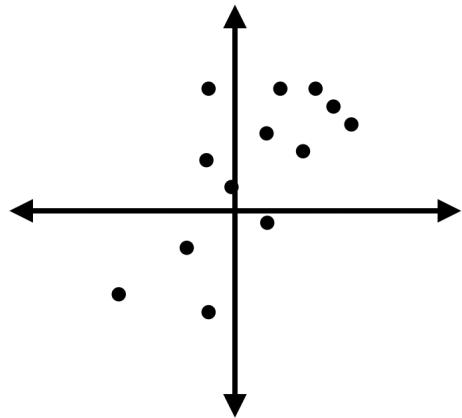
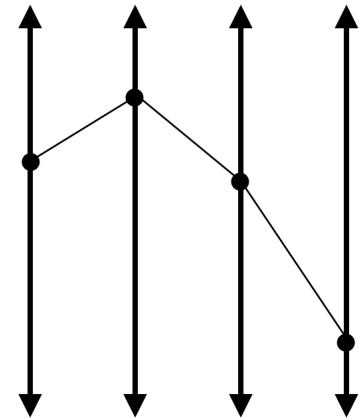


Biplots

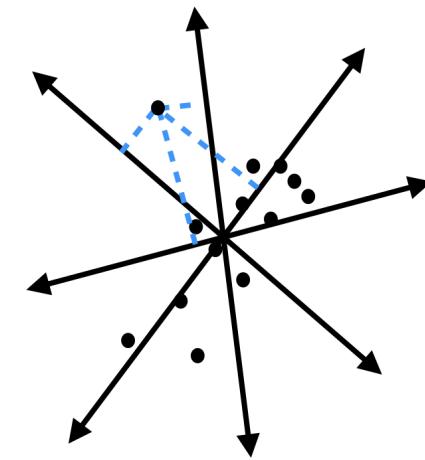
Axis positions



scatterplot



parallel coordinates plot

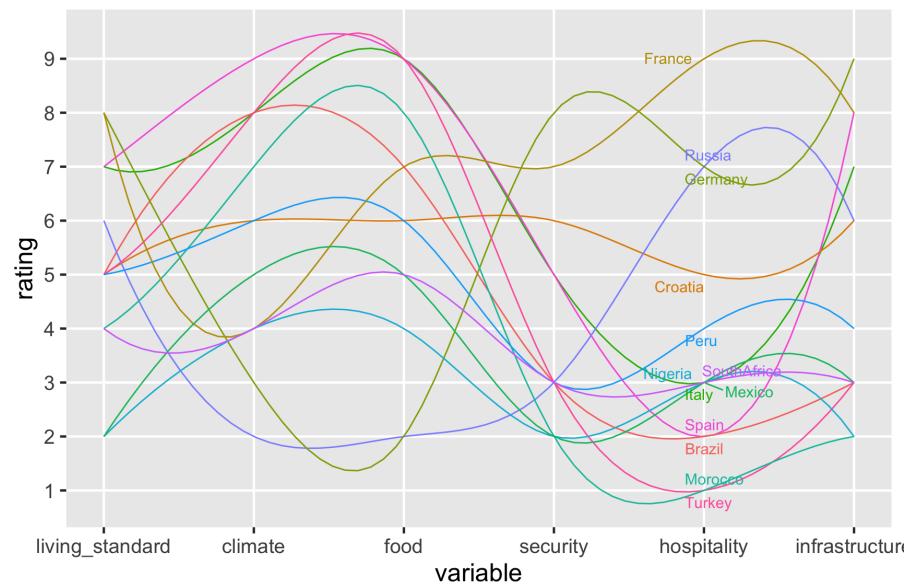


biplot

Multidimensional data

country	living_standard	climate	food	security	hospitality	infrastructure
Italy	7	8	9	5	3	7
Spain	7	9	9	5	2	8
Croatia	5	6	6	6	5	6
Brazil	5	8	7	3	2	3
Russia	6	2	2	3	7	6
Germany	8	3	2	8	7	9
Turkey	5	8	9	3	1	3
Morocco	4	7	8	2	1	2
Peru	5	6	6	3	4	4
Nigeria	2	4	4	2	3	2
France	8	4	7	7	9	8
Mexico	2	5	5	2	3	3
SouthAfrica	4	4	5	3	3	3

Parallel coordinates plot



Principal components analysis (PCA)

Goal: reduce dimensionality by finding a few linear combinations of the variables that capture most of the variance

Scaled data:

country	living_standard	climate	food	security	hospitality	infrastructure
Italy	0.90	1.04	1.20	0.5	-0.34	0.83
Spain	0.90	1.49	1.20	0.5	-0.74	1.23
Croatia	-0.12	0.14	-0.03	1.0	0.47	0.43
Brazil	-0.12	1.04	0.38	-0.5	-0.74	-0.77
Russia	0.39	-1.67	-1.68	-0.5	1.27	0.43
Germany	1.41	-1.22	-1.68	2.0	1.27	1.63
Turkey	-0.12	1.04	1.20	-0.5	-1.15	-0.77
Morocco	-0.63	0.59	0.79	-1.0	-1.15	-1.17
Peru	-0.12	0.14	-0.03	-0.5	0.06	-0.37
Nigeria	-1.64	-0.76	-0.85	-1.0	-0.34	-1.17
France	1.41	-0.76	0.38	1.5	2.08	1.23
Mexico	-1.64	-0.31	-0.44	-1.0	-0.34	-0.77
SouthAfrica	-0.63	-0.76	-0.44	-0.5	-0.34	-0.77

Variance of each scaled variable is 1 (by definition)

living_standard climate food security hospitality infrastructure

1

1

1

1

1

1

Principal components analysis (PCA)

Goal: reduce dimensionality by finding linear combinations of the variables that capture most of the variance

Variance of transformed data (principal component scores):

```
pca <- prcomp(ratings[,2:7], scale. = TRUE)
summary(pca)
```

```
## Importance of components:
##            PC1      PC2      PC3      PC4      PC5      PC6
## Standard deviation   1.854  1.4497  0.43959  0.39052  0.27517  0.19778
## Proportion of Variance 0.573  0.3503  0.03221  0.02542  0.01262  0.00652
## Cumulative Proportion 0.573  0.9232  0.95544  0.98086  0.99348  1.00000
```

```
pca$sdev^2
```

```
## [1] 3.43786752 2.10155050 0.19323758 0.15250902 0.07571813 0.03911725
```

```
sum(pca$sdev^2)
```

```
## [1] 6
```

Principal component loadings

also called principal component vectors, rotation

```
##          PC1     PC2     PC3     PC4     PC5     PC6
## living_standard -0.429  0.364  0.112 -0.673  0.466 -0.028
## climate         0.270  0.585 -0.210  0.036 -0.149 -0.719
## food            0.221  0.596  0.610  0.212 -0.077  0.417
## security        -0.475  0.244 -0.282  0.676  0.419  0.049
## hospitality     -0.484 -0.216  0.636  0.170 -0.213 -0.490
## infrastructure  -0.484  0.252 -0.297 -0.121 -0.731  0.256
```

Principal component scores

transformed data

```
pca$x
```

	PC1	PC2	PC3	PC4	PC5	PC6
[1,]	-0.3144044	2.0601117	0.009911069	-0.134077221	-0.15376381	0.130652255
[2,]	-0.1906992	2.5129060	-0.460310855	-0.234777788	-0.42763828	0.105956954
[3,]	-0.8282724	0.2714445	-0.175485202	0.780869173	-0.06824131	-0.178206062
[4,]	1.3859102	0.6388749	-0.104159743	-0.174708937	0.27192740	-0.445492153
[5,]	-1.5751250	-2.1218269	0.192930485	-0.852708398	-0.23467966	-0.047352038
[6,]	-3.6593276	-0.5746893	-0.849587594	0.024146583	0.34231023	0.030370889
[7,]	1.7629449	1.2165411	0.140877845	-0.069171948	0.29445443	0.094958158
[8,]	2.2000833	0.2988664	0.187885131	-0.119877663	0.23935404	0.135613994
[9,]	0.4671581	-0.2087754	0.228627707	-0.205534924	-0.02580749	-0.259717450
[10,]	1.5177586	-2.0201792	-0.131605174	0.305058184	-0.07781592	0.058312049
[11,]	-3.0431311	0.5179856	1.082495626	0.323774969	0.02733982	0.039265671
[12,]	1.5368589	-1.4095662	-0.094520836	0.360188174	-0.46936463	0.007202792
[13,]	0.7402455	-1.1816931	-0.027058461	-0.003180204	0.28192520	0.328434940

