PHUONG Final project DSCI 478

2025-4-30

```
library(tidyverse)

# Load the dataset
df <- read_csv("sesladder.csv")
```

```
## Rows: 365 Columns: 11
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (4): country, pid, date, sex
## dbl  (6): age, stairs, Qhouse, Qfood, Qmoney, Qthings
## time (1): time
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

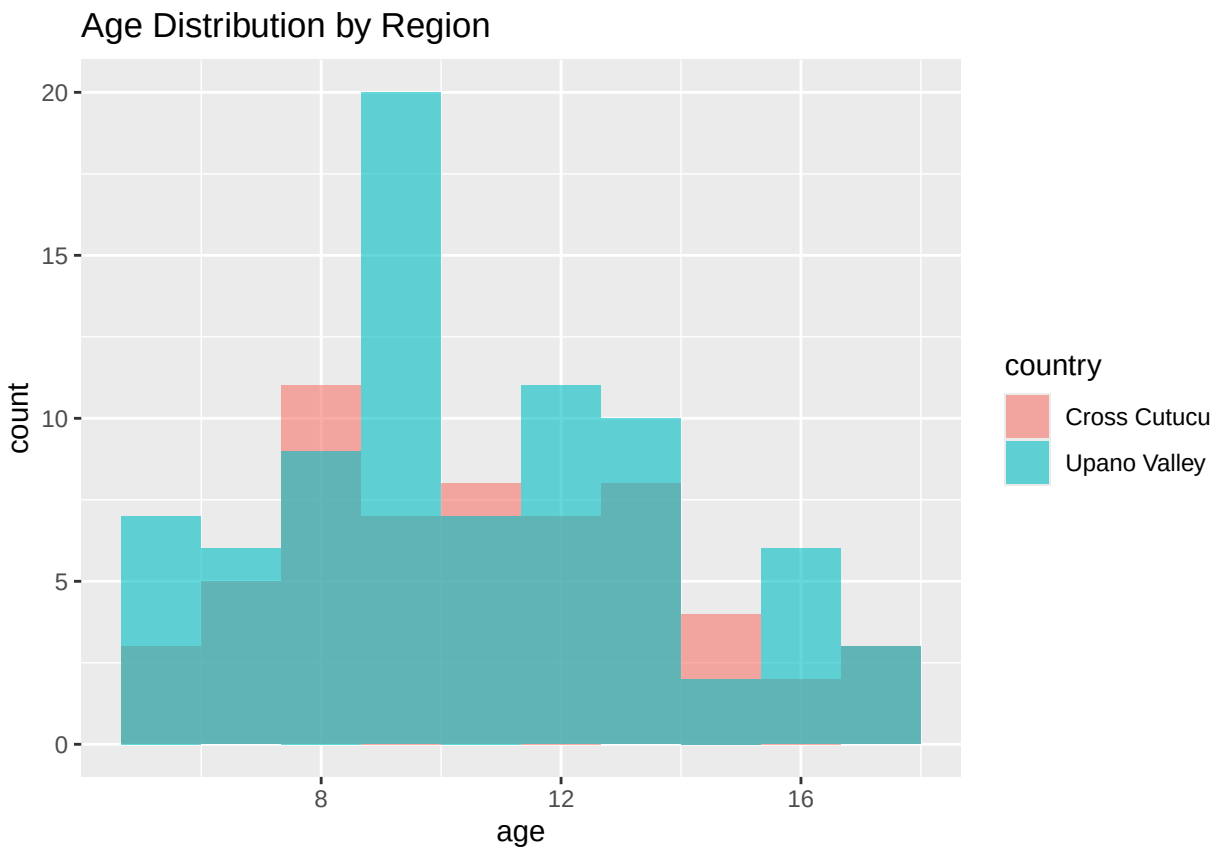
```
# Filter for just Cross Cutucu and Upano Valley
df_subset <- df %>%
  filter(country %in% c("Cross Cutucu", "Upano Valley"))
```

