

Vital Signs Report

Robert Jackson

5/31/2018

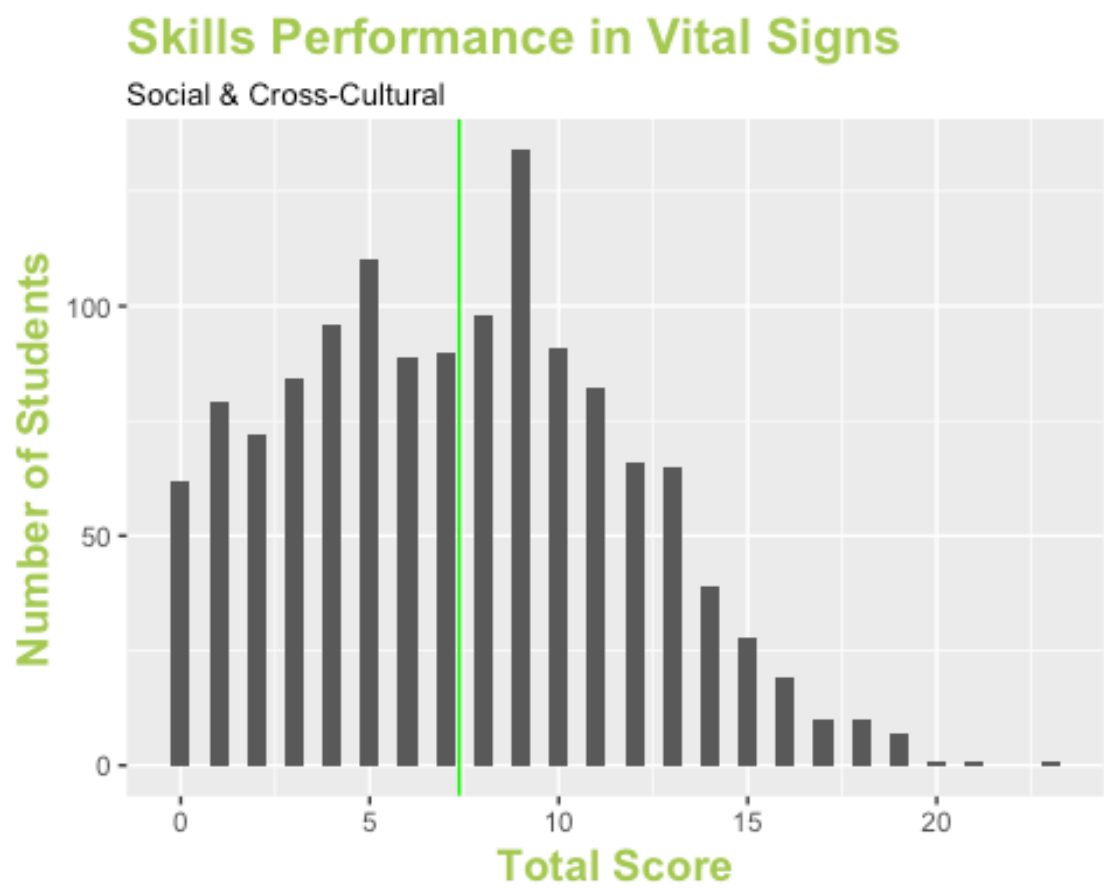
Part 1:

```
VS_Skills2 <- VS_Skills2[!duplicated(VS_Skills2), ]
```

Summary Statistics of Vital Signs - 2017 - 18 [table 1]

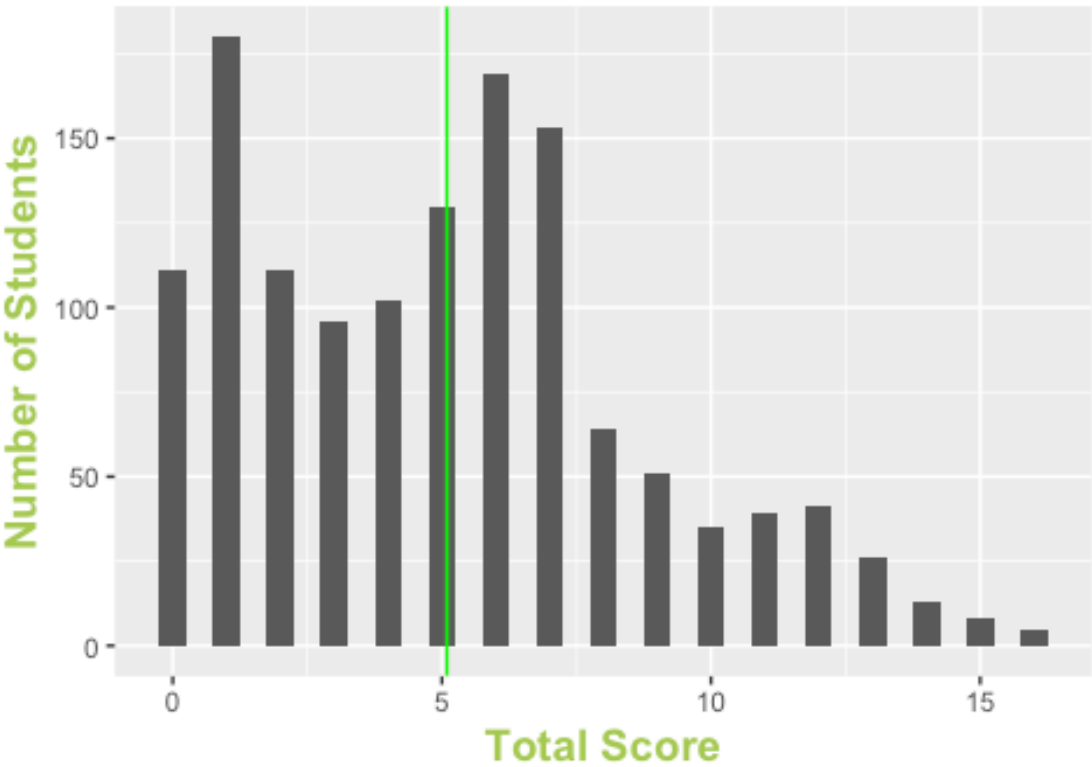
##	Social	Productivity	Leadership	Initiative	Flexibility
## Mean	7.368066	5.093703	8.412294	3.336582	4.829085
## SD	4.422737	3.645105	4.748217	2.567597	5.218485
## Median	7.000000	5.000000	7.000000	2.000000	2.000000
## Min	0.000000	0.000000	0.000000	0.000000	0.000000
## Max	23.000000	16.000000	24.000000	13.000000	22.000000
## N	1334.000000	1334.000000	1334.000000	1334.000000	1334.000000
##	Total.Good.Decisions	episodes	Grades		
## Mean	29.28936	5.845577	7.967016		
## SD	19.17128	3.777394	2.446509		
## Median	25.00000	5.000000	7.000000		
## Min	1.00000	1.000000	3.000000		
## Max	91.00000	12.000000	13.000000		
## N	1334.00000	1334.000000	1334.000000		

Histograms of Skill Point Distribution [Cumulative]



