

## Learning PowerBI: Struggles from Scratch

Using PowerBI turned out to be more complicated than I expected.

### 1. Data Acquisition

I initially wanted to scrape data using Python, but the results were unusable when imported into PowerBI. I tried AI-assisted scraping with similar results. Eventually, I used data from government websites and public databases, which took more time but provided reliable quality.

### 2. Data Cleaning: The Challenge of Dirty Data

The data had many issues - missing values, duplicates, formatting errors, and incorrect data types. Many "(blank)" values appeared in reports unexpectedly. I learned to use Power Query for cleaning: filling missing values, removing duplicates, and standardizing formats. This process showed me that 80% of data analysis involves cleaning data, while visualization is comparatively easier.

### 3. Visualization: From Chaos to Clarity

My initial reports were confusing due to unclear titles, illogical charts, and distracting colors. I improved by first defining the analysis goal, like "income disparities across ethnic groups" for population reports or "changes in energy usage" for energy reports. Clear goals made chart creation simpler.

This project taught me that data analysis requires telling a clear story, not just presenting numbers. For example, noticing lower incomes in Latino communities led to questions about education and job opportunities.

Areas for improvement:

1. Need to learn more about DAX formulas and advanced modeling
2. Limited by available data sources
3. Report design could be more professional in color schemes and layout