

I need to master the eda using python

Great! Mastering Exploratory Data Analysis (EDA) using Python is a valuable skill for any data analyst, engineer, or scientist. Here's a structured path to help you go from beginner to advanced.

#### 1. Fundamentals to Know First

Before diving into EDA, make sure you're comfortable with:

- Python basics: loops, functions, conditionals, list/dict/set
- Numpy: for numerical operations
- Pandas: for data manipulation
- Matplotlib / Seaborn: for visualizations

If you're rusty on any of these, I can provide crash guides or exercises.

# **2. Essential Libraries for EDA**

Library	Purpose
pandas	Data cleaning and manipulation
питру	Numerical operations
matplotlib	Basic plotting
seaborn	Statistical visualization
plotly	Interactive plots
missingno	Visualizing missing data
scipy.stats	Statistical analysis
scipy.stats	Statistical analysis

### 📶 5. EDA Process Dreakdown

### ✓ Step 1: Load and Inspect Data

```
import pandas as pd
df.head()
df.info()
df.describe()
df.columns
df.shape
```

## Step 2: Understand the Data

- Identify data types ( df.dtypes )
- Value counts: df['column'].value\_counts()
- Check for duplicates: df.duplicated().sum()

## Step 3: Handle Missing Data

```
df.dropna() # or df.fillna()
```

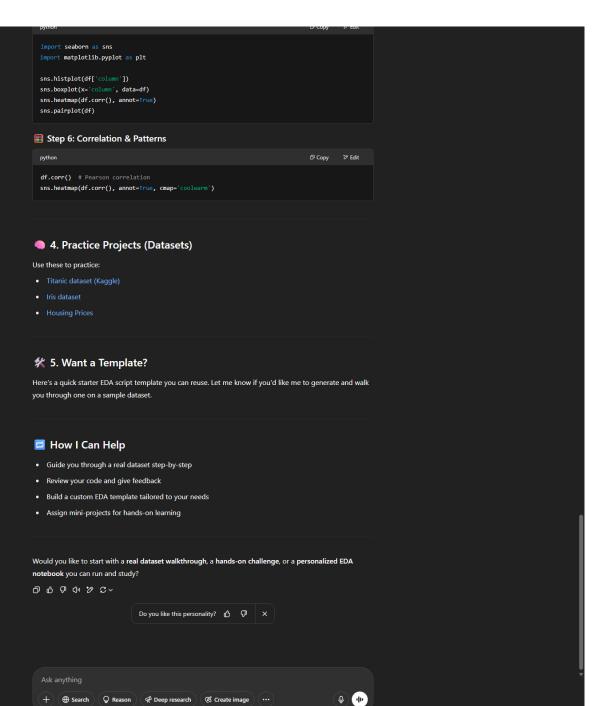
### Use missingno for visualizing:

```
import missingno as msno
msno.matrix(df)
```

## **▲** Step 4: Summary Stats and Distribution

```
python
df.describe()
df['column'].hist()
df['column'].value_counts().plot(kind='bar')
```

### Step 5: Data Visualization



ChatGPT can make mistakes. Check important info.