Study Notes

Here are your concise, visual study notes on EDA!

■ Key Concepts

- **EDA (Exploratory Data Analysis)**: Analyzing data to understand its characteristics, relationsh...
- **Graphical Techniques**: Visual tools (charts, plots) used to explore and present data patterns.
- **Hidden Insights**: Non-obvious patterns, relationships, or anomalies discovered in data.
- **Hypothesis Generation**: Forming educated guesses or theories based on data observations.

■ Important Points

- ■ **Main Goal**: Get a better understanding of the data to uncover hidden insights and relationsh...
- ■ **Typical Method**: Primarily involves using graphical techniques for visualization.
- ■ **Outcome**: Helps in formulating hypotheses and developing predictive models.
- ■ **Key Info**: Summarizes characteristics like mean, median, mode, standard deviation, min/max v...
- ■ **Purpose**: Identify patterns or trends within the data for informed decision-making.

■ Quick Facts

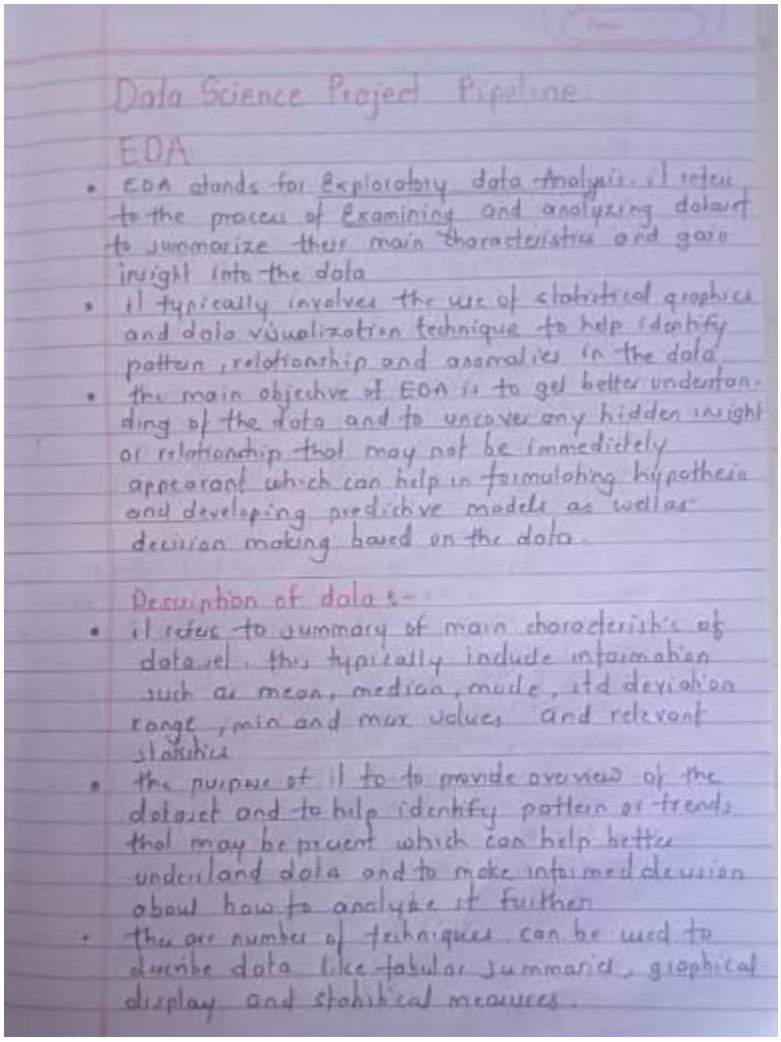
- EDA is a crucial first step before formal modeling.
- It focuses on making data visible and understandable.
- Helps to make data-driven decisions and refine strategies.

■ Memory Tips

- **EDA = E**xplore **D**ata **A**ctively! ■■■■■
- Think of EDA as "Data Detective Work": You're looking for clues (patterns) and stories (insights)...

