**GitHub Repo:** https://github.com/csadlo/GunOwnersAndCrime

|  |  |
| --- | --- |
| **Project Title** | **Gun Ownership and Crime**  **Is there a correlation between gun ownership and crime?** |
| **Team Members** | Matt Dooley, Glenda Decapia, Chris Sadlo, Carrie O’Connor-Walker |
| **Project Proposal/ Outline** | We will compare two datasets, violent crime statistics by state and gun ownership by state |
| **Research Questions**  **to Answer** | * Is there a relationship between murder and non-negligent manslaughter and gun ownership? * Is there a relationship between gun ownership and aggravated assault? * Is there a relationship between robbery and gun ownership? * Is there a relationship between motor vehicle theft and gun ownership? * Is there a relationship between burglary and gun ownership?   For each question, make a scatterplot (or other visualization), with each datapoint being a state:  x-axis would be gun ownership per capita  y-axis would be types of crime  For each question, do a brief summary of the correlation or lack of correlation that you observe.  ***Possible visualizations:***  Scatterplot/Linear Regression--Matt  Overlaying heatmaps, gun ownership/types of crime of crime-- Carrie  [Bubble chart?](https://python-graph-gallery.com/bubble-plot/) Would need to think of a third variable-- Glenda  Null Hypothesis Testing -- Chris |
| **Datasets to be used** | **Violent Crime:** [**https://www.ucrdatatool.gov/Search/Crime/State/StatebyState.cfm**](https://www.ucrdatatool.gov/Search/Crime/State/StatebyState.cfm)  **Gun Ownership:**[**https://www.kaggle.com/solorzano/gun-ownership-by-state**](https://www.kaggle.com/solorzano/gun-ownership-by-state) |
| **Rough breakdown of tasks** | **Rough Breakdown of Tasks**  **Chris puts data into excel workbook**  **Use the jupyter notebook from github repo and add your own code in your individual branch**  **Cross-reference crime data from the ucrdatatool with data from an independent source**  **Use matplotlib to create 2 visualizations per question**  **\*Include at least one linear regression**  ***Optional--use an API to find more data***  **Once a process is developed, process more states.**  **Create a write-up for each question** |
|  | *To review workflow, pull/push/merge etc:* [*Link*](https://umn.bootcampcontent.com/University-of-Minnesota-Boot-Camp/uofm-stp-data-pt-06-2020-u-c/blob/master/03-Projects/Git-Jupyter-Notebook-Workflow.md) |

|  |
| --- |
| **Project 1**  **Research Questions to answer:**  **Ideas:**   1. **Have well child checks decreased during COVID-19 outbreak?** 2. **Have ER visits decreased during COVID outbreak?** 3. **Have annual checkups decreased during COVID outbreak** 4. **If so, are decreases different or consistent across genders? Race? Socioeconomic status?**   **Datasets to be used:**  [**https://www.ahrq.gov/data/hcup/index.html**](https://www.ahrq.gov/data/hcup/index.html)   * [**https://www.hcup-us.ahrq.gov/nisoverview.jsp**](https://www.hcup-us.ahrq.gov/nisoverview.jsp)   **ER Visit DATA CDC:** [**https://www.cdc.gov/nchs/hus/contents2018.htm?search=Emergency\_department\_visits,**](https://www.cdc.gov/nchs/hus/contents2018.htm?search=Emergency_department_visits,)  [**https://www.statista.com/statistics/590346/us-hospitals-most-emergency-room-visits/**](https://www.statista.com/statistics/590346/us-hospitals-most-emergency-room-visits/)  [**https://www.reddit.com/r/dataisbeautiful/comments/hso3sf/oc\_blue\_counties\_in\_red\_states\_are\_getting\_hit/**](https://www.reddit.com/r/dataisbeautiful/comments/hso3sf/oc_blue_counties_in_red_states_are_getting_hit/)  [**https://www.reddit.com/r/dataisbeautiful/comments/hso3sf/oc\_blue\_counties\_in\_red\_states\_are\_getting\_hit/fycfwou/**](https://www.reddit.com/r/dataisbeautiful/comments/hso3sf/oc_blue_counties_in_red_states_are_getting_hit/fycfwou/)  **Rough breakdown of tasks:** |
| **Project 2**  **Research Question**  **Is there a correlation between the banning of leaded gasoline and a drop in violent crime?**  **Is there a correlation between the banning of leaded paint and a drop in violent crime?**  **How about leaded ammunition against waterfowl bans and the effect on the water tables?**  **Can any correlation be seen at the national level? Was there a ban (gasoline, paint etc.) at the federal level at one point? Is the correlation stronger on a state-by-state basis?**  **By looking across all state datasets, can we use a sudden drop in violent crime across many states to predict (extrapolate backwards in time) when a federal ban might have gone into effect? How accurate are we in that prediction?**  **What other items USED to contain lead that we are unaware of?**  **I believe burning coal releases lead dust...**  **Uniform Crime Reporting Statistics**  [***https://www.ucrdatatool.gov/Search/Crime/State/StatebyState.cfm***](https://www.ucrdatatool.gov/Search/Crime/State/StatebyState.cfm)  ***US Childhood Lead Surveillance Data by State***  [***https://www.cdc.gov/nceh/lead/data/state.htm***](https://www.cdc.gov/nceh/lead/data/state.htm)  [***https://www.cdc.gov/nceh/lead/data/national.htm***](https://www.cdc.gov/nceh/lead/data/national.htm)  ***Gun Ownership by State***  [***https://www.kaggle.com/solorzano/gun-ownership-by-state***](https://www.kaggle.com/solorzano/gun-ownership-by-state)  ***Alternate Source to cross-reference***  [***http://www.disastercenter.com/crime/uscrime.htm***](http://www.disastercenter.com/crime/uscrime.htm)  ***Site to find when certain bills were passed in different states***  [***https://www.ncsl.org/research/telecommunications-and-information-technology/ncsl-50-state-searchable-bill-tracking-databases.aspx***](https://www.ncsl.org/research/telecommunications-and-information-technology/ncsl-50-state-searchable-bill-tracking-databases.aspx)  [**https://pubmed.ncbi.nlm.nih.gov/24076506/**](https://pubmed.ncbi.nlm.nih.gov/24076506/)  [**https://en.wikipedia.org/wiki/Lead%E2%80%93crime\_hypothesis**](https://en.wikipedia.org/wiki/Lead%E2%80%93crime_hypothesis)  **Rough Breakdown of Tasks**  **Create a Jupyter Notebook titled “Cleanup”**  **Import necessary libraries and API keys**  **Partition code into sections w/team members names**  **Pick a single state to focus on and its dataset.**  **Save the data into an unaltered csv file and save into the github repository. Mark it as unclean.**  **Use pandas to load, clean reformat if needed, and save data to a csv in one Jupyter Notebook. Save the ‘clean data’ under an appropriate title in the github repository.**  **Create a new Jupter Notebook called “Read/Analyze” to read and analyze the clean data**  **Partition code into sections w/team members names**  **Cross-reference crime data from the ucrdatatool with data from an independent source**  **Use matplotlib to create 2 visualizations per question**  **\*Include at least one linear regression**  ***Optional--use an API to find more data***  **Once a process is developed, process more states.**  **Create a write-up for each question** |
|  |