CGT270
Midterm Part II

Data Visualization Challenge

Student:

<mark><Ria Rajan></mark>

Professor

Dr. Vetria Byrd

TA

Subia Ansari

Course:

CGT 27000 Data Visualization

Term:

Fall 2021

Contents

Halloween Visualization

This in-class assignment is to create data visualizations using data collected about trick-or-treaters in Cincinnati, OH. You should create two (2) visualizations, this can be a collection of charts or a dashboard, whatever is necessary to the story or analysis that is shown in your visualizations. Make sure you READ and FOLLOW ALL Instructions. The goal is to demonstrate your understanding of the data visualization process.

Data Description

The data is available in two formats

- Halloween data for Excel 2020 is a crosstab table which is ideal for creating visualizations in Excel. Numbers in the data file for Excel are **cumulative**.

- Halloween data for Tableau 2020" is unpivoted which is ideal for creating visualizations in Tableau. Numbers in the data file for Tableau are **not cumulative**.

- The data has been collected since 2008.

- The numbers in the table are cumulative totals of the number of trick-or-treaters who visited one house each year.

- The numbers are measured at 30-minute intervals, except for the last 15-minute interval.

- The trick-or-treat count was recorded in 30-minute intervals except for the last 15-minute interval.

- The night of trick-or-treating has always been on October 31st each year (some neighborhoods change the night of trick-or-treating).

- Official tick or treat hours are from 6 PM to 8 PM, but there are often "stragglers" past 8 PM that are not turned away. These stragglers are counted in the 8PM – 8:15 PM time slot. There has never been a trick-or-treater past 8:15 PM.

- The type of candy did not vary year-by-year. It is always a general mix of candy purchased in bulk variety bags.

Location of home

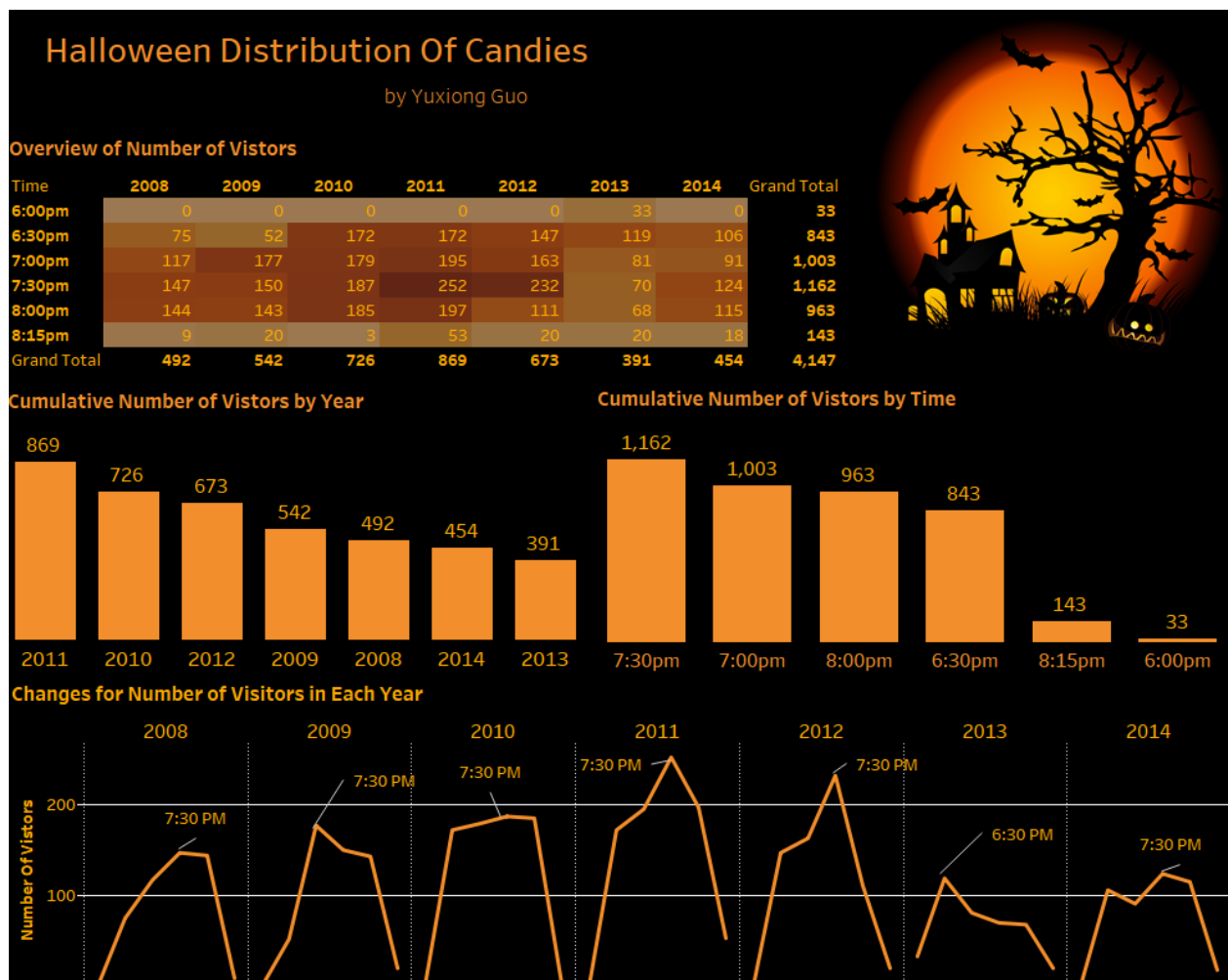Neighborhood: East Walnut Hills/Evanston
City, State: Cincinnati, Ohio
Zip code: 45207

