

STA5221 - Homework 3

Huong Tran

2022-11-18

**Exercise 1:**

Intuitively, all variables are positively correlated, we can expect that the entries of covariance and correlation matrix are positive. The transformation that are considered in this exercise will maintain the direction of correlation and covariance, therefore in the following models, we are sure that coefficients in the first PC will have the same signs and it can be interpreted as the weighted mean of each performance. Therefore, countries with larger scores in the first PC will have a higher rank.

**i. Time taken to complete the race:**

*Best 3 countries by PCA with Covariance matrix (TIME)*

Country	Prin1
WSamoa	132.895
Mauritiu	88.014
Cookis	60.388

*Best 3 countries by PCA with Correlation matrix (TIME)*

Country	Prin1
WSamoa	8.33288
Cookis	6.07728
Mauritiu	4.23385

**ii. Speed:**

*Best 3 countries by PCA with Covariance matrix (SPEED)*

Country	Prin1
USA	1.49962
USSR	1.45505
GDR	1.35933

*Best 3 countries by PCA with Correlation Matrix (SPEED)*

Country	Prin1
USSR	3.82147
GDR	3.79442
USA	3.70676

**iii. Time taken to run one foot distance:**

*Best 3 countries by PCA with Covariance matrix (FOOT/DISTANCE)*

Country	Prin1
WSamoa	0.066556
Mauritiu	0.039695
Cookis	0.030676

*Best 3 countries by PCA with Correlation Matrix (FOOT/DISTANCE)*

Country	Prin1
WSamoa	8.33288
Cookis	6.07728
Mauritiu	4.23385

For i, variance of marathon is much large that other variables, i.e, the large amount of total variability is caused by the variability of marathon, therefore the first principle give excessive weight for this variable, therefore using correlation matrix in this case is better, since it standardized variance of every variable, and give a unit-free measure. However, by using correlation matrix, we are actually forcing all variables to have equal variance and loose the information of variability of variance in each groups of athletes.

One of the reason for the above issue is because each of variable has different footing scale, for example, marathon athletes will need more time to finish the race than an athlete of 100 meters run. Therefore the nature of running strategies of each runners should be taken into account. And ii actually consider speed, measured by meter/seconds, which represent well the strategies of each runner, for example a marathon athlete usually is slower because he needs to have a long term persistence, while a 100 meter runner should run faster. Covariance matrix of this new data set is better, in the sense that none of the variance is a significantly large. The first principle using covariance matrix explains 83.7% of variability among data. Using correlation matrix does not really improve the percentage of variability that the first principle component can explain, which is 83.56%, we also loose the information of different variance between each of variables. In addition, the variance of 7 variables are appropriately equal, therefore result when using correlation matrix will be nearly identical.

Data measured time taken to run one foot distance actually has the same meaning of the above dataset, putting variable in the same scale. In fact, an athlete has faster speed means that he needs shorter time to run 1 foot. However, in this dataset, the variance of marathon is 10 times bigger than variance of other variables, which cause the same issue of using raw data. Again, standardizing help improve the weighted mean, but we have to trade off the information in variability.

In conclusion, the best criteria is using speed. Because PCA with covariance matrix give good approximation for weighted mean of 7 variables, and still keep all information of variability.

Note that, correlation matrix using raw data and correlation matrix for time taken to run one foot distance is identical, it is because we actually did a linear transformation in each of the variables:

$$\text{corr}(aX, bY) = \frac{\text{Cov}(aX, bY)}{\sqrt{\text{Var}(aX)}\sqrt{\text{Var}(bY)}} = \frac{a\text{Cov}(X, Y)b}{\sqrt{a^2\text{Cov}(X)}\sqrt{b^2\text{Cov}(Y)}} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}} = \text{Corr}(X, Y)$$

Also, this linear transformation makes the ranking of time taken to run one foot distance and that of time taken to complete the race are the identical.

