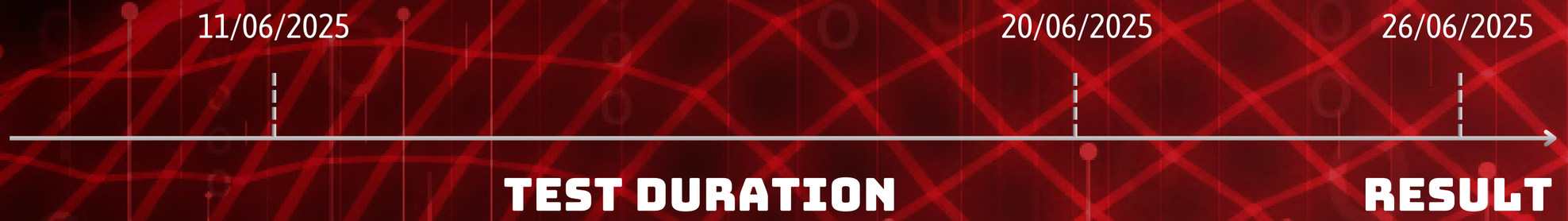
BUSINESS INTELLIGENCE 9

ROUND 01: CHESSBOARD



BUSINESS INTELLIGENCE
SEASON 9
TRÍ TUỆ KINH DOANH

BUSINESS INTELLIGENCE 9

TABLES OF CONTENT

BUSINESS INTELLIGENCE 9

INFORMATION

ROUND 01: CHESSBOARD

BUSINESS INTELLIGENCE
SEASON 3
TRÍ TUỆ KINH DOANH

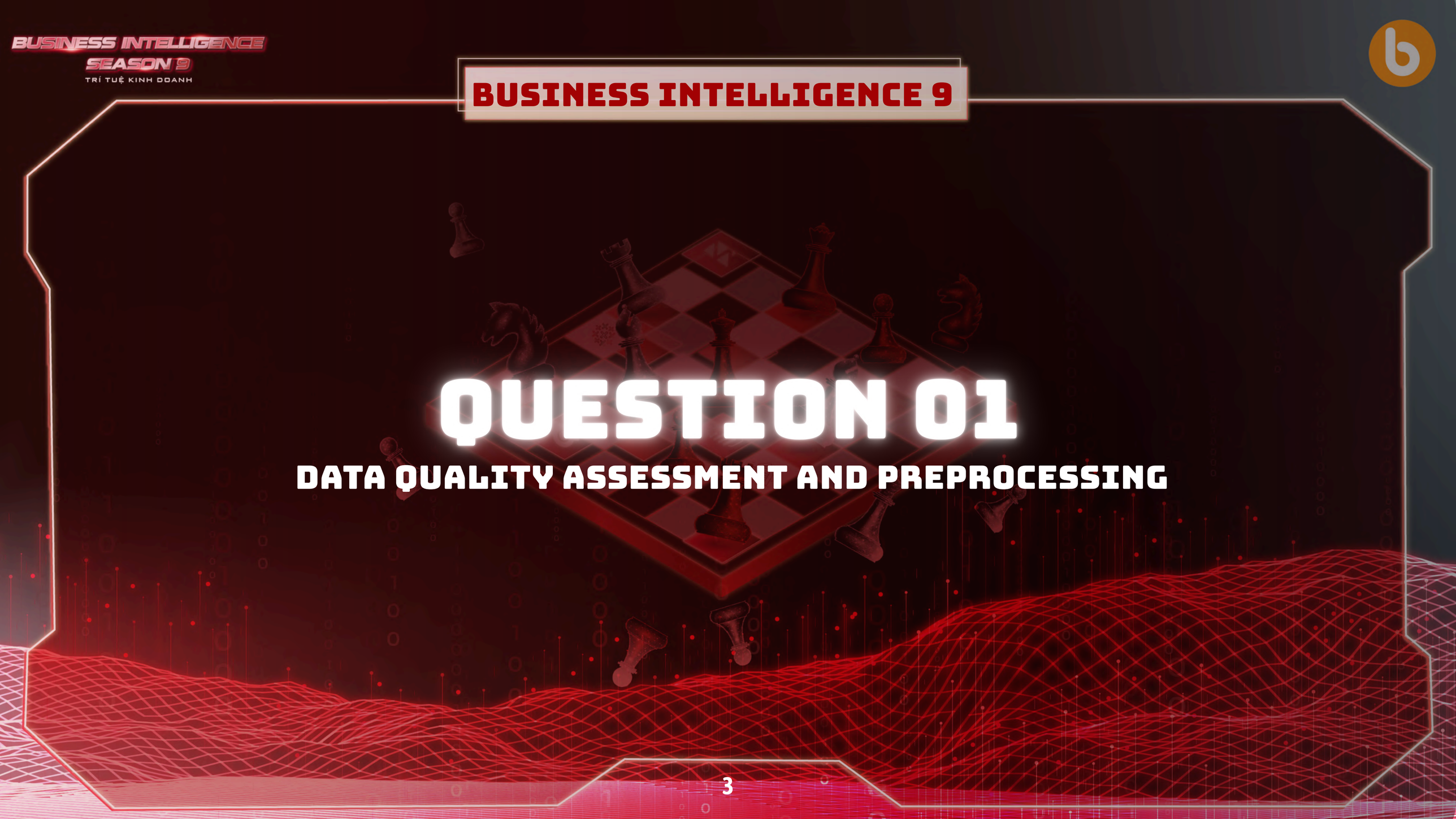
INFOMATION

Link to Dataset Round 01: Dataset Round 01

Using the provided dataset along with your insights, please assist ITB CLUB in addressing the following questions.

