data_cleaning

Team NO.8

```
#import all the population datasets
for (number in c(2005:2017)) {
  file <- paste(paste("Raw-Data/population/pop",number,sep="-",collapse=NULL),"csv",sep=".",collapse = N
  filename <- paste("pop",number,sep="_",collapse=NULL)
  assign(filename,read.csv(file))
}
```

```
#create a empty dataset to store all the population datasets
population <- tibble()

#keep the columns I need; rename them; add a year indicator for each dataset; append all the datasets
for (number in c(2005:2017)) {
  pop <- paste("pop",number,sep="_",collapse=NULL)
  population <- get(pop) %>%
    select(.,ID=GEO.id2, County=GEO.display.label,Population=HD01_VD01) %>%
    mutate(.,Year=number) %>%
    bind_rows(.,population)
}
```

```
## Warning in bind_rows_(x, .id): Unequal factor levels: coercing to character

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector
```

```
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
```

```r
#import all the Gini coefficients datasets
for (number in c(2006:2017)) {
  file <- paste(paste("Raw-Data/Gini/Gini",number,sep="-",collapse=NULL),"csv",sep=".",collapse = NULL)
  filename <- paste("Gini",number,sep="_",collapse=NULL)
  assign(filename,read.csv(file))
}
```

```r
#create a empty dataset to store all the population datasets
gini <- tibble()

#keep the columns I need; rename them; add a year indicator for each dataset; append all the datasets
for (number in c(2006:2017)) {
  GC <- paste("Gini",number,sep="_",collapse=NULL)
  gini <- get(GC) %>%
    select(.,ID=GEO.id2, County=GEO.display.label,Gini=HD01_VD01) %>%
    mutate(.,Year=number) %>%
    bind_rows(.,gini)
}
```

```
## Warning in bind_rows_(x, .id): Unequal factor levels: coercing to character

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
```

```
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
```

```r
#import all the unemployment rate datasets
for (number in c(2005:2017)) {
  file <- paste(paste("Raw-Data/unemployment/unemp",number,sep="-",collapse=NULL),"csv",sep=".",collaps
  filename <- paste("unemp",number,sep="_",collapse=NULL)
  assign(filename,read.csv(file))
}
```

```r
#create a empty dataset to store all the unemployment rate datasets
unemployment <- tibble()

#keep the columns I need; rename them; add a year indicator for each dataset; append all the datasets
for (number in c(2005:2017)) {
  unemp <- paste("unemp",number,sep="_",collapse=NULL)
  unemployment <- get(unemp) %>%
    select(.,ID=GEO.id2, County=GEO.display.label,UnemploymentRate=HC04_EST_VC01) %>%
    mutate(.,Year=number) %>%
    bind_rows(.,unemployment)
}
```

```
## Warning in bind_rows_(x, .id): Unequal factor levels: coercing to character

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector
```

```
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
```

```
#import all the median income datasets
for (number in c(2005:2017)) {
  file <- paste(paste("Raw-Data/income/inc",number,sep="-",collapse=NULL),"csv",sep=".",collapse = NULL)
  filename <- paste("inc",number,sep="_",collapse=NULL)
  assign(filename,read.csv(file))
}
```

```
#create a empty dataset to store all the median income datasets
income <- tibble()

#keep the columns I need; rename them; add a year indicator for each dataset; append all the datasets
for (number in c(2005:2017)) {
  inc <- paste("inc",number,sep="_",collapse=NULL)
  income <- get(inc) %>%
    select(.,ID=GEO.id2, County=GEO.display.label,MedianIncome=HC02_EST_VC02) %>%
    mutate(.,Year=number) %>%
    bind_rows(.,income)
}
```

```
## Warning in bind_rows_(x, .id): Unequal factor levels: coercing to character

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
```

```
## coercing into character vector
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
```

```
#import all the travel time to work datasets
for (number in c(2005:2017)) {
  file <- paste(paste("Raw-Data/travel_time_to_work/tran",number,sep="-",collapse=NULL),"csv",sep=".",c
  filename <- paste("tran",number,sep="_",collapse=NULL)
  assign(filename,read.csv(file))
}
```

```
#create a empty dataset to store all the travel time to work datasets
transportation <- tibble()

#keep the columns I need; rename them; add a year indicator for each dataset; append all the datasets
for (number in c(2005:2017)) {
  tran <- paste("tran",number,sep="_",collapse=NULL)

  if (number ==2005) {
    transportation <- get(tran) %>%
    select(.,ID=GEO.id2, County=GEO.display.label,TravelTime=HC01_EST_VC122) %>%
    mutate(.,Year=number) %>%
    bind_rows(.,transportation)
  }
  if (number %in% c(2006:2009)) {
    transportation <- get(tran) %>%
