TITLE OF YOUR PROJECT

Group 8

November 12th, 2024

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

««««< HEAD - John K. Samson (#########) - John Darnielle (#########) - Craig Finn (#########) - Joel Plaskett (########) - Ezra Furman (#########) ======= - Aditya Chauhan (169027493) - John Darnielle (#########) - Craig Finn (#########) - Joel Plaskett (########) - 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

General context, very brief data descriptions, techniques used, and general conclusions, all contained within a single, concise paragraph.

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

«Data Set 1»

```
## # A tibble: 84,601 x 33
##    Basin Number NameYear Name    ObservYear Month MonthName   Day  Hour Minute Identifier status lati
##    <chr>  <int>    <int> <chr>        <int> <int> <fct>     <int> <int>  <int> <chr>      <chr>  <chr
##  1 AL         1     1851 UNNAMED       1851     6 Jun          25     0      0 <NA>       HU     28.0
##  2 AL         1     1851 UNNAMED       1851     6 Jun          25     6      0 <NA>       HU     28.0
##  3 AL         1     1851 UNNAMED       1851     6 Jun          25    12      0 <NA>       HU     28.0
##  4 AL         1     1851 UNNAMED       1851     6 Jun          25    18      0 <NA>       HU     28.1
##  5 AL         1     1851 UNNAMED       1851     6 Jun          25    21      0 L          HU     28.1
##  6 AL         1     1851 UNNAMED       1851     6 Jun          26     0      0 <NA>       HU     28.1
##  7 AL         1     1851 UNNAMED       1851     6 Jun          26     6      0 <NA>       TS     28.3
##  8 AL         1     1851 UNNAMED       1851     6 Jun          26    12      0 <NA>       TS     28.4
##  9 AL         1     1851 UNNAMED       1851     6 Jun          26    18      0 <NA>       TS     28.0
## 10 AL         1     1851 UNNAMED       1851     6 Jun          27     0      0 <NA>       TS     29.0
## # i 84,591 more rows
## # i 17 more variables: NE_extend_34 <dbl>, SE_extend_34 <dbl>, SW_extend_34 <dbl>, NW_extend_34 <dbl
## #   SE_extend_50 <dbl>, SW_extend_50 <dbl>, NW_extend_50 <dbl>, NE_extend_64 <dbl>, SE_extend_64 <dbl
## #   NW_extend_64 <dbl>, r_max_wind <dbl>, lat <dbl>, lon <dbl>, date <dttm>, category <ord>
```

The data come from <> and describe <>.

In order to clean the data, we «steps to clean the data, concise but precise enough that a reader could follow your steps without seeing your code».

«Data Set 2»

```
## # A tibble: 184 x 4
##     year region    name  value
##    <int> <chr>     <chr> <dbl>
##  1  1978 Antarctic min    7.28
##  2  1978 Antarctic max   17.8
##  3  1978 Arctic    min   10.2
##  4  1978 Arctic    max   14.6
##  5  1979 Antarctic min    2.91
##  6  1979 Antarctic max   18.4
##  7  1979 Arctic    min    6.90
##  8  1979 Arctic    max   16.6
##  9  1980 Antarctic min    2.52
## 10  1980 Antarctic max   19.1
## # i 174 more rows
```

The data come from <> and detail <>.

In order to clean the data, we «steps to clean the data, concise but precise enough that a reader could follow your steps without seeing your code»

