Principal Components on the Diversity Data*

Employees at Canadian corporations filled out questionnaires about their jobs. Questionnaires employed 5-point scales, where 5 indicates the highest level of the trait or opinion being assessed (like job satisfaction) and 1 indicating the lowest level. The wording of the questions was varied so that sometimes a one indicated higher satisfaction (for example, strong disagreement with "I hate my job."), but the numbers were switched around so that in the data file, larger numbers always indicate more. Data consist of answers to

- Ten questions about commitment to the organization, with higher numbers indicating more commitment.
- Five questions about relations with colleagues at work, with higher numbers indicating better relations.
- Twelve questions about relations with management, in particular the respondent's immediate boss. Higher numbers indicate better relations.
- Six questions about fair opportunities for advancement, with higher numbers indicating more fairness.
- Four questions about job satisfaction, with higher numbers indicating more satisfaction.
- Three questions about senior management's commitment to diversity, with higher numbers indicating more commitment. These seem to be on a six-point scale instead of five.
- Gender: 0=Male, 1=Female
- Visible Minority status: 0=No, 1=Yes
- Education level, numbered 1-7. The exact meanings of the numbers are unknown, but surely higher numbers must indicate more education, mostly.
- Marital status: 1=never married, 2=married, 3=divorced or separated, 4=widowed. This is a guess, but I'm fairly confident.
- Age in years
- Born outside Canada: 0=No, 1=Yes

These are real data from a consulting job. There are two data sets of size n=500, randomly sampled from around 16,000 questionnaires. The idea is to arrive at conclusions and predictions based on the exploratory (training) sample, and then test them out on the replication sample.

There are lots of interesting things one could do with these data. The objective here is to reduce the 40 questions to a smaller number of linear combinations.

```
> rm(list=ls())
> # install.packages("readxl", dependencies = TRUE) # Only need to do this once
> library(readxl)
> # Download the data and put it in your working directory.
> # https://www.utstat.toronto.edu/brunner/data/legal/DiversityExplore.xlsx
> # The replication data are in DiversityReplic.xlsx
> ddata = read_excel("DiversityExplore.xlsx") # Read local copy

> ddata = as.data.frame(ddata) # Instead of a "tibble"
> dim(ddata); head(ddata)
[1] 500  47
  id Com1 Com2 Com3 Com4 Com5 Com6 Com7 Com8 Com9 Com10 RelC1 RelC2 RelC3 RelC4 RelC5
1  1    4    4    5    3    4    2    3    3    2     3     4     4     4     2     4
2  2    5    5    5    5    5    4    5    5    4     5     5     5     5     5     5
3  3    5    5    5    5    5    5    5    5    5     5     5     5     5     5     5
4  4    4    4    4    4    4    4    4    5    2     4     3     2     2     2     1
5  5    5    4    4    3    3    3    5    5    4     4     3     4     4     4     5
6  6    2    4    4    4    2    1    3    4    3     4     5     5     5     4     5
  RelM1 RelM2 RelM3 RelM4 RelM5 RelM6 RelM7 RelM8 RelM9 RelM10 RelM11 RelM12 Fair1
1     2     5     4     5     5     5     5     4     4      4      5      5     4
2     3     5     5     5     5     5     3     4     4      5      5      5     5
3     4     5     3     5     3     2     3     5     5      5      5      5     5
4     1     2     2     3     3     1     1     2     3      3      3      3     3
5     3     4     4     4     4     4     3     3     5      5      5      5     4
6     3     3     3     3     4     2     3     3     3      2      4      4     4
  Fair2 Fair3 Fair4 Fair5 Fair6 Sat1 Sat2 Sat3 Sat4 SM1 SM2 SM3 Gender VisMinority
1     1     4     1     2     1    2    2    3    4   3   1   2      0           1
2     4     5     5     5     5    5    4    5    5   4   5   5      1        <NA>
3     2     5     1     5     2    5    4    4    1   6   6   3      1           0
4     3     3     1     4     3    2    3    3    3   6   6   6      1           0
5     5     4     2     3     2    4    5    5    5   3   3   3      1           1