Being a corner house on the neighborhood border likely increases the number of trick-or-treaters.

Example

Here' an example of how previous Halloween data have been visualized. Be creative!

Halloween Distribution Of Candies
by Yuxiong Guo

Overview of Number of Vistors

Time	2008	2009	2010	2011	2012	2013	2014	Grand Total
6:00pm	0	0	0	0	0	33	0	33
6:30pm	75	52	172	172	147	119	106	843
7:00pm	117	177	179	195	163	81	91	1,003
7:30pm	147	150	187	252	232	70	124	1,162
8:00pm	144	143	185	197	111	68	115	963
8:15pm	9	20	3	53	20	20	18	143
Grand Total	492	542	726	869	673	391	454	4,147

Cumulative Number of Vistors by Year

Year	Number
2011	869
2010	726
2012	673
2009	542
2008	492
2014	454
2013	391

Cumulative Number of Vistors by Time

Time	Number
7:30pm	1,162
7:00pm	1,003
8:00pm	963
6:30pm	843
8:15pm	143
6:00pm	33

Changes for Number of Visitors in Each Year

2008 2009 2010 2011 2012 2013 2014

Number Of Vistors

The Assignment

There are multiple parts to this assignment. Make sure you read the entire assignment before starting.

Determine a story or goal to support the two (2) visualizations you will create using the Halloween data provided. Your two visualization MUST be different chart types. **This means DO NOT create two bar charts or two-line charts or two of the same chart types!** Challenge yourself. This is your time to show what you know.

Examples (these are examples):

- Homeowner dashboard summarizing Halloween
- Forecast future trick-or-treaters or estimate future candy needed
- Explore variation of the number of trick-or-treaters year by year
- [Be creative and think of other things you could do](#)

Data Visualization Process

Show your understanding of the data visualization process.

Acquire

The Data

Year	6pm	6:30pm	7pm	7:30pm	8pm	Total (8:15pm)
2020	11	55	107	155	211	219
2019	0	117	262	406	483	523
2018	18	191	342	497	589	600
2017	41	190	357	549	710	776
2016	22	160	386	612	759	822
2015	13	148	336	523	667	747
2014	0	106	197	321	436	454
2013	33	152	233	303	371	391
2012	0	147	310	542	653	673
2011	0	172	367	619	816	869
2010	0	172	351	538	723	726
2009	0	52	229	379	522	542
2008	0	75	192	339	483	492

Excel and Tableau versions of the data are provided in Brightspace. **Choose one (1) to work with**.

- HalloweenExcel

- HalloweenTableau

Parse & Mine

Use this page to provide a parsing of the data. For quantitative fields list some basic statistical procedures that can be performed in the space below. To be clear, you are to list the procedure (you are not required to actually do any calculations here).

