410-57, Assignment #7
Anamitra Bhattacharyya

INTRODUCTION

This assignment specifically deals with factor analysis to identify and segment sectors in the stock market data. The raw data set being used in this assignment comprises the daily returns from 20 individual stocks from a variety of market sectors. The overall goal of the assignment is to use statistical methods in SAS such as principal factor analysis, maximum likelihood expectation, with and without rotation to perform a segmentation analysis to derive the number of common factors to build a valid and appropriate model.

RESULTS

1. Data Prep

In this step of we are keeping 4 sectors and these specific stocks associated with them and dropping others:

a) Banking *e.g.* BAC, JPM, WFC
b) Oil Field Services *e.g.* BHI, HAL, SLB
c) Oil Refining *e.g.* CVX, HES, XOM
d) Industrial *e.g.* DD, DOW, HUN

The data is sorted in ascending order for the 12 stocks in the 4 sectors; the returns of the 12 stocks that are kept is shown in the output table below:

Obs	return_BAC	return_BHI	return_CVX	return_DD	return_DOW	return_HAL	return_HES	return_HUN	return_JPM	return_SLB	return_WFC	return_XOM
1	.	.	.	.	.	.	.	.	.	.	.	.
2	0.001723	0.009946	-0.001723	0.010906	0.005357	0.028008	0.010222	-0.008073	-0.000858	-0.007590	0.004562	0.000232531
3	0.082555	-0.013874	-0.009850	-0.006829	0.006324	-0.016074	-0.024015	-0.005079	0.020672	-0.021653	0.015978	-.003027130
4	-0.020817	0.008621	-0.007267	-0.014234	0.005954	0.012080	-0.020699	0.008114	-0.009009	-0.004269	-0.002761	-.007490672
5	0.014458	0.006223	0.010836	0.008435	-0.000330	0.011370	0.008472	-0.006079	-0.001698	0.015227	0.012363	0.004454350
6	0.055828	0.007148	-0.003935	0.015176	0.021864	0.026497	0.028757	0.035932	0.021024	0.027658	0.003747	0.002569795
7	0.035559	-0.034068	-0.011899	0.003388	0.014421	-0.026497	-0.010644	0.037522	0.016779	-0.008374	0.007115	-.007494180
8	-0.011713	-0.038990	-0.026325	0.016772	0.035322	-0.018543	-0.016181	0.040709	0.005169	-0.006578	-0.000338	-.004004245
9	-0.026867	-0.005607	0.010613	0.006218	-0.016724	-0.023010	-0.008101	0.016187	-0.025561	-0.024840	0.000000	0.001650749
10	-0.019863	-0.006686	0.005921	0.002888	0.019178	-0.002360	0.014745	-0.032641	-0.028521	-0.005161	0.007067	0.009497638

Conclusion: Created a sorted list of 12 stocks from the 4 sectors noted above (in date ascending order). The daily stock prices and calculated a log of the ratio of today's price against yesterday's price.

2. Principal Factor Analysis without rotation

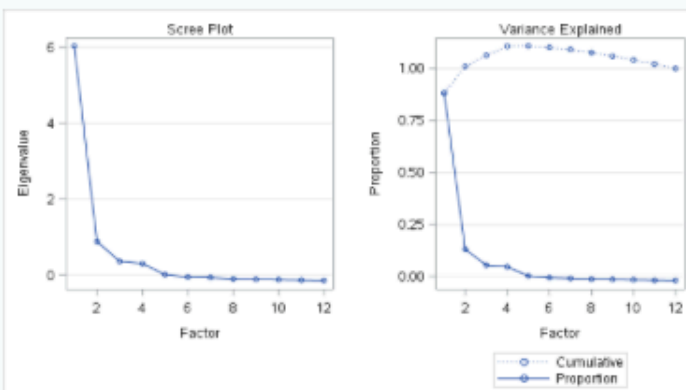
This section of the assignment uses the data set that was created in (1) above; a principal factor analysis (PFA) is performed without any factor rotation. The output is shown below. From the reduced correlation matrix we can see that of the 12 variables chosen there should be 12 factors that fall out from this, but SAS retains 2 primary factors. The criterion that SAS uses to determine the number of factors to retain is the PROPORTION method, such that a sufficient number of factors is retained so that the cumulative proportion of the variance explained is >1. To this end, factor 1 explains >80% of the variance and with the factor 2, this pushes the proportion of variance to over 1.0. Thus, SAS selects factors 1 and 2. The scree plot shows graphically that there is an elbow at factor=2 and the cumulative variance reveals that the first two factors push the variance >1 satisfying the default criteria for SAS for factor selection. Interestingly, We do not necessarily get the 4 factors selected coincident with the 4 original sectors selected.

The FACTOR Procedure
Initial Factor Method: Principal Factors

Prior Communality Estimates: SMC											
return_BAC	return_BHI	return_CVX	return_DD	return_DOW	return_HAL	return_HES	return_HUN	return_JPM	return_SLB	return_WFC	return_XOM
0.58577906	0.61046627	0.64179539	0.54681402	0.47197670	0.64986770	0.49976057	0.39225125	0.58034671	0.68269067	0.57372531	0.62696933

Eigenvalues of the Reduced Correlation Matrix: Total = 6.86244298 Average = 0.57187025				
	Eigenvalue	Difference	Proportion	Cumulative
1	6.04732583	5.16261770	0.8812	0.8812
2	0.88470813	0.52262870	0.1289	1.0101
3	0.36207942	0.05735386	0.0528	1.0629
4	0.30472556	0.29429115	0.0444	1.1073
5	0.01043441	0.06365245	0.0015	1.1088
6	-.05321803	0.01517115	-0.0078	1.1011
7	-.06838918	0.03291807	-0.0100	1.0911
8	-.10130725	0.01600696	-0.0148	1.0763
9	-.11731422	0.00866270	-0.0171	1.0593
10	-.12597692	0.01040221	-0.0184	1.0409
11	-.13637913	0.00786652	-0.0199	1.0210
12	-.14424565		-0.0210	1.0000

2 factors will be retained by the PROPORTION criterion.



