health_is_political

R Markdown

```
library(tidyverse)
```

```
## -- Attaching packages ---------------------------------------------------------------------------
```

```
## v ggplot2 3.2.1     v purrr   0.3.3
## v tibble  2.1.3     v dplyr   0.8.3
## v tidyr   1.0.2     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.4.0
```

```
## -- Conflicts ------------------------------------------------------------------------------------
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(tidycensus)
library(readxl)
library(corrr)
library(corrplot)
```

```
## corrplot 0.84 loaded
```

```
library(tidycensus)
library(boot)
library(broom)
library(lubridate)
```

```
##
## Attaching package: 'lubridate'
```

```
## The following object is masked from 'package:base':
##
##     date
```

```
library(gridExtra)
```

```
##
## Attaching package: 'gridExtra'
```

```
## The following object is masked from 'package:dplyr':
##
##     combine
```

```
######################################## Paths
#Source: https://data.opendatasoft.com/explore/dataset/usa-2016-presidential-election-by-county%40publi
#https://www.nytimes.com/elections/2016/results/president
#https://openpsych.net/paper/12
voter_info_data <- "~/data_health_voting/usa-2016-presidential-election-by-county.csv"

##Alternative source for data: https://electionlab.mit.edu/data
#DOI here: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/VOQCHQ
vote_county_data <- "~/data_health_voting/countypres_2000-2016.csv"
```

```
#Source: https://data.world/niccolley/us-zipcode-to-county-state
zip_county_data <- "~/data_health_voting/ZIP-COUNTY-FIPS_2016-12.csv"

#Source: https://www.countyhealthrankings.org/
county_health_data_2019 <- "~/data_health_voting/2019_county_health_data.xlsx"
county_health_data_2010 <- "~/data_health_voting/2010_county_health_data.xlsx"


#Diabetes, Obesity, physical inactivity data
#Source CDC: https://gis.cdc.gov/grasp/diabetes/DiabetesAtlas.html#, https://www.cdc.gov/diabetes/data/
#How they were calculated: https://www.cdc.gov/diabetes/data/statistics/faqs.html
diabetes_atlas_data <- "~/data_health_voting/Diabetes_2004_2016.xlsx"
obesity_atlas_data <- "~/data_health_voting/Obesity_2004_2016.xlsx"
phys_inact_atlas_data <- "~/data_health_voting/Physical_inactivity_2004_2016.xlsx"

#Crude data
diabetes_crude_atlas_data <- "~/data_health_voting/Diabetes_Crude_2006_2016.xlsx"
obesity_crude_atlas_data <- "~/data_health_voting/obesity_crude.xlsx"
phys_inact_crude_atlas_data <- "~/data_health_voting/physical_inactivity_crude.xlsx"

#GBD paths, mortality
respiratory_diseases_data <- "~/data_health_voting/IHME_USA_COUNTY_RESP_DISEASE_MORTALITY_1980_2014_NATI
infectious_diseases_data <- "~/data_health_voting/IHME_USA_COUNTY_INFECT_DIS_MORT_1980_2014_NATIONAL_Y2
cvd_diseases_data <- "~/data_health_voting/IHME_USA_COUNTY_CVD_MORTALITY_RATES_1980_2014_NATIONAL_Y2017
le_mortaliity_risk_data <- "~/data_health_voting/IHME_USA_COUNTY_LE_MORTALITY_RISK_1980_2014_NATIONAL_Y
cancer_data <- "~/data_health_voting/IHME_USA_COUNTY_CANCER_MORTALITY_RATES_1980_2014_NATIONAL_Y2017M01
deaths_of_despair_data <- "~/data_health_voting/IHME_USA_COUNTY_USE_INJ_MORTALITY_1980_2014_NATIONAL_Y2
sum_deaths_data <- "~/data_health_voting/IHME_USA_COUNTY_MORTALITY_RATES_1980_2014_NATIONAL_Y2016M12D13

#medicare healthcare costs
medicare_costs_data <- "~/data_health_voting/State County All Table 2017.xlsx"

#Opioid Prescribing Rate: 2017 opioid prescribing rates
cdc_prescribing_opioid_data <- "~/data_health_voting/cdc_opioid_prescribing_rates.xlsx"

#Census Median Age and population
pop_age_data <- "~/data_health_voting/population_and_age.csv"

#Unemployment rate (get it from BLS): Source: https://www.bls.gov/lau/#cntyaa
bls_path <- "~/data_health_voting/labor_force_bls.xlsx"

#SAHIE data
#Sources: https://www.census.gov/content/dam/Census/library/publications/2019/demo/p30-05.pdf
#https://www.census.gov/data-tools/demo/sahie/#/?s_iprcat=0&map_yearSelector=2008&s_year=2017,2016,2015
SAHIE_data_1 <- "~/data_health_voting/SAHIE_08JAN20_13_49_43_32.csv"
SAHIE_data_2 <- "~/data_health_voting/SAHIE_08JAN20_13_55_04_87.csv"
SAHIE_data_3 <- "~/data_health_voting/SAHIE_08JAN20_13_58_30_44.csv"

#County Health Rankings data
chr_trends_data <- "~/data_health_voting/CHR_TRENDS_CSV_2019.csv"

#County health rankings, processed data
chr_2019_data <- "~/data_health_voting/2019 County Health Rankings Data.xls"
```

```r
#Medicaid variables (from acs5)
medicaid_var_data <- "~/data_health_voting/medicaid_variables.csv"

#ACS variables
acs_data <- "~/data_health_voting/acs_variables.csv"
acs_vars_info <- "~/data_health_voting/acs_vairables.xlsx"

#categories
county_health_category_data <- "~/data_health_voting/county_health_variable_categories.xlsx"

#Compare states on Medicaid expansion- Medicaid data: https://data.medicaid.gov/widgets/n5ce-jxme
medicaid_expansion_data <- "~/data_health_voting/State_Medicaid_and_CHIP_Applications__Eligibility_Dete

#State election data
state_election_data <- "~/data_health_voting/1976-2016-president.csv"

#Preexisting conditions by state
preexiting_state_data <- "~/data_health_voting/pre-existing-conditions-by-congressional-district.xlsx"

########################################States
flip_states <- c("Michigan", "Pennsylvania", "Wisconsin", "Maine")
battle_states <- c("Arizona", "Florida", "Maine", "Michigan", "Minnesota", "Nebraska", "New Hampshire",

states_obama_won_by_lessthan_10 <-
  c("Nevada", "Colorado", "Virginia", "Florida", "Michigan", "Minnesota", "Wisconsin", "Iowa", "New Hamp

#####################################Read in data
#zip
zip_county <- read_csv(zip_county_data)
```

```
## Parsed with column specification:
## cols(
##   ZIP = col_double(),
##   COUNTYNAME = col_character(),
##   STATE = col_character(),
##   STCOUNTYFP = col_double(),
##   CLASSFP = col_character()
## )
```

```r
#voter
voter_info <-
  read_delim(voter_info_data, delim = ";") %>%
  select(-c(`Geo Shape`, `Lat Bins`, `Lon Bins`, `Precip Bins`, `Elevation Bins`))
```

```
## Parsed with column specification:
## cols(
##   .default = col_double(),
##   State = col_character(),
##   ST = col_character(),
##   Fips = col_character(),
##   County = col_character(),
##   `Temp Bins` = col_character(),
##   `Lat Bins` = col_logical(),
##   `Lon Bins` = col_character(),
```

```
##   `Precip Bins` = col_character(),
##   `Elevation Bins` = col_character(),
##   `Geo Shape` = col_character(),
##   `Name 16` = col_character(),
##   `Name Prev` = col_character(),
##   `Statecode Prev` = col_character()
## )
```

```
## See spec(...) for full column specifications.
```

```
vote_county <- read_csv(vote_county_data)
```

```
## Parsed with column specification:
## cols(
##   year = col_double(),
##   state = col_character(),
##   state_po = col_character(),
##   county = col_character(),
##   FIPS = col_double(),
##   office = col_character(),
##   candidate = col_character(),
##   party = col_character(),
##   candidatevotes = col_double(),
##   totalvotes = col_double(),
##   version = col_double()
## )
```

```
#county health covariates
county_health_2019 <- read_xlsx(county_health_data_2019)
county_health_2010 <- read_xlsx(county_health_data_2010)
```

