Data Analytics Bootcamp – Project #3

**Group Members:** Cristina Bardan, Antonio Aramoni, Daniel Orta, Purvi Patel

**TOPIC**: Analysis of sentencing data from Cook County State’s Attorney’s Office (CCSAO)

**DATA SETS**: <https://datacatalog.cookcountyil.gov/Courts/Sentencing/tg8v-tm6u>

# Proposal: Topic & Rational

In early 2018, the State’s Attorney (judicial prosecutor’s office) in Cook County, Illinois – where the city of Chicago is located – made public a series of unfiltered datasets on all felony criminal cases processed in Cook County, as part of an initiative to increase transparency in the criminal justice system. The case-level datasets contain anonymized information about every felony case processed by the State’s Attorney’s Office (SAO) dating back to roughly 2010, and have been provided for free so that any person or group can analyze the data and share the results publicly.

There are 4 datasets, representing the four stages of interaction with the SAO during the legal process:

1. Intake – initial investigations, potential cases for prosecution
2. Initiation – start of the legal process for cases that the SAO decides to prosecute
3. Disposition – the results of the fact-finding process and final outcome of the case
4. Sentencing – penalties imposed on cases found “guilty”

For this project, the group proposes to focus on analysis of ONE of the datasets (***Sentencing***), with a focus on specific topics related to sentencing outcomes, such as:

* A demographic analysis of persons found “guilty” of a felony charge (i.e. age, race, gender)
* An analysis of the types of sentences imposed, possibly disaggregated by type of offense
* An evaluation of the total time needed in the judicial process to arrive at the final sentence
* An evaluation of the number of guilty findings and types of sentences grouped by the court or district where the case was processed

The project will take advantage of an already-organized data set, publicly available for free in both downloadable format (csv) and through an API url. Since there is a strong interest by the SAO in Cook County to encourage the public to analyze the data, and since the SAO does not have the resources to do this analysis directly, this project would also present an opportunity to provide a public service by sharing the final product.

**PROPOSAL: Process (NEED TO CLARIFY)**

1. The initial analysis and data transformation will be done with Python coding in Pandas, with the original data loaded into Pandas via the API for the Sentencing dataset.
2. Once data analysis has been completed, the final dataframe will be loaded into an SQL database.
3. The cleaned data will be made available from the SQL database, accessed through Flask (RestAPI)
4. **Running queries in SQL??? Or is all of that going to be done in Pandas?**
5. Javascript code, and use of the \_\_\_\_ libraries, will be used to generate data visualizations, with the final dashboard published publicly via GitHub.

**GitHub repository:** <https://github.com/cckuqui/CCSAO-Sentencing>

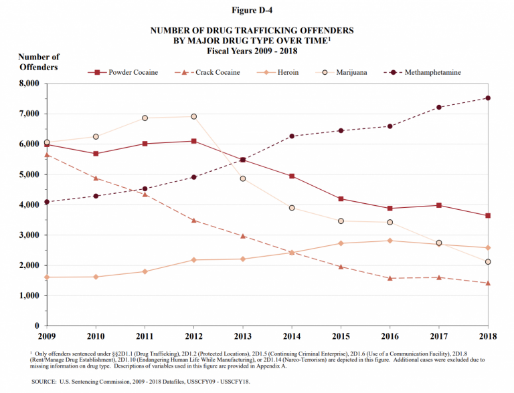
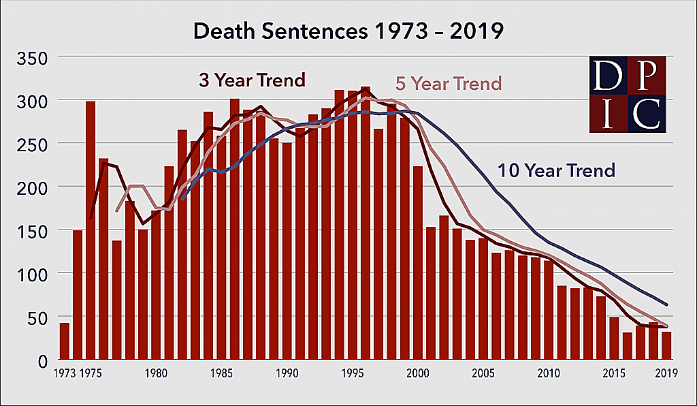
# **Inspiration for Project Design**

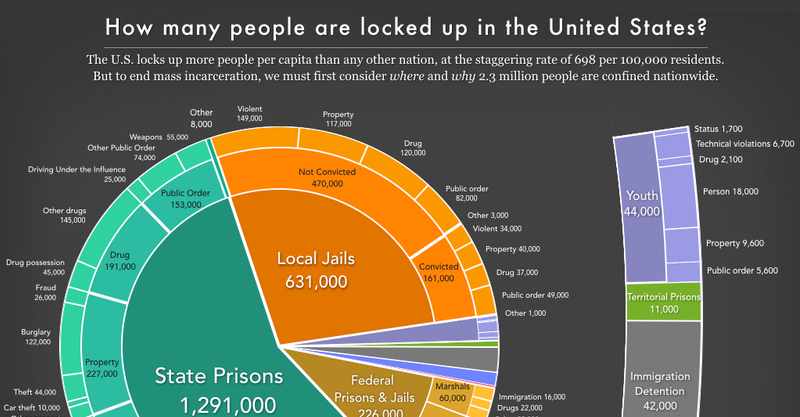
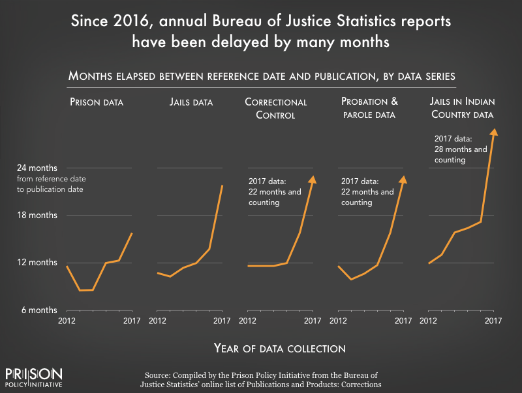
**Inspiration for the final product layout:**

