



iJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 6 Issue: IV Month of publication: April 2018

DOI: <http://doi.org/10.22214/ijraset.2018.4761>

www.ijraset.com

Call: ☎ 08813907089

E-mail ID: ijraset@gmail.com

Building Analytics Dashboard for Visualizing Data

Aishwarya Nair¹, Amruta Pandit², Ketaki Naik³, Payal Rajput⁴, Priyanka Andhalkar⁵

^{1, 2, 3, 4, 5}Dept of Information Technology, Bharati Vidyapeeth's College of Engineering for Women, Pune, India

Abstract: *Dealing with data, nowadays is the most crucial task to perform when it comes to company level, wherein we understand their users and make possible efforts to make them love your product. So there is a need to represent this data in effective way, so that it is easy to understand and interpret results from it. Dashboards is the most prominent and efficient option in the market due to its interface viewing capability. From responsive layouts, interactive charts and innovative color pallets to labeled icons for navigation and metro-based design, today's dashboards are setting new standards in usability and interaction design which enables business intelligence to evolve and user specific actions in order to enhance user experience.*

Keywords: *Big Data, Data Analytics, Visualizing tools, business Intelligence, dashboard*

I. INTRODUCTION

Data analytics deals with extracting similar patterns from the large data sets. Data analytics helps you to understand the data before working on it and you can refine the data to extract information that allow you to avail this data to give you and your organization a competitive advantage. Data visualization uses images to convey knowledge. For obtaining better data context, focusing on the right information is a key point in business strategy. The main objective of data visualization is to convey information clearly and effectively by using statistical graphics, plots and information graphic. Using dots, lines, or bars numerical data can visually communicate a quantitative message. Effective visualization is when users are able to analyze and make sense of the information available. Complex data is then made easier to understand. The project aims at generating analytics graphs including bar graph, line graph, pie chart for the provided dataset. Applying queries on database a resultant dataset is generated. Further, this resultant dataset is processed using aggregation queries. Then this resultant set is processed using JS Libraries to generate query specific Analytic Graphs. The dashboard aims at generating various different types of Graphs according to its need.

II. BACKGROUND

Earthmiles is a Health app where they put their health products for sale and through this app they also encourage various health related exercises such as running, walking, yoga, etc., they give some points based on this exercises. The user can purchase the products hosted on the app and get some discounts by using the points. For the track of user activity company is collecting the user information and putting it in database. The data of ever instance that user is spending on this app is been collected. As the company was a startup company earlier it was easy to manage the user data but it's becoming popular which resulted in the increase in user's and as a result increasing user data. So to manage all the user information and to track the user's activities the visual data analysis became the key need.

III. NEED

As the company is growing the database which contains sensitive and secrete user data that is being collected is also growing and at the same time such a big data handling and analysis is also important, which cannot be done manually as there is increase in number of users every day. There are many techniques now days available like Google analytics but as the data is secrete and sensitive we cannot share our company data with any other company as there can be risk of outsourcing of data by them. At the same time if we think of purchasing customized software's company cannot bare the amount to purchase the software and its annual license renewal charges. To tackle all the above difficulties the solution of designing an personalized dashboard was proposed which will be owned by the company so there will be no threat to data and this will be one time investment which is profitable for the company. At the same time company specific changes could be easily managed as per the upgrading in near future.

IV. FEATURES

A dashboard is a tool that can be used to easily understand and analyse user information in a one page format. Following are some important features of dashboard:

A. Customisable

Dashboards can be customized according to the need of user in order to meet their expectations.

B. Drill into Detail

User can access detailed data using graphs and view it yearly, monthly or daily basis according to its need.

C. Perceptive Data Representation

Reading a lot of information and then interpreting results is tedious task instead viewing graphs and understanding it is simple and easier as well.

V. METHODOLOGY

Multiple sets of refined information is used in a single page layout which allow companies to understand user behavior and data patterns and take action accordingly.