■ Reference Images

Image 1



Data Science Project Pipeline

EDA

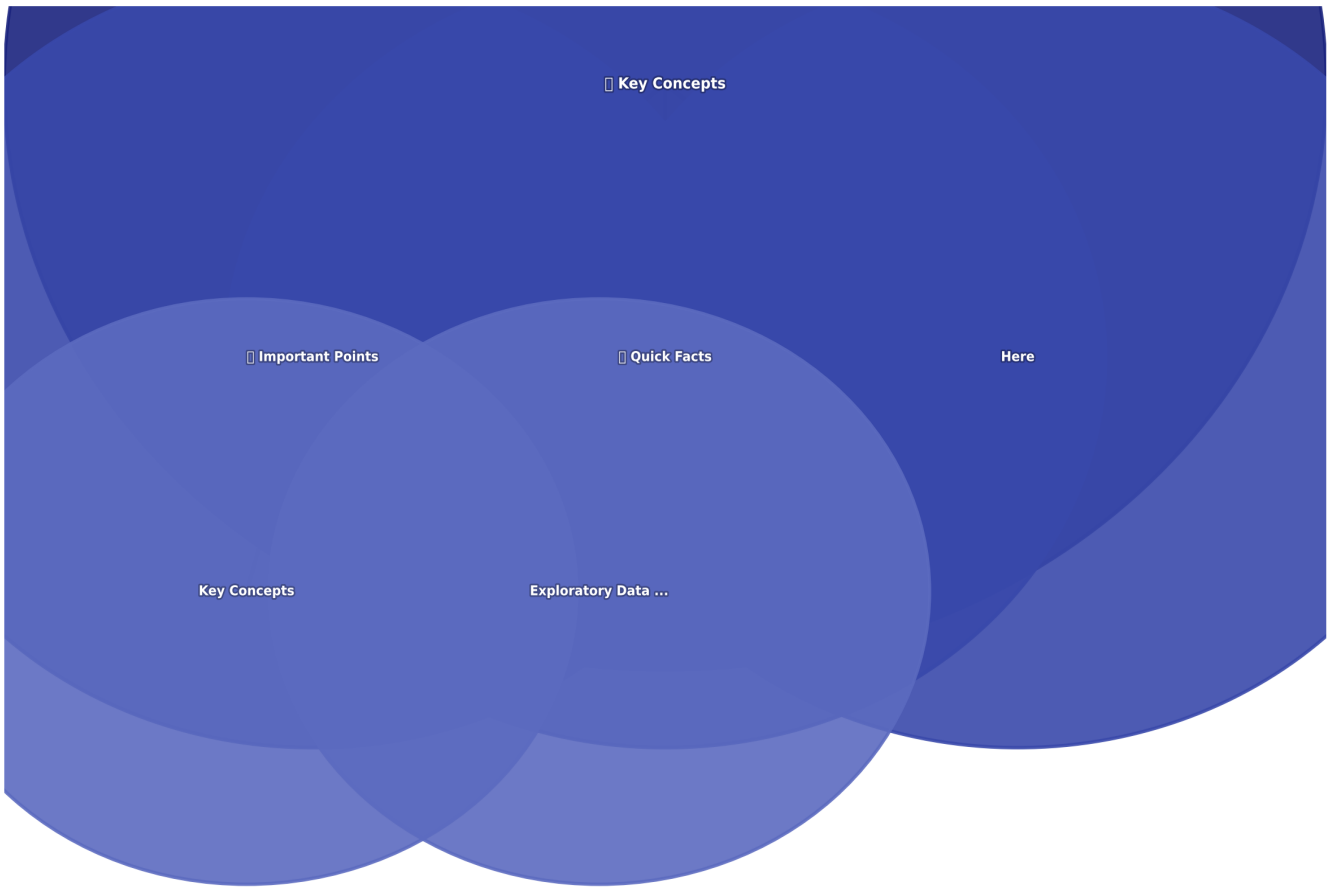
- EDA stands for Exploratory data Analysis. It refers to the process of Examining and analyzing dataset to summarize their main characteristics and gain insight into the data
- It typically involves the use of statistical graphics and data visualization technique to help identify pattern, relationship and anomalies in the data
- The main objective of EDA is to get better understanding of the data and to uncover any hidden insight or relationship that may not be immediately apparent which can help in formulating hypothesis and developing predictive models as well as decision making based on the data.