6     4     4     3     2     3    4    4    4    4   4   3   3      0           1
  EDUCLevel MaritalStatus Age CAN_Foreign_Born
1         4             2  28                0
2         4             2  41                0
3         5             3  45                0
4         3             2  39                0
5         7             1  26                0
6         3             2  59                0
> Sigma_hat = cor(quest); round(Sigma_hat,3)

        Com1  Com2  Com3  Com4  Com5  Com6  Com7  Com8  Com9 Com10 RelC1 RelC2 RelC3
Com1   1.000 0.598 0.407 0.600 0.721 0.643 0.650 0.538 0.443 0.402 0.261 0.277 0.293
Com2   0.598 1.000 0.404 0.474 0.507 0.463 0.456 0.358 0.400 0.414 0.312 0.276 0.308
Com3   0.407 0.404 1.000 0.476 0.429 0.454 0.400 0.299 0.147 0.395 0.203 0.208 0.209
Com4   0.600 0.474 0.476 1.000 0.678 0.650 0.602 0.584 0.365 0.408 0.263 0.277 0.357
Com5   0.721 0.507 0.429 0.678 1.000 0.747 0.668 0.581 0.394 0.367 0.271 0.270 0.334
Com6   0.643 0.463 0.454 0.650 0.747 1.000 0.696 0.560 0.372 0.314 0.233 0.227 0.318
Com7   0.650 0.456 0.400 0.602 0.668 0.696 1.000 0.651 0.351 0.339 0.270 0.251 0.332
Com8   0.538 0.358 0.299 0.584 0.581 0.560 0.651 1.000 0.479 0.285 0.233 0.279 0.318
Com9   0.443 0.400 0.147 0.365 0.394 0.372 0.351 0.479 1.000 0.244 0.245 0.267 0.253
Com10  0.402 0.414 0.395 0.408 0.367 0.314 0.339 0.285 0.244 1.000 0.167 0.176 0.164
RelC1  0.261 0.312 0.203 0.263 0.271 0.233 0.270 0.233 0.245 0.167 1.000 0.795 0.596
RelC2  0.277 0.276 0.208 0.277 0.270 0.227 0.251 0.279 0.267 0.176 0.795 1.000 0.598
RelC3  0.293 0.308 0.209 0.357 0.334 0.318 0.332 0.318 0.253 0.164 0.596 0.598 1.000
RelC4  0.259 0.271 0.163 0.286 0.266 0.251 0.259 0.215 0.237 0.156 0.483 0.533 0.651
RelC5  0.233 0.228 0.174 0.264 0.290 0.230 0.267 0.213 0.190 0.206 0.407 0.467 0.530
RelM1  0.250 0.223 0.133 0.330 0.298 0.311 0.260 0.189 0.187 0.131 0.240 0.263 0.347
RelM2  0.369 0.278 0.212 0.378 0.385 0.352 0.318 0.285 0.204 0.160 0.302 0.320 0.399
RelM3  0.316 0.236 0.216 0.344 0.316 0.337 0.308 0.271 0.223 0.171 0.288 0.334 0.399
RelM4  0.303 0.229 0.236 0.379 0.333 0.360 0.307 0.276 0.187 0.200 0.291 0.308 0.426
RelM5  0.347 0.278 0.225 0.375 0.323 0.320 0.284 0.263 0.217 0.194 0.299 0.321 0.308
RelM6  0.297 0.195 0.198 0.338 0.307 0.259 0.272 0.269 0.194 0.123 0.234 0.283 0.364
```

RelM7	0.270	0.211	0.207	0.328	0.303	0.267	0.299	0.250	0.193	0.135	0.251	0.309	0.341
RelM8	0.307	0.252	0.184	0.323	0.311	0.271	0.325	0.289	0.226	0.122	0.276	0.304	0.331
RelM9	0.343	0.223	0.212	0.385	0.333	0.352	0.356	0.339	0.212	0.152	0.364	0.408	0.438
RelM10	0.304	0.228	0.147	0.302	0.299	0.293	0.258	0.215	0.197	0.078	0.299	0.313	0.371
RelM11	0.335	0.291	0.271	0.377	0.326	0.317	0.330	0.274	0.171	0.207	0.379	0.428	0.435
RelM12	0.320	0.189	0.212	0.359	0.329	0.365	0.321	0.300	0.209	0.121	0.337	0.392	0.387
Fair1	0.419	0.307	0.273	0.414	0.434	0.416	0.436	0.400	0.287	0.245	0.326	0.372	0.393
Fair2	0.398	0.280	0.223	0.423	0.440	0.428	0.423	0.342	0.324	0.214	0.311	0.297	0.322
Fair3	0.360	0.329	0.180	0.330	0.348	0.347	0.331	0.279	0.222	0.199	0.359	0.390	0.419
Fair4	0.196	0.163	0.160	0.216	0.228	0.260	0.218	0.148	0.223	0.083	0.222	0.214	0.243
Fair5	0.259	0.232	0.106	0.292	0.266	0.318	0.257	0.224	0.177	0.113	0.215	0.208	0.304
Fair6	0.402	0.327	0.240	0.409	0.455	0.426	0.394	0.345	0.308	0.173	0.316	0.340	0.387
Sat1	0.279	0.270	0.237	0.342	0.351	0.387	0.312	0.260	0.219	0.127	0.262	0.266	0.288
Sat2	0.306	0.283	0.217	0.400	0.342	0.347	0.341	0.303	0.233	0.166	0.311	0.342	0.355
Sat3	0.315	0.256	0.206	0.347	0.288	0.338	0.331	0.257	0.188	0.131	0.275	0.301	0.351
Sat4	0.314	0.282	0.188	0.310	0.281	0.318	0.272	0.270	0.168	0.134	0.242	0.276	0.369
SM1	0.199	0.149	0.173	0.256	0.224	0.240	0.192	0.205	0.236	0.178	0.180	0.198	0.150
SM2	0.221	0.241	0.158	0.295	0.247	0.261	0.215	0.213	0.232	0.159	0.240	0.205	0.215
SM3	0.101	0.154	0.121	0.214	0.157	0.132	0.141	0.139	0.202	0.154	0.184	0.148	0.146

	RelC4	RelC5	RelM1	RelM2	RelM3	RelM4	RelM5	RelM6	RelM7	RelM8	RelM9	RelM10	RelM11
Com1	0.259	0.233	0.250	0.369	0.316	0.303	0.347	0.297	0.270	0.307	0.343	0.304	0.335
Com2	0.271	0.228	0.223	0.278	0.236	0.229	0.278	0.195	0.211	0.252	0.223	0.228	0.291
Com3	0.163	0.174	0.133	0.212	0.216	0.236	0.225	0.198	0.207	0.184	0.212	0.147	0.271
Com4	0.286	0.264	0.330	0.378	0.344	0.379	0.375	0.338	0.328	0.323	0.385	0.302	0.377
