DS104-10 LESSON 10 FINAL PROJECT:

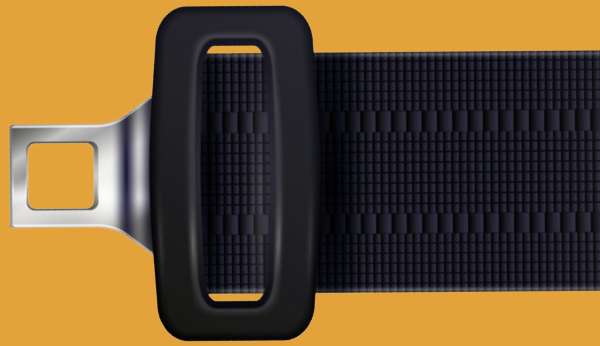
MOTOR VEHICLE CRASHES IN TEXAS IN 2013

HEATHER WALKER - 2022-10-23

TABLE OF CONTENTS

FINDING ADDITIONAL DATA

01



My first step in understanding the data was to open the .CSV file in Microsoft Excel and take note of the *dimensions* included in the dataset.

Upon inspection, there were several *dimensions* that were abbreviated or assigned an ID number with no key.

GETTING CONTEXT

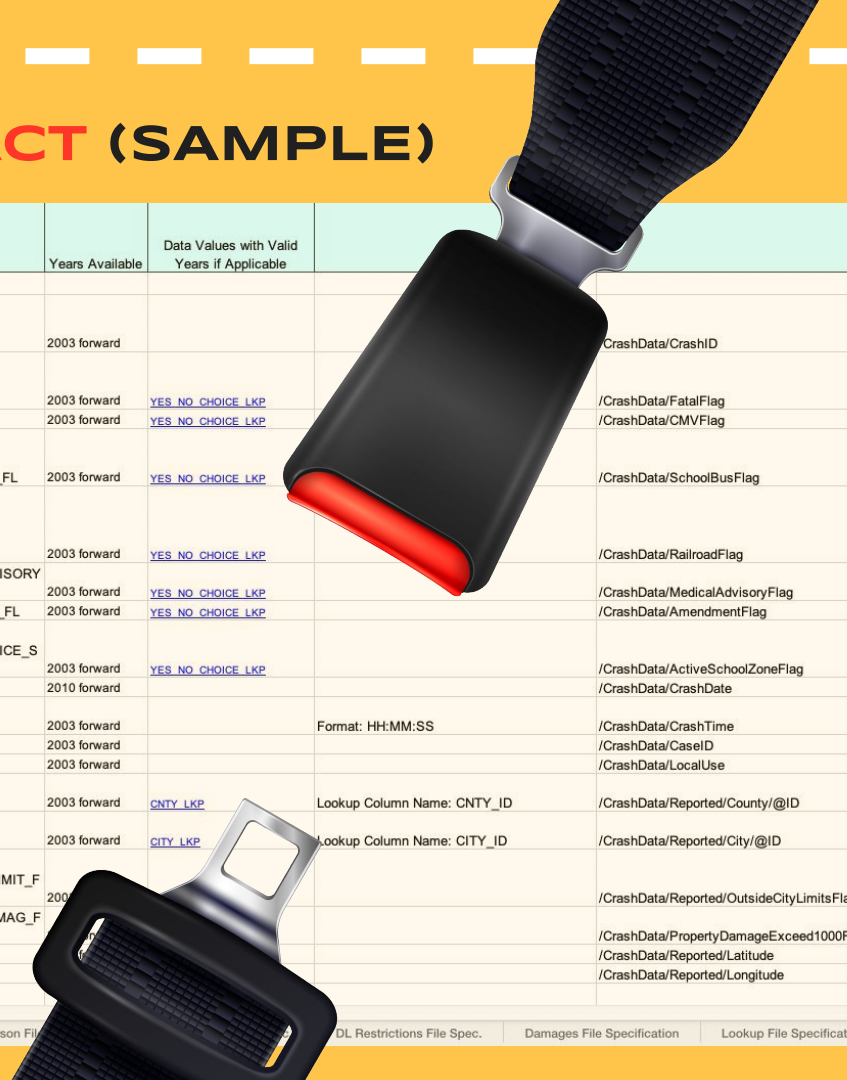
By searching for Texas Department of Transportation data for the year 2013, I was able to find a reference document that included a reference guide for abbreviations.