Use the Tab key to add more rows to the table below.

Variable	Data type	Statistical Method (where applicable)
Date	Date	Max, min, range
Date and Time	String	Max, min, range
Count	Int	Max, min, mean, mode
Day of Week	String	Max, min, mode
Time	String	Max, min, median, mode, range

Represent
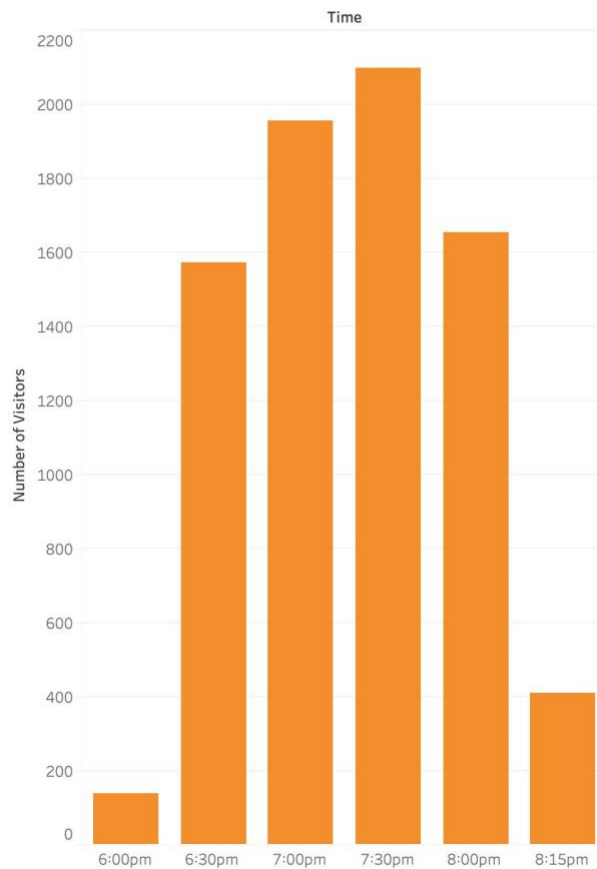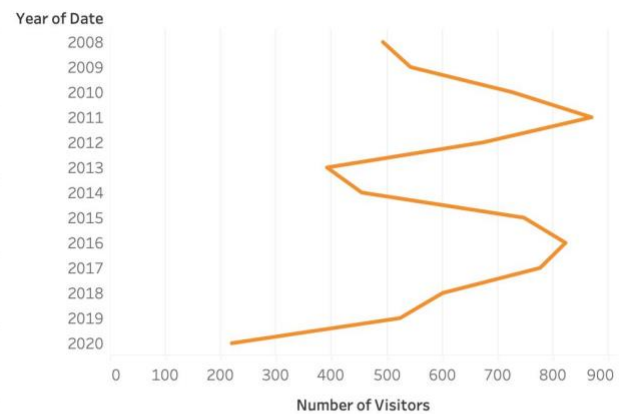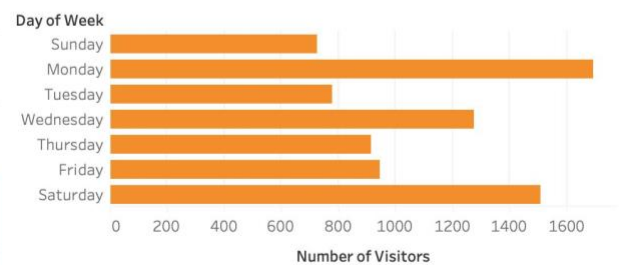
Figure 1. Total number of visitors per year, week, and time

