



Analysis and Design of Visualization of Educational Institution Database using Power BI Tool

By Mandava Geetha Bhargava, K. Tara Phani Surya Kiran
& Duvvada Rajeswara Rao

Koneru Lakshmaiah Educational Foundation

Abstract- Visualization of data set is a process of making understand the significance of data through visual context and part of data analytics where it's executed after the data correction. Nowadays visualization is more useful in business intelligence and Analytics in every field, There are different techniques for visualizing the datasets, it may be in dynamic or interactive nature, and datasets can be visualized in different types of visuals insights, This paper deals with the interactive visualization of educational institution database using Microsoft Power BI Tool with different modules and this paper focuses on process model, operations of Microsoft Power BI, types of data sources available in Tool and its different related types of visual insights or context.

Keywords: *analytics, visualization, business intelligence, insights.*

GJCST-C Classification: *1.6.8*



Strictly as per the compliance and regulations of:



Analysis and Design of Visualization of Educational Institution Database using Power BI Tool

Mandava Geetha Bhargava ^α, K. Tara Phani Surya Kiran ^σ & Duvvada Rajeswara Rao ^ρ

Abstract- Visualization of data set is a process of making understand the significance of data through visual context and part of data analytics where it's executed after the data correction. Nowadays visualization is more useful in business intelligence and Analytics in every field, There are different techniques for visualizing the datasets, it may be in dynamic or interactive nature, and datasets can be visualized in different types of visuals insights, This paper deals with the interactive visualization of educational institution database using Microsoft Power BI Tool with different modules and this paper focuses on process model, operations of Microsoft Power BI, types of data sources available in Tool and its different related types of visual insights or context.

Keywords: analytics, visualization, business intelligence, insights.

I. INTRODUCTION

Data Visualization is a process of making understand the significance of data through visual context, and it is a part of analytics, there are several techniques to visualize the data such as Interactive and Dynamic in nature and coming to visual context, there are a number of things such as plots, graphs, slicers, stacked column charts, Histogram, Bar Charts, tables, matrix and other forms of visual contexts; In this paper we focussed on interactive data visualization through Microsoft Power BI tool, Microsoft Power BI is a suite of business intelligence and analytics tool for analyze data and share insights and gets answers quickly with the help of interactive data visualization using dashboard available on every device such as Applications, Desktops, Mobiles...etc. With the help of visuals and filters, the user or person gets convenient and easier to understand the data and it has an architecture of five main components as discussed below and follows Power BI Services, Power BI Gateways, Power BI Desktop, Power BI Apps and Power BI Connectors as shown in figure 01. Power BI Services is the main component in the architecture where published reports are made into dashboards to share in the organization, Power BI Gateways is another main component in the architecture where it handles to get data operation from different data sources by means of

connectors and protocols, Power BI Desktop is the component in Architecture where the data is analyzed and transformed through some procedure using tools and made to report on the web by means of several visuals, tools and publish feature.

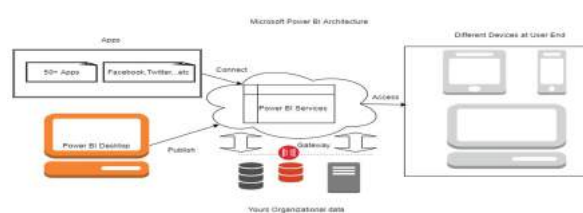


Fig. 1: Architecture of Microsoft Power BI

Power BI Apps are the crucial components at user side where viewing and accessing of dashboard through some applications such as Power Apps, Mobile Power BI...etc., Power BI connectors leads crucial role in getting data from the database and other sources using connector application such as database engines, Azure Consumption Insight Connector...etc. The general operations of Microsoft Power BI are as follows: 1) Get the Data from Required Data Source 2) Analyse the data by means of connectors and gateways of organization 3) Build the Report by means of Different Visuals and Filters 4) Publish the Report into web through Power BI Desktop 5) Edit the report if any changes are needed and make shareable by means of publishing on to web option for creating embed URL 6) Access the report data from different applications of Microsoft such as Power Apps, Mobile Power BI 7) Refresh the data using different gateways of Organization for updating the dashboard. In this paper, we discussed on process model and Visuals of the Power BI tool and interactive data visualization technique for analysis and design of educational institution data visualization using Microsoft Power BI tool.