Link to Data Dictionary Round 01: Data Dictionary Round 01

Instruction using ITB'S Online Exam Platform: Instruction Website Round 01



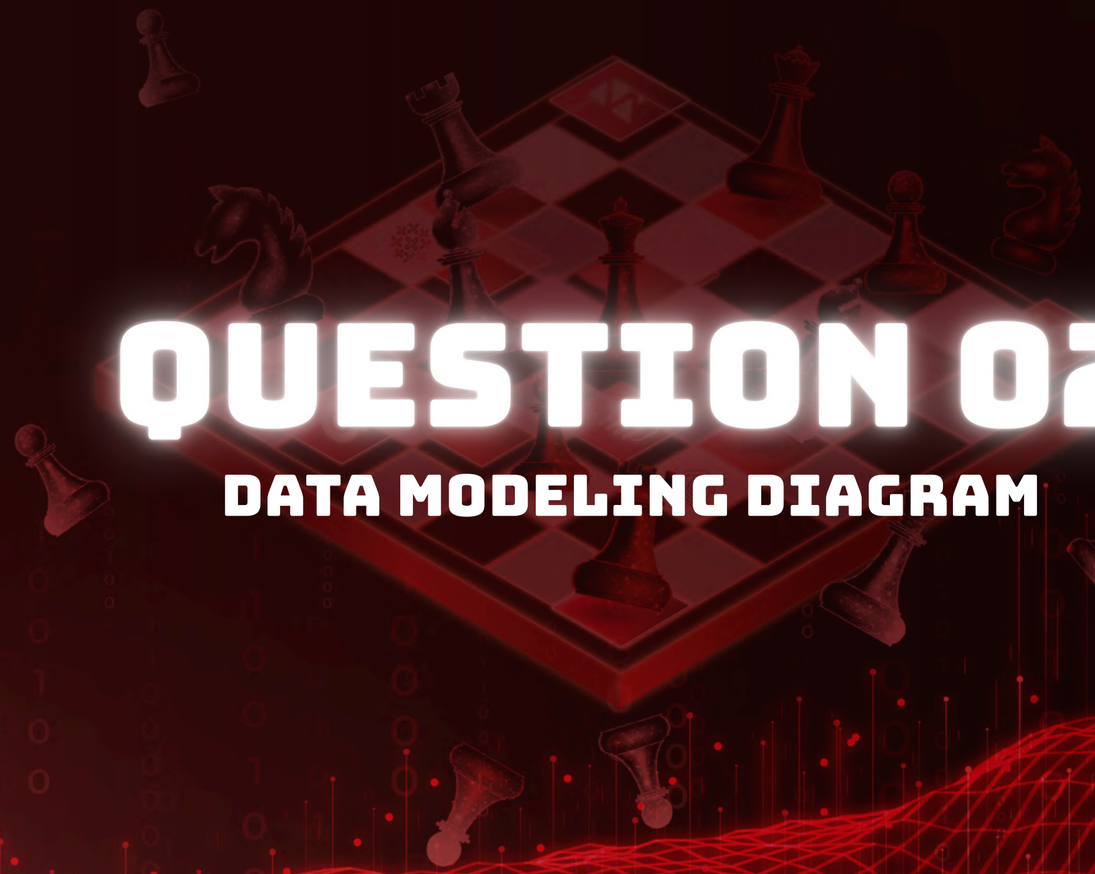
QUESTION 01

DATA QUALITY ASSESSMENT AND PREPROCESSING

QUESTION 01

OBJECTIVES

1. Identify data redundancy within the dataset to prevent unnecessary data duplication.
2. Detect and resolve inconsistencies in data labeling to enhance data integrity and uniformity.
3. Detect and correct violations of logical data constraints to maintain data validity.
4. Identify missing values and apply appropriate imputation methods to ensure dataset completeness.
5. Diagnose cross-table data inconsistencies and analyze their root causes to improve data coherence.