«Data Set 3»

```
## # A tibble: 110 x 7
##    country                 aware_no aware_alittle aware_moderate aware_alot aware_refuse aware_base
##    <chr>                      <dbl>         <dbl>          <dbl>      <dbl>        <dbl>      <dbl>
##  1 Albania                     9.99          43.0           33.1      10.6         3.25         329
##  2 Algeria                    23.2           37.6           26.0      11.1         2.15        1066
##  3 Angola                     19.1           43.0           23.5      11.1         3.29         727
##  4 Argentina                   8.25          37.4           43.8       9.88        0.696       1257
##  5 Armenia                     9.18          51.4           25.8      11.4         2.14         306
##  6 Asian.&.Pacific.Islands    13.3           40.1           24.8      19.5         2.26        1916
##  7 Australia                   1.13          26.4           48.1      23.9         0.407       1012
##  8 Austria                     1.34          16.4           48.6      32.7         0.985       1120
##  9 Azerbaijan                  6.19          42.4           35.9      13.2         2.20         541
## 10 Bangladesh                 24.5           36.2           16.2      18.9         4.22         910
## # i 100 more rows
```

The data come from <> and detail <>.

In order to clean the data, we «steps to clean the data, concise but precise enough that a reader could follow your steps without seeing your code»

«Data Set 4»

```
## # A tibble: 252 x 3
## # Groups:   country [252]
##    country             continent     total_cases
##    <chr>               <chr>               <dbl>
##  1 Afghanistan         Asia              8284992
##  2 Africa              <NA>            285057531
##  3 Albania             Europe            3444855
##  4 Algeria             Africa           10188569
##  5 American Samoa      Oceania                 0
##  6 Andorra             Europe             650844
##  7 Angola              Africa            1275899
##  8 Anguilla            North America        1005
##  9 Antigua and Barbuda North America       22209
## 10 Argentina           South America           NA
## # i 242 more rows
```

«Data Set 5»

```
## # A tibble: 165 x 11
## # Groups:   country [165]
##    country  year life_ladder log_gdp_per_capita social_support healthy_life_expecta~1 freedom_to_mak
##    <chr>   <dbl>       <dbl>              <dbl>          <dbl>                  <dbl>
##  1 Afghan~  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 Argent~  2023        6.39               9.99          0.892                   67.3
##  6 Armenia  2023        5.68               9.73          0.819                   68.2
##  7 Austra~  2023        7.02              10.8           0.896                   71.2
```

```
##  8 Austria  2023        6.64              10.9          0.874                  71.4
##  9 Azerba~  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: healthy_life_expectancy_at_birth, 2: freedom_to_make_life_choices, 3: perc
## # i 2 more variables: positive_affect <dbl>, negative_affect <dbl>
```

Combining the Data

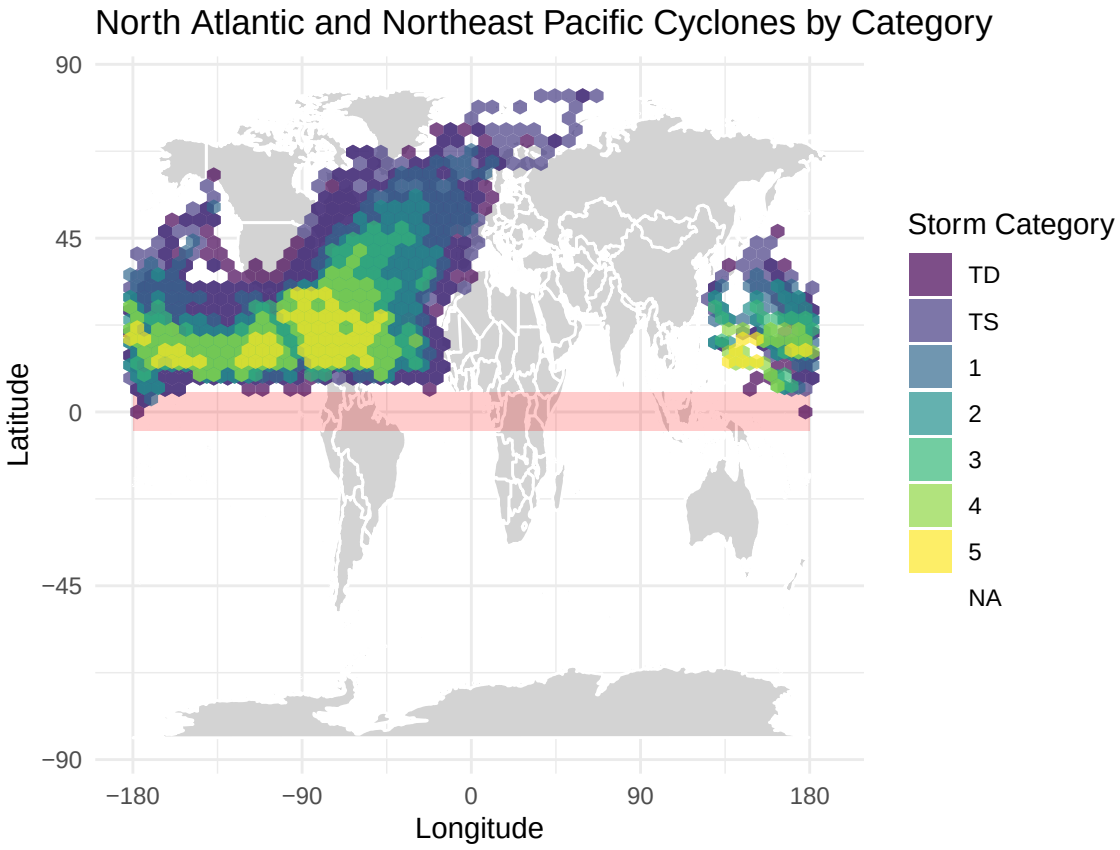
Explain how any combinations of data were performed. Explain what kind of join was needed, whether columns had to be modified (for example, matching "country" names.)

