

World Population Data Analytics Using IBM Cognos

Project Report

Aniket Sinha

19BME0359

VIT Vellore

Table of Contents

Sl No.	Title	Page
1	Introduction <ul style="list-style-type: none">■ Overview■ Purpose	2
2	Literature Survey <ul style="list-style-type: none">■ Existing Problems■ Proposed Solutions	3
3	Theoretical Analysis <ul style="list-style-type: none">■ Block Diagram	5
4	Flowchart	6
5	Result	7
6	Advantages and Disadvantages	11
7	Applications	13

Introduction

Overview:

This project is to analyse the World Population estimates given by the UN.

The 2019 Revision of World Population Prospects is the twenty-sixth round of official United Nations population estimates and projections that have been prepared by the Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat.

List of sources of empirical data used or considered and the methods applied in revising past estimates of population and components of demographic change (fertility, child, adult and overall mortality, international migration) are presented in tabular form for each demographic component and country or area for the period 1950-2020.

Purpose:

The purpose of this study is find and analyse United Nations population estimates and projections and present in a visual format for better understanding.

We will be using IBM Cosgnos analytics to visualize and analyse the dataset.

We will be creating different types of charts and tables for visualisation.

After the analysis we will be able to find various patterns and details in our data which will be helpful to come up with solutions.

Literature Survey

Existing Problems

1.The amount of data being collected

With today's data-driven organizations and the introduction of big data, risk managers and other employees are often overwhelmed with the amount of data that is collected. An organization may receive information on every incident and interaction that takes place on a daily basis, leaving analysts with thousands of interlocking data sets.

2. Collecting meaningful and real-time data

With so much data available, it's difficult to dig down and access the insights that are needed most. When employees are overwhelmed, they may not fully analyse data or only focus on the measures that are easiest to collect instead of those that truly add value. In addition, if an employee has to manually sift through data, it can be impossible to gain real-time insights on what is currently happening. Outdated data can have significant negative impacts on decision-making.

3. Visual representation of data

To be understood and impactful, data often needs to be visually presented in graphs or charts. While these tools are incredibly useful, it's difficult to build them manually. Taking the time to pull information from multiple areas and put it into a reporting tool is frustrating and time-consuming.

4. Data from multiple sources

The next issue is trying to analyse data across multiple, disjointed sources. Different pieces of data are often housed in different systems. Employees may not always realize this, leading to incomplete or inaccurate analysis. Manually combining data is time-consuming and can limit insights to what is easily viewed.

Proposed Solutions:

