BUSINESS INTELLIGENCE AND ANALYTICS

EXAM PREPARATION

dr. Edona Elshan
13th of March 2024
e.elshan@vu.nl

Your feedback is needed!

Evalueer via **https://app.evalytics.nl/#**

evm-686

Example Questions

Single Choice Questions

1. **Data Quality Dimension:** Which of the following is not a commonly recognized dimension of data quality?
   - A) Accuracy
   - B) Completeness
   - C) Colorfulness
   - D) Consistency
2. **Data Mining Process:** Which of the following best describes data mining?
   - A) A process to enter data into the system.
   - B) A process to extract patterns from large datasets.
   - C) A process to create backups of data.
   - D) A process to delete old data.
3. **Importance of Data Quality:** Poor data quality can lead to which of the following consequences?
   - A) More accurate business decisions.
   - B) Improved customer satisfaction.
   - C) Inefficient business processes and faulty decision-making.
   - D) Reduced costs of data management.
4. **ACID Properties in Databases:** Which of the following is not part of the ACID properties of a relational database system?
   - A) Atomicity
   - B) Consistency
   - C) Durability
   - D) Flexibility

Open Questions (1) - Enhancing Data Quality at HighFidelity Retail

HighFidelity Retail has established a strong presence in both brick-and-mortar and e-commerce spaces. As they expanded, the management noticed that discrepancies in their data are starting to impact their operational efficiency and decision-making processes. Issues such as inconsistent product categorization across platforms, errors in sales data due to manual entry, and the lack of a unified view of inventory levels have led to customer complaints, misinformed stock decisions, and an overall distrust in the data provided for analysis.

(a) Define data quality within the context of retail operations, considering the dimensions outlined by the DAMA UK Working Group on "Data Quality Dimensions" (2013), and analyze how each dimension directly impacts the functioning of a retail business. (4 points)

(b) Discuss the ramifications of HighFidelity Retail's data quality issues on different levels of operations, and refer to the impacts outlined by Redman (1998) in your explanation. (6 points)

(c) Develop a multi-faceted strategy for HighFidelity Retail to address their data quality issues. Your plan should include steps to clean dirty data, fill in missing values, resolve inconsistencies, integrate non-integrated data, and correct data in the wrong format, with an emphasis on both the roles and responsibilities within the organization, as noted by Weber et al. (2009). (7 points)

Example Answer to Open Question No. 1

A) Data quality in retail operations is the degree to which data sets meet the requirements of the retail processes they support. According to DAMA UK, data quality dimensions include:

- **Accuracy**: Ensuring data correctly reflects real-world attributes; for HighFidelity Retail, this means product descriptions and pricing are precise and reliable.
- **Completeness**: No vital data is missing; essential for maintaining comprehensive inventory records.
- **Consistency**: Data is uniform across all systems; critical for synchronizing online and offline sales data.
- **Timeliness**: Data is up to date; affects dynamic pricing and inventory restocking.

Each dimension plays a critical role in the functioning of a retail business. Inaccurate data leads to mispriced items, eroding profits and customer trust. Incomplete data can result in stockouts, damaging customer relations. Inconsistency across sales platforms can confuse customers and staff, leading to operational inefficiencies. Untimely data can delay critical actions, such as restocking or promotional offerings.

Example Answer to Open Question No. 1

B) HighFidelity Retail's data quality issues could lead to operational inefficiencies and poor decision-making. Redman (1998) suggests that the implications can be :

- **Operational Impact**: Errors in sales data disrupt inventory management, creating a direct financial cost through lost sales or excess inventory.
- **Tactical Impact**: Inconsistent data hinders the ability to conduct accurate sales analyses, misguiding marketing campaigns and resource allocation.
- **Strategic Impact**: If management cannot trust the data, strategic decisions become guesswork, potentially stunting growth and adaptation to market trends.

Example Answer to Open Question No. 1

c) HighFidelity Retail's strategy should include a multi-pronged approach to improve data quality:

- **Clean Dirty Data**: Implement data cleansing procedures to correct inaccuracies in sales data and standardize product categorizations.
- **Fill Missing Values**: Establish mandatory fields in data entry points and utilize data imputation techniques where appropriate.
- **Resolve Inconsistencies**: Integrate databases to ensure a unified view of inventory and sales, using master data management practices.
- **Format Correction**: Standardize input formats and validate data at the point of entry.

Roles and responsibilities must be well-defined, following Weber et al.'s framework:

- **Data Quality Lead**: Oversees data quality initiatives and reports progress to executives.
- **Data Stewards**: Monitor data quality, coordinate with different departments, and maintain data standards.
- **IT Department**: Implements and maintains data management systems.
- **Staff Training**: Regular workshops on data entry procedures and the importance of data quality.

Open Questions (2) - Relational Database Strategy for CloudTech

CloudTech Software has experienced rapid growth thanks to its cloud storage solutions. However, their current database system has been strained under the increasing load and is not keeping pace with the dynamic CRM needs. The limitations are becoming more evident as the company seeks to provide more personalized services and complex data analysis for customer usage patterns. The IT team has pointed out that the current system's inability to integrate with new data analytics tools and provide real-time insights is a significant hindrance.

