Problem Set 1

Santiago Taborga

2025-01-29

1. I chose the country of Madagascar. There are 1200 respondents and the interviews were conducted between 2022-04-26 and 2022-05-22.

2. Respondents have a broad set of backgrounds. People's age ranges from 18 to 88 years old, with an mean age of 38.77 and a median age of 36. There are 601 men and 599 women. 47.08% of people speak Malagasy officiel, 52.83% of people speak Malagasy avec spécificité régionale, and 0.08% of people speak another language. 99.58% of people are Black/African, 0.17% of people are White/European, and 0.25% of people are Coloured/Mixed Race.

```r
tab1 <-
  madagascar %>%
  select(Q2, Q100, Q101, Per) %>%
  pivot_longer(cols = -Per) %>%
  group_by(name, value) %>%
  summarise(Percent = round(sum(Per), 2)) %>%
  mutate(name = case_when(name == "Q100" ~ "Gender",
                          name == "Q101" ~ "Race",
                          name == "Q2" ~ "Language"),
         value = case_when((name == "Gender" & value == 1) ~ "Men",
                           (name == "Gender" & value == 2) ~ "Women",
                           (name == "Race" & value == 1) ~ "Black/African",
                           (name == "Race" & value == 2) ~ "White/European",
                           (name == "Race" & value == 3) ~ "Coloured/Mixed Race",
                           (name == "Language" & value == 420) ~ "Malagasy officiel",
                           (name == "Language" & value == 421) ~ "Malagasy avec spécificité régionale",
                           (name == "Language" & value == 9995) ~ "Other"))
colnames(tab1)[which(names(tab1) == "value")] <- "Value"
colnames(tab1)[which(names(tab1) == "name")] <- "Variable"

tab1 %>% kable(caption = "Descriptive Stats")
```

Table 1: Descriptive Stats

Variable	Value	Percent
Gender	Men	50.08
Gender	Women	49.92
Race	Black/African	99.58
Race	White/European	0.17
Race	Coloured/Mixed Race	0.25
Language	Malagasy officiel	47.08
Language	Malagasy avec spécificité régionale	52.83

Variable	Value	Percent
Language	Other	0.08

3.

```
tab2 <-
  madagascar %>%
  count(Q78A, name = "Number") %>%
  mutate(Percent = round((100 * Number/sum(Number)), 2),
         Q78A = case_when(Q78A == 1 ~ "Very negative",
                          Q78A == 2 ~ "Somewhat negative",
                          Q78A == 3 ~ "Neither positive nor negative",
                          Q78A == 4 ~ "Somewhat positive",
                          Q78A == 5 ~ "Very positive",
                          Q78A == 8 ~ "Refused",
                          Q78A == 9 ~ "Don't know",
                          Q78A == -1 ~ "Missing"))

colnames(tab2)[which(names(tab2) == "Q78A")] <- "Influence"
```

The economic and political influence of China in Madagascar is most often seen as "Somewhat positive". The numbers show that 12.17% of people found it to be very negative, 20.67% somewhat negative, 11.5% neither positive nor negative, 29.08% somewhat positive, and 5.17% very positive.

```
tab2 <- kable(tab2, caption = "Political and Economic Influence: China")
tab2
```

Table 2: Political and Economic Influence: China

Influence	Number	Percent
Very negative	146	12.17
Somewhat negative	248	20.67
Neither positive nor negative	138	11.50
Somewhat positive	349	29.08
Very positive	62	5.17
Refused	1	0.08
Don't know	256	21.33

4.

```
tab2 <-
  madagascar %>%
  count(Q78B, name = "Number") %>%
  mutate(Percent = round((100 * Number/sum(Number)), 2),
         Q78B = case_when(Q78B == 1 ~ "Very negative",
                          Q78B == 2 ~ "Somewhat negative",
                          Q78B == 3 ~ "Neither positive nor negative",
                          Q78B == 4 ~ "Somewhat positive",
                          Q78B == 5 ~ "Very positive",
                          Q78B == 8 ~ "Refused",
```

```
                Q78B == 9 ~ "Don't know",
                Q78B == -1 ~ "Missing"))

colnames(tab2)[which(names(tab2) == "Q78B")] <- "Influence"
```

The economic and political influence of China in Madagascar is most often seen as "Don't know". The numbers show that 3.83% of people found it to be very negative, 8.17% somewhat negative, 12% neither positive nor negative, 30.75% somewhat positive, and 6.08% very positive.

```
tab2 <- kable(tab2, caption = "Political and Economic Influence: United States")
tab2
```

Table 3: Political and Economic Influence: United States

Influence	Number	Percent
Very negative	46	3.83
Somewhat negative	98	8.17
Neither positive nor negative	144	12.00
Somewhat positive	369	30.75
Very positive	73	6.08
Refused	3	0.25
Don't know	467	38.92

5. The t-test below is equal to -11.387. With 5% significance for a two-tailed t-test, I reject the null ($|t| = 11.387 > 1.96$). People view the economic and political influence of China significantly more negatively than the economic and political influence of the United States.

```
madagascar <-
  madagascar %>%
  mutate(
    across(
      Q78A:Q78B,
      ~ if_else(.x %in% 1:5, .x, NA)
      )
  )
t.test(madagascar$Q78A, madagascar$Q78B, paired = TRUE)
```

```
##
##  Paired t-test
##
## data:  madagascar$Q78A and madagascar$Q78B
## t = -11.387, df = 721, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.6203086 -0.4378631
## sample estimates:
## mean of the differences
##               -0.5290859
```