SALES ANALYSIS AND PREDICTION DASHBOARD USING POWER BI

Monika Raje*1, Priyanshi Jain*2, Vikrant Chole*3

*1,2,3Department of Computer Science and Engineering, G. H Raisoni Institute of Engineering & Technology, Nagpur Maharashtra, India.

ABSTRACT

Business Intelligence can be defined as a solution that helps companies make wise decisions in business processes. It can be a tool to effectively organize and convert data into information, which is the basis of knowledge in decision-making. With the help of the right business technology solution, the company can play a significant role in monitoring the business performance on time, quickly responding to the external business environment. This project aims to create a sales dashboard that displays monthly sales performance by looking at data and visualizing them in power BI. We have upgraded the Power BI Marketing Dashboard with advertising data and predicted the expected sales through investment inputs of different marketing channels using multiple linear regression in Python and visualized that in the Power BI Interactive Dashboard. Based on the results of the evaluation, analysis, and prediction this sales dashboard meets the requirements of the sales department. The process of data delivery becomes smoother between visualized data and users than before. By viewing the sales dashboard, users can easily clarify sales performance in the current month.

Keywords: Linear Regression, Sales Prediction, Sales Analysis, Business Intelligence, Power BI, Sales Dashboard.

I.   INTRODUCTION

In the case of data, a large amount of data is generated in the daily business process. When faced with big, random data, companies have a hard time understanding and applying them. In addition, responding quickly to the external environment of the business requires making timely decisions in the right way. Therefore, business intelligence has evolved into a very important solution for mining data, converting data from information to information, which contributes to decision-making. By simplifying smart business software, some parts of the business can rely heavily on support from the IT department. For example, to generally monitor the performance of sales in the sales department, employees may make reports that are easier to view information than to ask their technical colleagues to provide them. It can save time and improve job efficiency to create more value for sales performance. In other words, companies should use the right technical solution to solve business problems. Business analysis is significant to any business to make sure that the business's needs are decisive and help embrace improvements and effective decision-making. This is expected to happen by bridging the gap between IT and business strategies confirming the core processes and delivering the required benefits to the businesses and their stakeholders. Business analysis enables data-driven decisions and communicating complex technical details more easily. Business strategies and tactics are the main defined topics that are advisable and drive upon business changes whenever required. Business analysis brings benefits with the power of data, making the businesses become analytically driven and helps to gain a new approach for making data-driven decisions. It brings deeper insights into business and is considered an imperative facet for nurturing business transformation and innovation into a customer-centric operation. Designing a sales dashboard using Power BI for a sales department to reflect sales performance will explore ways in which a dataset can empower business and build a report based on the sales and will provide interactive visualizations and Business Intelligence Capabilities With An Interface Simple Enough For End Users to understand their reports And dashboards.

II.   AIM & OBJECTIVE

The purpose of this project is to design a sales dashboard for the sales department using Power BI to demonstrate monthly sales performance which will provide data and user-related information. The content displayed should be fully visualized and appealing. In addition, as information are no longer relevant to decision-makers, the overall construction process should be transparent within the department. The final dashboard can be considered as a customized version. Users can still follow the design process or modify it if a new need arises. This project aims to seek ways to improve overall sales via

analyzing the data and generate sales performance dashboards for better decision making. This was done by identifying the highest & lowest selling products while drilling down into different dimensions to understand the hidden patterns, thus the objective is to build the most effective predictive model with optimum accuracy and evaluation measures to predict future sales and to draft valid analysis by evaluating the customer buying behavior to improve their future sales.

III.   RELATED WORK

1. 'Walmart's Sales Data Analysis '

In this study, an analysis of the data collected in the retail store and predicting future store-related strategies are performed. The result of various sequences of events such as weather, holidays, etc. It can change the status of different departments and therefore also study these results and assess its impact on sales.

2. 'Applying machine learning algorithms in sales prediction'

This is a thesis where several different machine learning processes are used to obtain better, more relevant results, which are further evaluated by predictive activities. Use four algorithms, a merging process, etc. Feature options have also been implemented using different strategies.

3. 'Sales Prediction System Using Machine Learning

In this paper, the aim is to obtain relevant results for predicting future sales or company requirements through strategies such as Clustering Models and sales forecasting measures. The strength of algorithmic methods is measured and used appropriately in further research.

