Group #5

Motivation:
Last election results are very interesting, and several states have surprising result . Most of the media had the wrong prectication. Moreover, in the last few years, the western countries have had several elections that have unexpected results too. We want to know if the same social values exist among western countries that would influence the election results.

Problem:
Can we use the data under following categories: Economics, Political affiliation (also known as political party), and Background or Social Media - to train a model which can use to predict the election results? Does the western world have similar social values that affect the elections? Is there a correlation between the results of elections and economics, background, or social media?

We expect it is complicated project which involves a lot of features, data collecting, and pre-processing. We expect that our model can give us some idea about people's preferences on the candidates and their raised issues. If our model can also work on other western countries, then we may be able to conclude that the social values are same among some countries.

Proposed Solution:
We want to use our data to see if we can create a model to predict whether or not the republican candidate will win the election. We randomly picked Republican, our choice has nothing to do with our own affiliations. We are separating the election into only Republican and Democratic. A description of all our desired attributes/features is at the end of the document.

We want to use SVM to create our model because the data is not linearly separable. It has more dependent features.

Datasets:
Presidential election results
https://data.opendatasoft.com/explore/dataset/usa-2016-presidential-election-by-county%40public/table/
115th Congress Members and Party/Affiliations
https://data.opendatasoft.com/explore/dataset/us-115th-congress-members%40public/table/
Census Survey 2015
https://data.world/uscensusbureau/acs-2015-5-e-placeofbirth/workspace/file?filename=00_README.md
inflation data
https://www.inflation.eu/inflation-rates/united-states/historic-inflation/cpi-inflation-united-states.aspx
Average income by state :
https://www.census.gov/content/dam/Census/library/publications/2017/acs/acsbr16-02.pdf
Poverty rate :
https://www.statista.com/statistics/233093/us-poverty-rate-by-state/

employment, employment, and population:
https://www.bls.gov/news.release/pdf/srgune.pdf
party affiliation by state:
http://www.pewforum.org/religious-landscape-study/compare/party-affiliation/by/state/

Note: Datasets are also provided on GitHub

Data:

Each row belongs to a state and year

Column/Attribute	Description	Have data?
State	The US state. Not used in the data model	--
year	Election Year for the data	--
Republican Senators	The percentage of republican senators in capitol hill	s
Republican Representatives	The percentage of republican representatives in capitol hill	s
Republican Legislators in state house legislature	percentage	s
Money Spent on Republican Campaign as percentage	The amount spent on the campaign out of the total spent (rep / (rep+dem))	a
Voter Affiliation	Number of voters who affiliate with the party at the time of the voting registration (2015)	unavailable
Republican Second Term	National candidate is running a second term (will be the same for each state for one year)	removed
Poverty	Poverty rate per year	a
inflation	Inflation rate for the year	a
Avg age	Average age for the state (may need ranges instead) -- removed	removed
Median household income	(self explanatory) -- State/US	a
Foreign born	Percentage population for the state - 2016	pending
Stock market	NASDAQ 1 year ago % change	a
Unemployment rate	percentage	a
High School Graduate	Rate (%) -- national reports >2010	m
College Graduate	Rate (%) --2016	m

Republican Candidate win (?)	Binary, TARGET column	s
Rural vs urban (Or population density)	% population rural (or pop/mi2) -- for last stage	pending
GDP per capita	State GDP per capita / US GDP per capita	s