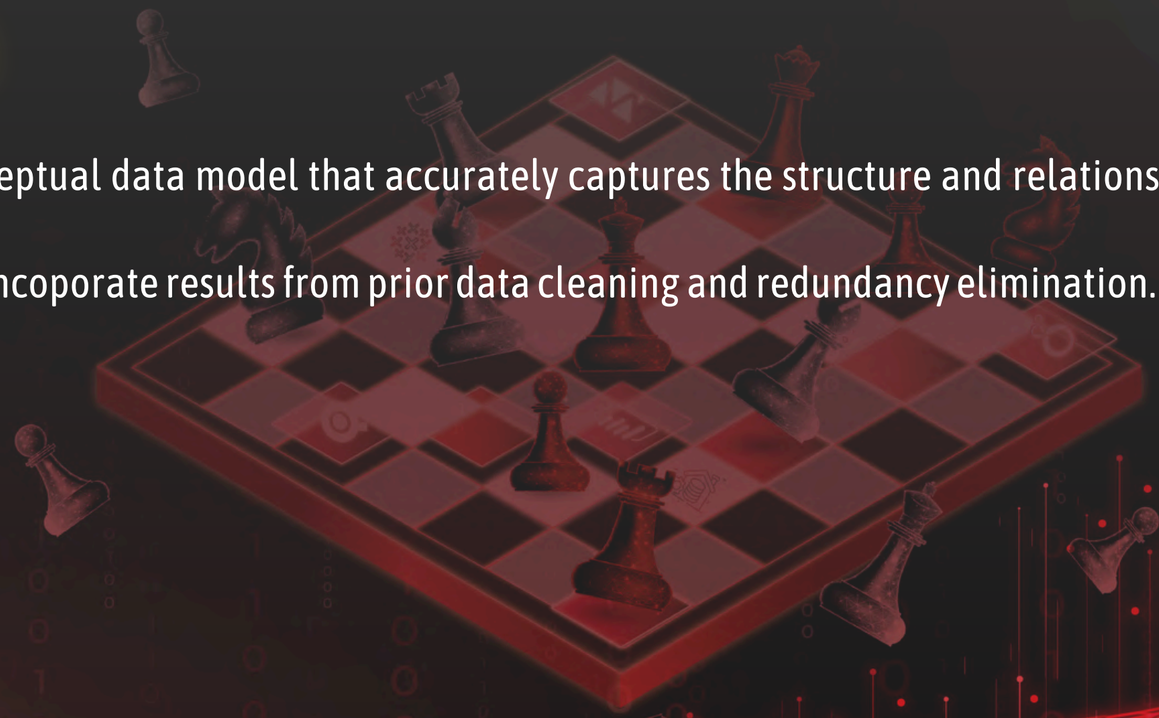
QUESTION 02

DATA MODELING DIAGRAM

QUESTION 02

OBJECTIVES

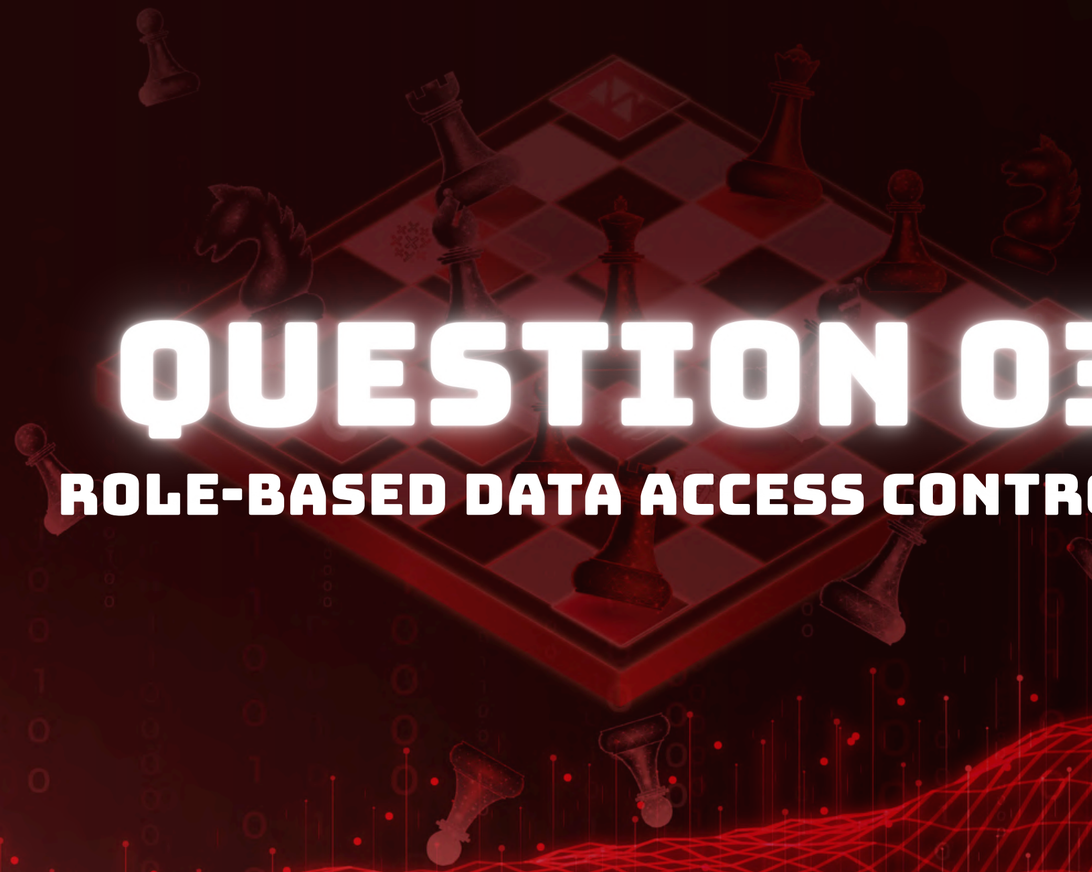
1. Design a conceptual data model that accurately captures the structure and relationships among the provided data entities, incoporate results from prior data cleaning and redundancy elimination.



QUESTION 02

INSTRUCTIONS

1. Develop an Entity-Relationship (ER) diagram or class diagram representing the cleaned data model, where redundant columns and inconsistent labels have been resolved.
2. Identify and represent key entities.
3. Specify and label all relevant attributes of each entity based on the cleaned and normalized tables.
4. Provide clear justification for each relationship type (one-to-one, one-to-many) with logical reasoning based on business context and data structure.
5. Use proper notation to indicate primary keys, foreign keys, cardinality, and relationship types.



QUESTION 03

ROLE-BASED DATA ACCESS CONTROLS

QUESTION 03

OBJECTIVES

There are several functional departments, including the Board of Directors (BOD), Human Resources (HR), Finance & Accounting, Marketing, and Sales. The Sales department comprises multiple hierarchical roles, including CRM Lead, Store Managers, Regional Managers, and Sales Operations. Different departments and roles require varying levels of data access to perform their duties while ensuring compliance and minimizing risk.

1. Propose a comprehensive strategy for defining role-based data access permissions across departments and roles to ensure both data security and operational efficiency.

QUESTION 03

OBJECTIVES

Below is an example of data usage needs by role.