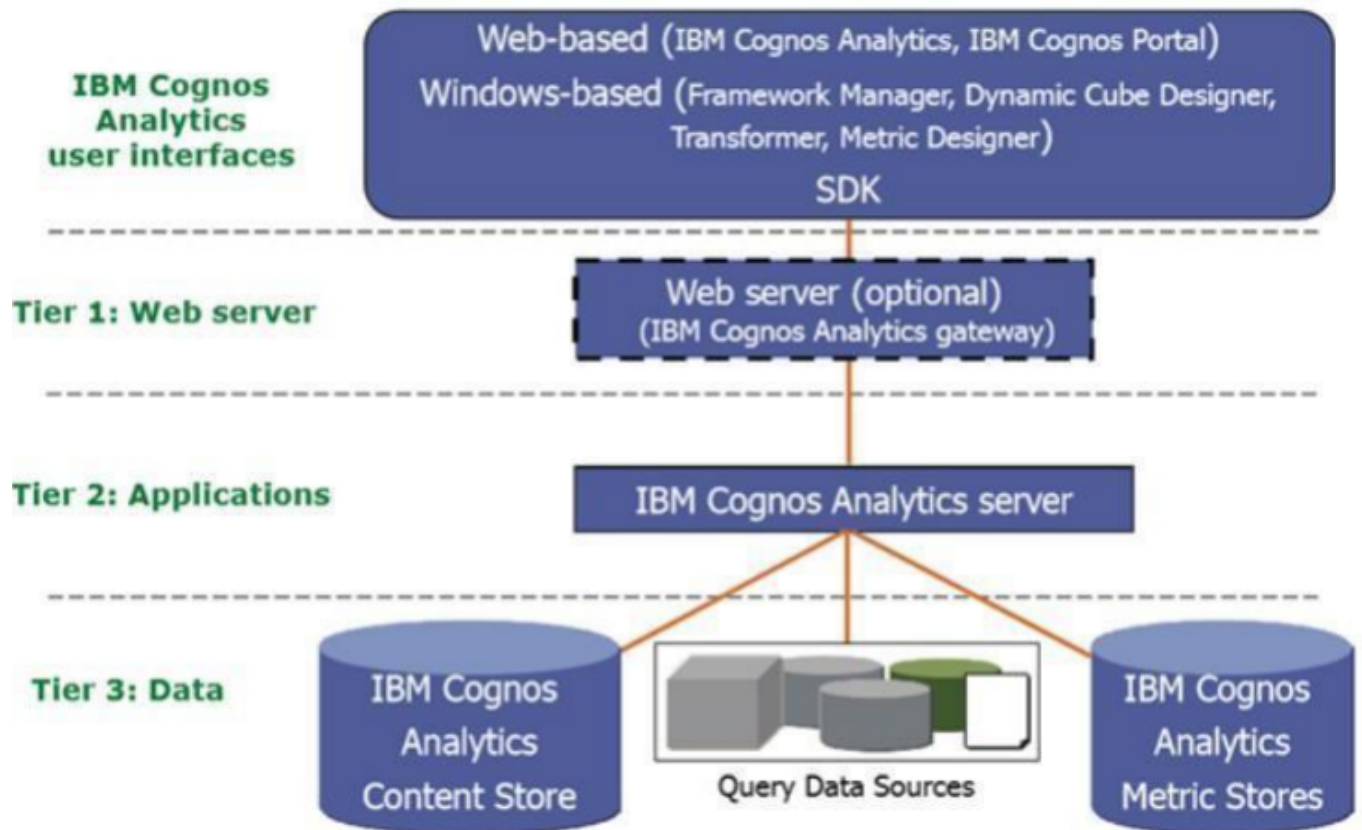
IBM Cognos Analytics

- There is a need for a data system that automatically collects and organizes information. Manually performing this process is far too time-consuming and unnecessary in today's environment. An automated system will allow employees to use the time spent processing data to act on it instead.
- A data system that collects, organizes and automatically alerts users of trends will help solve this issue. Employees can input their goals and easily create a report that provides the answers to their most important questions. With real-time reports and alerts, decision-makers can be confident they are basing any choices on complete and accurate information.
- Strong data systems enable report building at the click of a button. Employees and decision-makers will have access to the real-time information they need in an appealing and educational format.
- With a comprehensive and centralized system, employees will have access to all types of information in one location. Not only does this free up time spent accessing multiple sources, it allows cross-comparisons and ensures data is complete

