

Impact of Climate Change on Mountain Gorillas

Sarah Theriot

2024-10-20

Introduction:

Mountain gorillas (*Gorilla beringei beringei*) are critically endangered and face many serious threats, one being climate change. In this project, I will explore how climate change affects their populations, stress levels, and habitat use. By analyzing this data, I hope to understand these impacts and suggest ways to help protect these amazing animals. If we are able to identify specific causes to dwindling gorilla populations then we can come up with specific ways to help them.

Problem Statement:

One of the biggest issues gorillas face is loss of their habitat. This is directly causing their growing extinction. Loss of habitat happens for one of two reasons, deforestation and “natural causes” (i.e. environmental changes). These “natural causes” are unfortunately caused by climate change, which is caused by the increase in greenhouse gases. Therefore, when looking at this project it is important to look out for data that will give us a better insight into all of that.

Research Questions:

1. How do mountain gorillas use their habitat, and how has this changed in response to climate change?
2. What are the long-term environmental trends in the Democratic Republic of Congo (DRC)?
3. How do temperature and rainfall changes correlate with cortisol levels (stress) in gorillas?
4. What is the relationship between sodium acquisition (a dietary need) and stress levels?
5. How can these insights contribute to more effective conservation efforts?

Addressing the Problem Statement:

Approach:

I will analyze three different datasets:

1. Dataset on the social and ecological factors affecting mountain gorillas
2. Dataset on how mountain gorillas utilize their changing habitats to acquire necessary sodium
3. Dataset on climate change throughout the years in the Democratic Republic of Congo

I hope that by using these datasets I may be ²able to discover patterns and trends in changes throughout the years. I also hope to point out why it is important to take climate change seriously and how loss of the mountain gorillas home could have detrimental effects to their species. This would be the very first step in a long research process and plan. I would start

```
# Required Packages:
```

```
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
```

```
library(ggplot2)
```

```
library(readr)
```

```
library(readxl)
```

```
library(tidyr)
```

Data Analysis

Step 1: Load the Dataset

```
## Load the necessary datasets
```

```
stress_data <- read_csv("C:/Users/sarah/Desktop/MSDS/Statistics for Data Science/Final Project/Envir_Gr
```

```
## Rows: 1094 Columns: 23
```

```
## -- Column specification -----
```

```
## Delimiter: ","
```

```
## chr (5): Month, Gorilla, Group, Sex, GrpType
```

```
## dbl (18): Year, Age, HomeR70, HomeR70_excl, HomeR70_accum_overlap, HomeR_acc...
```

```
##
```

```
## 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.
```

```
habitat_data <- read_excel("C:/Users/sarah/Desktop/MSDS/Statistics for Data Science/Final Project/Sodium
```

```
## Warning: Expecting numeric in Q1705 / R1705C17: got 'NA'
```

```
## Warning: Expecting numeric in R1705 / R1705C18: got 'NA'
```

```
## Warning: Expecting numeric in Q1706 / R1706C17: got 'NA'
```

```
## Warning: Expecting numeric in R1706 / R1706C18: got 'NA'
```

```
## Warning: Expecting numeric in Q1707 / R1707C17: got 'NA'
```

Warning: Expecting numeric in R1707 / R1707C18: got 'NA'

Warning: Expecting numeric in Q1725 / R1725C17: got 'NA'

Warning: Expecting numeric in R1725 / R1725C18: got 'NA'

Warning: Expecting numeric in Q1726 / R1726C17: got 'NA'

Warning: Expecting numeric in R1726 / R1726C18: got 'NA'

Warning: Expecting numeric in Q1727 / R1727C17: got 'NA'

Warning: Expecting numeric in R1727 / R1727C18: got 'NA'

Warning: Expecting numeric in Q1728 / R1728C17: got 'NA'

Warning: Expecting numeric in R1728 / R1728C18: got 'NA'

Warning: Expecting numeric in Q1729 / R1729C17: got 'NA'

Warning: Expecting numeric in R1729 / R1729C18: got 'NA'

Warning: Expecting numeric in Q1730 / R1730C17: got 'NA'

Warning: Expecting numeric in R1730 / R1730C18: got 'NA'

Warning: Expecting numeric in Q1731 / R1731C17: got 'NA'

Warning: Expecting numeric in R1731 / R1731C18: got 'NA'

Warning: Expecting numeric in Q1732 / R1732C17: got 'NA'

Warning: Expecting numeric in R1732 / R1732C18: got 'NA'

Warning: Expecting numeric in Q1847 / R1847C17: got 'NA'

Warning: Expecting numeric in R1847 / R1847C18: got 'NA'

Warning: Expecting numeric in Q1848 / R1848C17: got 'NA'

Warning: Expecting numeric in R1848 / R1848C18: got 'NA'

Warning: Expecting numeric in Q1849 / R1849C17: got 'NA'

Warning: Expecting numeric in R1849 / R1849C18: got 'NA'

Warning: Expecting numeric in Q1850 / R1850C17: got 'NA'

Warning: Expecting numeric in R1850 / R1850C18: got 'NA'

Warning: Expecting numeric in Q1885 / R1885C17: got 'NA'

Warning: Expecting numeric in R1885 / R1885C18: got 'NA'

Warning: Expecting numeric in Q1886 / R1886C17: got 'NA'

Warning: Expecting numeric in R1886 / R1886C18: got 'NA'

Warning: Expecting numeric in Q1887 / R1887C17: got 'NA'

Warning: Expecting numeric in R1887 / R1887C18: got 'NA'

Warning: Expecting numeric in Q1888 / R1888C17: got 'NA'

Warning: Expecting numeric in R1888 / R1888C18: got 'NA'

Warning: Expecting numeric in Q1889 / R1889C17: got 'NA'

Warning: Expecting numeric in R1889 / R1889C18: got 'NA'

Warning: Expecting numeric in Q1890 / R1890C17: got 'NA'

Warning: Expecting numeric in R1890 / R1890C18: got 'NA'

Warning: Expecting numeric in Q1902 / R1902C17: got 'NA'

Warning: Expecting numeric in R1902 / R1902C18: got 'NA'

Warning: Expecting numeric in Q1903 / R1903C17: got 'NA'

Warning: Expecting numeric in R1903 / R1903C18: got 'NA'

Warning: Expecting numeric in Q1904 / R1904C17: got 'NA'

Warning: Expecting numeric in R1904 / R1904C18: got 'NA'

Warning: Expecting numeric in Q1922 / R1922C17: got 'NA'

Warning: Expecting numeric in R1922 / R1922C18: got 'NA'

Warning: Expecting numeric in Q1923 / R1923C17: got 'NA'

Warning: Expecting numeric in R1923 / R1923C18: got 'NA'

Warning: Expecting numeric in Q1924 / R1924C17: got 'NA'

Warning: Expecting numeric in R1924 / R1924C18: got 'NA'

Warning: Expecting numeric in Q1982 / R1982C17: got 'NA'

Warning: Expecting numeric in R1982 / R1982C18: got 'NA'

Warning: Expecting numeric in Q1983 / R1983C17: got 'NA'

Warning: Expecting numeric in R1983 / R1983C18: got 'NA'

Warning: Expecting numeric in Q1984 / R1984C17: got 'NA'

Warning: Expecting numeric in R1984 / R1984C18: got 'NA'

Warning: Expecting numeric in Q1985 / R1985C17: got 'NA'