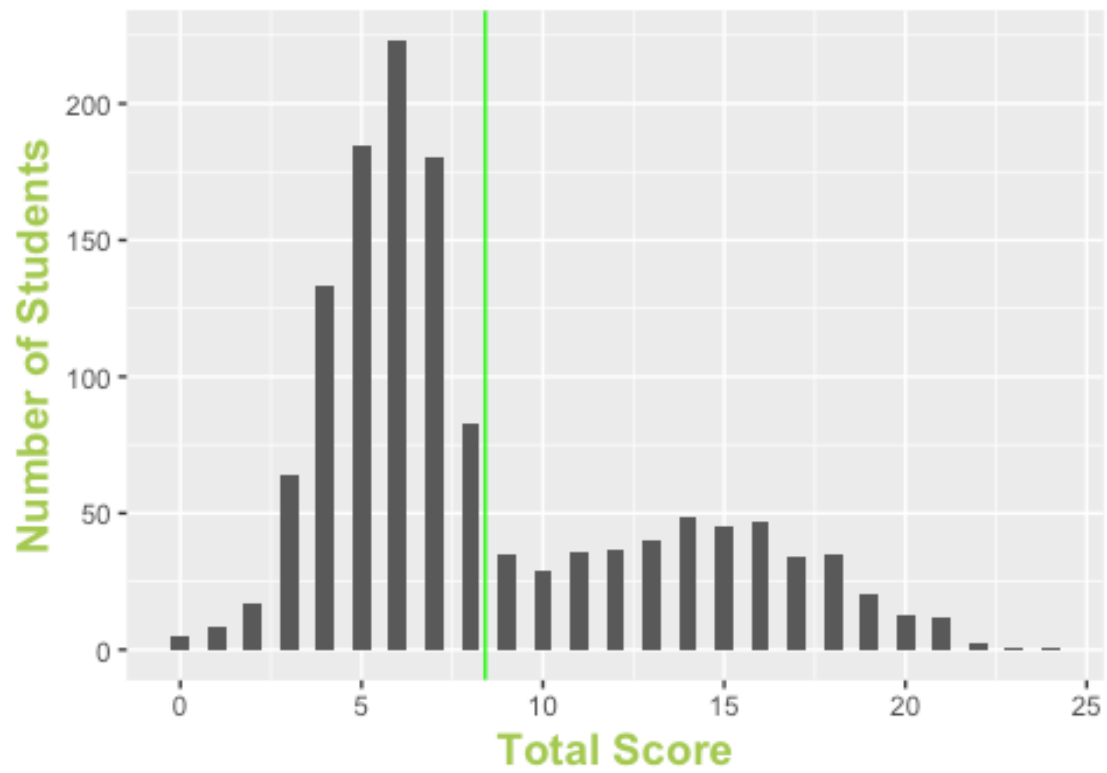
Skills Performance in Vital Signs

Productivity & Accountability



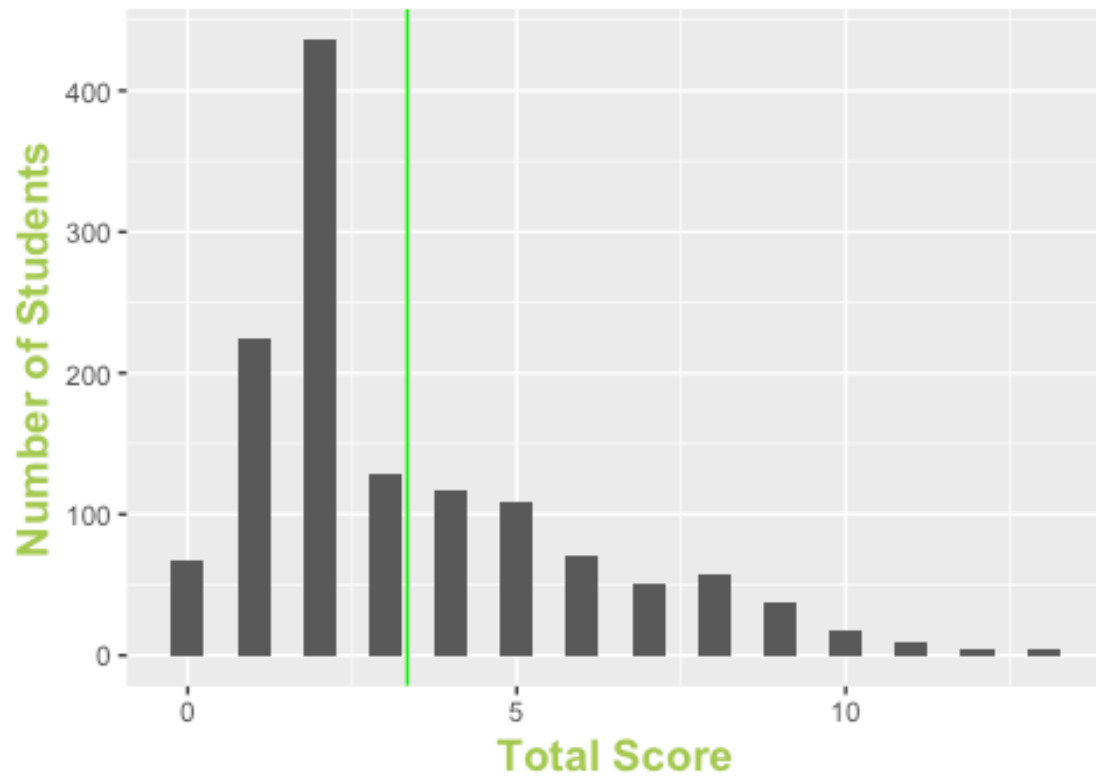
Skills Performance in Vital Signs

Leadership & Responsibility



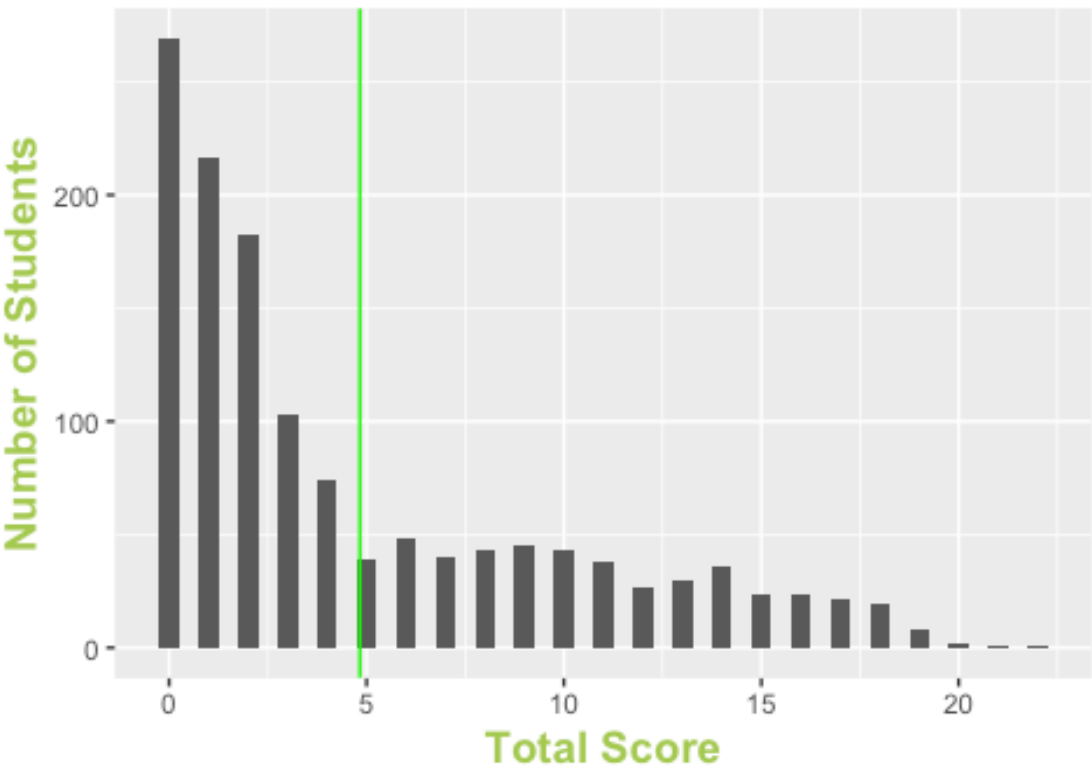
Skills Performance in Vital Signs

Initiative & Self-Direction



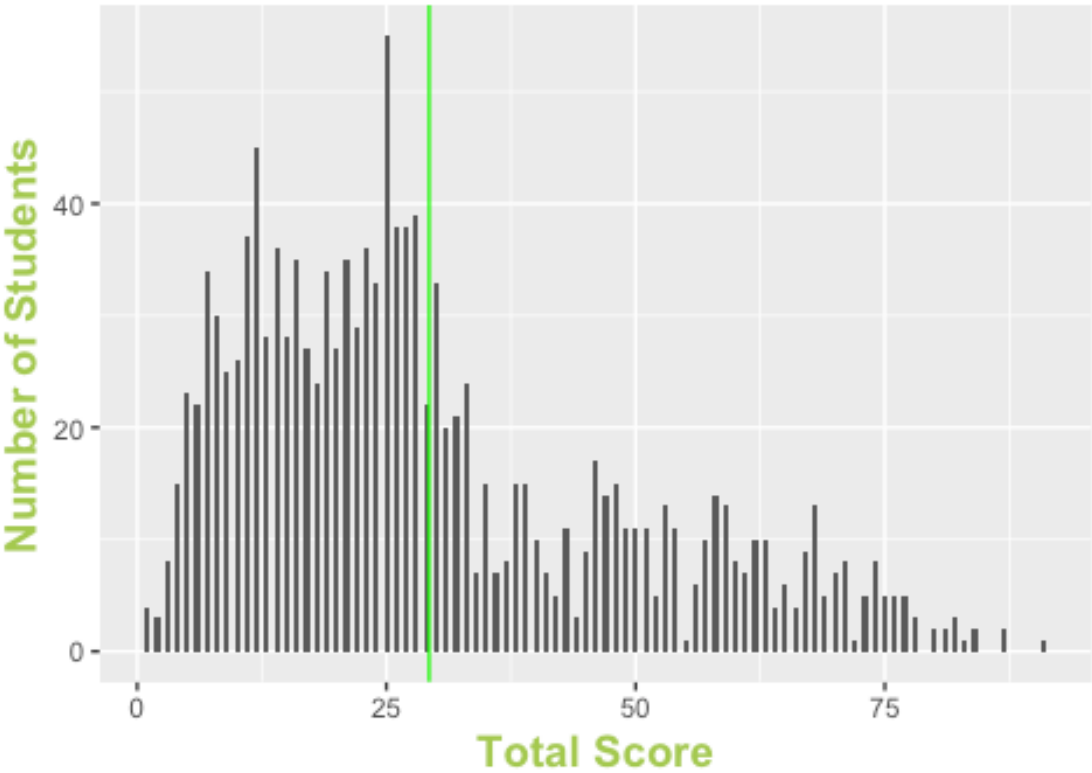
Skills Performance in Vital Signs

Flexibility & Adaptability

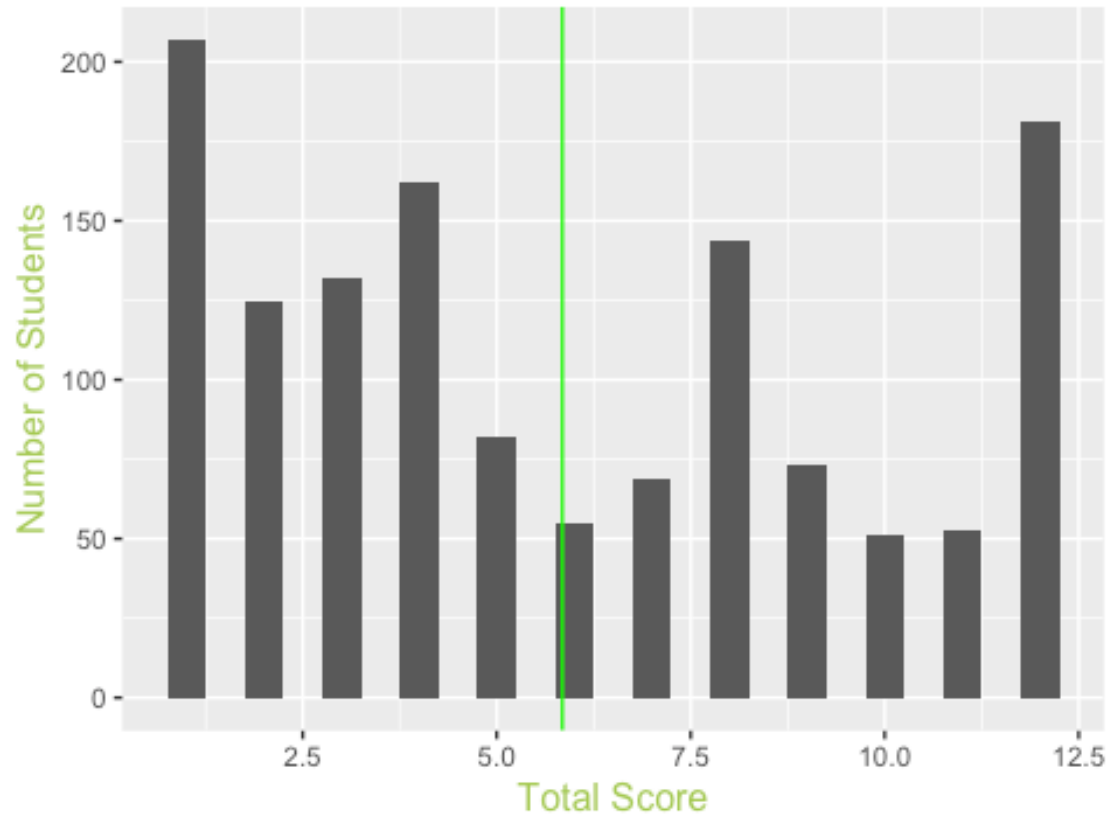


Performance in Vital Signs

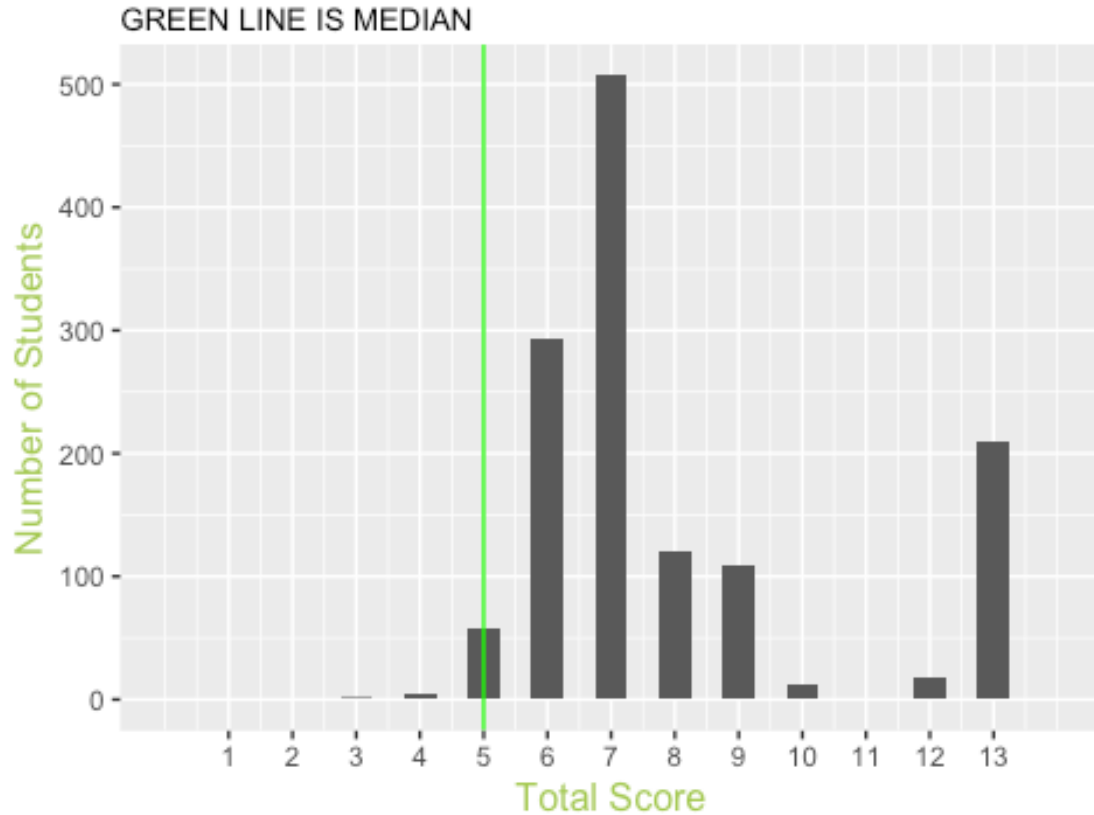
Total Good Decision



Distribution of Episodes Played in Vital Signs



Distribution of Student Grade Levels who Play V



Summary Statistics of Skill Point Distribution [By Percentage] *The additional variables takes into account the particular episodes played by students and which points were available for them to score.*

Motified Summary Statistics of Vital Signs Including Percentage - 2017 - 18 [table 2]

##	SocialR	ProdR	LeadR	InitiativeR
FlexR				
## Mean	0.6292242	0.5091815	0.7020609	0.6369561
0.4372367				
## SD	0.2087652	0.2122707	0.1572709	0.2575319
0.2523438				
## Median	0.6666667	0.5263158	0.7037037	0.6000000
0.4666667				
## Min	0.0000000	0.0000000	0.0000000	0.0000000
0.0000000				
## Max	1.0000000	1.0000000	1.0000000	1.0000000
1.0000000				
## N	1334.0000000	1334.0000000	1334.0000000	1334.0000000
1334.0000000				
##	Points_Earned_Social	Points_Earned_Productivity		
## Mean	8.886807	6.433283		

```

## SD          5.143690          4.246314
## Median      9.000000          6.000000
## Min         0.000000          0.000000
## Max         23.000000         17.000000
## N           1334.000000        1334.000000
##           Points_Earned_Leadership Points_Earned_Flexibility
## Mean              10.259370          6.931034
## SD                5.494747          6.106559
## Median            8.000000          6.000000
## Min              0.000000          0.000000
## Max              25.000000         23.000000
## N                1334.000000        1334.000000
##           Points_Earned_Initiative Points_Available_Social
## Mean              4.639430         13.670915
## SD                3.272056          6.466866
## Median            4.000000         14.000000