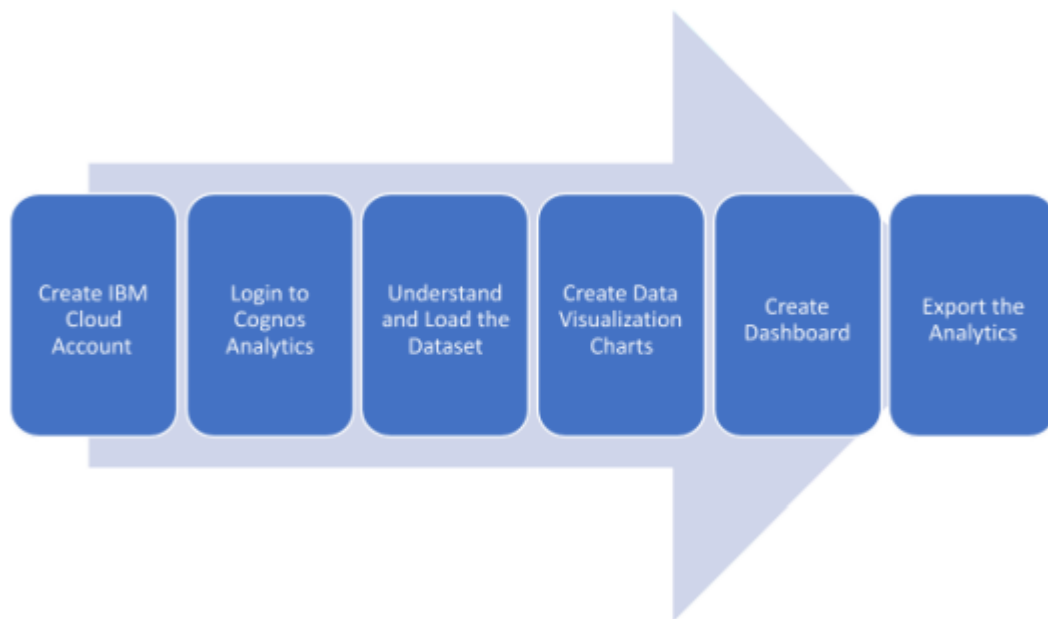
Theoretical Analysis

Block Diagram

IBM Cognos Analytics architecture (high level)



Flowchart:



Result

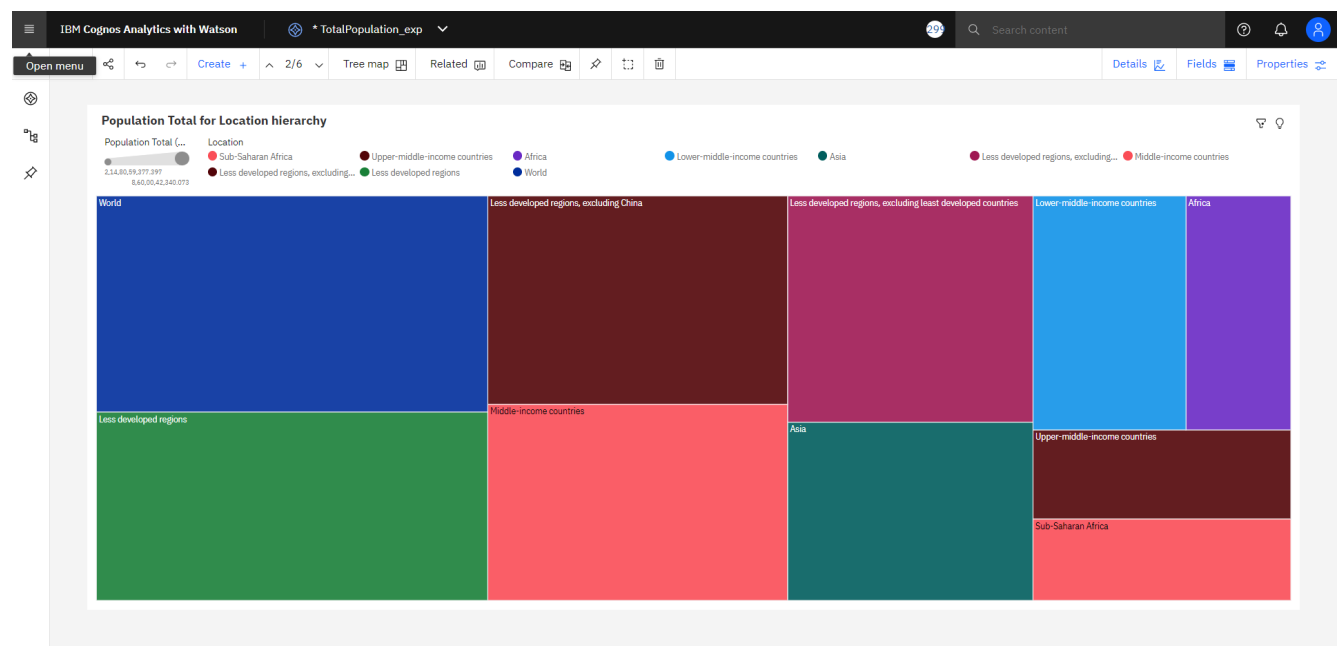
Top10 Pop Total By Location Using Tree Map

This give us a visualization of the Top 10 most populated regions in the world.