<https://demos.creative-tim.com/material-dashboard-dark/examples/dashboard.html>

**3 or 4 screenshots of relevant, “inspiring” visualizations:**

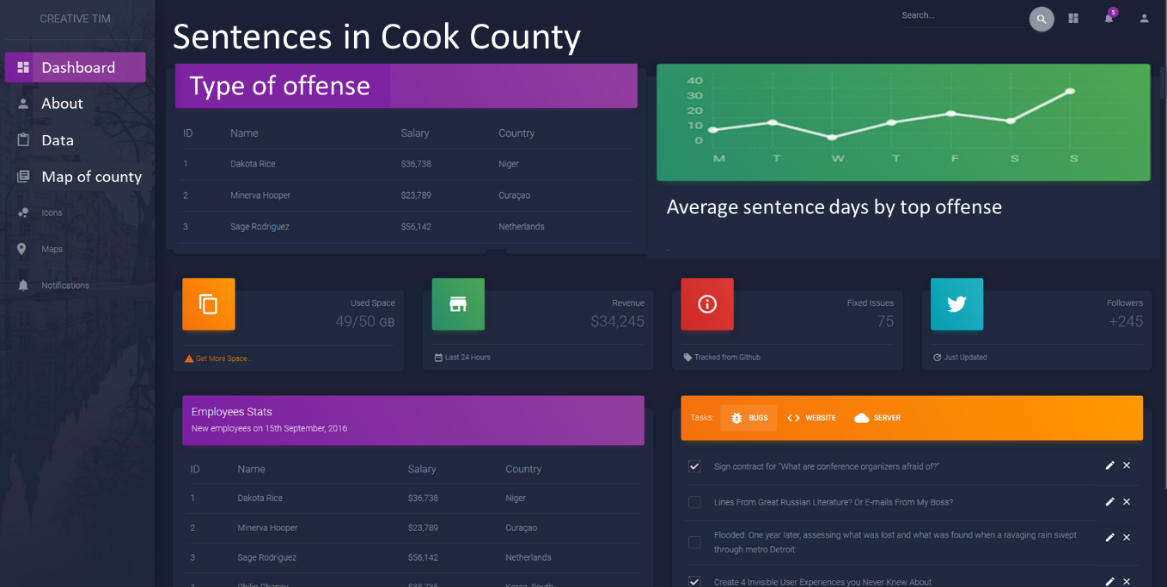
generated by other organizations, with topics similar to the subject matter of our proposed project

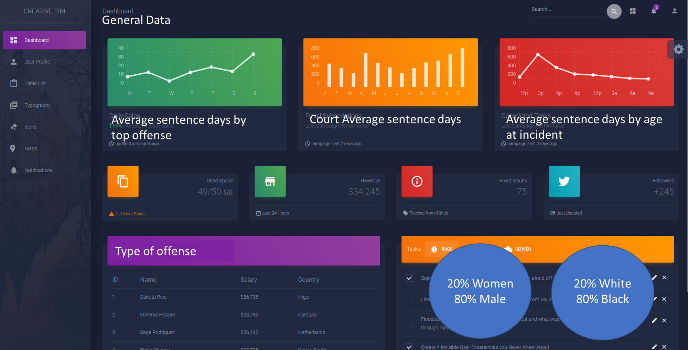
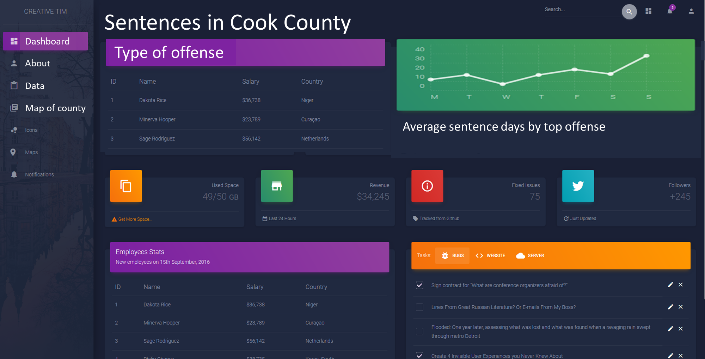
 

**VISION FOR FINAL PRODUCT**

**“Sketch” of ideal visuals:**



**“Sketch” of final product pages:**



**DATASET DETAILS**

**Dataset Link:** <https://datacatalog.cookcountyil.gov/Courts/Sentencing/tg8v-tm6u>

**Screenshots of the metadata:**