[Scaled data] x [PC loadings] = [PC scores]

[13 x 6][6 x 6] = [13 x 6]

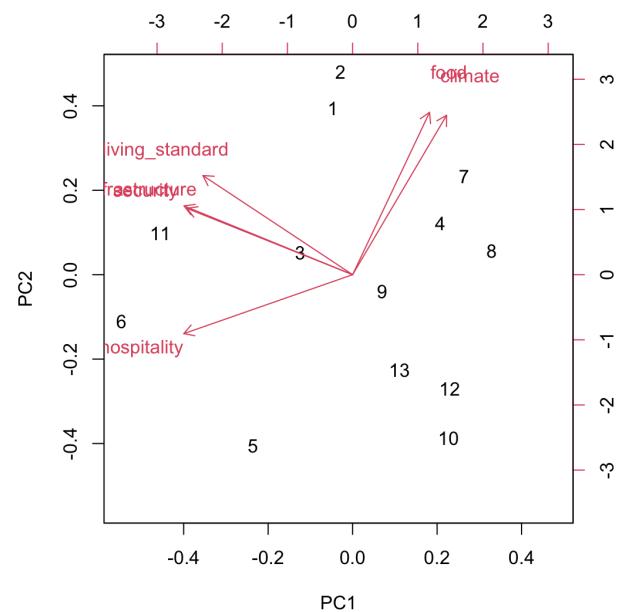
Principal component scores

PC1 = -0.43 x living_standard + 0.27 x climate + 0.22 x food + -0.48 x security + -0.48 x hospitality + -0.48 x infrastructure

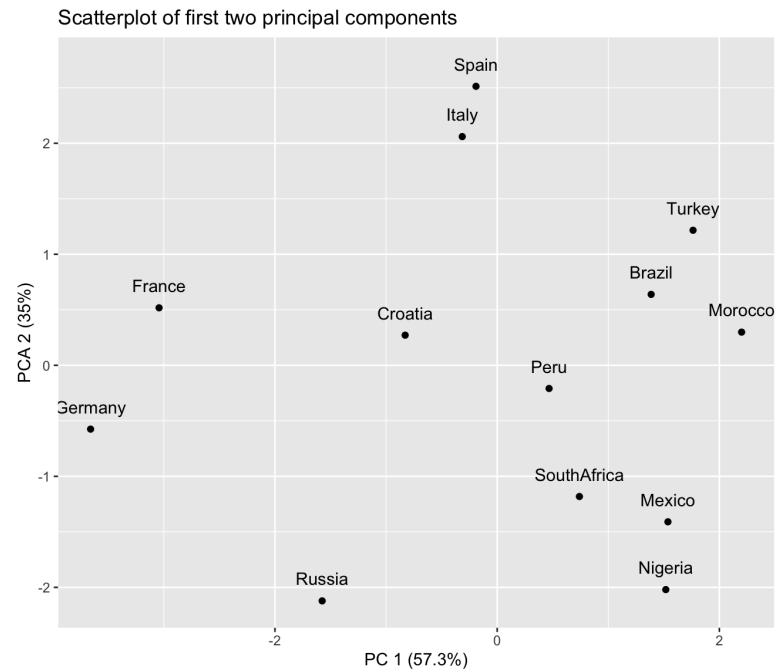
PC2 = 0.36 x living_standard + 0.59 x climate + 0.6 x food + 0.24 x security + -0.22 x hospitality + 0.25 x infrastructure

etc.

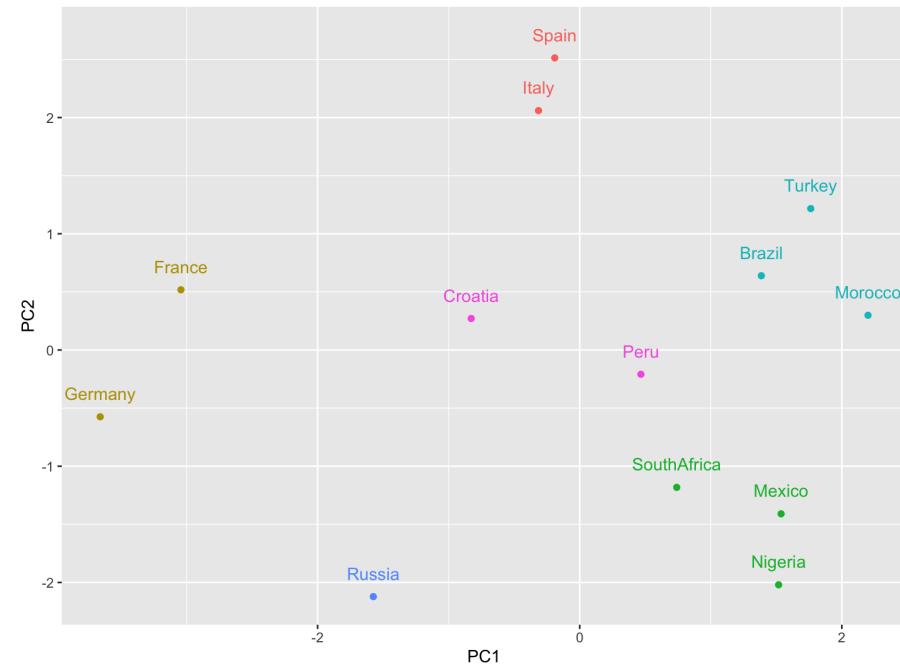
Biplot (default plot)



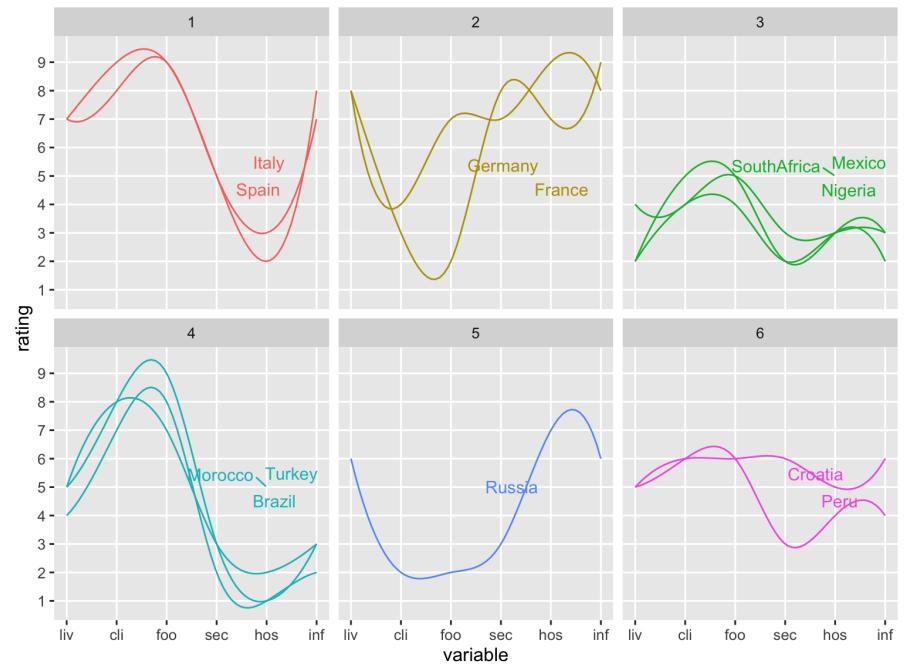
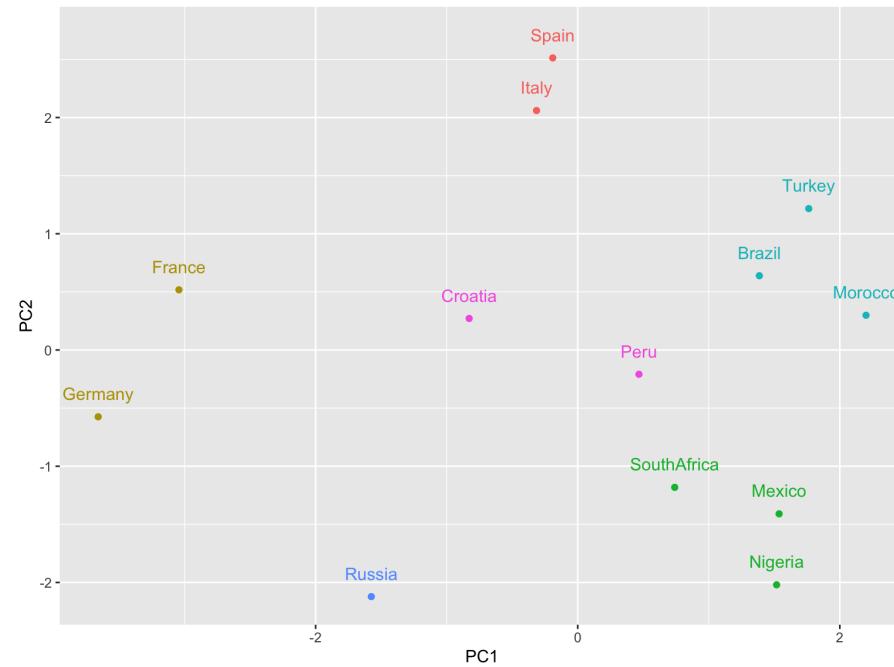
Biplot



