Analysis of Global Crises and Their Impact on Interational Well-Being

Group 8

November 12, 2024

List your group members, including their student numbers, here:

«««< HEAD

- Aditya Chauhan (169027493)
- Andy Ferrer Lorenzo (169081356)
- Oscar Saul (169057882)
- Taras Zakydalsky (169099616)
- Ezra Furman (#########) »»»> 5f7fee6584554860c6964f24cabaeb60755a72fe

You **must** be in a group in MyLS in order to see the DropBox used for submission. Even if you're alone, you must join a group by yourself.

You **must** be in a group with people from the same section as you. MyLS does not allow for groups including students from both Data100A and Data100B.

Instructions

You are encouraged to remove this instruction section prior to submission.

It is recommended that you follow the structure of this template. The text is all placeholder - you are free to change any/all wording as you please, but it is very helpful for the grading process if you keep the same structure. Anything in <> definitely needs to be changed, but you are free to change any/all sentences!

Note that all of the code is *hidden* by default. This file will be graded based on the insights, not the code.

You will only submit the PDF version of this document. To knit to PDF, you'll need to run `install.packages("tinytex")` in the console, followed by `tinytex::install_tinytex()` (DO NOT PUT THESE COMMANDS IN AN RMD FILE!!!). If you encounter errors in "Knit to PDF", you can "knit to html" and then print the html file to PDF using your operating system's PDF view (e.g. Adobe Acrobat). Only standalone PDF files will be accepted by MyLS.

Abstract

Introduction

Climate change is something that has been studied. Here's some relevant information about the context of our study.

If needed, this paragraph is more information about the context.

In this report, we are going to explore some aspects climate change and the impact and/or perceptions of it by using exploratory techniques. We'll explore <> using <>.

By the end of this report, we will have shown ...

Data Description

«Cyclones Data»

```
## Simple feature collection with 84601 features and 34 fields
## Geometry type: POINT
## Dimension:     XY
## Bounding box:  xmin: -180 ymin: 0.4 xmax: 179.9 ymax: 83
## Geodetic CRS:  WGS 84
## # A tibble: 84,601 x 35
##    Basin Number NameYear Name    ObservYear Month   Day  Hour Minute Identifier status latitude long
##  * <chr>  <int>    <int> <chr>        <int> <int> <int> <int>  <int> <chr>      <chr>  <chr>    <chr
##  1 AL         1     1851 UNNAMED       1851     6    25     0      0 <NA>       HU     28.0N    94.8
##  2 AL         1     1851 UNNAMED       1851     6    25     6      0 <NA>       HU     28.0N    95.4
##  3 AL         1     1851 UNNAMED       1851     6    25    12      0 <NA>       HU     28.0N    96.0
##  4 AL         1     1851 UNNAMED       1851     6    25    18      0 <NA>       HU     28.1N    96.5
##  5 AL         1     1851 UNNAMED       1851     6    25    21      0 L          HU     28.2N    96.8
##  6 AL         1     1851 UNNAMED       1851     6    26     0      0 <NA>       HU     28.2N    97.0
##  7 AL         1     1851 UNNAMED       1851     6    26     6      0 <NA>       TS     28.3N    97.6
##  8 AL         1     1851 UNNAMED       1851     6    26    12      0 <NA>       TS     28.4N    98.3
##  9 AL         1     1851 UNNAMED       1851     6    26    18      0 <NA>       TS     28.6N    98.9
## 10 AL         1     1851 UNNAMED       1851     6    27     0      0 <NA>       TS     29.0N    99.4
## # i 84,591 more rows
## # i 11 more variables: SE_extend_64 <dbl>, SW_extend_64 <dbl>, NW_extend_64 <dbl>, r_max_wind <dbl>,
```

The data set `cyclone_data` provided by the National Hurricane Center and the Central Pacific Hurricane Center records hurricane entries in the Atlantic and North Pacific basins from 1851 to 2022 and 1949 to 2022, respectively. The data frame includes detailed time stamps, coordinates, and categories, which are vital for analysis.

The preliminary step in supporting the analysis involved converting the coordinates into the Cartesian coordinate system. Next, we used the `rnaturalearth` package and `geolocation` to convert the coordinates into country identifiers and determine each storm's distance from the equator. As shown in Figure 1 in the appendix, cyclones are now distributed across the North Atlantic and Northeast Pacific basins.

«World Covid 2020»

```
## # A tibble: 252 x 6
## # Groups:   country [250]
##    country             continent     total_cases hemisphere distance_to_equator_km
##    <chr>               <chr>               <dbl> <chr>                      <dbl>
##  1 Afghanistan         Asia              8284992 North                      4017. (((7405225 448779
##  2 Africa              <NA>            285057531 <NA>                         NA
##  3 Albania             Europe            3444855 North                      5035. (((2153184 514139
##  4 Algeria             Africa           10188569 North                      3302. (((954738.6 443035
##  5 American Samoa      Oceania                 0 South                      1609. (((-19005161 -161
```

```
##  6 Andorra             Europe            650844 North                    5243. (((189917.1 52366
##  7 Angola              Africa           1275899 South                    1386. (((1455252 -516503
##  8 Anguilla            North America       1005 North                    2064. (((-7013264 206351
##  9 Antigua and Barbuda North America      22209 <NA>                       NA
## 10 Argentina           South America         NA South                    4386. (((-6412992 -3527
## # i 242 more rows
```

«World Happiness Census»

Combining the Data

Explain how any combinations of data were performed. Explain what kind of join was needed, whether columns had to be modified (for example, matching "country" names.)

Exploratory Data Analysis

To achieve our goals, we explored the data by...

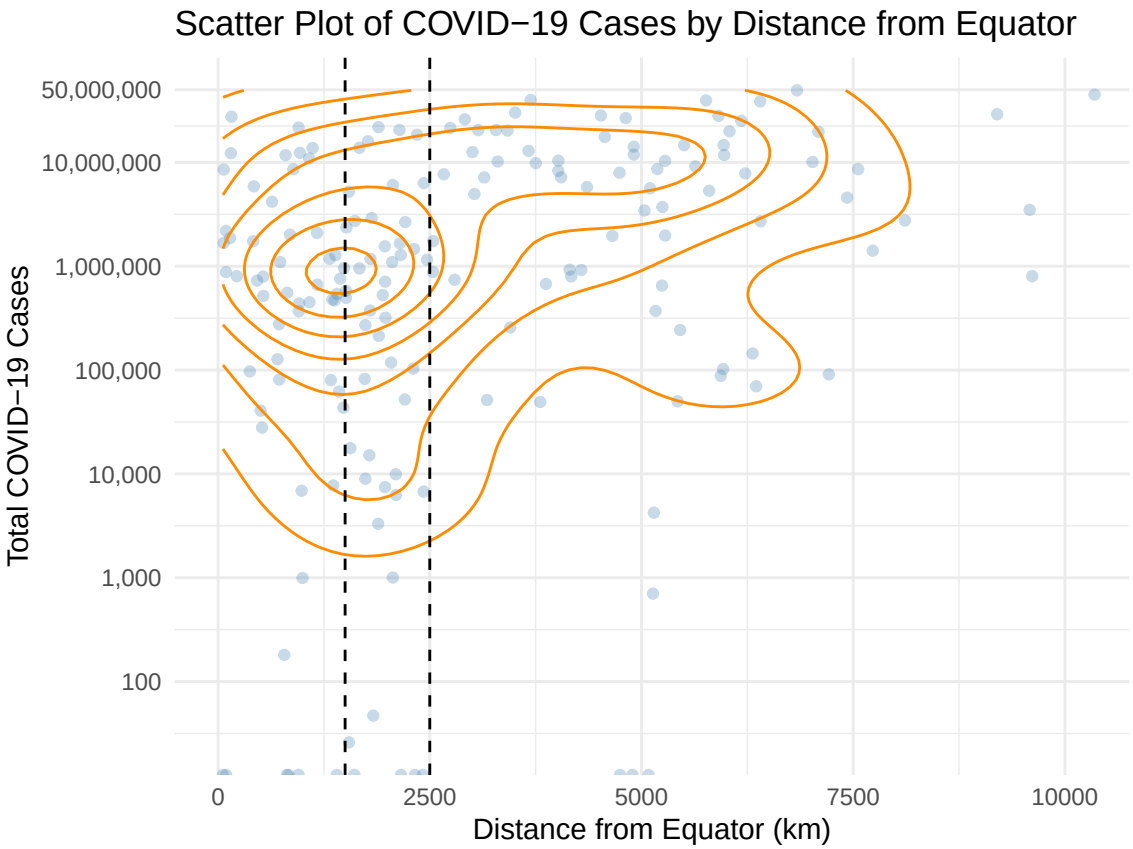
We explored many aspects of the data, but will demonstrate three. These are «Social Support VS Life Expectancy», «Perception of Corruption VS Covid Cases», and «Freedom to Make Life Choices & Climate Awareness (OR DISTANCE VS COVID)»

