Final

Saamia

7/9/2020

# clearing environment  
rm(list=ls(all=TRUE))  
# loading libraries  
library(tidycensus)  
library(rio)  
library(data.table)  
library(tidyverse)

## ── Attaching packages ────────

## ✓ ggplot2 3.3.1 ✓ purrr 0.3.4  
## ✓ tibble 3.0.1 ✓ dplyr 1.0.0  
## ✓ tidyr 1.1.0 ✓ stringr 1.4.0  
## ✓ readr 1.3.1 ✓ forcats 0.5.0

## ── Conflicts ─────────────────  
## x dplyr::between() masks data.table::between()  
## x dplyr::filter() masks stats::filter()  
## x dplyr::first() masks data.table::first()  
## x dplyr::lag() masks stats::lag()  
## x dplyr::last() masks data.table::last()  
## x purrr::transpose() masks data.table::transpose()

library(WDI)  
library(labelled)  
library(googlesheets4)  
library(varhandle)  
library(ggrepel)  
library(geosphere)  
library(rgeos)

## Loading required package: sp

## rgeos version: 0.5-3, (SVN revision 634)  
## GEOS runtime version: 3.7.2-CAPI-1.11.2   
## Linking to sp version: 1.4-1   
## Polygon checking: TRUE

library(viridis)

## Loading required package: viridisLite

library(mapview)  
library(rnaturalearth)  
library(rnaturalearthdata)  
library(devtools)

## Loading required package: usethis

library(remotes)

##   
## Attaching package: 'remotes'

## The following objects are masked from 'package:devtools':  
##   
## dev\_package\_deps, install\_bioc, install\_bitbucket, install\_cran,  
## install\_deps, install\_dev, install\_git, install\_github,  
## install\_gitlab, install\_local, install\_svn, install\_url,  
## install\_version, update\_packages

## The following object is masked from 'package:usethis':  
##   
## git\_credentials

library(raster)

##   
## Attaching package: 'raster'

## The following object is masked from 'package:dplyr':  
##   
## select

## The following object is masked from 'package:tidyr':  
##   
## extract

## The following object is masked from 'package:data.table':  
##   
## shift

library(sp)  
library(sf)

## Linking to GEOS 3.7.2, GDAL 2.4.2, PROJ 5.2.0

library(Imap)

##   
## Attaching package: 'Imap'

## The following object is masked from 'package:purrr':  
##   
## imap

library(pdftools)

## Using poppler version 0.73.0

library(tidyr)  
library(tidytext)  
library(dplyr)  
library(stringr)  
library(ggplot2)  
library(rvest)

## Loading required package: xml2

##   
## Attaching package: 'rvest'

## The following object is masked from 'package:purrr':  
##   
## pluck

## The following object is masked from 'package:readr':  
##   
## guess\_encoding

library(xml2)  
  
# connecting api key  
census\_api\_key("c27810c8b6634b4df374ba6ef4045a7dc16103c6",  
 install = TRUE,  
 overwrite = TRUE)

## Your original .Renviron will be backed up and stored in your R HOME directory if needed.

## Your API key has been stored in your .Renviron and can be accessed by Sys.getenv("CENSUS\_API\_KEY").   
## To use now, restart R or run `readRenviron("~/.Renviron")`

## [1] "c27810c8b6634b4df374ba6ef4045a7dc16103c6"

readRenviron("~/.Renviron")

1. GINI variable = B19083\_001

v10 <- load\_variables(year = 2010,  
 'acs5')  
# gini variable B19083\_001   
v15 <- load\_variables(year = 2015,  
 'acs5')  
# gini variable B19083\_001   
  
  
# state level data for both years   
#gini <- get\_acs(geography = 'state', variables = 'B19083\_001')

# inequality\_wide <- inequality\_panel %>%  
# pivot\_wider(id\_cols = c(''),  
# names\_from = ,   
# values\_from = )

# inequality\_long <-inequality\_wide %>%  
# pivot\_longer(cols = ,   
# names\_to = ,  
# names\_prefix = ,  
# values\_to = ,   
# values\_drop\_na = FALSE)

# count(inequality\_panel)  
# count(inequality\_long)

# inequality\_collapsed <- inequality\_long %>%  
# group\_by() %>%  
# summarize(across(where(is.numeric), sum)) %>%  
# select(-c())

gdp\_current = WDI(country = 'all',   
 indicator = c('NY.GDP.MKTP.CD'),   
 start = 2006, end = 2007, extra = FALSE, cache = NULL)

9.Chose 2015 because that is the year that deflator = 100

deflator\_data = WDI(country = 'all', indicator = c('NY.GDP.DEFL.ZS'),   
 start = 2006,  
 end = 2016,   
 extra = FALSE, cache = NULL)  
setnames(deflator\_data, 'NY.GDP.DEFL.ZS', 'deflator')  
usd\_deflator = subset(deflator\_data, country == 'United States')  
subset(usd\_deflator, deflator==100)

## iso2c country deflator year  
## 2785 US United States 100 2015

rm(deflator\_data)  
usd\_deflator$country <- NULL  
usd\_deflator$iso2c <- NULL  
gdp\_deflated = left\_join(gdp\_current,   
 usd\_deflator,  
 by = c('year'))  
gdp\_deflated$deflated\_amount = gdp\_deflated$NY.GDP.MKTP.CD/(gdp\_deflated$deflator/100)  
head(gdp\_deflated)

