

01

Katerina Bischel

2025-07-02

Background

This project visualizes results from the Fundación Instituto de Tecnología de Alimentos (I.T.A.) study conducted in 2006, focused on the Pilcomayo River basin in Chuquisaca, Bolivia. The study analyzed metals in water, soil, sediments, fish, human and animal blood, and plants.

Source: Muestreo de aguas, suelos, vegetales, sangre humana y animales, peces y sedimentos en puntos seleccionados en la cuenca del Río Pilcomayo en Chuquisaca (2006)

Libraries

```
library(readr)
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.4.1
```

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##   filter, lag

## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

Data Visualizations

Water Samples

```
water <- read_csv("data/ITA_water_2006.csv")

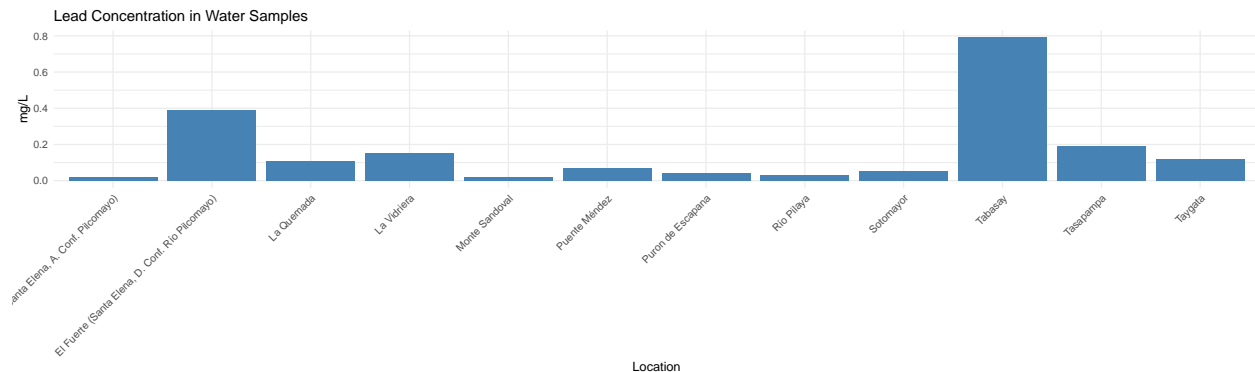
## Rows: 12 Columns: 14
## -- Column specification -----
## Delimiter: ","
## chr   (3): Municipality, Location, Sampling Date
## dbl  (10): Point No., pH, Pb (mg/l), Hg (mg/l), As (mg/l), Cd (mg/l), Zn (mg...
## time  (1): Time
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
head(water)
```

```
## # A tibble: 6 x 14
##   Municipality 'Point No.' Location      Time 'Sampling Date'    pH 'Pb (mg/l)'
##   <chr>         <dbl> <chr>         <tim> <chr>         <dbl> <dbl>
## 1 Culpina      1 Río Pilaya  13:33 09-10-2005    8.8    0.03
## 2 Las Carreras 2 Puro de Esc~ 09:07 10-10-2005    8.4    0.04
## 3 Las Carreras 3 Monte Sandov~ 15:40 10-10-2005    8.1    0.02
## 4 Camargo      4 La Vidriera  09:30 11-10-2005    7.8    0.15
## 5 Camargo      5 La Quemada  14:05 11-10-2005    8      0.11
## 6 Incahuasi    6 El Fuerte (S~ 10:50 12-10-2005    8.4    0.02
## # i 7 more variables: 'Hg (mg/l)' <dbl>, 'As (mg/l)' <dbl>, 'Cd (mg/l)' <dbl>,
## #   'Zn (mg/l)' <dbl>, 'X (UTM)' <dbl>, 'Y (UTM)' <dbl>, 'Elevation (m)' <dbl>
```

Lead in Water

```
ggplot(water, aes(x = Location, y = `Pb (mg/l)`)) +
  geom_col(fill = "steelblue") +
  labs(
    title = "Lead Concentration in Water Samples",
    x = "Location",
    y = "mg/L"
  ) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



Soil Samples