    select(.,ID=GEO.id2, County=GEO.display.label,TravelTime=HC01_EST_VC104) %>%
    mutate(.,Year=number) %>%
    bind_rows(.,transportation)
  }
  if (number %in% c(2010:2012)) {
    transportation <- get(tran) %>%
    select(.,ID=GEO.id2, County=GEO.display.label,TravelTime=HC01_EST_VC120) %>%
    mutate(.,Year=number) %>%
    bind_rows(.,transportation)
  }
  if (number %in% c(2013:2017)) {
    transportation <- get(tran) %>%
    select(.,ID=GEO.id2, County=GEO.display.label,TravelTime=HC01_EST_VC118) %>%
    mutate(.,Year=number) %>%
    bind_rows(.,transportation)
  }
}
```

```
## Warning in bind_rows_(x, .id): Unequal factor levels: coercing to character

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector
```

```
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
```

```
#import all the poverty rate datasets
for (number in c(2005:2017)) {
  file <- paste(paste("Raw-Data/poverty/poverty",number,sep="-",collapse=NULL),"csv",sep=".",collapse =
  filename <- paste("poverty",number,sep="_",collapse=NULL)
  assign(filename,read.csv(file))
}
```

```r
#create a empty dataset to store all the poverty rate datasets
poverty <- tibble()

#keep the columns I need; rename them; add a year indicator for each dataset; append all the datasets
for (number in c(2005:2017)) {
  pov <- paste("poverty",number,sep="_",collapse=NULL)
  poverty <- get(pov) %>%
    select(.,ID=GEO.id2, County=GEO.display.label,PovertyRate=HC03_EST_VC01) %>%
    mutate(.,Year=number) %>%
    bind_rows(.,poverty)
}
```

```
## Warning in bind_rows_(x, .id): Unequal factor levels: coercing to character

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
```

```
## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
```

```
#import all the housing costs without mortgage datasets
for (number in c(2005:2017)) {
  file <- paste(paste("Raw-Data/housing_costs_no_mortgage/costs",number,sep="-",collapse=NULL),"csv",sep
  filename <- paste("costs",number,sep="_",collapse=NULL)
  assign(filename,read.csv(file))
}
```

```
#create a empty dataset to store all the housing costs without mortgage datasets
costs <- tibble()

#keep the columns I need; rename them; add a year indicator for each dataset; append all the datasets
for (number in c(2005:2017)) {
  cost <- paste("costs",number,sep="_",collapse=NULL)

  if (number %in% c(2010:2014)) {
    costs <- get(cost) %>%
    select(.,ID=GEO.id2, County=GEO.display.label,HousingCosts=HC01_EST_VC42) %>%
    mutate(.,Year=number) %>%
    bind_rows(.,costs)
  } else {
    costs <- get(cost) %>%
    select(.,ID=GEO.id2, County=GEO.display.label,HousingCosts=HC01_EST_VC39) %>%
    mutate(.,Year=number) %>%
    bind_rows(.,costs)
  }
}
```

```
## Warning in bind_rows_(x, .id): Unequal factor levels: coercing to character

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
```

```
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
```

```r
#import all the housing costs with mortgage datasets
for (number in c(2005:2017)) {
  file <- paste(paste("Raw-Data/housing_costs_mortgage/mortgage",number,sep="-",collapse=NULL),"csv",sep
  filename <- paste("mortgage",number,sep="_",collapse=NULL)
  assign(filename,read.csv(file))
}
```

```r
#create a empty dataset to store all the housing costs with mortgage datasets
mortgage <- tibble()

#keep the columns I need; rename them; add a year indicator for each dataset; append all the datasets
for (number in c(2005:2017)) {
  mort <- paste("mortgage",number,sep="_",collapse=NULL)

  if (number <=2009) {
    mortgage <- get(mort) %>%
    select(.,ID=GEO.id2, County=GEO.display.label,MortgageCosts=HC01_EST_VC47) %>%
    mutate(.,Year=number) %>%
    bind_rows(.,mortgage)
  }
  if (number %in% c(2010:2014)) {
    mortgage <- get(mort) %>%
    select(.,ID=GEO.id2, County=GEO.display.label,MortgageCosts=HC01_EST_VC51) %>%
    mutate(.,Year=number) %>%
    bind_rows(.,mortgage)
  }
  if (number >=2015) {
    mortgage <- get(mort) %>%
    select(.,ID=GEO.id2, County=GEO.display.label,MortgageCosts=HC01_EST_VC48) %>%
    mutate(.,Year=number) %>%
    bind_rows(.,mortgage)
  }
}
```

```
## Warning in bind_rows_(x, .id): Unequal factor levels: coercing to character

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
```

```
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
```

```
#import all the fertility datasets
for (number in c(2005:2017)) {
  file <- paste(paste("Raw-Data/fertility/fertility",number,sep="-",collapse=NULL),"csv",sep=".",collaps
  filename <- paste("fertility",number,sep="_",collapse=NULL)
  assign(filename,read.csv(file))
}
```

```
#create a empty dataset to store all the poverty rate datasets
