**CITY BIKE DATA ANALYSIS**

**USING IBM CLOUD COGNO’S DASHBOARD**

An Internship Project Report

Submitted By

(BATCH NO: CSE\_AIML\_D15)

Y.HEMASRI (18481A05N4)

                              V.MANJUSHA (18481A05M0)

S.PREM KUMAR NAIDU (18481A05I9)

            V.VENKATA NAGESWARARAO (18481A05M1)

**INDEX**

**TITLE PAGE NO**

**CHAPTER 1: INTRODUCTION 3-4**

* 1. OVERVIEW 3-4
  2. PURPOSE 4

**CHAPTER 2: LITERATURE SURVEY 5**

* 1. EXISTING PROBLEM 5
  2. PROPOSED SOLUTION 5

**CHAPTER 3: THEORTICAL ANALYSIS 6-8**

3.1 BLOCK DIAGRAM 6

3.2 HARDWARE/SOFTWARE DESIGNING 7-8

**CHAPTER 4: EXPERIMENTAL INVESTIGATIONS 9-10**

**CHAPTER 5: FLOW CHART 11-12**

**CHAPTER 6: RESULTS 13**

**CHAPTER 7: ADVANTAGES AND DISADVANTAGES 14**

**CHAPTER 8: APLLICATIONS 15**

**CHAPTER 9: CONCLUSION 15**

**CHAPTER 10: FUTURE SCOPE 16**

**CHAPTER 11: BIBILOGRAPHY 17**

**CHAPTER12: APPENDIX 17**

**CHAPTER 1**

**INTRODUCTION**

* 1. **Overview:**

The goal of this analysis is to create an operating report of Citi Bike for the year of 2018. Let’s understand the data we’re working with and give a brief overview of what each feature represents or should represent.

1. Trip Duration (seconds) - How long a trip lasted

2. Start Time and Date - Self-explanatory

3. Stop Time and Date - Self-explanatory

4. Start Station Name - Self-explanatory

5. End Station Name - Self-explanatory

6. Station ID - Unique identifier for each station

7. Station Lat/Long - Coordinates

8. Bike ID - unique identifier for each bike

9. User Type (Customer = 24-hour pass or 3-day pass user; Subscriber = Annual Member) - Customers are usually tourists, subscribers are usually NYC residents

10. Gender (Zero=unknown; 1=male; 2=female) - Usually unknown for customers since they often sign up at a kiosk

11. Year of Birth - Self-entered, not validated by an ID.

The final aim of City Bike Data Analysis using IBM Cloud Congo’s dashboard is to create the visualizations in a dashboard and export it.

To accomplish this, we have to complete all the activities and tasks listed below

* IBM Cloud Account
* Login to Cogno’s Analytics
* Working with the Dataset
  + Understand the Dataset
  + Loading the Dataset
* Data visualization charts
  + Problem Statement 1: Number of trips
  + Problem Statement 2: Percentage of Subscribers and Customers
  + Problem Statement 3: Bike Usage
  + Problem Statement 4: Age Group differentiation by Bike
  + Problem Statement 5: Top 10 Start Station Names with Customer age group
* Dashboard Creation
* Export the Analytics

**1.2 Purpose:**

By city bike data analysis using IBM Cogno’s dashboard we will:

* Know fundamental concepts and can work on IBM Cogno’s Analysis.
* Gain a broad understanding of plotting different graphs.
* Able to create meaningful dashboards.

**CHAPTER 2**

**LITERATURE SURVEY**

* 1. **Existing Problem:**

Collecting the data of the bikes is more difficult. Creating the dashboard and visualization the data is more difficult. This bike-sharing service used data and AI to manage and maintain 12,000 bicycles. This global digital agency used IBM Cloud Pak for Data to build a dashboard for eliminating data silos and integrating AI predictive analytics models help.

Finding the top bike used with respects to trip duration is difficult. Data points is difficult to identify every single estimations of filed in all purposes of the bike data analysis.

* 1. **Proposed Solution:**

In IBM Cogno’s Analytics, dashboards provide data discovery capabilities. On a dashboard, you can visually explore and interact with your data to identify the key insights for improving data-driven decisions. You can perform data discovery and assemble the information that is most relevant to you in one place, without IT assistance or formal training.

IBM Cogno’s Analytics is a fast, flexible, and complete business intelligence and analytics solution that organizations use to improve decision quality and accelerate decision making. IBM Cogno’s Analytics integrates reporting, modelling, analysis, dashboards, stories, and event. Use a word cloud visualization when you want to see a text-based.

