**Final Project**

Kyla R. Prows

IU School of Informatics and Indiana University-Purdue University Indianapolis

IN-NEWM- N328: Visualizing Information, Section 22809

Dr. Khairi Reda

April 30, 2022

**Final Project**

The first thing I did was look at the headers in my csv file to get an idea of what data was collected. I observed that the following data points were being collected for each tip left: total bill amount, gender of tipper, mealtime, day of week, smoker/non-smoker, and size. In my mind, I imagined that, as an owner of a restaurant, one might want to see the attributes of what type of customer was tipping the most. At first, I sketched out a rough version of what my visualization might look like:

A piece of paper with writing on it

Description automatically generated with medium confidence

I thought that the y-axis would be the total bill, while the x-axis would show the percentage of tip per tipper attribute. I quickly dismissed this idea, but it got me thinking deeper about how to display valuable information to the restauranter that would be easy to quickly digest. This led me to my second concept:

Chart, bar chart

Description automatically generated

In this iteration, I wanted the y-axis to show the average percent of total bill that was tipped. I wanted my x-axis to show groupings of attributes and then have a tooltip on hover show up to provide even more information. To me, this would provide a large amount of information at quick glance, but then allow for a deeper dive if the viewer wanted to interact more with the data. I opted not to include the size of the party and if the tipper was a smoker or not because I didn’t feel like these data points added much value in being able to offer meal specials at specific times to increase tip revenue.

**Design Choices**

I chose to go with a bar chart because I could map my data’s position to categorical attributes, while mapping channels to a quantitative attribute. Due to the categories of information I was going to display, I believe a bar chart was the best tool. Since my data points weren’t really related to time per se, I didn’t feel a scatterplot would be a good choice.

I chose colors that I believed my imaginary restaurant would have. I wanted them to have a slight nod to money, so I chose a green palette with enough variance in color that they were easily distinguishable from each other. I also placed it on a soft mint-green background to eliminate a high-contrast white background.

As the numbers increase, the bars go higher from bottom to top. This was also a choice I made to help the restaurant subconsciously know if the restaurant was performing well. The “higher” the tips, the “higher” the amount of money the restaurant is bringing in.

**Data Interaction**

I was interested to know which gender tipped more; which meal was more often tipped upon; and which day of the week was most often tipped. I felt like this information would be useful to a restauranter to know so that they could offer “date night” specials vs. “ladies lunch specials” at the time and day that would most likely lead to higher tipping. I also wanted to know more specifics on the restaurant’s overall tip information, such as who, when, and how much. The first image below shows that a man’s average tip was 16% of his total bill, and equaled a little over $3.00. It also shows men tend to tip more at dinner versus lunch, and tip most often on Saturdays.

A picture containing chart

Description automatically generated

This second image shows that the average tip for the restaurant was 16%, coming in at $3.00. It also shows that dinners and Sundays are the times when most tips occur.

Chart, bar chart

Description automatically generated