Com5	0.266	0.290	0.298	0.385	0.316	0.333	0.323	0.307	0.303	0.311	0.333	0.299	0.326
Com6	0.251	0.230	0.311	0.352	0.337	0.360	0.320	0.259	0.267	0.271	0.352	0.293	0.317
Com7	0.259	0.267	0.260	0.318	0.308	0.307	0.284	0.272	0.299	0.325	0.356	0.258	0.330
Com8	0.215	0.213	0.189	0.285	0.271	0.276	0.263	0.269	0.250	0.289	0.339	0.215	0.274
Com9	0.237	0.190	0.187	0.204	0.223	0.187	0.217	0.194	0.193	0.226	0.212	0.197	0.171
Com10	0.156	0.206	0.131	0.160	0.171	0.200	0.194	0.123	0.135	0.122	0.152	0.078	0.207
RelC1	0.483	0.407	0.240	0.302	0.288	0.291	0.299	0.234	0.251	0.276	0.364	0.299	0.379
RelC2	0.533	0.467	0.263	0.320	0.334	0.308	0.321	0.283	0.309	0.304	0.408	0.313	0.428
RelC3	0.651	0.530	0.347	0.399	0.399	0.426	0.308	0.364	0.341	0.331	0.438	0.371	0.435
RelC4	1.000	0.655	0.364	0.401	0.347	0.367	0.329	0.352	0.336	0.310	0.373	0.396	0.395
RelC5	0.655	1.000	0.388	0.378	0.351	0.358	0.366	0.399	0.402	0.334	0.345	0.369	0.421
RelM1	0.364	0.388	1.000	0.711	0.611	0.621	0.555	0.575	0.570	0.518	0.529	0.590	0.484
RelM2	0.401	0.378	0.711	1.000	0.663	0.702	0.676	0.647	0.595	0.574	0.616	0.612	0.604
RelM3	0.347	0.351	0.611	0.663	1.000	0.755	0.609	0.638	0.610	0.622	0.678	0.622	0.647
RelM4	0.367	0.358	0.621	0.702	0.755	1.000	0.606	0.679	0.617	0.606	0.670	0.635	0.649
RelM5	0.329	0.366	0.555	0.676	0.609	0.606	1.000	0.628	0.574	0.536	0.598	0.540	0.588
RelM6	0.352	0.399	0.575	0.647	0.638	0.679	0.628	1.000	0.765	0.599	0.635	0.605	0.546
RelM7	0.336	0.402	0.570	0.595	0.610	0.617	0.574	0.765	1.000	0.607	0.596	0.565	0.520
RelM8	0.310	0.334	0.518	0.574	0.622	0.606	0.536	0.599	0.607	1.000	0.671	0.573	0.567
RelM9	0.373	0.345	0.529	0.616	0.678	0.670	0.598	0.635	0.596	0.671	1.000	0.650	0.735
RelM10	0.396	0.369	0.590	0.612	0.622	0.635	0.540	0.605	0.565	0.573	0.650	1.000	0.638
RelM11	0.395	0.421	0.484	0.604	0.647	0.649	0.588	0.546	0.520	0.567	0.735	0.638	1.000
RelM12	0.348	0.358	0.532	0.621	0.662	0.704	0.567	0.600	0.565	0.591	0.789	0.619	0.682
Fair1	0.366	0.334	0.459	0.512	0.479	0.459	0.508	0.511	0.532	0.454	0.534	0.466	0.480
Fair2	0.317	0.272	0.406	0.400	0.440	0.424	0.342	0.427	0.452	0.372	0.456	0.430	0.392
Fair3	0.392	0.372	0.445	0.446	0.393	0.396	0.388	0.402	0.406	0.372	0.451	0.417	0.391
Fair4	0.273	0.191	0.232	0.229	0.256	0.264	0.190	0.268	0.222	0.199	0.281	0.266	0.191
Fair5	0.329	0.284	0.397	0.364	0.361	0.357	0.282	0.380	0.362	0.328	0.385	0.387	0.306
Fair6	0.366	0.331	0.453	0.463	0.459	0.450	0.425	0.502	0.476	0.418	0.493	0.448	0.409
Sat1	0.260	0.209	0.312	0.345	0.355	0.349	0.261	0.325	0.353	0.333	0.394	0.288	0.279
Sat2	0.311	0.304	0.348	0.360	0.341	0.325	0.337	0.377	0.412	0.386	0.434	0.400	0.304
Sat3	0.301	0.336	0.356	0.385	0.392	0.353	0.350	0.417	0.440	0.424	0.461	0.404	0.336
Sat4	0.337	0.340	0.322	0.356	0.327	0.303	0.423	0.402	0.407	0.387	0.392	0.353	0.285
SM1	0.153	0.157	0.228	0.233	0.190	0.211	0.273	0.239	0.202	0.189	0.270	0.202	0.198
SM2	0.223	0.185	0.234	0.270	0.238	0.250	0.296	0.245	0.255	0.249	0.310	0.254	0.275
SM3	0.158	0.140	0.125	0.164	0.123	0.138	0.199	0.151	0.109	0.143	0.208	0.127	0.209

	RelM12	Fair1	Fair2	Fair3	Fair4	Fair5	Fair6	Sat1	Sat2	Sat3	Sat4	SM1	SM2
Com1	0.320	0.419	0.398	0.360	0.196	0.259	0.402	0.279	0.306	0.315	0.314	0.199	0.221
Com2	0.189	0.307	0.280	0.329	0.163	0.232	0.327	0.270	0.283	0.256	0.282	0.149	0.241
Com3	0.212	0.273	0.223	0.180	0.160	0.106	0.240	0.237	0.217	0.206	0.188	0.173	0.158
Com4	0.359	0.414	0.423	0.330	0.216	0.292	0.409	0.342	0.400	0.347	0.310	0.256	0.295
Com5	0.329	0.434	0.440	0.348	0.228	0.266	0.455	0.351	0.342	0.288	0.281	0.224	0.247
Com6	0.365	0.416	0.428	0.347	0.260	0.318	0.426	0.387	0.347	0.338	0.318	0.240	0.261

Com7	0.321	0.436	0.423	0.331	0.218	0.257	0.394	0.312	0.341	0.331	0.272	0.192	0.215
Com8	0.300	0.400	0.342	0.279	0.148	0.224	0.345	0.260	0.303	0.257	0.270	0.205	0.213
Com9	0.209	0.287	0.324	0.222	0.223	0.177	0.308	0.219	0.233	0.188	0.168	0.236	0.232
Com10	0.121	0.245	0.214	0.199	0.083	0.113	0.173	0.127	0.166	0.131	0.134	0.178	0.159
RelC1	0.337	0.326	0.311	0.359	0.222	0.215	0.316	0.262	0.311	0.275	0.242	0.180	0.240
RelC2	0.392	0.372	0.297	0.390	0.214	0.208	0.340	0.266	0.342	0.301	0.276	0.198	0.205
RelC3	0.387	0.393	0.322	0.419	0.243	0.304	0.387	0.288	0.355	0.351	0.369	0.150	0.215
RelC4	0.348	0.366	0.317	0.392	0.273	0.329	0.366	0.260	0.311	0.301	0.337	0.153	0.223
