PCA Recap

PCA represents your data along a new set of variables - the principal components (PCs):

- The PCs are eigenvectors of $\mathbf{X}^T\mathbf{X}$
  - Where X is a centered or standardized data matrix.
- The data projected onto the PCs is uncorrelated
  - This means the scores are orthogonal too!
- The **variance explained** by a PC is given by the **corresponding eigenvalue**
  - Small Eigenvalues ⇒ not much variance (information) along that component.
- Total amount of variance in data is given by sum of eigenvalues.
  - Equivalently by the sum of the variances of your original variables.

Covariance Matrix of PCs

Since the **new data** (after projection onto principal components) is **completely uncorrelated** (orthogonal), the **covariance matrix is diagonal**.

$$cov(PCs) = \mathbf{\Sigma} = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \lambda_p \end{pmatrix}$$

PCA Recap

- PCA is no more than a matrix factorization, which gives you new basis vectors and scores/coordinates for each observation:

$$\mathbf{X}^{\mathrm{T}}=\mathbf{V}\mathbf{S}^{\mathrm{T}}$$

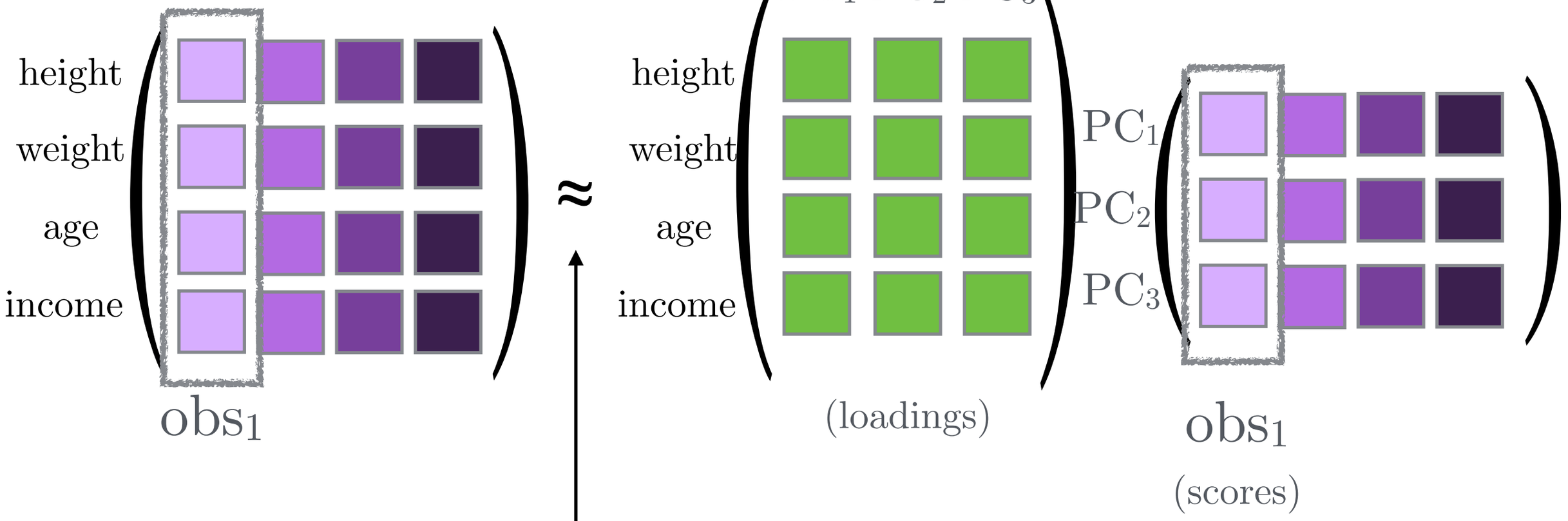
  - $\mathbf{V}$ is the matrix of eigenvectors (principal components)
  - $\mathbf{S}^{\mathrm{T}}$ is the matrix having the coordinates for each observation as columns
  - Covariance PCA: the data matrix $\mathbf{X}$ contains centered data
  - Correlation PCA: the data matrix $\mathbf{X}$ contains standardized data.

- Above equation same as writing $\mathbf{X}=\mathbf{S}\mathbf{V}^{\mathrm{T}}$.

