```
---
title: "Econ 151 HW 1"
author: "Casey Shu"
date: "2023-09-20"
output:
  html_document: default
  pdf_document: default
---
```

QUESTION 1

This is the code to calculate the mean, min, max, sd, and median of 1998 and 2008 quarterly expenditures, as well as the table associated with it.

```{r}
# Load required libraries
library(pacman)
library(dplyr)
library(tidyverse)

# Use pacman to ensure tidyverse packages are installed and loaded
pacman::p_load(tidyverse)
setwd("/Users/caseyshu/Desktop/Econ 151 HW 1")
enigh98 <- readRDS('enigh98.rds')
enigh08 <- readRDS('enigh08.rds')
mean98 <- mean(enigh98$quarterly_exp_norm)
min98 <- min(enigh98$quarterly_exp_norm)
max98 <- max(enigh98$quarterly_exp_norm)
sd98 <- sd(enigh98$quarterly_exp_norm)
median98 <- median(enigh98$quarterly_exp_norm)
mean08 <- mean(enigh08$quarterly_exp)
min08 <- min(enigh08$quarterly_exp)
max08 <- max(enigh08$quarterly_exp)
sd08 <- sd(enigh08$quarterly_exp)
median08 <- median(enigh08$quarterly_exp)
QuarterlyExpendituresSummary<- data.frame(
  Year = c("2008", "1998"),
  Mean = c(mean08, mean98),
  Min = c(min08, min98),
  Max = c(max08, max98),
  SD = c(sd08, sd98),
  Median = c(median08, median98)
)
print(QuarterlyExpendituresSummary)
```

```{r}
totalConsumption98 <- mean(enigh98$quarterly_exp_norm)
totalConsumption08 <- mean(enigh08$quarterly_exp)
AvgAnnualRateofGrowthofRealConsumption <- ((totalConsumption08 /
totalConsumption98)^(1/10)) - 1
print(AvgAnnualRateofGrowthofRealConsumption)
```
-0.03879378 is the growth rate.

QUESTION 3

```{r}
consumption98USD <- enigh98
consumption08USD <- enigh08

consumption98USD$quarterly_exp <- consumption98USD$quarterly_exp/8.96

consumption08USD$quarterly_exp <- consumption08USD$quarterly_exp/10.39
```

```
numberhouseholds98 <- length(consumption98USD$quarterly_exp)
numberhouseholds08 <- length(consumption08USD$quarterly_exp)

N98 <- numberhouseholds98 * 4
N08 <- numberhouseholds08 * 4


P098_1_dollar <- (1/N98) * sum(consumption98USD$quarterly_exp < 90) * 4
P098_3_dollar <- (1/N98) * sum(consumption98USD$quarterly_exp < 270) * 4

P008_1_dollar <- (1/N08) * sum(consumption08USD$quarterly_exp < 90) * 4
P008_3_dollar <- (1/N08) * sum(consumption08USD$quarterly_exp < 270) * 4

print(P098_1_dollar)
print(P098_3_dollar)
print(P008_1_dollar)
print(P008_3_dollar)
```

```
The poverty headcount ratio for 1 dollar in 1998 is 0.02198809 and for 2008 it is
0.039549. The poverty headcount ratio for 3 dollars in 1998 is 0.1149794 and for 2008 it
is 0.1644406. The headcount ratio increased for both 1 and 3 dollar thresholds during this
period.

QUESTION 4

```{r}
consumption98USD <- enigh98
consumption08USD <- enigh08

consumption98USD$quarterly_exp <- consumption98USD$quarterly_exp/8.96
consumption08USD$quarterly_exp <- consumption08USD$quarterly_exp/10.39

#make new tables so we can subtract later
poverty_90_98 <- consumption98USD
poverty_270_98 <- consumption98USD

poverty_90_08 <- consumption08USD
poverty_270_08 <- consumption08USD

#filter out the people under the poverty line
poverty_90_98 <- subset(poverty_90_98, quarterly_exp < 90)
poverty_270_98 <- subset(poverty_270_98, quarterly_exp < 270)

poverty_90_08 <- subset(poverty_90_08, quarterly_exp < 90)
poverty_270_08 <- subset(poverty_270_08, quarterly_exp < 270)

#subtract the incomes from the poverty line and divide by poverty line
poverty_90_98$quarterly_exp <- ((90 - poverty_90_98$quarterly_exp) / 90)
poverty_270_98$quarterly_exp <- ((270 - poverty_270_98$quarterly_exp) /270)

poverty_90_08$quarterly_exp <- ((90 - poverty_90_08$quarterly_exp) / 90)
poverty_270_08$quarterly_exp <- ((270 - poverty_270_08$quarterly_exp) /270)

#total households
total_households98 <- length(enigh98$household_id)
total_households08 <- length(enigh08$household_id)

#sum each income and divide by number of people
P_gap_90_98 <- (sum(poverty_90_98$quarterly_exp) / total_households98)
P_gap_270_98 <- (sum(poverty_270_98$quarterly_exp) / total_households98)

P_gap_90_08 <- (sum(poverty_90_08$quarterly_exp) / total_households08)
P_gap_270_08 <- (sum(poverty_270_08$quarterly_exp) / total_households08)
```
```

```
print(P_gap_90_98)
print(P_gap_270_98)
print(P_gap_90_08)
print(P_gap_270_08)
```
The poverty gap under 90 dollars is 0.008932724 in 1998 and 0.01522148 in 2008. The poverty gap under 270 dollars is 0.04603279 in 1998 and 0.07030276 in 2008. The poverty gap overall increased from 1998 to 2008.