Details:

Population Total ranges from over 2.1 billion, when **Location** is Sub-Saharan Africa, to over 8.6 billion, when **Location** is World.

For **Population Total**, the most significant values of **Location** are World and Less developed regions, whose respective **Population Total** values add up to approximately 16 billion, or 32.8 % of the total.



Pop Total by Time Using Line Chart

This gives a relation between total population and time.

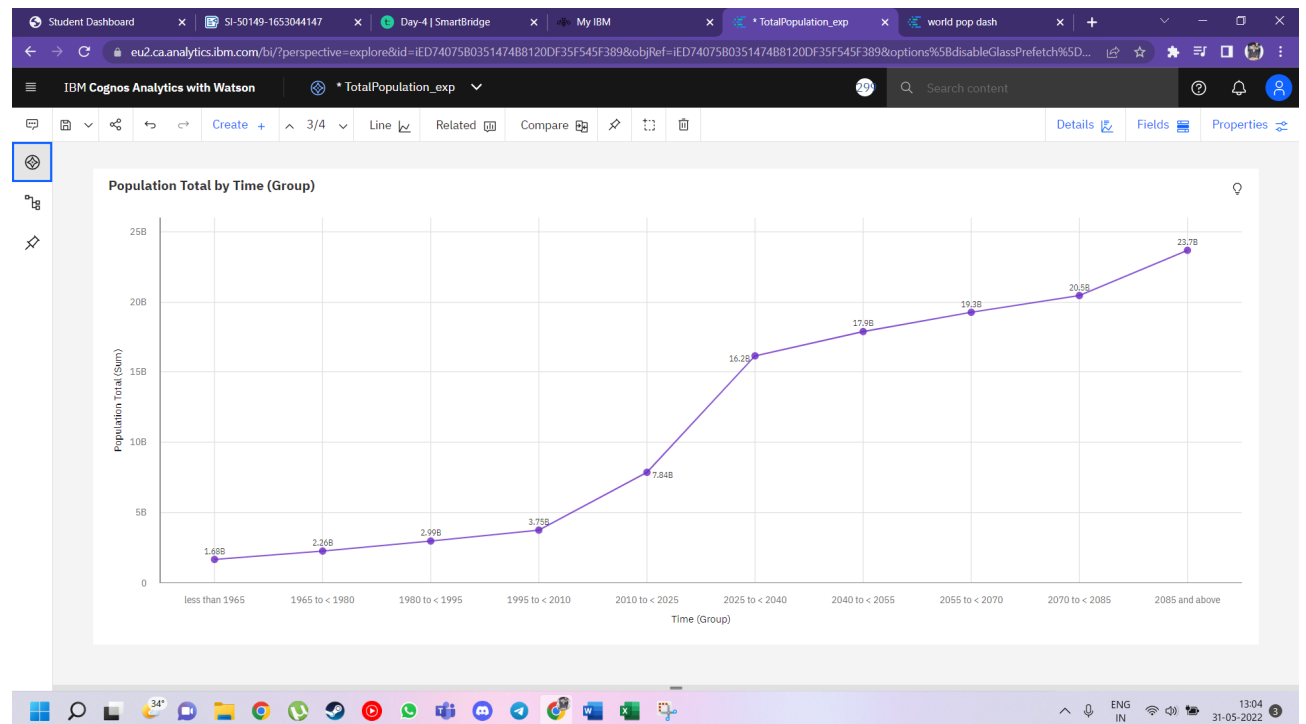
Details:

Over all values of **Time (Group)**, the sum of **Population Total** is approximately 116 billion.