Illustrated on next two slides for **reference**

$$\mathbf{X}^{\mathrm{T}} = \mathbf{V}\mathbf{S}^{\mathrm{T}}$$

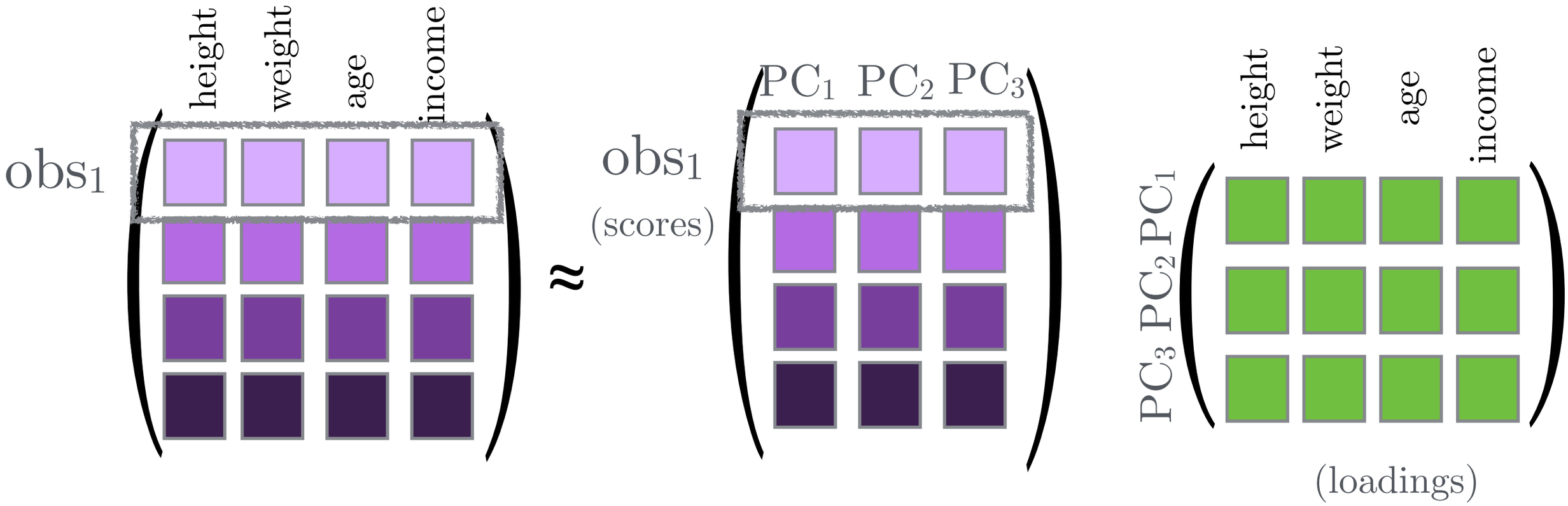
data e-vectors scores



Approximately equal why?

$$\mathbf{X} = \mathbf{S}\mathbf{V}^{\mathrm{T}}$$

data scores e-vectors



Covariance vs. Correlation PCA

- We use correlation PCA in cases where the scales of our data are drastically different.

- We use covariance PCA when our variables are on the same or similar scales

- We can use other forms of standardization (range/minmax standardization) to transform our data prior to covariance PCA

- **THERE IS NO MATHEMATICAL RELATIONSHIP BETWEEN COVARIANCE AND CORRELATION PCA.** Given one set of loadings and scores, I cannot re-create the other set. Each must be computed independently.

Factor Analysis via PCA Rotations

Factor Analysis

What is it?

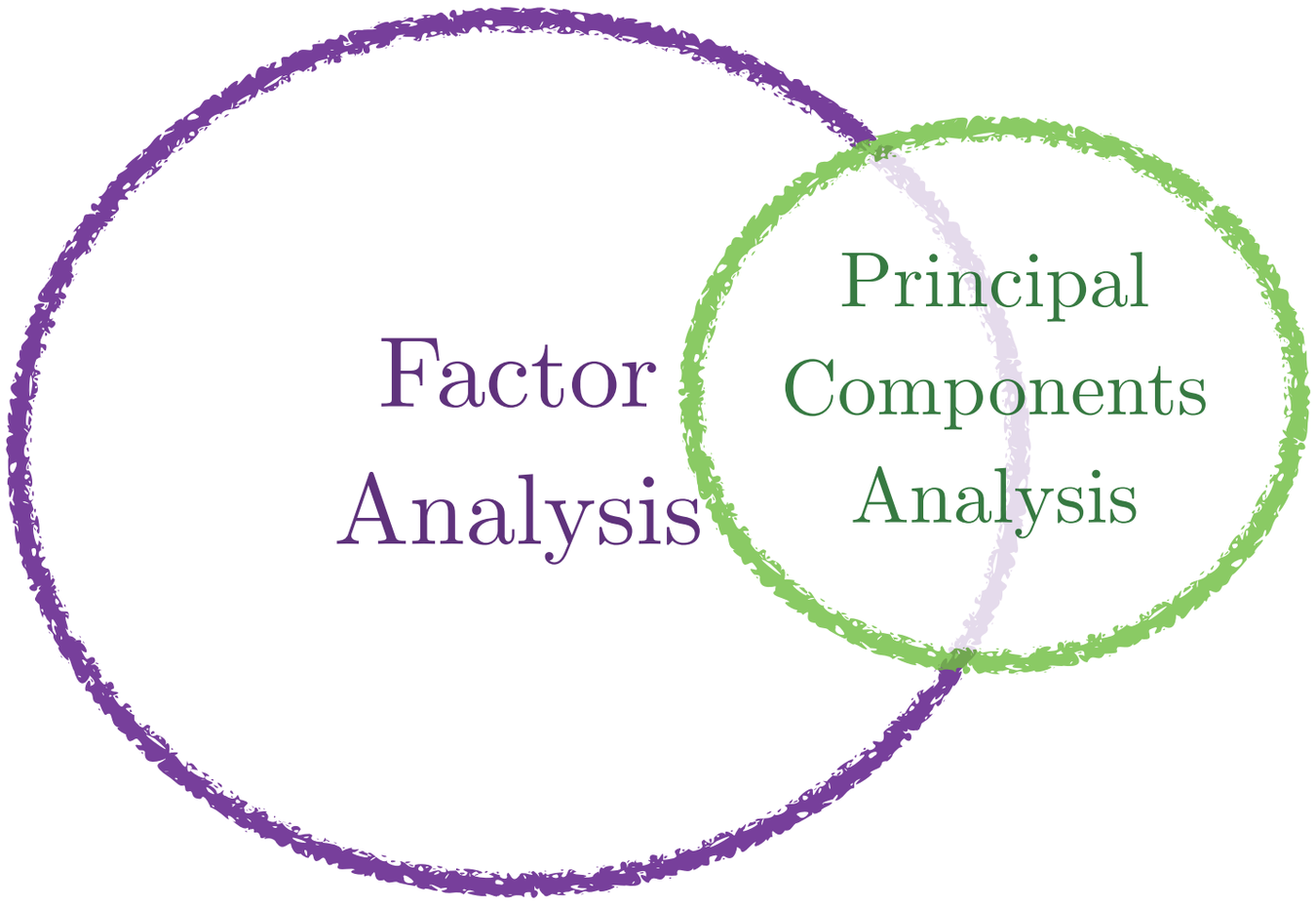
- **Factor Analysis** is the art of extracting *latent (=hidden)* variables from data.