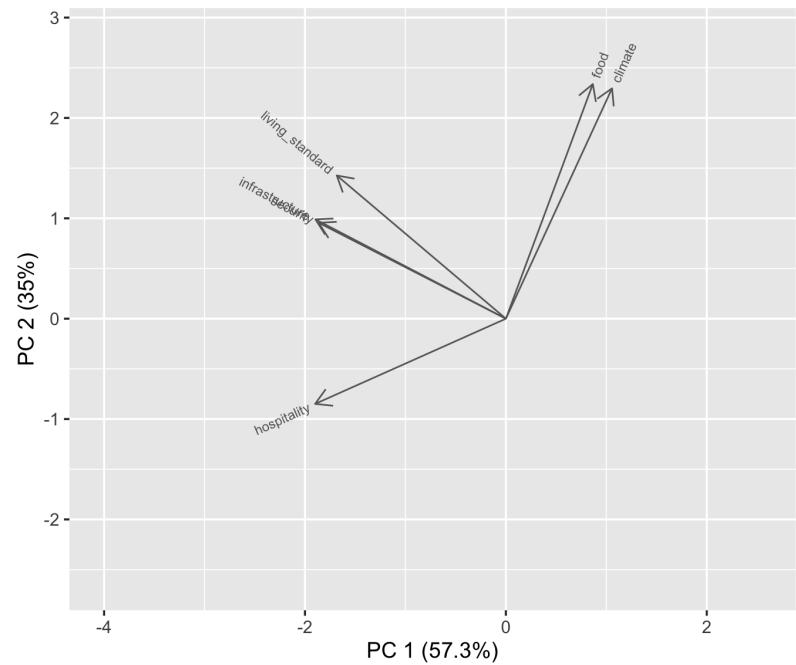
Clusters



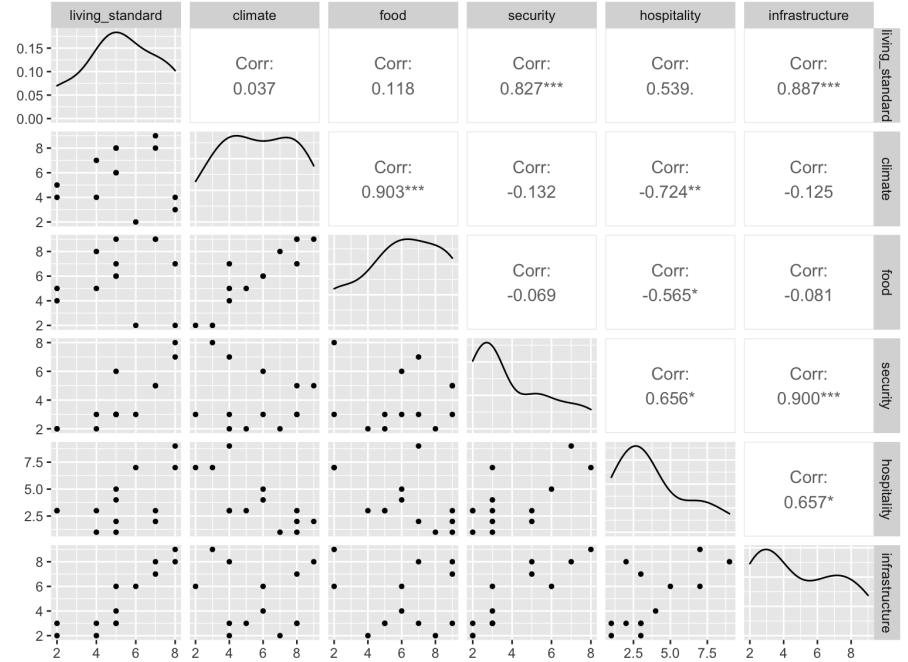
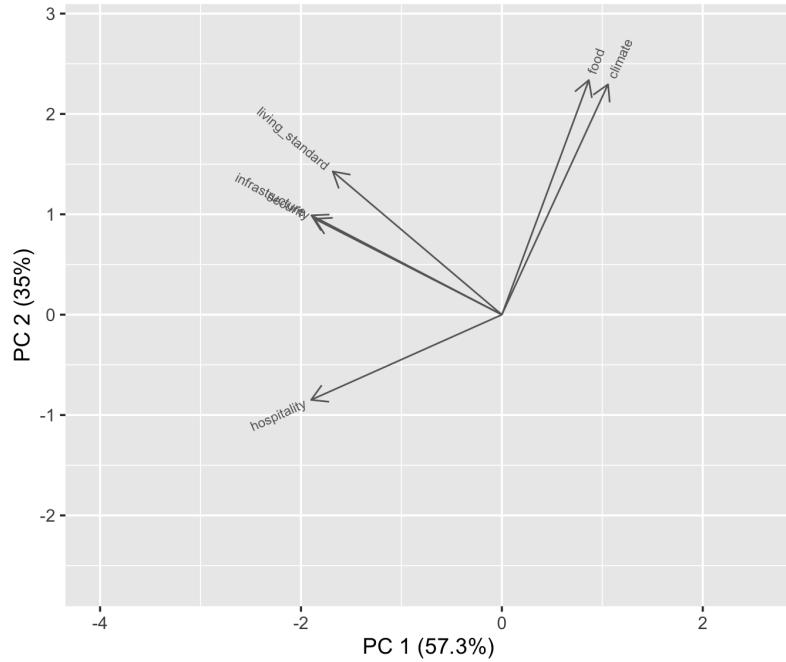
Clusters



