

PSC 400

SYRACUSE UNIVERSITY

DATA ANALYTICS

FOR POLITICAL

SCIENCE

FINDING AND CLEANING DATA,

CONTINUED

REMINDER

- **Thursday: no class**
 - **Use the time to find data sources for your research project instead**

QOG DATA

- <https://www.gu.se/en/quality-government/qog-data/data-downloads/basic-dataset>
- We want to download csv file of cross-sectional data: `qog_bas_cs_jan24.csv`
- Codebook: `codebook_bas_jan24.pdf`

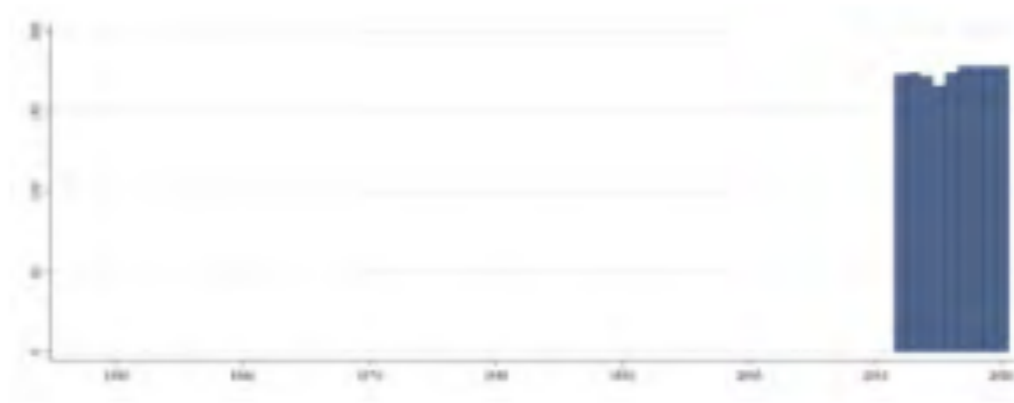
EXERCISE: QOG

4.94.1 Corruption Perceptions Index (ti_cpi)

Corruption Perceptions Index. Scale of 0-100 where a 0 equals the highest level of perceived corruption and 100 equals the lowest level of perceived corruption.



Min. Year:2017 Max. Year: 2017
N: 178



Min. Year:2012 Max. Year: 2020
N: 178 n: 1571 \bar{N} : 175 \bar{T} : 9

- Create a new variable `ti_cpi_max10` where 0 is the highest level of corruption and 10 is the lowest

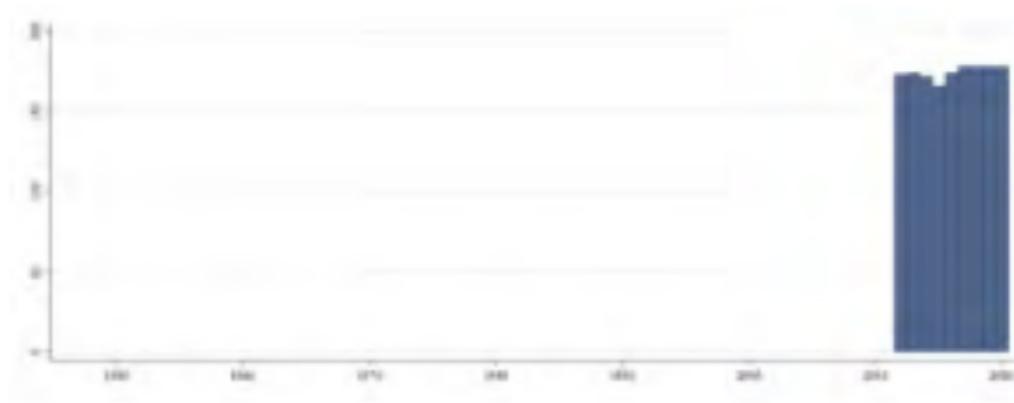
EXERCISE: QOQ

4.94.1 Corruption Perceptions Index (ti_cpi)

Corruption Perceptions Index. Scale of 0-100 where a 0 equals the highest level of perceived corruption and 100 equals the lowest level of perceived corruption.