Texas Department of Transportation public extract file specification (download link)

TEXAS D.O.T. PUBLIC EXTRACT (SAMPLE)

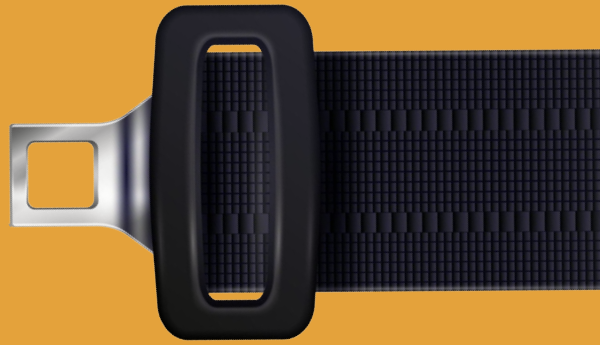
Column Header	Description	Database Data Type	Extract File Data Type	Max Characters	Source	Years Available	Data Values with Valid Years if Applicable		
	CR-3 Reported Data Fields								
Crash_ID	Crash ID – System-generated unique identifying number for a crash	INTEGER	Integer		CRASH_FACT.CRASH_ID	2003 forward			CrashData/CrashID
Crash_Fatal_Fl	Fatal Crash Identifier - Indicates that the crash involved one or more fatalities	CHAR(1)	String	1	CRASH_FACT.CRASH_FATAL_FL	2003 forward	YES_NO_CHOICE_LKP		/CrashData/FatalFlag
Cmv_Involv_Fl	CMV Crash Identifier	CHAR(1)	String	1	CRASH_FACT.CRASH_CMV_FL	2003 forward	YES_NO_CHOICE_LKP		/CrashData/CMVFlag
Schl_Bus_Fl	School Bus Crash Identifier - Indicates whether the crash involved a school bus	CHAR(1)	String	1	CRASH_FACT.CRASH_SCHOOL_BUS_FL	2003 forward	YES_NO_CHOICE_LKP		/CrashData/SchoolBusFlag
Rr_Relat_Fl	Railroad Crash Identifier - Indicates whether the crash involved a train or railroad crossing.	CHAR(1)	String	1	dorse	2003 forward	YES_NO_CHOICE_LKP		/CrashData/RailroadFlag
Medical_Advisory_Fl	Medical Advisory Board (MAB) Crash Identifier	CHAR(1)	String	1	CRASH_FACT.CRASH_MEDICAL_ADVISORY_FL	2003 forward	YES_NO_CHOICE_LKP		/CrashData/MedicalAdvisoryFlag
Amend_Supp_Fl	Supplement Crash Identifier	CHAR(1)	String	1	CRASH_FACT.CRASH_AMEND_SUPP_FL	2003 forward	YES_NO_CHOICE_LKP		/CrashData/AmendmentFlag
Active_School_Zone_Fl	Active School Zone Crash Identifier - Indicates whether the crash involved a school zone.	CHAR(1)	String	3	YES_NO_CHOICE_LKP.YES_NO_CHOICE_SHORT_DESC	2003 forward	YES_NO_CHOICE_LKP		/CrashData/ActiveSchoolZoneFlag
Crash Date	Crash Date	DATE			CRASH_FACT.CRASH_DATE	2010 forward			/CrashData/CrashDate
Crash_Time	Crash Time - Time crash occurred	TIME	String	8	CRASH_FACT.CRASH_TIME	2003 forward		Format: HH:MM:SS	/CrashData/CrashTime
Case_ID	Case ID	VARCHAR(20)	String	20	CRASH_FACT.CASE_ID	2003 forward			/CrashData/CaseID
Local_Use	Local Use	VARCHAR(20)	String	20	CRASH_FACT.LOCAL_USE	2003 forward			/CrashData/LocalUse
Rpt_CRIS_Cnty_ID	County Name - The county in which the crash was located.	INTEGER	Integer		CRASH_FACT.RPT_CRIS_CNTY_ID	2003 forward	CNTY_LKP	Lookup Column Name: CNTY_ID	/CrashData/Reported/County/@ID
Rpt_City_iD	City Name - The city in which the crash was located if applicable.	INTEGER	Integer		CRASH_FACT.RPT_CITY_ID	2003 forward	CITY_LKP	Lookup Column Name: CITY_ID	/CrashData/Reported/City/@ID
Rpt_Outside_City_Limit_Fl	Outside City Limit - Indicates if the officer reported the crash location outside a city's limits	CHAR(1)	String	1	CRASH_FACT.RPT_OUTSIDE_CITY_LIMIT_FL	200			/CrashData/Reported/OutsideCityLimitsFla
Thousand_Damage_Fl	$1,000 Damage to any One Person's Property	CHAR(1)	String	1	CRASH_FACT.CRASH_THOUSAND_DMAG_FL				/CrashData/PropertyDamageExceed1000F
Rpt_Latitude	Latitude	DECIMAL(11,8)	Decimal(8)		POINT_LKP.LATITUDE				/CrashData/Reported/Latitude
Rpt_Longitude	Longitude	DECIMAL(11,8)	Decimal(8)		POINT_LKP.LONGITUDE				/CrashData/Reported/Longitude
	Roadway System (road on which								

