

# Hackathon on US elections 2020?

Agnes Gaspart, Romà Masana

February 2020

## Contents

<a href="#">1 Introduction</a>	1
<a href="#">2 Databases</a>	1

## 1 Introduction

We are planning to hold a hackathon of prediction contest at around 4th of April (a Saturday) during the whole day at PS111. I thought the topic of trying to predict the **US presidential election for 2020**, but Agnes and Prof. Kang say this is going to be complicated, if not impossible, and that another topic should be chosen. Agnes advocates for predicting and growing our models on **Emotion prediction** which I actually think is quite interesting. Agnes claims she already has the dataset and that all we have to do is run it.

However, when we asked a passer-by (by the name of *Hunter the sharp*) he said that the topic of the US election would be *spicy*, and that could be an attractive in order to draw the attention of the mainstream IIT studentdom.

For me we can do it about any other topic. But my passion lies on the US elections of 2020. We asked Prof. Kang and she was amused by the fact that we were going to attempt that. She said we won't obtain any rigorous results. Maybe she is right. Or maybe she is not. ;)

## 2 Databases

Prof. Kang suggested the following places were to mine for relevant data:

1. 538 political website <https://fivethirtyeight.com/>
2. Form the US Census bureau, I downloaded here <https://www.census.gov/data/tables/2019/demo/income-poverty/p60-266.html> the **table A-2**, containing the **mean, median and other relevant percentiles** of the **income per household** from **1967 until 2018**.
3. from US Census bureau: **Historical poverty** tables: people and families : <https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-poverty-people.html> 1959 until 2018.
4. from US Census bureau: **Historical School enrollment** tables: <https://www.census.gov/data/tables/time-series/demo/school-enrollment/cps-historical-time-series.html>
5. Polls from primary elections since 1970. <https://fivethirtyeight.com/features/what-more-than-40-years-of-early-primary-polls-tell-us-about-2020-part-1/>

I downloaded all these databases:

1. TableV9 Poverty by region from 1959 to 2018
2. Table v13 Families below poverty from 1959 to 2018
3. Table v15 Age distribution of the poor from 1959 to 2018

4. Table v23 Poverty by nativity from 1993 to 2018
5. Table A2 total money income by race from 1967 to 2018.
6. Table A4 Household income dispersion measures 1967 to 2018.
7. Table A7 Real median and mean earnings of total workers and full time workers, female-male from 1960 to 2018.
8. Table A2 Population 3 years and older enrolled in some sort of school from 1947 to 2018.
9. Table 5B Population 18 and 19 yrs old in high school, that dropped out, or that are in college, or none of the previous but that graduated from high school. From 1967 to 2018.