RelC5	0.358	0.334	0.272	0.372	0.191	0.284	0.331	0.209	0.304	0.336	0.340	0.157	0.185
RelM1	0.532	0.459	0.406	0.445	0.232	0.397	0.453	0.312	0.348	0.356	0.322	0.228	0.234
RelM2	0.621	0.512	0.400	0.446	0.229	0.364	0.463	0.345	0.360	0.385	0.356	0.233	0.270
RelM3	0.662	0.479	0.440	0.393	0.256	0.361	0.459	0.355	0.341	0.392	0.327	0.190	0.238
RelM4	0.704	0.459	0.424	0.396	0.264	0.357	0.450	0.349	0.325	0.353	0.303	0.211	0.250
RelM5	0.567	0.508	0.342	0.388	0.190	0.282	0.425	0.261	0.337	0.350	0.423	0.273	0.296
RelM6	0.600	0.511	0.427	0.402	0.268	0.380	0.502	0.325	0.377	0.417	0.402	0.239	0.245
RelM7	0.565	0.532	0.452	0.406	0.222	0.362	0.476	0.353	0.412	0.440	0.407	0.202	0.255
RelM8	0.591	0.454	0.372	0.372	0.199	0.328	0.418	0.333	0.386	0.424	0.387	0.189	0.249
RelM9	0.789	0.534	0.456	0.451	0.281	0.385	0.493	0.394	0.434	0.461	0.392	0.270	0.310
RelM10	0.619	0.466	0.430	0.417	0.266	0.387	0.448	0.288	0.400	0.404	0.353	0.202	0.254
RelM11	0.682	0.480	0.392	0.391	0.191	0.306	0.409	0.279	0.304	0.336	0.285	0.198	0.275
RelM12	1.000	0.482	0.410	0.404	0.277	0.362	0.476	0.382	0.407	0.409	0.351	0.222	0.248
Fair1	0.482	1.000	0.590	0.556	0.287	0.445	0.590	0.336	0.548	0.482	0.473	0.315	0.335
Fair2	0.410	0.590	1.000	0.445	0.404	0.445	0.663	0.396	0.500	0.480	0.352	0.267	0.336
Fair3	0.404	0.556	0.445	1.000	0.355	0.454	0.510	0.344	0.430	0.406	0.410	0.244	0.300
Fair4	0.277	0.287	0.404	0.355	1.000	0.280	0.433	0.305	0.352	0.353	0.233	0.112	0.123
Fair5	0.362	0.445	0.445	0.454	0.280	1.000	0.593	0.324	0.379	0.377	0.296	0.185	0.232
Fair6	0.476	0.590	0.663	0.510	0.433	0.593	1.000	0.469	0.496	0.517	0.406	0.251	0.304
Sat1	0.382	0.336	0.396	0.344	0.305	0.324	0.469	1.000	0.578	0.630	0.432	0.165	0.198
Sat2	0.407	0.548	0.500	0.430	0.352	0.379	0.496	0.578	1.000	0.828	0.590	0.227	0.261
Sat3	0.409	0.482	0.480	0.406	0.353	0.377	0.517	0.630	0.828	1.000	0.606	0.202	0.280
Sat4	0.351	0.473	0.352	0.410	0.233	0.296	0.406	0.432	0.590	0.606	1.000	0.232	0.279
SM1	0.222	0.315	0.267	0.244	0.112	0.185	0.251	0.165	0.227	0.202	0.232	1.000	0.732
SM2	0.248	0.335	0.336	0.300	0.123	0.232	0.304	0.198	0.261	0.280	0.279	0.732	1.000
SM3	0.156	0.172	0.144	0.119	0.083	0.070	0.175	0.120	0.195	0.174	0.183	0.613	0.553

	SM3
Com1	0.101
Com2	0.154
Com3	0.121
Com4	0.214
Com5	0.157
Com6	0.132
Com7	0.141
Com8	0.139
Com9	0.202
Com10	0.154
RelC1	0.184
RelC2	0.148
RelC3	0.146
RelC4	0.158
RelC5	0.140
RelM1	0.125
RelM2	0.164
RelM3	0.123
RelM4	0.138
RelM5	0.199
RelM6	0.151
RelM7	0.109
RelM8	0.143
RelM9	0.208
RelM10	0.127
RelM11	0.209
RelM12	0.156
Fair1	0.172
Fair2	0.144
Fair3	0.119
Fair4	0.083
Fair5	0.070

```
Fair6   0.175
Sat1    0.120
Sat2    0.195
Sat3    0.174
Sat4    0.183
SM1     0.613
SM2     0.553
SM3     1.000

> eigenSigma = eigen(Sigma_hat); ls(eigenSigma)
[1] "values"  "vectors"
> eigenSigma$values
 [1] 14.9024012  3.4419990  2.0981709  1.9522247  1.9047772  1.1634482  1.0084527
 [8]  0.9387470  0.8322382  0.7987125  0.7269172  0.6780117  0.6376820  0.5990323
[15]  0.5528033  0.5087344  0.4880584  0.4747871  0.4679052  0.4437166  0.4131693
[22]  0.3845121  0.3784929  0.3567195  0.3457392  0.3305095  0.3247698  0.2996803
[29]  0.2900113  0.2752323  0.2681413  0.2498753  0.2285067  0.2161325  0.2016001
[36]  0.1932987  0.1838430  0.1618372  0.1535005  0.1256088
> lambda_hat = eigenSigma$values
> lambda_hat/40 # Proportions of explained variance
 [1] 0.372560030 0.086049975 0.052454272 0.048805617 0.047619431 0.029086206
 [7] 0.025211317 0.023468675 0.020805954 0.019967812 0.018172930 0.016950292
[13] 0.015942050 0.014975808 0.013820082 0.012718359 0.012201460 0.011869677
[19] 0.011697629 0.011092915 0.010329233 0.009612802 0.009462323 0.008917986
[25] 0.008643479 0.008262737 0.008119246 0.007492007 0.007250282 0.006880807
[31] 0.006703532 0.006246883 0.005712668 0.005403312 0.005040002 0.004832469
[37] 0.004596076 0.004045931 0.003837513 0.003140219
> cumsum(lambda_hat/40) # Cumulative sum
 [1] 0.3725600 0.4586100 0.5110643 0.5598699 0.6074893 0.6365755 0.6617868 0.6852555
 [9] 0.7060615 0.7260293 0.7442022 0.7611525 0.7770946 0.7920704 0.8058905 0.8186088
[17] 0.8308103 0.8426799 0.8543776 0.8654705 0.8757997 0.8854125 0.8948749 0.9037928
[25] 0.9124363 0.9206991 0.9288183 0.9363103 0.9435606 0.9504414 0.9571449 0.9633918
[33] 0.9691045 0.9745078 0.9795478 0.9843803 0.9889763 0.9930223 0.9968598 1.0000000

> # Principal Components