For **Population Total**, the most significant values of **Time (Group)** are 2085 and above, 2070 to < 2085, 2055 to < 2070, 2040 to < 2055, and 2025 to < 2040, whose respective **Population Total** values add up to approximately 97 billion, or 84 % of the total.

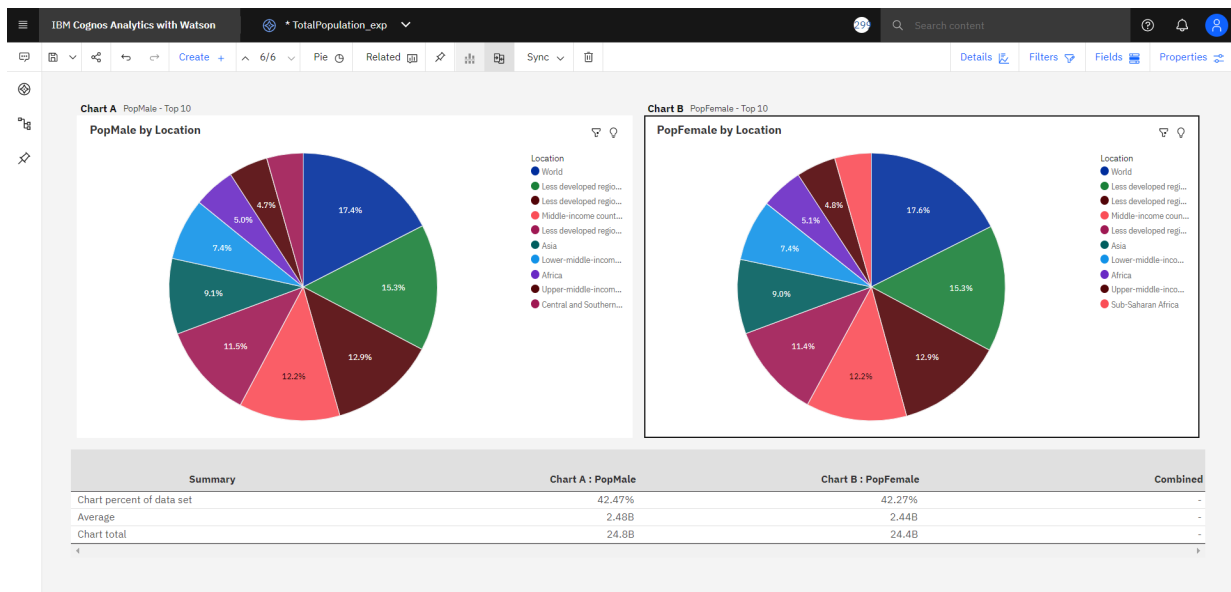
Population Total ranges from nearly 1.7 billion, when **Time (Group)** is less than 1965, to approximately 24 billion, when **Time (Group)** is 2085 and above.

Population Total is unusually high when **Time (Group)** is 2085 and above.