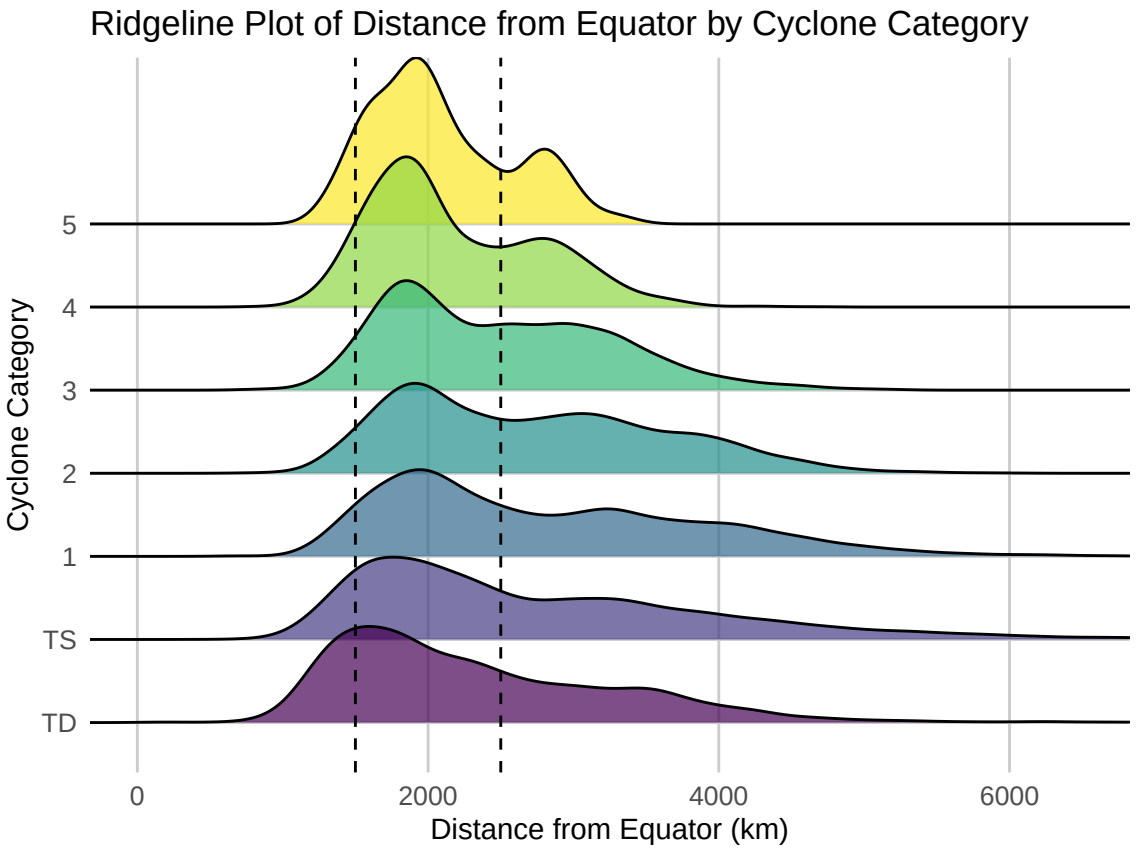
The analysis highlighted an interesting finding, as shown in @ref(fig): a significant concentration of COVID-19 cases appears within the 1500-2500 km range from the equator, suggesting a strong relationship between COVID-19 cases and cyclone activity within a margin of approximately 9 degrees of longitude.

```
## Warning in scale_y_continuous(labels = scales::comma, trans = "log10", breaks = c(1, : log-10 transfo
## log-10 transformation introduced infinite values.
```

```
## Warning: Removed 41 rows containing non-finite outside the scale range (`stat_density2d()`).
```

```
## Warning: Removed 28 rows containing missing values or values outside the scale range (`geom_point()`)
```


Scatter Plot of COVID−19 Cases by Distance from Equator
Total COVID−19 Cases
50,000,000
10,000,000
1,000,000
100,000
10,000
1,000
100
0
2500
5000
7500
10000
Distance from Equator (km)



A report conducted by the National Librarbty of Medicine 1 highlights a strong correlation between COVID-19 cases and proximity to the equator, suggesting that certain geographic and environmental factors may influence virus spread and outbreak. This observation signifies a trend where regions near the equator are subject to unique vulnerabilities, not only experiencing increased COVID-19 cases but also facing intensified cyclone activity (as shown in the figure above), volcanic activity, elevated ocean currents, seasonal flooding, and the presence of natural carbon sinks.

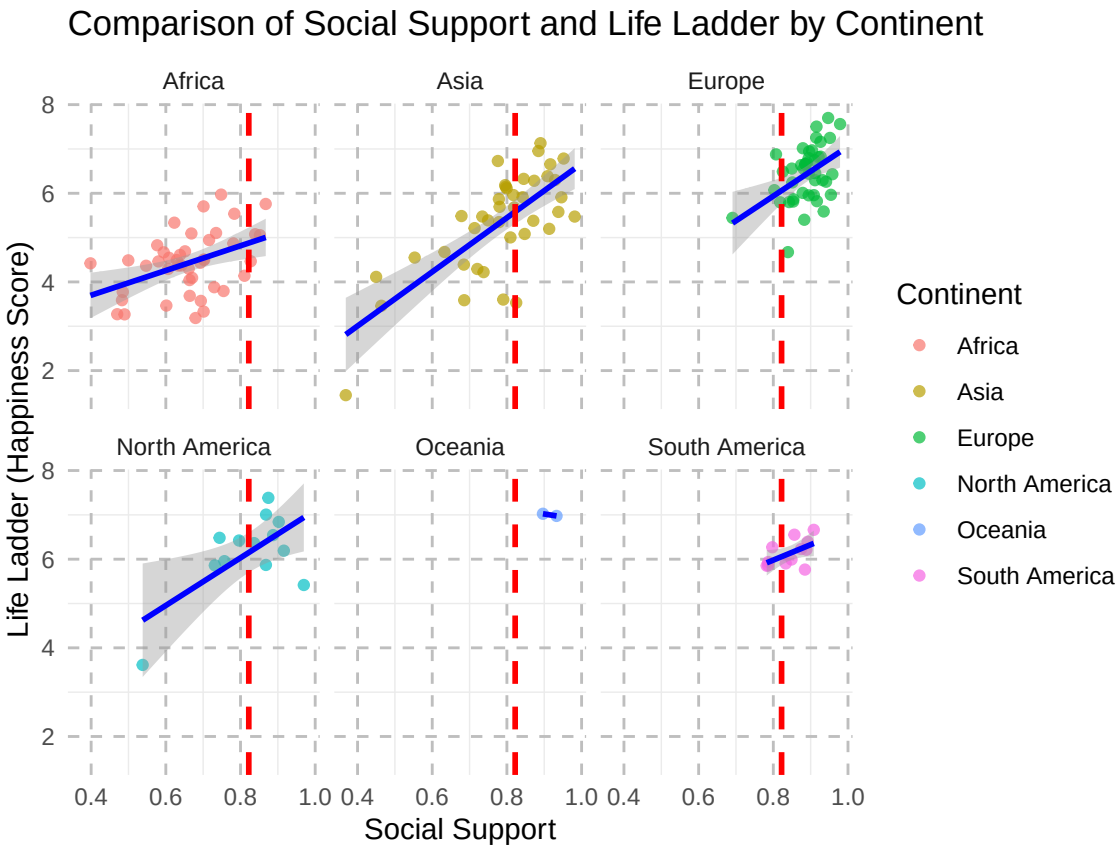
Applying this insight into the analysis, it is evident that there are strong correlations to geological crises such as cyclones and COVID-19, and key indicators like GDP, life satisfaction, social support, and the overall well-being of countries.These crises profoundly impact not only economic stability but also the social and health of a nation, as evidenced by happiness levels in Figure 5 of the appendix. This reinforces the insight that crisis have a strong effect on the affected regions economic and social dynamics.