**CHAPTER 3**

**THEORTICAL ANALYSIS**

**3.1 Block Diagram:**

DATA VISUALIZATION CHARTS

LOADING THE DATASET INTO IBM COGNOS ANALYTICS

IBM COGNOS ANALYTICS

>Number of trips

>Customer and Subscriber with gender

>Bike Usage

>Age group differentiation by bike

>Top 10 start station names with respect to customer age group

Creating a dashboard

IBM CLOUD ACCOUNT

**3.2 Hardware/software designing:**

**Software specifications:**

# IBM Cogno’s Analytics 11.1.x (11.1.5, 11.1.6, 11.1.7) Supported Software Environments.

|  |  |
| --- | --- |
| **REQUIREMENT** | **SPECIFICATION** |
| Java™ Runtime Environment (JRE) | An IBM JRE is provided as part of the install with IBM Cognos Analytics on all operating systems. |
| Database | 1.You must have one of the following databases available to store IBM Cognos data:   * Oracle * IBM Db2 * Microsoft SQL Server * Informix®   2. The Easy (previously Ready to Run!) option installs and configures an Informix database as the content store.  3. TCP/IP connectivity is required for all database types. |
| Web browser | For all Web browsers, the following must be enabled:   * cookies * JavaScript   For Microsoft Internet Explorer only, the following must be enabled:   * Run ActiveX controls and plug-ins * Script ActiveX controls marked safe for scripting * Active scripting |

**Hardware Specifications:**

|  |  |
| --- | --- |
| **REQUIREMENT** | **SPECIFICATIONS** |
| Operating system | Microsoft Windows  UNIX  Linux® |
| Processing | Minimum: 4 CPU cores for one user. For each deployment, a sizing exercise is highly recommended. |
| RAM | Minimum 10 GB. |
| Operating system specifications | File descriptor limit set to 8192 on UNIX and Linux |
| Disk space | A minimum of 7 GB of free space is required to install the software and 5 GB of free space on the drive that contains the temporary directory used by IBM Cognos components. |

**CHAPTER 4**

**EXPERIMENTAL INVESTIGATIONS**

Analysis or the investigation made while working on the solution:

While working on the solution we investigated on what is city bike analysis, IBM cloud, IBM Cogno’s analytics and how to build different visualizations and creating a dashboard and finally exporting it. The key role on investigation is collection of dataset.

**IBM Cloud Account**:

IBM Acquired soft layer, a public cloud platform, to serve as the foundation for its IaaS offering. In October 2016, IBM rolled the soft layer brand under its Blue mix brand of PaaS offerings, giving users to access both IaaS and PaaS resources from a single console. IBM cloud provides a full-stack, public cloud platform with various products in the catalog, including options for compute, storage, networking, end to end developer solutions for app development, testing and deployment, security databases, and cloud native services.

Creating the IBM cloud account by going to the IBM cloud login page and click create on IBM cloud account. Enter our IBM id and an ID is created based on the email that we enter. Completing the remaining fields with our information and click create account by this the account is created.

**Dataset collection**:

The data collection on crop production by:

* Articulate the problem early.
* Establish data collection.
* Check our data quickly.
* Format data to make it consistent.
* Reduce data.
* Complete data cleaning.
* Decompose data.
* Join transactional and attribute data.

**IBM Cogno’s Analytics:**

IBM Cogno’s analytic is a web-based integrated business intelligence suite by IBM. It provides a tool set for reporting, analytics, score carding and monitoring of events and metrics. Creating amazing meaningful dashboards using Cogno’s analytics.

**Creating a Dashboard:**

Dashboard track KPIs, metrics and other data points in one visual, central place. They give a high level view of work helping to make quick decisions and keeping everyone up to date. To store the dataset into the dashboard we need to import the data into it. Set up our excel dashboard file and create a table with the raw data. Analyze the data and build the dashboard customize with Macros, color and more.

**To export the Dashboard:**

To save the dashboard as image, select image. The select location for download dialog box opens. To save the dashboard as a Flash file, select MHT, then do one of the following. To save the dashboard as a PDF file, select PDF dashboard is exported and displayed in a browser window.

**CHAPTER 5**

**FLOW CHART**



Data Collection



Data

Visualization



Number of trips



Customers and subscribers with gender



Age group differentiation by bike



Bike Usage



Top 10 start station name with respect to customer age group



Creating dashboard

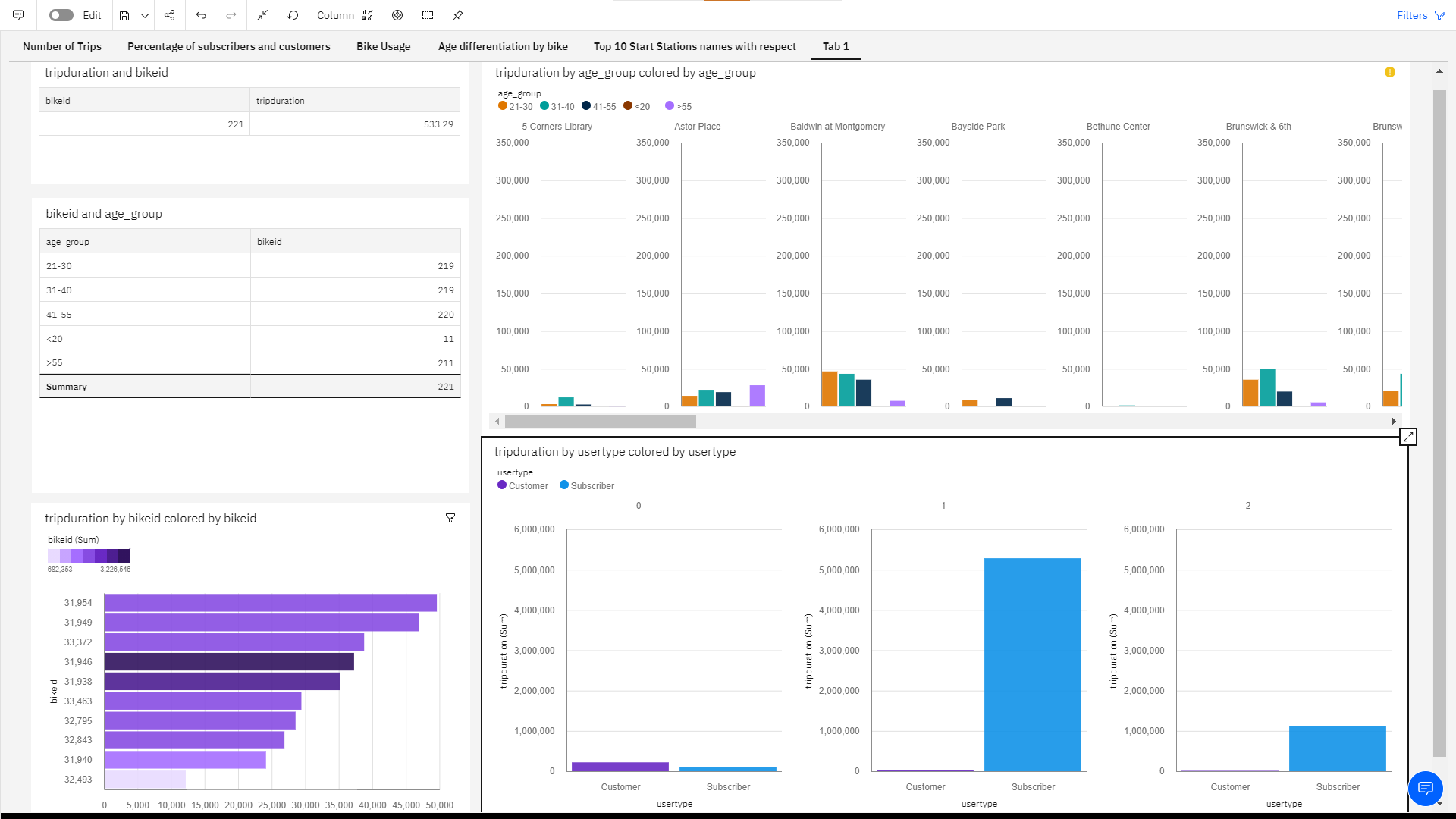


Exporting the dashboard

**CHAPTER 6**

**RESULTS**

**Final output of the project:**

****

**CHAPTER 7**

**ADVANTAGES AND DISADVANTAGES**

**Advantages:**

* 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.
* High performance data access across all sources.
* Ability to create complex, multi-page layouts using different data sources.

**Disadvantages:**

* Data source adapter error
* The permission level for a user cannot be modified
* Data grouping
* Custom color palettes
* Custom visualizations
* Insights in visualization

Lower c Lower costs—reduces maintenance due to complete report co

-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.

High performance data ac **CHAPTER 8**

**APLLICATIONS**

**The areas where this solution can be applied:**

* Query performance
* Aggregate view of data vs transactional view
* Complex SQL
* Normalized databases are typically tuned for simple queries

**CHAPTER 9**

**CONCLUSION**

**From this entire findings we know fundamental concepts and can work on IBM Cogno’s Analytics.**

* Gain a board understanding of plotting graphs.
* Able to create meaningful dashboards.
* Learn to build stunning dashboards with cogno’s analytics.
* To create data visualizations to understand.
* Total Number of Trips.
* Find the top bike used with respect to trip duration
* We will understand how a dashboard is different from a report, when to use both.
* We will understand the reporting interface.
* Calculating the number of bikes used by respective age groups.
* Top 10 Start Station Names with respect to Customer age group.