Introduction | Crash File Specification | Unit File Specification | Person File Specification | Charges File Specification | PrimaryPerson Fil | DL Restrictions File Spec. | Damages File Specification | Lookup File Specificat



PRELIMINARY DATA EXPLORATION

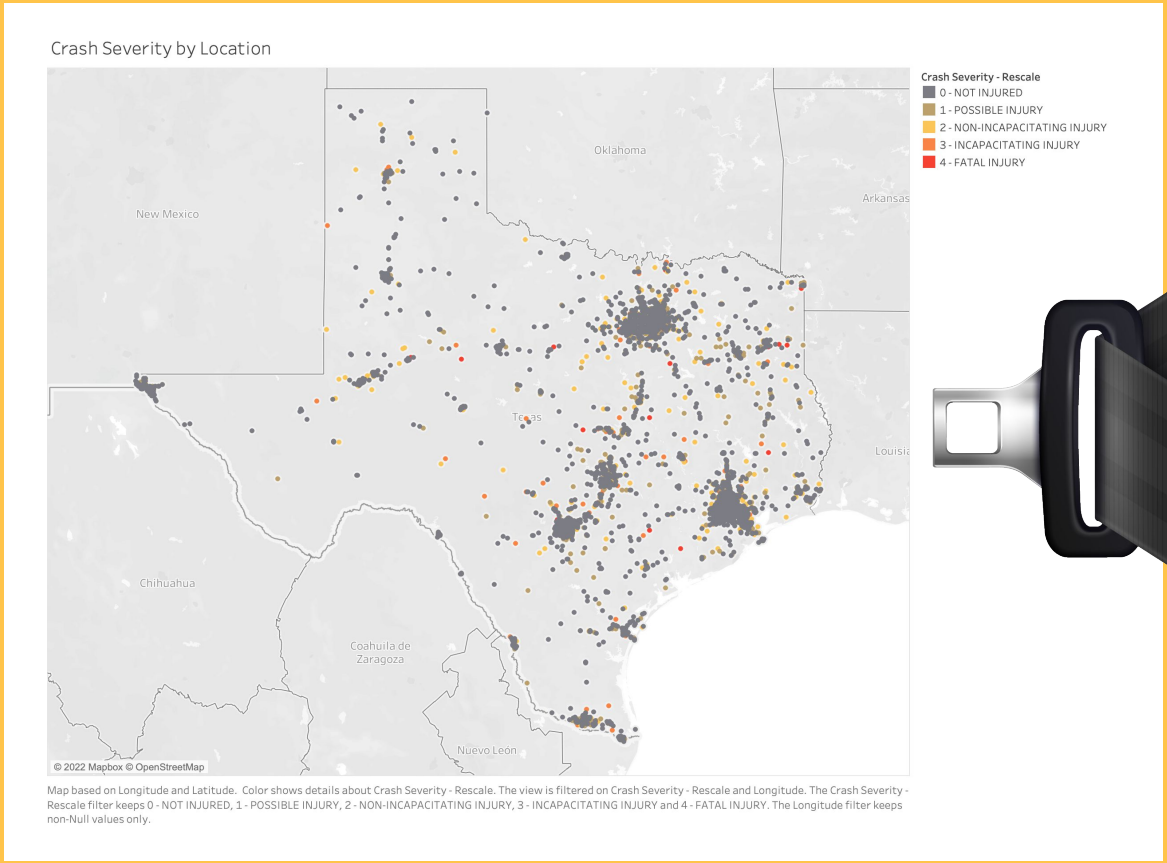
02



MAP

The first data visualization that made sense to create was a map of the crashes.

Here, I've also color coded by the severity of the worst injury sustained by anyone involved in the crash.