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

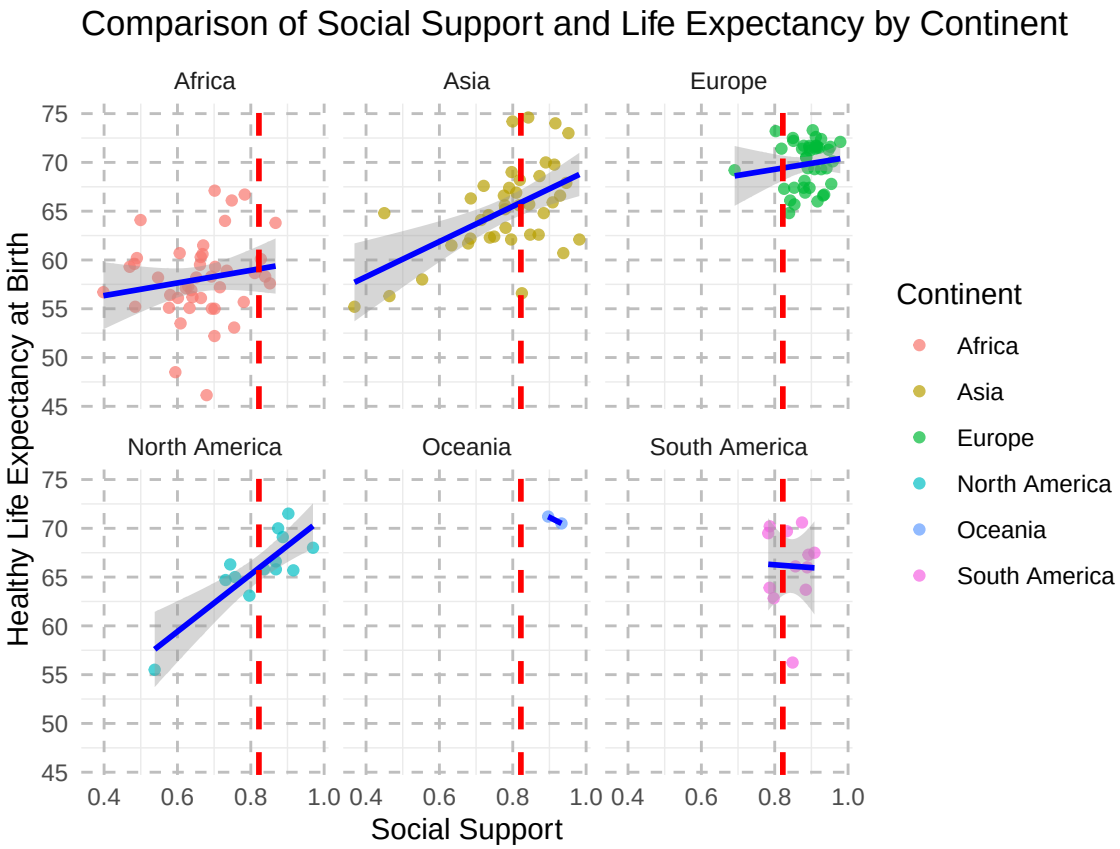
```
## Warning in qt((1 - level)/2, df): NaNs produced
```

```
## Warning: Removed 2 rows containing missing values or values outside the scale range (‘geom_point()‘)
```

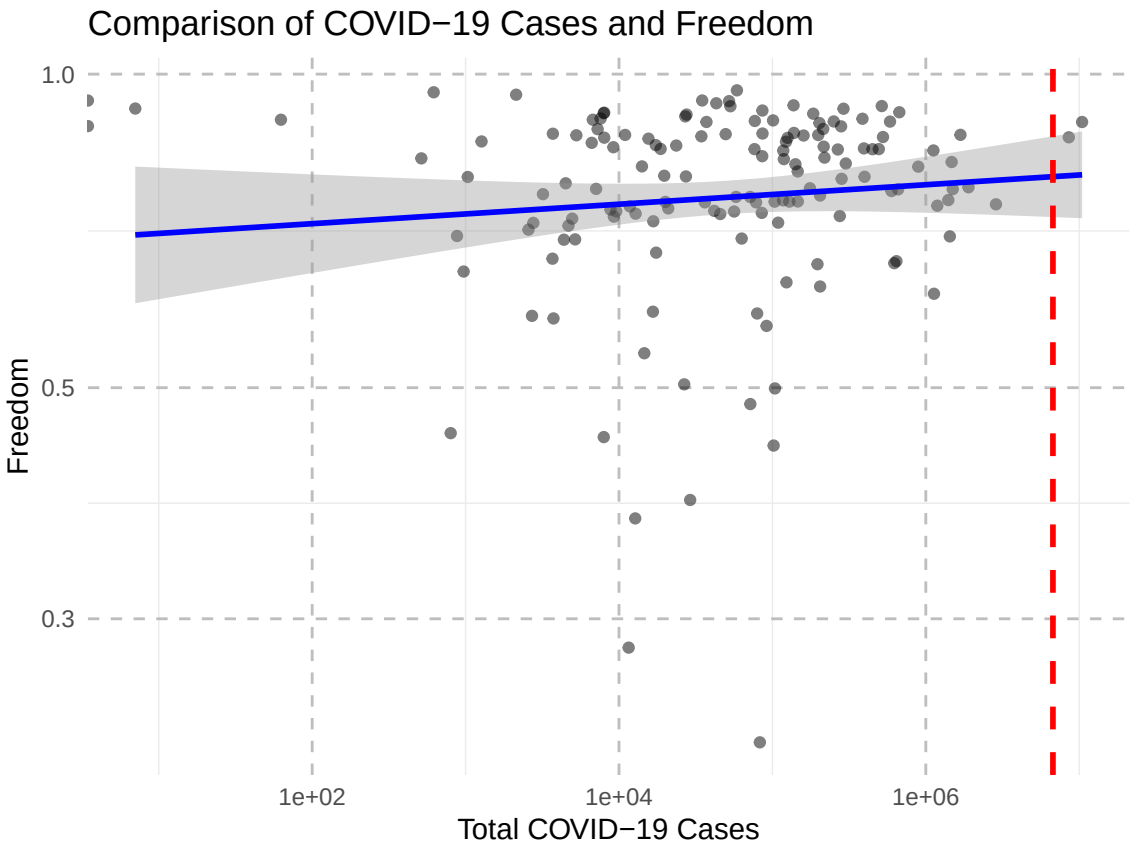
```
## Warning in max(ids, na.rm = TRUE): no non-missing arguments to max; returning -Inf
```

