

Allocating Resources after Hurricane Harvey

Table of Contents

Background and Scope.....	1
Import the Data.....	1
Cleaning the data.....	1
Two States Most Impacted by Harvey.....	2
Table of Events for Two Most Impacted States.....	3
Visualizations	4
Figure of Event Types.....	4
Analysis	5
Three Counties with Most Events in Texas.....	5
Three Counties with Most Events in Louisiana.....	6
Three Counties with Highest Damage Cost in Texas.....	6
Three Counties with Highest Property Cost in L.....	6
Conclusions and Recommendations	7

Background and Scope

The data file contains a list of weather events that occurred in the United States in 2017. The events have been categorized by type and the state in which they occurred. There is also information about

- damage costs
- number of resulting injuries
- location of some of the events

Import the Data

```
clear
% import the data using a function generated by the Import Tool
events = import2017stormdata("StormEvents_2017_finalProject.csv");
```

Cleaning the data

Some Property Costs and Crop Costs are missing, with the value NaN. Replacing the missing data with \$0 makes the data easier to work with. Since the insurance company is interested in the total damage cost, the property and crop costs can be added together. The data file contains many more weather events than the ones pertaining to Hurricane Harvey, so the results were filtered to show only the states that Harvey passed through.

```
% Set missing Property and Crop Cost to $0
events.Property_Cost(ismissing(events.Property_Cost)) = 0;
events.Crop_Cost(ismissing(events.Crop_Cost)) = 0;
% Add total damage to the table
events.Total_Damage = events.Property_Cost + events.Crop_Cost;
% Filter to Harvey-relevant States
harveystates = events(events.State == 'TEXAS' | events.State == 'ARKANSAS' | events.State == 'KENTUCKY');
```

Next, Harvey-related events lasted from August 17th to September 3rd.

```
% Filter to correct timeframe
harveydays = harveystates(isbetween(harveystates.Begin_Date_Time, '2017-08-17', '2017-09-03'),

% Hard remove states and event types that do not appear in the filtered
% table, so they do not affect later figures
harveydays.State = removecats(harveydays.State);
harveydays.Event_Type = removecats(harveydays.Event_Type);
```

Are all the events in this filtered table Harvey-related?

```
groupsummary(harveydays, "Event_Type")
```

ans = 12x2 table

	Event_Type	GroupCount
1	Flash Flood	222
2	Flood	24
3	Funnel Cloud	3
4	Hail	12
5	Heat	47
6	Heavy Rain	10
7	Hurricane	9
8	Storm Surge/Tide	10
9	Strong Wind	12
10	Thunderstorm Wind	47
11	Tornado	40
12	Tropical Storm	44

