**Fundamentals Of Project**

1 Create a problem statement.

2 Identify the data you want to analyse.

3 Explore and Clean the data.

4 Analyze the data to get valuable insights.

5 present the data in terms of reports/ Dashboards using visualization.

**1. Create a Problem Statement:**

The first step in any data analysis project is to clearly define the problem you want to solve or the question you want to answer. This problem statement provides a clear direction for your analysis and helps you understand the purpose and scope of your project. It's important to make sure the problem is well-defined and relevant to the goals of your analysis.

**2. Identify the Data You Want to Analyze:**

Once you have a clear problem statement, you need to identify the data sources that are relevant to your analysis. This might involve collecting data from various sources, such as databases, spreadsheets, APIs, or online sources. It's important to ensure that the data you gather is accurate, comprehensive, and appropriate for addressing your problem statement.

**3. Explore and Clean the Data:**

Before you can analyze the data, you need to explore its structure and content. This involves understanding the variables (columns) in the dataset, their types, and the relationships between them. Exploratory Data Analysis (EDA) techniques are used to gain insights into the data, identify patterns, outliers, missing values, and potential errors. Cleaning the data involves addressing these issues – filling in missing values, correcting errors, and transforming the data into a suitable format for analysis.

**4. Analyze the Data to Get Valuable Insights:**

With the cleaned and prepared data, you can now start performing various analyses to gain valuable insights. The type of analysis you conduct will depend on your problem statement and the nature of the data. This step may involve descriptive statistics (summarizing data), inferential statistics (drawing conclusions from a sample to a larger population), machine learning algorithms (predictive modeling), and more advanced techniques like clustering or time-series analysis. The goal is to extract meaningful information from the data that can help address the problem statement.

**5. Present the Data Using Visualization and Reports/Dashboards:**

Once you've derived insights from your data analysis, it's important to communicate those findings effectively. Visualizations are a powerful way to present complex information in a clear and understandable manner. You can create graphs, charts, diagrams, and other visual representations that highlight key trends, patterns, and relationships in the data. Additionally, creating reports or interactive dashboards allows you to share your findings with stakeholders in a structured and accessible way. Tools like Tableau, Power BI, or custom programming libraries can be used to build these visualizations and reports.

In summary, the process of data analysis involves moving from a well-defined problem statement through data identification, exploration, cleaning, analysis, and finally, effective presentation of insights. Each step is crucial for ensuring that your analysis is accurate, insightful, and capable of driving informed decision-making.

**DF column types**

**1. Categorical Columns:**

These are like different groups of toys. You have a group of toy cars, another group of toy animals, and so on. Each group is special and different from the others.

**2. Numerical Columns:**

Think of these like counting your toys. You can count how many cars you have, how many dolls, and so on. Some toys can have numbers that go on and on, like counting how many steps you take.

**3. Text Columns:**

Text is like stories we tell with words. Just like you talk about your day, in data, we write down words to tell about things. It's like a tiny diary for the computer.

**4. Date and Time Columns:**

Dates and times are like knowing when special events happen. It's like remembering when your birthday is or when it's time to eat lunch.

**5. Boolean Columns:**

Booleans are like having two magic buttons. One button says "yes" and the other says "no." It's like when you choose between two options, like picking your favorite color.

**6. Ordinal Columns:**

These are like lining up your toys in a row from your most favorite to your least favorite. You can tell which ones you like more, even though you don't use numbers.

**7. Ratio Columns:**

Ratios are like sharing something fairly. If you have twice as many cookies as your friend, that's a ratio. It's like when things are in equal parts.

**8. Interval Columns:**

Intervals are like measuring how hot or cold it is. You can tell if it's warmer or colder, but you don't need to know the exact numbers.

**9. Geospatial Columns:**

Geospatial is like knowing where you are on a map. It's like telling someone your address or where you went on a trip.

**10. Image Columns:**

Images are like drawings or pictures you make. You can look at them to see what things look like, just like looking at your favorite storybook.

**11. Audio Columns:**

Audio is like listening to your favorite songs or stories. It's like hearing sounds and voices coming from your toys.

**12. Mixed Data Columns:**

Mixed data is like having a treasure box with different things inside. You might find toys, stickers, and crayons all in one place.

Remember, all these types of columns help computers understand and organize information, just like how you organize your toys, drawings, and stories to make sense of everything!