NOAA Storm Database - worst cases

erickfis

2017 maio, 11

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: http://rpubs.com/erickfis/noaa

GitHub version, with code included and pdf version: https://github.com/erickfis/NOAA-Storm-Database

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.

The necessary transformations were:

- sanitized var names
- evaluated duration of events, however they are not useful
- 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 48875 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: **99.3% 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 21.058 fattalities**.

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

Table 1: Worst fatal occurrences, mean = 1.49 and median = 1

rank	event	day	state	county	fatalities
1	storm	2001-06-05	texas	harris	22

The single most fatal event was a **storm, that occurred in texas, harris, on 2001-06-05, killing 22 people.**

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

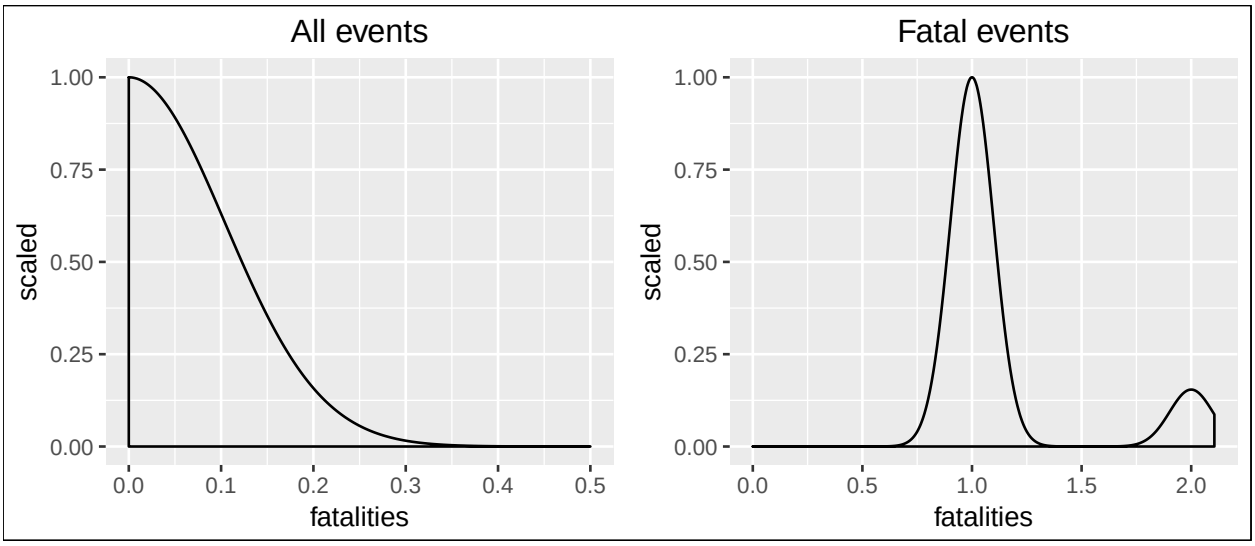

Figure 1: Population distribution for fatalities / occurrences

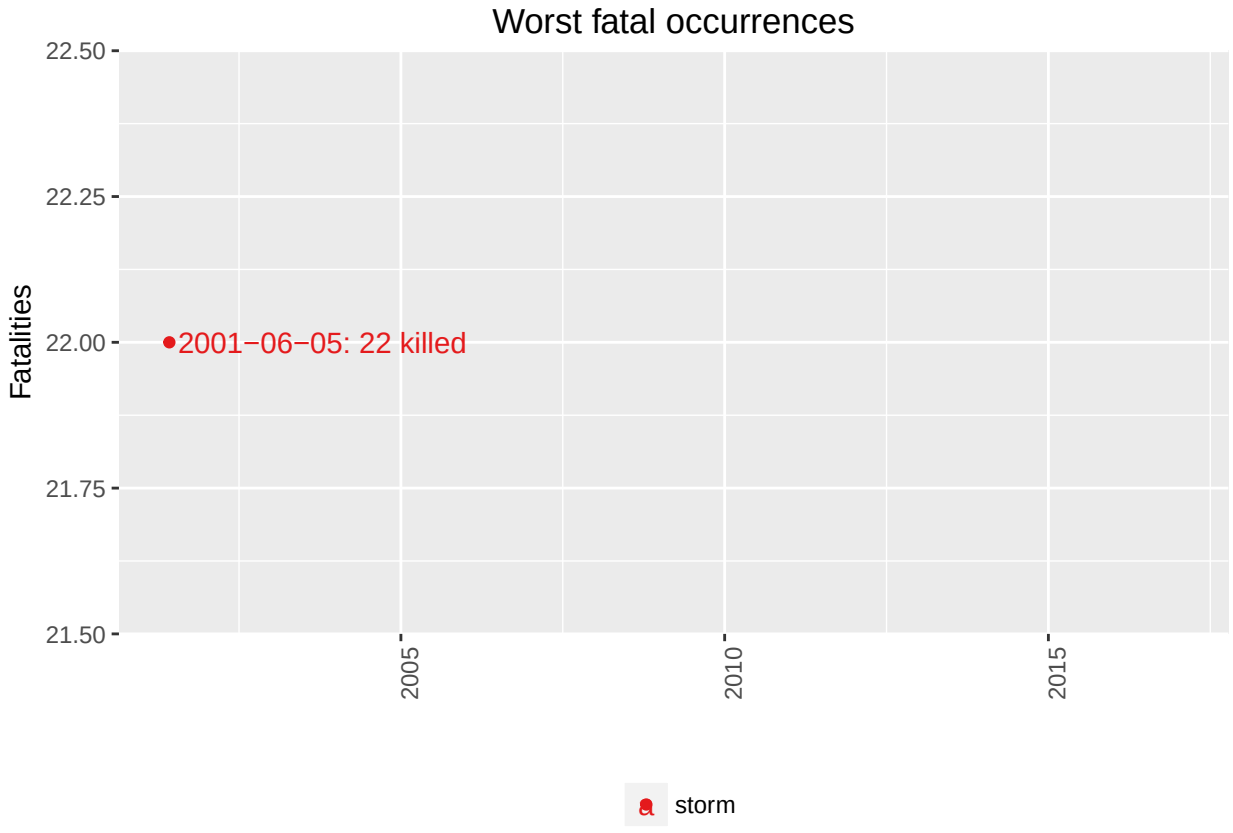

Figure 2: Worst fatal occurrences

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, mean = 29.31 and median = 12.5

rank	event	total
1	heat	166
2	flood	52
3	rip current	49
4	lightning	44
5	tornado	40
6	wind	36

