J.J.COLLEGE OF ENGINEERING AND TECHNOLOGY
Approved by AICTE | Affiliated to Anna University | ISO 9001 : 2008 Certified
(Sowdambikaa Group of Institutions)

Data Titans: Unearthing trends from LinkedIn influencers

Project Report

Submitted By

- Saveethaa VP
- Nivetha Rajam D
- Mothina P
- Selvaraani T
- Ratchana S

Repository Link:

https://github.com/mothinap/Data-Titans-Unearthing-Trends-from-LinkedIn-Influencers

Table of Contents

1.Introduction

1.1 Overview

The term "LinkedIn Influencers Data Analytics" describes the practise of using data analytics methods and instruments to obtain information and decide wisely regarding the actions and effects of LinkedIn Influencers on the platform. High-profile people who frequently contribute insightful articles, knowledge, and experience with their followers on LinkedIn are known as LinkedIn Influencers. When it comes to organisations, marketers, and individuals looking to enhance their LinkedIn presence, analysing their data might yield insightful information.

The measurement of the content's reach and engagement is a crucial component of LinkedIn Influencer Data Analytics. Monitoring data like the quantity of views, likes, comments, and shares on their postings is part of this. You can determine which kinds of material appeal to their audience the most by examining these indicators, and you can then adjust your content strategy appropriately. Sentiment analysis can also be used to better understand the audience's perceptions of the material, which is important for brand development and reputation management.

Additionally, trends and patterns pertaining to the subjects, sectors, and keywords that these influencers use the most frequently can be found with the help of LinkedIn Influencer Data Analytics. Marketing plans and content production can be guided by this data. Furthermore, demographic information regarding LinkedIn Influencers' followers might be useful for focusing on particular audience niches. In general, LinkedIn Influencers Data Analytics is a big aid to professionals and organisations in making data-driven choices to improve their visibility and influence on the platform and to create deep connections with their audience.

1.2 Purpose

The purpose of LinkedIn Influencers Data Analytics is to gain valuable insights and make data-driven decisions related to the activities and impact of LinkedIn Influencers. This analysis serves several important purposes:

Content Strategy Optimisation: People and companies can adjust their content strategy to produce posts that connect with their target audience by examining the performance of LinkedIn Influencers' content. The efficacy and quality of content can be increased by understanding the kinds of material that get the greatest interaction.

Improved audience targeting: Is made possible by knowing the geographic and demographic makeup of LinkedIn Influencer followers. The likelihood of success is increased by using this data to target the appropriate audience with pertinent content and marketing initiatives.

Reputation management: It is critical for LinkedIn Influencers and the companies they represent to keep an eye on the opinions and sentiment expressed in relation to their material. Analytics can be used to distinguish between favourable and unfavourable sentiments, enabling prompt replies and image-maintaining modifications.

Analysis of market and Trends: LinkedIn Influencers frequently talk about new subjects and trends in the market. Individuals and companies may keep up with the most recent developments in the market and obtain insightful knowledge that they can utilise to improve their own plans by evaluating their material.

Lastly, we will use the visualisations above to create a dashboard. By choosing any area or time from any of the charts, we can interact with the dashboard and see how the population changes.