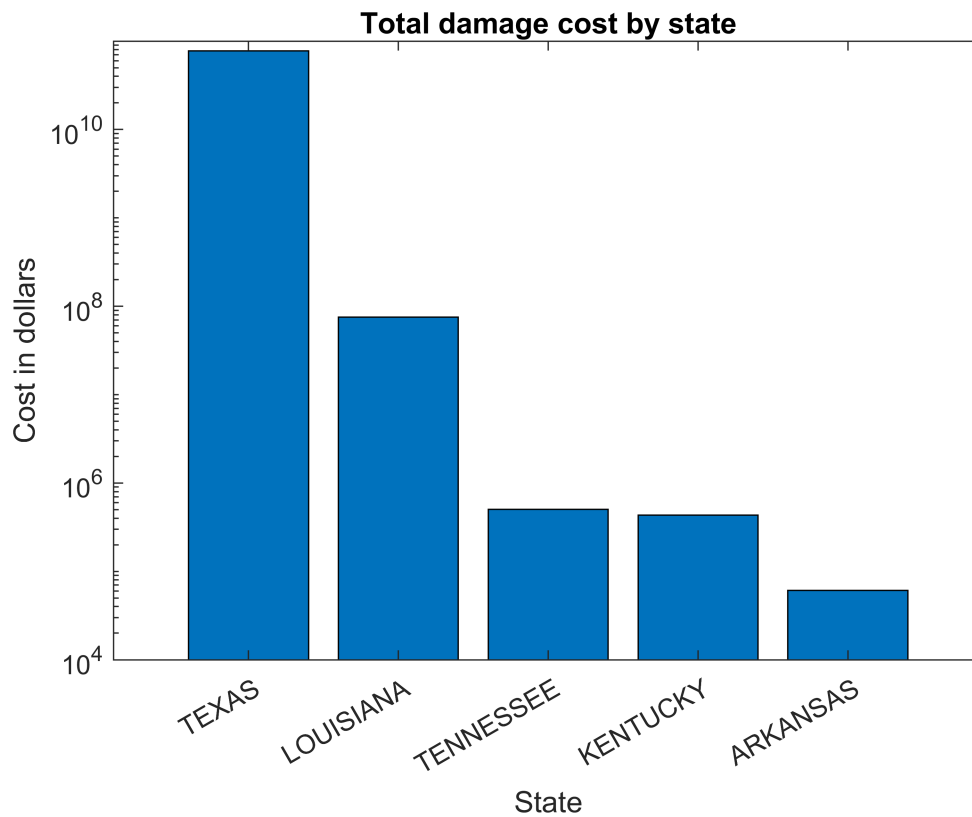
Most of the categories above are hurricane-related, but "Heat" is most likely unrelated. (Checking the dates and locations on those events confirm that these heat events happened while Harvey was not in the area.)

```
% Remove heat from table
harveydays = harveydays(harveydays.Event_Type ~= 'Heat',:);
```

Two States Most Impacted by Harvey

```
% Group Total Damage Costs by State
statecost = groupsummary(harveydays, "State", "sum", "Total_Damage");
% Order states by descending order of costs
[S_Sorted,idx1] = sort(statecost.sum_Total_Damage,'descend');
statecost.State = categorical(statecost.State(idx1));
statecost.State = reordercats(statecost.State,string(statecost.State));
% View the results using a bar chart
bar(statecost.State,S_Sorted);
% For better viewing, set y-axis to logarithmic scale
```

```
set(gca, 'YScale', 'log');
title("Total damage cost by state")
xlabel("State")
ylabel("Cost in dollars")
```



The two most impacted states were Texas and Louisiana!

Table of Events for Two Most Impacted States

```
txla = harveydays(harveydays.State == 'TEXAS' | harveydays.State == "LOUISIANA",:);
head(txla,25)
```

