Class09

Alyssa Hayashi

Here we analyze a candy dataset

```
candy<- read.csv("candy-data.csv",row.names=1)
candy
```

	chocolate	fruity	caramel	peanutyalmondy	nougat
100 Grand	1	0	1	0	0
3 Musketeers	1	0	0	0	1
One dime	0	0	0	0	0
One quarter	0	0	0	0	0
Air Heads	0	1	0	0	0
Almond Joy	1	0	0	1	0
Baby Ruth	1	0	1	1	1
Boston Baked Beans	0	0	0	1	0
Candy Corn	0	0	0	0	0
Caramel Apple Pops	0	1	1	0	0
Charleston Chew	1	0	0	0	1
Chewey Lemonhead Fruit Mix	0	1	0	0	0
Chiclets	0	1	0	0	0
Dots	0	1	0	0	0
Dum Dums	0	1	0	0	0
Fruit Chews	0	1	0	0	0
Fun Dip	0	1	0	0	0
Gobstopper	0	1	0	0	0
Haribo Gold Bears	0	1	0	0	0
Haribo Happy Cola	0	0	0	0	0
Haribo Sour Bears	0	1	0	0	0
Haribo Twin Snakes	0	1	0	0	0
Hershey's Kisses	1	0	0	0	0
Hershey's Krackel	1	0	0	0	0
Hershey's Milk Chocolate	1	0	0	0	0

Hershey's Special Dark	1	0	0	0	0
Jawbusters	0	1	0	0	0
Junior Mints	1	0	0	0	0
Kit Kat	1	0	0	0	0
Laffy Taffy	0	1	0	0	0
Lemonhead	0	1	0	0	0
Lifesavers big ring gummies	0	1	0	0	0
Peanut butter M&M's	1	0	0	1	0
M&M's	1	0	0	0	0
Mike & Ike	0	1	0	0	0
Milk Duds	1	0	1	0	0
Milky Way	1	0	1	0	1
Milky Way Midnight	1	0	1	0	1
Milky Way Simply Caramel	1	0	1	0	0
Mounds	1	0	0	0	0
Mr Good Bar	1	0	0	1	0
Nerds	0	1	0	0	0
Nestle Butterfinger	1	0	0	1	0
Nestle Crunch	1	0	0	0	0
Nik L Nip	0	1	0	0	0
Now & Later	0	1	0	0	0
Payday	0	0	0	1	1
Peanut M&Ms	1	0	0	1	0
Pixie Sticks	0	0	0	0	0
Pop Rocks	0	1	0	0	0
Red vines	0	1	0	0	0
Reese's Miniatures	1	0	0	1	0
Reese's Peanut Butter cup	1	0	0	1	0
Reese's pieces	1	0	0	1	0
Reese's stuffed with pieces	1	0	0	1	0
Ring pop	0	1	0	0	0
Rolo	1	0	1	0	0
Root Beer Barrels	0	0	0	0	0
Runts	0	1	0	0	0
Sixlets	1	0	0	0	0
Skittles original	0	1	0	0	0
Skittles wildberry	0	1	0	0	0
Nestle Smarties	1	0	0	0	0
Smarties candy	0	1	0	0	0
Snickers	1	0	1	1	1
Snickers Crisper	1	0	1	1	0
Sour Patch Kids	0	1	0	0	0
Sour Patch Tricksters	0	1	0	0	0

Starburst	0	1	0	0	0
Strawberry bon bons	0	1	0	0	0
Sugar Babies	0	0	1	0	0
Sugar Daddy	0	0	1	0	0
Super Bubble	0	1	0	0	0
Swedish Fish	0	1	0	0	0
Tootsie Pop	1	1	0	0	0
Tootsie Roll Juniors	1	0	0	0	0
Tootsie Roll Midgies	1	0	0	0	0
Tootsie Roll Snack Bars	1	0	0	0	0
Trolli Sour Bites	0	1	0	0	0
Twix	1	0	1	0	0
Twizzlers	0	1	0	0	0
Warheads	0	1	0	0	0
Welch's Fruit Snacks	0	1	0	0	0
Werther's Original Caramel	0	0	1	0	0
Whoppers	1	0	0	0	0

	crispedricewafer	hard	bar	pluribus	sugarpercent
100 Grand	1	0	1	0	0.732
3 Musketeers	0	0	1	0	0.604
One dime	0	0	0	0	0.011
One quarter	0	0	0	0	0.011
Air Heads	0	0	0	0	0.906
Almond Joy	0	0	1	0	0.465
Baby Ruth	0	0	1	0	0.604
Boston Baked Beans	0	0	0	1	0.313
Candy Corn	0	0	0	1	0.906
Caramel Apple Pops	0	0	0	0	0.604
Charleston Chew	0	0	1	0	0.604
Chewey Lemonhead Fruit Mix	0	0	0	1	0.732
Chiclets	0	0	0	1	0.046
Dots	0	0	0	1	0.732
Dum Dums	0	1	0	0	0.732
Fruit Chews	0	0	0	1	0.127
Fun Dip	0	1	0	0	0.732
Gobstopper	0	1	0	1	0.906
Haribo Gold Bears	0	0	0	1	0.465
Haribo Happy Cola	0	0	0	1	0.465
Haribo Sour Bears	0	0	0	1	0.465
Haribo Twin Snakes	0	0	0	1	0.465
Hershey's Kisses	0	0	0	1	0.127
Hershey's Krackel	1	0	1	0	0.430
Hershey's Milk Chocolate	0	0	1	0	0.430