# Exercise 1

## Women track record (TIME)

### PCA using Covariance matrix

#### The PRINCOMP Procedure

Observations	55
Variables	7

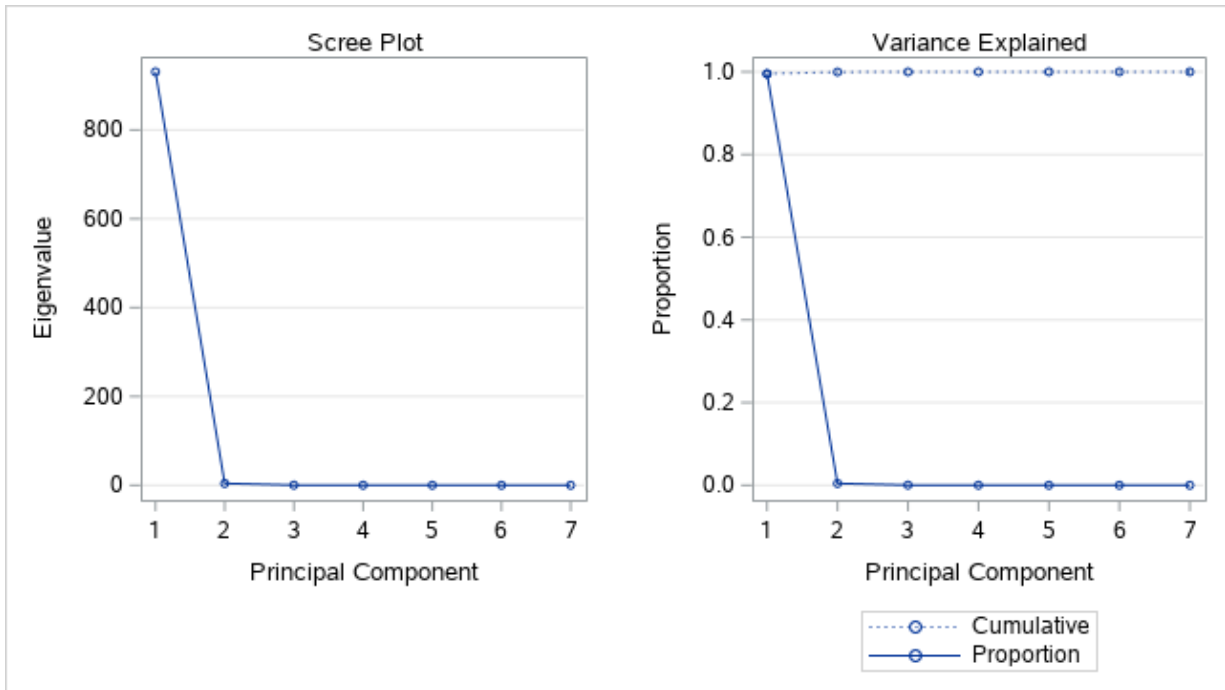
Simple Statistics							
	m100	m200	m400	m800	m1500	m3000	marathon
Mean	11.61854545	23.64163636	53.40581818	2.076363636	4.325454545	9.447636364	173.2532727
Std	0.45221028	1.11106016	2.67833675	0.108224110	0.332425532	0.824335790	30.4295449

Covariance Matrix							
	m100	m200	m400	m800	m1500	m3000	marathon
m100	0.2044941	0.4787135	1.0109549	0.0356131	0.1094933	0.2764854	9.4443604
m200	0.4787135	1.2344547	2.5501422	0.0870635	0.2579372	0.6501613	23.1786260
m400	1.0109549	2.5501422	7.1734877	0.2604123	0.7014529	1.7169047	57.4924621
m800	0.0356131	0.0870635	0.2604123	0.0117125	0.0324369	0.0770412	2.5663732
m1500	0.1094933	0.2579372	0.7014529	0.0324369	0.1105067	0.2655816	8.8807855
m3000	0.2764854	0.6501613	1.7169047	0.0770412	0.2655816	0.6795295	22.5716671
marathon	9.4443604	23.1786260	57.4924621	2.5663732	8.8807855	22.5716671	925.9572039

Total Variance	935.37138916
----------------	--------------

Eigenvalues of the Covariance Matrix				
	Eigenvalue	Difference	Proportion	Cumulative
1	930.865778	926.815068	0.9952	0.9952
2	4.050710	3.732028	0.0043	0.9995
3	0.318682	0.203468	0.0003	0.9999
4	0.115214	0.100924	0.0001	1.0000
5	0.014290	0.008433	0.0000	1.0000
6	0.005857	0.005000	0.0000	1.0000
7	0.000857		0.0000	1.0000

Eigenvectors							
	Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7
m100	0.010210	0.120418	0.326096	0.150066	0.925339	-.001656	0.016781
m200	0.025063	0.314711	0.879886	0.013991	-.353756	0.025278	0.012338
m400	0.062211	0.934172	-.327607	-.121781	0.013367	-.021835	-.025309
m800	0.002772	0.025565	-.037115	0.048685	-.015164	0.261819	0.962711
m1500	0.009578	0.038551	-.055357	0.339733	-.034288	0.899065	-.265416
m3000	0.024340	0.082216	-.087692	0.918837	-.130442	-.349305	0.040842
marathon	0.997349	-.069859	-.002240	-.020113	0.002138	-.000094	-.000027