## Min              0.000000          0.000000
## Max              13.000000         23.000000
## N                1334.000000        1334.000000
##           Points_Available_Productivity Points_Available_Leadership
## Mean              12.047976         15.104948
## SD                6.874228          8.151662
## Median            10.000000         11.000000
## Min              1.000000          2.000000
## Max              22.000000         27.000000
## N                1334.000000        1334.000000
##           Points_Available_Flexibility Points_Available_Initiative
## Mean              13.108696          8.178411
## SD                8.815575          5.788439
## Median            12.000000          8.000000
## Min              1.000000          0.000000
## Max              26.000000         18.000000
## N                1334.000000        1334.000000

```

Histogram of Skill Point Distribution [By Percentage]

This takes into account the particular episodes played by students and which points were available for them to score. The distribution is now a scale of 0 - 100%.

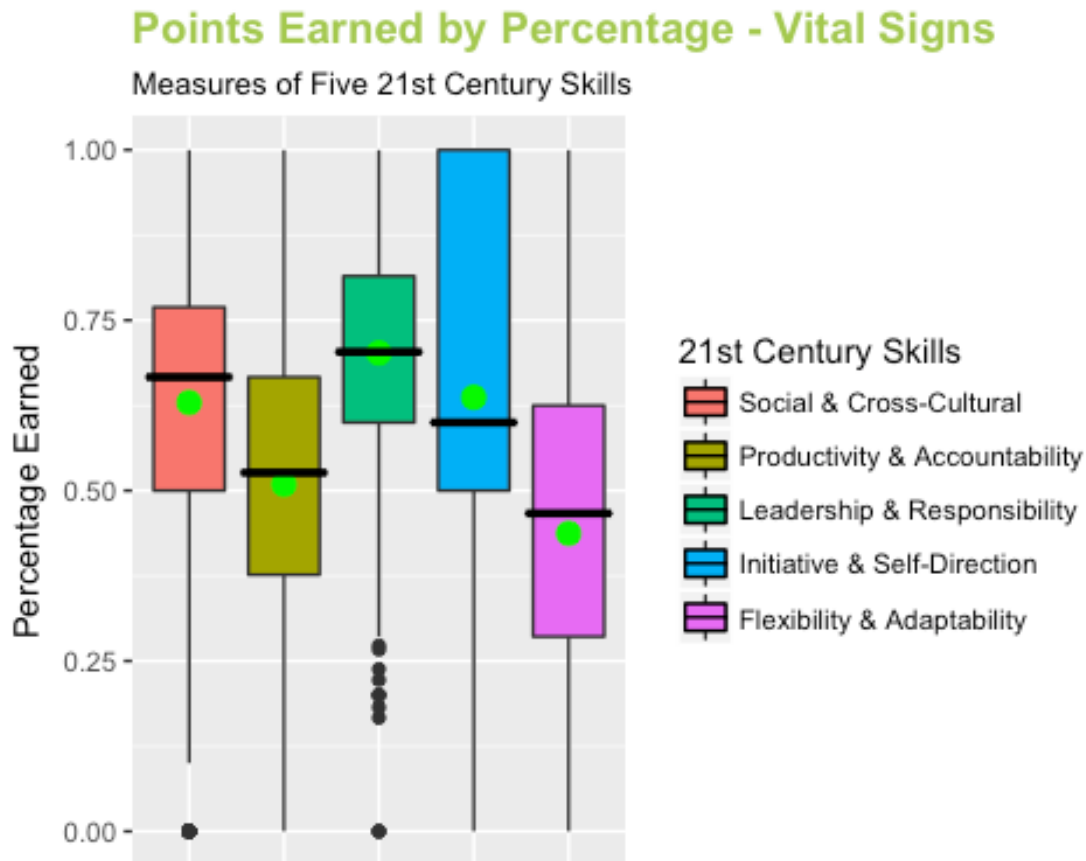
Meaning student could have earned 0% of the points available to them based on the episodes they played or they may have earned 100% of the point available to them in the episodes they've played.

Box Plot of 21st Century Skills by Percentage

```
## No id variables; using all as measure variables
```

```
## Warning: Removed 2 rows containing non-finite values (stat_boxplot).
```

```
## Warning: Removed 2 rows containing non-finite values (stat_summary).
## Warning: Removed 2 rows containing non-finite values (stat_summary).
```



We'll use this next section to outline how many student had an opportunity to perform each of the five 21st century skills

```
D1 <- dplyr::select(VSPer_Skills, contains("possible_"))
## No id variables; using all as measure variables
```

This boxplot is dynamic and analyzes what episodes of the game students reach and whether they encounter an opportunity to demonstrate one of the five 21st century skills. Since these skills don't appear evenly throughout the game its important to know in which ways students are not being exposed to

certain skill opportunities.