Read in 2012 and 2016 voter data

```
vote_county_2012_2016 <-
  vote_county %>%
  filter(
    year == 2016
  ) %>%
  select(-c(party, version)) %>%
  spread(candidate, candidatevotes) %>%
  mutate(
    prct_trump_2016 = `Donald Trump` / totalvotes,
    prct_clinton_2016 = `Hillary Clinton` / totalvotes
  ) %>%
  rename("totalvotes_2016" = totalvotes) %>%
  select(state:totalvotes_2016, prct_trump_2016, prct_clinton_2016) %>%
  left_join(
    vote_county %>%
    filter(
      year == 2012
    ) %>%
    select(-c(party, version)) %>%
    spread(candidate, candidatevotes) %>%
    mutate(
      prct_obama_2012 = `Barack Obama` / totalvotes,
      prct_romney_2012 = `Mitt Romney` / totalvotes
```

```
    ) %>%
    rename("totalvotes_2012" = totalvotes) %>%
    select(state:totalvotes_2012, prct_romney_2012, prct_obama_2012),
    by = c("state", "state_po", "county", "FIPS", "office")
  ) %>%
  mutate(FIPS = as.character(FIPS))
```

Preprocess county health data

#CDC Diabetes, obesity, physical inactivity

This is thee same data, except crude

#Mortality data from GBD

#Medicare Healthcare costs

#Opioid Prescribing Rate: 2017 opioid prescribing rates

```
opioid_rates <-
  read_xlsx(cdc_prescribing_opioid_data) %>%
  select("FIPS" = `State/County FIPS Code`, opioid_prescribing_rate = `2017`)
```

#Census Median Age and population

#Unemployment rate (get it from BLS): Source: https://www.bls.gov/lau/#cntyaa

```
bls_unemploy <-
  excel_sheets(bls_path)  %>%
  map_df( ~ read_xlsx(bls_path, sheet = ., skip = 1) %>%
      mutate_at(
        vars(State, County, Year),
        as.character
      )
  )
```

#SAHIE data

#County Health Rankings data

#County health rankings, processed data

Select county health rankings data that we will use

Select the CDC atlass data

```
cdc_atlas_2016 <-
  diabetes_atlas %>%
  filter(year == "2016") %>%
  select(CountyFIPS, `Diabetes %` = Percentage) %>%
  left_join(
    phys_inact_atlas %>%
    filter(year == "2016") %>%
    select(CountyFIPS, `Physical Inactive %` = Percentage),
    by = c("CountyFIPS")
  ) %>%
  left_join(
    obesity_atlas %>%
    filter(year == "2016") %>%
    select(CountyFIPS, `Obesity %` = Percentage),
```

```
    by = c("CountyFIPS")
  )
```

Select the GBD data

```
GBD_diseases <-
  respiratory_diseases %>%
  filter(!is.na(FIPS)) %>%
  select(FIPS, disease, `Mortality Rate, 2014*`) %>%
  spread(disease, `Mortality Rate, 2014*`) %>%
  mutate(FIPS = as.character(FIPS)) %>%
  left_join(
    infectious_diseases %>%
    filter(!is.na(FIPS)) %>%
    select(FIPS, disease, `Mortality Rate, 2014*`) %>%
    spread(disease, `Mortality Rate, 2014*`) %>%
    mutate(FIPS = as.character(FIPS)),
    by = c("FIPS")
  ) %>%
  left_join(
    cvd_diseases %>%
    filter(!is.na(FIPS)) %>%
    select(FIPS, disease, `Mortality Rate, 2014*`) %>%
    spread(disease, `Mortality Rate, 2014*`) %>%
    mutate(FIPS = as.character(FIPS)),
    by = c("FIPS")
  ) %>%
  left_join(
    cancer %>%
    filter(!is.na(FIPS)) %>%
    select(FIPS, disease, `Mortality Rate, 2014*`) %>%
    spread(disease, `Mortality Rate, 2014*`) %>%
    mutate(FIPS = as.character(FIPS)),
    by = c("FIPS")
  ) %>%
  left_join(
    deaths_of_despair %>%
    filter(!is.na(FIPS)) %>%
    select(FIPS, disease, `Mortality Rate, 2014*`) %>%
    spread(disease, `Mortality Rate, 2014*`) %>%
    mutate(FIPS = as.character(FIPS)),
    by = c("FIPS")
  ) %>%
  left_join(
    le_mortaliity_risk %>%
    filter(disease != "Life expectancy") %>%
    filter(!is.na(FIPS)) %>%
    select(FIPS, disease, `Mortality risk, 2014*`)  %>%
    spread(disease, `Mortality risk, 2014*`) %>%
    mutate(FIPS = as.character(FIPS)),
    by = c("FIPS")
  ) %>%
  mutate_at(vars(Asbestosis:`Mortality risk, age 65-85`), as.numeric)
```

Health costs variables to keep

```
health_costs_keep <-
  c(
    "State and County FIPS Code", "Part B Drugs Actual Costs", "Emergency Department Visits", "Imaging
    "Procedures Per Capita Actual Costs", "Hospice Per Capita Actual Costs", "Tests Per Capita Actual C
    "Actual Per Capita Costs", "Percent Eligible for Medicaid", "Percent Male", "Percent Female"
  )
```

Insurance rates

```
sahie_ins <-
  sahie_insurance %>%
  filter(Year == 2016) %>%
  select(ID, `Income Category`, `Uninsured: %`) %>%
  spread(`Income Category`, `Uninsured: %`) %>%
  rename(
    `Uninsured %: <= 138% of Poverty` = `<= 138% of Poverty`,
    `Uninsured %: <= 400% of Poverty` = `<= 400% of Poverty`,
    `Uninsured %: All Incomes` = `All Incomes`
  )
```

Unemployment

```
bls_unemploy_for <-
  bls_unemploy %>%
  filter(Year == 2016) %>%
  mutate(FIPS = substr(LAUS, 3, 7) %>% as.numeric() %>% as.character()) %>%
  select(FIPS, `Unemployment Rate`)
```

#Medicaid variables (from acs5)