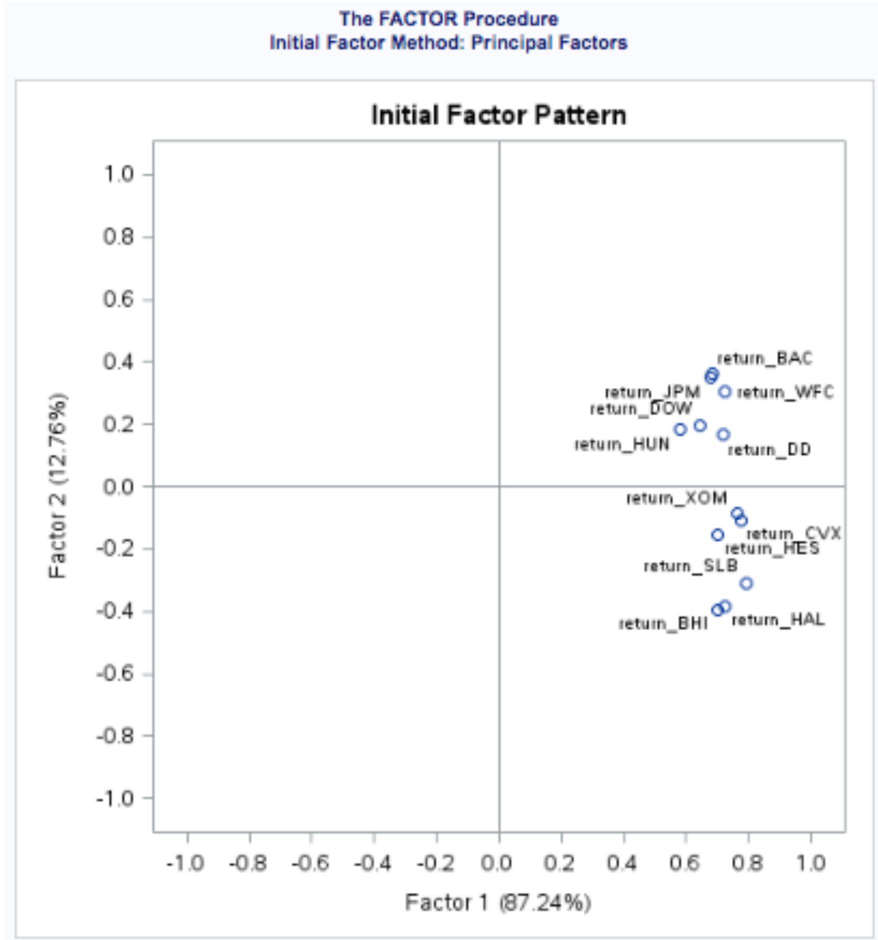
Factor Pattern		
	Factor1	Factor2
return_BAC	0.68475	0.36021
return_BHI	0.69984	-0.39498
return_CVX	0.77402	-0.10833
return_DD	0.71605	0.16703
return_DOW	0.64548	0.19801
return_HAL	0.72630	-0.38221
return_HES	0.70361	-0.15709
return_HUN	0.58030	0.18186
return_JPM	0.67874	0.34813
return_SLB	0.79382	-0.30815
return_WFC	0.72445	0.30517
return_XOM	0.76500	-0.08361

