

10/31/2021

538's Halloween Candy Rankings

Bryan Sam

With Halloween around the corner, both kids and adults love candy especially with brains! FiveThirtyEight has a dataset of information on different types of candy. I'll explore the dataset by asking these types of questions:

- Does a candy have peanut and almond flavor?
- Is there a fruity candy?
- How many types of candy out there?
- How much does it cost compare to other candies?
- Which candy do people prefer?

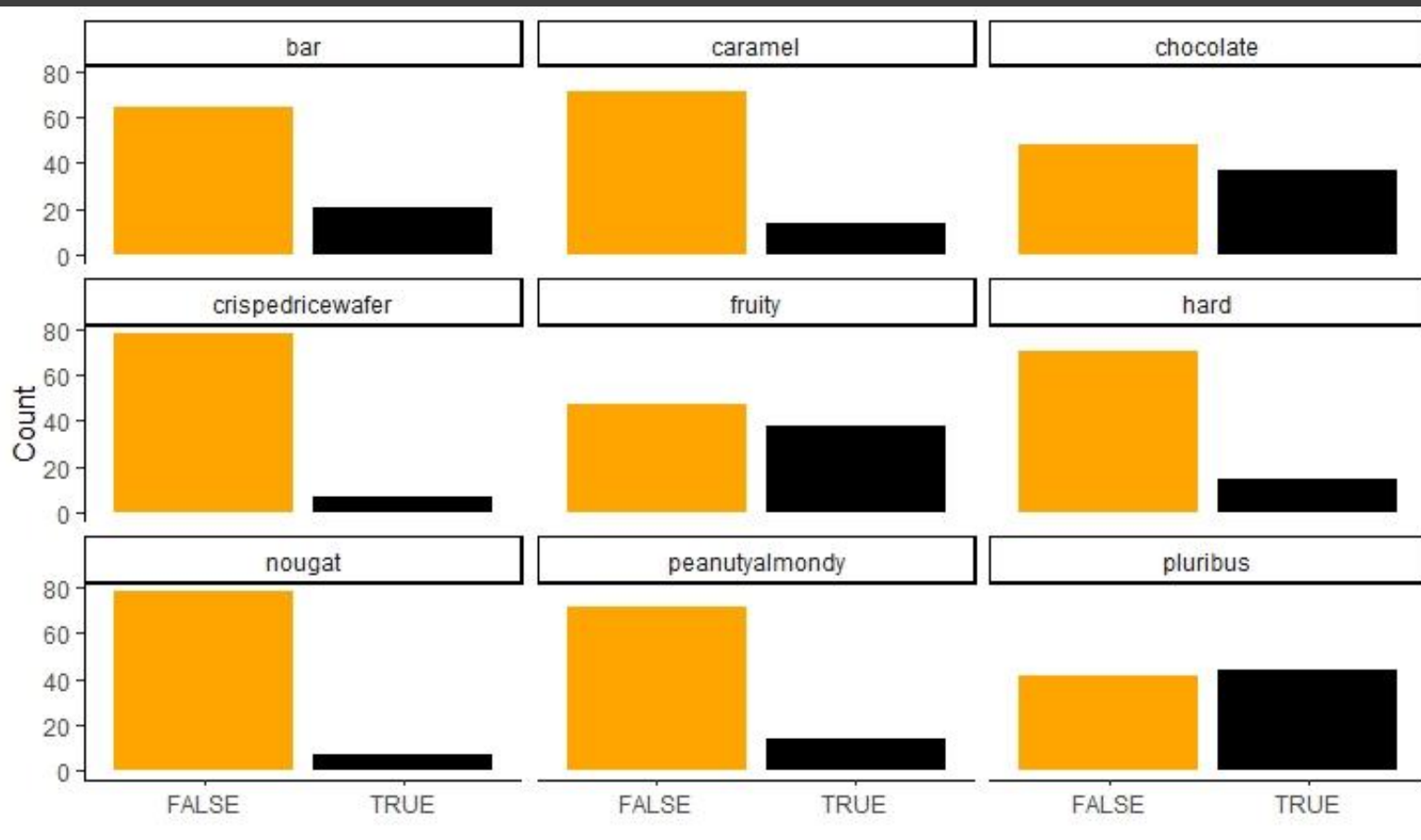


Halloween Candy Ranking Dataset

Attribute	Description
Competitorname	Candy Manufacturer
Chocolate	Does it contain chocolate?
Fruity	Is it fruit flavored?
Caramel	Is there caramel in the candy?
Peanutyalmondy	Does it contain peanuts, peanut butter or almonds?
Nougat	Does it contain nougat?
Crispedricewafer	Does it contain crisped rice, wafers, or a cookie component?
Hard	Is it a hard candy?
Bar	Is it a candy bar?
Pluribus	Is it one of many candies in a bag or box?
Sugarpercent	The percentile of sugar it falls under within the data set.
Pricepercent	The unit price percentile compared to the rest of the set.