- **Exploratory Factor Analysis** is when we find those factors without any *a priori* hypothesis. We just look for hidden patterns

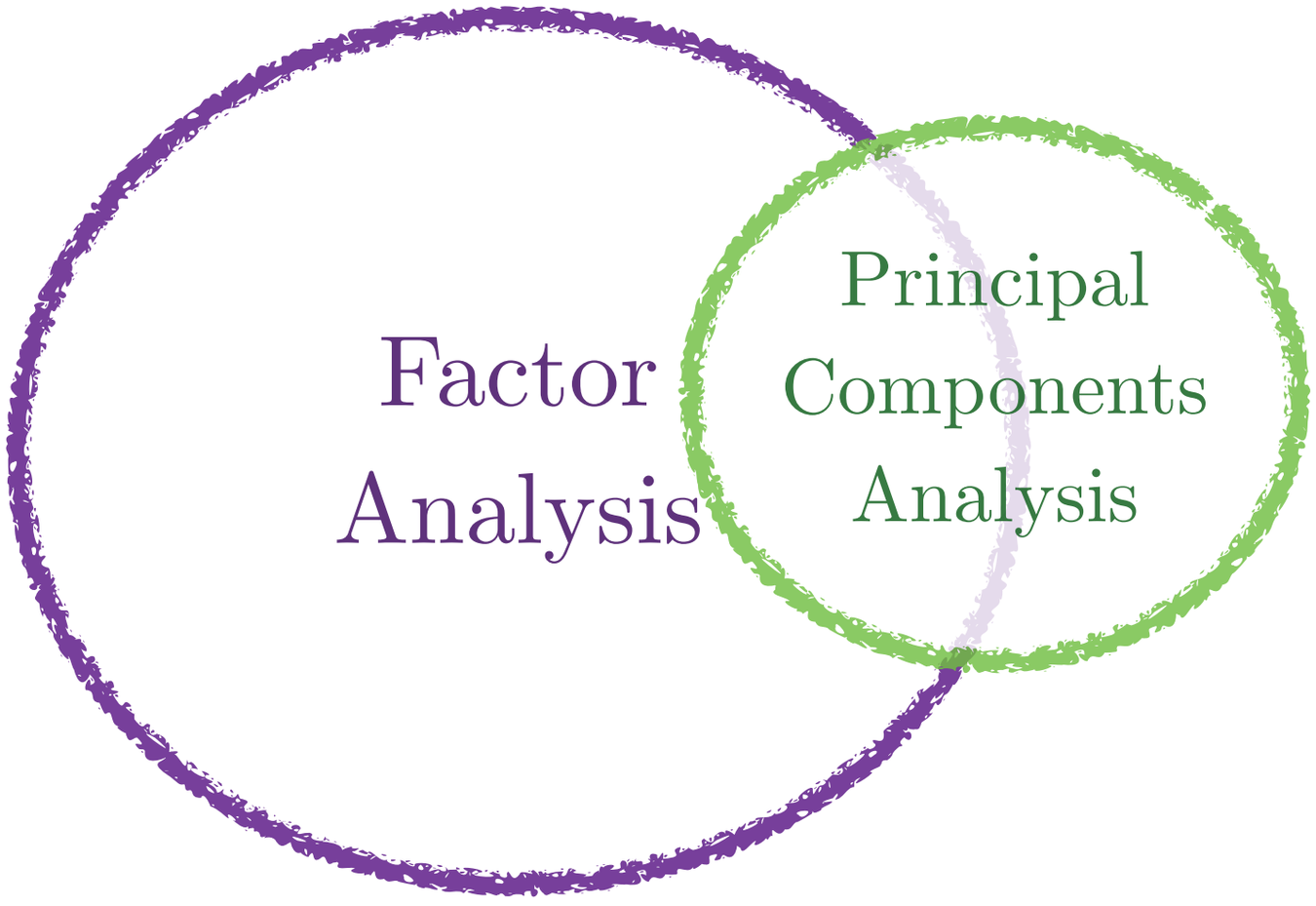
- **Confirmatory Factor Analysis** is integral to survey design, and checks whether your survey measures the things it was designed to measure.