Huong Tran - Assignment 3

**Exercise 1**  
**Women track record (TIME)**  
**PCA using Covariance matrix**  
**'The best 3 countries'**

country	Prin1
WSamoa	132.895
Mauritiu	88.014
Cookis	60.388

Huong Tran - Assignment 3

**Exercise 1**  
**Women track record (TIME)**  
**PCA using correlation matrix**

The PRINCOMP Procedure

Observations	55
Variables	7

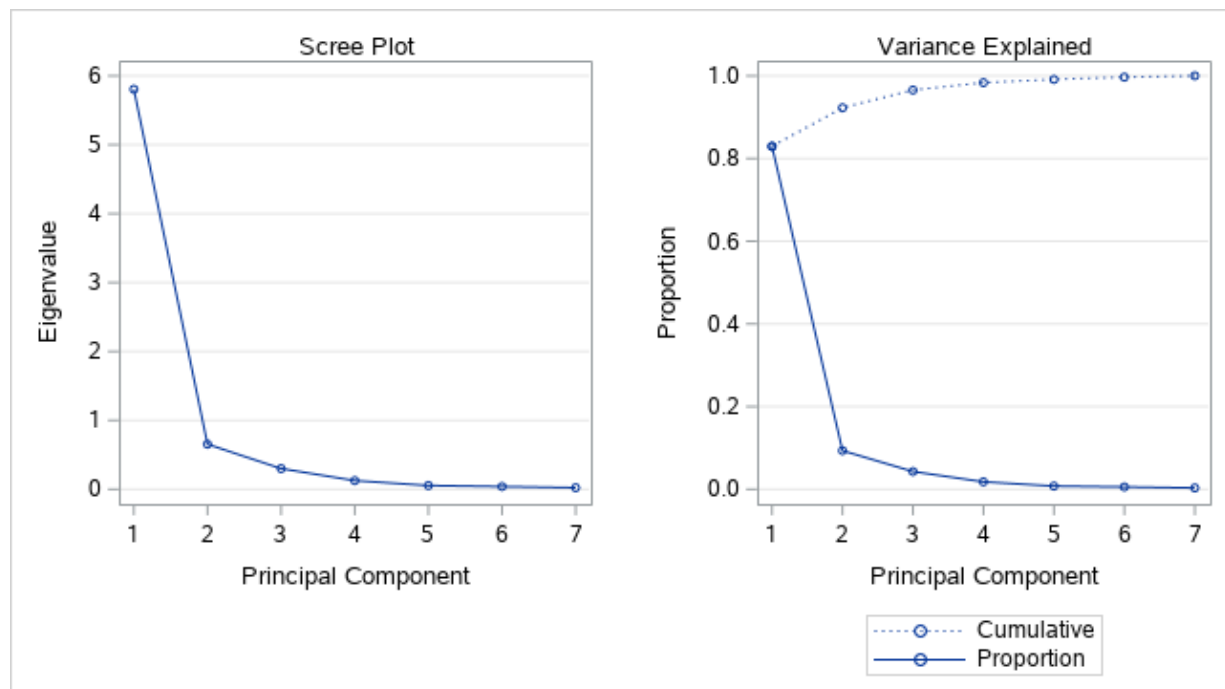
Simple Statistics							
	m100	m200	m400	m800	m1500	m3000	marathon
Mean	11.61854545	23.64163636	53.40581818	2.076363636	4.325454545	9.447636364	173.2532727
Std	0.45221028	1.11106016	2.67833675	0.108224110	0.332425532	0.824335790	30.4295449

Correlation Matrix							
	m100	m200	m400	m800	m1500	m3000	marathon
m100	1.0000	0.9528	0.8347	0.7277	0.7284	0.7417	0.6863
m200	0.9528	1.0000	0.8570	0.7241	0.6984	0.7099	0.6856
m400	0.8347	0.8570	1.0000	0.8984	0.7878	0.7776	0.7054

Correlation Matrix							
	m100	m200	m400	m800	m1500	m3000	marathon
m800	0.7277	0.7241	0.8984	1.0000	0.9016	0.8636	0.7793
m1500	0.7284	0.6984	0.7878	0.9016	1.0000	0.9692	0.8779
m3000	0.7417	0.7099	0.7776	0.8636	0.9692	1.0000	0.8998
marathon	0.6863	0.6856	0.7054	0.7793	0.8779	0.8998	1.0000

Eigenvalues of the Correlation Matrix				
	Eigenvalue	Difference	Proportion	Cumulative
1	5.80568576	5.15204024	0.8294	0.8294
2	0.65364552	0.35376309	0.0934	0.9228
3	0.29988243	0.17440494	0.0428	0.9656
4	0.12547749	0.07166058	0.0179	0.9835
5	0.05381692	0.01476763	0.0077	0.9912
6	0.03904928	0.01660668	0.0056	0.9968
7	0.02244260		0.0032	1.0000

Eigenvectors							
	Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7
m100	0.368356	0.490060	0.286012	-0.319386	0.231169	0.619825	0.052177
m200	0.365364	0.536580	0.229819	0.083302	0.041455	-0.710765	-0.109225
m400	0.381610	0.246538	-0.515367	0.347377	-0.572178	0.190946	0.208497
m800	0.384559	-0.155402	-0.584526	0.042076	0.620324	-0.019089	-0.315210
m1500	0.389104	-0.360409	-0.012912	-0.429539	0.030261	-0.231248	0.692562
m3000	0.388866	-0.347539	0.152728	-0.363120	-0.463355	0.009277	-0.598359
marathon	0.367004	-0.369208	0.484370	0.672497	0.130536	0.142281	0.069598



Huong Tran - Assignment 3

### Exercise 1

#### Women track record (TIME)

**PCA using correlation matrix**  
**'The best 3 countries'**

country	Prin1
WSamoa	8.33288
Cookis	6.07728
Mauritiu	4.23385

Huong Tran - Assignment 3

**Exercise 1**  
**Women record (SPEED)**  
**PCA using Covariance matrix**

The PRINCOMP Procedure

Observations	55
Variables	7

Simple Statistics							
	m100	m200	m400	m800	m1500	m3000	marathon
Mean	8.619563220	8.477681719	7.508259742	6.438315277	5.809893843	5.327650818	4.154344492
StD	0.331058113	0.391558869	0.375282162	0.329456996	0.403074818	0.416455353	0.567096453