Description of data :-

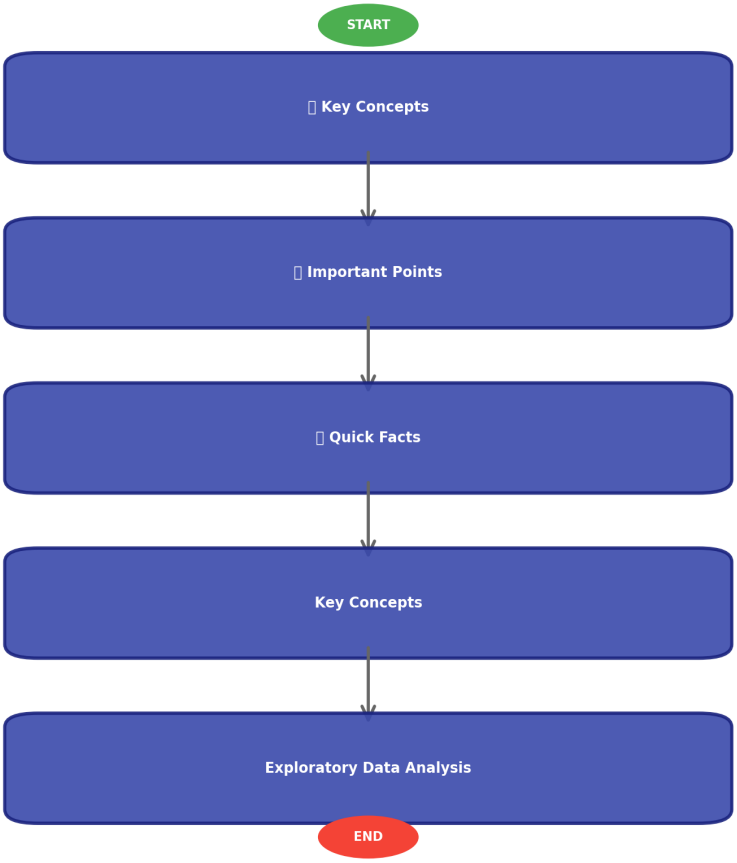
- It refers to summary of main characteristics of dataset. This typically include information such as mean, median, mode, std deviation range, min and max values and relevant statistics
- The purpose of it is to provide overview of the dataset and to help identify pattern or trends that may be present which can help better understand data and to make informed decision about how to analyse it further
- There are number of techniques can be used to describe data like tabular summaries, graphical display and statistical measures.

Visual Diagrams

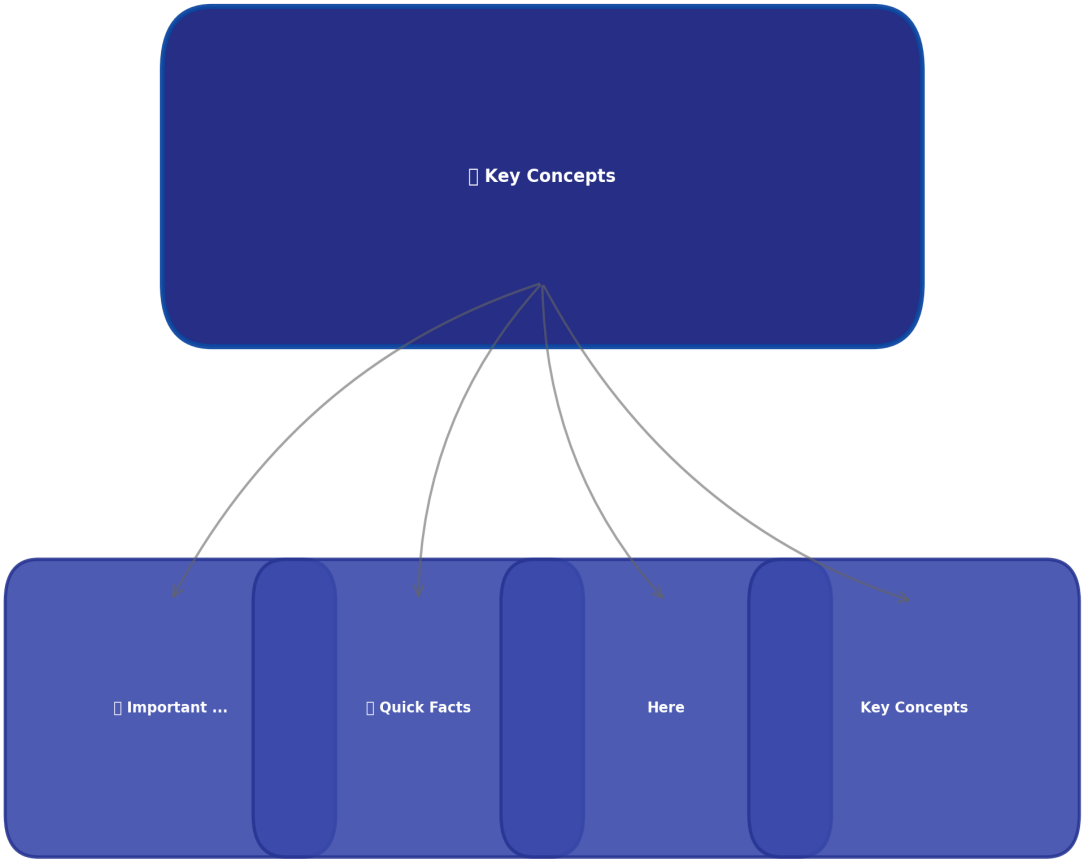
Concept Mind Map



Process Flowchart



Concept Hierarchy


Key Concepts
Important ...
Quick Facts
Here
Key Concepts