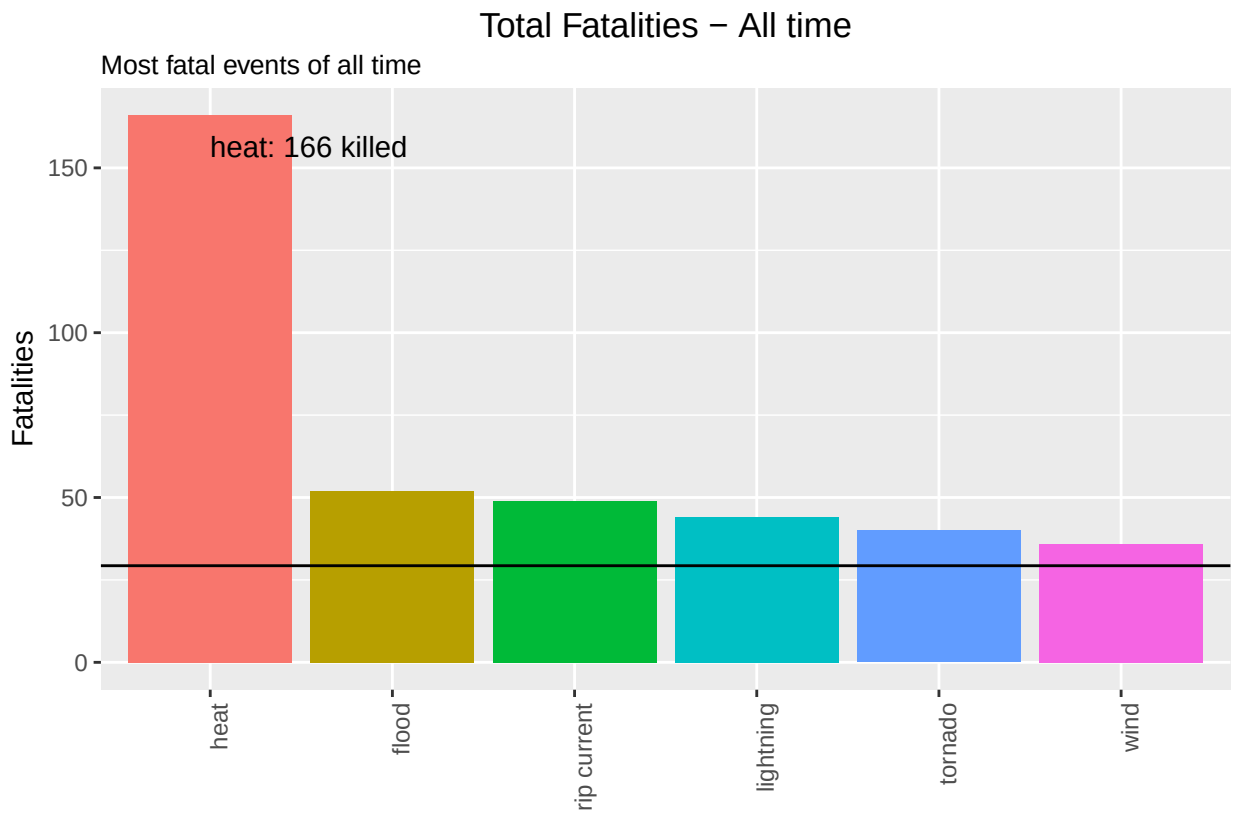

Figure 3: Total fatalities by event

The most fatal event along the time is the **heat. It has killed 166 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
16	hurricane	1
15	snow	2
14	rain	3
13	blizzard	4
12	fire	5
11	surf	6
10	fog	7
9	winter	7
8	avalanche	18
7	storm	29

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 88.694 injuries**.

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

Table 4: Worst injuring occurrences, mean = 4.01 and median = 2

rank	event	day	state	county	injuries
1	heat	2001-08-06	michigan	wayne	107

The single most injuring event was a **heat, that occurred in michigan, wayne, on 2001-08-06, injuring 107 people.**