II. PROCESS MODEL OF POWER BI

The Process model consists of Seven steps as discussed below and shown in figure 03, GET DATA, FETCH, PROCESS, ANALYZE, VISUALIZE, EDITING, WEB.

Author ^{α σ}: M.Tech Scholar, Dept of CSE, Koneru Lakshmaiah Educational Foundation, Guntur, A.P.

e-mail: geetha.bhargava@hotmail.com

Author ^ρ: Professor, Dept of CSE, Koneru Lakshmaiah Educational Foundation, Guntur, A.P.

- a) *Get Data*: In this step, the end user is going to get data from different sources such as files, databases, Microsoft Azure, Online services and other as shown in details of data sources.
- b) *Fetch*: In this step, the data which is selected through browsing data location or using queries by selecting types of data sources and connectors.
- c) *Process*: In this step, the data is truncated and edited using several operations while loading the data into the Power BI.
- d) *Analyze*: In this step, the data is analyzed using filters, conditional Queries and other operations such as adding column erroneous data.
- e) *Visualize*: In this step, the analyzed and processed data is visualized through means of visuals such as Microsoft Power BI Visuals and Custom Visuals like plots, graphs, slicer, KPI...Etc. After this process, the report is published to power BI services.
- f) *Editing*: In this step, the published report on to web is finalized after rectifying errors for any changes such as if any filters or visuals need to solve and after completion of the editing process, the report is made to publish on the web.
- g) *Web*: As said above step, it is a state where the report is converted into the dashboard, and it can be share-able via URLs, websites. Etc.



Fig. 2: Process Model of the Power BI Desktop

III. TYPES OF VISUALS IN POWER BI

The Microsoft Power BI visuals provide high-quality visualizations of data that you can use to extend Power BI. Microsoft Power BI contains over 20 different types of visual contexts, the framework to run them, and the testing infrastructure that enables you to build high-quality visuals. The framework provides all the interfaces you need to integrate fully with Power BI's selection, filtering, and other UI experiences. The code is written in Typescript, so it's easier to build and debug. Everything compiles down to JavaScript and runs in modern web browsers. The visuals are made using D3, but you can use your favorite technology like Canvas, SVG, or WebGL. This enables you everything you need to build custom visualizations for Power BI.

There are two types of visuals in power BI. 1) Microsoft Visuals, 2) Custom Visuals.

1) Microsoft Visuals

These are the official visuals from Microsoft, and some were built-in visuals in power BI installation package.

They are secured in nature and following are the official Visuals of Microsoft such as Stacked bar chart, line chart, Waterfall chart, Area Chart, Clustered Bar chart, Slicer, table, and Matrix...etc.

2) Custom Visuals

These visuals contexts are developed by the third party or the end users such as developers and these are shared among the users through the portal like office store and git Hub. They are unsecured in nature and following are the Custom visuals provided by third party and Microsoft officials such as dot plot, route map, flow map, journey chart, and scroller...etc.

IV. DETAILS OF DATA SOURCES

Table 1: Details of Data sources

| Data Format | Data Sources |
|-----------------|--|
| Files | Excel, Text/CSV, XML, JSON, Folder, Share Point Folder |
| Database | SQL SERVER, ACCESS, ORACLE, IBM DB2, IBM INFORMIX, IBM NETEZZA, MySQL, Postgre, Sysbase, Teradata, SAP, Google Bigquery, etc. |
| Azure | Azure SQL database, Azure SQL Data Warehouse, Azure Blob Storage, Azure Table Storage, Azure HD Insight(HDFS), etc. |
| Online Services | Power Bi service, Share point online list, Dynamics 365, Microsoft Exchange Online, Salesforce, Google Analytics, Facebook, Github, etc. |
| Other | Web, SharePoint list, OData Feed, Active Directory, Microsoft Exchange, Hadoop File(HDFS), R Script, ODBC, OLE DB, etc. |

V. LITERATURE REVIEWS

Marija Blagojevic et al. [1] studied on web-based intelligent report of e-learning system by using the technique of data mining and it deals about PDCA method such as (Plan, Do Check, Act) for improving the web-based intelligent reports of eLearning system by means of data mining techniques and concluded that their proposed system has an improvement since it predicts behavior patterns thus leading to the increase in count of participants and in there study ,it proved that their proposed system has improvements in terms of report system in the field of LMS (learning management system) or e-learning . Moreover, the development and implementation of new modules.