4. 'Intelligent Sales Prediction Using Machine Learning Techniques

This study presents an evaluation of the decisions to be taken from the experimental data and the insights gained from the data look. Use data mining techniques. The Gradient Boost algorithm is shown to show high accuracy in predicting future transactions.

5. Retail sales prediction and item recommendations'

This paper describes a sales forecasting system and a product recommendation system, used to benefit a group of retail stores. Consumer value data is used to accurately design individual sales.

6. 'Utilization of artificial neural networks and GAs'

In the study, the usage of deep neural network techniques is to know about their sales strategy regarding electronic components ahead in time. Some optimization algorithms are also used to maximize the efficiency of the system: like the Genetic Algorithm.

7. 'Bayesian learning for sales rate prediction'

In this paper, it is shown that from the prediction of the single one's rate of transactions, many vendors would benefit from it, which means the information obtained could be beneficial for the construction of a set-up that would estimate a large number of outputs. The prediction uses a neural network approach. Here they need to practice Bayesian learning to realize insights.

IV.   TOOLS USED

4.1 Power BI Desktop

Power BI is an analytical application for analyzing data and sharing knowledge in business. Power BI includes cloud integration that provides data storage capabilities such as data optimization, data availability, and interactive dashboards. Microsoft releases additional support features called Embedded Power BI on the Azure cloud platform. Power BI also can combine different databases, files, and web services so that it can quickly make changes or fix data and problems automatically. Power BI also ensures the security of publishing reports generated within the company and automatically manages the most up-to-date information. Power BI can also integrate all data into a company, whether cloud or in-house because Power BI has a gateway that allows you to connect to SQL Server data, analytics service models, and many other data sources on the dashboard.

AutoML is a feature of Power BI Dataflows that enables users to create, train, and plead machine learning models. The best part is that it is done 100% inside of Power BI. AutoML is designed for business analysts in mind. AutoML empowers business analysts, or traditional BI engineers, to learn the basics of machine learning, while at the same time developing a model for making useful predictions. It currently supports the

implementation of Binary Prediction, General Classification, and Regression models. Additional model types, such as Forecasting, and the ability to export ML models made in Power BI to Jupyter Notebooks can be expected in the future. This can be of great help to analysts, or developers who want to make an example of Power BI performance, but then refer you to a data scientist to refine and set a standard model.

4.2 Spyder

It is a free and open scientific environment written in Python and designed for data scientists, IT engineers, and data analysts. It features a mixture of the advanced editing, analysis, debugging, and profiling functionality of a comprehensive development tool with interactive execution, deep inspection, and pleasant visualization capabilities of a scientific package. Spyder integrates with a variety of prominent packages within the scientific Python stack, including NumPy, SciPy, Matplotlib, pandas, IPython, SymPy, and Cython, also as other open-source software.

4.3 Microsoft Excel

It is a spreadsheet developed by Microsoft for Windows, Mac OS, Android, and iOS. It features graphing tools, pivot tables, and a macro programming language called Visual Basic Application. Microsoft Excel has the common features of all spreadsheets, employing a grid of cells arranged in numbered rows and letter-named columns to arrange data manipulations like arithmetic operations. In addition, it can display data as graphs, histograms, and charts, with a limited three-dimensional graphical display.
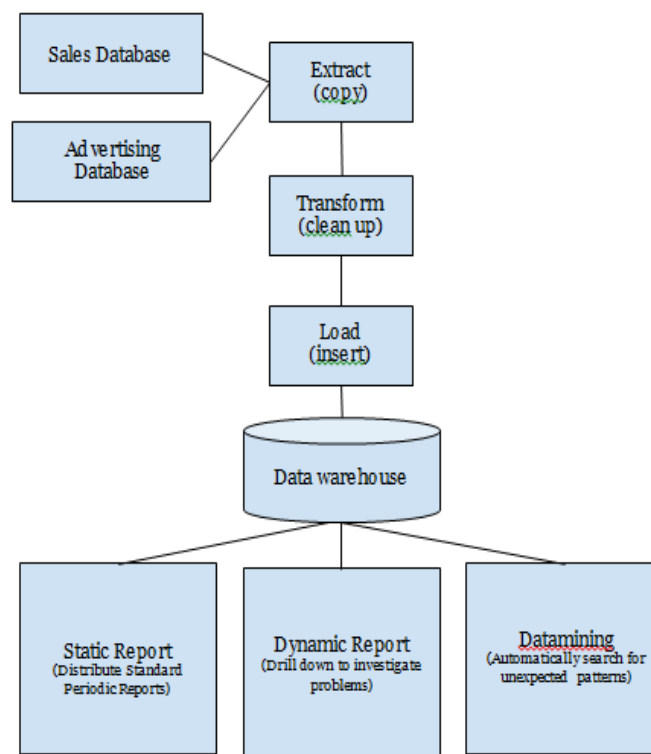
V.   METHODOLOGY