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

5.2.2 Most injuring in all time

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

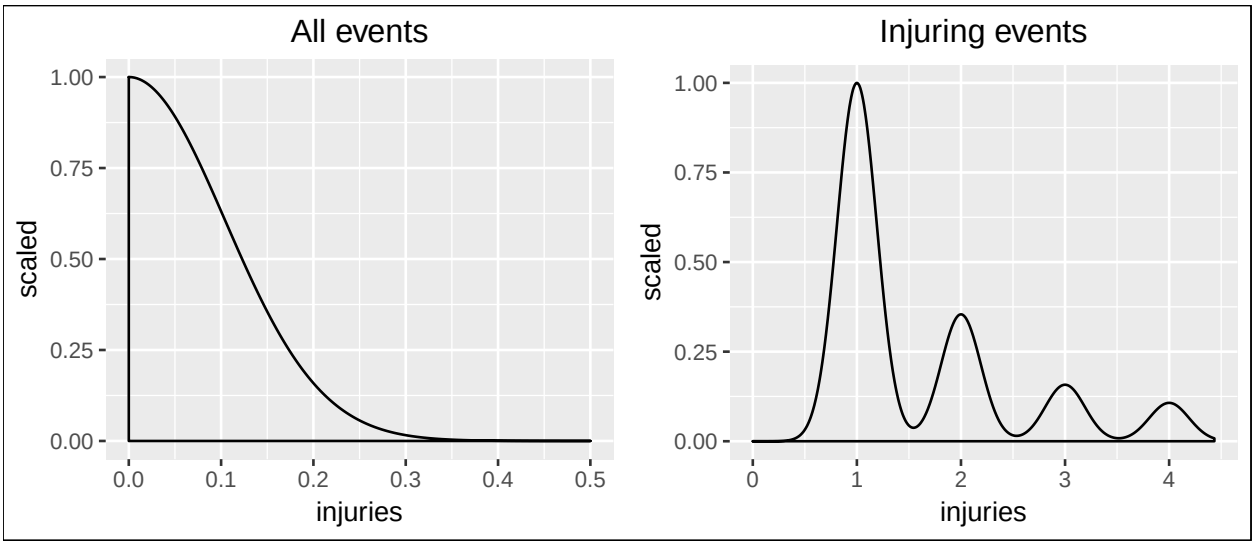

Figure 4: Population distribution for Injuries / occurrences

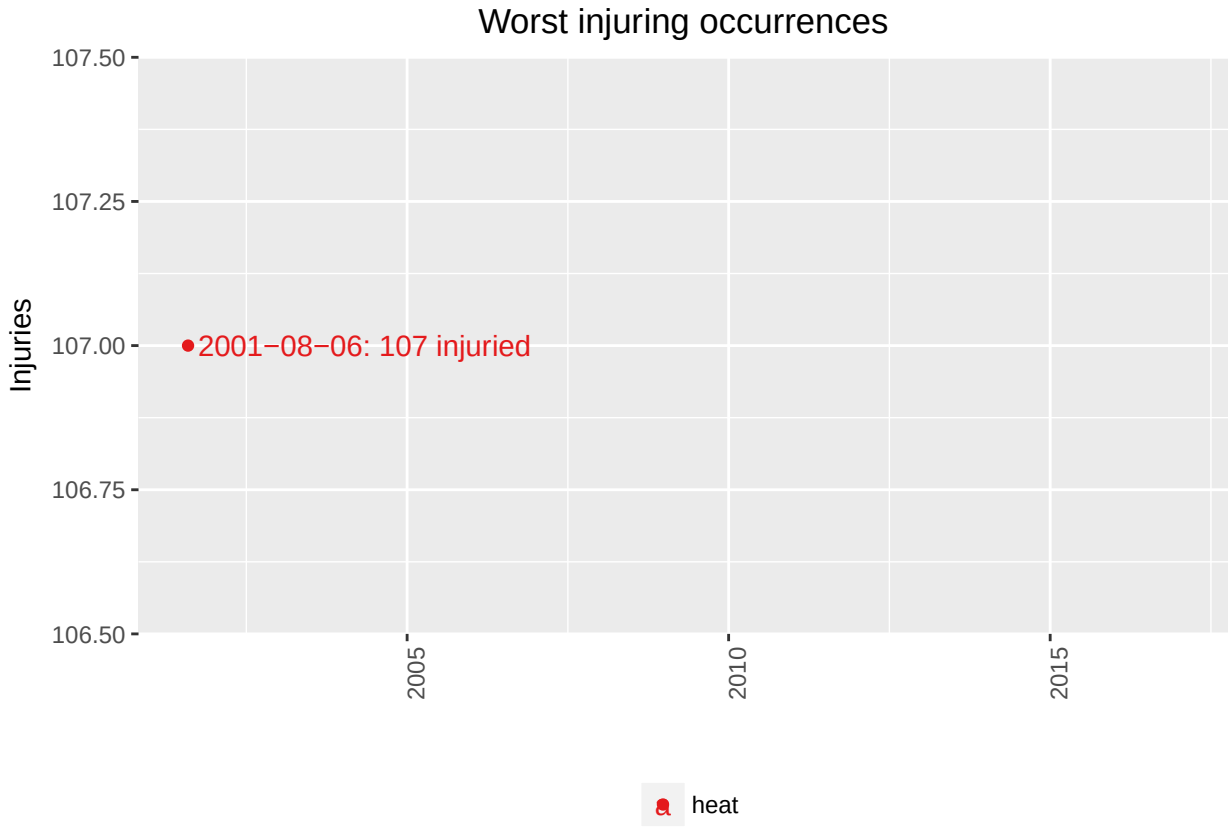

Figure 5: Worst injuring occurrences

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, mean = 143.21 and median = 32

rank	event	total
1	tornado	743
2	heat	445
3	wind	441
4	lightning	372
5	flood	277

The most injuring event along the time is the **tornado. It has injuried 743 people until now.**

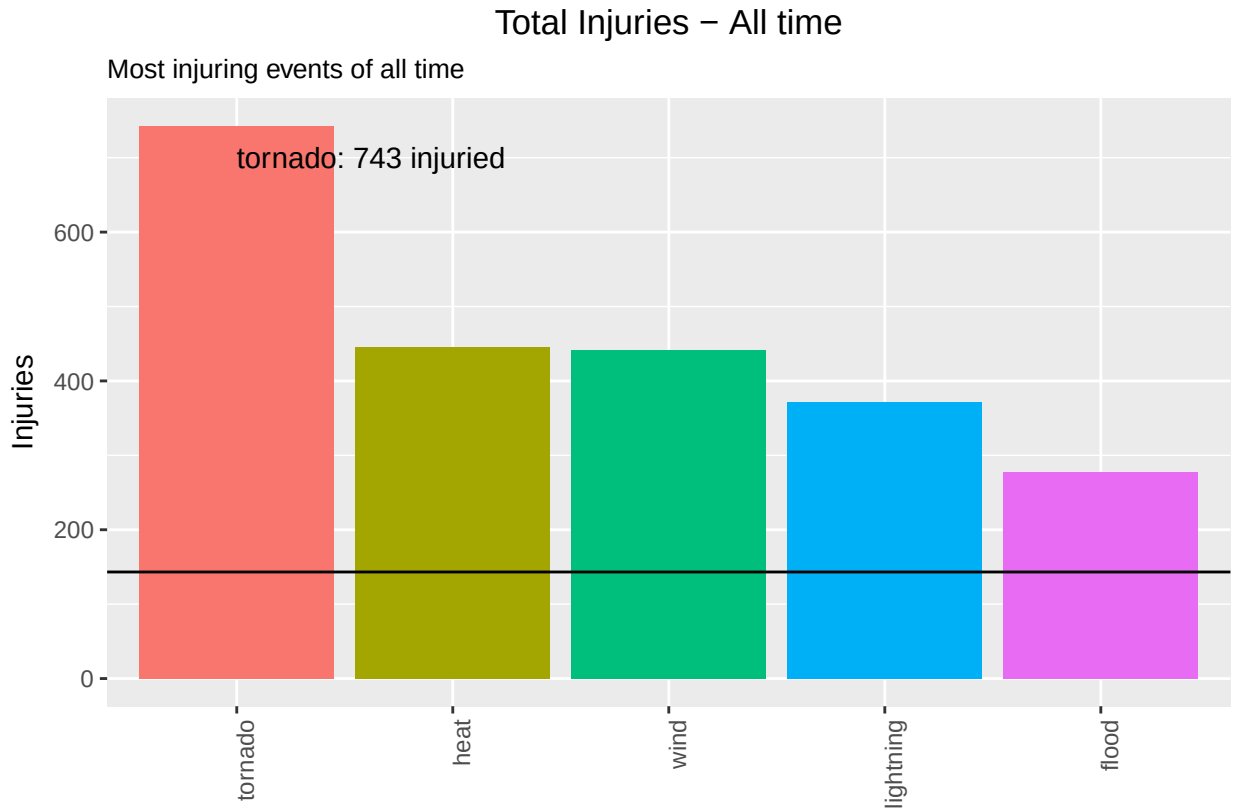

Figure 6: Total Injuries by event

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
19	dust devil	1

rank	event	total
18	rain	1
17	tide	1
16	other	4
15	blizzard	5
14	avalanche	6
13	snow	19
12	storm	19
11	surf	22
10	hail	32

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 **$69,470,000 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 $100,000,000 in losses**.