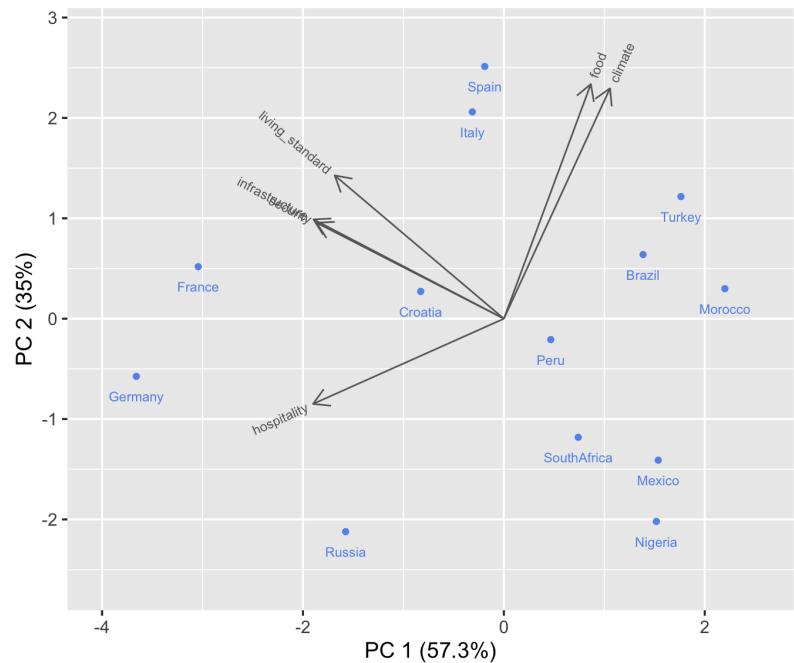
Rotation vectors only



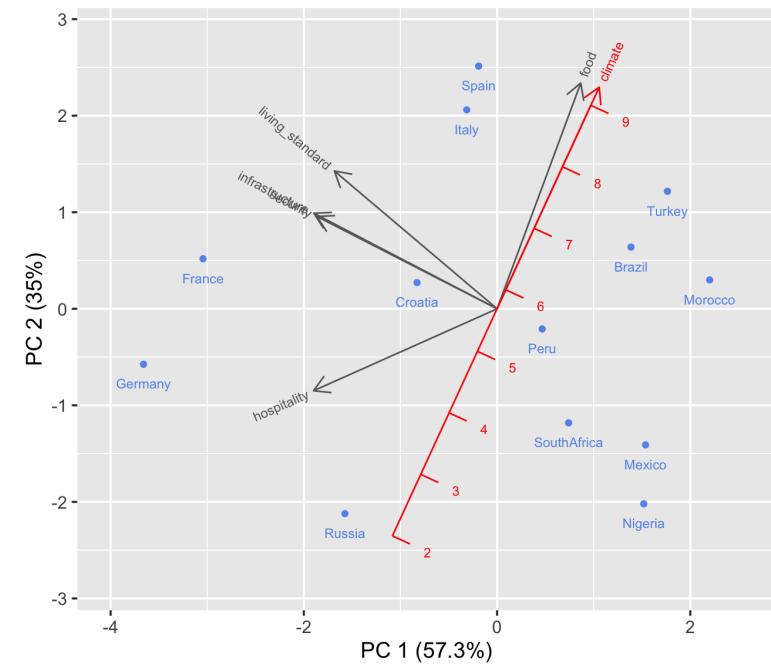
Rotation vectors only



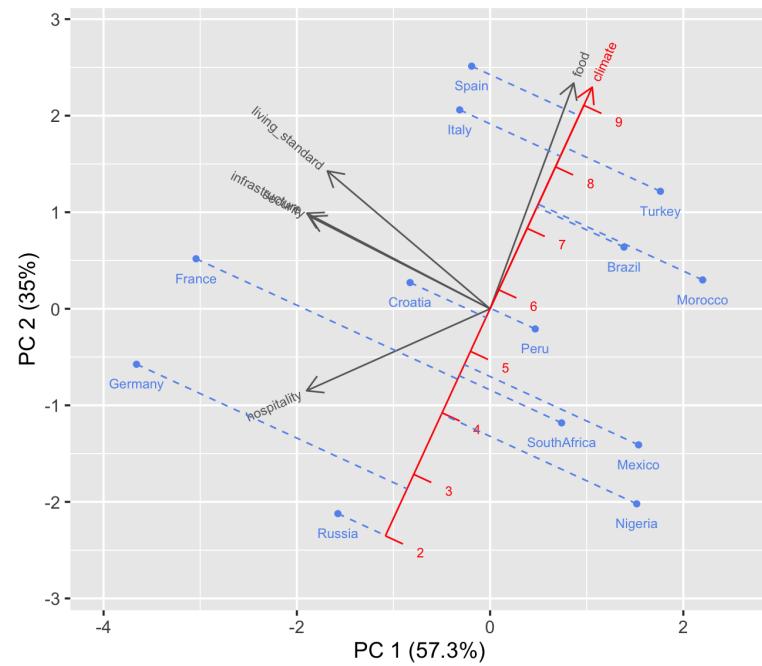
Biplot



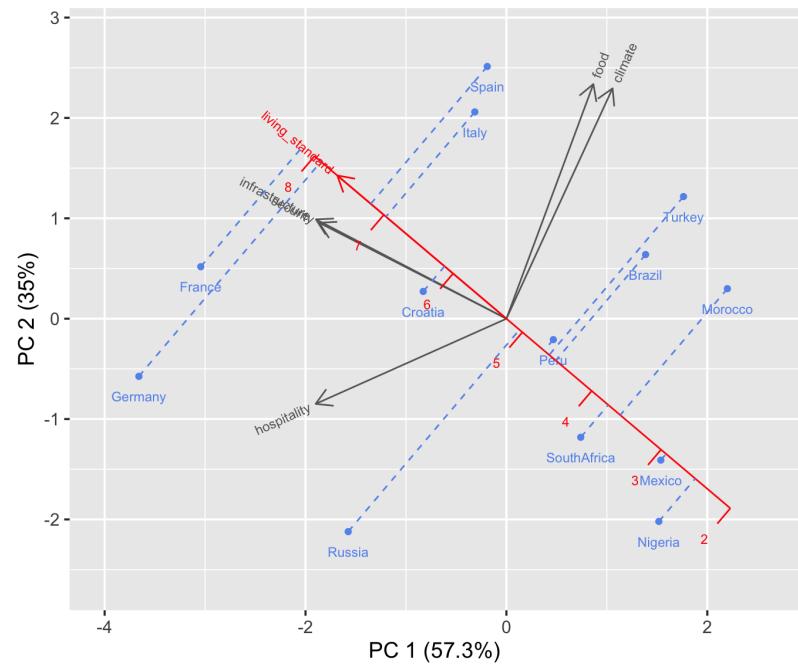
Biplot with calibrated axis



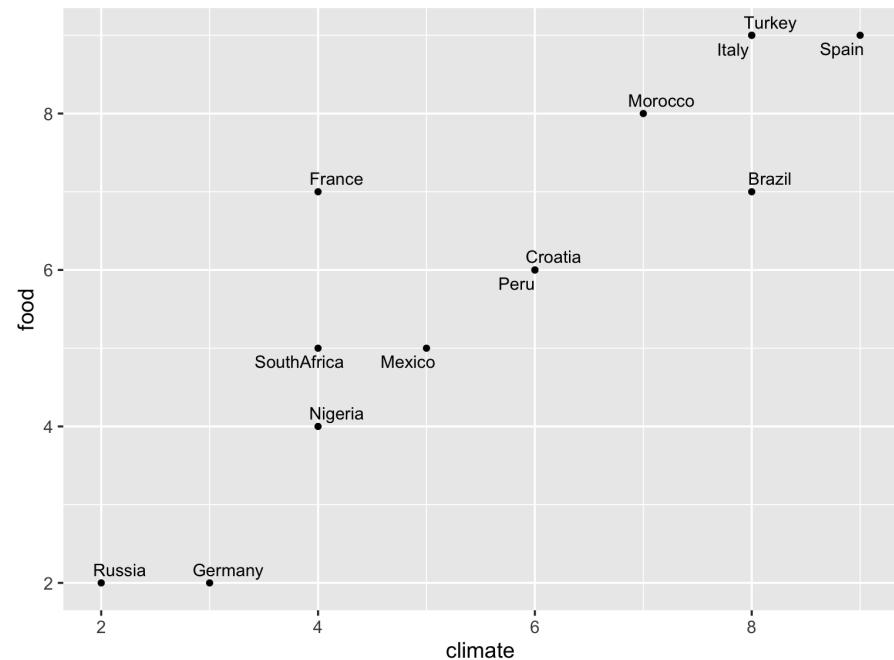
Biplot with calibrated axis and projection lines



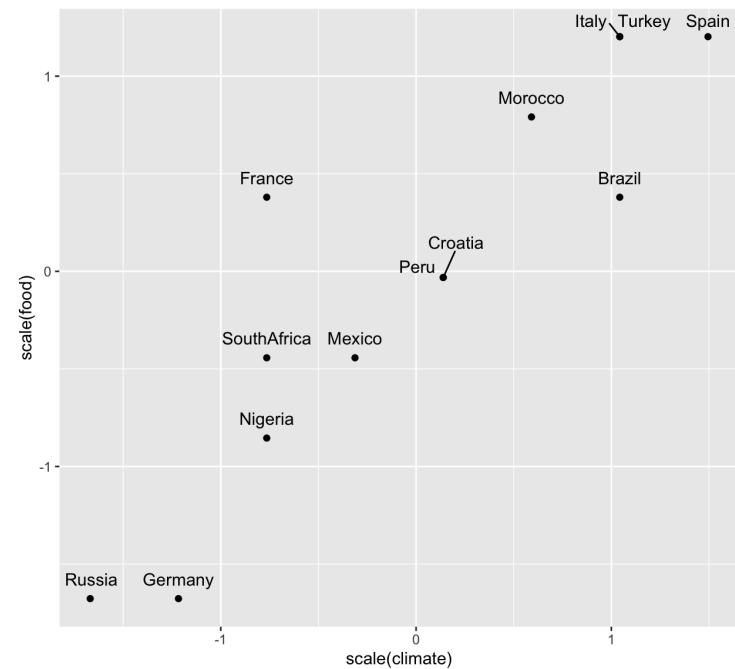
Biplot with calibrated axis and projection lines



