class10_Halloween.Mini.Project

Kalle Liimatta A59002114

Table of contents

1. Importing candy data

```
candy_file <- read.csv("candy-data.csv")
```

```
candy= data.frame(candy_file, row.names=1)
head(candy)
```

```
           chocolate fruity caramel peanutyalmondy nougat crispedricewafer
100 Grand          1      0       1              0      0                1
3 Musketeers       1      0       0              0      1                0
One dime           0      0       0              0      0                0
One quarter        0      0       0              0      0                0
Air Heads          0      1       0              0      0                0
Almond Joy         1      0       0              1      0                0
           hard bar pluribus sugarpercent pricepercent winpercent
100 Grand     0   1        0        0.732        0.860   66.97173
3 Musketeers  0   1        0        0.604        0.511   67.60294
One dime      0   0        0        0.011        0.116   32.26109
One quarter   0   0        0        0.011        0.511   46.11650
Air Heads     0   0        0        0.906        0.511   52.34146
Almond Joy    0   1        0        0.465        0.767   50.34755
```

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

```
nrow(candy)
```

```
[1] 85
```

85 different candy types

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

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

```
[1] 38
```

3

38 fruity candy types

2. What is your favorite candy?

```
candy["Twix", ]$winpercent
```

[1] 81.64291

Q3. What is your favorite candy in the dataset and what is it's winpercent value?

```
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
HersheyÕs Kisses	1	0	0	0	0
HersheyÕs Krackel	1	0	0	0	0
HersheyÕs Milk Chocolate	1	0	0	0	0
HersheyÕs 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&MÕs	1	0	0	1	0
M&MÕs	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
ReeseÕs Miniatures	1	0	0	1	0
ReeseÕs Peanut Butter cup	1	0	0	1	0
ReeseÕs pieces	1	0	0	1	0
ReeseÕs 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
WelchÕs Fruit Snacks	0	1	0	0	0
WertherÕs 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
HersheyÕs Kisses	0	0	0	1	0.127
HersheyÕs Krackel	1	0	1	0	0.430
HersheyÕs Milk Chocolate	0	0	1	0	0.430
HersheyÕs 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&MÕs	0	0	0	1	0.825
M&MÕs	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
ReeseÕs Miniatures	0	0	0	0	0.034
ReeseÕs Peanut Butter cup	0	0	0	0	0.720
ReeseÕs pieces	0	0	0	1	0.406
ReeseÕs 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
WelchÕs Fruit Snacks	0	0	0	1	0.313
WertherÕs 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
HersheyÕs Kisses	0.093	55.37545
HersheyÕs Krackel	0.918	62.28448
HersheyÕs Milk Chocolate	0.918	56.49050
HersheyÕs 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&MÕs	0.651	71.46505
M&MÕs	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
ReeseÕs Miniatures	0.279	81.86626
ReeseÕs Peanut Butter cup	0.651	84.18029
ReeseÕs pieces	0.651	73.43499
ReeseÕs 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
WelchÕs Fruit Snacks                       0.313   44.37552
WertherÕs Original Caramel                 0.267   41.90431
Whoppers                                   0.848   49.52411
```

```
candy["Almond Joy", ]$winpercent
```

[1] 50.34755

My favorite candy in the dataset is Almond Joy. Its winpercent value is 50.35.

Q4. What is the winpercent value for "Kit Kat"?

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

[1] 76.7686

76.77

Q5. What is the winpercent value for "Tootsie Roll Snack Bars"?

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

[1] 49.6535

49.65

```
#install.packages("skimr")
library("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	

```
#alterantive way to call functions if you only need one or two functions from a package
skimr::skim(candy)
```

Table 3: Data summary

Name	candy

11

Table 3: Data summary

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?

```
skimcandy <- skim(candy)
skimcandy$numeric.mean
```

```
 [1]  0.43529412  0.44705882  0.16470588  0.16470588  0.08235294  0.08235294
 [7]  0.17647059  0.24705882  0.51764706  0.47864705  0.46888235 50.31676381
```

```
skimcandy[12,]
```

Table 5: Data summary

Name	candy
Number of rows	85
Number of columns	12
Column type frequency:	
numeric	1
Group variables	None

Variable type: numeric

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100	hist
winpercent	0	1	50.32	14.71	22.45	39.14	47.83	59.86	84.18	

Winpercent looks to be on a different scale.

Q7. What do you think a zero and one represent for the candy$chocolate column?

Zero represents that this candy is not a chocolate candy (0=no chocolate), and one represents that it is a chocolate candy (1=yes chocolate)

Q8. Plot a histogram of winpercent values

```
hist(candy$winpercent)
```

Histogram of candy$winpercent
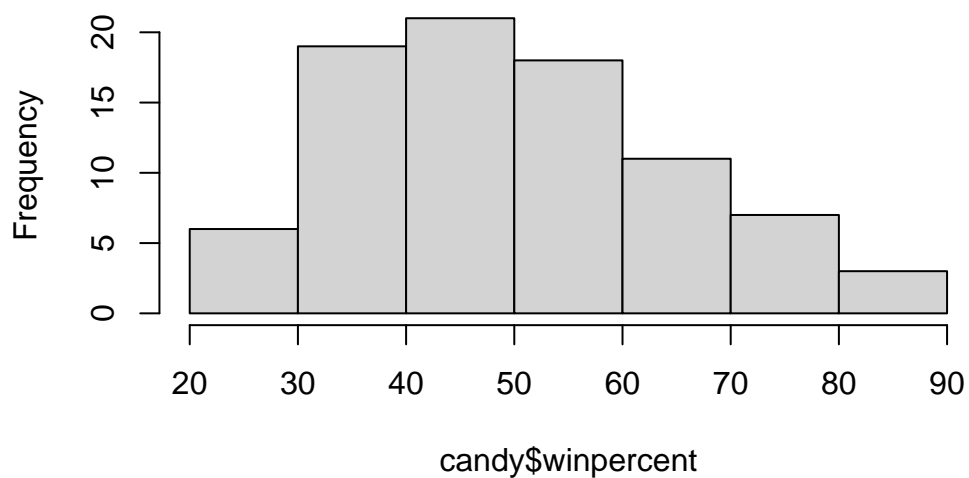

