

U.S. Department of Transportation's Fatality Analysis Reporting System (FARS) Visualizations Demo

Transcript of FARS Visualizations Demo Video

Speakers: James Hove

Hello, my name is James Hove. I'm a social scientist at the United States Department of Transportation. Today I will show you some Tableau visualizations we've put together at the Department as part of the Safety Data Initiative, and we'll be using data from the Fatality Analysis Reporting System otherwise known as FARS. So, that data set I just said is basically just the Department's flagship data set for reporting fatalities on U.S. roadways.

To start, we're going to be looking at the visualization in front of you, which is just looking at speeding related fatalities on U.S. roadways in 2016 specifically. So obviously already you can see there's places for improvement - we certainly could look at a larger data set or add in other data sets. But for the first step of this process we just went with the simplicity of looking at this limited data set. So, this tool, you know, it's meant to reach large audiences, so the hope is that not only more technical folks at DOTs, but also news organizations and other transportation stakeholders could find this useful. So, we're trying to make it a wide-ranging product but there's certainly room for more niche products that reach out to specific groups and target specific needs that aren't necessarily just the content within this visualization. So, the hope of the overall project was to make the data that exists at the Department right now, namely this data set, more accessible to users. And the hope is that these visualizations start us down that path of making this information a little bit more useful.

So, to start, I will jump right in. So, as I described this as a visualization dashboard that looks at traffic fatalities that are specifically speeding related in the United States. Right now, right off the bat you can see percentages, this is simply the percentage of traffic fatalities within a given state that are speeding related. You can already see, you know, information right in front of you just based on color, and you can choose to drill down as you please, but I'll start over here. The nice thing about Tableau is its dynamics so people will get information on data right away. The first thing here, is you can just hover over one of the text box here and get a lot more information about the visuals. So you hover here you'll get information about how the dashboard works, what to do what not to do. And then this is certainly another area for improvement you know making this dashboard as user-friendly as possible so that people who are new to this process can easily access this information. And maybe clean up some of the information so that it's geared towards someone who hasn't used the product before.

So, we'll get right into the actual use of the dashboard. As you can see, these are all data visualizations. So this is, what I'm looking at here, is time by hour of day, and this is once again, looking at the percentage of fatalities that were speeding related, and then this is a simple bar chart for the states represented above. This is another bar chart that looks into information based on what type of roadway these are occurring on. And then, up here we have our filters and what we use the most to narrow down the information; they're actually visualizations in and of themselves, so the darker color will indicate you know higher percentage of fatalities being speeding related. So these filters being dynamic in color can kind of help guide your research. So say we wanted to dive a little deeper and we saw that, well, it looks like, you know, occupants are more likely or likely to be speeding related, so you can click on that. And then, once you click on something it will actually update the entire visualization to represent specifically occupant fatalities, so otherwise someone who is inside of the motor vehicle. So you can click on that, and then you say, "okay, well also looks like the first, you know, six hours of the day are pretty heavy", so we can further filter down at that point and get a little bit more information. One other interesting thing is, you'll see based on these filtering you have dynamic text. So, you can hover over part of it or a

little piece and this text will update as you as you go to different pieces, and will tell you exactly what you're looking at. We tried to put this information in plain text so that it would be easily digestible as opposed to, you know, statistics at times sometimes get spit out of these are processes. So, that's a cool way to filter and you'll notice all the visuals updated just to show the information we now have highlighted, which would be occupant fatalities specifically between the hours of 12:00 to 6:00 a.m. or 12:00 to 5:59 to be more specific, and each of these will update as you as you look at them.

So, another thing we added to make it simpler for people so that if they don't remember what all the filters they've gotten into, they can simply click, "remove filters," and that will just very simply clear out all of your selections, so you can start over from square one, so that it's really simple and you know what you're doing. Similarly, you're going to have text like this that can also be dynamic. So, if say you didn't want to highlight here and you just want it all things in the a.m. to make it more convenient, you can also just click the headers for where they exist to filter down for more specific information to make it a lot easier on yourself.

So, I've already talked about tool tips, but it is important to know that there is also like dynamic text throughout the visualizations. So, for example, right here beneath the map you'll see this is telling you, you know, how many fatalities there were and how many are speeding related fatalities there are. So, you get this information, and then as you click on different filters that information will update so you always know exactly the universe of information you're looking at. So hopefully that helps to make the process a little bit easier.