How is it different from PCA?

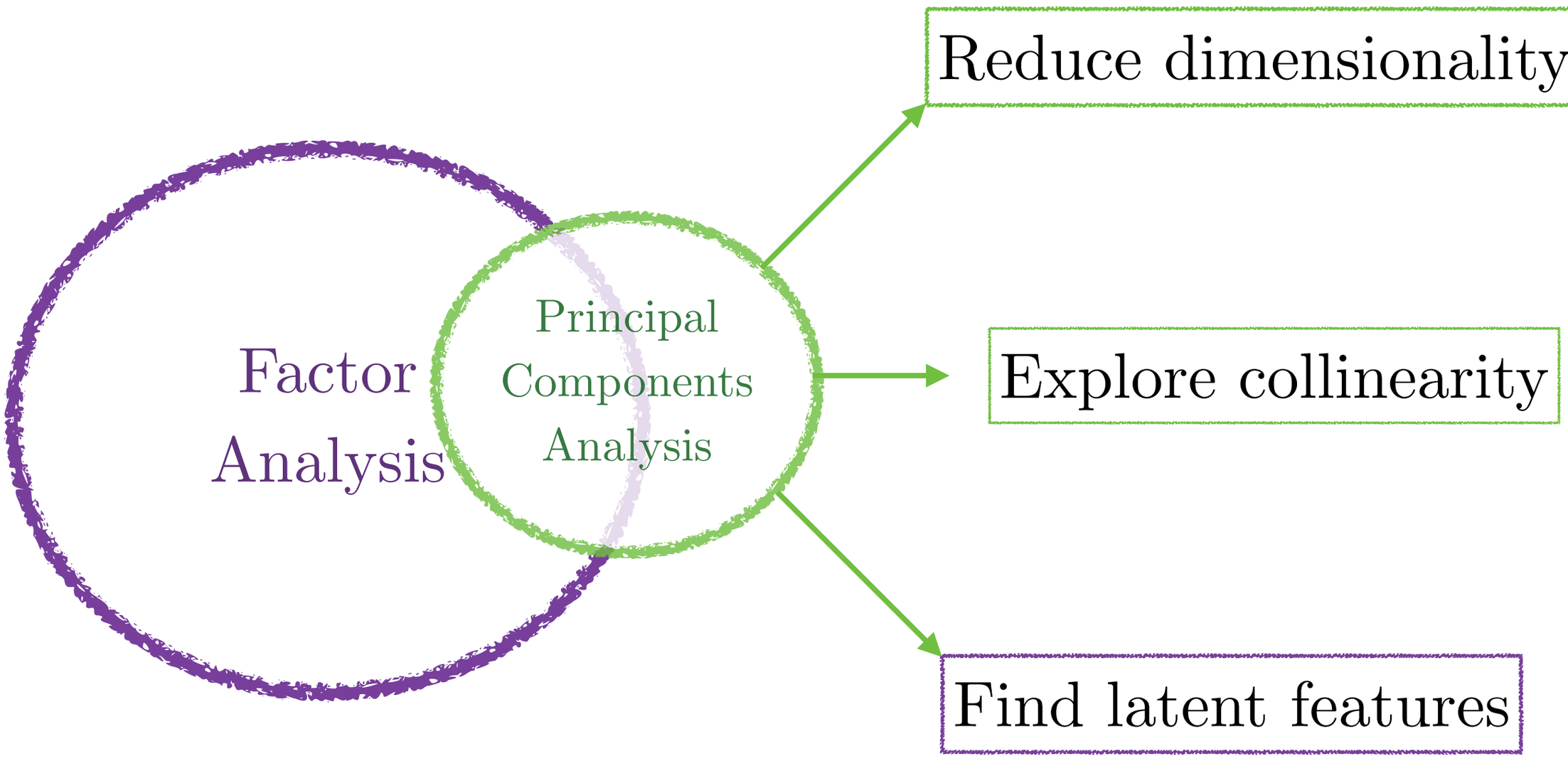
PCA is one method for factor analysis. Other methods exist, like Maximum Likelihood Estimation, that are beyond our scope.



When is PCA inside FA and when is it not?