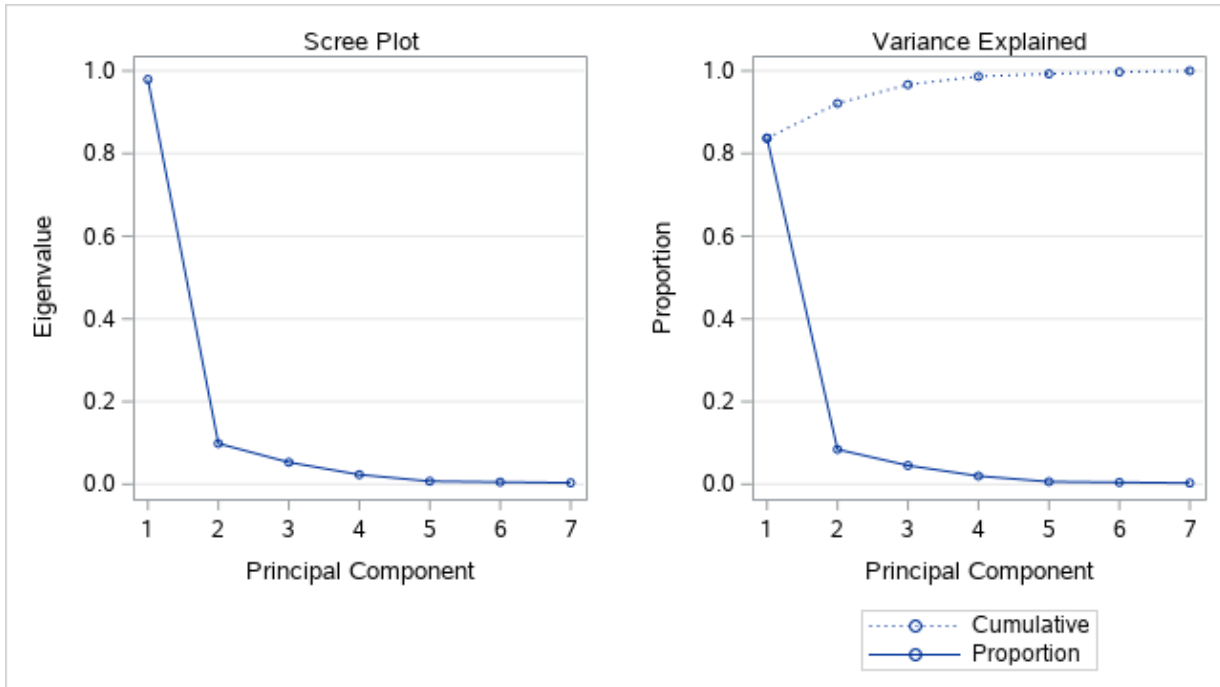
Covariance Matrix							
	m100	m200	m400	m800	m1500	m3000	marathon
m100	0.1095994742	0.1237832898	0.1038905488	0.0795460124	0.0991434510	0.1031927674	0.1348335245
m200	0.1237832898	0.1533183476	0.1264920245	0.0939865908	0.1136836024	0.1174087274	0.1582593575
m400	0.1038905488	0.1264920245	0.1408367011	0.1111527230	0.1217076143	0.1222224807	0.1518052167
m800	0.0795460124	0.0939865908	0.1111527230	0.1085419120	0.1220487003	0.1199176569	0.1468255042
m1500	0.0991434510	0.1136836024	0.1217076143	0.1220487003	0.1624693091	0.1617652421	0.1963723872
m3000	0.1031927674	0.1174087274	0.1222224807	0.1199176569	0.1617652421	0.1734350606	0.2097175807
marathon	0.1348335245	0.1582593575	0.1518052167	0.1468255042	0.1963723872	0.2097175807	0.3215983865

Total Variance	1.1697991911
----------------	--------------

Eigenvalues of the Covariance Matrix				
	Eigenvalue	Difference	Proportion	Cumulative
1	0.97910541	0.88053887	0.8370	0.8370
2	0.09856654	0.04551016	0.0843	0.9212
3	0.05305638	0.02981429	0.0454	0.9666
4	0.02324210	0.01603734	0.0199	0.9865
5	0.00720475	0.00198981	0.0062	0.9926
6	0.00521494	0.00180586	0.0045	0.9971
7	0.00340907		0.0029	1.0000

Eigenvectors							
	Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7
m100	0.290843	0.426967	-.250363	0.329132	0.155822	0.727883	-.089472
m200	0.341929	0.558190	-.320174	0.132018	0.099659	-.628763	0.215159
m400	0.338593	0.381781	0.320836	-.537041	-.476833	0.053809	-.343489
m800	0.305423	0.007925	0.475251	-.309251	0.504602	0.137220	0.558303
m1500	0.385868	-.197142	0.372461	0.362247	0.364538	-.225856	-.598658

Eigenvectors							
	Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7
<b>m3000</b>	0.399608	-.253973	0.214598	0.474764	-.590741	0.023273	0.393500
<b>marathon</b>	0.531023	-.506889	-.566770	-.365473	0.043953	0.039572	-.052738



