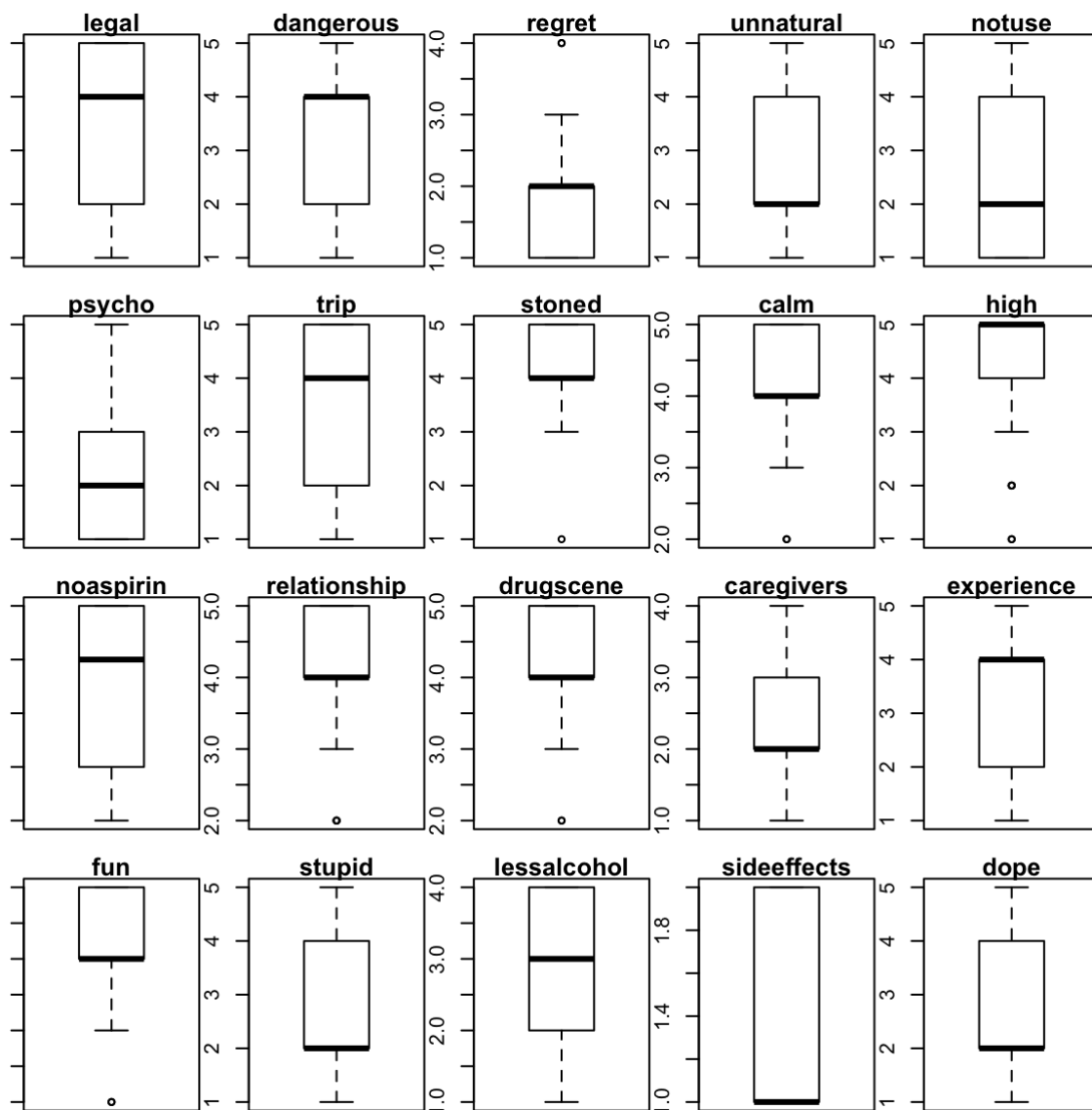


# R Notebook

We load in the data and change the data type for manipulation accordingly.

