

INTRODUCTION

a. Overview of the organization :-

About APSCHE :-

The state govt. has accordingly decided to fill gap by constituting a state council at higher education as recommended in the NEP of the government of India as recommended by the committee constituted by the UGC.

Thus, AP State Council of higher education (APSCHE) came into existence to advise the government in matters relating to higher education in the state and to oversee its development with perspective planning.

DATA ANALYTICS :-

Data analytics converts raw data into actionable insights. It includes a range of tools, techniques and processes used to find trends and solve problems by using data. Data analytics can shape business processes, improve decision-making and foster business growth.

Types of data analytics :-

1. Business Intelligence
2. Data visualization
3. Statistics
4. Predictive analysis.

- (2)
- 5. Data mining
 - 6. Medical diagnosis
 - 7. Descriptive statistic
 - 8. Prescriptive analytics
 - 9. Exploratory data analysis
 - 10. Regression analysis
 - 11. Statistical Inference
 - 12. Time series
 - 13. Forecasting
 - 14. Predictive modeling
 - 15. Cohort analysis
 - 16. Factor analysis
 - 17. Language Interpretation.

Importance of data analytics :-

Data analytics helps a business optimize its performance, perform more efficiently, maximize profit, or make more strategically - guided decisions. The techniques and processes of data analytics have been automated into mechanical processes and algorithms that work over raw data for human consumption.

→ Data analytics relies on a variety of software tools including spreadsheet, data visualization reporting tools, data mining, programs and open resources languages for the greatest data manipulation.

2.

LITERATURE SURVEY :-

As mentioned at the beginning of the post, data analysis methods can be divided into two big categories:

quantitative and qualitative. Each of these categories holds a powerful analytics value that changes depending on the scenario and type of data you are working with. Below, we will discuss IT methods that are divided into quantitative and qualitative approaches.

1. Cluster analysis:-

the action of grouping a set of data elements in a way that said elements are more similar to each other than to those in other group hence the term 'cluster'. Since there is no target variable when clustering the method is often used to find hidden patterns in the data. The approach is also used to provide additional context to a trend or data set.

2. Cohort analysis:-

This type of data analysis approach uses historical data to examine and compare a determined segment of user's behaviour, which can then be grouped with others with similar characteristics.

By using this methodology. It's possible to gain a wealth of insight into consumer needs.

Tablet and
Desktop traffic
68,733 users

Aug 13, 2017 - Aug 19, 2017

10,458. users

Aug 20, 2017 - Aug 26, 2017

10,444 users

Sep 3, 2017 - Sep 9, 2017

9,735 users.

	week 0	week 1	week 2	week 3
99.99%	3.41%	1.99%	1.55%	
3.35%	2.21%	1.67%	1.47%	1.08%
3.54%	1.96%	1.77%	1.55%	
3.78%	2.44%	2.41%		
4.04%	2.96%			
4.92%				

Mobile traffic

32,335 users

Aug 13, 2017 - Aug 19

5,302 users

Sep 3, 2017 - Sep 9, 2017

4,541 users.

99.97% 2.66% 0.90% 0.69%

2.81%	0.98%	0.64%	0.85%	0.81%
2.49%	1.01%	1.12%	0.75%	
2.33%	1.13%	0.84%		
3.28%	0.99%			
4.46%				

3. Regression analysis :-

Regression uses historical data to understand how a dependent variables value is affected when one or more independent variables change or stay the same. By understanding each variables relationship.

and how it developed in the past, you can anticipate possible outcomes.

④ Neural networks :-

The neural network forms the basis for the intelligent algorithms of machine learning. It is a form of analytics that attempts, with minimal intervention, to understand how the human brain would generate insights and predict values.

⑤ Factor analysis :-

The factor analysis also called "dimension reduction" is a type of data analysis used to describe variability among observed correlated variables in terms of a potentially lower number of unobserved variables called "factors".

⑥ Data mining :-

A method of data analysis that is the umbrella term for engineering metrics and insights for additional value direction, and context. By using exploratory statistical evaluation data mining aims to identify dependencies & relations.