## iso2c country NY.GDP.MKTP.CD year deflator  
## 1 1A Arab World 1.639412e+12 2007 88.25933  
## 2 1A Arab World 1.405832e+12 2006 85.95046  
## 3 S3 Caribbean small states 6.091144e+10 2007 88.25933  
## 4 S3 Caribbean small states 5.472063e+10 2006 85.95046  
## 5 B8 Central Europe and the Baltics 1.264970e+12 2007 88.25933  
## 6 B8 Central Europe and the Baltics 1.001499e+12 2006 85.95046  
## deflated\_amount  
## 1 1.857494e+12  
## 2 1.635631e+12  
## 3 6.901417e+10  
## 4 6.366531e+10  
## 5 1.433242e+12  
## 6 1.165205e+12

1. The 3 main components are the library, the UI, and server functions. The UI has inputs and outputs. The server function has 3 rules: storing directions as object, using render, and refering output object to input object.

pdf\_final <- pdf\_text(pdf = 'https://pdf.usaid.gov/pdf\_docs/PA00TNMG.pdf')

armeniatext <- as.data.frame(pdf\_final, stringsAsFactors = FALSE)

data("stop\_words")   
tidy\_armenia <- armeniatext %>%  
 unnest\_tokens(word, pdf\_final) %>%  
 anti\_join(stop\_words)

## Joining, by = "word"

14 Top 5 = armenia, political, corruption, governance, democracy

top\_5 <- tidy\_armenia %>%   
 count(word, sort = TRUE) %>%  
 filter(n>130)

hot100 <- 'https://www.billboard.com/charts/hot-100'  
hot100exam <- read\_html(hot100)

hot100exam

## {html\_document}  
## <html class="" lang="">  
## [1] <head>\n<meta http-equiv="Content-Type" content="text/html; charset=UTF-8 ...  
## [2] <body class="chart-page chart-page-" data-trackcategory="Charts-TheHot100 ...

body\_nodes <- hot100exam %>%  
 html\_node('body') %>%  
 html\_children()  
body\_nodes

## {xml\_nodeset (36)}  
## [1] <div class="header-wrapper ">\n<header id="site-header" class="site-head ...  
## [2] <div class="site-header\_\_placeholder"></div>  
## [3] <script>\n var PGM = window.PGM || {};\n PGM.config = PGM. ...  
## [4] <div class="chart-piano-overlay\_\_attachment-point"></div>  
## [5] <main id="main" class="page-content"><div id="charts" data-page-title="T ...  
## [6] <div class="ad\_desktop dfp-ad dfp-ad-promo " data-position="promo" data- ...  
## [7] <div class="ad-container footerboard footerboard--bottom">\n <div cla ...  
## [8] <footer id="site-footer" class="site-footer"><div class="container foote ...  
## [9] <div class="biz-modal">\n <div class="biz-modal\_\_content">\n < ...  
## [10] <script>\n window.CLARITY = window.CLARITY || [];\n</script>  
## [11] <div class="ad\_clarity" data-out-of-page="true" style="display: none;">< ...  
## [12] <script>\n var darkMatterCMD = function() {\n this.darkMatterC ...  
## [13] <script src="https://www.billboard.com/assets/1593527595/js/vendors\_/art ...  
## [14] <script src="https://www.billboard.com/assets/1593527595/js/vendors\_/clo ...  
## [15] <script src="https://www.billboard.com/assets/1593527595/js/vendors\_/rea ...  
## [16] <script src="https://www.billboard.com/assets/1593527595/js/vendors\_/rea ...  
## [17] <script src="https://www.billboard.com/assets/1593527595/js/vendors\_/rea ...  
## [18] <script src="https://www.billboard.com/assets/1593527595/js/vendors\_/rea ...  
## [19] <script src="https://www.billboard.com/assets/1593527595/js/default\_/art ...  
## [20] <script src="https://www.billboard.com/assets/1593527595/js/default\_/rea ...  
## ...

body\_nodes %>%   
 html\_children()

## {xml\_nodeset (9)}  
## [1] <header id="site-header" class="site-header " role="banner"><div class="s ...  
## [2] <div class="header-wrapper\_\_secondary-header">\n<nav class="site-header-l ...  
## [3] <div id="charts" data-page-title="THE HOT 100" data-chart-code="HSI" data ...  
## [4] <div class="footerboard-wrapper">\n <div class="ad\_desktop\_placeho ...  
## [5] <div class="container footer-content">\n\t\t\t\t\t<div class="cover-image ...  
## [6] <div class="container">\n\t\t<p class="copyright\_\_paragraph">© 2020 Billb ...  
## [7] <div class="container">\n\t\t<p class="station-identification">\n\t\t\tBi ...  
## [8] <div class="container">\n\t\t\n\n\n <div class="ad\_desktop dfp-ad dfp- ...  
## [9] <div class="biz-modal\_\_content">\n <button class="biz-modal\_\_close ...

rank <- hot100exam %>%  
 rvest::html\_nodes('body') %>%  
 xml2::xml\_find\_all("//span[contains(@class, 'chart-element\_\_rank\_\_number')]")  
  
artist <- hot100exam %>%   
 rvest::html\_nodes('body') %>%  
 xml2::xml\_find\_all("//span[contains(@class, 'chart-element\_\_information\_\_artist')]")  
  
title <- hot100exam %>%  
 rvest::html\_nodes('body') %>%  
 xml2::xml\_find\_all("//span[contains(@class, 'chart-element\_\_information\_\_song')]")  
  
last\_week <- hot100exam %>%  
 rvest::html\_nodes('body') %>%  
 xml2::xml\_find\_all("//span[contains(@class, 'chart-element\_\_information\_\_delta\_\_text text--last')]")

1. <https://github.com/simtiaz17/final>