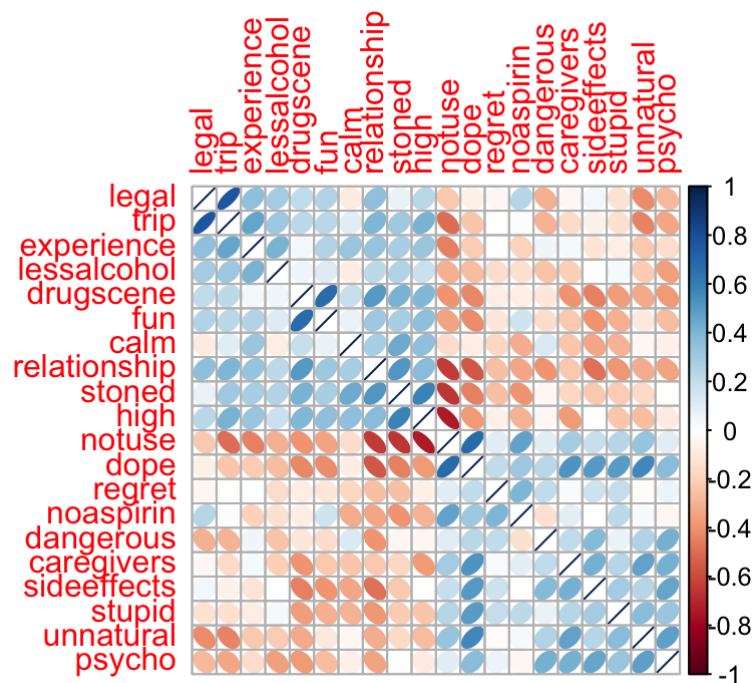
We first observe the data's univariate statistics.



It was observed that the 'regret' variable was skewed right, while variables 'fun', 'drugscene', 'relationship', 'high', 'calm', and 'stoned' were skewed to the left, with some outliers present. To transform the data, the square root of 'regret' was used, while the other variables mentioned above were squared. Through the transformation, outliers were addressed in all variables except that for 'stoned', although the skew was greatly reduced by the transformation.

We then check to see if some variables are linearly correlated.

```
## corrplot 0.84 loaded
```

