

EAS 345: Phase 01  
Systemic racism in America  
Aghose, 10/09/2020

**Author:** Akash Ghose.

**Area of research:** Social issues surrounding racial tension in America.

**Title of project:** Does systemic racism exist in America?

**Potential clients:** People who wish to view the data about social disparities in America

**Potential sponsors:** People who wish to inform others of the existence (or non-existence) of social disparities in America

**Potential data sources:**

***FBI Crime/Arrest data:***

<https://ucr.fbi.gov/crime-in-the-u.s/2019/crime-in-the-u.s.-2019/tables/table-49>

<https://ucr.fbi.gov/crime-in-the-u.s/2018/crime-in-the-u.s.-2018/tables/table-49>

\*\*Replace 'YEAR' with actual year, dating back until 1995.

***US Sentencing commission data:***

<https://www.ussc.gov/sites/default/files/pdf/research-and-publications/annual-reports-and-sourcebooks/2019/2019-Annual-Report-and-Sourcebook.pdf>

^Of interest here, 56% of federal offenders were Hispanic. Even though Hispanics make up a very small percentage of the US population. How is this possible?

***US census bureau population data for the last 10 years:***

<https://www.census.gov/data/tables/time-series/demo/popest/2010s-national-detail.html>

***Consumer Financial Protection Bureau data:***

<https://www.consumerfinance.gov/data-research/hmda/historic-data/>

**Goal:**

In the Oct. 7, 2020 Vice Presidential debate, Mike Pence declared that systematic racism does not exist in America. My goal with this project is to aggregate data to prove or disprove that notion. I will try to do so by looking at the public data available and attempt to compare race vs arrests vs population size, race vs severity of crime vs incarceration rate/time, race vs mortgage loans denied/accepted, race vs income/job opportunities, race vs educational opportunities. In the end, I hope to be able to use this data to paint a very clear and coherent picture about social disparities in America, and aggregate it all in a very clean and concise place for all to view.

Phase 02: Data Collection

Akash Ghose, 10/23/20

### Crime Related Data sources

Data source 01:

**Name of files:** 2019\_FBI\_arrests\_by\_race\_total.csv

2019\_FBI\_arrests\_by\_race\_under18.csv

2019\_FBI\_arrests\_by\_race\_18\_and\_over.csv

**Source:** <https://ucr.fbi.gov/crime-in-the-u.s/2019/crime-in-the-u.s.-2019/topic-pages/tables/table-49>

**Details:** As the name suggests, this contains data about arrests in 2019. It contains details such as the race of the perpetrators and the type of crime they were arrested for.

**\*\*Of note:** I currently only downloaded the data provided for 2019, because I am not sure I need more than one year's data. So, for the sake of cleanliness, I have limited the data. However, if in the future, I need/want to get more data, it can be obtained with little to no effort.

### Finance Related Data sources

Data source 01:

**Name of file:** NFWBS\_PUF\_2016\_data\_readable.csv

**Source:** <https://www.consumerfinance.gov/data-research/financial-well-being-survey-data/>

**Details:** This is the National Financial Wellbeing Survey data from a survey that was conducted in 2016. This contains details about respondents and respondents' financial well-being, including characteristics like income, age, race, savings, past financial experiences, financial skills, behaviors, attitudes ect.

\*\*Of note: The original file I downloaded was: NFWBS\_PUF\_2016\_data.csv. I used NFWBS\_PUF\_2016\_read\_in\_R.R to read the file and then write it into the more readable csv.

Data source 02:

**Name of file:** hmda\_2017\_nationwide\_all-records\_labels.csv

**Source:** [https://www.consumerfinance.gov/data-research/hmda/historic-data/?geo=nationwide&records=all-records&field\\_descriptions=labels](https://www.consumerfinance.gov/data-research/hmda/historic-data/?geo=nationwide&records=all-records&field_descriptions=labels)

**Details:** This contains all the mortgage applications filed in 2017. It contains data about the applications and applicants, including details such as applicants' demographics and whether the application was accepted or rejected.

### US Population Data source

Data source 01:

**Name of file:** US\_population\_est\_2010-2019.csv

**Source:** [https://www.census.gov/data/tables/time-series/demo/popest/2010s-national-detail.html#par\\_textimage\\_1537638156](https://www.census.gov/data/tables/time-series/demo/popest/2010s-national-detail.html#par_textimage_1537638156)

**Details:** Contains US population estimates from 2010-2019. Includes separation by race as well as the totals.