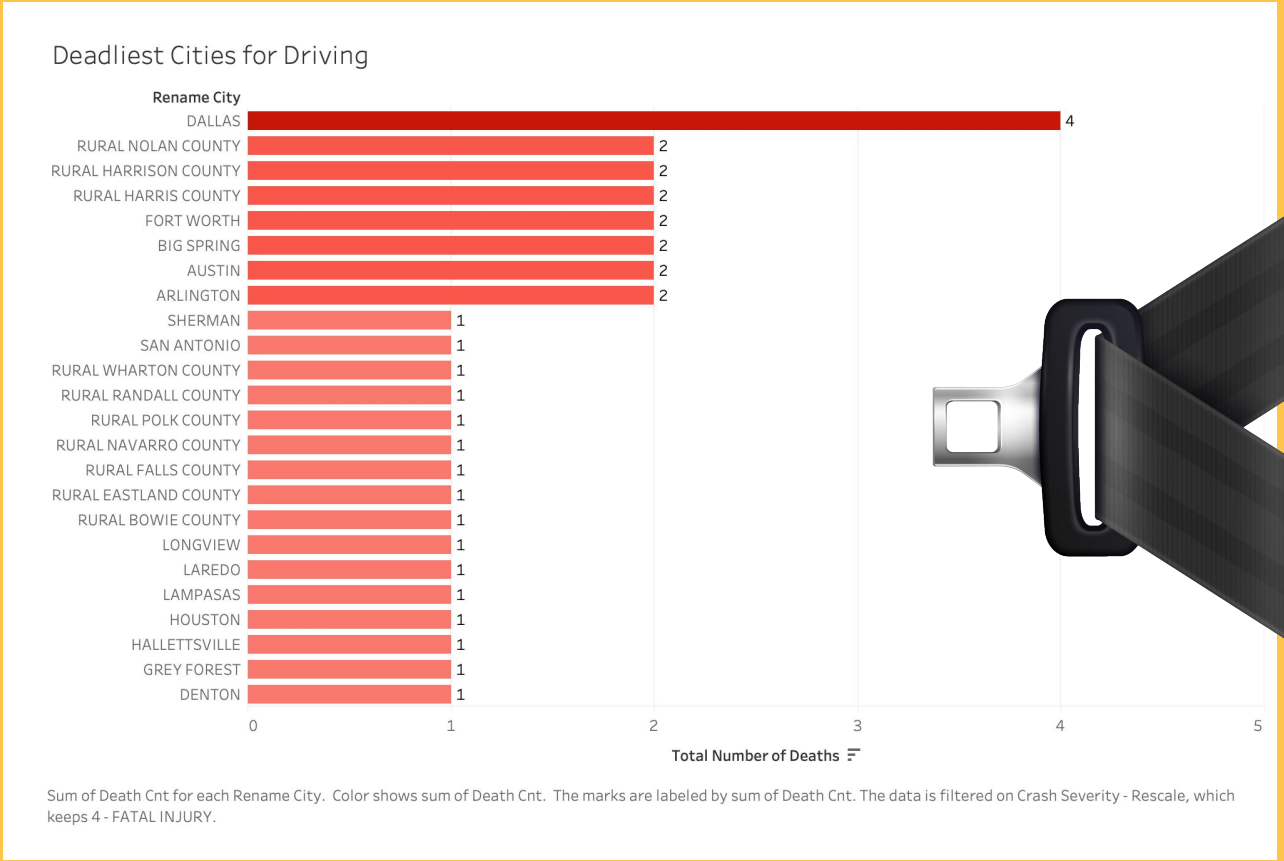


VIEW THE VISUALIZATION HERE (LINK)

FATALITIES BY CITY

This bar graph shows the total count of deaths per city.

I related the *City ID* dimension to the *City Id* column from the Texas DOT spreadsheet to create this visualization.

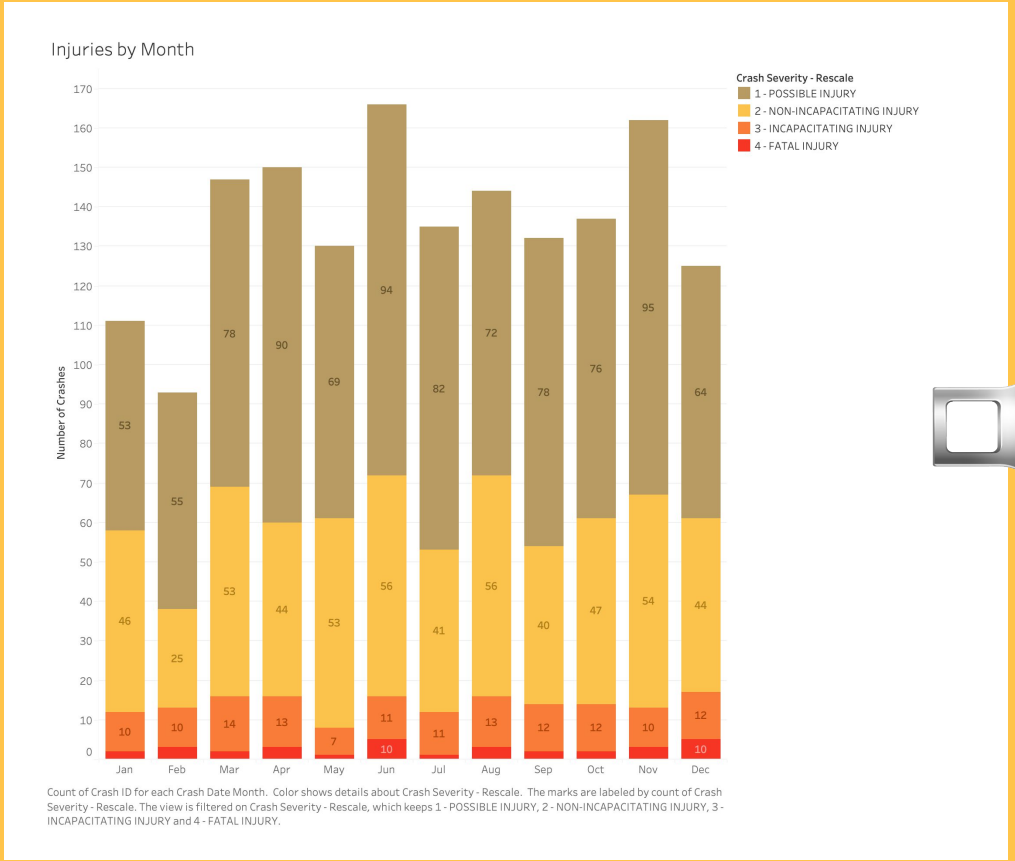


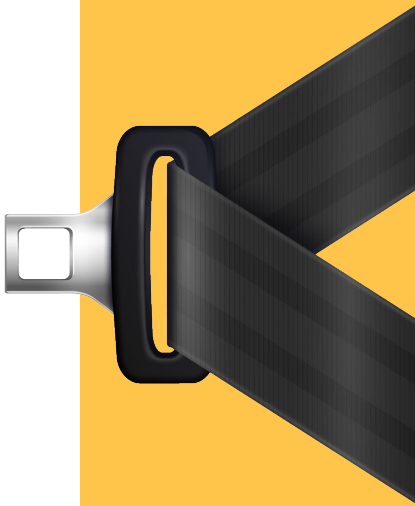
VIEW THE VISUALIZATION HERE (LINK)