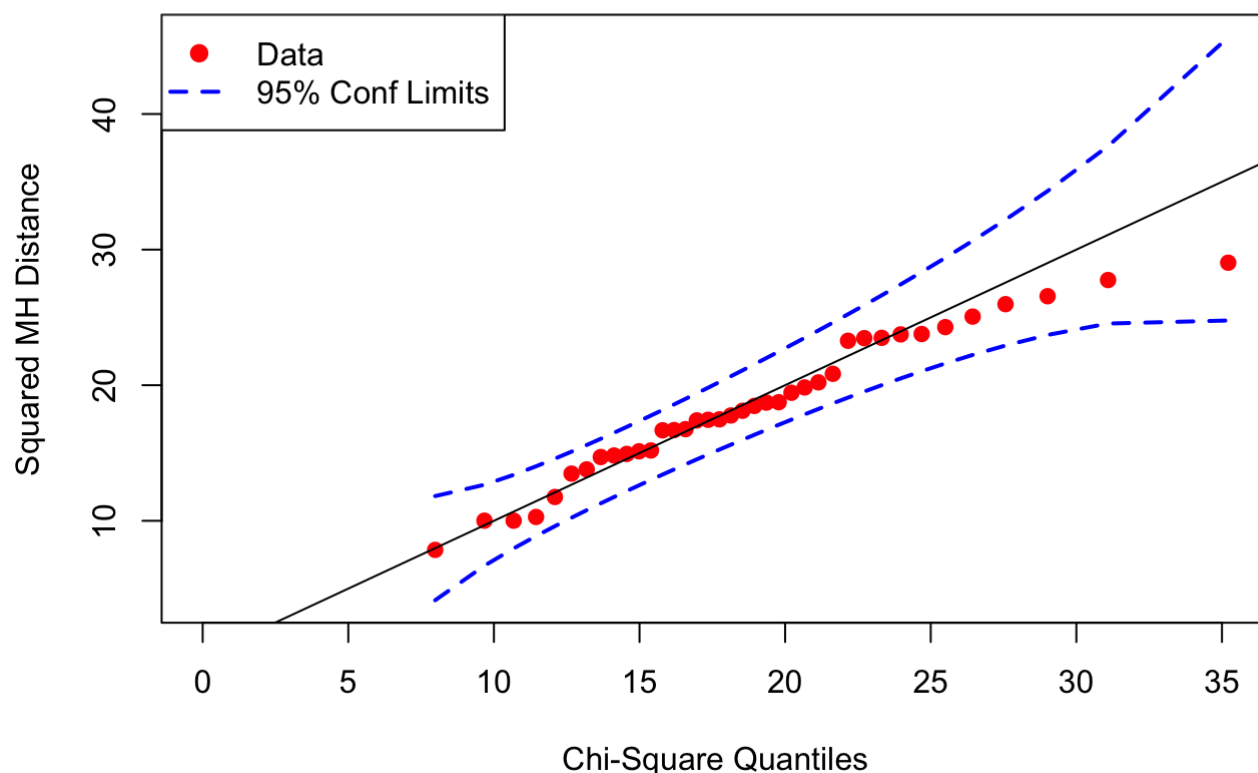


There are a number of pairs of variables with both positive (e.g. trip and legal, dope and notuse) and negative correlation (e.g. notuse and high, stoned and notuse) of 0.6 and above, which leads us to believe that PCA will be useful in combining certain variables.

We then load the function to graph a Chi-Square QQ Plot from the JDRS website.

Then test to see if the data has a multivariate normal distribution.

## Chi-Square Quantiles for Drug Attitude Data



Upon comparing this to the untransformed chi-square quantile plot (not shown but previously tested), the transformed data appears to follow the multivariate normal line better than the raw data. The left tail lies closer to the normal line, and the outliers on the right tail are brought closer to the normal line, though still not completely normal. We conclude that the data does not have a normal multivariate distribution, therefore in testing PCA, we cannot apply Parallel Testing.

Here we perform PCA analysis.

```
## Importance of components:
##
##          Comp.1    Comp.2    Comp.3    Comp.4    Comp.5
## Standard deviation  2.448351 1.5115008 1.32921022 1.20286033 1.1005956
## Proportion of Variance 0.315496 0.1202439 0.09298946 0.07615121 0.0637532
## Cumulative Proportion 0.315496 0.4357399 0.52872936 0.60488057 0.6686338
##
##          Comp.6    Comp.7    Comp.8    Comp.9
## Standard deviation  1.05215962 0.93774416 0.92494920 0.86768999
## Proportion of Variance 0.05826526 0.04628232 0.04502795 0.03962557
## Cumulative Proportion 0.72689903 0.77318135 0.81820930 0.85783487
##
##          Comp.10    Comp.11    Comp.12    Comp.13
## Standard deviation  0.75315452 0.68916611 0.67882175 0.5528159
## Proportion of Variance 0.02985483 0.02499736 0.02425258 0.0160845
## Cumulative Proportion 0.88768970 0.91268706 0.93693964 0.9530241
##
##          Comp.14    Comp.15    Comp.16    Comp.17
## Standard deviation  0.54095297 0.4731544 0.382632830 0.345344985
## Proportion of Variance 0.01540158 0.0117829 0.007705678 0.006277008
## Cumulative Proportion 0.96842572 0.9802086 0.987914299 0.994191307
##
##          Comp.18    Comp.19
## Standard deviation  0.255805531 0.211963892
## Proportion of Variance 0.003444025 0.002364668
## Cumulative Proportion 0.997635332 1.000000000
```

```
## Warning in if (loadings) {: the condition has length > 1 and only the first
## element will be used
```

```
##
## Loadings:
##
```