Huong Tran - Assignment 3

**Exercise 1**  
**Women record (SPEED)**  
**PCA using Covariance matrix**  
**'The best 3 countries'**

country	Prin1
<b>USA</b>	1.49962
<b>USSR</b>	1.45505
<b>GDR</b>	1.35933

Huong Tran - Assignment 3

**Exercise 1**  
**Women record (SPEED)**  
**PCA using Correlation matrix**

The PRINCOMP Procedure

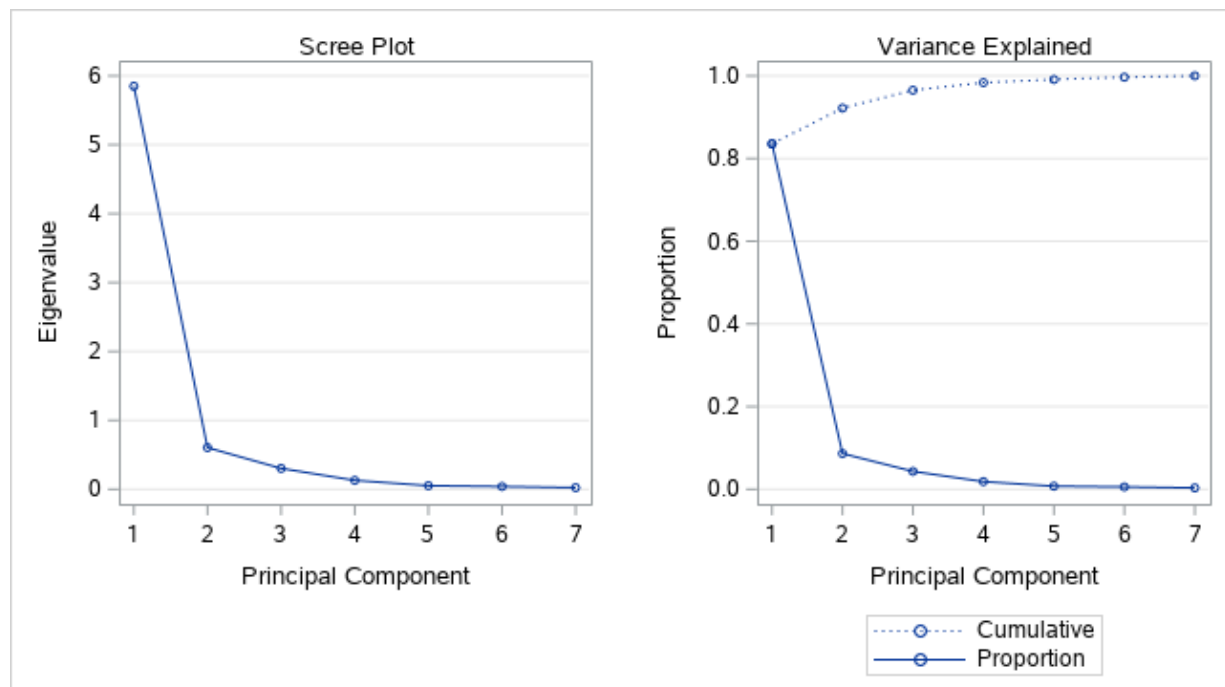
<b>Observations</b>	55
<b>Variables</b>	7

Simple Statistics							
	m100	m200	m400	m800	m1500	m3000	marathon
<b>Mean</b>	8.619563220	8.477681719	7.508259742	6.438315277	5.809893843	5.327650818	4.154344492
<b>Std</b>	0.331058113	0.391558869	0.375282162	0.329456996	0.403074818	0.416455353	0.567096453

Correlation Matrix							
	m100	m200	m400	m800	m1500	m3000	marathon
m100	1.0000	0.9549	0.8362	0.7293	0.7430	0.7485	0.7182
m200	0.9549	1.0000	0.8608	0.7286	0.7203	0.7200	0.7127
m400	0.8362	0.8608	1.0000	0.8990	0.8046	0.7820	0.7133
m800	0.7293	0.7286	0.8990	1.0000	0.9191	0.8740	0.7859
m1500	0.7430	0.7203	0.8046	0.9191	1.0000	0.9637	0.8591
m3000	0.7485	0.7200	0.7820	0.8740	0.9637	1.0000	0.8880
marathon	0.7182	0.7127	0.7133	0.7859	0.8591	0.8880	1.0000

Eigenvalues of the Correlation Matrix				
	Eigenvalue	Difference	Proportion	Cumulative
1	5.84910699	5.24441162	0.8356	0.8356
2	0.60469537	0.30211996	0.0864	0.9220
3	0.30257542	0.17343132	0.0432	0.9652
4	0.12914410	0.07731256	0.0184	0.9836
5	0.05183153	0.01232201	0.0074	0.9911
6	0.03950952	0.01637246	0.0056	0.9967
7	0.02313706		0.0033	1.0000

Eigenvectors							
	Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7
m100	0.369511	0.500733	0.250916	-0.301503	0.275564	0.613971	-0.077263
m200	0.367313	0.546731	0.178001	0.010791	-0.012271	-0.704313	0.195341
m400	0.381088	0.229391	-0.525643	0.368233	-0.529424	0.181939	-0.277244
m800	0.384336	-0.211661	-0.531472	0.120641	0.583117	0.007987	0.412771
m1500	0.389116	-0.353964	-0.050567	-0.427755	0.132373	-0.278060	-0.665527
m3000	0.386813	-0.357782	0.150441	-0.391221	-0.527335	0.105919	0.507327
marathon	0.366831	-0.315647	0.566925	0.650232	0.089412	0.072668	-0.091222



