

Data Visualization and Report Generation Tool: A Survey

^{1st} Ashwaq Khazi
Pune Institute of Computer Technology
Pune, India

^{2st} Atharva Bhawe
Pune Institute of Computer Technology
Pune, India

^{4th} R. N. Shriwas
Pune Institute of Computer Technology
Pune, India

^{3nd} Shakti Tripathi
Pune Institute of Computer Technology
Pune, India

Abstract—Business Intelligence tools are software designed to retrieve, analyze, transform and report data for business intelligence. The platform will allow the user to query the database by simply using the User Interface without having the knowledge of SQL. This platform can be used for data preprocessing. The platform allows the user to perform operations like data cleaning, data transformation and data reduction. It will also provide the functionality of exploratory data analysis with the help of graphical representations. Determine whether online users are authorized to access digital objects to preserve privacy. Implementation of a global system for storing and evaluating access control lists. The system will perform authorization checks based on role permissions and grant access accordingly. Query performance will be enhanced through the optimization of table joins. The software can convert excel databases and CSV files to the respective relational databases by keeping the relationships intact. The enterprises can connect their databases and perform automated report generation on them. Reports are broadcast to respective authority using micro-services.

Index Terms—Business Intelligence

I. INTRODUCTION

Business intelligence(BI) is the process of transforming data into actionable lessons that help enterprises make clever and tactical business decisions by leveraging on software and services. To provide customers with comprehensive intelligence on the state of the market, BI tools navigate and evaluate data sets and display empirical results in analyses, summaries, dashboards, graphs, charts, and maps [1]. Data visualizations are surprisingly common in our everyday life, but they often appear in the form of well-known charts and graphs. It can be used to discover unknown facts and trends. Good data visualizations are created when communication, data science, and design collide [2]. In this paper, we will explore the daily workflow of an BI Analyst and the challenges faced by BI Analyst . This paper further investigates the available solutions in the industry to tackle these challenges. We study how BI tools can help solve these issues, survey four different BI tools commercially available in the market, compare their features and study their usability. We identify the problems solved by

these platforms and the challenges faced by BI analysts that remain persistent.

II. BI WORKFLOW

The following diagram depicts the daily routine or workflow of an ML Engineer:

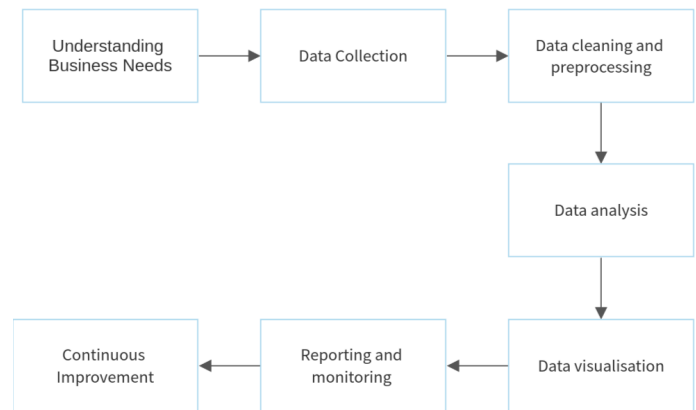


Fig. 1. Daily Routine of an BI Analyst

III. CHALLENGES IN BI ANALYST WORKFLOW

The most common issue faced by organizations is data quality. The quality of the data can significantly impact the accuracy of analysis. Missing or incomplete data, inconsistent data formats, and errors in data collection can all pose challenges to a BI Analyst’s workflow. Starlight attempts at solving this issue by robust data cleaning techniques and it also allows the client to contribute in the process. Another major issue faced by BI analysts is data integration. Integrating data from different sources can be challenging. Each system may use different data formats and structures, and it can be challenging to merge them into a single cohesive dataset. Starlight aims to solve this issue by implementing a Data Lake of relational, non relational, structured and unstructured data from multiple

data sources. It can also maintain relationship across entities of different sources.

