CGT270
Midterm Part II

Data Visualization Challenge

Student:

Jiaxiang Li

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.

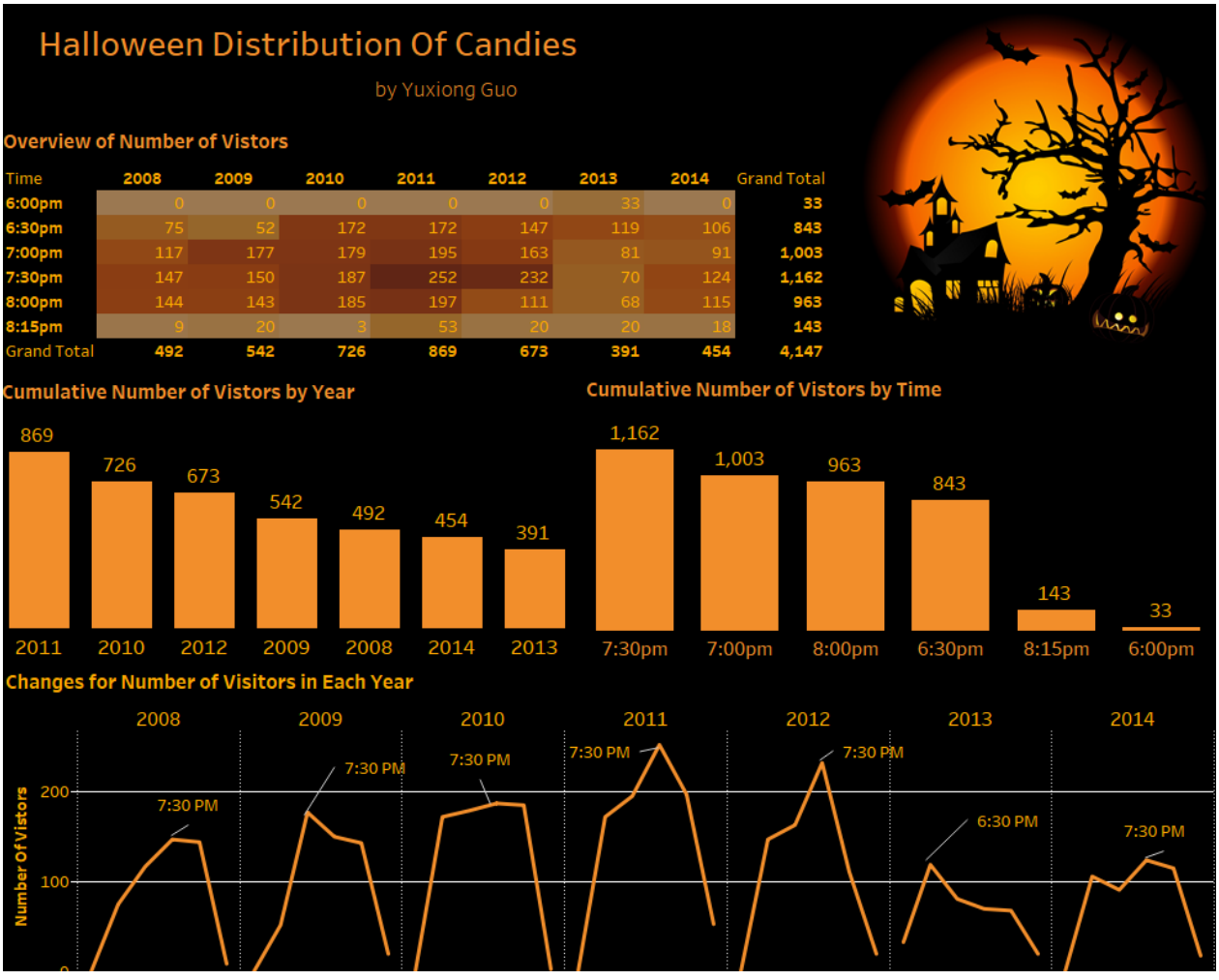
Location of home

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)
Date	Date	String length
Date and Time	Date & Time	String length
Count	integer	Average, median, max, min, standard deviation
Day of Week	string	String length
Time	String	String length