Time with Most Visitors

Day of Week with Most Visitors

Year with Most Visitors
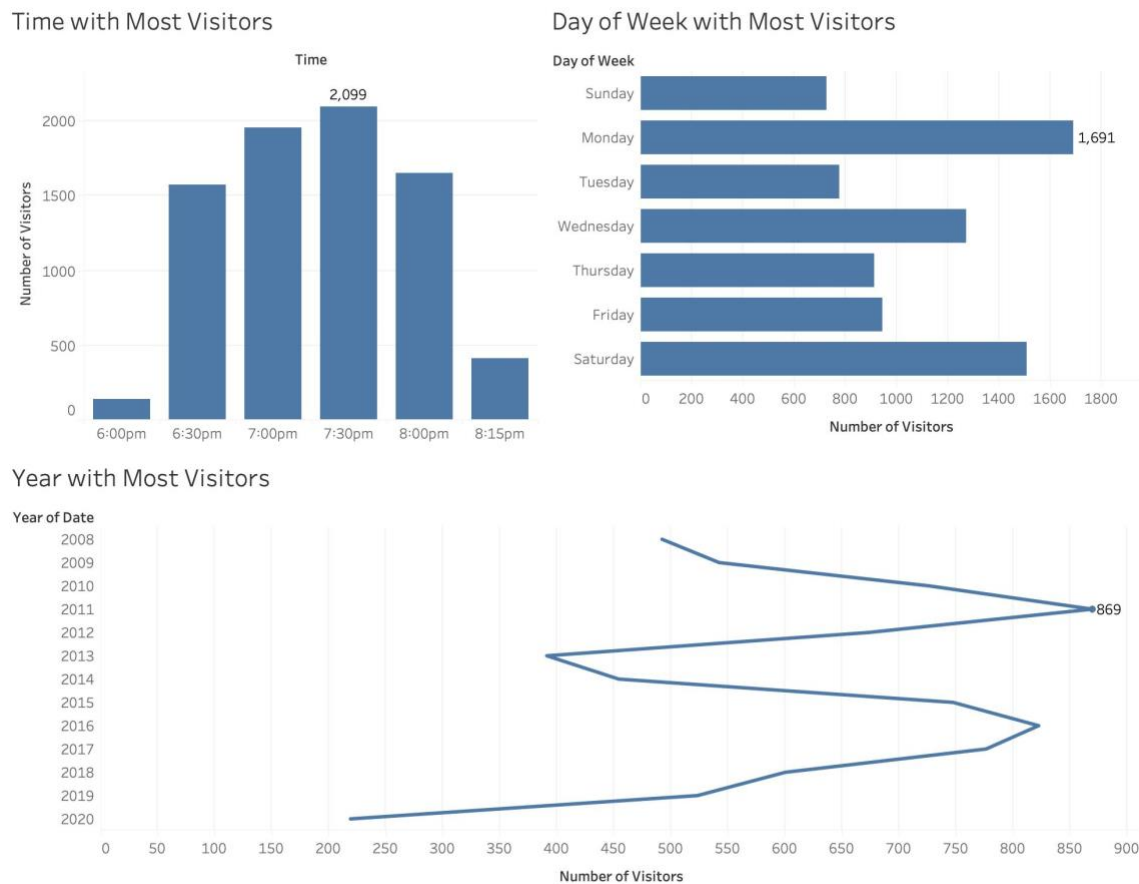
Figure 2. The year, week, and time with the highest number of visitors

Helpful Tip: Utilize the space that you have. Do NOT create a tiny visualization that is unreadable. Remember, the purpose of visualization is insight, but all insight is lost if it cannot be seen.

Use this page format for visualization that require a landscape layout. Remove this text and replace the figure with your own visualization.

Figure # Don't forget the figure caption. Source: https://towardsdatascience.com/my-top-10-most-fascinating-data-visualizations-from-2020-22a91b23e981

Filter

In this page show the data you used to create your visualizations.

Figure 1

The data I used was:

- Count (Number of visitors)
- Time
- Date (Year)
- Day of week

Figure 2

The data I used was:

- Count (Number of visitors)
- Time
- Date (Year)
- Day of week

Critique

Rate your visualizations (Figure 1 and Figure 2) using the link below

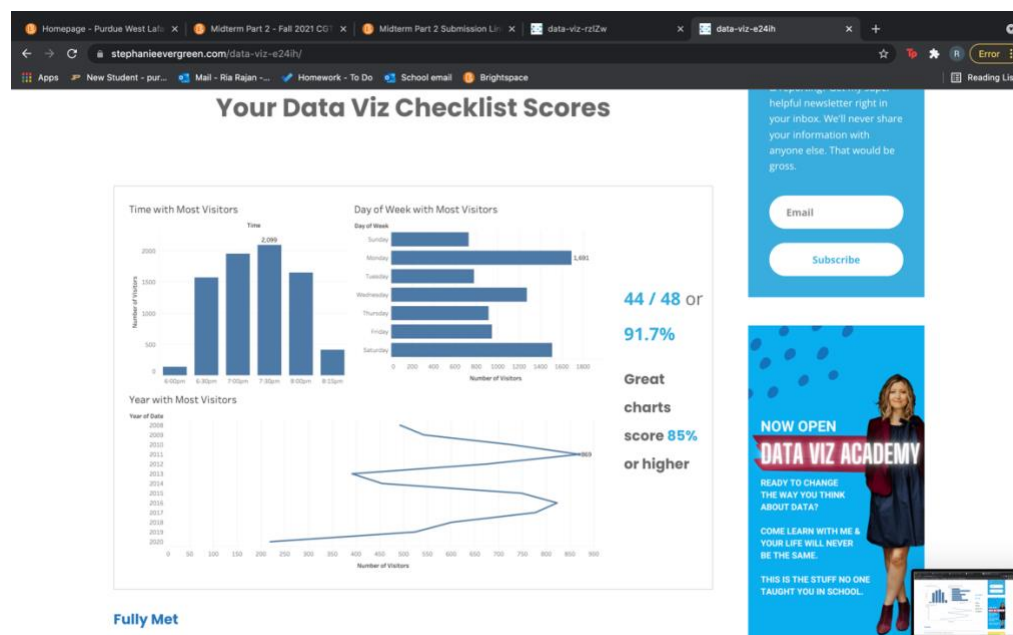https://stephanieevergreen.com/rate-your-visualization/
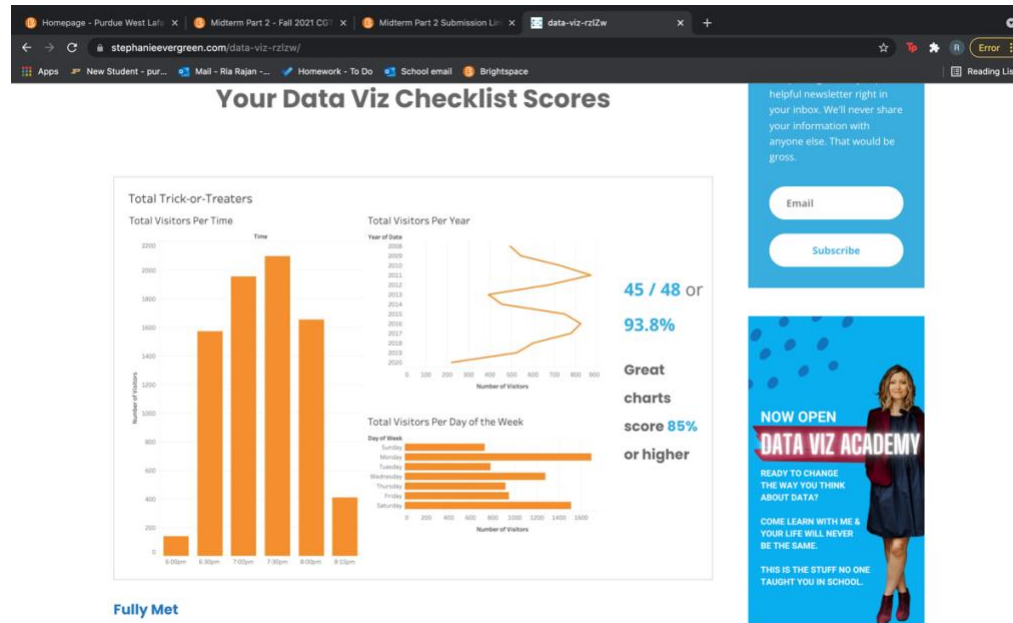
Figure 1 Rating
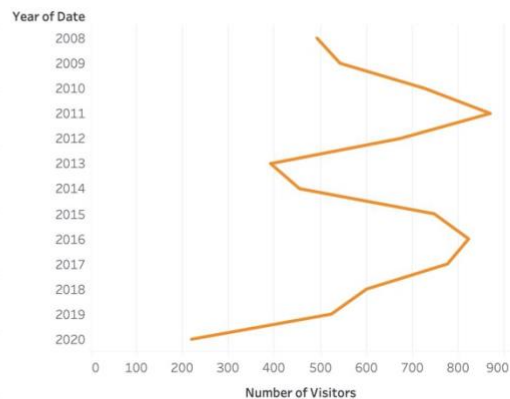