Huong Tran - Assignment 3



**Exercise 1**  
**Women record (SPEED)**  
**PCA using Correlation matrix**  
**'The best 3 countries'**

country	Prin1
USSR	3.82147
GDR	3.79442
USA	3.70676

Huong Tran - Assignment 3

**Exercise 1**  
**Women record (TIME taken to run 1 FOOT (s/foot))**  
**PCA using Covariance matrix**

The PRINCOMP Procedure

Observations	55
Variables	7

Simple Statistics							
	m100	m200	m400	m800	m1500	m3000	marathon
Mean	0.0355295120	0.0361480620	0.0408287480	0.0476214000	0.0606947782	0.0662846167	0.0751445492
Std	0.0013828591	0.0016988110	0.0020475884	0.0024821200	0.0046645951	0.0057835399	0.0131981024

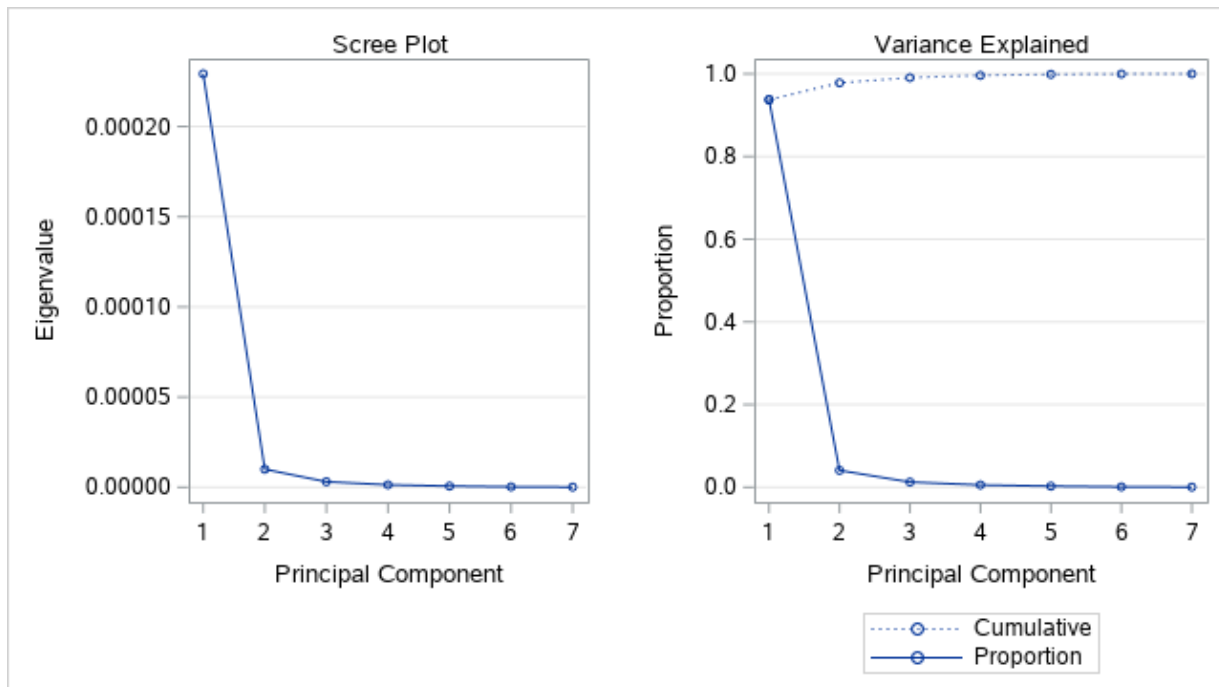
Covariance Matrix							
	m100	m200	m400	m800	m1500	m3000	marathon
m100	0.0000019123	0.0000022383	0.0000023635	0.0000024977	0.0000046983	0.0000059320	0.0000125264
m200	0.0000022383	0.0000028860	0.0000029809	0.0000030531	0.0000055340	0.0000069746	0.0000153713
m400	0.0000023635	0.0000029809	0.0000041926	0.0000045660	0.0000075248	0.0000092090	0.0000190636
m800	0.0000024977	0.0000030531	0.0000045660	0.0000061609	0.0000104390	0.0000123969	0.0000255290
m1500	0.0000046983	0.0000055340	0.0000075248	0.0000104390	0.0000217584	0.0000261461	0.0000540489
m3000	0.0000059320	0.0000069746	0.0000092090	0.0000123969	0.0000261461	0.0000334493	0.0000686862
marathon	0.0000125264	0.0000153713	0.0000190636	0.0000255290	0.0000540489	0.0000686862	0.0001741899

Total Variance	0.0002445495
----------------	--------------

Eigenvalues of the Covariance Matrix				
	Eigenvalue	Difference	Proportion	Cumulative
1	0.00022923	0.00021923	0.9373	0.9373
2	0.00000999	0.00000695	0.0409	0.9782
3	0.00000305	0.00000171	0.0125	0.9907
4	0.00000134	0.00000071	0.0055	0.9961
5	0.00000063	0.00000041	0.0026	0.9987
6	0.00000022	0.00000012	0.0009	0.9996
7	0.00000009		0.0004	1.0000

Eigenvectors							
	Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7
m100	0.066339	0.136004	0.373227	0.370001	0.256456	0.276208	0.747557
m200	0.080686	0.141093	0.539258	0.403980	0.266514	0.156398	-.651225