INJURIES BY MONTH

Question: Does time of year make a significant amount of difference in crashes or their severity?

This stacked bar graph shows the proportion of injuries sustained during crashes during certain months.

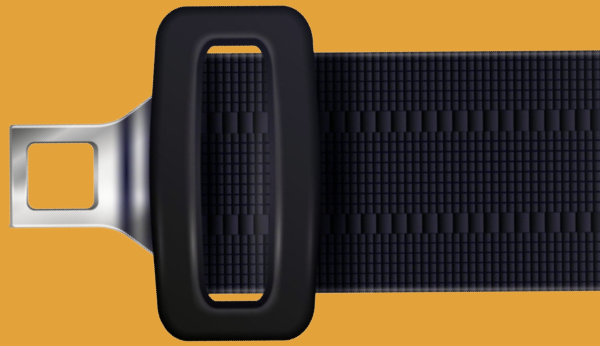




VIEW THE VISUALIZATION HERE (LINK)

MORE IN-DEPTH DATA EXPLORATION

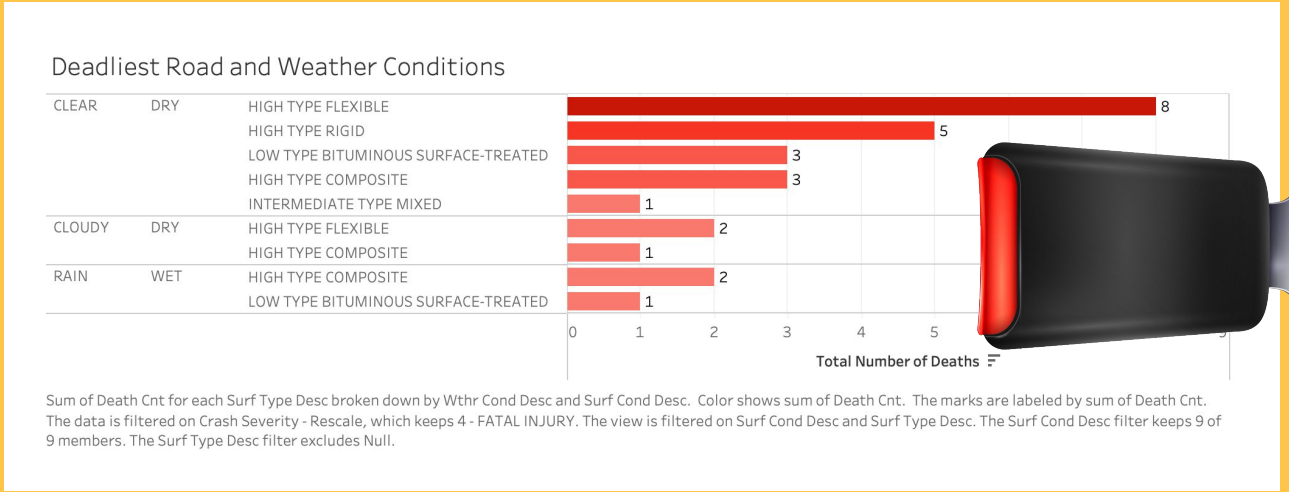
03



ROAD AND WEATHER CONDITIONS

Question: How does *Road Surface Type*, *Road Surface Condition*, and *Weather* affect crashes?