	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5	Comp.6	Comp.7	Comp.8
## dangerous	0.138	0.345	0.142	0.337	0.274	0.102	0.170	0.428
## regret	0.111	-0.100	-0.241	0.516	0.272	0.119	-0.041	-0.093
## unnatural	0.235	0.219	0.177	0.025	-0.455	-0.065	-0.195	0.101
## notuse	0.319	-0.254	0.141	-0.060	-0.052	0.280	-0.038	0.091
## psycho	0.211	0.391	0.102	0.129	-0.057	-0.239	-0.007	-0.160
## trip	-0.220	0.012	-0.434	0.081	-0.071	0.153	0.129	-0.382
## stoned	-0.268	0.318	0.065	0.035	-0.131	-0.092	-0.198	-0.093
## calm	-0.156	0.237	0.297	0.046	0.013	0.629	-0.266	-0.101
## high	-0.265	0.268	-0.085	0.249	0.122	-0.134	-0.223	-0.171
## noaspirin	0.145	-0.391	-0.173	0.384	-0.209	0.067	-0.053	-0.107
## relationship	-0.316	0.028	0.041	-0.140	-0.212	0.008	0.131	-0.266
## drugscene	-0.268	-0.124	0.205	0.334	-0.133	-0.159	-0.006	0.081
## caregivers	0.226	0.215	-0.091	-0.050	-0.414	0.028	0.521	-0.105
## experience	-0.180	0.230	-0.337	0.053	-0.157	0.451	0.190	0.276
## fun	-0.223	-0.128	0.083	0.403	-0.413	-0.155	0.108	0.299
## stupid	0.214	0.076	-0.292	0.016	-0.203	-0.120	-0.566	-0.012
## lessalcohol	-0.156	0.007	-0.438	-0.265	-0.082	-0.029	-0.221	0.525
## sideeffects	0.225	0.283	-0.296	0.048	0.205	-0.227	0.161	-0.023
## dope	0.324	0.075	-0.071	0.086	-0.195	0.256	-0.140	-0.169

```
##
```

	Comp.9	Comp.10	Comp.11	Comp.12	Comp.13	Comp.14	Comp.15
## dangerous	0.027	0.110	0.276	0.165	0.318	0.087	0.407
## regret	0.452	0.322	-0.246	-0.296	0.005	-0.002	-0.270
## unnatural	0.030	0.143	-0.513	-0.037	0.411	0.167	0.050
## notuse	-0.246	0.039	0.150	-0.106	-0.128	0.060	-0.075
## psycho	0.166	-0.404	0.139	-0.124	-0.291	0.524	-0.188
## trip	-0.238	-0.024	0.126	0.244	0.509	0.310	-0.169
## stoned	0.008	0.040	0.430	-0.482	0.229	-0.226	-0.152
## calm	-0.128	-0.046	0.001	-0.110	-0.074	0.028	-0.038
## high	-0.293	-0.097	-0.344	0.014	-0.186	-0.350	0.264
## noaspirin	-0.074	-0.325	0.171	-0.325	0.023	0.033	0.494
## relationship	0.274	0.322	-0.052	-0.038	-0.214	0.325	0.522
## drugscene	-0.145	0.472	0.274	0.170	-0.255	0.145	-0.158
## caregivers	0.074	0.162	0.129	-0.149	-0.042	-0.432	-0.021
## experience	0.208	-0.222	-0.075	0.179	-0.284	-0.021	-0.057
## fun	-0.155	-0.212	-0.133	0.100	-0.035	-0.024	-0.202
## stupid	0.274	0.045	0.298	0.438	-0.074	-0.189	0.072
## lessalcohol	-0.117	0.155	0.011	-0.387	-0.041	0.231	0.019
## sideeffects	-0.438	0.144	-0.061	-0.075	-0.200	0.082	0.043
## dope	-0.305	0.293	-0.013	0.069	-0.187	0.067	-0.031

```
##
```

	Comp.16	Comp.17	Comp.18	Comp.19
## dangerous	0.064	0.029	0.186	0.033
## regret	-0.093	0.029	0.107	0.066
## unnatural	0.005	-0.089	-0.306	0.137
## notuse	-0.164	-0.176	0.184	0.711
## psycho	0.154	0.160	0.079	0.116
## trip	0.006	0.115	0.056	0.178
## stoned	-0.026	-0.449	0.022	0.011
## calm	-0.338	0.398	-0.104	-0.178
## high	0.249	0.116	0.134	0.378
## noaspirin	0.048	0.025	-0.278	-0.119

