The Data Science Union x Diversity Inclusion Data Product (DIPD) Collab Event

[Title]

NYPD Arrest Discrimination Analysis

[Overview]

We find that the NYPD is one of the most active police departments in the US, where they go through many different conflicts with criminals throughout the state with a multitude of different crimes ranging from murder, mariujana possesion, robbery, arson, etc.. Through many instances of arrests in the state, we see some specific instances of discrimination of specific arrests where the perpatrator can be wrongly accused or even wrongly acquitted from their crimes.

Suppose you were the governor of NY, who has the responsibility of evaluating the city's crime system and finding instances of discrimination and bias based off a demographic of the civilians of NY. We have collected a sample of data from government officials in the NY of over 1,000,000 documented NY state arrests from the years 2006-2017. This dataset contains variables of each arrest such as the age group, gender, precinct (city district), race, and other descriptive factors.

In this competition, you're challenged to use dataset to establish a case for a certain demographic bias among the arrests in NY. You will need to provide visualizations, models, or tables of the data that can help support your teams' case as to why a certain demographic is being discriminated against through the NY crime system. You can use any variable in the dataset to help justify your case. You're welcome to use any outside data set as long as you include the source.

[Evaluation]

Submissions are evaluated based on the team's ability to make clear and concise points and support their points with the data through different methods of building visualizations and finding key statistics to prove that their case is sufficient.

Submission File

Each team will be asked to submit 2 files:

1. A PDF/HTML document that gives us the your teams' evaluation and analysis with code generated tables, plots, and statistics that helps support your teams points and ideas.
2. The script of code your team used to produce the supporting tables, plots, and statistics used in making your case.

The code for the project should be written in a statistical oriented software (R, Python, Stata, Tableau, etc.).

WINNER LICENSE TYPE: Open-Source

DATA ACCESS AND USE: Competition Use , Non-Commercial Purposes & Academic Research, Commercial Purposes

Competitions are open to residents of the United States and worldwide, except that if you are a resident of Crimea, Cuba, Iran, Syria, North Korea, Sudan, or are subject to U.S. export controls or sanctions, you may not enter the Competition. Other local rules and regulations may apply to you, so please check your local laws to ensure that you are eligible to participate in skills-based competitions. The Competition Sponsor reserves the right to award alternative Prizes where needed to comply with local laws.

ENTRY IN THIS COMPETITION CONSTITUTES YOUR ACCEPTANCE OF THESE OFFICIAL COMPETITION RULES.

The Competition named above is a skills-based competition to promote and further the field of data science. Your competition submissions must conform to the requirements stated on the Competition Website. Your Submissions will be scored based on the evaluation metric described on the Competition Website. Subject to compliance with the Competition Rules, Prizes, if any, will be awarded to participants with the best scores, based on the merits of the data science models submitted. See below for the complete Competition Rules.

[Competition Rules]

Team Limits

The maximum team size is 3.

Submission Limits

Each team may select up to 2 final submissions for judging.

Prize

1st - $100 Amazon Gift Card

2nd - $50 Amazon Gift Card

3rd - $25 Amazon Gift Card

Special prizes for best modeling, visualization, and use of outside resources.

Judge

Professors from Stats Department will be judges for this event.

[Competition Timeline]

Start Date: February 19th, 2020

Entry Deadline: February 27th, 2020 11:59 PM

Submission Open Date: March 23th 2020

End Date (Final Submission Deadline): March 29th 2020 11:59 PM (Last Day of of Spring Break)

Results Party : Week1 Spring (TBD)