NOAA Storm Database - worst cases

erickfis

2017 maio, 10

Contents

1 Introduction

In this study we have analysed the NOAA Storm Database in order to determine what are the worst natural catastrophic events, both in terms of public health and in economic impact.

The U.S. National Oceanic and Atmospheric Administration's (NOAA) storm database tracks characteristics of major storms and weather events in the United States, including when and where they occur, as well as estimates of any fatalities, injuries, and property damage.

The database currently contains data from January 1950 to January 2017, as entered by NOAA's National Weather Service (NWS).

The database can be found on:

https://www.ncdc.noaa.gov/stormevents/ftp.jsp

RPubs version, with fewer plots, for Coursera: http://rpubs.com/erickfis/noaa

2 Objective

The goal of this study is to answer the questions:

1. Across the United States, which types of events were the most harmful with respect to population health ever recorded in a single occurrence?
2. Which types of events caused most harm to population health along all those years?
3. Which types of events had the greatest economic consequences in a single occurrence?
4. Which types of events had the greatest economic consequences along all those years?
5. Which were the places that were subject to the greatest losses, both in terms of human health and economic losses.

3 Methods

To answer each one of those questions, we did a very simple **descriptive analysis** of data.

We used R tools to filter, sort and combine data, so we could get the total sum of fatalities, injuries and economic losses.

4 Data Processing

In order to answer our questions, the original database needed to be treated from its raw form to a more useful format.

4.0.1 New files to download and process data again:

```
StormEvents_details-ftp_v1.0_d2006_c20170405.csv.gz, StormEvents_details-ftp_v1.0_d2015_c20170216.cs
```

The necessary transformations were:

- sanitized var names
- evaluated duration of events
- evaluated damages values according to multipliers provided

- sanitized and grouped similar events: strong snow, heavy snow and light snow all became just “snow”
- sanitized county names

This database has 1419673 observations. Each observation corresponds to an event occurrence.

To determine the most harmful events to human health, we checked the variables related to human health, which are “fatalities” and “injuries”.

To determine the most harmful events to economy, we checked the variables related to economic measures, from “propdmg” through “cropdmgexp”.

Also, in order to analyse various occurrences of the same event, we measured the duration of the event, its magnitude and where the event occurred (state and county name).

This is a really big database whose data has been being registered by a lot of different people since 1950. Thus, as expected, there are variations on how people registered events.

For example, the string “snow” was used to register a lot of events. They are the same type of event, but count as different:

This is why we decided to filter those events: we grouped them by its common strings.

5 Human health: the most harmfull events

We have determined what events did more harm to human health.

There were occurrences that caused zero fatalities but a lot of injuries. The inverse is also true, so we did a separate analysis to fatal and non-fatal events.

5.1 Fatal Occurrences

5.1.1 Most fatal in a single occurrence

Most fatal in a single occurrence

In order to determine what were the most fatal events in a single occurrence, we need to see how fatalities are distributed along the occurrences.

Looking at this distribution, we can infer that the vast majority of those occurrences were not fatal at all: **99.2% occurrences didn’t caused any fatalities.**

On the other hand, fatal occurrences had to have at least 1 fatality.

Now, among the fatal occurrences, we are interested in the ones whose fatalities are beyond the confidence interval, ie. above 99% of the most common values.

Looking at this distribution, we can infer that **99.8% of the fatal occurrences caused up to 57.04 fattalities**.

Distribution plots

In this study, we looked on the 1% deadliest occurrences.

Table 1: Worst fatal occurrences

rank	event	magnitude	day	duration	state	countyname	fatalities
1	HURRICANE		2005-08-28	0S	LOUISIANA	ORLEANS	638
2	TORNADO		2011-05-22	0S	MISSOURI	JASPER	161
3	HURRICANE		2005-08-28	0S	LOUISIANA	LOWER.ST..BERNARD	140

rank	event	magnitude	day	duration	state	countyname	fatalities
4	TORNADO	0	1953-06-08	0S	MICHIGAN	GENESEE	116
5	TORNADO	0	1953-05-11	0S	TEXAS	MCLENNAN	114
6	HURRICANE		2005-08-28	0S	MISSISSIPPI	HARRISON	97
7	HEAT		1999-07-28	0S	ILLINOIS	COOK	93
8	TORNADO	0	1953-06-09	0S	MASSACHUSETTS	WORCESTER	90
9	TORNADO	0	1955-05-25	0S	KANSAS	COWLEY	75
10	HEAT		1999-07-04	0S	PENNSYLVANIA	PHILADELPHIA	58

The single most fatal event was a **HURRICANE, that occurred in LOUISIANA, ORLEANS, on 2005-08-28, killing 638 people.**

However, if we compare this single awful event to the mean of fatalities caused, we see that this is very unlikely to happen.

5.1.2 Most fatal in all time

Most fatal in all time

Notice that are several occurrences of the same type of event along the time.

Therefore, in order to know which is the worst type of event along all the years, we summed up the fatalities caused by each one of occurrences of this events.

Notice that we are interested only in the worst of them, ie, the ones which are above the mean.

Table 2: Total fatalities by event

rank	event	total	mean	median
1	TORNADO	5873	677.5714	184
2	HEAT	2854	677.5714	184
3	WIND	2304	677.5714	184
4	FLOOD	1944	677.5714	184
5	WINTER	1196	677.5714	184
6	HURRICANE	1128	677.5714	184
7	LIGHTNING	833	677.5714	184
8	RIP CURRENT	798	677.5714	184

The most fatal event along the time is the **TORNADO. It has killed 5873 people until now.**

5.1.3 Least fatal events

Just for curiosity, these are the less fatal among the fatal events:

Table 3: Least fatal events

rank	event	total
28	TROPICAL DEPRESSION	1
27	DENSE SMOKE	2
26	SLEET	2
25	WATERSPOUT	2

rank	event	total
24	COLD	4
23	DUST DEVIL	4
22	SLIDE	4
21	SNEAKERWAVE	14
20	HAIL	20
19	TIDE	22

5.2 Injuring Occurrences

5.2.1 Most injuring in a single occurrence

Most injuring in a single occurrence

In order to determine what were the most injuring events in a single occurrence, we need to see how injuries are distributed along the occurrences.

Looking at this distribution, we can infer that the vast majority of those occurrences were not injuring at all: **98.5% occurrences didn't caused any injuries**

On the other hand, injuring occurrences had to have at least 1 injury.

Now, among the injuring occurrences, we are interested in the ones whose harm is beyond the confidence interval, ie. above 99% of the most common values.

Looking at this distribution, we can infer that **99.8% of the injuring occurrences caused up to 500 injuries**.

Distribution plots

In this study, we looked on the 1% most injuring occurrences.

Table 4: Worst injuring occurrences

rank	event	magnitude	day	duration	state	countyname	injuries	mean
1	HURRICANE		2008-09-12	0S	TEXAS	HARRIS	2400	7.570462
2	TORNADO	0	1979-04-10	0S	TEXAS	WICHITA	1700	7.570462
3	TORNADO	0	1953-06-09	0S	MASSACHUSETTS	WORCESTER	1228	7.570462
4	TORNADO	0	1974-04-03	0S	OHIO	GREENE	1150	7.570462
5	TORNADO		2011-05-22	0S	MISSOURI	JASPER	1150	7.570462
6	FLOOD		1998-10-17	0S	TEXAS	COMAL	800	7.570462
7	TORNADO		2011-04-27	0S	ALABAMA	TUSCALOOSA	800	7.570462
8	TORNADO	0	1953-06-08	0S	MICHIGAN	GENESEE	785	7.570462
9	HURRICANE		2004-08-13	0S	FLORIDA	CHARLOTTE	700	7.570462
10	TORNADO		2011-04-27	0S	ALABAMA	JEFFERSON	700	7.570462
11	FLOOD		1998-10-17	0S	TEXAS	BEXAR	600	7.570462
12	TORNADO	0	1953-05-11	0S	TEXAS	MCLENNAN	597	7.570462
13	TORNADO	0	1965-04-11	0S	INDIANA	HOWARD	560	7.570462
14	HEAT		2007-08-04	0S	MISSOURI	ST..LOUIS	519	7.570462
15	TORNADO	0	1966-03-03	0S	MISSISSIPPI	HINDS	504	7.570462