```
library(ggplot2)
ggplot(candy) + aes(winpercent) + geom_histogram(bins = 10, col="red", fill="purple")
```
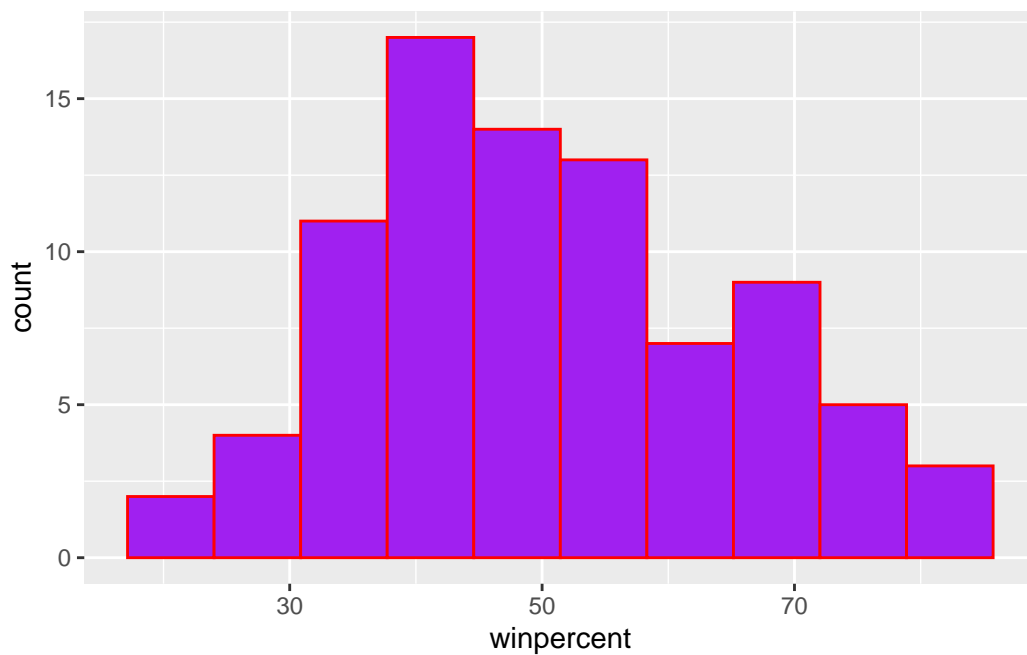
Q9. Is the distribution of winpercent values symmetrical?

No, it is slightly skewed to the left.

Q10. Is the center of the distribution above or below 50%?

Below

Q11. On average is chocolate candy higher or lower ranked than fruit candy?

```
mean(candy$winpercent[as.logical(candy$chocolate)])
```

[1] 60.92153

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

[1] 44.11974

On average, chocolate candy is higher ranked than fruit candy.

```
chocolate.inds <-as.logical(candy$chocolate)
chocolate.win <-candy[chocolate.inds,]$winpercent
mean(chocolate.win)
```

[1] 60.92153

```
fruity.inds <-as.logical(candy$fruity)
fruity.win <-candy[fruity.inds,]$winpercent
mean(fruity.win)
```

[1] 44.11974

Q12. Is this difference statistically significant?

```
t.test(candy$winpercent[as.logical(candy$chocolate)], candy$winpercent[as.logical(candy$fr
```

```
    Welch Two Sample t-test

data:  candy$winpercent[as.logical(candy$chocolate)] and candy$winpercent[as.logical(candy$fr
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
```

Yes, this difference is significantly different as indicated by the p-value < 0.05.

3. Overall Candy Rankings

Q13. What are the five least liked candy types in this set?

```
library(dplyr)
```

```
Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

    filter, lag

The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union
```

```
candy %>% arrange(winpercent) %>% head(5)
```

	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

	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

	winpercent
Nik L Nip	22.44534
Boston Baked Beans	23.41782
Chiclets	24.52499
Super Bubble	27.30386
Jawbusters	28.12744

Least liked: Nik L Nip, Boston Baked Beans, Chiclets, Super Bubble, Jawbusters

```
#alternative method
inds<- order(candy$winpercent)
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
Sugar Daddy	0	0	1	0	0
One dime	0	0	0	0	0
Sugar Babies	0	0	1	0	0
Haribo Happy Cola	0	0	0	0	0
Caramel Apple Pops	0	1	1	0	0
Strawberry bon bons	0	1	0	0	0
Sixlets	1	0	0	0	0
Ring pop	0	1	0	0	0
Chewey Lemonhead Fruit Mix	0	1	0	0	0

Red vines	0	1	0	0	0
Pixie Sticks	0	0	0	0	0
Nestle Smarties	1	0	0	0	0
Candy Corn	0	0	0	0	0
Charleston Chew	1	0	0	0	1
Warheads	0	1	0	0	0
Lemonhead	0	1	0	0	0
Fun Dip	0	1	0	0	0
Now & Later	0	1	0	0	0
Dum Dums	0	1	0	0	0
Pop Rocks	0	1	0	0	0
Laffy Taffy	0	1	0	0	0
WertherÕs Original Caramel	0	0	1	0	0
Haribo Twin Snakes	0	1	0	0	0
Dots	0	1	0	0	0
Runts	0	1	0	0	0
Tootsie Roll Juniors	1	0	0	0	0
Fruit Chews	0	1	0	0	0
WelchÕs Fruit Snacks	0	1	0	0	0
Twizzlers	0	1	0	0	0
Tootsie Roll Midgies	1	0	0	0	0
Smarties candy	0	1	0	0	0
One quarter	0	0	0	0	0
Payday	0	0	0	1	1
Mike & Ike	0	1	0	0	0
Gobstopper	0	1	0	0	0
Trolli Sour Bites	0	1	0	0	0
Mounds	1	0	0	0	0
Tootsie Pop	1	1	0	0	0
Whoppers	1	0	0	0	0
Tootsie Roll Snack Bars	1	0	0	0	0
Almond Joy	1	0	0	1	0
Haribo Sour Bears	0	1	0	0	0
Air Heads	0	1	0	0	0
Sour Patch Tricksters	0	1	0	0	0
Lifesavers big ring gummies	0	1	0	0	0
Mr Good Bar	1	0	0	1	0
Swedish Fish	0	1	0	0	0
Milk Duds	1	0	1	0	0
Skittles wildberry	0	1	0	0	0
Nerds	0	1	0	0	0
HersheyÕs Kisses	1	0	0	0	0
HersheyÕs Milk Chocolate	1	0	0	0	0

Baby Ruth	1	0	1	1	1
Haribo Gold Bears	0	1	0	0	0
Junior Mints	1	0	0	0	0
HersheyÕs Special Dark	1	0	0	0	0
Snickers Crisper	1	0	1	1	0
Sour Patch Kids	0	1	0	0	0
Milky Way Midnight	1	0	1	0	1
HersheyÕs Krackel	1	0	0	0	0
Skittles original	0	1	0	0	0
Milky Way Simply Caramel	1	0	1	0	0
Rolo	1	0	1	0	0
Nestle Crunch	1	0	0	0	0
M&MÕs	1	0	0	0	0
100 Grand	1	0	1	0	0
Starburst	0	1	0	0	0
3 Musketeers	1	0	0	0	1
Peanut M&Ms	1	0	0	1	0
Nestle Butterfinger	1	0	0	1	0
Peanut butter M&MÕs	1	0	0	1	0
ReeseÕs stuffed with pieces	1	0	0	1	0
Milky Way	1	0	1	0	1
ReeseÕs pieces	1	0	0	1	0
Snickers	1	0	1	1	1
Kit Kat	1	0	0	0	0
Twix	1	0	1	0	0
ReeseÕs Miniatures	1	0	0	1	0
ReeseÕs Peanut Butter cup	1	0	0	1	0

	crispedricewafer	hard	bar	pluribus	sugarpercent
Nik L Nip	0	0	0	1	0.197
Boston Baked Beans	0	0	0	1	0.313
Chiclets	0	0	0	1	0.046
Super Bubble	0	0	0	0	0.162
Jawbusters	0	1	0	1	0.093
Root Beer Barrels	0	1	0	1	0.732
Sugar Daddy	0	0	0	0	0.418
One dime	0	0	0	0	0.011
Sugar Babies	0	0	0	1	0.965
Haribo Happy Cola	0	0	0	1	0.465
Caramel Apple Pops	0	0	0	0	0.604
Strawberry bon bons	0	1	0	1	0.569
Sixlets	0	0	0	1	0.220
Ring pop	0	1	0	0	0.732
Chewey Lemonhead Fruit Mix	0	0	0	1	0.732