> Z = scale(quest) # Standardize columns (This is a 500 x 40 matrix)
> C_hat = eigenSigma$vectors
> Y_hat = Z %*% C_hat # Sample principal components
> # Looking at the variance-covariance matrix of the principal components,
> round(var(Y_hat), 4) # Should equal D
        [,1]  [,2]   [,3]   [,4]   [,5]   [,6]   [,7]   [,8]   [,9]  [,10]  [,11]
 [1,] 14.9024 0.000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
 [2,]  0.0000 3.442 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
 [3,]  0.0000 0.000 2.0982 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
 [4,]  0.0000 0.000 0.0000 1.9522 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
 [5,]  0.0000 0.000 0.0000 0.0000 1.9048 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
 [6,]  0.0000 0.000 0.0000 0.0000 0.0000 1.1634 0.0000 0.0000 0.0000 0.0000 0.0000
 [7,]  0.0000 0.000 0.0000 0.0000 0.0000 0.0000 1.0085 0.0000 0.0000 0.0000 0.0000
 [8,]  0.0000 0.000 0.0000 0.0000 0.0000 0.0000 0.0000 0.9387 0.0000 0.0000 0.0000
 [9,]  0.0000 0.000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.8322 0.0000 0.0000
[10,]  0.0000 0.000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.7987 0.0000
[11,]  0.0000 0.000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.7269
[12,]  0.0000 0.000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000

Skipping ...

[34,] 0.2161 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
[35,] 0.0000 0.2016 0.0000 0.0000 0.0000 0.0000 0.0000
[36,] 0.0000 0.0000 0.1933 0.0000 0.0000 0.0000 0.0000
[37,] 0.0000 0.0000 0.0000 0.1838 0.0000 0.0000 0.0000
[38,] 0.0000 0.0000 0.0000 0.0000 0.1618 0.0000 0.0000
[39,] 0.0000 0.0000 0.0000 0.0000 0.0000 0.1535 0.0000
[40,] 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.1256
```

```
> # I would say either keep just one, or keep 7
> y = Y_hat[,1:7] # The first 7 components explain about 70% of the variance.
> zy = cor(Z,y); round(zy,3)
         [,1]   [,2]   [,3]   [,4]   [,5]   [,6]   [,7]
Com1   -0.604 -0.522  0.231 -0.101  0.011  0.019  0.007
Com2   -0.499 -0.463  0.042 -0.134 -0.051 -0.075  0.218
Com3   -0.406 -0.396  0.162 -0.086 -0.073 -0.309  0.262
Com4   -0.640 -0.460  0.187 -0.014 -0.040 -0.072  0.003
Com5   -0.628 -0.534  0.220 -0.069  0.017  0.053 -0.032
Com6   -0.618 -0.512  0.221 -0.004  0.070  0.039 -0.072
Com7   -0.598 -0.512  0.202 -0.078  0.048  0.016 -0.133
Com8   -0.534 -0.474  0.166 -0.064 -0.017  0.057 -0.321
Com9   -0.425 -0.371 -0.027 -0.013 -0.098  0.258 -0.279
Com10  -0.339 -0.416  0.111 -0.093 -0.177 -0.153  0.435
RelC1  -0.517 -0.045 -0.489 -0.402 -0.164 -0.032 -0.195
RelC2  -0.550 -0.007 -0.475 -0.426 -0.150 -0.058 -0.195
RelC3  -0.610  0.002 -0.383 -0.403 -0.063 -0.003 -0.037
RelC4  -0.568  0.066 -0.417 -0.375 -0.079  0.087  0.166
RelC5  -0.550  0.097 -0.312 -0.340 -0.093 -0.027  0.267
RelM1  -0.669  0.320  0.154  0.014 -0.026  0.128  0.184
RelM2  -0.744  0.289  0.213 -0.036 -0.097  0.013  0.085
RelM3  -0.731  0.334  0.241 -0.055 -0.063 -0.004 -0.052
RelM4  -0.737  0.335  0.264 -0.066 -0.105  0.002 -0.011
RelM5  -0.692  0.250  0.190  0.016 -0.179 -0.104  0.095
RelM6  -0.719  0.359  0.187  0.052 -0.018 -0.002  0.074
RelM7  -0.704  0.327  0.153  0.048  0.033 -0.051  0.086
RelM8  -0.684  0.288  0.186  0.012 -0.030 -0.118 -0.125
RelM9  -0.776  0.298  0.121  0.005 -0.078 -0.068 -0.223
RelM10 -0.698  0.351  0.126 -0.021 -0.026  0.057 -0.036
RelM11 -0.714  0.261  0.143 -0.163 -0.215 -0.086 -0.077
RelM12 -0.736  0.310  0.162 -0.038 -0.053 -0.051 -0.225
Fair1  -0.732 -0.011 -0.043  0.132  0.101  0.108  0.088
Fair2  -0.667 -0.084 -0.064  0.191  0.215  0.295 -0.014
Fair3  -0.640  0.014 -0.190  0.013  0.125  0.219  0.202
Fair4  -0.417 -0.023 -0.190  0.082  0.308  0.293 -0.028
Fair5  -0.543  0.065 -0.096  0.128  0.250  0.404  0.205
Fair6  -0.716 -0.025 -0.096  0.163  0.253  0.304  0.056
Sat1   -0.553 -0.064 -0.145  0.196  0.378 -0.265 -0.142
Sat2   -0.638 -0.046 -0.285  0.269  0.397 -0.289 -0.054
Sat3   -0.639  0.023 -0.257  0.284  0.416 -0.332 -0.044
Sat4   -0.578  0.009 -0.231  0.204  0.231 -0.344  0.104
SM1    -0.386 -0.125 -0.234  0.544 -0.542  0.065  0.010
SM2    -0.444 -0.108 -0.257  0.502 -0.483  0.066  0.018
SM3    -0.284 -0.115 -0.279  0.452 -0.561 -0.080 -0.053
>
> # This is hard to look at. Sign of an eigenvector is arbitrary.
