CGT270
Midterm Part II

Data Visualization Challenge

Student:

Joy Gao

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.

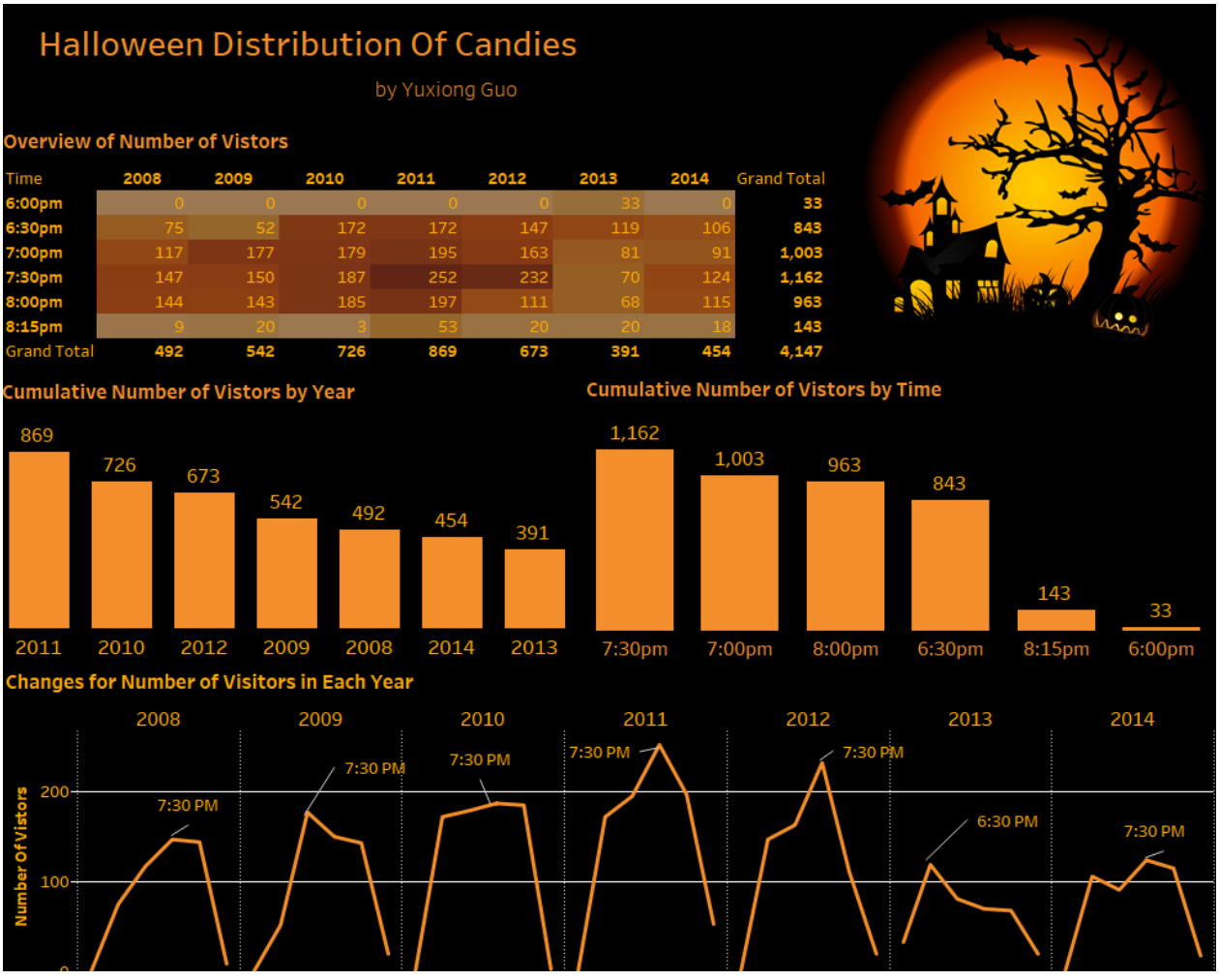
Location of home

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!