Daniel J.Power [2]studied the data-driven decision support system and it deals with data-driven decision support system and its advantages at Business Intelligence and concluded that mainframe-based decision support systems would need to be updated or replaced by web-based or web-enabled systems the accessibility reach for data-driven decision support systems are open source software's, new hardware's, web technologies, etc.

Zhijun Ren [3] studied the delivering of a comprehensive Business Intelligence solution using Microsoft Business Intelligence stack and it deals about features and advantages of business intelligence stack of Microsoft and concluded that by integrating several technologies such as database, connectors, SharePoint

servers, and business intelligence tools will lead to Faster delivery of comprehensive business intelligence solution within an enterprise

Guangzhi Zheng et al. [4] studied on business intelligence to healthcare informatics Curriculum and their paper deals with the preliminary analysis of integration of Business Intelligence with Healthcare Information Technology and concluded that Business Intelligence had been a neglected part in many healthcare information technology programs yet both the industry and academia have realized the importance of Business Intelligence

Michelle Hoda Wilkerson et al. [5]done work on youth reasoning with interactive data visualization and it deals with the youth understanding of data by interactive data visualization, they concluded that supporting learners in the coordination of any resources they choose to leverage is more likely helpful than supporting a particular approach or sequence of resource use

Yuri Vanessa Nieto et al. [6]done a work on academic decision-making model for higher education institutions with the help of learning analytics .it deals with modeling and construction of software architecture for creating and categorizing indicators and they concluded that proposed software architecture has benefit of providing integration of learning analytics indicators and supports decision making in universities.

VI. METHODOLOGY

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, Online Services and other, the detailed description of data sources as shown in details of data sources 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 and 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, As the option is seen at the left side of the power BI tool, if we required any conditional columns or to add new data in datasets by means of edit queries function we can proceed it and next step is to make report by means of clicking on report option and white empty sheet appears on the screen where we create our own report. The analyzed and corrected data is visualized by means of different Visuals such as Custom Visuals and Power BI Visuals such as stacked bar chart, stacked column chart and by means of Publish option the report is published on the Power BI Services. We have to log in to power bi services, if required we can do any editing operations otherwise the report is made into the dashboard by means of publish on the web

functionality and embed link is generated, the embedded link can be used in any websites for viewing dashboards by all end users as well as embedded link can be shared to any individuals. In this methodology, there are two types of modules as follows Faculty Module and Student Module. The Faculty Modules contains Student Internal Marks Progress and Placement Progression submodules and Student Modules contains only one module for viewing their data and there are explained as follows.

Pseudo Code Algorithms:

Pseudo Code Algorithm 1:

Start

Get 'P_i'

View 'P_i'

Initialise P_{at}, P_{t1}, P_{t2}, P_{ha}, P_q, P_{ch}, P_{ca}

Processing of P_i by means of string operations

Edit the dataset 'P_i' for adding new conditional columns P_{ap}, P_{am}, P_t, P_{t1pe}, P_{t2pe}, P_{qp}, P_{hap} using edit query function

Create new column P_{ap}

$P_{ap} = P_{ca} / P_{ch} * 100;$

Create new column P_{am}

P_{am} = If P_{ap} == 100, 5

Else If P_{ap} >= 95,

Else If P_{ap} >= 90,

Else If P_{ap} >= 85,

Else If P_{ap} >= 80,

Else If P_{ap} >= 75,

Else 0

Create new column P_{t1}

P_{t1} = If P_{t1} < P_{t2}, ((P_{t2} * 0.75) + (P_{t1} * 0.25))

Else ((P_{t2} * 0.25) + (P_{t1} * 0.75))

