Lab 9 - Halloween Candy Mini-Project

Shreyas Sankaranarayanan

Project Context

In this report we will be conducting and documenting our analysis on the candy-data csv provided by the new organization FiveThirtyEight. We will be trying to determine what exactly makes a candy popular.

In order to do so we must first import the data:

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

Data Analysis

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

1

```
diffcandy <- nrow(candy.data)
diffcandy
```

[1] 85

Q2. How many fruity candy types are in the data set?

```
sum(candy.data$fruity)
```

[1] 38

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

```
candy.data["Welch's Fruit Snacks",]$winpercent
```

[1] 44.37552

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

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

[1] 76.7686

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

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

[1] 49.6535

Sidenote: Attempting to use skimr on the csv

```
library(skimr)
skimcandy <- skim(candy.data)
skimcandy
```

Table 1: Data summary

Name	candy.data
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	

```
min(candy.data$winpercent)
```

```
[1] 22.44534
```

Q6. Is there any variable/column that looks to be on a different scale to the majority of the other columns in the dataset?

When analyzing the skimcandy dataset which utilizes the **skimr** package to create a comprehensive view of the csv. This value seems to be the **winpercent** variable/column.

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

A zero represents that there is no chocolate in the corresponding candy while a one indicates there is

Q8. Plot a histogram of winpercent values

```
library(ggplot2)
ggplot(candy.data, aes(x = candy.data[,1], candy.data$winpercent))
```

```
Warning: Use of `candy.data$winpercent` is discouraged.
i Use `winpercent` instead.
```
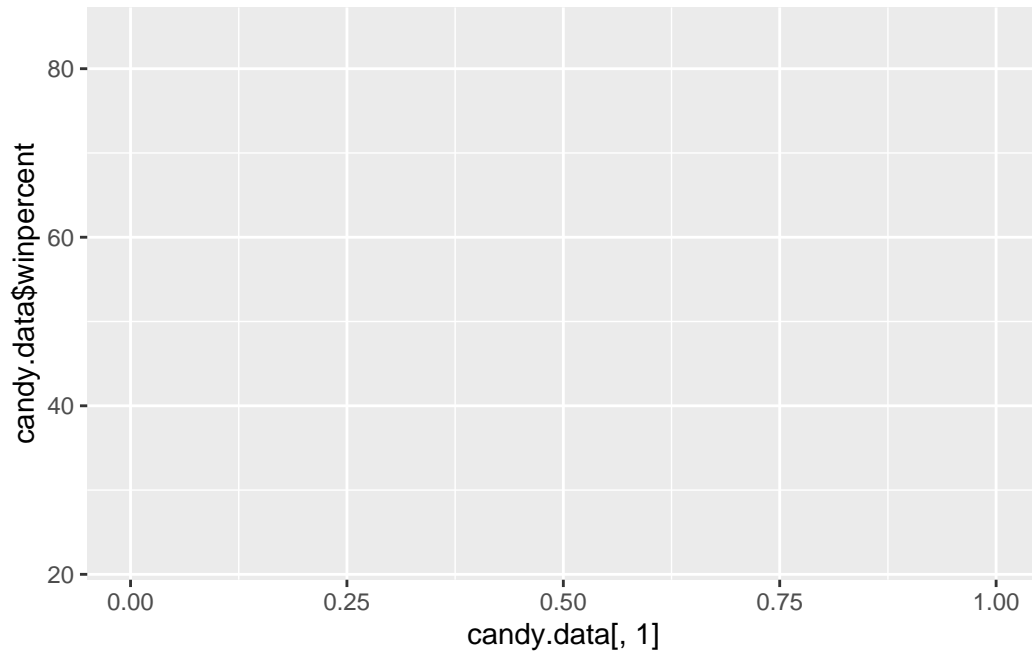


```
hist(candy.data$winpercent)
hist(candy.data$winpercent)
```

Histogram of candy.data$winpercent

candy.data$winpercent
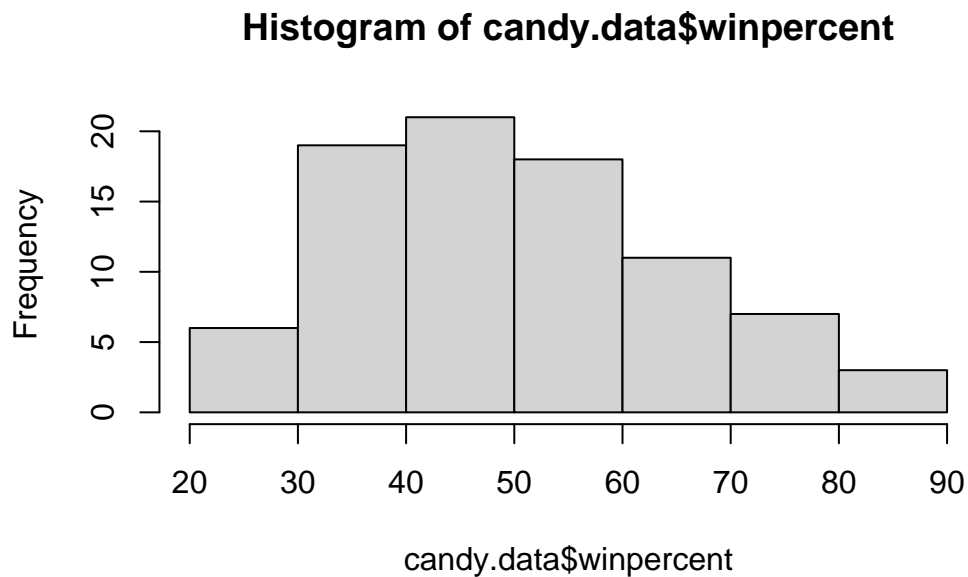
Q9. Is the distribution of winpercent values symmetrical?

No, it is not as seen above

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

Below 50%

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

```
avgchocrank <- candy.data$winpercent[which(candy.data$chocolate == 1)]
avgfruitrank <- candy.data$winpercent[which(candy.data$fruity == 1)]
mean(avgchocrank)
```

[1] 60.92153

```
mean(avgfruitrank)
```

[1] 44.11974

Q12. Is this difference statistically singificant?

```
t.test(avgfruitrank,avgchocrank)
```

```
	Welch Two Sample t-test

data:  avgfruitrank and avgchocrank
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:
 -22.15795 -11.44563
sample estimates:
mean of x mean of y
 44.11974  60.92153
```

The p-value is significantly lower than 0.05 showing that the difference is statistically significant.

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

```
head(candy.data[order(candy.data$winpercent, decreasing = FALSE),], 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?

```r
head(candy.data[order(candy.data$winpercent, decreasing = TRUE),], n=5)
```

	chocolate	fruity	caramel	peanutyalmondy	nougat
Reese's Peanut Butter cup	1	0	0	1	0
Reese's 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
Reese's Peanut Butter cup	0	0	0	0	0.720
Reese's 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
Reese's Peanut Butter cup	0.651	84.18029
Reese's Miniatures	0.279	81.86626
Twix	0.906	81.64291
Kit Kat	0.511	76.76860
Snickers	0.651	76.67378

Q15. Make a first barplot of candy ranking based on winpercent values.