Warning: Expecting numeric in R1985 / R1985C18: got 'NA'

Warning: Expecting numeric in Q1986 / R1986C17: got 'NA'

Warning: Expecting numeric in R1986 / R1986C18: got 'NA'

Warning: Expecting numeric in Q1987 / R1987C17: got 'NA'

Warning: Expecting numeric in R1987 / R1987C18: got 'NA'

Warning: Expecting numeric in Q1996 / R1996C17: got 'NA'

Warning: Expecting numeric in R1996 / R1996C18: got 'NA'

Warning: Expecting numeric in Q1997 / R1997C17: got 'NA'

Warning: Expecting numeric in R1997 / R1997C18: got 'NA'

Warning: Expecting numeric in Q1998 / R1998C17: got 'NA'

Warning: Expecting numeric in R1998 / R1998C18: got 'NA'

Warning: Expecting numeric in Q1999 / R1999C17: got 'NA'

Warning: Expecting numeric in R1999 / R1999C18: got 'NA'

Warning: Expecting numeric in Q2000 / R2000C17: got 'NA'

Warning: Expecting numeric in R2000 / R2000C18: got 'NA'

Warning: Expecting numeric in Q2001 / R2001C17: got 'NA'

Warning: Expecting numeric in R2001 / R2001C18: got 'NA'

Warning: Expecting numeric in Q2002 / R2002C17: got 'NA'

Warning: Expecting numeric in R2002 / R2002C18: got 'NA'

Warning: Expecting numeric in Q2025 / R2025C17: got 'NA'

Warning: Expecting numeric in R2025 / R2025C18: got 'NA'

Warning: Expecting numeric in Q2026 / R2026C17: got 'NA'

Warning: Expecting numeric in R2026 / R2026C18: got 'NA'

Warning: Expecting numeric in Q2027 / R2027C17: got 'NA'

Warning: Expecting numeric in R2027 / R2027C18: got 'NA'

Warning: Expecting numeric in Q2028 / R2028C17: got 'NA'

Warning: Expecting numeric in R2028 / R2028C18: got 'NA'

Warning: Expecting numeric in Q2040 / R2040C17: got 'NA'

Warning: Expecting numeric in R2040 / R2040C18: got 'NA'

Warning: Expecting numeric in Q2041 / R2041C17: got 'NA'

Warning: Expecting numeric in R2041 / R2041C18: got 'NA'

Warning: Expecting numeric in Q2042 / R2042C17: got 'NA'

Warning: Expecting numeric in R2042 / R2042C18: got 'NA'

Warning: Expecting numeric in Q2044 / R2044C17: got 'NA'

Warning: Expecting numeric in R2044 / R2044C18: got 'NA'

Warning: Expecting numeric in Q2045 / R2045C17: got 'NA'

Warning: Expecting numeric in R2045 / R2045C18: got 'NA'

Warning: Expecting numeric in Q2046 / R2046C17: got 'NA'

Warning: Expecting numeric in R2046 / R2046C18: got 'NA'

Warning: Expecting numeric in Q2060 / R2060C17: got 'NA'

Warning: Expecting numeric in R2060 / R2060C18: got 'NA'

Warning: Expecting numeric in Q2063 / R2063C17: got 'NA'

Warning: Expecting numeric in R2063 / R2063C18: got 'NA'

Warning: Expecting numeric in Q2112 / R2112C17: got 'NA'

Warning: Expecting numeric in R2112 / R2112C18: got 'NA'

Warning: Expecting numeric in Q2113 / R2113C17: got 'NA'

Warning: Expecting numeric in R2113 / R2113C18: got 'NA'

Warning: Expecting numeric in Q2114 / R2114C17: got 'NA'

Warning: Expecting numeric in R2114 / R2114C18: got 'NA'

Warning: Expecting numeric in Q2115 / R2115C17: got 'NA'

Warning: Expecting numeric in R2115 / R2115C18: got 'NA'

Warning: Expecting numeric in Q2116 / R2116C17: got 'NA'

Warning: Expecting numeric in R2116 / R2116C18: got 'NA'

Warning: Expecting numeric in Q2193 / R2193C17: got 'NA'

Warning: Expecting numeric in R2193 / R2193C18: got 'NA'

Warning: Expecting numeric in Q2194 / R2194C17: got 'NA'

Warning: Expecting numeric in R2194 / R2194C18: got 'NA'

Warning: Expecting numeric in Q2195 / R2195C17: got 'NA'

Warning: Expecting numeric in R2195 / R2195C18: got 'NA'

Warning: Expecting numeric in Q2269 / R2269C17: got 'NA'

Warning: Expecting numeric in R2269 / R2269C18: got 'NA'

Warning: Expecting numeric in Q2270 / R2270C17: got 'NA'

Warning: Expecting numeric in R2270 / R2270C18: got 'NA'

Warning: Expecting numeric in Q2271 / R2271C17: got 'NA'

Warning: Expecting numeric in R2271 / R2271C18: got 'NA'

Warning: Expecting numeric in Q2272 / R2272C17: got 'NA'

Warning: Expecting numeric in R2272 / R2272C18: got 'NA'

Warning: Expecting numeric in Q2273 / R2273C17: got 'NA'

Warning: Expecting numeric in R2273 / R2273C18: got 'NA'

Warning: Expecting numeric in Q2274 / R2274C17: got 'NA'

Warning: Expecting numeric in R2274 / R2274C18: got 'NA'

Warning: Expecting numeric in Q2275 / R2275C17: got 'NA'

Warning: Expecting numeric in R2275 / R2275C18: got 'NA'

Warning: Expecting numeric in Q2276 / R2276C17: got 'NA'

Warning: Expecting numeric in R2276 / R2276C18: got 'NA'

Warning: Expecting numeric in Q2277 / R2277C17: got 'NA'

Warning: Expecting numeric in R2277 / R2277C18: got 'NA'

Warning: Expecting numeric in Q2435 / R2435C17: got 'NA'

Warning: Expecting numeric in R2435 / R2435C18: got 'NA'

Warning: Expecting numeric in Q2436 / R2436C17: got 'NA'

Warning: Expecting numeric in R2436 / R2436C18: got 'NA'

Warning: Expecting numeric in Q2437 / R2437C17: got 'NA'

Warning: Expecting numeric in R2437 / R2437C18: got 'NA'

Warning: Expecting numeric in Q2438 / R2438C17: got 'NA'

Warning: Expecting numeric in R2438 / R2438C18: got 'NA'

Warning: Expecting numeric in Q2490 / R2490C17: got 'NA'

Warning: Expecting numeric in R2490 / R2490C18: got 'NA'

Warning: Expecting numeric in Q2491 / R2491C17: got 'NA'

Warning: Expecting numeric in R2491 / R2491C18: got 'NA'

Warning: Expecting numeric in Q2492 / R2492C17: got 'NA'

Warning: Expecting numeric in R2492 / R2492C18: got 'NA'

```
emissions_data <- read_csv("C:/Users/sarah/Desktop/MSDS/Statistics for Data Science/Final Project/ghg-emissions.csv")
```

```
## Warning: One or more parsing issues, call 'problems()' on your data frame for details,
## e.g.:
##   dat <- vroom(...)
##   problems(dat)

## Rows: 6 Columns: 34
## -- Column specification -----
## Delimiter: ","
## chr  (2): Sector, unit
## dbl (32): 1990, 1991, 1992, 1993, 1994, 1995, 1996, 1997, 1998, 1999, 2000, ...
##
## 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.
```