Create new column P_t

P_t = P_{at} + P_{ha} + P_q + P_{am} + P_{t1}

Create new column P_{fs}

P_{fs} = If P_t >= 16, Promotion

Else If P_t < 16, Detention

Else If P_{ap} < 65, Detention

Else If P_{ap} > 75 && P_{ap} <= 100, Promotion

Else If P_{ap} >= 65 && P_t >= 16 && P_{ap} < 75,

Condonation

Else Error

Create new column P_{ape}

$P_{ape} = P_{at} / 5 * 100$

Create new Column P_{t1pe}

$P_{t1pe} = P_{t1} / 20 * 100$

Create new Column P_{t2pe}

$P_{t2pe} = P_{t2} / 20 * 100$

Create new Column P_{qpe}

$P_{qpe} = P_{q} / 5 * 100$

Create new Column P_{hape}

$P_{hape} = P_{ha} / 5 * 100$

Visualize the above-considered variables using different Visuals

Publish the Report to power bi services

notations

P_i = Dataset of Students Internal marks from different data sources

P_{at} = Variable name of Assignment Test marks

P_{t1} = Variable name of Test 1 marks

P_{t2} = Variable name of Test 2 marks

P_{ha} = Variable name of Home Assignment marks

P_q = Variable name of Quiz exam

P_{ch} = Variable name of classes held

P_{ca} = Variable name of Classes Attended

P_{ape} = Variable name of assignment test performance

P_{am} = Variable name of Attendance marks

P_{ap} = Variable name of attendance percentage

P_t = Variable name of the total percentage

P_{t1} = Variable name of test total marks

P_{fs} = Variable name of Final Status

X_{t1pe} = Variable name of Test 1 Performance

P_{t2pe} = Variable name of Test 2 Performance

P_{qp} = Variable name of quiz performance

P_{hap} = Variable name of Home Assignment Performance

Pseudo Code Algorithm 2:

Start

Get 'P_{ii}'

View 'P_{ii}'

Initialise P_{id}, P_{na}, P_{up}, P_{qf}, P_{br}, P_{st}, P_{cn}, P_{lp}

Processing of P_{ii} by means of string operations

Visualize the above-considered variables using different Visuals

Publish the Report to power bi services

notations

P_{id} = Variable name of id number

P_{na} = Variable name of the name
 P_{up} = Variable name of UG/PG
 P_{Qr} = Variable name of Qualification
 P_{br} = Variable name of Branch
 P_{st} = Variable name of Status
 P_{cn} = Variable name of Company name
 P_{lp} = Variable name of LPA

Pseudo Code Algorithm 3:

Start

Get ' P_{iii} '

View ' P_{iii} '

Initialise P_{ap} , P_{am} , P_t , P_{tm} , P_{fs} , P_{ape} , P_{t1pe} , P_{t2pe} , P_{qp} , P_{hap} , P_{fs} , of similar variables to Algorithm 1

Processing of x by means of string operations

Visualize the above-considered variables using different Visuals

Publish the Report to power bi services

As shown in the Pseudo code Algorithms 1,2, and 3 for different modules, Firstly the tool Programme is started and required data such as P_i , P_{ii} , P_{iii} in Algorithms 1,2 and 3 are captured into the tool from different data sources via getting data function and dataset is viewed if any corrections to be done for selected data and then data is initialized with multiple columns such as P_{at} , P_{t1} , P_{t2} , P_{na} , P_q , P_{ch} , P_{ca} , P_{Api} , P_{am} , P_{id} , P_{na} , P_{up} , P_{Qr} , P_{br} , P_{st} , P_{cn} , P_{lp} as shown in Pseudo-Code 1,2,and 3 respectively where different datasets contain a different number of columns. If required Uploaded dataset is processed by means of string operations, otherwise by means of queries conditional columns are created in datasets using edit queries as shown in pseudo code 1 and 2 with mathematical and conditional operations and then the dataset is Visualized through Microsoft Power Bi Visuals, by visualization of datasets report process is completed and then it is published to power BI Services by means of Publish function in the tool. After Publishing the report into the services of Power BI, the report is made into the dashboard by means of clicking publish on to web function in the file tab and then generated link can be shareable to anyone, it can be share to individuals of organization by means of Uniform Resource Locator (URL) directly or indirectly by means of SharePoint, Website,..Etc., as shown in the Results and discussion section.