```r
ggplot(candy.data) +
  aes(candy.data$winpercent, rownames(candy.data)) +
  geom_col()
```
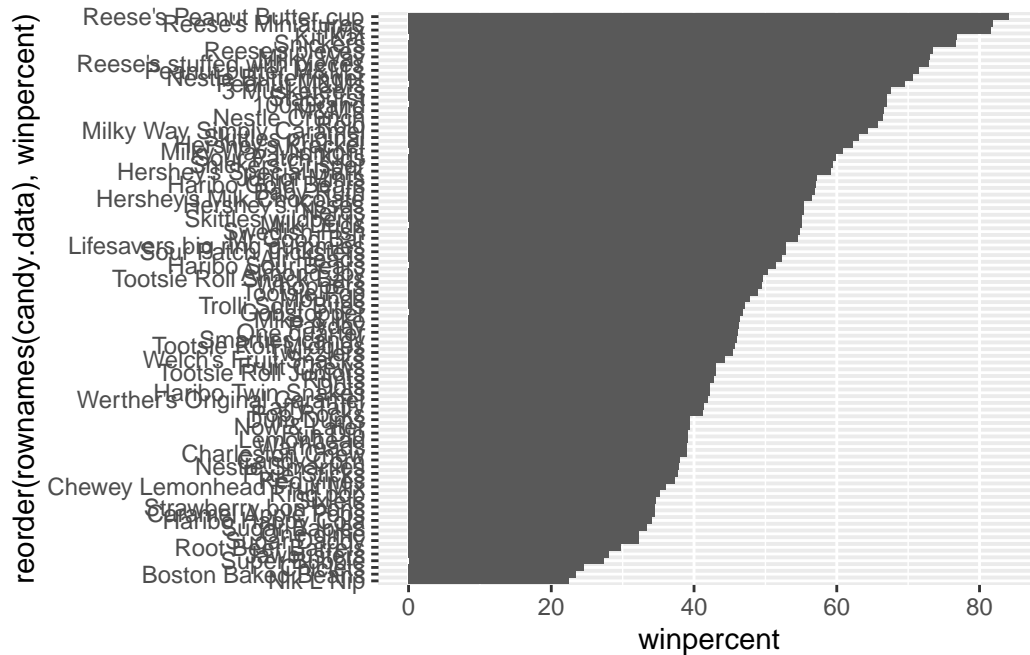
```
Warning: Use of `candy.data$winpercent` is discouraged.
i Use `winpercent` instead.
```
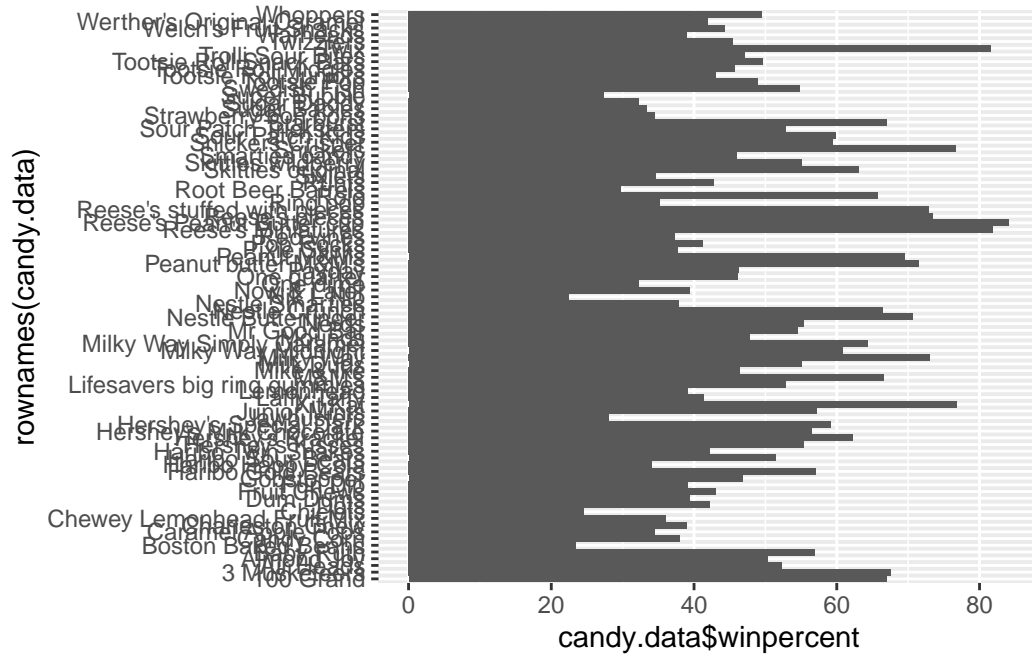
Q16. This is quite ugly, use the reorder() function to get the bars sorted by winpercent?

```
ggplot(candy.data) +
  aes(winpercent, reorder(rownames(candy.data), winpercent)) +
  geom_col()
```
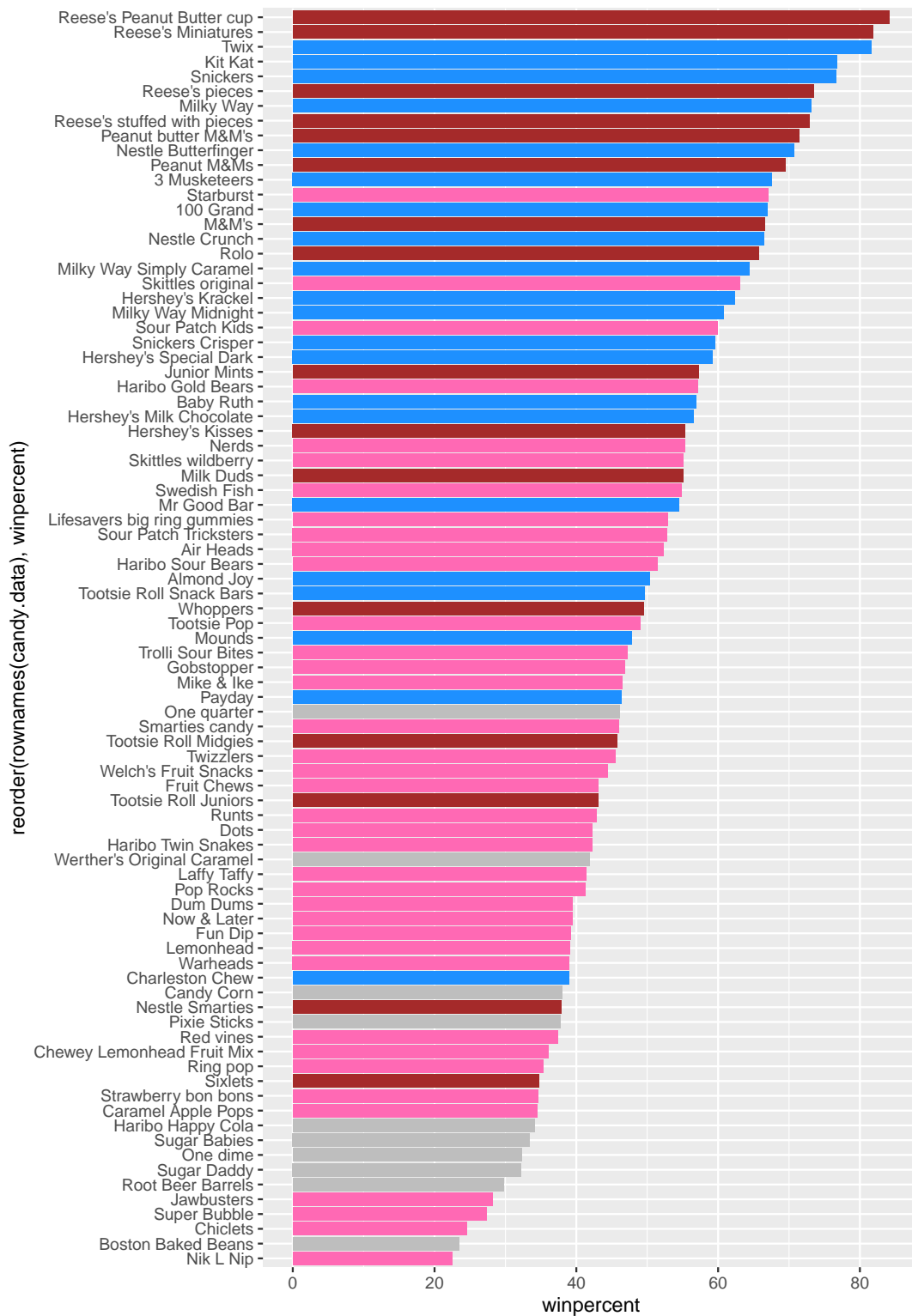
```
ggsave("mybarplot.png", width = 10, height = 10)
```

Let's make it cleaner