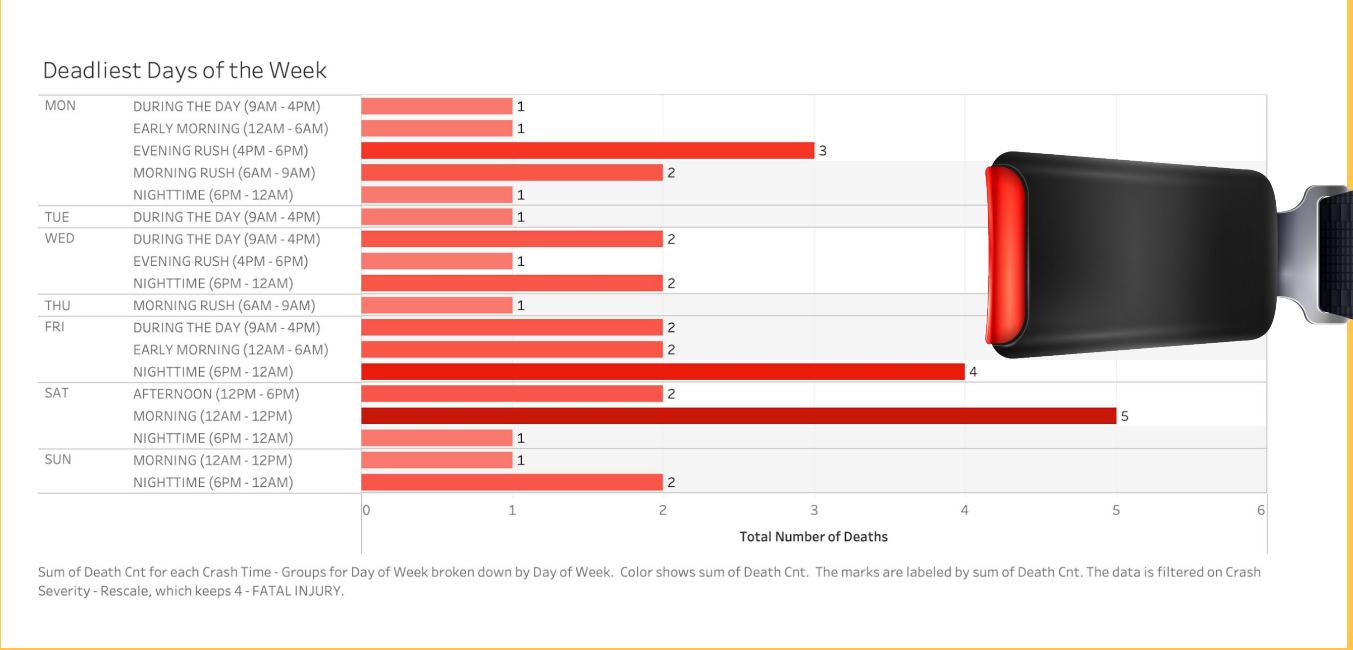
This is a more detailed grouping of the road conditions vs. weather from the dataset.



VIEW THE VISUALIZATION HERE (LINK)

DEADLIEST DAYS OF THE WEEK

Creating this visualization required a little more behind-the-scenes action, from converting the time from string to time format, and then grouping by time of day. The time of day groups are also different based on weekday vs. weekend.