VII. RESULTS AND DISCUSSION

After publishing the dashboard to the power BI services it appears as shown in below figure and we can have some editing operations if any filters or visuals are not properly accessible then they can be rectified here before publishing on to the web. Not only above discussed modules, we can have dashboards of different modules of every educational institution, it may be related to staffs, Infrastructure and other Amenities of

institutions...Etc. following are the results of dashboards of discussed modules of educational Institutions.

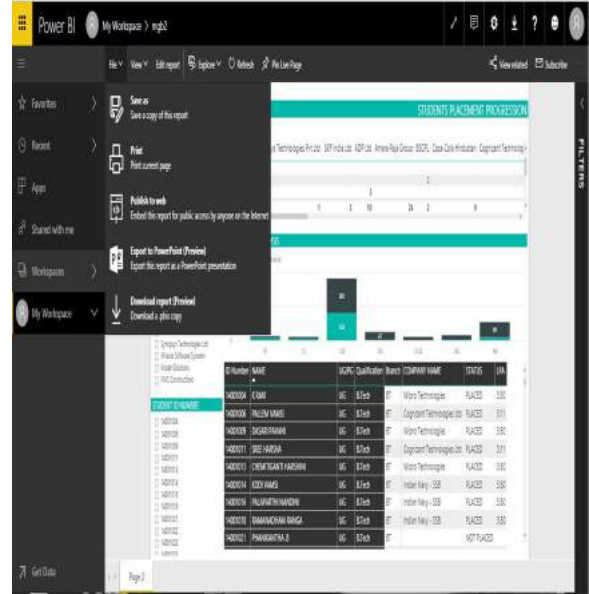


Fig. 3: View of publishing to web option A report in the Power BI Services

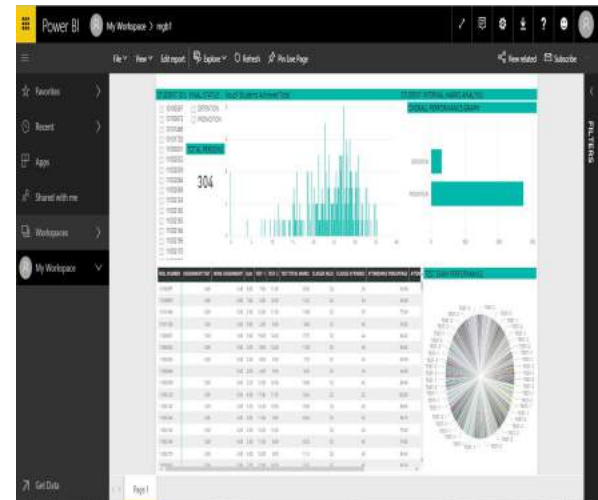


Fig. 4: View of Student Internal Marks Analysis

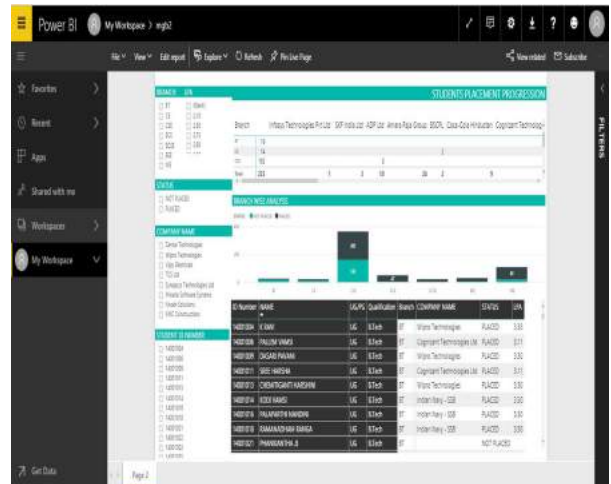


Fig. 5: View of Student Placement Progression

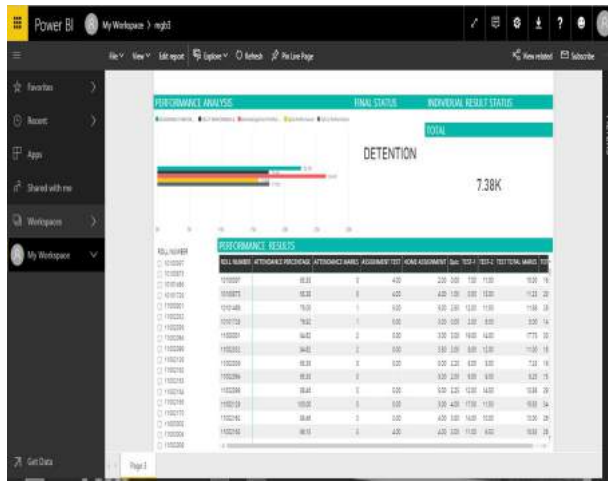


Fig. 6: View of Student Modules for viewing in the Power BI Services Purpose

After editing operations, the reports in the Power BI Services are shareable and kept into the web by means of publishing on to web option.