Hershey's Special Dark	0	0	1	0	0.430
Jawbusters	0	1	0	1	0.093
Junior Mints	0	0	0	1	0.197
Kit Kat	1	0	1	0	0.313
Laffy Taffy	0	0	0	0	0.220
Lemonhead	0	1	0	0	0.046
Lifesavers big ring gummies	0	0	0	0	0.267
Peanut butter M&M's	0	0	0	1	0.825
M&M's	0	0	0	1	0.825
Mike & Ike	0	0	0	1	0.872
Milk Duds	0	0	0	1	0.302
Milky Way	0	0	1	0	0.604
Milky Way Midnight	0	0	1	0	0.732
Milky Way Simply Caramel	0	0	1	0	0.965
Mounds	0	0	1	0	0.313
Mr Good Bar	0	0	1	0	0.313
Nerds	0	1	0	1	0.848
Nestle Butterfinger	0	0	1	0	0.604
Nestle Crunch	1	0	1	0	0.313
Nik L Nip	0	0	0	1	0.197
Now & Later	0	0	0	1	0.220
Payday	0	0	1	0	0.465
Peanut M&Ms	0	0	0	1	0.593
Pixie Sticks	0	0	0	1	0.093
Pop Rocks	0	1	0	1	0.604
Red vines	0	0	0	1	0.581
Reese's Miniatures	0	0	0	0	0.034
Reese's Peanut Butter cup	0	0	0	0	0.720
Reese's pieces	0	0	0	1	0.406
Reese's stuffed with pieces	0	0	0	0	0.988
Ring pop	0	1	0	0	0.732
Rolo	0	0	0	1	0.860
Root Beer Barrels	0	1	0	1	0.732
Runts	0	1	0	1	0.872
Sixlets	0	0	0	1	0.220
Skittles original	0	0	0	1	0.941
Skittles wildberry	0	0	0	1	0.941
Nestle Smarties	0	0	0	1	0.267
Smarties candy	0	1	0	1	0.267
Snickers	0	0	1	0	0.546
Snickers Crisper	1	0	1	0	0.604
Sour Patch Kids	0	0	0	1	0.069
Sour Patch Tricksters	0	0	0	1	0.069

Starburst	0	0	0	1	0.151
Strawberry bon bons	0	1	0	1	0.569
Sugar Babies	0	0	0	1	0.965
Sugar Daddy	0	0	0	0	0.418
Super Bubble	0	0	0	0	0.162
Swedish Fish	0	0	0	1	0.604
Tootsie Pop	0	1	0	0	0.604
Tootsie Roll Juniors	0	0	0	0	0.313
Tootsie Roll Midgies	0	0	0	1	0.174
Tootsie Roll Snack Bars	0	0	1	0	0.465
Trolli Sour Bites	0	0	0	1	0.313
Twix	1	0	1	0	0.546
Twizzlers	0	0	0	0	0.220
Warheads	0	1	0	0	0.093
Welch's Fruit Snacks	0	0	0	1	0.313
Werther's Original Caramel	0	1	0	0	0.186
Whoppers	1	0	0	1	0.872

	pricepercent	winpercent
100 Grand	0.860	66.97173
3 Musketeers	0.511	67.60294
One dime	0.116	32.26109
One quarter	0.511	46.11650
Air Heads	0.511	52.34146
Almond Joy	0.767	50.34755
Baby Ruth	0.767	56.91455
Boston Baked Beans	0.511	23.41782
Candy Corn	0.325	38.01096
Caramel Apple Pops	0.325	34.51768
Charleston Chew	0.511	38.97504
Chewey Lemonhead Fruit Mix	0.511	36.01763
Chiclets	0.325	24.52499
Dots	0.511	42.27208
Dum Dums	0.034	39.46056
Fruit Chews	0.034	43.08892
Fun Dip	0.325	39.18550
Gobstopper	0.453	46.78335
Haribo Gold Bears	0.465	57.11974
Haribo Happy Cola	0.465	34.15896
Haribo Sour Bears	0.465	51.41243
Haribo Twin Snakes	0.465	42.17877
Hershey's Kisses	0.093	55.37545
Hershey's Krackel	0.918	62.28448
Hershey's Milk Chocolate	0.918	56.49050

Hershey's Special Dark	0.918	59.23612
Jawbusters	0.511	28.12744
Junior Mints	0.511	57.21925
Kit Kat	0.511	76.76860
Laffy Taffy	0.116	41.38956
Lemonhead	0.104	39.14106
Lifesavers big ring gummies	0.279	52.91139
Peanut butter M&M's	0.651	71.46505
M&M's	0.651	66.57458
Mike & Ike	0.325	46.41172
Milk Duds	0.511	55.06407
Milky Way	0.651	73.09956
Milky Way Midnight	0.441	60.80070
Milky Way Simply Caramel	0.860	64.35334
Mounds	0.860	47.82975
Mr Good Bar	0.918	54.52645
Nerds	0.325	55.35405
Nestle Butterfinger	0.767	70.73564
Nestle Crunch	0.767	66.47068
Nik L Nip	0.976	22.44534
Now & Later	0.325	39.44680
Payday	0.767	46.29660
Peanut M&Ms	0.651	69.48379
Pixie Sticks	0.023	37.72234
Pop Rocks	0.837	41.26551
Red vines	0.116	37.34852
Reese's Miniatures	0.279	81.86626
Reese's Peanut Butter cup	0.651	84.18029
Reese's pieces	0.651	73.43499
Reese's stuffed with pieces	0.651	72.88790
Ring pop	0.965	35.29076
Rolo	0.860	65.71629
Root Beer Barrels	0.069	29.70369
Runts	0.279	42.84914
Sixlets	0.081	34.72200
Skittles original	0.220	63.08514
Skittles wildberry	0.220	55.10370
Nestle Smarties	0.976	37.88719
Smarties candy	0.116	45.99583
Snickers	0.651	76.67378
Snickers Crisper	0.651	59.52925
Sour Patch Kids	0.116	59.86400
Sour Patch Tricksters	0.116	52.82595

```
Starburst                       0.220   67.03763
Strawberry bon bons             0.058   34.57899
Sugar Babies                    0.767   33.43755
Sugar Daddy                     0.325   32.23100
Super Bubble                    0.116   27.30386
Swedish Fish                    0.755   54.86111
Tootsie Pop                     0.325   48.98265
Tootsie Roll Juniors            0.511   43.06890
Tootsie Roll Midgies            0.011   45.73675
Tootsie Roll Snack Bars         0.325   49.65350
Trolli Sour Bites               0.255   47.17323
Twix                            0.906   81.64291
Twizzlers                       0.116   45.46628
Warheads                        0.116   39.01190
Welch's Fruit Snacks            0.313   44.37552
Werther's Original Caramel      0.267   41.90431
Whoppers                        0.848   49.52411
```