Red vines	0	0	0	1	0.581
Pixie Sticks	0	0	0	1	0.093
Nestle Smarties	0	0	0	1	0.267
Candy Corn	0	0	0	1	0.906
Charleston Chew	0	0	1	0	0.604
Warheads	0	1	0	0	0.093
Lemonhead	0	1	0	0	0.046
Fun Dip	0	1	0	0	0.732
Now & Later	0	0	0	1	0.220
Dum Dums	0	1	0	0	0.732
Pop Rocks	0	1	0	1	0.604
Laffy Taffy	0	0	0	0	0.220
WertherÕs Original Caramel	0	1	0	0	0.186
Haribo Twin Snakes	0	0	0	1	0.465
Dots	0	0	0	1	0.732
Runts	0	1	0	1	0.872
Tootsie Roll Juniors	0	0	0	0	0.313
Fruit Chews	0	0	0	1	0.127
WelchÕs Fruit Snacks	0	0	0	1	0.313
Twizzlers	0	0	0	0	0.220
Tootsie Roll Midgies	0	0	0	1	0.174
Smarties candy	0	1	0	1	0.267
One quarter	0	0	0	0	0.011
Payday	0	0	1	0	0.465
Mike & Ike	0	0	0	1	0.872
Gobstopper	0	1	0	1	0.906
Trolli Sour Bites	0	0	0	1	0.313
Mounds	0	0	1	0	0.313
Tootsie Pop	0	1	0	0	0.604
Whoppers	1	0	0	1	0.872
Tootsie Roll Snack Bars	0	0	1	0	0.465
Almond Joy	0	0	1	0	0.465
Haribo Sour Bears	0	0	0	1	0.465
Air Heads	0	0	0	0	0.906
Sour Patch Tricksters	0	0	0	1	0.069
Lifesavers big ring gummies	0	0	0	0	0.267
Mr Good Bar	0	0	1	0	0.313
Swedish Fish	0	0	0	1	0.604
Milk Duds	0	0	0	1	0.302
Skittles wildberry	0	0	0	1	0.941
Nerds	0	1	0	1	0.848
HersheyÕs Kisses	0	0	0	1	0.127
HersheyÕs Milk Chocolate	0	0	1	0	0.430

Baby Ruth	0	0	1	0	0.604
Haribo Gold Bears	0	0	0	1	0.465
Junior Mints	0	0	0	1	0.197
HersheyÕs Special Dark	0	0	1	0	0.430
Snickers Crisper	1	0	1	0	0.604
Sour Patch Kids	0	0	0	1	0.069
Milky Way Midnight	0	0	1	0	0.732
HersheyÕs Krackel	1	0	1	0	0.430
Skittles original	0	0	0	1	0.941
Milky Way Simply Caramel	0	0	1	0	0.965
Rolo	0	0	0	1	0.860
Nestle Crunch	1	0	1	0	0.313
M&MÕs	0	0	0	1	0.825
100 Grand	1	0	1	0	0.732
Starburst	0	0	0	1	0.151
3 Musketeers	0	0	1	0	0.604
Peanut M&Ms	0	0	0	1	0.593
Nestle Butterfinger	0	0	1	0	0.604
Peanut butter M&MÕs	0	0	0	1	0.825
ReeseÕs stuffed with pieces	0	0	0	0	0.988
Milky Way	0	0	1	0	0.604
ReeseÕs pieces	0	0	0	1	0.406
Snickers	0	0	1	0	0.546
Kit Kat	1	0	1	0	0.313
Twix	1	0	1	0	0.546
ReeseÕs Miniatures	0	0	0	0	0.034
ReeseÕs Peanut Butter cup	0	0	0	0	0.720

	pricepercent	winpercent
Nik L Nip	0.976	22.44534
Boston Baked Beans	0.511	23.41782
Chiclets	0.325	24.52499
Super Bubble	0.116	27.30386
Jawbusters	0.511	28.12744
Root Beer Barrels	0.069	29.70369
Sugar Daddy	0.325	32.23100
One dime	0.116	32.26109
Sugar Babies	0.767	33.43755
Haribo Happy Cola	0.465	34.15896
Caramel Apple Pops	0.325	34.51768
Strawberry bon bons	0.058	34.57899
Sixlets	0.081	34.72200
Ring pop	0.965	35.29076
Chewey Lemonhead Fruit Mix	0.511	36.01763

Red vines	0.116	37.34852
Pixie Sticks	0.023	37.72234
Nestle Smarties	0.976	37.88719
Candy Corn	0.325	38.01096
Charleston Chew	0.511	38.97504
Warheads	0.116	39.01190
Lemonhead	0.104	39.14106
Fun Dip	0.325	39.18550
Now & Later	0.325	39.44680
Dum Dums	0.034	39.46056
Pop Rocks	0.837	41.26551
Laffy Taffy	0.116	41.38956
WertherÕs Original Caramel	0.267	41.90431
Haribo Twin Snakes	0.465	42.17877
Dots	0.511	42.27208
Runts	0.279	42.84914
Tootsie Roll Juniors	0.511	43.06890
Fruit Chews	0.034	43.08892
WelchÕs Fruit Snacks	0.313	44.37552
Twizzlers	0.116	45.46628
Tootsie Roll Midgies	0.011	45.73675
Smarties candy	0.116	45.99583
One quarter	0.511	46.11650
Payday	0.767	46.29660
Mike & Ike	0.325	46.41172
Gobstopper	0.453	46.78335
Trolli Sour Bites	0.255	47.17323
Mounds	0.860	47.82975
Tootsie Pop	0.325	48.98265
Whoppers	0.848	49.52411
Tootsie Roll Snack Bars	0.325	49.65350
Almond Joy	0.767	50.34755
Haribo Sour Bears	0.465	51.41243
Air Heads	0.511	52.34146
Sour Patch Tricksters	0.116	52.82595
Lifesavers big ring gummies	0.279	52.91139
Mr Good Bar	0.918	54.52645
Swedish Fish	0.755	54.86111
Milk Duds	0.511	55.06407
Skittles wildberry	0.220	55.10370
Nerds	0.325	55.35405
HersheyÕs Kisses	0.093	55.37545
HersheyÕs Milk Chocolate	0.918	56.49050

```
Baby Ruth                       0.767    56.91455
Haribo Gold Bears               0.465    57.11974
Junior Mints                    0.511    57.21925
HersheyÕs Special Dark          0.918    59.23612
Snickers Crisper                0.651    59.52925
Sour Patch Kids                 0.116    59.86400
Milky Way Midnight              0.441    60.80070
HersheyÕs Krackel               0.918    62.28448
Skittles original               0.220    63.08514
Milky Way Simply Caramel        0.860    64.35334
Rolo                            0.860    65.71629
Nestle Crunch                   0.767    66.47068
M&MÕs                           0.651    66.57458
100 Grand                       0.860    66.97173
Starburst                       0.220    67.03763
3 Musketeers                    0.511    67.60294
Peanut M&Ms                     0.651    69.48379
Nestle Butterfinger             0.767    70.73564
Peanut butter M&MÕs             0.651    71.46505
ReeseÕs stuffed with pieces     0.651    72.88790
Milky Way                       0.651    73.09956
ReeseÕs pieces                  0.651    73.43499
Snickers                        0.651    76.67378
Kit Kat                         0.511    76.76860
Twix                            0.906    81.64291
ReeseÕs Miniatures              0.279    81.86626
ReeseÕs Peanut Butter cup       0.651    84.18029
```

```r
head(candy[inds,], n=5)
```

```
                    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
                    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
                winpercent
Nik L Nip           22.44534
Boston Baked Beans  23.41782
Chiclets            24.52499
Super Bubble        27.30386
Jawbusters          28.12744
```