```
my_cols=rep("grey", nrow(candy.data))
my_cols[as.logical(candy.data$chocolate)] = "brown"
my_cols[as.logical(candy.data$bar)] = "dodgerblue"
my_cols[as.logical(candy.data$fruity)] = "hotpink"
ggplot(candy.data) +
  aes(winpercent, reorder(rownames(candy.data), winpercent)) +
  geom_col(fill = my_cols)
```

Q17. What is the worst ranked chocolate candy?

According to the plot above and the code below, Sixlets are the worst ranked chocolate candy

```
wccand <- subset(candy.data, candy.data$chocolate == 1)
sortwccand <- wccand[order(wccand$winpercent, decreasing = FALSE),]

head(sortwccand, n=1)
```

```
        chocolate fruity caramel peanutyalmondy nougat crispedricewafer hard
Sixlets         1      0       0             0      0                 0    0
        bar pluribus sugarpercent pricepercent winpercent
Sixlets   0        1         0.22        0.081     34.722
```

Q18. What is the best ranked fruity candy?

According to the plot above and the code below, Starburst are the best ranked fruity candy

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?

Reese's Miniatures is highest ranked for winpercent but also is one of the cheapest.

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

NiklNip as it is the least like but the most expensive

```
ord <- order(candy.data$pricepercent, decreasing = TRUE)
head( candy.data[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
```

```
wccand2 <- subset(candy.data, candy.data$fruity == 1)
sortwccand <- wccand2[order(wccand2$winpercent, decreasing = TRUE),]

head(sortwccand, n=1)
```

```
          chocolate fruity caramel peanutyalmondy nougat crispedricewafer hard
Starburst         0      1       0              0      0                0    0
         bar pluribus sugarpercent pricepercent winpercent
Starburst   0        1        0.151         0.22   67.03763
```

```r
library(ggrepel)

# How about a plot of price vs win
ggplot(candy.data) +
  aes(winpercent, pricepercent, label=rownames(candy.data)) +
  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
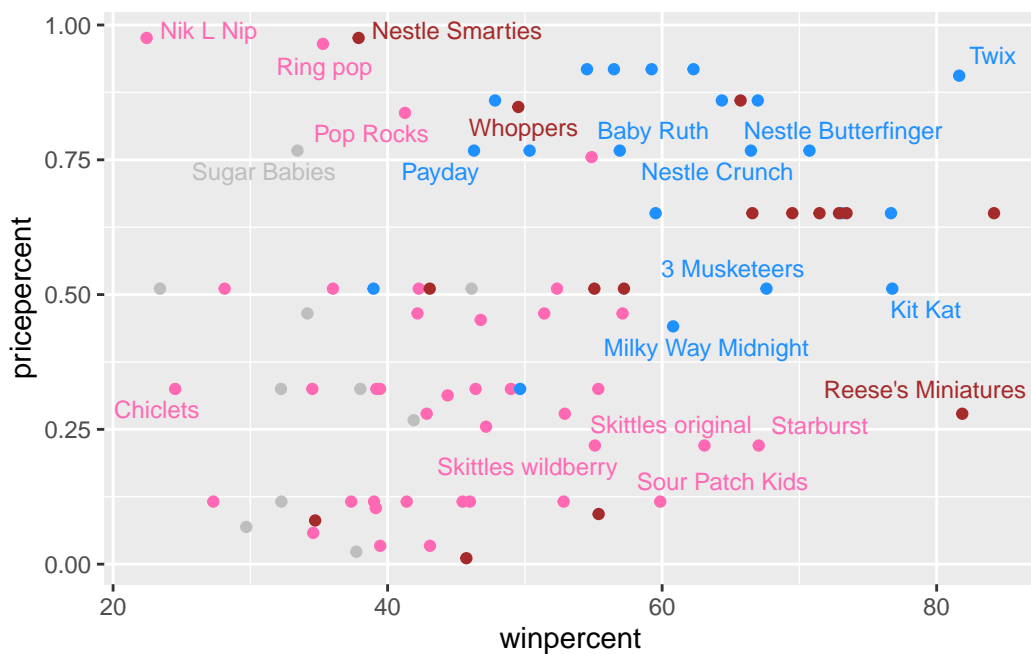
Analyszing correlation structure

```
library(corrplot)
```

```
corrplot 0.92 loaded
```

```
## corrplot 0.90 loaded
cij <- cor(candy.data
corrplot(cij)
```
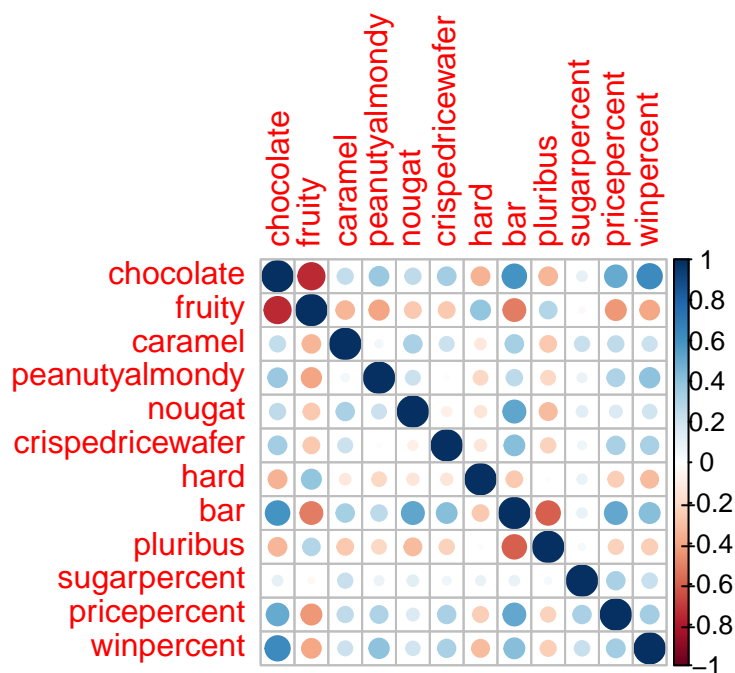

Q22. Examining this plot what two variables are anti-correlated (i.e. have minus values)?

Chocolate and fruity are most anticorrelated > Q23. Similarly, what two variables are most positively correlated?

Chocolate and winpercent are most positively correlated

13

6. Conducting Principle Component Analysis

We will perform a PCA of the candy. Key-question: Do we need to scale for PCA? Answer: Yes, as we have values that are rated on diff scales.