#tidycensus data (how we pulled ACS, medicaiid data etc)

```
# library(tidycensus)
#
# census_api_key("Insert your API Key", install = TRUE, overwrite=TRUE)
# pop_age <- get_acs(geography = "county", variables = c("B01003_001", "B01002_001"), year = 2016)
# pop_age %>% write_csv("~/Downloads/population_and_age.csv")
#
# medicare_vars <-
#   load_variables(2016, "acs5", cache = TRUE) %>%
#   filter(str_detect(concept, "MEDICAID")) %>%
#   pull(name)
#
# medicaid <- get_acs(geography = "county", variables = medicare_vars, year = 2016)
# medicaid %>% write_csv("~/Downloads/medicaid_variables.csv")
#
# acs_vars <- read_xlsx("~/Downloads/acs_vairables.xlsx", sheet = 1)
# acs <- get_acs(geography = "county", variables = acs_vars$name, year = 2016)
# acs %>% write_csv("~/Downloads/acs_variables.csv")
```

#ACS variables

ACS filtering

All disease data- put it together

```
all_disease <-
  all_crude_cdc_atlas %>%
```

```
  filter(year == 2016) %>%
  select(-year) %>%
  rename("FIPS" = CountyFIPS) %>%
  full_join(
    county_health_ranks %>%
      select(-County),
    by = c("FIPS")
  ) %>%
  full_join(
    GBD_diseases,
    by = c("FIPS")
  ) %>%
  full_join(
    opioid_rates %>%
      mutate(FIPS = as.character(FIPS)),
    by = c("FIPS")
  )
```

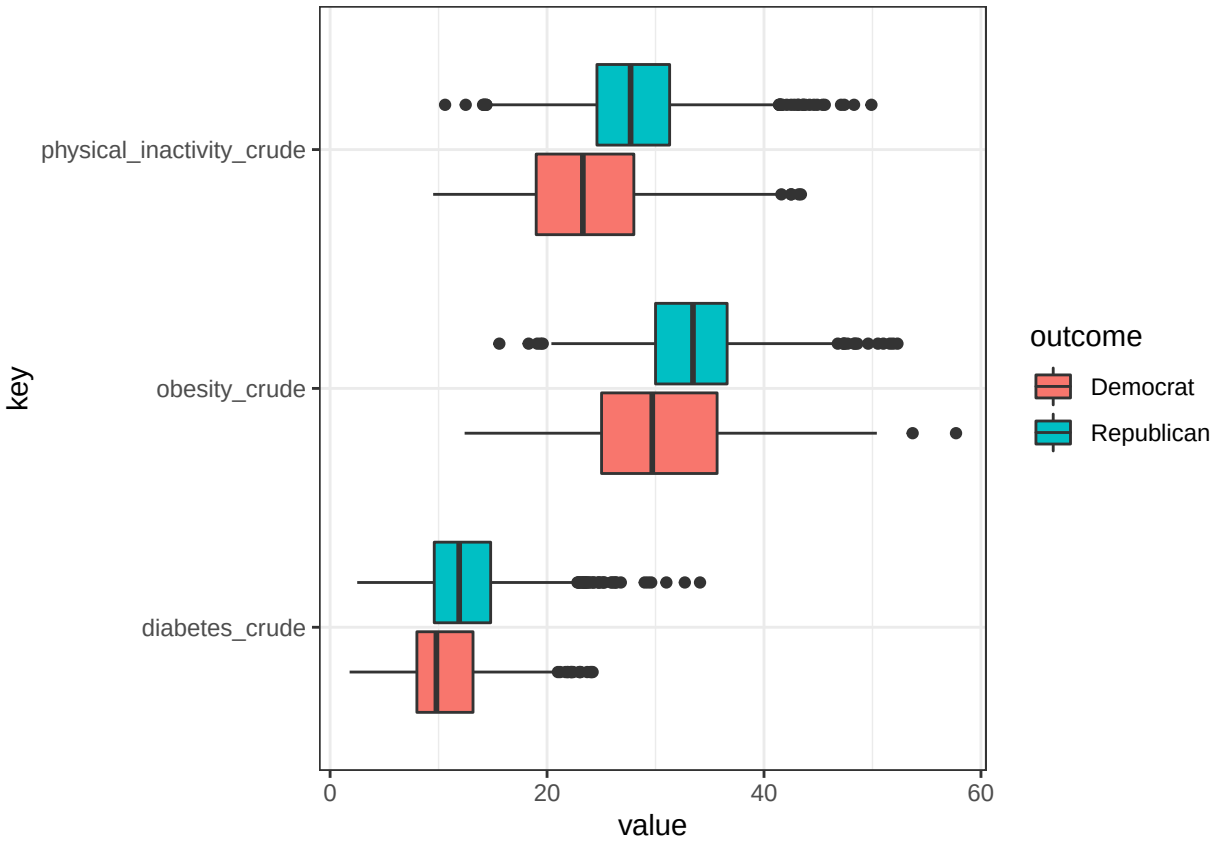
#Correlate all of the public health data with voting

```
vote_county_2012_2016  <-
  vote_county_2012_2016 %>%
  mutate(
    rep_margin_change = (prct_trump_2016 - prct_clinton_2016) - (prct_romney_2012 - prct_obama_2012),
    outcome = if_else(prct_trump_2016 > prct_clinton_2016, "Republican", "Democrat")
  )
```

#Final table for disease variables

#all insurance

Boxplots of CDC data



Categories

```
health_factors_lifestyle <-
  all_disease %>%
  select(
    diabetes_crude:physical_inactivity_crude, `% Fair/Poor`, `% LBW`, `% Smokers`,
    `% Excessive Drinking`, `Teen Birth Rate`, `% Screened`, `% Vaccinated`, `Graduation Rate`,
    `% Some College`, `% Children in Poverty`, `% Single-Parent Households`, `% Severe Housing Problems`

  )

health_access <-
  all_disease %>%
  select(
    `% With Access`, `PCP Rate`, `Dentist Rate`, `MHP Rate`,
  )

#to_fix <- c(`Preventable Hosp. Rate`, )

death_variables <-
  all_disease %>%
  select(
    `Years of Potential Life Lost Rate` : `YPLL Rate (White)`, `Life Expectancy`,
    `Life Expectancy (Black)`, `Life Expectancy (White)`, `Age-Adjusted Mortality`
  )

death_rates <- c("Violent Crime Rate", "Injury Death Rate")
```

```
Other <-
  all_disease %>%
  select(`Physically Unhealthy Days`, `Mentally Unhealthy Days`, `Food Environment Index`, `Income Ratio


GBD_categories <-
  respiratory_diseases %>%
  distinct(disease) %>%
  mutate(category = "Respiratory diseases") %>%
  rbind(
    infectious_diseases %>%
    distinct(disease) %>%
    mutate(category = "Infectious diseases")
  ) %>%
  rbind(
    cvd_diseases %>%
    distinct(disease) %>%
    mutate(category = "Cardiovascular diseases")
  ) %>%
  rbind(
    cancer %>%
    distinct(disease) %>%
    mutate(category = "Cancers")
  ) %>%
  rbind(
    deaths_of_despair %>%
    distinct(disease) %>%
    mutate(category = "Deaths of Despair")
  ) %>%
  rbind(
    le_mortaliity_risk %>%
    distinct(disease) %>%
    mutate(category = "Life expectancy and Mortality")
  )

county_health_category <- read_xlsx(county_health_category_data)
```

