

CGT270

Midterm Part II

Data Visualization Challenge

Student:

Nathan Rusk

Professor

Dr. Vetria Byrd

TA

Subia Ansari

Course:

CGT 27000 Data Visualization

Term:

Fall 2021

Contents

Halloween Visualization	3
Data Description	3
Location of home	4
Example	4
The Assignment	5
Data Visualization Process	6
Acquire	6
The Data	6
Parse & Mine	7
Represent	8
Filter	10
Critique	11
Refine	13
What's the story?	15

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

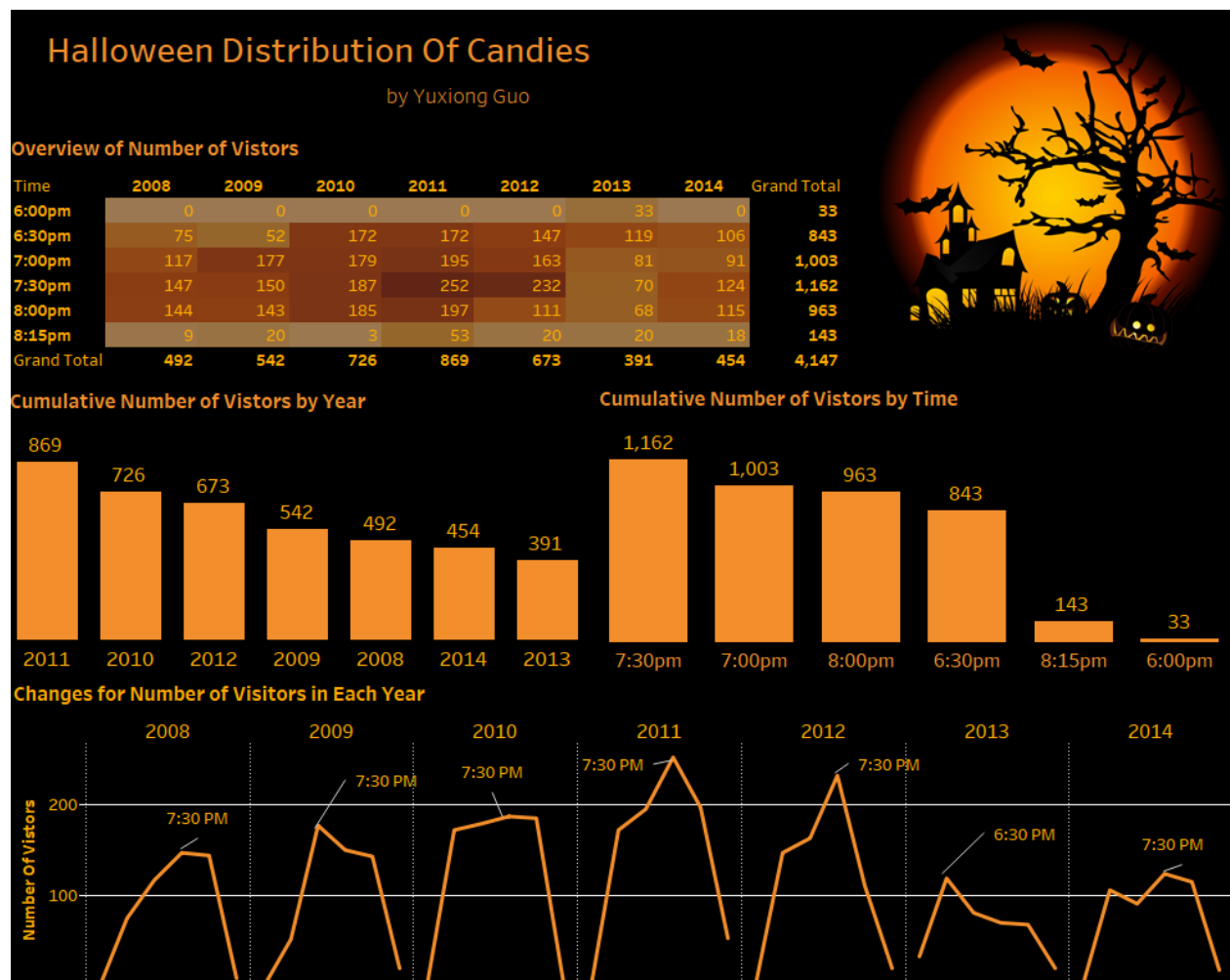
City, State: Cincinnati, Ohio

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!