Q14. What are the top 5 all time favorite candy types out of this set?

```
candy %>% arrange(desc(winpercent)) %>% head(5)
```

```
                        chocolate fruity caramel peanutyalmondy nougat
ReeseŌs Peanut Butter cup       1      0       0              1      0
ReeseŌs Miniatures              1      0       0              1      0
Twix                            1      0       1              0      0
Kit Kat                         1      0       0              0      0
Snickers                        1      0       1              1      1
                        crispedricewafer hard bar pluribus sugarpercent
ReeseŌs Peanut Butter cup               0    0   0        0        0.720
ReeseŌs Miniatures                      0    0   0        0        0.034
Twix                                    1    0   1        0        0.546
Kit Kat                                 1    0   1        0        0.313
Snickers                                0    0   1        0        0.546
                        pricepercent winpercent
ReeseŌs Peanut Butter cup      0.651   84.18029
ReeseŌs Miniatures             0.279   81.86626
Twix                           0.906   81.64291
Kit Kat                        0.511   76.76860
Snickers                       0.651   76.67378
```
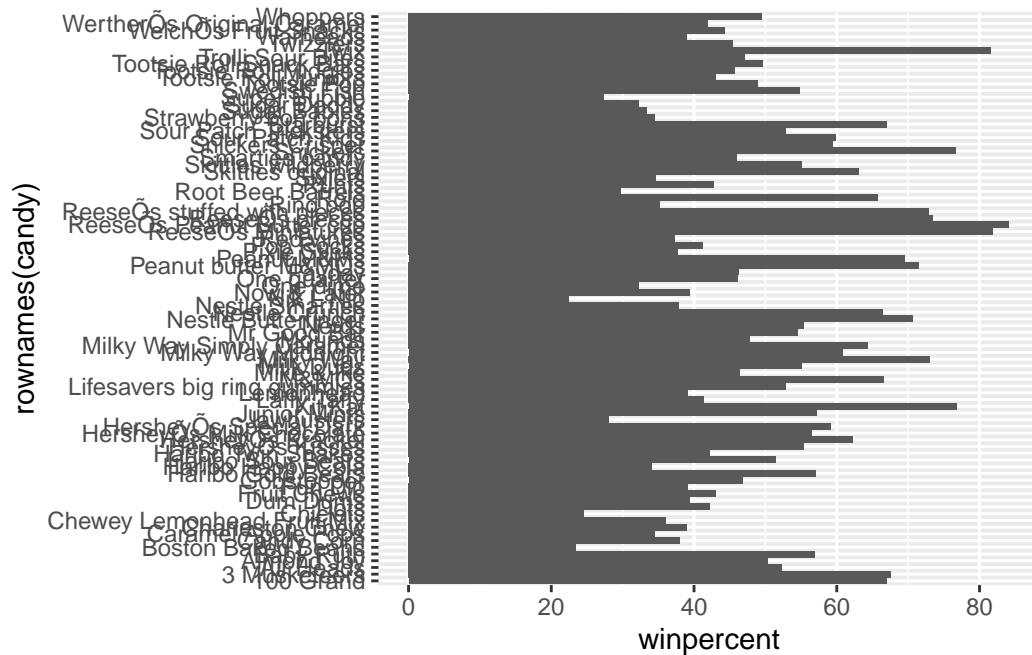
Favorite candies: Reese's Peanut Butter Cup, Reese's Miniatures, Twix, Kit Kat, Snickers
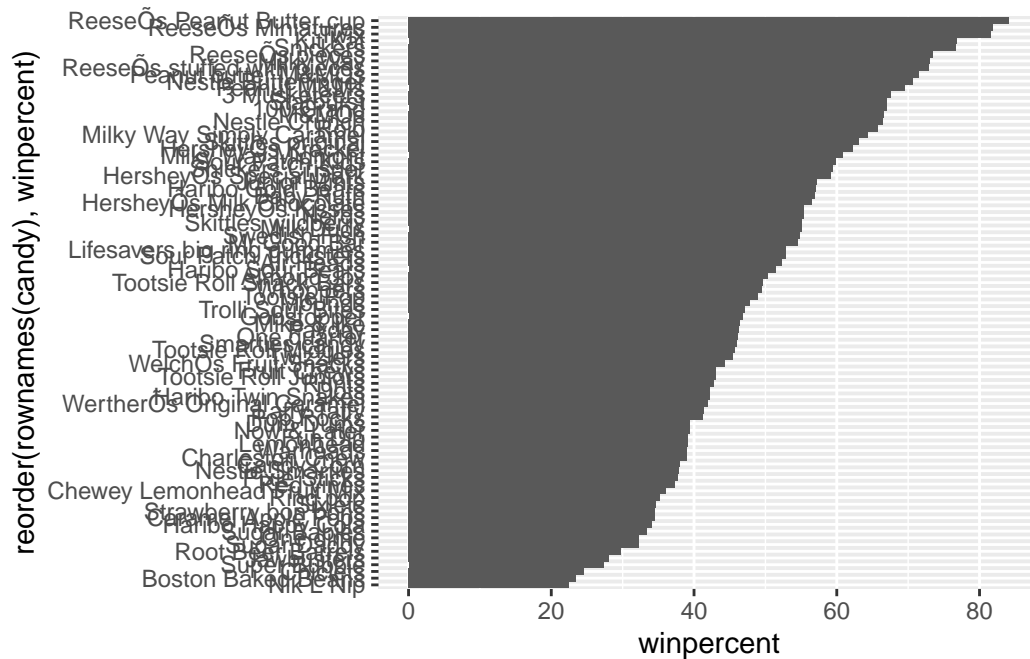
Q15. Make a first barplot of candy ranking based on winpercent values.