Step 2: Inspect and Clean Datasets

Inspect the Stress Dataset

```
# Check the first few rows and structure of the STRESS dataset
head(stress_data)
```

```
## # A tibble: 6 x 23
##   Month      Year Gorilla Group Sex      Age HomeR70 HomeR70_excl
##   <chr>      <dbl> <chr>  <chr> <chr> <dbl>    <dbl>      <dbl>
## 1 September  2011 AFR    PAB   F     11.4    964.      964.
## 2 October   2011 AFR    PAB   F     11.4    359.      275.
## 3 November  2011 AFR    PAB   F     11.5    722.      520.
## 4 December  2011 AFR    PAB   F     11.6    433.      429.
## 5 January   2012 AFR    PAB   F     11.7    558.      553.
## 6 February  2012 AFR    PAB   F     11.8    155.      155.
## # i 15 more variables: HomeR70_accum_overlap <dbl>,
## #   HomeR_accum_overlap_index <dbl>, N_Grp_HomeR70_overlap <dbl>,
## #   N_B_Total <dbl>, 'Mean_CORT_(ng/g)_B_Total' <dbl>,
## #   Monthly_rain_mm_Bisate <dbl>, Monthly_minTemp_IESA <dbl>,
## #   Monthly_maxTemp_IESA <dbl>, Monthly_Altitude <dbl>,
## #   'Aud&Vis_Interactions_inclLSB' <dbl>,
## #   'Aud&Vis_Interactions_inclLSB_stressful' <dbl>, ...
```

```
str(stress_data)
```

```
## spc_tbl_ [1,094 x 23] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
##   $ Month                                     : chr [1:1094] "September" "October" "November" "December"
##   $ Year                                       : num [1:1094] 2011 2011 2011 2011 2012 ...
##   $ Gorilla                                   : chr [1:1094] "AFR" "AFR" "AFR" "AFR" ...
##   $ Group                                     : chr [1:1094] "PAB" "PAB" "PAB" "PAB" ...
##   $ Sex                                       : chr [1:1094] "F" "F" "F" "F" ...
##   $ Age                                       : num [1:1094] 11.4 11.4 11.5 11.6 11.7 ...
##   $ HomeR70                                  : num [1:1094] 964 359 722 433 558 ...
```

```

## $ HomeR70_excl : num [1:1094] 964 275 520 429 553 ...
## $ HomeR70_accum_overlap : num [1:1094] 44.79 83.9 292.42 32.11 5.68 ...
## $ HomeR_accum_overlap_index : num [1:1094] 2.57e-09 7.61e-08 9.05e-08 2.17e-08 2.46e-09
## $ N_Grp_HomeR70_overlap : num [1:1094] 2 1 4 2 1 0 0 0 1 0 ...
## $ N_B_Total : num [1:1094] 4 8 3 3 4 4 4 3 3 5 ...
## $ Mean_CORT_(ng/g)_B_Total : num [1:1094] 32.2 24.1 27.2 19.8 23.7 ...
## $ Monthly_rain_mm_Bisate : num [1:1094] 185.5 216.8 222.8 129.6 39.8 ...
## $ Monthly_minTemp_IESA : num [1:1094] 10.7 10.5 10.7 10.7 7.9 8 9.3 9.1 9.3 10.4 .
## $ Monthly_maxTemp_IESA : num [1:1094] 20.9 20.8 21.3 21.8 23.6 23.2 23.1 21.8 21.9
## $ Monthly_Altitude : num [1:1094] 3398 3392 3098 3371 3347 ...
## $ Aud&Vis_Interactions_inclLSB : num [1:1094] 0 0 2 0 0 0 0 1 0 1 ...
## $ Aud&Vis_Interactions_inclLSB_stressful : num [1:1094] 0 0 2 0 0 0 0 1 0 1 ...
## $ Mean_GrpSize_mo_rounded : num [1:1094] 45 45 45 45 45 45 45 43 43 44 ...
## $ Mean_N_SB_mo_rounded : num [1:1094] 5 5 5 5 5 5 5 4 4 4 ...
## $ GrpType : chr [1:1094] "Multi-Male" "Multi-Male" "Multi-Male" "Mult
## $ Mean_N_AF_mo_rounded : num [1:1094] 12 13 13 13 13 13 13 13 13 13 ...
## - attr(*, "spec")=
## .. cols(
## .. Month = col_character(),
## .. Year = col_double(),
## .. Gorilla = col_character(),
## .. Group = col_character(),
## .. Sex = col_character(),
## .. Age = col_double(),
## .. HomeR70 = col_double(),
## .. HomeR70_excl = col_double(),
## .. HomeR70_accum_overlap = col_double(),
## .. HomeR_accum_overlap_index = col_double(),
## .. N_Grp_HomeR70_overlap = col_double(),
## .. N_B_Total = col_double(),
## .. 'Mean_CORT_(ng/g)_B_Total' = col_double(),
## .. Monthly_rain_mm_Bisate = col_double(),
## .. Monthly_minTemp_IESA = col_double(),
## .. Monthly_maxTemp_IESA = col_double(),
## .. Monthly_Altitude = col_double(),
## .. 'Aud&Vis_Interactions_inclLSB' = col_double(),
## .. 'Aud&Vis_Interactions_inclLSB_stressful' = col_double(),
## .. Mean_GrpSize_mo_rounded = col_double(),
## .. Mean_N_SB_mo_rounded = col_double(),
## .. GrpType = col_character(),
## .. Mean_N_AF_mo_rounded = col_double()
## .. )
## - attr(*, "problems")=<externalptr>

```

```
colnames(stress_data)
```

```

## [1] "Month"
## [2] "Year"
## [3] "Gorilla"
## [4] "Group"
## [5] "Sex"
## [6] "Age"
## [7] "HomeR70"
## [8] "HomeR70_excl"

```

```
## [9] "HomeR70_accum_overlap"
## [10] "HomeR_accum_overlap_index"
## [11] "N_Grp_HomeR70_overlap"
## [12] "N_B_Total"
## [13] "Mean_CORT_(ng/g)_B_Total"
## [14] "Monthly_rain_mm_Bisate"
## [15] "Monthly_minTemp_IESA"
## [16] "Monthly_maxTemp_IESA"
## [17] "Monthly_Altitude"
## [18] "Aud&Vis_Interactions_inclLSB"
## [19] "Aud&Vis_Interactions_inclLSB_stressful"
## [20] "Mean_GrpSize_mo_rounded"
## [21] "Mean_N_SB_mo_rounded"
## [22] "GrpType"
## [23] "Mean_N_AF_mo_rounded"
```