Min. Year:2017 Max. Year: 2017
N: 178



Min. Year:2012 Max. Year: 2020
N: 178 n: 1571 \bar{N} : 175 \bar{T} : 9

- Create a new variable `ti_cpi_reverse` where 100 is the *highest* level of corruption and 0 is the lowest

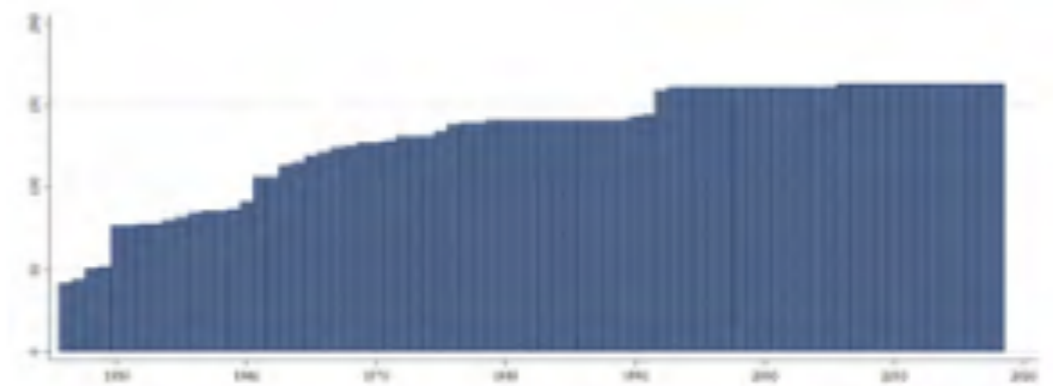
EXERCISE: QOQ

4.70.1 Real GDP per Capita (mad_gdppc)

Real GDP per capita in 2011 US dollars, multiple benchmarks.



Min. Year:2017 Max. Year: 2017
N: 163



Min. Year:1946 Max. Year: 2018
N: 175 n: 9559 \bar{N} : 131 \bar{T} : 55

- Plot the density of mad_gdppc
- Create a new variable mad_gdppc_log that is the logged value of mad_gdppc (function: log)
- Plot the density of mad_gdppc_logrr

EXERCISE: QOG

4.41.1 Colonial Origin (ht_colonial)

This is a tenfold classification of the former colonial ruler of the country. Following Bernard et al. (2004), we have excluded the British settler colonies (the US, Canada, Australia, Israel and New Zealand), and exclusively focused on “Western overseas” colonialism. This implies that only Western colonizers (e.g. excluding Japanese colonialism), and only countries located in the non-Western hemisphere “overseas” (e.g. excluding Ireland & Malta), have been coded. Each country that has been colonized since 1700 is coded. In cases of several colonial powers, the last one is counted, if it lasted for 10 years or longer. The categories are the following:

0. Never colonized by a Western overseas colonial power
1. Dutch
2. Spanish
3. Italian
4. US
5. British
6. French
7. Portuguese
8. Belgian
9. British-French
10. Australian

EXERCISE: QOQ

4.41.1 Colonial Origin (ht_colonial)

This is a tenfold classification of the former colonial ruler of the country. Following Bernard et al. (2004), we have excluded the British settler colonies (the US, Canada, Australia, Israel and New Zealand), and exclusively focused on “Western overseas” colonialism. This implies that only Western colonizers (e.g. excluding Japanese colonialism), and only countries located in the non-Western hemisphere “overseas” (e.g. excluding Ireland & Malta), have been coded. Each country that has been colonized since 1700 is coded. In cases of several colonial powers, the last one is counted, if it lasted for 10 years or longer. The categories are the following:

0. Never colonized by a Western overseas colonial power
1. Dutch
2. Spanish
3. Italian
4. US
5. British
6. French
7. Portuguese
8. Belgian
9. British-French
10. Australian

- **Create a variable “colonized” that is 1 if the country was ever colonized by a Western colonial power, and 0 if not**

EXERCISE: QO G

4.41.1 Colonial Origin (ht_colonial)

This is a tenfold classification of the former colonial ruler of the country. Following Bernard et al. (2004), we have excluded the British settler colonies (the US, Canada, Australia, Israel and New Zealand), and exclusively focused on “Western overseas” colonialism. This implies that only Western colonizers (e.g. excluding Japanese colonialism), and only countries located in the non-Western hemisphere “overseas” (e.g. excluding Ireland & Malta), have been coded. Each country that has been colonized since 1700 is coded. In cases of several colonial powers, the last one is counted, if it lasted for 10 years or longer. The categories are the following:

0. Never colonized by a Western overseas colonial power
1. Dutch
2. Spanish
3. Italian
4. US
5. British
6. French
7. Portuguese
8. Belgian
9. British-French
10. Australian

- Create a variable “colonized2” that is “colonized” if the country was ever colonized by a Western colonial power, and “not colonized” if not

EXERCISE: QOQ

- **Create a variable of the COVID fatality rate in 2020 (# of deaths/#number of cases)**
- **Create a binary variable that indicates whether a country is a democracy or not**
 - **Democracy: Revised Combined Polity Score of 6 or greater**
- **Compare COVID fatality rates between democracies and non-democracies**