2.Literature Survey

2.1 Existing Problem

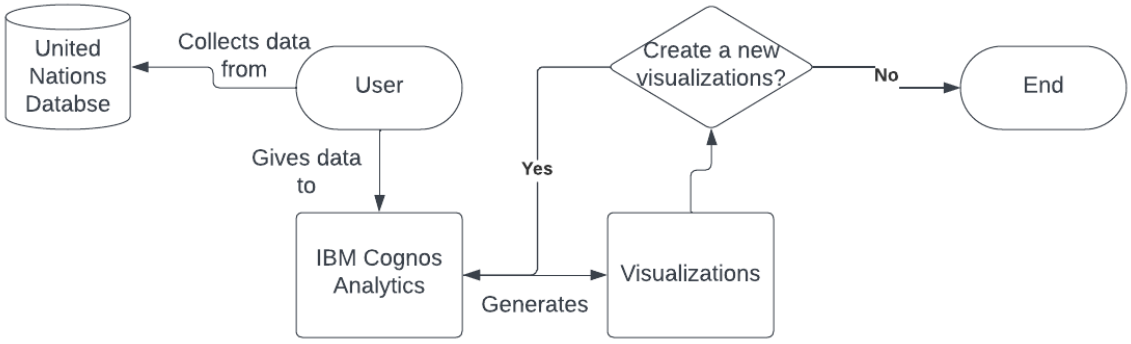
- **Data collection:** Selected influencers' LinkedIn profiles are used by the algorithm to gather data.
- **Metrics tracking:** It keeps track of important metrics on influencer postings, like views, likes, comments, and shares.
- **Sentiment analysis:** The programme evaluates the attitude of the audience towards content from influencers.
- **Demographic data:** It collects data, such as demographics, on the followers of the influencers.
- **Content analysis:** The system assesses the performance of influencers as well as the kinds of content they post.
- **Trend identification:** It recognises popular subjects covered by influencers and trends in the sector.
- **Reporting:** Reports on audience engagement and influencer performance are generated by the current system.
- **Dashboard:** To see statistics and insights on particular influencers, users can go to a dashboard.

2.2 Proposed System

- To clean and retrieve reliable data, we can execute data preparation using IBM Cognos Analytics.
- Since graphs and charts are easier to understand than tabular data, we can use IBM Cognos Analytics to build a variety of charts to help us understand the data.
- It is possible to generate multiple dashboards and reports that summarise the company's interests on a single page. Decision-making can be aided by the automatic insights offered by IBM Cognos Analytics.

3.Theoritical Analysis

3.1 Block Diagram



3.2 Hardware/Software Designing

Hardware Requirements:

- Processor: Intel(R) Core(TM) i7-8565U CPU @ 1.80GHz 1.99 GHz
- Installed RAM: 8.00 GB (7.89 GB usable)
- System type: 64-bit operating system, x64-based processor Proper Internet

Software Requirements:

Cognos Analytics from IBM: IBM offers a web-based comprehensive business intelligence suite called Cognos Analytics. It offers a suite of tools for score carding, analytics, reporting, and event and metric tracking. The software is made up of several parts that are intended to satisfy the various information needs of an organisation.

Google Chrome: Is a cross-platform web browser that was created by the company. Using the most recent version of the Chrome browser is crucial if you want to utilise all of IBM Cognos Analytics' features.

4.Experimental Investigation

Dataset Name: Linkedin Influencers Data.csv

Dataset__Source:https://www.kaggle.com/datasets/shreyasajal/linkedin-influencers-data.csv