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

```
pca$x
```

```
                                   PC1            PC2          PC3          PC4
100 Grand                  -3.81986175 -0.5935787670 -2.186308676 -2.37159574
3 Musketeers               -2.79602364 -1.5196062111  1.412198551  0.69943868
One dime                    1.20258363  0.1718120657  2.060771178 -1.20067824
One quarter                 0.44865378  0.4519735621  1.476492844 -1.00177141
Air Heads                   0.70289922 -0.5731343263 -0.929389343  0.41245656
Almond Joy                 -2.46833834  0.7035501120  0.858108916  0.57249739
Baby Ruth                  -4.10531223 -2.1000967736  1.347834706  1.85505225
Boston Baked Beans          0.71385813  1.2098216537  0.941899950  1.10219913
Candy Corn                  1.01357204  0.2834319621 -0.840681586  0.83664703
Caramel Apple Pops          0.81049645 -1.6960889498 -0.207020586 -0.30186567
Charleston Chew            -2.15436587 -1.9304213037  1.675469334  0.46999498
Chewey Lemonhead Fruit Mix  1.65268482  0.0726434944 -0.909617411  0.58609915
Chiclets                    2.38180817  0.4430926071  1.000422079 -0.59998577
Dots                        1.51249936  0.1623958592 -0.967135199  0.63622661
Dum Dums                    2.14430933 -1.8388386160 -0.385372660 -0.14796280
Fruit Chews                 2.26133763  0.5818322520  0.978626618 -0.39164187
Fun Dip                     1.82383348 -1.7828662094 -0.719415821 -0.08544003
Gobstopper                  1.96047812 -1.0584680267 -1.873874385  0.84237208
Haribo Gold Bears           1.33360746  0.5892699921 -0.431929774  0.33530766
```

```
Haribo Happy Cola              1.11167365   0.6257697808   0.054459647   0.16024129
Haribo Sour Bears              1.46152952   0.5073691482  -0.379443632   0.28956535
Haribo Twin Snakes             1.66849016   0.3748646265  -0.294528131   0.21556045
Hershey's Kisses               0.37722675   1.5654519145   1.104739528  -0.30451907
Hershey's Krackel             -3.04788356   0.6850792787  -1.154357778  -2.79294516
Hershey's Milk Chocolate      -2.11696417   0.2504568891   0.218316614  -0.64942872
Hershey's Special Dark        -2.17850376   0.2898570052   0.193067056  -0.62742342
Jawbusters                     2.62491587  -0.6343671618   0.114043053  -0.54172092
Junior Mints                  -0.16010610   1.6194428347   0.442156347  -0.08935729
Kit Kat                       -2.87086546   0.9069655335  -0.545771148  -2.94691419
Laffy Taffy                    1.65450042  -0.2379605144   1.217408326  -0.81578254
Lemonhead                      2.33564695  -1.2553404646   1.125823900  -1.18755633
Lifesavers big ring gummies    1.19528766  -0.0783610246   0.814040659  -0.61506538
Peanut butter M&M's           -1.52223814   1.9291395890  -0.815897653   2.27060871
M&M's                         -0.76747561   1.2573539136  -1.260658369   0.98037043
Mike & Ike                     1.57487290   0.0664259746  -1.114406454   0.84284942
Milk Duds                     -0.76836937   0.4192793946  -0.137573021  -0.01222490
Milky Way                     -3.69272218  -2.4933313173   0.843423990   0.70792542
Milky Way Midnight            -3.23036513  -2.8201031327   0.902884388   0.75904764
Milky Way Simply Caramel      -3.04936226  -1.1774777304  -1.382617058   0.15488883
Mounds                        -1.81292795   0.2120726312   0.636094539  -0.91126847
Mr Good Bar                   -2.67327849   0.9217207344   0.997161433   0.40634715
Nerds                          1.93426895  -0.9133307225  -1.670281710   0.79359684
Nestle Butterfinger           -2.97855081   0.8798835368   0.348599786   0.94918411
Nestle Crunch                 -2.92740488   0.8119013154  -0.747159803  -2.97250665
Nik L Nip                      1.63985272   0.4210217322  -0.083217936  -0.24015596
Now & Later                    1.98070982   0.5117150919   0.460099768  -0.21340533
Payday                        -2.39180556  -1.4839637512   2.091687409   1.65680787
Peanut M&Ms                   -1.38897069   2.0947188031  -0.260214925   1.89874716
Pixie Sticks                   1.67042227   0.8969792365   1.394703254  -0.48031281
Pop Rocks                      1.76879348  -0.8060325640  -1.567639814   0.42017068
Red vines                      2.12406849   0.1366822960  -0.115183020   0.27721045
Reese's Miniatures            -1.55210251   1.9287569793   1.884620322   0.48605155
Reese's Peanut Butter cup     -2.28427985   1.4648923293  -0.156138940   1.63994489
Reese's pieces                -1.40590761   2.3077984818   0.136661895   1.64348077
Reese's stuffed with pieces   -2.13382398   1.0787289654  -0.673152403   1.96066081
Ring pop                       1.19274412  -1.7069749284  -1.423826969   0.02570015
Rolo                          -1.61259322   0.1773734932  -1.931879747   1.00697641
Root Beer Barrels              2.10440254  -0.8711340556  -0.594335570   0.36204105
Runts                          2.25699185  -1.1223199934  -1.557678507   0.71996790
Sixlets                        0.81799664   1.1888290122   1.093105891  -0.33001942
Skittles original              1.29259129   0.2263705137  -1.306145308   1.05900061
Skittles wildberry             1.47148517   0.1118354559  -1.232745536   0.99503181
```

```
Nestle Smarties            -0.27556563   1.3792344137  -0.080047831  -0.03345921
Smarties candy              2.60115214  -0.6047947520   0.003482896  -0.21938457
Snickers                   -4.39576792  -1.7919312516   1.434654778   1.89861891
Snickers Crisper           -4.01457335  -0.0347673522  -1.089868643  -1.42309414
Sour Patch Kids             1.81551769   0.8879445215   0.863881832  -0.32795118
Sour Patch Tricksters       1.97326660   0.7869473239   0.928605869  -0.38435897
Starburst                   1.50658493   0.9437290830   0.487658690  -0.12150268
Strawberry bon bons         2.80647837  -1.0331193111  -0.524069119   0.13960292
Sugar Babies               -0.01900559  -0.8219542293  -1.802826526   0.92212945
Sugar Daddy                 0.19642038  -1.2073694698   0.520140143  -0.59663517
Super Bubble                1.99242820  -0.3915898648   1.481310204  -1.01767057
Swedish Fish                1.00547407   0.5003327040  -1.068588828   0.59499272
Tootsie Pop                 0.84734171  -1.1060686710  -0.480874078  -0.23675349
Tootsie Roll Juniors       -0.40463667   0.5848580362   0.836999949  -0.59623669
Tootsie Roll Midgies        0.66730732   1.3709464980   1.179339290  -0.32789249
Tootsie Roll Snack Bars    -1.31149842   0.0009721286   0.885976952  -0.78242207
Trolli Sour Bites           1.85048456   0.5304055168   0.254559391  -0.02435072
Twix                       -4.12909044  -0.2180299573  -1.943536689  -2.52917855
Twizzlers                   1.56312584  -0.1794588354   1.179917535  -0.78310886
Warheads                    2.30707033  -1.2940268825   1.004249910  -1.11380521
Welch's Fruit Snacks        1.84808801   0.5022006184   0.213204782  -0.03387254
Werther's Original Caramel  0.68420363  -2.0146385440   0.506488679  -0.97208672
Whoppers                   -1.42549552   1.3654147702  -2.759982292  -1.23030133
                                   PC5            PC6           PC7           PC8
100 Grand                  -0.66236243  -0.545218405  -0.14340559   0.577224229
3 Musketeers               -0.16006665   0.382588420   2.02155526  -1.502574960
One dime                   -0.26977985   0.094950530  -0.77220066   0.155622104
One quarter                -0.05093737   0.428354037  -0.59089199   0.359134372
Air Heads                   0.33108524   0.188791601   0.96808076   1.012393252
Almond Joy                  1.43200435   1.022273480  -0.66604596   0.608261288
Baby Ruth                  -0.53157179   0.026152894  -0.27477854   0.617961302
Boston Baked Beans         -0.01286555   1.147068667  -1.66290853   1.017203801
Candy Corn                 -0.92992691   0.482570575  -0.17268382  -0.433778134
Caramel Apple Pops         -1.26916612  -0.641471733  -0.77030933   1.705704289