Eigenvectors							
	Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7
<b>m400</b>	0.102122	0.262679	0.544259	-.157018	-.340019	-.685500	0.119061
<b>m800</b>	0.136346	0.325091	0.259925	-.626661	-.228073	0.601848	-.034979
<b>m1500</b>	0.285080	0.503461	-.250791	-.265503	0.687012	-.244646	-.005567
<b>m3000</b>	0.360056	0.537036	-.374529	0.457580	-.472843	0.085127	-.038385
<b>marathon</b>	0.865540	-.495005	0.054367	-.051676	0.001964	-.004510	0.012675



Huong Tran - Assignment 3

**Exercise 1**  
**Women record (TIME taken to run 1 FOOT (s/foot))**  
**PCA using Covariance matrix**  
**'The best 3 countries'**

country	Prin1
<b>WSamoa</b>	0.066556
<b>Mauritiu</b>	0.039695
<b>Cookis</b>	0.030676

Huong Tran - Assignment 3

**Exercise 1**  
**Women record (TIME taken to run 1 FOOT (s/foot))**  
**PCA using correlation matrix**

The PRINCOMP Procedure

<b>Observations</b>	55
<b>Variables</b>	7

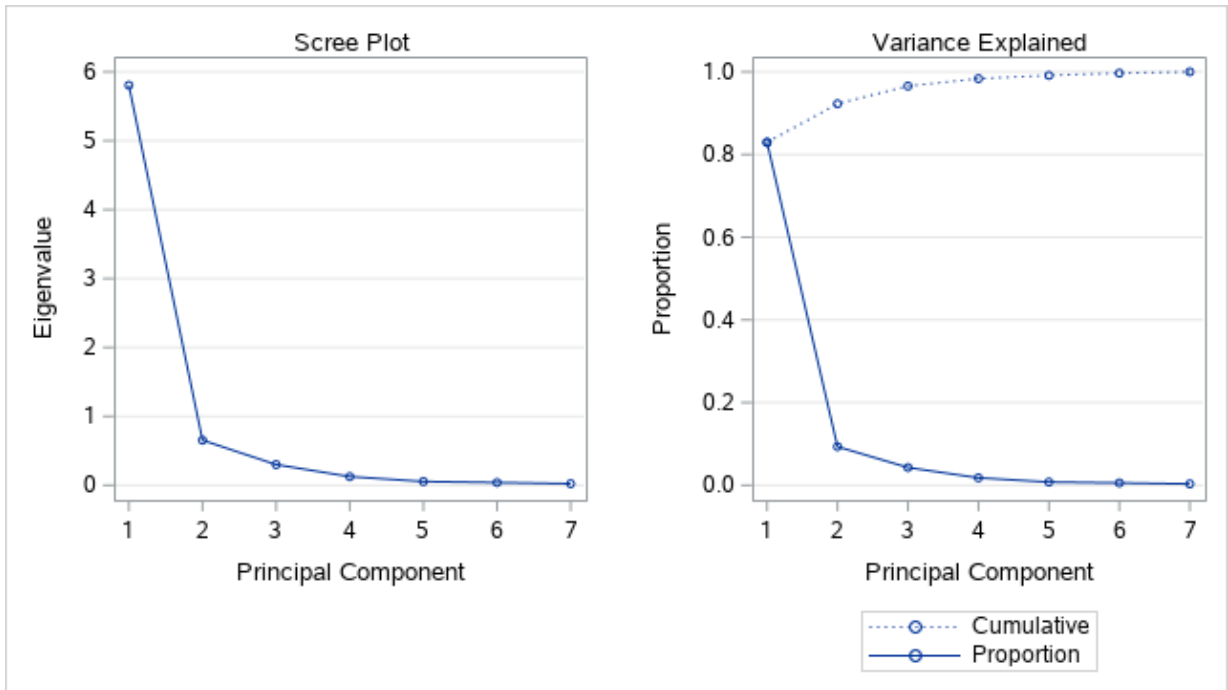
Simple Statistics							
	m100	m200	m400	m800	m1500	m3000	marathon

Simple Statistics							
	m100	m200	m400	m800	m1500	m3000	marathon
Mean	0.0355295120	0.0361480620	0.0408287480	0.0476214000	0.0606947782	0.0662846167	0.0751445492
Std	0.0013828591	0.0016988110	0.0020475884	0.0024821200	0.0046645951	0.0057835399	0.0131981024

Correlation Matrix							
	m100	m200	m400	m800	m1500	m3000	marathon
m100	1.0000	0.9528	0.8347	0.7277	0.7284	0.7417	0.6863
m200	0.9528	1.0000	0.8570	0.7241	0.6984	0.7099	0.6856
m400	0.8347	0.8570	1.0000	0.8984	0.7878	0.7776	0.7054
m800	0.7277	0.7241	0.8984	1.0000	0.9016	0.8636	0.7793
m1500	0.7284	0.6984	0.7878	0.9016	1.0000	0.9692	0.8779
m3000	0.7417	0.7099	0.7776	0.8636	0.9692	1.0000	0.8998
marathon	0.6863	0.6856	0.7054	0.7793	0.8779	0.8998	1.0000

Eigenvalues of the Correlation Matrix				
	Eigenvalue	Difference	Proportion	Cumulative
1	5.80568576	5.15204024	0.8294	0.8294
2	0.65364552	0.35376309	0.0934	0.9228
3	0.29988243	0.17440494	0.0428	0.9656
4	0.12547749	0.07166058	0.0179	0.9835
5	0.05381692	0.01476763	0.0077	0.9912
6	0.03904928	0.01660668	0.0056	0.9968
7	0.02244260		0.0032	1.0000

