Arham Ullah Khan
Professor Jordan Wirfs-Brock
CS-215

My Data Journey: A Personal Manifesto

Data? Data is more than just numbers. For me, data isn't just a collection of numbers or facts. It's like an unexplored world filled with hidden gems waiting to be discovered. Reflecting on the DIKW pyramid, data is the raw material, like the rough drafts in a file cabinet metaphor. But once sorted, labeled, and understood, it transforms into a story with meaning and purpose. Like turning a chaotic photo album into a beautifully narrated travelogue, data becomes valuable when given context and narrative. And a data scientist? That's a modern-day explorer. It's not just about crunching numbers; it's about asking "why" and "how," diving deep into the ocean of data to uncover truths and trends. This course has shown me that it's as much about the journey – the process of discovery and learning – as it is about the destination or the final analysis.

Data work is a balancing act. Sure, I've had to get cozy with Python and wrestle with SQL queries, but that's just the start. It's also about thinking critically, like a detective piecing together clues, and communicating findings like a storyteller. Thinking about all the projects, especially the EV prevalence nutrition label: It wasn't just about the analysis; it was about telling its story and making it matter. Hence, my advice to the others comes in the form of lessons from the trenches:

- Stay curious. Every dataset has a story.
- Get your basics right – stats and coding are your foundation. Consider expanding your knowledge to other fields and courses like Statistics and Machine Learning
- Ethics aren't optional. Consider the impact of your work.
- Never stop learning. The field's always evolving, and so should you.

Data's great at answering "what" and "how." It's like a flashlight shining on patterns and trends. But sometimes, it struggles with the "why" – the deep, human part of the story. That's where you need more than just data; you need empathy and understanding. Reflecting on the past, I think about my curiosity, my concern for ethical issues, and my belief that data can be a force for

good. It's not just about the analysis; it's about understanding our world a little better and maybe making it a bit better too. What if a new project comes my way in the future? We'll consider this: Let's say I'm tackling a project on the impact of urban green spaces on mental health. Here's my game plan:

- Collect data responsibly, respecting privacy.
- Analyze for trends – how do green spaces correlate with mental well-being?
- Share insights in a meaningful way, not just statistical jargon.
- Think about what this means for urban planning and public health.

It is necessary to think not only about the code and statistical analysis, but about impacts, outcomes, applications, and ethical concerns. Data is inherently biased, and going in with that knowledge helps make our work all the more ethical. To me, data science is a canvas where numbers meet narratives. It's about finding the story in the spreadsheet, the human behind the histogram. This class has been a journey – one that's shown me the power of data, the importance of asking the right questions, and the responsibility that comes with this knowledge. I am looking forward to continuing this journey in my career hopefully.