**PROG 8460**

**Web Analytics and Business Intelligence**

**Assignment 1**

**Task 2**

1. **What are the 3 Vs of Data and explain each one in detail**?

**Volume:** This refers to the amount of data being generated. With the rise of the internet, social media, and IoT (Internet of Things), the volume of data produced is enormous. For example, companies like Facebook or Google handle petabytes of data daily, which requires special tools to store and analyze.

**Velocity**: This is the speed at which data is generated and processed. In today's world, data is created in real-time, such as live social media feeds or financial transactions. Businesses need to analyze this data quickly to make timely decisions, like detecting fraud in banking transactions.

**Variety**: This refers to the different types of data available. Data can come in various formats, such as structured data (like databases), unstructured data (like emails and videos), and semi-structured data (like XML files). Businesses must be able to manage and analyze all these different formats to gain insights.

1. **List capabilities of Business Intelligence systems.**

* Data Integration: Combining data from different sources for a unified view.
* Data Analysis: Analyzing data to find trends and patterns.
* Reporting: Creating reports and dashboards to visualize data insights.
* Performance Metrics: Tracking key performance indicators (KPIs) to measure success.
* Data Mining: Discovering hidden patterns in large datasets to predict future trends.

1. **Different types of data with examples for each type.**

* **Structured Data:** This is organized data that fits into a table format, like a database. Example: Customer records in a relational database.
* **Unstructured Data:** This is data that doesn't have a predefined format. Example: Social media posts, emails, and videos.
* **Semi-structured Data:** This data doesn't fit strictly into a table but has some organizational properties. Example: XML files or JSON files.

1. **Define data visualization.**

Data visualization is the graphical representation of information and data. By using visual elements like charts, graphs, and maps, data visualization helps to make complex data easier to understand and identify patterns, trends, and insights.

1. **What is a KPI and provide an example**

KPI (Key Performance Indicator): is a measurable value that demonstrates how effectively a company is achieving its business objectives. For example, the Customer Retention Rate measures the percentage of customers who stay with a company over a specific period. A high retention rate indicates customer satisfaction and loyalty.

1. **What is a BI system?**

A BI (Business Intelligence) system: is a technology-driven process for analyzing data and presenting actionable information to help business managers make informed decisions.

1. **What are the 5 C's of Data for data preparation and the purpose of each?**

* **Cleaning:** Removing errors and inconsistencies in the data to ensure accuracy.
* **Consolidation:** Combining data from different sources to create a single, comprehensive dataset.
* **Conversion:** Transforming data into a format suitable for analysis, such as changing text to numeric values.
* **Curation:** Organizing and maintaining data to ensure it is relevant and easily accessible.
* **Collaboration:** Involving different teams and stakeholders to ensure data is correctly interpreted and used.

1. **What are some Key Success Factors of a Successful BI Program and explain each factor?**

* **Clear Objectives:** Setting specific goals for what the BI program should achieve ensures everyone is aligned and focused.
* **User Engagement:** Involving end-users in the BI process helps ensure the tools and reports meet their needs, leading to higher usage rates.
* **Data Quality:** Ensuring high-quality data is crucial for accurate analysis and decision-making. Poor data leads to poor outcomes.
* **Management Support:** Strong support from management is essential for securing resources and fostering a data-driven culture.
* **Scalable Infrastructure:** A BI system must be able to grow with the organization, accommodating increasing data volumes and user demands.
* **Training and Support:** Providing training for users ensures they can effectively utilize BI tools, leading to better insights and decisions.
* **Continuous Improvement:** Regularly reviewing and improving the BI process helps adapt to changing business needs and enhances overall effectiveness.