Variance Explained by Each Factor	
Factor1	Factor2
6.0473258	0.8847081

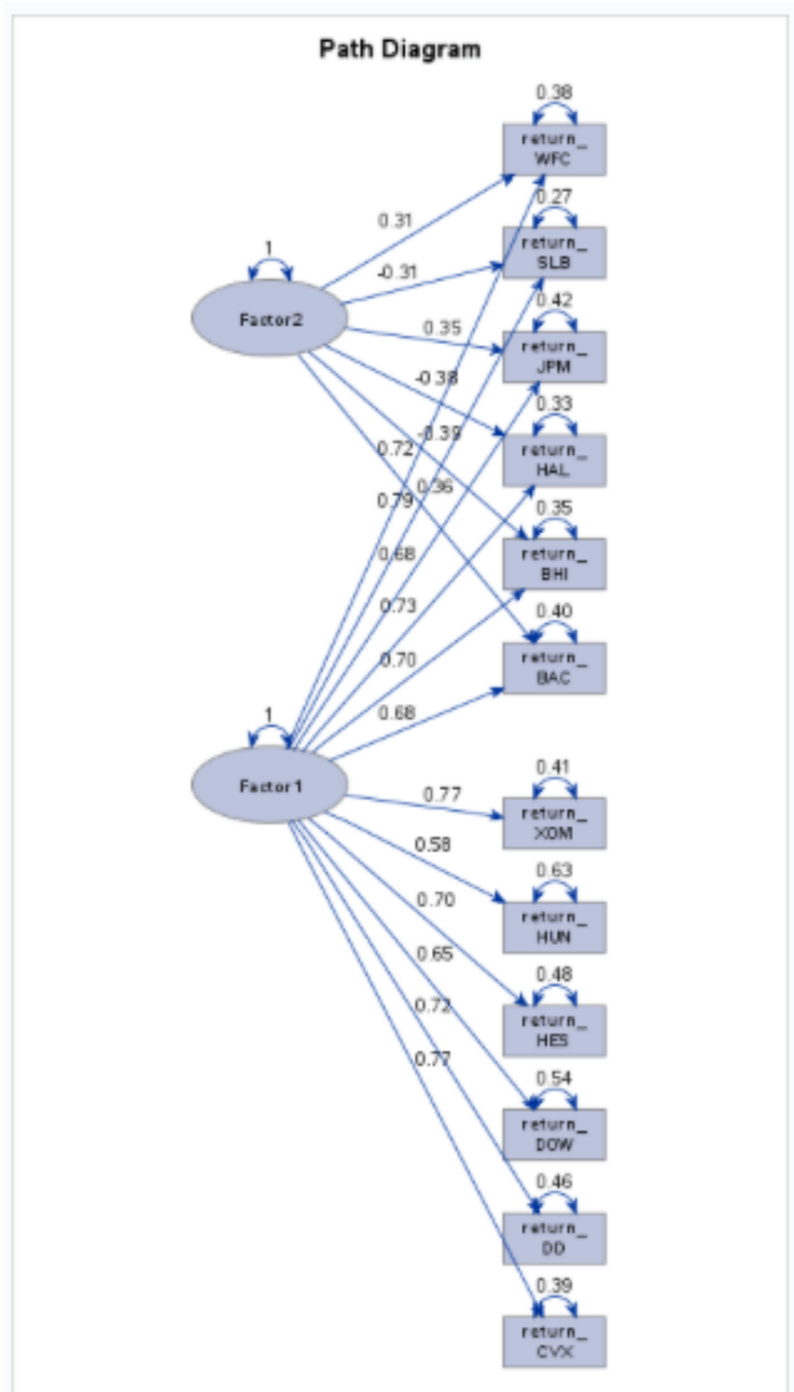
Final Communality Estimates: Total = 6.932034											
return_BAC	return_BHI	return_CVX	return_DD	return_DOW	return_HAL	return_HES	return_HUN	return_JPM	return_SLB	return_WFC	return_XOM
0.59863104	0.64577915	0.61083713	0.54062043	0.45584934	0.67359085	0.51974549	0.36982204	0.58188382	0.72509913	0.61795857	0.59221697

The factor pattern table shows that for the factor loading values for factor 1 are all large and significant compared to factor 2, where only some are positive. Thus, factor 1 has the most significant contribution to the variance *versus* factor 2.

To understand the factor loadings further the loadings from factor 1 are plotted against factor 2, the results from this output is shown below. Looking at the output plot, the stocks are approximately in a line with the banking and industrial stocks in the positive half of the plot and oil refining and oil services in the negative half of the plot. This is consistent with the factor 1 have high loadings and factor 2 having a combination of positive and negative values, see in the loading tables above.



The segregation of the sectors and their respective stocks for factors 1 and 2 is further illustrated in the path diagram shown below.



Conclusion: In this step SAS retained 2 out of 12 factors on the basis of proportion of variance explained by each being cumulatively >1. Plotting factors 1 versus 2 segregated the stock sectors such that banking and industrials were in the positive half of the plot, while oil related stocks were in the negative half of the plot.

3. Application of Varimax rotation to PFA

A comparison of the results with and without Varimax rotation is shown below.

Rotation of factors (Varimax)

The FACTOR Procedure
Rotation Method: Varimax

Orthogonal Transformation Matrix		
	1	2
1	0.70781	0.70640
2	0.70640	-0.70781

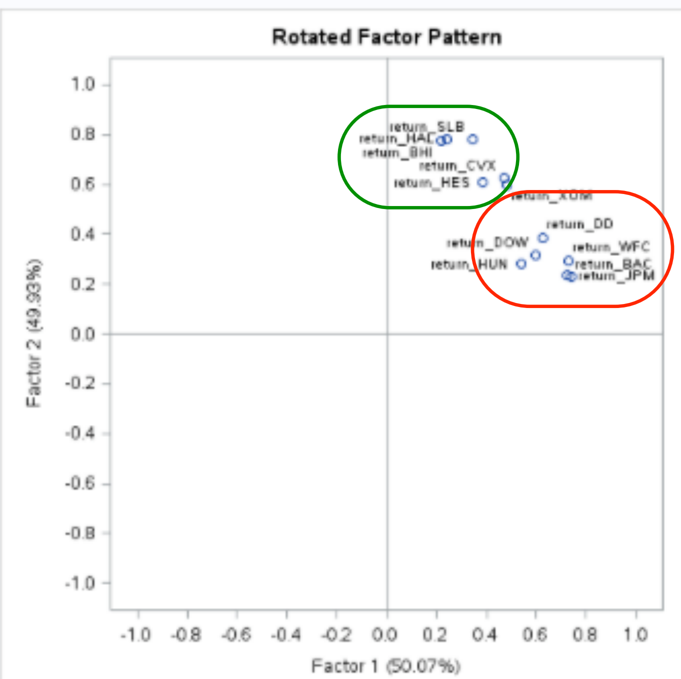
Rotated Factor Pattern		
	Factor1	Factor2
return_BAC	0.73912	0.22875
return_BHI	0.21634	0.77394
return_CVX	0.47133	0.62344
return_DD	0.62482	0.38759
return_DOW	0.59675	0.31582
return_HAL	0.24408	0.78359
return_HES	0.38705	0.60822
return_HUN	0.53921	0.28120
return_JPM	0.72634	0.23305
return_SLB	0.34419	0.77886
return_WFC	0.72835	0.29575
return_XOM	0.48241	0.59958