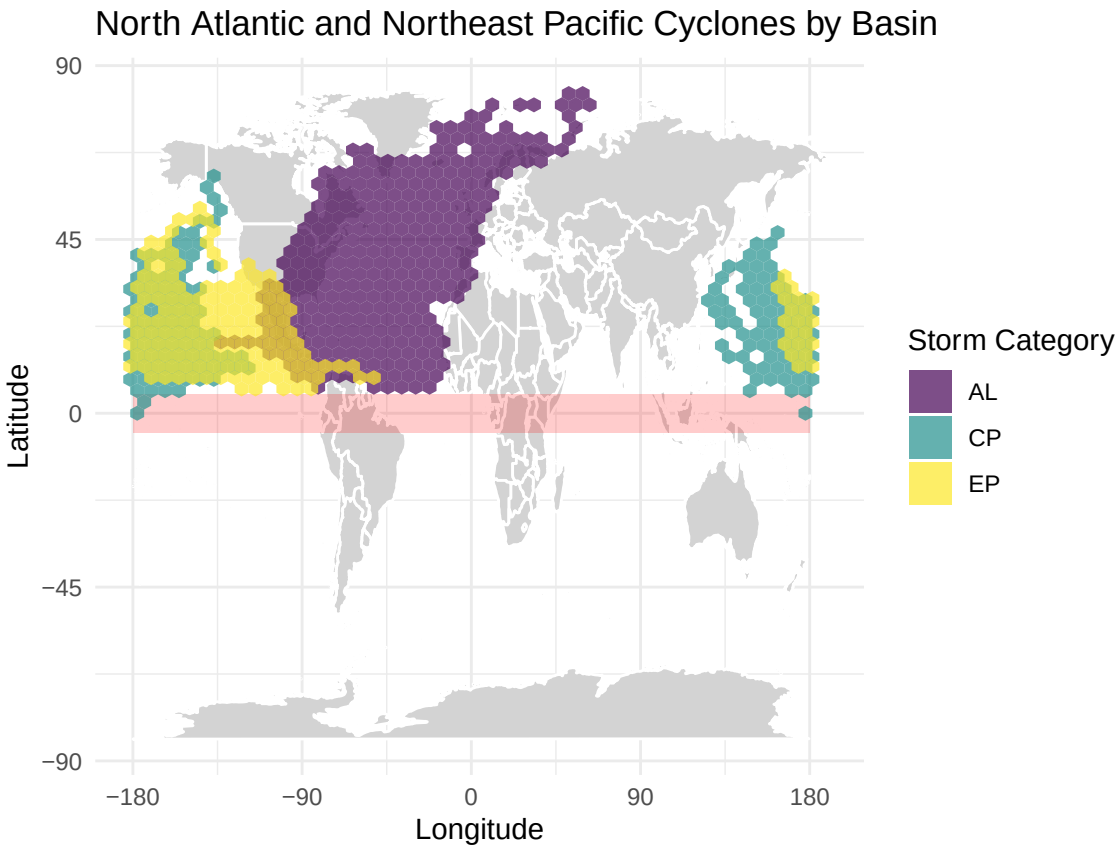
Exploratory Data Analysis