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)
Month and day	Date	
Year	Integer	Median
Time	String	
Count	Integer	Mean, min, max

Represent

Remove this text before submission: You can use any visualization tool you are most comfortable with. Replace the image below with your first visualization.

How to replace this figure: Right-click on the figure below, select Change Picture → From a File. Locate your figure.

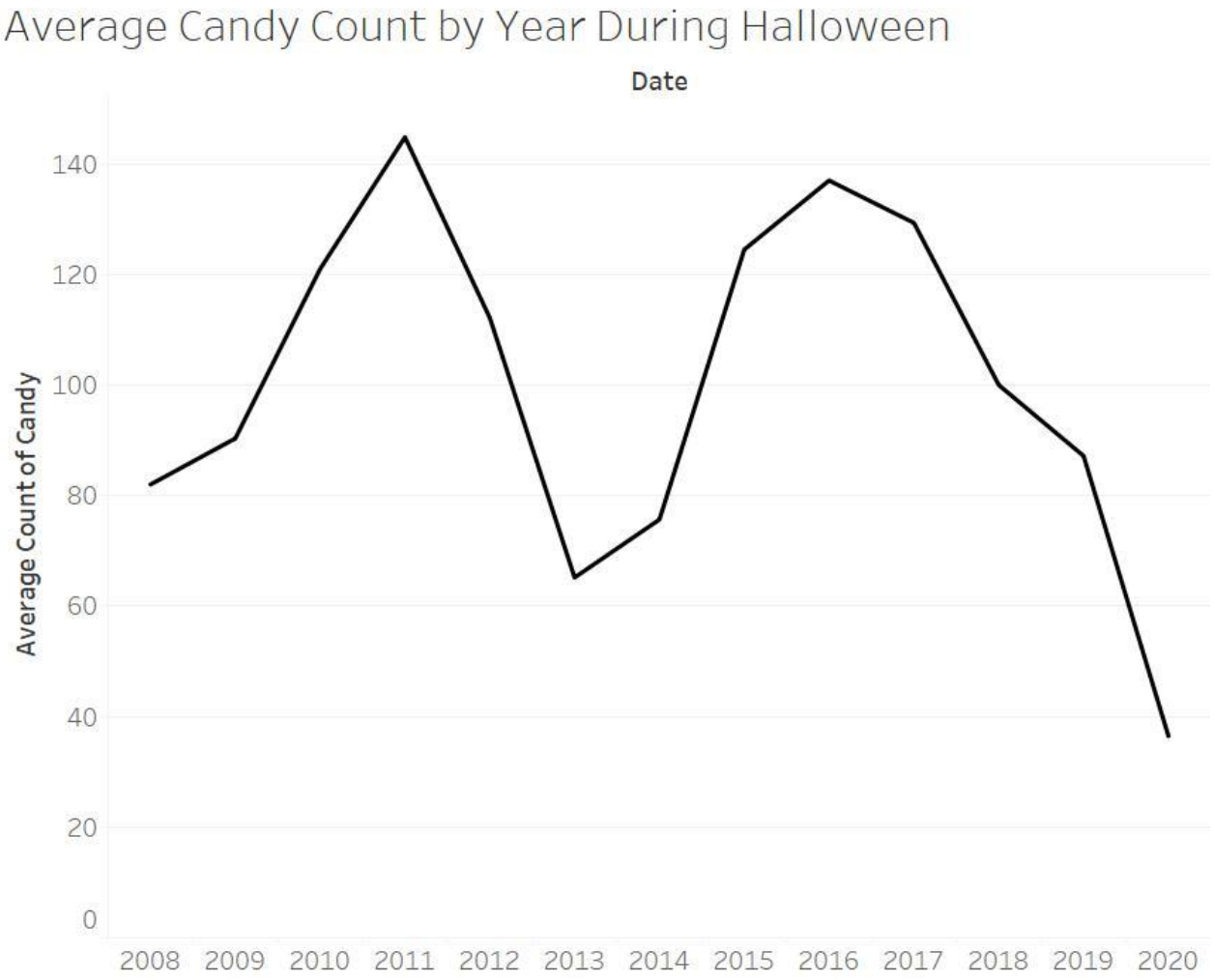


Figure 1. This visualization demonstrates the average count of candy Trick-or-Treaters received during Halloween every year.