The single most injuring event was a **HURRICANE, that occurred in TEXAS, HARRIS, on 2008-09-12, injuring 2400 people.**

However, if we compare this single awful event to the mean of injuries caused, we see that this is very unlikely

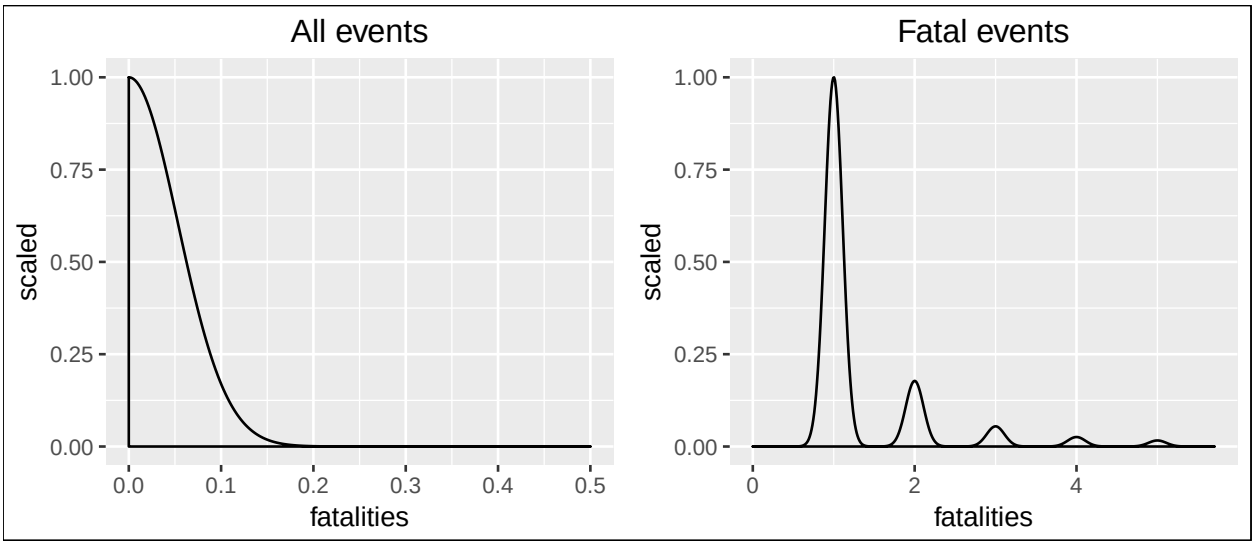

Figure 1: Population distribution for fatalities / occurrences

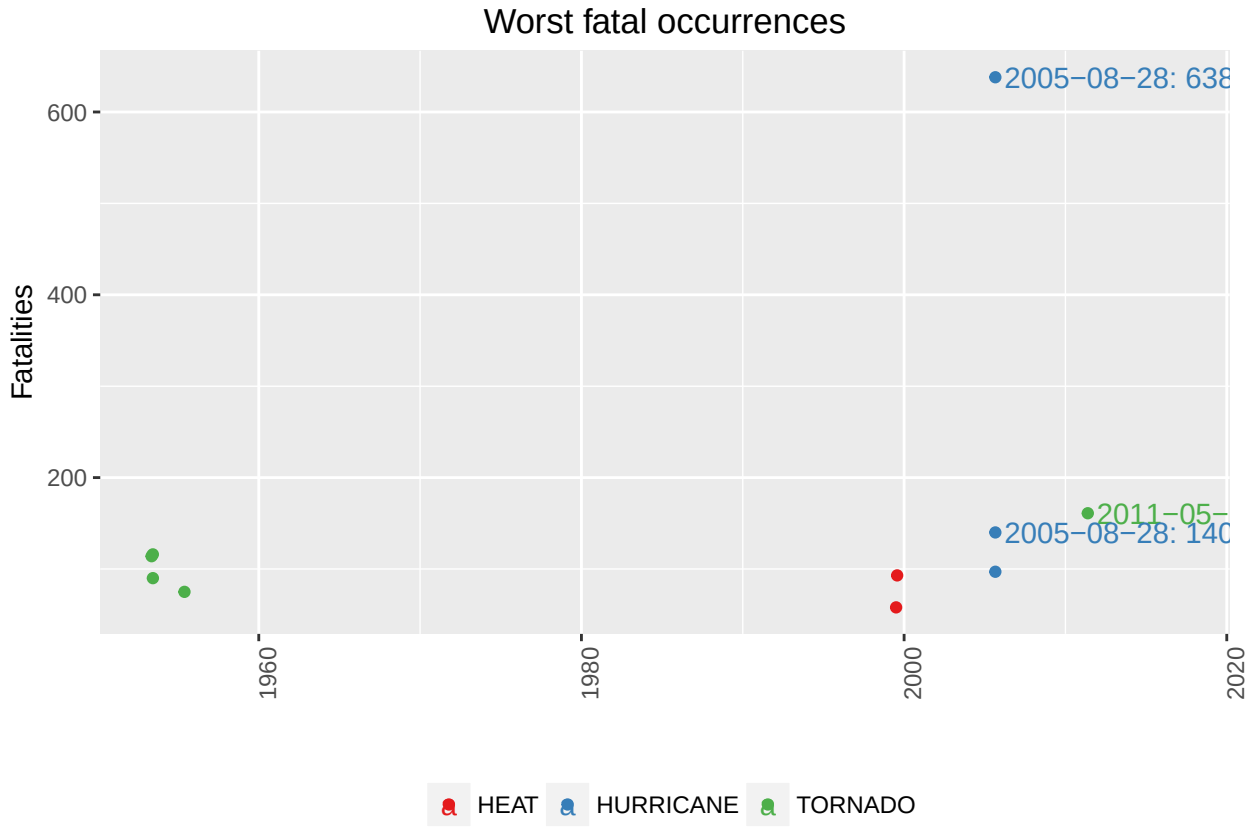

Figure 2: Worst fatal occurrences

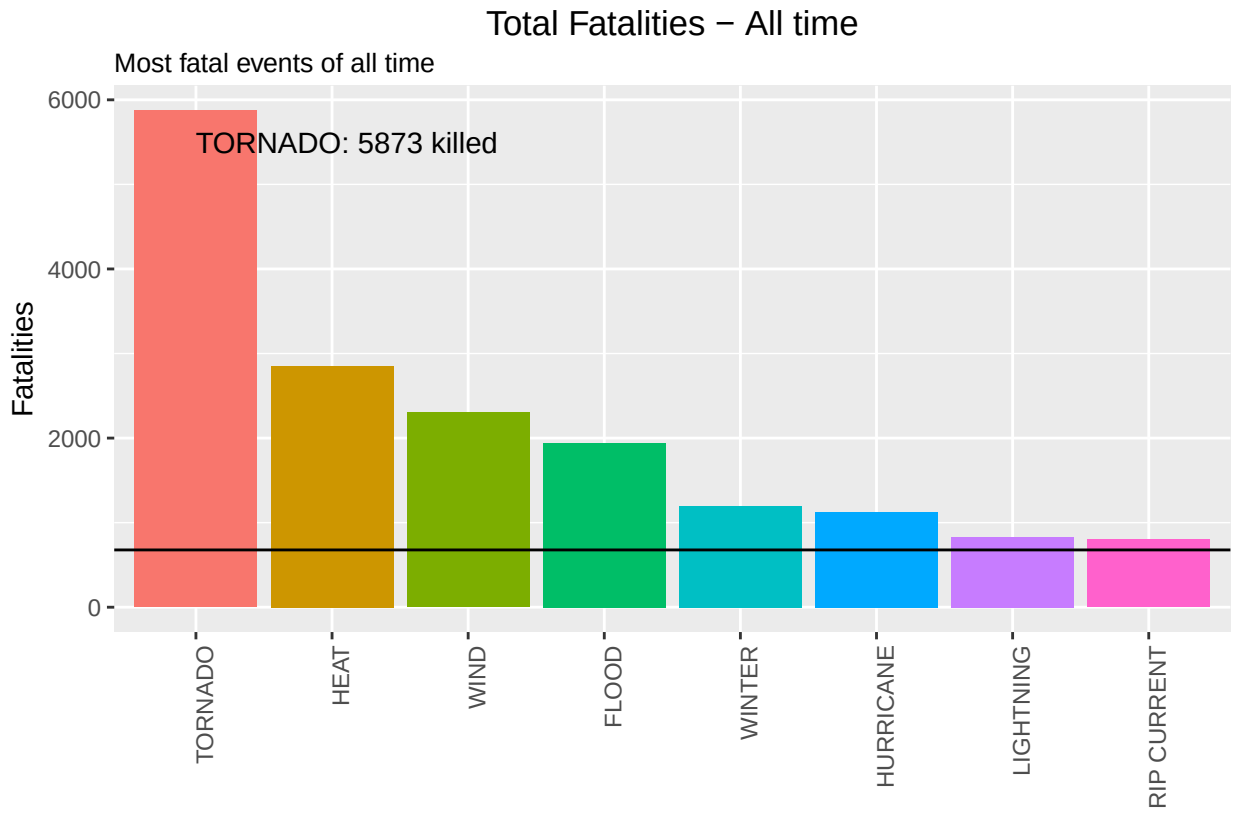

Figure 3: Total fatalities by event

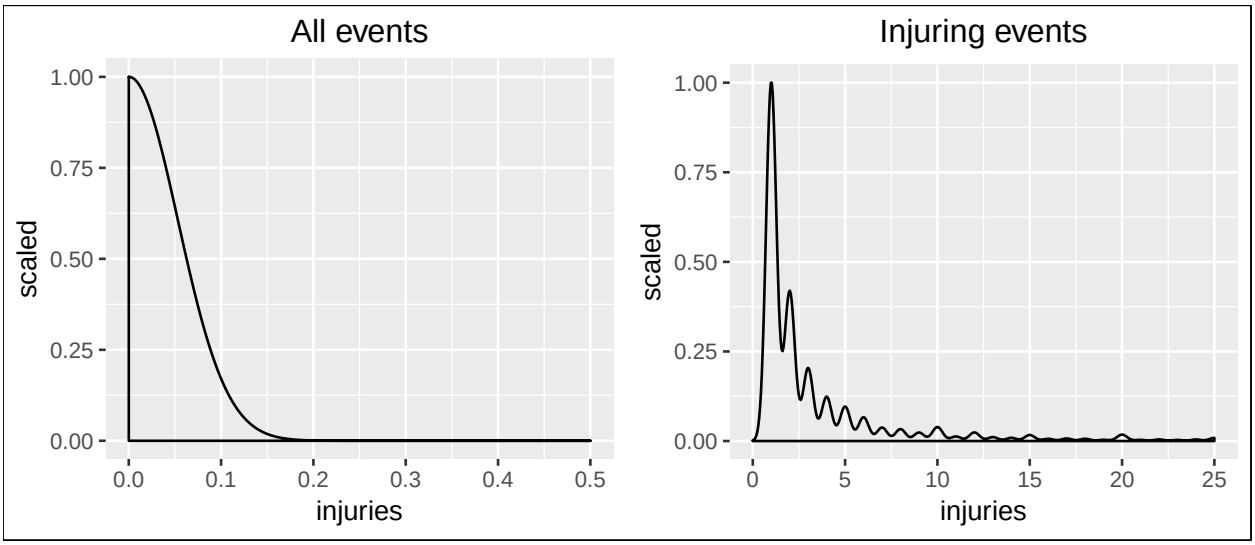