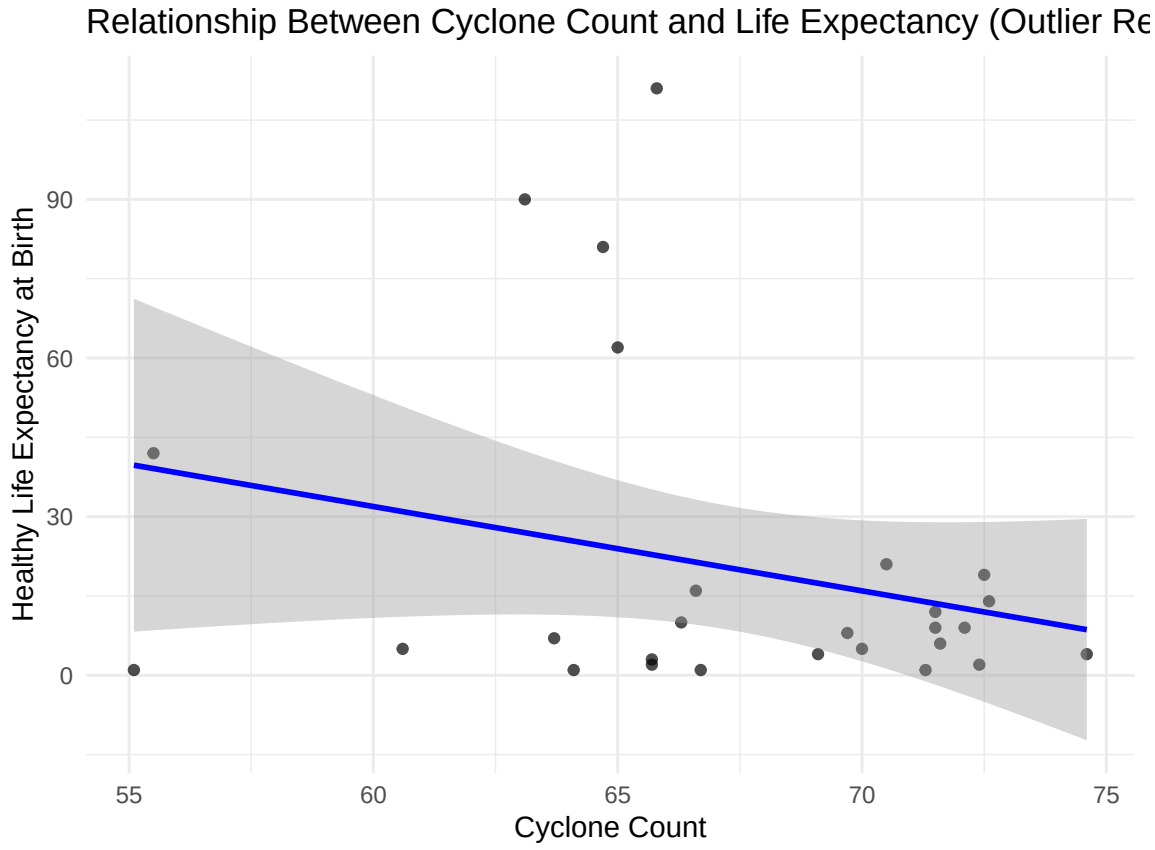



```
## Warning: Removed 3 rows containing non-finite outside the scale range (`stat_smooth()`).

## Warning in qt((1 - level)/2, df): NaNs produced

## Warning: Removed 3 rows containing missing values or values outside the scale range (`geom_point()`)

## Warning in max(ids, na.rm = TRUE): no non-missing arguments to max; returning -Inf
```

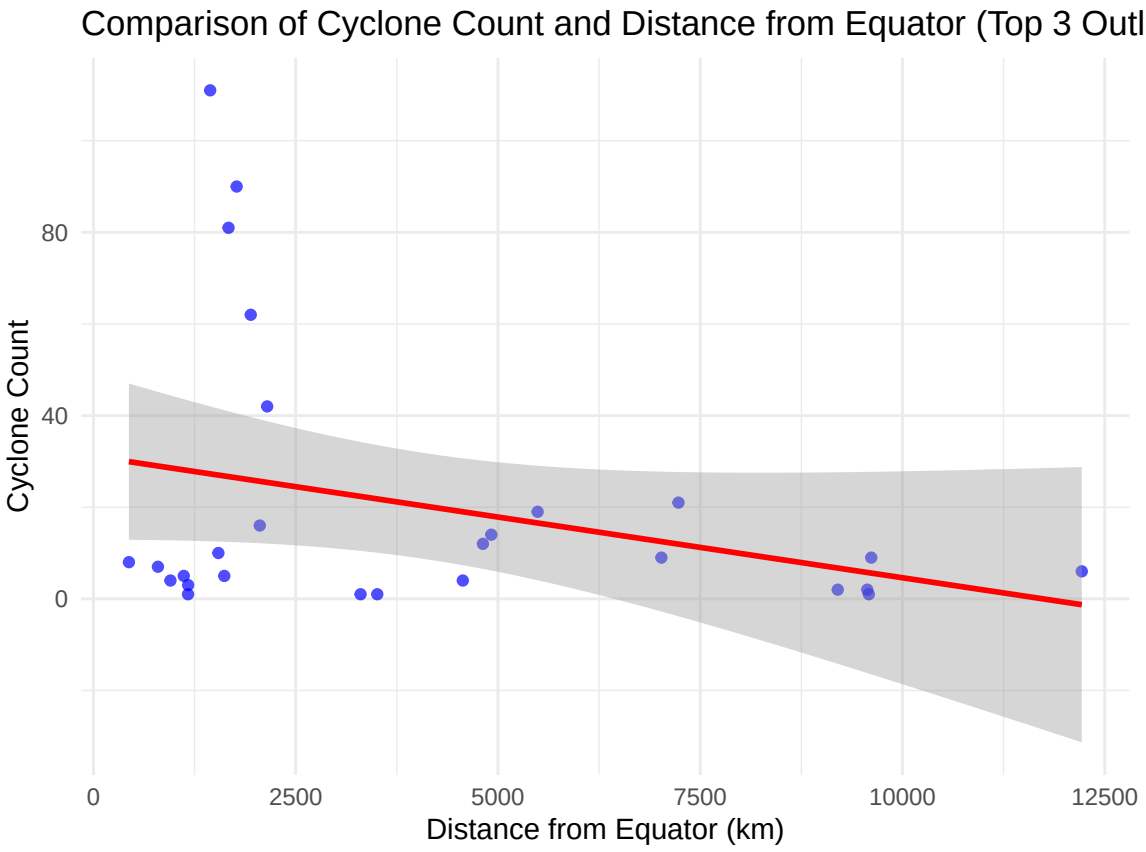



```
## Warning in scale_x_log10(): log-10 transformation introduced infinite values.

## Warning in scale_x_log10(): log-10 transformation introduced infinite values.

## Warning: Removed 9 rows containing non-finite outside the scale range (`stat_smooth()`).

## Warning: Removed 7 rows containing missing values or values outside the scale range (`geom_point()`)
```