Charleston Chew            -0.48839246   1.490541025   1.23825892  -1.445444979
Chewey Lemonhead Fruit Mix -0.64216946   0.867070020   0.51175007   0.500258867
Chiclets                   -0.82031918   0.835939962  -0.06670709   0.514147274
Dots                       -0.57043885   0.625011322   0.68287986   0.487777458
Dum Dums                    1.60464382  -0.712637531  -0.06881335  -0.229647214
Fruit Chews                -0.64943194  -0.515343269   0.62911226   0.295017216
Fun Dip                     1.64564645  -0.061325594  -0.22205535  -0.058799069
Gobstopper                  0.97535806   0.006638151   0.09322099  -0.594393205
Haribo Gold Bears          -0.41418743  -0.076745232   0.97316112   0.470072508
```

```
Haribo Happy Cola              -0.96450884   0.897177682  -0.57777493  -0.279999961
Haribo Sour Bears              -0.47964306   0.144138214   0.81700174   0.481462045
Haribo Twin Snakes             -0.58554144   0.501497845   0.56435691   0.499888780
Hershey's Kisses               -0.65415252  -0.956497265  -0.15459157  -1.475048333
Hershey's Krackel               0.64433919   0.815979037   0.72869731  -0.058197261
Hershey's Milk Chocolate        0.45257976   1.365951323   0.09727415  -0.537271619
Hershey's Special Dark          0.48406856   1.259690716   0.17239789  -0.542750795
Jawbusters                      0.74872541   0.777635266  -0.86953189  -0.404938915
Junior Mints                   -0.56772916  -0.100807494  -0.27701048  -1.244289684
Kit Kat                         0.74560410  -0.651964126   1.26789202  -0.308264629
Laffy Taffy                     0.12742282  -0.323401158   0.50907401   0.902891431
Lemonhead                       1.59348201  -0.612577970  -0.46974919  -0.088239458
Lifesavers big ring gummies     0.28553918  -0.405905368   0.76717408   0.968451468
Peanut butter M&M's             0.70534888  -0.535487043  -0.31945944  -0.055419751
M&M's                          -0.42260375  -0.093849007   0.23580574  -1.272394704
Mike & Ike                     -0.54748868   0.068855836   0.96217007   0.350296998
Milk Duds                      -1.94930706  -1.088570940  -1.46793514  -0.532609515
Milky Way                      -1.43541753  -0.603255019   0.91522507  -0.708957591
Milky Way Midnight             -1.60495490  -0.577210751   0.75050728  -0.825932767
Milky Way Simply Caramel       -0.81154750  -0.095588630  -0.56663709   0.057913957
Mounds                          0.34136052   1.562117376  -0.17156397  -0.536908890
Mr Good Bar                     1.49882941   1.178267833  -0.70645139   0.710404328
Nerds                           1.05269810  -0.612484674   0.36162582  -0.677966908
Nestle Butterfinger             1.66950147   0.246674393  -0.03583585   0.547351886
Nestle Crunch                   0.66634611   0.310195315   0.85793736  -0.137897493
Nik L Nip                      -0.74139722   2.364290586  -0.37097983   0.877307522
Now & Later                    -0.64458895   0.275284807   0.43215999   0.459054334
Payday                          0.57021053   1.685803617   0.71659189   0.920899230
Peanut M&Ms                     0.67649842  -0.481270608  -0.49445201  -0.017712739
Pixie Sticks                   -1.00053741  -0.249857693  -0.45261662  -0.491657159
Pop Rocks                       0.96236780   1.036365445  -0.40726527  -0.314719629
Red vines                      -0.69083254  -0.068691578   0.66734570   0.288409391
Reese's Miniatures              1.52874138  -1.896976303  -0.34129661   0.391227548
Reese's Peanut Butter cup       1.62984822  -1.101118658  -0.10711621   0.504507095
Reese's pieces                  0.71687442  -0.652294629  -0.48369702   0.001608461
Reese's stuffed with pieces     1.50741757  -0.638134777  -0.27656579   0.488051492
Ring pop                        1.69809389   1.498435385  -0.64909630   0.323514628
Rolo                           -1.75944281  -0.678446043  -1.06073297  -0.430807299
Root Beer Barrels               0.42962062  -0.109312144  -1.19510153  -1.555611980
Runts                           0.90293644  -0.227471967   0.05500446  -0.683423916
Sixlets                        -0.89038600  -0.174585492  -0.66527067  -1.454385221
Skittles original              -0.37037605  -0.800924438   1.50687697   0.245536582
Skittles wildberry             -0.46191313  -0.492027764   1.28849431   0.261464394
```

```
Nestle Smarties          -0.71703419  1.677902657 -1.00236322 -0.943766974
Smarties candy            0.89831080 -0.766691658 -0.09224917 -0.697074269
Snickers                 -0.32409247 -0.999567686  0.29375023  0.519082351
Snickers Crisper          0.28905186 -0.982077142 -0.99809936  1.715122168
Sour Patch Kids          -0.44613213 -0.989653570  1.01684337  0.317967197
Sour Patch Tricksters    -0.52684967 -0.717267988  0.82427320  0.332012368
Starburst                -0.34591253 -1.030380238  1.20373631  0.352584423
Strawberry bon bons       0.76654988 -0.423292666 -0.21835972 -0.752170933
Sugar Babies             -2.27338367  0.557052968 -1.67505130  0.548087948
Sugar Daddy              -1.58729466 -0.485682924 -1.85240975  0.941435233
Super Bubble             -0.03565422  0.216125641  0.09347559  0.939439208
Swedish Fish             -0.39241423  0.662591135  0.83851208  0.624071063
Tootsie Pop               1.88703233 -0.633721587 -0.18826747 -1.043909239
Tootsie Roll Juniors      0.05449315  0.394706045 -0.68468874 -0.662902544
Tootsie Roll Midgies     -0.77589776 -0.759442617 -0.35278919 -1.510639321
Tootsie Roll Snack Bars   0.28510912  0.328392863  0.22536722 -0.875755689
Trolli Sour Bites        -0.56414212 -0.168850396  0.72703429  0.389139607
Twix                     -0.49203516 -1.029755076  0.13814854  0.601927243
Twizzlers                 0.17417770 -0.481177950  0.62061848  0.894755897
Warheads                  1.59506307 -0.576609554 -0.45482313 -0.087796992
Welch's Fruit Snacks     -0.58742724  0.067119205  0.62144223  0.428665533
Werther's Original Caramel 0.00698052 -1.343145309 -2.51219819 -0.091968746
Whoppers                 -0.46170658  0.678567795  0.16799768 -0.639281286
                                       PC9        PC10        PC11
100 Grand                 0.379148201 -0.154099543  0.141903843
3 Musketeers              0.023832701 -0.155239068 -0.517927160
One dime                 -0.453988975 -0.943783623 -0.515870807
One quarter              -0.638889823  0.095579648 -0.986719399
Air Heads                -1.383012190 -0.573723493 -0.514453684
Almond Joy               -0.139059851 -0.312778705  1.061148660
Baby Ruth                 1.143067977  0.264099876 -0.009085882
Boston Baked Beans        1.396216830 -0.973188567  0.102073964
Candy Corn               -0.173119296 -1.580718107 -0.026369468
Caramel Apple Pops       -0.916479426 -0.201217892  0.098522113
Charleston Chew           0.265940391 -0.852594237 -0.299018234
Chewey Lemonhead Fruit Mix -0.056700856 -0.274516517  0.133783407
Chiclets                  0.600474625  0.299142845  0.263138381
Dots                     -0.109595061 -0.122162631  0.085957526
Dum Dums                 -0.385161810 -0.858986682  0.026635152
Fruit Chews               0.431017129  0.094469913  0.405402342
Fun Dip                  -0.432729164 -0.348623806 -0.240086957
Gobstopper                0.401434380  0.160051400  0.173108273
Haribo Gold Bears        -0.021735861  0.618649969 -0.035818163
```

```
Haribo Happy Cola              0.174940628 -0.664580531 -0.210044044
Haribo Sour Bears              0.026531168  0.479623974  0.007823919
Haribo Twin Snakes            0.104620713  0.254698661  0.078430922
Hershey's Kisses               0.015418665 -0.009459574  0.108560740
Hershey's Krackel              0.589905081 -0.259828939 -0.295461506
Hershey's Milk Chocolate      -1.211270861  0.556641044  0.566669950
Hershey's Special Dark        -1.234490726  0.623522410  0.545675013
Jawbusters                     1.175121068  1.211984587  0.107699541
Junior Mints                  -0.125729538  0.657349544 -0.278384695
Kit Kat                        0.627262392 -0.428229606 -0.052463331