No rotation of factors (Step 2)

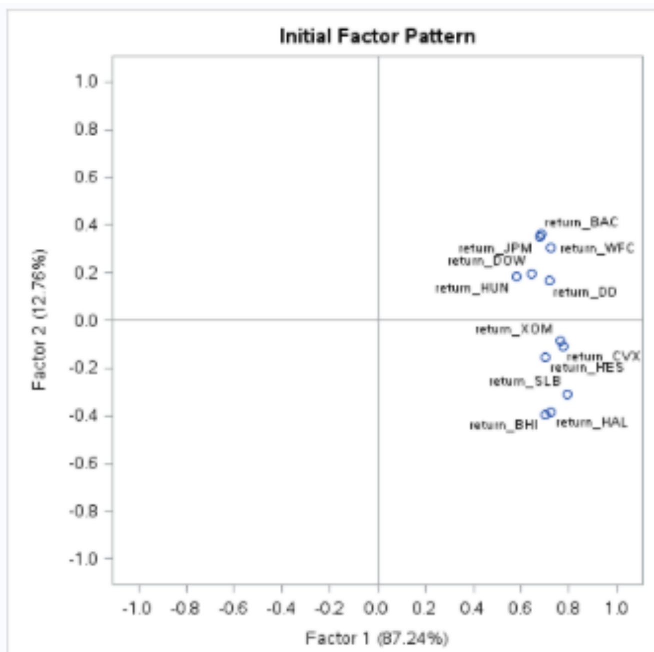
Factor Pattern		
	Factor1	Factor2
return_BAC	0.68475	0.36021
return_BHI	0.69984	-0.39498
return_CVX	0.77402	-0.10833
return_DD	0.71605	0.16703
return_DOW	0.64548	0.19801
return_HAL	0.72630	-0.38221
return_HES	0.70361	-0.15709
return_HUN	0.58030	0.18186
return_JPM	0.67874	0.34813
return_SLB	0.79382	-0.30815
return_WFC	0.72445	0.30517
return_XOM	0.76500	-0.08361

Rotation of factors (Varimax)

The FACTOR Procedure
Rotation Method: Varimax



No rotation of factors (Step 2)



Did SAS retain the same number of factors?
Although SAS did not change the number of factors retained, which remains the same at 2, the loading values did change with the Varimax rotation, see results table above.

What components of the PROC FACTOR output did the rotation change? Did we obtain a 'simple structure' from our factor rotation? Did we increase the interpretability using the factor rotation?
With the absence of rotation the values for factor 1 were much higher than factor 2, and the factor 2 loadings were sometimes negative. In the presence of rotation, see table above left, the loadings were all positive for both factors. In addition, as we analyze the stocks between factors 1 and 2, sometimes the loadings for some stocks are higher for factor 1 while other stocks are higher with factor 2. For example, for BAC the loading value for factor 1 > factor 2; however, for BHI the loading for factor 2 > factor 1. Either way the loadings per stock are being maximized between the two factors. Plotting and comparing the loadings between factor 1 and 2, see plots above, one can see that with the rotation a simple structure is obtained which increases the interpretability of factors 1 and 2. Specifically, looking at the rotated factor 1 *versus* factor 2 plot (colored rings) the banking and industrial stocks are clearly and discretely segmented (clustered, red ring) from the oil industry sector stocks (green).

Conclusion: The Varimax rotation resulted in a simpler structure being generated, which facilitated a better interpretation of the stock data when comparing the results without factor rotation. With the rotation, discrete clusters were apparent on the plots of factor 1 *versus* factor 2, resulting in a segmentation of stocks falling in to these sectors (compare green and red rings above).

4. Using Maximum Likelihood Estimation to estimate common factors
In this section of the assignment a principal factor analysis is used with squared multiple correlations for the prior communality estimates (*i.e.* PRIORS=SMC), premised on a common factor model.

How many common factors does ML Factor Analysis suggest? How does ML Factor Analysis arrive at this number of factors, and in general how do we interpret the output from a ML Factor Analysis?

Initially SAS uses and presents the results from the PROC PROPORTIONS method, producing the following correlation matrix (see below).

The FACTOR Procedure
Initial Factor Method: Maximum Likelihood

Prior Communality Estimates: SMC											
return_BAC	return_BHI	return_CVX	return_DD	return_DOW	return_HAL	return_HES	return_HUN	return_JPM	return_SLB	return_WFC	return_XO
0.58577906	0.61046627	0.64179539	0.54681402	0.47197670	0.64986770	0.49976057	0.39225125	0.58034671	0.68269067	0.57372531	0.6269693

Preliminary Eigenvalues: Total = 16.9350893 Average = 1.41125745				
	Eigenvalue	Difference	Proportion	Cumulative
1	14.9446192	12.7338755	0.8825	0.8825
2	2.2107436	1.3691513	0.1305	1.0130
3	0.8415924	0.1303533	0.0497	1.0627
4	0.7112391	0.6898126	0.0420	1.1047
5	0.0214265	0.1469738	0.0013	1.1060
6	-0.1255473	0.0709849	-0.0074	1.0986
7	-0.1965322	0.0239638	-0.0116	1.0869
8	-0.2204960	0.0345938	-0.0130	1.0739
9	-0.2550898	0.0286557	-0.0151	1.0589
10	-0.2837455	0.0587253	-0.0168	1.0421
11	-0.3424708	0.0281790	-0.0202	1.0219
12	-0.3706498		-0.0219	1.0000

2 factors will be retained by the PROPORTION criterion.