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

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

```
ggplot(candy) + aes(winpercent, reorder(rownames(candy), winpercent)) + geom_col()
```

```
#this function will save the last plot that was made
ggsave("mybarplot.png")
```

Saving 5.5 x 3.5 in image

```
#can also tell it what plot to save, how tall, the resolution, etc. (see help page for ggs
```

```
#make all candies black as default. Then overwrite with different colors based on certain
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"
```

```
ggplot(candy) +
  aes(winpercent, reorder(rownames(candy),winpercent)) +
  geom_col(fill=my_cols)
```

26
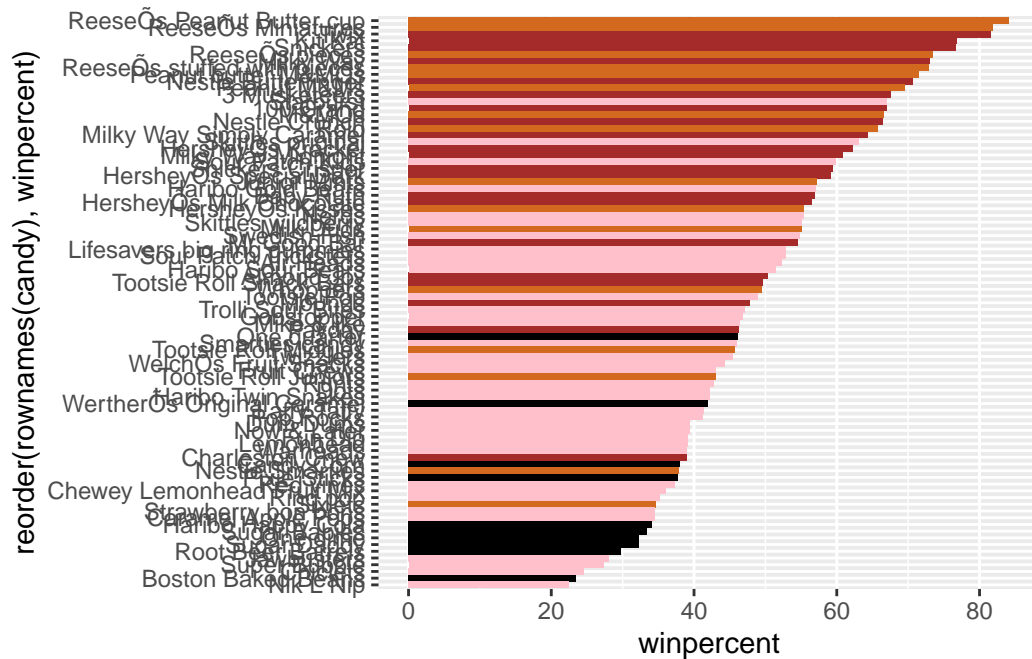
Q17. What is the worst ranked chocolate candy?

```
candy %>% filter(chocolate==1) %>% arrange(winpercent)
```

	chocolate	fruity	caramel	peanutyalmondy	nougat
Sixlets	1	0	0	0	0
Nestle Smarties	1	0	0	0	0
Charleston Chew	1	0	0	0	1
Tootsie Roll Juniors	1	0	0	0	0
Tootsie Roll Midgies	1	0	0	0	0
Mounds	1	0	0	0	0
Tootsie Pop	1	1	0	0	0
Whoppers	1	0	0	0	0
Tootsie Roll Snack Bars	1	0	0	0	0
Almond Joy	1	0	0	1	0
Mr Good Bar	1	0	0	1	0
Milk Duds	1	0	1	0	0
HersheyÕs Kisses	1	0	0	0	0
HersheyÕs Milk Chocolate	1	0	0	0	0
Baby Ruth	1	0	1	1	1
Junior Mints	1	0	0	0	0

HersheyÕs Special Dark	1	0	0	0	0
Snickers Crisper	1	0	1	1	0
Milky Way Midnight	1	0	1	0	1
HersheyÕs Krackel	1	0	0	0	0
Milky Way Simply Caramel	1	0	1	0	0
Rolo	1	0	1	0	0
Nestle Crunch	1	0	0	0	0
M&MÕs	1	0	0	0	0
100 Grand	1	0	1	0	0
3 Musketeers	1	0	0	0	1
Peanut M&Ms	1	0	0	1	0
Nestle Butterfinger	1	0	0	1	0
Peanut butter M&MÕs	1	0	0	1	0
ReeseÕs stuffed with pieces	1	0	0	1	0
Milky Way	1	0	1	0	1
ReeseÕs pieces	1	0	0	1	0
Snickers	1	0	1	1	1
Kit Kat	1	0	0	0	0
Twix	1	0	1	0	0
ReeseÕs Miniatures	1	0	0	1	0
ReeseÕs Peanut Butter cup	1	0	0	1	0

	crispedricewafer	hard	bar	pluribus	sugarpercent
Sixlets	0	0	0	1	0.220
Nestle Smarties	0	0	0	1	0.267
Charleston Chew	0	0	1	0	0.604
Tootsie Roll Juniors	0	0	0	0	0.313
Tootsie Roll Midgies	0	0	0	1	0.174
Mounds	0	0	1	0	0.313
Tootsie Pop	0	1	0	0	0.604
Whoppers	1	0	0	1	0.872
Tootsie Roll Snack Bars	0	0	1	0	0.465
Almond Joy	0	0	1	0	0.465
Mr Good Bar	0	0	1	0	0.313
Milk Duds	0	0	0	1	0.302
HersheyÕs Kisses	0	0	0	1	0.127
HersheyÕs Milk Chocolate	0	0	1	0	0.430
Baby Ruth	0	0	1	0	0.604
Junior Mints	0	0	0	1	0.197
HersheyÕs Special Dark	0	0	1	0	0.430
Snickers Crisper	1	0	1	0	0.604
Milky Way Midnight	0	0	1	0	0.732
HersheyÕs Krackel	1	0	1	0	0.430
Milky Way Simply Caramel	0	0	1	0	0.965

Rolo	0	0	0	1	0.860
Nestle Crunch	1	0	1	0	0.313
M&MÕs	0	0	0	1	0.825
100 Grand	1	0	1	0	0.732
3 Musketeers	0	0	1	0	0.604
Peanut M&Ms	0	0	0	1	0.593
Nestle Butterfinger	0	0	1	0	0.604
Peanut butter M&MÕs	0	0	0	1	0.825
ReeseÕs stuffed with pieces	0	0	0	0	0.988
Milky Way	0	0	1	0	0.604
ReeseÕs pieces	0	0	0	1	0.406
Snickers	0	0	1	0	0.546
Kit Kat	1	0	1	0	0.313
Twix	1	0	1	0	0.546
ReeseÕs Miniatures	0	0	0	0	0.034
ReeseÕs Peanut Butter cup	0	0	0	0	0.720

	pricepercent	winpercent
Sixlets	0.081	34.72200
Nestle Smarties	0.976	37.88719
Charleston Chew	0.511	38.97504
Tootsie Roll Juniors	0.511	43.06890
Tootsie Roll Midgies	0.011	45.73675
Mounds	0.860	47.82975
Tootsie Pop	0.325	48.98265
Whoppers	0.848	49.52411
Tootsie Roll Snack Bars	0.325	49.65350
Almond Joy	0.767	50.34755
Mr Good Bar	0.918	54.52645
Milk Duds	0.511	55.06407
HersheyÕs Kisses	0.093	55.37545
HersheyÕs Milk Chocolate	0.918	56.49050
Baby Ruth	0.767	56.91455
Junior Mints	0.511	57.21925
HersheyÕs Special Dark	0.918	59.23612
Snickers Crisper	0.651	59.52925
Milky Way Midnight	0.441	60.80070
HersheyÕs Krackel	0.918	62.28448
Milky Way Simply Caramel	0.860	64.35334
Rolo	0.860	65.71629
Nestle Crunch	0.767	66.47068
M&MÕs	0.651	66.57458
100 Grand	0.860	66.97173
3 Musketeers	0.511	67.60294

```
Peanut M&Ms                      0.651   69.48379
Nestle Butterfinger              0.767   70.73564
Peanut butter M&MÕs              0.651   71.46505
ReeseÕs stuffed with pieces      0.651   72.88790
Milky Way                        0.651   73.09956
ReeseÕs pieces                   0.651   73.43499
Snickers                         0.651   76.67378
Kit Kat                          0.511   76.76860
Twix                             0.906   81.64291
ReeseÕs Miniatures               0.279   81.86626
ReeseÕs Peanut Butter cup        0.651   84.18029
```