Figure 4: Population distribution for Injuries / occurrences

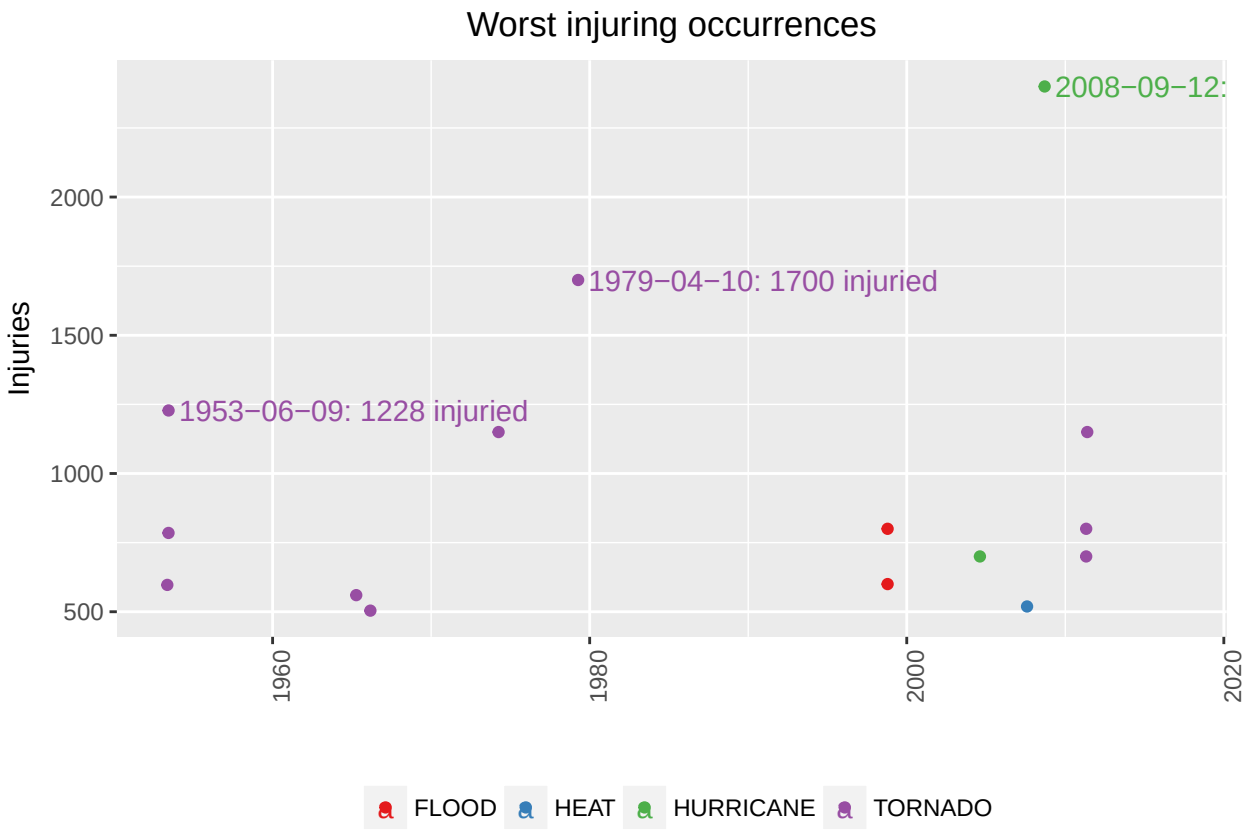

Figure 5: Worst injuring occurrences

to happen.

5.2.2 Most injuring in all time

Most injuring in all time

Notice that are several occurrences of the same type of event along the time.

Therefore, in order to know which is the worst type of event along all the years, we summed up the injuries caused by each one of occurrences of this events.

Notice that we are interested only in the worst of them, ie, the ones which are above the mean.

