

Christian Taylor
Joseph Weingold
Bill Hansen

Harvey Data Mining Report

We decided to investigate the Hurricane Harvey data because of its topical nature and relief efforts going on right at this very moment. The programming language we used is R, and the IDE we worked in is RStudio. Upon seeing the three different CSV files given, we decided to mainly work with the “needs.csv” and “shelters.csv.” From here, we looked further into narrowing the data.

We wanted to focus mainly on the Houston area for our analysis. After throwing out the null entries we were left with 91 for needs and 193 for shelters. We used someone’s Google API to determine the narrowed latitude and longitude range for the Houston area. The approximate location of central Houston is 29.76, -95.37. Thus we decided the range for latitude as (29.5,30.0) and the range for longitude as (-96.0,-95.0).

The color-coordinated Map of Needs displays the 91 entries on a segmented map of Texas. Red (41) signifies a location where both supplies and volunteers were needed; green (44) signifies a location where only supplies were needed; orange (5) signifies a location where only volunteers were needed; purple (1) signifies a location where neither were needed. While there was a need for volunteers, there was a greater need for supplies considering 44 locations were supplies only.

The color-coordinated Map of Shelters displays the 193 entries on a segmented map of Texas and Louisiana. Red (109) signifies a shelter where they were no longer accepting victims; blue (84) signifies a shelter where they were still accepting victims. It is confirmed from the map that the shelters in the Houston metropolitan area rapidly filled, while some of the more outlying shelters still held space. It is also confirmed from the map that although the outlying shelters had space, they most often needed both supplies and volunteers.

The bar graphs correspond to the respective Map of Needs. We can observe that 46 locations needed volunteers and 85 locations needed supplies in total. When limiting our observations within our latitude and longitude ranges, 22 locations needed volunteers and 38 locations needed supplies. When going outside our ranges, 16 locations needed volunteers and 20 needed supplies. It is evident, then, that supplies were in higher demand than volunteers both within our restricted Houston area as well as outside.

Overall, we can conclude that the focus of relief preparation and efforts needs to be on supplies more than volunteers. Our data tells the story of a city and area that did not evacuate all of their people, yet were flooded internally and cut off from much of their supplies.