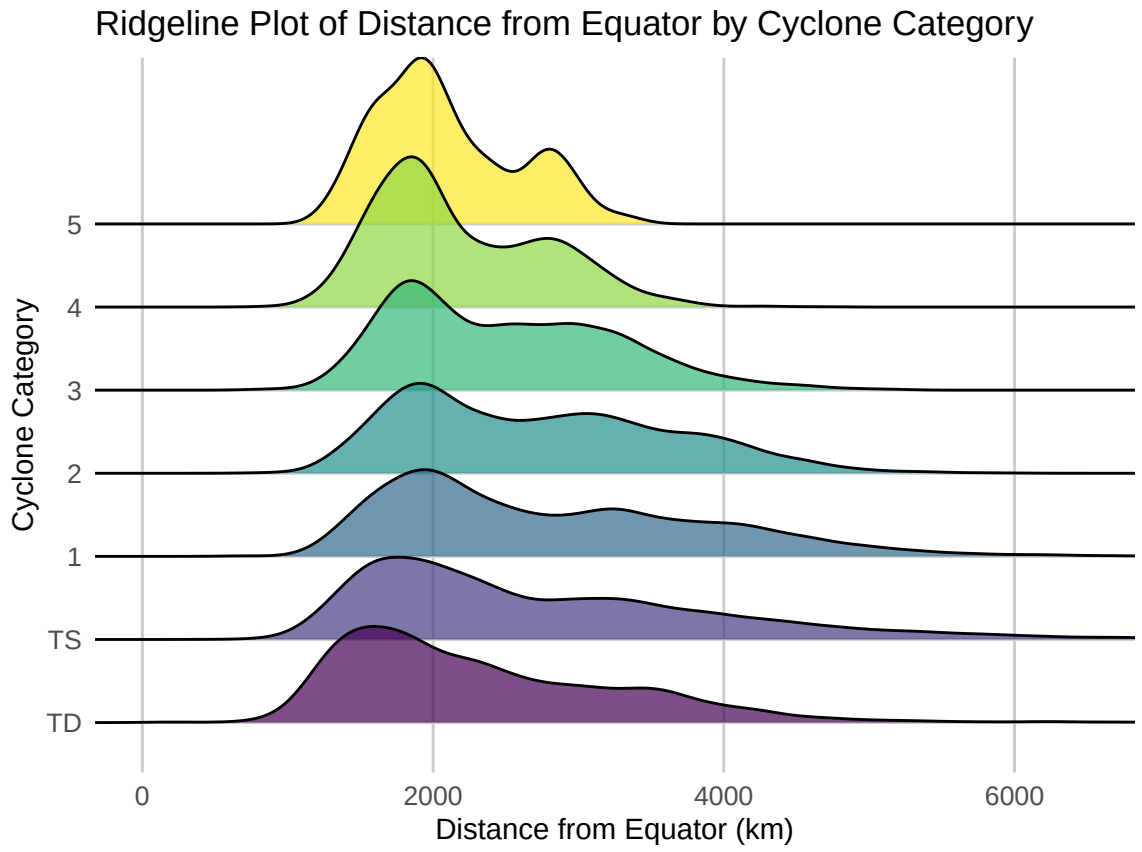

Comparison of COVID−19 Cases and Freedom
1.0
0.5
0.3
Freedom
1e+02
1e+04
1e+06
Total COVID−19 Cases



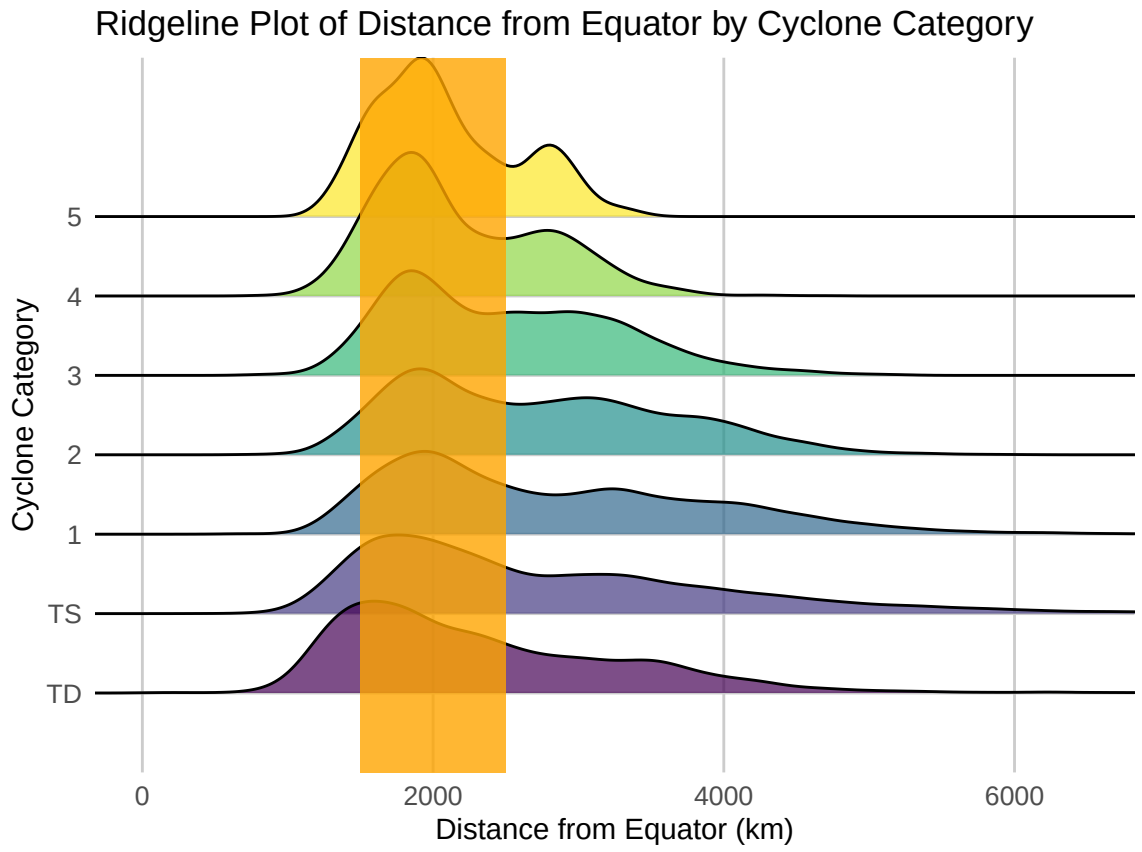

```
##  [1] "country"                    "year"                       "life_ladder"
##  [8] "generosity"                 "perceptions_of_corruption"  "positive_affect"
```



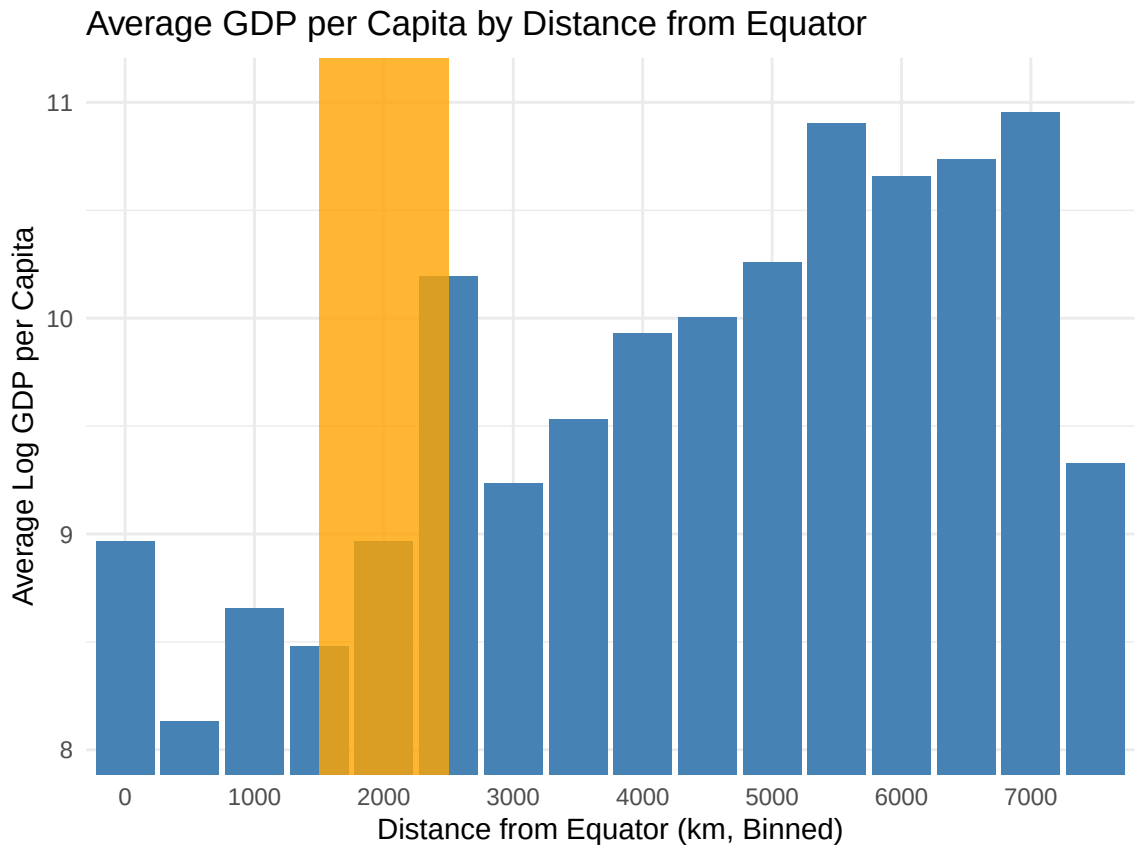

```
## Warning in geom_density_ridges(scale = 2, alpha = 0.7, color = "black", : Ignoring unknown parameters
```