All disease (can write)

```
all_disease <-
  all_crude_cdc_atlas %>%
  filter(year == 2016) %>%
  select(-year) %>%
  rename("FIPS" = CountyFIPS) %>%
  full_join(
    county_health_ranks %>%
      select(-County),
    by = c("FIPS")
  ) %>%
  full_join(
    GBD_diseases,
    by = c("FIPS")
  ) %>%
  full_join(
```

```
    opioid_rates %>%
      mutate(FIPS = as.character(FIPS)),
    by = c("FIPS")
  )

#all_disease %>% write_csv("~/Downloads/all_disease_health_political.csv")
```

```
healthcare_var_cats <-
  medicaid_var %>%
  select(-GEOID) %>%
  names() %>%
  as.tibble() %>%
  mutate(category = "Insurance and Healthcare cost") %>%
  rbind(
    no_hs_25_64_ins %>%
    select(-GEOID) %>%
    names() %>%
    as.tibble() %>%
    mutate(category = "Insurance and Healthcare cost")
  ) %>%
  rbind(
    yes_hs_25_64_ins %>%
    select(-GEOID) %>%
    names() %>%
    as.tibble() %>%
    mutate(category = "Insurance and Healthcare cost")
  ) %>%
  rbind(
    bachelor_25_64_ins %>%
    select(-GEOID) %>%
    names() %>%
    as.tibble() %>%
    mutate(category = "Insurance and Healthcare cost")
  ) %>%
  rbind(
    wnh_dis %>%
    select(-GEOID) %>%
    names() %>%
    as.tibble() %>%
    mutate(category = "Insurance and Healthcare cost")
  ) %>%
  rbind(
    w_dis %>%
    select(-GEOID) %>%
    names() %>%
    as.tibble() %>%
    mutate(category = "Insurance and Healthcare cost")
  ) %>%
  rbind(
    sahie_ins %>%
    select(-ID) %>%
    names() %>%
    as.tibble() %>%
    mutate(category = "Insurance and Healthcare cost")
```

```
  ) %>%
  rbind(
    health_costs %>%
    filter(year == 2016) %>%
    select(health_costs_keep) %>%
    rename("GEOID" = `State and County FIPS Code`) %>%
    select(-GEOID) %>%
    names() %>%
    as.tibble() %>%
    mutate(category = "Insurance and Healthcare cost")
  )
```

```
## Warning: `as.tibble()` is deprecated, use `as_tibble()` (but mind the new semantics).
## This warning is displayed once per session.
```

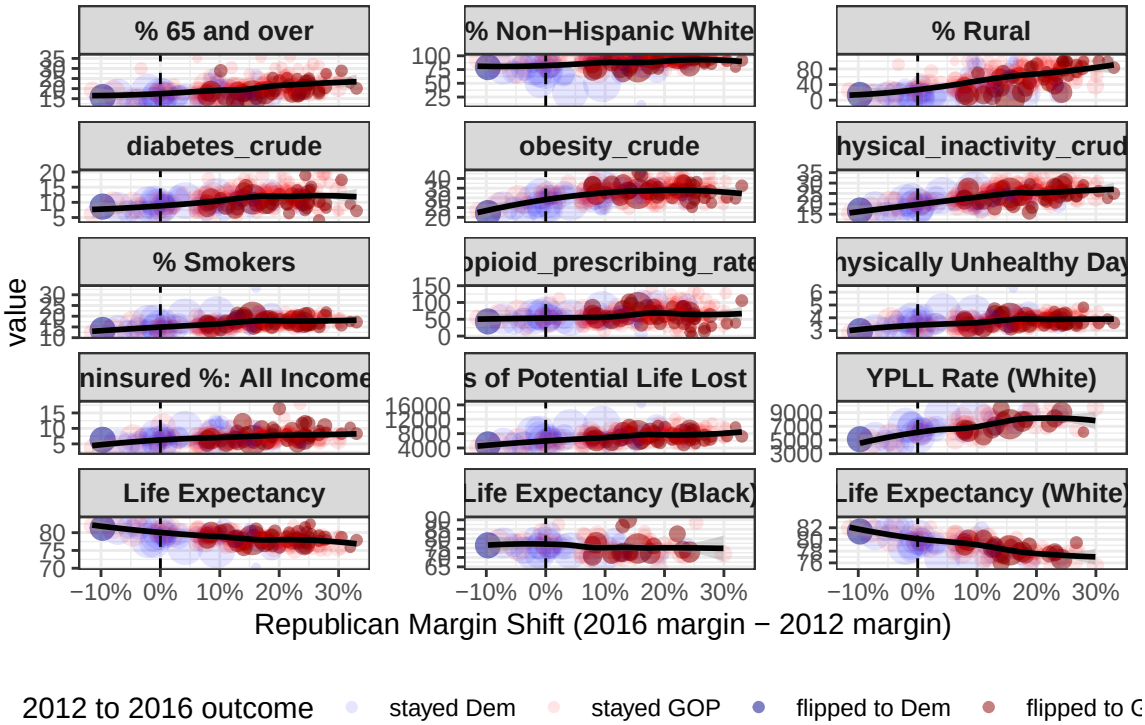
```
all_cats <-
  GBD_categories %>%
  rbind(county_health_category %>% rename("disease" = Disease, "category" = Category)) %>%
  rbind(tibble(disease = "opioid_prescribing_rate", category = "Health Behaviors")) %>%
  rbind(tibble(disease = "diabetes_crude", category = "Health Behaviors")) %>%
  rbind(tibble(disease = "obesity_crude", category = "Health Behaviors")) %>%
  rbind(tibble(disease = "physical_inactivity_crude", category = "Health Behaviors"))
```

#Weighted correlation Table

#Correlation Table (unweighted)

#T tests and quartile differences table Same Analysis as correlations, except now, we compare republican and democrat counties (overall)