⑦ Time series analysis :-

As its name suggests, time series

(6)

is used to analyze a set of data points collected over a specified period of time. The time series analysis is not uniquely used for purpose of collecting data.

⑧ Decision trees :-

The decision tree analysis aims to act as a support to make smart and strategies decisions by visually displaying potential outcomes, consequences and costs in a tree - like model, researchers and company users can easily evaluate all factors involved.

⑨ Conjoint analysis :-

Last but not least, we have conjoint analysis this approach is usually used in surveys to understand how individuals value different attributes of a product or service and it is one of the most effective methods to extract consumer preferences.

⑩ Correspondence Analysis :-

Also known as reciprocal averaging, correspondence analysis is a method used to analyze the relationship b/w categorical variables presented within a contingency table.

B. Qualitative methods :-

Qualitative data analysis methods are defined as the observation of non-numerical data that is gathered and procedure using method of observation such as interviews, Focus groups.

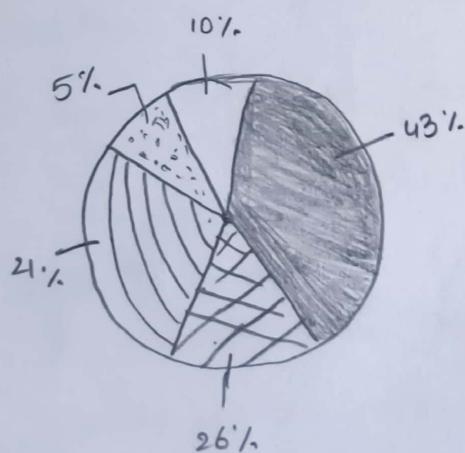
(12) Text analysis :-

Text analysis also known in the industry as text mining, works by taking large sets of textual data and arranging them in a way that makes it easier to manage.

(13) Content Analysis :-

This is a straightforward and very popular method that examines the presence and frequency of certain words, concepts and subjects in different content formats.

Distribution of transportation related cost :-



- Transportation costs
- Waterhousing costs
- Inventory carrying costs
- Administrative costs
- Order processing costs.

⑯ Thematic Analysis :-

Very similar to context analysis thematic analysis also helps in identifying and interpreting patterns in qualitative data with the main difference being that the first one can also be applied.

⑰ Narrative Analysis :-

A bit more complex in nature than the two previous ones, narrative analysis is used to explore the meaning behind the stories that people tell and most importantly how they tell them.

⑱ Discourse Analysis :-

Discourse analysis is used to understand the meaning behind any type of written, verbal or symbolic discourse based on its political, social or cultural context. It mixes the analysis of languages and situations together.