Figure 1. Process of creating a report on Power BI

The following defines the steps of this methodology as illustrated in the above block diagram.

1. First, collect all the databases and copy them for that first clean and transform the dataset to make it more usable.
2. For creating reports, use filters to drill down based on product and employees.
3. Determine sales by revenue, number of units sold, individual product revenue, and revenue generated under each supervisor.
4. Diagnose daily growth in revenue and create a drill-down table with relevant columns.
5. Incorporate an easy button for better UX.

6.   Publish and share the dashboard with the Power BI service.

5.1 Extract Data

The dataset for sales Analysis contains Employee data, Electronic Product data, and Sales data. The datasets for sales prediction were obtained from GitHub.This Advertising dataset contains 347 rows of sales records with advertising details from January 2015 to March 2019.

5.2 Transform Data

Sales Analysis database can be directly loaded but for prediction of sales, application of a machine learning algorithm was needed. So, data can't be used in its normal form as the data needs to be devised before applying it in machine learning models. Therefore, before loading the data into Power BI it is needed to be transformed first. This method is used to solve problems that are not yet known by the knowledge extractor. Proper formatted and cleaned data is necessary for preprocessing.

In the advertising dataset, there was a linear relationship between all three of those channels (T.v, Radio, and Youtube) ultimately this will be able to use in the equation i.e predicted sales is the sum of multiple of those channels and the coefficient.

$y = b0 + b1*x1 + ... + bn*xn$

where y is the predicted sales, x1,...,xn are the channels, and b is the coefficient.

To isolate those coefficients and to find the value of predicted sales use of linear regression was done using python script in Power BI.

5.3 Algorithm used

Linear Regression is the most common and widely used Machine Learning algorithm. It is used for establishing a linear relationship between the target or dependent variable and the response or independent variables. The linear regression model is based upon the following equation:

$yb0 + b1x1 + b2x2 + b3x3 + ... + bnxn$

where y is the target variable, b0 is the intercept, x1, x2, x3,....xn are independent variables, and b1, b2, b3,..., bn are their respective coefficients.

The main aim of this algorithm is to find the best fit line between the target variable and the independent variables of the data. It is achieved by finding the foremost optimal values for all b. With the best fit, it is meant that the predicted value should be very close to the actual values and have minimum error. Error is the distance between the data points to the fitted regression line and generally can be calculated by using the following equation:

$Error = y1 - y2$

where y1 is the actual value and y2 is the predicted value.

5.4 Load Data

For the sales analysis dashboard, an Electronic product sales database was loaded in Power BI which consisted of employee data, electronics product data, and sales data. And for the sales prediction dashboard, after transforming the advertising database the database was loaded in power BI which consisted of Date, T.V, Radio, and newspaper data.