ans = 25×25 table

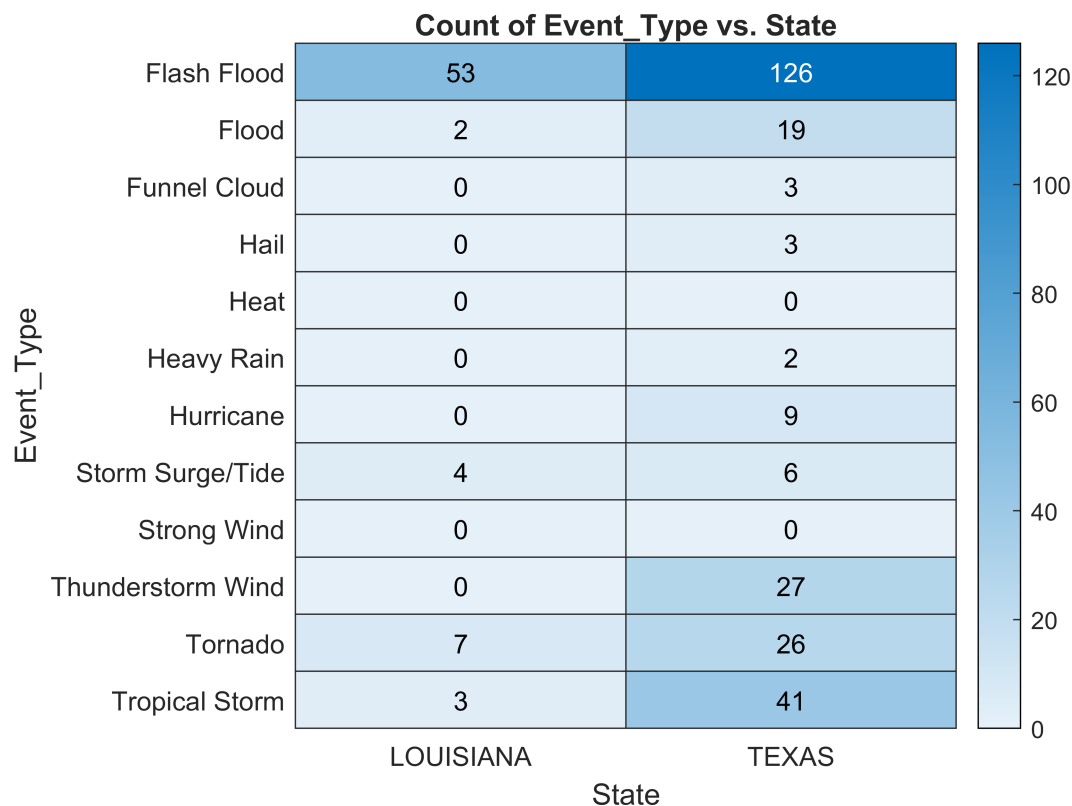
	EpisodeID	Event_ID	State	Year	Month	Event_Type	CZ_Name
1	119753	723472	TEXAS	2017	August	Tropical Storm	MONTGOMERY
2	119753	723473	TEXAS	2017	August	Tropical Storm	FORT BEND
3	119753	723449	TEXAS	2017	August	Tropical Storm	GALVESTON
4	119753	723474	TEXAS	2017	August	Tropical Storm	SAN JACINTO
5	119753	723475	TEXAS	2017	August	Tropical Storm	WALKER
6	119753	723648	TEXAS	2017	August	Tropical Storm	POLK
7	120011	719146	TEXAS	2017	August	Flash Flood	EL PASO
8	120012	719147	TEXAS	2017	August	Thunderstorm Wind	EL PASO

	EpisodeID	Event_ID	State	Year	Month	Event_Type	CZ_Name
9	120012	719148	TEXAS	2017	August	Flash Flood	EL PASO
10	119746	719493	TEXAS	2017	August	Flash Flood	HARDIN
11	119746	719496	TEXAS	2017	August	Flash Flood	JASPER
12	119746	719497	TEXAS	2017	August	Flash Flood	NEWTON
13	119753	720340	TEXAS	2017	August	Flash Flood	FORT BEND
14	119826	718436	TEXAS	2017	August	Thunderstorm Wind	MIDLAND
15	117836	708282	TEXAS	2017	August	Thunderstorm Wind	BRISCOE
16	119746	719740	TEXAS	2017	August	Thunderstorm Wind	JASPER
17	119746	720010	TEXAS	2017	August	Flood	ORANGE
18	119753	720464	TEXAS	2017	August	Flash Flood	MONTGOMERY
19	119853	718515	TEXAS	2017	August	Thunderstorm Wind	ECTOR
20	119746	719342	TEXAS	2017	August	Flash Flood	JEFFERSON
21	119753	720344	TEXAS	2017	August	Flash Flood	GALVESTON
22	118032	709519	TEXAS	2017	August	Thunderstorm Wind	SWISHER
23	118916	714375	LOUISIANA	2017	August	Tropical Storm	SABINE
24	118330	711050	TEXAS	2017	August	Flash Flood	ANGELINA
25	118330	711054	TEXAS	2017	August	Flash Flood	ANGELINA

Visualizations

Figure of Event Types

```
topstates = harveydays(harveydays.State == 'TEXAS' | harveydays.State == 'LOUISIANA',:);
topstates.State = removecats(topstates.State);
heatmap(topstates, "State", "Event_Type")
```



Flash flooding was the most frequent in both states, especially in Texas.

Analysis

Texas and Louisiana had the most damage costs by far. But not all areas in these two states were affected equally. Below is an analysis of which counties in each state had the most events and damage costs.