Number of Columns: 10

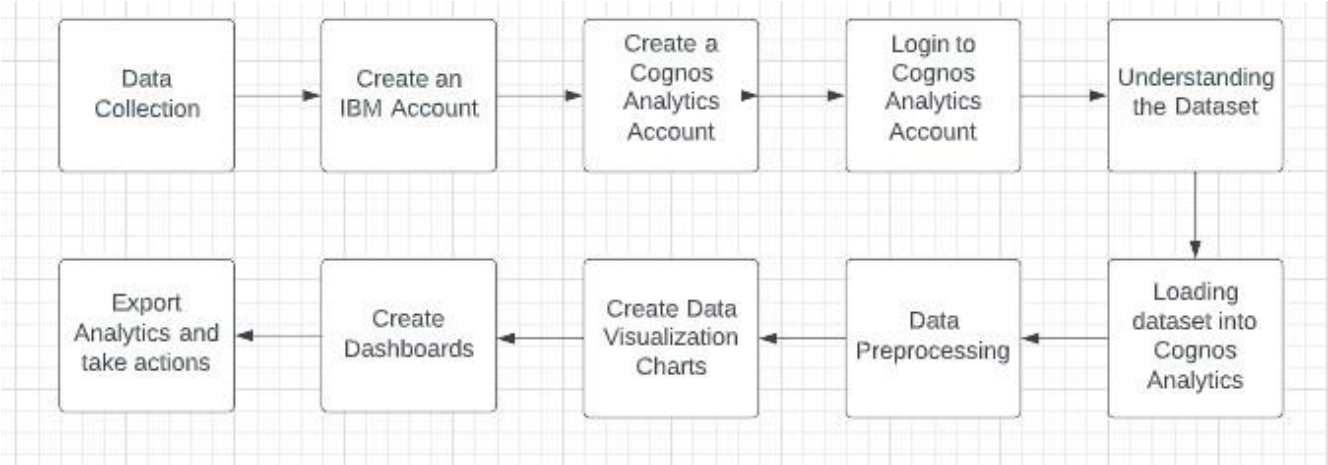
Number of Rows: 280932

Overview of the Dataset:

The following characteristics are included in the data set:

- **LocID (numeric):** A location's numerical code that adheres to ISO 3166-1 for regions and nations.
- **Location (string):** The area, country, subregion, or region's name
- **Varid (numeric):** Variant identification number
- **Version (string):** Name of projection version (the most common is Medium)
- **Time (string):** Label designating the data's duration (e.g., 1950–1955) or a single year (e.g., 1950).
- **Midperiod (numeric):** A number that indicates the middle of the data set; the month is indicated by the decimal (e.g. 1950.5)
- **Popmale (numerical):** The total number of men in thousands
- **Pofemale (numerical):** The number of females in thousands
- **Poptotal:** The total population expressed in thousands
- **Popdensity (numeric):** The number of people per thousand