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	Categorized by year in my case
Measure values	integer	Number of people
count	integer	Sum of Count

Represent

Most Popular Times to Trick-or-Treat

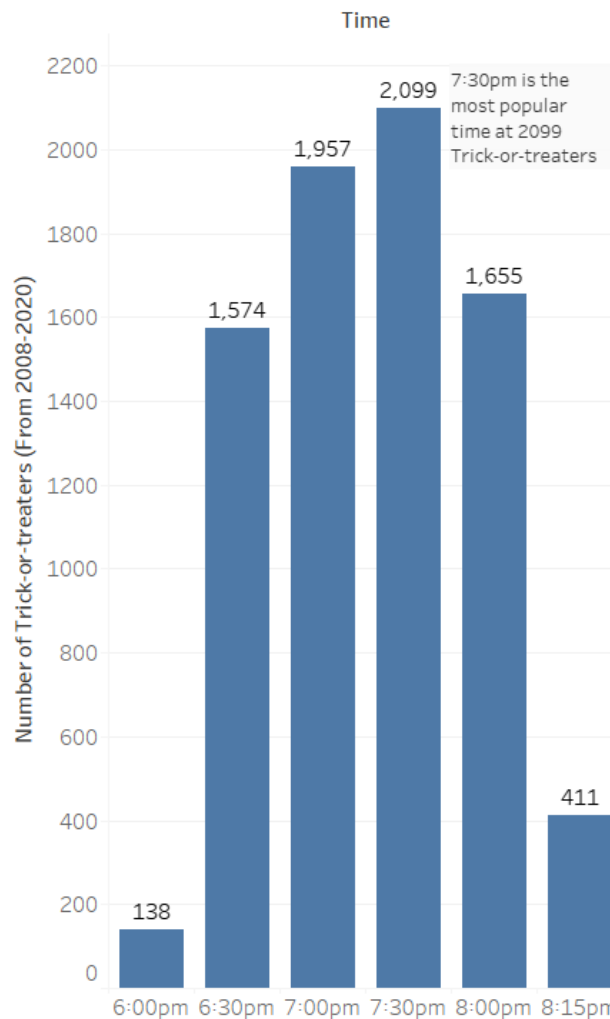


Figure 1. This figure shows a bar chart of the number of trick-or-treaters and the most popular times to trick-or-treat during the trick-or-treating time window.

