

# Analyzing Hurricane Harvey

This script analyzes the effects of Hurricane Harvey in the US; particularly the two states which were most affected. It has been divided into four sections: Background, Visualization, Analysis and Conclusion. The background deals with importing the dataset we have and finding out the states which were most impacted by this storm. The visualization section has a few plots to understand the kind of events that we are dealing with, and their locations in the two states. Next, we analyze each state by grouping it into counties and counting the total number of events in each county, along with finding out which counties in each state had recorded the highest property damage cost. Finally, we summarize our findings and make a suggestion regarding the counties to be prioritized.

## Table of Contents

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

## Background and Scope

### Importing the Data

We import the file with the necessary variables; while discarding columns not needed for this analysis.

```
Storms_2017 = importfile("StormEvents_2017_finalProject.csv");  
Storms_2017 = Storms_2017(2:end, :)
```

```
Storms_2017 = 57005x13 table
```

	State	Year	Month	Event_Type	CZ_Name	Begin_Date_Time	End_Date_Time
1	NEW JER...	2017	April	Thunderstorm Wind	GLOUCESTER	2017-04-06 15:09...	2017-04-06 15:...
2	FLORIDA	2017	April	Tornado	LEE	2017-04-06 09:30...	2017-04-06 09:...
3	OHIO	2017	April	Thunderstorm Wind	GREENE	2017-04-05 17:49...	2017-04-05 17:...
4	OHIO	2017	April	Flood	CLERMONT	2017-04-16 17:59...	2017-04-16 19:...

	State	Year	Month	Event_Type	CZ_Name	Begin_Date_Time	End_Date_Time
5	NEBRASKA	2017	April	Hail	CASS	2017-04-15 15:50...	2017-04-15 15:...
6	INDIANA	2017	April	Flash Flood	SWITZERLAND	2017-04-29 09:15...	2017-04-29 11:...
7	VIRGINIA	2017	April	Thunderstorm Wind	WESTMOREL...	2017-04-21 19:15...	2017-04-21 19:...
8	GULF OF...	2017	October	Marine Thunders...	ATCHAFALA...	2017-10-22 10:15...	2017-10-22 10:...
9	OHIO	2017	April	Flash Flood	CLERMONT	2017-04-29 09:45...	2017-04-29 11:...
10	NEBRASKA	2017	April	Thunderstorm Wind	BURT	2017-04-15 18:55...	2017-04-15 18:...
11	ARKANSAS	2017	April	Hail	FRANKLIN	2017-04-26 07:57...	2017-04-26 07:...
12	OKLAHOMA	2017	October	Hail	KIOWA	2017-10-21 15:20...	2017-10-21 15:...
13	ATLANTI...	2017	October	Marine Strong W...	DE BAY WA...	2017-10-24 02:24...	2017-10-24 02:...
14	ATLANTI...	2017	October	Marine High Wind	DE BAY WA...	2017-10-24 03:36...	2017-10-24 03:...

⋮

Now, as given to us in the assignment introduction, we are going to focus on states that were most affected by Harvey: **Arkansas, Kentucky, Louisiana, Mississippi, North Carolina, Tennessee, and Texas**. We will also narrow our dates between August 17 and September 3rd, so as to focus solely on Harvey-related storms.

```
Harvey = Storms_2017(ismember(Storms_2017.State,{'ARKANSAS','KENTUCKY','LOUISIANA','MISSISSIPPI','NORTH CAROLINA','TENNESSEE','TEXAS'}))
Harvey = Harvey((Harvey.Begin_Date_Time >= '17-Aug-2017') & (Harvey.End_Date_Time <= '03-Sep-2017'))
```

## Two States Most Impacted by Harvey

We wish to find the two states most affected by Hurricane Harvey in terms of Property Cost. We will do so by grouping the table by State and then sorting the column 'Property\_Cost' of the table.

```
property_cost_by_state = grpstats(Harvey, {'State'}, 'sum', "DataVars", 'Property_Cost')
```

```
property_cost_by_state = 7x3 table
```