PCA has many use cases



Recall this example from Day 1

$$\mathbf{B} = \begin{array}{l} \\ \text{“cat”} \\ \text{“dog”} \\ \text{“tired”} \\ \text{“injured”} \\ \text{“ankle”} \\ \text{“sprained”} \end{array} \begin{array}{c} \begin{array}{cccc} \textit{doc}1 & \textit{doc}2 & \textit{doc}3 & \textit{doc}4 \end{array} \\ \begin{pmatrix} 1 & 2 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \end{array}$$

$$\mathbf{B} \approx \begin{array}{l} \\ \text{“cat”} \\ \text{“dog”} \\ \text{“tired”} \\ \text{“injured”} \\ \text{“ankle”} \\ \text{“sprained”} \end{array} \begin{array}{c} \begin{array}{cc} \textit{Factor}1 & \textit{Factor}2 \end{array} \\ \begin{pmatrix} 1.0 & 0 \\ 1.6 & 0 \\ 0.4 & 0.4 \\ 0 & 0.8 \\ 0 & 0.8 \\ 0 & 0.8 \end{pmatrix} \end{array} \begin{array}{c} \begin{array}{cccc} \textit{doc}1 & \textit{doc}2 & \textit{doc}3 & \textit{doc}4 \end{array} \\ \begin{pmatrix} 1.0 & 1.7 & 0 & 0.0 \\ 0 & 0.1 & 0.9 & 1.1 \end{pmatrix} \end{array}$$

Factor Analysis

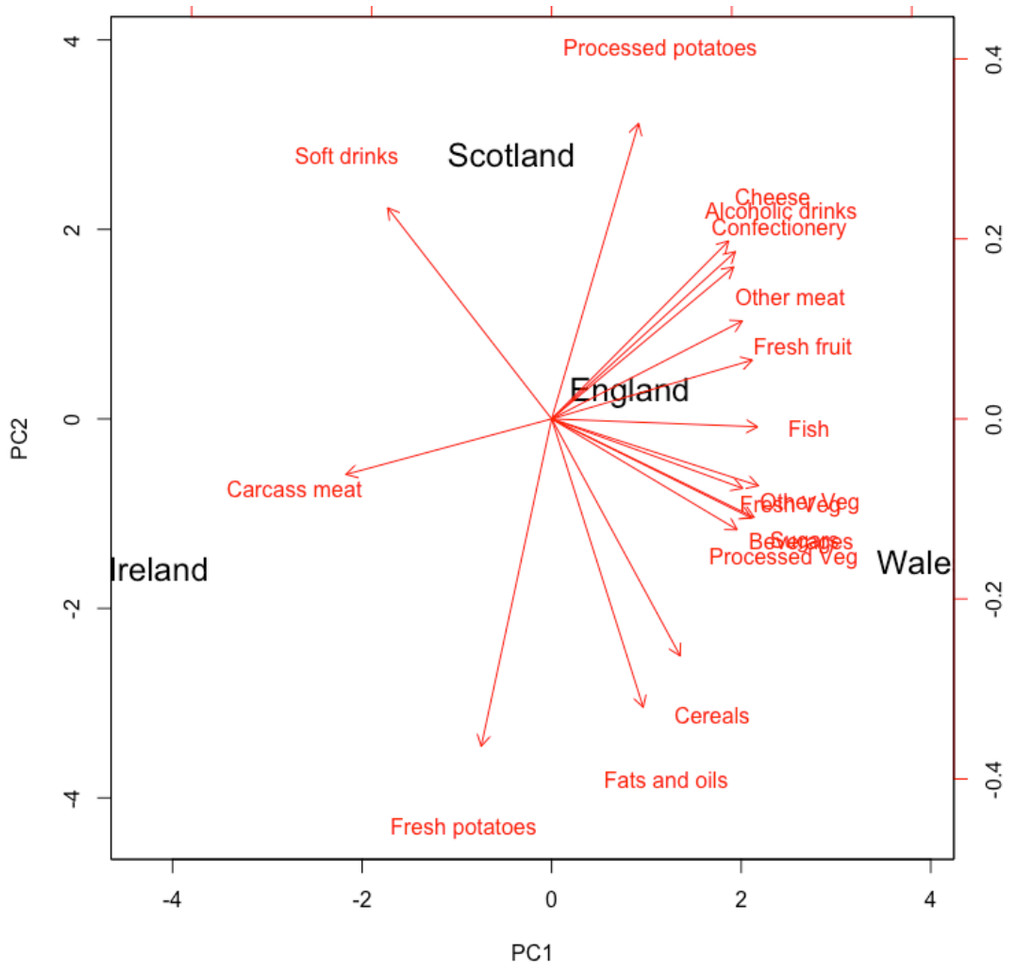
$$\mathbf{B} \approx \begin{array}{c} \\ \text{“cat”} \\ \text{“dog”} \\ \text{“tired”} \\ \text{“injured”} \\ \text{“ankle”} \\ \text{“sprained”} \end{array} \begin{array}{c} \begin{array}{cc} \textit{Factor1} & \textit{Factor2} \end{array} \\ \begin{pmatrix} 1.0 & 0 \\ 1.6 & 0 \\ 0.4 & 0.4 \\ 0 & 0.8 \\ 0 & 0.8 \\ 0 & 0.8 \end{pmatrix} \end{array} \begin{array}{c} \begin{array}{cccc} \textit{doc1} & \textit{doc2} & \textit{doc3} & \textit{doc4} \end{array} \\ \begin{pmatrix} 1.0 & 1.7 & 0 & 0.0 \\ 0 & 0.1 & 0.9 & 1.1 \end{pmatrix} \end{array}$$