5.Flow Chart



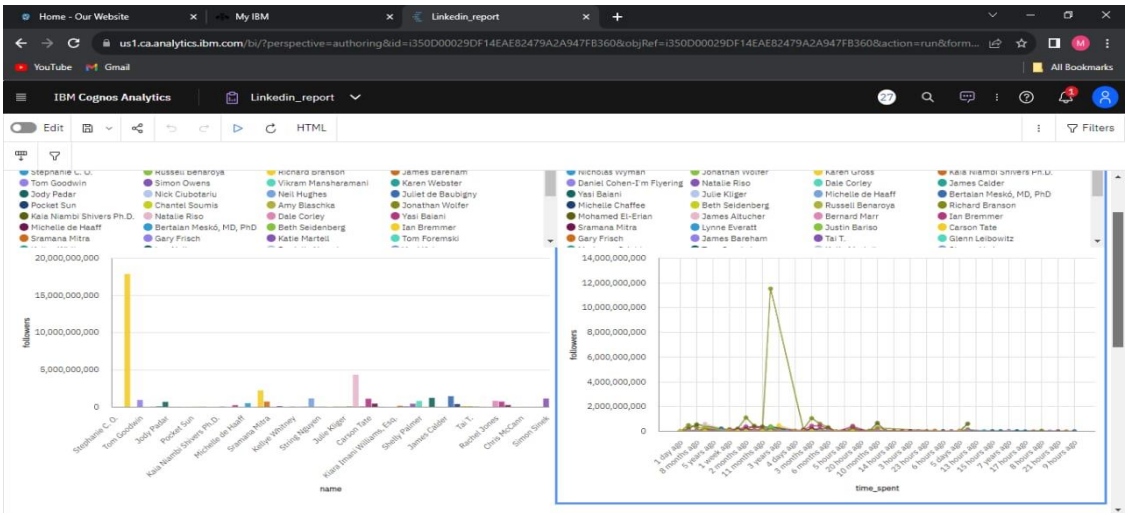
6.Results

Preparing and Preprocessing Data

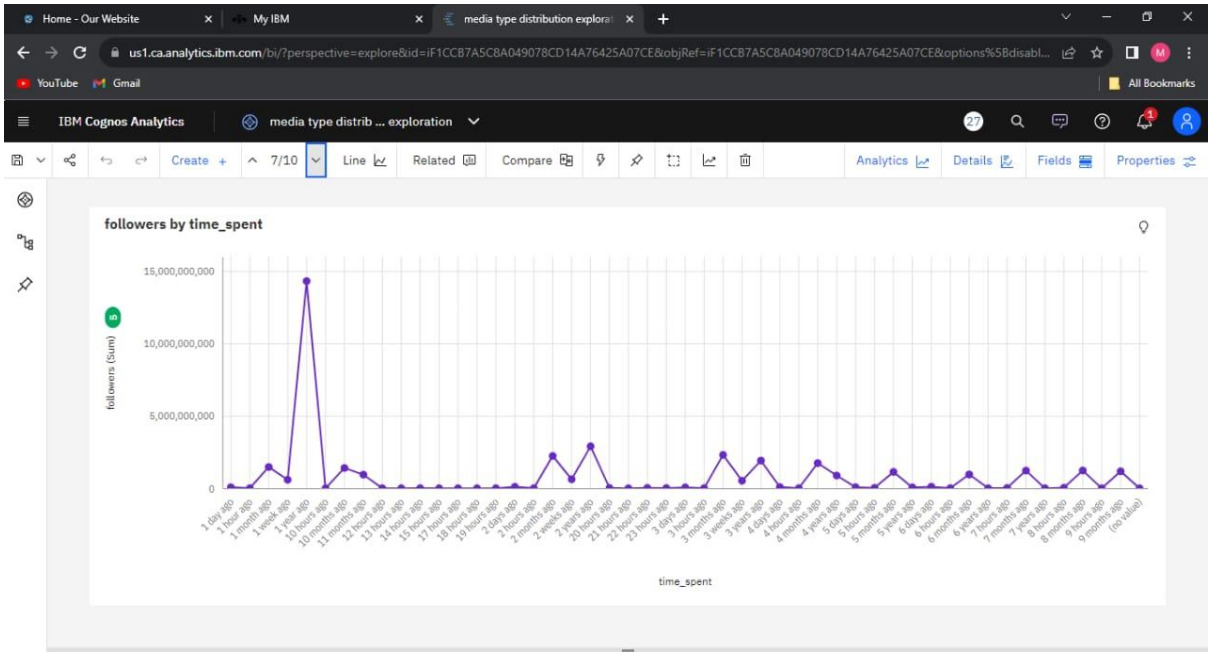
- ❖ Take out the current column. total pop.
- ❖ Establish A Fresh Calculated Field Pop Male + Pop Female equals Pop Total; format the information.
- ❖ Data types for Pop Male and Pop Female format fields.