Iteration	Criterion	Ridge	Change	Communalities											
1	0.6510876	0.0000	0.0766	0.61329	0.67722	0.59574	0.52016	0.43403	0.68819	0.52149	0.34962	0.59394	0.74783	0.65036	0.58326
2	0.6468795	0.0000	0.0233	0.62141	0.68257	0.58380	0.50908	0.42409	0.71149	0.51403	0.34494	0.60820	0.75240	0.65204	0.57178
3	0.6466363	0.0000	0.0044	0.62579	0.68570	0.57998	0.50612	0.42109	0.71292	0.51264	0.34336	0.61064	0.75401	0.65312	0.56870
4	0.6466159	0.0000	0.0017	0.62697	0.68582	0.57927	0.50513	0.42017	0.71351	0.51229	0.34293	0.61230	0.75408	0.65322	0.56796
5	0.6466133	0.0000	0.0006	0.62759	0.68586	0.57907	0.50478	0.41980	0.71345	0.51224	0.34275	0.61278	0.75412	0.65329	0.56779

Since a different method is being used (MLE) the eigenvalues are different compared to previously. Again using this initial output the first 2 factors explain the largest proportion of the variance, which leads to the initial recommendation from PROPORTION that 2 factors will be retained.

From a modeling perspective what does ML Factor Analysis provide that Principal Factor Analysis does not?

The ML factor analysis provides hypothesis testing (*e.g.* Chi-squared) that was not available in PFA, and one also makes assumptions about the normality of the distribution. Specifically, the hypothesis testing results shown below indicate that the p-value for null hypothesis (no common factors) is low so we reject it, suggesting common factors are required. In addition, the p-value for another hypothesis test – null hypothesis states 2 factors are required – is also low, implying > 2 factors are required.

Significance Tests Based on 501 Observations			
Test	DF	Chi-Square	Pr > ChiSq
H0: No common factors	66	3656.2617	<.0001
HA: At least one common factor			
H0: 2 Factors are sufficient	43	319.3192	<.0001
HA: More factors are needed			

Chi-Square without Bartlett's Correction	323.30664
Akaike's Information Criterion	237.30664
Schwarz's Bayesian Criterion	55.99257
Tucker and Lewis's Reliability Coefficient	0.88187

Squared Canonical Correlations	
Factor1	Factor2
0.94176593	0.73146692

The MLE method does perform some additional weighting on the variances for factors 1 and 2, the loadings for factors 1 and 2 are shown below (as well as the weighted/unweighted variances). Before performing any rotations, factor 1 appears to possess the more significant values for its composite loadings, whereas the factor 2 loading fluctuate from positive to negative values.

MLE without Varimax rotation

Factor Pattern		
	Factor1	Factor2
return_BAC	0.67169	0.42017
return_BHI	0.73979	-0.37223
return_CVX	0.76083	-0.01315
return_DD	0.68596	0.18472
return_DOW	0.61384	0.20710
return_HAL	0.76435	-0.35943
return_HES	0.70939	-0.09484
return_HUN	0.55701	0.18007
return_JPM	0.66686	0.41017
return_SLB	0.82615	-0.26755
return_WFC	0.71228	0.38208
return_XOM	0.75342	0.01106

Variance Explained by Each Factor		
Factor	Weighted	Unweighted
Factor1	16.1720778	6.02460756
Factor2	2.7239360	0.94882513

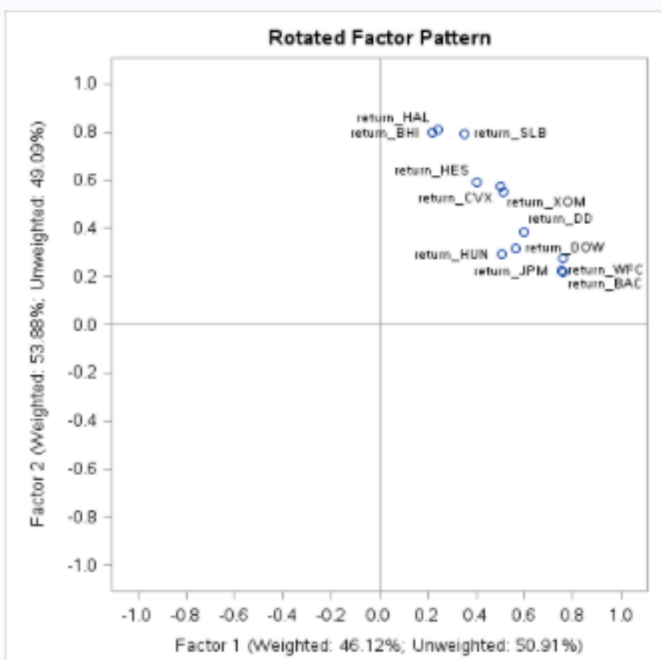
MLE with rotation

Rotated Factor Pattern		
	Factor1	Factor2
return_BAC	0.76122	0.21969
return_BHI	0.21664	0.79932
return_CVX	0.49806	0.57530
return_DD	0.59542	0.38748
return_DOW	0.56395	0.31884
return_HAL	0.24256	0.80907
return_HES	0.40289	0.59153
return_HUN	0.50588	0.29457
return_JPM	0.75054	0.22277
return_SLB	0.35223	0.79376
return_WFC	0.75994	0.27534
return_XOM	0.51113	0.55362