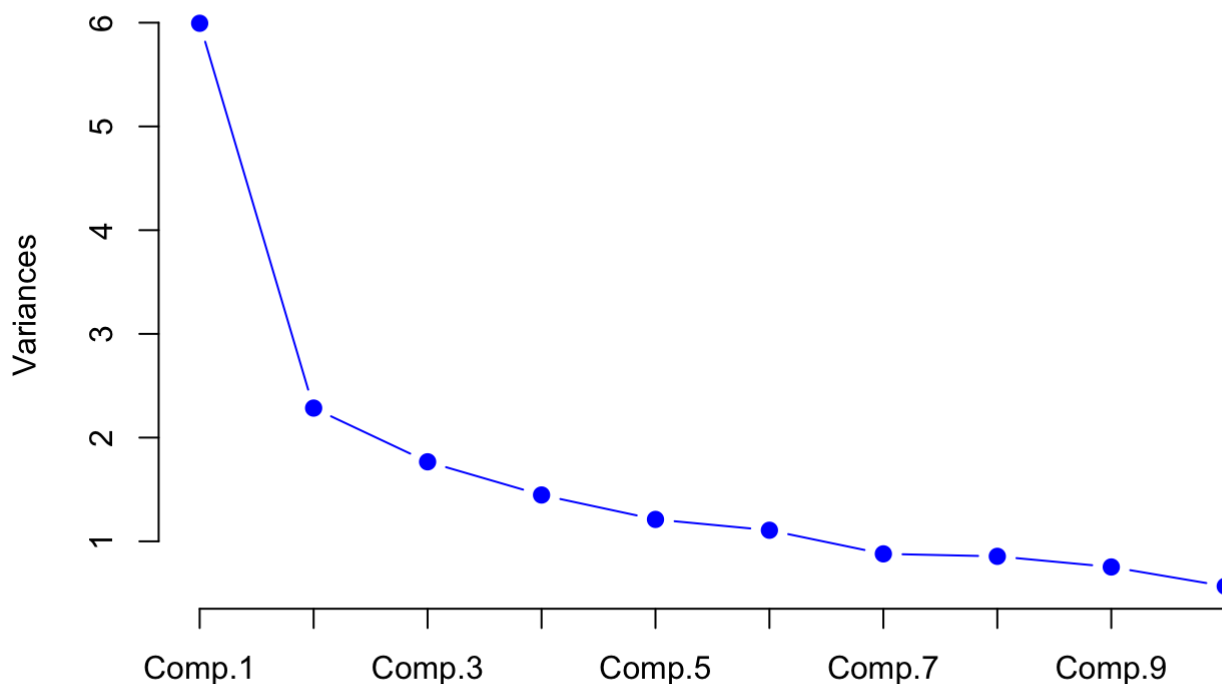
```
## relationship -0.231 -0.137  0.245  0.075
## drugscene   0.185  0.125 -0.434  0.109
## caregivers  0.025  0.412  0.021  0.085
## experience  0.112 -0.374 -0.285  0.098
## fun         -0.342 -0.028  0.434 -0.152
## stupid      -0.264  0.096  0.031  0.053
## lessalcohol 0.130  0.325  0.127  0.007
## sideeffects -0.534 -0.168 -0.251 -0.117
## dope        0.410 -0.243  0.335 -0.402
```

The first six components have an eigenvalue above 1, however it requires 8 parameters to cover 80% of the variance in data. Choosing six components will obtain 72.6% of the variance, which is a relatively small difference from 80%.