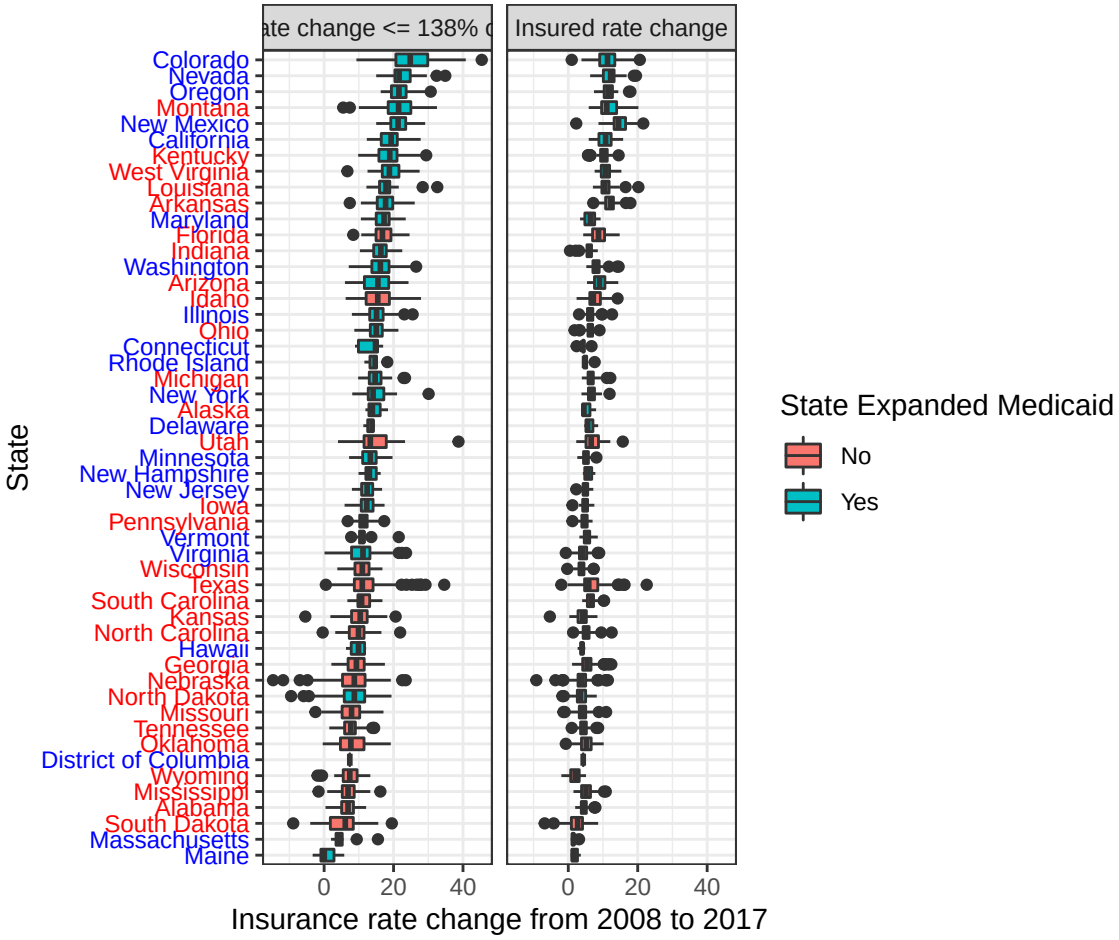
#Plots of republican margin shift and different variables



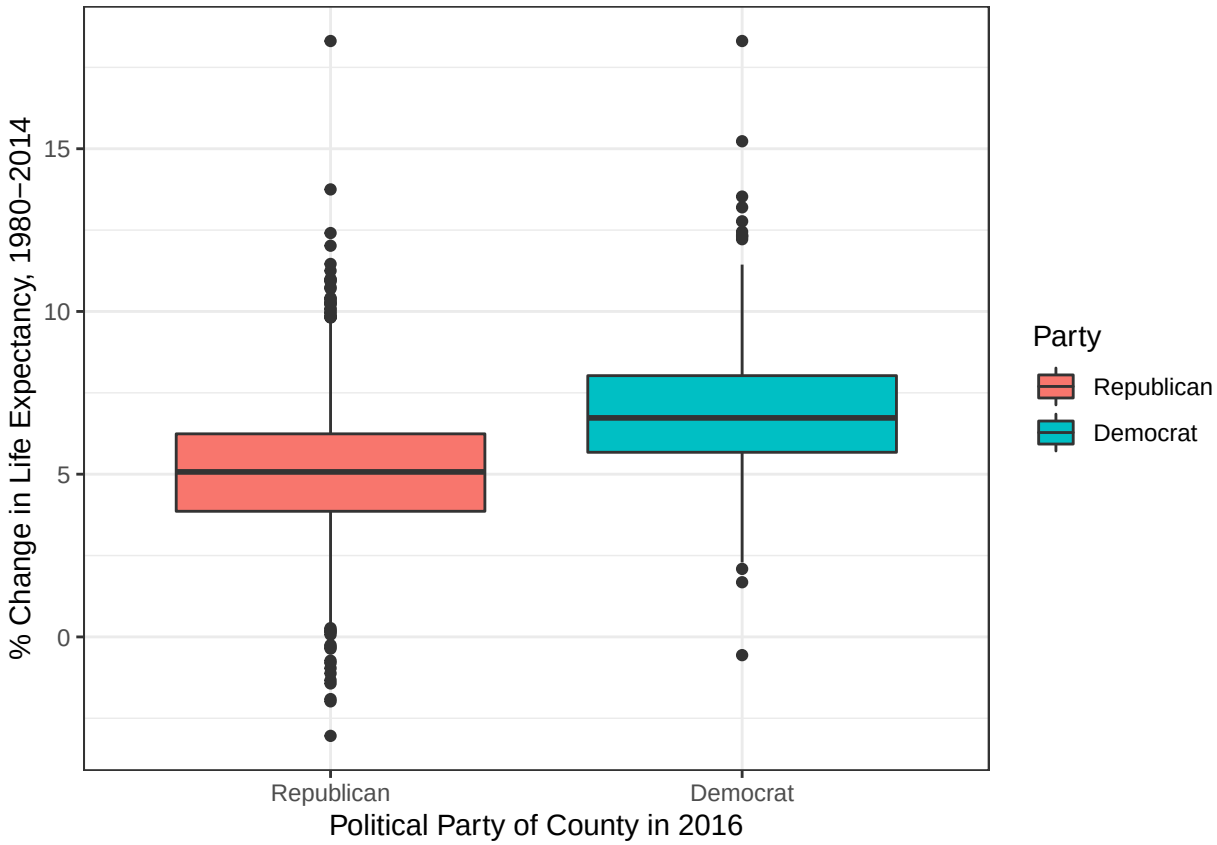
#Compare states on Medicaid expansion- Medicaid data: https://data.medicaid.gov/widgets/n5ce-jxme

