**CSSELEC3/CS0009**

FUNDAMENTAL OF ANALYTICS

LAB SUMMATIVE

3

THE VALUE OF DATA: TRANSACTIONAL INFORMATION AND EFFECTIVE DATA VISUALIZATION

|  |  |
| --- | --- |
| Name of Student | Name of Professor |
| Data Performed | Date Submitted |

1. **OBJECTIVES**

At the end of this exercise, students must be able to:

* Identify the concepts of valuing the organization’s data
* Determine the different types of data
* Perform simple analysis for understanding your dataset
* Identify analytical questions for the organization
* Identify the different challenges in preparing data
* Identify the different foundation of analytics and its differences
* Understand the scope of data mining and analytics
* Understand how analytics works on the different area such as the descriptive, predictive and prescriptive analytics

1. **BACKGROUND INFORMATION**

Corporations have been the greatest beneficiaries from this data revolution. In 2006, oil and energy companies dominated the list of top six most valuable firms in the world, but in 2016, the list is dominated by data firms like Alphabet, Apple, Facebook, Amazon and Microsoft. Platform companies and data-aggregators capitalize on individual data by selling to advertisement networks and marketers looking to target specific segments, influence buyer behaviour and make dynamic pricing decisions.

Governments are increasingly becoming data savvy and leveraging open data for improving quality of life of its citizens by better design and targeting of welfare schemes, data driven policies, and improving participative governance.

The world produces 2.5 quintillion bytes a day, and 90% of all data has been produced in just the last two years. To assess the value of data for stakeholders, it is important to differentiate between data, content and information. Data is simple, raw and unorganized facts and consists of basic subscriber information (BSI), transactional data and content data. Transactional data is information related to communication such as IP addresses, device information used by subscribers to communicate and content data is the substance, purport or meaning of a communication. BSI and transaction data are together known as non-content data and were traditionally lower in the value hierarchy as compared to content data, but the same has changed in recent times and non-content data has been seen to provide critical insights. When this data is organized, processed and given a context, it can be termed as information. It is this information that is leveraged by corporations and is critical in decision making.

**Article Reference:** <https://www.weforum.org/agenda/2017/09/the-value-of-data/>

Data visualization is the graphical representation of information and data. By using visual elements like charts, graphs, and maps, data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data.

In the world of Big Data, data visualization tools and technologies are essential to analyze massive amounts of information and make data-driven decisions.

Our eyes are drawn to colors and patterns. We can quickly identify red from blue, square from circle. Our culture is visual, including everything from art and advertisements to TV and movies.

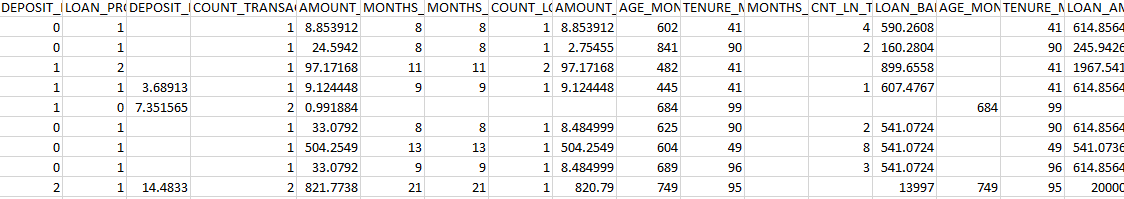
Data visualization is another form of visual art that grabs our interest and keeps our eyes on the message. When we see a chart, we quickly see trends and outliers. If we can see something, we internalize it quickly. It’s storytelling with a purpose. If you’ve ever stared at a massive spreadsheet of data and couldn’t see a trend, you know how much more effective a visualization can be.

**Article Reference:** <https://www.tableau.com/learn/articles/data-visualization>

1. **INSTRUCTIONS**

**Part 1:**

1. Open your FEU Canvas and download the **Transactional\_Information.csv** file from **[M3-TECHNICAL]**.
2. Evaluate each field of the given dataset.
3. Identify what type of data each field is and explain your answer.
4. Formulate 10 analytical questions*(reports)* that can be answered based on the given dataset.
5. What are the anomalies present on the given data? Provide basic steps in solving those anomalies.



**Part 2:**

1. Open FEU Canvas and download the **StudentsPerformance.xlsx** and **heartattack.csv** datasets from Chapter 6 folder on your Files tab.
2. Use Microsoft Excel or any chart editor application to generate at least 5 data visualization for each dataset.
3. Draw insights for each visualization.
4. Prepare a PowerPoint presentation for your data visualization.
5. Upload the Powerpoint presentation and the excel files.
6. **ASSESSMENT**

|  |  |
| --- | --- |
| Department | Computer Science |
| Subject Code | CSSELEC3/CS0009 |
| Description | Fundamentals of Analytics |
| Term/Academic Year | 2 |

|  |  |
| --- | --- |
| Topic | Implementing Analytic Solutions, Putting It Together – Integrated View of Customers |
| Lab Activity No | 3 |
| Lab Activity | The Value of Data: Transactional Information and Effective Data Visualization |
| CLO | **2, 3** |

**Note: The following rubrics/metrics will be used to grade students’ output in the lab exercise 3.**

|  |  |  |
| --- | --- | --- |
| **Criteria** | **Description** | **Score** |
| Accuracy | Able to identify the correct type of data for each field | 40% |
| Completeness | Able to provide the expected answers and analysis | 30% |
| Timeliness | Activity was submitted before the deadline | 10% |
| Analysis of the issues | Identifies the best solution for each problem | 20% |
| **Total** |  | 100% |