A. Architecture

The project aims at generating analytics graphs including bar graph, line graph and pie chart for the provided dataset. Applying queries on database a resultant dataset is generated. Furthermore, this resultant dataset is processed using aggregation queries. Then this resultant set is processed using JS Libraries to generate query specific Analytic Graphs.

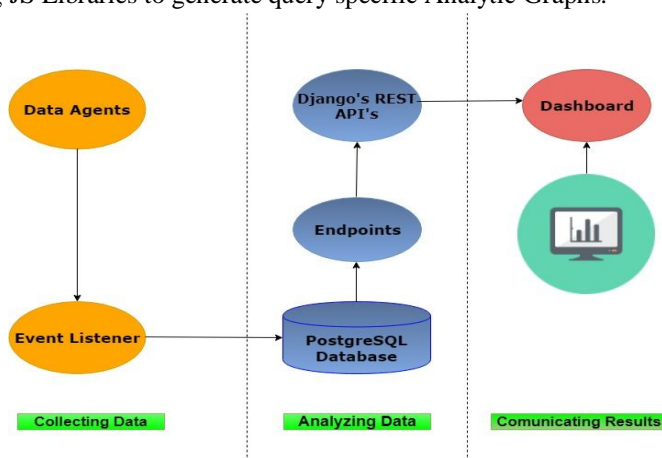


Fig. 1. Architecture of entire methodology i.e. from where data is collected, it is then analyzed and at last results are interpreted.

Input will be a huge database in the PostgreSQL form. Now the expected output of this particular module is to obtain parameter specific data from the entire db, as needed for graphs. To obtain this parameter specific data, we need query on the db accordingly. This output will be in the form of array for which we will be using python Django framework using which we are querying in our db.

B. Module Split Up

- 1) *Querying Required Data*; Input will be a huge Database in the PostgreSQL form. Now the expected output of this particular module is to obtain parameter specific data from the entire db, as needed for graphs. To obtain this parameter specific data, we need query on the db accordingly. This output will be in the form of array for which we will be using python django framework using which we are querying in our db.
- 2) *Identify Endpoints*: he endpoints for the required graphs are considered and plotted. Endpoints are generated using the resultant set obtained by querying dataset. Then these endpoints are accessed using API's. API response will be JSON format.
- 3) *Formation of Graphs*
 - a) *New User's Record*: This graph depicts New User's Growth in the App. Further user can also view it according to his need i.e. particular day, weekly growth, monthly growth or yearly growth. So the drop box will contain following parameters according to which the graph can be customized further (Today, Week, Month, Year).
 - b) *Active User's Record*: This graph depicts Active User's Record in the App. Further user can also view it according to his need i.e. particular day, weekly growth, monthly growth or yearly growth. So the drop box will contain following parameters according to which the graph can be customized further (Today, Week, Month, Year).
 - c) *Sessions*

- i) *Part 1: Sessions:* This graph depicts number of sessions made in the App. Further user can also view it according to his need i.e. particular day, weekly growth, monthly growth or yearly growth. So the drop box will contain following parameters according to which the graph can be customized further (Today, Week, Month, Year).
- ii) *Part 2: Session Length:* This graph helps in learning the session length. It helps in analyzing what is the maximum or minimum time sessions were made which in turn results in knowing the Customer interest on the particular page.
- d) *Top Versions:* This graph helps to learn what number of crowd is using which version of App the session length. It helps in analyzing what is the maximum or minimum time sessions were made which in turn results in knowing the Customer interest on the particular page. Further user can also view it according to his need i.e. particular day, weekly growth, monthly growth or yearly growth. So the drop box will contain following parameters according to which the graph can be customized further (Today, Week, Month, Year).
- e) *Page Views:* This graph helps in determining the Page that is viewed most by the customer's. Further user can also view it according to his need i.e. particular day, weekly growth, monthly growth or yearly growth. So the drop box will contain following parameters according to which the graph can be customized further (Today, Week, Month, Year). As there are number of pages so the user is also provided with the option to select a particular page and give the analytic performance of that particular page, so drop box contents will be the entire pages list.
- f) *Activity Diagram:* This pie chart helps to know the community of people according to their interests.
- g) *Age Groups:* This pie chart helps in identifying the crowd categorized as Age Groups.
- h) *Gender-Wise Groups:* This pie chart helps in identifying the crowd categorized as Genders.
- i) *Event Records:* day, weekly growth, monthly growth or yearly growth. So the drop box will contain following parameters according to which the graph can be customized further (Today, Week, Month, Year). As there are number of custom events so the user is also provided with the option to select a particular event and give the analytic performance of that particular event, so drop box contents will be the entire events list.
- j) *OS User's:* This graph evaluates to understand the population of customer using the particular Operating System.
- k) *Customizing Dashboard:* The dashboard is customized according to its need and in the navigation bar under various parameters graphs are created.