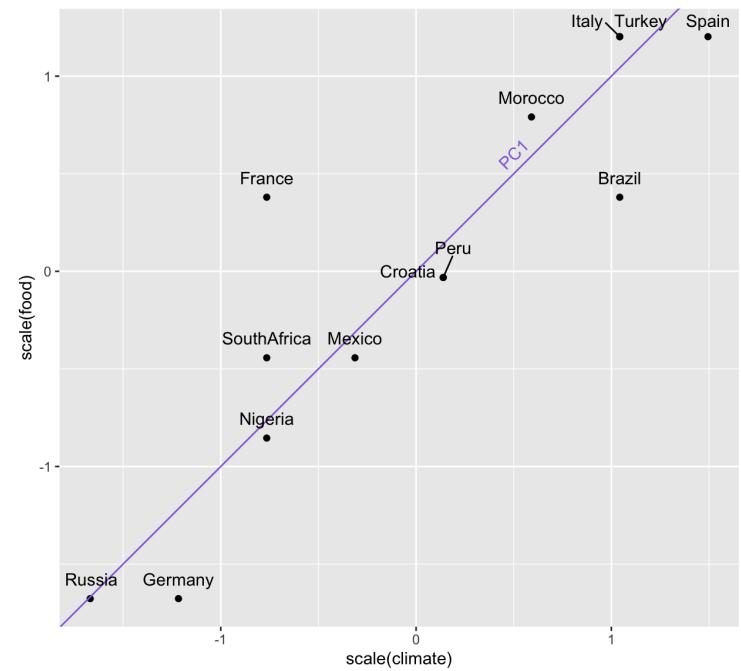
Back to scatterplot (two variables)



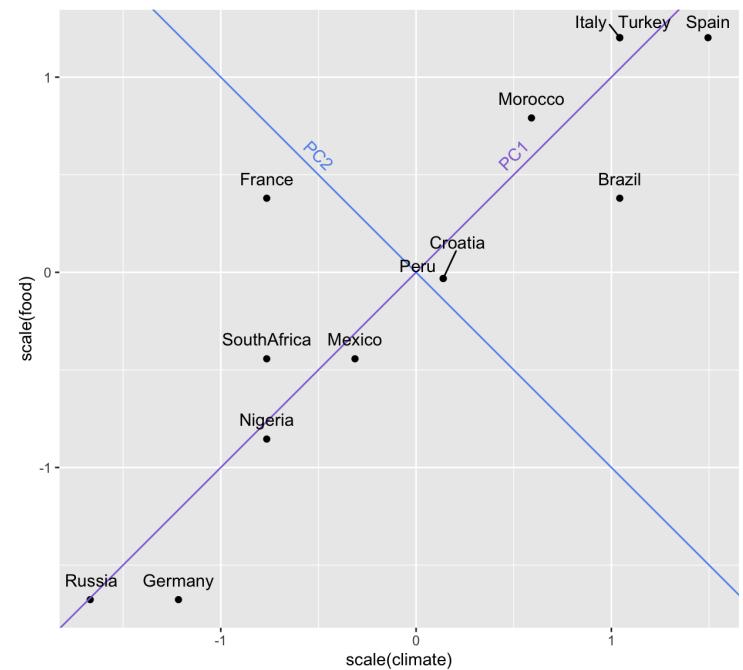
Scatterplot of scaled data



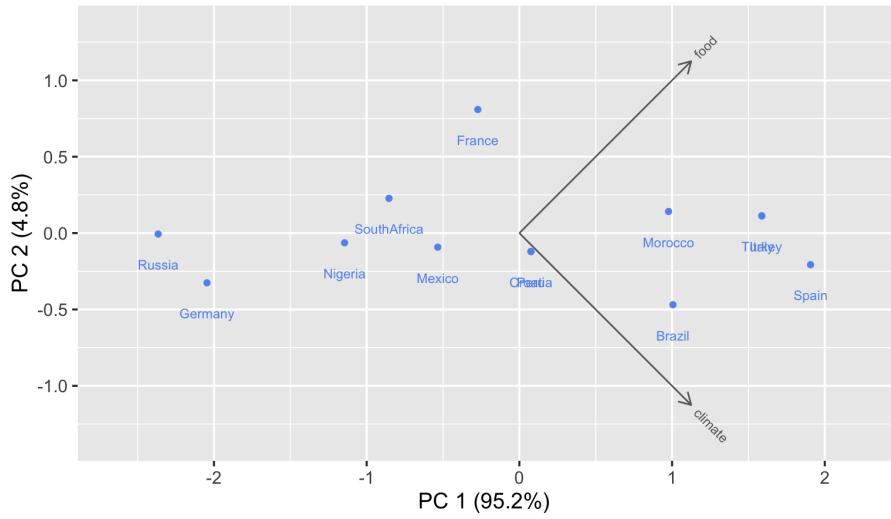
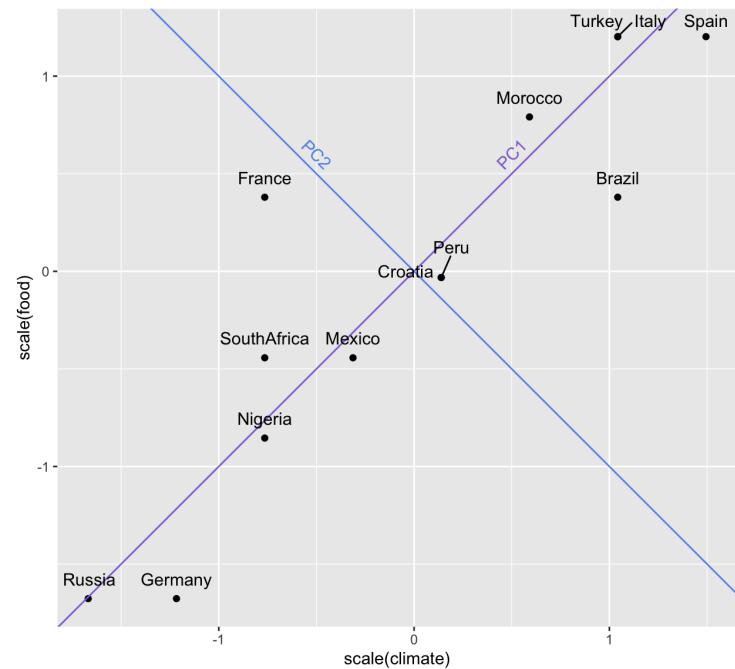
Scatterplot of scaled data