So one of the main reasons we wanted to put this here too was so that people could do regional analysis. So, say if a state DOT or local government was interested in looking at information that affected them, or even a local news outlet, they could easily use this map or this bar chart to suss out specific information. So I'll start by the bar chart down here. The bar chart is mostly put together so that, unlike the map, you can see them kind of ranked next to each other, so if you wanted to look at the top or bottom or specific subject you can see if there's any regional trends. So let's say we want to grab just the top few...you can highlight over the first few bars and you'll see where those are located. It looks like, in this case, it's mostly the north, northeast here, and then also Alaska and Hawaii. And once again, even the coloring, like I mentioned, is dynamic so now you have these colors updating in relation to each other. So, although these used to be darker red, now when you're just comparing these seven examples you'll see that they're only a few of them are dark red because New Hampshire and D.C. stand out so significantly up above the other options. So, all this is dynamic so you can continue to drill down to contribute to see trends as you go.

So, another few things in the map. I will just click anywhere here to filter...it takes a second sometimes if you do a lot...but, so there we removed the filters. And now we also have it set up so that you can look at, not just the map of the whole state, but you can actually look at individual incidents. So, while this is looking at all fatalities, we have a map that if you just click the crashes button here, that will show just crashes. So this, they could have multiple fatalities, but each one of these dots is a crash that occurred in 2016 that was speeding related. So this is really interesting if you want to look at a more granular level, or say if you were a local government and wanted to see what's specific to you, you can zoom in and zoom out, which I will show a little bit later. So, I just wanted to show that another thing associated with that, is that you can hover. And once again, the tooltip here I discussed, see if you hover, you can get really detailed information about exactly what happened, when it happened, which can be really helpful. And then even more helpful, if you click on one you'll get the option to do a Google Street View—so you can actually click the street view and it will pop up in a web browser and show you the

street view of that exact incident, which, once again, can be really helpful if you're doing a crash investigation and that kind of thing. I can see if I can get one to load here, but yeah, so this review can be really useful if you're doing specific research on specific incidences if you're if you're going away from the data approach. So you'll see this an exact point an exact roadway where if a fatality did occur during 2016, so definitely useful if you're interested in that side of things.

So, as we come back to the map here, I'll go back to the state information. So if you switch back it'll just bring it right back to the state, and then one more cool feature is that if you're interested in say, a region, and you really want to dive in deeper and the statewide isn't granular enough for you, you can also like hover over a few states or just one state, whichever you choose, and you can actually click for a further breakdown similar to the Google Street View. So if you click that breakdown, you'll likely get brought into a new dashboard altogether, which is linked and you can go back and forth between the dashboards, but now you'll see it's all filtered for what I had selected and now it's at the county level, so you can even dive a little bit deeper and try to find trends. So here we see, it looks like Florida and Georgia have kind of relatively low speeding related fatalities, whereas once you get up into North Carolina, South Carolina it goes up a little bit so you can start to ask yourself questions about why that might be. Then lastly on this one, I'll even show you a further breakdown just so this simplicity of it. If you, say you want to click on a specific county that looks kind of high, it'll zoom right in there and it looks like this one only really has one fatality, so it might not be a very good data point, but it allows you to zoom in and you can see exactly where it is, see where the roadways are, so that should help you make it a little bit more useful.

So that's... I will wrap it up here with this visualization, because we have one other one. And you can go back to the main dashboard, which is still filtered, and you can still look at all these different visualizations with the filter, but obviously there is certainly room for improvement or ways to build off of what we've done. A few would be, maybe, as I mentioned earlier, making it a little bit more user-friendly. I know right now it is a little complicated and all the instructions aren't 100% clear, so making something that's a little bit more user focused, possibly adding more data sources or new data sources even expanding as a whole. Obviously speeding related crashes in just 2016 isn't necessarily that large of a sample, so you could expand past that and get a little bit more information. Certainly can modify functionality, maybe there's other variables that people find more important, so there's definitely a lot of different directions we could go with this from here so. But, I will leave it at that for this visualization and show you quickly the other visualization we have put together at the Department.

So another one we have is here. This one is in relation to and it is a little...it is a little bit different dimensions that it will appear online, because I downloaded onto my desktop for this presentation. But as you can see... so based on the title, this is pedestrian traffic fatalities and it is a longer range, so this is a newer visualization that we added a few more things to, but basically what we're looking at are pedestrian fatalities/traffic fatalities so any fatality that was a pedestrian from 2011 to 2017 in the United States. So, with a much larger sample, mostly because it is a smaller size of data points as well. So here we were looking at that, and there's a little bit of some more functionality. I won't talk about the functionality I talked about in the previous ones, but I'll show you some other things we've done and maybe some areas to improve so. As you can see here, we've been looking at it by year now as well, because we have more than just one year of data and then also we added some visualizations within that text that pops up. So here, now if you hover over an individual year you're actually going to get data that shows you by month. So within that year, you'll see that it seems to spike what seems to be the end of the year, the winter months possibly, so maybe something to look into. So that's there, and then the hypertext still exists so you get these, you know, sentences that make things a little bit clearer.