Department / Role	Needs
Board of Directors (BOD)	High-level performance trends, brand health KPIs, and strategic benchmarks.
HR/People & reward analytics	Store-level performance metrics tied to employee incentives and rewards.
Finance and Accounting	Transaction details, data for P&L, budget forecasting.
Marketing	Brand health indicators, competitive positioning, and customer segmentation.
CRM Lead	Customer loyalty metrics (e.g., NPS).
Sales Operation	Cross-region analytics to identify operational efficiencies and gaps.
Regional Managers	Performance trends and comparisons for all stores within their region.
Store Managers	Real-time metrics specific to their store (e.g., sales, inventory, staffing).

QUESTION 03

INSTRUCTIONS

1. Role Definition and Permission Assignment: Define appropriate roles and responsibilities for each department and job function.
2. Data Access Matrix Development: Construct a detailed data access matrix that aligns with organizational requirements and confidentiality obligations, ensuring that data accessibility corresponds to the principle of least privilege and operational needs:

- Rows: Data entities/tables.
- Columns: Designated roles.
- Access Levels: Granular access control levels based on the roles and functions mentioned.

QUESTION 03

INSTRUCTIONS

3. Final Justification per Role: For each role, include a brief but well-reasoned justification explaining why the specific data access level is appropriate. Justifications should consider:

- Operational responsibilities of the role.
- Sensitivity of the data.
- Hierarchical level within the organization.
- Potential risks associated with the misuse or overexposure of data.

QUESTION 04

COMPETITIVE LANDSCAPE ANALYSIS AND BRAND POSITIONING FOR HIGHLAND COFFEE

QUESTION 04

OBJECTIVES

1. Conduct a comprehensive competitive landscape analysis of Highland Coffee and key competitors using brand funnel, customer perception, and market presence data. This task aims to generate actionable insights into brand performance across various customer engagement stages and dimensions of brand health, enabling strategic positioning and growth planning.

QUESTION 05

CUSTOMER CHURN ANALYSIS AND DASHBOARD FOR HIGHLAND COFFEE

QUESTION 05

OBJECTIVES

1. Analyze customer churn patterns for Highlands Coffee by identifying key brand perception, segmentation behavioral, companionship, need states, and demographic factors through advanced data visualization and interpretation.

QUESTION 05

INSTRUCTIONS

1. Definition of Churn: A customer is defined as churned if they had activity in the past 3 months (P3M) but no visit in the past 1 month (P1M).
2. Churn Logic Implementation: Accurately identify churned customers based on P3M and P1M activity, ensuring logical integrity in the churn flag.
3. Churn Interactive Dashboard: Create an interactive dashboard to visualize churn rates across multiple dimensions. Ensure all charts are clearly labeled, well-formatted, and equipped with tooltips for ease of interpretation.

QUESTION 05

INSTRUCTIONS

4. Multi-Dimensional Visualization – Use grouped bar charts or pie charts to show churn rates by:

- Brand perception (Comprehension, NPS#P3M#Group, NPS#P3M).
- Brand image (Attribute).
- Segmentation variables (Segmentation).
- Weekday & Daypart (Weekday#end, Dayofweek, Daypart).
- Companion behavior (Companion#group).
- Need states (NeedstateGroups, Needstates).
- Demographic segments (Age, Gender, Occupation, MPI).

5. Hierarchical Drill-downs – Enable multi-level exploration using multi-chart linked drill-downs or sunburst charts:

- NPS#P3M#Group → NPS#P3M.
- Weekday#end → Dayofweek.
- NeedstateGroup → Needstates.
- Age#Group#2 → Age#group → Age.
- MPI#detail → MPI → MPI#Mean.
- Occupation#group → Occupation.

QUESTION 05

INSTRUCTIONS

6. Analysis and Recommendation: Analyze key patterns and insights revealed through the dashboard. Discuss potential root causes, business implications, and provide actionable recommendations to reduce churn.

QUESTION 06

CUSTOMER SEGMENTATION AND CHURN PREDICTION

QUESTION 06

OBJECTIVES

1. Develop a data-driven approach that integrates customer segmentation and churn prediction to identify high-risk customers. The goal is to enable Highlands Coffee to implement proactive, segment-specific retention strategies that maximize business value.