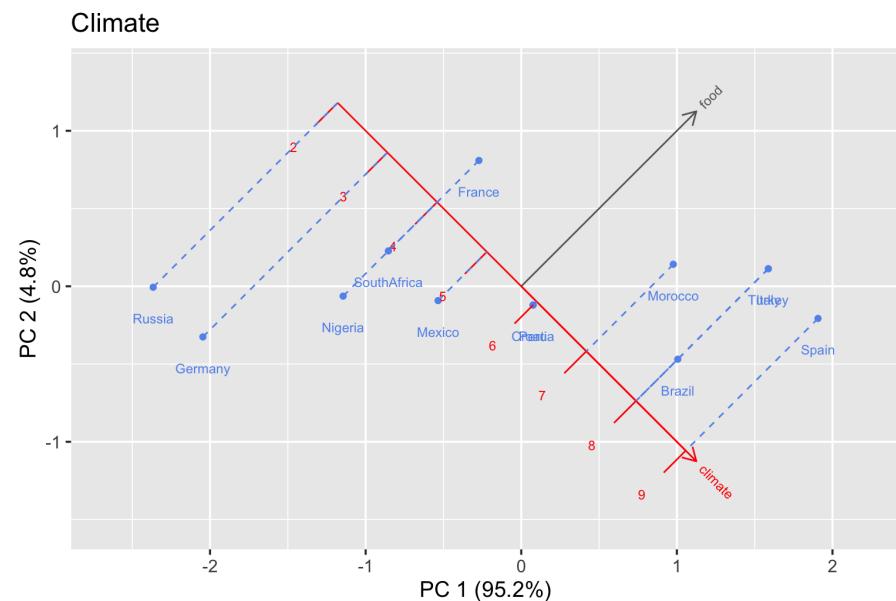
Scatterplot of scaled data



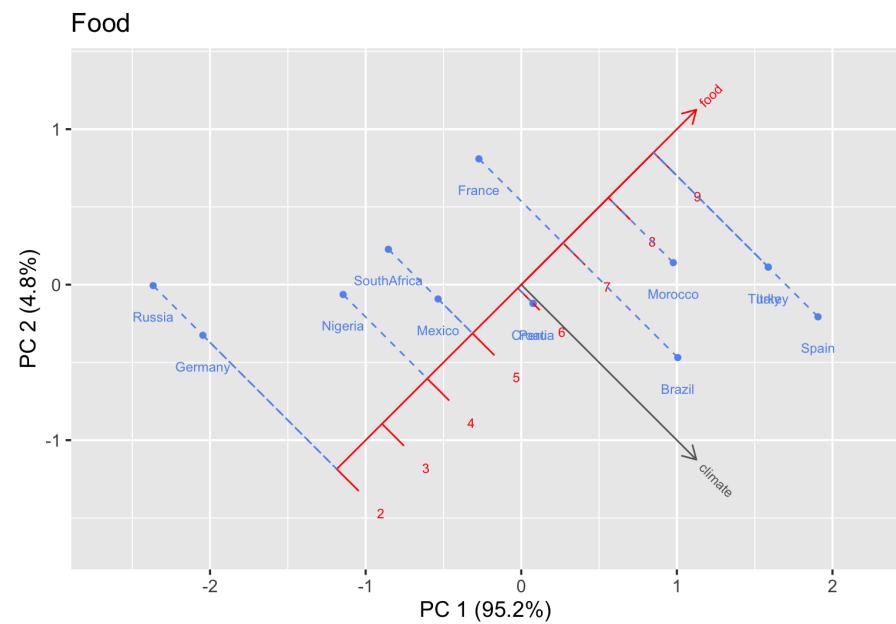
PCA with two variables



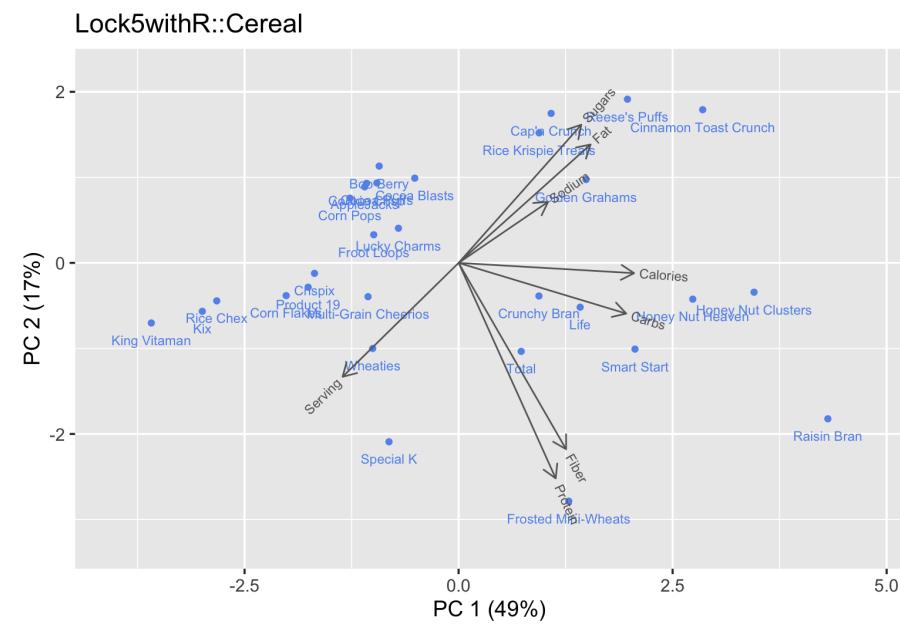
PCA with two variables



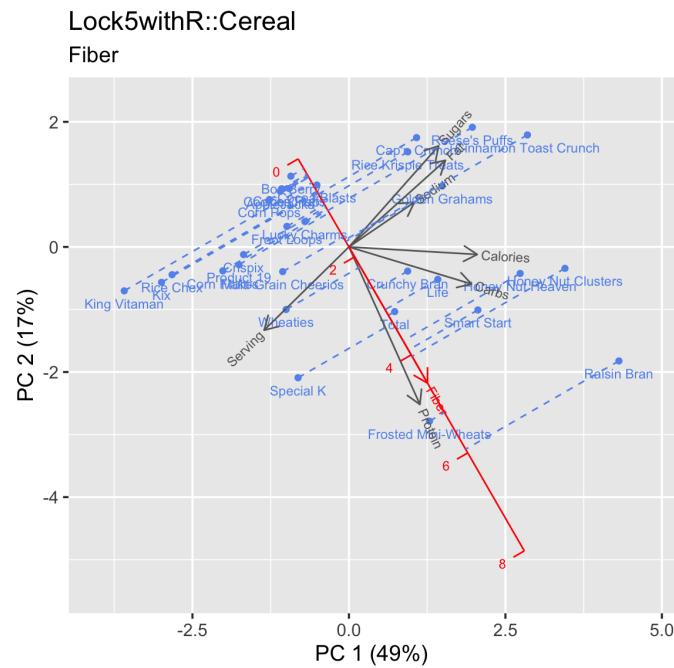
PCA with two variables



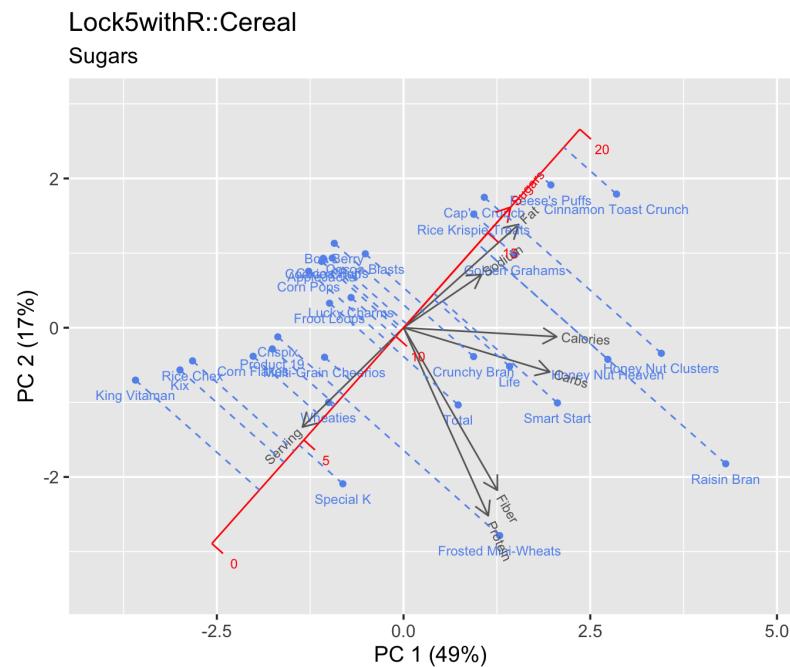
Cereal



Fiber

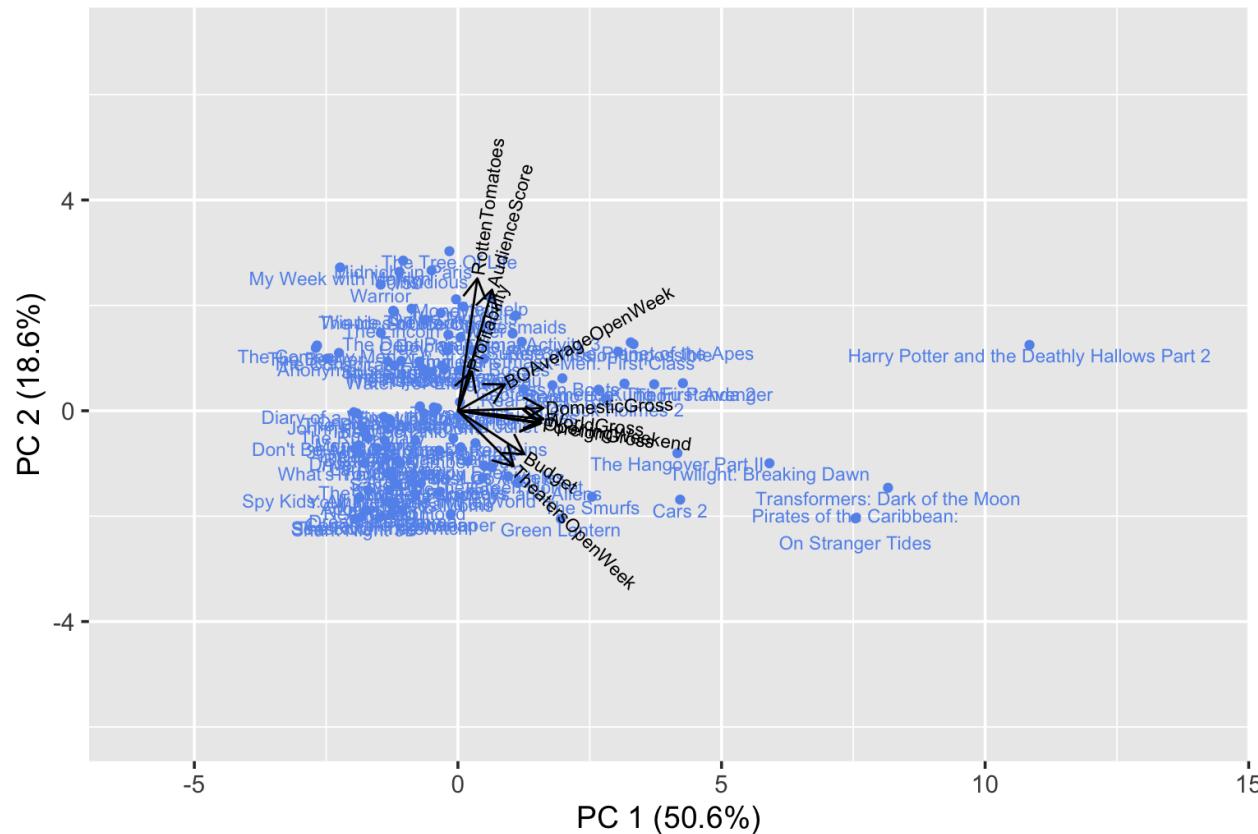