Winpercent	The overall win percentage according to 269,000 matchups.

Preview of the Dataset

```
Rows: 85
Columns: 13
$ competitorname <chr> "100 Grand", "3 Musketeers", "One dime", "One quar...
$ chocolate      <lgl> TRUE, TRUE, FALSE, FALSE, FALSE, TRUE, TRUE, FALSE...
$ fruity         <lgl> FALSE, FALSE, FALSE, FALSE, TRUE, FALSE, FALSE, FA...
$ caramel        <lgl> TRUE, FALSE, FALSE, FALSE, FALSE, FALSE, TRUE, FAL...
$ peanutyalmondy <lgl> FALSE, FALSE, FALSE, FALSE, FALSE, TRUE, TRUE, TRU...
$ nougat         <lgl> FALSE, TRUE, FALSE, FALSE, FALSE, FALSE, TRUE, FAL...
$ crispedricewafer <lgl> TRUE, FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, FA...
$ hard           <lgl> FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, F...
$ bar            <lgl> TRUE, TRUE, FALSE, FALSE, FALSE, TRUE, TRUE, FALSE...
$ pluribus       <lgl> FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, T...
$ sugarpercent   <dbl> 0.732, 0.604, 0.011, 0.011, 0.906, 0.465, 0.604, 0...
$ pricepercent   <dbl> 0.860, 0.511, 0.116, 0.511, 0.511, 0.767, 0.767, 0...
$ winpercent     <dbl> 66.97173, 67.60294, 32.26109, 46.11650, 52.34146, ...
```

