

|  |
| --- |
| Tornados and Trailer Parks |
|  |
| May 25  Coursera Capstone Project  Authored by: Glenn Meisenheimer |



|  |
| --- |
| Tornados and Trailer ParksOr the Safest Place in the USA to Build a Trailer Park Introduction  It is a well-known fact that when a tornado touches down the first thing it does is ask for directions to the nearest trailer park. This is an unfortunate thing for people who live in those trailer parks (as I did when I was a kid in Kansas).  But there is a dataset out there that contains statistical information on every tornado in the US from 1950 to 2015. So it occurred to me that this data might be useful in determining the best place to build a trailer park in the US.  In addition this information could prove to be invaluable to people who already own mobile homes and are looking for the safest place to park them.  So this study is designed to answer two questions for two different groups of people:   * For mobile home park developers:, “Where is the safest place in the US to build a trailer park?” * For mobile home owners:, “What is the safest trailer park in the US in which to park my mobile home?”   The Data |
| The data for this project came originally from the National Oceanographic and Atmospheric Assoaciation (NOAA) which keeps records on all severe storm activity in the United States and makes the resulting data available through their website. This particular dataset was downloaded from NOAA, cleaned up and made available as a .csv file from kaggle.com.I took that .csv file and imported it into a pandas dataframe which you can see a portion of above. The fields in the dataframe are: om = tornado number within a given year  yr = year  mo = month  dy = day  date = date in MM/DD/YYYY format  time = time of tornado 'birth'  tz = time zone  st = State  stf = a numeric code for the state (we'll drop this column as we don't need it)  stn = the number of this tornado in it's home state in a given year  mag = magnitude of tornado on the EF scale (EF0 - EF5)  inj = number of injuries  fat = number of fatalities  loss = property loss where 1 <=$50,2<=$500, 3 <=$5,000,4<=$50,000,5<=$500,000 and so on.  closs = crop loss  slat = starting latitude  slon = starting longitude  elat = ending latitude  elon = ending longitude (not to be confused with Elon Musk)  len = length in miles  wid = width in yards  fc = a composite code that contains info on multiple states affected. Earlier entries lack this so we will drop it.  For the sake of the map plots I created a smaller dataset – a subset of the data in the .csv file. This smaller dataset contains only the starting latitudes and longitudes of each tornado. But there are a lot of tornados in the data therefore before plotting we will use k-means clustering to group them into clusters of 120 tornados and only plot the center points of those clusters. I didn’t bother with label data for the points since they are going to be clustered before being plotted on the map.  Methodology  Let’s see what the data points look like if we put them all in a matplotlib scatter chart:    Just as we expected that is pretty dense data. It’s hard to infer anything from it. So let’s use k-means clustering to group and calculate the centerpoints of 500 clusters each containing about 120 tornados:  After running k-means we take the centers of each of the clusters and create yet another dataset. But this time the data won’t be so dense. A plot of that data looks like this:    Good, that data makes a lot more sense. Let’s create a folium map of the US and plot the latitudes and longitudes of each of these cluster centers on the folium map:    That looks more presentable. But I’m still a little nervous about making a decision based on this data alone. The problem is that there is a certain amount of inherent randomness in the way the k-means clustering algorithm works. In the end different tornados can wind up in different clusters each time k-means is run which means that the cluster centers might move around a little each time we run the algorithm.  Generally speaking it is true that the areas that have the greatest cluster density have the most tornados and areas with sparse cluster density have fewer tornados. But I think I would feel more comfortable about my decision to park my trailer somewhere if we could add a couple other parameters to this map.  Specifically I’d like to see where all the really big (EF5) tornados are and some indication of where the most tornado fatalities have been. That data is in our original dataset so let’s pull it out and overlay it on this map of cluster centers.    Results  That’s better. That map gives us a better feel for where the high risk areas are and where the low risk areas are. Generally speaking it looks like we should avoid the Midwest and the South. The Pacific Northwest looks a lot better but the area of lowest activity seems to be in South central Nevada.  In fact I would say that the area about halfway between Las Vegas and Sparks would be the safest place to build a trailer park or to park my own mobile home. And there is a small town there too. If we zoom into the map we find the town of Tonopah, Nevada.  So that answers the first of our two questions. Tonopah, Nevada is the safest place to build a mobile home park. But to answer the second question, “Where is the safest trailer park in the US to park my mobile home?”, that’s going to require a Foursquare search to determine. We need to search Tonopah, Nevada for trailer parks.  And this is where we get into trouble: |
| Discussion It turns out that Tonopah, Nevada is a small town in the middle of a big desert. Foursquare limits the search radius to about 75 miles. So in total they only list 18 venues in the 75 mile radius surrounding Tonopah. And none of those are trailer parks. But there is still Google. When we Google “mobile home parks in Tonopah, Nevada” we find that there are two RV parks and YES there is one trailer park in Tonopah – the safest trailer park in America.  It would have been nice if Foursquare allowed a larger search area for isolated communities like Tonopah. In fact it would have been nice if the trailer park we found in Tonopah had been registered with Foursquare.  But although there were shortcomings in the data and hiccups in our methodology, once we got the data all visualized properly the answers to our initial two questions were pretty clear.  But please consider these caveats:  The cluster centers are to a degree subject to the whimsy of fate because you can never be sure which tornados k-means will group into which clusters and that will make a difference in where the cluster centers fall. But areas of high tornadic density will always have denser center points and will carry heavier risks than areas of sparse density.  Using the EF5 rating is problematical because an EF5 rating implies that the storm had winds over 300 mph. Yet it is extremely difficult to measure wind speed inside a tornado from a distance. You have to get right up to the base of them with not one but two doppler radar units looking from different angles so that their data can be triangulated. Therefore the EF5 rating is only given after evaluating the damage the tornado left in it’s wake. This means that if you have a tornado with 300 mph winds out in the middle of farmland and it leaves no damage in it’s wake it gets a rating of EF0 not EF5. So the EF5 rating by itself is not really a good enough indicator with which to assess risk.  And as for fatalities that is problematic as well. People who live in Tornado Alley and know what these storms are capable of are typically better prepared for tornados than somebody living in a trailer somewhere that doesn’t get nearly as many tornados. So fatalities can actually be higher in areas of lower tornadic activity.  But nonetheless, if these three parameters are taken together and plotted on a map it is pretty clear which areas are high risk and which are safer. It would be safer to park your mobile home anywhere in the Pacific Northwest than it would be to park it in Oklahoma.  Conclusion  This study has been a bit whimsical. The whole “tornados and trailer parks” correlation is a bit of a joke. But when I was a kid I lived in a trailer park in Kansas. When I was in second grade I was home alone one night when a tornado came through. There was no place for me to run so I had to just hunker down in the trailer which was rocking back and forth wildly in the wind. I could hear the tornado as it roared past. Or at least I convinced myself that I could hear it. It was a terrifying experience to go through.  So in conclusion let me state that there are trailer parks in America which are safer than others. And based on the analysis done in this paper I believe we have found the safest trailer park in the United States.  The safest trailer park in America from the viewpoint of tornados is:  Jensen's Trailer Park  789 Central St  Tonopah, NV 89049  Phone: (775) 482-3737  So there you have it campers! Enjoy a carefree tomorrow in beautiful Tonopah. Just watch out for the rattle snakes... |
|  |
|  |