Table 5: Total injuries by event

rank	event	total	mean	median
1	TORNADO	94614	5200.419	316
2	HEAT	15436	5200.419	316
3	WIND	13439	5200.419	316
4	FLOOD	8806	5200.419	316
5	WINTER	8224	5200.419	316

The most injuring event along the time is the **TORNADO. It has injuried 94614 people until now.**

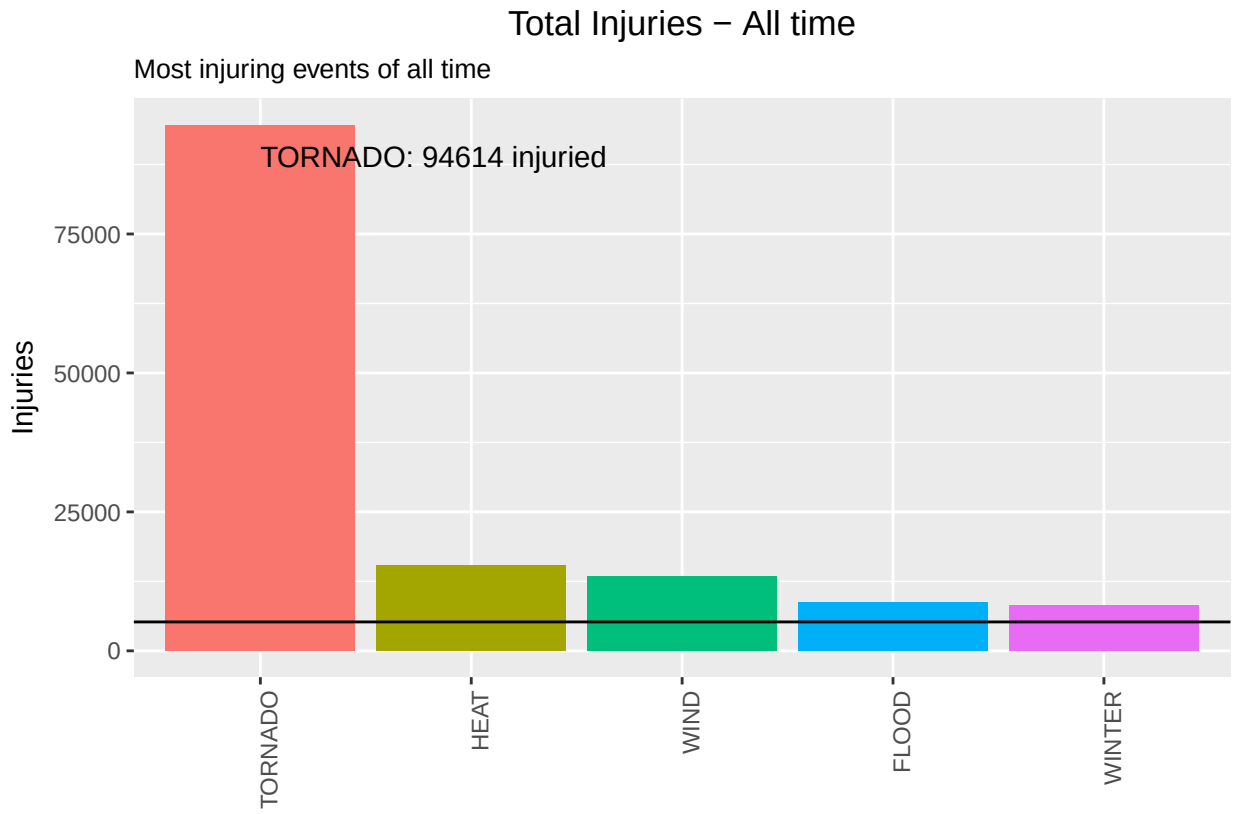

Figure 6: Total Injuries by event

5.2.3 Least injuring events

Just for curiosity, lets show now what are the less injuring among the injuring events:

Table 6: Least injuring events

rank	event	total
31	FUNNEL	3
30	TROPICAL DEPRESSION	3
29	WATERSPOUT	3
28	OTHER	4
27	DROUGHT	8
26	SLEET	10
25	SNEAKERWAVE	11
24	SLIDE	13
23	COLD	15
22	DENSE SMOKE	17

6 Economy: the the most harmfull events

We have determined what events did more harm to economy, both in terms of property and crops damage.

There were events that causes zero property damage but a lot of crop damage. The inverse is also true, so we did a separate analysis to property VS crop damaging events.

6.1 Property losses

6.1.1 Most Property Damaging event in a single occurrence

In order to determine what were the most property damaging events in a single occurrence, we need to see how damages are distributed along the occurrences.

Looking at this distribution, we can infer that 99.8% of the occurrences caused less than **$40,000,000 in losses**.

On the other hand, damaging occurrences had to have damages above zero.

Now, among the damaging occurrences, we are interested in the ones whose damages are above 99.8% of the most common values.

Looking at this distribution, we can infer that **99.8% of the damaging occurrences caused up to $128,960,700 in losses**.

Distribution plots

In this study, we looked on the 1% most harmful occurrences.

Table 7: Worst property damaging occurrences, mean = $1,137,697 and median = $10,000

rank	event	magnitude	day	duration	state	countyname	value
1	TIDE		2005-08-29	0S	LOUISIANA	ORLEANS	$17,900
2	HURRICANE		2005-10-24	0S	FLORIDA	COASTAL.PALM.BEACH	$10,000
3	FLOOD		2012-10-29	0S	NEW JERSEY	EASTERN.OCEAN	$7,500,0

rank	event	magnitude	day	duration	state	countyname	value
4	TIDE		2005-08-29	0S	MISSISSIPPI	HARRISON	$5,630,0
5	STORM		2001-06-05	0S	TEXAS	HARRIS	$5,030,0
6	FLOOD		2012-10-28	0S	NEW JERSEY	EASTERN.MONMOUTH	$5,000,0
7	FLOOD		2012-10-28	0S	NEW JERSEY	WESTERN.MONMOUTH	$5,000,0
8	HURRICANE		2004-09-13	0S	FLORIDA	COASTAL.ESCAMBIA	$4,000,0
9	TIDE		2008-09-12	0S	TEXAS	GALVESTON	$4,000,0
10	HURRICANE		2005-08-28	0S	LOUISIANA	ORLEANS	$3,560,0
11	TIDE		2005-08-29	0S	MISSISSIPPI	HANCOCK	$3,380,0
12	TIDE		2005-08-29	0S	LOUISIANA	ST..TAMMANY	$3,030,0
13	TIDE		2005-08-29	0S	LOUISIANA	LOWER.PLAQUEMINES	$3,030,0
14	TIDE		2005-08-29	0S	LOUISIANA	LOWER.ST..BERNARD	$3,020,0
15	TIDE		2005-08-29	0S	LOUISIANA	UPPER.ST..BERNARD	$3,020,0
16	FLOOD		1997-04-18	0S	NORTH DAKOTA	GRAND.FORKS	$3,000,0
17	HURRICANE		1999-09-15	0S	NORTH CAROLINA	ALAMANCE	$3,000,0
18	HURRICANE		2004-08-13	0S	FLORIDA	CHARLOTTE	$3,000,0
19	TIDE		2008-09-12	0S	TEXAS	HARRIS	$3,000,0
20	HURRICANE		2005-08-28	0S	MISSISSIPPI	HARRISON	$2,940,0

The single most economic damaging event to properties was a **TIDE, that occurred in LOUISIANA, ORLEANS, on 2005-08-29, causing U$ $17,900,000,000 in losses**.

6.1.2 Most Property Damaging event in all time

Notice that are several occurrences of the same type of event along the time.

Therefore, in order to know which is the worst type of event along all the years, we summed up the losses caused by each one of occurrences of this events.

Notice that we are interested only in the worst of them, ie, the ones which are above the mean.

Table 8: Total losses by event

rank	event	total	mean	median
1	HURRICANE	$87,005,170,310	$11,737,195,233	$229,971,300
2	FLOOD	$82,920,597,380	$11,737,195,233	$229,971,300
3	TORNADO	$63,648,478,192	$11,737,195,233	$229,971,300
4	TIDE	$54,155,102,600	$11,737,195,233	$229,971,300
5	HAIL	$25,360,544,274	$11,737,195,233	$229,971,300
6	WIND	$24,868,014,778	$11,737,195,233	$229,971,300
7	STORM	$16,753,590,360	$11,737,195,233	$229,971,300

The most property damaging event along the time is the **HURRICANE. It has caused $87,005,170,310 in losses.**

6.1.3 Least property damaging events

Just for curiosity, these are the less damaging events:

