



Big Data for the Enterprise

Meznah deajj alotaibi¹, Areej Abdulla Alshehri², Ayshah Yahya Ghazwani³, Rwan Abdullah Darwesh⁴ and Samar Abdulla Alzahrani⁵

Department of Management Information Systems, Imam Abdulrahman Bin Faisal University (IAU),
Dammam, Kingdom of Saudi Arabia

ABSTRACT

This paper discusses big data for enterprise; an explanation is included on what it is and its benefit. Also, it mentions how Facebook takes advantage of big data and some useful tools to use for it like Tableau and Rapid Miner.

Keywords: *Big Data, Enterprise, Useful Tools, Tableau and Rapid Miner.*

1 INTRODUCTION

Big data describes the large amount of data both structured and unstructured that flood businesses daily. We can articulate the definition of big data as the four V's:

- Volume: the sheer amount of data being generated.
- Variety: structured, unstructured or semi-structured data like text or images.
- Velocity: time-sensitive processes that must be dealt with in a timely manner.
- Veracity: uncertainty of data.

Enterprises today gather more data than ever before. Majority of it is difficult to analyze, but nevertheless the insights contained within this data can be very valuable. Both analysts and business leaders agree that applying and implementing analytics to big data is crucial for organizations in their quest to stay competitive. One of the most popular social networking sites today is Facebook, which was launched in 2004 and comprise approximately 2 billion accounts. With the obsession of social media, the number of people on Facebook has grown tremendously producing a large amount of data by the minute. Facebook uses data mining to track users' behavior, whether it is their wall posts, their favorite movies or books, and clicks and likes, they analyze each and every bit of their data to offer them better services with each time users log in. Now, there are new technologies emerging for exploring and analyzing data that are more advanced and easy to use for not only organizations but everyone else who is interested in big data and data mining. Some of these technologies are data miner and tableau, which are

data visualization and analysis tools that helps in data discovery, identifying patterns and gaining insights.

2 LITERATURE REVIEW

The amount of data in the world right now is massive and analyzing it will become a vital key for competition, productivity growth, innovation and success. Researching and integrating data is essential to companies in any field, that's why the demand for data specialists or scientists is increasing. Lisa-Christina Winter who is a Digital Marketing and Data Science Specialist says: "I see a bright future in the field of big data for 2018;" also, "The big data movement will not cease to surprise us with amazing new applications and results anytime soon." [1] Big data can also be used In Healthcare, for example: Preventing Medication Errors, like taking the wrong medication or the wrong dosage. Ben Rogojan who is a data engineer at Healthentic says: "With the massive amount of digital medical images available, data scientists and big data professionals have the ability to make doctors work more efficient," [1] Another field that big data can be useful in is banking, it can help in Fraud Detection & Prevention, Customer Segmentation, and Enhanced Compliance Reporting. Sam Kumar, Global Head of Analytics, Standard Chartered Bank says: "Data offers the intelligence to make sure what the client is being offered by the bank has value and relevance to their choices." [2]

3 RELATED WORK

Impact of Big data analytics on enterprise:

Nowadays enterprises are investing in technologies that collect valuable data for decision making and forecasting the future of the business. Big data analytics have huge impact on spotting new and potential trends. Moreover, they help uncover the hidden patterns behind the company critical performance in the industry it is part of and customers' analysis. A study published by European Business Application Research Center (BARC) reveals interesting numbers on the improvements companies achieved after investing in big data technologies. The study shows that strategic management improved by 69% [1], and better interpreting of customer needs and behavior by 52% [3] (BARC - Business Application Research Center, 2018). If a company owns valuable amount of big data on certain aspect such as numbers of highest user engagement area in a social network, this probably creates a business opportunity to introduce these users as potential customers for advertisement. Many social networks now like Facebook and Instagram now have presented for users customized content of advertisements and prediction of what could the user do. These improvements are one of the results of big data analytics.

4 RELATED WORK

Tableau: can help anyone see and understand the data.

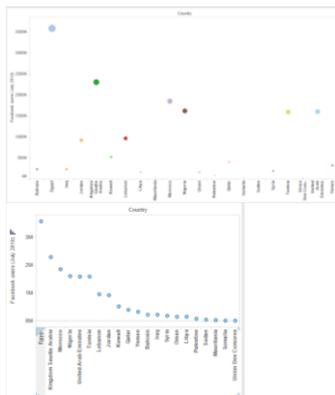


Fig. 1. Here we use a CLASSIFICATION: Of the Arab countries used for Facebook and are arranged in descending order. It is easy to read the information available here.