Range of Points Available - Vital Signs

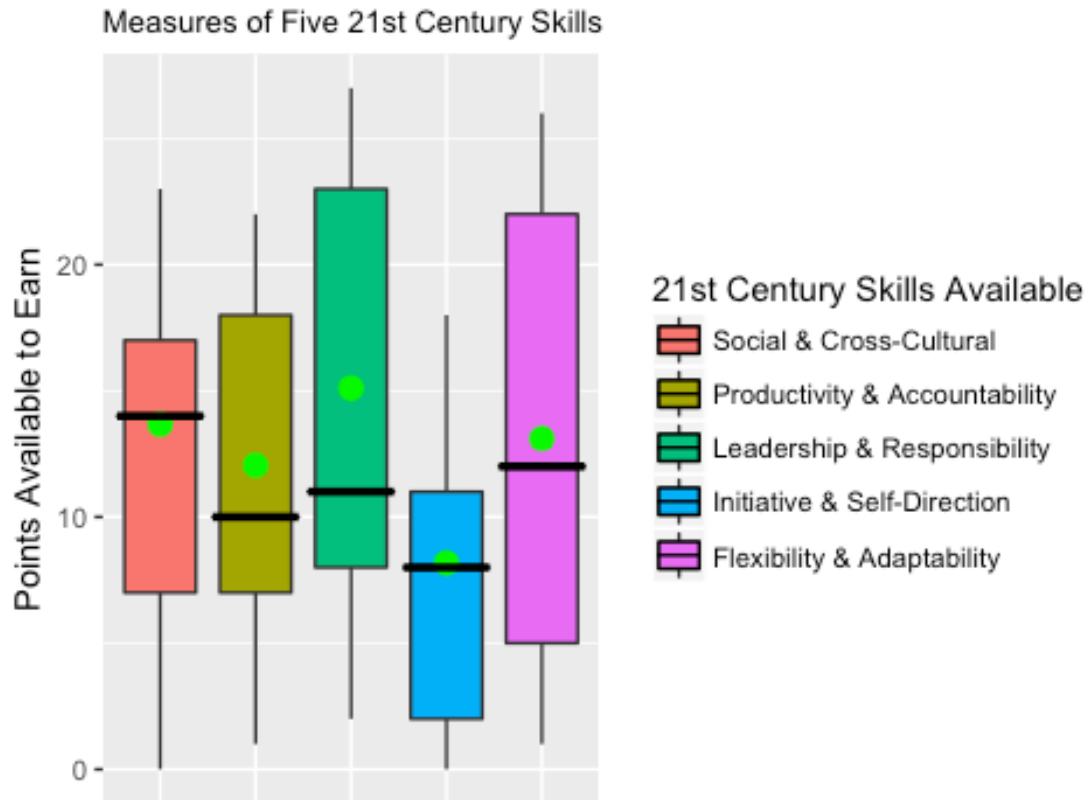


Table of Points Available to Earn - Vital Signs

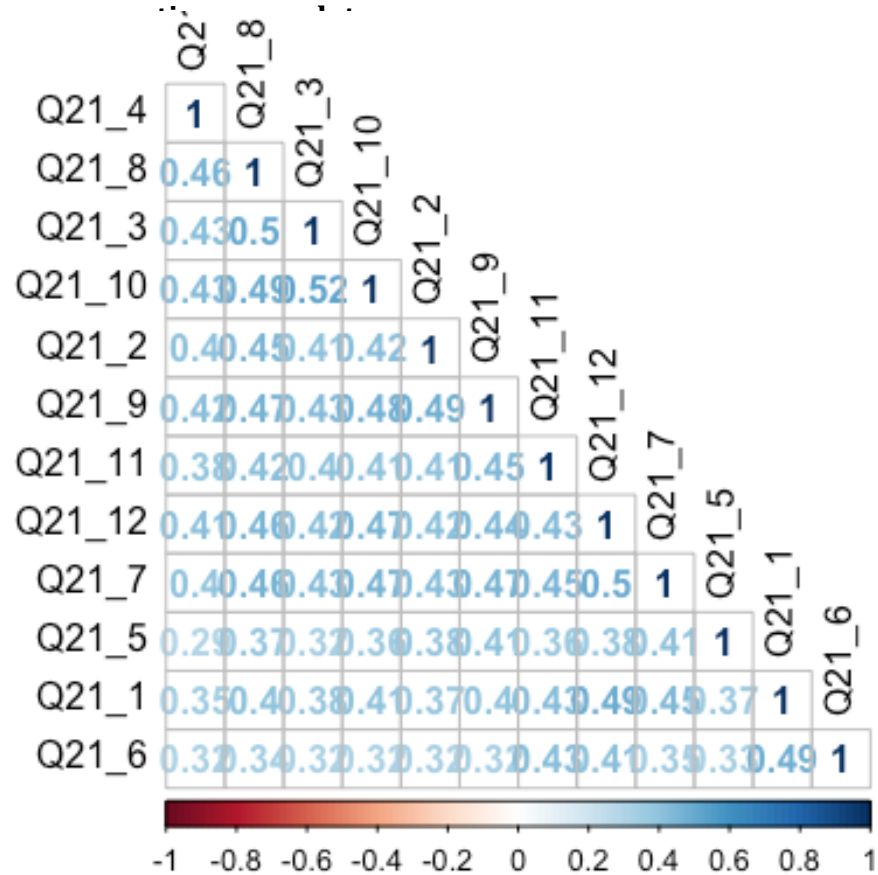
These are the count of students who have encountered an opportunity to demonstrate a particular 21st century skill.

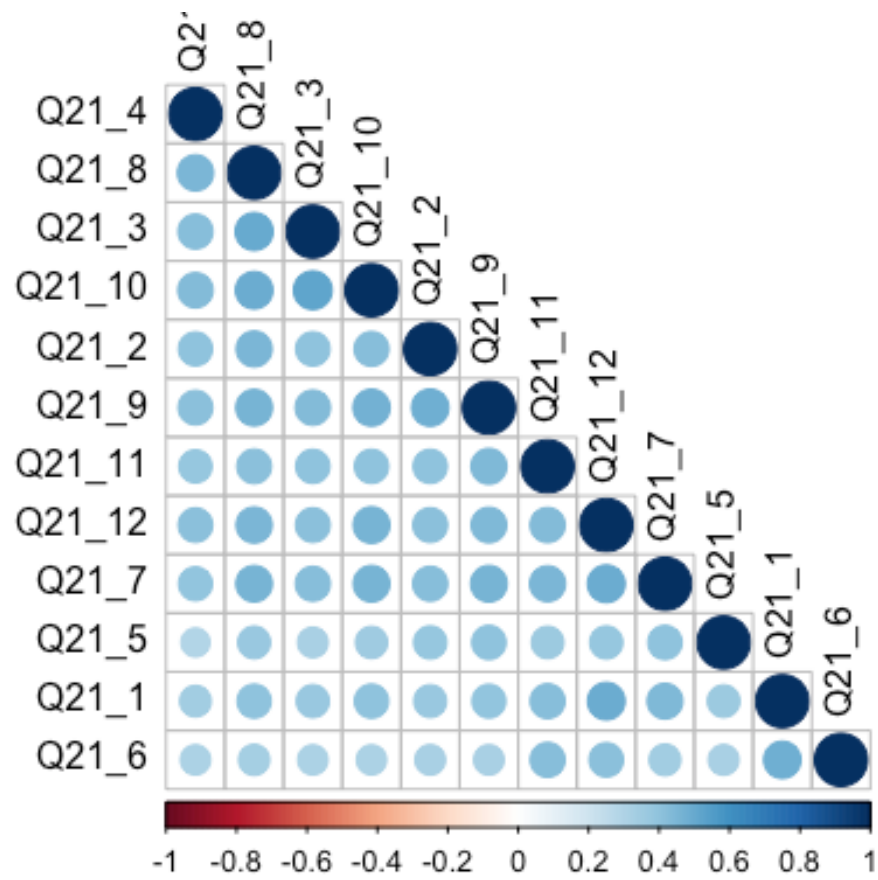
```
##      Points_Available_Social Points_Available_Productivity
##                               1333                               1334
##      Points_Available_Leadership Points_Available_Initiative
##                               1334                               1333
##      Points_Available_Flexibility
##                               1334
```

Part Two:

We're now going to add the survey data. Go through the same EDA. Then finally we'll look at the correlation between the skills in VS and the measures in the survey data.