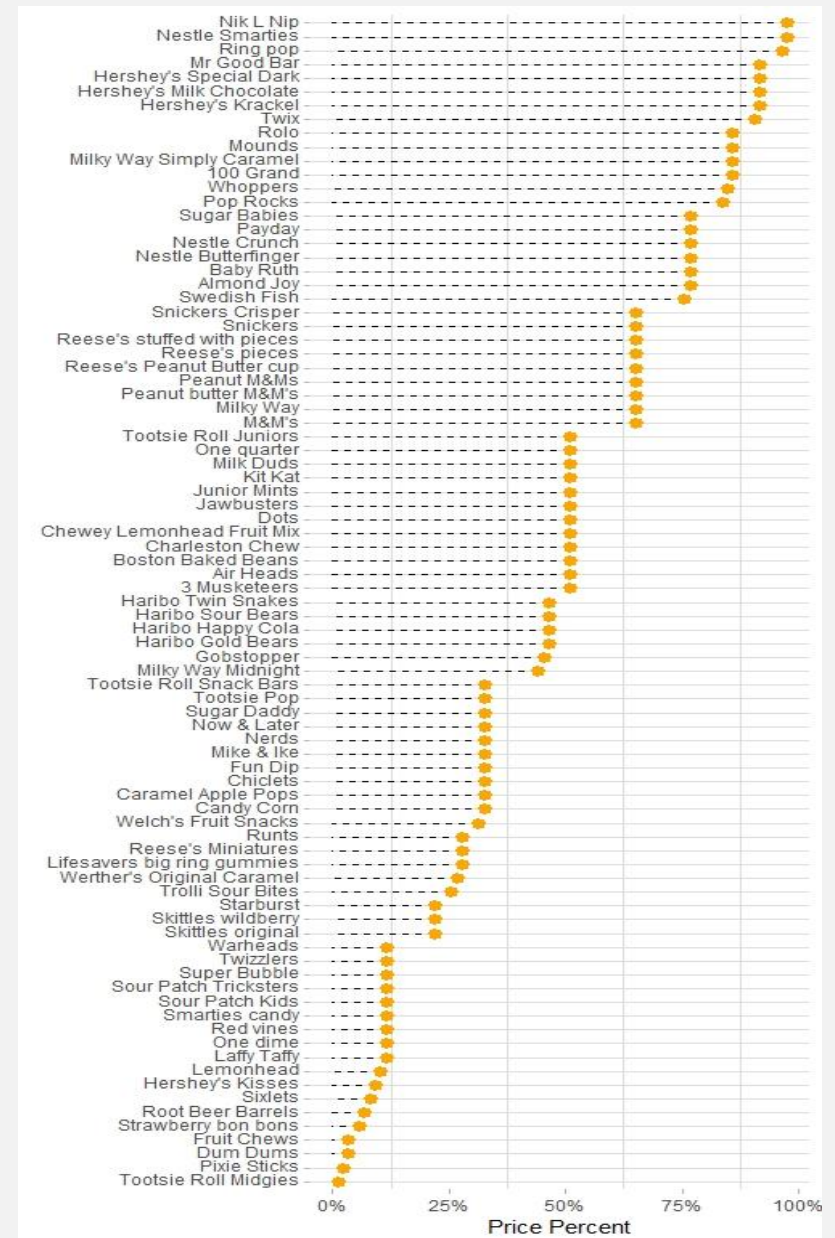


Distributions of Categorical Variables

- Shows the number of TRUE's and FALSE's for each type of candy in the dataset

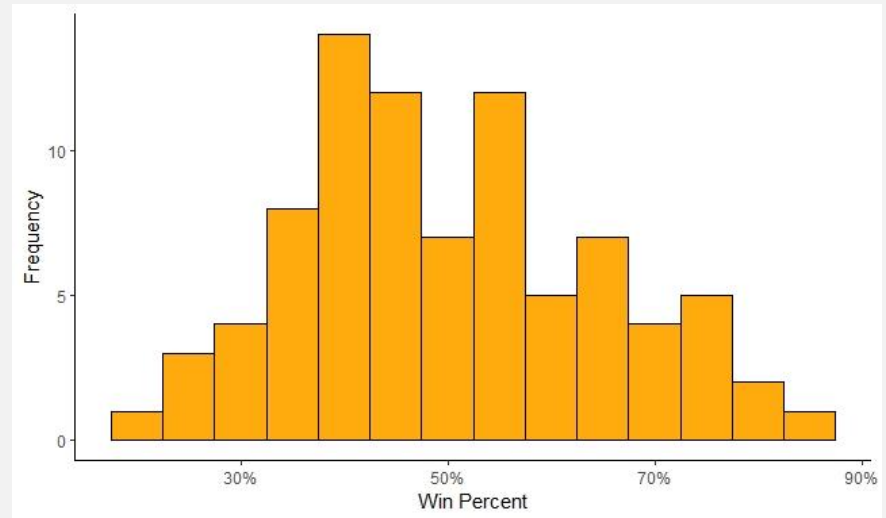
Price Percent Comparison

- The *pricepercent* variable is the percentile rank of the candy's price against to other candies in the dataset.
- The lollipop chart shows which candy is the most and least expensive.
- Most of the candies share the same ranking meaning a handful of candies are around the same price.