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

Table 7: Worst property damaging occurrences, mean = $942,545 and median = $10,000

	rank	event	day	state	county	value
1	1	storm	2001-06-05	texas	harris	$5,030,000,000
2	2	hail	2001-04-10	missouri	st.louis	$400,000,000
3	3	hail	2001-04-10	nebraska	douglas	$300,000,000
4	4	hail	2001-04-10	missouri	st.louis	$300,000,000
5	5	hail	2001-04-30	nebraska	douglas	$200,000,000
6	6	flood	2001-04-01	minnesota	anoka	$200,000,000
7	7	hail	2001-05-06	texas	bexar	$120,000,000
8	8	flood	2001-10-11	arkansas	columbia	$120,000,000
9	9	hail	2001-06-09	north.dakota	morton	$113,000,000
10	10	hail	2001-06-09	north.dakota	burleigh	$113,000,000
NA	NA	NA	NA	NA	NA	NA

	rank	event	day	state	county	value
NA.1	NA	NA	NA	NA	NA	NA
NA.2	NA	NA	NA	NA	NA	NA
NA.3	NA	NA	NA	NA	NA	NA
NA.4	NA	NA	NA	NA	NA	NA
NA.5	NA	NA	NA	NA	NA	NA
NA.6	NA	NA	NA	NA	NA	NA
NA.7	NA	NA	NA	NA	NA	NA
NA.8	NA	NA	NA	NA	NA	NA
NA.9	NA	NA	NA	NA	NA	NA

The single most economic damaging event to properties was a **storm, that occurred in texas, harris, on 2001-06-05, causing U$ $5,030,000,000 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 property losses by event, mean = $527,726,193 and median = $20,599,000

rank	event	total
1	storm	$5,195,880,000
2	hail	$2,368,341,440
3	flood	$1,237,204,000
4	tornado	$630,086,900

The most property damaging event along the time is the **storm. It has caused $5,195,880,000 in losses.**

6.1.3 Least property damaging events

Just for curiosity, these are the less damaging events:

Table 9: Least property damaging events

rank	event	total
19	dust devil	$3,000
18	avalanche	$44,000
17	rip current	$60,000
16	volcanic ash	$500,000
15	surf	$622,500
14	hurricane	$650,000
13	fog	$1,283,000
12	tide	$1,460,000
11	blizzard	$8,145,000

rank	event	total
10	winter	$20,599,000

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 $130,250,000 in losses.**

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

Table 10: Worst crops damaging occurrences, mean = $1,301,381 and median = $15,000

	rank	event	day	state	county	value
1	1	drought	2001-12-01	texas	parmer	$420,000,000
2	2	drought	2001-07-01	michigan	wayne	$150,000,000
NA	NA	NA	NA	NA	NA	NA
NA.1	NA	NA	NA	NA	NA	NA
NA.2	NA	NA	NA	NA	NA	NA
NA.3	NA	NA	NA	NA	NA	NA
NA.4	NA	NA	NA	NA	NA	NA
NA.5	NA	NA	NA	NA	NA	NA
NA.6	NA	NA	NA	NA	NA	NA
NA.7	NA	NA	NA	NA	NA	NA
NA.8	NA	NA	NA	NA	NA	NA
NA.9	NA	NA	NA	NA	NA	NA
NA.10	NA	NA	NA	NA	NA	NA
NA.11	NA	NA	NA	NA	NA	NA
NA.12	NA	NA	NA	NA	NA	NA
NA.13	NA	NA	NA	NA	NA	NA
NA.14	NA	NA	NA	NA	NA	NA
NA.15	NA	NA	NA	NA	NA	NA
NA.16	NA	NA	NA	NA	NA	NA
NA.17	NA	NA	NA	NA	NA	NA

The single most economic damaging event to crops was a **drought, that occurred in texas, parmer, on 2001-12-01, causing U$ $420,000,000 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.

