HW 01 - Halloween Candy

Nathan Kim

Tue, Jan 28. Due Wed, Jan 29 at 11:59p

```
library(tidyverse)
library(fivethirtyeight)
```

Exercise 1

```
candy_rankings
```

```
## # A tibble: 85 x 13
##    competitorname chocolate fruity caramel peanutyalmondy nougat
##    <chr>          <lgl>     <lgl>  <lgl>   <lgl>          <lgl>
##  1 100 Grand      TRUE      FALSE  TRUE    FALSE          FALSE
##  2 3 Musketeers   TRUE      FALSE  FALSE   FALSE          TRUE
##  3 One dime       FALSE     FALSE  FALSE   FALSE          FALSE
##  4 One quarter    FALSE     FALSE  FALSE   FALSE          FALSE
##  5 Air Heads      FALSE     TRUE   FALSE   FALSE          FALSE
##  6 Almond Joy     TRUE      FALSE  FALSE   TRUE           FALSE
##  7 Baby Ruth      TRUE      FALSE  TRUE    TRUE           TRUE
##  8 Boston Baked ~ FALSE     FALSE  FALSE   TRUE           FALSE
##  9 Candy Corn     FALSE     FALSE  FALSE   FALSE          FALSE
## 10 Caramel Apple~ FALSE     TRUE   TRUE    FALSE          FALSE
## # ... with 75 more rows, and 7 more variables: crispedricewafer <lgl>,
## #   hard <lgl>, bar <lgl>, pluribus <lgl>, sugarpercent <dbl>,
## #   pricepercent <dbl>, winpercent <dbl>
```

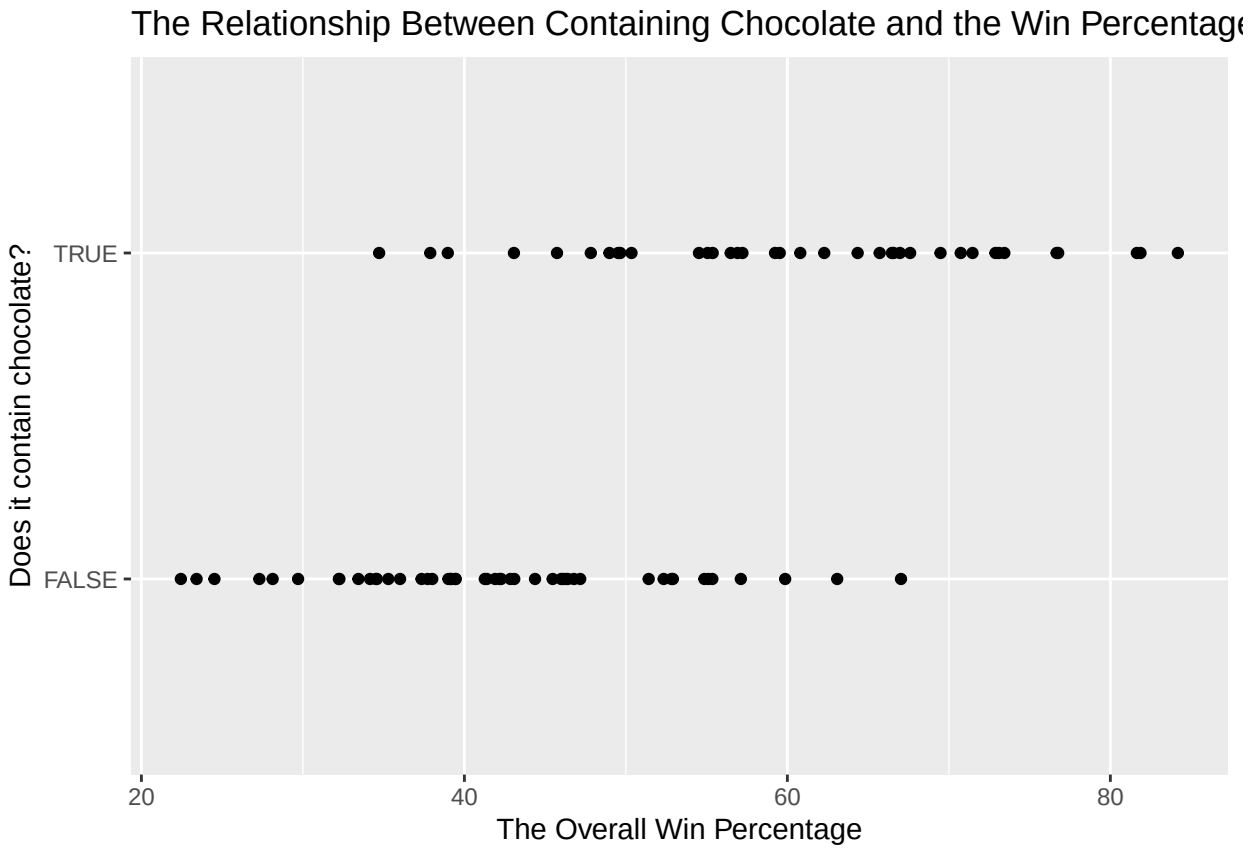
There are 13 variables and 85 observations

Exercise 2

2 categorical variables are chocolate and fruity. 2 numerical variables are sugarpercent and pricepercent.