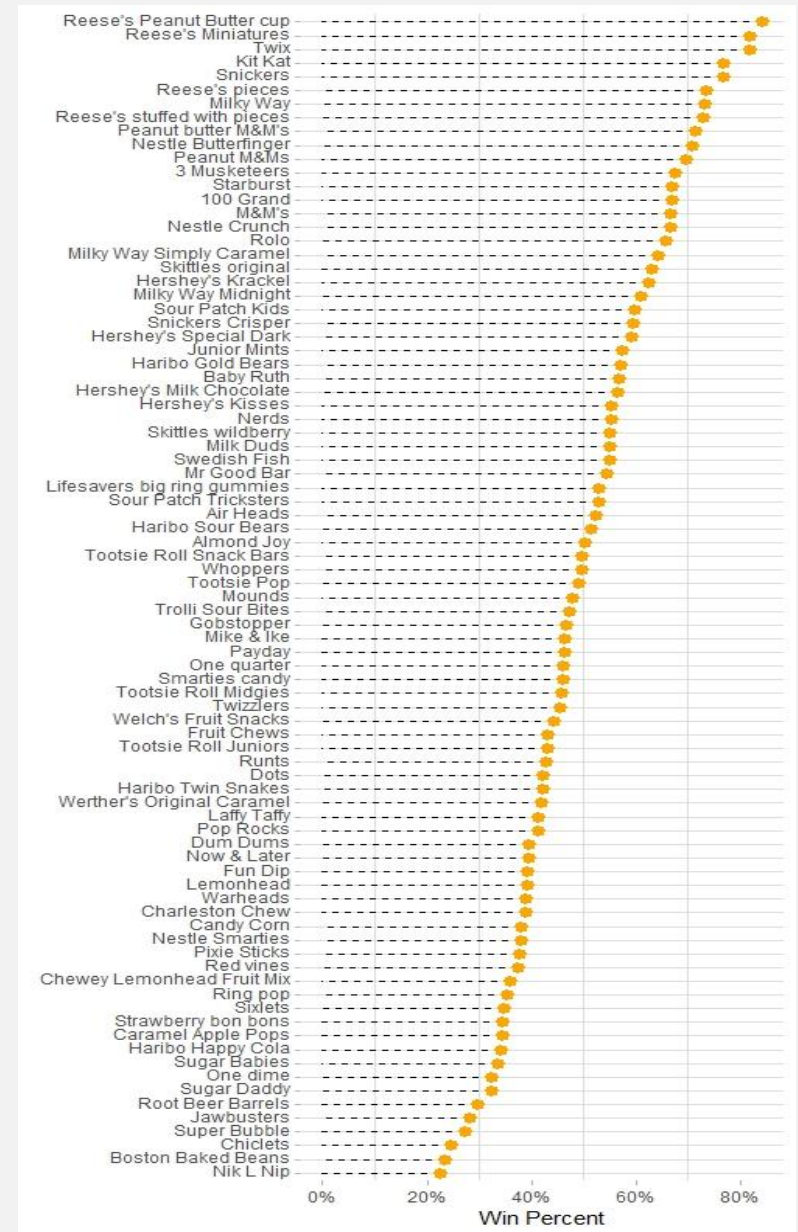
Win Percent Distribution

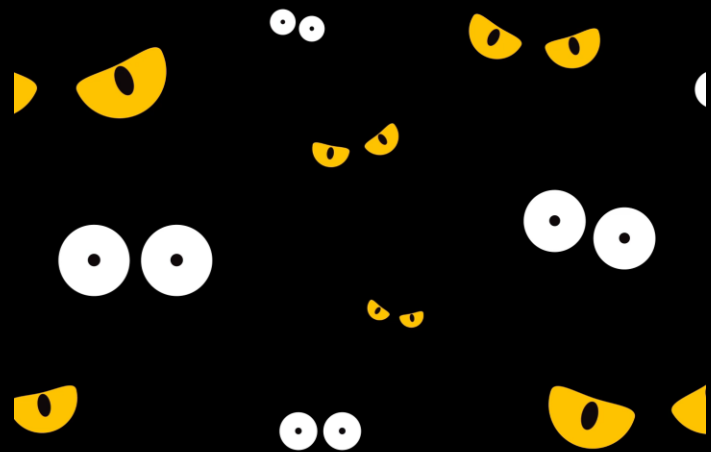
- The *winpercent* variable is the percentage of people who prefer this candy over another randomly chosen candy in the dataset.
- The histogram chart shows an approximate symmetrical distribution in the dataset.
- Most of it seems to center between 40% to 45%.



Win Percent Comparison

- After looking at the histogram, another lollipop chart shows the candy rankings according to people's preference.
- Reese's Peanut Butter Cups are people's favorite out of all candies in the dataset.
- Whereas Nik L Nip is people's least favorite candy.





Thank You!
Happy
Halloween!