Represent

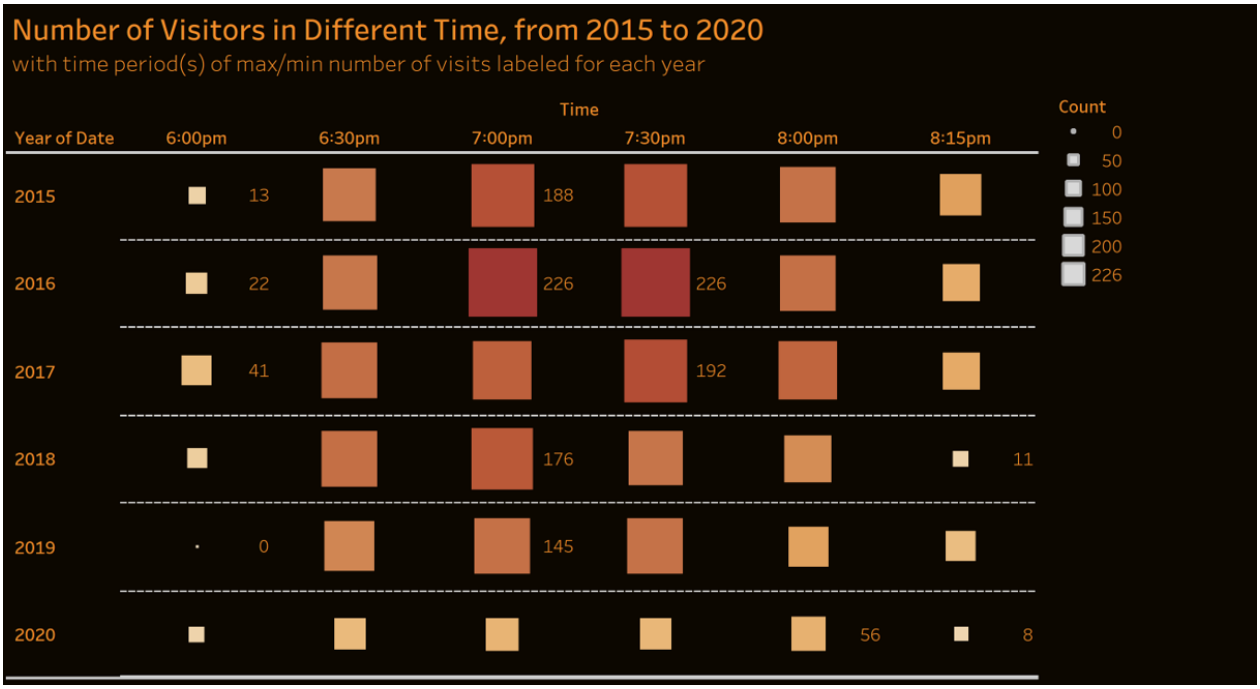


Figure 1. Number of Visitors in Different Time, from year 2015 to 2020, with max/min of each year labeled. The differences are presented in square size and color.