Maintaining data security and privacy is critical and challenging to be faced by BI analysts, and it's essential to ensure that sensitive information is not accessible to unauthorized individuals. It can be challenging to balance the need for data accessibility with the need for data security. Starlight maintains accessibility and security across its microservices by implementing security standards of Zanzibar[6]

The ability to communicate insights effectively is essential. It can be challenging to present complex data in a way that is accessible to non-technical stakeholders. BI analysts must be able to communicate their findings in a clear and concise manner. Starlight has various charts for visualising data in order to find and share insights. It also provides a feature of sharing dashboards where in multiple clients can create modify and update the graphs and the respective changes will be visible to all the clients using that dashboard in real time. The BI tool that we are building also generates reports from these dashboards which can be mailed to respective accounts.

BI analysts must adopt new tools and software as well as keep up with technological advancements. It can be difficult to stay current with fashion, especially given how quickly technology is developing.

As BI analysts work with different data sets and analytics tools, they need to continuously learn and develop new skills to keep up with the latest trends and technologies in the industry. This requires a commitment to ongoing learning and professional development.

Finally, a BI analyst has to do a lot of repetitive tasks like extracting data transforming data, and loading. This is an iterative process and takes up a lot of time.

IV. BI TOOL

V. AVAILABLE TOOLS

In this section, we explore the existing BI tools and platforms. The BI platforms selected for this survey are : (1) PowerBI (2) Tableau (3) QlikView (4) SAP BusinessObjects

A. PowerBI

Power BI is Microsoft's most recent Self-service BI offering. Microsoft Power BI solution gives business user oriented data analysis and visualization capabilities to upgrade decision-making process and the business visions. Power BI is a cloud-based Self-service BI solution, which means you can build and deploy solution immediately with data from cloud and on-premises data sources, systems and function [3]. In Power BI, the different types of data are fetched by means of getting data function from different data sources and the different data sources are Files, Database, Azure and Online Services . After selecting the Data sources, we have to get the data by means of queries (if the data source is other than files) or selecting files from folders. The data will be loaded in the Power BI tool [5]. Before making the report in the power BI tool, the uploaded data should be Analysed and Corrected for error freed data through edit queries function

in data part of Power BI tool and we can have relations between different datasets by means of relationships part. The analyzed and corrected data is visualized by means of different Visuals such as Custom Visuals and Power BI Visuals. [5]

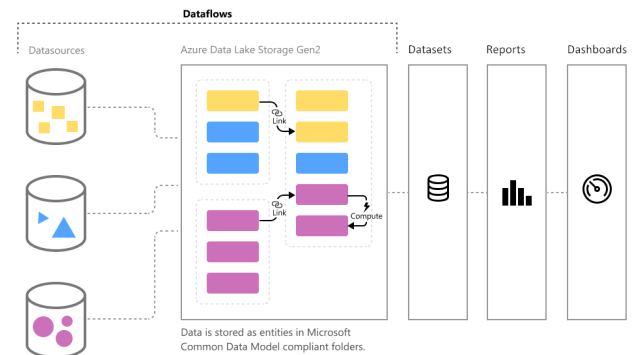


Fig. 2. PowerBI

However, PowerBI reports and dashboards can be only shared with same email domain users or with those listed in office 365 tenant. PowerBI doesn't handle large data sources properly. Users also report that it is complex to use and expensive.

B. Tableau

A visualization and dashboarding software, Tableau gives the user a unique capability to blend the data, explore the data by query the data at different level of aggregation, the exploration and discovery leads to appropriate question which triggers a round of further probing and analysis. Tableau enables the user to present the story in the data in a very compelling way at different levels of aggregation. The ease of the menu driven point and click and drag and drop feature of the software has democratized the visualization and the dashboarding skill [4].

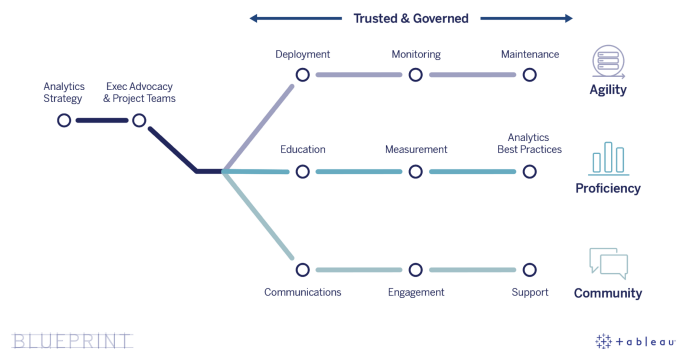


Fig. 3. Tableau's Blueprint

Tableau

However, Tableau does not offer the option to automatically refresh reports using scheduling. Tableau isn't a fully functional open tool. Any new visuals must be created from scratch rather than imported. Tableau is solely a tool for visualizing data. You are able to perform very basic preprocessing using Tableau Desktop.