The first component alone was able to account for 31.5% of the data, and at most 8 components are required to account for most of the data, showing a successful application of PCA.

```
## Comp.1  Comp.2  Comp.3  Comp.4  Comp.5  Comp.6  Comp.7  Comp.8  Comp.9
##    5.99    2.28    1.77    1.45    1.21    1.11    0.88    0.86    0.75
## Comp.10 Comp.11 Comp.12 Comp.13 Comp.14 Comp.15 Comp.16 Comp.17 Comp.18
##    0.57    0.47    0.46    0.31    0.29    0.22    0.15    0.12    0.07
## Comp.19
##    0.04
```

## Scree Plot of Raw Drug Attitude Data



The screeplot shows only one elbow after the first component, however following this plot will only account for 31.5% of the variance, thus we follow the results from the eigenvalue and total variance criteria.

As stated above, the Parallel Method is not used because the data does not display a multivariate normal distribution.

We then study the composition of the components by variable.

##	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5	Comp.6
## dangerous	0.14	0.35	0.14	0.34	0.27	0.10
## regret	0.11	-0.10	-0.24	0.52	0.27	0.12
## unnatural	0.24	0.22	0.18	0.02	-0.45	-0.07
## notuse	0.32	-0.25	0.14	-0.06	-0.05	0.28
## psycho	0.21	0.39	0.10	0.13	-0.06	-0.24
## trip	-0.22	0.01	-0.43	0.08	-0.07	0.15
## stoned	-0.27	0.32	0.07	0.04	-0.13	-0.09
## calm	-0.16	0.24	0.30	0.05	0.01	0.63
## high	-0.26	0.27	-0.09	0.25	0.12	-0.13
## noaspirin	0.15	-0.39	-0.17	0.38	-0.21	0.07
## relationship	-0.32	0.03	0.04	-0.14	-0.21	0.01
## drugscene	-0.27	-0.12	0.21	0.33	-0.13	-0.16
## caregivers	0.23	0.21	-0.09	-0.05	-0.41	0.03
## experience	-0.18	0.23	-0.34	0.05	-0.16	0.45
## fun	-0.22	-0.13	0.08	0.40	-0.41	-0.16
## stupid	0.21	0.08	-0.29	0.02	-0.20	-0.12
## lessalcohol	-0.16	0.01	-0.44	-0.26	-0.08	-0.03
## sideeffects	0.22	0.28	-0.30	0.05	0.21	-0.23
## dope	0.32	0.08	-0.07	0.09	-0.20	0.26

#### Component 1: Social Considerations of Doing Drugs

Positive: notuse, dope Negative: stoned, high, relationship, drugscene

Variables notuse and dope both avoid drugs, saying that even if offered by close friends, doing drugs is a dumb idea. Variables stoned, high, relationship, and drugscene enjoy drugs saying that the drug scene and relationships involving drugs are more enjoyable. This component accounts for how strong ones view towards the social benefits of doing drugs

Component 2: Health Risks of Doing Drugs positive: dangerous, psycho, stoned, high, sideeffects negative: not use, no aspirin

While all positive, dangerous, psycho, and sideeffects avoid drugs, while psycho and stoned enjoy it. Variables notuse and noaspirin also avoid drugs. These all point towards the health implications of doing drugs.

Component 3: Practical Considerations when Doing Drugs positive: calm negative: regret, trip, experience, stupid, lessalcohol, side effects

The positively ranked variable calm and negatively ranked variables trip, and lessalcohol advocate for alcohol, while negatively ranked variables regret, experience, stupid, and side effects avoid it. All variables however point to drug use or avoidance for practical reasons, citing its use for calming oneself and mildness in comparison with alcohol, as well as consequences of drug use including saying things one does not intend to, experiencing harmful side effects of drugs, and lawmakers needing personal experience before creating drug-related legislation.

Component 4: Enthusiasm Towards Drugs 1 positive: dangerous, regret, high, no aspirin, drugscene, fun negative: less alcohol

Positively ranked variables dangerous, regret, and noaspirin avoid drugs, whereas high, drugscene, fun, and negatively ranked variable lessalcohol advocates for alcohol. These all express general enthusiasm towards the idea of doing drugs.