	State	GroupCount	sum_Property_Cost
1 ARKANSAS	ARKANSAS	52	61000
2 KENTUCKY	KENTUCKY	20	435000
3 LOUISIANA	LOUISIANA	85	75277000
4 MISSISSIPPI	MISSISS...	39	915000
5 NORTH CAROLINA	NORTH C...	59	12338500
6 TENNESSEE	TENNESSEE	46	504000
7 TEXAS	TEXAS	271	7.7427e+10

```
property_cost_by_state = sortrows(property_cost_by_state, 'sum_Property_Cost', 'descend')
max_two_states = property_cost_by_state.State(1:2)
```

```
max_two_states = 2x1 categorical
```

TEXAS  
LOUISIANA

So, from this, we have our findings that the two states most affected in property by the storm Harvey are **Texas** and **Louisiana**.

## Table of Events for Two Most Impacted States

Now, we will make a table of the events of Harvey just for these states to be able to study them in more detail.

```
Texas_and_Louisiana = Harvey(ismember(Harvey.State,{'TEXAS','LOUISIANA'}),:)
```

```
Texas_and_Louisiana = 356×13 table
```

	State	Year	Month	Event_Type	CZ_Name	Begin_Date_Time	End_Date_Time
1	TEXAS	2017	August	Tropical Storm	MONTGOMERY	2017-08-25 12:00...	2017-08-30 00:...
2	TEXAS	2017	August	Tropical Storm	FORT BEND	2017-08-26 00:00...	2017-08-30 00:...
3	TEXAS	2017	August	Tropical Storm	GALVESTON	2017-08-25 12:00...	2017-08-30 00:...
4	TEXAS	2017	August	Tropical Storm	SAN JACINTO	2017-08-25 12:00...	2017-08-30 00:...
5	TEXAS	2017	August	Tropical Storm	WALKER	2017-08-25 12:00...	2017-08-30 00:...
6	TEXAS	2017	August	Tropical Storm	POLK	2017-08-25 12:00...	2017-08-30 00:...
7	TEXAS	2017	August	Flash Flood	EL PASO	2017-08-23 16:15...	2017-08-23 17:...
8	TEXAS	2017	August	Thunderstorm Wind	EL PASO	2017-08-25 18:10...	2017-08-25 18:...
9	TEXAS	2017	August	Flash Flood	EL PASO	2017-08-25 18:48...	2017-08-25 20:...
10	TEXAS	2017	August	Flash Flood	HARDIN	2017-08-27 12:40...	2017-08-30 16:...
11	TEXAS	2017	August	Flash Flood	JASPER	2017-08-29 22:29...	2017-08-30 16:...
12	TEXAS	2017	August	Flash Flood	NEWTON	2017-08-29 22:29...	2017-08-30 16:...
13	TEXAS	2017	August	Flash Flood	FORT BEND	2017-08-26 00:45...	2017-08-26 02:...
14	TEXAS	2017	August	Thunderstorm Wind	MIDLAND	2017-08-22 20:59...	2017-08-22 20:...