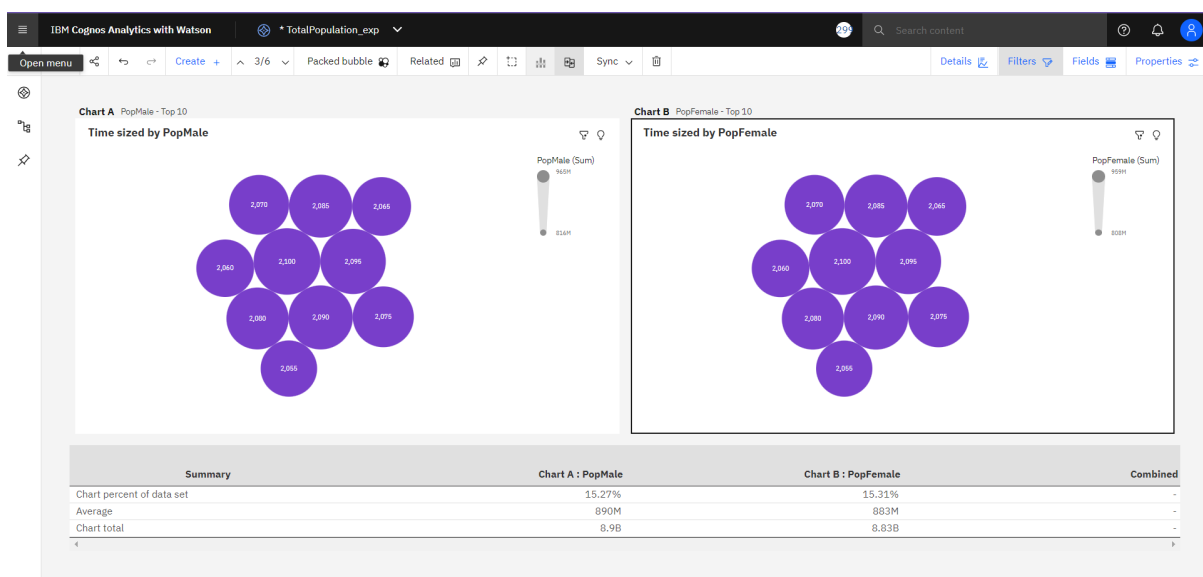


Pop Male by Location and Pop Female by Location Using Pie Charts

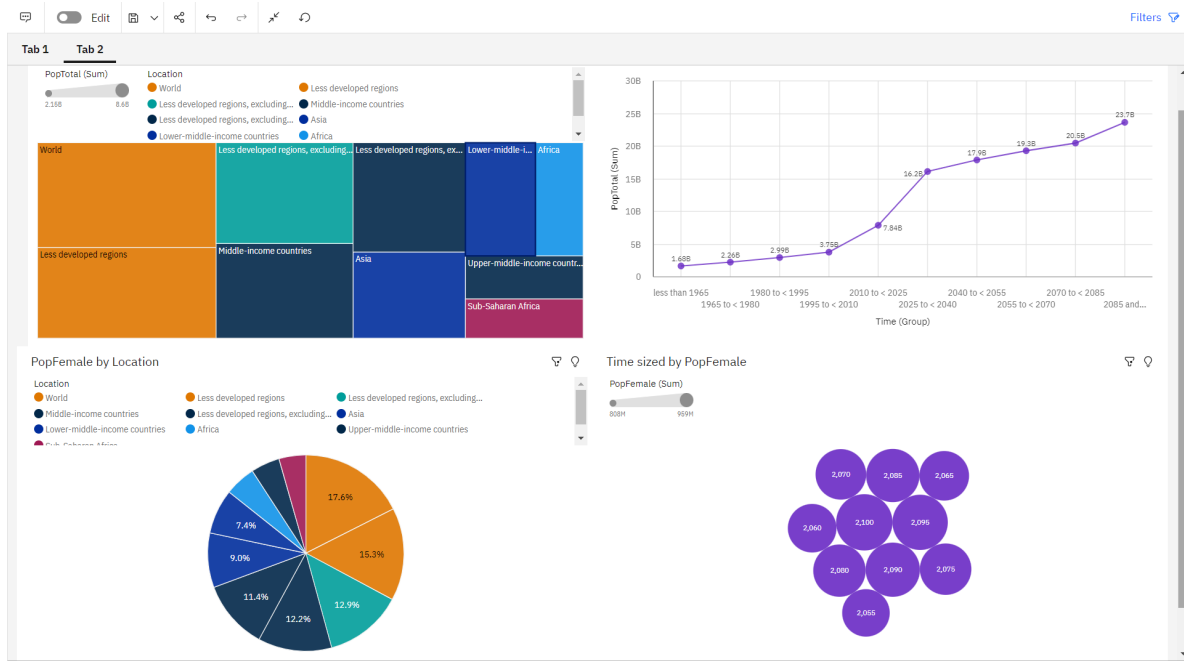
This visualization compares pie charts of Male Population by Locations and Female Population by Locations