```
df_subset %>%
  group_by(country) %>%
  summarise(
    mean_age = mean(age, na.rm = TRUE),
    percent_male = mean(sex == "M", na.rm = TRUE),
    Qhouse = mean(Qhouse, na.rm = TRUE),
    Qfood = mean(Qfood, na.rm = TRUE),
    Qmoney = mean(Qmoney, na.rm = TRUE),
    Qthings = mean(Qthings, na.rm = TRUE)
  )
```

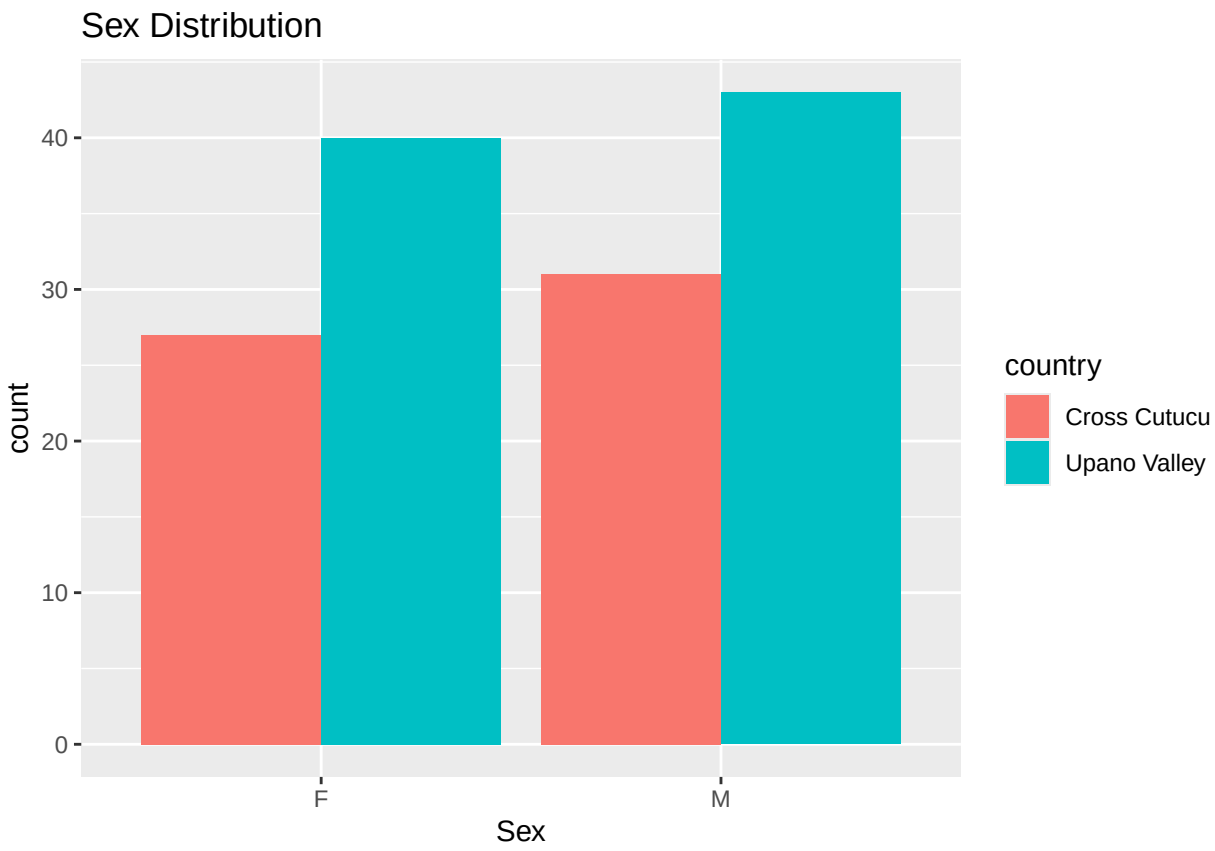
```
## # A tibble: 2 x 7
##   country      mean_age percent_male Qhouse Qfood Qmoney Qthings
##   <chr>           <dbl>        <dbl>  <dbl> <dbl>  <dbl>   <dbl>
## 1 Cross Cutucu     10.8        0.534   1.79  2.18   2.34    2
## 2 Upano Valley     10.6        0.518   1.70  1.94   2.11    1.96
```

```
ggplot(df_subset, aes(x = age, fill = country)) +
  geom_histogram(position = "identity", alpha = 0.6, bins = 10) +
  labs(title = "Age Distribution by Region")
```