#State election data

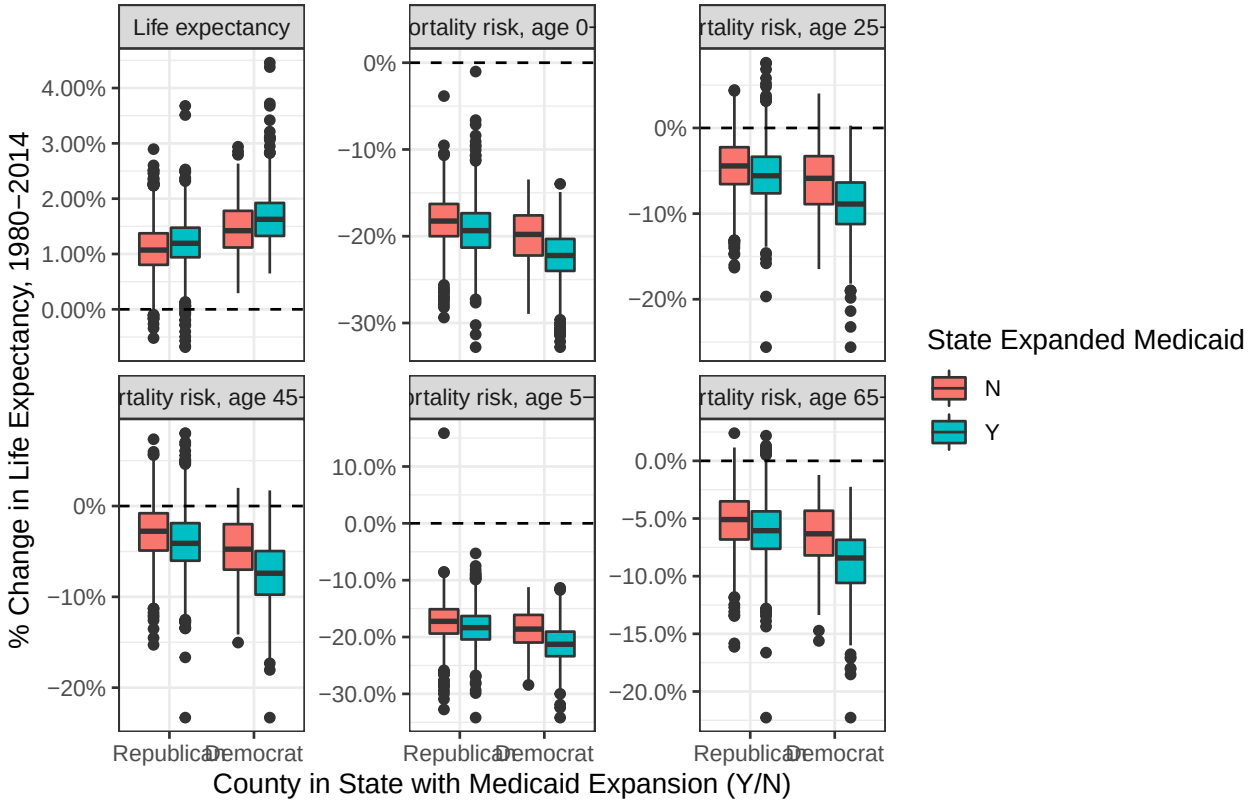
#Plot Insurance rate changes based on medicaid expansion and state presidential voting



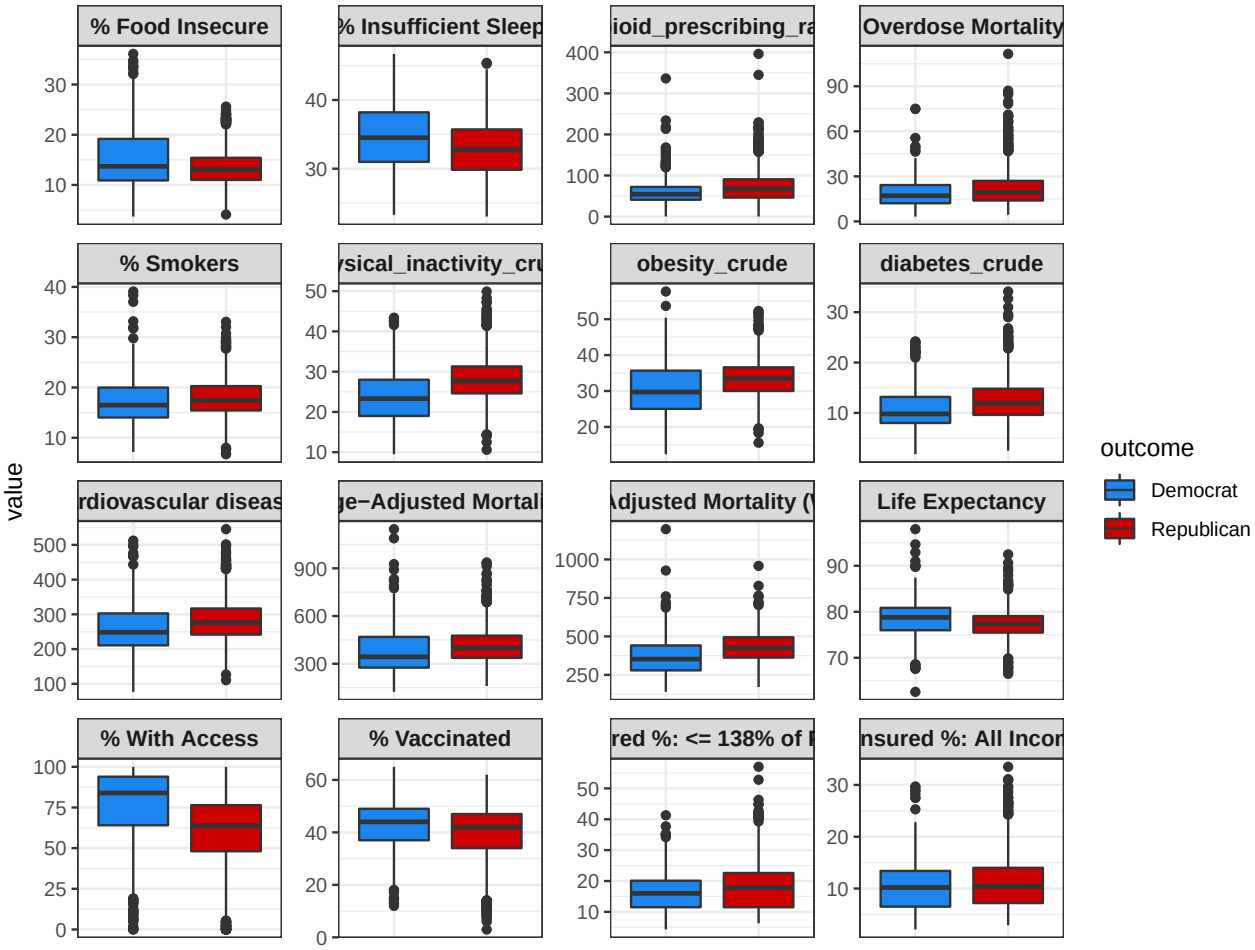
#Plot life expectancy change by party over 1980 to 2014 period