⑲ Grounded theory Analysis :-

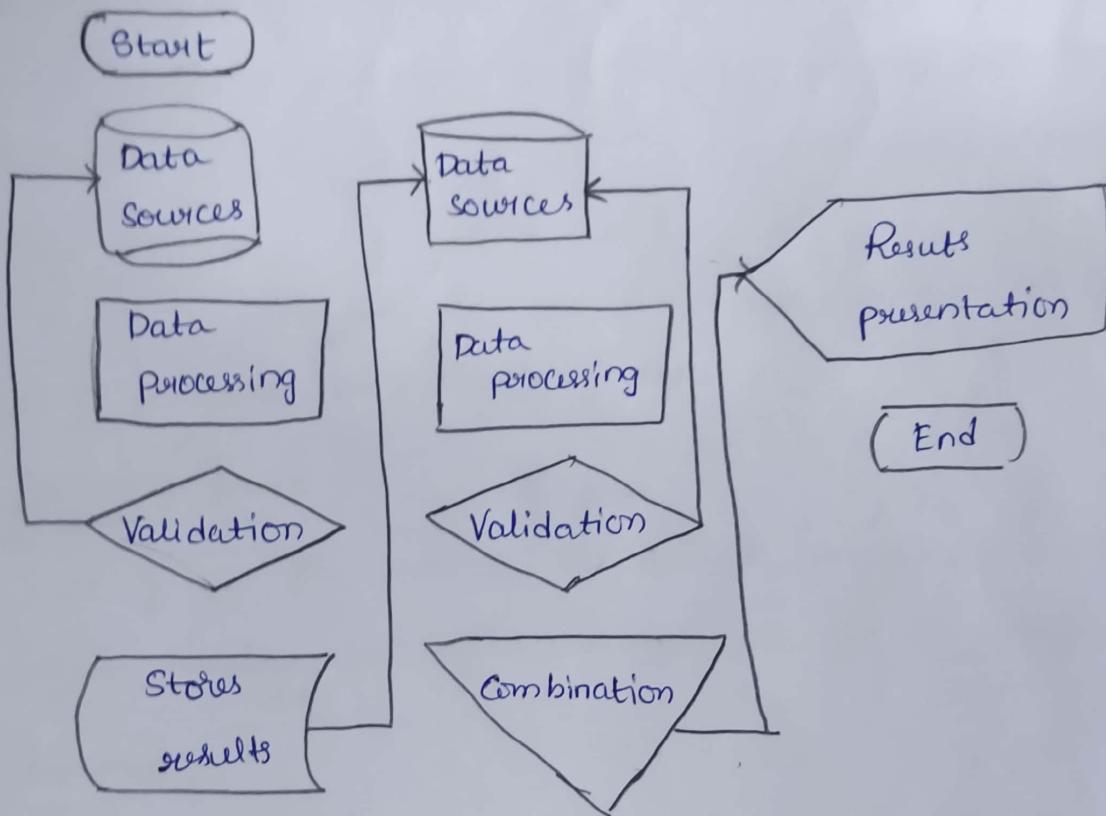
Traditionally, researchers decide on a method and hypothesis and start to collect the data to prove that hypothesis. The grounded theory is the only method that doesn't require an initial research question.

3

9

Theoretical Analysis :-

a) Block diagram :-



b) Hardware and software requirements :-

To be used efficiently - all computer software needs certain hardware components or other software resources to be present on a computer. These requirements is known as "System requirements" and are often used as a guideline as opposed to an absolute rule.

Most software defines two sets of higher processing power and resources.

→ Hardware requirements :-

The most common set of requirements defined by any operating system or software applications is the physical computer resources, also known as hardware.

* Architecture :-

All computers operating systems are designed for a particular computer architecture.

* Processing power :-

The power of the central processing unit (CPU) is a fundamental system requirement for any software.

* memory :-

All software, when run, resides in the random access memory (RAM) of a computer.

* secondary storage :-

Data storage device requirements vary, depending on the size of software installation, temporary files created and maintained.

* Peripherals :-

Some software applications need to make extensive and / or special use of some peripherals, demanding the higher performance or functionality of such peripherals.

→ Software requirements :-

Software requirements deal with defining software resource requirements and prerequisites that need to be installed on a computer to provide optimal functioning of an application.

* platform :-

A computing platform describes some sort of framework, either in hardware or software, which allows software to run.

* APIs and drivers :-

Software making extensive use of special hardware devices, like high end display adapters, needs special API or newer device drivers.

* web browser :-

Most web applications and software depend heavily on web technologies to make use of the default browser installed on the system. Microsoft Edge is a frequent choice of software running on Microsoft Windows.

4.

Result :-

Data analysis and findings :-

Data analysis is the most crucial part of any research. Data analysis :-

Summarize collected data. It involves the interpretation of data gathered through the use of analytical and logical reasoning to determine patterns, relationships or trends.

Analyzing data :-

- * Visualize your data eg:- charts, tables, and graphs to mention a few.
- * Identify patterns, correlations and trends
- * Test your hypothesis
- * Let your data tell a story.

Report the results :-

- * Communicate and interpret the results
- * Conclude and recommended.
- * Your targeted audience must understand your result.

→ Tips :-

- * Use more data sets and samples
- * Use accessible and understandable data analytical tool.
- * Do not delegate your data analysis.

- * Analyze cleaned data.
- * Understand your results.

Advantages and Disadvantages :-

Analytics solutions are not difficult to implement, however, they are costly, and the ROI is not immediate.

