**Task 2 - BI and Data related Questions [30 marks]**

1. **What are the 3 Vs of Data and explain each one in detail? [4 marks]**

The 3 Vs of Data are used to define the challenges and characteristics of Data. They are used understand the complexity of managing and extracting value from massive data sets. Volume, Velocity and Variety.

1. Volume  
Volume refers to the size or amount of data being collected and stored. With the rise of technologies like the internet, social media, IoT, and mobile devices, data is being generated at an unprecedented rate.

2. Velocity  
Velocity refers to the speed at which data is being generated, transmitted, and processed. In today's digital age, data is produced in real-time or near-real-time, creating a need for fast processing and decision-making capabilities.

3. Variety  
Variety refers to the different types, formats, and sources of data that organizations are handling. Unlike traditional data, which was primarily structured and came from relational databases, modern data is much more diverse like unstructured data.

1. **List capabilities of Business Intelligence systems. [5 marks]**

BI, Business Intelligence, systems are powerful tools that help organizations collect, integrate, analyze, and present business data. The goal of BI is to enable informed decision-making by turning data into actionable insights. Here are some capabilities of BI systems:

1. Data Analysis
2. Reports
3. Data Visualization
4. Dashboards
5. Data Integration

1. **Different types of data with example for each type. [3 marks]**
2. Structured Data  
   Structured data is organized and formatted in a way that can be easily stored and accessed. It is usually stored in relational databases with defined schemas (tables with rows and columns).

Examples:

* Database tables with columns (Order, Date, ShipAddress, Price) and a example row:
* 10, ’10/10/2024’, ‘Conestoga Ave 500, ON, N2J O5N’, 500,50.

2. Semi-structured Data  
Semi-structured data doesn’t have a fixed schema like structured data, but it contains tags or markers to separate different elements, making it more flexible. This type of data often resides in formats like JSON or XML.

Examples:

* JSON (JavaScript Object Notation): Widely used in web APIs and configuration files, this format stores data in a key-value pair structure.json.
* { "Order": 10,
* "Date": "10/10/2024",
* "ShipAddress": "Conestoga Ave 500, ON, N2J O5N",
* "Price": 500,50}

3. Unstructured Data  
Unstructured data lacks a predefined format or organizational schema. It can be text-heavy or contain rich media files like images, videos, and audio. Unstructured data is difficult to organize and analyze using traditional methods.

Examples:

* Social Media Posts: A Tweet from Twiter may have text and images in a unstructured way, which are highly unstructured.
* ”Just ordered some new shoes #FIRE #NEEDED”

4. Metadata  
Metadata is "data about data." It describes or provides information about other data, helping to understand its structure, content, or characteristics.

Examples:

* File Properties Data: Information about a file stored on a computer system, such as file name, size, creation date, and author.
  + Example: Filename: test.pdf, Size: 223MB, Created: 2024-10-10, Author: Lucas Alves

1. **Define data visualization. [2 marks]**

Data visualization is a graphical representation of data and information using visual elements like charts, graphs, maps, and dashboards. It helps simplify complex data, making it easier to understand. Transforming raw data into visual formats enables quicker decision-making. Effective data visualization enhances data storytelling and supports better communication of insights across an organization.

1. **What is a KPI and provide an example (apart from one that has been already provided within the lecture) [2.5 marks]**

Key Performance Indicator (KPI) is a measurable value that indicates how effectively an organization or individual is achieving specific business objectives. KPIs help businesses monitor progress toward their goals, identify areas for improvement, and make data-driven decisions. They can be financial, operational, or customer-focused metrics depending on the organization’s objectives.

Net Profit Margin is a financial KPI that measures the percentage of revenue that remains as profit after all expenses, taxes, and costs have been deducted. It reflects how efficiently a company converts revenue into actual profit and provides insights into overall profitability.

Formula: Net Profit Margin = (Net Profit / Total Revenue) × 100

* Net Profit is the income left after all operating expenses, interest, taxes, and other costs have been subtracted from total revenue.
* Total Revenue is the total income generated from sales or services.

Example:

If a company has a net profit of $50,000 and total revenue of $500,000:

Net Profit Margin = (50,000 / 500,000) × 100 = 10 %

1. **What is a BI system? [1 marks]**

A BI system is a solution that collects, analyzes and presents data to help organizations make informed decisions. BI systems use tools like dashboards, reports, and data visualization to transform raw data into actionable insights. These systems enable companies to track performance, identify trends and improve decision-making across various business functions.

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

The 5 C's of Data for data preparation provide a structured approach to ensure data quality and readiness for analysis. These steps are crucial for transforming raw data into a format suitable for generating insights. The 5 C's are: Capture, Clean, Combine, Calculate, and Control.

**Capture**: Quickly establish scalable connections to various data sources while limiting the number of individuals responsible for data acquisition. This approach prevents bottlenecks and shields downstream reporting from disruptions caused by source system changes.

**Clean**: Maintain high data quality by creating a dedicated repository that serves as the authoritative source. This involves filtering irrelevant data and ensuring completeness, which fosters user confidence in insights generated. Clean data is critical for accurate visualization and decision-making.

**Combine**: Integrating data from multiple sources enhances analysis but can be technically challenging. Centralizing data in a unified repository simplifies this process and facilitates more insightful visualizations.

**Calculate**: Centralized calculations ensure consistency and accuracy across reports, reducing errors associated with independent calculations. This fosters greater confidence in the results.

**Control**: Implementing these processes enhances control over data management, enabling organizations to be agile and responsive to business needs. This leads to quicker decision-making and more impactful data visualizations, aligning insights with strategic objectives.

1. **What are some Key Success Factors of a Successful BI Program and explain each factor? [7.5 marks]**

**Clear Business Objectives**: Establishing clear and aligned business objectives is essential for guiding the BI program. Organizations must define their goals, such as improving efficiency or driving revenue, to prioritize initiatives and ensure alignment with their overall strategy.

**Data Quality and Governance**: High-quality data is fundamental for accurate insights. Implementing governance policies ensures data consistency, accuracy, and reliability, establishing standards for data entry and regular cleansing.

**User Adoption and Training**: Successful BI relies on user adoption. Providing tailored training and resources helps users effectively utilize BI tools, fostering a culture of data-driven decision-making.

**Data Integration**: Centralizing data from various sources is crucial for comprehensive analysis. Effective integration allows users to analyze structured and unstructured data in context, revealing valuable insights.

**Performance Measurement and Metrics**: Establishing clear metrics to measure the BI program's performance is critical. Tracking KPIs related to usage and outcomes helps identify improvement areas and demonstrates the program's value.

**References**

OnStrategy. (n.d.). *27 examples of key performance indicators*. OnStrategy. <https://onstrategyhq.com/resources/27-examples-of-key-performance-indicators/>

Database Trends and Applications. (2021, March 24). *The five C’s of data visualization*. DBTA. <https://www.dbta.com/Editorial/Trends-and-Applications/The-Five-Cs-of-Data-Visualization-106821.aspx>

Yellowfin. (2023, March 13). *Top 5 keys to a successful business intelligence program*. Yellowfin. <https://www.yellowfinbi.com/blog/top-5-keys-to-a-successful-business-intelligence-program-macquarie-uni>