AI Tools for Data Analytics

Activity 1: Identify Tasks for AI

1. What aspects of your current data workflow are the most challenging or problematic?
2. Which data tasks consume the most time and resources in your current workflow?
3. Are there repetitive tasks that you find yourself doing over and over again?
4. Which data tasks do you enjoy the least?
5. Are there specific tasks within your workflow that often result in errors or inconsistencies?

Activity 2: Prep Data with ChatGPT

If you DO NOT have the paid version of ChatGPT, you have two options for this activity. You can either complete the activity below using the free ChatGPT 3.5, or you can complete the evaluation checklist I had provided for you to evaluate if ChatGPT 4 is worth using for your particular job and use cases.

Prompts for you to write into ChatGPT will appear *italicized like this*.

1. Sign up for a ChatGPT account here if you don't already have one. You can use the free version if you want.
2. Download data from the course's Github repo called "Uncleaned_DS_Jobs.csv". This file has data science related job postings.
3. Start a new chat on ChatGPT. If you have a paid version of ChatGPT, make sure to use ChatGPT 4 in this chat.
4. In the new chat, attach your downloaded data. Don't submit the chat yet.
   a. If you don't have the paid plan for ChatGPT, copy and paste 20 rows or so into the chat box and preface it to let ChatGPT know you are pasting data here.
5. In the chat box, provide this prompt:

   Your role as a data analyst is to clean my dirty data set using accepted practices and strategies for cleaning data.

   First, I want you to review and identify issues in the data set. Identify any common data quality issues within the data set like missing values, duplicates, and outliers.

6. Review what it identifies and then prompt it:

> *Give me a proposed plan to clean the data.*

7. Review the proposed plan. If it sounds reasonable to you, prompt it:

   > *Go ahead and execute this plan.*

8. There's a lot of useful info in the job description that can be extracted out, like the skills they list, the seniority level, etc. Prompt ChatGPT to:
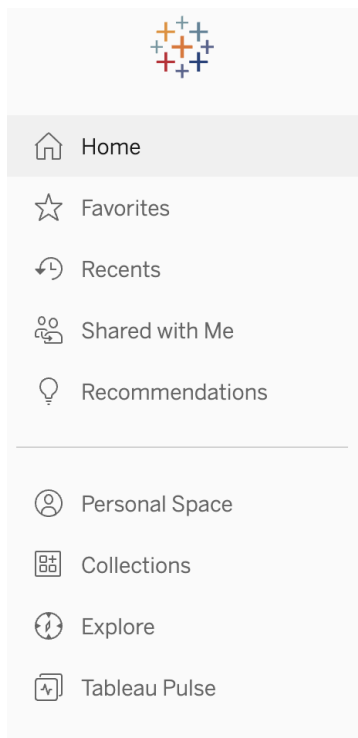
   > *I want you to extract skills, tools, and seniority levels from the job descriptions and other relevant columns to create new columns to represent these values. Each skill can be its own column.*

9. Optional: do some exploratory data analysis. You can use [prompts from here](#) for inspiration.
10. Now that you're done, you can ask ChatGPT to return you the file it prepared:

    > *Please attach a csv with the cleaned data.*

Activity 3

1. If you don't already have Tableau Cloud, you can start a free 2 weeks trial [here](#).
2. Log in to [Tableau Cloud](#)
3. On the left hand side menu, select "Tableau Pulse" (you'll find it below "Explore")

4. Click on the "New Metrics Definition" button on the right hand side.

5. Select "Superstore Datasource" as the source
6. Name this metric "Profits by Category"
7. Under measure, select the "Profit" measure
8. Keep the default aggregation to Sum, and sparkline values as Running total

Definition	Insights

Data Source

🗄 **Superstore Datasource**

Name *

Profits by Category

Description

Tracks profits and shows breakdown by category.

Value   **Clear Selections**

Measure *

⊞ Profit   ⌄

Aggregation

Sum   ⌄

Show sparkline values to date as ⓘ

Running total   ⌄

9. Under Time dimension, select "Order Date"
10. Under the Options section, add these adjustable filters and any other filters you would like to see profit breakdown by::
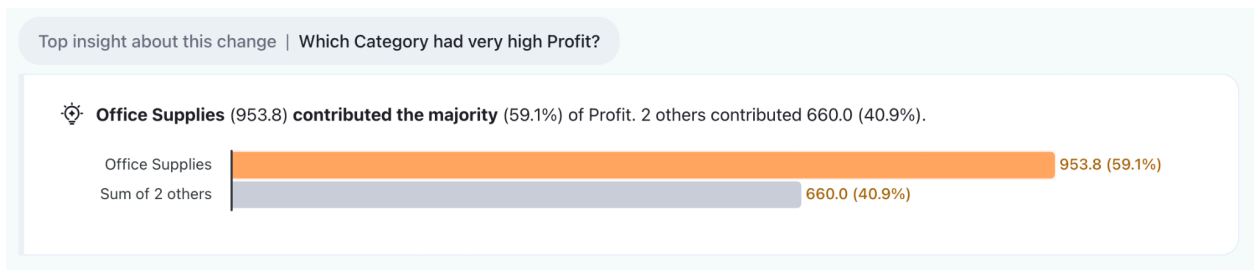
Options

Adjustable metric filters ⓘ *

Category	✕
Product Name	✕
Country/Region	✕

11. Click "Save Definition"
12. You should be redirected to your metrics page showing profits. Play around with the adjustable filters and suggested exploratory questions on your page.
13. You'll see options for questions you can explore and create visualizations for. Select the one for "Which category had high profit?"

Top insight about this change | Which Category had very high Profit?

Office Supplies (953.8) **contributed the majority** (59.1%) of Profit. 2 others contributed 660.0 (40.9%).

| Office Supplies | | 953.8 (59.1%) |
| Sum of 2 others | | 660.0 (40.9%) |

Resources

[AI Tool Evaluation Checklist](#)

[The chat I demoed on ChatGPT](#)

BONUS: Here's [my custom GPT](#) to help you prepare for your next data presentation.

Adam Mico made this cool custom GPT called [Tableau Doctor](#).

Great ideas for [data analysis prompts](#)

[Data science prompts](#)

If you want more practice with ChatGPT for data analytics, [here's Luke's course and resources](#).

If you want to try out Alteryx's cloud platform for 2 weeks, you can start a trial [here](#).