Variance Explained by Each Factor		
Factor	Weighted	Unweighted
Factor1	8.7156851	3.55022275
Factor2	10.1803287	3.42320994

The rotated factor pattern with MLE and Varimax is shown below:

The FACTOR Procedure
Rotation Method: Varimax



Interestingly, the factor pattern + rotation has segmented and extracted the stocks into the financial, industrial, and 2 oil industry sectors.

Conclusion: The MLE methodology has added a more nuanced view of the stock segmentation. Although suggesting a 2-factor analysis initially, the hypothesis testing indicates that > 2 factors are required, an aspect of the analysis not presented in PFA. Together with the rotation, the MLE is able to elicit separation of the 4 sectors in the starting stock portfolio.

5. Using ML factor analysis PRIORS=MAX option

In this section of the assignment as above an MLE is used with a MAX option for the prior communality estimates (*i.e.* PRIORS=MAX). This change in option (from SMC to MAX) sets the prior communality estimate for each variable to its maximum absolute correlation with any other variable. The initial output from this in SAS is shown below, and suggests 5 factors need chosen based on the PROPORTIONS criterion, at which stage the cumulative proportion will be >1.

The FACTOR Procedure
Initial Factor Method: Maximum Likelihood

Prior Communality Estimates: MAX											
return_BAC	return_BHI	return_CVX	return_DD	return_DOW	return_HAL	return_HES	return_HUN	return_JPM	return_SLB	return_WFC	return_XOM
0.71024895	0.72841318	0.74601189	0.62490691	0.62490691	0.74435397	0.60777225	0.53432283	0.71024895	0.74435397	0.65005778	0.74601189

Preliminary Eigenvalues: Total = 27.1688165 Average = 2.26406804				
	Eigenvalue	Difference	Proportion	Cumulative
1	20.5690725	17.1824218	0.7571	0.7571
2	3.3866507	1.7997066	0.1247	0.8817
3	1.5869441	0.2731426	0.0584	0.9401
4	1.3138014	0.9921696	0.0484	0.9885
5	0.3216318	0.1535855	0.0118	1.0003
6	0.1680463	0.1038877	0.0062	1.0065
7	0.0641586	0.0272768	0.0024	1.0089
8	0.0368818	0.0424458	0.0014	1.0102
9	-0.0055640	0.0327906	-0.0002	1.0100
10	-0.0383547	0.0588509	-0.0014	1.0086
11	-0.0972055	0.0400409	-0.0036	1.0051
12	-0.1372464		-0.0051	1.0000

5 factors will be retained by the PROPORTION criterion.

The hypothesis-testing table for the MLE is displayed below and indicates common factors are required (low p-value for NH), and a significant p-value for the null hypothesis test that 5-factors are sufficient. However, this does negate the premise that 4-factors, for example may also be sufficient.

Significance Tests Based on 501 Observations			
Test	DF	Chi-Square	Pr > ChiSq
H0: No common factors	66	3656.2617	<.0001
HA: At least one common factor			
H0: 5 Factors are sufficient	16	10.9169	0.8146
HA: More factors are needed			

Chi-Square without Bartlett's Correction	11.098156
Akaike's Information Criterion	-20.901844
Schwarz's Bayesian Criterion	-88.367542
Tucker and Lewis's Reliability Coefficient	1.005840

The AIC is lower than previous (step 4), which is good. After the Varimax rotation is performed the following loadings are obtained for the MLE with PRIORS=MAX option.

The FACTOR Procedure
Rotation Method: Varimax

Orthogonal Transformation Matrix					
	1	2	3	4	5
1	0.59957	0.48473	0.43530	0.46400	0.02768
2	-0.69192	0.66028	0.28501	-0.06233	-0.01247
3	-0.39835	-0.43867	0.21591	0.77390	-0.05782
4	0.02987	-0.36599	0.82461	-0.42644	-0.05762
5	-0.04671	-0.05183	0.05170	0.00658	0.99620