Inspect the Habitat Dataset

```
# Check the first few rows and structure of the HABITAT dataset
head(habitat_data)
```

```
## # A tibble: 6 x 18
##   resp habitat z.rank z.group_size repr.state z.daily.time log.duration
##   <dbl> <chr>   <dbl>         <dbl> <chr>         <dbl>         <dbl>
## 1 -2.63 Hagenia -0.268         -1.09 lac          -1.29         -5.19
## 2 -1.73 Hagenia -0.268         -1.09 lac          -1.29         -4.86
## 3  0.402 Hagenia -0.268         -1.09 lac          -1.29         -2.79
## 4 -2.44 Hagenia -0.268         -1.09 lac          -1.29         -3.47
## 5 -1.11 Hagenia -0.268         -1.09 lac          -1.29         -4.24
## 6 -1.53 Hagenia -0.268         -1.09 lac          -1.29         -3.43
## # i 11 more variables: ac.term <dbl>, focal <chr>, group <chr>, date <dbl>,
## #   habitat.b <dbl>, habitat.c <dbl>, habitat.d <dbl>, habitat.e <dbl>,
## #   habitat.f <dbl>, repr.state.b <dbl>, repr.state.c <dbl>
```

```
str(habitat_data)
```

```
## tibble [3,565 x 18] (S3: tbl_df/tbl/data.frame)
## $ resp      : num [1:3565] -2.632 -1.731 0.402 -2.436 -1.11 ...
## $ habitat   : chr [1:3565] "Hagenia" "Hagenia" "Hagenia" "Hagenia" ...
## $ z.rank    : num [1:3565] -0.268 -0.268 -0.268 -0.268 -0.268 ...
## $ z.group_size: num [1:3565] -1.09 -1.09 -1.09 -1.09 -1.09 ...
## $ repr.state : chr [1:3565] "lac" "lac" "lac" "lac" ...
## $ z.daily.time: num [1:3565] -1.29 -1.29 -1.29 -1.29 -1.29 ...
## $ log.duration: num [1:3565] -5.19 -4.86 -2.79 -3.47 -4.24 ...
## $ ac.term    : num [1:3565] 0.418 0.318 0.245 0.606 0.299 ...
## $ focal     : chr [1:3565] "GIN" "GIN" "GIN" "GIN" ...
## $ group      : chr [1:3565] "BWE" "BWE" "BWE" "BWE" ...
## $ date       : num [1:3565] 40100 40100 40100 40100 40100 40100 40100 40100 40100 40100 ...
## $ habitat.b  : num [1:3565] 0 0 0 0 0 0 0 0 0 0 ...
## $ habitat.c  : num [1:3565] 0 0 0 0 0 0 0 0 0 0 ...
## $ habitat.d  : num [1:3565] 1 1 1 1 1 1 1 1 1 1 ...
```

```
## $ habitat.e : num [1:3565] 0 0 0 0 0 0 0 0 0 0 ...
## $ habitat.f : num [1:3565] 0 0 0 0 0 0 0 0 0 0 ...
## $ repr.state.b: num [1:3565] 1 1 1 1 1 1 1 1 1 1 ...
## $ repr.state.c: num [1:3565] 0 0 0 0 0 0 0 0 0 0 ...
```

Inspect Emissions Dataset

```
# Check the first few rows and structure of the EMISSIONS dataset
head(emissions_data)
```

```
## # A tibble: 6 x 34
##   Sector unit '1990' '1991' '1992' '1993' '1994' '1995' '1996' '1997' '1998'
##   <chr>   <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 Energy MtCO~ 6.83 5.48 5.26 6.44 7.59 7.56 8.11 8.29 8.18
## 2 Industri~ MtCO~ 0.22 0.14 0.12 0.12 0.14 0.18 0.23 0.23 0.28
## 3 Agricult~ MtCO~ 25.1 24.9 25.0 25.0 24.9 24.7 19.4 19.0 19.9
## 4 Waste MtCO~ 6.22 6.46 6.7 6.93 7.17 7.4 7.62 7.84 8.07
## 5 Land-Use~ MtCO~ 398. 398. 398. 398. 398. 398. 393. 400. 401.
## 6 Data sou~ Clim~ NA NA NA NA NA NA NA NA NA
## # i 23 more variables: '1999' <dbl>, '2000' <dbl>, '2001' <dbl>, '2002' <dbl>,
## # '2003' <dbl>, '2004' <dbl>, '2005' <dbl>, '2006' <dbl>, '2007' <dbl>,
## # '2008' <dbl>, '2009' <dbl>, '2010' <dbl>, '2011' <dbl>, '2012' <dbl>,
## # '2013' <dbl>, '2014' <dbl>, '2015' <dbl>, '2016' <dbl>, '2017' <dbl>,
## # '2018' <dbl>, '2019' <dbl>, '2020' <dbl>, '2021' <dbl>
```

```
str(emissions_data)
```

```
## spc_tbl_ [6 x 34] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
## $ Sector: chr [1:6] "Energy" "Industrial Processes" "Agriculture" "Waste" ...
## $ unit : chr [1:6] "MtCO2e" "MtCO2e" "MtCO2e" "MtCO2e" ...
## $ 1990 : num [1:6] 6.83 0.22 25.13 6.22 398.44 ...
## $ 1991 : num [1:6] 5.48 0.14 24.94 6.46 398.44 ...
## $ 1992 : num [1:6] 5.26 0.12 25.05 6.7 398.44 ...
## $ 1993 : num [1:6] 6.44 0.12 24.95 6.93 398.44 ...
## $ 1994 : num [1:6] 7.59 0.14 24.89 7.17 398.44 ...
## $ 1995 : num [1:6] 7.56 0.18 24.73 7.4 398.44 ...
## $ 1996 : num [1:6] 8.11 0.23 19.43 7.62 393.27 ...
## $ 1997 : num [1:6] 8.29 0.23 18.99 7.84 399.94 ...
## $ 1998 : num [1:6] 8.18 0.28 19.87 8.07 400.82 ...
## $ 1999 : num [1:6] 7.87 0.34 18.07 8.29 394.81 ...
## $ 2000 : num [1:6] 7.33 0.39 17.7 8.53 391.68 ...
## $ 2001 : num [1:6] 7.52 0.43 18.93 8.76 386.89 ...
## $ 2002 : num [1:6] 7.65 0.47 23.92 9.02 396.01 ...
## $ 2003 : num [1:6] 8.44 0.52 26.21 9.35 400.2 ...
## $ 2004 : num [1:6] 8.33 0.56 28.08 9.63 401.98 ...
## $ 2005 : num [1:6] 8.57 0.63 27.44 9.92 401.77 ...
## $ 2006 : num [1:6] 8.74 0.7 28.11 10.26 402.28 ...
## $ 2007 : num [1:6] 9.25 0.78 24.73 10.61 396.67 ...
## $ 2008 : num [1:6] 9.46 0.81 26.05 10.95 400.39 ...
## $ 2009 : num [1:6] 9.37 0.91 25.84 11.29 400.04 ...
## $ 2010 : num [1:6] 9.54 1 27.22 11.64 402.16 ...
```

```

## $ 2011 : num [1:6] 10.1 1.14 25.52 12.04 638.39 ...
## $ 2012 : num [1:6] 8.26 1.27 24.89 12.45 635.5 ...
## $ 2013 : num [1:6] 10.06 1.44 26.48 12.86 634.89 ...
## $ 2014 : num [1:6] 12.12 1.54 25.81 13.26 632.09 ...
## $ 2015 : num [1:6] 10.33 1.72 26.86 13.67 633.61 ...
## $ 2016 : num [1:6] 9.71 1.84 27.01 14.13 633.41 ...
## $ 2017 : num [1:6] 10.75 2.27 27.72 14.6 633.9 ...
## $ 2018 : num [1:6] 9.55 2.55 26.22 15.07 629.08 ...
## $ 2019 : num [1:6] 9.95 2.81 26.9 15.54 632.42 ...
## $ 2020 : num [1:6] 10.25 2.95 27.25 16 631.36 ...
## $ 2021 : num [1:6] 10.24 3.19 27.82 16.54 629.44 ...
## - attr(*, "spec")=
## .. cols(
## ..   Sector = col_character(),
## ..   unit = col_character(),
## ..   '1990' = col_double(),
## ..   '1991' = col_double(),
## ..   '1992' = col_double(),
## ..   '1993' = col_double(),
## ..   '1994' = col_double(),
## ..   '1995' = col_double(),
## ..   '1996' = col_double(),
## ..   '1997' = col_double(),
## ..   '1998' = col_double(),
## ..   '1999' = col_double(),
## ..   '2000' = col_double(),
## ..   '2001' = col_double(),
## ..   '2002' = col_double(),
## ..   '2003' = col_double(),
## ..   '2004' = col_double(),
## ..   '2005' = col_double(),
## ..   '2006' = col_double(),
## ..   '2007' = col_double(),
## ..   '2008' = col_double(),
## ..   '2009' = col_double(),
## ..   '2010' = col_double(),
## ..   '2011' = col_double(),
## ..   '2012' = col_double(),
## ..   '2013' = col_double(),
## ..   '2014' = col_double(),
## ..   '2015' = col_double(),
## ..   '2016' = col_double(),
## ..   '2017' = col_double(),
## ..   '2018' = col_double(),
## ..   '2019' = col_double(),
## ..   '2020' = col_double(),
## ..   '2021' = col_double()
## .. )
## - attr(*, "problems")=<externalptr>