- ❖ PopMale, PopFemale, and PopTotal are double datatypes in this instance. However, population does not exist in binary or fractional forms. creating an integer out of their data kinds.

Data visualization Charts

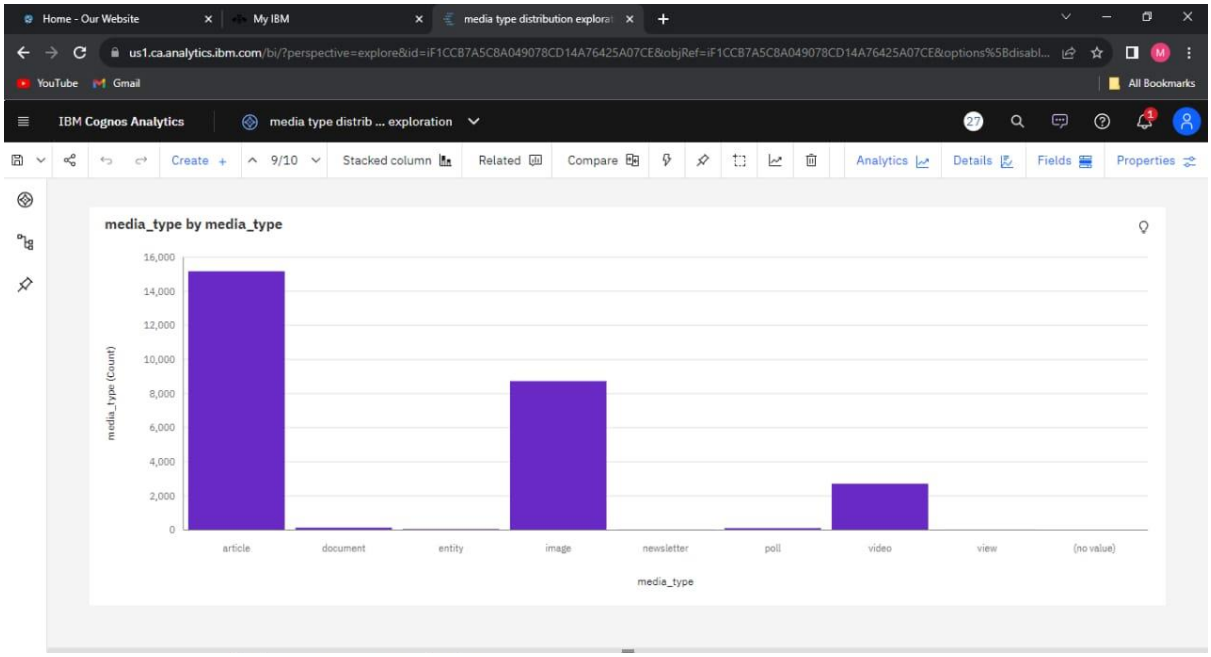
1. Top10 Pop Total by Location Using Tree Map