Figure 2 Rating

Refine

In this part of the visualization challenge, you should identify one or more characteristics of the visualizations you created (Figure 1 and Figure 2) and update the figures. Include an updated version of each Figure below. In the figure caption, state what changes were made.
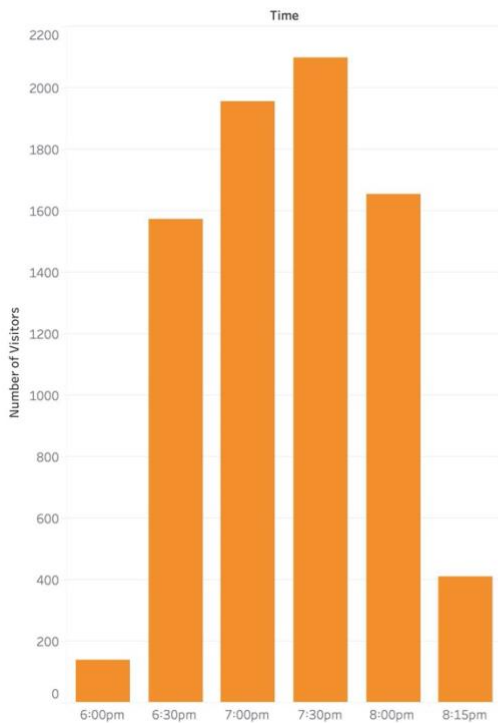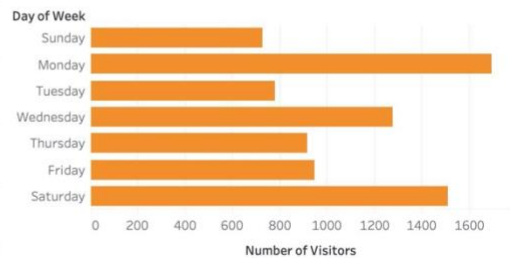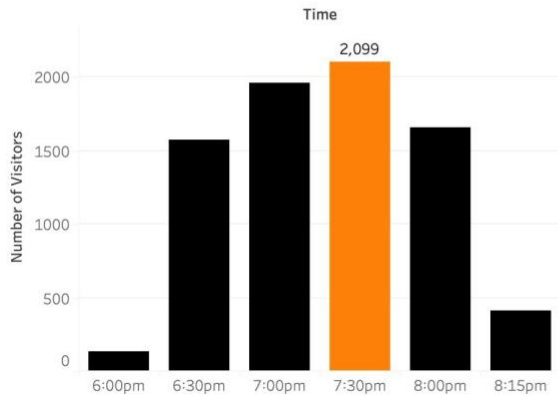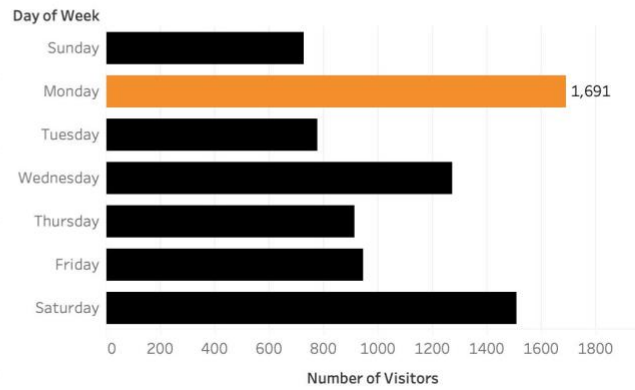

Figure 1 Refined. Changed each visualization to fit the entire view so that the layout flowed better and there weren't any uneven empty spaces.

