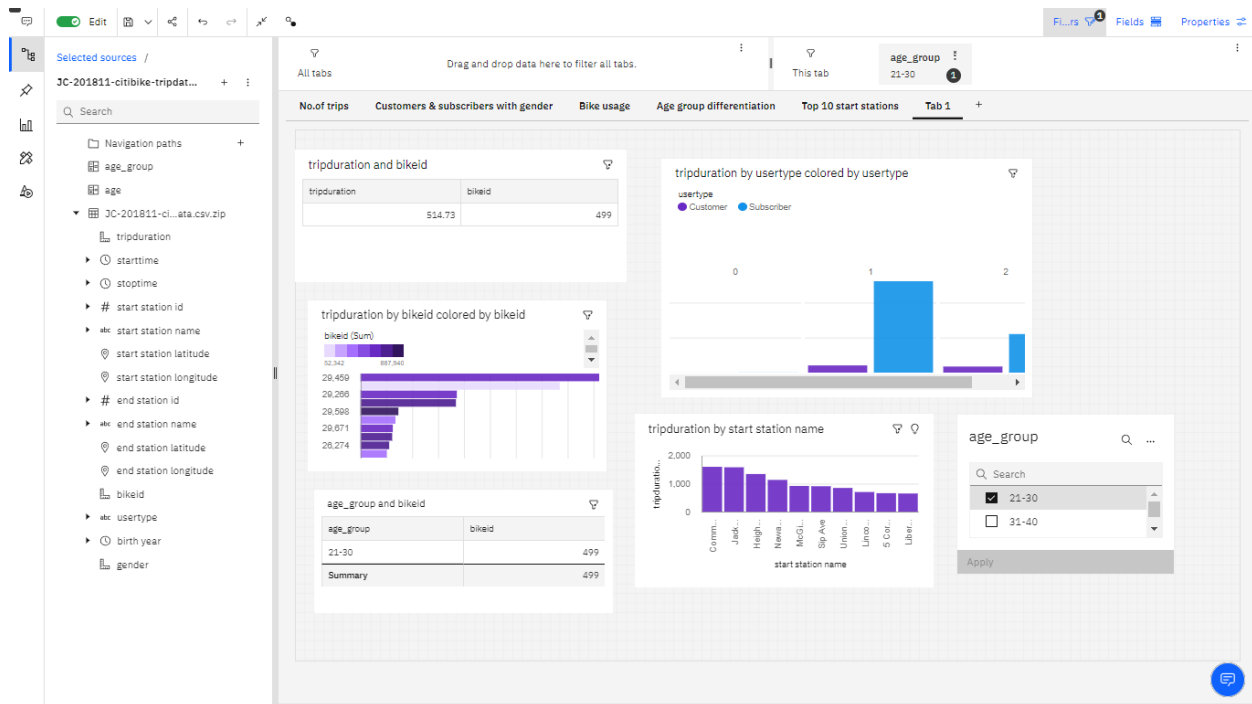


Fig: Snapshot of data visualization on Dashboard

7.ADVANTAGES AND DISADVANTAGES

Advantages:

There are various advantages to integrating with Cognos 10.2.2 for generating reports. Some of these are:

- Lower costs—reduces maintenance due to complete report coverage and a zero-footprint environment.
- Faster results—shortens reporting time due to seamless integration and adaptive authoring.
- Improved decision making—reports and dashboards present data in easily-understood formats.
- Adaptive authoring automatically adjusts report layout when objects are added, moved, or removed.
- Ability to work with data using familiar business terms.
- Ability to use a variety of charts—crosstabs, bar or 3D bar, pie or doughnut, line, gauge, funnel, scatter, dot density, waterfall, and so forth.

- Ability to create complex, multi-page layouts using different data sources.
- High performance data access across all sources.
- Complete connectivity regardless of environment.
- Open architecture that leverages XML, SOAP, and WSDL.
- Multiple export formats—Excel, Portable Document Format (PDF), Extensible Markup Language (XML), Hypertext Markup Language (HTML), and Comma Separated Value (CSV).
- Multilingual capabilities automatically deliver reports in the users' working language.
- Ability to integrate seamlessly with the Selling and Fulfillment Foundation, without the user having to log in to the application again.

Disadvantages:

Along with the benefits of IBM Cognos Analytics mentioned above, there are a few drawbacks to know about, as well.

Some of the disadvantages are:

- **Total Cost of Ownership (TCO)** is more significant than other tools
- Minimal forecast capabilities
- Investment in Cognos R&D by IBM is declining
- Won't work smoothly with large data sets having many parameters
- Cross-browser compatibility 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 relationship between the average ride and the age of the rider, however there seems to be a better avenue of exploration 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 population. The observations in our population of interest not being independent, we may not be able to infer similar conclusion. However, additional analysis based on day and time of rides may proved worth while to investigate. Also, as future analysis, it would be interesting to explore any geographical relationships among the various bike docking stations.

10.FUTURE SCOPE

The future of Big Data Visualization dramatically 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 time using immersive visuals.

IBM Cognos provides a good vary of options and might be considered as an enterprise software 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

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