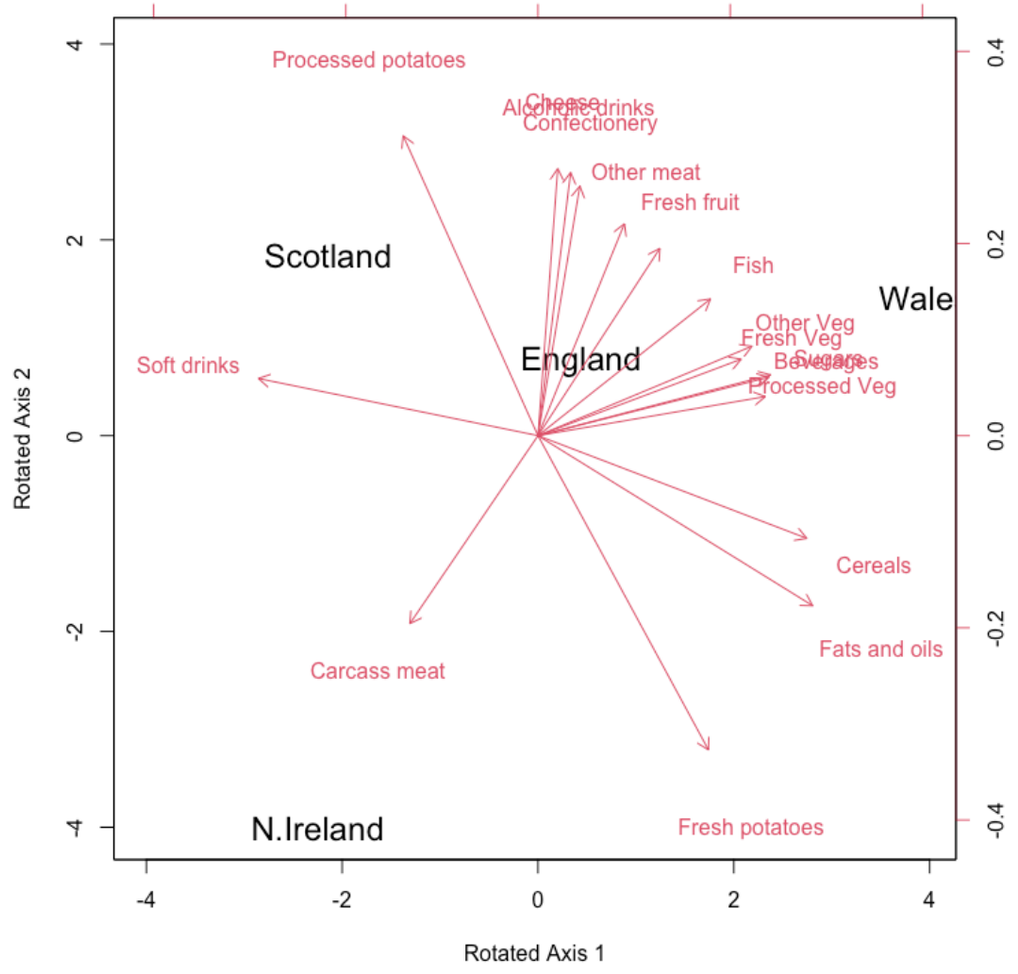
What made this example so easy to interpret?

Sparsity of the loadings!

Rotations => Sparsity

- We can achieve more sparse loadings matrices by rotating our principal component axes to line up with our original variables
- We do this ***after*** we have reduced the dimensionality of our data, in the projected data.





Rotations of PCA Axes

- Once the number of principal components are decided, we can **rotate the axes** so they more closely line up with groups of original variables.
  - **Does NOT change the total amount of variance explained** by *all* the factors together
  - Only **changes the proportion of the variance** explained by each factor
  - **Goal** is to make solution **more interpretable**
    - i.e. make **loadings more sparse!**

IPIP Test Data

International Personality Item Pool

Personality test designed to measure
5 latent features of personality with
Likert scale response questions

Agreeable

A1	I feel little concern for others.
A2	I am interested in people.
A3	I insult people.
A4	I sympathize with others feelings.
A5	I am not interested in other peoples
A6	I have a soft heart.
A7	I am not really interested in others.
A8	I take time out for others.
A9	I feel others emotions.
A10	I make people feel at ease.

Conscientious

C1	I am always prepared.
C2	I leave my belongings around.
C3	I pay attention to details.
C4	I make a mess of things.
C5	I get chores done right away.
C6	I often forget to put things back in their
C7	I like order.
C8	I shirk my duties.
C9	I follow a schedule.
C10	I am exacting in my work.

Extroverted

E1	I am the life of the party.
E2	I dont talk a lot.
E3	I feel comfortable around people.
E4	I keep in the background.
E5	I start conversations.
E6	I have little to say.
E7	I talk to a lot of different people at parties.
E8	I dont like to draw attention to myself.
E9	I dont mind being the center of attention.
E10	I am quiet around strangers.