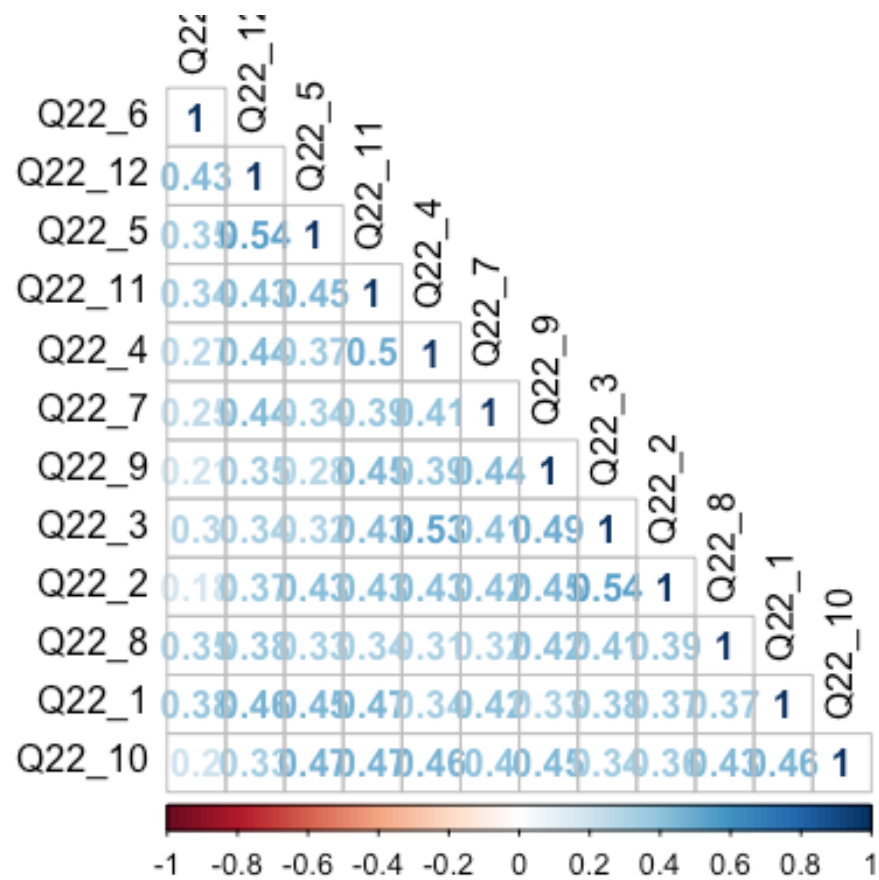
Working through Survey 1

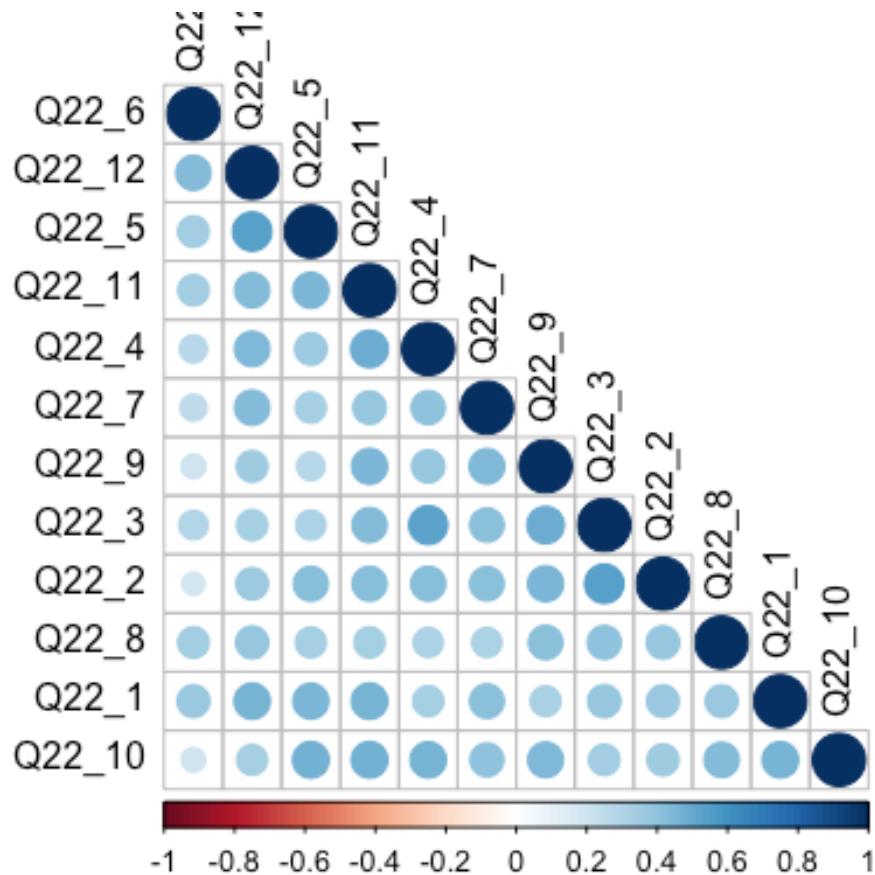




Survey 2

Corplot of the Milwaukee Survey Questions





Part Three: Creating the 2 cumulative measures from the External Instruments & Time to merge External and internal metrics

Renaming the Variables for ease of use.

```
SurveyM2 <- dplyr::rename(
  SurveyM,
  "Users" = "Your Game Username",
  "Q1" = "I can set an example for others.",
  "Q2" = "I can contribute to group efforts.",
  "Q3" = "I can take an active role in learning.",
  "Q4" = "I compare possible solutions to find the best one.",
  "Q5" = "I can do things on my own.",
  "Q6" = "I usually have more than one source of information before
making a decision.",
  "Q7" = "I take steps to accomplish goals.",
  "Q8" = "I can be in charge of a group of classmates.",
```

```
"Q9" = "I compare different ideas when thinking about a topic.",
"Q10" = "When faced with a decision, I realize that some choices are
better than others.",
"Q11" = "I use what I know already to solve new problems.",
"Q12" = "I think about all the information I have about making new
decisions."
)
```

Summary Statistics of External Measures - 2017 - 18 [table 3]

LCSDI	CTDMPS	L_C	SD_I	CT
Mean 17.003609	16.675448	7.465948	7.743241	4.668354
SD 3.186401	3.230627	1.678105	1.567088	1.128781
Median 17.666667	17.500000	8.000000	8.000000	4.500000
Min 5.166667	5.166667	2.333333	2.333333	1.500000
Max 20.666667	20.666667	9.333333	9.333333	6.000000
N 2150.000000	2150.000000	2150.000000	DM PS	Mean 4.853287
SD 1.081932	1.089240	Median 5.000000	5.000000	Min 1.500000
Max 6.000000	6.000000	N 2150.000000	2150.000000	

```
write.csv(MasterFrame1, file = "AllVariables.VS.6.5.2018.csv")
```

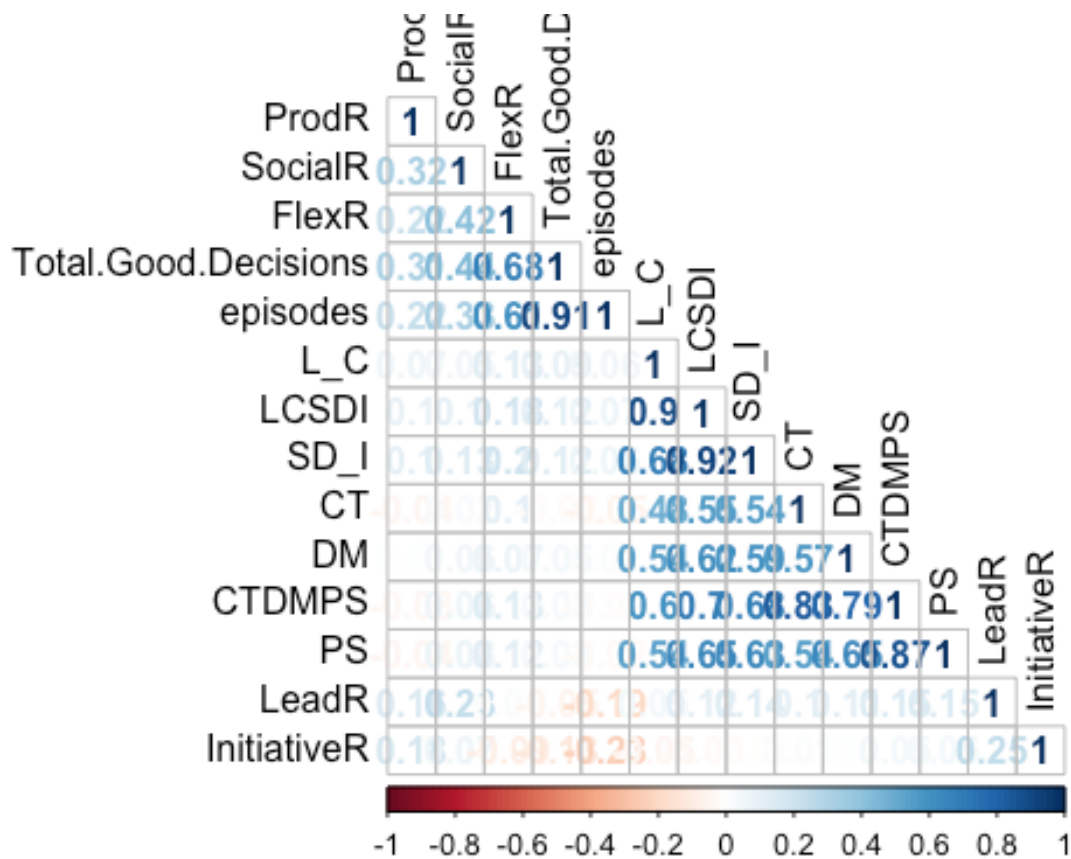
```
MasterFrame1 <- dplyr::select(MasterFrame1, Users, Total.Good.Decisions,
episodes, SocialR, ProdR, LeadR, InitiativeR, FlexR, Response_ID, Student, LCSDI,
CTDMPS, L_C, SD_I, CT, DM, PS)
```

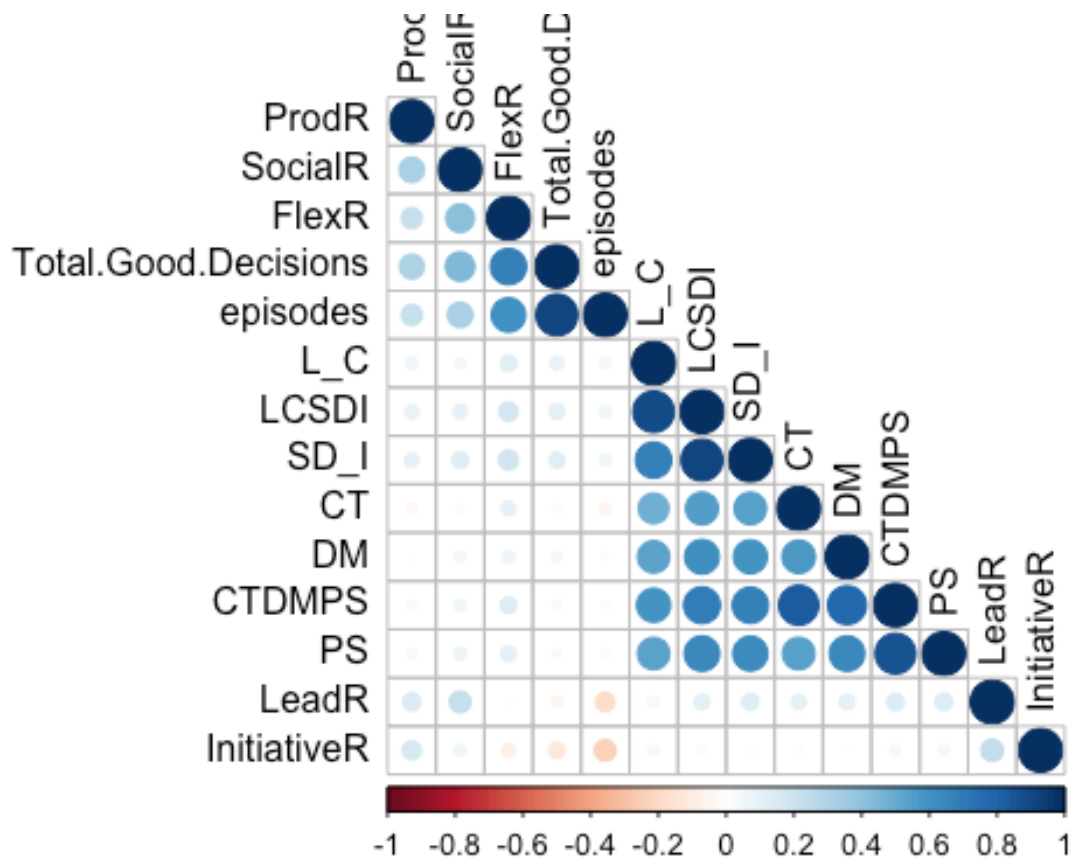
Summary Statistics of Vital Signs & Survey Data 2017 - 2018 [table 4]