Three Counties with Most Events in Texas

```
% Texas events only
topstate1 = harveydays(harveydays.State == 'TEXAS',:);
% remove empty categories
topstate1.CZ_Name = removecats(topstate1.CZ_Name);
% Group Events by County
txevents = groupsummary(topstate1, "CZ_Name");
% Rename column title
txevents.Properties.VariableNames{2} = 'No. of Events';
% Order counties by number of events
txevents = sortrows(txevents, 'No. of Events', 'descend');
head(txevents,3)
```

ans = 3x2 table

	CZ_Name	No. of Events
1	HARRIS	20
2	GALVESTON	17
3	FORT BEND	13

Harris, Galveston, and Fort Bend Counties had the most recorded events in Texas.

Three Counties with Most Events in Louisiana

```
% Louisiana events only
topstate2 = harveydays(harveydays.State == 'LOUISIANA',:);
% remove empty categories
topstate2.CZ_Name = removecats(topstate2.CZ_Name);
% Group Events by County
laevents = groupsummary(topstate2, "CZ_Name");
% Rename column title
laevents.Properties.VariableNames{2} = 'No. of Events';
% Order counties by number of events
laevents = sortrows(laevents, 'No. of Events', 'descend');
head(laevents,3)
```

ans = 3x2 table

	CZ_Name	No. of Events
1	NATCHITOC...	20
2	SABINE	14
3	RED RIVER	8

Natchitoches, Sabine, and Red River Counties had the most recorded events in Louisiana.

Three Counties with Highest Damage Cost in Texas

Reordering the rows of the previous two tables, it can be seen that the most affected areas cost-wise are Galveston, Fort Bend, and Montgomery Counties for Texas, and Calcasieu, Beauregard, and Acadia Counties for Louisiana. While two of the counties in Texas with the most events are also at the top of the list for damage costs, the list for Louisiana is completely different!

```
% Group Total Damage costs by County
txcountycost = groupsummary(topstate1, "CZ_Name", 'sum', 'Total_Damage');
% Order counties by damage cost
txcountycost = sortrows(txcountycost, 'sum_Total_Damage', 'descend');
txcountycost = removevars(txcountycost, 'GroupCount');
head(txcountycost,3)
```

ans = 3x2 table

	CZ_Name	sum_Total_Damage
1	GALVESTON	2.0000e+10
2	FORT BEND	1.6004e+10
3	MONTGOMERY	1.4000e+10

Three Counties with Highest Property Cost in Louisiana

```
% Group Total Damage costs by County
lacountycost = groupsummary(topstate2, "CZ_Name", 'sum', 'Total_Damage');
% Order counties by damage cost
lacountycost = sortrows(lacountycost, 'sum_Total_Damage', 'descend');
```

```
lacountycost = removevars(lacountycost, 'GroupCount');
head(lacountycost,3)
```

```
ans = 3x2 table
```

	CZ_Name	sum_Total_Damage
1	CALCASIEU	60000000
2	BEAUREGARD	15000000
3	ACADIA	200000

Conclusions and Recommendations

Most of the events recorded, over 60%, were related to flooding or flash flooding. Of these, 145 events happened in Texas! But since the number of Events are simply the records of weather events that already happened, the amount of damage cost deserves the more attention when considering where to send out contractors. Harris County had the most events recorded, but missed the top 3 when it came to damage costs. However, it still stands at fourth place with \$10 billion worth of damage costs. The top 11 counties for damage cost in Texas actually all have at least \$1 billion in damage costs, whereas the top county in Louisiana has a relatively low \$60 million. For the insurance company, it is recommended that the order of priority to send contractors to are as follows:

```
head(txcountycost,11)
```

```
ans = 11x2 table
```

	CZ_Name	sum_Total_Damage
1	GALVESTON	2.0000e+10
2	FORT BEND	1.6004e+10
3	MONTGOMERY	1.4000e+10
4	HARRIS	1.0001e+10
5	JEFFERSON	3.0000e+09
6	BRAZORIA	2.0009e+09
7	ARANSAS	1.9500e+09
8	ORANGE	1.5000e+09
9	NUECES	1.3000e+09
10	WALKER	1.2000e+09
11	LIBERTY	1.0000e+09

Galveston, Fort Bend, Montgomery, and Harris Counties especially have an urgent need for storm recovery, all totaling more than \$10 billion.