Sugars



Hollywood 2011

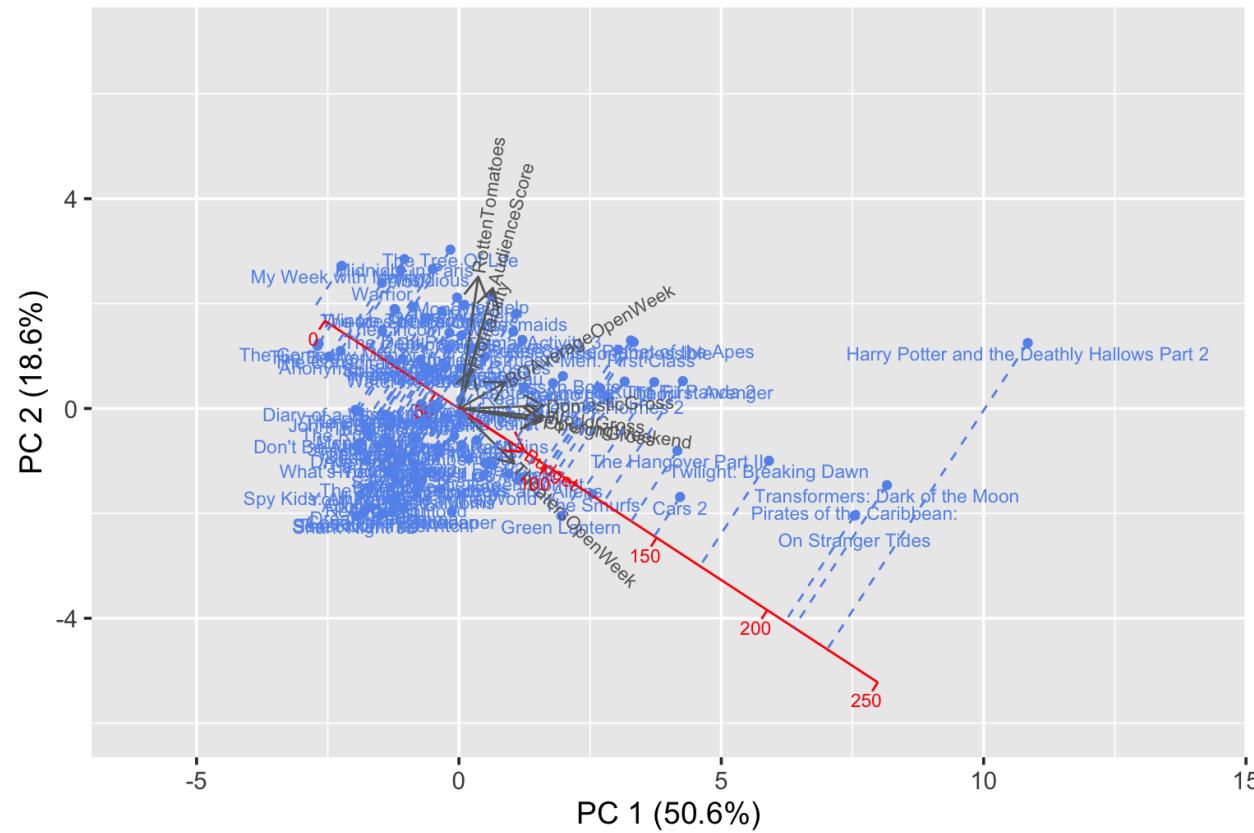
Lock5withR::HollywoodMovies2011



Hollywood 2011

Lock5withR::HollywoodMovies2011

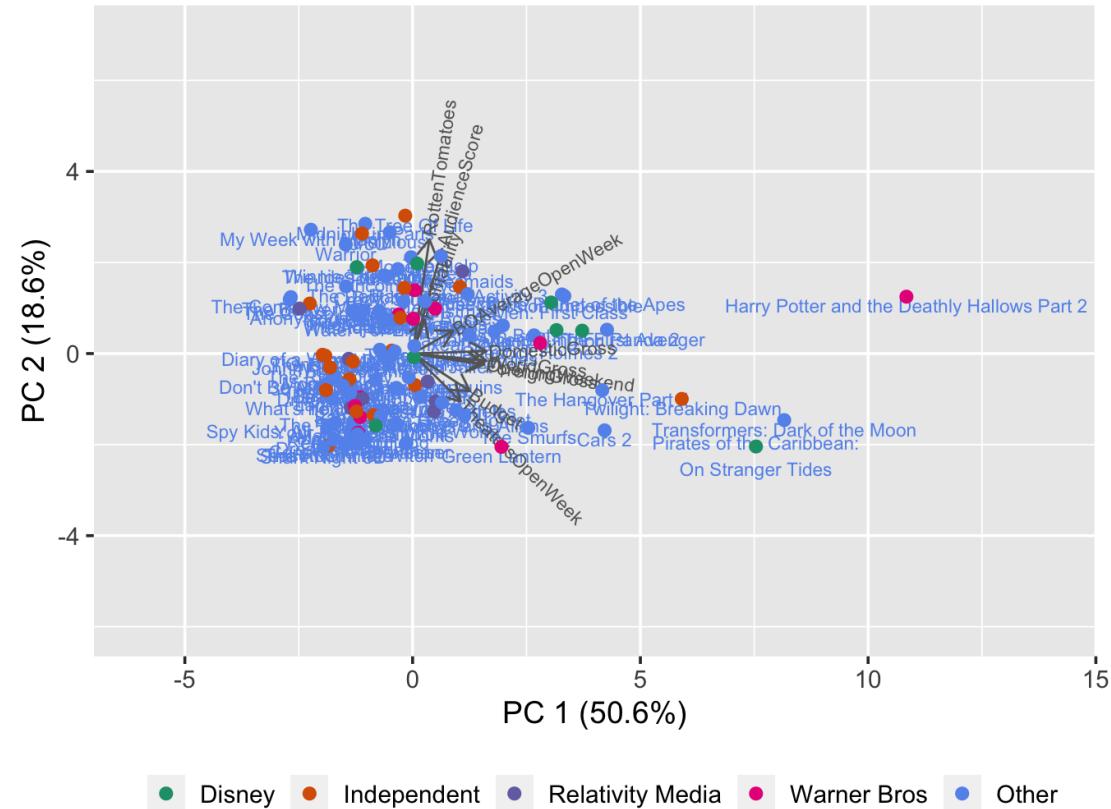
Budget



Hollywood 2011

Lock5withR::HollywoodMovies2011

Lead Studio



Hollywood 2011

```
pr <- prcomp(na.omit(hollywood), scale. = TRUE)
pr$rotation
```

