Data Analysis and visualization Lab

I've prepared sample data that mirrors what executive education professionals typically work with. But if you have your own data : enrollment numbers, participant feedback, program evaluations, financial data : you can use that instead.
Link to the data : https://github.com/rihanamsaddek/CMED2025/blob/main/enrollment_data.csv

The Three Best Tools for Data Analysis

ChatGPT Plus with Code Interpreter Best overall for complex analysis

- Can execute Python code
- Creates visualizations
- Handles large datasets well

Claude

- Excellent for combining data analysis with written insights
- Great for creating narrative reports from data
- Strong at finding unexpected patterns

Google Gemini

- Best if your data lives in Google Sheets
- Native integration with Google Workspace
- Good for collaborative analysis

STARTER PROMPT #1: Initial Data Exploration

I've uploaded program enrollment data from an executive education center.

Before we do any analysis, please:

1. Show me the first 10 rows so I can see the structure

2. Tell me how many rows and columns there are

3. Identify any missing data or quality issues

4. Summarize what time period this data covers

5. List all the unique programs in the dataset

See what it's doing? It's:

- Reading the CSV
- Showing us the structure
- Identifying data quality issues (missing values, inconsistencies)
- Giving us context about what we're working with

This is like a research assistant doing the preliminary review before diving into analysis.

Common data issues it might find:

- Missing values in certain columns
- Inconsistent naming (e.g., 'Executive Leadership' vs 'Exec Leadership')
- Date formatting problems
- Outliers or suspicious values

And it tells you all of this proactively

Here's what's revolutionary about AI-powered data analysis: **You don't need to know Excel formulas, SQL, or Python. You just ask questions in plain English.**

Let me show you with progressively more complex questions.

SIMPLE QUESTIONS: Descriptive Statistics

STARTER PROMPT #2: Basic Summary Statistics

Give me summary statistics for the enrollment data:
- Total enrollments across all programs and years
- Average enrollments per program per year
- Which program has the highest total enrollment?
- Which program has the lowest?
- What's the overall trend - are enrollments growing or declining?

ADVANCED: Insight Discovery

STARTER PROMPT #3: Finding Hidden Patterns

Analyze this data for insights I might not have thought to look for:
1. Are there correlations between program type and enrollment trends?
2. Do certain regions show different patterns?
3. Is there a relationship between corporate vs. individual enrollments and program success?
4. Are there any anomalies or outliers I should investigate?
5. What's the most surprising finding in this data?

Be specific and explain what you found and why it matters.

This is where it gets really powerful. I didn't tell it what to find : I asked it to discover insights.

Look at what it's doing:

- *Running correlation analyses*
- *Identifying patterns across multiple variables*
- *Flagging anomalies*
- *Suggesting implications*

VISUALIZATION DEMO

PROMPT #4: Basic Visualizations

Create visualizations to help me present these enrollment trends:

1. A line chart showing total enrollments by year across all programs

2. A bar chart comparing total enrollments by program (all years combined)

3. A stacked bar chart showing enrollment by participant type (corporate vs individual) over time

4. A heat map showing enrollments by program and quarter to identify seasonal patterns

Make the charts clear, professional, and suitable for an executive presentation.

Common refinements you can ask for:

- Different colors or color schemes
- Larger fonts
- Different chart types ('make this a pie chart instead')
- Add trendlines
- Highlight specific data points
- Adjust axis scales
- Remove gridlines for cleaner look