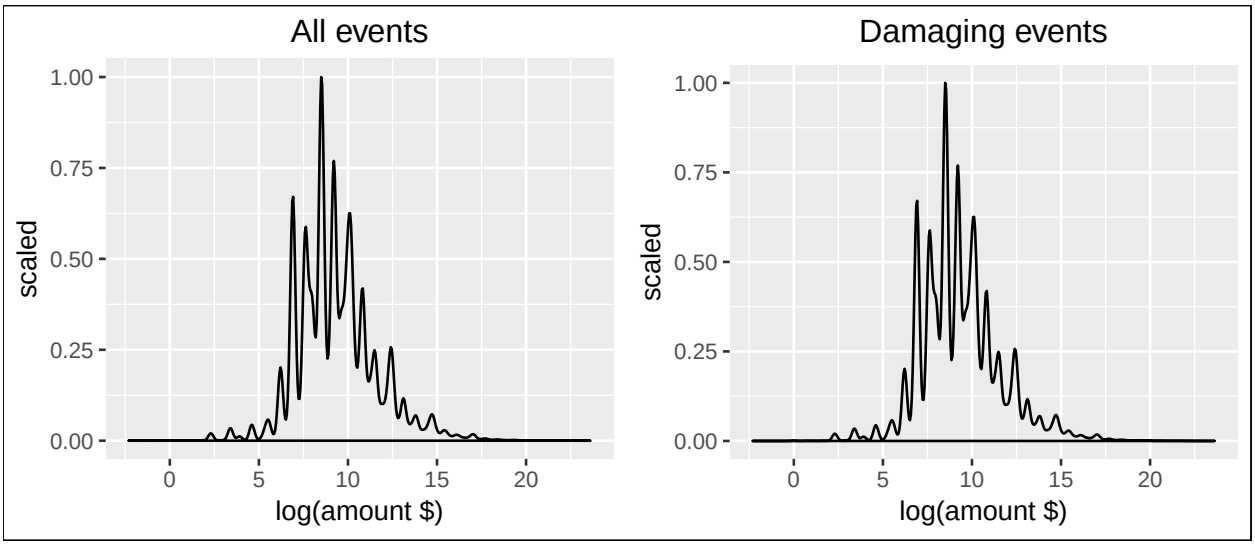

Figure 7: Population distribution for losses / occurrences

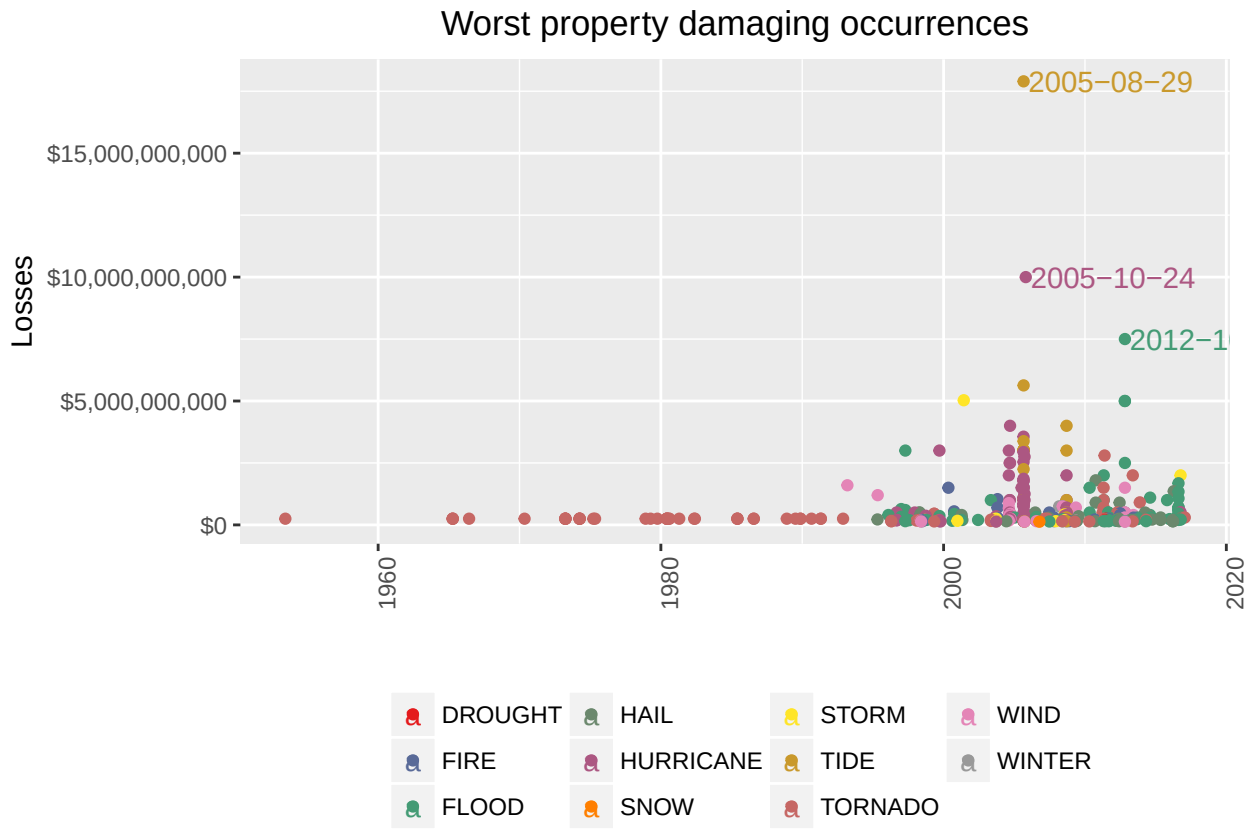

Figure 8: Worst property damaging occurrences

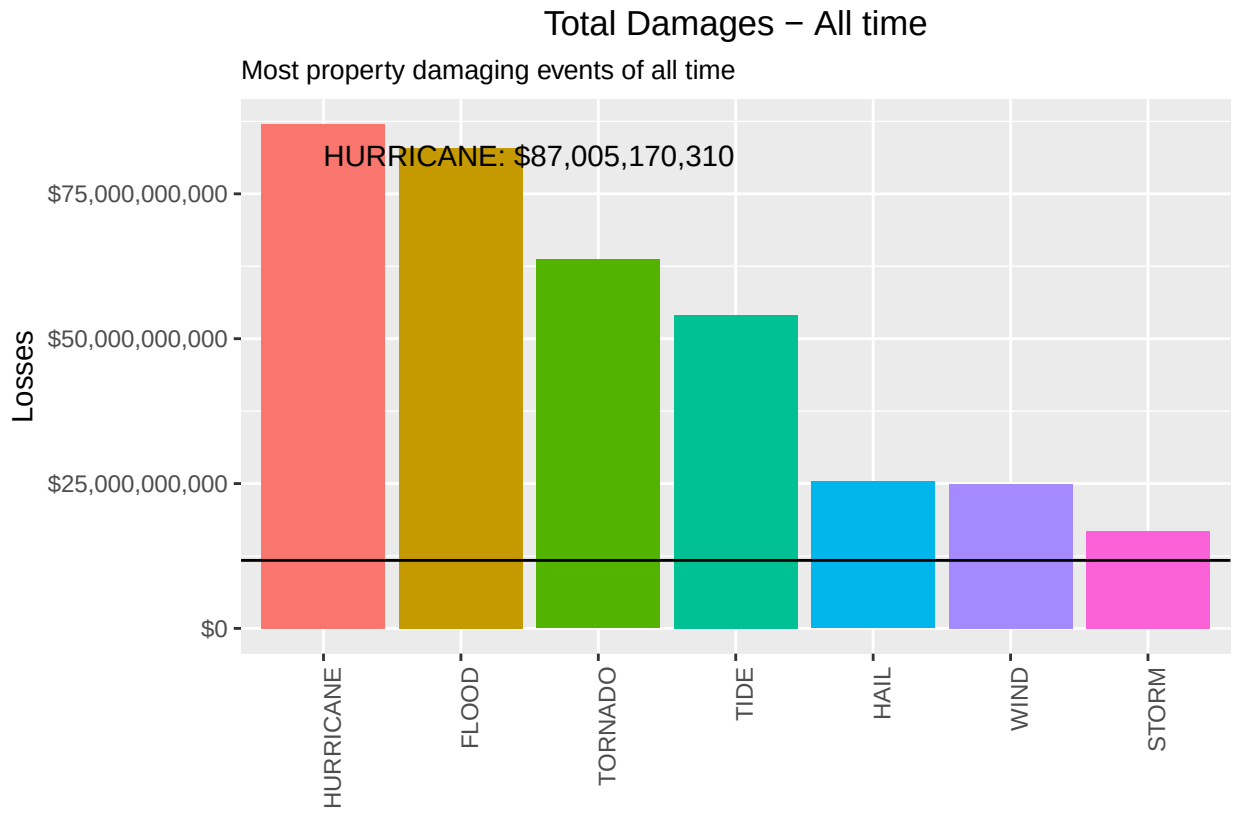

Figure 9: Total Property Damages by event

Table 9: Least property damaging events

rank	event	total
32	OTHER	$1,000
31	FUNNEL	$123,100
30	DENSE SMOKE	$130,000
29	RIP CURRENT	$163,000
28	VOLCANIC ASH	$500,000
27	DUST DEVIL	$1,147,430
26	SEICHE	$1,402,000
25	SLEET	$3,084,000
24	AVALANCHE	$4,058,050
23	WATERSPOUT	$5,748,200

6.2 Crop losses

6.2.1 Most Crop Damaging event in a single occurrence

In order to determine what were the most crop damaging events in a single occurrence, we need to see how damages are distributed along the occurrences.