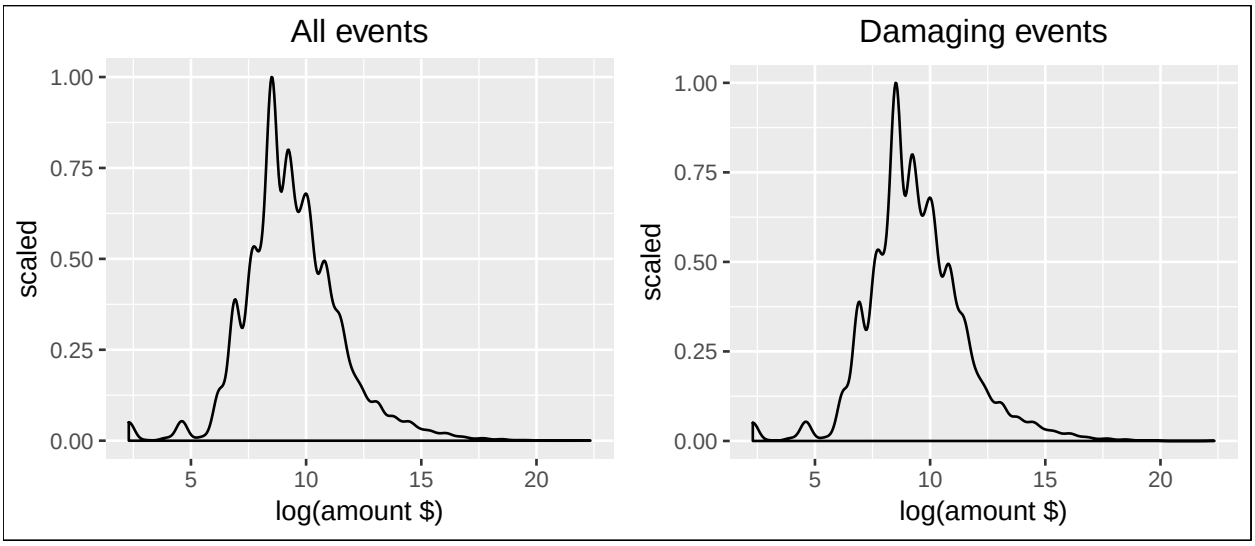

Figure 7: Population distribution for losses / occurrences

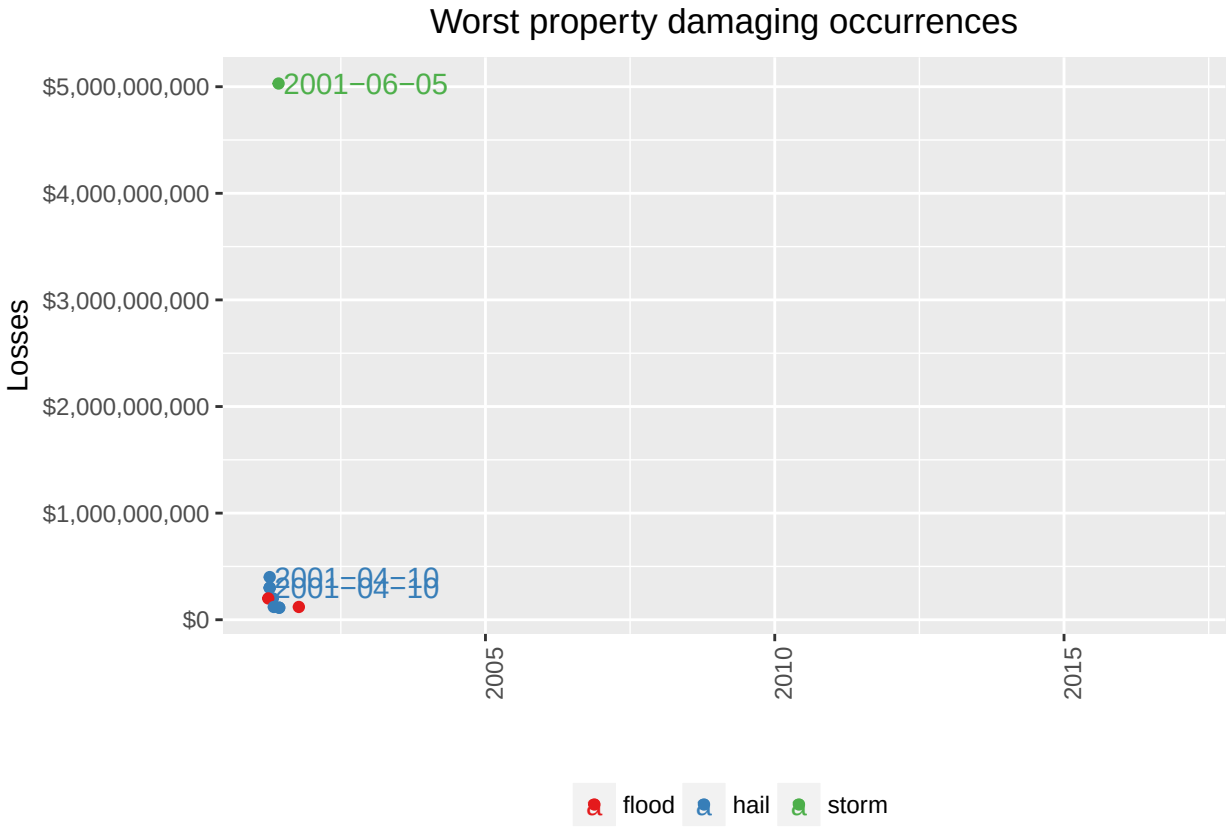

Figure 8: Worst property damaging occurrences

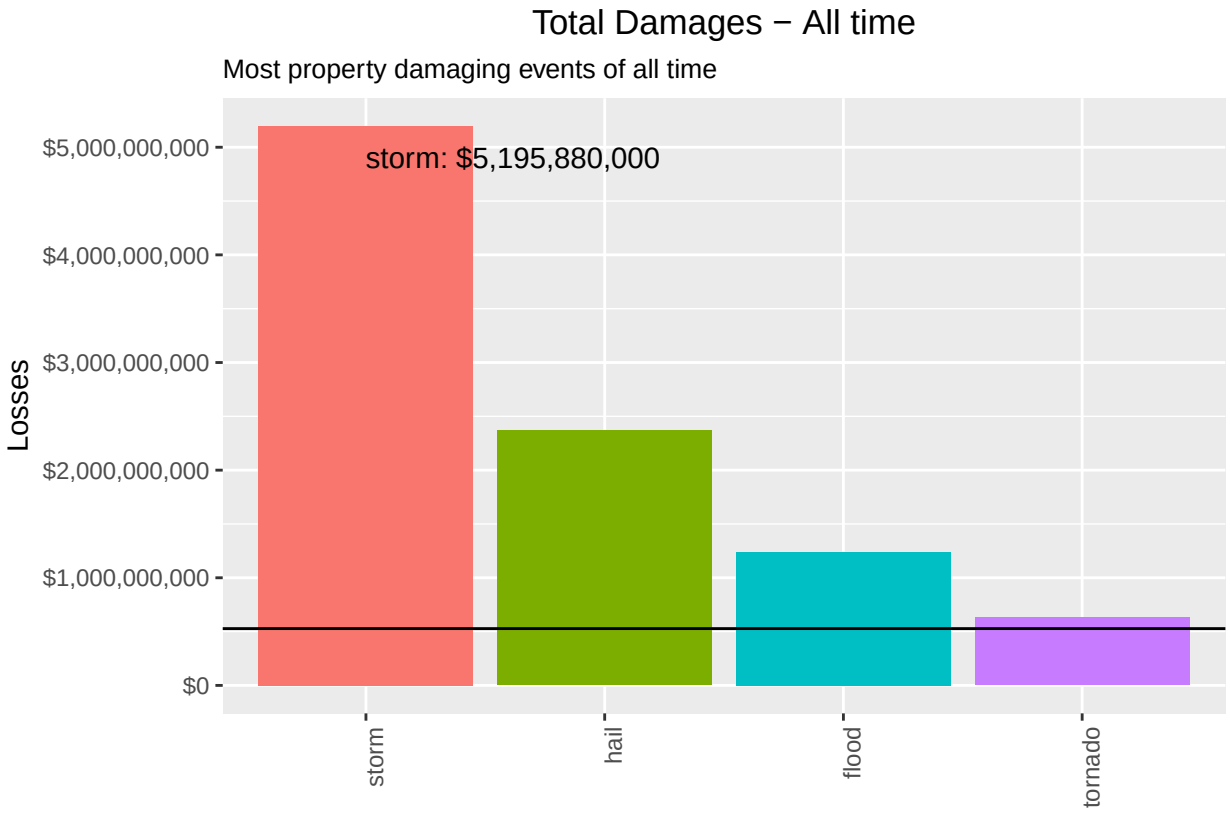

Figure 9: Total Property Damages by event

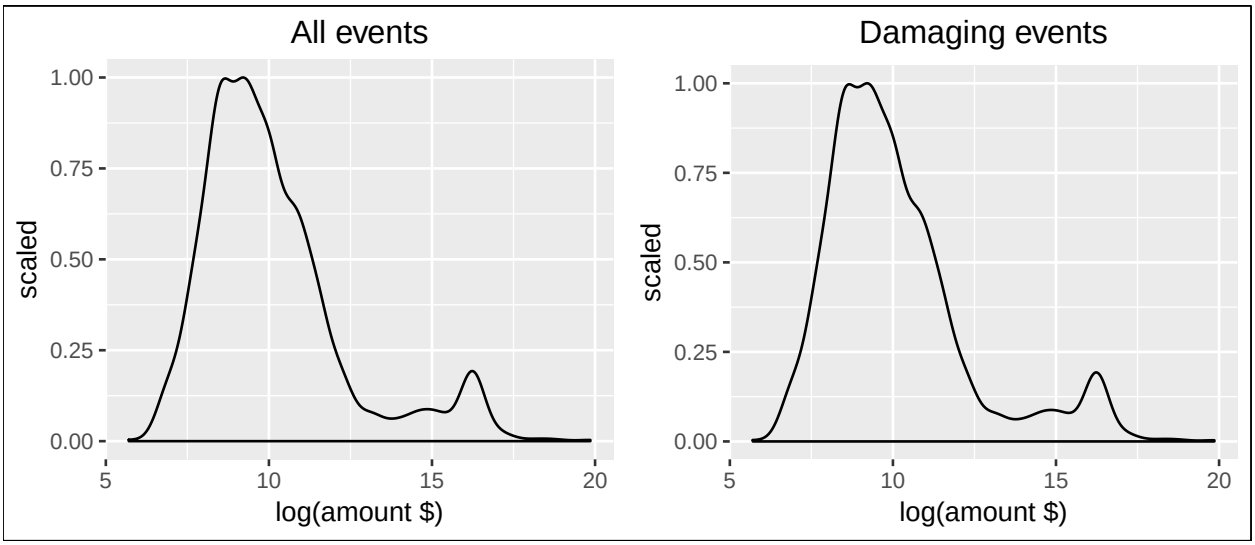

Figure 10: Population distribution for losses / occurrences

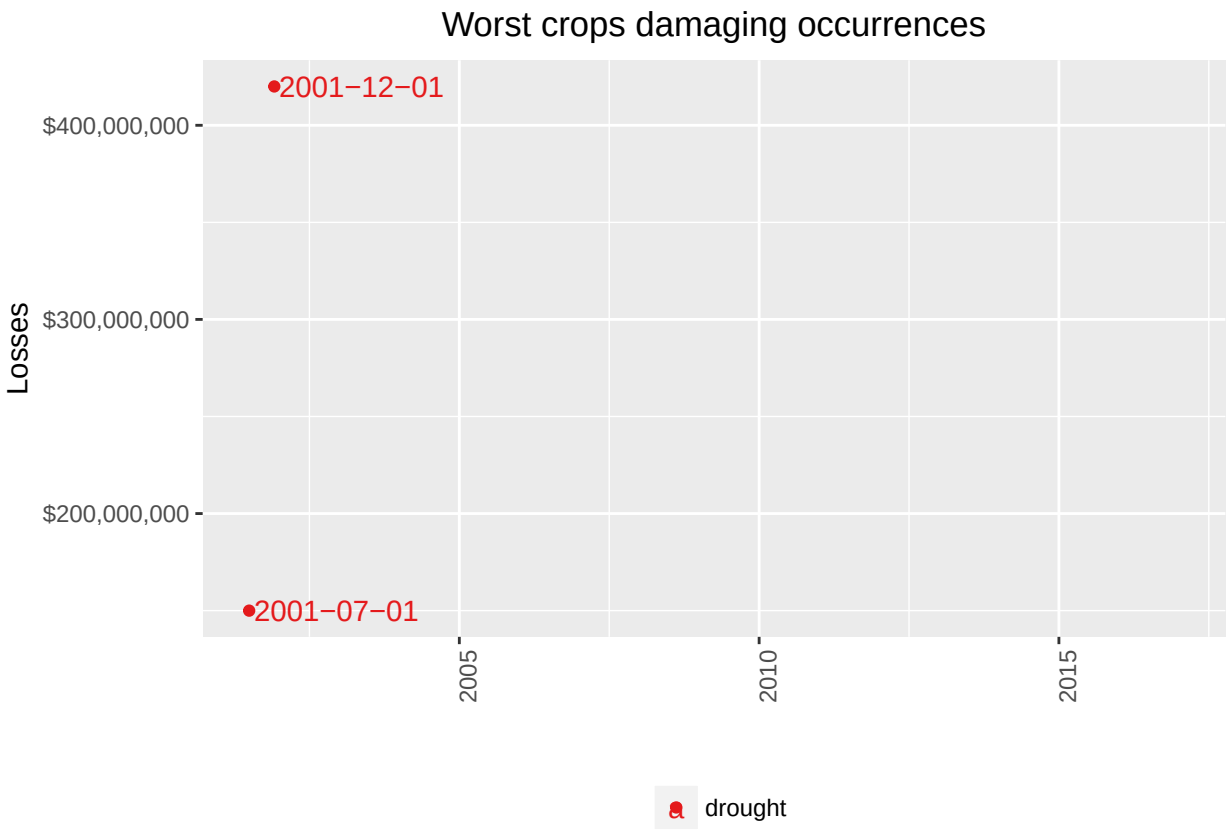