Q1. How many different candy types are in this dataset? Q2. How many fruity candy types are in the dataset?

```
nrow(candy)
```

```
[1] 85
```

```
sum(candy$fruity)
```

```
[1] 38
```

Q3. What is your favorite candy in the dataset and what is it's winpercent value? Q4. What is the winpercent value for "Kit Kat"? Q5. What is the winpercent value for "Tootsie Roll Snack Bars"?

```
#View(candy)
candy["Twix", ]$winpercent
```

```
[1] 81.64291
```

```
candy["Kit Kat",]$winpercent
```

```
[1] 76.7686
```

```
candy["Tootsie Roll Snack Bars", ]$winpercent
```

```
[1] 49.6535
```

```
x<- c(5, 3,4,1)
sort(x)
```

```
[1] 1 3 4 5
```

```
order(x)
```

```
[1] 4 2 3 1
```

```
inds<- order(candy$winpercent)
head(candy[inds,])
```

```
                  chocolate fruity caramel peanutyalmondy nougat
Nik L Nip                 0      1       0              0      0
Boston Baked Beans        0      0       0              1      0
Chiclets                  0      1       0              0      0
Super Bubble              0      1       0              0      0
Jawbusters                0      1       0              0      0
Root Beer Barrels         0      0       0              0      0
                  crispedricewafer hard bar pluribus sugarpercent pricepercent
Nik L Nip                        0    0   0        1        0.197        0.976
Boston Baked Beans               0    0   0        1        0.313        0.511
Chiclets                         0    0   0        1        0.046        0.325
Super Bubble                     0    0   0        0        0.162        0.116
Jawbusters                       0    1   0        1        0.093        0.511
Root Beer Barrels                0    1   0        1        0.732        0.069
                  winpercent
Nik L Nip           22.44534
```

```
Boston Baked Beans   23.41782
Chiclets             24.52499
Super Bubble         27.30386
Jawbusters           28.12744
Root Beer Barrels    29.70369
```

```
skimr::skim(candy)
```

Table 1: Data summary

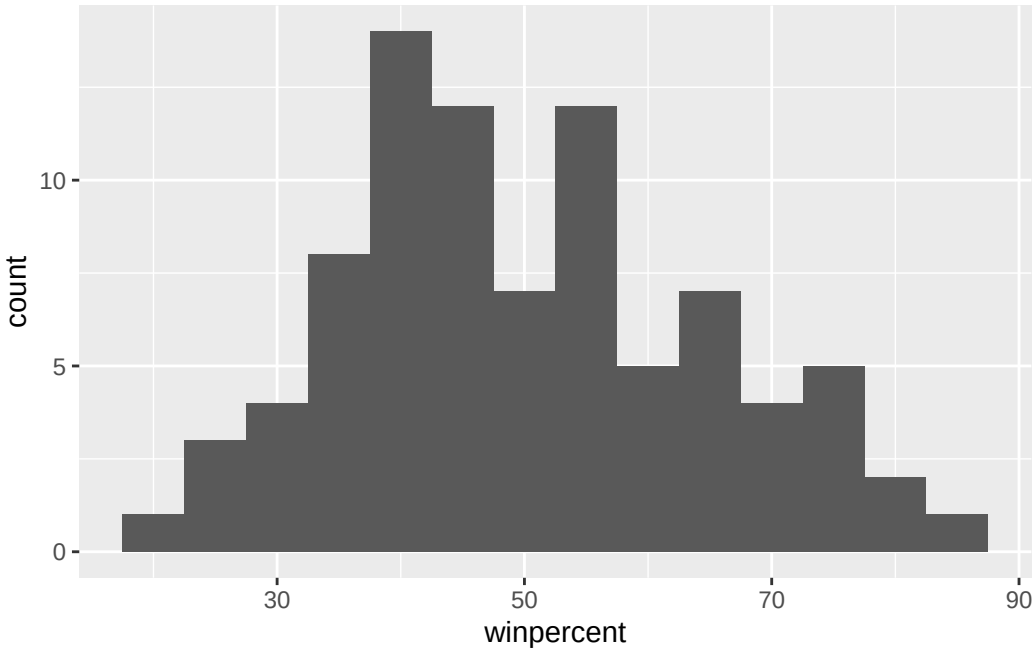
Name	candy
Number of rows	85
Number of columns	12
_____	
Column type frequency:	
numeric	12
_____	
Group variables	None