Component 5: Enthusiasm Towards Drugs 2 positive: dangerous, regret negative: unnatural, caregivers, fun

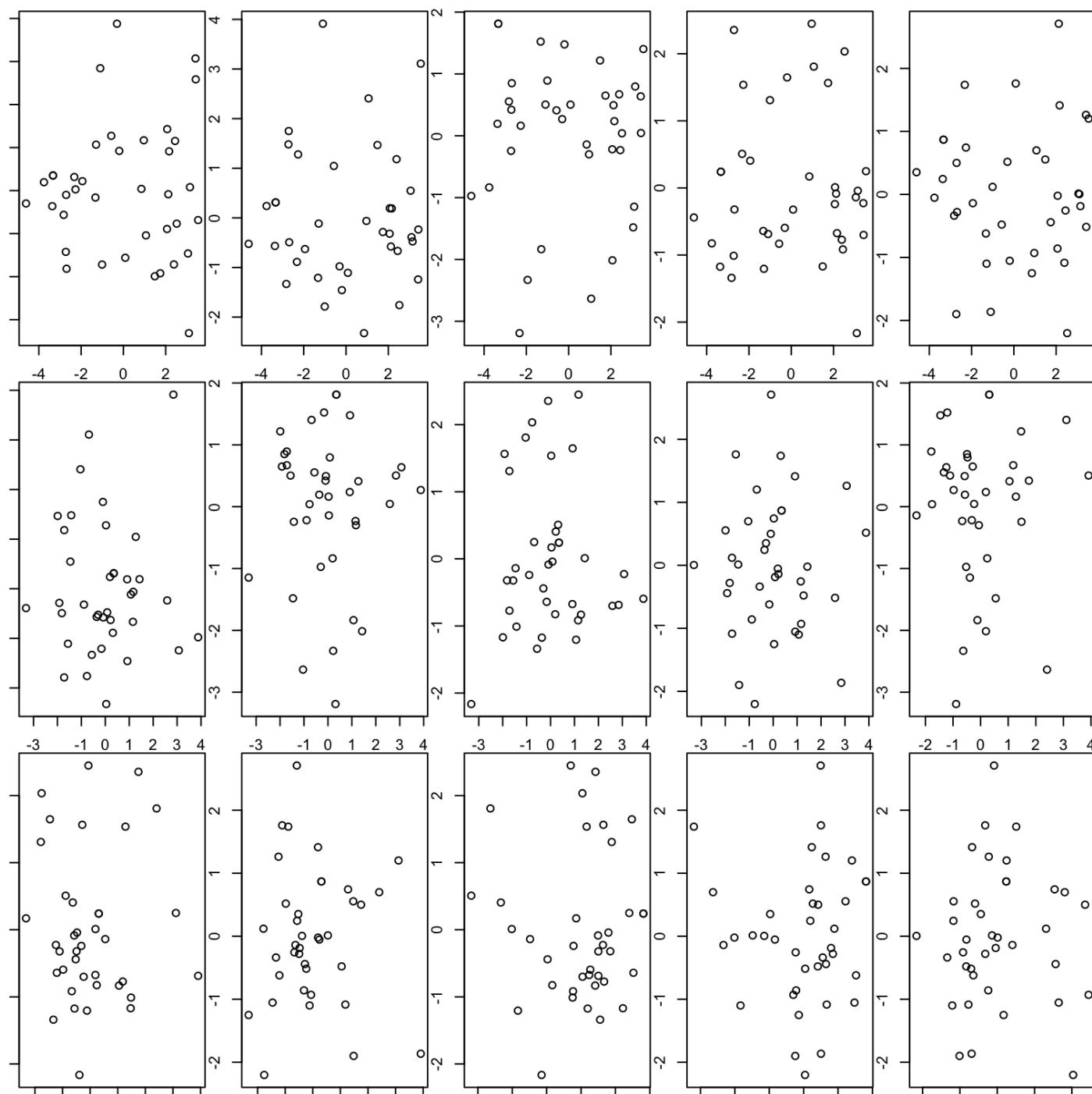
Positively ranked variables dangerous and regret, as well as negatively ranked variables unnatural and caregivers avoid drugs, while negatively ranked variable fun advocates for it. These also point towards general feelings of enthusiasm or fear of drugs.