Figure 11: Worst crops damaging occurrences

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 crops losses by event, mean = $165,157,100 and median = $21,501,000

rank	event	total
1	drought	$1,273,860,000
2	hail	$270,386,300

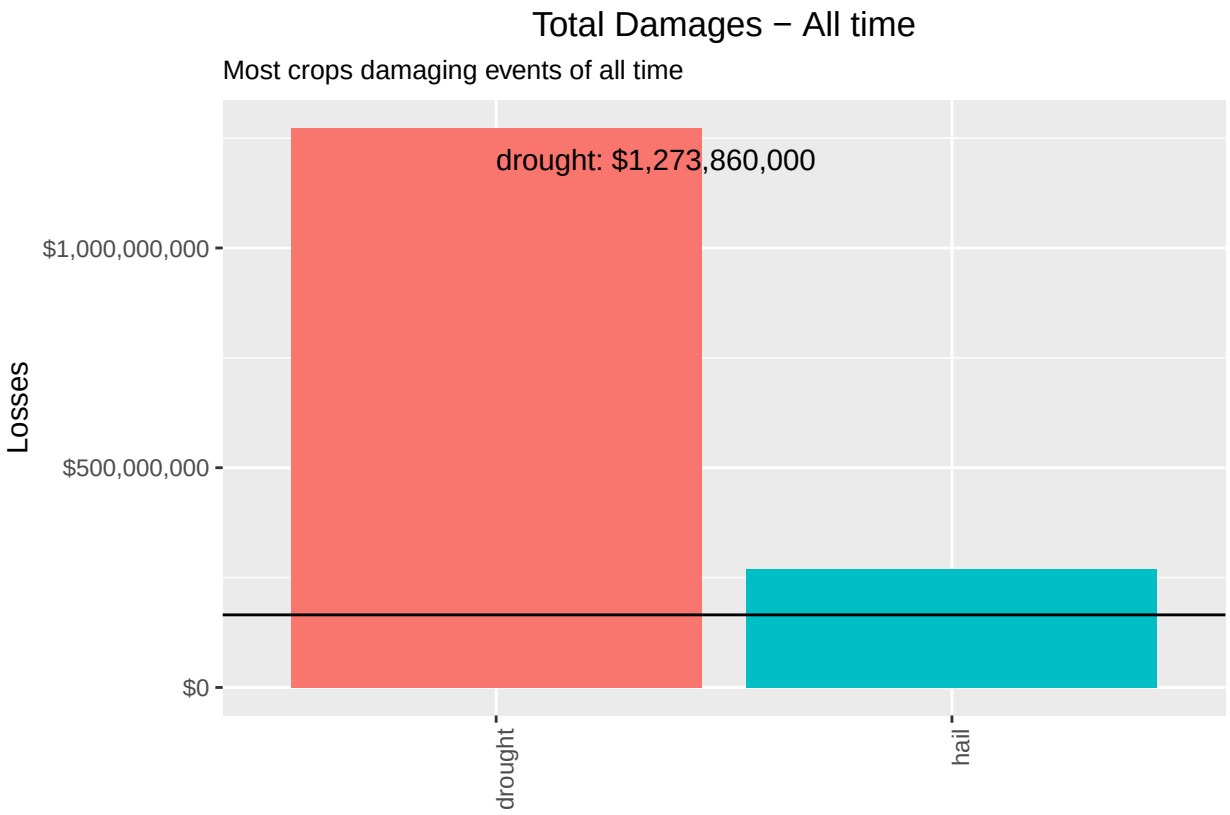

Figure 12: Total Crop Damages by event

The most crop damaging event along the time is the **drought. It has caused $1,273,860,000 in losses.**