```

```
> # Sign of an eigenvector is arbitrary.
> y = -Y_hat[,1:7]
> zy = cor(Z,y); round(zy,3)

        [,1]   [,2]   [,3]   [,4]   [,5]   [,6]   [,7]
Com1   0.604  0.522 -0.231  0.101 -0.011 -0.019 -0.007
Com2   0.499  0.463 -0.042  0.134  0.051  0.075 -0.218
Com3   0.406  0.396 -0.162  0.086  0.073  0.309 -0.262
Com4   0.640  0.460 -0.187  0.014  0.040  0.072 -0.003
Com5   0.628  0.534 -0.220  0.069 -0.017 -0.053  0.032
Com6   0.618  0.512 -0.221  0.004 -0.070 -0.039  0.072
Com7   0.598  0.512 -0.202  0.078 -0.048 -0.016  0.133
Com8   0.534  0.474 -0.166  0.064  0.017 -0.057  0.321
Com9   0.425  0.371  0.027  0.013  0.098 -0.258  0.279
Com10  0.339  0.416 -0.111  0.093  0.177  0.153 -0.435
RelC1  0.517  0.045  0.489  0.402  0.164  0.032  0.195
RelC2  0.550  0.007  0.475  0.426  0.150  0.058  0.195
RelC3  0.610 -0.002  0.383  0.403  0.063  0.003  0.037
RelC4  0.568 -0.066  0.417  0.375  0.079 -0.087 -0.166
RelC5  0.550 -0.097  0.312  0.340  0.093  0.027 -0.267
RelM1  0.669 -0.320 -0.154 -0.014  0.026 -0.128 -0.184
RelM2  0.744 -0.289 -0.213  0.036  0.097 -0.013 -0.085
RelM3  0.731 -0.334 -0.241  0.055  0.063  0.004  0.052
RelM4  0.737 -0.335 -0.264  0.066  0.105 -0.002  0.011
RelM5  0.692 -0.250 -0.190 -0.016  0.179  0.104 -0.095
RelM6  0.719 -0.359 -0.187 -0.052  0.018  0.002 -0.074
RelM7  0.704 -0.327 -0.153 -0.048 -0.033  0.051 -0.086
RelM8  0.684 -0.288 -0.186 -0.012  0.030  0.118  0.125
RelM9  0.776 -0.298 -0.121 -0.005  0.078  0.068  0.223
RelM10 0.698 -0.351 -0.126  0.021  0.026 -0.057  0.036
RelM11 0.714 -0.261 -0.143  0.163  0.215  0.086  0.077
RelM12 0.736 -0.310 -0.162  0.038  0.053  0.051  0.225
Fair1  0.732  0.011  0.043 -0.132 -0.101 -0.108 -0.088
Fair2  0.667  0.084  0.064 -0.191 -0.215 -0.295  0.014
Fair3  0.640 -0.014  0.190 -0.013 -0.125 -0.219 -0.202
Fair4  0.417  0.023  0.190 -0.082 -0.308 -0.293  0.028
Fair5  0.543 -0.065  0.096 -0.128 -0.250 -0.404 -0.205
Fair6  0.716  0.025  0.096 -0.163 -0.253 -0.304 -0.056
Sat1   0.553  0.064  0.145 -0.196 -0.378  0.265  0.142
Sat2   0.638  0.046  0.285 -0.269 -0.397  0.289  0.054
Sat3   0.639 -0.023  0.257 -0.284 -0.416  0.332  0.044
Sat4   0.578 -0.009  0.231 -0.204 -0.231  0.344 -0.104
SM1    0.386  0.125  0.234 -0.544  0.542 -0.065 -0.010
SM2    0.444  0.108  0.257 -0.502  0.483 -0.066 -0.018
SM3    0.284  0.115  0.279 -0.452  0.561  0.080  0.053
```

```
> # Now principal components the easy way
> # help(prcomp) # Better than princomp
> pc = prcomp(quest, scale = T)
> ls(pc) # What's in pc?