```
soil <- read_csv("data/ITA_soil_2006.csv")
```

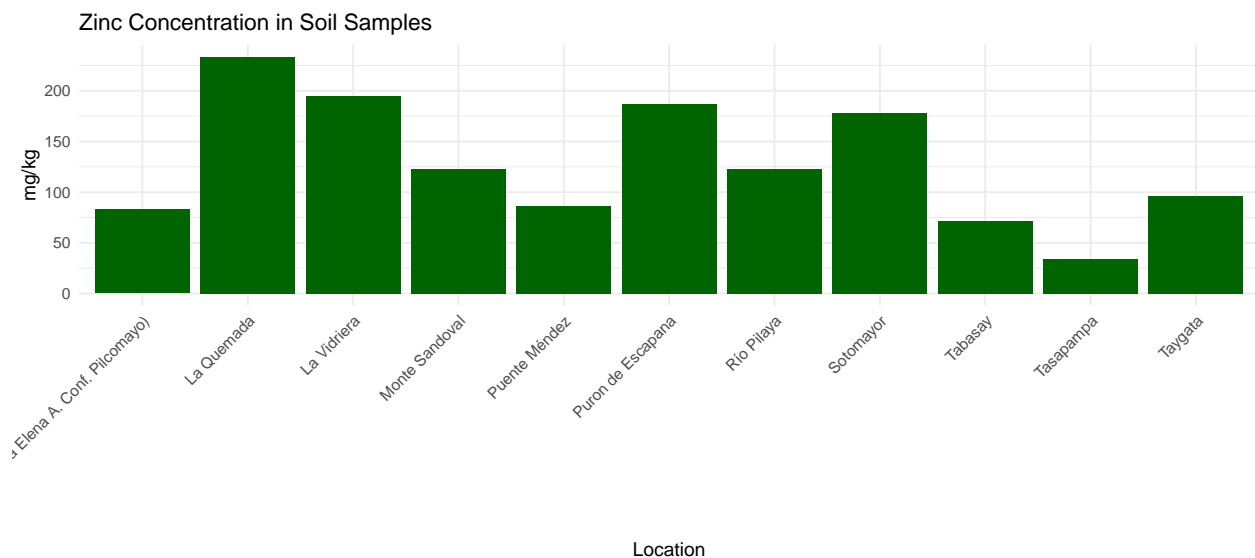
```
## Rows: 11 Columns: 15
## -- Column specification -----
## Delimiter: ","
## chr   (4): Municipality, Location, Sample Type, Date
## dbl   (10): Point No., Pb (mg/kg), Sb (mg/kg), Hg (mg/kg), As (mg/kg), Cd (mg...
## time  (1): Sampling Time
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
head(soil)
```

```
## # A tibble: 6 x 15
##   Municipality 'Point No.' Location   'Sample Type' 'Sampling Time' 'Pb (mg/kg)'
##   <chr>         <dbl> <chr>         <chr>         <time>         <dbl>
## 1 Culpina      1 Río Pilaya Soil          14:20         36.2
## 2 Las Carreras 2 Puron de ~ Soil          10:00         63.8
## 3 Las Carreras 3 Monte San~ Soil          15:10         44.2
## 4 Camargo      4 La Vidrie~ Soil          10:00         46.9
## 5 Camargo      5 La Quemada Soil          14:30         64.9
## 6 Incahuasi    6 El Fuerte~ Soil          12:50         32.7
## # i 9 more variables: 'Sb (mg/kg)' <dbl>, 'Hg (mg/kg)' <dbl>,
## #   'As (mg/kg)' <dbl>, 'Cd (mg/kg)' <dbl>, 'Zn (mg/kg)' <dbl>,
## #   'X (UTM)' <dbl>, 'Y (UTM)' <dbl>, 'Elevation (m)' <dbl>, Date <chr>
```

Zinc in Soil

```
ggplot(soil, aes(x = Location, y = `Zn (mg/kg)`)) +
  geom_col(fill = "darkgreen") +
  labs(
    title = "Zinc Concentration in Soil Samples",
    x = "Location",
    y = "mg/kg"
  ) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



Sediment Samples

```
sediment <- read_csv("data/ITA_sed_2006.csv")
```

```
## Rows: 12 Columns: 15
## -- Column specification -----
## Delimiter: ",",
```