fertility <- tibble()

#keep the columns I need; rename them; add a year indicator for each dataset; append all the datasets
for (number in c(2005:2017)) {
  fer <- paste("fertility",number,sep="_",collapse=NULL)
```

```
  fertility <- get(fer) %>%
    select(.,ID=GEO.id2, County=GEO.display.label,Fertility=HC04_EST_VC01) %>%
    mutate(.,Year=number) %>%
    bind_rows(.,fertility)
}
```

```
## Warning in bind_rows_(x, .id): Unequal factor levels: coercing to character

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
```

```
## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
```

```
#import all the bachelor rate datasets
for (number in c(2005:2017)) {
  file <- paste(paste("Raw-Data/education_attainment/bachelor",number,sep="-",collapse=NULL),"csv",sep=
  filename <- paste("bachelor",number,sep="_",collapse=NULL)
  assign(filename,read.csv(file))
}
```

```
#create a empty dataset to store all the bachelor rate datasets
bachelor <- tibble()

#keep the columns I need; rename them; add a year indicator for each dataset; append all the datasets
for (number in c(2005:2017)) {
  bac <- paste("bachelor",number,sep="_",collapse=NULL)

  if (number ==2005) {
    bachelor <- get(bac) %>%
    select(.,ID=GEO.id2, County=GEO.display.label,BachelorRate=HC02_EST_VC19) %>%
    mutate(.,Year=number) %>%
    bind_rows(.,bachelor)
  }
  if (number %in% c(2006:2009)) {
    bachelor <- get(bac) %>%
    select(.,ID=GEO.id2, County=GEO.display.label,BachelorRate=HC02_EST_VC12) %>%
    mutate(.,Year=number) %>%
    bind_rows(.,bachelor)
  }
  if (number %in% c(2010:2014)) {
    bachelor <- get(bac) %>%
    select(.,ID=GEO.id2, County=GEO.display.label,BachelorRate=HC02_EST_VC13) %>%
    mutate(.,Year=number) %>%
    bind_rows(.,bachelor)
  }
  if (number >= 2015) {
    bachelor <- get(bac) %>%
    select(.,ID=GEO.id2, County=GEO.display.label,BachelorRate=HC02_EST_VC14) %>%
    mutate(.,Year=number) %>%
    bind_rows(.,bachelor)
  }
}
```

```
## Warning in bind_rows_(x, .id): Unequal factor levels: coercing to character

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector

## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector
```

```
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
```

```
## Warning in bind_rows_(x, .id): binding factor and character vector,
## coercing into character vector
```

```
## Warning in bind_rows_(x, .id): binding character and factor vector,
## coercing into character vector
```

```
#join all the datasets
data <-
population %>%
  #inner_join(.,gini,by=c("ID","Year","County")) %>%
  inner_join(., unemployment, by=c("ID","Year","County")) %>%
  inner_join(.,income, by=c("ID","Year","County")) %>%
  #inner_join(.,transportation, by=c("ID","Year","County")) %>%
  inner_join(.,poverty, by=c("ID","Year","County")) %>%
  inner_join(.,costs, by=c("ID","Year","County")) %>%
  inner_join(.,mortgage, by=c("ID","Year","County")) %>%
  inner_join(.,fertility, by=c("ID","Year","County")) %>%
  inner_join(.,bachelor, by=c("ID","Year","County")) %>%
  arrange(.,Year)
```

```
#add variables to record the change by year
data2 <-
data %>%
  group_by(ID,County) %>%
  mutate(.,PreviousPopulation=lag(Population, n=1, order_by = Year),
         PopulationGrowth=(Population-PreviousPopulation)/PreviousPopulation,
         PreviousUnemploymentRate=lag(UnemploymentRate,n=1,order_by = Year),
         UnemploymentChange=(UnemploymentRate-PreviousUnemploymentRate)/PreviousUnemploymentRate,
         PreviousIncome=lag(MedianIncome,n=1,order_by=Year),
         MedianIncomeChange=(MedianIncome-PreviousIncome)/PreviousIncome,
         PreviousPoverty=lag(PovertyRate,n=1,order_by=Year),
         PovertyRateChange=(PovertyRate-PreviousPoverty)/PreviousPoverty,
         PreviousHousingCosts=lag(HousingCosts,n=1,order_by=Year),
         HousingCostsChange=(HousingCosts-PreviousHousingCosts)/PreviousHousingCosts,
         PreviousMortgageCosts=lag(MortgageCosts,n=1,order_by=Year),
         MortgageCostsChange=(MortgageCosts-PreviousMortgageCosts)/PreviousMortgageCosts,
         PreviousFertility=lag(Fertility,n=1,order_by=Year),
         FertilityChange=(Fertility-PreviousFertility)/PreviousFertility,
         PreviousBachelor=lag(BachelorRate,n=1,order_by=Year),
         BachelorRateChange=(BachelorRate-PreviousBachelor)/PreviousBachelor,
         ) %>%
  ungroup %>%
  select(-PreviousBachelor,-PreviousFertility,-PreviousIncome,-PreviousHousingCosts,-PreviousPoverty,-P
  filter(.,!Year==2005) %>%
  arrange(.,Year)
```

```
data3 <- data2 %>%
  inner_join(.,gini,by=c("ID","Year","County"))
```

```
write_csv(data3,path ="project_data.csv")
```