C. Looker

You can explore, share, and visualise data using the cloud-based business intelligence (BI) tool Looker, which enables you to make better business decisions. The Google Cloud Platform now includes Looker. Anyone in your company can use it to quickly analyse and glean insights from your datasets.

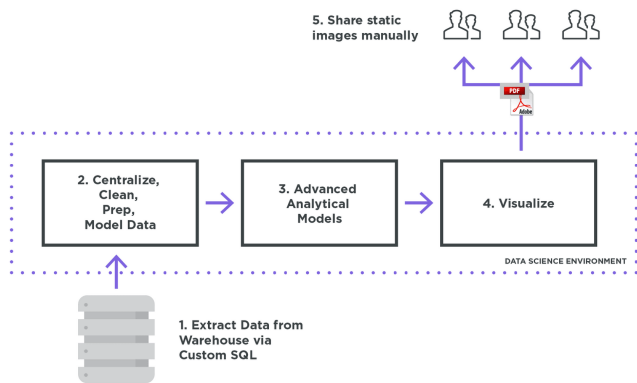


Fig. 4. Looker Workflow

Looker can be slow at times. The initial setup for Looker can be tough. Looker is an expensive platform depending on level of business.

D. QlikView

You can create interactive guided analytics applications and dashboards quickly with QlikView. You can discover your own insights by posing and addressing your own questions. You can make decisions with the help of your coworkers. For data integration, conversational analytics, and turning unstructured data into a knowledge base, use QlikView, a business intelligence tool. Ad hoc query features are available in QlikView, which enables quick decision-making based on readily available data. This program uses an associative model and can look for information in any direction.

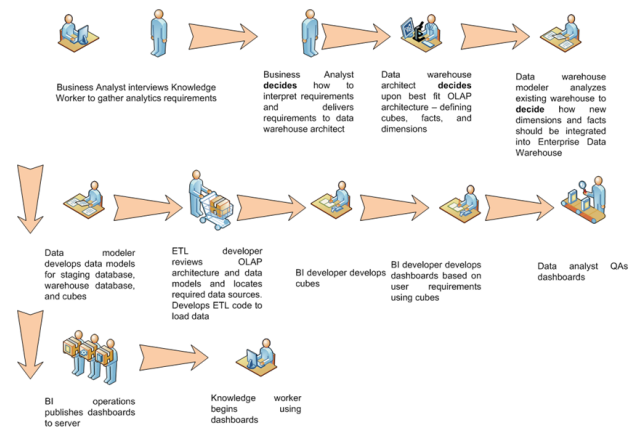


Fig. 5. QlikView Workflow

Since Qlikview needs more RAM, using it is more computationally expensive. With QlikView, real-time data analysis is occasionally impossible. It's slower to use because it lacks more recent drag-and-drop technology. It needs a lot of additional features, all of which must be paid for separately. Because of this, using QlikView is more expensive. Its dashboard is more cluttered and has a less user-friendly interface. Application development, embedding or integrating with other software, and even scaling present additional challenges for users.

VI. COMPARISON OF DIFFERENT BI TOOLS

The table given below compares different features that are present/absent in the surveyed BI platforms.

TABLE I
 COMPARISON OF DIFFERENT BI PLATFORMS

Feature \ Tool	PowerBI	Tableau
Data Visualization	Moderate	Strong
Data Modelling	Moderate	Moderate
Data Integration	Strong	Moderate
Collabroration	Moderate	Moderate
Mobile Support	Yes	Yes
Advanced Analytics	Moderate	Moderate
Pricing	Subscription Based	Subscription Based
Deployment	Cloud Based	On-Premise/Cloud Based
Vendor	Microsoft	Tableau Software

VII. RESEARCH GAP ANALYSIS

The main hurdles to the widespread adoption of BI tools are the high cost, complexity to use, extensive infrastructure, and limited scalability. Users get overwhelmed by the UI many existing BI tools offer and their complexities. We recognize that there is currently no BI tool that can simultaneously visualize, report, maintains data quality, security and integration that is affordable, scalable, easy to use, and does not require extensive infrastructure. There is a need for such an BI tool with intuitive UI dashboards. The examination of the various BI tools points

to a new research area for developing a completely automated user interface-based BI tool that can be used by domain experts and developers.