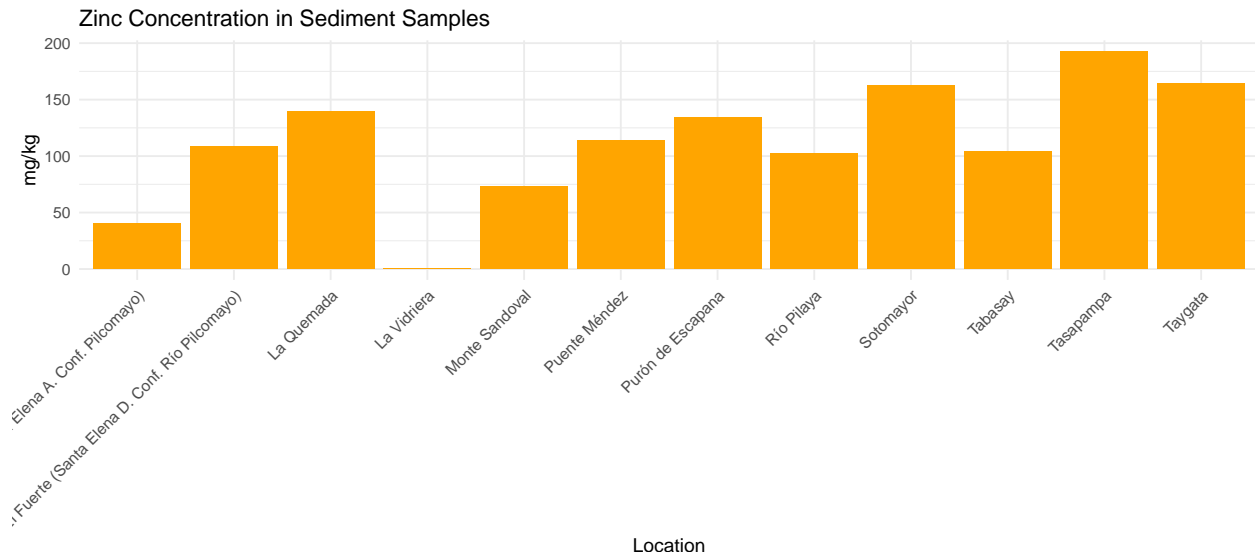
```
## chr (4): Municipality, Location, Sample Type, Date
## dbl (10): Point No., Pb (mg/kg), Sb (mg/kg), Hg (mg/kg), As (mg/kg), Cd (mg...
## time (1): Sampling Time
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
head(sediment)
```

```
## # A tibble: 6 x 15
##   Municipality 'Point No.' Location 'Sample Type' 'Sampling Time' 'Pb (mg/kg)'
##   <chr>         <dbl> <chr>      <chr>         <time>         <dbl>
## 1 Culpina             1 Río Pilaya Sediment    14:55          30.5
## 2 Las Carreras        2 Purón de ~ Sediment    09:24          56.4
## 3 Las Carreras        3 Monte San~ Sediment    15:45          20.2
## 4 Camargo             4 La Vidrie~ Sediment    09:30          26.7
## 5 Camargo             5 La Quemada Sediment    09:30          31.7
## 6 Incahuasi           6 El Fuerte~ Sediment    10:45          14.7
## # i 9 more variables: 'Sb (mg/kg)' <dbl>, 'Hg (mg/kg)' <dbl>,
## #   'As (mg/kg)' <dbl>, 'Cd (mg/kg)' <dbl>, 'Zn (mg/kg)' <dbl>,
## #   'X (UTM)' <dbl>, 'Y (UTM)' <dbl>, 'Elevation (m)' <dbl>, Date <chr>
```

Mercury in Sediments

```
ggplot(sediment, aes(x = Location, y = `Zn (mg/kg)`)) +
  geom_col(fill = "orange") +
  labs(
    title = "Zinc Concentration in Sediment Samples",
    x = "Location",
    y = "mg/kg"
  ) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



Vegetation Samples