Source: https://infogram.com/blog/do-this-not-that-data-visualization-before-and-after-examples/

<Remove this text before submitting your work: replace the image below with your second visualization.>

How to replace this figure: Right-click on the figure below, select Change Picture → From a File. Locate your figure.

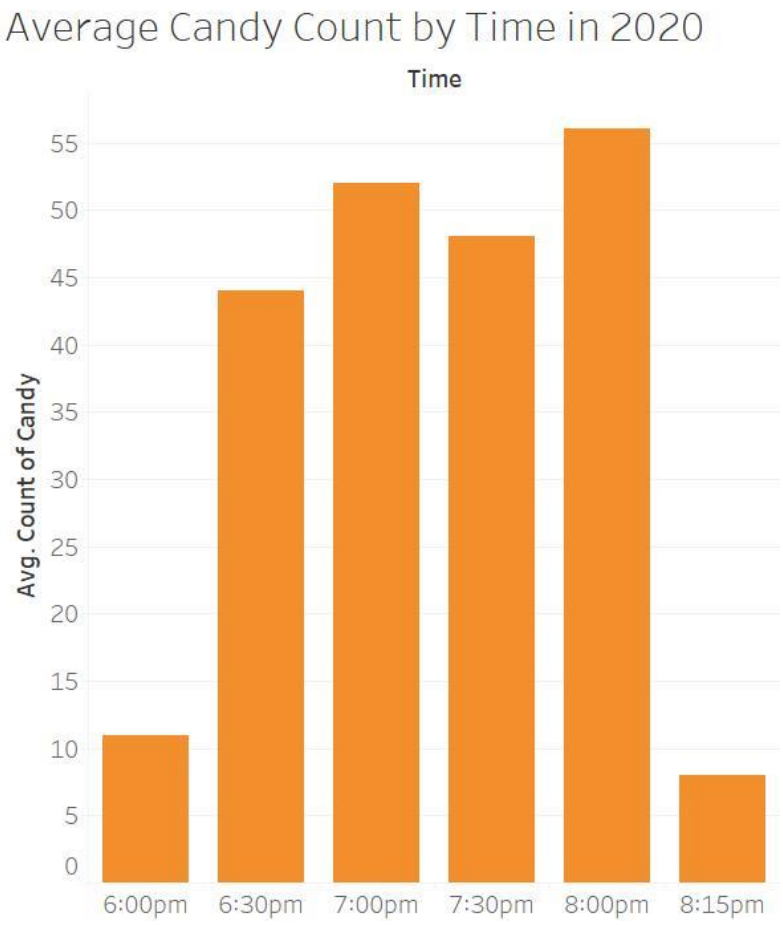


Figure 2 This visualization demonstrates average candy count by time during Trick-or-Treating on Halloween in 2020.

Source: https://infogram.com/blog/do-this-not-that-data-visualization-before-and-after-examples/

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 the first visualization, I used the variables: date and count.

For date, I included the years from 2008 to 2020.

For count, I used the average count during the years from 2008 to 2020.

Figure 2

In the second visualization, I used the variables: time and count.

For time, I used 30-minute increments from 6:00 pm to 8:15 pm.

I also filtered out date by excluding all the years except 2020.