VIII. PROPOSED METHODOLOGY

Creating a software which is a web based application. The client registers or logs themselves in to access their dashboard. This dashboard is completely customisable to the client. They can add graphs and charts of data in the form of widgets. They can also perform complex queries through UI. All of his is handled by NextJS which is the frontend technology along with Nivo charts as a charting library and ChakraUI for UI components. Reports are generated based on a template created by the client. This template must contain respective queries for the dynamic fields of the report. Client can visualize the ELT process and also create their own transformation pipelines. REST architecture is followed for communication between the frontend and backend. The backend is golang microservices which is responsible for handling the security and business logic for the application. Fivetran is used to sync data from various sources of client into Snowflake, which acts as a Data Lake. The microservices are deployed on Kubernetes cluster in order to make scaling of the application simpler.

IX. CONCLUSION

This paper is based on the demonstration of a data visualization and report generation tool. This project utilizes various technologies to achieve our goals such as Next.js for serving users with a UI and Golang as the primary language for the development of other services that are required in the backend. This project offers any organization or individual extensive potential by diversifying access to business intelligence and supporting data-driven decision-making, whereby they can easily provide data, build reports or have them automatically created, aggregated in dashboards, and shared with ease. Business analysts can leverage powerful analysis capabilities for easy data exploration, discovery, and report generation. Data can be viewed and analyzed in one place with live dashboards and reports. In addition, work must be done to analyze various other features to integrate many more graphs in the visualization feature. Further, user collaboration can be added as a feature in the project. Data sources can also be extended to include social media such as Twitter, Instagram, etc. Moreover, this project can be used as an API service. We recognize that there is currently no BI platform that is affordable, scalable, easy to use, and does not require extensive infrastructure. There is a need for such an BI platform with intuitive UI and extensive support for Data Visualization. The examination of the various platforms points to a new research area for developing a completely automated user interface-based BI tool that can be used by domain experts and developers.

REFERENCES

- [1] Meet Joshi , Ashwini Dubbawar, 2021, Review on Business Intelligence, Its Tools and Techniques, and Advantages and Disadvantages, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH TECHNOLOGY (IJERT) Volume 10, Issue 12 (December 2021)
- [2] M. Islam and S. Jin, "An Overview of Data Visualization," 2019 International Conference on Information Science and Communications Technologies (ICISCT), Tashkent, Uzbekistan, 2019, pp. 1-7, doi: 10.1109/ICISCT47635.2019.9012031.
- [3] Amrapali Bansal, A. K. Upadhyay, "Microsoft Power BI" International Journal of Soft Computing and Engineering (IJSC) ISSN: 2231-2307, Volume-7 Issue-3, July 2017
- [4] Brijesh R. Sharma, Sachin Deshmukh, "Data Visualization for Accelerated Business Intelligence in the Indian Health Care Sector using Tableau" International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878 (Online), Volume-9 Issue-4 November 2020
- [5] Mandava Geetha Bhargava, K. Tara Phani Surya Kiran Duvvada Rajeswara Rao, "Analysis and Design of Visualization of Educational Institution Database using Power BI Tool" Global Journal of Computer Science and Technology Volume 1 Issue 4 Version 1.0 Year 2018 Online ISSN: 0975-4172 Print ISSN: 0975-4350
- [6] Ruoming Pang, Ramon Caceres, Mike Burrows, Zhifeng Chen, Pratik Dave, Nathan Germer, Alexander Golynski, Kevin Graney, Nina Kang, Lea Kissner, Jeffrey L. Korn, Abhishek Parmar, Christina D. Richards, Mengzhi Wang (2019). Zanzibar: Google's Consistent, Global Authorization System. In 2019 USENIX Annual Technical Conference (USENIX ATC '19).