Variable type: numeric

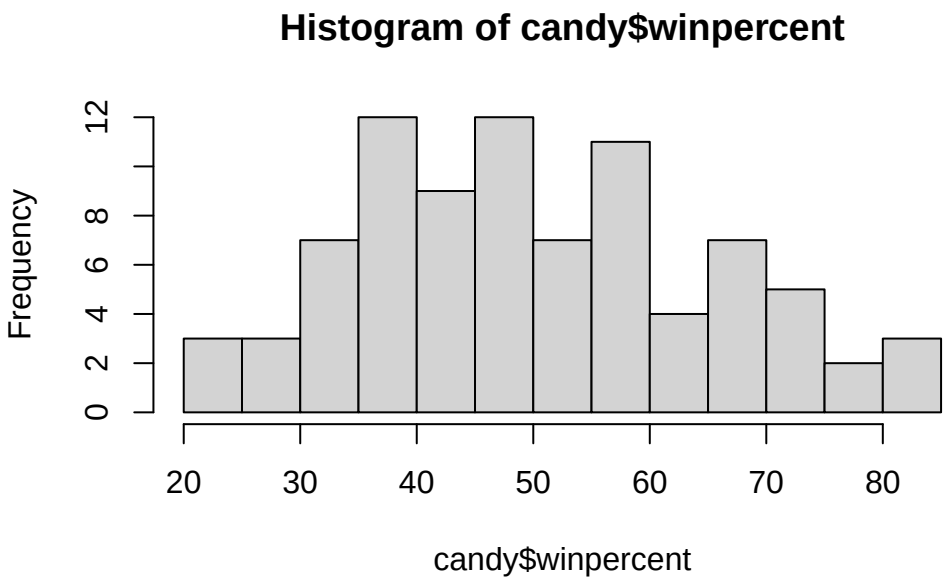
skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100	hist
chocolate	0	1	0.44	0.50	0.00	0.00	0.00	1.00	1.00	
fruity	0	1	0.45	0.50	0.00	0.00	0.00	1.00	1.00	
caramel	0	1	0.16	0.37	0.00	0.00	0.00	0.00	1.00	
peanutyalmondy	0	1	0.16	0.37	0.00	0.00	0.00	0.00	1.00	
nougat	0	1	0.08	0.28	0.00	0.00	0.00	0.00	1.00	
crispedricewafer	0	1	0.08	0.28	0.00	0.00	0.00	0.00	1.00	
hard	0	1	0.18	0.38	0.00	0.00	0.00	0.00	1.00	
bar	0	1	0.25	0.43	0.00	0.00	0.00	0.00	1.00	
pluribus	0	1	0.52	0.50	0.00	0.00	1.00	1.00	1.00	
sugarpercent	0	1	0.48	0.28	0.01	0.22	0.47	0.73	0.99	
pricepercent	0	1	0.47	0.29	0.01	0.26	0.47	0.65	0.98	
winpercent	0	1	50.32	14.71	22.45	39.14	47.83	59.86	84.18	

Q6. Is there any variable/column that looks to be on a different scale to the majority of the other columns in the dataset? Q7. What do you think a zero and one represent for the candy$chocolate column? Q8. Plot a histogram of winpercent values

```
library(ggplot2)
ggplot(candy, aes(winpercent))+ geom_histogram(binwidth = 5)
```




```
hist(candy$winpercent, breaks=15)
```



Q9. Is the distribution of winpercent values symmetrical? No it is not. It is skewed. Q10. Is the center of the distribution above or below 50%? The center is below 50% Q11. On average is chocolate candy higher or lower ranked than fruit candy?

First find all the chocolate candy and their $winpercent values

```
choc<- candy$winpercent[as.logical(candy$chocolate)]

choc
```

```
 [1] 66.97173 67.60294 50.34755 56.91455 38.97504 55.37545 62.28448 56.49050
 [9] 59.23612 57.21925 76.76860 71.46505 66.57458 55.06407 73.09956 60.80070
[17] 64.35334 47.82975 54.52645 70.73564 66.47068 69.48379 81.86626 84.18029
[25] 73.43499 72.88790 65.71629 34.72200 37.88719 76.67378 59.52925 48.98265
[33] 43.06890 45.73675 49.65350 81.64291 49.52411
```

Next summarize these values into one number

```
summary(choc)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
```

```
  34.72   50.35   60.80   60.92   70.74   84.18
```

Then do the same for fruity candy and compare numbers

```
fruity<- candy$winpercent[as.logical(candy$fruity)]
fruity
```

```
 [1] 52.34146 34.51768 36.01763 24.52499 42.27208 39.46056 43.08892 39.18550
 [9] 46.78335 57.11974 51.41243 42.17877 28.12744 41.38956 39.14106 52.91139
[17] 46.41172 55.35405 22.44534 39.44680 41.26551 37.34852 35.29076 42.84914
[25] 63.08514 55.10370 45.99583 59.86400 52.82595 67.03763 34.57899 27.30386
[33] 54.86111 48.98265 47.17323 45.46628 39.01190 44.37552
```

```
summary(fruity)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  22.45   39.04   42.97   44.12   52.11   67.04
```

On average chocolate candy is higher ranked than fruity candy.

Q12. Is this difference statistically significant?

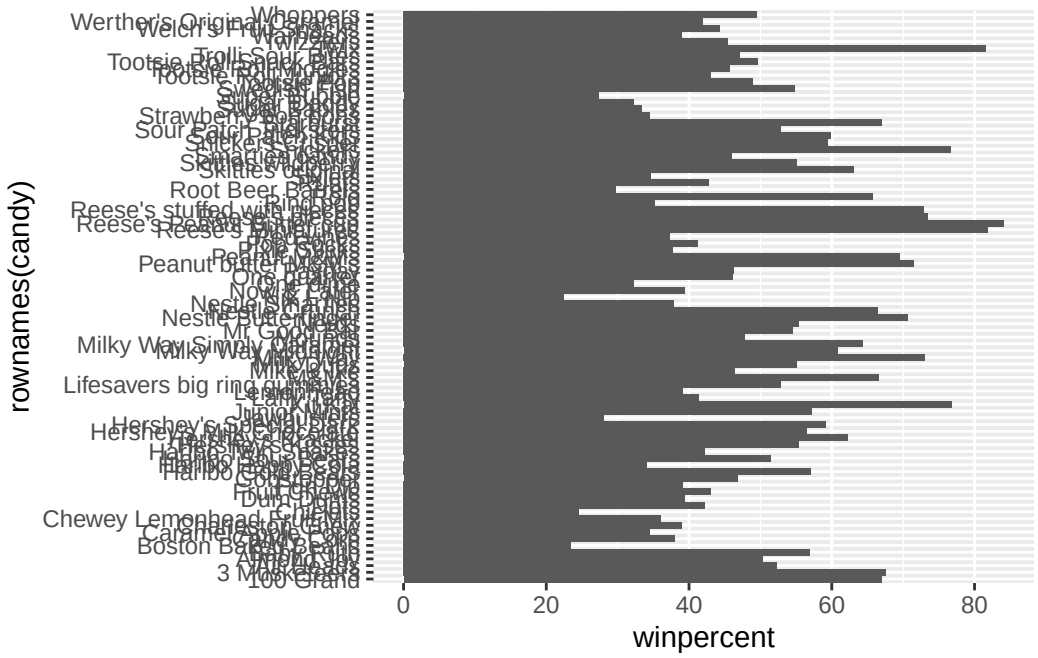
```
t.test(choc, fruity)
```

```

    Welch Two Sample t-test

data:  choc and fruity
t = 6.2582, df = 68.882, p-value = 2.871e-08
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 11.44563 22.15795
sample estimates:
mean of x mean of y
 60.92153  44.11974
```

The difference between chocolate and fruity candy is significant.