Advantages :-

- * Reduced Expenses
- * Augmented Revenue
- * Enhanced customer service
- * Agility Expanded.
- * Faster time - to - market
- * More innovation
- * Detection of Fraud.
- * Improved and Increased productivity.

Disadvantages :-

- * Privacy and security concerns
- * Unreliable data quality.
- * cybersecurity Risks
- * legal and Regulatory Issues.
- * Hardware Needs
- * costs.

⑥

Applications :-

Data analytics can be applied in business in various ways. It can be used to analyze customer behaviour, optimize marketing campaigns, improve supply chain management, enhance product development, predict demand, identify fraud.

→ Security

→ Transportation

→ Risk management

→ Delivery

→ customer interactions

→ Health care

→ Education

→ Marketing and data advertising.

→ Travel.

→ Internet / web search.

⑦

Conclusion :-

Executive Summary

This report about my internship programs with smarbridge - IBM cognos. In this report, I have discussed about every major aspect

15

which I observed and practised during my internship programme.

Registering in smartbridge in data analytics through APSCHE - LMS portal has landed me into its pool of courses which are mainly focusing on the emerging technologies in the world. Out of many choices I opted for virtual internship sponsored by IBM academy.

The main objective of this internship is to identify trends and patterns, making data-driven decisions. Finding correlations and relationships. Detecting anomalies, improving performance and predictive modeling.

In detailed analysis of the organization technical and managerial skills, basic aspects that I have developed and the processes and procedures which I have undergone through during my internship program have been evaluated in this report.

②

Future Scope :-

Data analytics and data analytics bootcamp is a huge industry and is predicted to keep growing. It is expected to touch US\$ • 11.87 billion by 2020 as it keeps growing at a steady pace. This industry will disrupt the market causing a great shift in it and bringing several job opportunities with it.

Is data analyst a good career?

The average salary for an analyst in the United States is just under \$ 75,000 per year. The field is growing rapidly with a projected 23 percent increase in job opportunities from 2021 - 2031, far more than the average 5 percent, for any other industry. Companies such as Amazon, Facebook, and Google are actively seeking data analysts. This is a promising career choice with great potential for growth and development.

INTRODUCTION :-

SOCIAL MEDIA

Social media refers to a variety of technology that facilitate the sharing of ideas and information among their users. From Facebook and Instagram to X platform and YouTube more than 4.7 billion people use social media, equal to roughly 60% of the world's population. In early 2023, 94.8% of users accessed chat and messaging apps and websites followed closely by social platforms, with 94.6% of users.

KEY POINTS :-

- * Social media is digital technology that allows the sharing of ideas and information, including text and visuals, through virtual networks and communities.
- * Social media typically features user generated content that lends itself to engagement via likes, shares, comments, and discussion.
- * More than 4.7 billion people around the world use social media

* The largest social media platforms worldwide are facebook, youtube, whatsapp, instagram and we chat.

Advantages :-

- ① Connectivity :- Social media allows people to stay connected with friends and family, regardless of geographical distances.
- ② Information sharing :- It's a platform for sharing news, information and ideas quickly and easily.
- ③ Business Opportunities :- Social media is a powerful tool for marketing and promoting business and products.
- ④ Entertainment :- It provides a source of entertainment such as videos, games and memes.
- ⑤ Awareness and Activism :- It can raise awareness about important social and political issues and facilitates activism.

Disadvantages :-

- ① Privacy concerns :- Social media can lead to privacy breaches and the unauthorized sharing of personal information.
- ② Addiction :- Excessive use can lead to addiction and impact mental health negatively.
- ③ Cyberbullying :- It can be a platform for harassment, cyberbullying and online abuse.
- ④ Fake News :- Spread of false information and misinformation is a concern on social media.
- ⑤ Time Wasting :- Excessive use can lead to procrastination and reduced productivity.

IMPORTANCE OF SOCIAL MEDIA

IBM points out that with prevalence of social media : "News of a great product can spread like wild fire. And news about a bad product - or a bad experience with a

Customer service buzz - can spread just as quickly. Consumers are now holding organizations to account for their brand promises and sharing their experiences with friends, co-workers and public at large*.