Critique

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

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

Figure 1 Rating

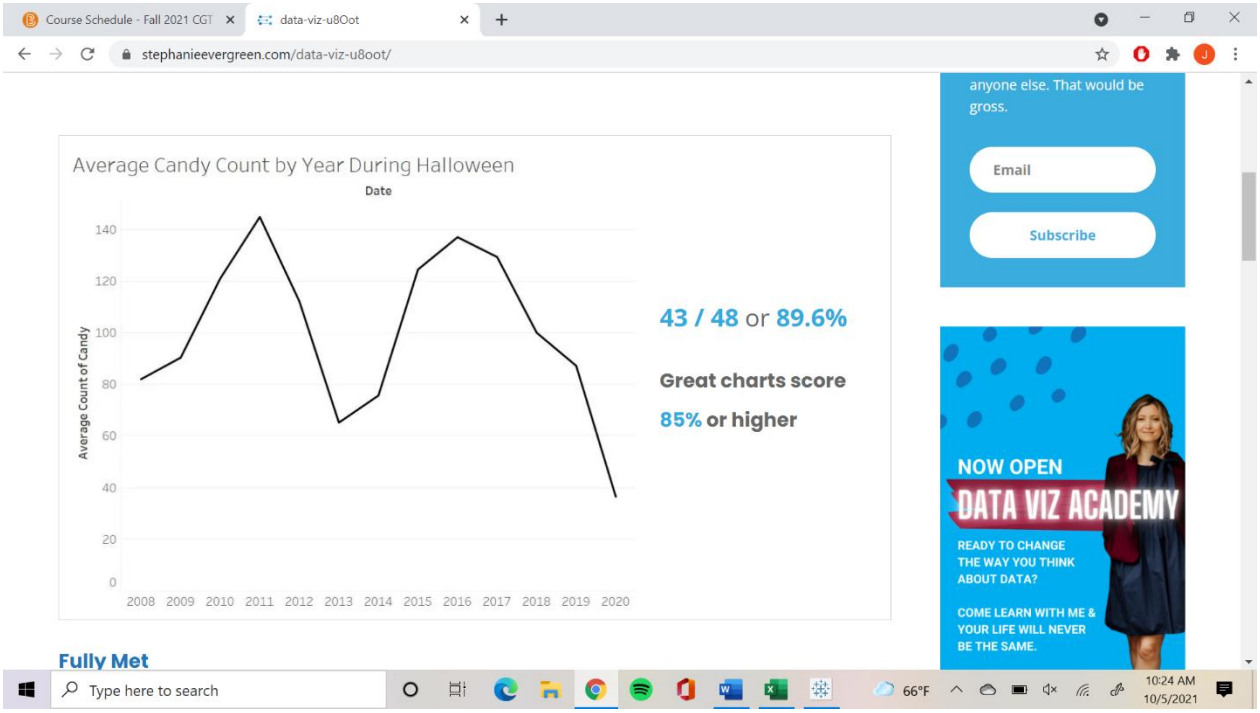
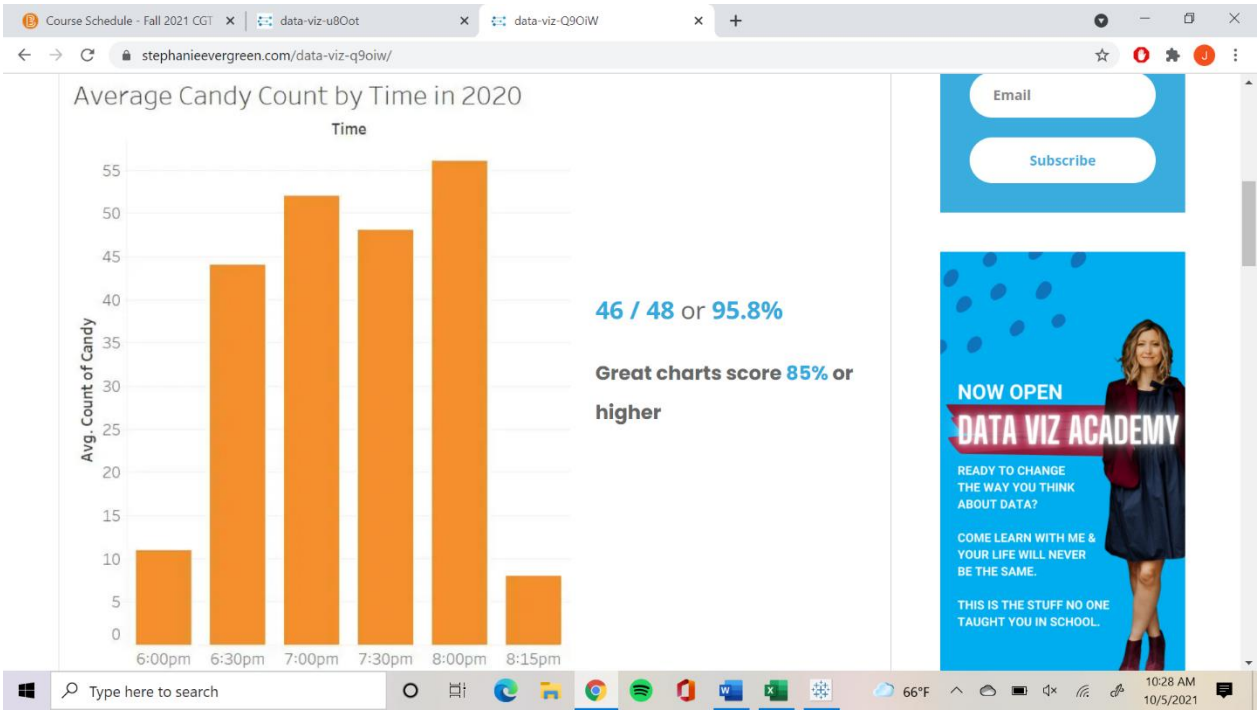