[1] "center"   "rotation" "scale"    "sdev"     "x"
>     # center has the sample means before standardization
>     # rotation is C-hat
>     # scale has the standard deviations before standardization
>     # sdev has estimated standard deviations of the components
>     # x is a matrix of the principal components: Y-hat
>
> pc$sdev^2 # Eigenvalues
 [1] 14.9024012  3.4419990  2.0981709  1.9522247  1.9047772  1.1634482  1.0084527
 [8]  0.9387470  0.8322382  0.7987125  0.7269172  0.6780117  0.6376820  0.5990323
[15]  0.5528033  0.5087344  0.4880584  0.4747871  0.4679052  0.4437166  0.4131693
[22]  0.3845121  0.3784929  0.3567195  0.3457392  0.3305095  0.3247698  0.2996803
[29]  0.2900113  0.2752323  0.2681413  0.2498753  0.2285067  0.2161325  0.2016001
[36]  0.1932987  0.1838430  0.1618372  0.1535005  0.1256088
> lambda_hat # For comparison
 [1] 14.9024012  3.4419990  2.0981709  1.9522247  1.9047772  1.1634482  1.0084527
 [8]  0.9387470  0.8322382  0.7987125  0.7269172  0.6780117  0.6376820  0.5990323
[15]  0.5528033  0.5087344  0.4880584  0.4747871  0.4679052  0.4437166  0.4131693
[22]  0.3845121  0.3784929  0.3567195  0.3457392  0.3305095  0.3247698  0.2996803
[29]  0.2900113  0.2752323  0.2681413  0.2498753  0.2285067  0.2161325  0.2016001
[36]  0.1932987  0.1838430  0.1618372  0.1535005  0.1256088
> dim(pc$x) # x is Y-hat
[1] 500  40


> pc7 = prcomp(quest, scale = T, rank = 7) # Retain seven principal components
> round(pc7$rotation, 4) # Just the first 7 columns of C-hat
          PC1     PC2     PC3     PC4     PC5     PC6     PC7
Com1   0.1564 -0.2814 -0.1597  0.0723 -0.0079  0.0179 -0.0066
Com2   0.1292 -0.2498 -0.0291  0.0956  0.0371 -0.0697 -0.2170
Com3   0.1051 -0.2135 -0.1119  0.0613  0.0528 -0.2864 -0.2610
Com4   0.1659 -0.2478 -0.1292  0.0100  0.0287 -0.0665 -0.0029
Com5   0.1627 -0.2878 -0.1522  0.0496 -0.0123  0.0492  0.0318
Com6   0.1600 -0.2761 -0.1529  0.0031 -0.0508  0.0360  0.0712
Com7   0.1549 -0.2758 -0.1396  0.0560 -0.0348  0.0152  0.1328
Com8   0.1383 -0.2554 -0.1148  0.0459  0.0123  0.0528  0.3199
Com9   0.1100 -0.2001  0.0188  0.0091  0.0712  0.2389  0.2777
Com10  0.0879 -0.2244 -0.0764  0.0667  0.1281 -0.1420 -0.4327
RelC1  0.1338 -0.0242  0.3373  0.2874  0.1187 -0.0295  0.1939
RelC2  0.1424 -0.0036  0.3281  0.3046  0.1084 -0.0538  0.1939
RelC3  0.1581  0.0011  0.2643  0.2885  0.0454 -0.0026  0.0367
RelC4  0.1472  0.0358  0.2877  0.2680  0.0570  0.0806 -0.1651
RelC5  0.1424  0.0521  0.2151  0.2435  0.0672 -0.0249 -0.2656
RelM1  0.1733  0.1727 -0.1066 -0.0102  0.0190  0.1183 -0.1835
RelM2  0.1927  0.1556 -0.1470  0.0258  0.0706  0.0125 -0.0842
RelM3  0.1895  0.1799 -0.1667  0.0393  0.0455 -0.0038  0.0519
RelM4  0.1910  0.1806 -0.1826  0.0473  0.0763  0.0019  0.0112
RelM5  0.1794  0.1346 -0.1315 -0.0113  0.1296 -0.0965 -0.0943
RelM6  0.1862  0.1933 -0.1293 -0.0375  0.0130 -0.0020 -0.0738
RelM7  0.1823  0.1761 -0.1057 -0.0345 -0.0242 -0.0471 -0.0853
RelM8  0.1771  0.1550 -0.1283 -0.0088  0.0221 -0.1098  0.1247
RelM9  0.2009  0.1605 -0.0838 -0.0035  0.0562 -0.0633  0.2216
RelM10 0.1809  0.1890 -0.0867  0.0153  0.0192  0.0526  0.0355
RelM11 0.1849  0.1408 -0.0984  0.1166  0.1557 -0.0797  0.0765
RelM12 0.1906  0.1669 -0.1118  0.0269  0.0383 -0.0475  0.2239
Fair1  0.1897 -0.0059  0.0298 -0.0941 -0.0729  0.1006 -0.0874
Fair2  0.1729 -0.0451  0.0444 -0.1365 -0.1558  0.2733  0.0144
Fair3  0.1659  0.0074  0.1309 -0.0093 -0.0905  0.2032 -0.2012
Fair4  0.1079 -0.0124  0.1314 -0.0588 -0.2232  0.2719  0.0279
Fair5  0.1408  0.0349  0.0665 -0.0916 -0.1808  0.3741 -0.2044
Fair6  0.1855 -0.0137  0.0661 -0.1166 -0.1832  0.2821 -0.0559
Sat1   0.1433 -0.0347  0.1001 -0.1403 -0.2742 -0.2460  0.1418
Sat2   0.1654 -0.0247  0.1966 -0.1929 -0.2879 -0.2675  0.0537
```

```
Sat3    0.1654  0.0124  0.1774 -0.2034 -0.3016 -0.3074  0.0437
Sat4    0.1496  0.0046  0.1592 -0.1458 -0.1676 -0.3185 -0.1036
SM1     0.1000 -0.0674  0.1613 -0.3891  0.3930  0.0602 -0.0095
SM2     0.1150 -0.0582  0.1775 -0.3594  0.3502  0.0609 -0.0175