```

Clean the Stress Dataset

```
# Check for NA's
summary(stress_data)
```

```
##      Month      Year      Gorilla      Group
## Length:1094    Min.   :2011    Length:1094    Length:1094
## Class :character 1st Qu.:2011    Class :character  Class :character
## Mode  :character Median :2012    Mode  :character  Mode  :character
##              Mean  :2012
##              3rd Qu.:2012
##              Max.   :2012
##      Sex      Age      Homer70      Homer70_excl
## Length:1094    Min.   : 0.93    Min.   : 21.15    Min.   : 0.00
## Class :character 1st Qu.: 5.07    1st Qu.: 137.46    1st Qu.: 58.53
## Mode  :character Median :12.54    Median : 201.17    Median : 111.16
##              Mean  :13.17    Mean  : 266.40    Mean  : 161.98
##              3rd Qu.:19.30    3rd Qu.: 359.24    3rd Qu.: 196.31
##              Max.   :33.97    Max.   :1293.33    Max.   :1230.78
## Homer70_accum_overlap Homer_accum_overlap_index N_Grp_Homer70_overlap
## Min.   : 0.00      Min.   :0.000e+00      Min.   :0.000
## 1st Qu.: 44.79      1st Qu.:4.095e-08      1st Qu.:2.000
## Median : 121.44      Median :2.390e-07      Median :4.000
## Mean   : 282.12      Mean   :2.344e-07      Mean   :3.204
## 3rd Qu.: 403.39      3rd Qu.:3.720e-07      3rd Qu.:5.000
## Max.   :1959.35      Max.   :1.030e-06      Max.   :7.000
## N_B_Total      Mean_CORT_(ng/g)_B_Total      Monthly_rain_mm_Bisate
## Min.   : 3.000      Min.   : 9.341      Min.   : 10.5
## 1st Qu.: 3.000      1st Qu.: 20.804      1st Qu.:120.4
## Median : 4.000      Median : 25.050      Median :157.6
## Mean   : 4.037      Mean   : 26.698      Mean   :158.9
## 3rd Qu.: 5.000      3rd Qu.: 29.880      3rd Qu.:208.6
## Max.   :11.000      Max.   :102.085      Max.   :340.3
## Monthly_minTemp_IESA Monthly_maxTemp_IESA Monthly_Altitude
## Min.   : 7.900      Min.   :19.80      Min.   :2852
## 1st Qu.: 9.300      1st Qu.:20.80      1st Qu.:2944
## Median :10.400      Median :21.80      Median :3038
## Mean   : 9.871      Mean   :21.72      Mean   :3090
## 3rd Qu.:10.700      3rd Qu.:22.10      3rd Qu.:3188
## Max.   :11.100      Max.   :23.60      Max.   :3651
## Aud&Vis_Interactions_inclLSB Aud&Vis_Interactions_inclLSB_stressful
## Min.   :0.0000      Min.   :0.0000
## 1st Qu.:0.0000      1st Qu.:0.0000
## Median :0.0000      Median :0.0000
## Mean   :0.7157      Mean   :0.5969
## 3rd Qu.:1.0000      3rd Qu.:1.0000
## Max.   :5.0000      Max.   :5.0000
## Mean_GrpSize_mo_rounded Mean_N_SB_mo_rounded      GrpType
## Min.   : 3.00      Min.   :1.000      Length:1094
## 1st Qu.: 9.00      1st Qu.:2.000      Class :character
## Median :11.00      Median :2.000      Mode  :character
## Mean   :17.34      Mean   :2.509
## 3rd Qu.:14.00      3rd Qu.:3.000
```

```
## Max.      :47.00          Max.      :6.000
## Mean_N_AF_mo_rounded
## Min.      : 1.000
## 1st Qu.: 3.000
## Median : 5.000
## Mean      : 5.825
## 3rd Qu.: 6.000
## Max.      :13.000
```

```
sum(is.na(stress_data))
```

```
## [1] 0
```

Clean the Habitat Dataset

```
# Check for NA's
summary(habitat_data)
```

```
##      resp      habitat      z.rank      z.group_size
## Min.      :-4.3621 Length:3565 Min.      :-1.4602 Min.      :-1.2222
## 1st Qu.: -1.5734 Class :character 1st Qu.: -0.8639 1st Qu.: -1.0893
## Median : -0.6579 Mode  :character Median : -0.1037 Median : 0.9038
## Mean      : -0.6408 Mean      : 0.0000 Mean      : 0.0000
## 3rd Qu.: 0.1148 3rd Qu.: 0.7757 3rd Qu.: 0.9038
## Max.      : 6.3994 Max.      : 1.5210 Max.      : 1.0366
##
## repr.state      z.daily.time      log.duration      ac.term
## Length:3565 Min.      :-2.26842 Min.      :-6.802 Min.      :-8.045272
## Class :character 1st Qu.: -0.78042 1st Qu.: -4.128 1st Qu.: -0.500168
## Mode  :character Median : 0.02284 Median : -3.507 Median : -0.002277
## Mean      : 0.00000 Mean      : -3.506 Mean      : 0.000000
## 3rd Qu.: 0.77343 3rd Qu.: -2.856 3rd Qu.: 0.463325
## Max.      : 2.93301 Max.      : -1.017 Max.      : 8.277364
##
## focal      group      date      habitat.b
## Length:3565 Length:3565 Min.      :40095 Min.      :0.0000
## Class :character Class :character 1st Qu.:40221 1st Qu.:0.0000
## Mode  :character Mode  :character Median :40339 Median :0.0000
## Mean      :40327 Mean      :0.1481
## 3rd Qu.:40437 3rd Qu.:0.0000
## Max.      :40540 Max.      :1.0000
##
## habitat.c      habitat.d      habitat.e      habitat.f
## Min.      :0.00000 Min.      :0.0000 Min.      :0.0000 Min.      :0.0000
## 1st Qu.:0.00000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.0000
## Median :0.00000 Median :0.0000 Median :0.0000 Median :0.0000
## Mean      :0.01178 Mean      :0.3495 Mean      :0.2376 Mean      :0.2432
## 3rd Qu.:0.00000 3rd Qu.:1.0000 3rd Qu.:0.0000 3rd Qu.:0.0000
## Max.      :1.00000 Max.      :1.0000 Max.      :1.0000 Max.      :1.0000
##
## repr.state.b      repr.state.c
```



```
## Min. :0.0000 Min. :0.0000
## 1st Qu.:0.0000 1st Qu.:0.0000
## Median :1.0000 Median :0.0000
## Mean :0.5646 Mean :0.1863
## 3rd Qu.:1.0000 3rd Qu.:0.0000
## Max. :1.0000 Max. :1.0000
## NA's :76 NA's :76
```

```
sum(is.na(habitat_data))
```

```
## [1] 152
```

```
# Identify which columns have NA's
colSums(is.na(habitat_data))
```

```
##      resp      habitat      z.rank z.group_size repr.state z.daily.time
##      0          0          0          0          0          0
## log.duration ac.term      focal      group      date      habitat.b
##      0          0          0          0          0          0
## habitat.c    habitat.d    habitat.e    habitat.f repr.state.b repr.state.c
##      0          0          0          0          76          76
```

```
# Delete columns repr.state (reproductive state)
# We have chosen to delete these columns due to the missing (NA) values. While there are risks to removing columns, we have chosen to delete these columns due to the missing (NA) values.
habitat_data <- habitat_data[ , !names(habitat_data) %in% c("repr.state.b", "repr.state.c")]
```

```
# Verify those column names are deleted
colnames(habitat_data)
```

```
## [1] "resp"      "habitat"    "z.rank"     "z.group_size" "repr.state"
## [6] "z.daily.time" "log.duration" "ac.term"    "focal"        "group"
## [11] "date"       "habitat.b"  "habitat.c"  "habitat.d"    "habitat.e"
## [16] "habitat.f"
```