On the other hand, damaging occurrences had to have damages above zero.

Now, among the damaging occurrences, we are interested in the ones whose damages are above 99% of the most common values.

Looking at this distribution, we can infer that **99.8% of the damaging occurrences caused up to $197,450,000 in losses.**

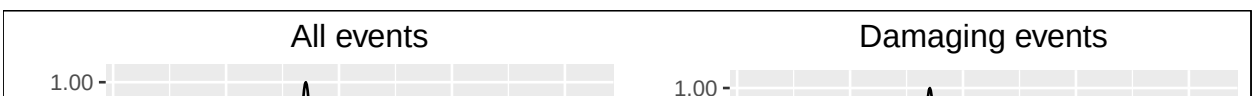

Table 10: Worst crops damaging occurrences, mean = $1,783,319 and median = $20,000

rank	event	magnitude	day	duration	state	countyname
1	DROUGHT		2014-12-01	0S	CALIFORNIA	NORTHERN.SAN.JOAQUIN.VALLE
2	DROUGHT		2011-06-01	0S	TEXAS	LUBBOCK
3	DROUGHT		2006-01-01	0S	TEXAS	MONTAGUE
4	DROUGHT		2007-06-01	0S	MISSISSIPPI	WARREN
5	COLD		2007-01-11	0S	CALIFORNIA	SE.S.J..VALLEY
6	DROUGHT		2000-11-01	0S	TEXAS	PARMER
7	DROUGHT		1998-07-06	0S	OKLAHOMA	CHOCTAW
8	DROUGHT		1999-07-01	0S	PENNSYLVANIA	POTTER
9	HURRICANE		1999-09-15	0S	NORTH CAROLINA	ALAMANCE
10	FLOOD		2000-10-03	0S	FLORIDA	COASTAL.DADE
11	FLOOD		2007-07-01	0S	MISSOURI	HENRY
12	WIND		1998-12-20	0S	CALIFORNIA	SOUTHERN.SAN.JOAQUIN.VALLE
13	DROUGHT		1998-12-01	0S	TEXAS	YOAKUM
14	HURRICANE		2005-08-25	0S	FLORIDA	COASTAL.DADE
15	DROUGHT		2001-12-01	0S	TEXAS	PARMER
16	DROUGHT		2007-09-01	0S	GEORGIA	BALDWIN
17	DROUGHT		2006-02-01	0S	TEXAS	FANNIN
18	COLD		2010-01-10	0S	FLORIDA	INLAND.COLLIER.COUNTY
19	COLD		2010-01-10	0S	FLORIDA	INLAND.MIAMI.DADE
20	DROUGHT		1998-12-01	0S	TEXAS	ANDREWS

The single most economic damaging event to crops was a **DROUGHT, that occurred in CALIFORNIA, NORTHERN.SAN.JOAQUIN.VALLEY, on 2014-12-01, causing U$ $1,500,000,000 in losses.**

6.2.2 Most Crop Damaging event in all time

Notice that are several occurrences of the same type of event along the time.

Therefore, in order to know which is the worst type of event along all the years, we summed up the losses caused by each one of occurrences of this events.

Notice that we are interested only in the worst of them, ie, the ones which are above the mean.

Table 11: Total losses by event

rank	event	total	mean	median
1	DROUGHT	$27,454,862,620	$2,957,587,920	$450,448,110
2	FLOOD	$7,750,252,370	$2,957,587,920	$450,448,110
3	HURRICANE	$5,341,874,800	$2,957,587,920	$450,448,110
4	COLD	$4,919,893,200	$2,957,587,920	$450,448,110
5	WIND	$3,679,520,230	$2,957,587,920	$450,448,110
6	HAIL	$3,657,650,043	$2,957,587,920	$450,448,110

The most crop damaging event along the time is the **DROUGHT. It has caused $27,454,862,620 in losses.**