- Spot trends related to offerings and brands
- Understand conversations - what is being said and how it is being received.
- Derive customer sentiment towards products and services.

Business problems associated with Social media :-

① Reputation Management :- Negative comments and reviews on social media can quickly damage a company's reputation. Effective reputation management involves monitoring online mentions, responding to negative feedback, and promoting positive interaction.

② Privacy concerns :- Business must adhere to data protection law and ensure that the personal information they collect from users is used.

3. content management :- Consistently creating and sharing high-quality content on multiple social media platforms can be time-consuming and challenging. Business need to develop a content strategy and schedule to keep their audience engaged.

④ competition :- The crowded nature of social media means that businesses must compete for users' attentions. This necessitates creativity in its content creation and potentially paid advertising to reach the target audience effectively.

⑤ Algorithm changes :- Social media platforms frequently update their algorithms, which can affect the visibility of posts. Business need to adapt and optimize their content strategies accordingly.



Social media platforms are online services or websites that allow users to create, share, and interact with content and other users.

① Facebook :- One of the largest and most widely used social networking sites, where users can connect with friends and family, share photos and post updates.

② Twitter :- A microblogging platform where users can share short posts or "tweets" and engage in real-time conversations.

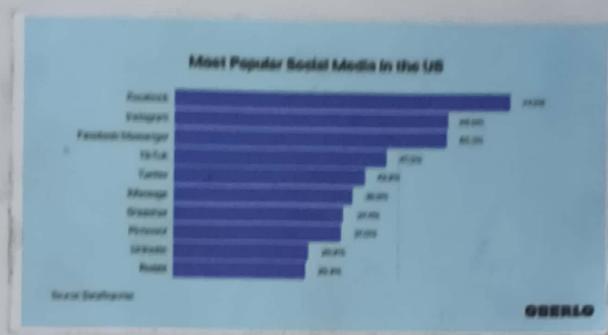
③ Instagram :- A photo and video-sharing platform that focuses on visual content and is popular for its use of hashtags and filters.

④ LinkedIn :- A professional networking site for job seekers, recruiters, and business professionals to connect and share content.



(6) Youtube :- A video-sharing platform where user can upload, watch, and interact with videos on a wide range of topics.

(6) Snapchat :- A multimedia messaging app that allows users to send self-destructing photos and videos, as well as post stories that disappear after 24 hours.



(7) Whatsapp :- A messaging app for text, voice, and video communication, as well as sharing media with contacts or groups.



Business Requirements for Social Media

1. Clear Objectives :- Define specific measurable goals for your social media efforts, whether it's increasing brand awareness, driving website traffic or boosting sales.
2. Target Audience Understanding :- Thoroughly understand your target audience demography, interests and behaviours to tailor your content and engagement strategies.
3. Content Strategy :- Develop a content plan that includes the type of content you'll create, a content calendar and a consistent brand choice.
4. Visual Branding :- Maintain a consistent visual identity across all social media profiles, including profile pictures, cover photos and post visuals.
5. Analytics and Measurement :- Use analytics tools to track key performance indicators and regularly assess the effectiveness of your social media strategy.

6. Advertising strategy :- If applicable develop a paid advertising strategy on social media platforms, including budgeting and targeting options.

Impact of Social media :-

Social Impact :-

1. Communication :- Social media has revolutionized the way people communicate connecting individuals across the globe and enabling instant communication.

2. Social Activism :- Social media has been a catalyst for social and political movements helping people organize and advocate for change.

3. Mental health :- Excessive use has raised concern about its impact on mental health contributing to issues like cyberbullying, addiction and anxiety.

Business Impact :-

1. E-commerce :- Social platforms increasingly support e-commerce, allowing business to sell products directly to consumers.

2. Influencer marketing :- Influencers on social media have become key for brand endorsement and product promotions.
3. Competitive Advantage :- A strong social media presence can give business a competitive advantages and build customer loyalty.
4. Data and Analytics :- Social media providers provides data for market research and customer insights, helping business make data-driven decisions.