Eigenvectors							
	Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7
m100	0.368356	0.490060	0.286012	-.319386	0.231169	0.619825	0.052177
m200	0.365364	0.536580	0.229819	0.083302	0.041455	-.710765	-.109225
m400	0.381610	0.246538	-.515367	0.347377	-.572178	0.190946	0.208497
m800	0.384559	-.155402	-.584526	0.042076	0.620324	-.019089	-.315210
m1500	0.389104	-.360409	-.012912	-.429539	0.030261	-.231248	0.692562
m3000	0.388866	-.347539	0.152728	-.363120	-.463355	0.009277	-.598359
marathon	0.367004	-.369208	0.484370	0.672497	0.130536	0.142281	0.069598



Huong Tran - Assignment 3

**Exercise 1**  
**Women record (TIME taken to run 1 FOOT (s/foot))**  
**PCA using correlation matrix**  
**'The best 3 countries'**

country	Prin1
WSamoa	8.33288
Cookis	6.07728
Mauritiu	4.23385

Huong Tran - Assignment 3

```

footnote2 j = r height= 8pt italic "Huong Tran - Assignment 3";
data track;
infile '/home/u59404828/sasuser.v94/STA5221/HW3/women-track.txt' delimiter="," firstobs=2;
input x1 x2 x3 x4 x5 x6 x7 country $;
rename x1 = m100 x2 = m200 x3 = m400 x4 = m800 x5 = m1500 x6 = m3000 x7 = marathon;
run;

proc print data = track;
title1 "Exercise 1";
title2 "Women track record (TIME)";
run;

proc princomp data = track cov out = pc_time_var;
var m100 m200 m400 m800 m1500 m3000 marathon;
title3 "PCA using Covariance matrix";
run;
proc sort data=pc_time_var;
by DESCENDING prin1;
proc print data = pc_time_var (obs = 3);
id country;
var prin1;
title4 'The best 3 countries';
run;

proc princomp data = track out=pc_time_cor;
var m100 m200 m400 m800 m1500 m3000 marathon;
title3 "PCA using correlation matrix";
run;
proc sort data=pc_time_cor;
by DESCENDING prin1;
proc print data = pc_time_cor (obs = 3);
id country;
var prin1;
title4 'The best 3 countries';
run;

*new dataset which all variables measured in seconds;
data time_same_unit;
set track;
m800=m800*60;
m1500=m1500*60;
m3000=m3000*60;
marathon=marathon*60;
x1=100/m100;
x2=200/m200;
x3=400/m400;
x4=800/m800;
x5=1500/m1500;
x6=3000/m3000;
x7=42195/marathon;
run;

* Dataset with measured in SPEED;
data speed;
set time_same_unit;
keep country x1-x7;
rename x1 = m100 x2 = m200 x3 = m400 x4 = m800 x5 = m1500 x6 = m3000 x7 = marathon;
proc print data=speed;
format m100 m200 m400 m800 m1500 m3000 marathon 6.2;
title2 "Women record (SPEED)";
run;

```

```
proc princomp data=speed cov out=pc_speed_var;
var m100 m200 m400 m800 m1500 m3000 marathon;
title3 "PCA using Covariance matrix";
run;

proc sort data=pc_speed_var;
by DESCENDING prin1;
proc print data = pc_speed_var (obs = 3);
id country;
var prin1;
title4 'The best 3 countries';
run;

proc princomp data = speed out=pc_speed_cor;
var m100 m200 m400 m800 m1500 m3000 marathon;
title3 "PCA using Correlation matrix";
run;

proc sort data=pc_speed_cor;
by DESCENDING prin1;
proc print data = pc_speed_cor (obs = 3);
id country;
var prin1;
title4 'The best 3 countries';
run;

* Dataset with measured by TIME/FOOT (s/ft);
data track_foot;
set time_same_unit;
keep country m100 m200 m400 m800 m1500 m3000 marathon;
m100 = m100 * 0.3058/100;
m200 = m200 * 0.3058/200;
m400 = m400 * 0.3058/400;
m800 = m800 * 0.3058/800;
m1500 = m1500 * 0.3508/1500;
m3000 = m3000 * 0.3508/3000;
marathon = marathon/138336;
proc print data = track_foot;
format m100 m200 m400 m800 m1500 m3000 marathon 6.3;
title2 "Women record (TIME taken to run 1 FOOT (s/foot))";
run;

proc princomp data = track_foot cov out=pc_time_foot_cov;
var m100 m200 m400 m800 m1500 m3000 marathon;
title3 "PCA using Covariance matrix";
run;

proc sort data=pc_time_foot_cov;
by DESCENDING prin1;
proc print data = pc_time_foot_cov (obs = 3);
id country;
var prin1;
title4 'The best 3 countries';
run;

proc princomp data = track_foot out=pc_time_foot_cor;
var m100 m200 m400 m800 m1500 m3000 marathon;
title3 "PCA using correlation matrix";
run;

proc sort data=pc_time_foot_cor;
by DESCENDING prin1;
proc print data = pc_time_foot_cor (obs = 3);
id country;
var prin1;
title4 'The best 3 countries';
```

**run;**