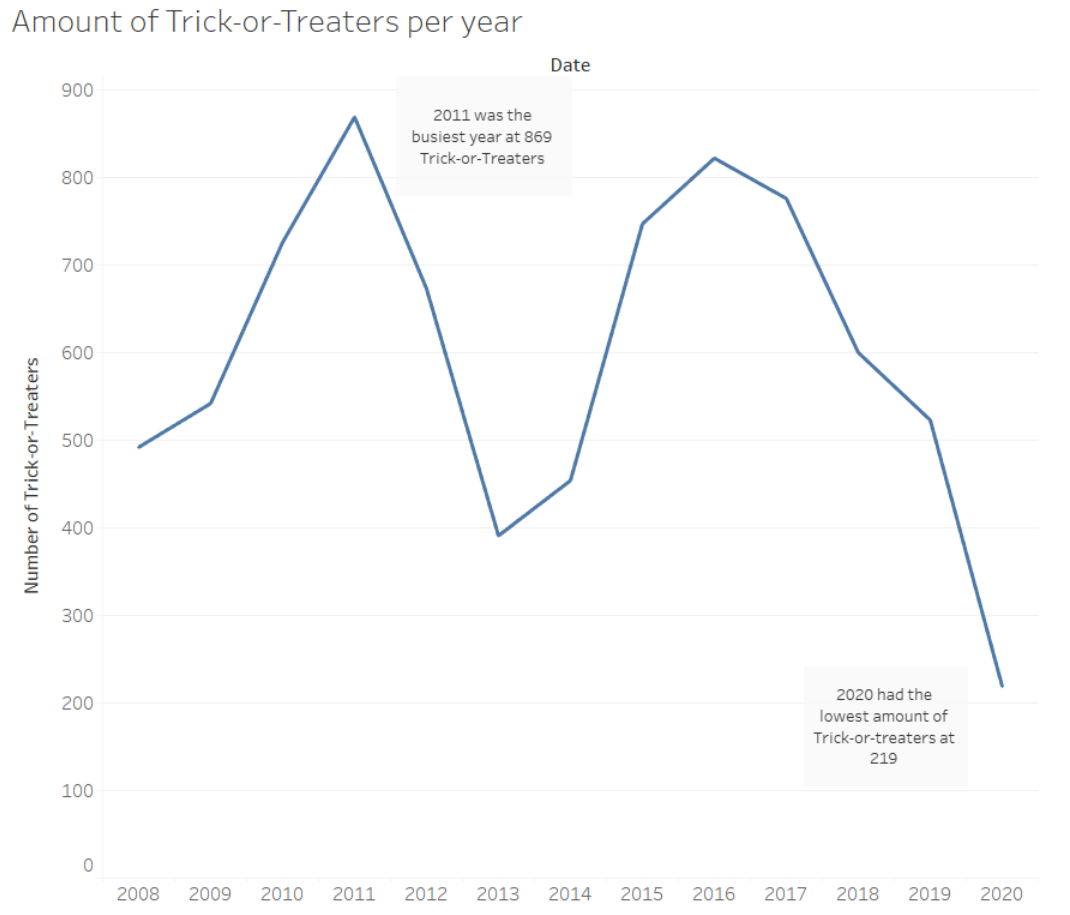


Figure 2. This figure shows a line graph of the number of trick-or-treaters per each of the recorded years and the most and least popular years to trick-or-treat.

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.

Filter

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

Figure 1

In figure 1 I used the time of day (Time) and the number of people (measure values) to show the number of Trick-or-Treaters per time slot.

Figure 2

In figure 2 I used the year (Date) and the number of people (sum of Count) to show the number of Trick-or-Treaters per year.