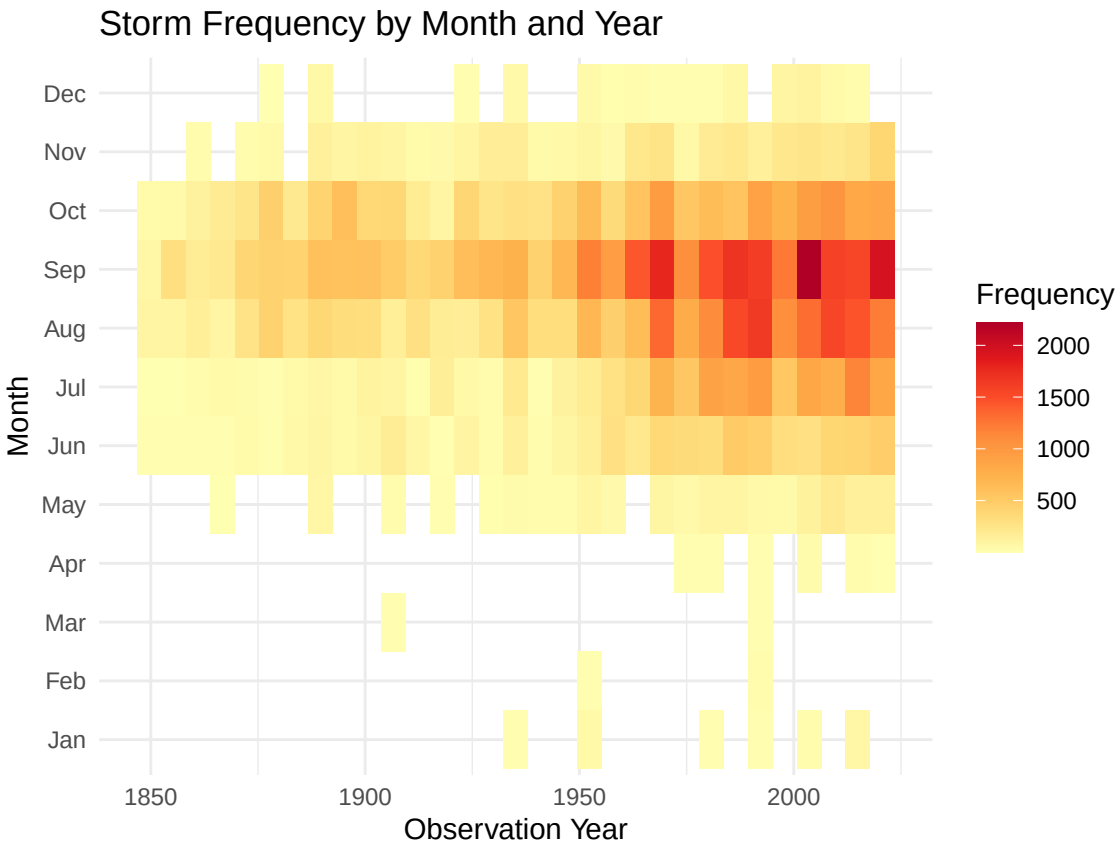
To achieve our goals, we explored the data by...