```
veg <- read_csv("data/ITA_veg_2006.csv")
```

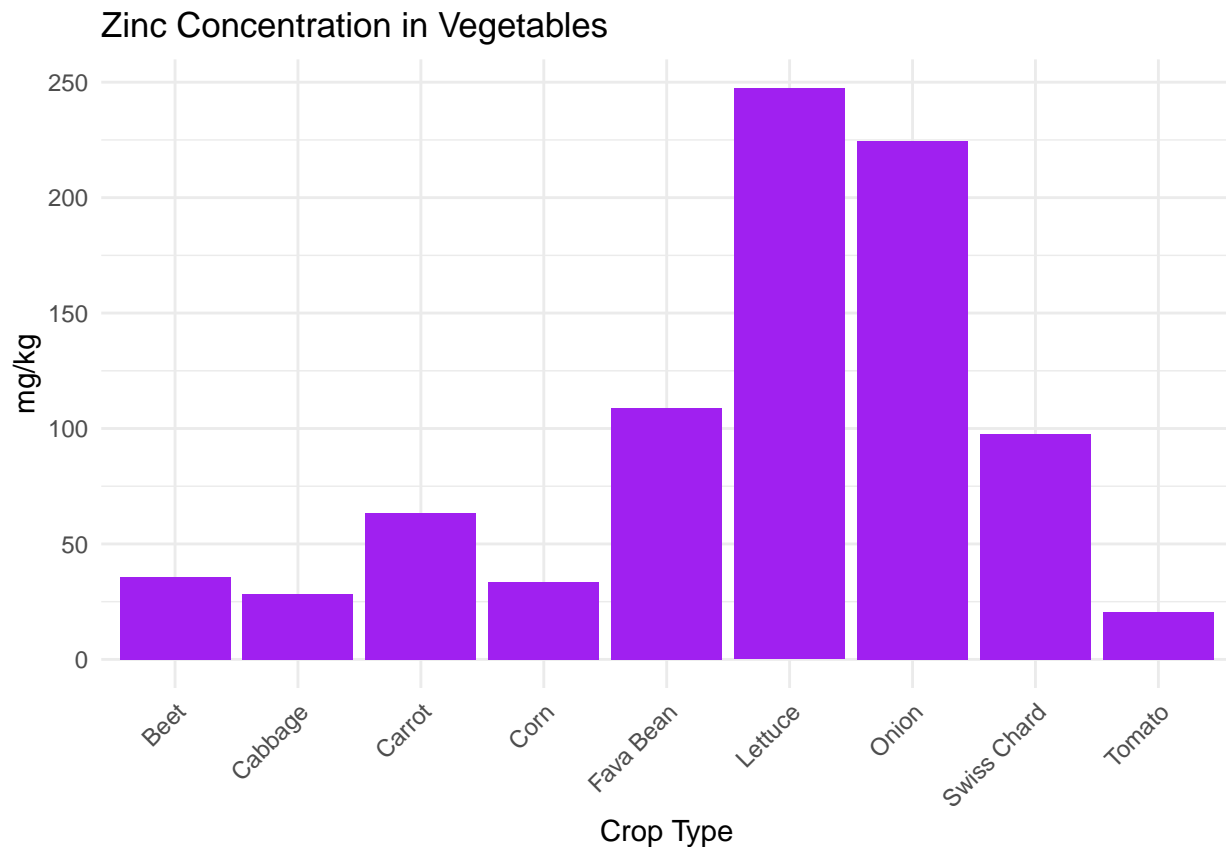
```
## Rows: 21 Columns: 14
## -- Column specification -----
## Delimiter: ","
## chr (5): Municipality, Location, Crop, Date, Hg (mg/kg)
## dbl (9): Point No., X (UTM), Y (UTM), Elevation (m), Pb (mg/kg), Sb (mg/kg),...
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
head(veg)
```

```
## # A tibble: 6 x 14
##   Municipality 'Point No.' Location   'X (UTM)' 'Y (UTM)' 'Elevation (m)' Crop
##   <chr>         <dbl> <chr>         <dbl>     <dbl>         <dbl> <chr>
## 1 Culpina             1 Río Pilaya  372486.  7660544.         814 Onion
## 2 Culpina             1 Río Pilaya  372486.  7660544.         814 Toma~
## 3 Las Carreras        2 Purón de E~ 267173.  7621425.        2431 Fava~
## 4 Las Carreras        2 Purón de E~ 267173.  7621425.        2431 Onion
## 5 Las Carreras        3 Monte Sand~ 271255.  7656927         2311 Onion
## 6 Las Carreras        3 Monte Sand~ 271255.  7656927         2311 Carr~
## # i 7 more variables: Date <chr>, 'Pb (mg/kg)' <dbl>, 'Sb (mg/kg)' <dbl>,
## #   'Hg (mg/kg)' <chr>, 'As (mg/kg)' <dbl>, 'Cd (mg/kg)' <dbl>,
## #   'Zn (mg/kg)' <dbl>
```

Lead in Vegetables

```
ggplot(veg, aes(x = Crop, y = `Zn (mg/kg)`) +
  geom_col(fill = "purple") +
  labs(
    title = "Zinc Concentration in Vegetables",
    x = "Crop Type",
    y = "mg/kg"
  ) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



Fish Samples

```
fish <- read_csv("data/ITA_fish_2006.csv")
```

```
## Rows: 24 Columns: 13
## -- Column specification -----
## Delimiter: ","
## chr  (4): Municipality, Location, Sample Type, Sample Date
## dbl  (8): Point No., Pb (mg/kg), As (mg/kg), Cd (mg/kg), Point, X (UTM), Y (...
## time (1): Sample 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.
```