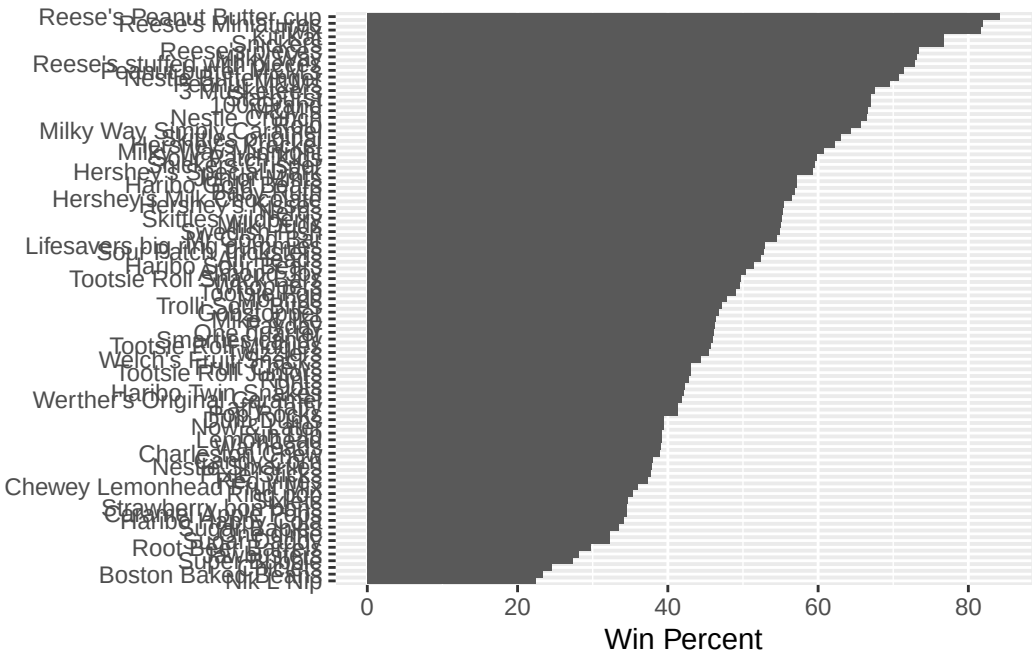
Q13. What are the five least liked candy types in this set? Q14. What are the top 5 all time favorite candy types out of this set?

```
library(ggplot2)
ggplot(candy)+
  aes(winpercent, rownames(candy))+
  geom_col()
```