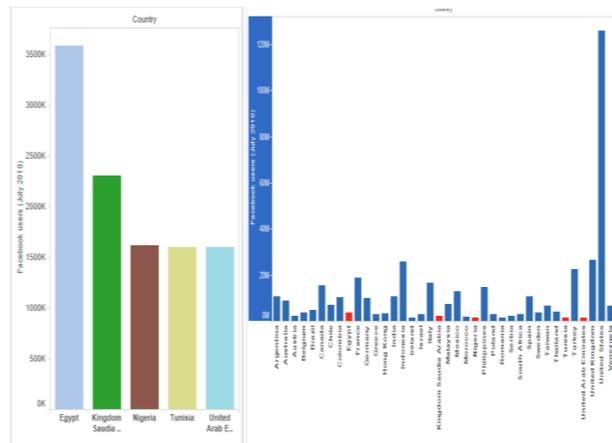


Fig. 2. REPORT OPERATION: Requires hypothesis validation by asking questions.

"Most Facebook users are from non-Arab countries".

Do Arab countries not use Facebook or are they some countries use it, and some do not? Do non-Arab countries use Facebook more? After the analysis, these questions were answered. The second chart the blue color for Non-Arab countries and the red for the Arab countries.

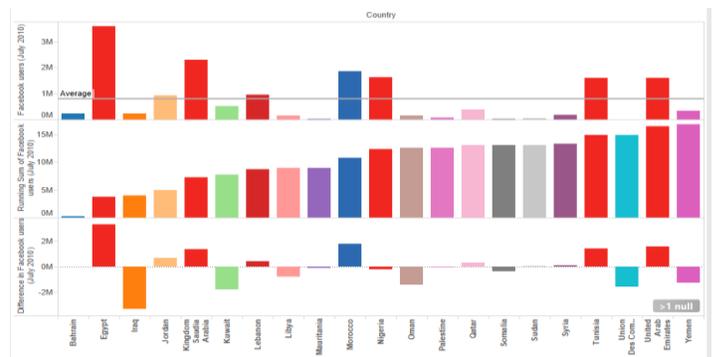


Fig. 3. STATISTICAL ANALYSIS: We make a comparison between the average in the first chart and running total in the second, the last chart we use a difference.

RapidMiner:

1. CLUSTERING: This operator uses K-means algorithm to perform clustering in RapidMiner by clustering groups together which are similar to each other.

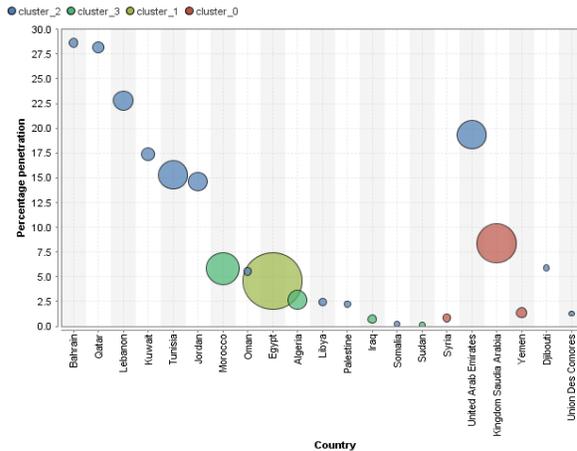


Fig. 4. Arab Nations.

The cluster divided the Arab Nations into four groups that is similar to each other in Facebook users size and each group distinguishing itself by different colors.

1. **BLENDING:** We choose from Blending the Filter operator. It can manipulate the data. For example, Select, Remove, and Delete rows cases for the data. After manipulating the data, then we make them in ascending or descending order.

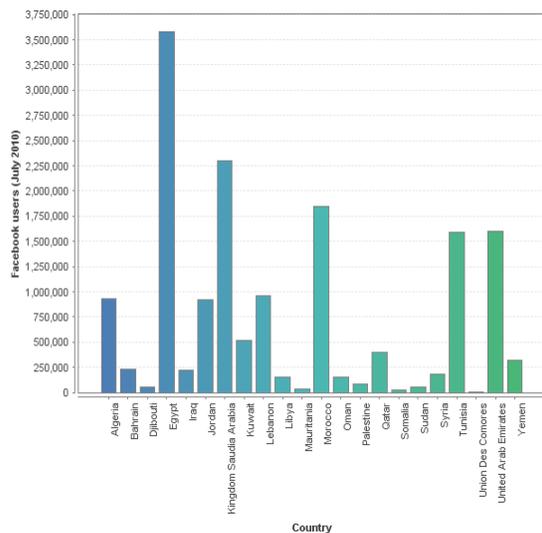


Fig. 5. Arab Nations are in Ascending Order.