DATA VISUALIZATION

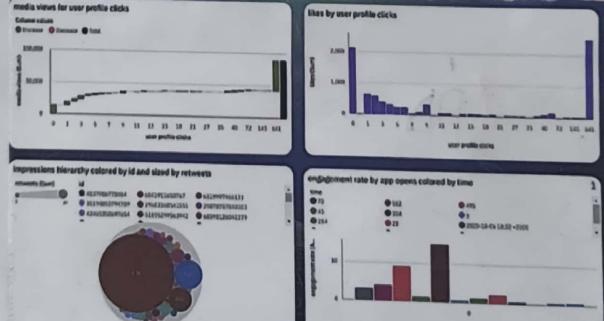
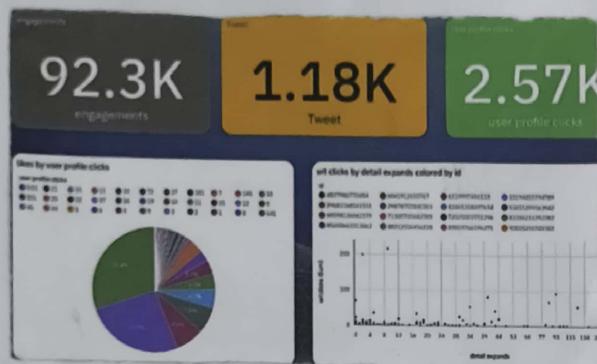
Data and information visualization is the practice of designing and creating easy to communicate and easy to understand graphic and visual representation of a large amount of complex quantitative and qualitative data and information with the help of static dynamic or interactive visual items. Typically based on data and information collected from a certain domain of expertise, these visualizations are intended for a broader audience to help them visually explore and discover quickly understand, interpret and gain important insights into otherwise difficult to identify structures, relationships, correlation, local and global patterns, trends, variations, constancy, cluster, outliers and unusual groupings within data. When intended for the general public to convey a concise version of known, specific information in a clear and engaging manner presentational or explanatory visualization.

Dashboard :-

A dashboard is a graphical user interface (GUI) that displays information and data in an organized, easy to read format. Dashboards are often used to provide real-time monitoring and analysis of data, and are typically designed for a specific purpose or use case. Dashboards can be used in a variety of settings, such as business, finance, manufacturing, healthcare, and many other industries.

In our dashboard we have provided 8 visualisations: KPI chart of tweets, pie chart for likes by user profile clicks and coloured id, line chart for the clicks by details expand, coloured by id, waterfall chart for media views for user profile clicks, column chart for likes by user specific clicks, summary chart shows that engagement and hierarchy bubble chart shows impressions hierarchy colored by id and sized by tweets.

✓ VISUALIZATION OF DASHBOARD



STORY :-

A data story is a way of presenting data and analytics in a narrative format, with the goal of making the information more engaging and easier to understand. A data story typically includes a clear introduction that sets the stage and explains the context for the data. Data stories can be told using a variety of mediums such as reports, presentation, interactive visualizations and videos.