Laffy Taffy                   -0.694575541 -0.358253574 -0.196159662
Lemonhead                      0.133629280  0.441715610 -0.165939908
Lifesavers big ring gummies  -0.856144679  0.130909678 -0.425911984
Peanut butter M&M's            0.267660757 -0.221930007 -0.095957385
M&M's                         -0.712293695  0.050005905 -0.359922697
Mike & Ike                    -0.220487198 -0.593471988  0.252731956
Milk Duds                     -0.136911155  1.039501593  0.120288784
Milky Way                      0.004765911  0.843304841 -0.327048378
Milky Way Midnight             0.046247663 -0.050365745 -0.014682753
Milky Way Simply Caramel      -1.628250067  0.337544026  1.024026651
Mounds                        -1.038014053  0.444568191  0.664244259
Mr Good Bar                   -0.083279402  0.319687965  0.860817575
Nerds                          0.395546909  0.241504770  0.214795902
Nestle Butterfinger           -0.418486899 -0.056067346  0.931659519
Nestle Crunch                  0.670459908 -0.224206267 -0.210210679
Nik L Nip                      0.390203352  1.144571263 -0.293659295
Now & Later                    0.340332791  0.362285847  0.182098894
Payday                         1.592516886 -0.586376631 -0.517961675
Peanut M&Ms                    0.463012748  0.130263439 -0.124891428
Pixie Sticks                   0.506958178 -0.721036477  0.100340132
Pop Rocks                      0.614743430  1.229233468 -0.196822517
Red vines                      0.116010128 -0.683310675  0.459814126
Reese's Miniatures            -0.201912762  0.339292365 -0.538620165
Reese's Peanut Butter cup     -0.813354976 -0.127457497 -0.769615618
Reese's pieces                 0.573552037  0.549292897 -0.190638431
Reese's stuffed with pieces  -0.924164172 -0.865127171 -0.632341326
Ring pop                      -0.509522417  0.693685961 -0.801535980
Rolo                          -0.716390425  0.955937401 -0.177541179
Root Beer Barrels              0.748159981 -1.361887249  0.275586696
Runts                          0.490713003 -0.186267132  0.357472263
Sixlets                        0.120551003 -0.694412596  0.295248736
Skittles original             -0.396609466 -0.492989386  0.235345367
Skittles wildberry            -0.329109954 -0.687411685  0.296377074
```

```
Nestle Smarties             -0.095850762  0.891841882 -0.546823295
Smarties candy               0.957784710  0.645048933  0.369028845
Snickers                     1.040500931  0.639420129 -0.064039216
Snickers Crisper             1.597773064 -0.876877339  0.668926998
Sour Patch Kids              0.319739180  0.748914211  0.190355889
Sour Patch Tricksters        0.379260358  0.577472316  0.244173747
Starburst                    0.178115491  0.966910650  0.055007728
Strawberry bon bons          0.831798717 -0.257396978  0.567295681
Sugar Babies                -0.204221618 -0.392704656 -0.026264253
Sugar Daddy                 -0.751459910 -0.659787353 -0.269135682
Super Bubble                -0.530802993 -0.601257076 -0.099471654
Swedish Fish                -0.159360361  0.838989403 -0.260036175
Tootsie Pop                 -0.721118894  0.326836317 -0.137911107
Tootsie Roll Juniors        -1.149669881 -0.284072969 -0.704615875
Tootsie Roll Midgies         0.074811904 -0.471079914  0.266947340
Tootsie Roll Snack Bars     -1.078720446 -0.723988086  1.173413519
Trolli Sour Bites            0.215399333  0.265589321  0.205354764
Twix                         0.390371217  0.606070076 -0.048120393
Twizzlers                   -0.729052624 -0.258947465 -0.227333151
Warheads                     0.096482954  0.378764818 -0.167107014
Welch's Fruit Snacks         0.229115315  0.300496305  0.173167709
Werther's Original Caramel   0.028457521  0.448967578 -0.298604763
Whoppers                     1.212121185 -1.054026246 -1.220426934
                                    PC12
100 Grand                    0.0646988318
3 Musketeers                -0.1839471662
One dime                    -0.5107477914
One quarter                 -0.9582719109
Air Heads                    0.0442992399
Almond Joy                   0.1889347102
Baby Ruth                    0.5798508261
Boston Baked Beans          -0.1012981497
Candy Corn                  -0.6799350975
Caramel Apple Pops           0.6962504166
Charleston Chew              0.5494017342
Chewey Lemonhead Fruit Mix   0.3105973331
Chiclets                     0.5337581991
Dots                         0.1503797578
Dum Dums                     0.1849692597
Fruit Chews                  0.1399898762
Fun Dip                      0.1237990873
Gobstopper                  -0.1944414618
Haribo Gold Bears           -0.2638818244
```

```
Haribo Happy Cola          -0.6879064447
Haribo Sour Bears          -0.1176800541
Haribo Twin Snakes          0.1188546962
Hershey's Kisses            0.1984492942
Hershey's Krackel          -0.0817067486
Hershey's Milk Chocolate   -0.3573702410
Hershey's Special Dark     -0.4277036644
Jawbusters                  0.1337584661
Junior Mints                0.0649489483
Kit Kat                    -0.3769186469
Laffy Taffy                 0.3026020451
Lemonhead                   0.0619011170
Lifesavers big ring gummies -0.0228900695
Peanut butter M&M's         0.1200218111
M&M's                      -0.1023879726
Mike & Ike                  0.1113757406
Milk Duds                   0.3400578947
Milky Way                  -0.1552479006
Milky Way Midnight          0.2304632793
Milky Way Simply Caramel   -0.2533060125
Mounds                     -0.1415021722
Mr Good Bar                 0.0210418054
Nerds                      -0.3936975444
Nestle Butterfinger        -0.3100685916
Nestle Crunch              -0.1731325150
Nik L Nip                   0.4597035356
Now & Later                 0.1806417362
Payday                     -0.7729766006
Peanut M&Ms                 0.1319358415
Pixie Sticks               -0.7378511713
Pop Rocks                  -0.1936682261
Red vines                   0.3438211129
Reese's Miniatures         -0.0689054242
Reese's Peanut Butter cup  -0.1005437626
Reese's pieces             -0.0005861225
Reese's stuffed with pieces  0.2335945192
Ring pop                    0.0735408387
Rolo                        0.0787870205
Root Beer Barrels          -0.7082327026
Runts                      -0.0585635801
Sixlets                     0.7459023430
Skittles original          -0.2795750540
Skittles wildberry         -0.0751177198
```

```
Nestle Smarties               0.4628839641
Smarties candy               -0.2022435492
Snickers                      0.0911696493
Snickers Crisper              0.6306015108
Sour Patch Kids              -0.3186623415
Sour Patch Tricksters        -0.1383715363
Starburst                    -0.5130782116
Strawberry bon bons           0.1543709171
Sugar Babies                 -0.4541934821
Sugar Daddy                  -0.2885121237
Super Bubble                  0.6537195364
Swedish Fish                 -0.2507351031
Tootsie Pop                   1.2506316262
Tootsie Roll Juniors          0.5695856300
Tootsie Roll Midgies          0.4724506497
Tootsie Roll Snack Bars      -0.0373596827
Trolli Sour Bites             0.0146955266
Twix                         -0.3530484324
Twizzlers                     0.1981702880
Warheads                      0.0702649429
Welch's Fruit Snacks          0.0727669700
Werther's Original Caramel   -0.8335356607
Whoppers                      0.7201599382
```

```r
plot(pca$x[,1], pca$x[,2], col = my_cols, pch = 16)
```
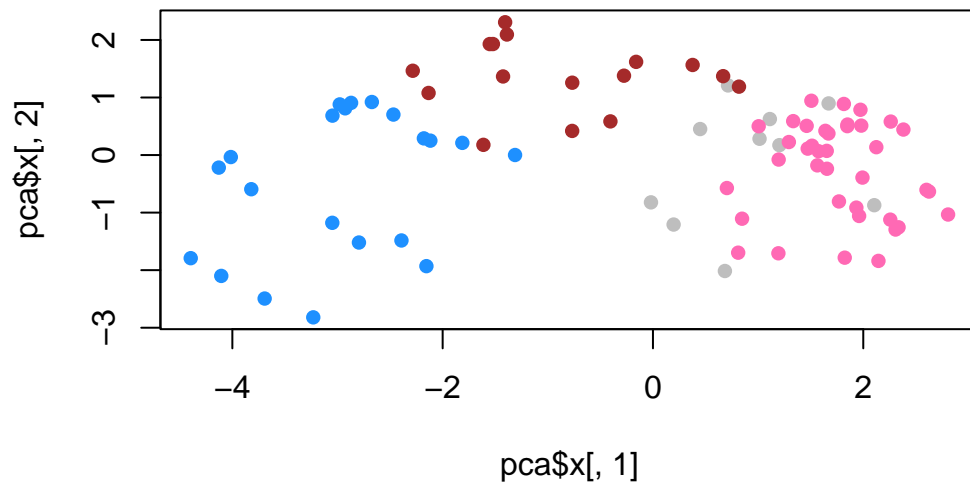
```
my_data <- cbind(candy.data, pca$x[,1:3])
p <- ggplot(my_data) +
        aes(x=PC1, y=PC2,
            size=winpercent/100,
            text=rownames(my_data),
            label=rownames(my_data)) +
        geom_point(col=my_cols)
p
```
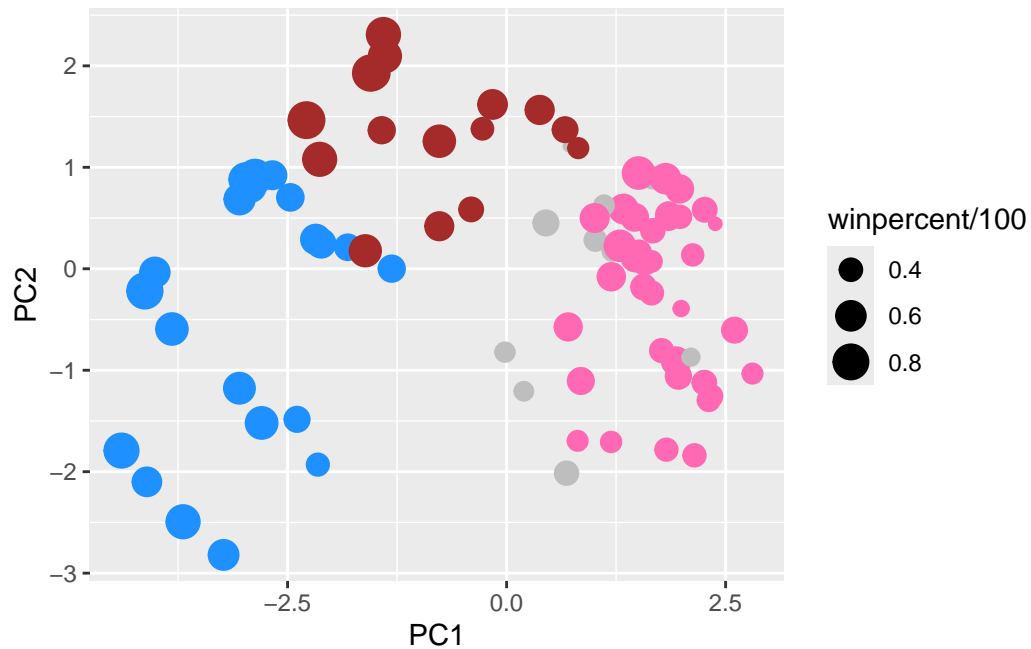
Let's add repel to this:

```
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 (brown), chocolate other (blue), fruity (p
        caption="Data from 538")
```