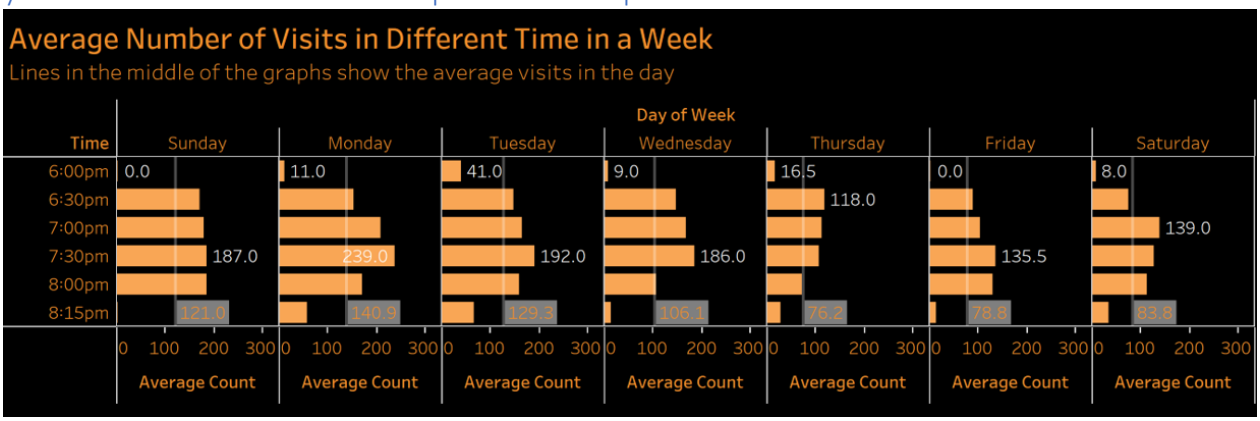