5.5 Create Reports in Power BI

Using the donut chart, the revenue by the supervisor is displayed containing four supervisors along with revenue in percentage and number as shown in figure 2. A drill-down option is provided to get the employee-wise data. So, data can be easily drilled down to the next level. Revenue by date is displayed using the area chart which is containing revenue date-wise as shown in Figure 3.
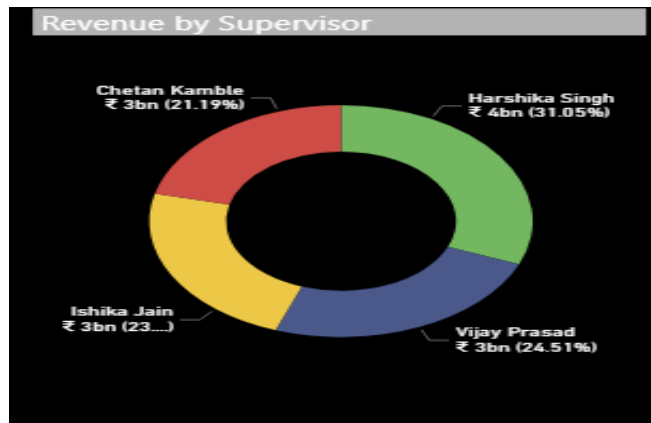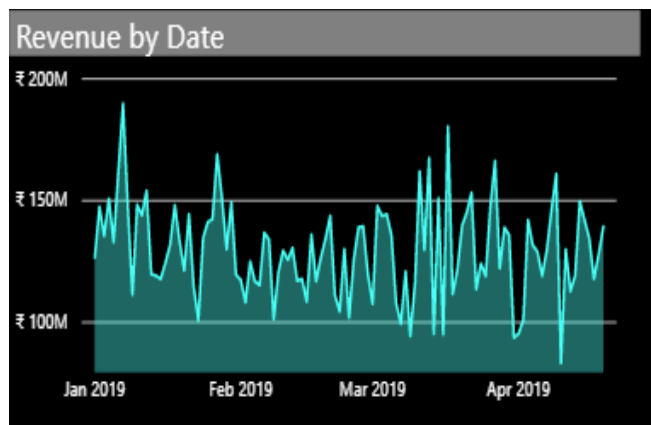
Figure 2. Revenue by supervisor

Figure 3. Revenue by Date

Revenue by Employee name chart is shown in figure 4. Employees who are booking the sale product. So, the largest area block is the employee who is providing the maximum revenue. Figure 5 chart is the product-wise revenue chart and laptop/ computer is giving the highest revenue and revenue is in million.
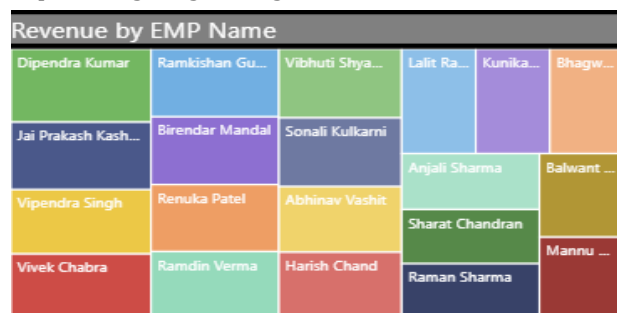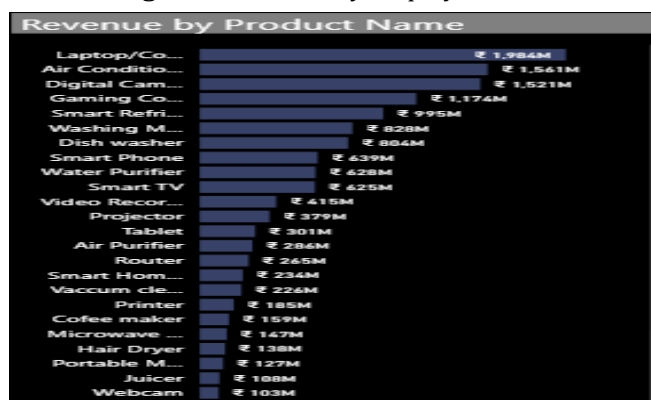

Figure 4. Revenue by Employee name

Figure 5. Revenue by product name

VI.   RESULT

6.1 Sales Analysis

In this Dashboard, the total revenue, the number of units sold, Revenue by date using area chart, Revenue by product name using a clustered bar chart, and also, we have used a treemap to display revenue by employee name and donut chart to display revenue by the supervisor is displayed. On the header, the use of cards is done to select attributes and a next button to navigate to the next dashboard.