VI. LITERATURE SURVEY

A. Social Set Visualizer: A Set Theoretical Approach to Big Social Data Analytics of Real-World Events

It proposes and illustrates an alternate holistic approach to big social data analytics, social set analysis (SSA), which is based on the sociology of associations, mathematics of set theory, and advanced visual analytics of event studies[1]. Here, they have used Social Set Visualizer (SoSeVi) to create the analytic dashboard. They have collected big data which comes from social media, in this case Facebook, based on user interactions. It targets corporate social responsibility (CSR) researchers and practitioners. They have used D3.js visual analytics libraries. The advantages of this over the project are that the graphs are created easily which help in analysis of the user interactions and the use of D3.js is used to generate the graphs. Its disadvantages can be that big data is used. Also the SSA software which facilitates quantitative or qualitative analysis of social networks, describing features of a network either through numerical or visual representation approach is used. Another limitation is the lack of exposition of the full range of set theoretical approaches beyond the classical "crisp sets" discussed in the paper (for example: fuzzy sets, rough sets, random sets, Bayesian sets). Limited space devoted to the technical aspects of the IT-artifact.

B. Keshif: Rapid and Expressive Tabular Data Exploration for Novices

General purpose graphical interfaces for data explorations are typically based on manual visualizations and interactions specifications. While designing manual specification can be very expressive, it demands high efforts to make effective decisions, therefore reducing exploratory speed. Instead principled automated designs can increase exploratory speed, decreasing the learning efforts, which helps in avoiding ineffective decisions, and therefore better support data analytics novices. Keshif, a new systematic design for tabular data exploration. To summarize a given dataset, Keshif aggregates records by value within attributes summaries, and visualize aggregate characteristics using a consistent design based on data types[3]. The advantages are that it is effective, consistent out-of-the-box visualizations with zero effort. No design tools or wizards are needed. Keshif automated, minimalist model also supports simple drag-and-drop authoring of dashboards from raw data and it creates interactive charts. It is a web based technology which is not needed to download, load data and create charts and dashboards online. The disadvantages can be that Keshif free version is not useful for industrial purpose so need to purchase its annual package. It doesn't guarantee data security as everything is on the network. It allows only limited amount of data to be loaded at a time.

C. Applications of Excel VBA in Data Calculations and Charts Creation of Mineral Resources Economics

Mineral resource economics usually needs some calculating and analyzing work with a large amount of data of mineral resources stored in Excel files. Excel VBA provides a fast and easy way to finish the calculations work and create various charts to help analyzing work. Excel VBA is instructed at first. Then based on Excel VBA, some examples for calculating annual increment of mineral resources, creating batch charts from data and generating a comparing chart from resources data are illustrated in details. Finally, the conclusion that Excel VBA programming can save lots of time and reduce errors is made[4]. has the advantages such as it has a very user friendly development environment. The user can implement numerical methods as well as automating tasks such as formatting or data organization in VBA and guide the calculation using any desired intermediate results reported back to the spreadsheet. Automation of usually implemented tasks can be done by macros which contains code of that task and it is used to control many aspects of the host applications, including manipulating user interface features. The disadvantages with respect to the project would be that small changes in the excel would make changes in the macro code manually. It is supported by Microsoft Office applications only. Excel is not freely available for commercial use, it needs to be licensed and also has to be renewed after every particular interval.