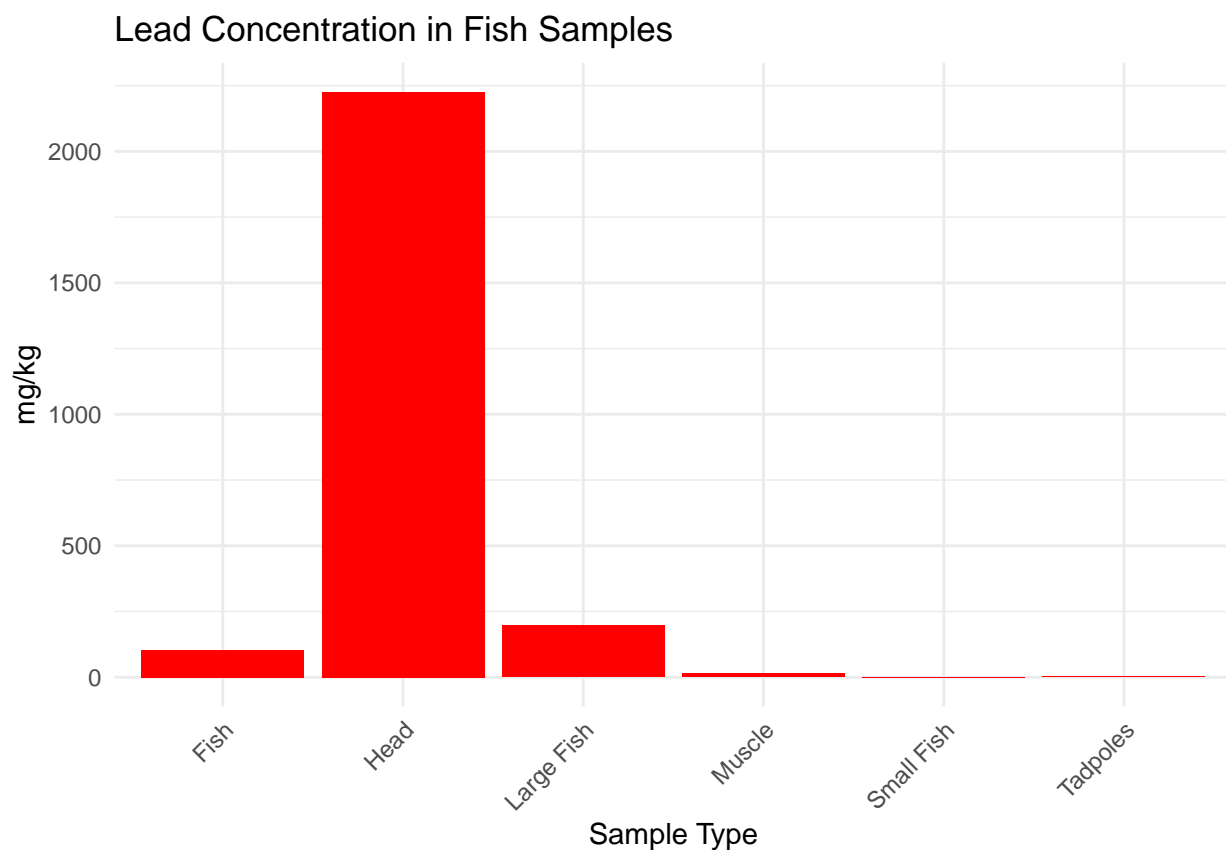
```
head(fish)
```

```
## # A tibble: 6 x 13
##   Municipality 'Point No.' Location   'Sample Type' 'Sample Date' 'Sample Time'
##   <chr>         <dbl> <chr>         <chr>         <chr>         <time>
## 1 Culpina      1 Río Pilaya Head         09-10-2005    13:33
## 2 Culpina      1 Río Pilaya Muscle       09-10-2005    13:33
## 3 Las Carreras 2 Purón de E~ Head         10-10-2005    09:00
## 4 Las Carreras 2 Purón de E~ Muscle       10-10-2005    09:00
## 5 Las Carreras 3 Monte Sand~ Fish         10-10-2005    15:25
## 6 Las Carreras 3 Monte Sand~ Fish         10-10-2005    15:25
## # i 7 more variables: 'Pb (mg/kg)' <dbl>, 'As (mg/kg)' <dbl>,
```

```
## # 'Cd (mg/kg)' <dbl>, Point <dbl>, 'X (UTM)' <dbl>, 'Y (UTM)' <dbl>,
## # 'Elevation (m)' <dbl>
```

Lead in Fish Muscle Tissue

```
ggplot(fish, aes(x = `Sample Type`, y = `Pb (mg/kg)`)) +
  geom_col(fill = "red") +
  labs(
    title = "Lead Concentration in Fish Samples",
    x = "Sample Type",
    y = "mg/kg"
  ) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



Human Blood Samples