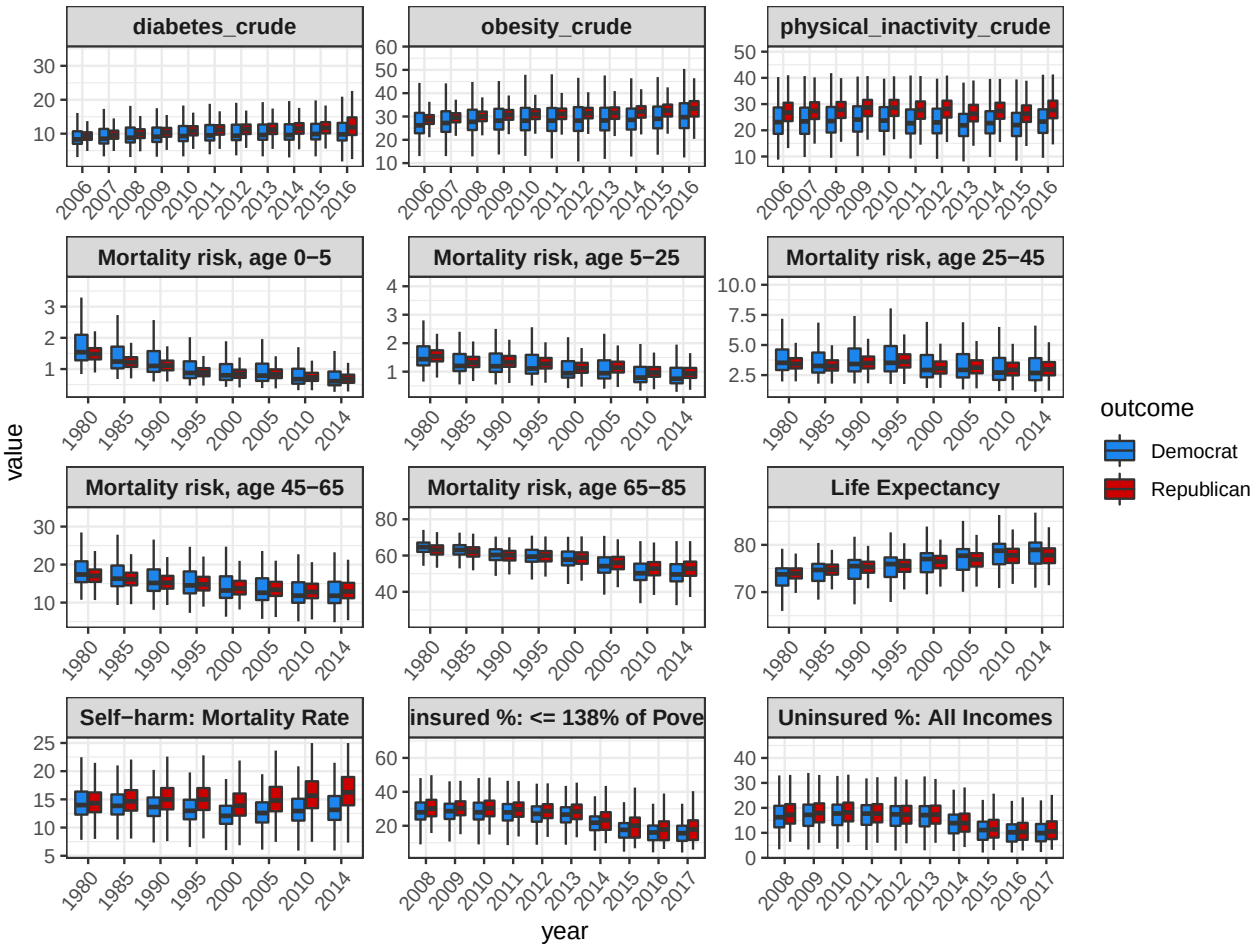
#Mortality risk plots



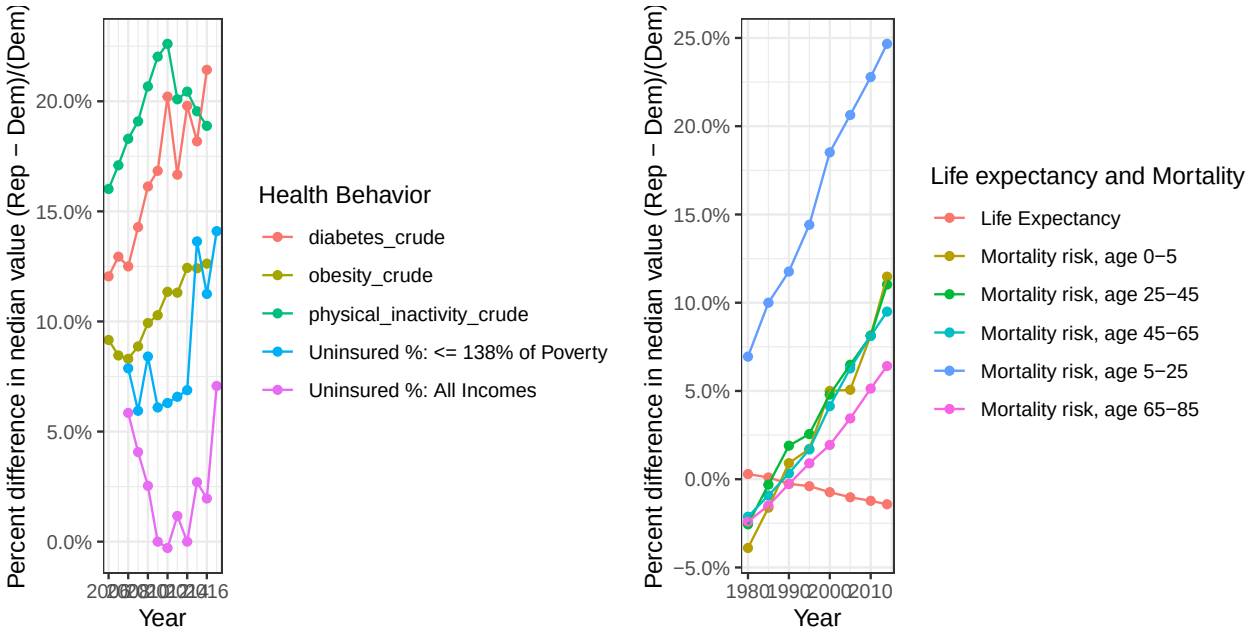
#Plot boxplots of key variables by political party



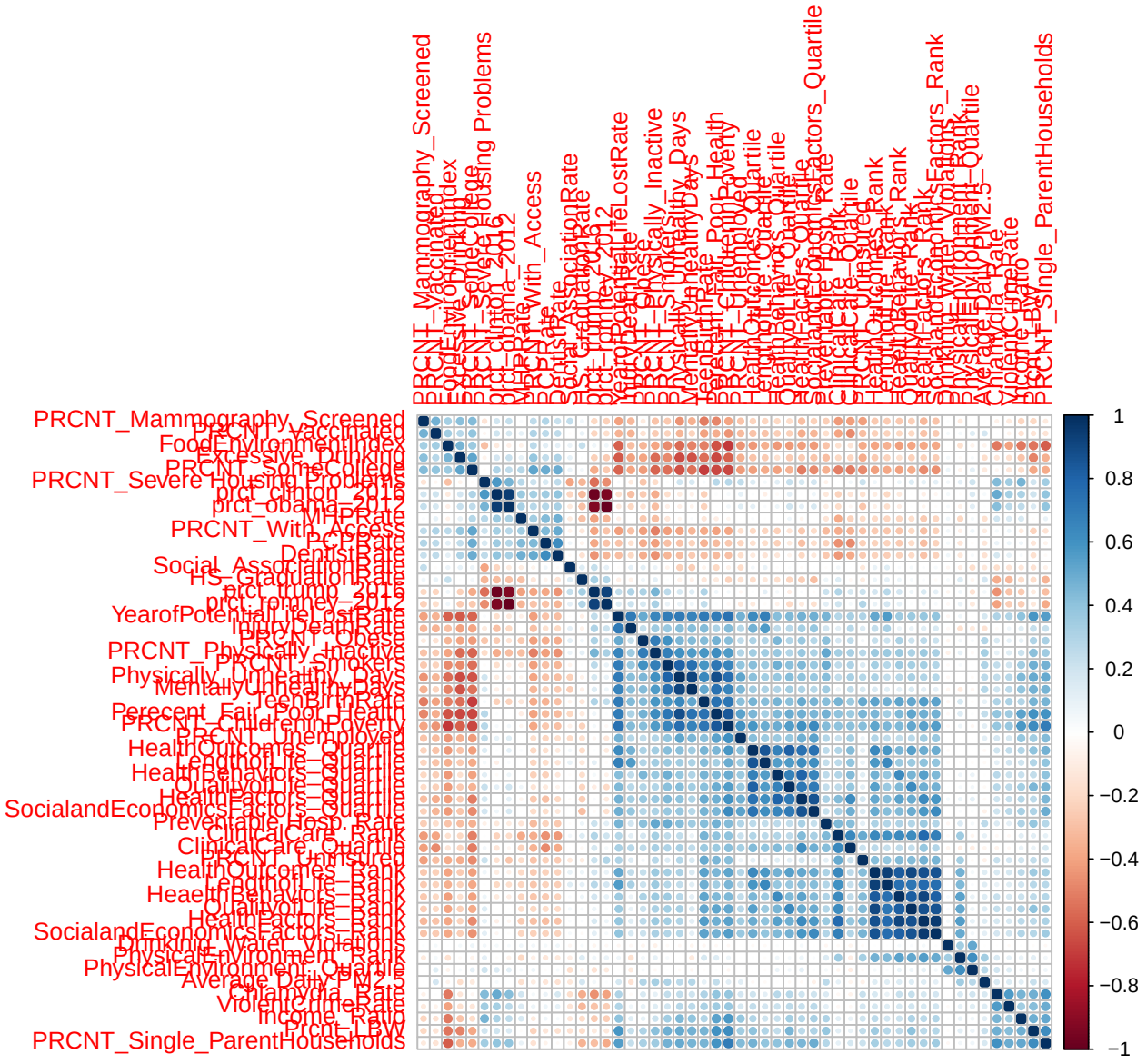
#Plot key variable changes over time Variables that we can capture dynamics for: obesity, physical inactivity, diabetes, Uninsurance in different groups, GDB:Mortality risk of different ages,Deaths of despair over time, Cardiovascularrr disease over time