QUESTION 5

```{r}
poverty_90_98_p2 <- consumption98USD
poverty_270_98_p2 <- consumption98USD

poverty_90_08_p2 <- consumption08USD
poverty_270_08_p2 <- consumption08USD

#filter out the people under the poverty line
poverty_90_98_p2 <- subset(poverty_90_98_p2, quarterly_exp < 90)
poverty_270_98_p2 <- subset(poverty_270_98_p2, quarterly_exp < 270)

poverty_90_08_p2 <- subset(poverty_90_08_p2, quarterly_exp < 90)
poverty_270_08_p2 <- subset(poverty_270_08_p2, quarterly_exp < 270)

poverty_90_98_p2$quarterly_exp <- ((90 - poverty_90_98_p2$quarterly_exp) / 90)
poverty_270_98_p2$quarterly_exp <- ((270 - poverty_270_98_p2$quarterly_exp) /270)

poverty_90_08_p2$quarterly_exp <- ((90 - poverty_90_08_p2$quarterly_exp) / 90)
poverty_270_08_p2$quarterly_exp <- ((270 - poverty_270_08_p2$quarterly_exp) /270)

P2_gap_90_98 <- sum((poverty_90_98_p2$quarterly_exp)^2) / total_households98
P2_gap_270_98 <- sum((poverty_270_98_p2$quarterly_exp)^2) / total_households98

P2_gap_90_08 <- sum((poverty_90_08_p2$quarterly_exp)^2) / total_households08
P2_gap_270_08 <- sum((poverty_270_08_p2$quarterly_exp)^2) / total_households08

print(P2_gap_90_98)
print(P2_gap_270_98)
print(P2_gap_90_08)
print(P2_gap_270_08)
```
P2 was 0.5236758 in 1998 for 90 dollars and 0.008179018 in 2008. P2 was 0.02636182 in 1998 for 90 dollars and 0.04160283 in 2008. For both poverty lines, P2 increased from 1998 to 2008.

QUESTION 6
```{r}
all_sums98 <- lapply(enigh98$quarterly_exp_norm,
                     function(x){enigh98$quarterly_exp_norm - x}) %>%
  set_names(enigh98$household_id) %>%
  bind_rows()

#isolate lower triangle 98
lower_triangle98 <- all_sums98[lower.tri(all_sums98)]
abs_lower_tri98 <- abs(lower_triangle98)
sum_triangle98 <- sum(abs_lower_tri98)

giniN98 <- length(enigh98$household_id)
giniX98 <- mean(enigh98$quarterly_exp_norm, na.rm = TRUE)

gini98 <- (1/((giniN98)^2 * giniX98)) * sum_triangle98
gini_98_pop <- (1/(N98 * (N98 - 1)*giniX98)) * sum_triangle98
```

```
print(gini98)
```
```{r}
all_sums08 <- lapply(enigh08$quarterly_exp,
                     function(x){enigh08$quarterly_exp - x}) %>%
  set_names(enigh08$household_id) %>%
  bind_rows()

#isolate lower triangle 98
lower_triangle08 <- all_sums08[lower.tri(all_sums08)]
abs_lower_tri08 <- abs(lower_triangle08)
sum_triangle08 <- sum(abs_lower_tri08)

giniN08 <- length(enigh08$household_id)
giniX08 <- mean(enigh08$quarterly_exp, na.rm = TRUE)

gini08 <- (1/((giniN08)^2 * giniX08)) * sum_triangle08
gini_08_pop <- (1/(N08 * (N08 - 1)*giniX08)) * sum_triangle08
print(gini08)
```
According to the gini coefficient, the inequlity did increase from 0.4791075 to 0.541611 in this period.