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

Halloween Candy PCA Space
Colored by type: chocolate bar (brown), chocolate other (blue), fruity (pink)
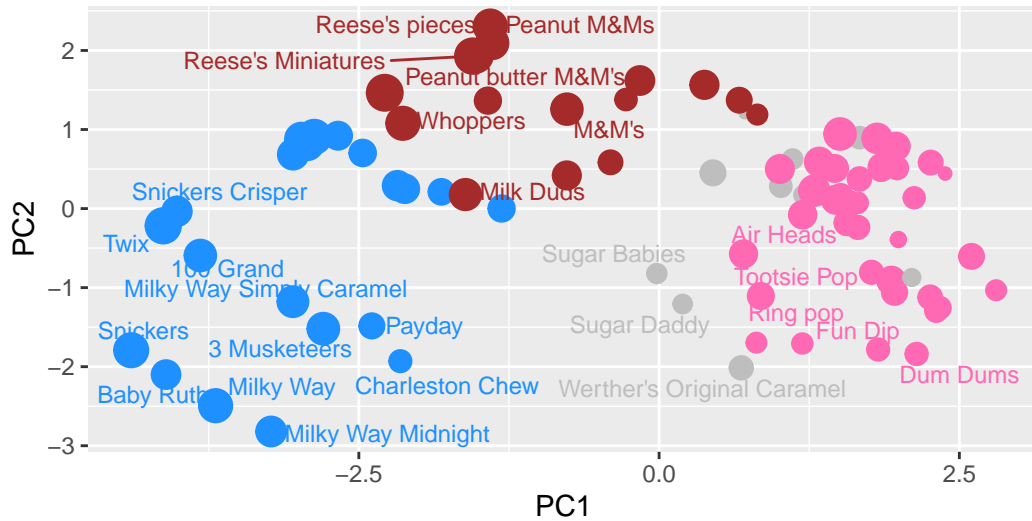
Data from 538

```
library(plotly)
ggplotly(p)
```