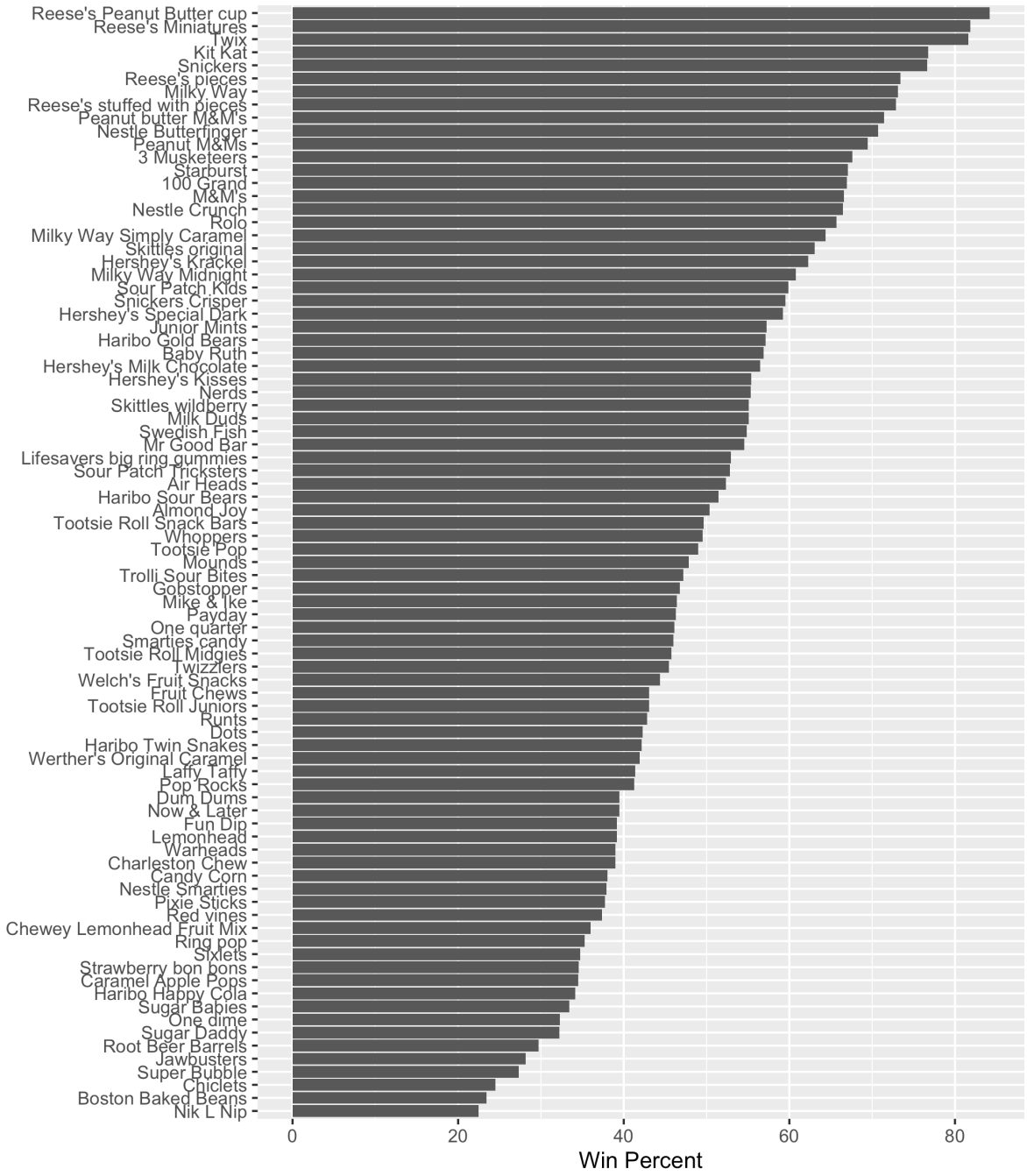


Q16. This is quite ugly, use the reorder() function to get the bars sorted by winpercent?

```
library(ggplot2)
ggplot(candy)+
  aes(winpercent, reorder(rownames(candy), winpercent))+
  geom_col() + labs(y=NULL, x= "Win Percent")
```




```
ggsave('barplot1.png',width=7, height=8)
```



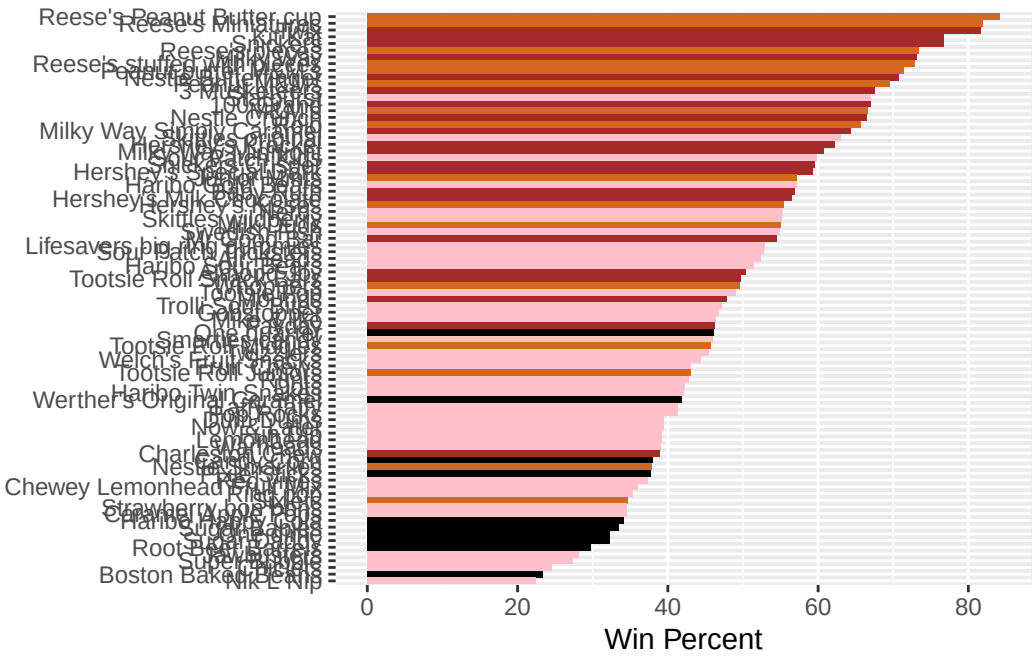
You can insert any image using this markdown syntax

Add some color to our ggplot. We need to make a sutom color vector

```
#start with all black colors
my_cols=rep("black", nrow(candy))
my_cols[as.logical(candy$chocolate)] = "chocolate"
my_cols[as.logical(candy$bar)] = "brown"
my_cols[as.logical(candy$fruity)] = "pink"
my_cols
```

```
 [1] "brown"     "brown"     "black"     "black"     "pink"      "brown"
 [7] "brown"     "black"     "black"     "pink"      "brown"     "pink"
[13] "pink"      "pink"      "pink"      "pink"      "pink"      "pink"
[19] "pink"      "black"     "pink"      "pink"      "chocolate" "brown"
[25] "brown"     "brown"     "pink"      "chocolate" "brown"     "pink"
[31] "pink"      "pink"      "chocolate" "chocolate" "pink"      "chocolate"
[37] "brown"     "brown"     "brown"     "brown"     "brown"     "pink"
[43] "brown"     "brown"     "pink"      "pink"      "brown"     "chocolate"
[49] "black"     "pink"      "pink"      "chocolate" "chocolate" "chocolate"
[55] "chocolate" "pink"      "chocolate" "black"     "pink"      "chocolate"
[61] "pink"      "pink"      "chocolate" "pink"      "brown"     "brown"
[67] "pink"      "pink"      "pink"      "pink"      "black"     "black"
[73] "pink"      "pink"      "pink"      "chocolate" "chocolate" "brown"
[79] "pink"      "brown"     "pink"      "pink"      "pink"      "black"
[85] "chocolate"
```

```
library(ggplot2)
ggplot(candy)+
  aes(winpercent, reorder(rownames(candy), winpercent))+
  geom_col(fill=my_cols) + labs(y=NULL, x= "Win Percent")
```



Q17. What is the worst ranked chocolate candy? sixlets Q18. What is the best ranked fruity candy? Skittles

```
candy$pricepercent
```