Sixlets

Q18. What is the best ranked fruity candy?

```r
candy %>% filter(fruity==1) %>% arrange(desc(winpercent))
```

	chocolate	fruity	caramel	peanutyalmondy	nougat
Starburst	0	1	0	0	0
Skittles original	0	1	0	0	0
Sour Patch Kids	0	1	0	0	0
Haribo Gold Bears	0	1	0	0	0
Nerds	0	1	0	0	0
Skittles wildberry	0	1	0	0	0
Swedish Fish	0	1	0	0	0
Lifesavers big ring gummies	0	1	0	0	0
Sour Patch Tricksters	0	1	0	0	0
Air Heads	0	1	0	0	0
Haribo Sour Bears	0	1	0	0	0
Tootsie Pop	1	1	0	0	0
Trolli Sour Bites	0	1	0	0	0
Gobstopper	0	1	0	0	0
Mike & Ike	0	1	0	0	0
Smarties candy	0	1	0	0	0
Twizzlers	0	1	0	0	0
WelchÕs Fruit Snacks	0	1	0	0	0
Fruit Chews	0	1	0	0	0
Runts	0	1	0	0	0
Dots	0	1	0	0	0
Haribo Twin Snakes	0	1	0	0	0

Laffy Taffy	0	1	0	0	0
Pop Rocks	0	1	0	0	0
Dum Dums	0	1	0	0	0
Now & Later	0	1	0	0	0
Fun Dip	0	1	0	0	0
Lemonhead	0	1	0	0	0
Warheads	0	1	0	0	0
Red vines	0	1	0	0	0
Chewey Lemonhead Fruit Mix	0	1	0	0	0
Ring pop	0	1	0	0	0
Strawberry bon bons	0	1	0	0	0
Caramel Apple Pops	0	1	1	0	0
Jawbusters	0	1	0	0	0
Super Bubble	0	1	0	0	0
Chiclets	0	1	0	0	0
Nik L Nip	0	1	0	0	0

	crispedricewafer	hard	bar	pluribus	sugarpercent
Starburst	0	0	0	1	0.151
Skittles original	0	0	0	1	0.941
Sour Patch Kids	0	0	0	1	0.069
Haribo Gold Bears	0	0	0	1	0.465
Nerds	0	1	0	1	0.848
Skittles wildberry	0	0	0	1	0.941
Swedish Fish	0	0	0	1	0.604
Lifesavers big ring gummies	0	0	0	0	0.267
Sour Patch Tricksters	0	0	0	1	0.069
Air Heads	0	0	0	0	0.906
Haribo Sour Bears	0	0	0	1	0.465
Tootsie Pop	0	1	0	0	0.604
Trolli Sour Bites	0	0	0	1	0.313
Gobstopper	0	1	0	1	0.906
Mike & Ike	0	0	0	1	0.872
Smarties candy	0	1	0	1	0.267
Twizzlers	0	0	0	0	0.220
WelchÕs Fruit Snacks	0	0	0	1	0.313
Fruit Chews	0	0	0	1	0.127
Runts	0	1	0	1	0.872
Dots	0	0	0	1	0.732
Haribo Twin Snakes	0	0	0	1	0.465
Laffy Taffy	0	0	0	0	0.220
Pop Rocks	0	1	0	1	0.604
Dum Dums	0	1	0	0	0.732
Now & Later	0	0	0	1	0.220

Fun Dip	0	1	0	0	0.732
Lemonhead	0	1	0	0	0.046
Warheads	0	1	0	0	0.093
Red vines	0	0	0	1	0.581
Chewey Lemonhead Fruit Mix	0	0	0	1	0.732
Ring pop	0	1	0	0	0.732
Strawberry bon bons	0	1	0	1	0.569
Caramel Apple Pops	0	0	0	0	0.604
Jawbusters	0	1	0	1	0.093
Super Bubble	0	0	0	0	0.162
Chiclets	0	0	0	1	0.046
Nik L Nip	0	0	0	1	0.197

	pricepercent	winpercent
Starburst	0.220	67.03763
Skittles original	0.220	63.08514
Sour Patch Kids	0.116	59.86400
Haribo Gold Bears	0.465	57.11974
Nerds	0.325	55.35405
Skittles wildberry	0.220	55.10370
Swedish Fish	0.755	54.86111
Lifesavers big ring gummies	0.279	52.91139
Sour Patch Tricksters	0.116	52.82595
Air Heads	0.511	52.34146
Haribo Sour Bears	0.465	51.41243
Tootsie Pop	0.325	48.98265
Trolli Sour Bites	0.255	47.17323
Gobstopper	0.453	46.78335
Mike & Ike	0.325	46.41172
Smarties candy	0.116	45.99583
Twizzlers	0.116	45.46628
WelchÕs Fruit Snacks	0.313	44.37552
Fruit Chews	0.034	43.08892
Runts	0.279	42.84914
Dots	0.511	42.27208
Haribo Twin Snakes	0.465	42.17877
Laffy Taffy	0.116	41.38956
Pop Rocks	0.837	41.26551
Dum Dums	0.034	39.46056
Now & Later	0.325	39.44680
Fun Dip	0.325	39.18550
Lemonhead	0.104	39.14106
Warheads	0.116	39.01190
Red vines	0.116	37.34852

```
Chewey Lemonhead Fruit Mix          0.511    36.01763
Ring pop                            0.965    35.29076
Strawberry bon bons                 0.058    34.57899
Caramel Apple Pops                  0.325    34.51768
Jawbusters                          0.511    28.12744
Super Bubble                        0.116    27.30386
Chiclets                            0.325    24.52499
Nik L Nip                           0.976    22.44534
```
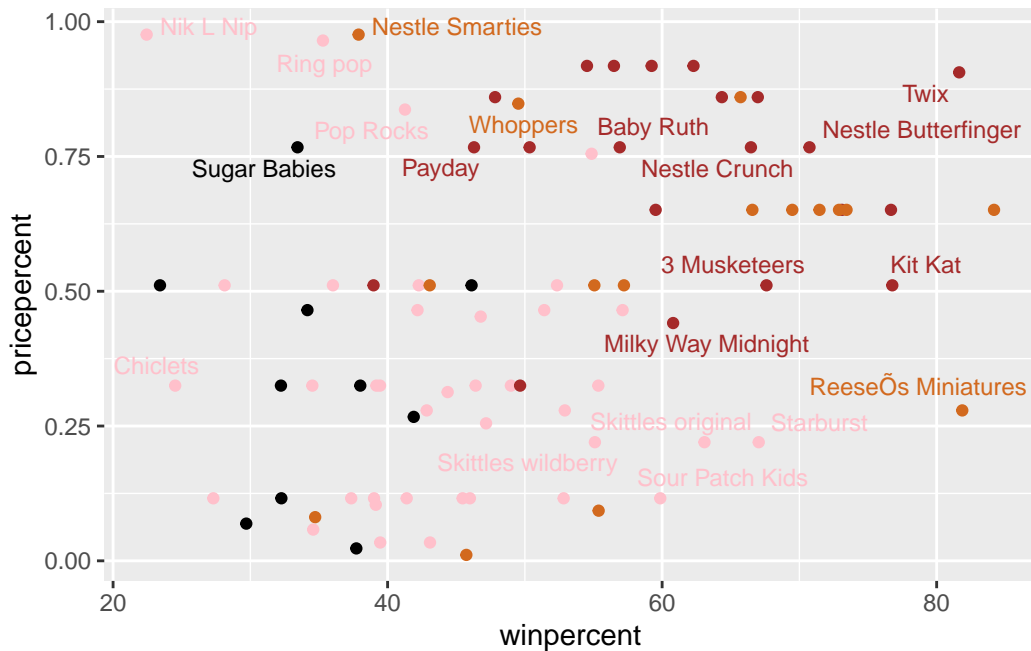
Starburst

4. Taking a look at pricepercent

```
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
```

Q19. Which candy type is the highest ranked in terms of winpercent for the least money - i.e. offers the most bang for your buck?

Fruity candies are generally the best bang for your buck. Reese's miniatures is the best bang for your buck more specifically.

Q20. What are the top 5 most expensive candy types in the dataset and of these which is the least popular?

```r
exp <- candy %>% arrange(desc(pricepercent)) %>% head(5)
exp %>% arrange(winpercent)
```

	chocolate	fruity	caramel	peanutyalmondy	nougat
Nik L Nip	0	1	0	0	0
Ring pop	0	1	0	0	0
Nestle Smarties	1	0	0	0	0
HersheyÕs Milk Chocolate	1	0	0	0	0
HersheyÕs Krackel	1	0	0	0	0
	crispedricewafer	hard	bar	pluribus	sugarpercent
Nik L Nip	0	0	0	1	0.197
Ring pop	0	1	0	0	0.732

```
Nestle Smarties                          0    0    0         1        0.267
HersheyÕs Milk Chocolate                 0    0    1         0        0.430
HersheyÕs Krackel                        1    0    1         0        0.430
                     pricepercent winpercent
Nik L Nip                   0.976    22.44534
Ring pop                    0.965    35.29076
Nestle Smarties             0.976    37.88719
HersheyÕs Milk Chocolate    0.918    56.49050
HersheyÕs Krackel           0.918    62.28448
```

Most expensive: Nik L Nip, Ring pop, Nestle Smarties, Hershey's Milk, Hershey's Krackel

Of these, Nik L Nip is the least popular.

```
#alternative method
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
HersheyÕs Krackel           0.918    62.28448
HersheyÕs Milk Chocolate    0.918    56.49050
```

Q21 –> Optional

5 Exploring the correlation structure

```
#install.packages("corrplot")
library("corrplot")
```

corrplot 0.92 loaded

```
cij <- cor(candy)
corrplot(cij)
```
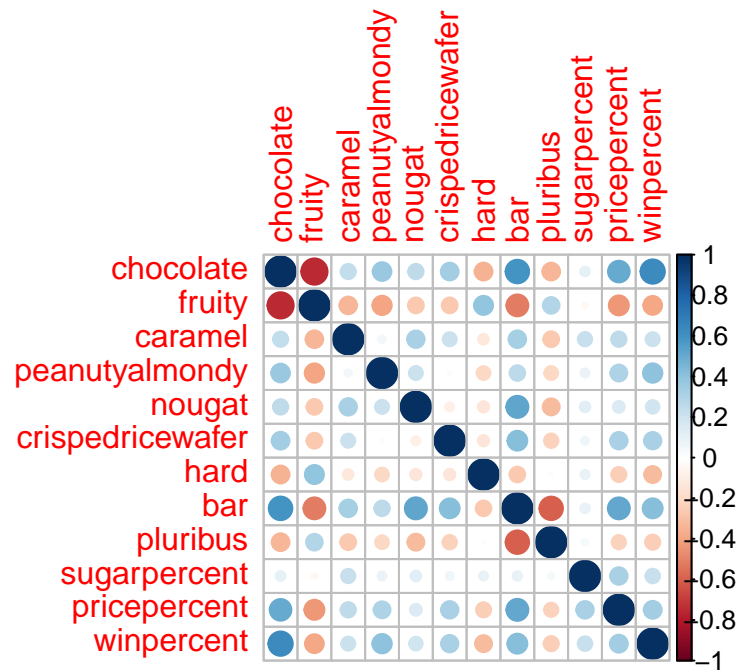
Q22. Examining this plot what two variables are anti-correlated (i.e. have minus values)?

Fruity and chocolate

Q23. Similarly, what two variables are most positively correlated?

Winpercent and chocolate

6. Principal Component Analysis

Need to scale our PCA plot, otherwise winpercent will dominate everything.

The main function that is always there for us is "prcomp()". It has an important argument that is set to "scale=FALSE".

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

```
Importance of components:
                  PC1     PC2     PC3     PC4     PC5     PC6     PC7
```

36

```
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
```

My PCA plot (aka PC1 vs PC2 score plot)

```
plot(pca$x[,1], pca$x[,2], col=my_cols, pch=16)
```
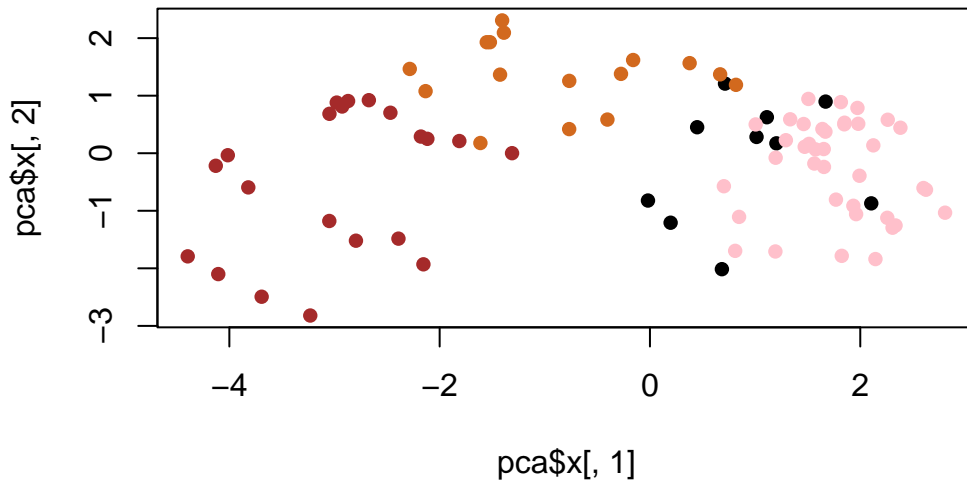


```
# Make a new data-frame with our PCA results and candy data
my_data <- cbind(candy, pca$x[,1:3])
head(my_data)
```

```
              chocolate fruity caramel peanutyalmondy nougat crispedricewafer
100 Grand             1      0       1              0      0                1
3 Musketeers          1      0       0              0      1                0
One dime              0      0       0              0      0                0
One quarter           0      0       0              0      0                0
```

```
Air Heads              0      1        0              0        0                0
Almond Joy             1      0        0              1        0                0
          hard bar pluribus sugarpercent pricepercent winpercent        PC1
100 Grand     0   1        0        0.732        0.860   66.97173 -3.8198617
3 Musketeers  0   1        0        0.604        0.511   67.60294 -2.7960236
One dime      0   0        0        0.011        0.116   32.26109  1.2025836
One quarter   0   0        0        0.011        0.511   46.11650  0.4486538
Air Heads     0   0        0        0.906        0.511   52.34146  0.7028992
Almond Joy    0   1        0        0.465        0.767   50.34755 -2.4683383
                  PC2         PC3
100 Grand    -0.5935788  2.1863087
3 Musketeers -1.5196062 -1.4121986
One dime      0.1718121 -2.0607712
One quarter   0.4519736 -1.4764928
Air Heads    -0.5731343  0.9293893
Almond Joy    0.7035501 -0.8581089
```