Figure 6. Sales Analysis Dashboard

6.2 Sales Prediction

This sales prediction model is made using multiple linear regression in python and is visualized into an interactive dashboard using power bi and advertising dataset. In this dashboard, users have to enter the investment inputs of T. V, Radio, and YouTube advertising and it predicts the expected sales. In addition to that, one Q&A section is added where users can ask any question regarding the data.
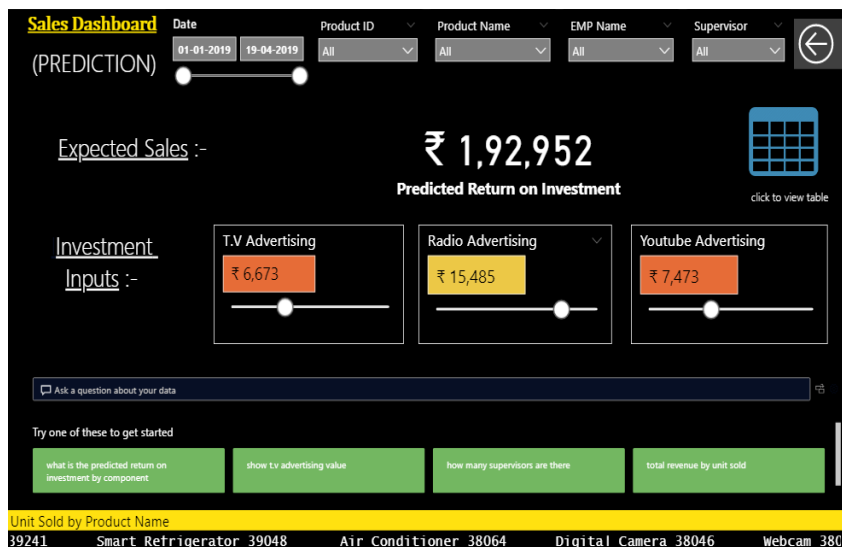

Figure 7. Sales Prediction Dashboard

VII.   CONCLUSION

Based on the discussion of results and the implementation of Power BI, the conclusions are:

1.   The construction of business intelligence dashboards can be used as a reference in decision-making processes.
2.   The data cleaning process is the most important stage of producing the right information.
3.   User involvement from process and design in result verification will greatly improve the quality of the

required dashboard.

4. According to this case study, the recommendations for continuous improvement are:

- The pre-built database needs to be upgraded to make it function effectively and efficiently to effectively support information needs in many companies owned by the company.

- Advanced dashboard upgrades will also be done for business components such as financing, hiring, manufacturing, and other businesses.

- It uses Power BI at a higher level (guessing and interpretation) by the machine learning should be done at a level of advanced development so that it can cover sophisticated modeling and data integration conditions and ultimately empower organizations that make data-based decisions on all aspects of a business.

5. Sales analytics via a consistent interpretation offers a wide range of actionable knowledge to businesses where it brings a competitive edge. This actionable knowledge needs more support from predictive analytics to formulate solid business strategies to win the current competitive business environment.

6. Predictive analytics allows businesses to foresee future events to effectively implement the formulated/designed action plans to enjoy competitive advantages and gain more profits.

7. The sales analytics revealed lots of useful insights on the sales performance, market basket analysis, and customer analysis which enlighten ways to understand the current situation.

8. The predictive models are evaluated and compared using the measures such as accuracy to get the best fit model. In the evaluation process, the multiple linear regression model has been identified as the best-fit model for this problem with different channels.

In conclusion, the performed predictive analytics results can support identifying the potential customers and predict the sales to a targeted mailing activity that promotes sales to a greater level.

VIII.   REFERENCES

[1]   Research Paper- Building A Sales Dashboard for A Sales Department by Using Power BI by Lan Huang Lan Huang

[2]   "Data Strategy: How to Profit from a world of Big Data, Analytics and The Internet of things" by Bernard Marr

[3]   "Big Data Demystified: How to use Big Data, Data Science and AI to Make Better Business Decisions and Gain Competitive Advantage" by David Stephenson PhD

[4]   "The Development of Performance Dashboard Visualization with Power Bi as Platform" by Surlisa Widjaja, Tuga Mauritius

[5]   "Descriptive and Predictive Analytics on Adventure Works Cycle: A Corporate Decision Making" by Yew Liong Lim, Raheem Mafas

[6]   Delen, D., Sharda, R. & Turban, E. 2014. Business Intelligence: A Managerial Perspective on Analytics. 3rd edition. England: Pearson Education Limited.

[7]   PT. Indomobil Sukses Internasional, "Annual Report 2017 PT Indomobil Sukses Internasional Tbk .," www.indomobil.com, vol. 2017, 2017.