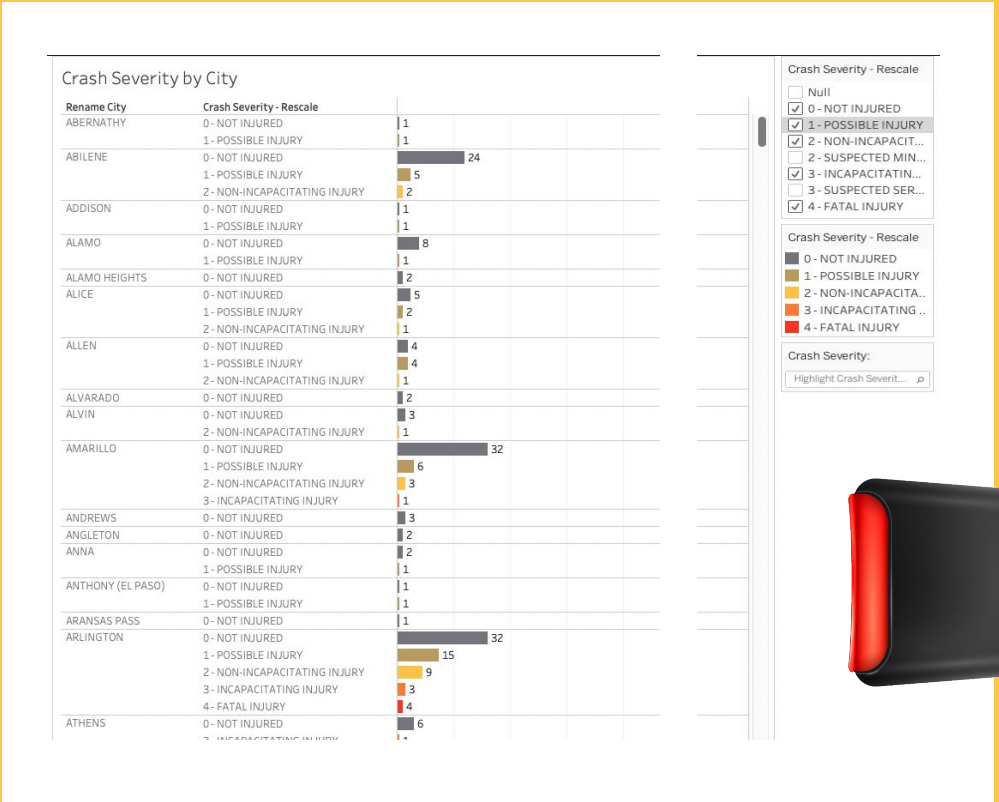


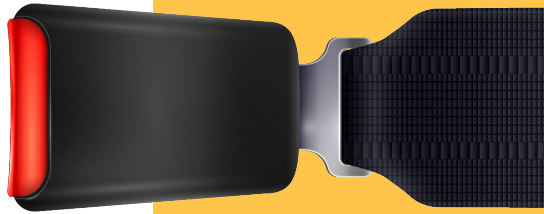
VIEW THE VISUALIZATION HERE (LINK)

CRASH SEVERITY BY CITY

Here, Texas cities have a tally of the count of crashes of each severity level.

(Critique: needs more filtering / specificity, currently an unwieldy size)

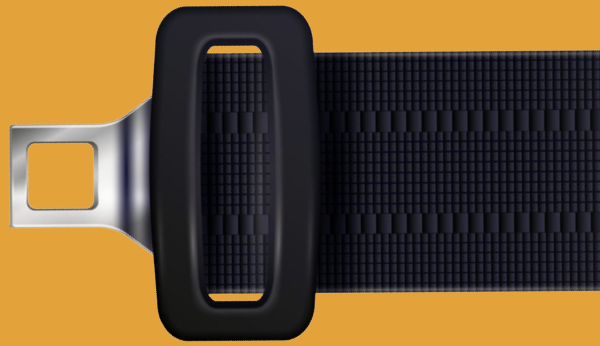




[VIEW THE VISUALIZATION HERE (LINK)]

FINAL THOUGHTS

04



I have a bad habit of getting excited about the more nuanced, detailed, or complex analyses available with such a large set of data.

An example: one of the more complex tasks I created for myself was wanting to "rescale" the crash severity numbering for color coding and filtering. I accomplished this by writing a function, after researching the correct syntax for concatenating measure names and strings.

WHAT I LEARNED:

- Creating a data extract
- Creating various relationships, joins, and aliases
- Creating custom color palettes
- Writing more complex functions in Tableau
- Using Parameters and Highlighters
- Exploring Pages as a way to add detail to a viz
- Start simple. It's easy to get off on a troubleshooting or more-detailed-visualization-option tangent!

BE SURE TO CHECK OUT THE INTERACTIVE PROJECT

LIVE ON TABLEAU PUBLIC!

INCLUDES:

- Filters
- Highlighters
- interactive Tooltips

THANKS!

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik**