Q) Explain
(a) Data quality
(b) Data Integrity
(c) Data Noise
(d) Outlers
(e) Missing and duplicate values.

(a) **Data quality :-**

→ The high quality data is the data which allows the designers to perform operations like decisions, analysis, planning etc...

→ A high quality data must contain five R's : Relevancy, Recency, Range, Robustness and Reliability.

(b) **Data Integrity :-**

→ Data Integrity is the process of maintaining consistency and accuracy while transmitting the data.

→ Data Integrity is used to check whether the data is corrupted (or) not

(c) **Data Noise :-**

→ Data noise is the difference of measured value and true value.

DN = measured value - true value.

→ When data has large additional information, its said to be noisy.

(d) **Outlers :-**

→ The data is said to be outlier, if the data is from outside the dataset.

→ Through, the data contain fluctuations by small (or) large amount of values.
→ So, it's neccessary to eliminate outliers from data set.

(e) Missing and duplicate values:-

→ The data which doesn't occur in daten set is called missing values
→ The data which occurs more than once in a data set is called as duplicate values.

Q) Discuss in detail the management of data for analysis.

A) The data management is the process of controlling, securing, delivering & improving the value of data.

→ Functions of data management are:-

1) Managing datawarehouse.
2) Data quality management.
3) Database maintainance management such as RDBMS, NOSQL.
4) Controlling and monitoring the security of data, privacy and deletion.
5) Managing records, documents and contents.
6) Creating architecture of data, modelling data and data analysis.
7) Integration of data & application.
8) Maintaining business Intelligence.

Q) Various constraints and influences will have an effect on data architecture design. Justify

A) Various factors that influence the data architecture design are :-

1) Enterprise requirements :-

→ It includes elements such as transaction reliability, transparent data management, system expansion etc...

→ It also includes conversion of records, images into useful information because it helps in taking business decisions.

2) Technology drivers :-

→ These are suggested by the database architecture designs.

→ In some cases, they are also derived from existing frameworks.

3) Economics :-

→ It's possible that some solutions though they are optional may not be potential because of their cost.

4) Business Policies :-

→ It includes various government laws that is required to make business operations.

5) Data processing needs :-

→ It includes accurate transactions performed in high volumes, data warehouses for supporting information systems.