How do the original variables contribute to our PCs? For this we look at the loading component of our results object i.e. `pca$rotation`:

```
pca$rotation
```

	PC1	PC2	PC3	PC4	PC5
chocolate	-0.4019466	0.21404160	0.01601358	-0.016673032	0.066035846
fruity	0.3683883	-0.18304666	-0.13765612	-0.004479829	0.143535325
caramel	-0.2299709	-0.40349894	-0.13294166	-0.024889542	-0.507301501
peanutyalmondy	-0.2407155	0.22446919	0.18272802	0.466784287	0.399930245
nougat	-0.2268102	-0.47016599	0.33970244	0.299581403	-0.188852418
crispedricewafer	-0.2215182	0.09719527	-0.36485542	-0.605594730	0.034652316
hard	0.2111587	-0.43262603	-0.20295368	-0.032249660	0.574557816
bar	-0.3947433	-0.22255618	0.10696092	-0.186914549	0.077794806
pluribus	0.2600041	0.36920922	-0.26813772	0.287246604	-0.392796479
sugarpercent	-0.1083088	-0.23647379	-0.65509692	0.433896248	0.007469103
pricepercent	-0.3207361	0.05883628	-0.33048843	0.063557149	0.043358887
winpercent	-0.3298035	0.21115347	-0.13531766	0.117930997	0.168755073
	PC6	PC7	PC8	PC9	PC10

```
chocolate       -0.09018950 -0.08360642 -0.49084856 -0.151651568  0.107661356
fruity          -0.04266105  0.46147889  0.39805802 -0.001248306  0.362062502
caramel         -0.40346502 -0.44274741  0.26963447  0.019186442  0.229799010
peanutyalmondy  -0.09416259 -0.25710489  0.45771445  0.381068550 -0.145912362
nougat           0.09012643  0.36663902 -0.18793955  0.385278987  0.011323453
crispedricewafer -0.09007640  0.13077042  0.13567736  0.511634999 -0.264810144
hard            -0.12767365 -0.31933477 -0.38881683  0.258154433  0.220779142
bar              0.25307332  0.24192992 -0.02982691  0.091872886 -0.003232321
pluribus         0.03184932  0.04066352 -0.28652547  0.529954405  0.199303452
sugarpercent     0.02737834  0.14721840 -0.04114076 -0.217685759 -0.488103337
pricepercent     0.62908570 -0.14308215  0.16722078 -0.048991557  0.507716043
winpercent      -0.56947283  0.40260385 -0.02936405 -0.124440117  0.358431235
                      PC11        PC12
chocolate        0.10045278  0.69784924
fruity           0.17494902  0.50624242
caramel          0.13515820  0.07548984
peanutyalmondy   0.11244275  0.12972756
nougat          -0.38954473  0.09223698
crispedricewafer -0.22615618  0.11727369
hard             0.01342330 -0.10430092
bar              0.74956878 -0.22010569
pluribus         0.27971527 -0.06169246
sugarpercent     0.05373286  0.04733985
pricepercent    -0.26396582 -0.06698291
winpercent      -0.11251626 -0.37693153
```

```r
par(mar=c(8,4,2,2))
barplot(pca$rotation[,1], las=2, ylab="PC1 Contribution")
```
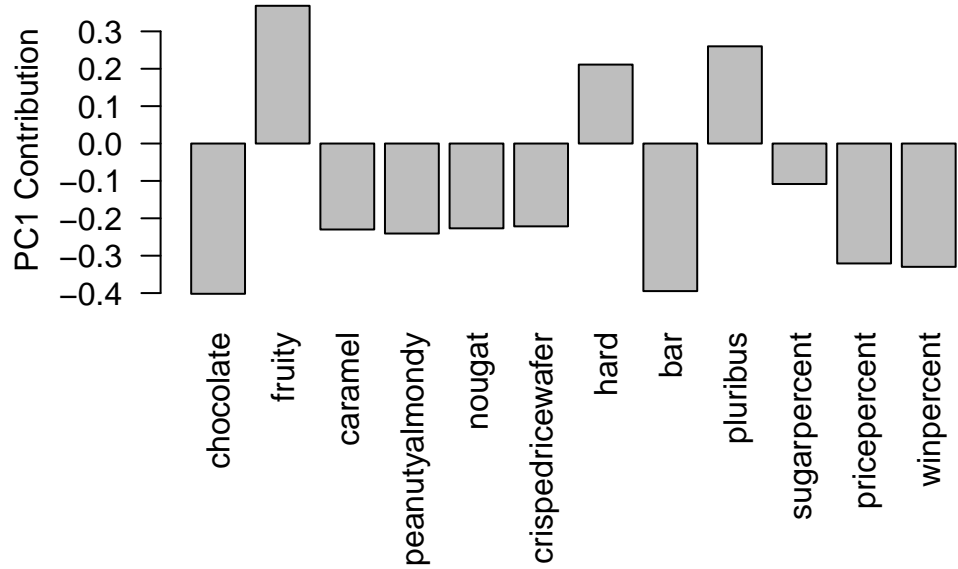
```
res = as.data.frame(pca$rotation)
ggplot(res) + aes(PC1, reorder(row.names(res), PC1)) + geom_col()
```
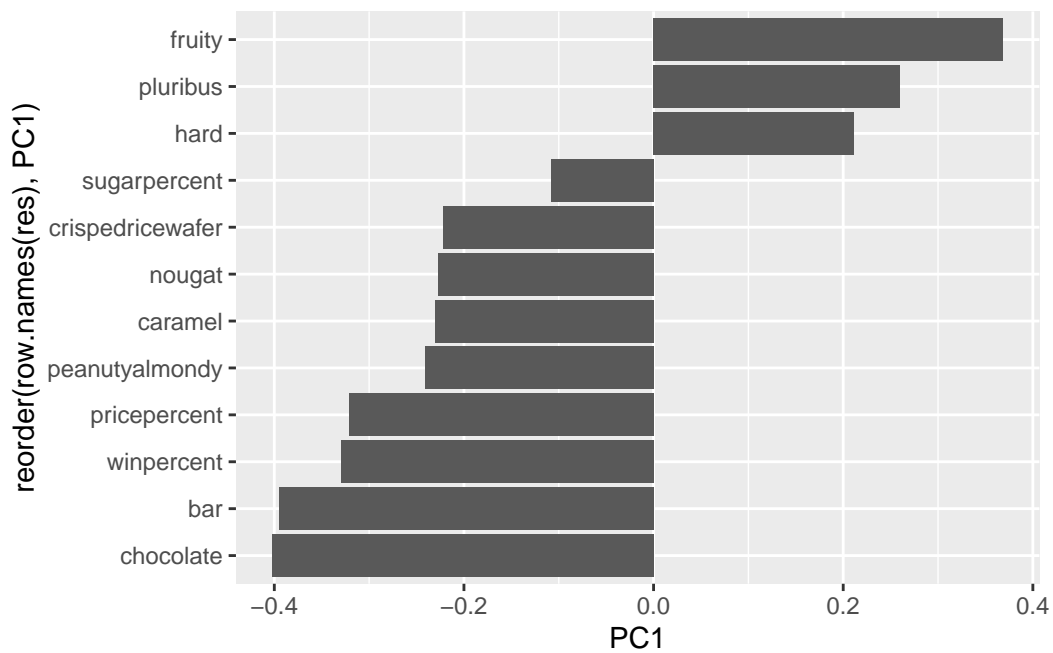
Q24. What original variables are picked up strongly by PC1 in the positive direction? Do these make sense to you?

Fruity, hard, and pluribus are picked up strongly by PC1. Makes sense because these are generally correlated with fruity candy while the other stuff is more correlated to chocolate candy.