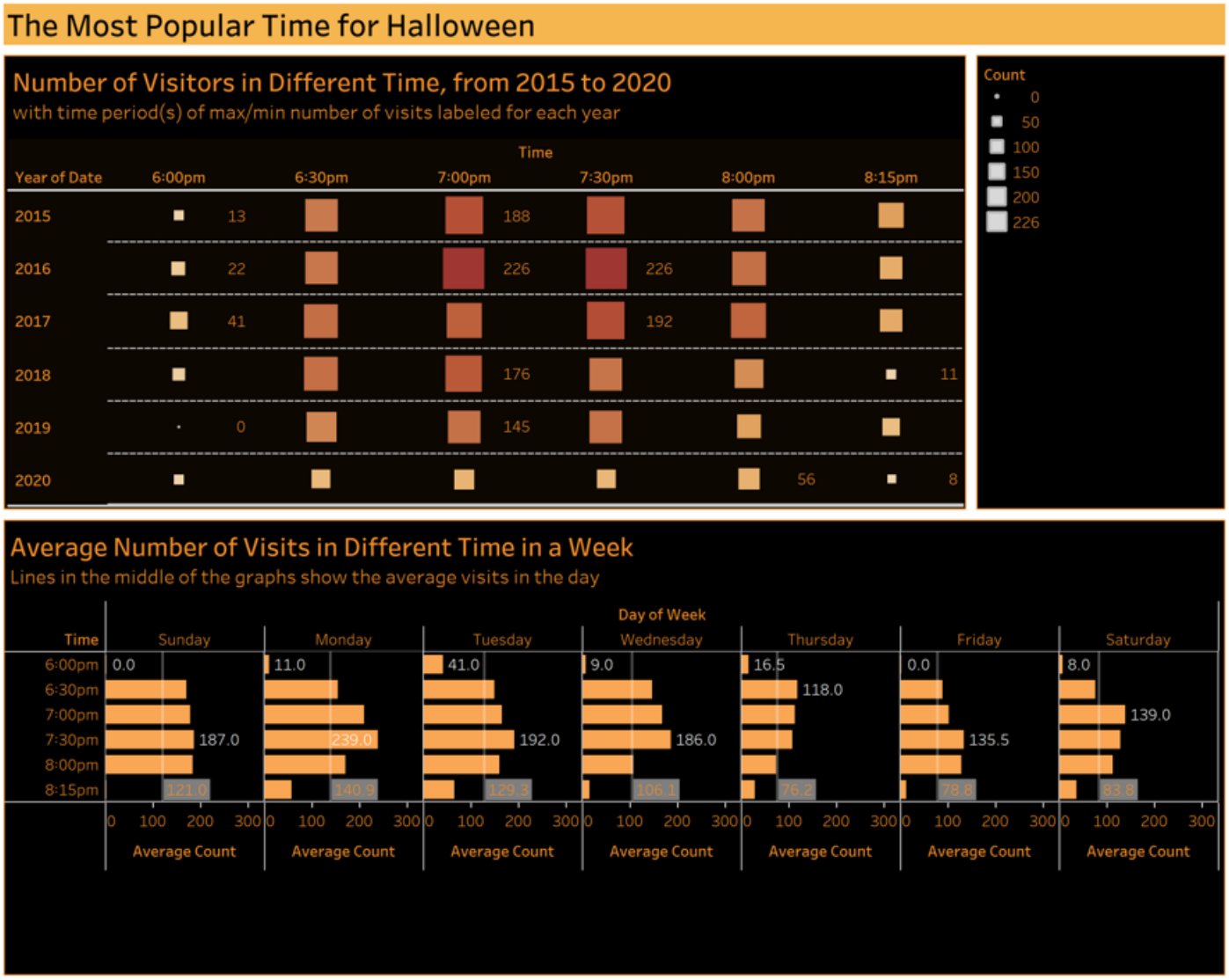