Neurotic

N1	I get stressed out easily.
N2	I am relaxed most of the time.
N3	I worry about things.
N4	I seldom feel blue.
N5	I am easily disturbed.
N6	I get upset easily.
N7	I change my mood a lot.
N8	I have frequent mood swings.
N9	I get irritated easily.
N10	I often feel blue.

Open

O1	I have a rich vocabulary.
O2	I have difficulty understanding abstract
O3	I have a vivid imagination.
O4	I am not interested in abstract ideas.
O5	I have excellent ideas.
O6	I do not have a good imagination.
O7	I am quick to understand things.
O8	I use difficult words.
O9	I spend time reflecting on things.
O10	I am full of ideas.

race Chosen from a drop down menu. 1=Mixed Race, 2=Arctic (Siberian, Eskimo), 3=Caucasian (European), 4=Caucasian (Indian), 5=Caucasian (Middle East), 6=Caucasian (North African, Other), 7=Indigenous Australian, 8=Native American, 9=North East Asian (Mongol, Tibetan, Korean Japanese, etc), 10=Pacific (Polynesian, Micronesian, etc), 11=South East Asian (Chinese, Thai, Malay, Filipino, etc), 12=West African, Bushmen, Ethiopian, 13=Other (0=missed)
age entered as text (individuals reporting age < 13 were not recorded)
engnat Response to "is English your native language?". 1=yes, 2=no (0=missed)
gender Chosen from a drop down menu. 1=Male, 2=Female, 3=Other (0=missed)
hand "What hand do you use to write with?". 1=Right, 2=Left, 3=Both (0=missed)

On this page users were also asked to confirm that their answers were accurate and could be used for research. Participants who did not were not recorded).

Some values were calculated from technical information.

country The participant's technical location. ISO country code.
source How the participant came to the test. Based on HTTP Referer. 1=from another page on the test website, 2=from google, 3=from facebook, 4=from any url with ".edu" in its domain name (e.g. xxx.edu, xxx.edu.au), 6=other source, or HTTP Referer not provided.

```
cas;
caslib _all_ assign;

proc factor data=public.ipip;
var e1--o10;
run;
```

The **default proc factor is NO different mathematically from proc princomp**. The only difference is that SAS scales the Principal Component columns by the standard deviations ($\sqrt{\lambda_i}$).

The first factor is precisely the same direction as the first PC by default.

The FACTOR Procedure
Initial Factor Method: Principal Components
Prior Communality Estimates: ONE

Eigenvalues of the Correlation Matrix: Total = 50 Average = 1

	Eigenvalue	Difference	Proportion	Cumulative
1	8.04739	3.43104	0.16090	0.16090
2	4.61635	0.86800	0.09230	0.25330
3	3.74835	0.19487	0.07500	0.32820
4	3.55349	0.78998	0.07110	0.39930
5	2.76351	1.18244	0.05530	0.45460
6	1.58108	0.25046	0.03160	0.48620
7	1.33062	0.27699	0.02660	0.51280
8	1.05363	0.08727	0.02110	0.53390
9	0.96636	0.04017	0.01930	0.55320
10	0.92619	0.02937	0.01850	0.57170
11	0.89682	0.03318	0.01790	0.58970
12	0.86364	0.01677	0.01730	0.60600

8 factors will be retained by the MINEIGEN criterion.

		Factor Pattern							
		Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
E1	I am the life of the party.	-0.521	0.277	-0.292	-0.135	0.251	0.142	0.007	0.072
E2	I dont talk a lot.	0.510	-0.359	0.270	0.142	-0.216	0.198	-0.013	0.142
E3	I feel comfortable around people.	-0.710	0.158	-0.116	-0.218	0.113	0.099	0.071	0.045
E4	I keep in the background.	0.584	-0.203	0.314	0.178	-0.228	0.199	0.044	-0.001
E5	I start conversations.	-0.658	0.319	-0.164	-0.125	0.218	-0.005	0.066	-0.007

Min eigen criteria: The default for choosing the number of factors is to only choose factors with an associated eigenvalue greater or equal to 1. This is because the correlation PCA is used so you want each factor to explain at least as much variance as one of your original variables!

Specify number of factors = 5 because **5 latent constructs** we are trying to extract.

```
proc factor data=public.ipip out=ipipfactors nfactors=5;
var e1--o10;
run;
```

Variance Explained by Each Factor

Factor1	Factor2	Factor3	Factor4	Factor5
8.0473909	4.6163522	3.7483528	3.5534873	2.7635116

Eigenvalues of the Correlation Matrix: Total = 50 Average = 1

	Eigenvalue	Difference	Proportion	Cumulative
1	8.04739	3.43104	0.16090	0.16090
2	4.61635	0.86800	0.09230	0.25330
3	3.74835	0.19487	0.07500	0.32820
4	3.55349	0.78998	0.07110	0.39930
5	2.76351	1.18244	0.05530	0.45460

Communalities tell us how much of the variance in *each individual variable* is explained by the factor model.