[8]   L. Moss and S. Atre, Business Intelligence Roadmap. 2003.

[9]   D. Delen and H. Demirkan, "Data, information and analytics as services," Decis. Support Syst., vol. 55, no. 1, pp. 359–363, 2013.

[10]   D. Appelbaum, A. Kogan, M. Vasarhelyi, and Z. Yan, "Impact of business analytics and enterprise systems on managerial accounting," Int. J. Account. Inf. Syst., vol. 25, no. March, pp. 29–44, 2017.

[11]   Braun, V. & Clarke, V. 2006. Using Thematic Analysis in Psychology. Article from ResearchGate. Accessed 21 January 2019.

[12]   https://www.researchgate.net/publication/235356393_Using_Thematic_Analysis_in_Psychology

[13]   Brinkmann, S. 2013. Qualitative Interviewing. Book from ProQuest. Accessed 20 January 2019. https://ebookcentral.proquest.com/lib/laurea/detail.action?docID=1274289

[14]   Castleberry, A. & Nolen, A. 2018. Thematic analysis of qualitative research data: Is it as easy as it

sounds? Article from ScienceDirect. Accessed 23 January 2019.

[15]   https://www.sciencedirect.com/science/article/pii/S1877129717300606

[16]   Clark, D. 2017. Beginning Power BI: A Practical Guide to Self-Service Data Analytics with Excel 2016 and Power BI Desktop. Book from ProQuest. Accessed 2 January 2019. https://ebookcentral.proquest.com/lib/laurea/detail.action?docID=4815394&query=clark

[17]   Mckay, R. 2018. Data Visualization Example: A Look into Modern Visual Innovation. Posted 3 October. Accessed 26 December 2018. https://www.tapclicks.com/innovative-data-visualization-examples/

[18]   Microsoft. No date. Building Blocks of Power BI. Accessed 20 December 2018. https://docs.microsoft.com/en-us/learn/modules/get-started-with-power-bi/3-buildingblocks-of-power-bi

[19]   Microsoft. 2016. Overview of Service Manager OLAP cubes for advanced analytics. Accessed 3 January 2019. https://docs.microsoft.com/en-us/system-center/scsm/olap-cubes-overview?view=sc-sm1807

[20]   Microsoft. 2018. Intro to the dashboard for power BI designers. Accessed 5 January 2018. https://docs.microsoft.com/en-us/power-bi/service-dashboards

[21]   M. Mitri, "Teaching Tip Active Learning via a Sample Database: The Case of Microsoft's Adventure Works," Journal of Information Systems Education, vol. 26(3), 2015.

[22]   "A Guide to the Business Analysis Body of Knowledge: Draft Material for Review and Feedback," International Institute of Business Analysis, 2006

[23]   Knaflic C., "Storytelling With Data: A Data Visualization Guide for Business Professionals", Wiley, 2015.

[24]   D. M. Velumani, "A Study on Consumer Buying Behaviour towards Nokia Mobile in Erode District," IOSR Journal of Business and Management (IOSR-JBM), Vol. 16, No. 12, pp. 52-54, 2014.

[25]   D. L. I. A. M. Asoc. Prof. Elīna Gaile-Sarkane, "Consumer behavior-changing: methods of evaluation," Trends of Economics and Management, Vol. 3, No.4, 2013

[26]   Shafique U., Qaiser H., "A Comparative Study of Data Mining Process Models (KDD, CRISP-DM, and SEMMA)," International Journal of Innovation and Scientific Research, Vol. 12, No. 1, pp. 217-222, 2014.

[27]   J. H. Rüdiger Wirth, "CRISP-DM: Towards a Standard Process Model for Data Mining", IFIP International Conference on Digital Forensics, 2017

[28]   Vente J, Waal A.D, Willers C., "Specializing CRISP-DM for Evidence Mining," SpringerLink.

[29]   M. Faber, "Visualizing Student Research though a Multi-Conditional Branching - Choose Your Adventure-Style Tutorial," Semantic Scholar, 2015. Accessed on Jun 19, 2019. [Online]. Available: https://www.semanticscholar.org//paper/Visualizing-Student-Research-through-a-Branching-%22-Faber/10684e7187e06ccadb11e185ab 1d58c8cf90a35e