Figure 2. Average number of visits in different time of a week. The number of time each day in a week is Halloween from 2008 to 2020 is different, so the visits in each day is calculated as average. Average counts in each day are labeled as line.



The dashboard shows both graphs with legends, presenting an overview of visits in each circumstances in the past few years.

Filter

Figure 1

The first graph used Time as column and Year as rows, with a filter in year implemented, filtering year 2015 to 2020 to present the general trend of visiting counts in different time but without the graph being packed with information and too redundant for viewers who want to see a clearer image.

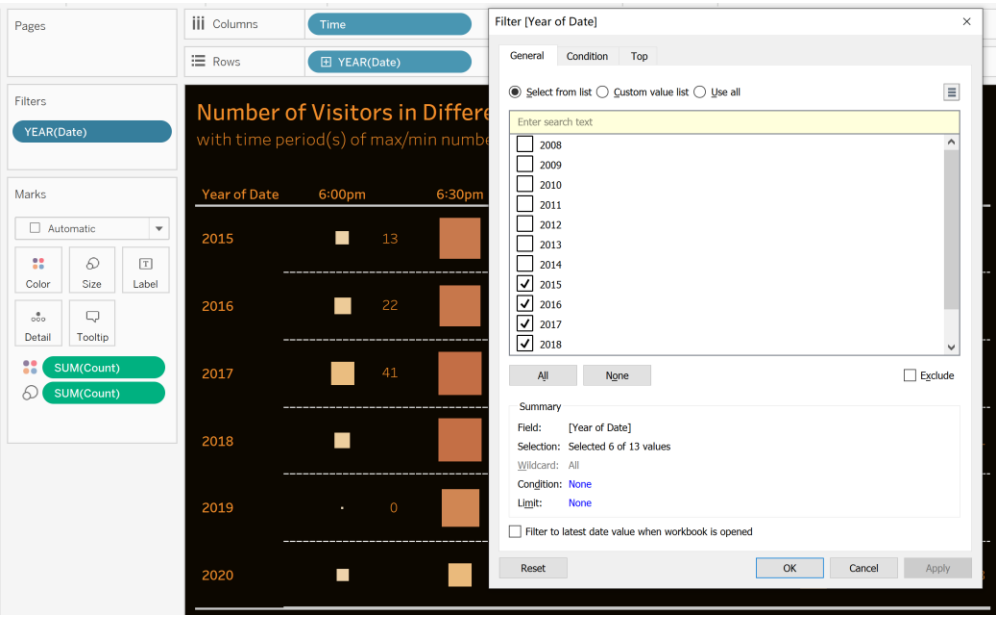
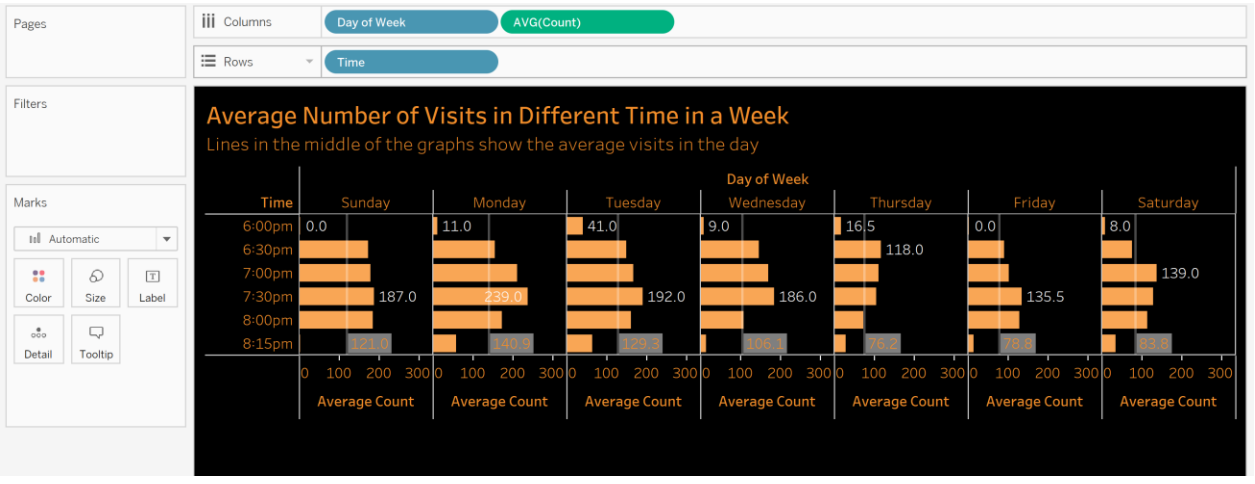


