**KU Data Analytics Bootcamp -Trilogy/2U -Final Project**

**Presentation Date: 07/11/2020**

**Team: JuST Data Members: Jemi, Sanita, Tom**

**Data Sourcing and Cleansing Process**

Raw data was sourced from a set of US Census Data website files originally in either Old- version-XL, Comma-Delimited or Text files (ref below):

Raw data was derived from (4) sets of data files including:

1980-2014 Turnout Rates,

2000, 2004, 2008, 2012, 2016 Primary Election data,

2000, 2004, 2008, 2012, 2016 General Election data, and

1996-2016 Registered Voter data.

This data was restructured and standardized by importing it into XL, any differences in format were corrected and any missing data was derived from rate of Population values. Lastly, the data files were merged providing “by Year”, “by State” and “total US” numbers.

At this point the (5) selected years 2000, 2004, 2008, 2012, 2016 were chosen to maximize usefulness (presidential election data on 4yr cycle) and to ensure a manageable amount of data. It was then passed to subsequent python processes.

Data Tables All

<https://www.census.gov/topics/public-sector/voting/data/tables.html>

Data Tables 9999

<https://www.census.gov/topics/public-sector/voting/data/tables.2014.html>

Voting and Registration 9999

<https://www.census.gov/data/tables/time-series/demo/voting-and-registration/p20-577.html>

<https://www.census.gov/data/tables/2010/demo/voting-and-registration/voting-registration-2010-election.html>

<https://www.census.gov/data/tables/2008/demo/voting-and-registration/p20-562-rv.html>

Census tbl04

2016 <https://www2.census.gov/programs-surveys/cps/tables/p20/580/table04a.xlsx>

2014 <https://www2.census.gov/programs-surveys/cps/tables/p20/577/table04a.xls>

2012 <https://www2.census.gov/programs-surveys/cps/tables/p20/568/table04a.xls>

2010 <https://www2.census.gov/programs-surveys/cps/tables/p20/voting-registration-2010-election/table4a_2010.xls>

2008 <https://www2.census.gov/programs-surveys/cps/tables/p20/562-rv/table-04a.xls>

2006 <https://www2.census.gov/programs-surveys/cps/tables/p20/557/tab04a.xls>

2004 <https://www2.census.gov/programs-surveys/cps/tables/p20/556/tab04c.xls>

2002 <https://www2.census.gov/programs-surveys/cps/tables/p20/552/tab04c.xls>

2000 <https://www2.census.gov/programs-surveys/cps/tables/p20/542/tab04c.xls>

1998 <https://www2.census.gov/programs-surveys/cps/tables/p20/523-rv/tab04.txt>

1996 <https://www2.census.gov/programs-surveys/cps/tables/p20/504/tab4a.txt>