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
11	winter	$60,000
10	hurricane	$450,000
9	lightning	$2,007,800

rank	event	total
8	storm	$2,200,000
7	tornado	$7,410,500
6	rain	$21,501,000
5	flood	$43,042,500
4	cold	$47,850,000
3	wind	$147,960,000
2	hail	$270,386,300

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

rank	state	county	fatalities	injuries	prop.dmg	crop.dmg
1	texas	harris	42	0	$5,079,324,000	$0
2	illinois	cook	37	9	$37,640,000	$0
3	pennsylvania	philadelphia	25	0	$0	$0
4	florida	escambia	8	5	$515,000	$0
5	guam	guam	8	11	$475,000	$0
6	missouri	st.louis.	8	127	$0	$0
7	missouri	st.louis	7	83	$850,385,000	$0
8	mississippi	pontotoc	6	43	$28,045,850	$0
9	oklahoma	oklahoma	6	3	$2,652,000	$0
10	puerto.rico	northeast	6	0	$5,325,000	$1,000,000

The county with the biggest fatality count is **harris, in texas, with 42 people killed.**

7.2 Worst injuries count

Table 14: Total injuries by county

rank	state	county	fatalities	injuries	prop.dmg	crop.dmg
1	missouri	st.louis.	8	127	$0	$0
2	michigan	wayne	0	107	$95,000	$150,000,000
3	california	san.diego.mountains	0	106	$0	$0
4	missouri	st.louis	7	83	$850,385,000	$0
5	michigan	oakland	1	80	$1,298,000	$0
6	texas	travis	2	70	$3,065,000	$100,000
7	maryland	prince.george's	2	58	$100,294,000	$0
8	south.carolina	horry	2	49	$9,133,000	$50,000

rank	state	county	fatalities	injuries	prop.dmg	crop.dmg
9	mississippi	pontotoc	6	43	$28,045,850	$0
10	mississippi	bolivar	0	36	$8,283,000	$0

The county with the biggest injuries count is **st.louis., in missouri, with 127 people injuried.**

7.3 Worst property losses

Table 15: Total property losses by county

rank	state	county	fatalities	injuries	prop.dmg	crop.dmg
1	texas	harris	42	0	$5,079,324,000	$0
2	missouri	st.louis	7	83	$850,385,000	$0
3	nebraska	douglas	1	1	$501,177,000	$0
4	missouri	st.charles	1	0	$205,250,000	$0
5	minnesota	anoka	4	1	$200,100,000	$0
6	texas	bexar	0	16	$120,410,000	$30,020,000
7	arkansas	columbia	0	0	$120,017,000	$0
8	north.dakota	morton	0	0	$113,200,000	$0
9	north.dakota	burleigh	0	0	$113,050,000	$0
10	texas	lubbock	0	1	$101,018,000	$15,635,000

The county with the biggest property losses is **harris, in texas, with $5,079,324,000 in losses.**

7.4 Worst crops losses

Table 16: Total crops losses by county

rank	state	county	fatalities	injuries	prop.dmg	crop.dmg
1	texas	parmer	0	0	$413,000	$420,035,000
2	michigan	wayne	0	107	$95,000	$150,000,000
3	florida	inland.palm.beach	0	0	$0	$106,000,000
4	washington	yakima	0	2	$395,000	$70,000,000
5	texas	hockley	0	0	$830,000	$36,820,000
6	florida	inland.collier	0	0	$0	$34,030,000
7	california	ecentralsj	0	10	$20,000	$30,800,000
8	texas	bexar	0	16	$120,410,000	$30,020,000
9	north.dakota	traill	0	0	$1,904,500	$30,000,000
10	florida	inland.dade	0	0	$0	$25,000,000

The county with the biggest croperty losses is **parmer, in texas, with $420,035,000 in losses.**