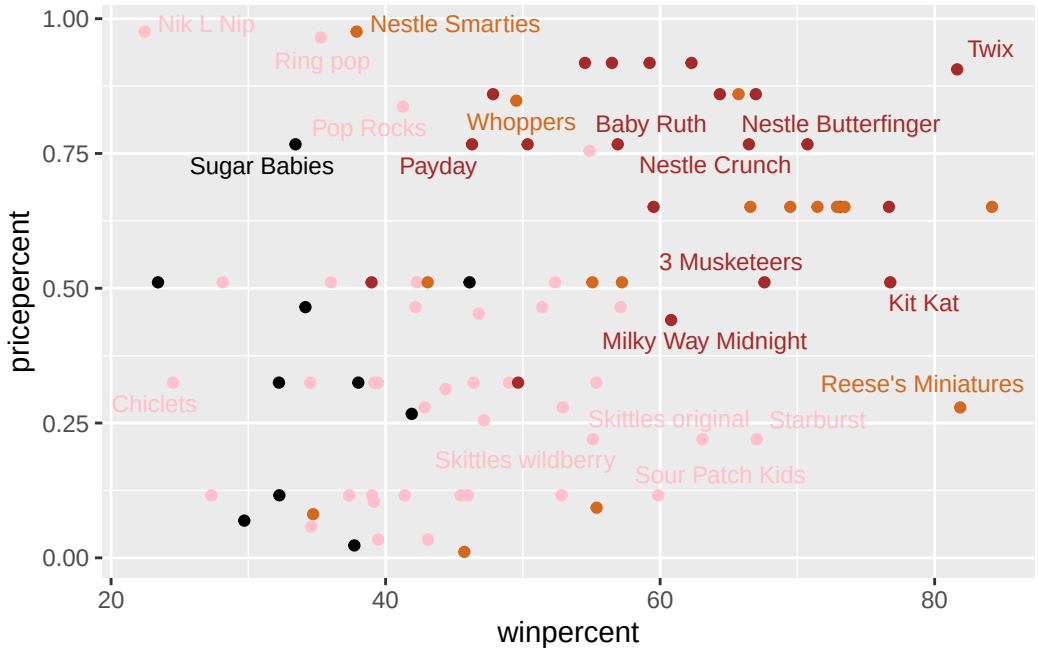
```
 [1] 0.860 0.511 0.116 0.511 0.511 0.767 0.767 0.511 0.325 0.325 0.511 0.511
[13] 0.325 0.511 0.034 0.034 0.325 0.453 0.465 0.465 0.465 0.465 0.093 0.918
[25] 0.918 0.918 0.511 0.511 0.511 0.116 0.104 0.279 0.651 0.651 0.325 0.511
[37] 0.651 0.441 0.860 0.860 0.918 0.325 0.767 0.767 0.976 0.325 0.767 0.651
[49] 0.023 0.837 0.116 0.279 0.651 0.651 0.651 0.965 0.860 0.069 0.279 0.081
[61] 0.220 0.220 0.976 0.116 0.651 0.651 0.116 0.116 0.220 0.058 0.767 0.325
[73] 0.116 0.755 0.325 0.511 0.011 0.325 0.255 0.906 0.116 0.116 0.313 0.267
[85] 0.848
```

If we want to see what is good candy to buy in terms of winpercent and pricepercent we can plot these two variable and then see the best candy foor the least amount of money

```
library(ggplot2)
library(ggrepel)

# How about a plot of price vs win
```

```
ggplot(candy) +
  aes(winpercent, pricepercent, label=rownames(candy)) +
  geom_point(col=my_cols) +
  geom_text_repel(col=my_cols, size=3.3, max.overlaps = 5)
```

```
Warning: ggrepel: 65 unlabeled data points (too many overlaps). Consider
increasing max.overlaps
```




```
  geom_text_repel(col=my_cols, size=3.3, max.overlaps = 5)
```

```
geom_text_repel: parse = FALSE, na.rm = FALSE, box.padding = 0.25, point.padding = 1e-06, mi
stat_identity: na.rm = FALSE
position_identity
```

```
ord <- order(candy$pricepercent, decreasing = TRUE)
head( candy[ord,c(11,12)], n=5 )
```

```
pricepercent winpercent
```

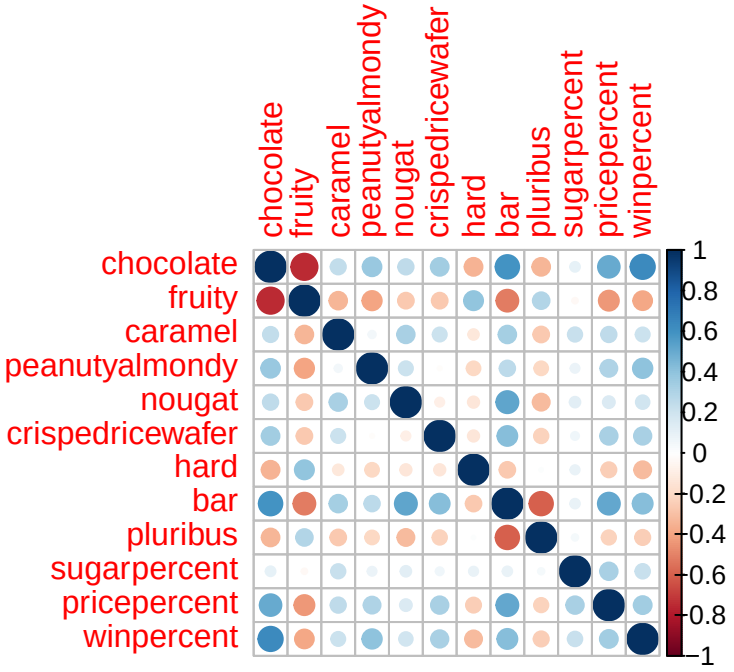
```
Nik L Nip                     0.976   22.44534
Nestle Smarties               0.976   37.88719
Ring pop                      0.965   35.29076
Hershey's Krackel             0.918   62.28448
Hershey's Milk Chocolate      0.918   56.49050
```

5. Explorng the Correlation Structure

```
library(corrplot)
```

```
corrplot 0.92 loaded
```

```
cij<-cor(candy)
corrplot(cij)
```



Q22. Examining this plot what two variables are anti-correlated (i.e. have minus values)? Q23. Similarly, what two variables are most positively correlated?