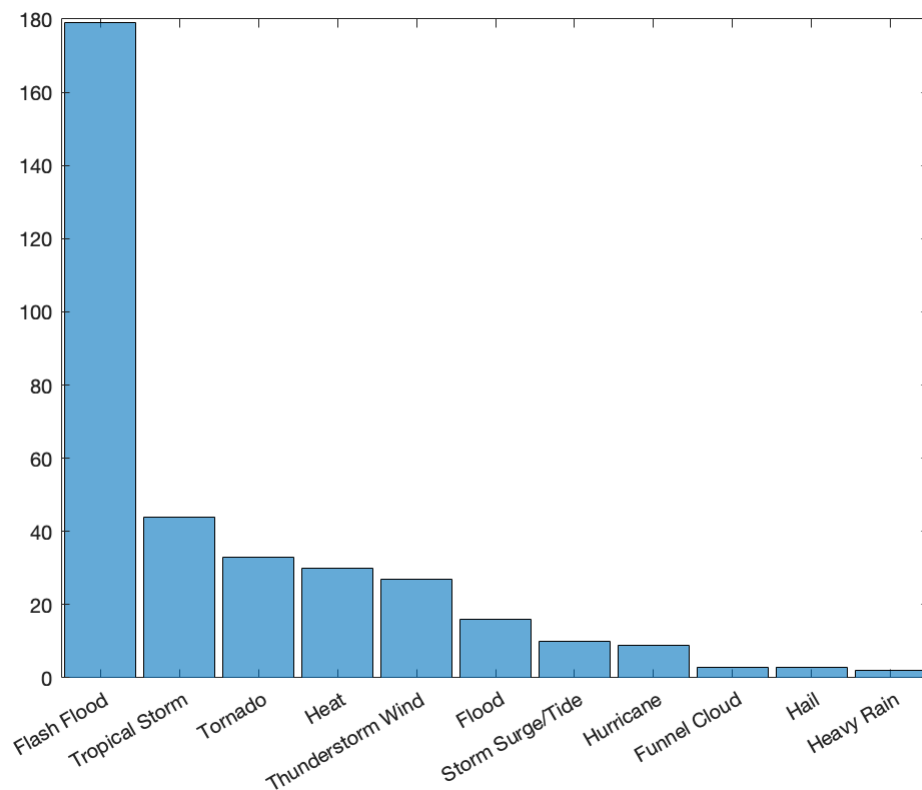
⋮

## Visualizations

### Figure of Event Types

First we want to understand and get a sense of what kind of events we are dealing with. So, we will make a histogram of the event types in the two states.

```
h = histogram(Texas_and_Louisiana.Event_Type,"DisplayOrder",'descend','NumDisplayBins',
```



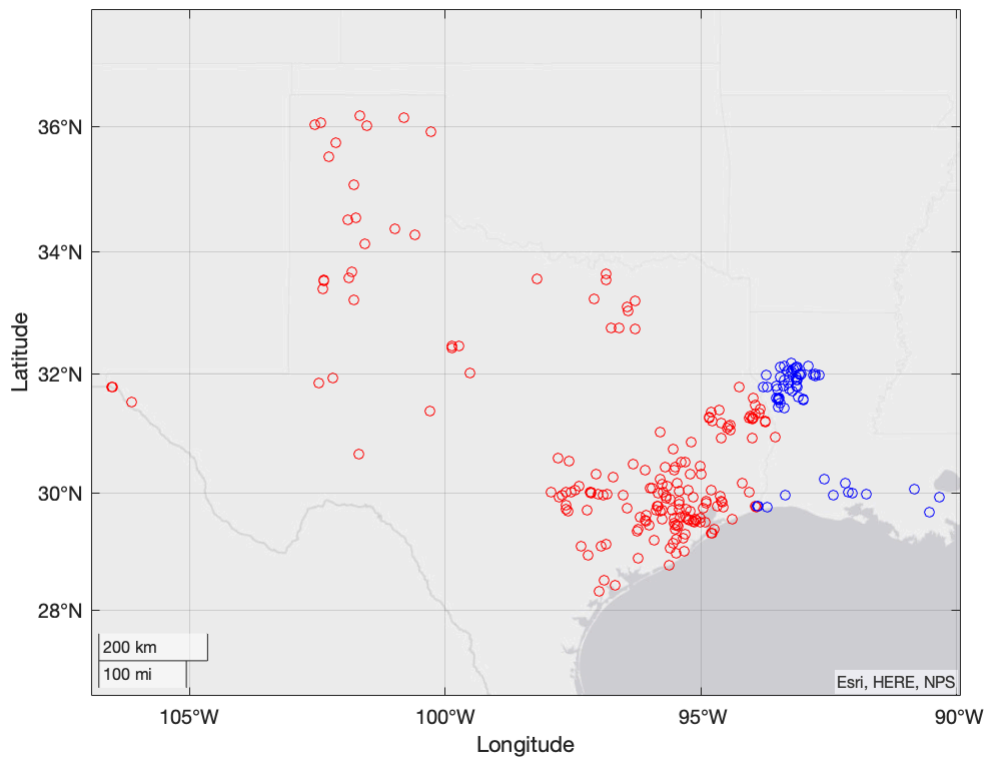
Clearly, Flash Floods are the most common storm occurrences due to Hurricane Harvey in Texas and Louisiana.