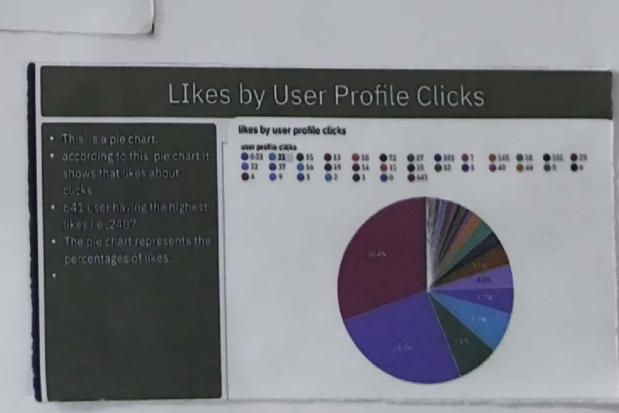
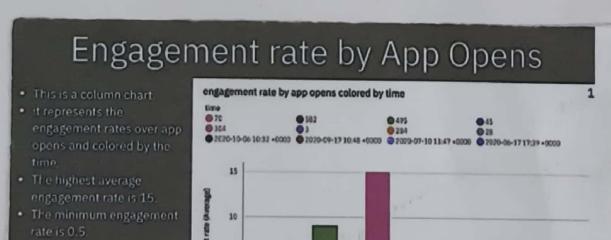
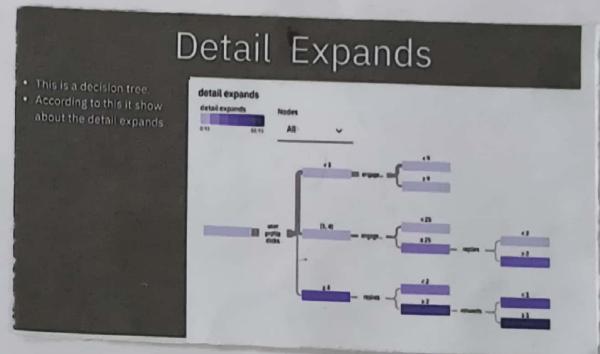
In our story we have taken 8 scenes the first scene is about the logo of social media and second scene is created using pie chart. it shows likes by URL clicks by profile clicks. The next scene created using line chart is represents URL clicks by details expands. The other scene is created using column chart it represents the engagement rates by App opens. The next scene created using KPI by tweets. This is about our story.

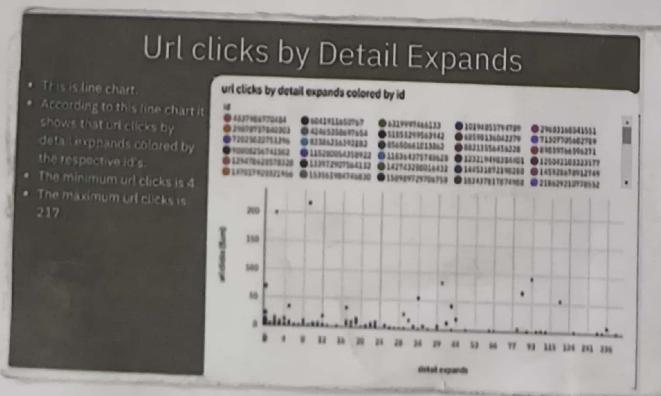
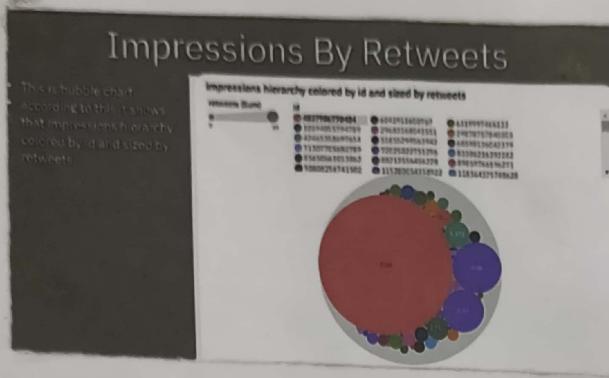
VISUALIZATIONS OF STORY

Tweet, id ,time and date

The table shows the information about tweet id, time and date.

Tweet	ID	Time	Date
a float-integer x < 25.0, i...	52798534736462533	2020-07-13 23:09 +0000	2020-07-13
a float-integer x > 26.0, i...	38843133023621599	2020-09-19 18:58 +0000	2020-09-19
a float-number x > 26.0, i...	13163108002047923	2020-09-17 14:23 +0000	2020-09-17
a float-number x > 26.0, i...	25250499256358339	2020-09-01 22:37 +0000	2020-09-01
a float-number x > 26.0, i...	646535363364913	2020-10-13 14:35 +0000	2020-10-13
a float-number x > 26.0, i...	5254442546488839	2020-07-14 14:23 +0000	2020-07-14
a float-in float-histogram p...	426627208529670	2020-06-17 19:39 +0000	2020-06-17
a float-in float-histogram p...	581050494717136218	2020-10-17 03:11 +0000	2020-10-17
a float-in float-histogram p...	884792338444350255	2020-07-31 14:39 +0000	2020-07-31
a float-in float-histogram p...	1335011954746820	2020-09-29 02:31 +0000	2020-09-29
a float-in float-histogram p...	79367648957123990	2020-08-09 18:41 +0000	2020-08-09



