Short Notes — Data Bias, Missing Data Strategies, KPIs vs Metrics

☑️ Data Bias

Data bias occurs when the collected data does not fairly represent the real population or business reality, leading to misleading analysis and wrong decisions. Bias usually enters during data collection, sampling, or measurement.
Common types include:

- Sampling bias — data collected from only one group
- Selection bias — excluding important segments
- Measurement bias — wrong or inconsistent measurement method
- Reporting bias — only positive results recorded

Real example:A food delivery company analyzes customer satisfaction using only app feedback forms. Since mostly unhappy customers submit feedback, the dataset is biased toward negative experiences and does not represent all customers.

Business impact:Biased data leads to wrong conclusions, poor targeting, and faulty strategy decisions.
Best practice:Always check who is missing from the dataset and how the data was collected before trusting results.

☑️ Missing Data Strategies

Missing data is common in real datasets due to system errors, optional fields, or integration gaps. The handling strategy depends on column type and business meaning.
Main strategies:

1. Deletion — Remove rows or columnsUsed when missing values are very few and not important. Risk: data loss.
2. Imputation — Fill missing values

- Numeric → mean or median
- Categorical → mode or “Unknown”
- Time series → forward/backward fill

3. Business-rule fillingFill based on domain logic.Example: missing sales value → treat as 0 if it means no recorded sale.

4. Leave as nullUsed when value cannot be safely guessed (like missing dates of events). Analysis then excludes those rows only when needed.

Real example:In a sales dataset, missing discount field may be filled with 0, but missing order_date should not be fabricated.

Best practice:Choose method based on data type and business context — not one rule for all.

☑️ KPIs vs Metrics

Both KPIs and metrics measure performance, but they are not the same.
Metrics are general measurements that track activity or performance. KPIs (Key Performance Indicators) are critical metrics directly tied to business goals and decision-making.
Metrics answer: What are we measuring?KPIs answer: What matters most for success?

Real examples in an online store:
Metrics:

- Website visits
- Page views
- Cart additions
- Email opens

KPIs:

- Conversion rate
- Revenue per customer
- Customer retention rate
- Monthly profit

All KPIs are metrics, but not all metrics are KPIs.

Business difference:Teams monitor many metrics daily, but leadership focuses only on KPIs for decisions.

Best practice:KPIs must be goal-linked, measurable, and actionable — not just easy to track.