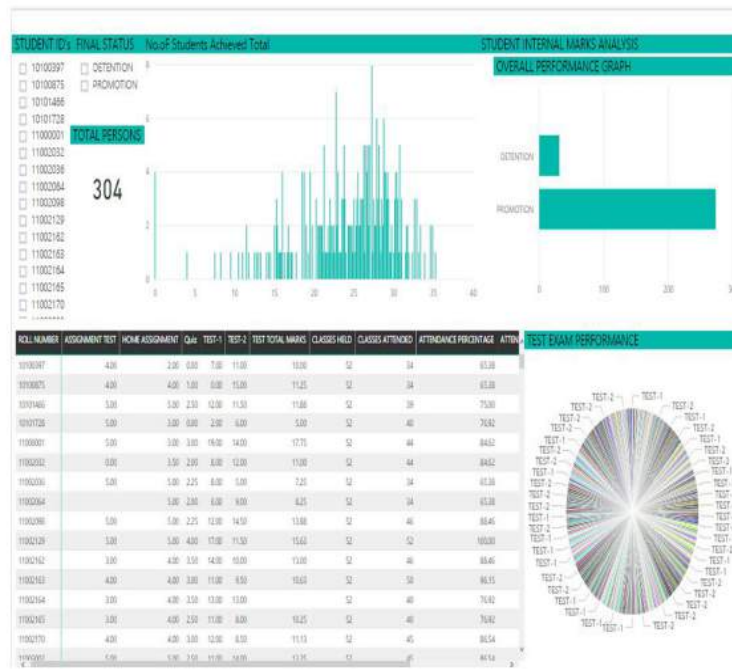


Fig. 7: View of Student Internal Analysis Dashboard in the web

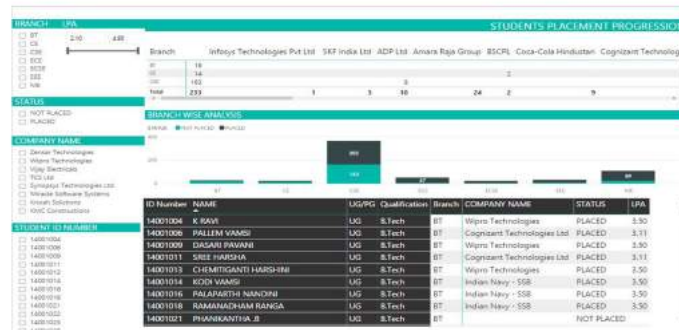


Fig. 8: View of Student Placement Progression Dashboard in the web

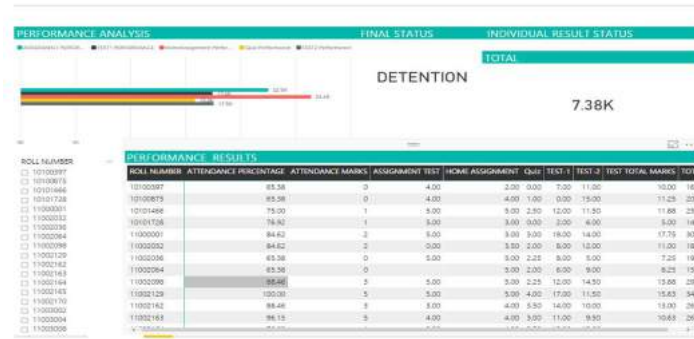


Fig. 9: View of Student view module Dashboard in the web

VIII. CONCLUSION AND FURTHER WORK

Visualization of data is a very important thing in every organization and Business development, by means of visualization the end user or user can read the data easier and convenient to understand the data. In this paper we have visualized the educational institution database into three modules such as Placement Progression, Students Internal Marks and module of viewing purpose from student side by means of analyze and rectification of captured data from different data sources, The dashboards which are created using Power BI tool can be viewed by means of Mobile Applications such as Power Apps and Mobile Power BI ,Web such as by means of website or shareable link As a further work we can make other modules on Infrastructure related amenities such as quality of infrastructure provided, Staff-related issues such as feedbacks, work status, and other issues .

REFERENCES RÉFÉRENCES REFERENCIAS

- Marija Blagojevic, Zivadin Micic, A web-based Intelligent report e-learning system using data mining technique, Computers and Electrical Engineering, Elsevier 2013; 39: 465-474.
- Daniel J. Power, Understanding Data-Driven Decision Support Systems, Information Systems management 2008; 25: 2, 149-154
- Zhijun Ren, Delivering a Comprehensive BI solution with Microsoft Business Intelligence Stack, International Conference on Challenges in Environmental Science and Computer Engineering 2010; 183, 278-281.
- Guangzhi Zheng, Chi Zhang, Lei Li, Bringing Business Intelligence to HealthCare Informatics Curriculum: A Preliminary Investigation, SIGCSE '14 Proceedings of the 45th ACM technical symposium on computer science education 2014, 205-210.