#Plot of differences between dem and republican values over time (Median)



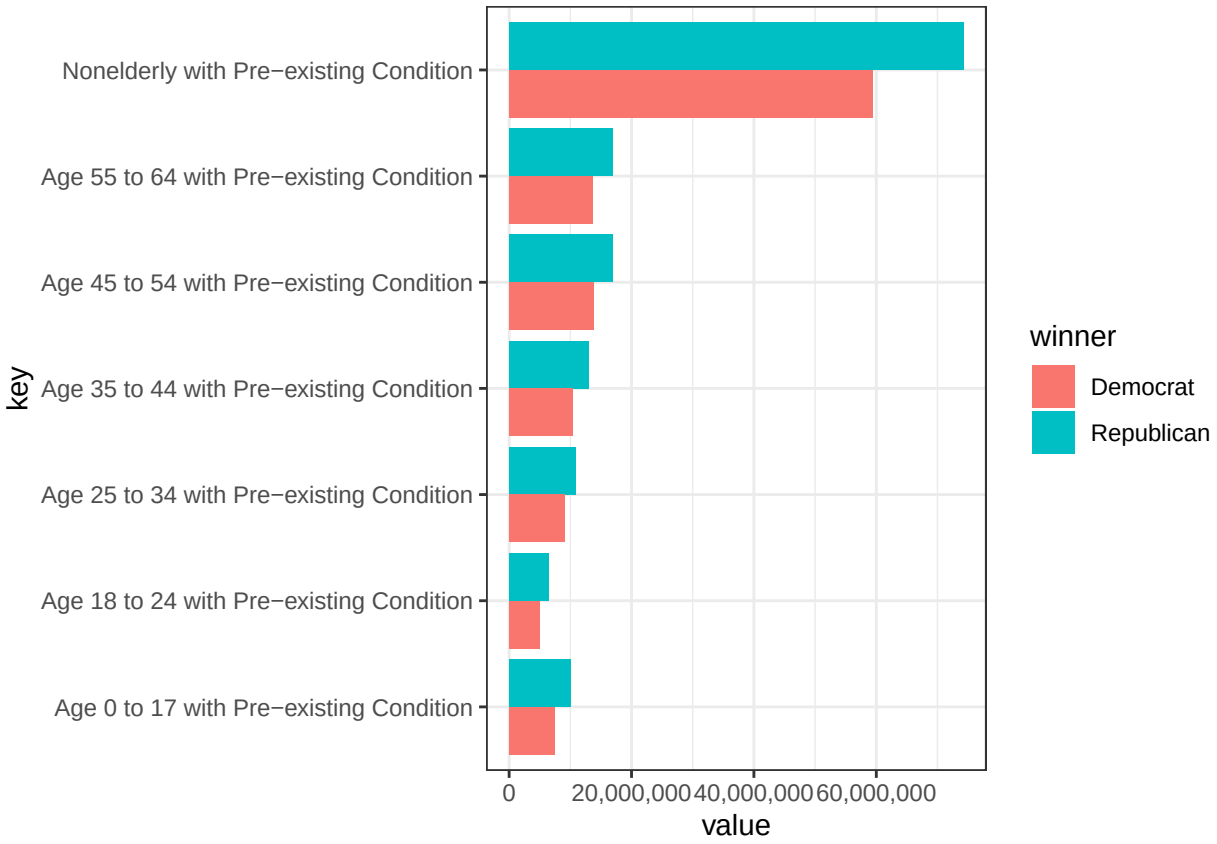
Correlation plot for exploration



#Preexisting conditions by state

```
preexiting_state <- read_xlsx(preexiting_state_data , sheet = 2, skip = 2)

preexiting_state %>%
  filter(State != "National") %>%
  left_join(state_data, by = c("State" = "state")) %>%
  gather(-State, -winner, key = "key", value = "value") %>%
  group_by(winner, key) %>%
  summarise(value = sum(value, na.rm = TRUE)) %>%
  ungroup() %>%
  filter(key != "Percent of Nonelderly with Pre-existing Condition") %>%
  ggplot(aes(key, value, fill = winner)) +
  geom_col(position = "dodge") +
  scale_y_continuous(labels = scales::comma) +
  coord_flip() +
  theme_bw()
```



Absolute preexiting conditions by political party (state presidential)

```
preexiting_state %>%
  filter(State != "National") %>%
  left_join(state_data, by = c("State" = "state")) %>%
  select(State, winner, key = `Nonelderly with Pre-existing Condition`) %>%
  group_by(winner) %>%
  summarise(key = sum(key, na.rm = TRUE)) %>%
  ungroup()
```

```
## # A tibble: 2 x 2
##   winner          key
##   <chr>         <dbl>
## 1 Democrat   59424600
## 2 Republican 74269800
```

Percent preexiting conditions by political party (state presidential)

```
preexiting_state %>%
  filter(State != "National") %>%
  left_join(state_data, by = c("State" = "state")) %>%
  select(State, winner, key = `Percent of Nonelderly with Pre-existing Condition`) %>%
  group_by(winner) %>%
  summarise(key = mean(key, na.rm = TRUE)) %>%
  ungroup()
```

```
## # A tibble: 2 x 2
##   winner       key
##   <chr>      <dbl>
```

```
## 1 Democrat   0.513
## 2 Republican 0.502
```