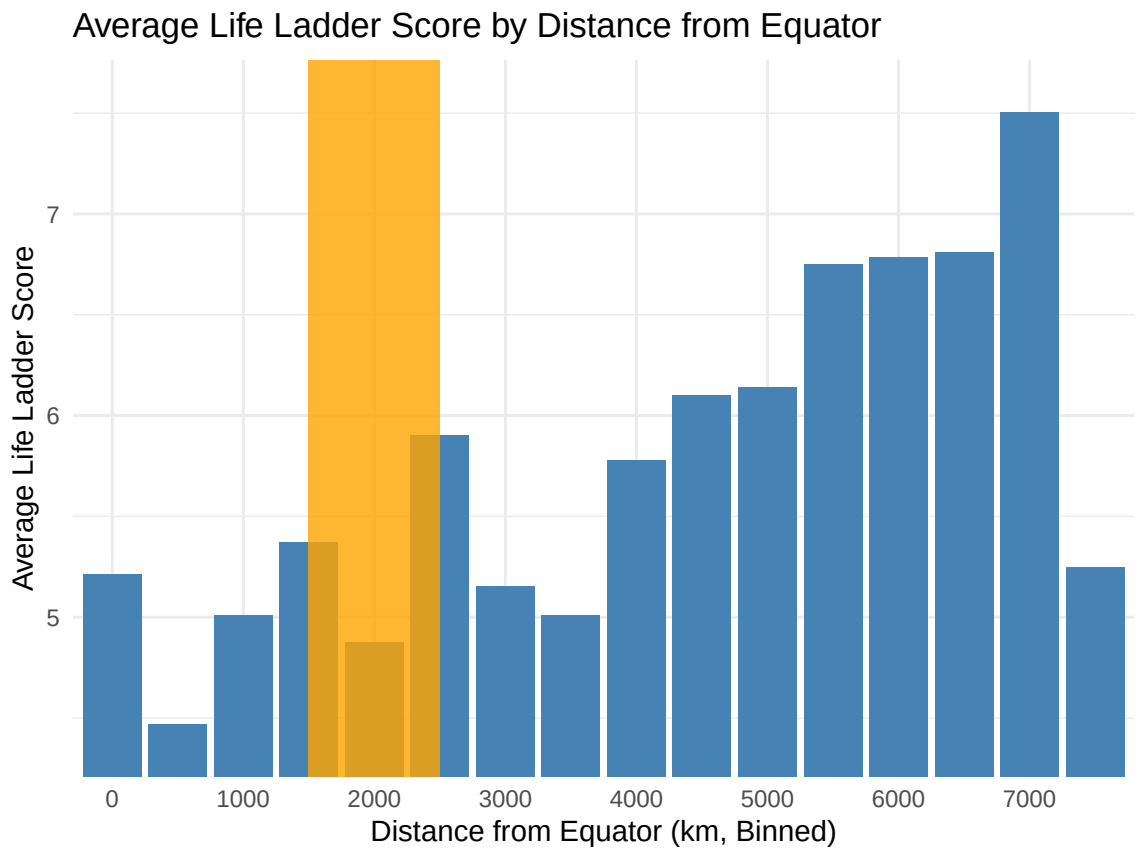



```
##  [1] "country"                    "year"                       "life_ladder"
##  [8] "generosity"                 "perceptions_of_corruption"  "positive_affect"
```




```
## # A tibble: 165 x 16
## # Groups:   country [165]
##    country      year life_ladder log_gdp_per_capita social_support healthy_life_expectancy_at_birth
##    <chr>       <dbl>       <dbl>              <dbl>          <dbl>                             <dbl>
##  1 Afghanistan  2023        1.45              NA             0.368                              55.2
##  2 Albania      2023        5.44               9.69          0.691                              69.2
##  3 Algeria      2022        5.54               9.32          0.783                              66.7
##  4 Angola       2014        3.79               9.01          0.755                              53.1
##  5 Argentina    2023        6.39               9.99          0.892                              67.3
##  6 Armenia      2023        5.68               9.73          0.819                              68.2
##  7 Australia    2023        7.02              10.8           0.896                              71.2
##  8 Austria      2023        6.64              10.9           0.874                              71.4
##  9 Azerbaijan   2023        5.21               9.64          0.713                              64.1
## 10 Bahrain      2023        5.96              10.9           0.817                              65.6
## # i 155 more rows
## # i abbreviated names: 1: freedom_to_make_life_choices, 2: perceptions_of_corruption, 3: distance_to_
## # i 2 more variables: distance_to_equator_km.y <dbl>, geometry.y <MULTIPOLYGON [°]>
```


Average GDP per Capita by Distance from Equator
Average Log GDP per Capita
11
10
9
8
0
1000
2000
3000
4000
5000
6000
7000
Distance from Equator (km, Binned)



This insight is supported by the summary statistics in table @ref(tab:summary_stats)

«««< HEAD

The next insight that we found is shown in @ref(fig:insight2).

Finally, @ref(fig:insight3) shows ...

Conclusion and Future Work

Overall, we found <>.

A second paragraph about our findings.

The next steps in this analysis are...

The limitations of this analysis are as follows. (Do not simply list potential issues with sampling, but relate them to your analysis and how they affect your conclusions. An honest and complete acknowledgement of the limitations makes the analysis more trustworthy.)

References

I am not strict about MLA or APA style or anything like that. For this report, I would much rather have your citations be easy to match to your insights.

The easiest way is to use Rmd's footnote syntax. This will put a number beside the word where the footnote appears, and the full text of the footnote at the bottom of the page (pdf) or end of the document (html). The syntax is:[1], where I suggest that you put in something like this[2] to make references for this assignment.

Alternatively, you could make a list of citations with their main arguments and why they're relevent to your insights, methods, etc.

The link above also references "bibtex" files. These are also extremely convenient, but have a steep learning curve and they make it difficult to tie them to an insight. If you use bibtext, then make sure that you provide a sentence to describe the source and it's relevance when you cite it - don't just add citations to the end of a sentence (this is common practice in academia, but I want to know that your citations are directly relevant for this assignmnet).