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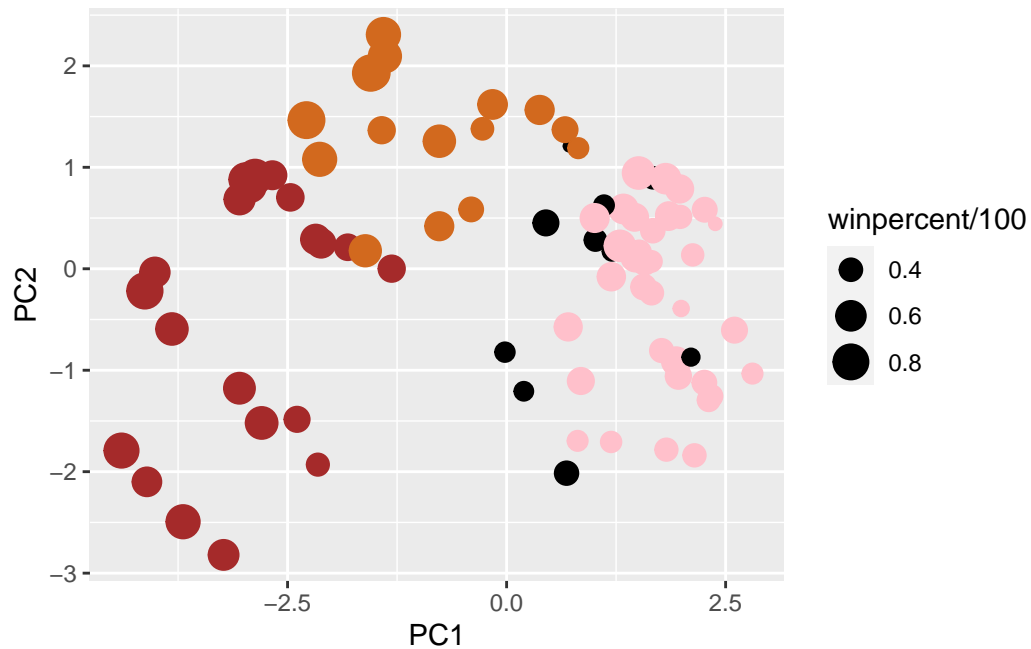
```
p + geom_text_repel(size=3.3, col=my_cols, max.overlaps = 7)  +
  theme(legend.position = "none") +
  labs(title="Halloween Candy PCA Space",
       subtitle="Colored by type: chocolate bar (dark brown), chocolate other (light brown
       caption="Data from 538")
```

Warning: ggrepel: 60 unlabeled data points (too many overlaps). Consider
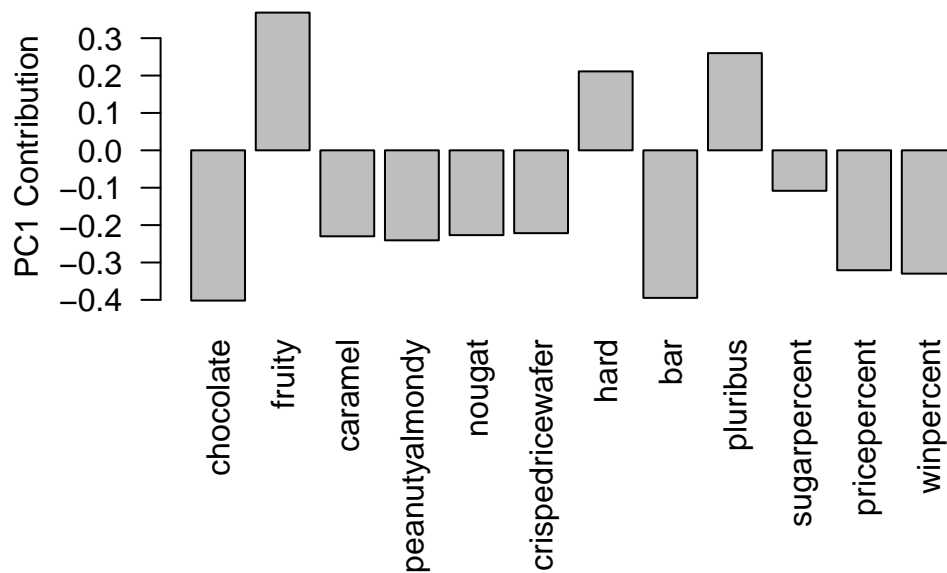increasing max.overlaps

Halloween Candy PCA Space
Colored by type: chocolate bar (dark brown), chocolate other (light brown),
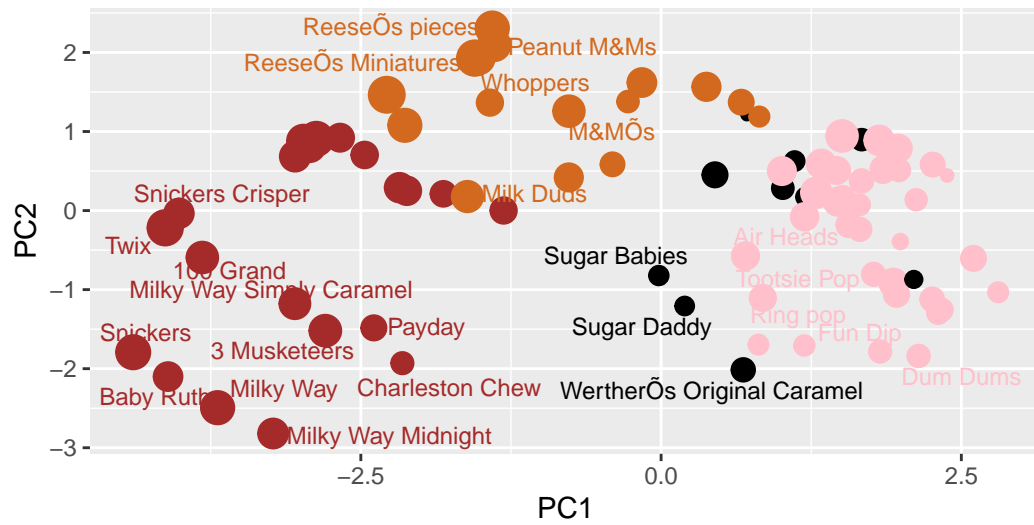
Data from 538

```
par(mar=c(8,4,2,2))
barplot(pca$rotation[,1], las=2, ylab="PC1 Contribution")
```

Q24. What original variables are picked up strongly by PC1 in the positive direction? Do these make sense to you?

Fruity, hard, and pluribus. These make sense as fruity candies often come in packs of multiple candies and are often hard.