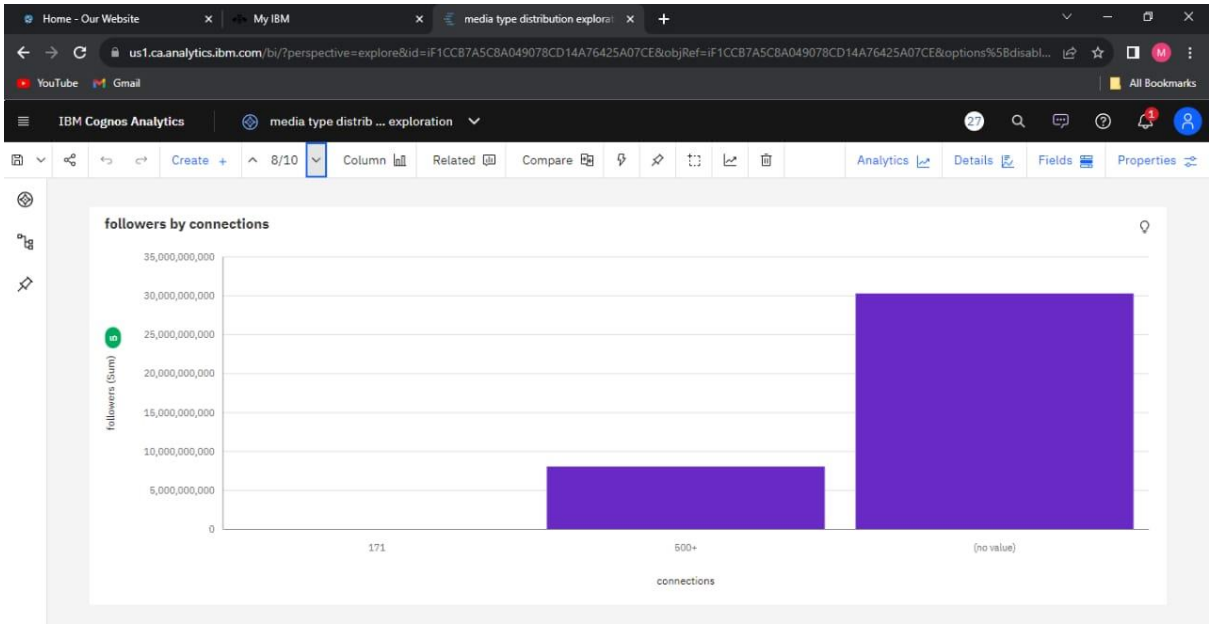
2. Pop Total By Time Using Line Chart



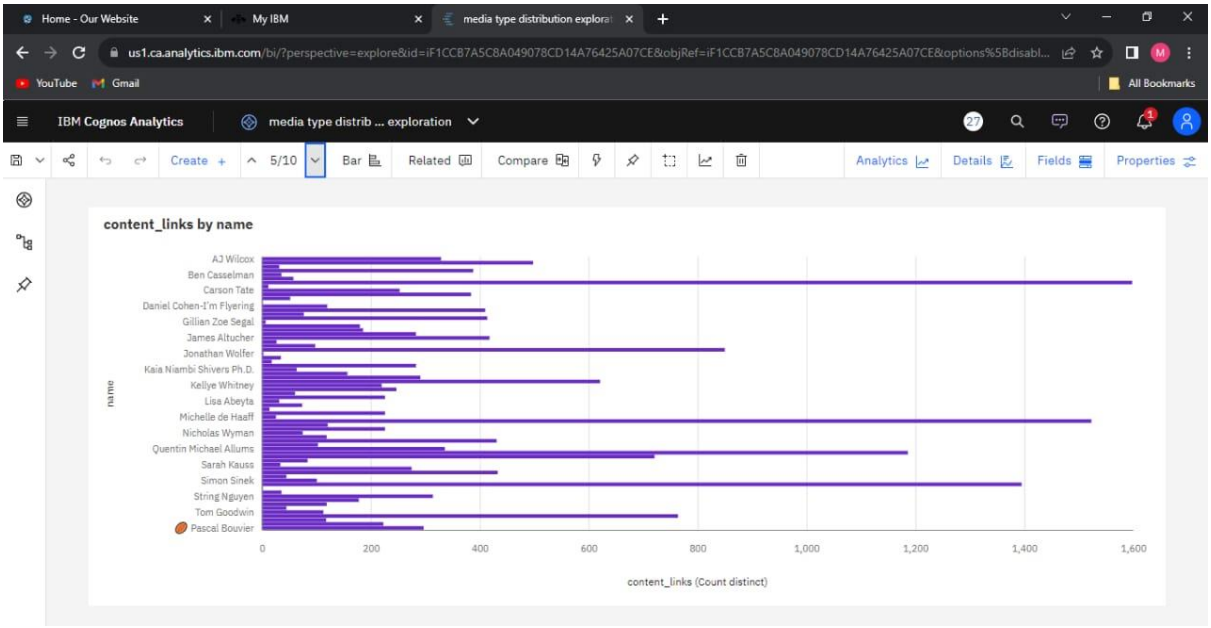