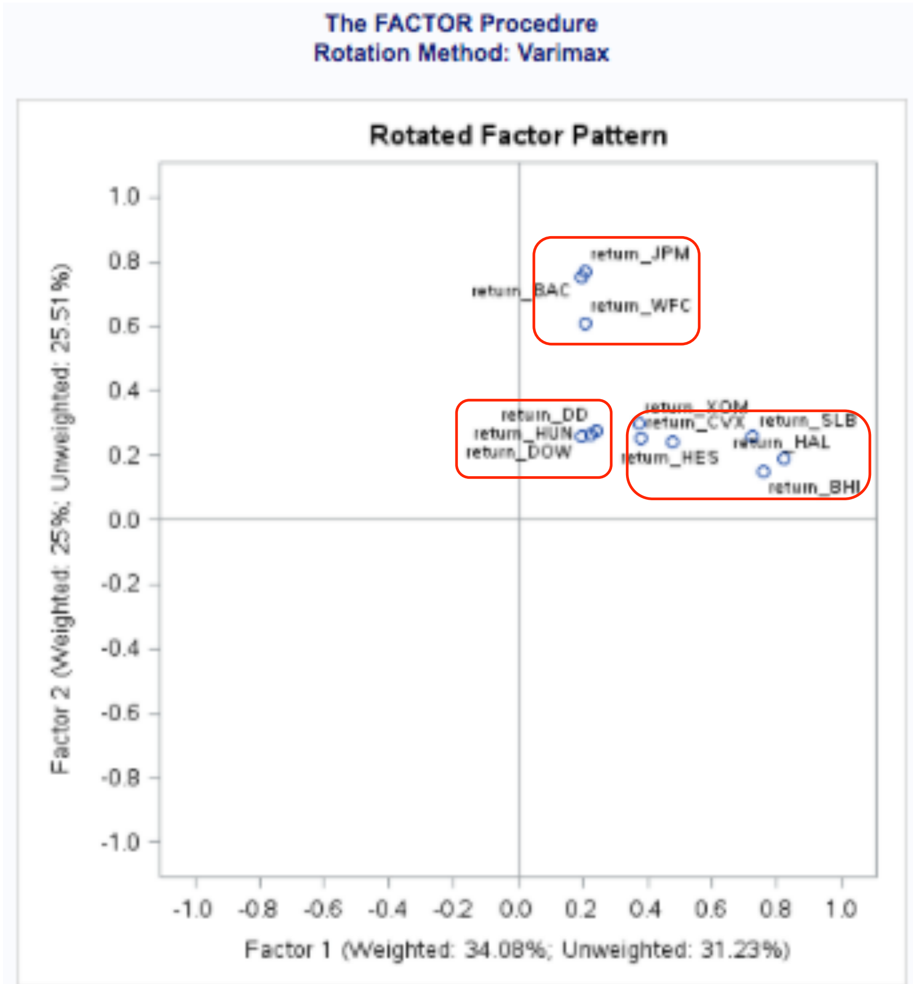
Rotated Factor Pattern					
	Factor1	Factor2	Factor3	Factor4	Factor5
return_BAC	0.19300	0.75425	0.26803	0.17215	0.09285
return_BHI	0.75597	0.14970	0.18684	0.24628	-0.01722
return_CVX	0.37688	0.25354	0.26440	0.70383	0.02658
return_DD	0.24372	0.27524	0.66859	0.31138	-0.13337
return_DOW	0.19396	0.25931	0.64481	0.23505	-0.00701
return_HAL	0.82071	0.18978	0.20801	0.16916	-0.00609
return_HES	0.47834	0.23976	0.25785	0.40900	0.24903
return_HUN	0.22592	0.26677	0.60996	0.06709	0.16770
return_JPM	0.20547	0.77151	0.22874	0.17842	-0.03102
return_SLB	0.72537	0.25575	0.24707	0.30301	0.05701
return_WFC	0.20847	0.61032	0.35934	0.29285	-0.00631
return_XOM	0.37166	0.29603	0.24083	0.66560	-0.02404

Looking at the factor loadings, there are two important points:

1. The other factors 1-4, has a maximum loading for at least one stock. For example,
   a) Factor 1 has a maximum loading for BHI, HAL and SLB
   b) Factor 2 has a maximum loading for BAC, JPM and WFC
   c) Factor 3 has a maximum loading for DOW and HUN
   d) Factor 4 has a maximum loading for CVX and XOM

2. Factor 5 has low values for all its loadings.

Since factor 5 has low loadings for all its stock loadings, it is debatable if this factor is significant to be included in the model. Using a rotation with 5 factors produces 10 rotation pattern diagrams, one is shown for factor 1 *versus* factor 2, using the MLE with PRIORS=MAX option and a Varimax rotation (see below). In this figure the sectors become segmented/clustered to different extents using the different combination of factors.



As noted above, while using this approach it was clear from the loadings from factor 5 that this factor was not required for the model. Another model was created using NFACTOR=4, see below. Note how the loading for the

The FACTOR Procedure
Initial Factor Method: Maximum Likelihood

Prior Communality Estimates: MAX											
return_BAC	return_BHI	return_CVX	return_DD	return_DOW	return_HAL	return_HES	return_HUN	return_JPM	return_SLB	return_WFC	return_XOM
0.71024895	0.72841318	0.74601189	0.62490691	0.62490691	0.74435397	0.60777225	0.53432283	0.71024895	0.74435397	0.65005778	0.74601189

Preliminary Eigenvalues: Total = 27.1688165 Average = 2.26406804				
	Eigenvalue	Difference	Proportion	Cumulative
1	20.5690725	17.1824218	0.7571	0.7571
2	3.3866507	1.7997066	0.1247	0.8817
3	1.5869441	0.2731426	0.0584	0.9401
4	1.3138014	0.9921696	0.0484	0.9885
5	0.3216318	0.1535855	0.0118	1.0003
6	0.1680463	0.1038877	0.0062	1.0065
7	0.0641586	0.0272768	0.0024	1.0089
8	0.0368818	0.0424458	0.0014	1.0102
9	-0.0055640	0.0327906	-0.0002	1.0100
10	-0.0383547	0.0588509	-0.0014	1.0086
11	-0.0972055	0.0400409	-0.0036	1.0051
12	-0.1372464		-0.0051	1.0000

4 factors will be retained by the NFACTOR criterion.