## Figure of Event Locations

We will plot the locations of the storms in Texas and Louisiana, and we will mark them with different colours to differentiate between the two.

```
Texas = Texas_and_Louisiana(Texas_and_Louisiana.State == 'TEXAS', :);
Louisiana = Texas_and_Louisiana(Texas_and_Louisiana.State == 'LOUISIANA', :);

geoscatter(Texas.Begin_Lat, Texas.Begin_Lon, 25, "red")
hold on
geoscatter(Louisiana.Begin_Lat, Louisiana.Begin_Lon, 25, "blue")
hold off
```



## Analysis

### Three Counties with Most Events in Texas

Let us analyze the event distribution in Texas by counties, and identify the three counties with most events in Texas.

```
counties_texas = groupcounts(Texas, 'CZ_Name');
counties_texas = sortrows(counties_texas, 'GroupCount', 'descend');
max_three_counties_texas = counties_texas.CZ_Name(1:3)
```

```
max_three_counties_texas = 3x1 categorical
HARRIS
GALVESTON
FORT BEND
```

Hence, we have the names of the counties which had the most events in Texas.

### Three Counties with Most Events in Louisiana

We apply the same procedure to the state of Louisiana.

```
counties_ls = groupcounts(Louisiana, 'CZ_Name');
counties_ls = sortrows(counties_ls, 'GroupCount', 'descend');
max_three_counties_ls = counties_ls.CZ_Name(1:3)
```

```
max_three_counties_ls = 3x1 categorical
```

## Three Counties with Highest Property Cost in Texas

Now, we find out the three counties that reported the highest property costs in Texas.

```
counties_by_cost_tx = grpstats(Texas, "CZ_Name", "sum", "DataVars", "Property_Cost");  
counties_by_cost_tx = sortrows(counties_by_cost_tx, "sum_Property_Cost", "descend");  
max_three_cost_tx = counties_by_cost_tx(1:3, [1 3])
```

max\_three\_cost\_tx = 3×2 table

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

## Three Counties with Highest Property Cost in Louisiana

We do the same for Louisiana.

```
counties_by_cost_ls = grpstats(Louisiana, "CZ_Name", "sum", "DataVars", "Property_Cost");  
counties_by_cost_ls = sortrows(counties_by_cost_ls, "sum_Property_Cost", "descend");  
max_three_cost_ls = counties_by_cost_ls(1:3, [1 3])
```

max\_three\_cost\_ls = 3×2 table

	CZ_Name	sum_Property_Cost
1 CALCASIEU	CALCASIEU	60000000
2 BEAUREGARD	BEAUREGARD	15000000
3 ACADIA	ACADIA	200000

## Conclusions and Recommendations

Notice that we narrowed down the dataset to two of the most impacted states. Within the states, we identified the counties with the most events and those with the highest amounts of property costs. As can be seen, in Texas, the property costs are much higher than those in Louisiana, by several orders of magnitude. So, I would recommend the company to send its people to these counties first, followed by some others from Texas which have a very high property cost. I would prioritize the counties of Louisiana after the most damaged ones in Texas.