Points_Earned_Social	Points_Available_Social
Mean 9.779412	14.51961
SD 5.701124	6.91008
Median 10.000000	15.00000
Min 0.000000	1.00000
Max 22.000000	23.00000
N 612.000000	612.00000
Points_Earned_Productivity	Points_Available_Productivity
Mean 7.055556	12.973856
SD 4.469228	7.236204
Median 7.000000	12.000000
Min 0.000000	2.000000
Max 17.000000	22.000000
N 612.000000	612.000000
Points_Earned_Leadership	Points_Available_Leadership
Mean 11.207516	16.344771
SD 5.982332	8.652621
Median 9.500000	17.000000
Min 2.000000	5.000000
Max 25.000000	27.000000
N 612.000000	612.000000
Points_Earned_Initiative	Points_Available_Initiative
Mean 5.351307	9.367647
SD 3.580148	6.106834
Median 5.000000	9.000000
Min 0.000000	1.000000
Max 13.000000	18.000000
N 612.000000	612.000000
Points_Earned_Flexibility	Points_Available_Flexibility
Mean 8.173203	14.596405
SD 6.648732	9.159552
Median 8.000000	14.000000
Min 0.000000	1.000000
Max 23.000000	26.000000
N 612.000000	612.000000
Total.Good.Decisions	episodes
Grades	SocialR
ProdR	Mean 35.69444
Mean 7.223856	0.6460814
SD 0.5223357	SD 20.63714
4.023458	1.709860
0.2003075	0.2019923
Median 31.00000	6.000000
7.000000	0.6666667
0.5454545	Min 2.00000
1.000000	5.000000
0.000000	0.000000
0.000000	Max 87.00000
12.000000	13.000000
1.000000	1.000000
N 612.00000	612.000000
612.000000	612.000000
LeadR	InitiativeR
FlexR	LCSDI
CTDMPS	Mean 0.7032195
SD 0.1426991	0.2417112

0.2493611 3.124828 3.102452 Median 0.7037037 0.5833333 0.5000000
 18.333333 17.500000 Min 0.2000000 0.0000000 0.0000000 5.166667 5.166667
 Max 1.0000000 1.0000000 1.0000000 20.666667 20.666667 N 612.0000000
 612.0000000 612.0000000 612.0000000 612.0000000 L_C SD_I CT DM PS Mean
 7.624339 7.872114 4.723684 4.944542 5.095238 SD 1.623520 1.531734 1.100192
 1.023966 1.055322 Median 8.000000 8.333333 5.000000 5.000000 5.500000 Min
 2.333333 2.333333 1.500000 1.500000 1.500000 Max 9.333333 9.333333
 6.000000 6.000000 6.000000 N 612.000000 612.000000 612.000000 612.000000
 612.000000

Correlation plot of the External measures and the Internal measures from VS





Part 4: PCA/ Factor Analysis

```
M3 <- dplyr::select(M2, -Total.Good.Decisions, - episodes)
```

```
## [1] "sdev"      "rotation" "center"   "scale"    "x"
```

```
##          PC1          PC2          PC3          PC4
```

```
PC5
```

```
## SocialR    -0.063957141 -0.583574506  0.154750068  0.2861317
```

```
0.08498418
```

```
## ProdR      -0.029387599 -0.530364323 -0.005369975 -0.4710536 -
```

```
0.22626496
```

```
## LeadR      -0.081233370 -0.331009163 -0.506083973  0.2151349
```

```
0.68128738
```

```
## InitiativeR -0.007073981 -0.230435134 -0.655204124 -0.1645391 -
```

```
0.50616687
```

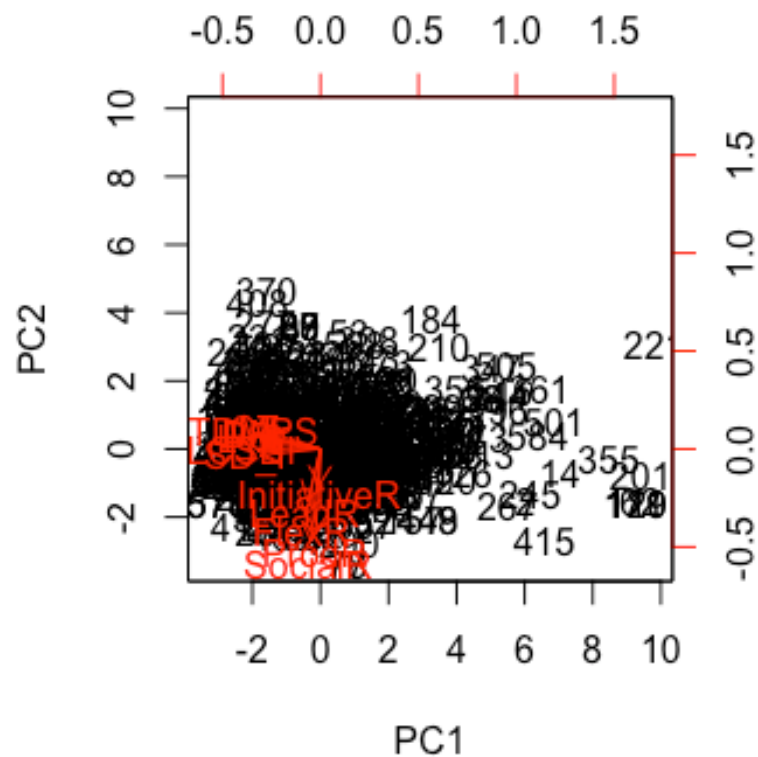
```
## FlexR      -0.095295922 -0.425240828  0.473634440  0.2955227 -
```

```
0.21948292
```

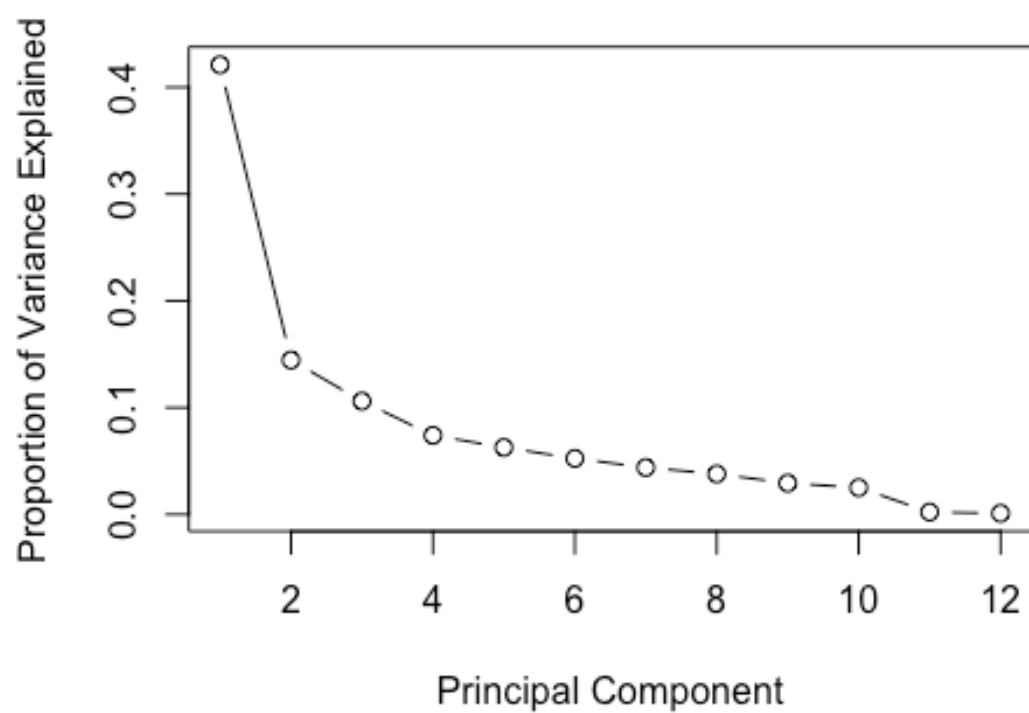
```
## LCSDI      -0.403524618 -0.008444935  0.108484343 -0.3404451
```

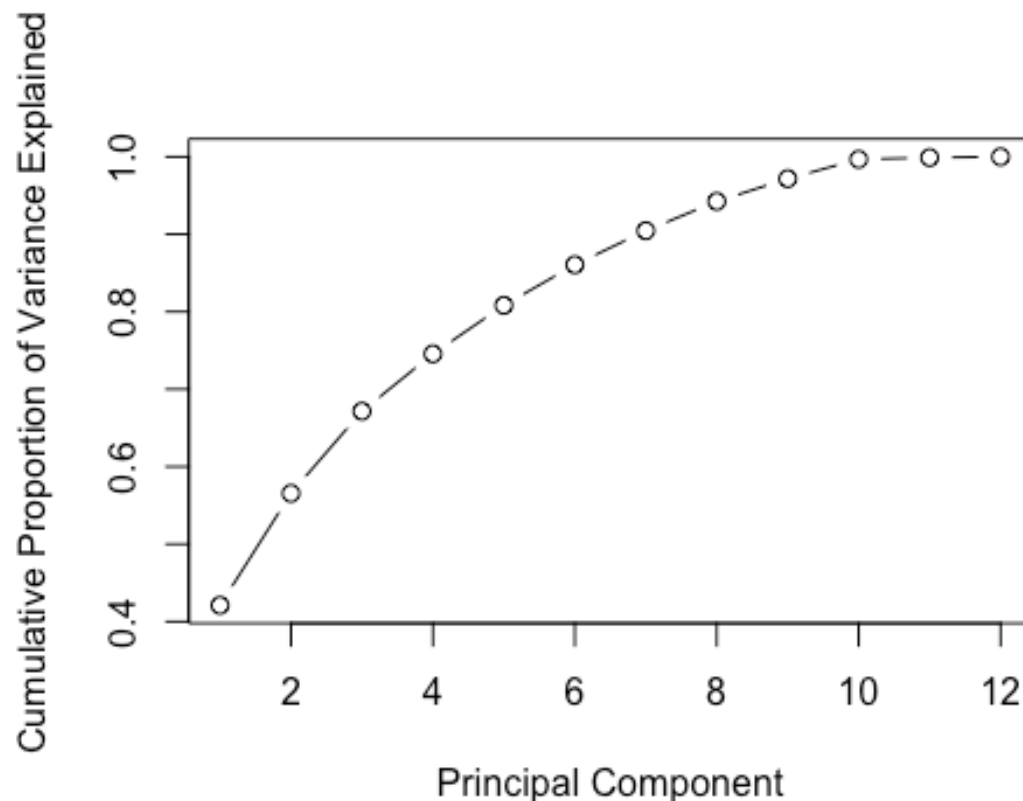
0.17514817				
## CTDMP5	-0.409393313	0.091068304	-0.106015054	0.2585721 -
0.18001923				
## L_C	-0.354979475	0.035948240	0.132445189	-0.4044855
0.16610443				
## SD_I	-0.383800865	-0.041016635	0.076641876	-0.2319667
0.13983783				
## CT	-0.337271237	0.108205764	-0.078654933	0.2776632 -
0.21311376				
## DM	-0.355864875	0.085245360	-0.070815416	0.1344264 -
0.09762329				
## PS	-0.367662085	0.065898791	-0.093034366	0.2005009 -
0.08874023				
##	PC6	PC7	PC8	PC9
PC10				
## SocialR	0.07925342	-0.671748068	0.28942994	-0.02213044
0.02123244				
## ProdR	-0.63674729	0.120062186	-0.09977762	-0.11277129
0.04667925				
## LeadR	-0.05274325	0.322364139	-0.06431972	0.05753089
0.05516684				
## InitiativeR	0.47297488	-0.014170824	0.04167388	0.08896486 -
0.02034588				
## FlexR	0.32410060	0.531529735	-0.20566677	0.13874788
0.02970926				
## LCSDI	0.19823684	0.001113726	0.06059578	-0.01994537 -
0.07983192				
## CTDMP5	-0.15172809	0.014910666	-0.01388630	-0.12754330
0.10136398				
## L_C	0.21009669	-0.011874972	0.16546993	0.19897104
0.62271397				
## SD_I	0.15875210	0.011164289	0.00660373	-0.19229916 -
0.70981455				
## CT	-0.24762333	0.273792200	0.68217722	-0.05960603 -
0.01070190				
## DM	-0.25232347	-0.211317497	-0.32131747	0.75372617 -
0.15426584				
## PS	0.00984595	-0.166101830	-0.50813672	-0.53984051
0.24669940				
##	PC11	PC12		
## SocialR	0.006563060	-0.005397045		
## ProdR	-0.002239230	0.003862368		
## LeadR	-0.001157151	0.009248979		
## InitiativeR	-0.017444034	-0.001842573		
## FlexR	-0.015580159	0.003564157		
## LCSDI	-0.076917176	-0.789360400		
## CTDMP5	0.810402418	-0.082870098		
## L_C	0.037086964	0.404421623		
## SD_I	0.059932649	0.449443342		
## CT	-0.374160529	0.023020646		