	PC1	PC2	PC3	PC4	PC5
## RottenTomatoes	-0.09740034	0.66364616	-0.19591402	-0.11554808	0.15661675
## AudienceScore	-0.17101525	0.60708105	-0.26340086	-0.15201657	-0.04790960
## TheatersOpenWeek	-0.27930281	-0.27803073	-0.12150764	-0.52944876	0.55208645
## BOAverageOpenWeek	-0.23539966	0.13247814	0.22134113	0.73789025	0.53017695
## DomesticGross	-0.42842553	0.01406371	0.03891897	-0.04302708	-0.09567339
## ForeignGross	-0.41316522	-0.06474888	0.04362208	0.07986600	-0.43213937
## WorldGross	-0.42966005	-0.04169669	0.04337005	0.04332699	-0.33812649
## Budget	-0.33269334	-0.21768140	-0.32175958	-0.01649217	0.25424730
## Profitability	-0.07009632	0.19777849	0.84270016	-0.35765235	0.10675419
## OpeningWeekend	-0.41639386	-0.05953566	0.09752105	0.02114850	-0.03198934
	PC6	PC7	PC8	PC9	
## RottenTomatoes	-0.050239658	0.56193138	-0.39433320	-0.005908157	
## AudienceScore	0.027504670	-0.53761006	0.45241077	-0.113865754	
## TheatersOpenWeek	0.337112564	0.20860345	0.29098428	-0.054563471	
## BOAverageOpenWeek	0.080752520	0.03892043	0.21048241	0.011679654	
## DomesticGross	0.194107111	-0.18851204	-0.22039673	0.787359977	
## ForeignGross	-0.096557510	0.37252598	0.32325076	-0.248230618	
## WorldGross	-0.007521864	0.20592541	0.16074878	0.071702080	
## Budget	-0.791593825	-0.15985166	-0.14521806	-0.015013727	
## Profitability	-0.318381018	-0.06243242	0.03055516	-0.017805948	
## OpeningWeekend	0.318604554	-0.32250629	-0.55670754	-0.544654268	
	PC10				
## RottenTomatoes	-0.0000007526450				
## AudienceScore	0.0000001634105				
## TheatersOpenWeek	-0.0000024778204				
## BOAverageOpenWeek	0.0000009422844				
## DomesticGross	0.2490171774883				
## ForeignGross	0.5620211224982				
## WorldGross	-0.7887475535115				
## Budget	0.0000015252330				
## Profitability	-0.0000002728879				
## OpeningWeekend	0.0000010704130				

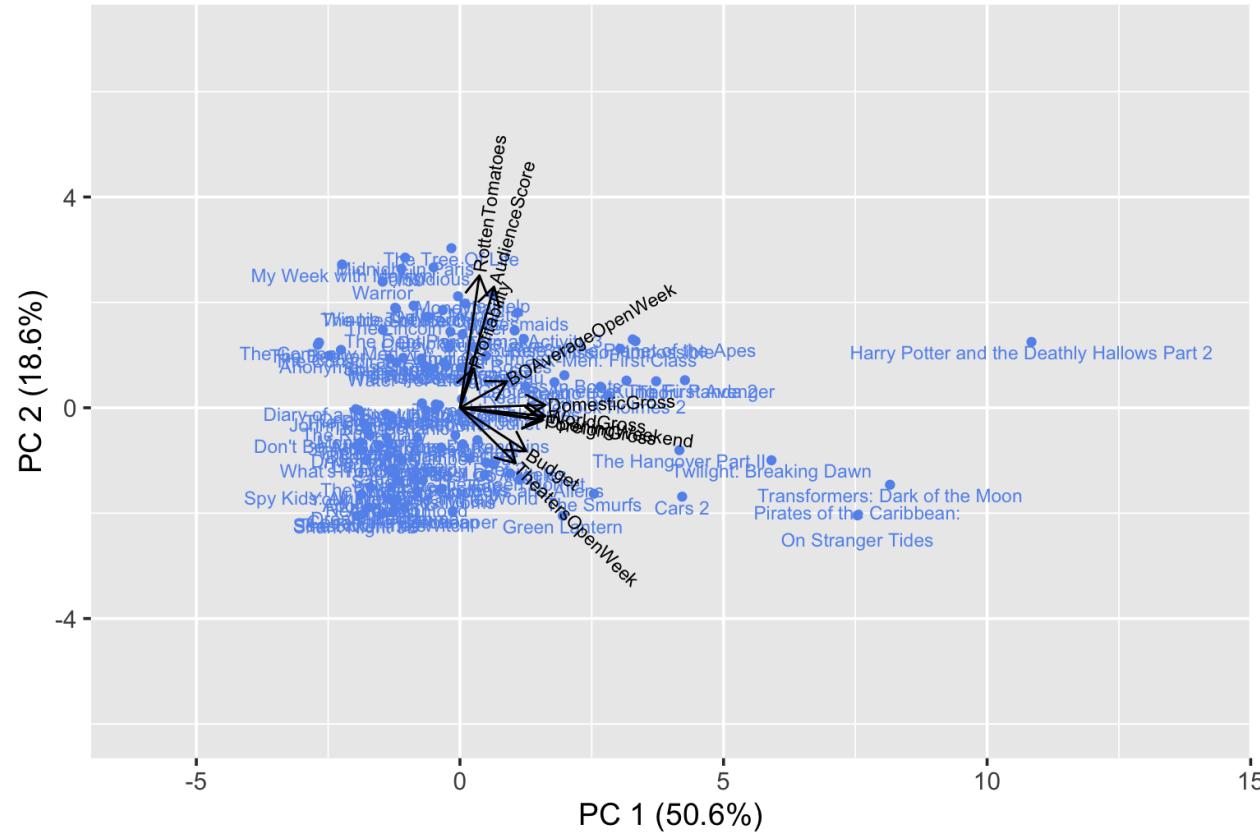
Hollywood 2011

```
prin <- princomp(na.omit(hollywood), cor = TRUE)
prin$loadings
```

```
##  
## Loadings:  
##          Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7 Comp.8  
## RottenTomatoes    0.664  0.196  0.116  0.157           0.562  0.394  
## AudienceScore    0.171  0.607  0.263  0.152          -0.538 -0.452  
## TheatersOpenWeek 0.279 -0.278  0.122  0.529  0.552 -0.337  0.209 -0.291  
## BOAverageOpenWeek 0.235  0.132 -0.221 -0.738  0.530           -0.210  
## DomesticGross     0.428                   -0.194 -0.189  0.220  
## ForeignGross      0.413                   -0.432  0.373 -0.323  
## WorldGross         0.430                   -0.338  0.206 -0.161  
## Budget            0.333 -0.218  0.322       0.254  0.792 -0.160  0.145  
## Profitability     0.198 -0.843  0.358  0.107  0.318  
## OpeningWeekend    0.416                   -0.319 -0.323  0.557  
##          Comp.9 Comp.10  
## RottenTomatoes    0.114  
## AudienceScore     0.114  
## TheatersOpenWeek  
## BOAverageOpenWeek  
## DomesticGross     -0.787 -0.249  
## ForeignGross       0.248 -0.562  
## WorldGross         0.789  
## Budget  
## Profitability  
## OpeningWeekend    0.545  
##  
##          Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7 Comp.8 Comp.9  
## SS loadings       1.0    1.0    1.0    1.0    1.0    1.0    1.0    1.0    1.0  
## Proportion Var    0.1    0.1    0.1    0.1    0.1    0.1    0.1    0.1    0.1  
## Cumulative Var   0.1    0.2    0.3    0.4    0.5    0.6    0.7    0.8    0.9  
##          Comp.10  
## SS loadings       1.0  
## Proportion Var    0.1  
## Cumulative Var   1.0
```

Hollywood 2011

Lock5withR::HollywoodMovies2011



Hollywood 2011

Lock5withR::HollywoodMovies2011

