Introduction To Data Analysis







Why Should We Analyze Data?

- Understand what is happening in a dataset
- Discover trends, patterns, and outliers
- Make smarter, data-driven decisions
- Solve real-world problems using numbers
- Example: "Which NYC borough has the most expensive Airbnb listings?"

What Tools to Use

- Python easy-to-learn programming language
- Pandas library for working with tables (DataFrames)
- Matplotlib library for simple plots and charts
- Jupyter Notebook / Google Colab interactive coding environment

The Data Analysis Life Cycle

Step 1	Step 2	Step 3	Step 4	Step 5	Step 6
				9	848
Business Issue Understanding	Data Understanding	Data Preparation	Exploratory Analysis and	Validation	Visualization and Presentation
Define business objectives	Collect initial data	Gather data from	Modeling	Evaluate results	Communicate results
	Identify data requirements	multiple sources	Develop	Review process	
Gather required information		Cleanse methodology Determine	Determine	Determine best	
	Determine data availability			next steps	method to present insights based on analysis and audience
Determine appropri- ate analysis metod		Format	important variables	Results are valid → proceed to step 6	
	Explore data and characteristics	Blend	Build model		Craft a compelling story
Clarify scope of work		Sample	Assess model	Results are invalid revisit steps 1-4	
Identify deliverables					Make recommendations

Let's Analyze Some Data