## DM	-0.172429466	0.016663833
## PS	-0.402541699	0.059440878



## Importance of components:						
##	PC1	PC2	PC3	PC4	PC5	PC6
PC7						
## Standard deviation	2.248	1.3169	1.1282	0.9424	0.86721	0.79303
0.72589						
## Proportion of Variance	0.421	0.1445	0.1061	0.0740	0.06267	0.05241
0.04391						
## Cumulative Proportion	0.421	0.5655	0.6716	0.7456	0.80827	0.86067
0.90458						
##	PC8	PC9	PC10	PC11	PC12	
## Standard deviation	0.67415	0.59273	0.54823	0.16217	0.11104	
## Proportion of Variance	0.03787	0.02928	0.02505	0.00219	0.00103	
## Cumulative Proportion	0.94246	0.97173	0.99678	0.99897	1.00000	
## [1]	0.420996485	0.144520418	0.106073831	0.074004067	0.062671008	
## [6]	0.052408225	0.043910277	0.037872631	0.029277359	0.025046774	
## [11]	0.002191481	0.001027445	NA	NA	NA	
## [16]	NA	NA	NA	NA	NA	





Rerunning the PCA with a different package

Welcome! Related Books: `Practical Guide To Cluster Analysis in R`
at <https://goo.gl/13EFCZ>

Results for the Principal Component Analysis (PCA)
The analysis was performed on 543 individuals, described by 12
variables

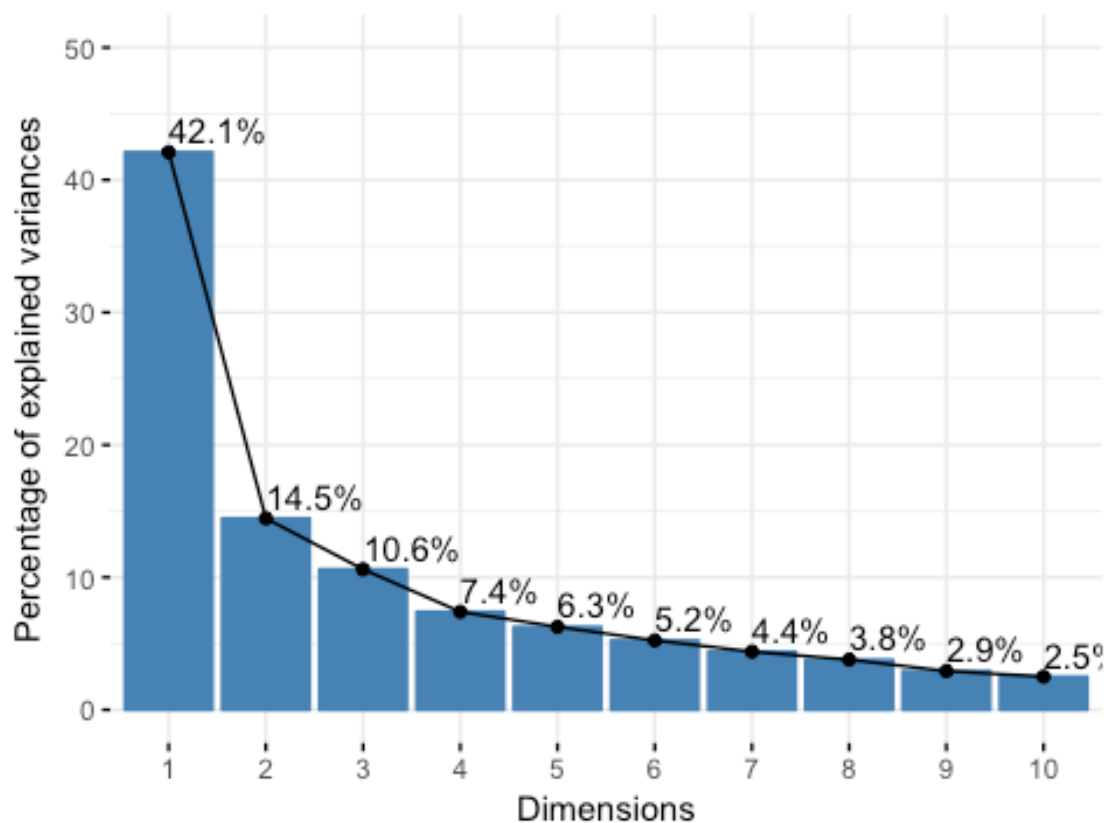
*The results are available in the following objects:

	name	description
## 1	"\$eig"	"eigenvalues"
## 2	"\$var"	"results for the variables"
## 3	"\$var\$coord"	"coord. for the variables"
## 4	"\$var\$cor"	"correlations variables - dimensions"
## 5	"\$var\$cos2"	"cos2 for the variables"
## 6	"\$var\$contrib"	"contributions of the variables"
## 7	"\$ind"	"results for the individuals"
## 8	"\$ind\$coord"	"coord. for the individuals"
## 9	"\$ind\$cos2"	"cos2 for the individuals"
## 10	"\$ind\$contrib"	"contributions of the individuals"
## 11	"\$call"	"summary statistics"
## 12	"\$call\$centre"	"mean of the variables"


```
## 13 "$call$ecart.type" "standard error of the variables"
## 14 "$call$row.w"      "weights for the individuals"
## 15 "$call$col.w"      "weights for the variables"

##      eigenvalue variance.percent cumulative.variance.percent
## Dim.1  5.05195782      42.0996485      42.09965
## Dim.2  1.73424501      14.4520418      56.55169
## Dim.3  1.27288597      10.6073831      67.15907
## Dim.4  0.88804880       7.4004067      74.55948
## Dim.5  0.75205209       6.2671008      80.82658
## Dim.6  0.62889870       5.2408225      86.06740
## Dim.7  0.52692332       4.3910277      90.45843
## Dim.8  0.45447157       3.7872631      94.24569
## Dim.9  0.35132831       2.9277359      97.17343
## Dim.10 0.30056129       2.5046774      99.67811
## Dim.11 0.02629777       0.2191481      99.89726
## Dim.12 0.01232934       0.1027445     100.00000
```