```
human <- read_csv("data/ITA_human_2006.csv")
```

```
## Rows: 110 Columns: 17
## -- Column specification -----
## Delimiter: ","
## chr (4): Municipality, Location, Sex, Sampling Date
## dbl (13): Point No., Age, As [µg/dl], Pb [µg/dl] Children, Pb [µg/dl] Adults...
##
## 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.
```

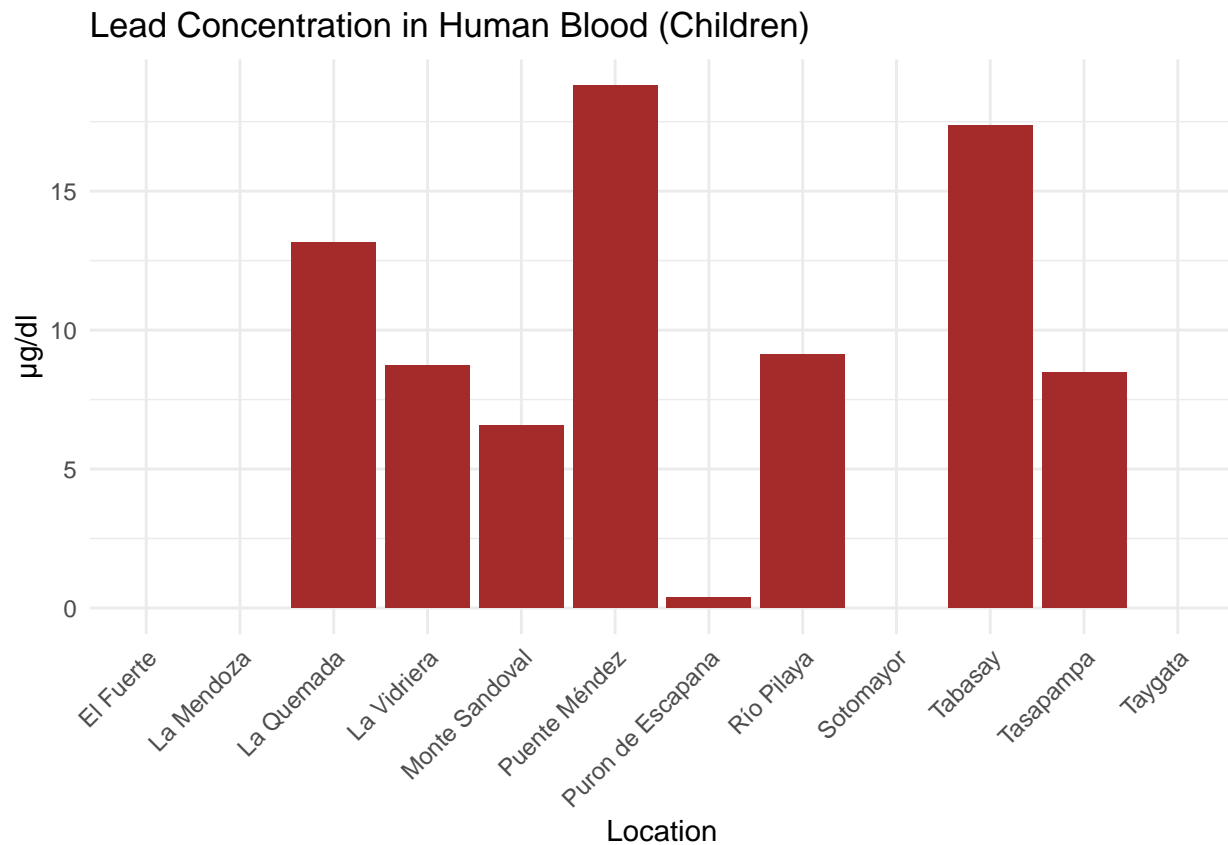
```
head(human)
```

```
## # A tibble: 6 x 17
##   Municipality 'Point No.' Location      Age Sex   'As [µg/dl]'
##   <chr>         <dbl> <chr>      <dbl> <chr>   <dbl>
## 1 Culpina             1 Río Pilaya    39 F     0.27
## 2 Culpina             1 Río Pilaya    55 F     0.25
## 3 Culpina             1 Río Pilaya    49 F     0.22
## 4 Culpina             1 Río Pilaya    30 F     0.28
## 5 Culpina             1 Río Pilaya    15 F     0.28
## 6 Culpina             1 Río Pilaya    13 F     0.25
## # i 11 more variables: 'Pb [µg/dl] Children' <dbl>, 'Pb [µg/dl] Adults' <dbl>,
## #   'Cd [µg/dl]' <dbl>, 'Sampling Date' <chr>, 'X (UTM)' <dbl>,
## #   'Y (UTM)' <dbl>, 'Elevation (m)' <dbl>, 'Mean As [µg/dl]' <dbl>,
## #   'Mean Pb [µg/dl] Children' <dbl>, 'Mean Pb [µg/dl] Adults' <dbl>,
## #   'Mean Cd [µg/dl]' <dbl>
```

Lead in Child Human Blood

```
ggplot(human, aes(x = Location, y = `Pb [µg/dl] Children`)) +
  geom_col(fill = "brown") +
  labs(
    title = "Lead Concentration in Human Blood (Children)",
    x = "Location",
    y = "µg/dl"
  ) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

```
## Warning: Removed 97 rows containing missing values or values outside the scale range
## ('geom_col()').
```

