AI-Powered Data Analysis: From Installation to Webapp

Building Your First Data Analysis Webapp with AI

Ka Wai Tsang

Data Science Program

September 25, 2025

Course Outline

1. Introduction
2. Part 1: Environment Setup
3. Part 2: The Data Analysis Mindset
4. Part 3: Building Your Analysis Webapp
5. Part 4: Data Storytelling and Presentation
6. Part 5: AI Prompting Best Practices

What You'll Build: Your First Data Analysis Webapp

Course Focus: Hands-On Webapp Development

- **Real Skills**: Build an actual webapp you can share and use
- **AI-Powered**: Learn to work with AI as your coding partner
- **Practical**: Create something that demonstrates your abilities
- **Extensible**: Start simple, add features as you learn

What You'll Learn

Technical Skills:

- Setting up development environment
- Using AI coding assistants effectively
- Building React web applications
- Data visualization and analysis
- Deploying web applications online

Data Analysis Skills:

- Systematic data exploration
- Statistical analysis methods
- Data storytelling and presentation
- AI-guided analytical thinking
- Professional data reporting

Step 1: Installing Cursor AI Editor

Why Cursor?

- **AI-First**: Built specifically for AI-assisted coding
- **Student-Friendly**: Designed for learning and experimentation
- **Powerful**: Handles complex projects while remaining accessible
- **Modern**: Uses the latest AI models for code generation

Download and Install

1. Visit: `https://cursor.sh/`
2. Download for your operating system
3. Run installer as Administrator
4. Launch and sign in with GitHub account

Step 2: Python Environment Setup

Python Installation

- Download Python 3.9+ from `python.org`
- **Critical**: Check "Add Python to PATH" during installation
- Verify: Open terminal and type `python --version`

Package Installation (China-Optimized)

Recommended for China: Set Tsinghua Mirror

- `pip config set global.index-url https://pypi.tuna.tsinghua.edu.cn/simple/`
- `pip install pandas numpy matplotlib plotly streamlit jupyter`

Step 3: Project Setup

Create Your First Project

1. Create folder: `C:/Users/[YourName]/Desktop/Demo1`
2. Add your data file (CSV, PDF, etc.)
3. Open folder in Cursor: `File → Open Folder`
4. Test AI: Ask "Create a webapp to present the content from my document"

Alternative Project

1. Create folder: `C:/Users/[YourName]/Desktop/Demo2`
2. Copy `UM_C19_2021.csv` to this folder
3. Open in Cursor and start analyzing!

Before We Code: The Right Way to Analyze Data

Data Analysis is Like Being a Detective

1. **Survey the Scene**: Understand what data you have
2. **Look for Clues**: Identify patterns and issues
3. **Follow Evidence**: Analyze relationships systematically
4. **Draw Conclusions**: Interpret findings meaningfully

The Golden Rule

Never start analyzing data without first understanding what you're working with!

Step 1: Data Understanding with AI

AI Prompt for Data Overview

```
"I have a new dataset called 'UM_C19_2021.csv'. Before I analyze it, help
me understand what I'm working with. What questions should I ask about
this data first?"
```

What AI Should Help You Discover

- What does this dataset represent?
- How many records and columns?
- What are the column names and meanings?
- What time period does it cover?
- What is the main purpose?

Step 2: Data Quality Assessment

AI Prompt for Quality Check

```
"I want to check if my data has any quality issues. How should I approach
looking for missing values, duplicates, or other problems?"
```

What to Investigate

- Missing values and their patterns
- Duplicate records
- Data validation (do values make sense?)
- Date ranges and logical consistency

From Analysis to Webapp: The Natural Progression

Why Build a Webapp?

- **Interactive Learning**: Change parameters and see results instantly
- **Visual Understanding**: Charts and graphs make concepts clearer
- **Portfolio Piece**: Something tangible to show your skills
- **Real-World Skills**: Modern data analysis happens in applications

Your Webapp Will Include

- Data upload and preview
- Interactive statistical analysis
- Beautiful visualizations
- Professional presentation of results
- Extensible architecture for new features

Your First Webapp: Simple and Achievable

Start with Something Basic

```
"I want to create a simple webapp that uploads a CSV file and shows basic
statistics and a chart.  Can you help me build this with Streamlit?"
```

What This Creates

- **File Upload**: Simple CSV file upload
- **Data Preview**: Show first few rows of data
- **Basic Statistics**: Mean, median, standard deviation
- **Simple Chart**: One visualization (like a histogram)
- **Clean Interface**: Easy to understand and use

The Art of Data Storytelling

Data Presentation is Like Being a Storyteller

1. **Know Your Audience**: Who are you presenting to?
2. **Structure Your Narrative**: Beginning, middle, and end
3. **Choose Right Evidence**: Which data supports your story?
4. **Make It Memorable**: How will audience remember key points?

The Golden Rule

Your data should tell a story, not just show numbers!

Planning Your Data Presentation

AI Prompt for Audience Analysis

```
"I want to present my COVID-19 data analysis findings. Help me think about
who my audience might be and what they would care about most."
```

Consider These Questions

- Who needs to see this data? (Administrators, public health officials, students?)
- How familiar is your audience with data analysis?
- What aspects would be most relevant to them?
- What decisions might they make based on your presentation?

Choosing the Right Visualizations

AI Prompt for Chart Selection

```
"I want to show different aspects of my COVID-19 data. Help me think about
what types of charts would best communicate each key message."
```

Match Charts to Messages

- **Trends over time**: Line charts, area charts
- **Comparisons between groups**: Bar charts, grouped bar charts
- **Relationships**: Scatter plots, correlation heatmaps
- **Distributions**: Histograms, box plots

Effective AI Prompting for Data Analysis

Prompt Structure That Works

1. **Start with Understanding**: "Help me understand what I'm looking at..."
2. **Ask for Guidance**: "What should I consider when..."
3. **Request Systematic Approach**: "How should I approach..."
4. **Seek Interpretation Help**: "What does this tell me about..."

Prompt Examples for Each Phase

Data Understanding

- "What questions should I ask about this dataset first?"
- "How can I assess the quality of my data?"
- "What should I look for when examining the structure?"

Data Cleaning

- "How should I approach handling missing values?"
- "What's the best way to identify outliers?"
- "How do I decide whether to remove or fix data issues?"

Remember: The Goal is Learning, Not Perfection

Key Success Principles

- **Start Simple**: Get basic functionality working first
- **Iterate Frequently**: Make small improvements regularly
- **Learn from Mistakes**: Every error is a learning opportunity
- **Ask for Help**: Use AI, communities, and classmates
- **Have Fun**: Enjoy the process of building something new

Your Journey Starts Now

You now have all the tools and knowledge to build your first data analysis webapp. The only thing left is to start coding!

Thank You and Good Luck!

Questions?

Remember:

- Use AI as your coding partner
- Start with understanding your data
- Build iteratively and systematically
- Focus on creating something you can share
- Have fun learning!

Now go build something amazing!