D. Graphic Visualization of Specific Dashboards in Transport Technologies

This paper describes a web application that visualizes the results of monitoring traffic in city centers. Interactive graphs have been created. Instead of using existing JavaScript libraries like Chart.js, D3.js and Flot, they have created their own JavaScript library called “inCharts”. This library basically uses Canvas technique for rendering process. The library supports three basic types of charts- pie charts and line charts. These charts are designed to show specific data from information for traffic monitoring[5].

Also a system of interactive objects and an event manager were created. The manager will control any interactivity. It also supports animations. Its advantages are that a new JavaScript library called “inCharts” is created. The complexity to use this application is very low. It is created using simple technologies like JavaScript and HTML 5 which makes it very easy to understand. And the disadvantages can be that it is very simple and only supports three types of graphs i.e. bar graph, line graph and pie charts.

Since canvas element is used, its adjustment needs to be done by CSS while redrawing it results in deformation of data and loss of quality.

E. A Web Application for Content Based Geographic Image Retrieval

A web application is presented for detecting locations with content-based similarity on high resolution ortho-imagery via histogram matching methods. For this application, a web interface utilizing Bing Maps for determining query image was constructed via ‘Python Flask Framework’[2].

This web interface uses the center location of the Bing Maps extent, and the image closest to the center location is used as a query for histogram based matching with other 29106 images in the collection and presenting the user defined number of similar images. In order to determine general performance of the application in the dataset, 600 of the images were labeled into 6 classes and with 100 images per class. The application was also implemented by using these labeled images and each of them came into use as a query image. The results were calculated as MAP (Mean Average Precision) of each class containing 100 images for each histogram matching method used. The advantages are that Flask provides simplicity, flexibility and fine-grained control. It is un-opinionated i.e. it lets you decide how you want to implement things. Useful for developing small and simple applications. As it is a micro framework, it needs less memory space. And the disadvantages are that it is a new micro framework so many templates are missing which would be needed. It does not support large web applications.

VII. TECHNOLOGIES USED

- A. Python (programming language for backend)
- B. Django (backend web framework)
- C. PostgreSQL (RDBMS Database engine)
- D. JavaScript (web frontend language)
- E. HTML & CSS for web
- F. Chart.js library for building graphs

VIII. CONCLUSION

A. Hence, following key objectives are achieved

- 1) Faster reporting, analysis or planning



- 2) Fast and effective decision-making
- 3) On demand, accurate and relevant information in line with business priorities
- 4) Focused identification of problems, inefficiencies or negative trends for immediate action and improved performance
- 5) Improved operational efficiency
- 6) Increased competitive advantage
- 7) Increased revenues

REFERENCES

- [1] Benjamin Flesch, Ravi Vatrpu, Raghava Rao Mukkamala, Abid Hussain "Social Set Visualizer: A Set Theoretical Approach to Big Social Data Analytics of Real-World Events" 2015 IEEE International Conference on Big Data (Big Data).
- [2] Naime Celik, Hasan Ogul "A web application for content based geographic image retrieval" IEEE 2017.
- [3] Mehmet Adil Yalçın, Niklas Elmqvist, Senior Member, IEEE, and Benjamin B. Bederson "Keshif: Rapid and Expressive Tabular Data Exploration for Novices"
- [4] Liu Pengpeng, Du Ruiqing : "Applications of Excel VBA in Data Calculation and Charts Creation of Mineral Resource Economics"
- [5] Pavel Pokorný, Kamil Stokláška "Graphics Visualization of Specific Dashboards in Transport Technologies" 2016 Third International Conference on Mathematics and Computers in Sciences and in Industry.



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)