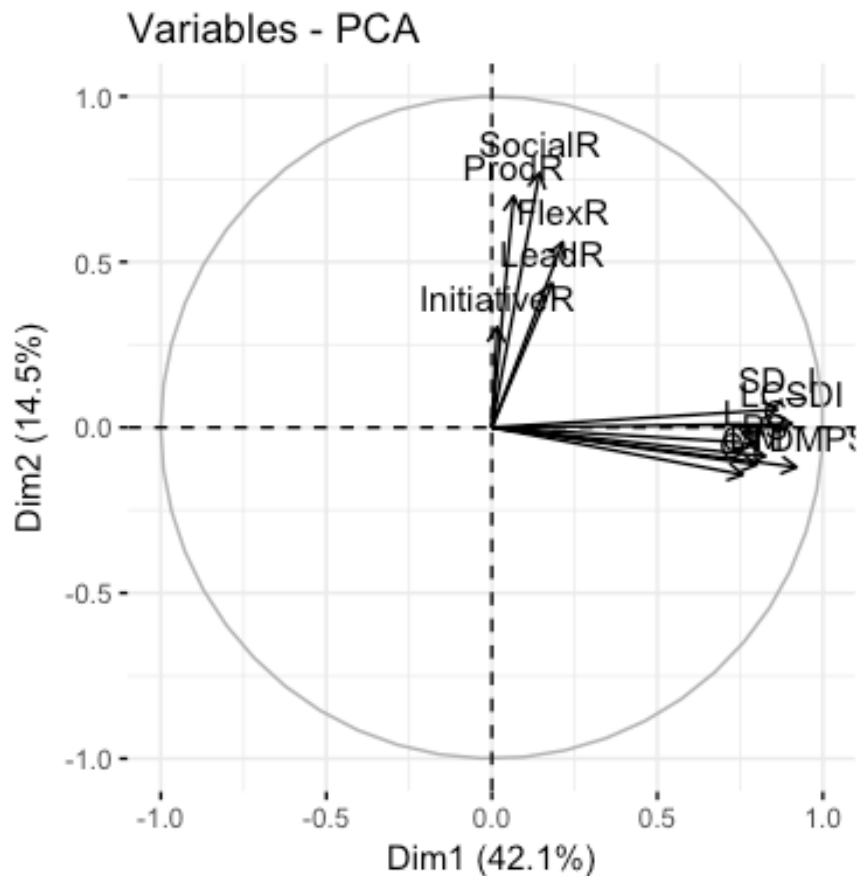
Scree plot

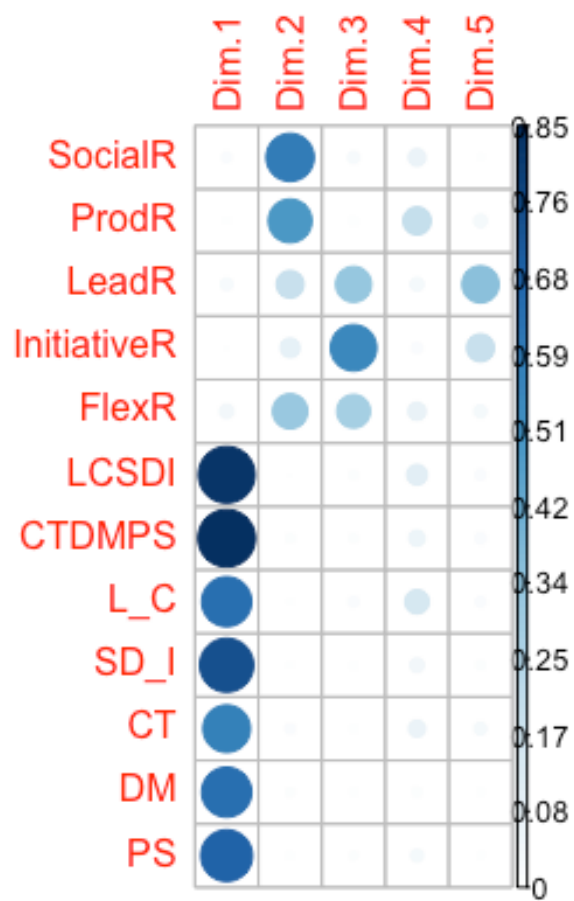


```
## Principal Component Analysis Results for variables
## =====
##   Name      Description
## 1 "$coord"  "Coordinates for the variables"
## 2 "$cor"    "Correlations between variables and dimensions"
```

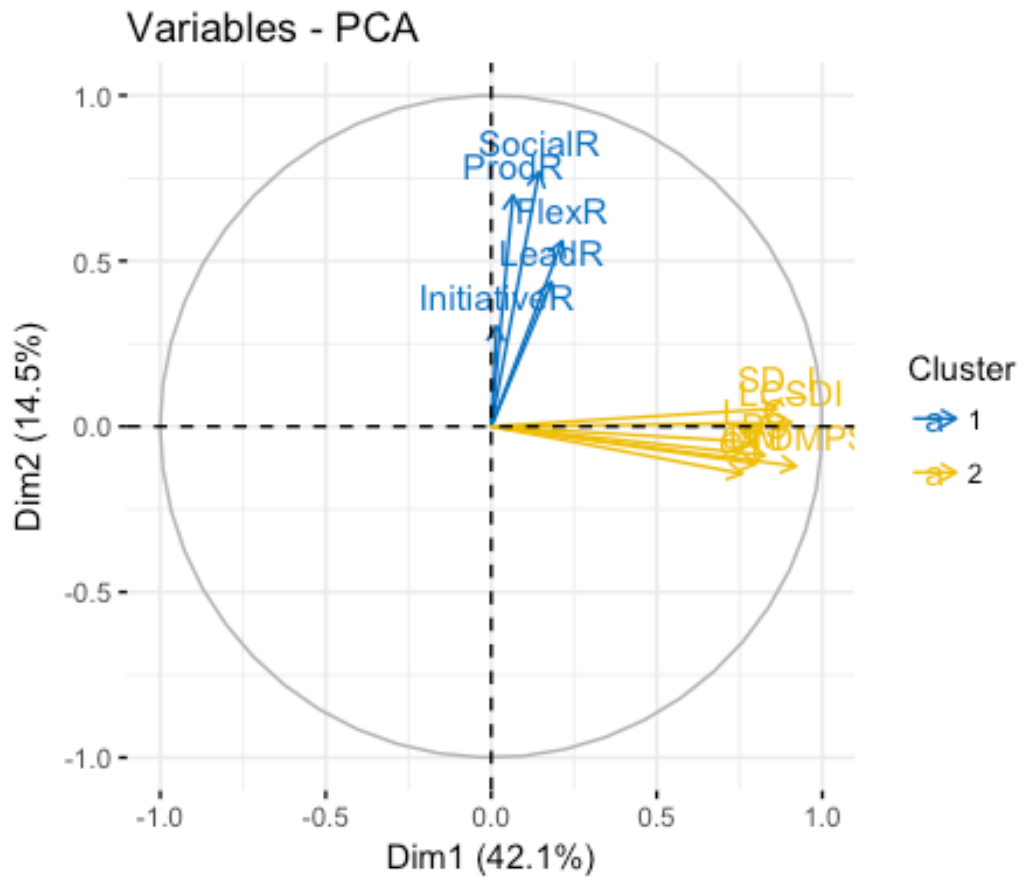
```
## 3 "$cos2"      "Cos2 for the variables"
## 4 "$contrib"   "contributions of the variables"

##          Dim.1      Dim.2      Dim.3      Dim.4
Dim.5
## SocialR      0.14375366  0.76851357 -0.174592507  0.2696400 -
0.07369908
## ProdR        0.06605322  0.69844068  0.006058526 -0.4439037
0.19621909
## LeadR        0.18258468  0.43590841  0.570975318  0.2027352 -
0.59081880
## InitiativeR  0.01589988  0.30346173  0.739216025 -0.1550556
0.43895266
## FlexR        0.21419246  0.56000278 -0.534365025  0.2784898
0.19033764
## LCSDI        0.90698456  0.01112120 -0.122394475 -0.3208230 -
0.15189013
## CTDMPS       0.92017537 -0.11992852  0.119608568  0.2436689
0.15611437
## L_C          0.79787177 -0.04734050 -0.149427640 -0.3811723 -
0.14404732
## SD_I         0.86265234  0.05401511 -0.086469088 -0.2185969 -
0.12126867
```





How the variables cluster together



```
## $quanti
## correlation p.value
## CTDMP 0.9201754 1.749361e-222
## LCSDI 0.9069846 2.543441e-205
## SD_I 0.8626523 2.820533e-162
## PS 0.8263779 4.909111e-137
## DM 0.7998618 4.826213e-122
## L_C 0.7978718 5.212894e-121
## CT 0.7580697 1.679972e-102
## FlexR 0.2141925 4.697420e-07
## LeadR 0.1825847 1.861506e-05
## SocialR 0.1437537 7.807099e-04

## $quanti
## correlation p.value
## SocialR 0.7685136 5.384773e-107
## ProdR 0.6984407 1.227815e-80
## FlexR 0.5600028 3.791730e-46
## LeadR 0.4359084 1.365392e-26
## InitiativeR 0.3034617 4.968027e-13
## PS -0.0867826 4.323898e-02
```

```
## DM          -0.1122602  8.839992e-03
## CTDMPs      -0.1199285  5.137501e-03
## CT          -0.1424970  8.688620e-04
```

Part 5: Regression

Regression LCSDI

the variables that make up majority of PC1 as it relates to LCSDI (External1) I will run a regression on the variables that have over 75% correlation to PC1

```
##
## Call:
## lm(formula = LCSDI ~ +FlexR + Total.Good.Decisions + episodes,
##     data = M2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -11.9456  -1.6096   0.5131   2.3567   4.8975
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    16.41690     0.28725  57.152 < 2e-16 ***
## FlexR           2.28589     0.70548   3.240  0.00127 **
## Total.Good.Decisions 0.02943     0.01664   1.769  0.07748 .
## episodes       -0.17515     0.07960  -2.200  0.02821 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.024 on 539 degrees of freedom
## Multiple R-squared:  0.04148,    Adjusted R-squared:  0.03614
## F-statistic: 7.775 on 3 and 539 DF,  p-value: 4.33e-05
##
## Call:
## lm(formula = LCSDI ~ SocialR + FlexR + Total.Good.Decisions +
##     episodes, data = M2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -11.9593  -1.6117   0.5079   2.3485   4.9375
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    16.30385     0.45728  35.654 < 2e-16 ***
## SocialR         0.23931     0.75276   0.318  0.75067
```

```
## FlexR                2.24654    0.71684    3.134    0.00182 **
## Total.Good.Decisions  0.02775    0.01747    1.589    0.11274
## episodes              -0.16960    0.08156   -2.079    0.03805 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.027 on 538 degrees of freedom
## Multiple R-squared:  0.04166,    Adjusted R-squared:  0.03453
## F-statistic: 5.847 on 4 and 538 DF,  p-value: 0.0001305
```

the variables that make up majority of PC1 as it relates to CTDMPs (External2) I will run a regression on the variables that have over 75% correlation to PC1

```
##
## Call:
## lm(formula = CTDMPs ~ +FlexR + Total.Good.Decisions + episodes,
##     data = M2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12.0053  -1.5096   0.3567   2.2744   4.9770
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    16.44989    0.28980   56.763  < 2e-16 ***
## FlexR           2.42426    0.71175    3.406 0.000708 ***
## Total.Good.Decisions  0.02593    0.01678    1.545 0.122980
## episodes       -0.22877    0.08031   -2.849 0.004560 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.051 on 539 degrees of freedom
## Multiple R-squared:  0.03765,    Adjusted R-squared:  0.03229
## F-statistic: 7.029 on 3 and 539 DF,  p-value: 0.0001212

##
## Call:
## lm(formula = CTDMPs ~ SocialR + FlexR + Total.Good.Decisions +
##     episodes, data = M2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12.0109  -1.5081   0.3574   2.2761   4.9734
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    16.45955    0.46139   35.674  < 2e-16 ***
## SocialR        -0.02045    0.75952   -0.027 0.978534
## FlexR           2.42762    0.72328    3.356 0.000845 ***
## Total.Good.Decisions  0.02607    0.01762    1.479 0.139646
```

```
## episodes          -0.22924    0.08229  -2.786 0.005530 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.054 on 538 degrees of freedom
## Multiple R-squared:  0.03765,    Adjusted R-squared:  0.0305
## F-statistic: 5.262 on 4 and 538 DF,  p-value: 0.0003654
```