Summer Overview

Website: https://github.com/tylerharter/sum19-city

Presenting Results:

- CS 301 focus: write code to create some plots
- This summer, we'll do more around the plots
  - what is the question?
  - present plot
  - highlight key observations + interpret
- Emphasis on narrative/argument
  - organize plots hierarchically and sequentially to tell a story
  - explicitly state our conclusions
- The ultimate goal is to discover actionable things. Nobody will act unless you communicate in a clear and compelling way.

Format

- Each team chooses their own topic (related to City of Madison or USGS). This is very open-ended; as long as you produce interesting results, you have a lot of latitude to explore what interests you
- Concurrently write both a 2-3 and 8-10 page report on that topic
- Every week, add plots/text to your working document. Use grayscale, large font, labels, small numbers (choose units carefully).
- Must write code to generate plots, preferably in Python (no Excel!)
- Show to the group (everybody should participate in feedback -- no cells or other work!)
- Meetings mostly driven by you! I'll occasionally demo various tools as needed.

Example topics

- https://github.com/tylerharter/s19-city/blob/master/final-presentations/README.md
- Where should bus stations/routes be added?
- Where should fire stations be created?
- How does weather/climate affect spending?
- What kind of variation can be expected for salaries?
- How might the property tax base grow/shrink over time?
- Deep dive on specific agency?

Data

- https://github.com/tylerharter/sum19-city/tree/master/data
- http://data-cityofmadison.opendata.arcgis.com/
- other?

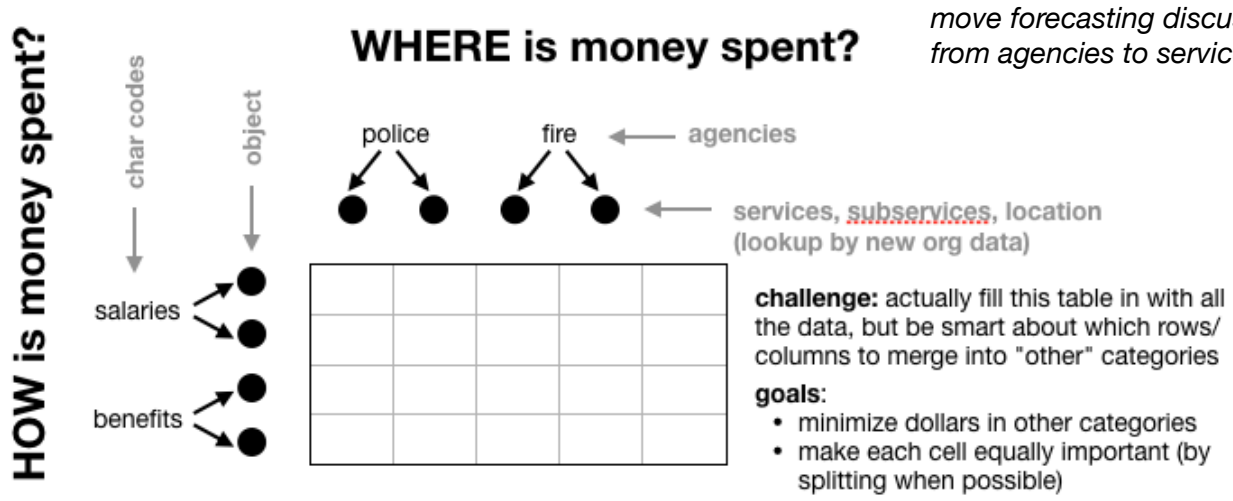
Introductions

- Name? Year in school? Prior related experience? Topics of interest?

Report format

- We'll use Latex (you write Latex "code", then use pdflatex to generate a PDF)
- Install Windows: https://miktex.org/, Install Mac: http://www.tug.org/mactex/
- Latex does typesetting, chooses layout for you
- Images will be in .eps format

Example: how much did each agency spend in 2018?



move forecasting discussion from agencies to services

Data: spend.csv

GROUP BY in Pandas: `df.groupby("COL1").sum()["COL2"]`

Grayscale: `DATA.plot.bar(colormap="gray")`

EPS: `ax.get_figure().savefig("name.eps", bbox_inches="tight")`

Add to example.tex file:

```
\begin{figure}[t]
  \includegraphics[width=\columnwidth]{name.eps}
  \caption{{\bf Title.}  Details.}
\end{figure}
```

example.tex to example.pdf: `pdflatex example.tex` (run in terminal)

Next Week

- Bring a couple plots in a PDF generated by Latex
- Think about forming teams
- Identify a 2-3 topics you might want to pursue; list as many interesting questions as you can for each topic
- **Optional** (for those interested): get Python Machine Learning by Sebastian Raschka (consider using sklearn for your projects)