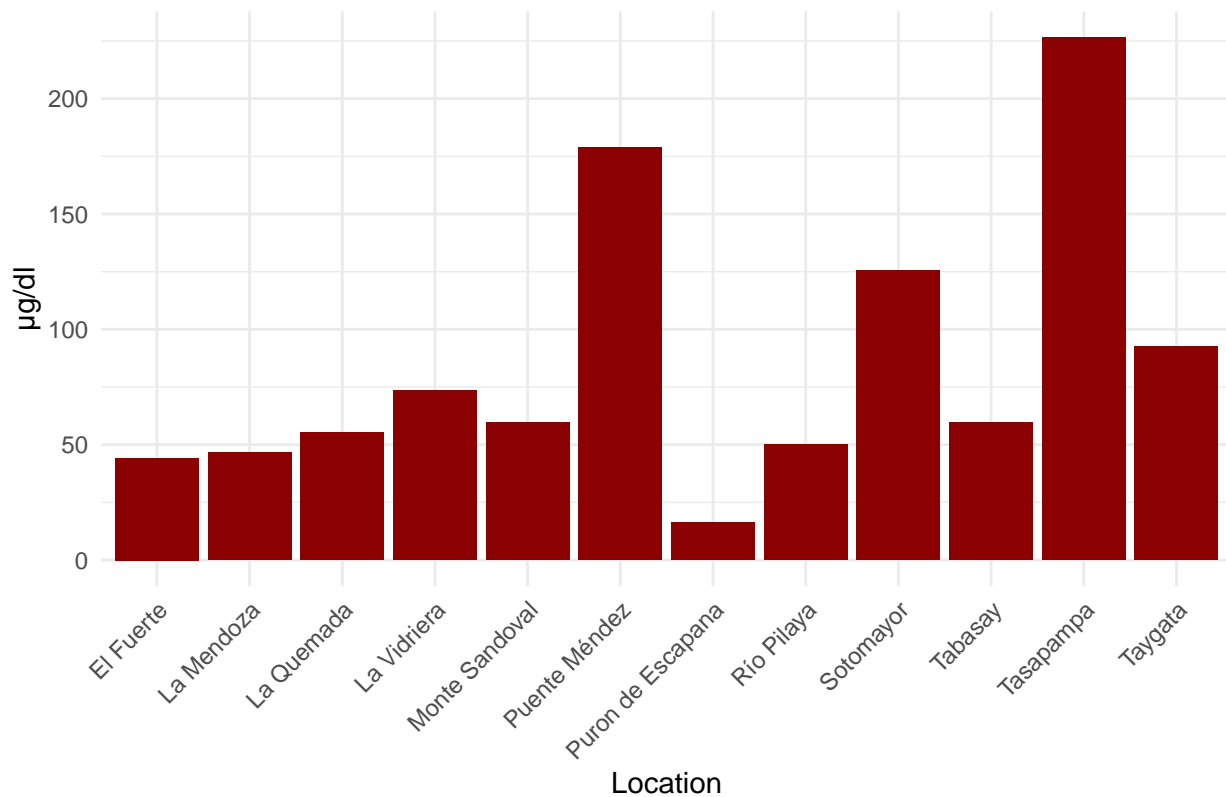



Lead in Adult Human Blood

```
ggplot(human, aes(x = Location, y = `Pb [µg/dl] Adults`)) +
  geom_col(fill = "darkred") +
  labs(
    title = "Lead Concentration in Human Blood (Adults)",
    x = "Location",
    y = "µg/dl"
  ) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

```
## Warning: Removed 13 rows containing missing values or values outside the scale range
## ('geom_col()').
```

Lead Concentration in Human Blood (Adults)



Animal Blood Samples