Exercise 3

```
ggplot(data=candy_rankings, mapping = aes(x=winpercent, y=chocolate)) +
  geom_point()+
  labs(title = "The Relationship Between Containing Chocolate and the Win Percentage",x="The Overall Wi
```



If the candy contains chocolate, it is more likely to win than candy that does not contain chocolate

Exercise 4

```
options(digits=4)
```

```
candy_rankings%>%
  filter(winpercent<32.26)%>%
  arrange(desc(winpercent))%>%
  select(competitorname, winpercent)
```

```
## # A tibble: 7 x 2
##   competitorname     winpercent
##   <chr>                   <dbl>
## 1 Sugar Daddy              32.2
## 2 Root Beer Barrels        29.7
## 3 Jawbusters               28.1
## 4 Super Bubble             27.3
## 5 Chiclets                 24.5
## 6 Boston Baked Beans       23.4
## 7 Nik L Nip                22.4
```

Exercise 5

```
candy_rankings%>%
  group_by(chocolate)%>%
  summarise(mean=mean(winpercent))
```

```
## # A tibble: 2 x 2
##   chocolate  mean
##   <lgl>     <dbl>
## 1 FALSE      42.1
## 2 TRUE       60.9
```

Exercise 6

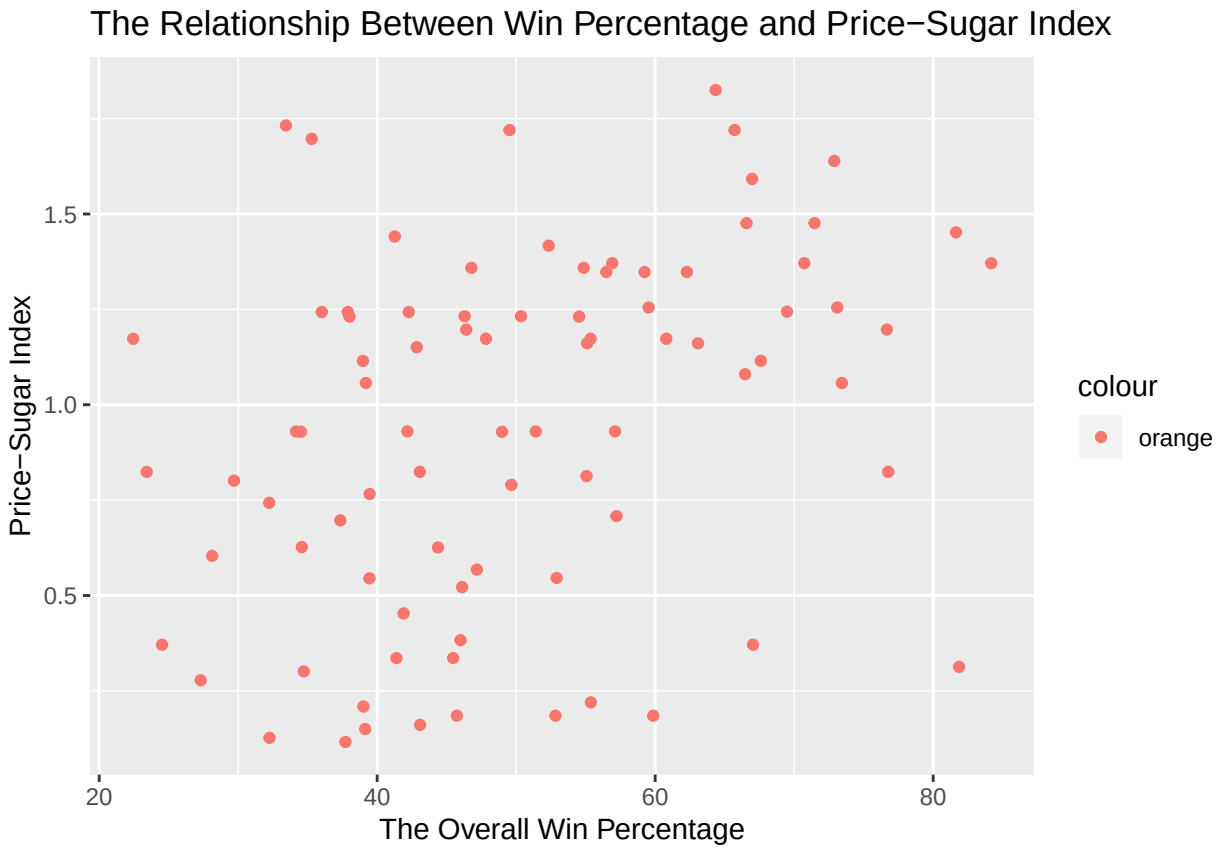
```
candy_rankings%>%
  filter(fruity)%>%
  select(competitorname, winpercent)%>%
  arrange(desc(winpercent))%>%
  slice(1:3)
```

```
## # A tibble: 3 x 2
##   competitorname    winpercent
##   <chr>                  <dbl>
## 1 Starburst               67.0
## 2 Skittles original       63.1
## 3 Sour Patch Kids         59.9
```