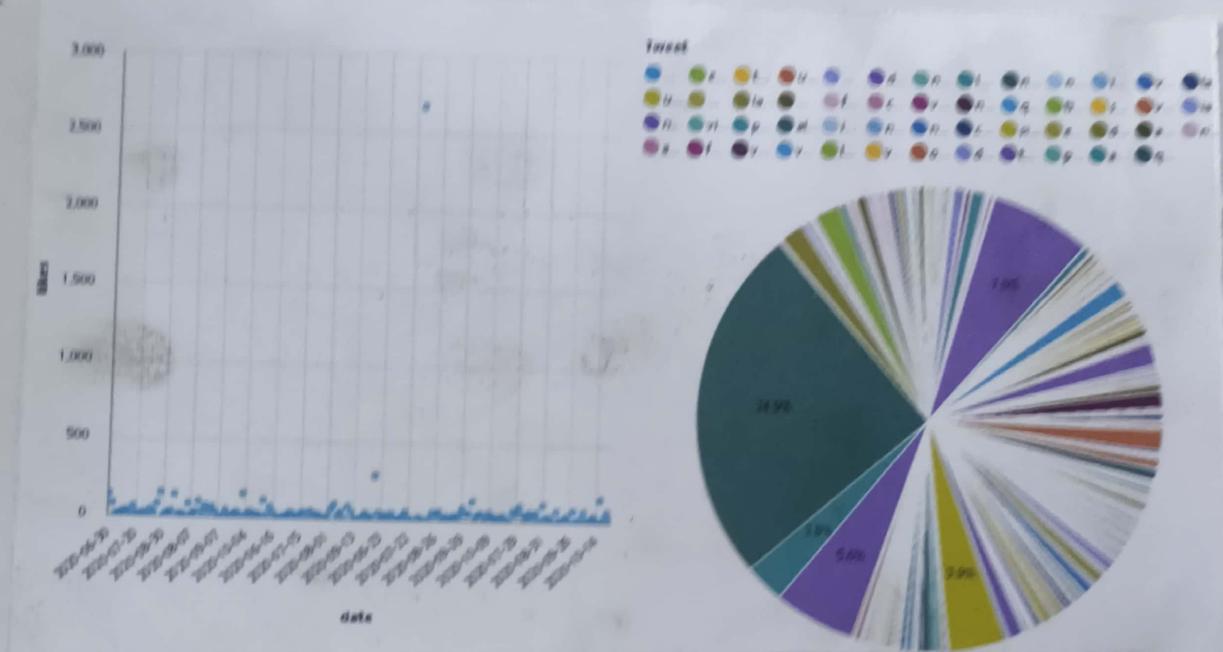


REPORT :-

A report is a document that presents information in a specific format and layout, usually based on data from a database or other data source. A report in IBM cognos can contain various elements, such as tables, charts, graphs and images as well as text and data elements. There are several different types of reports available in IBM cognos, including list reports, crosstab reports, chart reports and report studio reports.

In our report on a comprehensive analysis of social media we provided four visualizations. The report was composed by line chart shows date by likes, The pie chart was created by using tweets id's . The column chart is created using media views. The hierarchy bubble chart was created using impressions.

Visualizations Of Report



Conclusion :-

This report about my Internship program with smartbridge - IBM cognos . In this report , I have discussed about every major aspect which I observed and perceived during my internship programme.

Registering in Smartbridge in data analytics through APSCIE - LMS Portal has landed me into the pool of courses which are mainly focusing on the emerging technologies in the world . Out of many choices I opted for virtual Internship sponsored by IBM academy .

The main objective of this internship is to identifying trends and patterns , making data - driven decisions . Finding correlations , relations detecting anomalies , improving performance and predictive modeling .

In detailed analysis of the organization , technical and managerial skills , basis aspects that I have developed and the processes and procedures which I have undergone through during my internship programme have been evaluated in this report .