Figure 2



The second figure filtered Day of Week and the average of visiting counts as columns, and Time as rows. There is no additional filter because there are only 7 days in a week and each day is important for viewers to learn about the popularity of trick-or-treating activity in different circumstances.

Critique

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

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

Figure 1 Rating

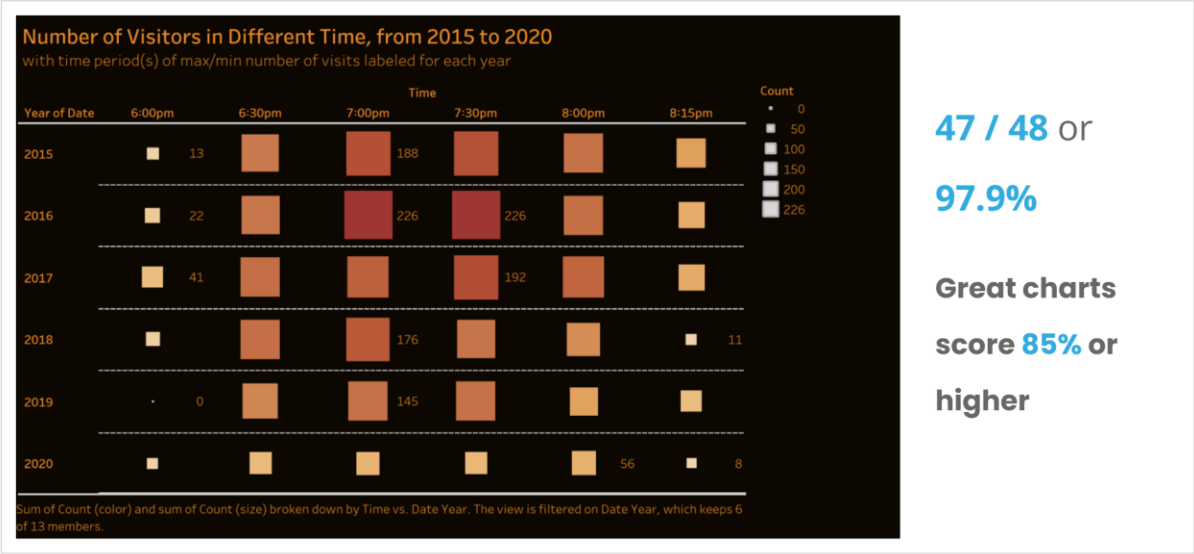
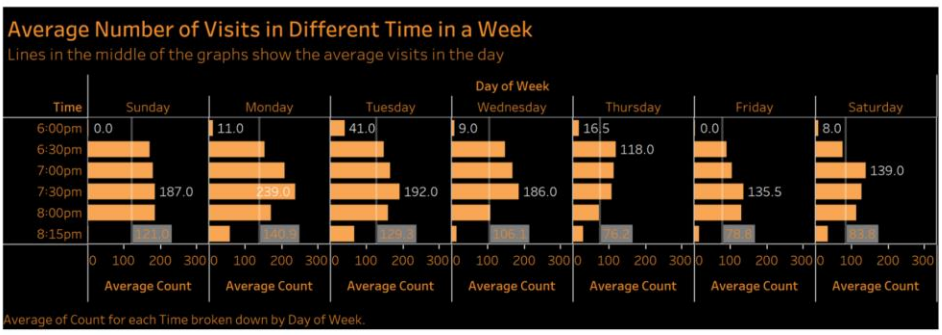