3. Bar graph for media type



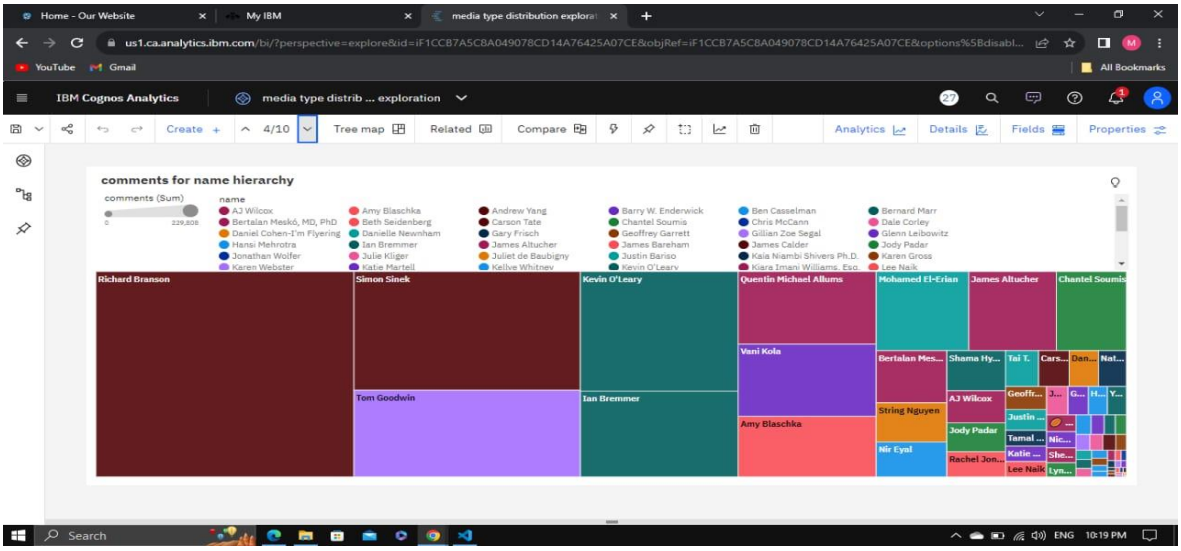
4. Bar chart for followers by connections

