

For Part 1 you will create a set of slides that you will present in class and that will be saved as a .PDF for the Part 1 submission with the following information (be brief):

1. Title: Crime and Health in Denver
2. Team members:
 - Naureen Bharwani
 - Cody Thornton
 - Sean Mulligan
3. Description: 2-3 one sentence project descriptions - what interesting questions do you intend to answer? Crime Related Questions:

- Has Denver's Star Program had any effect on the data?
- Covid has increased crime nationwide. Within Denver, has this increase in crime changed in any way from previous crime data patterns? For example are their changes in the relational attribute/object patterns in the types of crimes, their locations, their time of occurrence
- Does living in a specific neighborhood that has the most crime affect life expectancy? Does this correlate to ethnicity/race?

Hate Crime Related Questions:

- Which kinds of hate crimes are most and least common?
- What group(race/ethnicity/religion/etc.) receive the most hate crimes against them? Which receives the least?
- Does a specific year/month/location correlate with the amounts of hate crimes being reported?

Health Related Questions:

- Does Life Expectancy have positive or negative correlation with crime and/or a specific neighborhood?
- Which neighborhoods have the most obese population? Which neighborhoods have the least obese population?
- How does the obesity of a neighborhood correlate to the crime being reported? Is there more crime or less crime? Does this increase the hate crime being reported? What ethnicity/race is most common in these neighborhoods?

4. Prior Work: What prior work has been done on your idea
 - The City of Denver has put together this compendium of datasets - <https://www.denvergov.org/opendata/>
 - Searched for previous studies on health and crime in Denver

Crime Related Datasets:

1. Colorado Crime Statistics - <https://coloradocrimestats.state.co.us/tops/>

2. Denver Crime Data - <https://www.kaggle.com/paultimothymooney/denver-crime-data/version/25>
3. Previous Student Analysis of crime in Denver - <https://data.world/stoltzmaniac/raw-crime-data-for-denver-colorado>
4. Specific Crime Data - <https://data.world/denver/marijuana-related-crime>
5. Previous Student Analysis of crime in Denver - <https://github.com/the4thparadigm/denvercrime>

Health Related Datasets:

1. Global Life Expectancy - <https://www.kaggle.com/kumaraarshi/life-expectancy-who>

5. Datasets:

a. List of datasets to use:

- Descriptions below in part (a) will match to links found in part (b)
- Datasets 1 - 4 (crime related datasets) will be merged into a mega crime dataset, dataset 5 will be used for supporting primary demographics, datasets 6 - 8 (comprising of health datasets) will be used as individual points of comparison, and dataset 9 (real estate datasets) will be used for supporting . . .

1. Crime in the City of Denver
 2. Hate Crimes in the City of Denver
 3. Traffic Accidents in the City of Denver
 4. Police Pedestrian Stops and Vehicle Stops in the City of Denver
 5. American Community Survey Nbrhd in the City of Denver (2013-2017)
 6. Life Expectancy 2010 - 2015 in the City of Denver
 7. Adult Obesity 2014-2016 in the City of Denver
 8. Adult Diabetes 2010-2016 in the City of Denver
 9. Real Property Sales and Transfers in the City of Denver
- b. Where found (URL and who is supplying the data, e.g., NASA) All of the datasets have been sourced the City of the Denver
1. <https://www.denvergov.org/opendata/dataset/city-and-county-of-denver-crime>
 2. <https://www.denvergov.org/opendata/dataset/hate-crimes>
 3. <https://www.denvergov.org/opendata/dataset/city-and-county-of-denver-traffic-accidents>
 4. <https://www.denvergov.org/opendata/dataset/city-and-county-of-denver-police-pedestrian-stops-and-vehicle-stops>
 5. <https://www.denvergov.org/opendata/dataset/city-and-county-of-denver-american-community-survey-nbrhd-2013-2017>
 6. <https://www.denvergov.org/opendata/dataset/city-and-county-of-denver-life-expectancy-2010-2015>

7. <https://www.denvergov.org/opendata/dataset/city-and-county-of-denver-adult-obesity-2014-2016>
8. <https://www.denvergov.org/opendata/dataset/city-and-county-of-denver-adult-diabetes-2010-2016>
9. <https://www.denvergov.org/opendata/dataset/city-and-county-of-denver-real-property-sales-and-transfers>
- c. Whether you'll have it downloaded (on who's machine)
 - We will utilize Github for version control. As such, all datasets will be simultaneously available to all team members via a shared repository. To maintain the repository and prevent merge conflicts, we will utilize the Kanban style app Trello for project management. Our base file type will be comma separated value(CSV) files.
6. Proposed work/Data Preprocessing: what do you need to do?
 1. Data Cleaning:
 - Review all datasets for simple data comparison to create sets of sub-domains(attribute values).
 - Verify data quality (accuracy, completeness, consistency, timeliness, believability, interpretability)
 - Data cleaning routines work to “clean” the data by filling in missing values, smoothing noisy data, identifying or removing outliers, and resolving inconsistencies.
 - Note: As the datasets themselves appear to be of high quality, much of the cleaning effort will be done in preparation for the integration phase.
 2. Data integration:
 - Entity Identification Problem
 - Redundancy and Correlation Analysis
 - Tuple Duplication
 - Data Reduction
7. List of tool(s) you intend to use
 - a. NumPy
 - b. Panadas
 - c. matplotlib
 - d. GitHub
 - e. Trello
 - f. Tableau
8. Evaluation: How you can evaluate your results
 -
 -
 -