Figure 2 Rating



Refine

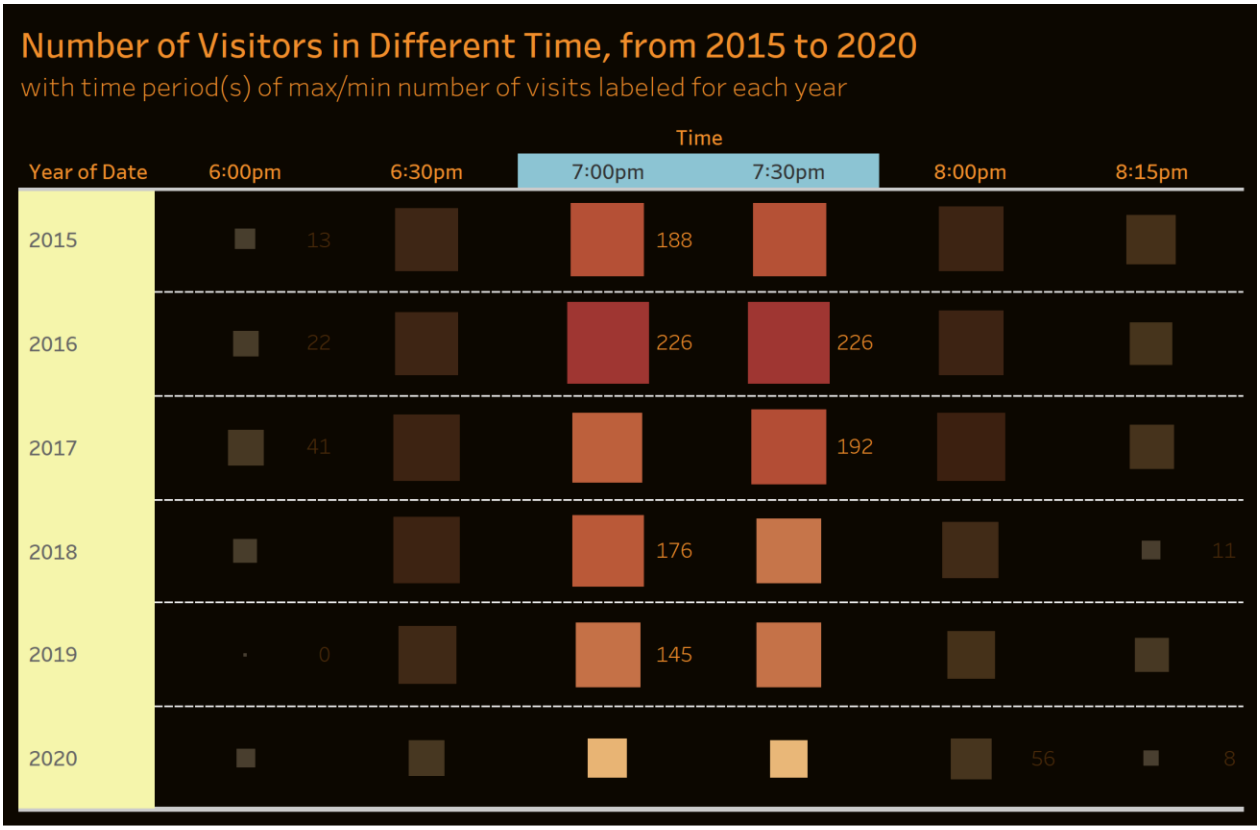


Figure 1 Refined. Highlighted the most important part of the figure (time periods with most visitors).

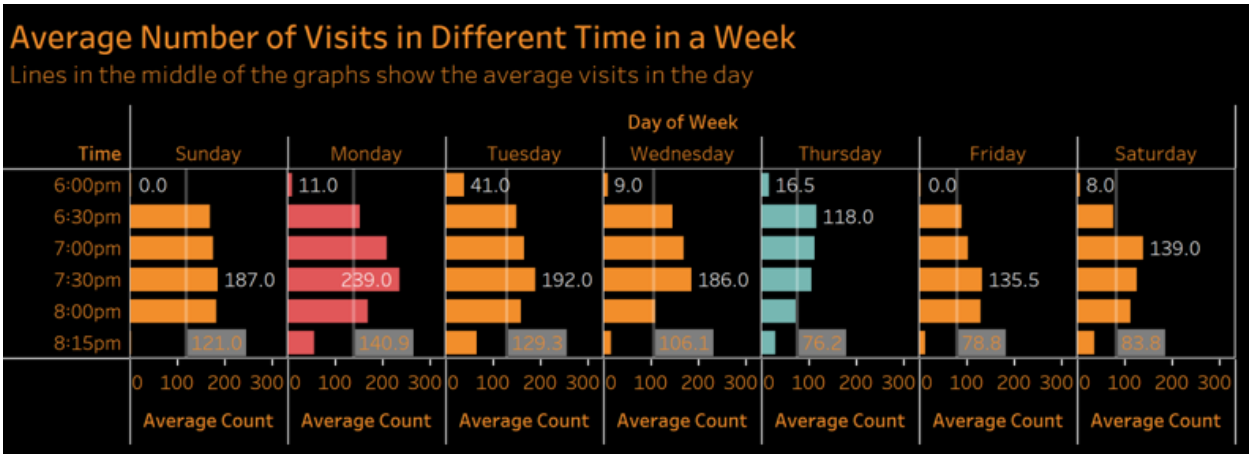


Figure 2 Refined. Changes colors for some of the more important days of the week to highlight its significance.

Use this page if your visualizations require a landscape layout. Remove this page if it is not needed.