And here, as you can see, we also have a heat map. So this is a heat map of gender and age. So you're saying that, you know, "males from 51 to 55 are an especially high category of pedestrian fatalities," so maybe that's something to look into. And the filtering is all the same with the same remove filter options and just like the other one you also have these links to give us feedback or to directly access the source so people can feel more connected with the data. Another thing, I'll show you in the heat map here is that not only did we provide a heat map over here, but we also added some buttons where you can change the actual visualization you're looking at, so maybe you're not interested in what we gave you so maybe we're interested in, you know, month of year, hour of day possibly, right? So right now we're in month of year by year, let's do month of year by hour of day. So if you click that it'll just update here and I'll actually give you your heat map and then right away we can see those hours, those commuting hours in the winter seem especially dangerous; you know, 579 pedestrian fatalities versus three months earlier/four months earlier it's only 67. So you can start to look at those trends and find interest there and dive a little deeper. And once again, we have that the same option for pulling up a completely different map or a completely different visualization right here at the bottom. So right now it's looking at vehicle type, how many pedestrians are killed by individual vehicle types, but say you wanted to look at road condition or light condition instead—road type and light condition, that's also an option. So this gives you a little bit more functionality than we previously allowed, and in the previous visualization tryin to build on that.

And then a really large change we made, one of the last things I'll discuss here, is that we actually allowed, in the last visualization here remember we looked at the percentage of fatalities that were speeding related, right now we're looking at just the number of pedestrian fatalities. So instead of a percentage it's just a raw number, but actually we have three different options you can choose from for what number the dashboard represents. So like the other one, if I click this it will give us the percentage of all fatalities that were pedestrians...so this allows you to change the entire dashboard to represent a different measure, which is very interesting and can lead to different insights so it really gives you more functionality. Once again, there's areas to improve in ways we communicate this information or how the user accesses it but we're definitely trying to give as much information as possible here. And then, so this obviously lights up completely differently now there are things that, you know, as a percentage are completely different, and then even more interestingly we integrated census data as well into the visualization so now we can look at pedestrian fatalities per capita as well. Obviously some of the visualizations which we don't have per capita data on will go away and change, but if you click on that we're going to get completely different values which might be more useful. So you know, point pedestrians are dying per population size might give us better information on what to drill down on.

So, as I mentioned some visualizations went away when we did this, but we actually just changed them out to look at some of the top U.S. cities considering pedestrian fatalities are especially bad in U.S. cities. We kind of gave some cities here and as you'll notice the color represents how many fatalities there were, and then that/the circle size is giving you that per capita value. So you know, red meaning a lot of pedestrian fatalities, but as you can see in New York it's a smaller circle so maybe not as high per capita; whereas, Detroit there's not as many pedestrian deaths, but the per capita value is quite high. So, this is just another way to look at the data, and you'll see that the scatterplot at the bottom is actually linked to these—so as you hover, you can kind of see where that point lies in in the grand scheme of things...so another way to view the data.

So once again, this is kind of a starting point for a way to bring the data to people and allow them to better access it. So, we're definitely interested in ways to improve and how to go from here but this is our first stab at bringing one of, you know, our most important data sets and now with this one also

including census information, to better communicate some of our insights over some possible insights we haven't considered yet to people. And then once again, I'll show you before we close, the crashes here as well so you can filter between the types you're looking at.

And then, the last thing I will say before I sign off here is...as the crash map loads here—there's a lot of data points for pedestrians especially over this many years so it takes a while to load...but here you can see the actual distribution, which is quite extensive, once again giving you information, and once again you can filter for a street view. But, one last thing I'll show is the original map, you might be wondering, the original map we were trying to make sure that the insights were evenly conveyed, and we noticed in the first map it's kind of hard to distinguish between some of the smaller states and to actually (and especially D.C.) to see the relative distribution of those, so we actually changed to this hex map so that every state is roughly in the location it would be in a normal map—but having it in this hex shape will allow the standardization of sizes you can actually see every state. We were hoping this would give a better representation and give a better way to actually view those insights. So certainly, we're open to moving forward on improving this and making the whole experience better and to bringing this data and more data to people more effectively. So, I will sign off there and I appreciate you taking the time to listen.

###

Contact Information:

The Safety Data Initiative Team

U.S. Department of Transportation

Email: USDOTSafetyDataInitiative@dot.gov