The hypothesis test results (see below) suggest that (a) common factors are required and (b) using a 4-factor model should be significant (null hypothesis p-value= 0.6211). The AIC value has become much smaller (*e.g.* -26 with the 4-factor model *versus* -20 with the 5-factor model).

Significance Tests Based on 501 Observations			
Test	DF	Chi-Square	Pr > ChiSq
H0: No common factors	66	3656.2617	<.0001
HA: At least one common factor			
H0: 4 Factors are sufficient	24	21.2978	0.6211
HA: More factors are needed			

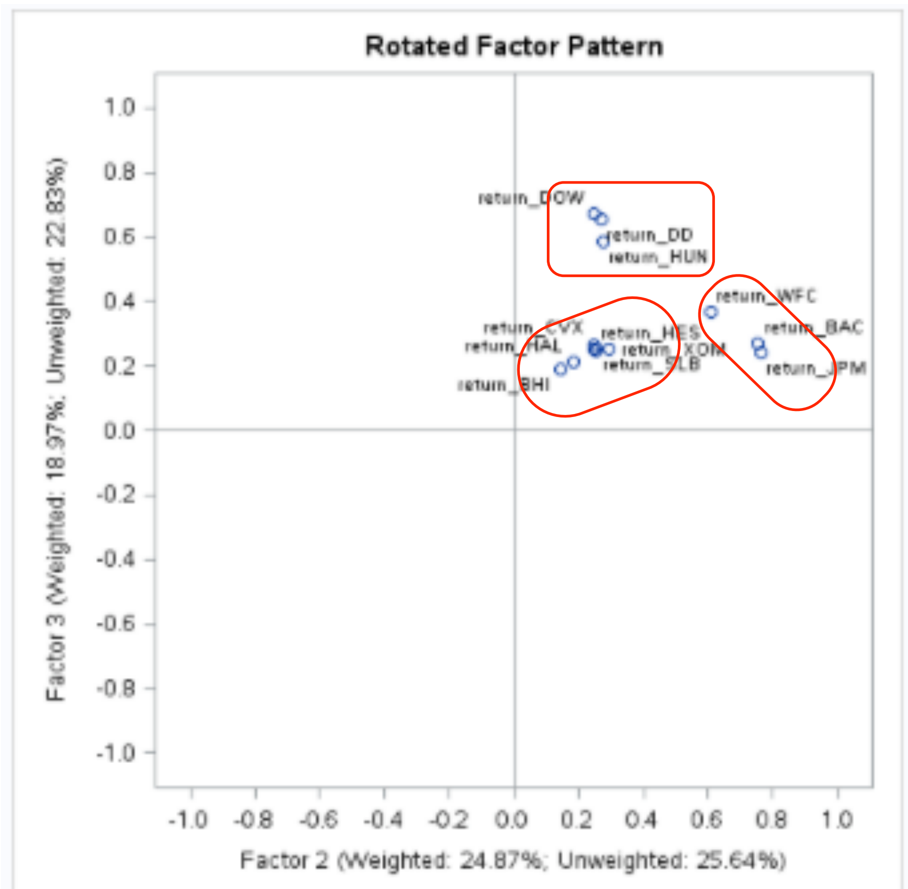
Chi-Square without Bartlett's Correction	21.62209
Akaike's Information Criterion	-26.37791
Schwarz's Bayesian Criterion	-127.57645
Tucker and Lewis's Reliability Coefficient	1.00207

Squared Canonical Correlations			
Factor1	Factor2	Factor3	Factor4
0.95364555	0.77149229	0.62960830	0.55735866

The factor loadings with the Varimax rotation shown in the table below reveals that all the factors have positive values and each factor has at least one stock with a high loading value.

Rotated Factor Pattern				
	Factor1	Factor2	Factor3	Factor4
return_BAC	0.19757	0.75423	0.27205	0.17115
return_BHI	0.75560	0.14546	0.19113	0.24328
return_CVX	0.37940	0.24873	0.26653	0.71355
return_DD	0.24525	0.27169	0.65220	0.30797
return_DOW	0.19069	0.24648	0.67301	0.22569
return_HAL	0.81890	0.18590	0.21348	0.16413
return_HES	0.48670	0.24951	0.25485	0.39568
return_HUN	0.23250	0.27340	0.58577	0.07364
return_JPM	0.20509	0.76170	0.23935	0.17940
return_SLB	0.72981	0.25515	0.24982	0.29668
return_WFC	0.20922	0.60742	0.36668	0.29113
return_XOM	0.37625	0.29148	0.25235	0.65233

Once rotation is added to the model (example below shows factor 2 plotted *versus* factor 3), one can again see the graphical segmentation of the stocks in the market sectors.



Conclusion: The ML Factor Analysis shows a significant and valid model with 4 common factors, the hypothesis-testing component testing aspect of the output shown earlier reveals this. When the PRIORS option was set to MAX, the results indicate that the prior estimates of the communalities are very sensitive to the estimation of the common factors.

CONCLUSION

The use of the maximum likelihood expectation (MLE) model with its various options, including its hypothesis testing component, and significance testing, was able to produce a valid segmentation model of the 4-sector stock portfolio. While the Principal component analysis (PCA) started off with 2 factors, the MLE approach with SAS generated a 4-common factor model that was statistically valid and created through hypothesis testing of the factor number. The use of orthogonal factor rotation (Varimax) allowed for a 'simple model' representation to be produced, so that the rotated factors remain uncorrelated and the communalities are preserved. The final model produced a graphical factor pattern that displayed a teasing apart of the various stocks into their 4 different market sector families.