Final Project | Personal Data Manifesto & Data Cookbook

Data surrounds our digital lives and has a massive influence on choices, regulations, and creativity. However, facts alone do not constitute truth. It is gathered, organized, and analyzed by humans. My data principle is based on the concept that data should be treated as narratives, with information, diligence, and an individual's viewpoint. Inspired by Giorgia Lupi's Data Humanism, this manifesto explains my four basic principles for dealing with data and discusses the ethics, problems, and power of data in today's world.

Principle 1: Data Is Not Neutral

Over the past few weeks, we have discussed in the Data Science class many examples, such as Matthew Connelly's Declassification Engine, which exploited data to show patterns in government visibility. This demonstrated that data can support or undermine power hierarchies, based on how it is used.

The Key: As a data scientist, I constantly ask, "Who created this data?" Who is this for? Which voices are missing? Context, bias, and effect must all be considered when gathering, analyzing, and visualizing data.

Principle 2: Data Science = Technical + Ethical Thinking

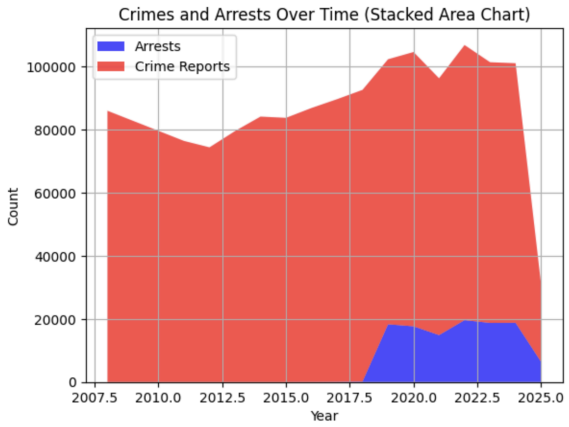
One of the most interesting things I learned this semester was that critical thinking is equally vital as coding. When I was studying arrest data and displaying crime trends (such as the stacked area chart I created), I realized how false graphics could influence public opinion.

Implementing skills include Python, Pandas, Matplotlib, and SQL; ethical literacy and skepticism are required for interpretation.

Code Snippet:

```
plt.stackplot(df['Year'], df['Crime Reports'], df['Arrests'], labels=['Crime Reports', 'Arrests'],
colors=['red', 'blue'])
plt.legend(loc='upper left')
plt.title('Crimes and Arrests Over Time')
```

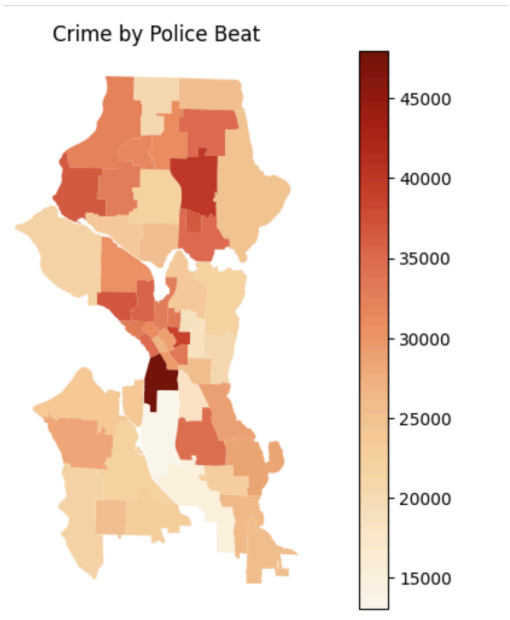
Visual:



Principle 3: Data Is a Process, Not Just an Outcome

One of my class assignments, such as mapping crime reports or evaluating COVID patterns, taught me that data processing is iterative. The work includes cleaning messy datasets, managing missing numbers, and re-evaluating visuals.

Advice for new data scientists: acknowledge uncertainty. Data research is rarely a linear process, and sometimes the most interesting ideas emerge from error or unusual trends.



Crime Trend in Seattle (2008-2025):

Principle 4: Data Should Be Communicative, Not Just Correct

Some of the most difficult issues I had were ensuring that my visualizations conveyed knowledge rather than simply looking "cool." After observing students present projects on themes such as European League clubs, flood crisis, food, and stock market crises, I realized the importance of strong storytelling using charts.

Inspiration: Giorgia Lupi's focus on personal data stories and how simple hand-drawn data can be impactful inspired me to put people over appearances.

Conclusion:

Data science is more than just knowing technologies; it's also about ethically applying those techniques to reveal figures, challenge systems, and share discoveries. As I continue on my data journey, these four concepts guide me: data is not neutral, data science is both technical and trustworthy, data work is iterative, and communication is more important than complexity.