(a) Describe the structure and benefits of relational databases for managing CRM data, with a focus on the essential characteristics that make them suitable for such a task. (4 points)

(b) Examine the constraints of CloudTech Software's current operational database system for business intelligence and the advantages of transitioning to a relational database for this purpose. Discuss how a relational database could improve the company's reporting and decision-making capabilities, supported by Wixom & Watson's (2001) findings on Data Warehousing Success. (6 points)

(c) Provide examples of how SQL can be used to extract insights from customer data, suggesting specific queries that would help CloudTech Software understand and predict customer storage needs. Reference appropriate algorithms and methods from data mining tasks as presented in the CRISP-DM process (2024). (7 points)

Open Questions (2) - Relational Database Strategy for CloudTech

CloudTech Software has experienced rapid growth thanks to its cloud storage solutions. However, their current database system has been strained under the increasing load and is not keeping pace with the dynamic CRM needs. The limitations are becoming more evident as the company seeks to provide more personalized services and complex data analysis for customer usage patterns. The IT team has pointed out that the current system's inability to integrate with new data analytics tools and provide real-time insights is a significant hindrance.

(a) Describe the structure and benefits of relational databases for managing CRM data, with a focus on the essential characteristics that make them suitable for such a task. (4 points)

(b) Examine the constraints of CloudTech Software's current operational database system for business intelligence and the advantages of transitioning to a relational database for this purpose. Discuss how a relational database could improve the company's reporting and decision-making capabilities, supported by Wixom & Watson's (2001) findings on Data Warehousing Success. (6 points)

Example Answer to Open Question No. 2

A) Relational databases are structured in a way that enables the storage of data in tables with predefined relationships between them, which is essential for CRM systems that need to manage complex and interrelated customer data. The key benefits of relational databases include:

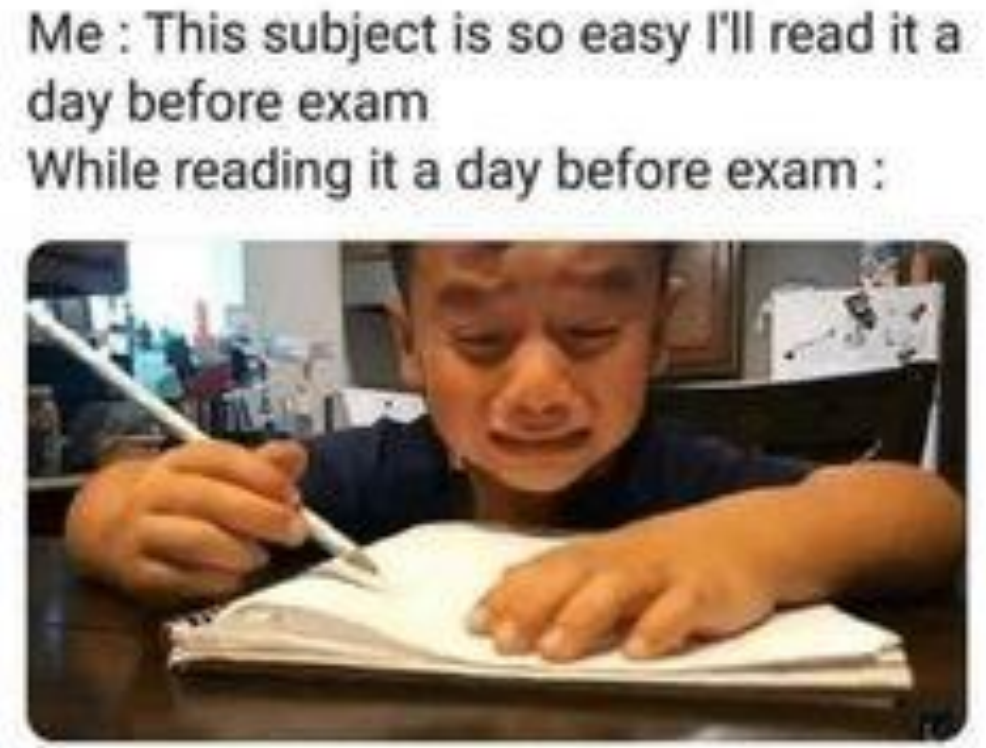
- **Data Integrity**: Relational databases use primary and foreign keys to ensure data remains consistent across different tables, which is vital for accurate customer relationship management.
- **Flexibility**: They allow for complex queries and reporting, making it easier to track customer interactions and understand customer behavior.
- **Scalability**: As CloudTech Software grows, relational databases can scale to manage increasing amounts of data without compromising performance.
- **Security**: These databases provide robust security features, crucial for protecting sensitive customer data in the CRM.

Example Answer to Open Question No. 2

B) CloudTech Software's current operational database system may be limited by its inability to handle large volumes of diverse data efficiently and may not support complex data analytics tools necessary for business intelligence (BI). According to Wixom & Watson (2001), successful data warehousing, which relies on a robust relational database, significantly improves an organization's BI capabilities. Transitioning to a relational database could offer:

- **Improved Reporting**: Relational databases facilitate comprehensive reporting by allowing complex queries that combine data from multiple sources.
- **Enhanced Decision-Making**: By enabling real-time analytics and insights, relational databases support informed decision-making based on up-to-date customer data.
- **Better Integration**: Relational databases can easily integrate with modern BI tools, enhancing CloudTech Software's ability to personalize services.

Questions?



(Hopefully not you)

Some last words from me

Thank you for showing up &
actively participating

BUSINESS INTELLIGENCE AND ANALYTICS

EXAM PREPARATION

dr. Edona Elshan
13th of March 2024
e.elshan@vu.nl