Critique

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

<https://stephanieevergreen.com/rate-your-visualization/>

Figure 1 Rating

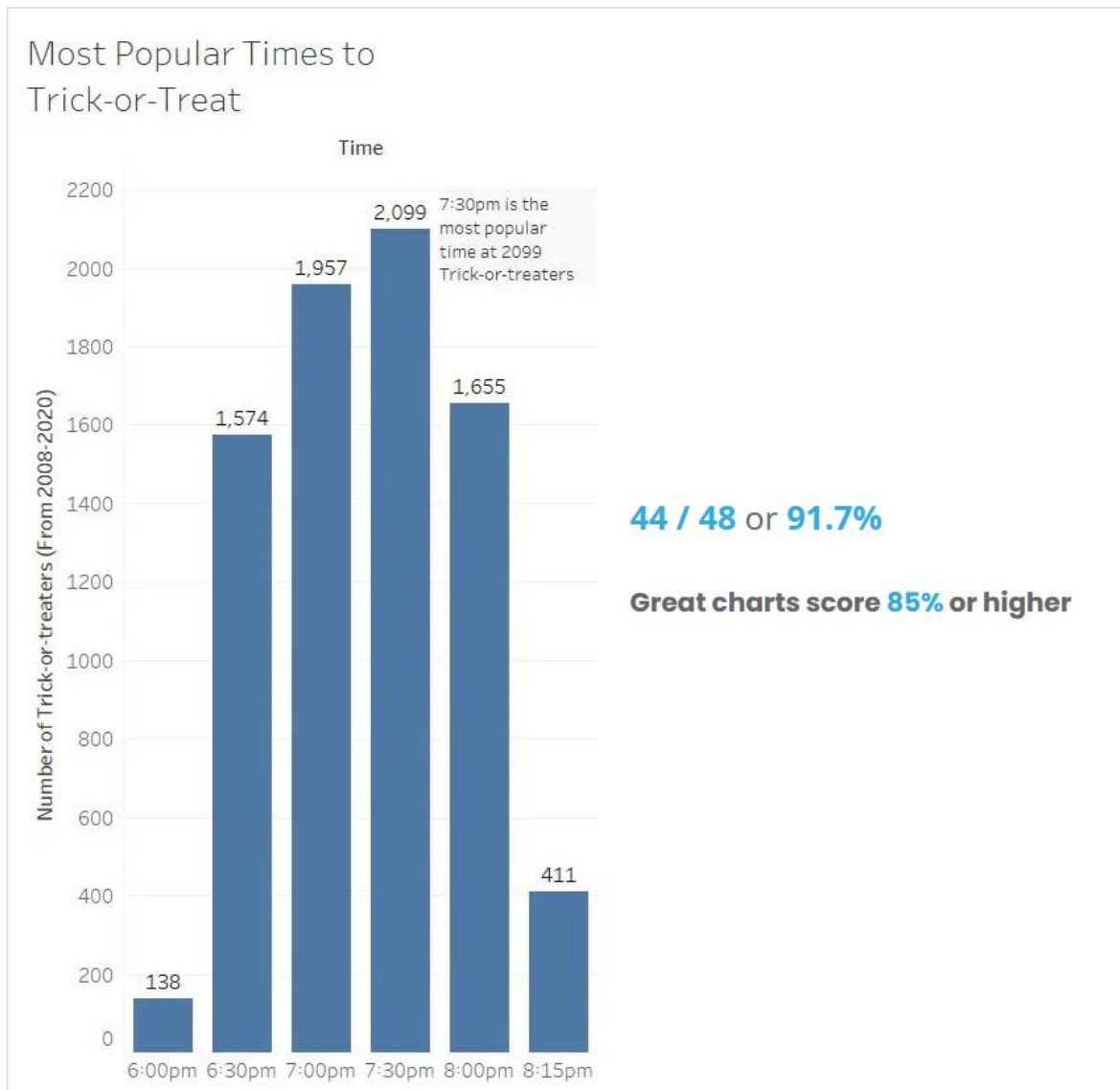
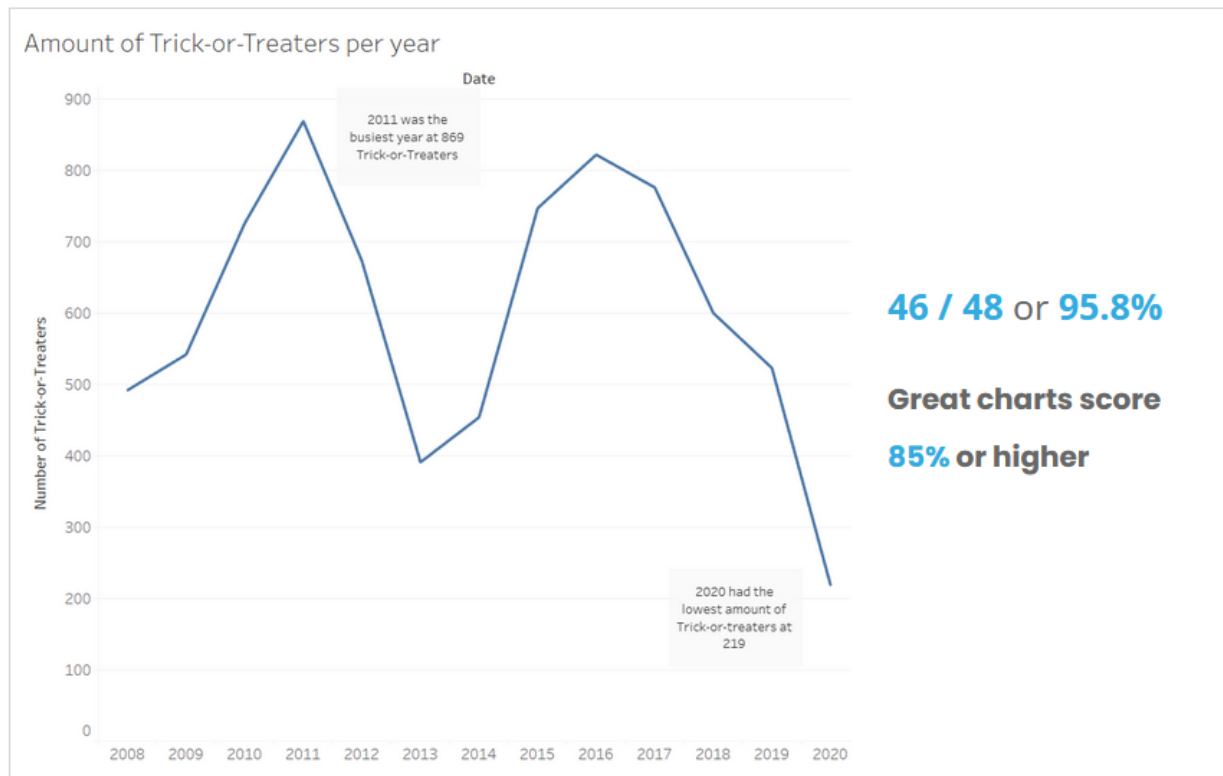


Figure 2 Rating



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.