Clean the Emissions Dataset

```
# Check for NA's
summary(emissions_data)
```

```
##      Sector      unit      1990      1991
## Length:6      Length:6      Min. : 0.22 Min. : 0.14
## Class :character Class :character 1st Qu.: 6.22 1st Qu.: 5.48
## Mode :character Mode :character Median : 6.83 Median : 6.46
##      Mean : 87.37 Mean : 87.09
##      3rd Qu.: 25.13 3rd Qu.: 24.94
##      Max. :398.44 Max. :398.44
##      NA's :1      NA's :1
##      1992      1993      1994      1995
## Min. : 0.12 Min. : 0.12 Min. : 0.14 Min. : 0.18
## 1st Qu.: 5.26 1st Qu.: 6.44 1st Qu.: 7.17 1st Qu.: 7.40
```

##	Median :	6.70	Median :	6.93	Median :	7.59	Median :	7.56
##	Mean :	87.11	Mean :	87.38	Mean :	87.65	Mean :	87.66
##	3rd Qu.:	25.05	3rd Qu.:	24.95	3rd Qu.:	24.89	3rd Qu.:	24.73
##	Max. :	:398.44	Max. :	:398.44	Max. :	:398.44	Max. :	:398.44
##	NA's :	:1	NA's :	:1	NA's :	:1	NA's :	:1
##		1996		1997		1998		1999
##	Min. :	0.23	Min. :	0.23	Min. :	0.28	Min. :	0.34
##	1st Qu.:	7.62	1st Qu.:	7.84	1st Qu.:	8.07	1st Qu.:	7.87
##	Median :	8.11	Median :	8.29	Median :	8.18	Median :	8.29
##	Mean :	85.73	Mean :	87.06	Mean :	87.44	Mean :	85.88
##	3rd Qu.:	19.43	3rd Qu.:	18.99	3rd Qu.:	19.87	3rd Qu.:	18.07
##	Max. :	:393.27	Max. :	:399.94	Max. :	:400.82	Max. :	:394.81
##	NA's :	:1	NA's :	:1	NA's :	:1	NA's :	:1
##		2000		2001		2002		2003
##	Min. :	0.39	Min. :	0.43	Min. :	0.47	Min. :	0.52
##	1st Qu.:	7.33	1st Qu.:	7.52	1st Qu.:	7.65	1st Qu.:	8.44
##	Median :	8.53	Median :	8.76	Median :	9.02	Median :	9.35
##	Mean :	85.13	Mean :	84.51	Mean :	87.41	Mean :	88.94
##	3rd Qu.:	17.70	3rd Qu.:	18.93	3rd Qu.:	23.92	3rd Qu.:	26.21
##	Max. :	:391.68	Max. :	:386.89	Max. :	:396.01	Max. :	:400.20
##	NA's :	:1	NA's :	:1	NA's :	:1	NA's :	:1
##		2004		2005		2006		2007
##	Min. :	0.56	Min. :	0.63	Min. :	0.70	Min. :	0.78
##	1st Qu.:	8.33	1st Qu.:	8.57	1st Qu.:	8.74	1st Qu.:	9.25
##	Median :	9.63	Median :	9.92	Median :	10.26	Median :	10.61
##	Mean :	89.72	Mean :	89.67	Mean :	90.02	Mean :	88.41
##	3rd Qu.:	28.08	3rd Qu.:	27.44	3rd Qu.:	28.11	3rd Qu.:	24.73
##	Max. :	:401.98	Max. :	:401.77	Max. :	:402.28	Max. :	:396.67
##	NA's :	:1	NA's :	:1	NA's :	:1	NA's :	:1
##		2008		2009		2010		2011
##	Min. :	0.81	Min. :	0.91	Min. :	1.00	Min. :	1.14
##	1st Qu.:	9.46	1st Qu.:	9.37	1st Qu.:	9.54	1st Qu.:	10.10
##	Median :	10.95	Median :	11.29	Median :	11.64	Median :	12.04
##	Mean :	89.53	Mean :	89.49	Mean :	90.31	Mean :	:137.44
##	3rd Qu.:	26.05	3rd Qu.:	25.84	3rd Qu.:	27.22	3rd Qu.:	25.52
##	Max. :	:400.39	Max. :	:400.04	Max. :	:402.16	Max. :	:638.39
##	NA's :	:1	NA's :	:1	NA's :	:1	NA's :	:1
##		2012		2013		2014		2015
##	Min. :	1.27	Min. :	1.44	Min. :	1.54	Min. :	1.72
##	1st Qu.:	8.26	1st Qu.:	10.06	1st Qu.:	12.12	1st Qu.:	10.33
##	Median :	12.45	Median :	12.86	Median :	13.26	Median :	13.67
##	Mean :	:136.47	Mean :	:137.15	Mean :	:136.96	Mean :	:137.24
##	3rd Qu.:	24.89	3rd Qu.:	26.48	3rd Qu.:	25.81	3rd Qu.:	26.86
##	Max. :	:635.50	Max. :	:634.89	Max. :	:632.09	Max. :	:633.61
##	NA's :	:1	NA's :	:1	NA's :	:1	NA's :	:1
##		2016		2017		2018		2019
##	Min. :	1.84	Min. :	2.27	Min. :	2.55	Min. :	2.81
##	1st Qu.:	9.71	1st Qu.:	10.75	1st Qu.:	9.55	1st Qu.:	9.95
##	Median :	14.13	Median :	14.60	Median :	15.07	Median :	15.54
##	Mean :	:137.22	Mean :	:137.85	Mean :	:136.49	Mean :	:137.52
##	3rd Qu.:	27.01	3rd Qu.:	27.72	3rd Qu.:	26.22	3rd Qu.:	26.90
##	Max. :	:633.41	Max. :	:633.90	Max. :	:629.08	Max. :	:632.42
##	NA's :	:1	NA's :	:1	NA's :	:1	NA's :	:1
##		2020		2021				

```
## Min.    : 2.95    Min.    : 3.19
## 1st Qu.: 10.25    1st Qu.: 10.24
## Median : 16.00    Median : 16.54
## Mean   :137.56    Mean   :137.45
## 3rd Qu.: 27.25    3rd Qu.: 27.82
## Max.   :631.36    Max.   :629.44
## NA's   :1         NA's   :1
```

```
sum(is.na(emissions_data))
```

```
## [1] 32
```