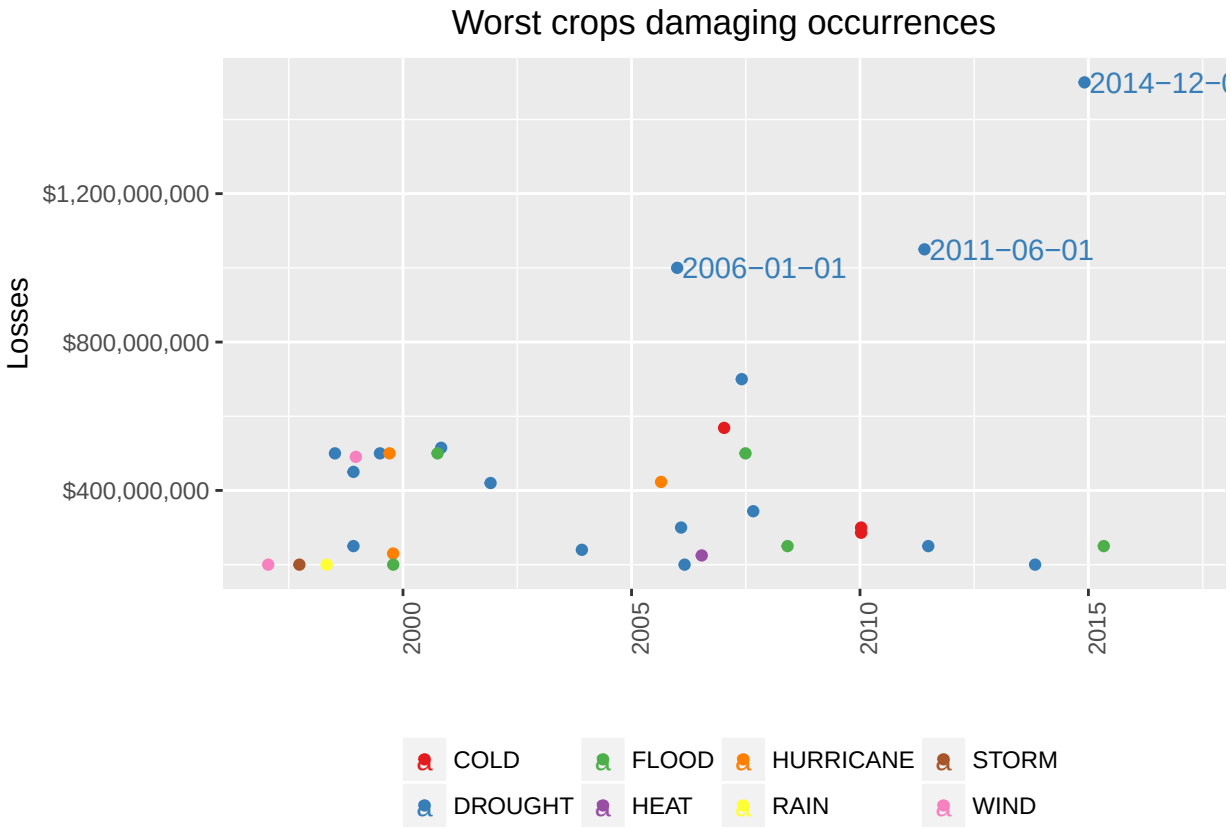

Figure 11: Worst crops damaging occurrences

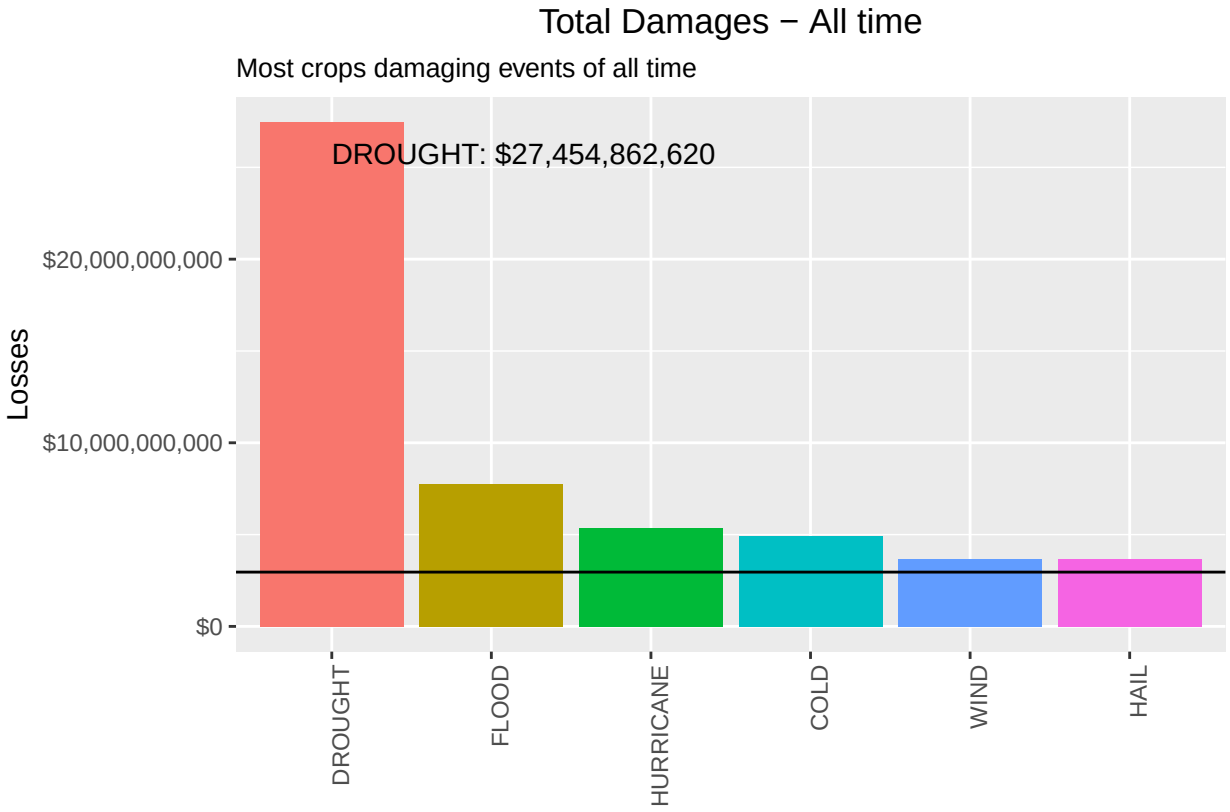

Figure 12: Total Crop Damages by event

6.2.3 Least crops damaging events

Just for curiosity, lets show now what are the less damaging among the events:

Table 12: Least crops damaging events

rank	event	total
19	SLIDE	$17,000
18	TSUNAMI	$20,000
17	TIDE	$955,000
16	BLIZZARD	$7,060,000
15	LIGHTNING	$7,422,640
14	DEBRIS FLOW	$20,001,500
13	WINTER	$46,924,000
12	SNOW	$91,145,900
11	FIRE	$447,668,860
10	TORNADO	$450,448,110

7 Most aflicted locations

We have determined what locations had the worst outcome from those events, both in terms of human health and economic losses.

Unfortunatelly, these has been the worst counties for living in:

7.1 Worst fatality count

Table 13: Total fatalities by county

Table 14: Total injuries by county

rank	state	countyname	fatalities	injuries	prop.dmg	crop.dmg
1	MISSOURI	ST..LOUIS	65	3144	$1,461,882,880	$10,500
2	TEXAS	HARRIS	216	2825	$10,890,439,870	$7,442,000
3	MISSOURI	ST..LOUIS.	118	2701	$79,552,000	$5,000
4	TEXAS	WICHITA	55	1853	$310,822,880	$0
5	TEXAS	DALLAS	149	1757	$1,946,192,730	$1,405,000
6	ALABAMA	JEFFERSON	117	1699	$2,037,082,100	$3,355,000
7	MASSACHUSETTS	WORCESTER	96	1292	$286,072,530	$0
8	OHIO	GREENE	40	1278	$289,967,757	$540,000
9	MISSOURI	JASPER	178	1273	$2,864,021,330	$46,475,500
10	OKLAHOMA	OKLAHOMA	79	1253	$1,356,088,290	$8,330,000

The county with the biggest injuries count is **ST..LOUIS, in MISSOURI, with 3144 people injuried.**

7.3 Worst property losses

Table 15: Total property losses by county

rank	state	countyname	fatalities	injuries	prop.dmg	crop.dmg
1	LOUISIANA	ORLEANS	649	99	$21,614,049,550	$0
2	TEXAS	HARRIS	216	2825	$10,890,439,870	$7,442,000
3	FLORIDA	COASTAL.PALM.BEACH	5	7	$10,828,630,000	$75,000,000
4	MISSISSIPPI	HARRISON	110	90	$8,870,659,460	$0
5	NEW JERSEY	EASTERN.OCEAN	16	112	$8,116,441,690	$10
6	NEW JERSEY	EASTERN.MONMOUTH	13	397	$6,527,278,550	$0
7	LOUISIANA	ST..TAMMANY	8	89	$5,677,622,950	$0
8	FLORIDA	COASTAL.ESCAMBIA	14	0	$5,632,695,000	$25,300,000
9	TEXAS	GALVESTON	43	259	$5,358,909,770	$109,602,000
10	NEW JERSEY	WESTERN.MONMOUTH	6	84	$5,267,488,450	$0

The county with the biggest property losses is **ORLEANS, in LOUISIANA, with $21,614,049,550 in losses.**

7.4 Worst crops losses

Table 16: Total crops losses by county

rank	state	countyname	fatalities	injuries	prop.dmg	crop.dmg
1	TEXAS	LUBBOCK	37	679	$2,007,426,360	$2,439,945,(
2	TEXAS	MONTAGUE	5	42	$118,971,700	$1,963,106,(
3	CALIFORNIA	NORTHERN.SAN.JOAQUIN.VALLEY	14	25	$5,948,500	$1,520,000,(
4	TEXAS	PARMER	1	23	$44,654,090	$1,181,360,(
5	FLORIDA	COASTAL.DADE	10	1	$693,020,000	$1,168,000,(
6	CALIFORNIA	SE.S.J..VALLEY	25	64	$5,742,300	$992,123,00
7	MISSISSIPPI	WARREN	43	341	$312,674,880	$728,657,00
8	CALIFORNIA	E.CENTRAL.S.J..VALLEY	43	73	$6,936,800	$547,162,00

rank	state	countyname	fatalities	injuries	prop.dmg	crop.dmg
9	CALIFORNIA	SOUTHERN.SAN.JOAQUIN.VALLEY	1	22	$18,657,000	$517,800,00
10	NORTH CAROLINA	ALAMANCE	2	8	$3,005,157,200	$503,166,00

The county with the biggest croperty losses is **LUBBOCK, in TEXAS, with $2,439,945,000 in losses.**

8 Results

8.1 Population Health

The single most fatal event was a **HURRICANE, that occurred in LOUISIANA, ORLEANS, on 2005-08-28, killing 638 people.**

The most fatal event along the time is the **TORNADO. It has killed 5873 people until now.**

The single most injuring event was a **HURRICANE, that occurred in TEXAS, HARRIS, on 2008-09-12, injuring 2400 people.**

The most injuring event along the time is the **TORNADO. It has injuried 94614 people until now.**

8.2 Economic Damages

The single most economic damaging event to properties was a **TIDE, that occurred in LOUISIANA, ORLEANS, on 2005-08-29, causing U$ $17,900,000,000 in losses**.

The most property damaging event along the time is the **HURRICANE. It has caused $87,005,170,310 in losses.**

The single most economic damaging event to crops was a **DROUGHT, that occurred in CALIFORNIA, NORTHERN.SAN.JOAQUIN.VALLEY, on 2014-12-01, causing U$ $1,500,000,000 in losses**.