QUESTION 06

INSTRUCTIONS

1. Data Preprocessing and Feature Engineering: Prepare and transform the dataset to support effective segmentation and prediction. Focus on constructing relevant features across the following dimensions:

- Perceptual factors: Comprehension, Brand_Likability, NPS#P3M, NPS#P3M#Group.
- Behavioral patterns: Visit, PPA, Spending, Visit#Dayofweek, Visit#Daypart
- Segmentation attributes: Segment.
- Companionship context: Companion#Group.
- Need states: NeedstateGroup, Needstates.
- Demographics: Age, Gender, Occupation, MPI.

QUESTION 06

INSTRUCTIONS

2. Customer Segmentation Modeling: Develop a clustering model to segment customers into distinct and meaningful groups:

- Select appropriate features for clustering and provide justification for inclusion.
- Determine the optimal number of clusters using appropriate techniques.
- Choose a suitable clustering algorithm and justify the selection.
- Train the segmentation model and analyze the resulting clusters.
- Create detailed segment profiles by identifying the key differentiating attributes of each group.

QUESTION 06

INSTRUCTIONS

3. Churn Prediction Modeling:

Build a predictive model to estimate churn risk across the full customer base:

- Target variable is a binary churn flag, where:
  - Churn = 1 if the customer had activity in the past 3 months (P3M) but no visit in the past 1 month (P1M).
  - Retained = 0 if the customer had at least one visit in the past 1 month (P1M).
- Select relevant features for prediction, including the predicted customer segment, with clear rationale.
- Choose an appropriate classification model and provide justification.
- Apply hyper-parameter tuning to improve model performance.
- Evaluate the model using appropriate metrics.
- Discuss key churn drivers identified by the model

QUESTION 06

INSTRUCTIONS

4. Limitations and Future work: Critically evaluate the analysis and modeling process by addressing:

- Limitations in data quality, feature coverage, or modeling assumptions.
- Potential biases or confounding variables that could affect results.
- Suggestions for future improvements, such as additional data collection, advanced modeling techniques, or alternative segmentation strategies.

DELIVERABLES

QUESTION 01

DELIVERABLES

1. A detailed identification of data redundancies, labeling inconsistencies, logical integrity violations, missing data, and cross-table inconsistencies.
2. A summary explaining the methods and criteria used for identifying and resolving each type of data quality issue.
3. An analysis of the root causes behind cross-table inconsistencies along with recommended corrective actions.
4. A cleaned and preprocessed dataset with redundant columns removed, inconsistent labels standardized, logical violations corrected, and missing values appropriately imputed.