Step 3: Condense the Data

Condense the Stress Dataset

```
clean_stress_data <- stress_data %>%
  distinct() %>%
  filter(!is.na(`Mean_CORT_(ng/g)_B_Total`)) %>%
  select(Month, Year, Gorilla, Group, Sex,
         `Mean_CORT_(ng/g)_B_Total`,
         Monthly_rain_mm_Bisate,
         Monthly_minTemp_IESA,
         Monthly_maxTemp_IESA)

# View summary
glimpse(clean_stress_data)
```

```
## Rows: 1,094
## Columns: 9
## $ Month          <chr> "September", "October", "November", "Decemb~
## $ Year            <dbl> 2011, 2011, 2011, 2011, 2012, 2012, 2012, 2~
## $ Gorilla         <chr> "AFR", "AFR", "AFR", "AFR", "AFR", "AFR", "~
## $ Group           <chr> "PAB", "PAB", "PAB", "PAB", "PAB", "PAB", "~
## $ Sex             <chr> "F", "F", "F", "F", "F", "F", "F", "F", "F"~
## $ `Mean_CORT_(ng/g)_B_Total` <dbl> 32.19544, 24.10109, 27.24359, 19.77390, 23.~
## $ Monthly_rain_mm_Bisate <dbl> 185.5, 216.8, 222.8, 129.6, 39.8, 157.6, 19~
## $ Monthly_minTemp_IESA <dbl> 10.7, 10.5, 10.7, 10.7, 7.9, 8.0, 9.3, 9.1,~
## $ Monthly_maxTemp_IESA <dbl> 20.9, 20.8, 21.3, 21.8, 23.6, 23.2, 23.1, 2~
```

Condense the Habitat Dataset

```
clean_habitat_data <- habitat_data %>%
  distinct() %>%
  filter(!is.na(resp)) %>%
  select(resp, habitat, z.rank, z.group_size, z.daily.time, log.duration, focal, group, date, habitat.b

# View summary
glimpse(clean_habitat_data)
```

```
## Rows: 3,563
## Columns: 14
## $ resp      <dbl> -2.6316206, -1.7305311, 0.4024652, -2.4362103, -1.1095561~
## $ habitat    <chr> "Hagenia", "Hagenia", "Hagenia", "Hagenia", "Hagenia", "H~
## $ z.rank     <dbl> -0.2676756, -0.2676756, -0.2676756, -0.2676756, -0.267675~
## $ z.group_size <dbl> -1.089282, -1.089282, -1.089282, -1.089282, -1.089282, -1~
## $ z.daily.time <dbl> -1.29397523, -1.29397523, -1.29397523, -1.29397523, -1.29~
## $ log.duration <dbl> -5.192957, -4.856485, -2.790526, -3.470190, -4.237445, -3~
## $ focal      <chr> "GIN", "GIN", "GIN", "GIN", "GIN", "GIN", "GIN", "GIN", "~
## $ group      <chr> "BWE", "BWE", "BWE", "BWE", "BWE", "BWE", "BWE", "BWE", "~
## $ date       <dbl> 40100, 40100, 40100, 40100, 40100, 40100, 40100, 40100, 4~
## $ habitat.b  <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ~
## $ habitat.c  <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ~
## $ habitat.d  <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, ~
## $ habitat.e  <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ~
## $ habitat.f  <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ~
```

Condense the Emissions Dataset

```
clean_emissions_data <- emissions_data %>%
  distinct() %>%
  filter(!is.na(Sector)) %>% # Filter out rows with NA in the Sector column
  select(Sector, unit, `1990`, `1991`, `1992`, `1993`, `1994`, `1995`, `1996`,
    `1997`, `1998`, `1999`, `2000`, `2001`, `2002`, `2003`, `2004`,
    `2005`, `2006`, `2007`, `2008`, `2009`, `2010`, `2011`, `2012`,
    `2013`, `2014`, `2015`, `2016`, `2017`, `2018`, `2019`, `2020`,
    `2021`) # Include all the years for emissions data

# View summary
glimpse(clean_emissions_data)
```

```
## Rows: 6
## Columns: 34
## $ Sector <chr> "Energy", "Industrial Processes", "Agriculture", "Waste", "Land~
## $ unit <chr> "MtCO2e", "MtCO2e", "MtCO2e", "MtCO2e", "MtCO2e", "Climate Watc~
## $ `1990` <dbl> 6.83, 0.22, 25.13, 6.22, 398.44, NA
## $ `1991` <dbl> 5.48, 0.14, 24.94, 6.46, 398.44, NA
## $ `1992` <dbl> 5.26, 0.12, 25.05, 6.70, 398.44, NA
## $ `1993` <dbl> 6.44, 0.12, 24.95, 6.93, 398.44, NA
## $ `1994` <dbl> 7.59, 0.14, 24.89, 7.17, 398.44, NA
## $ `1995` <dbl> 7.56, 0.18, 24.73, 7.40, 398.44, NA
## $ `1996` <dbl> 8.11, 0.23, 19.43, 7.62, 393.27, NA
## $ `1997` <dbl> 8.29, 0.23, 18.99, 7.84, 399.94, NA
## $ `1998` <dbl> 8.18, 0.28, 19.87, 8.07, 400.82, NA
## $ `1999` <dbl> 7.87, 0.34, 18.07, 8.29, 394.81, NA
## $ `2000` <dbl> 7.33, 0.39, 17.70, 8.53, 391.68, NA
## $ `2001` <dbl> 7.52, 0.43, 18.93, 8.76, 386.89, NA
## $ `2002` <dbl> 7.65, 0.47, 23.92, 9.02, 396.01, NA
## $ `2003` <dbl> 8.44, 0.52, 26.21, 9.35, 400.20, NA
## $ `2004` <dbl> 8.33, 0.56, 28.08, 9.63, 401.98, NA
## $ `2005` <dbl> 8.57, 0.63, 27.44, 9.92, 401.77, NA
## $ `2006` <dbl> 8.74, 0.70, 28.11, 10.26, 402.28, NA
```

```
## $ '2007' <dbl> 9.25, 0.78, 24.73, 10.61, 396.67, NA
## $ '2008' <dbl> 9.46, 0.81, 26.05, 10.95, 400.39, NA
## $ '2009' <dbl> 9.37, 0.91, 25.84, 11.29, 400.04, NA
## $ '2010' <dbl> 9.54, 1.00, 27.22, 11.64, 402.16, NA
## $ '2011' <dbl> 10.10, 1.14, 25.52, 12.04, 638.39, NA
## $ '2012' <dbl> 8.26, 1.27, 24.89, 12.45, 635.50, NA
## $ '2013' <dbl> 10.06, 1.44, 26.48, 12.86, 634.89, NA
## $ '2014' <dbl> 12.12, 1.54, 25.81, 13.26, 632.09, NA
## $ '2015' <dbl> 10.33, 1.72, 26.86, 13.67, 633.61, NA
## $ '2016' <dbl> 9.71, 1.84, 27.01, 14.13, 633.41, NA
## $ '2017' <dbl> 10.75, 2.27, 27.72, 14.60, 633.90, NA
## $ '2018' <dbl> 9.55, 2.55, 26.22, 15.07, 629.08, NA
## $ '2019' <dbl> 9.95, 2.81, 26.90, 15.54, 632.42, NA
## $ '2020' <dbl> 10.25, 2.95, 27.25, 16.00, 631.36, NA
## $ '2021' <dbl> 10.24, 3.19, 27.82, 16.54, 629.44, NA
```