The most crop damaging event along the time is the **DROUGHT. It has caused $27,454,862,620 in losses.**

8.3 Most aflicted locations

The county with the biggest fatality count is **ORLEANS, in LOUISIANA, with 649 people killed.**

The county with the biggest injuries count is **ST..LOUIS, in MISSOURI, with 3144 people injuried.**

The county with the biggest property losses is **ORLEANS, in LOUISIANA, with $21,614,049,550 in losses.**

The county with the biggest croperty losses is **LUBBOCK, in TEXAS, with $2,439,945,000 in losses.**

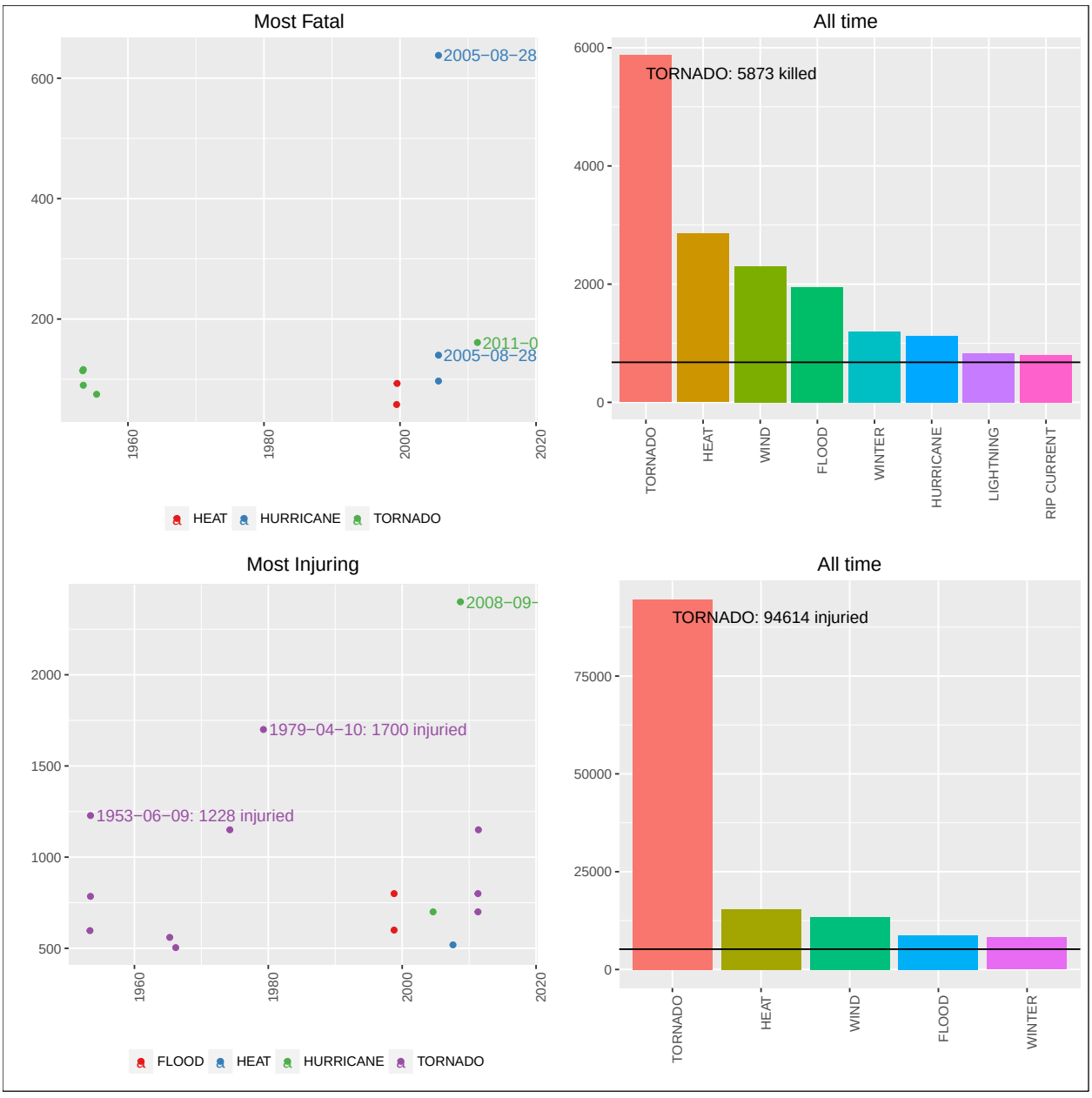

Figure 13: Population Health: fatalities and injuries

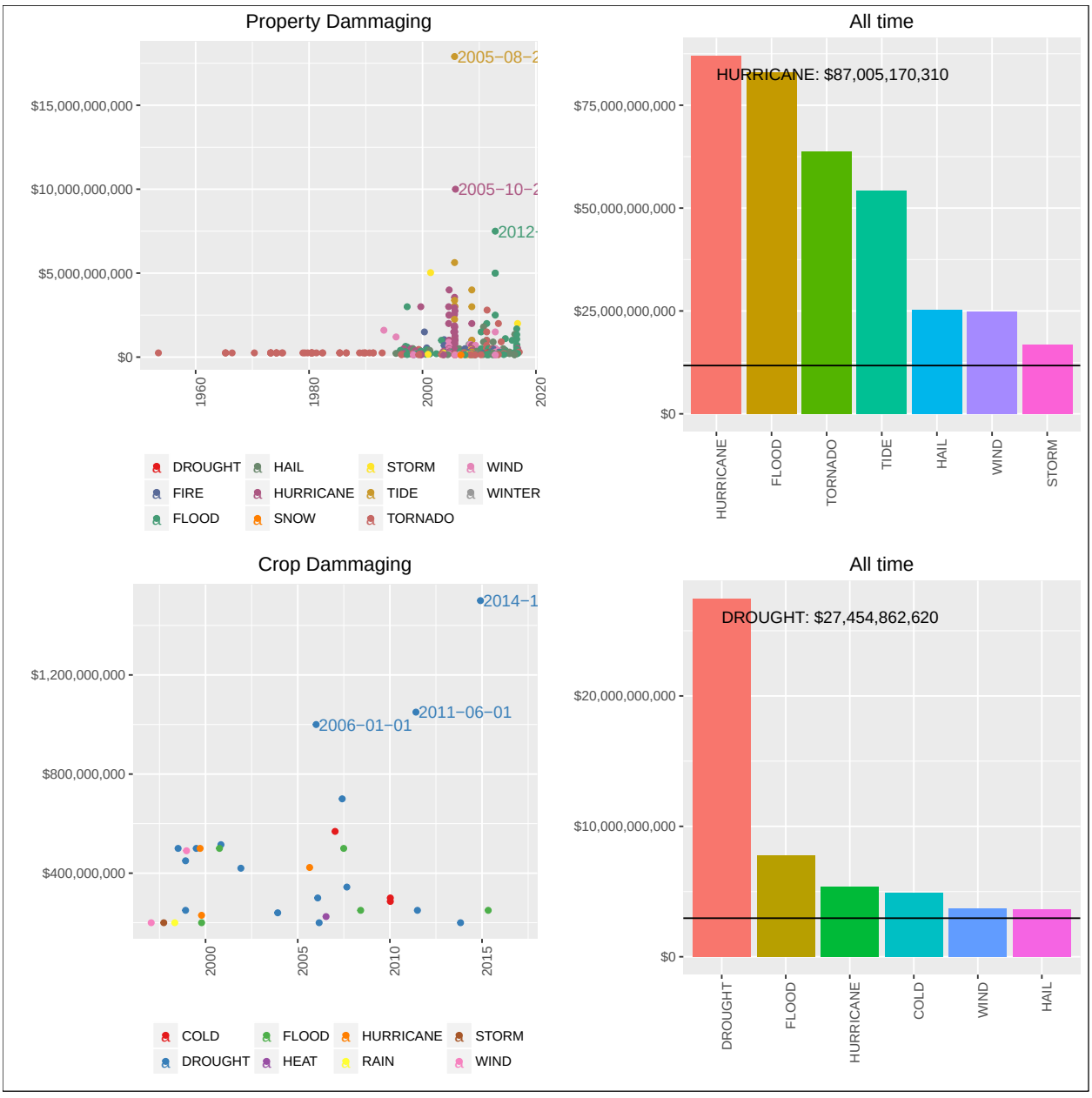

Figure 14: Economic Damages: property and crops

8.4 Distribution of data