

EDA_SEE - Questions

Q1. What is the purpose of the SEE stage?

- A. Modify the dataset
- B. Understand structure, patterns, and gaps
- C. Train models
- D. Remove outliers

Correct answer: B. Understand structure, patterns, and gaps

Q2. Why do we check data types during SEE?

- A. To compress the dataset
- B. To ensure columns behave correctly in KPIs
- C. To reduce memory
- D. To label rows

Correct answer: B. To ensure columns behave correctly in KPIs

Q3. What does a missing arrival timestamp imply?

- A. A forecasting issue
- B. A data-entry failure or missing scan
- C. A bad product
- D. A unit conversion problem

Correct answer: B. A data-entry failure or missing scan

Q4. Why are frequency checks important?

- A. They help create charts
- B. They show operational footprint consistency
- C. They help with model training
- D. They remove duplicates

Correct answer: B. They show operational footprint consistency

Q5. Why do we perform outlier checks in SEE?

- A. To delete outliers immediately
- B. To understand whether the extremes are real or errors
- C. To reduce dataset size
- D. To prepare for classification

Correct answer: B. To understand whether the extremes are real or errors