Figure 2 Rating


Course Schedule - Fall 2021 CGT
data-viz-u8Oot
data-viz-Q9OiW
stephanieevergreen.com/data-viz-q9oiw/
Average Candy Count by Time in 2020
Time
Avg. Count of Candy
55
50
45
40
35
30
25
20
15
10
5
0
6:00pm
6:30pm
7:00pm
7:30pm
8:00pm
8:15pm
46 / 48 or 95.8%
Great charts score 85% or higher
Email
Subscribe
NOW OPEN
DATA VIZ ACADEMY
READY TO CHANGE THE WAY YOU THINK ABOUT DATA?
COME LEARN WITH ME & YOUR LIFE WILL NEVER BE THE SAME.
THIS IS THE STUFF NO ONE TAUGHT YOU IN SCHOOL.
Type here to search
66°F
10:28 AM
10/5/2021

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.

Replace the picture with your visualization, remove this text before submission.

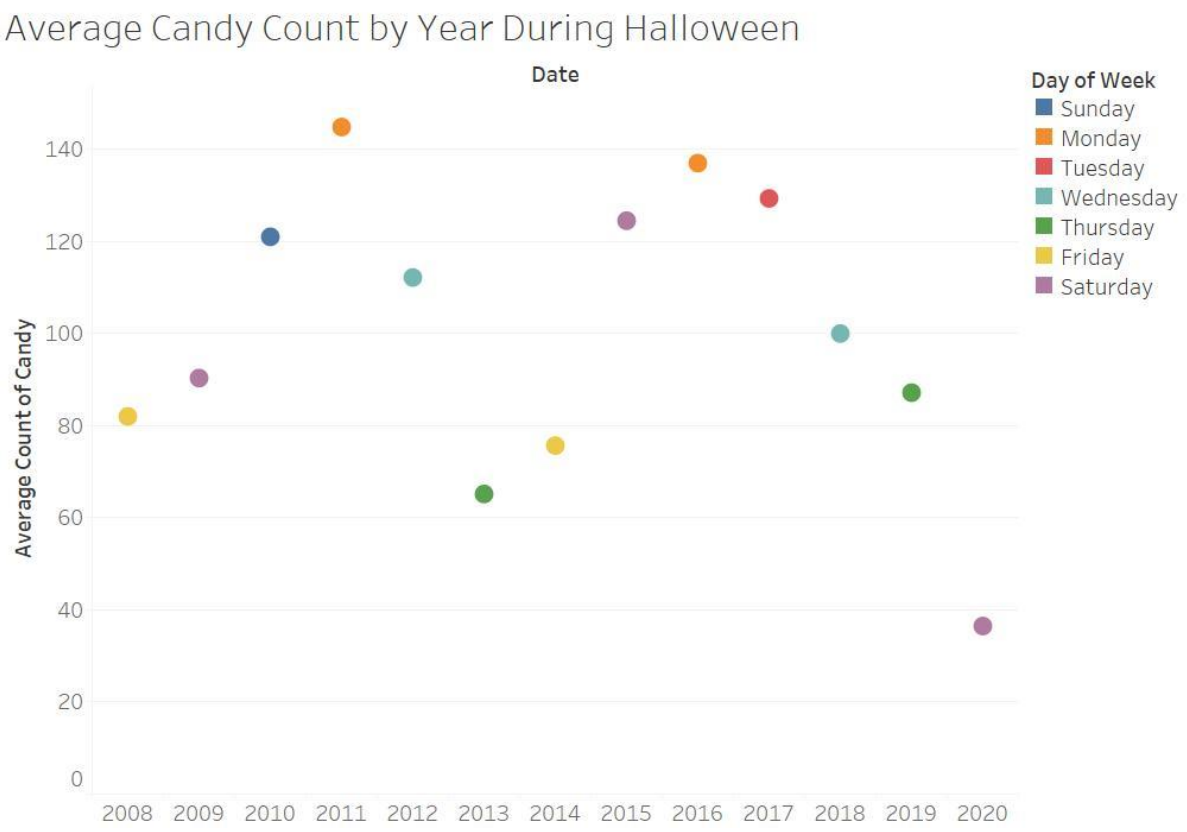


Figure 1 Refined

I chose to refine this visualization by adding the day of the week Halloween is because this could provide more insight on why certain values are higher or lower than others. I also chose to make it a scatterplot instead of a line chart because this makes it clearer that the data is only gathered from one day every year and is not continuously changing throughout the year. Source: https://infogram.com/blog/do-this-not-that-data-visualization-before-and-after-examples/

Replace the picture with your visualization, remove this text before submission.

Figure 2 Refined.

I changed the title to include the word 'Halloween' because this makes it clearer to the audience that this data is collected from Halloween. I also chose to include the day of the week Halloween was on so that there could be more information to reference when making assumptions.

Source: https://infogram.com/blog/do-this-not-that-data-visualization-before-and-after-examples/

What's the story?