Linear Regression:

Predicting Future Greenhouse Gas Emissions in the DRC

```
# Reshaping the data for modeling (emissions by year)
emissions_long <- clean_emissions_data %>%
  gather(key = "year", value = "emissions", `1990`:`2021`) %>%
  mutate(year = as.numeric(year))

# Linear model for greenhouse gas emissions prediction
emissions_lm <- lm(emissions ~ year, data = emissions_long)

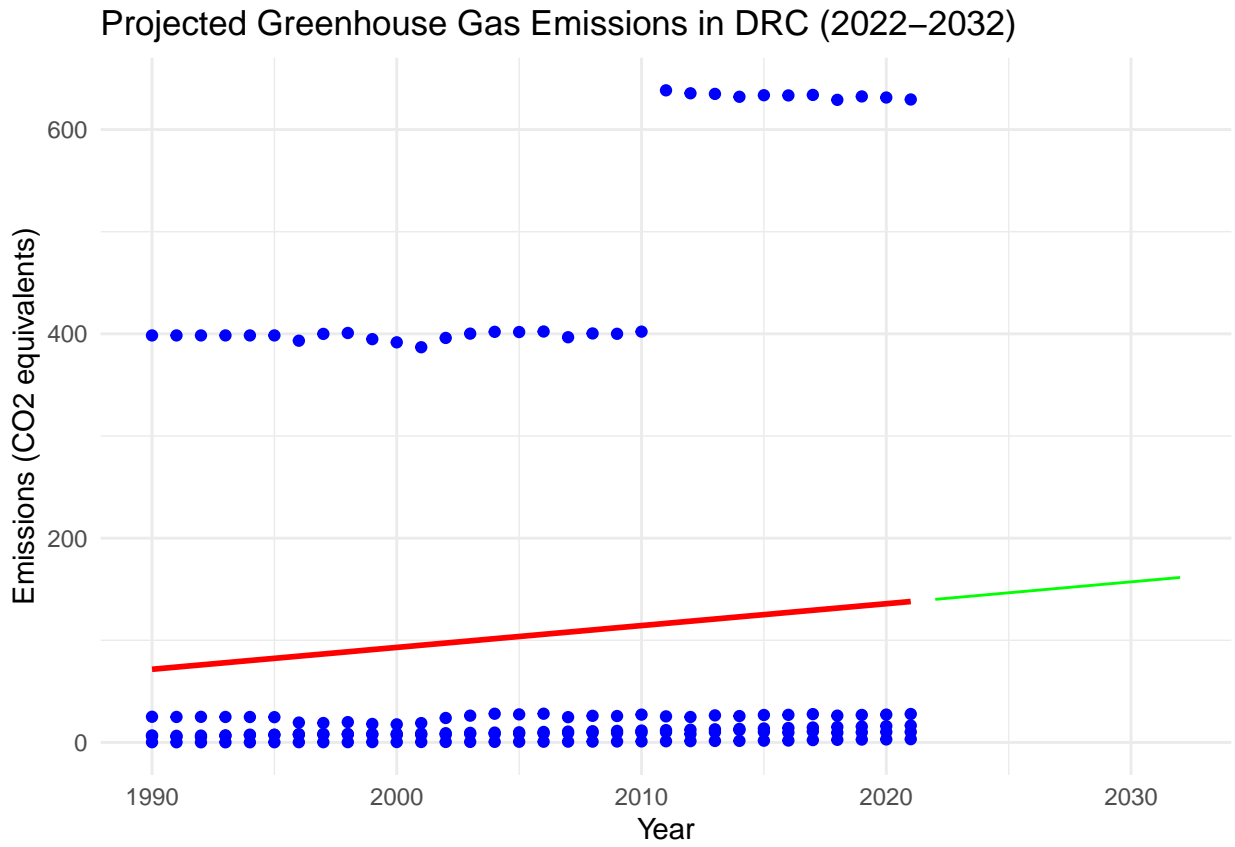
# Predict future emissions (next 10 years)
future_years <- data.frame(year = seq(2022, 2032, by = 1))
predicted_emissions <- predict(emissions_lm, newdata = future_years)

# Plot historical emissions and future predictions
ggplot(emissions_long, aes(x = year, y = emissions)) +
  geom_point(color = "blue") +
  geom_smooth(method = "lm", se = FALSE, color = "red") +
  geom_line(data = future_years, aes(x = year, y = predicted_emissions), color = "green") +
  labs(title = "Projected Greenhouse Gas Emissions in DRC (2022-2032)",
       x = "Year", y = "Emissions (CO2 equivalents)") +
  theme_minimal()
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```

```
## Warning: Removed 32 rows containing non-finite outside the scale range
## ('stat_smooth()').
```

```
## Warning: Removed 32 rows containing missing values or values outside the scale range
## ('geom_point()').
```



My Analysis:

Data Cleaning and Preparation:

I started off by removing irrelevant or missing data to ensure the datasets were ready for analysis and to assist with readability. One of the biggest issues we had to address was the missing values for reproductive state. We can assume that the NA in reproductive state is because they are male gorillas. Therefore, we just deleted that column as it did not affect the study of their habitat or stress levels.

Dataset Finds:

Stress Levels:

One of the biggest correlations I discovered is the strong correlation between temperature and cortisol levels in gorillas. As the temperatures rose, so did their cortisol levels. This tells us that with the rise of temperatures gorillas become more stressed and may potentially change their forage behaviors.

Habitat Use:

As previously mentioned, changes in gorillas environment can contribute to changes in foraging behaviors. I discovered the same thing with this data set.

Greenhouse Gases:

Based on the Linear Regression I performed 'Predicting Future Greenhouse Gas Emissions in the DRC', it appears that greenhouse gases will rise pretty steadily over the next 10 years. This is not great news for gorilla populations as we have learned gorillas cortisol levels increase when the temperature increases as well.

Implications:

The findings from this project really stress the urgency of addressing climate change in an effort to help the critically endangered species, the mountain gorilla. We have learned that gorillas stress levels increase due to rising temperatures, which in turn causes them to venture out of their home territory. Trends also show us that these rising temperatures will get even worse, so it is imperative we take action now to protect these animals.

Luckily, there are things we can do to help prevent the loss of this great species. We can work to reduce greenhouse gases as well as work to restore gorillas habitats and resources.

Limitations:

Data Gaps:

There were a couple of missing data values in our datasets that could potentially have had an effect on our results.

Regional Accuracy:

Although mountain gorillas do live in the DRC²³, this dataset was not specific on which part of the DRC these numbers are from, so we cannot be 100% certain these are the temperatures these gorillas are dealing with.