SM3     0.0736 -0.0620  0.1928 -0.3236  0.4062 -0.0739  0.0529
> head(pc7$x) # There should be 7 columns
            PC1        PC2         PC3        PC4        PC5        PC6         PC7
[1,] -2.5719438  3.8011128 -2.94534355  1.5277304  0.3738404 -1.9978874 -0.39489617
[2,]  4.8217546 -0.4496531  0.96123661  0.1024910 -0.7486379  0.5081813 -0.68676015
[3,]  2.5980433 -1.4661589  0.06945463  1.2610963  1.7768526  0.5199767  0.43999477
[4,] -6.0176716 -2.4164930 -1.65105390 -4.7049320  0.7849950  0.9358714  0.26234943
[5,]  0.3889181  1.2399474 -0.61418564 -0.4463095 -1.8600463 -0.8687935  0.63108469
[6,] -2.7582399  1.7312678  3.27706331  0.6921211 -0.8078170 -0.6490922 -0.08667956
>
> round(cor(quest,pc7$x),4) # Correlations of variables with components
          PC1     PC2     PC3     PC4     PC5     PC6     PC7
Com1   0.6036 -0.5221 -0.2314  0.1010 -0.0108  0.0193 -0.0066
Com2   0.4986 -0.4635 -0.0421  0.1336  0.0513 -0.0752 -0.2179
Com3   0.4056 -0.3960 -0.1620  0.0856  0.0729 -0.3090 -0.2621
Com4   0.6404 -0.4597 -0.1872  0.0140  0.0396 -0.0717 -0.0029
Com5   0.6283 -0.5339 -0.2204  0.0692 -0.0169  0.0531  0.0319
Com6   0.6178 -0.5122 -0.2214  0.0043 -0.0702  0.0388  0.0715
Com7   0.5979 -0.5116 -0.2022  0.0783 -0.0480  0.0164  0.1333
Com8   0.5340 -0.4739 -0.1663  0.0641  0.0169  0.0570  0.3213
Com9   0.4247 -0.3712  0.0273  0.0127  0.0983  0.2576  0.2789
Com10  0.3393 -0.4163 -0.1107  0.0932  0.1768 -0.1532 -0.4346
RelC1  0.5167 -0.0449  0.4886  0.4015  0.1638 -0.0318  0.1947
RelC2  0.5497 -0.0066  0.4752  0.4256  0.1497 -0.0580  0.1947
RelC3  0.6102  0.0020  0.3829  0.4032  0.0627 -0.0029  0.0369
RelC4  0.5684  0.0664  0.4168  0.3745  0.0786  0.0869 -0.1658
RelC5  0.5498  0.0966  0.3115  0.3403  0.0927 -0.0268 -0.2667
RelM1  0.6688  0.3203 -0.1544 -0.0143  0.0262  0.1277 -0.1843
RelM2  0.7438  0.2886 -0.2130  0.0360  0.0974  0.0134 -0.0846
RelM3  0.7314  0.3338 -0.2414  0.0550  0.0628 -0.0041  0.0521
RelM4  0.7373  0.3351 -0.2645  0.0661  0.1053  0.0021  0.0112
RelM5  0.6924  0.2498 -0.1905 -0.0157  0.1789 -0.1041 -0.0947
RelM6  0.7189  0.3586 -0.1872 -0.0524  0.0179 -0.0021 -0.0741
RelM7  0.7039  0.3267 -0.1530 -0.0481 -0.0334 -0.0508 -0.0857
RelM8  0.6837  0.2875 -0.1859 -0.0123  0.0305 -0.1184  0.1253
RelM9  0.7757  0.2978 -0.1214 -0.0049  0.0775 -0.0682  0.2225
RelM10 0.6983  0.3506 -0.1257  0.0213  0.0265  0.0568  0.0356
RelM11 0.7136  0.2612 -0.1425  0.1629  0.2149 -0.0859  0.0768
RelM12 0.7359  0.3097 -0.1619  0.0375  0.0528 -0.0513  0.2249
Fair1  0.7324 -0.0110  0.0431 -0.1315 -0.1006  0.1085 -0.0878
Fair2  0.6673 -0.0837  0.0643 -0.1907 -0.2150  0.2947  0.0144
Fair3  0.6404  0.0137  0.1896 -0.0130 -0.1249  0.2192 -0.2021
Fair4  0.4166 -0.0231  0.1904 -0.0822 -0.3081  0.2933  0.0280
Fair5  0.5434  0.0648  0.0963 -0.1279 -0.2495  0.4036 -0.2052
Fair6  0.7160 -0.0254  0.0957 -0.1629 -0.2529  0.3042 -0.0561
Sat1   0.5534 -0.0644  0.1451 -0.1961 -0.3785 -0.2654  0.1424
Sat2   0.6384 -0.0457  0.2848 -0.2695 -0.3973 -0.2886  0.0539
Sat3   0.6386  0.0230  0.2570 -0.2842 -0.4163 -0.3315  0.0438
Sat4   0.5776  0.0086  0.2306 -0.2038 -0.2313 -0.3436 -0.1040
SM1    0.3859 -0.1251  0.2337 -0.5437  0.5425  0.0649 -0.0096
SM2    0.4438 -0.1079  0.2571 -0.5022  0.4834  0.0657 -0.0176
SM3    0.2842 -0.1151  0.2792 -0.4521  0.5606 -0.0797  0.0531
```

```
> # Look at the a priori scale Commitment to the Organization
> pcOrg = prcomp(quest[,1:10], scale = T) # First 10 questions
> pcOrg$sdev^2  # Eigenvalues
 [1] 5.4007298 0.9906004 0.8413945 0.6100864 0.5470061 0.4211846 0.3858363 0.3158245
 [9] 0.2701758 0.2171614
> round(cor(quest[,1:10],pcOrg$x[,1:3]),4) # Correlations of variables with components
         PC1     PC2     PC3
Com1  0.8352 -0.0371  0.0532
Com2  0.6863  0.2103  0.3608
Com3  0.5881  0.5608 -0.1965
Com4  0.8111  0.0330 -0.1478
Com5  0.8548 -0.0805 -0.1674
Com6  0.8283 -0.0924 -0.2675
Com7  0.8149 -0.1472 -0.2289
Com8  0.7396 -0.3598 -0.0498
Com9  0.5554 -0.4146  0.5977
Com10 0.5405  0.5397  0.3691
>
> # To be continued.
```