QUESTION 7
```{r}
sorted_consumption98 <- enigh98[order(-enigh98$quarterly_exp_norm), ]
total_consumption98 <- sum(enigh98$quarterly_exp_norm)
total_households98 <- length(enigh98$household_id)
top_1_percent_count98 <- round(.01 * total_households98)
threshold_consumption <- sorted_consumption98$quarterly_exp_norm[top_1_percent_count98]
top_1_percent_consumption98 <-
sum(sorted_consumption98$quarterly_exp_norm[1:top_1_percent_count98])
top_1_percent_share98 <- (top_1_percent_consumption98 / total_consumption98) * 100
#this is the top 1 percent share in 1998
print(top_1_percent_share98)
```
```{r}
sorted_consumption08 <- enigh08[order(-enigh08$quarterly_exp), ]
total_consumption08 <- sum(enigh08$quarterly_exp)
total_households08 <- length(enigh08$household_id)
top_1_percent_count08 <- round(.01 * total_households08)
threshold_consumption <- sorted_consumption08$quarterly_exp[top_1_percent_count08]
top_1_percent_consumption08 <-
sum(sorted_consumption08$quarterly_exp[1:top_1_percent_count08])
top_1_percent_share08 <- (top_1_percent_consumption08 / total_consumption08) * 100
#this is the top 1 percent share in 2008
print(top_1_percent_share08)
```
The inequality grew by ~3% from 1998 to 2008 according to the 1% share.

QUESTION 8

```{r}
#top 20% in 1998
top_20_percent_count98 <- round(.2 * total_households98)
top_20_percent_consumption98 <-
sum(sorted_consumption98$quarterly_exp_norm[1:top_20_percent_count98])

#bottom 20% in 1998
sorted_consumption_ascending98 <- enigh98[order(enigh98$quarterly_exp_norm), ]
bottom_20_percent_count98 <- round(.2 * total_households98)
threshold_consumption98 <-
sorted_consumption_ascending98$quarterly_exp_norm[bottom_20_percent_count98]
bottom_20_percent_consumption98 <-
```

```
sum(sorted_consumption_ascending98$quarterly_exp_norm[1:bottom_20_percent_count98])

print(top_20_percent_consumption98 / bottom_20_percent_consumption98)
```
13.45245 is the 20/20 Kuznets ratio for 1998.

```{r}
#top 20% in 2008
top_20_percent_count08 <- round(.2 * total_households08)
top_20_percent_consumption08 <-
sum(sorted_consumption08$quarterly_exp[1:top_20_percent_count08])

#bottom 20% in 2008
sorted_consumption_ascending08 <- enigh08[order(enigh08$quarterly_exp), ]
bottom_20_percent_count08 <- round(.2 * total_households08)
bottom_20_percent_consumption08 <-
sum(sorted_consumption_ascending08$quarterly_exp[1:bottom_20_percent_count08])

print(top_20_percent_consumption08 / bottom_20_percent_consumption08)
```

The 20/20 Kuznets Ratio was 13.45245 in 1998, and 20.51476 in 2008. Thus, according to this measure, inequality grew over this time period.

QUESTION 9

Our results from questions 6,7, and 8 is not sensitive to how we measured inequality because all these ways of measuring inequality are painting the same picture for inequality in Mexico from 1998 to 2008, which is that inequality was growing largely over this time period.

QUESTION 10

```{r}
gini98 <- 0.4791075
gini08 <- 0.541611

mean_exp_98 <- mean(enigh98$quarterly_exp_norm)
mean_exp_08 <- mean(enigh08$quarterly_exp)

W98 <- mean_exp_98*(1-gini98)
W08 <- mean_exp_08*(1-gini08)

print(W98)
print(W08)
```
No matter what μ is, the welfare in 1998 is always great based off the gini coefficients we calculated in question 6. Thus, social welfare decreased from 1998 to 2008.

QUESTION 11

```{r}
log_consumption98 <- enigh98
#can only take the log of positive numbers, so have to filter out negatives
log_consumption98$quarterly_exp_norm <- ifelse(log_consumption98$quarterly_exp_norm <= 0,
1, log_consumption98$quarterly_exp_norm)
log_consumption98$quarterly_exp_norm <- log(log_consumption98$quarterly_exp_norm)
sum_ln_xi98 <- sum(log_consumption98$quarterly_exp_norm)
social_welfare98 <- (1/total_households98) * sum_ln_xi98

log_consumption08 <- enigh08
log_consumption08$quarterly_exp <- log(log_consumption08$quarterly_exp)
sum_ln_xi08 <- sum(log_consumption08$quarterly_exp)
social_welfare08 <- (1/total_households08) * sum_ln_xi08
```

```
print(social_welfare98)
print(social_welfare08)
```

The social welfare in 1998 was 9.444256 and in 2008 it was 8.915056. Thus, social welfare decreased from 1998 to 2008.

QUESTION 12

Based on questions in 10 and 11 our understanding of of social welfare in Mexico is not sensitive to how we define social welfare. This is similar to question 9, where our numbers and measures from 10 and 11 point in the same direction. They both show a decrease in social welfare over this time period. While both formulas capture the same trends, our team has a preference for the measure of social welfare that accounts for inequality using the Gini coefficient. We consider this measure a bit more illustrative because it ensures that welfare and inequality are not independent ideas - implying that societies with less inequality have a large sense of social welfare.