Component 6: Enthusiasm Towards Drugs 3 positive: notuse, calm, experience, dope

While all positive, notuse and dope avoid drugs while calm and experience advocate for it. They also offer general feelings of enthusiasm or fear of drugs.



Then the components were plotted pairwise against one another



They all show no clear correlation among the variables.

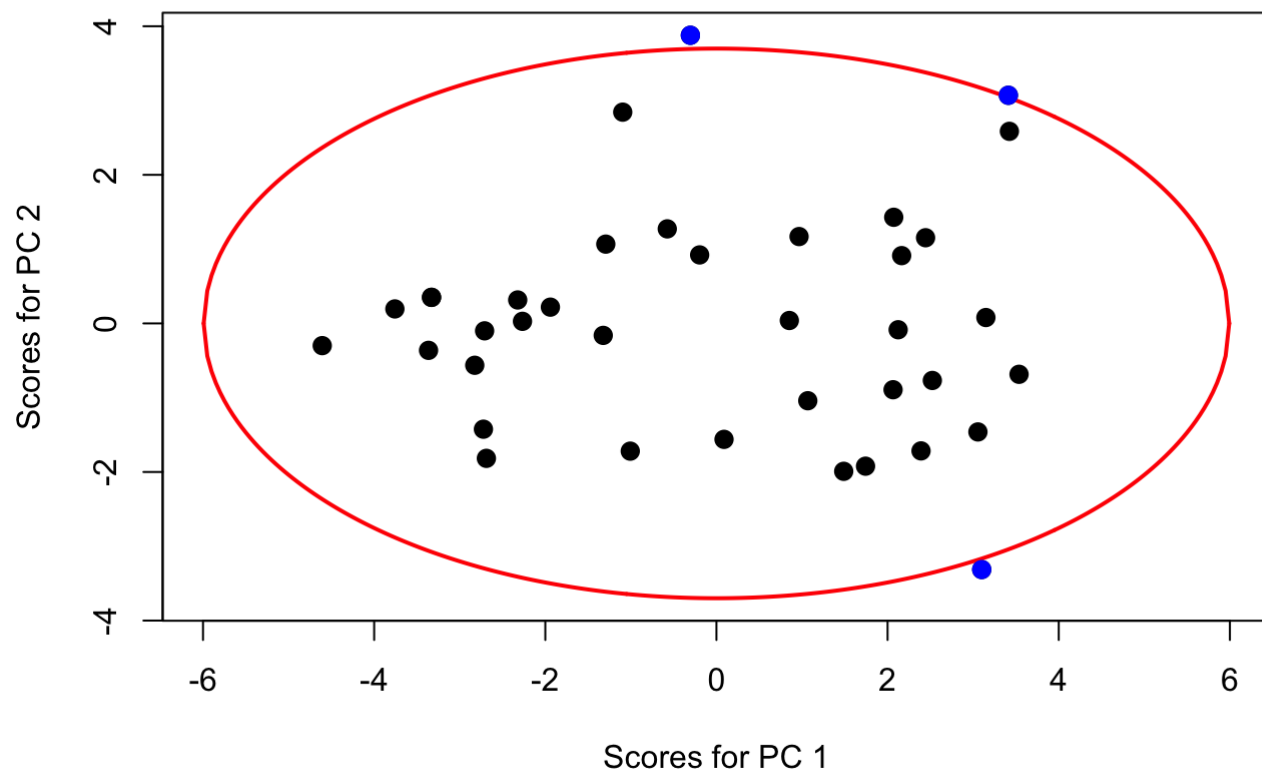
Load the function for creating a 95% confidence ellipse from the JDRS website.

We then create a 95% confidence ellipse plotting the first vs. second components.

```
## Warning in sqrt((5.99 - (y1vec^2)/x$sdev[comps[1]]^2) *
## x$sdev[comps[2]]^2): NaNs produced

## Warning in sqrt((5.99 - (y1vec^2)/x$sdev[comps[1]]^2) *
## x$sdev[comps[2]]^2): NaNs produced
```

### PC Score Plot with 95% CI Ellipse



We also create a biplot plotting the first vs. second components.