```
## Warning: Removed 2 rows containing non-finite outside the scale range
## (`stat_bin()`).
```

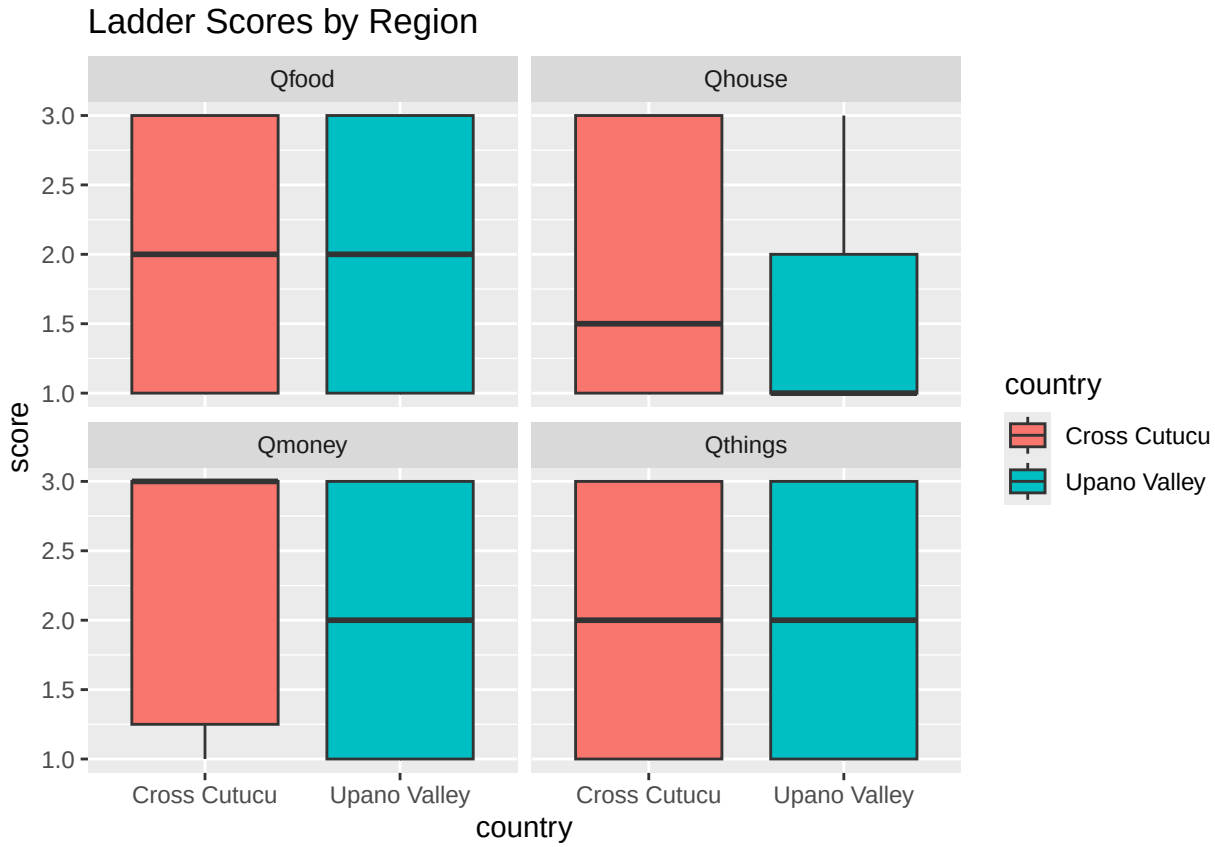



```
ggplot(df_subset, aes(x = sex, fill = country)) +
  geom_bar(position = "dodge") +
  labs(title = "Sex Distribution", x = "Sex")
```




```
df_subset %>%
  pivot_longer(cols = starts_with("Q"), names_to = "domain", values_to = "score") %>%
  ggplot(aes(x = country, y = score, fill = country)) +
  geom_boxplot() +
  facet_wrap(~domain) +
  labs(title = "Ladder Scores by Region")
```

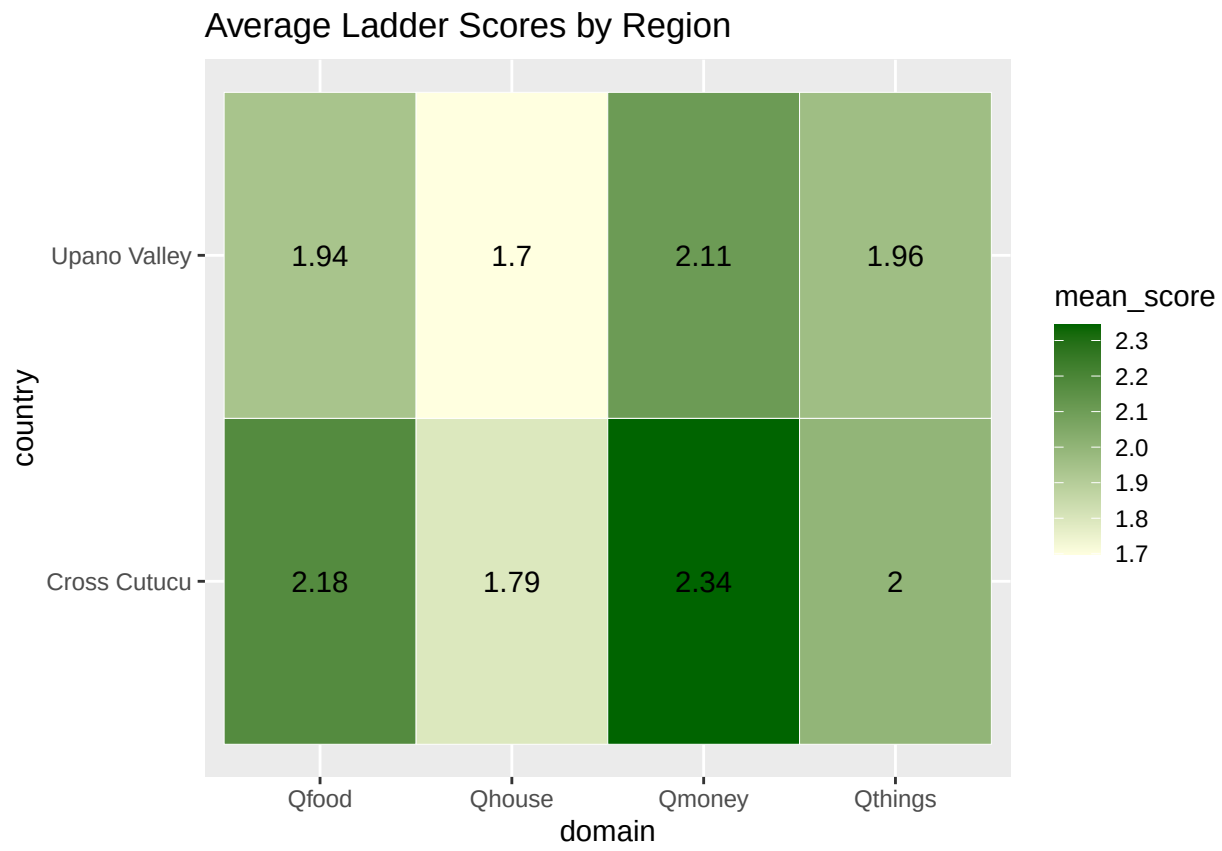
```
## Warning: Removed 4 rows containing non-finite outside the scale range
## (`stat_boxplot()`).
```