```
animal <- read_csv("data/ITA_animal_2006.csv")
```

```
## Rows: 77 Columns: 15
## -- Column specification -----
## Delimiter: ","
## chr   (4): Sample ID, Location, Sampling Date, Animal
## dbl   (10): Pb (mg/dL), Cd (mg/dL), As (mg/dL), X (UTM), Y (UTM), Elevation (...
## time  (1): Sampling 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.
```

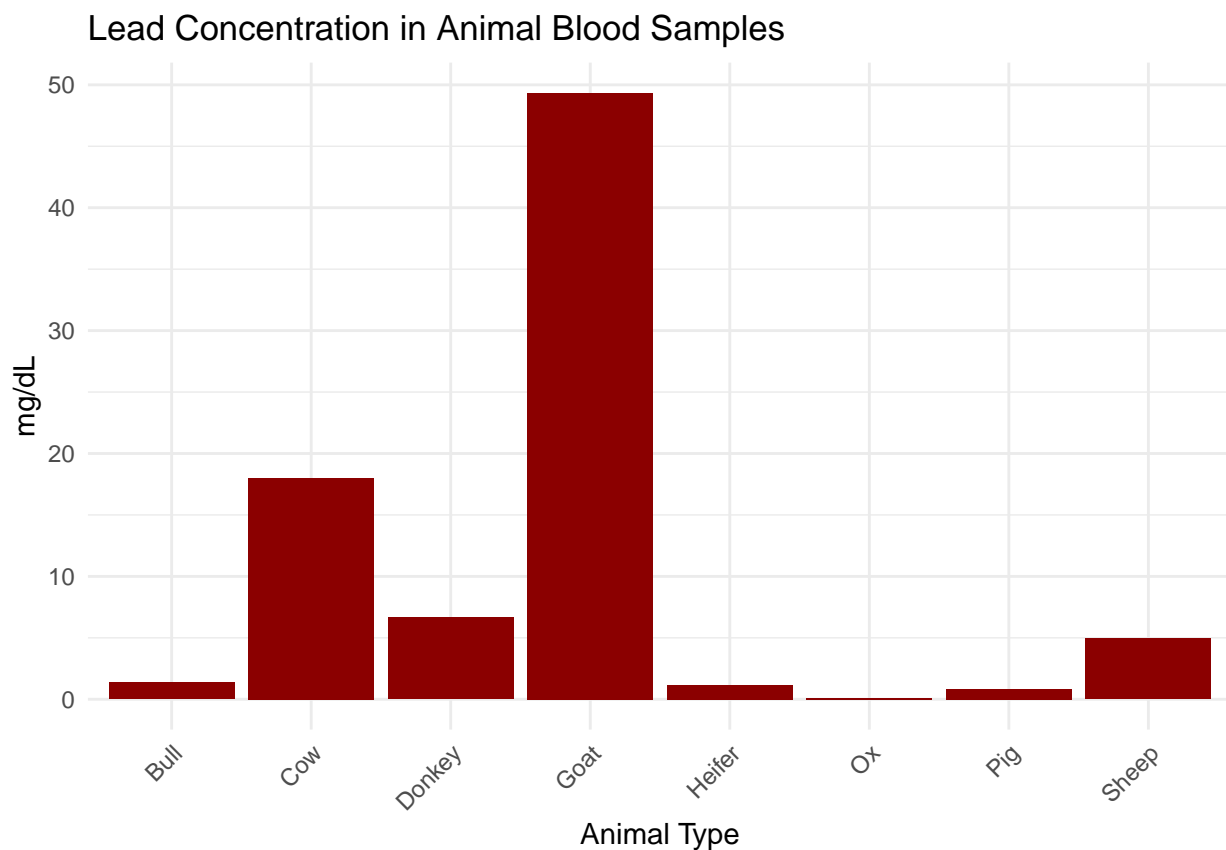
```
head(animal)
```

```
## # A tibble: 6 x 15
##   'Sample ID' 'Location' 'Sampling Date' 'Sampling Time' 'Pb (mg/dL)' 'Cd (mg/dL)'
##   <chr>       <chr>       <chr>         <time>         <dbl>       <dbl>
## 1 Blood No. 1 Río Pil~ 09-10-2005    10:42         0.2         0.18
## 2 Blood No. 2 Río Pil~ 09-10-2005    10:46         0.13         0.12
## 3 Blood No. 3 Río Pil~ 09-10-2005    10:50         0.04         0.37
## 4 Blood No. 4 Río Pil~ 09-10-2005    10:55         0.06         0.4
## 5 Blood No. 5 Río Pil~ 09-10-2005    11:01         0.2         0.18
## 6 Blood No. 6 Río Pil~ 09-10-2005    11:06         0.09         0.15
## # i 9 more variables: 'As (mg/dL)' <dbl>, 'X (UTM)' <dbl>, 'Y (UTM)' <dbl>,
```

```
## # 'Elevation (m)' <dbl>, 'Point No.' <dbl>, Animal <chr>,
## # 'Mean Pb (mg/dl)' <dbl>, 'Mean Cd (mg/dl)' <dbl>, 'Mean As (mg/dl)' <dbl>
```

Lead in Animal Blood

```
ggplot(animal, aes(x = Animal, y = `Pb (mg/dL)`)) +
  geom_col(fill = "darkred") +
  labs(
    title = "Lead Concentration in Animal Blood Samples",
    x = "Animal Type",
    y = "mg/dL"
  ) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



Conclusions

This is a preliminary visualization of the 2006 Pilcomayo River metals assessment. Further statistical summaries and comparisons to WHO or Bolivian environmental standards can be added as needed.