- Michelle Hoda Wilkerson, Vasiliki Laina, Youth Reasoning with Interactive Data Visualizations: A preliminary study, IDC'17 Proceedings of the 2017 conference on Interaction Design and Children 2017, 411-416.
- Yuri Vanesa Nieto, Vicente Garcia Diaz, Carlos Enrique Montenegro, Academic Decision Making model for higher education institutions using learning analytics, 4th International Symposium on computational and Business Intelligence 2016, 27-32.
- W. Villegas-Ch, S. Lujan-Mora, Analysis of data mining techniques applied to LMS for personalized education, World Engineering Education Conference 2017.
- Jayakody, J.R.K.C. and W.H.P. Allagalla, Identification of varying standard of student based on Moodle Pattern Identification Business Intelligence Tool, Proceedings of the International Symposium on ICT for sustainable development, 20125-25-29.
- Syed Mohd Ali, Noopur Gupta, Gopala Krishna Nayak, Rakesh Kumar Lenka, Big data Visualization: Tools and Challenges, Contemporary Computing and Informatics (IC31),2016,656-660.
- Jun Zhu, Eric Zhuang, Jian Fu, John Baranowski, Andrew Ford, James Shen, A framework based approach to utility big data analytics, IEEE transactions on power systems, Volume 31, Issue 3, 1-7.
- Microsoft Power BI Organization, Microsoft Power BI Website <https://powerbi.microsoft.com/en-us/what-is-power-bi/>. Accessed September 2017.
- Jianu, R., Rusu, A., Hu, Y., & Taggart, D. 2014. How to display group information on node-link diagrams: an evaluation. IEEE Transactions on Visualization & Computer Graphics, 20(11), 1530-1541.
- Walshaw, C. 2000. A multilevel algorithm for force-directed graph drawing. In Proceedings of Graph Drawing. Springer Berlin Heidelberg, 171-182.
- Nguyen, P. H., Xu, K., Walker, R., and Wong, B.W. 2015. Time Sets: Timeline visualization with set relations. Information Visualization. (Oct 2015) DOI=10.1177/1473871615605347.

15. Kucher, K., Schamp-Bjerede, T., Kerren, A., Paradis, C., & Sahlgren, M. (2016). Visual analysis of online social media to open up the investigation of stance phenomena. *Information Visualization*, 15(2), 93-116.