8 Results

8.1 Population Health

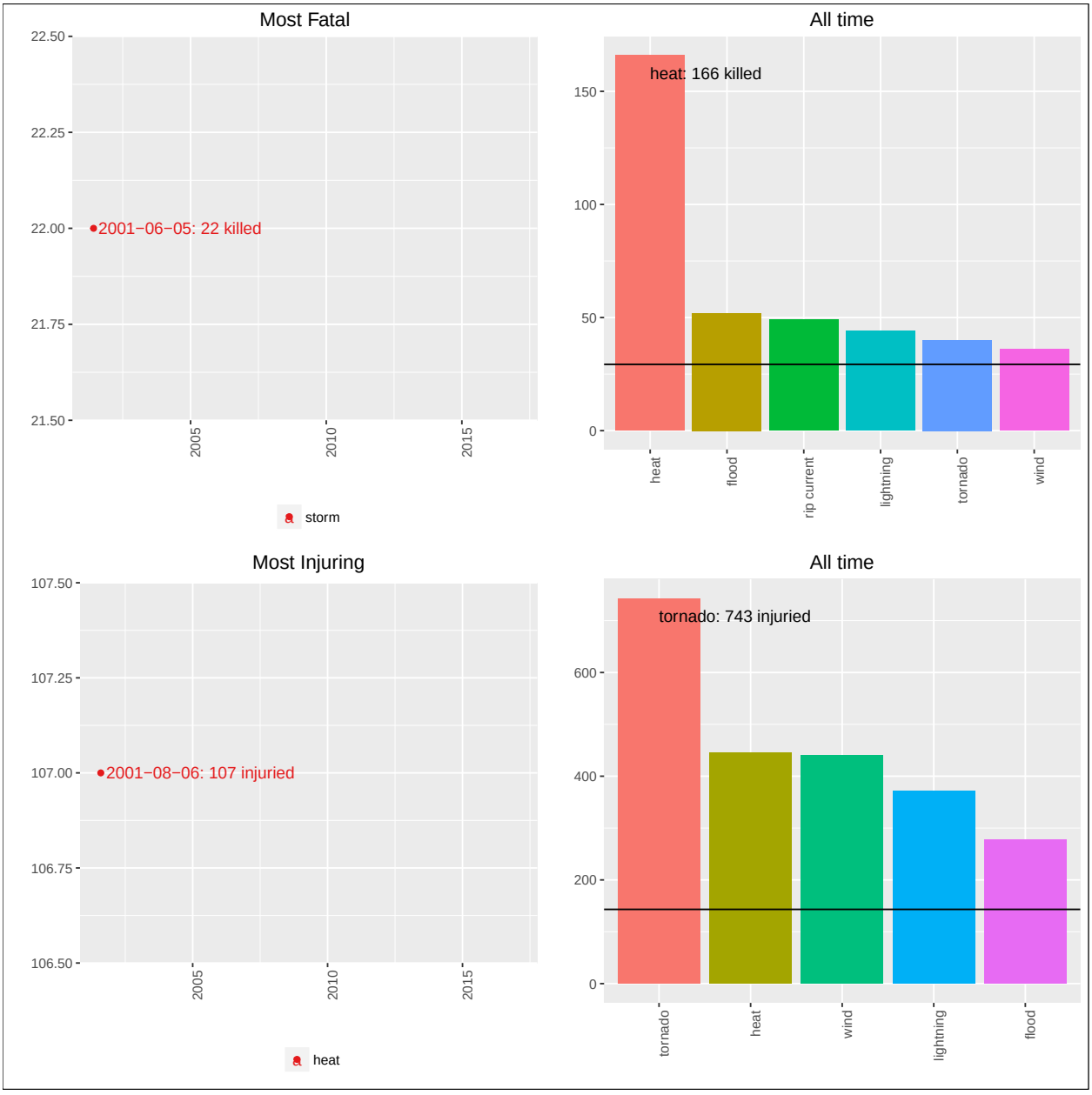

Figure 13: Population Health: fatalities and injuries

The single most fatal event was a **storm, that occurred in texas, harris, on 2001-06-05, killing 22 people.**

The most fatal event along the time is the **heat. It has killed 166 people until now.**

The single most injuring event was a **heat, that occurred in michigan, wayne, on 2001-08-06, injuring 107 people.**

The most injuring event along the time is the **tornado. It has injuried 743 people until now.**

8.2 Economic Damages

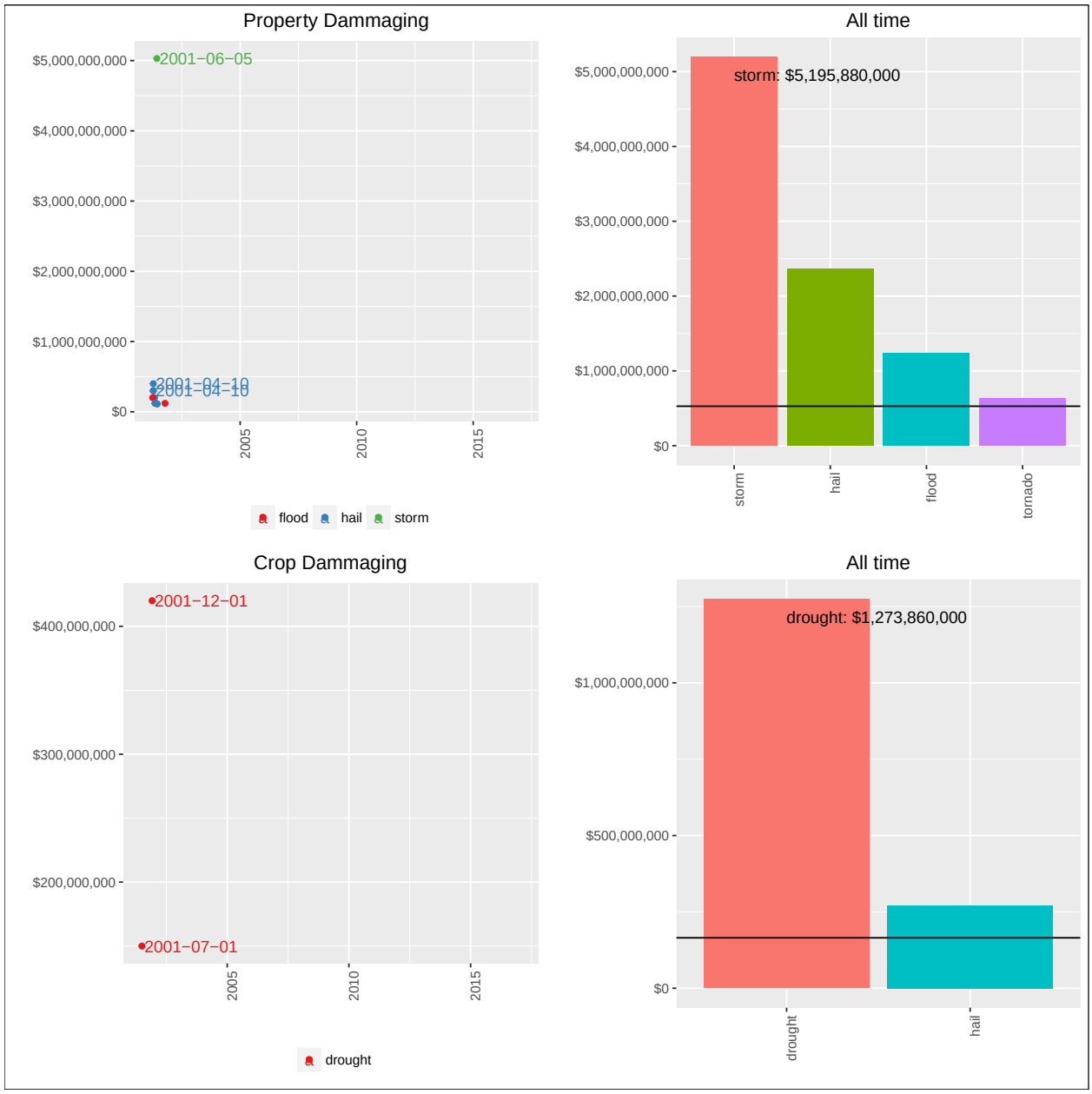

Figure 14: Economic Damages: property and crops

The single most economic damaging event to properties was a **storm, that occurred in texas, harris, on 2001-06-05, causing U$ $5,030,000,000 in losses**.

The most property damaging event along the time is the **storm. It has caused $5,195,880,000 in losses.**

The single most economic damaging event to crops was a **drought, that occurred in texas, parmer, on 2001-12-01, causing U$ $420,000,000 in losses**.

The most crop damaging event along the time is the **drought. It has caused $1,273,860,000 in losses.**

8.3 Most aflicted locations

The county with the biggest fatality count is **harris, in texas, with 42 people killed.**

The county with the biggest injuries count is **st.louis., in missouri, with 127 people injuried.**

The county with the biggest property losses is **harris, in texas, with $5,079,324,000 in losses.**

The county with the biggest croperty losses is **parmer, in texas, with $420,035,000 in losses.**

8.4 Distribution of data