Most Visitors per Date, Time, and Year

Time with Most Visitors

Day of Week with Most Visitors
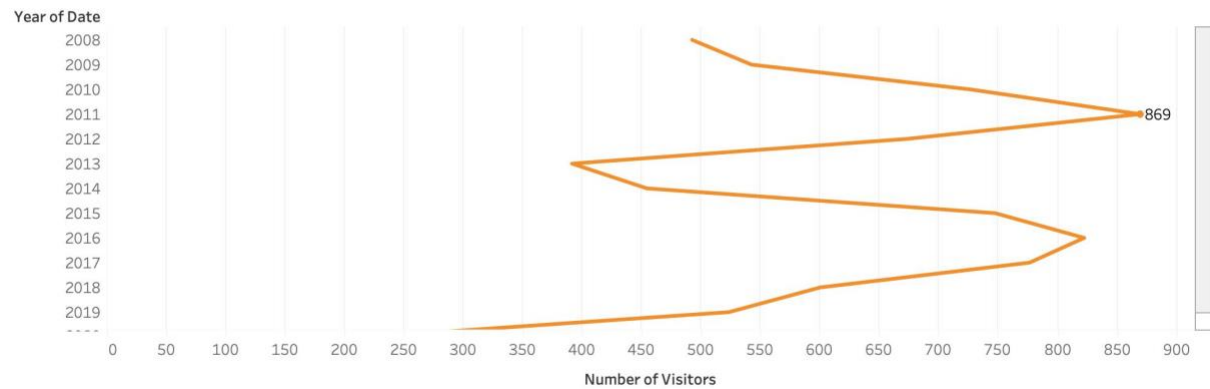

Year with Most Visitors

Figure 2 Refined. Added a dashboard title to better convey the message of the visualization. Highlighted columns with the most visitors to highlight specific what I want to point out to my audience.

What's the story?

Using complete sentences, answer the following questions:

1. What story did you tell using the data? *(Note: your story must be supported by your visualizations: Figure 1, Figure 1 Refined, Figure 2, Figure 2 Refined). (*500 words max, with complete sentences). Make sure you mention the figures (by name, for example, as seen Figure 1, Figure 1 Refined, etc.)

   a. In figure 1, my goal was to show all the data without highlighting any keep components just as it was. The reader could make any assumptions they wanted to because it was just the all the numbers as they came. With figure 2, my goal was to answer the question, what was the year, week, and time with the greatest number of visitors? In figure 1 refined, I fit the visualization to view so that the layout was easier to understand. In Figure 2 refined, I spent some time playing around with Tableau so that I could highlight the specific year day of the week in which the number of visitors was the greatest. I also added a data label, so the reader knows the exact number. My goal with this figure was to allow the reader to understand more about the data.

2. Who is your audience? (Use complete sentences)

   I think my audience could be anyone but more specifically, I think it would best adhere to Halloween lovers and especially trick-or-treaters in Ohio who want to know when they should go out to trick-or-treat and what they expect in terms of crowds.

3. List 3 assumptions you made while implementing the data visualization process?

   a. <assumption 1: complete sentences, 150 words max>

      i. 1 assumption I made was that trick-or-treaters usually did not leave the house until after 6:30 and the time slot of 6 was just a buffer period to be included.

   b. <assumption 1: complete sentences, 150 words max >

       i.   1 assumption I made was that the column date and time was maybe parsed from the original data set into date and time separately since I didn't use this variable.

   c.   <assumption 1: complete sentences, 150 words max >

       i.   1 assumption I made was that trick or treaters usually enjoy going out towards the end of the week.

Points will be taken off for incomplete sentences.

Bonus points for REALLY GOOD stories!

Checklist of what to submit:

- Save this file as LastnameFirstInitial_CGT270Fall2021_MidtermPartII.pdf

- Only submit one (1) file. All of your work should be contained in this file.

- Failure to follow these instructions will result in your work NOT being graded.

General Deductions (others made accordingly)

- No name on the first page of the document: -5 pts

- Altered template: -10 pts

- No figures included: -15 pts for each missing figure

- No figure captions: -10 pts for each missing caption

- Zip file submitted: See Checklist of what to submit (-80 pts)

- Late submissions: Will NOT be graded (-80 pts)

- Provided a link to visualizations instead of providing screenshot of the visualization: this will be treated as no figure, no figure caption (-25 pts)

- Failure to follow data visualization best practices (data visualization checklist): deductions made appropriately.

Keep in mind: one (1) second after the submission deadline is considered late.

Byrd Data Visualization Lab