# Natural Disaster Sentiment Analysis

**Scott Burstein**

**23 November, 2020**

# Setup

# Configure RStudio Rmd File Output Format

# Load the Libraries and Datasets

## Google Trends Package

gtrendsR CRAN Package (https://cran.r-project.org/web/packages/gtrendsR/gtrendsR.pdf) used to perform and display Google Trends Queries.

Google Trends data is accessible in R through the gTrendsR package, which was created by Philippe Masicotte (https://github.com/PMassicotte) and made available through CRAN (The Comprehensive R Archive Network) (https://cran.r-project.org/).

Google trends data for specified search queries is represented by a `hits` variable, which expresses the relative volume of Google searches for specific terms over geographical and time parameters.

`hits` - a numeric integer between 0 and 100 representing Google's weekly search volume as a proportion of the maximum search volume for the specified keyword within the given time and location bounds. It is worth noting that every Google trends query will *always* have at least one interest_over_time `hits` value of 100 for a given query.

## README

See README for more thorough description of the `FEMA_Declarations` data set, with content summary, column variable descriptions, acknowledgments, source description and licensing information.

## Primary Data Source

Kaggle Natural Disaster Data set (https://www.kaggle.com/headsortails/us-natural-disaster-declarations). The `FEMA_Declarations` data set is used for the bulk of this analysis and is also integrated with Google Trends data.

For later analysis, the `small_FEMA` data frame can be used to parse FEMA declarations that have occurred over the last 5 years in the states that have, in those 5 years, recorded the most FEMA declarations. It is created after `recent_FEMA_byState`.

# Filter for Event Types

## Generalizable Parameters for Event Filter

In order to produce research that is both reproducible and timely, certain parameters can be initialized once, and then referenced multiple times later. Since gTrendsR can easily return Google Trends information from the past five years, it is intuitive to create a `date_5y` variable that can be used to filter FEMA declarations by date, thus eliminating all those that occurred previously. The data set at large contains FEMA declarations from 1953 through the present. Throughout this research there will also be many variables that are reused. In portions of the analysis that copy methods used elsewhere, the assignment operator is used to allow for minimal rewriting in the event that certain functions are purposed for further analysis in the future.

---

# Research Questions Preview

**(1)** Do Google search patterns correlate with real-world events?

**(2)** Does interest in climate change increase regionally after the occurrence of a natural disaster?

**(3)** Do people modify their Google search behavior and become more environmentally conscious after a natural disaster?*

---

# Initial Analysis of FEMA Declarations

## All FEMA Declarations by State

`alltime_FEMA_byState` contains the number of FEMA declarations for every US state and territory since 1953.

```
## # A tibble: 59 x 2
## # Groups:   state [59]
##    state     n
##    <chr> <int>
##  1 CA      294
##  2 TX      280
##  3 OK      164
##  4 WA      156
##  5 FL      145
##  6 OR      119
##  7 NY      100
##  8 AZ       96
##  9 LA       91
## 10 NM       91
## # … with 49 more rows
```

Since the database was created in 1953, the following 10 states have had the most FEMA natural disaster emergency declarations called (as of 2020-11-23):

CA, TX, OK, WA, FL, OR, NY, AZ, LA, NM

# Recent FEMA Declarations by State

The date this file was last compiled: Current date ( `Sys.Date` ) is 2020-11-23.

The date 5 years before, which is the starting time bound for subsequent Google Trends queries shown below: Variable `date_5y` is 2015-11-23.

`recent_FEMA_byState` contains the number of FEMA declarations for every US state and territory from the last 5 years.

```
## # A tibble: 56 x 2
## # Groups:   state [56]
##    state     n
##    <chr> <int>
##  1 CA       75
##  2 WA       30
##  3 OR       29
##  4 OK       24
##  5 FL       23
##  6 TX       21
##  7 NV       20
##  8 AZ       19
##  9 LA       19
## 10 MS       19
## # … with 46 more rows
```

From 2015-11-23 to 2020-11-23, the following 10 states have had the most FEMA natural disaster emergency declarations called:

CA, WA, OR, OK, FL, TX, NV, AZ, LA, MS

In the past 5 years, California has recorded 75 FEMA declarations, outpacing all other states. Much of these can be attributed to the recent spell of wildfires which cause massive amounts of damage to property and endanger millions of Americans each year.

Florida has issued 23 FEMA declarations in the last 5 years, many of which are the result of tropical storms from the Atlantic that grow into hurricanes. These also cause billions of dollars in damage, widespread power outages, endanger the lives of local residents and also threaten to submerge the already eroding eastern coastline of Florida.

---

# Key Google Trends Summary Insights:

## Climate Change Google Searches by State:

A data frame arranged in descending order of all 50 states and Washington D.C., showing the relative hits count for the search term “climate change” over the last 5 years.

```
##                location hits
## 1               Vermont  100
## 2  District of Columbia   92
## 3                 Maine   73
## 4                Alaska   61
## 5                Oregon   57
## 6                Hawaii   57
## 7         New Hampshire   56
## 8         Massachusetts   56
## 9          Rhode Island   54
## 10             Colorado   53
## 11              Montana   52
## 12          Connecticut   51
## 13             Maryland   50
## 14                 Iowa   50
## 15           Washington   49
## 16           California   45
## 17            Minnesota   44
## 18             New York   44
## 19            Wisconsin   41
## 20             Virginia   41
## 21             Delaware   40
## 22           New Mexico   40
```

```
## 23          South Dakota   39
## 24                 Idaho   39
## 25            New Jersey   38
## 26               Wyoming   38
## 27               Indiana   36
## 28          Pennsylvania   36
## 29                  Utah   36
## 30        North Carolina   35
## 31              Illinois   35
## 32              Michigan   35
## 33               Arizona   34
## 34              Nebraska   33
## 35                Nevada   33
## 36                Kansas   32
## 37              Kentucky   31
## 38         West Virginia   31
## 39          North Dakota   30
## 40                  Ohio   30
## 41              Missouri   30
## 42               Florida   29
## 43              Arkansas   28
## 44              Oklahoma   26
## 45        South Carolina   26
## 46               Georgia   26
## 47             Louisiana   25
## 48                 Texas   25
## 49               Alabama   23
## 50             Tennessee   22
## 51           Mississippi   21
```

Vermont set the mark for most Google search hits - relative search volume for “climate change” over the past five years. Washington D.C. had 0.85 as much search volume for the same term. The top ten regions were:

VT, DC, ME, AK, OR, MA, NH, RI, HI, CO

# Climate Change Keywords

Visualizations, such as the faceted line plots below help illustrate important trends in sentiment towards climate change and associated topics. For example, interest in climate change, global warming, and the green new deal topics all spiked during the 2020 election season, as climate change became an important topic in the Presidential race.

```
##           date hits        keyword geo       time gprop category
## 1   2019-11-24   50 climate change  US today 12-m   web        0
## 2   2019-12-01   76 climate change  US today 12-m   web        0
## 3   2019-12-08   79 climate change  US today 12-m   web        0
```

```
## 4   2019-12-15   56 climate change  US today 12-m   web         0
## 5   2019-12-22   23 climate change  US today 12-m   web         0
## 6   2019-12-29   32 climate change  US today 12-m   web         0
## 7   2020-01-05   67 climate change  US today 12-m   web         0
## 8   2020-01-12   69 climate change  US today 12-m   web         0
## 9   2020-01-19   56 climate change  US today 12-m   web         0
## 10  2020-01-26   57 climate change  US today 12-m   web         0
## 11  2020-02-02   57 climate change  US today 12-m   web         0
## 12  2020-02-09   60 climate change  US today 12-m   web         0
## 13  2020-02-16   65 climate change  US today 12-m   web         0
## 14  2020-02-23   66 climate change  US today 12-m   web         0
## 15  2020-03-01   61 climate change  US today 12-m   web         0
## 16  2020-03-08   52 climate change  US today 12-m   web         0
## 17  2020-03-15   33 climate change  US today 12-m   web         0
## 18  2020-03-22   34 climate change  US today 12-m   web         0
## 19  2020-03-29   40 climate change  US today 12-m   web         0
## 20  2020-04-05   41 climate change  US today 12-m   web         0
## 21  2020-04-12   45 climate change  US today 12-m   web         0
## 22  2020-04-19   56 climate change  US today 12-m   web         0
## 23  2020-04-26   50 climate change  US today 12-m   web         0
## 24  2020-05-03   49 climate change  US today 12-m   web         0
## 25  2020-05-10   42 climate change  US today 12-m   web         0
## 26  2020-05-17   36 climate change  US today 12-m   web         0
## 27  2020-05-24   31 climate change  US today 12-m   web         0
## 28  2020-05-31   29 climate change  US today 12-m   web         0
## 29  2020-06-07   26 climate change  US today 12-m   web         0
## 30  2020-06-14   20 climate change  US today 12-m   web         0
## 31  2020-06-21   19 climate change  US today 12-m   web         0
## 32  2020-06-28   18 climate change  US today 12-m   web         0
## 33  2020-07-05   19 climate change  US today 12-m   web         0
## 34  2020-07-12   18 climate change  US today 12-m   web         0
## 35  2020-07-19   20 climate change  US today 12-m   web         0
## 36  2020-07-26   20 climate change  US today 12-m   web         0
## 37  2020-08-02   19 climate change  US today 12-m   web         0
## 38  2020-08-09   20 climate change  US today 12-m   web         0
## 39  2020-08-16   26 climate change  US today 12-m   web         0
## 40  2020-08-23   30 climate change  US today 12-m   web         0
## 41  2020-08-30   32 climate change  US today 12-m   web         0
## 42  2020-09-06   59 climate change  US today 12-m   web         0
## 43  2020-09-13  100 climate change  US today 12-m   web         0
## 44  2020-09-20   81 climate change  US today 12-m   web         0
## 45  2020-09-27   68 climate change  US today 12-m   web         0
## 46  2020-10-04   61 climate change  US today 12-m   web         0
## 47  2020-10-11   49 climate change  US today 12-m   web         0
## 48  2020-10-18   51 climate change  US today 12-m   web         0
## 49  2020-10-25   49 climate change  US today 12-m   web         0
## 50  2020-11-01   41 climate change  US today 12-m   web         0
```

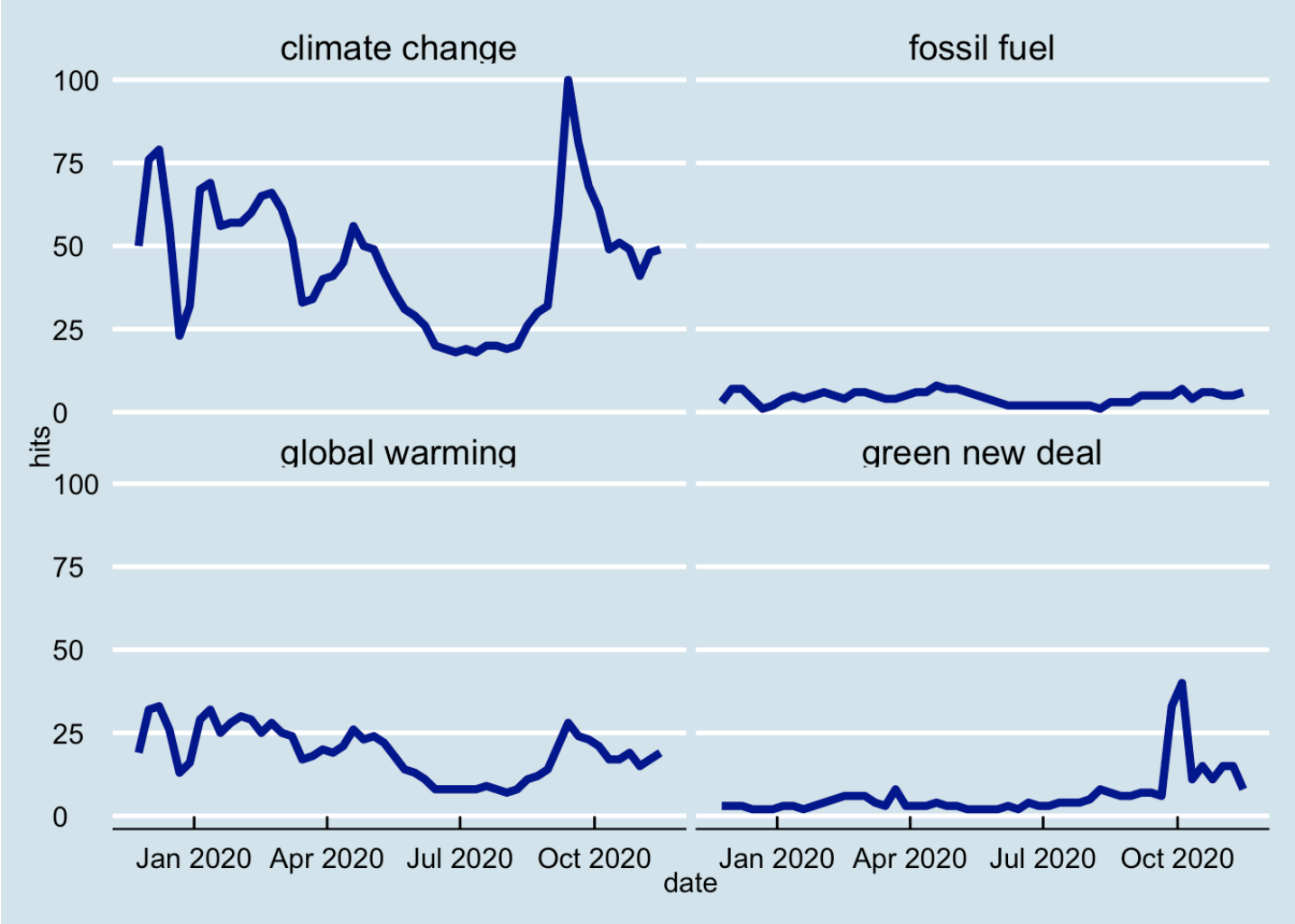
```
## 51  2020-11-08   48 climate change  US today 12-m   web         0
## 52  2020-11-15   49 climate change  US today 12-m   web         0
## 53  2019-11-24   19 global warming  US today 12-m   web         0
## 54  2019-12-01   32 global warming  US today 12-m   web         0
## 55  2019-12-08   33 global warming  US today 12-m   web         0
## 56  2019-12-15   26 global warming  US today 12-m   web         0
## 57  2019-12-22   13 global warming  US today 12-m   web         0
## 58  2019-12-29   16 global warming  US today 12-m   web         0
## 59  2020-01-05   29 global warming  US today 12-m   web         0
## 60  2020-01-12   32 global warming  US today 12-m   web         0
## 61  2020-01-19   25 global warming  US today 12-m   web         0
## 62  2020-01-26   28 global warming  US today 12-m   web         0
## 63  2020-02-02   30 global warming  US today 12-m   web         0
## 64  2020-02-09   29 global warming  US today 12-m   web         0
## 65  2020-02-16   25 global warming  US today 12-m   web         0
## 66  2020-02-23   28 global warming  US today 12-m   web         0
## 67  2020-03-01   25 global warming  US today 12-m   web         0
## 68  2020-03-08   24 global warming  US today 12-m   web         0
## 69  2020-03-15   17 global warming  US today 12-m   web         0
## 70  2020-03-22   18 global warming  US today 12-m   web         0
## 71  2020-03-29   20 global warming  US today 12-m   web         0
## 72  2020-04-05   19 global warming  US today 12-m   web         0
## 73  2020-04-12   21 global warming  US today 12-m   web         0
## 74  2020-04-19   26 global warming  US today 12-m   web         0
## 75  2020-04-26   23 global warming  US today 12-m   web         0
## 76  2020-05-03   24 global warming  US today 12-m   web         0
## 77  2020-05-10   22 global warming  US today 12-m   web         0
## 78  2020-05-17   18 global warming  US today 12-m   web         0
## 79  2020-05-24   14 global warming  US today 12-m   web         0
## 80  2020-05-31   13 global warming  US today 12-m   web         0
## 81  2020-06-07   11 global warming  US today 12-m   web         0
## 82  2020-06-14    8 global warming  US today 12-m   web         0
## 83  2020-06-21    8 global warming  US today 12-m   web         0
## 84  2020-06-28    8 global warming  US today 12-m   web         0
## 85  2020-07-05    8 global warming  US today 12-m   web         0
## 86  2020-07-12    8 global warming  US today 12-m   web         0
## 87  2020-07-19    9 global warming  US today 12-m   web         0
## 88  2020-07-26    8 global warming  US today 12-m   web         0
## 89  2020-08-02    7 global warming  US today 12-m   web         0
## 90  2020-08-09    8 global warming  US today 12-m   web         0
## 91  2020-08-16   11 global warming  US today 12-m   web         0
## 92  2020-08-23   12 global warming  US today 12-m   web         0
## 93  2020-08-30   14 global warming  US today 12-m   web         0
## 94  2020-09-06   21 global warming  US today 12-m   web         0
## 95  2020-09-13   28 global warming  US today 12-m   web         0
## 96  2020-09-20   24 global warming  US today 12-m   web         0
## 97  2020-09-27   23 global warming  US today 12-m   web         0
```

```
## 98  2020-10-04   21 global warming  US today 12-m   web         0
## 99  2020-10-11   17 global warming  US today 12-m   web         0
## 100 2020-10-18   17 global warming  US today 12-m   web         0
## 101 2020-10-25   19 global warming  US today 12-m   web         0
## 102 2020-11-01   15 global warming  US today 12-m   web         0
## 103 2020-11-08   17 global warming  US today 12-m   web         0
## 104 2020-11-15   19 global warming  US today 12-m   web         0
## 105 2019-11-24    3    fossil fuel  US today 12-m   web         0
## 106 2019-12-01    7    fossil fuel  US today 12-m   web         0
## 107 2019-12-08    7    fossil fuel  US today 12-m   web         0
## 108 2019-12-15    4    fossil fuel  US today 12-m   web         0
## 109 2019-12-22    1    fossil fuel  US today 12-m   web         0
## 110 2019-12-29    2    fossil fuel  US today 12-m   web         0
## 111 2020-01-05    4    fossil fuel  US today 12-m   web         0
## 112 2020-01-12    5    fossil fuel  US today 12-m   web         0
## 113 2020-01-19    4    fossil fuel  US today 12-m   web         0
## 114 2020-01-26    5    fossil fuel  US today 12-m   web         0
## 115 2020-02-02    6    fossil fuel  US today 12-m   web         0
## 116 2020-02-09    5    fossil fuel  US today 12-m   web         0
## 117 2020-02-16    4    fossil fuel  US today 12-m   web         0
## 118 2020-02-23    6    fossil fuel  US today 12-m   web         0
## 119 2020-03-01    6    fossil fuel  US today 12-m   web         0
## 120 2020-03-08    5    fossil fuel  US today 12-m   web         0
## 121 2020-03-15    4    fossil fuel  US today 12-m   web         0
## 122 2020-03-22    4    fossil fuel  US today 12-m   web         0
## 123 2020-03-29    5    fossil fuel  US today 12-m   web         0
## 124 2020-04-05    6    fossil fuel  US today 12-m   web         0
## 125 2020-04-12    6    fossil fuel  US today 12-m   web         0
## 126 2020-04-19    8    fossil fuel  US today 12-m   web         0
## 127 2020-04-26    7    fossil fuel  US today 12-m   web         0
## 128 2020-05-03    7    fossil fuel  US today 12-m   web         0
## 129 2020-05-10    6    fossil fuel  US today 12-m   web         0
## 130 2020-05-17    5    fossil fuel  US today 12-m   web         0
## 131 2020-05-24    4    fossil fuel  US today 12-m   web         0
## 132 2020-05-31    3    fossil fuel  US today 12-m   web         0
## 133 2020-06-07    2    fossil fuel  US today 12-m   web         0
## 134 2020-06-14    2    fossil fuel  US today 12-m   web         0
## 135 2020-06-21    2    fossil fuel  US today 12-m   web         0
## 136 2020-06-28    2    fossil fuel  US today 12-m   web         0
## 137 2020-07-05    2    fossil fuel  US today 12-m   web         0
## 138 2020-07-12    2    fossil fuel  US today 12-m   web         0
## 139 2020-07-19    2    fossil fuel  US today 12-m   web         0
## 140 2020-07-26    2    fossil fuel  US today 12-m   web         0
## 141 2020-08-02    2    fossil fuel  US today 12-m   web         0
## 142 2020-08-09    1    fossil fuel  US today 12-m   web         0
## 143 2020-08-16    3    fossil fuel  US today 12-m   web         0
## 144 2020-08-23    3    fossil fuel  US today 12-m   web         0
```

```
## 145 2020-08-30    3    fossil fuel  US today 12-m   web         0
## 146 2020-09-06    5    fossil fuel  US today 12-m   web         0
## 147 2020-09-13    5    fossil fuel  US today 12-m   web         0
## 148 2020-09-20    5    fossil fuel  US today 12-m   web         0
## 149 2020-09-27    5    fossil fuel  US today 12-m   web         0
## 150 2020-10-04    7    fossil fuel  US today 12-m   web         0
## 151 2020-10-11    4    fossil fuel  US today 12-m   web         0
## 152 2020-10-18    6    fossil fuel  US today 12-m   web         0
## 153 2020-10-25    6    fossil fuel  US today 12-m   web         0
## 154 2020-11-01    5    fossil fuel  US today 12-m   web         0
## 155 2020-11-08    5    fossil fuel  US today 12-m   web         0
## 156 2020-11-15    6    fossil fuel  US today 12-m   web         0
## 157 2019-11-24    3 green new deal  US today 12-m   web         0
## 158 2019-12-01    3 green new deal  US today 12-m   web         0
## 159 2019-12-08    3 green new deal  US today 12-m   web         0
## 160 2019-12-15    2 green new deal  US today 12-m   web         0
## 161 2019-12-22    2 green new deal  US today 12-m   web         0
## 162 2019-12-29    2 green new deal  US today 12-m   web         0
## 163 2020-01-05    3 green new deal  US today 12-m   web         0
## 164 2020-01-12    3 green new deal  US today 12-m   web         0
## 165 2020-01-19    2 green new deal  US today 12-m   web         0
## 166 2020-01-26    3 green new deal  US today 12-m   web         0
## 167 2020-02-02    4 green new deal  US today 12-m   web         0
## 168 2020-02-09    5 green new deal  US today 12-m   web         0
## 169 2020-02-16    6 green new deal  US today 12-m   web         0
## 170 2020-02-23    6 green new deal  US today 12-m   web         0
## 171 2020-03-01    6 green new deal  US today 12-m   web         0
## 172 2020-03-08    4 green new deal  US today 12-m   web         0
## 173 2020-03-15    3 green new deal  US today 12-m   web         0
## 174 2020-03-22    8 green new deal  US today 12-m   web         0
## 175 2020-03-29    3 green new deal  US today 12-m   web         0
## 176 2020-04-05    3 green new deal  US today 12-m   web         0
## 177 2020-04-12    3 green new deal  US today 12-m   web         0
## 178 2020-04-19    4 green new deal  US today 12-m   web         0
## 179 2020-04-26    3 green new deal  US today 12-m   web         0
## 180 2020-05-03    3 green new deal  US today 12-m   web         0
## 181 2020-05-10    2 green new deal  US today 12-m   web         0
## 182 2020-05-17    2 green new deal  US today 12-m   web         0
## 183 2020-05-24    2 green new deal  US today 12-m   web         0
## 184 2020-05-31    2 green new deal  US today 12-m   web         0
## 185 2020-06-07    3 green new deal  US today 12-m   web         0
## 186 2020-06-14    2 green new deal  US today 12-m   web         0
## 187 2020-06-21    4 green new deal  US today 12-m   web         0
## 188 2020-06-28    3 green new deal  US today 12-m   web         0
## 189 2020-07-05    3 green new deal  US today 12-m   web         0
## 190 2020-07-12    4 green new deal  US today 12-m   web         0
## 191 2020-07-19    4 green new deal  US today 12-m   web         0
```

```
## 192 2020-07-26    4 green new deal  US today 12-m   web         0
## 193 2020-08-02    5 green new deal  US today 12-m   web         0
## 194 2020-08-09    8 green new deal  US today 12-m   web         0
## 195 2020-08-16    7 green new deal  US today 12-m   web         0
## 196 2020-08-23    6 green new deal  US today 12-m   web         0
## 197 2020-08-30    6 green new deal  US today 12-m   web         0
## 198 2020-09-06    7 green new deal  US today 12-m   web         0
## 199 2020-09-13    7 green new deal  US today 12-m   web         0
## 200 2020-09-20    6 green new deal  US today 12-m   web         0
## 201 2020-09-27   33 green new deal  US today 12-m   web         0
## 202 2020-10-04   40 green new deal  US today 12-m   web         0
## 203 2020-10-11   11 green new deal  US today 12-m   web         0
## 204 2020-10-18   15 green new deal  US today 12-m   web         0
## 205 2020-10-25   11 green new deal  US today 12-m   web         0
## 206 2020-11-01   15 green new deal  US today 12-m   web         0
## 207 2020-11-08   15 green new deal  US today 12-m   web         0
## 208 2020-11-15    8 green new deal  US today 12-m   web         0
```

```
## Rows: 208
## Columns: 7
## $ date     <dttm> 2019-11-24, 2019-12-01, 2019-12-08, 2019-12-15, 2019-12-22,…
## $ hits     <int> 50, 76, 79, 56, 23, 32, 67, 69, 56, 57, 57, 60, 65, 66, 61, …
## $ keyword  <chr> "climate change", "climate change", "climate change", "clima…
## $ geo      <chr> "US", "US", "US", "US", "US", "US", "US", "US", "US", "US", …
## $ time     <chr> "today 12-m", "today 12-m", "today 12-m", "today 12-m", "tod…
## $ gprop    <chr> "web", "web", "web", "web", "web", "web", "web", "web", "web…
## $ category <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, …
```




Technical Resource Cited (https://martinctc.github.io/blog/vignette-google-trends-with-gtrendsr/)

Sentiment towards climate change can be difficult to track as individuals' search queries may contain different keywords. The faceted plot above also shows that volume for "climate change" searches greatly exceeds the "hits" value of synonymous terms, like "global warming" and also tangentially related topics, like "fossil fuel" and "green new deal". It is evident that out of the four terms plotted above, "climate change" is the most viable as an indicator of the general public's sentiment towards climate change. This is likely a result of it being the most used term in news outlets and research when referring to global changes in weather and the environment.

An interesting property of Google Trends data, which will be proven, is that in addition to tracking human sentiment towards longstanding topics, like climate change, it can also be used to detect one-time events

# Research Question 1

**Do Google search patterns correlate with real-world events?**

Causality between what happens in the real-world and Google search trends is a prerequisite for subsequent analysis. While it may seem intuitive, it is worth explaining the mechanisms by which society and events express themselves in the information age. A new adaptation of the proverbial question “does a tree really fall in the woods if no one is around to hear it?” may be “do current events really occur without being searched for on the internet?”. According to existing social science research - the short answer is no.

The use of Google’s search engine is so ingrained in the United States and global cultures, that it is the internet which aggregates and disseminates the vast majority of information humans use to interpret the world around them.

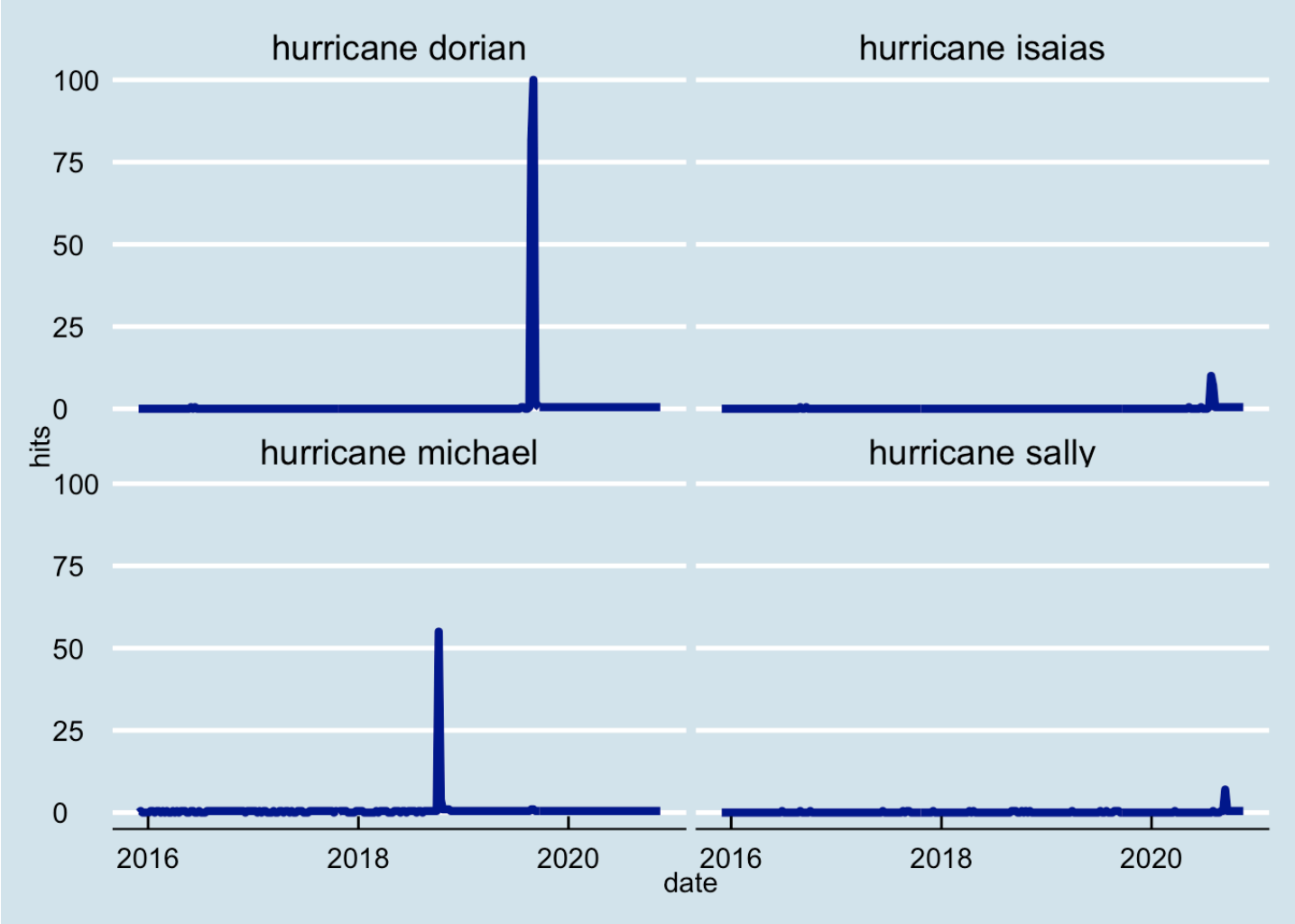
In order to illustrate the relationship between Google search patterns and the occurrence of natural disasters, it is necessary to view differences in interest over time for search queries pertaining to time-specific natural disasters.

Each fall, from August through October, two of the most devastating weather trends play out. Both of these phenomena lead to hundreds of deaths and billions of dollars in damages each year. A suitable method to track social interest in these disasters (as they happen) is to plot interest in the disaster name as a function of time. If there were to be a relationship, a clear spike in interest for the keyword would appear at the time of the event.

# Google Search Correlation with Florida Hurricane Events

The first of these two weather trends is tropical storm formation in the Atlantic, which ultimately leads to hurricanes in the southeastern United States. As global temperatures rise, the prevalence and intensity of tropical storm systems increase as well. Florida is subject to the most intense and frequent of these hurricanes.

```
## Rows: 1,040
## Columns: 7
## $ date     <dttm> 2015-11-29, 2015-12-06, 2015-12-13, 2015-12-20, 2015-12-27,…
## $ hits     <chr> "0", "<1", "0", "0", "0", "0", "<1", "<1", "0", "<1", "<1", …
## $ keyword  <chr> "hurricane michael", "hurricane michael", "hurricane michael…
## $ geo      <chr> "US-FL", "US-FL", "US-FL", "US-FL", "US-FL", "US-FL", "US-FL…
## $ time     <chr> "today+5-y", "today+5-y", "today+5-y", "today+5-y", "today+5…
## $ gprop    <chr> "web", "web", "web", "web", "web", "web", "web", "web", "web…
## $ category <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, …
```




It is thus verifiable that all of these events garnered peak interest on Google at the precise date that they occurred. The spike in Google query hits is so sharp, it seems as though human interest in these events is short-lived and mostly isolated at the exact moment of their occurrence. This directly confirms the hypothesis of research question 1: Google search volume for a natural disaster peaks precisely at the event date.

Hurricane Isaias and Hurricane Sally were two of the most devastating storms to hit the Gulf Coast and Southeastern United States in 2020. They were classified as Category 1 and Category 2 storms respectively, leaving more than $10.0 billion USD of damage in their wakes.

Conversely, Hurricane Dorian from 2019 and Hurricane Michael from 2018 were both Category 5 storms, significantly more powerful and inflicting even more damage. Hurricane Dorian is attributed to more than 100 deaths and $5.1 billion USD in damage alone. Hurricane Michael caused 74 deaths and led to $25.5 billion USD in damage.

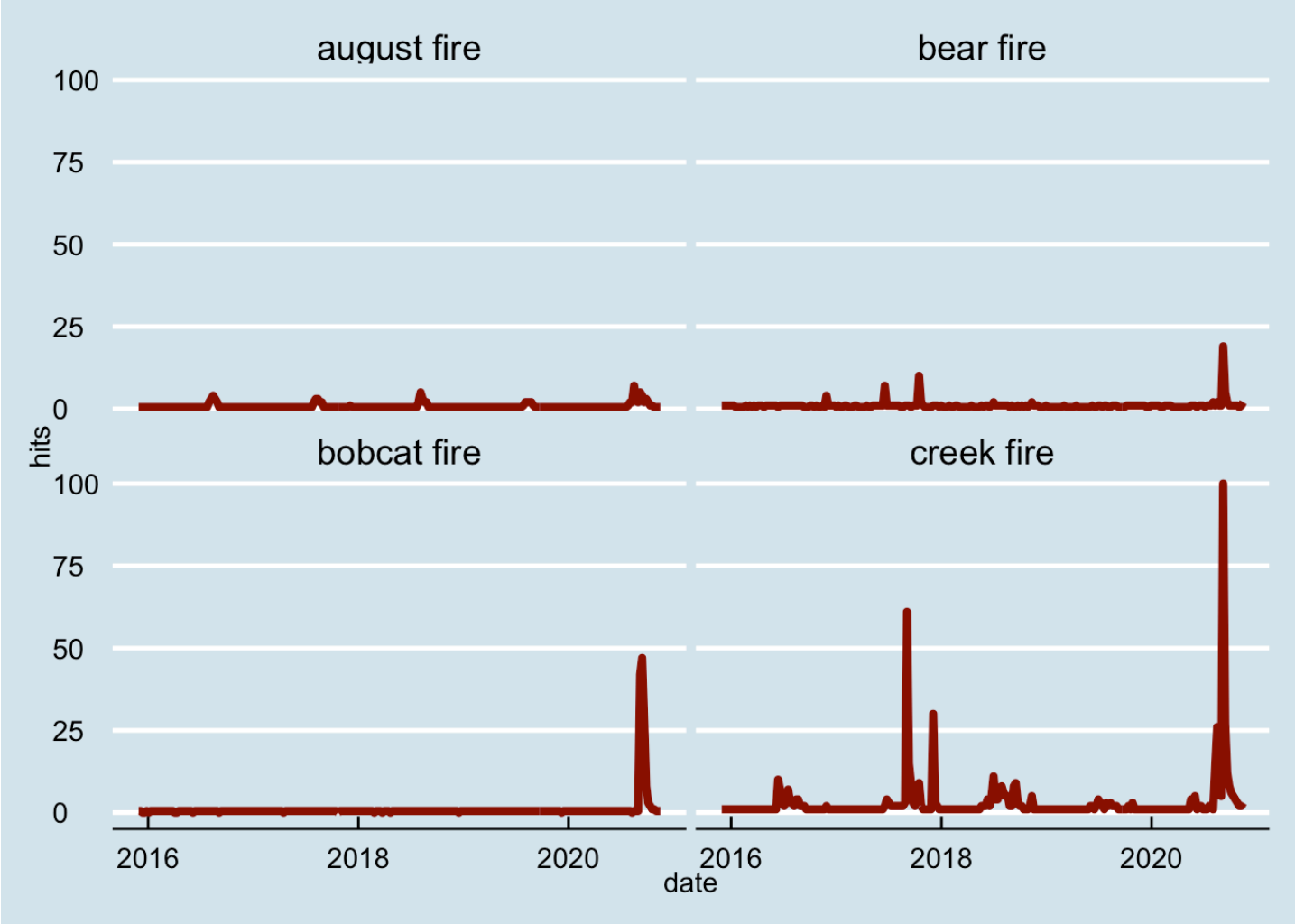
Google search interest for these category 5 hurricanes showed no more than ripples (hits ≤ 5) for the entirety of this past year. This would be unlikely if it were the case that event interest on Google was not time dependent; Hurricanes Dorian and Michael were significantly more consequential than Hurricanes Isaias and Sally.

When the data is analyzed carefully however, it is evident that even search hits for Hurricanes Dorian and Michael also increased fractionally during the 2020 hurricane season - around the time Google searches for active hurricanes spiked. It is perhaps the case that people do become more interested in past natural disasters in order to contextualize the events that they are currently experiencing.

# Google Search Correlation with California Wildfire Events

The second climate trend is increased temperatures in the western U.S., which lead to brush and forest fires in California. In 2018 alone, more than 1,670,000 acres of land burned, California's most destructive wildfire season to date.

```
## Rows: 1,040
## Columns: 7
## $ date     <dttm> 2015-11-29, 2015-12-06, 2015-12-13, 2015-12-20, 2015-12-27,…
## $ hits     <chr> "<1", "<1", "0", "0", "<1", "0", "<1", "<1", "<1", "<1", "<1…
## $ keyword  <chr> "bobcat fire", "bobcat fire", "bobcat fire", "bobcat fire", …
## $ geo      <chr> "US", "US", "US", "US", "US", "US", "US", "US", "US", "US", …
## $ time     <chr> "today+5-y", "today+5-y", "today+5-y", "today+5-y", "today+5…
## $ gprop    <chr> "web", "web", "web", "web", "web", "web", "web", "web", "web…
## $ category <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, …
```




Like the Google trends output for Hurricanes in Florida, wildfires in California show the same property of having peak interest on Google around the time of the event itself.

A key difference between wildfires in California and hurricanes in Florida is that wildfires in California tend to have a much longer duration, often lasting weeks. This is evident by the increased time periods of certain spikes in interest, which is even more evident when the `gtrends()` query is performed on a shorter duration (such as the past 1 year instead of 5 years).

Another notable feature of the line plots is that there appears to be multiple time periods that return high hits values for a given query. For example, “creek fire” most notably refers to the largest wildfire in recent California history that started on September 4th, 2020 and is still active at the date of submission for this paper as of November 22nd, 2020. However, due to the naming convention of California wildfires, it also returned Google “hits” values for the Creek Fire of greater Los Angeles in 2017 and smaller, similarly named fires in 2018. This makes it harder to infer when the exact peak in interest for an event occurred, as many of the keywords refer to multiple wildfires in recent California history.

An additional example of noise in this search query can be seen in the “august fire” graph, where, in addition to a peak at the start of the 2020 August Complex fire (August 16th, 2020), there are regular jumps visible each year in the month of August. This is surely not a coincidence, but rather of conflated search queries.

The faceted line plot for interest over time for California wildfires still confirms the hypothesis that Google search queries peak at the time of an event, as this was still an observable pattern in the data. All four example plots returned maximum hit values at days within the duration of each wildfire. However, it does prove the need to tread carefully. Depending on the nature of weather phenomena, or a weather phenomenon's naming convention, Google search queries for associated terms may appear noisy or misleading. Additionally, it becomes increasingly challenging to make claims like "local populations tend to/ to not search for past natural disaster events in order to contextualize impending events" as the data is more ambiguous.

# Transition to Climate Change Sentiment Analysis

Thus, it is also necessary to consider the impact (if any) which these events have on the general sentiment towards climate change, one of the most important factors which dictates intensity and frequency of these events. Many news sources and social networks address climate change and human action as a contributing factor to the destruction caused by such natural disasters.

---

# Research Question 2

**Does interest in climate change increase regionally after the occurrence of a natural disaster?**

# Context

By virtue of research question 1, it is reasonable to assume a relationship between natural disasters - physical events that shape critical aspects of the human experience - and Google search trends for these events. When FEMA declares an emergency, for a hurricane, wildfire, flood, tornado or other event, people who reside in the affected region want to know what is going on. This includes searching the internet for more information about the nature of the emergency, advice regarding how to protect their families and property, what to expect, and how to prepare. The Google trends hits value associated with the disaster's name thus fluctuates dramatically in the days leading up to, during, and after a natural disaster impacts a community.

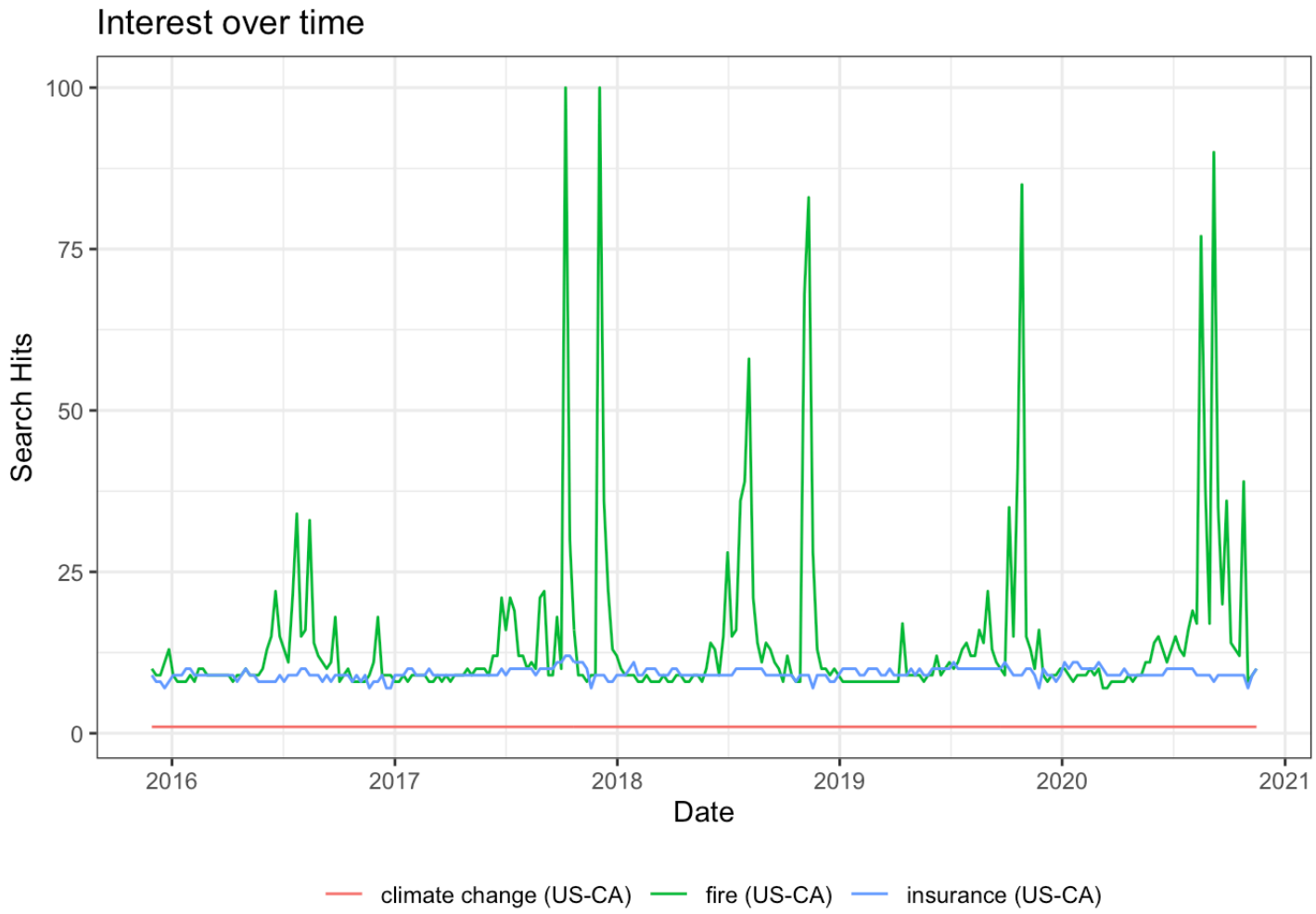
The next logical question to pose - at least from the perspective of this analysis - is whether or not humans have a tendency to search for topics like climate change after the occurrence of a local natural disaster. One of the most noticeable and consequential results of increased global warming is that extreme weather events, like storms, fires and floods, become more severe and can occur more frequently. This is often used as a rallying-cry by climate activists. Without timely action to minimize greenhouse gas emissions and protect our natural resources, there will likely be irreversible damage to the planet.

Therefore, it is critical to assess whether individuals in regions affected by natural disasters turn to Google in the days surrounding said event for information regarding climate change. Does the occurrence of a natural disaster ultimately lead to any change in internet behavior?

# Point of Qualification

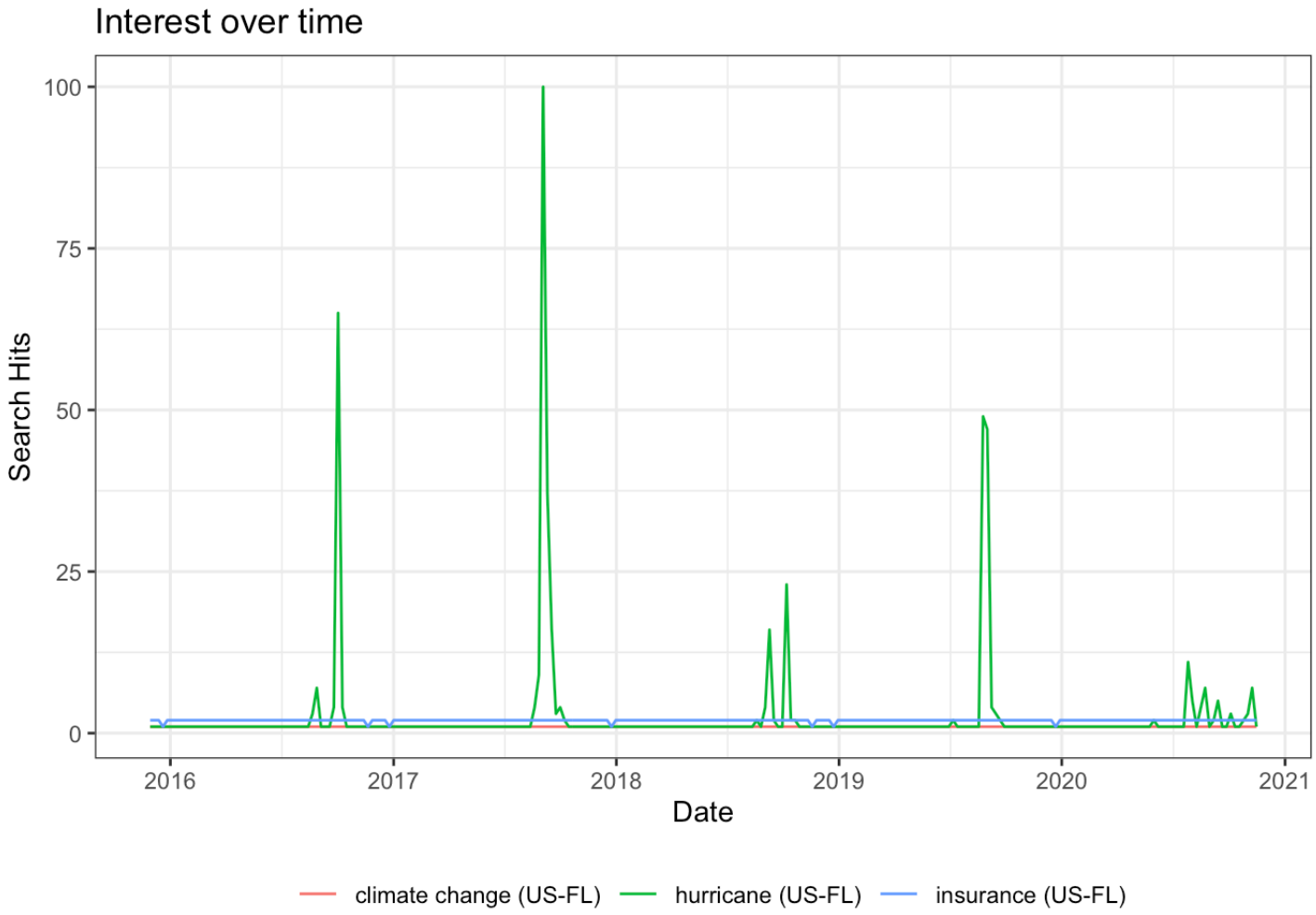
When using the gTrendsR package to analyze interest over time for “climate change” and “fire” keywords in California from the last five years, it is evident that climate change as a search term always receives fewer hits than fire. This could be the resulting confluence of many principles. First of all, some terms are just more prevalent in the public conscious than others. Fire often poses an immediate threat to the health and safety of individuals, whereas climate change, while still a real threat to society, is less immediate and obvious. Additionally, it is likely that “fire” as a keyword also tabulates search volume for search queries unrelated to wildfires, such as “How to start a fire?”. Conversely, individuals may be quite curious for topics like climate change and their connection to natural disasters, but use other terms in their query, such as “what causes natural disasters?”. These are inherent flaws of the study that must be stated. They serve as a point of qualification for any broad claims which this research seeks to make. Without the use of sophisticated machine learning models or fine-tuning the Google trends functionality, these are sources of error which must be accepted.

# California Interest Over Time Graph



This interesting output shows the utility of Google trends' output when displayed in a readable format. Interest for "insurance" is also plotted, as damage from wildfires often leads to an increase in insurance claims. However there does not seem to be a clear relationship on this graph between any of the three terms "fire", "insurance" or "climate change" in California over the past five years.

# Florida Interest Over Time Graph




The next visualization shows the Google trends interest over time for "hurricane", "insurance" and "climate change" plotted relative to one another. This time, the relative volume of "climate change" search queries *and* "insurance" search queries is dwarfed by the number of searches for "hurricane". There is again no clear relationship between any of these three keywords.

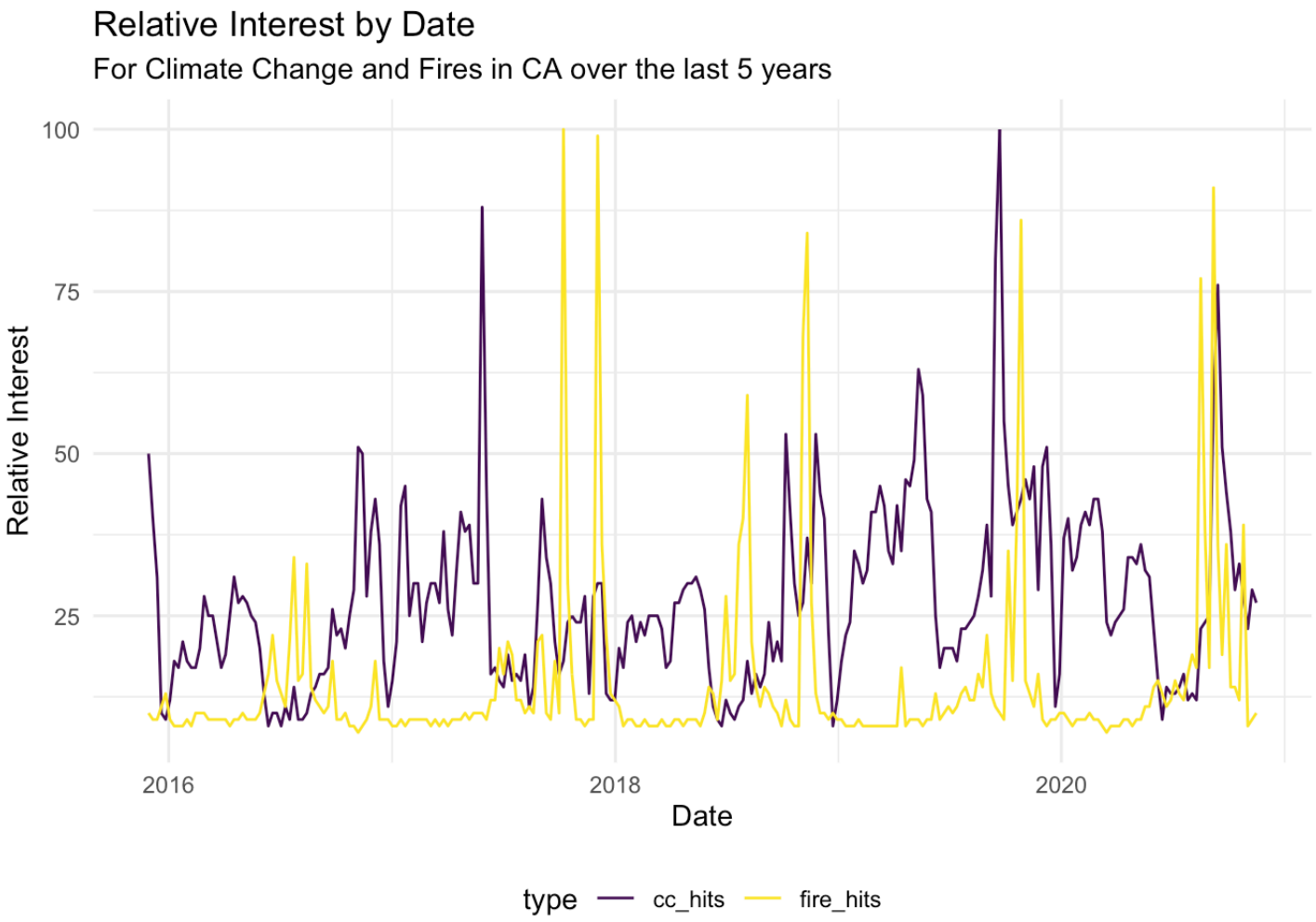
# Individual Keyword Queries Transformation

By transforming the data such that individual queries are processed for each of the three search terms, rather than calling `gtrends()` once on a vector of both "event type" and "climate change" keywords, it is possible to plot interest over time for each search term independently of the other By doing so, each search term will be plotted on the same proportional y-axis, where hits represent – by definition – the proportion of search

queries relative to its own maximum search volume within the given time and location parameters. This produces a more interesting output where anomalies in either of the search term's hits volume can be detected with ease, rather than having climate change search volume consistently at hits = <1. Insurance was dropped for simplicity.

# California Interest Over Time Transformation Graph

```
## Joining, by = "date"
```



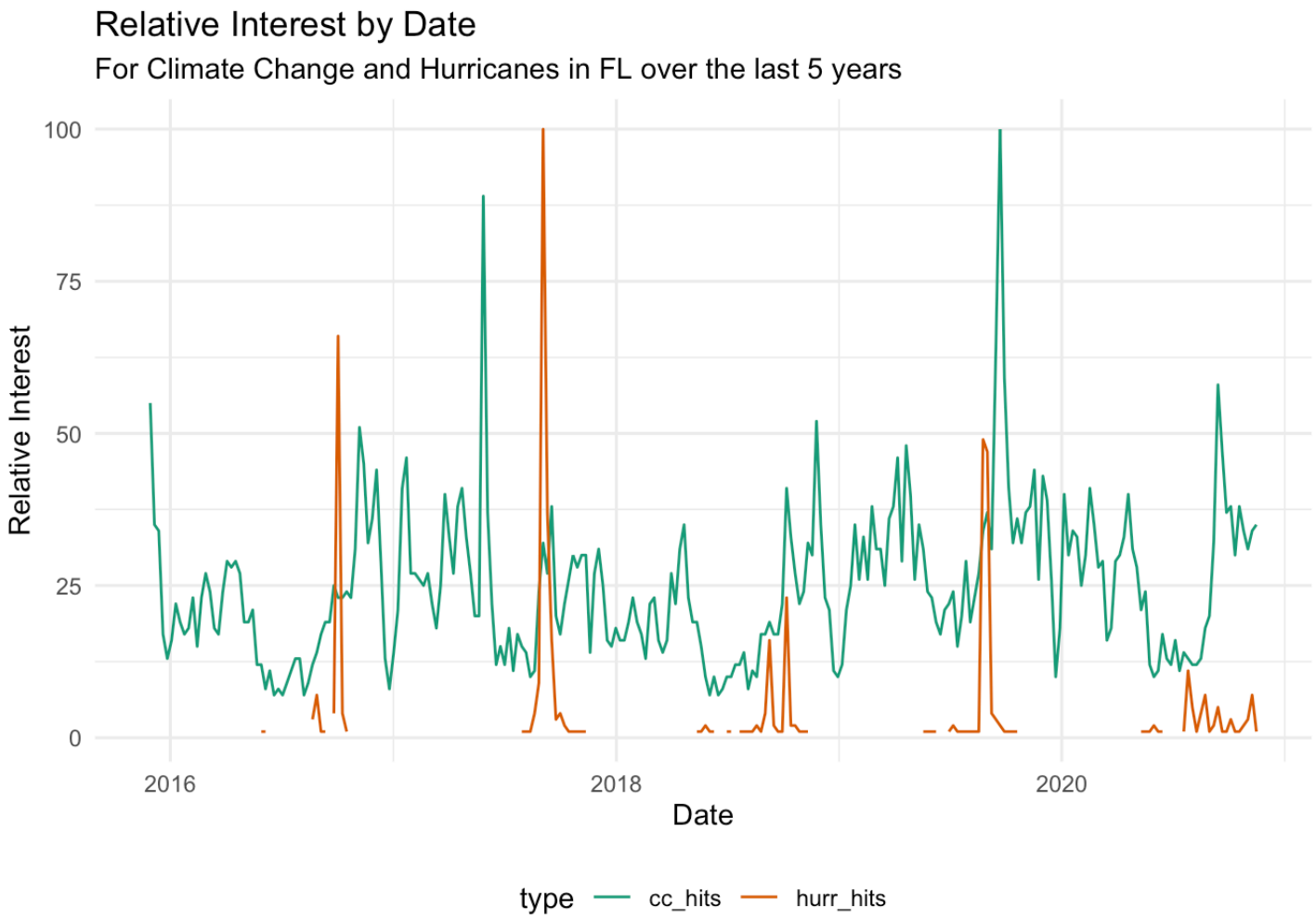

A transformed interest over time plot for wildfires and climate change in California over the past 5 years.

# Florida Interest Over Time Transformation Graph

```
## Warning in data.frame(as.integer(indiv_hurr_trend[["interest_over_time"]]
## [["hits"]])): NAs introduced by coercion
```

```
## Joining, by = "date"
```

```
## Warning: Removed 26 row(s) containing missing values (geom_path).
```




A transformed interest over time plot for hurricanes and climate change in Florida over the past 5 years.

# Combining FEMA data with Google Trends Hits Queries

By looking at the Google Trends change in search volume for climate change immediately after a natural disaster takes place, it is plausible to make conclusions about the presence (or absence) of fluctuations in local interest for climate change as a result of the natural disaster.

# Florida Hurricanes since 2015-11-23

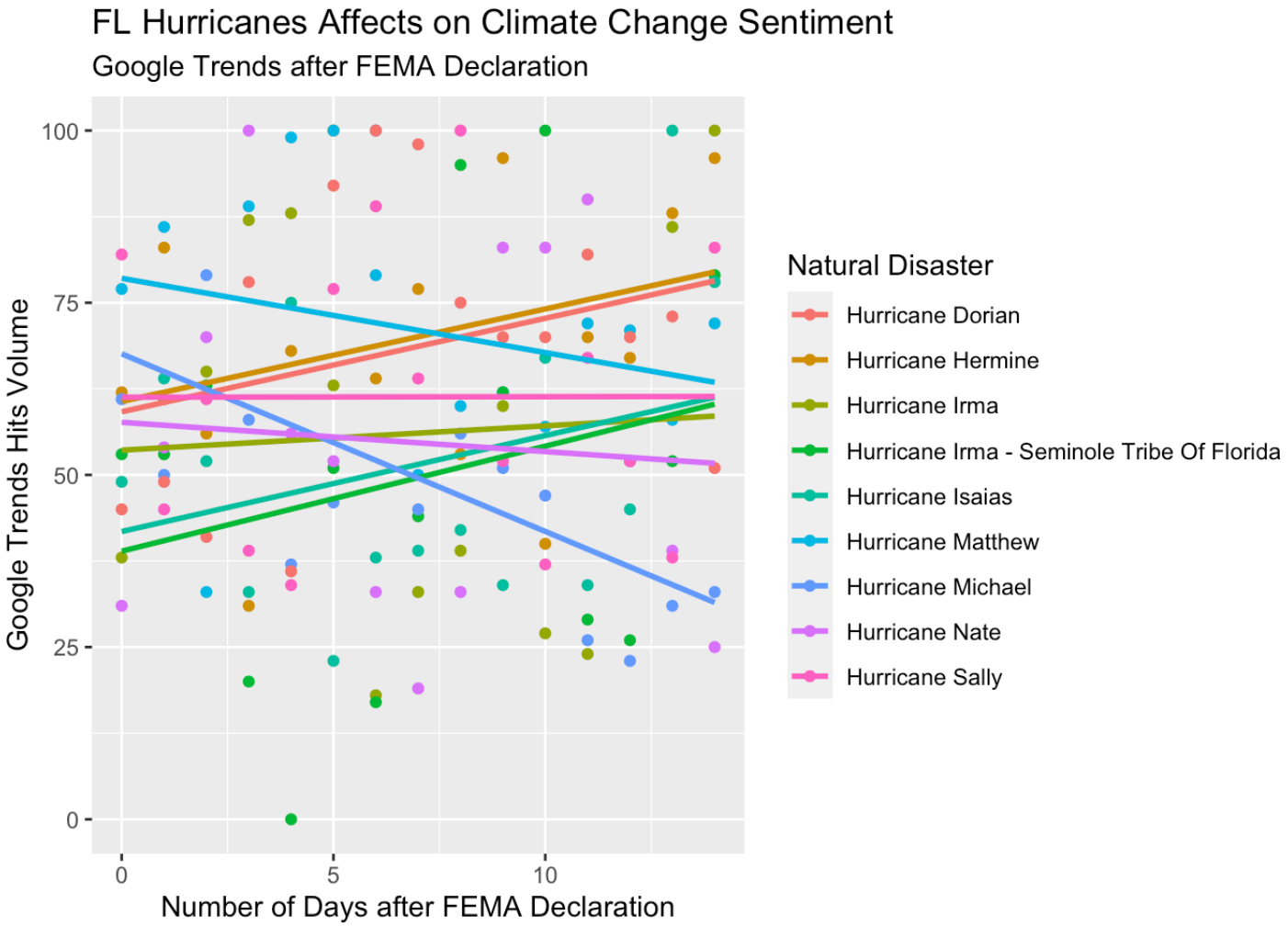
A list of all FEMA declarations for hurricanes in Florida occurring in the last 5 years.

```
## # A tibble: 9 x 2
##   declaration_title                          declaration_day
##   <chr>                                      <chr>
## 1 Hurricane Hermine                          2016-09-28
## 2 Hurricane Matthew                          2016-10-06
## 3 Hurricane Irma                             2017-09-05
## 4 Hurricane Irma - Seminole Tribe Of Florida 2017-09-27
## 5 Hurricane Nate                             2017-10-08
## 6 Hurricane Michael                          2018-10-09
## 7 Hurricane Dorian                           2019-08-30
## 8 Hurricane Isaias                           2020-08-01
## 9 Hurricane Sally                            2020-09-16
```

##Climate Change Sentiment after Hurricanes in Florida

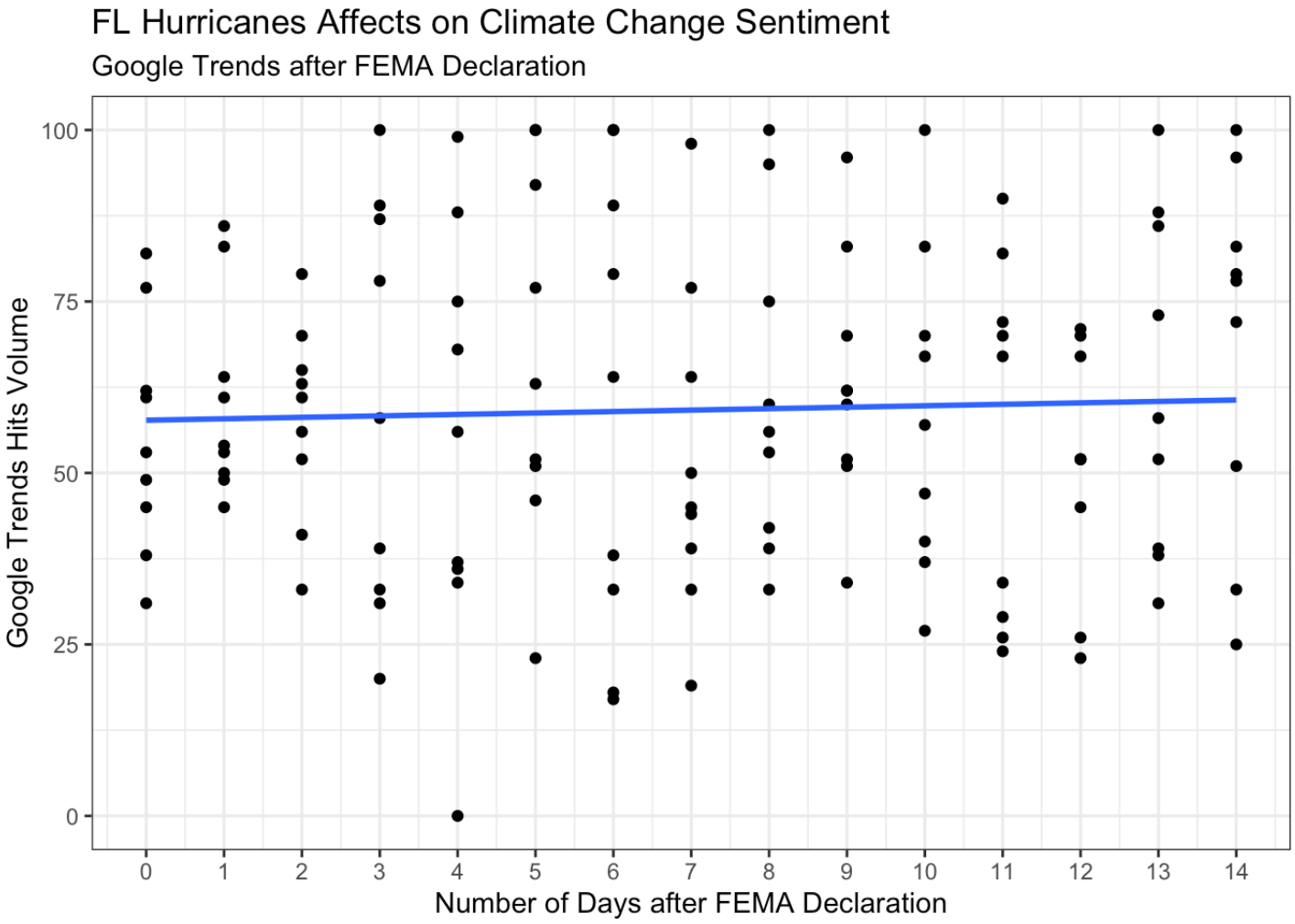
shows the change in search volume for “climate change” during the 14 days after each Florida hurricane in the `FEMA_Declarations` dataset.

```
## [1] 1
## [1] 2
## [1] 3
## [1] 4
## [1] 5
## [1] 6
## [1] 7
## [1] 8
## [1] 9
```




This visual shows the Google trends hits volume as a function of days after a natural disaster takes place, grouped by natural disaster. It is hard to extrapolate from this, as there are 9 events represented on the same graph, and the colors make it difficult to read.

A more intuitive method is to create one linear model that *ignores* the grouping by event, as the mean of all the slopes is an even more meaningful indicator of patterns in human behavior across different natural disasters.




Here, one linear model is visible, representing the average daily change in search volume for climate change in the two weeks following every Florida hurricane from the past 5 years. The null hypothesis is that it equals 0, because assuming there is no relationship between a natural disaster and climate change sentiment, there should be no change in daily search volume for climate change after a hurricane occurs. It appears to be slightly negative, but more analysis is necessary.

```
## # A tibble: 8 x 4
## # Groups:   event_name [8]
##   event_name                                event_date slope significance
##   <chr>                                     <date>     <dbl>        <dbl>
## 1 Hurricane Hermine                         2016-09-28  1.34        0.286
## 2 Hurricane Matthew                         2016-10-06 -1.08        0.345
## 3 Hurricane Irma                            2017-09-05  0.35        0.829
## 4 Hurricane Irma - Seminole Tribe Of Florida 2017-09-27  1.53        0.388
## 5 Hurricane Nate                            2017-10-08 -0.42        0.791
## 6 Hurricane Dorian                          2019-08-30  1.36        0.281
## 7 Hurricane Isaias                          2020-08-01  1.39        0.286
## 8 Hurricane Sally                           2020-09-16  0.01        0.996
```

Methodology from reprex package Reference (https://stackoverflow.com/questions/51355303/extract-slope-of-multiple-trend-lines-from-geom-smooth)

7 of the 9 hurricanes have linear models with significance levels greater than 0.2. These represent the slopes that are moderate in value, by filtering out all the slopes that were far too extreme to be realistic.

Next, it is possible to conduct a t-test to determine a p-value for the null hypothesis that the change in climate change hits immediately after a hurricane disaster will not change.

```
##
##  One Sample t-test
##
## data:  hurr_slopes_output$slope
## t = 1.5983, df = 7, p-value = 0.154
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.2685108  1.3885108
## sample estimates:
## mean of x
##      0.56
```

The p-value = 0.6881, which is greater than the significance level of 0.05, meaning that we fail to reject the null hypothesis that the change in climate change hits immediately after a hurricane in Florida will not change. Thus, there appears to be insufficient evidence to make a causal claim about the relationship between hurricanes in Florida and local Google trends sentiment for climate change.

---

# California Wildfires since 2015-11-23

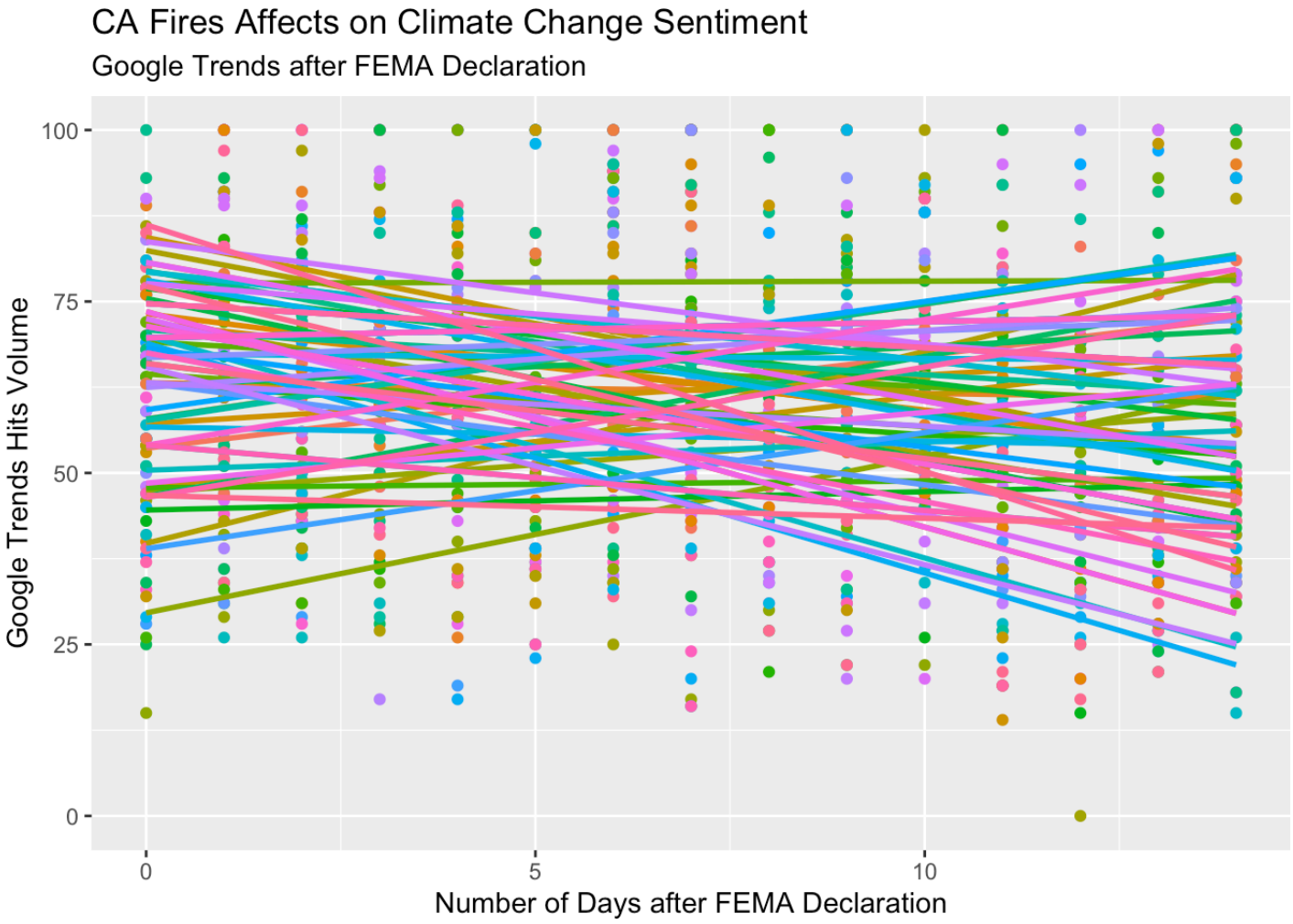
A list of all FEMA declarations for wildfires in California occurring in the last 5 years.

```
## # A tibble: 75 x 2
##    declaration_title declaration_day
##    <chr>             <chr>
##  1 Old Fire          2016-06-05
##  2 Border 3 Fire     2016-06-20
##  3 Fish Fire         2016-06-21
##  4 Erskine Fire      2016-06-24
##  5 Sage Fire         2016-07-10
##  6 Sand Fire         2016-07-24
##  7 Soberanes Fire    2016-07-29
##  8 Goose Fire        2016-07-31
##  9 Pilot Fire        2016-08-09
## 10 Clayton Fire      2016-08-14
## # … with 65 more rows
```

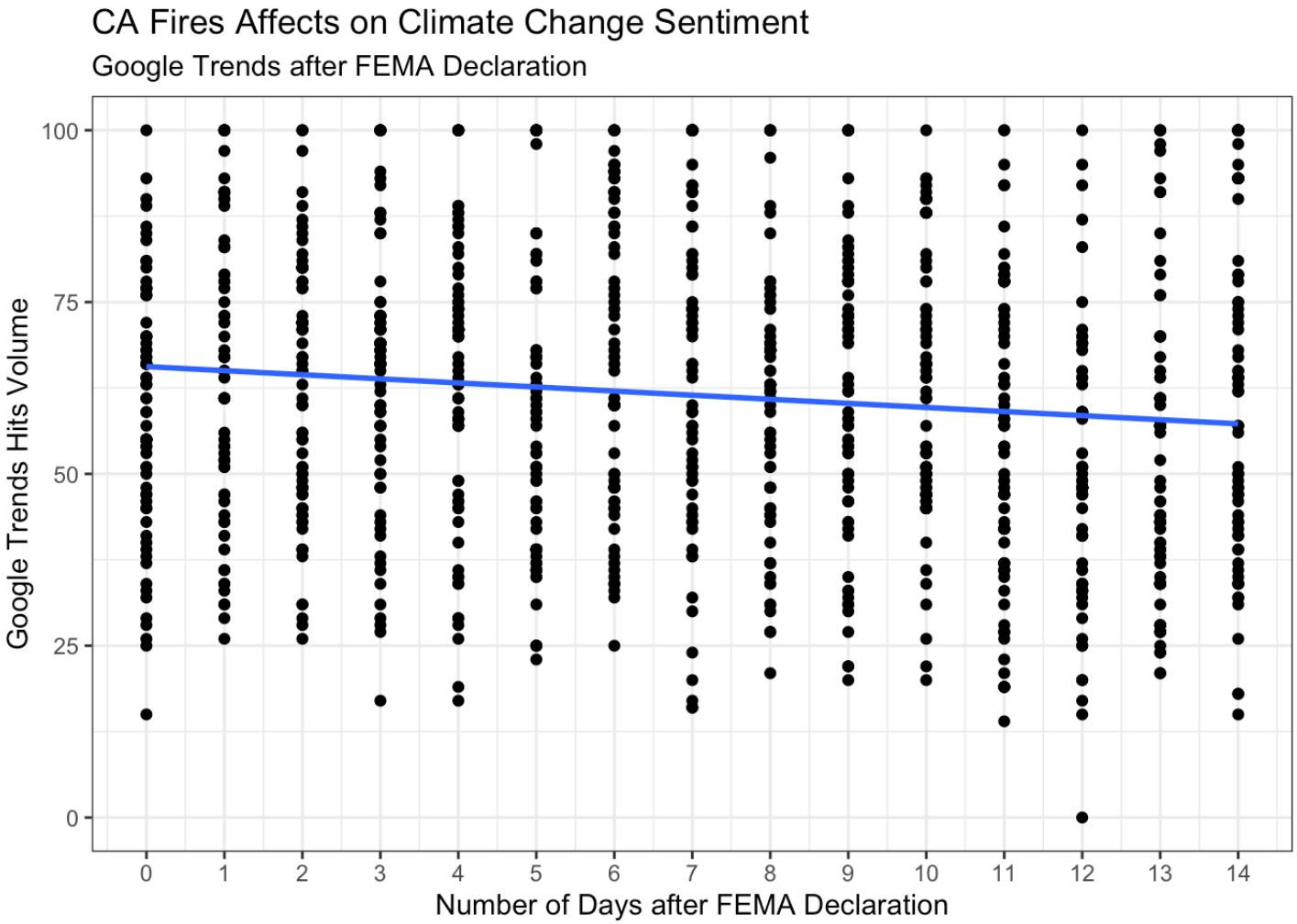
---

```
## [1] 1
## [1] 2
## [1] 3
## [1] 4
## [1] 5
## [1] 6
## [1] 7
## [1] 8
## [1] 9
## [1] 10
## [1] 11
## [1] 12
## [1] 13
## [1] 14
## [1] 15
## [1] 16
## [1] 17
## [1] 18
## [1] 19
## [1] 20
## [1] 21
## [1] 22
## [1] 23
## [1] 24
## [1] 25
## [1] 26
## [1] 27
## [1] 28
## [1] 29
## [1] 30
## [1] 31
## [1] 32
## [1] 33
## [1] 34
## [1] 35
## [1] 36
## [1] 37
## [1] 38
## [1] 39
## [1] 40
## [1] 41
## [1] 42
## [1] 43
## [1] 44
## [1] 45
## [1] 46
## [1] 47
```

```
## [1] 48
## [1] 49
## [1] 50
## [1] 51
## [1] 52
## [1] 53
## [1] 54
## [1] 55
## [1] 56
## [1] 57
## [1] 58
## [1] 59
## [1] 60
## [1] 61
## [1] 62
## [1] 63
## [1] 64
## [1] 65
## [1] 66
## [1] 67
## [1] 68
## [1] 69
## [1] 70
## [1] 71
## [1] 72
## [1] 73
## [1] 74
## [1] 75
```




This is not helpful as there are 75 independent fire events present, making it extremely difficult to infer anything from the visualization.




Here, one linear model is visible, representing the average daily change in search volume for climate change in the two weeks following every California fire from the past 5 years. The null hypothesis is that it equals 0, because assuming there is no relationship between a natural disaster and climate change sentiment, there should be no change in daily search volume for climate change after a natural disaster occurs. It appears to be slightly negative, but more analysis is necessary.

```
## # A tibble: 45 x 4
## # Groups:   event_name [45]
##    event_name     event_date  slope significance
##    <chr>          <date>      <dbl>        <dbl>
##  1 Border 3 Fire  2016-06-20 -1.43         0.245
##  2 Fish Fire      2016-06-21  0.27         0.844
##  3 Erskine Fire   2016-06-24 -0.75         0.582
##  4 Sage Fire      2016-07-10 -0.64         0.618
##  5 Soberanes Fire 2016-07-29  1.03         0.428
##  6 Goose Fire     2016-07-31  0.580        0.650
##  7 Clayton Fire   2016-08-14  0.82         0.568
##  8 Chimney Fire   2016-08-15 -1.91         0.208
##  9 Blue Cut Fire  2016-08-16 -0.16         0.912
## 10 Detwiler Fire  2017-07-18 -0.77         0.591
## # … with 35 more rows
```

41 of the 75 fires have linear models with significance levels greater than 0.2. These represent the slopes that are moderate in value, by filtering out all the slopes that were far too extreme to be realistic.

Next, it is possible to conduct a t-test to determine a p-value for the null hypothesis that the change in climate change hits immediately after a fire disaster will not change.

```
##
##  One Sample t-test
##
## data:  fire_slopes_output$slope
## t = -3.1895, df = 44, p-value = 0.002627
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.6473088 -0.1460246
## sample estimates:
##  mean of x
## -0.3966667
```

The p-value = 0.2888, which means we fail to reject the null hypothesis that the change in climate change hits immediately after a wildfire in California will not change. Thus, there appears to be insufficient evidence to make a causal claim about the relationship between wildfires in California and local Google trends sentiment for climate change.

Although we formally reject the null hypothesis since p-value is greater than 0.05, the p-value for the t-test on California wildfires was dramatically lower than the p-value for Florida Hurricanes. Additionally, the degrees of freedom was much higher, and more values were input into the test (40 rather than 6). This is very promising as it may indicate that the narrative around California wildfires is more focused on climate change as a root cause than the narrative surrounding hurricanes in Florida.

---

# Research Question 3 - Looking Forward

**Do people become more environmentally conscious after a natural disaster?**

The final question, which pushes beyond the scope of this inquiry, is whether or not humans' climate-actions change as a result of natural disasters. For example, in the days and weeks after a natural disaster, are more individuals likely to sign up for a climate petition? After a long and tumultuous wildfire season, are California residents more likely to buy an electric vehicle and reduce their own carbon footprint? Or are Florida residents more likely to reduce their waste and protect their coastlines after hurricanes cause damage to them? These are complex questions, but ultimately critical as we seek to understand how human behavior changes after local natural disasters threaten their ways of life.

Further research and data analysis will be necessary before such claims can be evaluated in a statistically meaningful way. Regardless, it is imperative that society considers the ways in which these natural disasters shape one's existence. Even more so, assuring that the public understands the relationship between human actions and climate change is absolutely critical to mitigating a future climate catastrophe. Without public support and trust, the environmental sciences as a discipline will never have the necessary support to push for significant changes in the government, economic, and social sectors.

---

# Conclusion

## Research Questions

Based on this analysis, it is clear there is a relationship between real-world events, like natural disasters, and Google search volume for said terms. When FEMA declares an emergency in a given region, more individuals search for these keywords in order to gather information.

In regards to assessing if there is a direct relationship between natural disasters and search volume for climate change, it is inconclusive whether or not there is a causal relationship. While there tends to be a slight decrease in search volume for climate change during the days after a natural disaster is declared, it is challenging to assert that this rise in Google trends hits for climate change is a result of the natural disaster itself. Rather, it could also be the product of other factors, like power outages, leading to fewer individuals accessing the internet, or also people instead searching for emergency information and being preoccupied with other matters. In a broader sense, perhaps the two weeks following a natural disaster are not even the best time to learn about the mechanisms and effects of climate change.

This being said, it was clear that the decreasing pattern for climate change hits over the 14 days following a wildfire in California was more significant, visually from the linear models outputs and quantitatively from the p-tests. This could mean that the relationship between natural disasters and climate change Google searches is stronger for California wildfires, where climate change is more readily attributed to wildfires as a cause, and

people are more likely to search for said terms when a wildfire occurs. It could also just mean that there are other confounding variables which impact an individual's ability to use Google in the two weeks following a FEMA declaration for a fire - such as the fire itself.

# What Else?

From a sociological perspective, it is also interesting to map the extent of these complex relationships between real events and human sentiment. While natural disasters and climate change were the focus of this analysis, other such connections could be examined as well using Google trends data, such as the relationship between Google search patterns and investment behavior, political affiliation, economic prosperity, race, religion, familial relationships, or even allegiance to a particular sports team. Google Trends is an incredibly powerful analytic engine that is underutilized in the social sciences.

---

# Bibliography

## Existing Scholarship on Climate Change Sentiment Analysis

1. CHOI, H. and VARIAN, H. (2012), Predicting the Present with Google Trends. Economic Record, 88: 2-9. https://doi.org/10.1111/j.1475-4932.2012.00809.x (https://doi.org/10.1111/j.1475-4932.2012.00809.x)
2. "What do Google searches tell us about our climate change fears?" NewsRx Health & Science, 3 Aug. 2014, p. 177. Gale Academic OneFile, https://link.gale.com/apps/doc/A378494502/AONE?u=duke_perkins&sid=AONE&xid=7d6f5991 (https://link.gale.com/apps/doc/A378494502/AONE?u=duke_perkins&sid=AONE&xid=7d6f5991). Accessed 16 Nov. 2020.
3. Pierre-louis, K., & Schwartz, J. (2020, August 21). Why Does California Have So Many Wildfires? Retrieved November 21, 2020, from https://www.nytimes.com/article/why-does-california-have-wildfires.html (https://www.nytimes.com/article/why-does-california-have-wildfires.html)
4. Machemer, T. (2020, August 21). In California, More Acres Have Burned in the 2020 Fire Season Than in All of 2019. Retrieved November 22, 2020, from https://www.smithsonianmag.com/smart-news/california-more-acres-have-burned-2020-all-2019-180975629/ (https://www.smithsonianmag.com/smart-news/california-more-acres-have-burned-2020-all-2019-180975629/)
5. Florida Hurricane List. (n.d.). Retrieved November 23, 2020, from https://www.nhc.noaa.gov/outreach/history/ (https://www.nhc.noaa.gov/outreach/history/)
6. Kurt Cobb (2019, March 18). Deep adaptation, post-sustainability and the possibility of societal collapse. Retrieved November 23, 2020, from https://www.resilience.org/stories/2019-03-17/deep-adaptation-post-sustainability-and-the-possibility-of-societal-collapse/

(https://www.resilience.org/stories/2019-03-17/deep-adaptation-post-sustainability-and-the-possibility-of-societal-collapse/)

# Technical Sources for Programming & Visualizations

7. https://cran.r-project.org/web/packages/gtrendsR/gtrendsR.pdf (https://cran.r-project.org/web/packages/gtrendsR/gtrendsR.pdf)
8. https://davetang.org/muse/2018/12/31/visualising-google-trends-results-with-r/ (https://davetang.org/muse/2018/12/31/visualising-google-trends-results-with-r/)
9. http://rstudio-pubs-static.s3.amazonaws.com/520533_f25fa191afa7476a995322a4393710ce.html (http://rstudio-pubs-static.s3.amazonaws.com/520533_f25fa191afa7476a995322a4393710ce.html)
10. https://stackoverflow.com/questions/51355303/extract-slope-of-multiple-trend-lines-from-geom-smooth (https://stackoverflow.com/questions/51355303/extract-slope-of-multiple-trend-lines-from-geom-smooth)
11. https://stackoverflow.com/questions/45515218/combine-rows-in-data-frame-containing-na-to-make-complete-row (https://stackoverflow.com/questions/45515218/combine-rows-in-data-frame-containing-na-to-make-complete-row)