The RapidMiner manipulates the excel sheet and filter them by using the Filter operator to choose 22 Arab Nations and sort it in ascending order.

2. **CLASSIFICATION (Prediction):** This operator can generate a decision tree model, which can be used for classification.

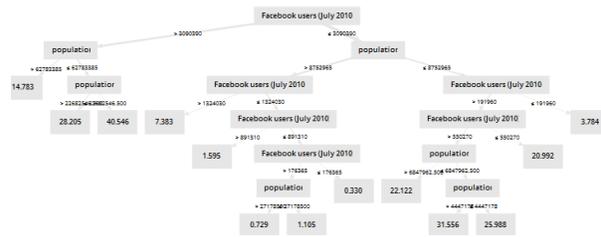


Fig. 6. Shows population by Facebook user in the countries.

The decision Tree shows how many population use Facebook around the world.

Figure [4] below shows the difference in a clear way. Which is almost the population around the world that use Facebook.

Important Factors for Prediction

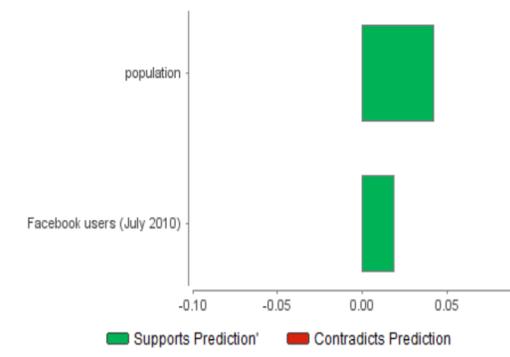


Fig. 7. population by Facebook users.

4 CONCLUSION

In conclusion, big data analytics are the drive of the improved decision making now in all business aspects. Enterprises are achieving better performance and understanding of business and customers' patterns. There are many tools used to help analyze the data in different patterns and shapes. Rapid Miner and Tableau are data visualizing tools that can extract what is relevant to the scope the business seek to interpret and get deeper into what knowledge it could provide. A methodology on such tools was practiced to demonstrate how data sheets can be viewed in multiple patterns. Each pattern can be extracted in a specific way and different specific data. These tools can provide coverage to all aspects to increase understanding of the subject matter and have a future vision. A data sheet on number of Arab users using Facebook was used in visual practice with

different operators that can aid in understanding. Therefore, such amount of data can be hard to read or understand without mining the relevant data and visualize it to have a closer and more interesting view to understand by different end users responsible in the decision making. It is recommended for enterprise dealing with huge amounts of data to follow up with last data mining and visualizing technologies as it is proved how it effective and efficient in all business operations. On the other hand, enterprises need to be careful on what big data they seek to collect since it is the critical ingredient. Moreover, the analysis tools must be up to what the business operates and what industry they work in since the rapidly changing ones provide constant supply of data that needs wise use of analysis and forecasts.

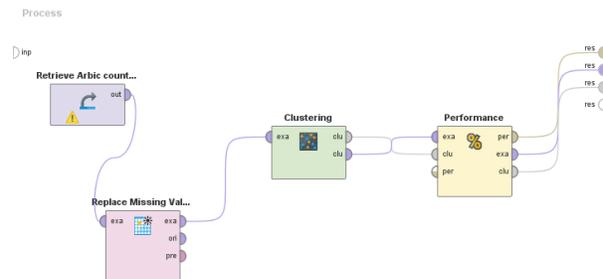
5 REFERENCES

- [1] BARC - Business Application Research Center. (2018). Big Data Use Cases 2015 - A BARC Research Study. [online] Available at: <http://barc-research.com/research/big-data-use-cases-2015/> [Accessed 10 Apr. 2018].
- [2] Bbc.com. (2018). BBC STORYWORKS |. [online] Available at: <http://www.bbc.com/storyworks/banking-on-innovation/bigdata> [Accessed 13 Apr. 2018].
- [3] Yuk, C. (2018). Big data in 2018: the experts have their say | Churchill Frank. [online] Churchill Frank. Available at: <https://www.churchillfrank.com/blog/future-of-big-data-2018/> [Accessed 13 Apr. 2018].

APPENDIX

Operators Design in RapidMiner:

o Clustering



o Blending



o Classification (prediction)