Apendix

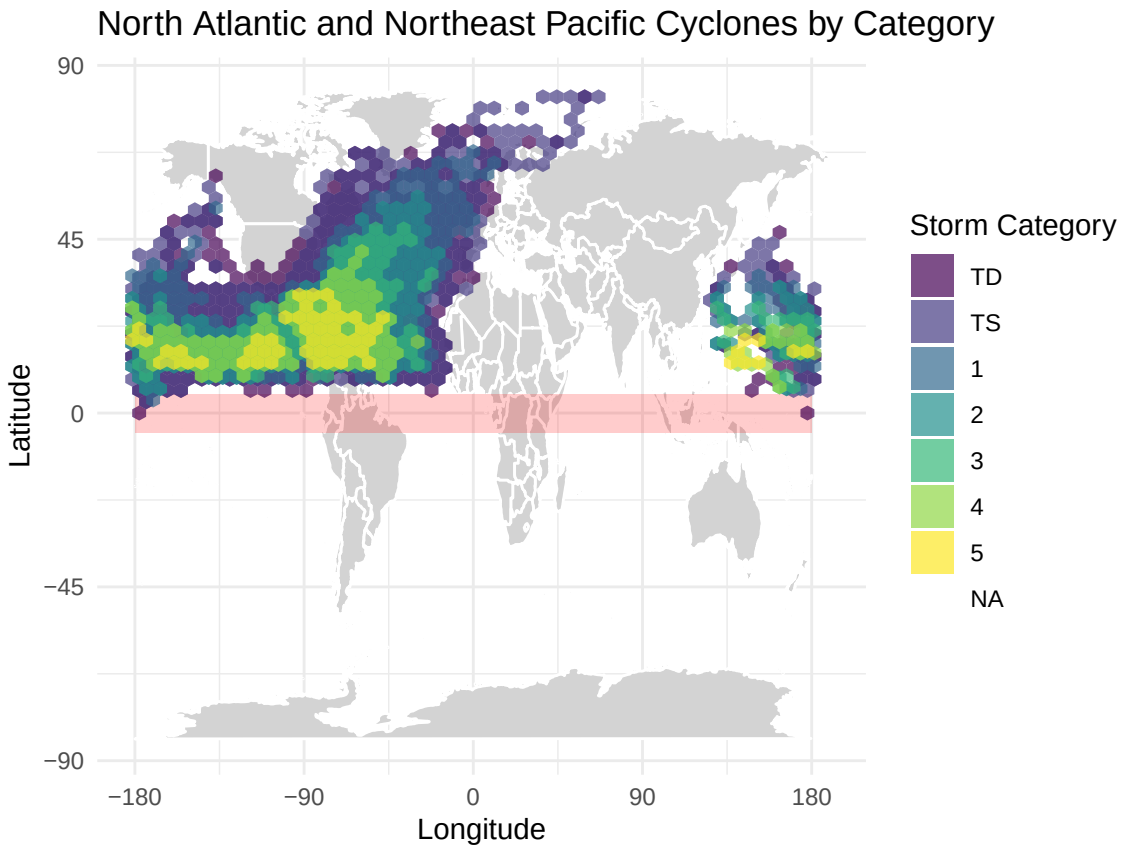


Figure 1: North Atlantic and Northeast Pacific Cyclones by Category

[1] See the source view to see this footnote
[2] The relevance to the insight is . . . . From <>, published on <>, url: <>