Amount of Trick-or-Treaters per year

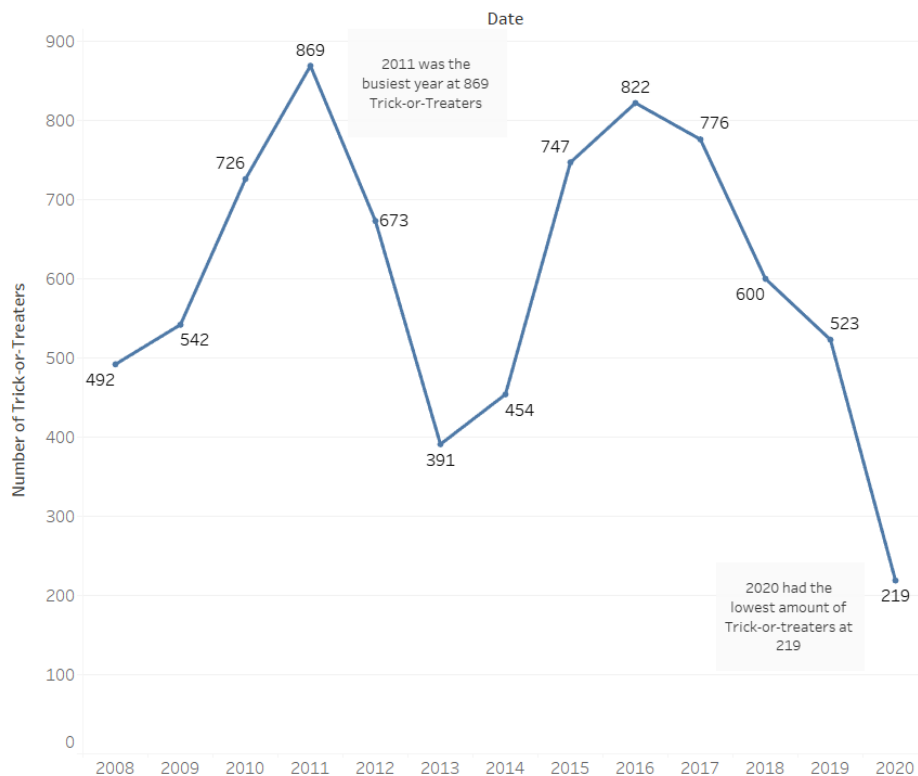


Figure 1 Refined. Updated the figure to include individual data labels.

Most Popular Times to Trick-or-Treat

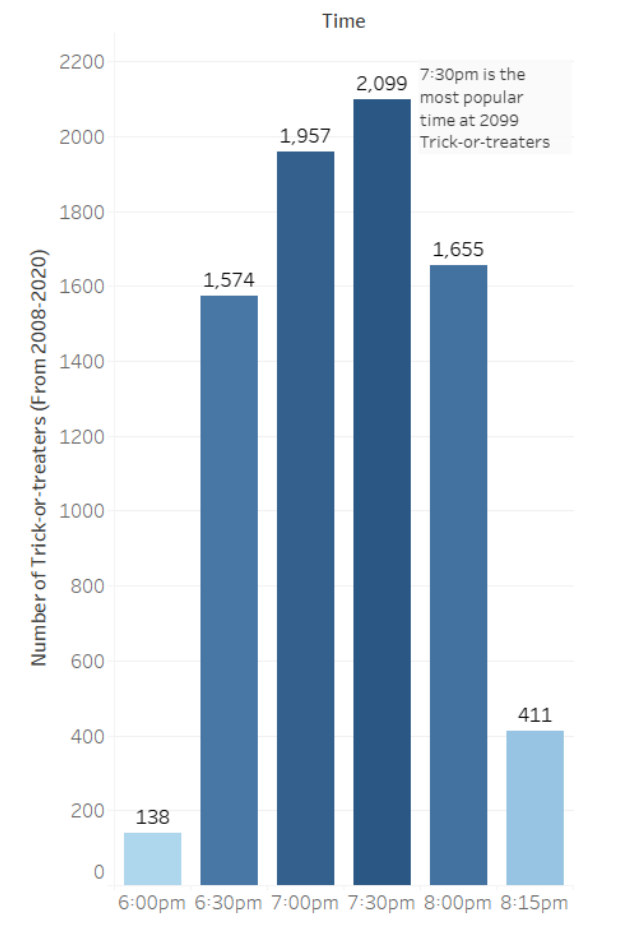


Figure 2 Refined. Changed the color to highlight the lowest to highest numbers of people. Darkest represents highest, lightest represents lowest.

What's the story?

Replace the text on this page with your story. The story should be no more than one-page. **If you go over the page limit, your story will NOT be read (-5 pts).** Single space Calibri Light (Body) font, font size 13. This is a very simple data set. There are only a few years of data broken down into 4 half-hour time blocks with cumulative totals.

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).*

The story is to visualize the number of people and the time slots and years where trick-or-treating is the busiest, or most popular. (Not much time to write more)

2. Who is your audience?

The audience would be anyone interested in knowing this data. Most likely it would be most interesting to the people from the specific location.

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

- a. The number of trick-or-treaters varies a lot. Per year and per time slot.
- b. 2020 was the least busiest year due to the pandemic.
- c. 7:30pm is the busiest time to trick-or-treat because it is in the middle of the time slot available to trick-or-treat

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