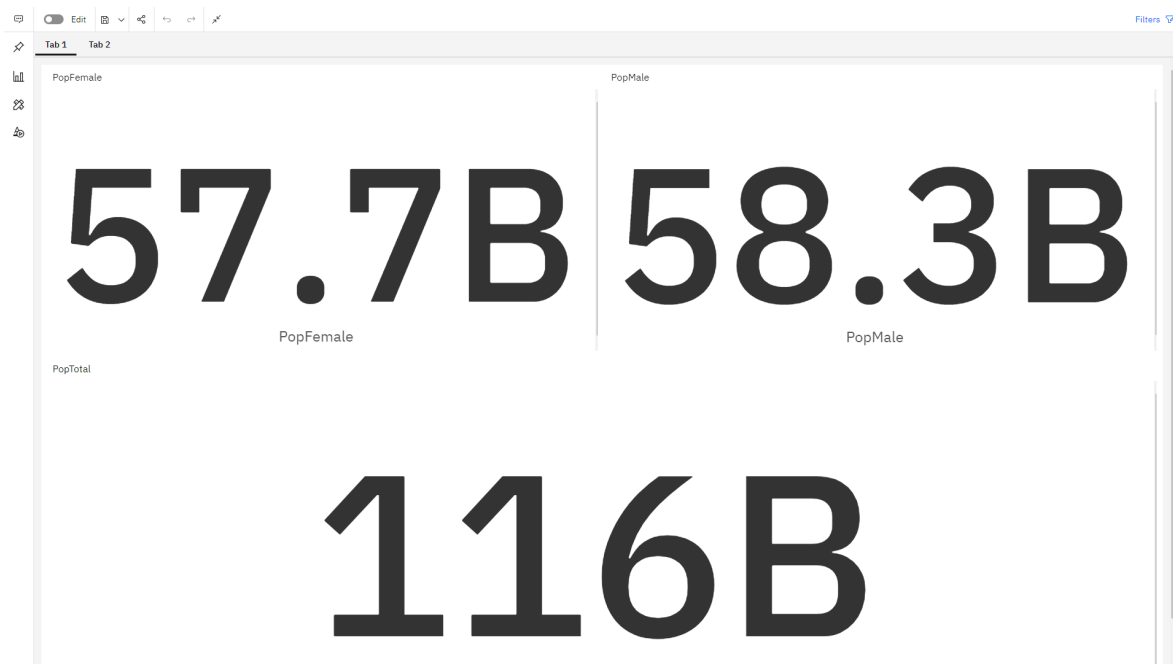
Pop Male by Time and Pop Female Using Packed Bubble Charts



Dashboard



Summary Card



Final Dashboard Link:

https://eu2.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.my_fol

Advantages

Advantages

- **Data analytics helps an organization make better decisions**

Analytics can help with transforming the data that is available into valuable information for executives so that better decisions can be made. This can be a source of competitive advantage if fewer poor decisions are made since poor decisions can have a negative impact on a number of areas including company growth and profitability.

- **Increase the efficiency of the work**

Analytics can help analyse large amounts of data quickly and display it in a formulated manner to help achieve specific organizational goals. It encourages a culture of efficiency and teamwork by allowing the managers to share the insights from the analytics results to the employees. The gaps and improvement areas within a company become evident and actions can be taken to increase the overall efficiency of the workplace thereby increasing productivity.

- **Personalization of products and services**

Analytics can help companies keep track of what kind of service, product, or content is preferred by the customer and then show the recommendations based on their preferences. For example, in social media, we usually see what we like to see, all of this is made possible due to the data collection and analytics that companies do. Data analytics can help provide targeted services to customers based on their individual requirements.

- **Improving quality of products and services**

Data analytics can help with enhancing the user experience by detecting and correcting errors or avoiding non-value-added tasks. For example, self-learning systems can use data to understand the way customers are interacting with the tools and make appropriate changes to improve user experience. In addition, data analytics can help with automated data cleansing and improving the quality of data and consecutively benefiting both customers and organizations.