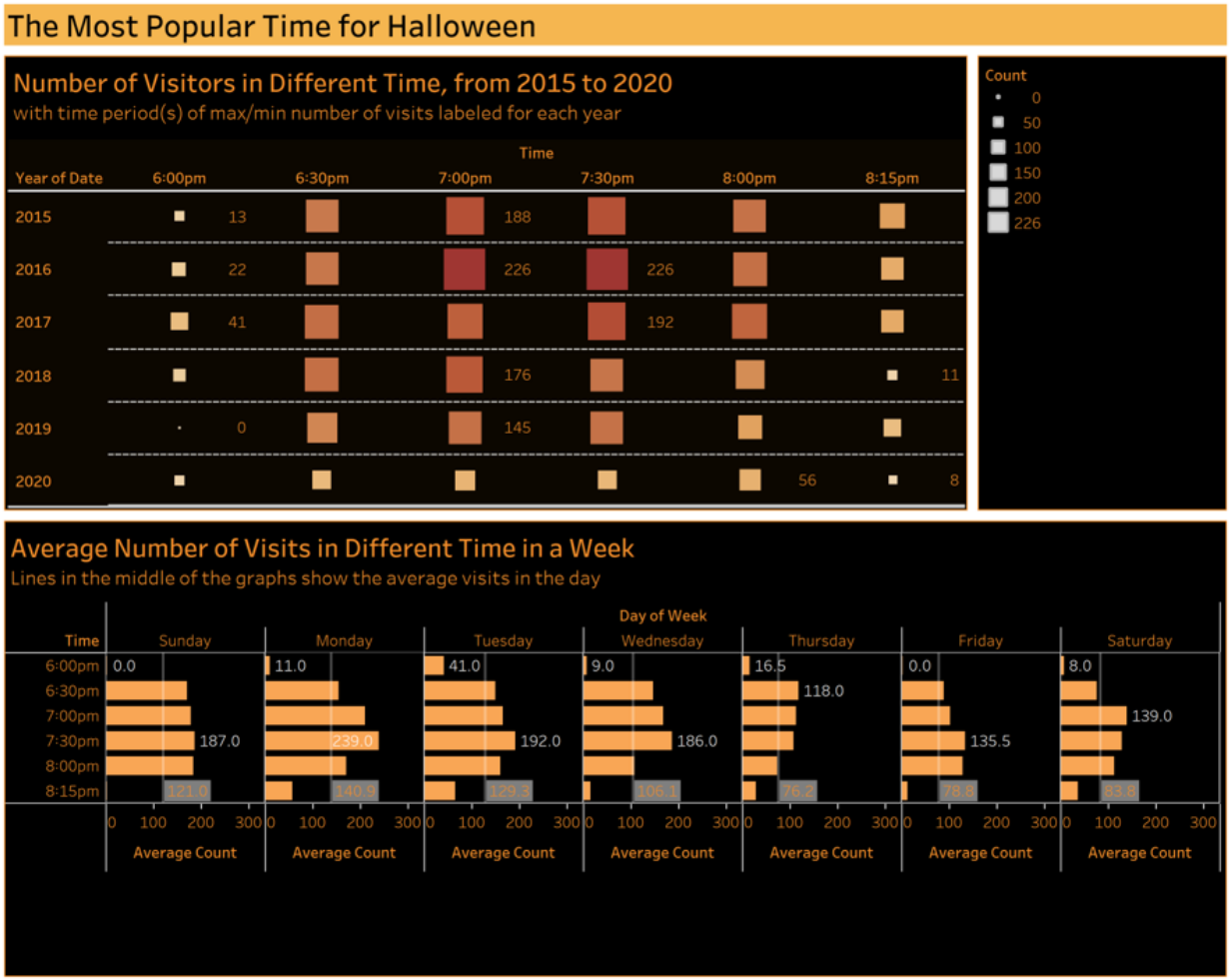


Figure # Figure overall refined.

What's the story?

Using complete sentences, answer the following questions:

1. The story is about the research of the day of the week and year in which Halloween's iconic trick-or-treat activity is most popular. With the data of visiting numbers in each year and each day of the week in hand, I explored about how number of visitors changed throughout the years, especially for the most recent years of 2015 to 2020. With average number of visitors in different time in years at hand, I understand that 7:00pm to 7:30pm is the prime time for the trick-or-treat party, and it is probably one of the best time to go out and find someone to prank and enjoy with. On the other hand, I ffind out that the average number of visits in Monday is, surprisingly, the largest, which means that most of the children and activity joiners come out and party during Monday, which is probably because of their long weekend and their playful heart that hasn't already realized that it's already weekday!
2. Mostly teenagers and people who are very interested in Halloween activities and want to find the best time to join the party.
3. List 3 assumptions you made while implementing the data visualization process?
    a. The best way to explore a best day for Halloween is by combining information in both years and weeks.
    b. The more frequently the activity is, the more popular the game is.

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