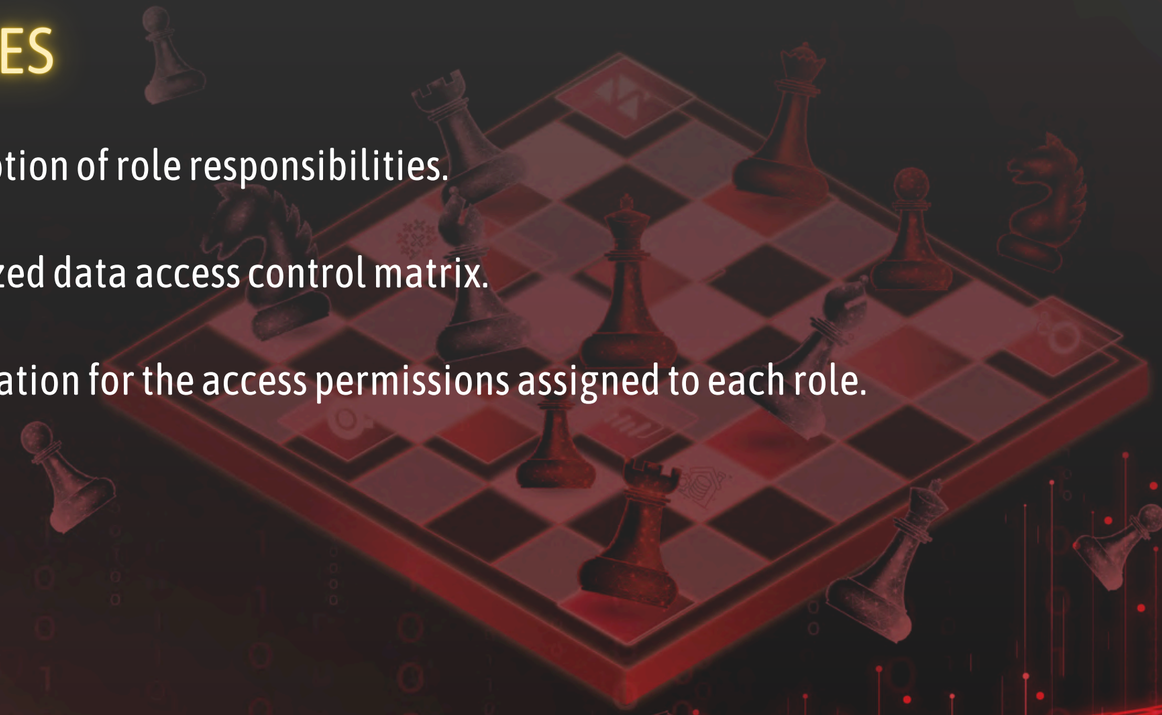
QUESTION 02

DELIVERABLES

1. A comprehensive and clearly labeled diagram of the cleaned data model, reflecting normalized structures without redundant or inconsistent attributes, and capturing key entity relationships.
2. Written explanation justifying the cardinalities of the relationships and their business logic.

QUESTION 03

DELIVERABLES



1. A brief description of role responsibilities.

2. A well-organized data access control matrix.

3. A final justification for the access permissions assigned to each role.

QUESTION 04

DELIVERABLES

1. An interactive competitive dashboard with supporting analysis that enables stakeholders to:

- Track Brand Performance Across the Customer Journey: Compare how Highland and competitors move customers from awareness to loyalty, identifying leakages and high-converting stages. (Implies: Stacked funnel chart by brand, conversion rate bar chart.).
- Diagnose Brand Perception Gaps: Visualize which attributes (e.g., "premium quality," "value") drive preference for competitors but lag for Highland. (Implies: Brand image attribute heatmap.).
- Map Competitive Positioning: Identify whitespace opportunities by plotting brands against key perception axes (e.g., price vs. convenience).
- Assess Market Strength vs. Loyalty: Prioritize markets where Highland's presence misaligns with customer retention (e.g., high share but low loyalty).
- Interactive filtering controls to explore Segment-Specific Insights.

QUESTION 04

DELIVERABLES

2. Interpretation and Strategic Recommendations for each Chart:

- Identify strengths/weaknesses across customer engagement stages (awareness → loyalty).
- Benchmark brand health (e.g., differentiation, relevance) against competitors.
- Uncover strategic opportunities for positioning and growth.

QUESTION 05

DELIVERABLES

1. Interactive Churn Dashboard: User-friendly, multi-dimensional dashboard with hierarchical drill-downs, segmented views, and visualized churn metrics.
2. Interpretation Report: Structured summary highlighting key patterns, high-risk segments, underlying churn drivers, and actionable, segment-specific strategic recommendations.

QUESTION 06

DELIVERABLES

1. Preprocessed Dataset: Cleaned and transformed dataset, including all engineered features used in modeling.
2. Customer Segmentation Model: Clustering model with a detailed analysis of segment profiles, highlighting key differentiating attributes.
3. Churn Prediction Model: Predictive model with comprehensive performance evaluation and discussion of key churn drivers.
4. Limitations and Future Work: Critical reflection on data or modeling constraints, potential biases, and suggestions for future enhancements or extensions.

SUBMISSION

DELIVERABLES

Prepare your submission in Docx/Word format, but submit it as a PDF file.

- Font: Times New Roman;
- Font size: 13 - Line spacing: 1.5;
- Language: Vietnamese or English.
- Use charts, tables, images, etc to illustrate your answers if necessary.
- Participants are allowed to use any tools during the competition.

File name format: [GroupName_LeaderName_BI9_R01]. *Example: [ITBClub_NguyenVanA_BI9_R01].*

Submit at Website: **ITB's Online Exam Platform**.

Due to: 23h59 20/6/2025.

CONTACT

If you have any questions about Round 01: Chessboard or about Business Intelligence Season 9, please contact ITB CLUB for more infomation.

ACADEMIC TEAM

businessintelligence.uel@gmail.com

COMPETITION GROUP

BUSINESS INTELLIGENCE COMPETITION