Business Intelligence

Name of the student:

Name of the professor:

Name of the course:

Name of the university:

Date:

Business Intelligence

**Introduction**

 Business Intelligence (BI) is a set of tools, architectures, and technologies that converts data into insightful information that could be used to take profitable business actions (Isik et al., 2013). BI is helpful for organizations of all the sizes that have the resources required to plan and execute a BI project.

The current paper gives an outline of a BI project for a major American organization, viz. UnitedHealth Group to facilitate good decision making. It details the data sources, statistical software, and visualization tools that will be used to complete analysis and prepare the visualizations.

**Organization Selected**

 UnitedHealth Group is a leading for-profit healthcare company based in the USA. It provides healthcare and insurance products targeted at multiple customer segments in the USA. The company has four main divisions catering to different sections of the population. Its ‘UnitedHealthcare Medicare and Retirement’ division focuses on providing services to individuals who are aged 65 and above.

The reason for selecting this organization is to present a BI framework regarding the role that data and analytics can play in improving the range and relevance of services offered by the organization.

**Data to Be Used For Creating the BI Solution**

 Data is crucial for successfully conducting any business intelligence project. The data collected will be used could be gathered from either primary or secondary sources. Primary data collected for the research process include interviews, data collection through interviews, and personal observation of the researchers (Wilcox et al., 2012).

Secondary data sources refer to the data that is already collected by other researchers for other purposes. In some cases, data will be collected by third parties such as monitoring agencies and government departments without any specific objective. It is intended to be used by other researchers as and when the need arises.

 The data that will be used to develop the BI framework will be selected from the publicly available sources. Some of the data sets that will be useful in the BI framework that will be developed and used for developing new solutions for the customers of UnitedHealthcare Medicare and Retirement are the rates of hospitalizations, rate of preventable hospitalizations, data on household drug abuse among the older adults, etc. Publicly available data from the U.S. Department of Health & Human Services will be used for the project (<https://healthdata.gov/>).

**BI Tools Selected**

 The BI framework should include the tools that will be used to analyze the data and get the crucial insights from it. Below are the analytics tools that will be used for conducting the data.

**Data Storage**

 The data that will be collected and used for the current project will be very large. It will not be possible to store the data in the form of excel sheets or simple text documents (in the CSV format). Database software that can handle large data sets is required to handle the data that will be downloaded and stored periodically. Hence, PostgreSQL will be used to store the downloaded data.

**R**

R is a free statistical language/package used for conducting statistical analysis. R is free to use and has the highest number of libraries to conduct all kinds of analysis such as regression analysis to understand the future trends and conjoint analysis to understand the general patterns in the data. It will be used for the project to conduct predictive analytics and calculating descriptive statistics.

**Visualization Tools**

 Visualization tools help in getting a better understanding of the insights gained from analysis. It will also help senior executives and persons without a complete understanding of the statistical techniques used to get critical insights into the findings. The visualization tool selected for the proposed BI framework is Tableau. This visualization software helps in preparing colorful reports and dashboards for conducting the analysis.

 Raw data and the results from the data analysis completed will be used or preparing visualizations in Tableau. Tableau will also be programmed to automatically connect with other data analysis tools selected for the project so that it can get automatic updates from the data analysis already completed.

**Outline of the Project**

 In the BI project that is proposed for the UnitedHealth Group, both predictive and descriptive analytics will be conducted to get the required insights.

**Predictive Analytics**

 Predictive analytics involves the application of analytic techniques, machine learning tools, and data modeling to guess future trends, patterns, and growth rates (Loshin, 2013). For the proposed BI project, it is proposed to use regression analysis to predict future trends in the healthcare market. Regression analysis involves predicting a dependent variable using a set of independent variables (Peng et al., 2010).

 One of the proposed regression models for the proposed BI project is given below:

y = A + Bx1 + Cx2 + Dx3 + ϵ.

**Where:**

y (Dependent Variable) – Number of insurance policies needed for the elderly in the near future

x1, x2, x3 – Independent (explanatory) variables

x1= Life expectancy of the elderly in the USA

x2= Rate of hospitalization among the elderly people in the USA

x3- Drug abuse among the elderly in American households

A – Intercept

B, C, D – Slopes

ϵ – Residual value in the regression equation (error).

 Estimating the dependent variable from the above equation will give an insight into the number and type of policies that should be developed by the UnitedHealthcare Medicare and Retirement division of the company in the future to cater to the needs of the market.

New independent variables will be added to the regression equation above to make the results more insightful. A few of the independent variables will also be removed from the equation while conducting the analysis in case any multicollinearity is found between the independent variables.

**Descriptive Analytics**

 Descriptive analytics involves providing a summary view of the past facts and figures in an understandable format for the readers (Bertrand and Goupil, 2000). Descriptive statistics not only help in summarizing the data in a meaningful form but also in preparing the data for further use in the future.

 Some of the statistical tools that will be used for descriptive analytics are: mean (averages), standard deviation, range, etc. Data downloaded from the public sources will be processed using both R and Tableau to calculate and provide a summary view of the insights gained in the form of graphs and visuals.

References

Bertrand, P. and Goupil, F. (2000). Descriptive statistics for symbolic data. *Analysis of Symbolic Data,* 106-124.

Isik, O., Jones, M.C., Sidorova, A. (2013). Business intelligence success: The roles of BI capabilities and decision environments. *Information & Management,* 50(1), 13-23.

Loshin, D. (2013). *Business intelligence: The savvy manager’s guide.* Waltham, MA: Elsevier.

Peng, C.J., Lee, K.L., and Ingersoll, G.M. (2010). An Introduction to Logistic Regression Analysis and Reporting. *The Journal of Educational Research,* 96(1), 3-14.

Wilcox, A.B., Gallagher, K.D., Boden-Albala, B., and Bakken, S.R. (2012). Research data collection methods: From paper to tablet computers. *Medical Care,* 50(7), 68-73.