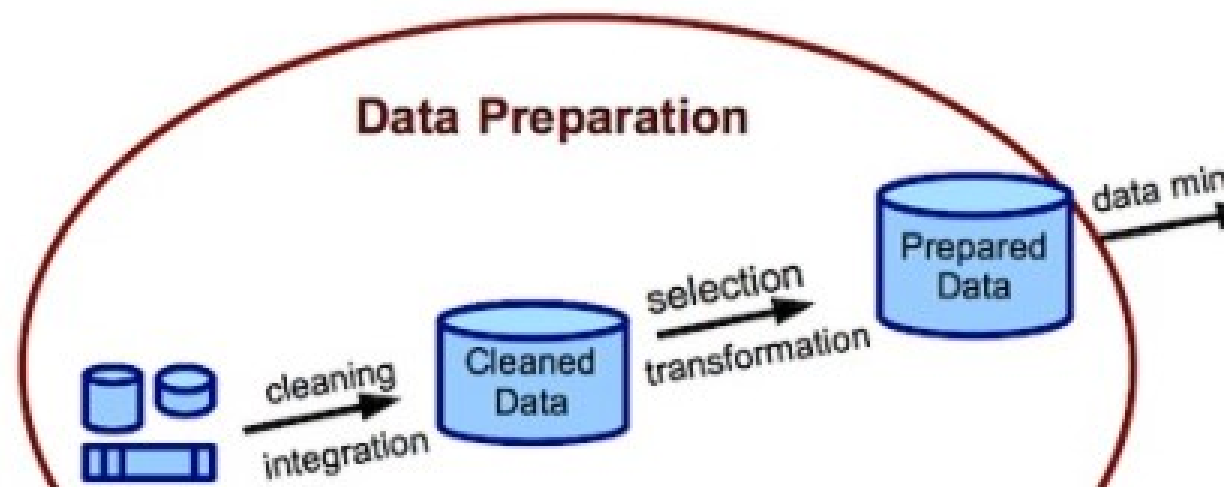


DATA PREPARATION

taroonreddy.com



Data preparation:

- Data preparation is the process of cleaning and transforming raw data prior to processing and analysis.
- It is an important step prior to processing and often involves reformatting data, making corrections to data and the combining of data sets to enrich data.
- Data preparation is often a lengthy undertaking for data professionals or business users, but it is essential as a prerequisite to put data in context in order to turn it into insights and eliminate bias resulting from poor data quality.
- The data preparation process usually includes standardizing data formats, enriching source data, and/or removing outliers.
- Data Preparation is the process of collecting, cleaning, and consolidating data into one file or data table, primarily for use in analysis.

Data preparation:

Data discretization:

It is a part of data reduction which contains particular importance especially for numerical data.

reference link

<https://towardsdatascience.com/an-introduction-to-discretization-in-data-science-55ef8c9775a2>

Data cleaning:

Manual data prep is error-prone, time-consuming and costly. Business decisions rely on analytics. But, if the data is inaccurate or incomplete, your analytics inform wrong businesses decisions. Bad analytics means poor business decisions.

Data integration:

Access data from any source — no matter the origin, format or narrative and integrating them together.

Intelligently and automatically extracting data from complex unstructured and semi-structured sources, like PDFs.

Increased access to data means less manual work, faster insights and faster time to value realized by your organization.

Data transformation:

Being able to quickly change the way data is summarized and presented enables business analysts and executives to quickly consider different perspectives and views of data. Monarch make it easy to package your clean and blended data for insightful reporting you can confidently share with the rest of your organization.

Data reduction:

It obtains reduced representation