On to PCA

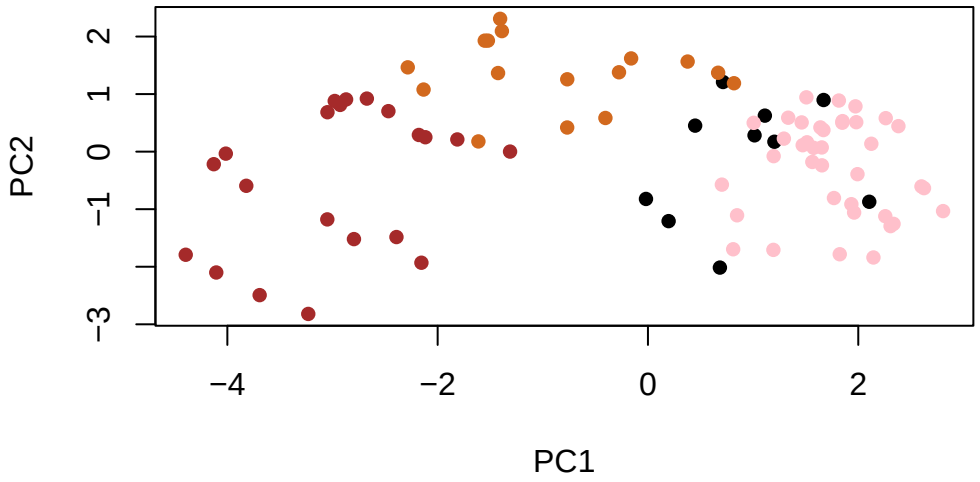
The main function for this is called 'prcom'

```
pca <- prcomp(candy, scale=T)
summary(pca)
```

```
Importance of components:
                          PC1    PC2    PC3     PC4    PC5     PC6     PC7
Standard deviation     2.0788 1.1378 1.1092 1.07533 0.9518 0.81923 0.81530
Proportion of Variance 0.3601 0.1079 0.1025 0.09636 0.0755 0.05593 0.05539
Cumulative Proportion  0.3601 0.4680 0.5705 0.66688 0.7424 0.79830 0.85369
                           PC8     PC9    PC10    PC11    PC12
Standard deviation     0.74530 0.67824 0.62349 0.43974 0.39760
Proportion of Variance 0.04629 0.03833 0.03239 0.01611 0.01317
Cumulative Proportion  0.89998 0.93832 0.97071 0.98683 1.00000
```

Plot my main PCA score with ggplot

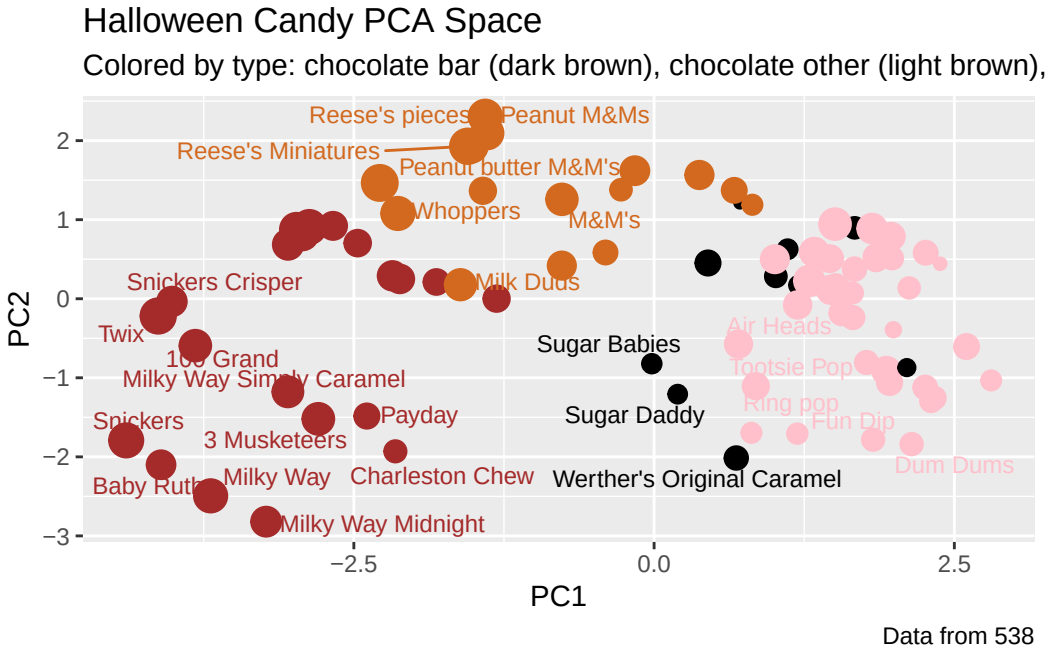
```
my_data <- cbind(candy, pca$x[,1:3])
plot(pca$x[,1:2], col=my_cols, pch=16)
```



```
p<-ggplot(my_data) +
        aes(x=PC1, y=PC2,
            size=winpercent/100,
            text=rownames(my_data),
            label=rownames(my_data)) +
        geom_point(col=my_cols)
```

```
library(ggrepel)

p + geom_text_repel(size=3.3, col=my_cols, max.overlaps = 7)  +
  theme(legend.position = "none") +
  labs(title="Halloween Candy PCA Space",
       subtitle="Colored by type: chocolate bar (dark brown), chocolate other (light brown
       caption="Data from 538")
```

```
Warning: ggrepel: 59 unlabeled data points (too many overlaps). Consider
increasing max.overlaps
```



```r
loadings<- as.data.frame(pca$rotation)

ggplot(loadings)+
  aes(PC1, reorder(rownames(loadings), PC1))+
  geom_col()
```