Exercise 7

```
candy_rankings<-candy_rankings%>%
  mutate(price_sugar_index=pricepercent+sugarpercent)

ggplot(data=candy_rankings, mapping = aes(x=winpercent, y=price_sugar_index, color="orange")) +
  geom_point()+
  labs(title = "The Relationship Between Win Percentage and Price-Sugar Index",x="The Overall Win Perce
```

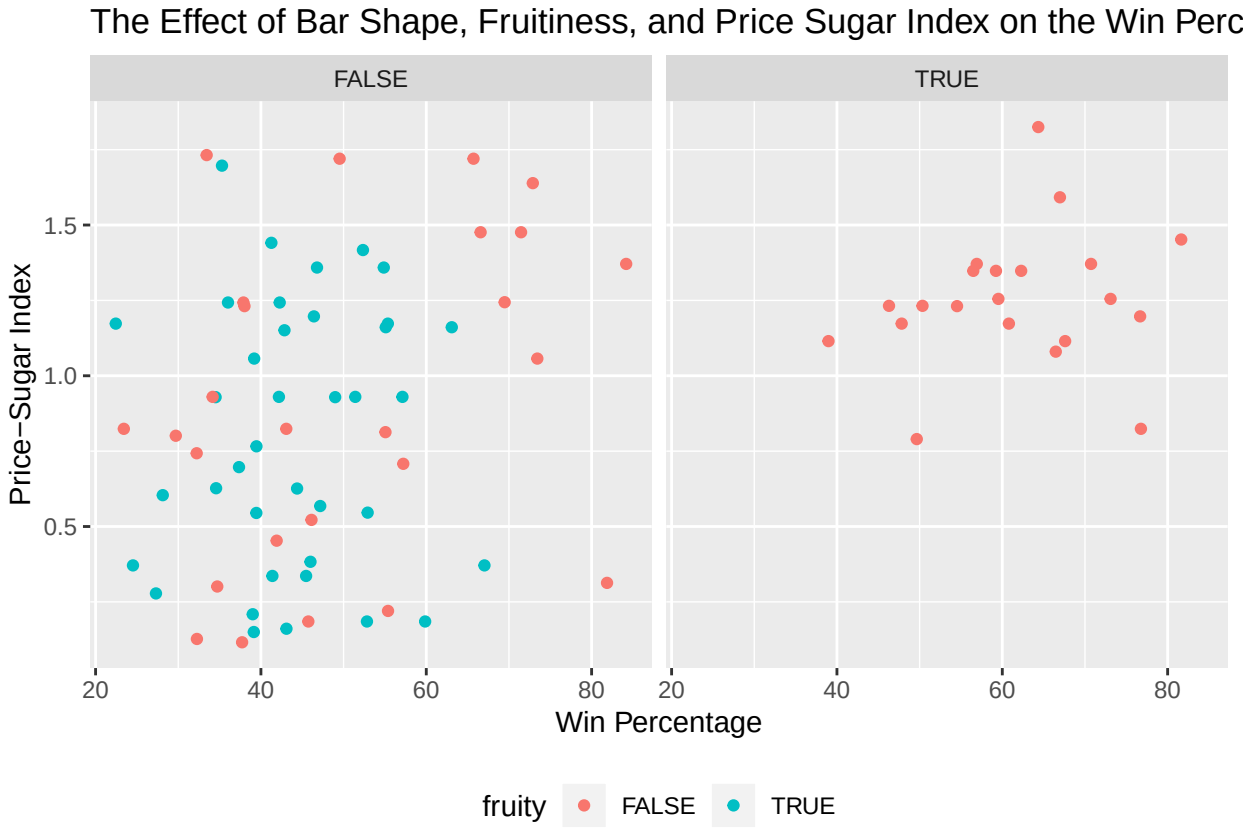


Typically, the more sugary and expensive the candy is, the more likely it is to win. This is most likely due to expensive candies being more high quality and considered to be tasty. The sugarpercent also contributes to the tastiness and popularity of candies. However, this is a fairly weak correlation.

Exercise 8

```
candy_rankings<-candy_rankings%>%
  mutate(price_sugar_index=pricepercent+sugarpercent)

ggplot(data=candy_rankings, mapping = aes(x=winpercent, y=price_sugar_index, color=fruity)) +
  geom_point()+
  facet_wrap(~ bar, ncol=2)+
  theme(legend.position = "bottom")+
  labs(title = "The Effect of Bar Shape, Fruitiness, and Price Sugar Index on the Win Percentage of Can
```



This plot shows that candies that are bar shaped tend to be more popular. Among those that are not bar shaped, candies that are not fruit flavored tend to be more popular than fruit flavored candies. Fruit flavored candies also tend to be cheaper or less sugary than non fruit flavored candies. Bar shaped candies also tend to be a bit more expensive and/or sugary.

Don't forget to knit to .pdf and upload on Gradescope! You must associate each question with a page and the "Overall section" with the first page.