1. My data tells the story that Trick-or-Treaters love to go Trick or Treating during the beginning of the week, particularly Monday, as seen in Figure 1 Refined. Since kids make up most of the trick-or-treaters, they seem to be well-rested from the weekend, so they have more energy to go to more houses. That is why more kids showed up to this house on the corner. By going to more houses, they receive more candy, therefore increasing the average count in 2011 and 2016, as seen in both Figure 1 and Figure 1 Refined. If Halloween is later in the week, kids may get more tired from school and after-school activities, so they want to sleep earlier. This means less candy and decreasing average count, as seen in Figure 1 and Figure 1 Refined. Covid may have affected 2020 significantly as more parents would want to keep their kids inside the house instead of going out to trick-or-treat. As a result, it is the lowest average count. I wanted to explore 2020 more so I specifically focused on Halloween in 2020. Figure 2 and Figure 2 Refined both show that Trick-or-Treaters like going around 6:30 to 8:00 the most. This is prime time because the evening is the best environment for spookiness. However, 8:15 is too late because many kids get tired and need to go to sleep soon. This is seen in my Figure 2 visualizations where the average candy count decreases significantly.
2. Who is your audience? My audience is mainly Trick-or-Treaters and their parents if the Trick-or-Treaters are young. This is because my visualizations can tell them the most popular times to go trick-or-treating. This can help them plan their Trick-or-Treating experience accordingly.
3. List 3 assumptions you made while implementing the data visualization process?
    a. I assumed the variable count in the dataset meant the average amount of candy Trick-or-Treaters received on Halloween.
    b. I assumed kids were the main Trick-or-Treaters who make up the dataset since they seem to be the age group that Trick-or-Treat the most. That is why the prime time to trick-or-treat is earlier around 6:00 to 8:15 pm.
    c. I assumed that count meant the number of individual pieces of candy that Trick-or-Treaters collected on Halloween.

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