```
ladder_means <- df_subset %>%
  group_by(country) %>%
  summarise(across(starts_with("Q"), mean, na.rm = TRUE)) %>%
  pivot_longer(-country, names_to = "domain", values_to = "mean_score")
```

```
## Warning: There was 1 warning in `summarise()`.
## i In argument: `across(starts_with("Q"), mean, na.rm = TRUE)`.
## i In group 1: `country = "Cross Cutucu"`.
## Caused by warning:
## ! The `...` argument of `across()` is deprecated as of dplyr 1.1.0.
## Supply arguments directly to `.fns` through an anonymous function instead.
##
##   # Previously
##   across(a:b, mean, na.rm = TRUE)
##
##   # Now
##   across(a:b, \(x) mean(x, na.rm = TRUE))
```

```
ggplot(ladder_means, aes(x = domain, y = country, fill = mean_score)) +
  geom_tile(color = "white") +
  geom_text(aes(label = round(mean_score, 2))) +
  scale_fill_gradient(low = "lightyellow", high = "darkgreen") +
  labs(title = "Average Ladder Scores by Region")
```



Exploring Socioeconomic Perceptions in Cross Cutucú and Upano Valley

This analysis explores socioeconomic perceptions among children from two rural regions in Ecuador: Cross Cutucú and Upano Valley. Using survey data from the `sesladder.csv` dataset, we compared responses to four "ladder" questions—housing quality, food access, financial stability, and material possessions—rated on a 0 to 3 scale. These self-reported scores provide insight into how children perceive their everyday living conditions.

The demographic profiles of the two groups are quite similar. The average age in Cross Cutucú is 10.8 years, while in Upano Valley it is 10.6 years. Gender distribution is also nearly identical, with approximately 53% of participants identifying as male in Cross Cutucú and 52% in Upano Valley. This similarity in demographic makeup allows for a meaningful comparison of perceived socioeconomic status between the two regions.

Across all four ladder domains, children from Cross Cutucú consistently reported slightly higher scores. The average housing score was 1.79 in Cross Cutucú compared to 1.70 in Upano Valley. Food access stood out most sharply, with Cross Cutucú averaging 2.18 versus 1.94 in Upano Valley. Perceived financial status (Qmoney) also followed this trend, with Cross Cutucú scoring 2.34 and Upano Valley 2.11. Material possessions (Qthings) were rated at 2.00 and 1.96 respectively.

These findings suggest that children in Cross Cutucú may perceive greater stability or support in their immediate environment, especially regarding access to food and financial resources. This aligns with what is known about the two areas. Cross Cutucú, while more geographically remote, includes Indigenous Shuar communities that often rely on strong kinship networks and traditional ecological knowledge for food and resource management. These systems may provide a buffer against food insecurity and economic hardship.

By contrast, Upano Valley, though geographically closer to more developed areas, has undergone more agricultural transformation and exposure to market dependency. These dynamics can lead to greater

variability in household resources and may explain the slightly lower ladder scores reported by children in this region.

Overall, while both regions face challenges common to rural communities in Ecuador, the data suggest that Cross Cutucú may offer a more resilient social and ecological foundation, at least in the eyes of the children who live there.