**CHAPTER 10**

**FUTURE SCOPE**

**Enhancements that can be made in the future:**

* Cognos is the one of the leading BI suites in the market for metadata modelling and reporting so learning this will be definitely helpful in our career growth in BI domain.
* IBM Cognos TM1 form 10 has been around for decent time and has officially experienced a few minor and real updates.
* IBM Cognos analytics leads to better decisions and improves company performance and profitability.
* We can scope the better job in future with easy experience.
* Total 709 companies are most often in computer software industry.
* Rightly so, a good majority of them focus on the strategic aspects of dashboard creation such as understanding our audience and purpose first and choosing the best charts to display our data visually for maximum readability and insights.
* The training industry quarterly further narrow scope for specific industries, audiences or purposes provide tips on the tailor dashboards for learning & development.

**CHAPTER 11**

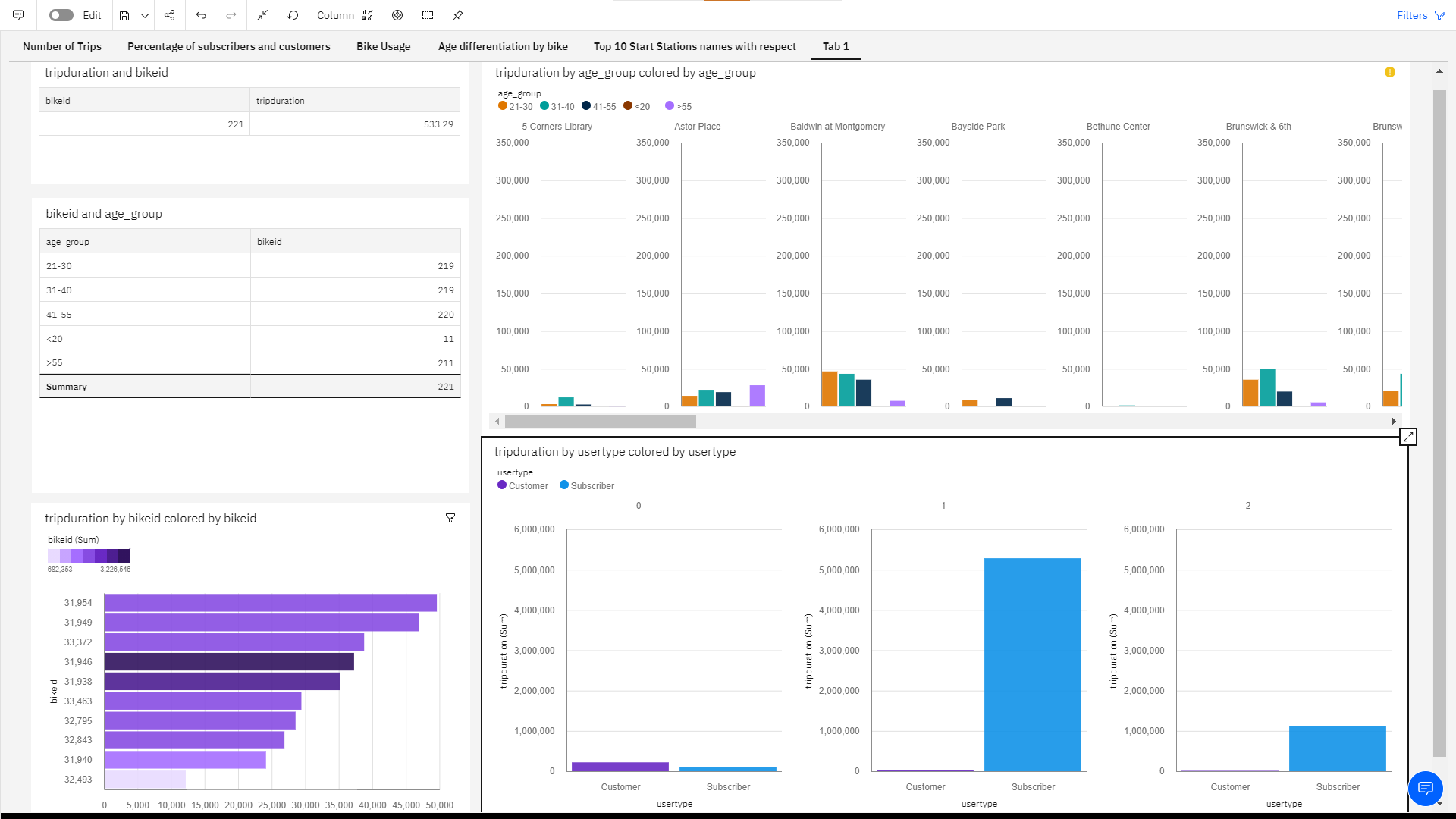
**BIBILOGRAPHY**

References of previous works or websites visited/books referred for analysis about the project, previous solution findings etc.

https://agrionline.nic.in/dash/dash.hmtl.

**CHAPTER12**

**APPENDIX**

****