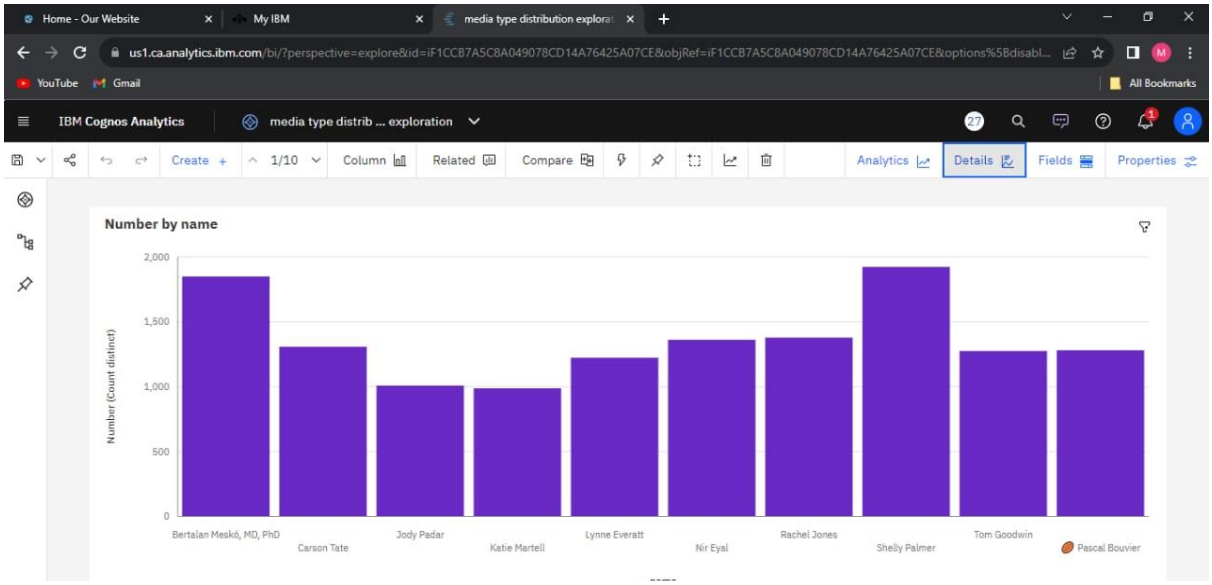


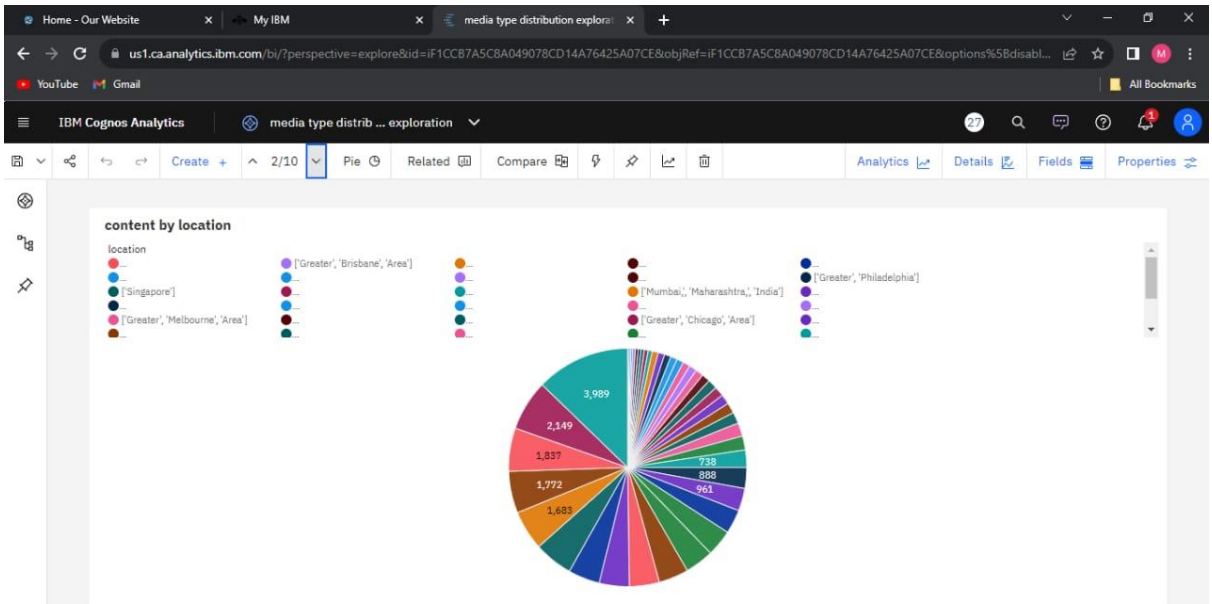
5. Line chart for content_links by name



6. Final overview using summary







7.Advantages

- LinkedIn Influencers as a Refinement of Content Strategy Businesses and individuals can improve their content strategies with the use of data analytics. Users can generate more engaging posts and successfully engage their audience by knowing what kinds of content get the most engagement.

- **Audience Targeting:** By offering details on the characteristics and inclinations of influencers' followers, it facilitates accurate audience targeting. By focusing marketing messages and efforts on particular target segments, this information helps to increase the relevancy of content.
- **Partnership prospects:** You may find prospects for cooperation by locating influencers whose material complements your goals. Influencer partnerships help you become more well-known and credible in your sector or specialty.

Disadvantages

- **Data Privacy Issues:** Information about LinkedIn influencers and their followers may be gathered and analysed, which may give rise to privacy issues. Working with user data requires adhering to ethical principles and data protection laws like GDPR, even when the data is publicly accessible on LinkedIn.
- **Data Accuracy and Reliability:** The information gathered from LinkedIn may not always be accurate or reliable. User interaction numbers may not always accurately reflect the emotions or intentions of the audience, and profiles may be out of date. This may result in biassed conclusions and poor data-driven decisions.
- **Platform improvements:** Data collecting techniques and analytics tools may be impacted by LinkedIn's regular algorithm and feature improvements. Adapting the analytics strategy and keeping up with these developments can be difficult and time-consuming.
- **Costs and Resources:** It can take a lot of resources to build and manage a LinkedIn Influencers Data Analytics system.

8.Applications

- The United Nations may use this to track the rate of population growth in a certain area and take appropriate action.
- This can also be used by International Welfare Organisations for the same goals as previously described.
- They can also collect the names of the most populous regions and population dense sites to assess if the amenities at those locations are up to the mark.

9.Conclusion

In this project, we have successfully collected data from United Nations website, understood the dataset and performed data preparation to obtain accurate data using IBM Cognos Analytics. Then we have created data visualizations and dashboards. These dashboards provided various insights to the user which can be applied in Decision Making. These insights on world population dataset can be utilized by various organizations for the welfare of the people.

10.Future Scope

This project can be fine grained to local level from the global level by collecting data from local regions. This data can be used by the local authorities to gain insights and take necessary actions. This will create even more impact as minor issues will also come into light. When global level is considered, these minor issues might be over shadowed by the other major issues.

Voice Assistance can be added to IBM Cognos Analytics platform to ease the process of creating visualizations.

11.Appendix

https://youtu.be/eSMW-_jqcTU?si=AnmmNstdHoumBNzo