Limitations

- **Lack of alignment within teams**

There is a lack of alignment between different teams or departments within an organization. Data analytics may be done by a select set of team members and the analysis done may be shared with a limited set of executives. However, the insights generated by these teams are either of not much value or are having limited impact on organizational metrics. This could be due to a “silos” way of working with each team only using their existing processes disconnected from other departments. The analytics team should be focussed on answering the right questions for the business and the results generated by data analytics teams needs to be properly communicated to the right employees to drive the right set of actions and behaviours so that it can have an positive impact on the organization.

- **Lack of commitment and patience**

Analytics solutions are not difficult to implement, however, they are costly, and the ROI is not immediate. Especially, if existing data is not available, it may take time to put processes and procedures in place to start collecting the data. By nature, the analytics models improve accuracy over time and require dedication to implement the solution. Since the business users do not see results immediately, they sometimes lose interest which results in loss of trust and the models fail. When an organization decides to implement data analytics methods, there needs to be a feedback loop and mechanism in place to understand what is working and what is not, and corrective actions are required to fix things that are broken. Without this closed loop system, senior management may decide that analytics is not working or much valuable and may abandon the entire exercise.

- **Privacy concerns**

Sometimes, data collection might breach the privacy of the customers as their information such as purchases, online transactions, and subscriptions are available to companies whose services they are using. Some companies might exchange those datasets with other companies for mutual benefit. Certain data collected can also be used against a person, country, or community. Organizations need to be cautious of what sort of data they are collecting from customers and ensure the security and confidentiality of the data. Only the data required for the analysis needs to be captured and if there is sensitive data, it needs to be anonymized so that sensitive data is protected.

Applications

1. Transportation

Data analytics can be applied to help in improving Transportation Systems and intelligence around them. The predictive method of the analysis helps find transport problems like Traffic or network congestions. It helps synchronize the vast amount of data and uses them to build and design plans and strategies to plan alternative routes, reduce congestions and traffics, which in turn reduces the number of accidents and mishappenings.

2. Logistics and Delivery

There are different logistic companies like DHL, FedEx, etc that uses data analytics to manage their overall operations. Using the applications of data analytics, they can figure out the best shipping routes, approximate delivery times, and also can track the real-time status of goods that are dispatched using GPS trackers. Data Analytics has made online shopping easier and more demandable.

3. Web Search or Internet Web Results

The web search engines like Yahoo, Bing, Duckduckgo, Google uses a set of data to give you when you search a data. Whenever you hit on the search button, the search engines use algorithms of data analytics to deliver the best-searched results within a limited time frame. The set of data that appears whenever we search for any information is obtained through data analytics.

4. Manufacturing

Data analytics helps the manufacturing industries maintain their overall working through certain tools like prediction analysis, regression analysis, budgeting, etc. The unit can figure out the number of products needed to be manufactured according to the data collected and analyzed from the demand samples and likewise in many

other operations increasing the operating capacity as well as the profitability.

5. Security

Data analyst provides utmost security to the organization, Security Analytics is a way to deal with online protection zeroed in on the examination of information to deliver proactive safety efforts. No business can foresee the future, particularly where security dangers are concerned, yet by sending security investigation apparatuses that can dissect security occasions it is conceivable to identify danger before it gets an opportunity to affect your framework and main concern.

6. Education

Data analytics applications in education are the most needed data analyst in the current scenario. It is mostly used in adaptive learning, new innovations, adaptive content, etc. Is the estimation, assortment, investigation, and detailing of information about students and their specific circumstance, for reasons for comprehension and streamlining learning and conditions in which it happens.

Conclusion

This way, with the help of diagrams, graphs, and maps we can understand given data. This understanding of data allows us to ask the right questions to reach our desired goals by optimizing methods. With this project, we learned how to upload and prepare data. We also statistical concepts which helped in calculations and plotting of graphs and maps to make a dashboard.