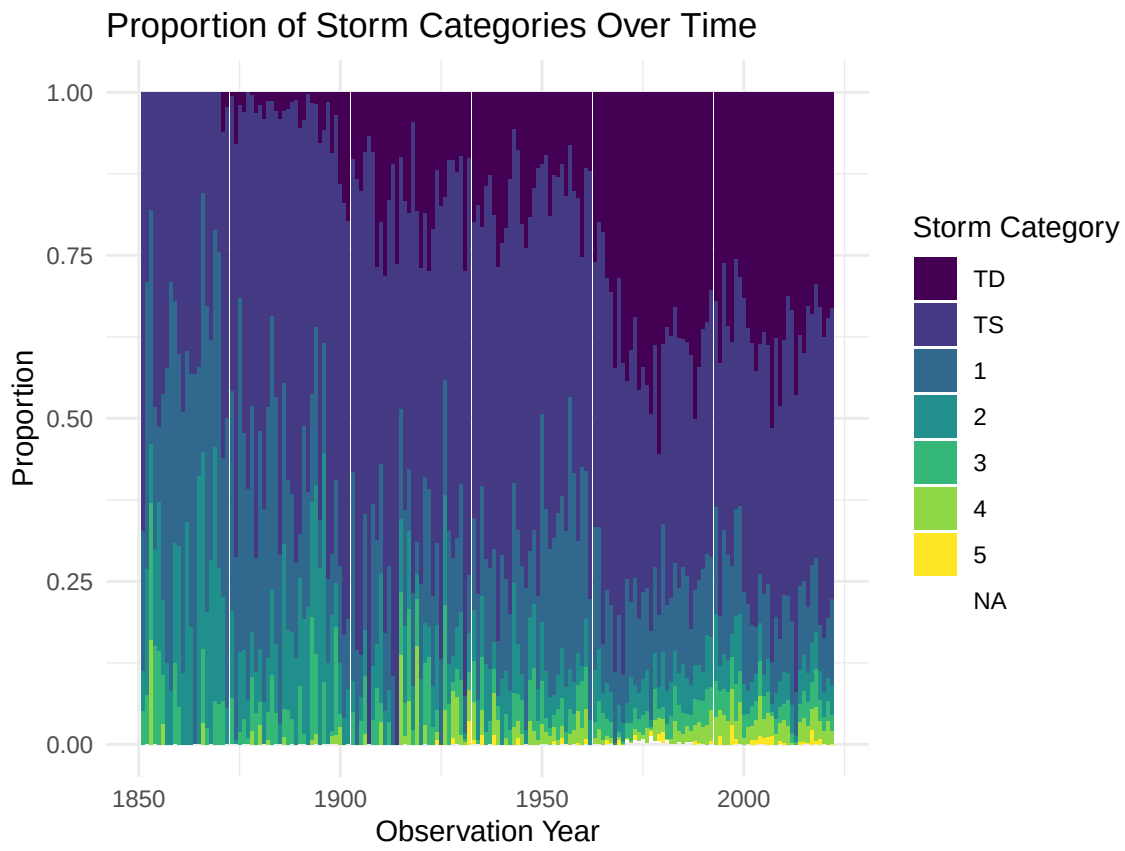
We explored many aspects of the data, but will demonstrate three. These are «insight 1», «insight 2», and <>

The first aspect that we found interesting is shown in @ref(fig:insight1). The insight should be specific to the data shown, not a general statement beyond the data (leave that for the conclusion).




North Atlantic and Northeast Pacific Cyclones by Basin
90
45
0
−45
−90
Latitude
−180
−90
0
90
180
Longitude
Storm Category
AL
CP
EP


Storm Frequency by Month and Year
Dec
Nov
Oct
Sep
Aug
Jul
Jun
May
Apr
Mar
Feb
Jan
Month
1850
1900
1950
2000
Observation Year
Frequency
2000
1500
1000
500



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

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).

[1] See the source view to see this footnote

[2] The relevance to the insight is ... . From <>, published on <>, url: <>