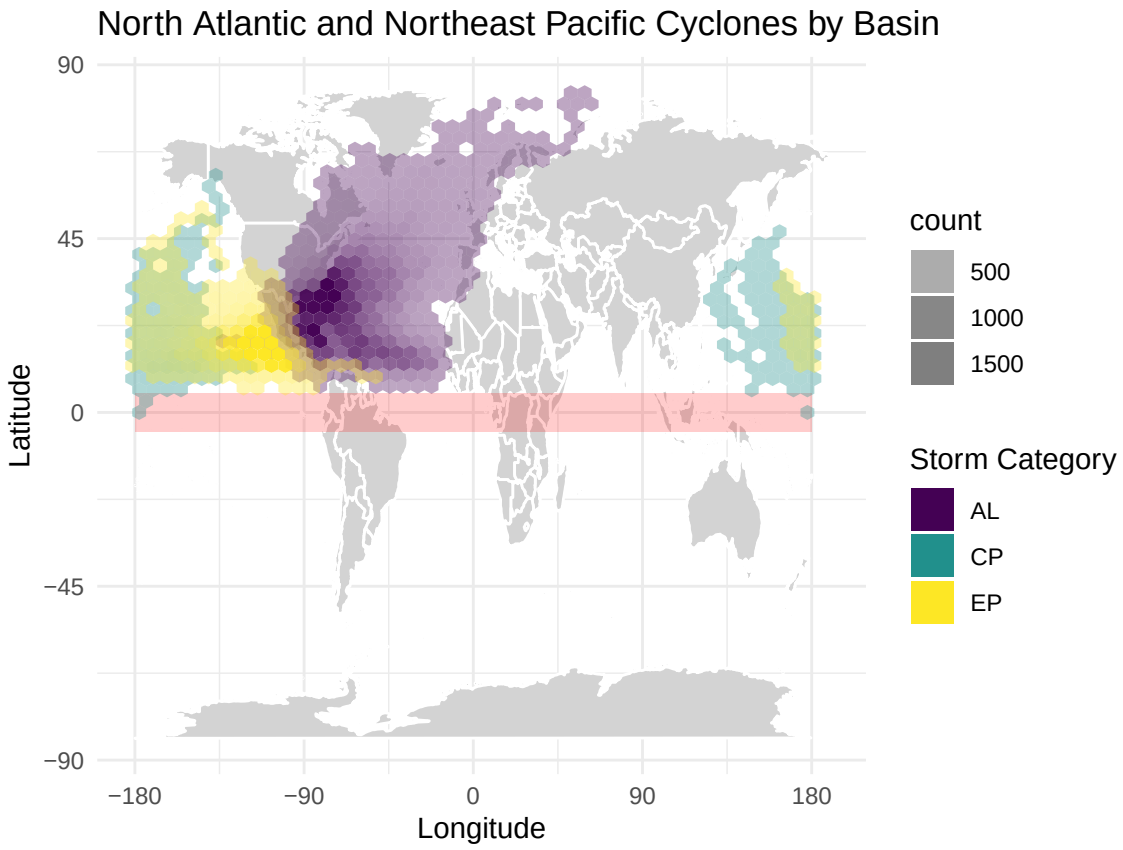


Figure 2: North Atlantic and Northeast Pacific Cyclones by Basin

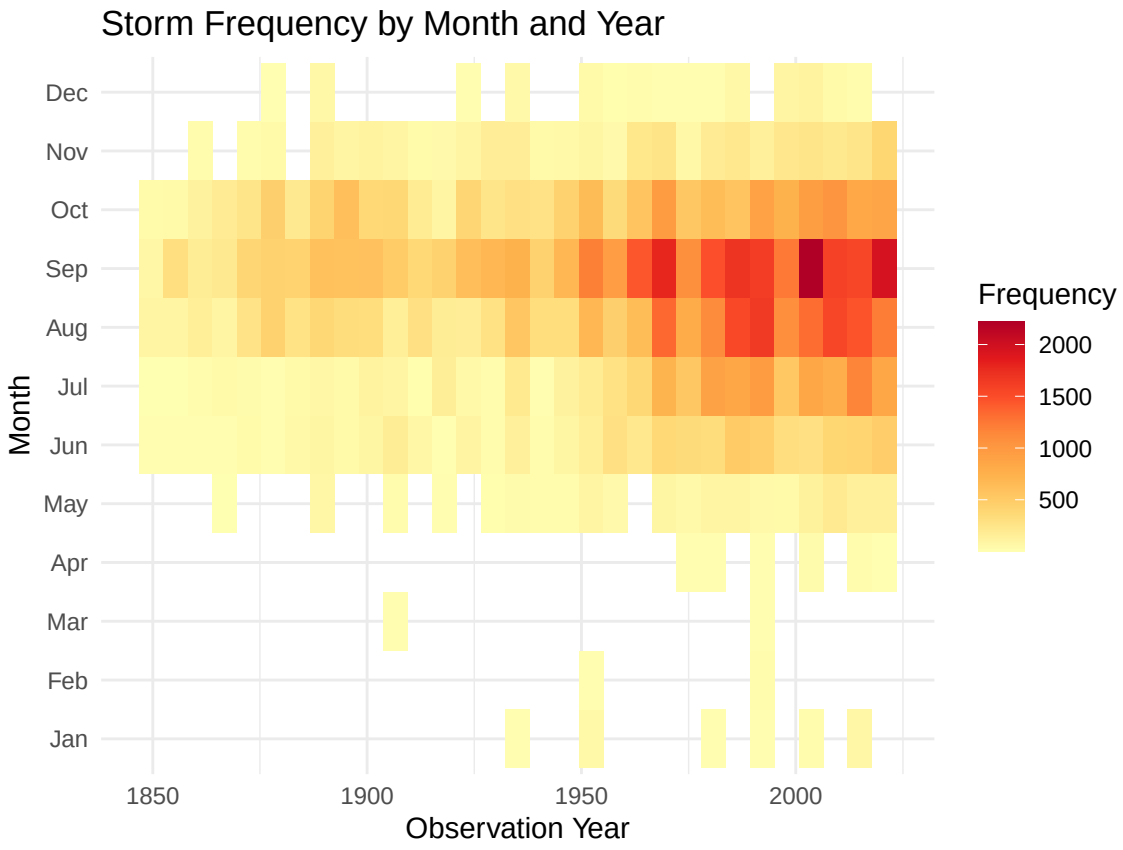


Figure 3: Storm Frequency by Month and Year

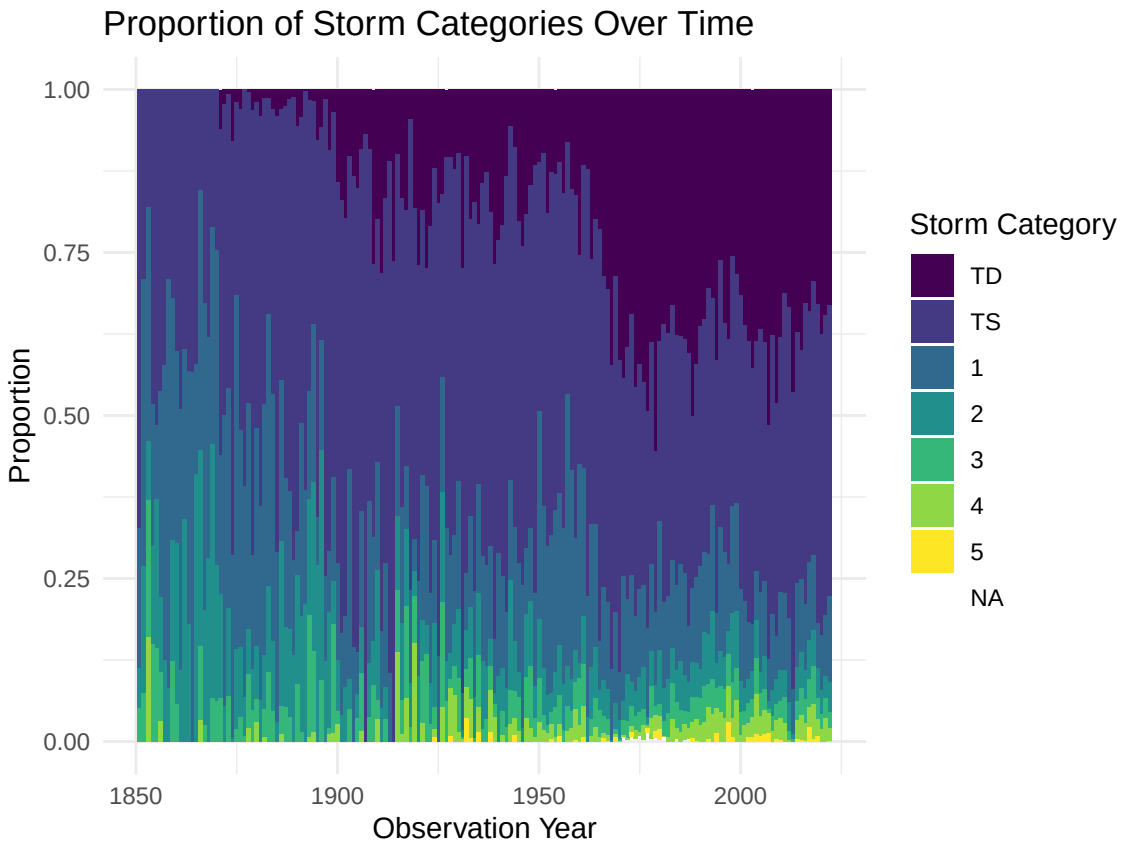


Figure 4: Proportion of Storm Categories Over Time

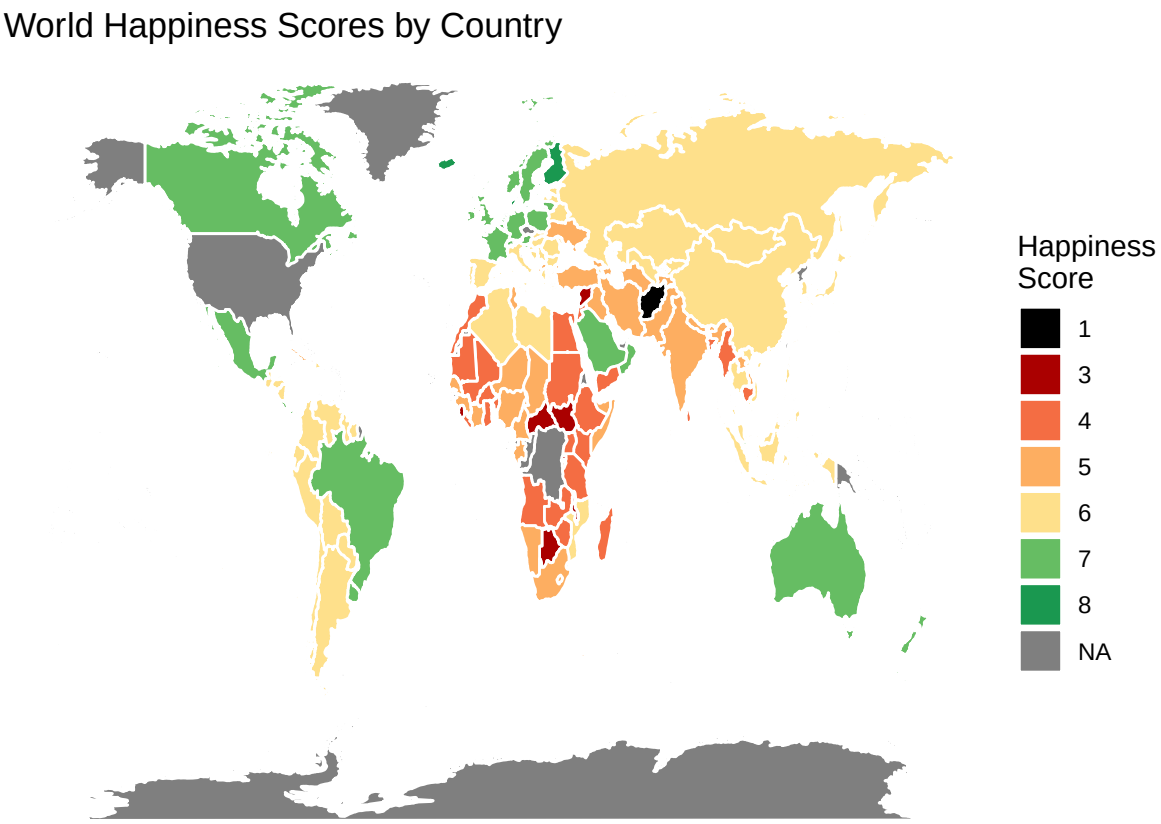


Figure 5: World Happiness Scores by Country