Final Communality Estimates: Total = 22.729095									
N10	A1	A2	A3	A4	A5	A6	A7	A8	A9
98101	0.26753440	0.46466315	0.35311688	0.64787177	0.51584341	0.45232441	0.55055927	0.43129670	0.56503510

You can think of this as the R^2 of predicting each variable with the factors determined by the model.

This is demonstrated on the next slide.

Final Communalities are the R^2 values that result from regression of an individual variable on all of the estimated factors. Measures how well each variable is explained in this reduced dimensional representation.

```
proc factor data=public.ipip out=ipipfactors nfactors=5;
var e1--o10;
run;
```

```
proc reg data = ipipfactors;
model e1 = factor1--factor5;
run;
```

From communalities table in proc factor

E1
0.51405133

From top table in proc reg

Root MSE	0.85933	R-Square	0.5141
Dependent Mean	2.62894	Adj R-Sq	0.5139
Coeff Var	32.68738		

Factor Pattern

		Factor1	Factor2	Factor3	Factor4	Factor5
E1	I am the life of the party.	-0.521	0.277	-0.292	-0.135	0.251
E2	I dont talk a lot.	0.510	-0.359	0.270	0.142	-0.216
E3	I feel comfortable around people.	-0.710	0.158	-0.116	-0.218	0.113
E4	I keep in the background.	0.584	-0.203	0.314	0.178	-0.228
E5	I start conversations.	-0.658	0.319	-0.164	-0.125	0.218
E6	I have little to say.	0.563	-0.273	0.200	-0.059	-0.121
E7	I talk to a lot of different people at parties.	-0.626	0.294	-0.237	-0.166	0.212
E8	I dont like to draw attention to myself.	0.379	-0.233	0.361	0.112	-0.238
E9	I dont mind being the center of attention.	-0.466	0.252	-0.337	-0.033	0.262
E10	I am quiet around strangers.	0.572	-0.160	0.300	0.156	-0.186
N1	I get stressed out easily.	0.440	0.489	0.245	0.022	0.232
N2	I am relaxed most of the time.	-0.367	-0.305	-0.202	-0.025	-0.240
N3	I worry about things.	0.349	0.465	0.347	0.072	0.170
N4	I seldom feel blue.	-0.288	-0.254	-0.100	-0.123	0.002
N5	I am easily disturbed.	0.405	0.406	0.075	-0.064	0.200
N6	I get upset easily.	0.473	0.543	0.167	-0.007	0.251
N7	I change my mood a lot.	0.461	0.528	0.016	0.075	0.265
N8	I have frequent mood swings.	0.492	0.536	0.020	0.068	0.273
N9	I get irritated easily.	0.503	0.424	0.048	0.088	0.375
N10	I often feel blue.	0.552	0.400	0.121	0.170	0.085

Table Continued From Previous Slide

		Factor1	Factor2	Factor3	Factor4	Factor5
A1	I feel little concern for others.	0.264	-0.181	-0.270	0.004	0.303
A2	I am interested in people.	-0.505	0.380	0.149	-0.093	-0.184
A3	I insult people.	0.281	0.114	-0.401	0.148	0.280
A4	I sympathize with others feelings.	-0.325	0.414	0.486	-0.117	-0.348
A5	I am not interested in other peoples problems.	0.380	-0.369	-0.333	0.131	0.328
A6	I have a soft heart.	-0.157	0.385	0.444	-0.161	-0.238
A7	I am not really interested in others.	0.532	-0.363	-0.217	0.147	0.260
A8	I take time out for others.	-0.377	0.300	0.367	-0.086	-0.238
A9	I feel others emotions.	-0.339	0.446	0.435	-0.065	-0.242
A10	I make people feel at ease.	-0.532	0.202	0.149	-0.056	-0.036
C1	I am always prepared.	-0.315	-0.265	0.364	0.166	0.351
C2	I leave my belongings around.	0.143	0.347	-0.359	0.063	-0.327
C3	I pay attention to details.	-0.210	-0.084	0.343	0.326	0.191
C4	I make a mess of things.	0.431	0.432	-0.293	0.017	-0.193
C5	I get chores done right away.	-0.310	-0.244	0.415	-0.058	0.386
C6	I often forget to put things back in their proper	0.263	0.362	-0.369	0.016	-0.321
C7	I like order.	-0.127	-0.151	0.440	0.120	0.369
C8	I shirk my duties.	0.394	0.240	-0.350	-0.022	-0.145
C9	I follow a schedule.	-0.257	-0.155	0.470	0.007	0.408
C10	I am exacting in my work.	-0.269	-0.113	0.335	0.280	0.279

Table Continued From Previous Slide

		Factor1	Factor2	Factor3	Factor4	Factor5
O1	I have a rich vocabulary.	-0.200	0.032	-0.080	0.619	-0.033
O2	I have difficulty understanding abstract ideas.	0.281	0.064	0.096	-0.549	0.191
O3	I have a vivid imagination.	-0.101	0.235	-0.066	0.543	-0.094
O4	I am not interested in abstract ideas.	0.206	-0.054	0.043	-0.473	0.256
O5	I have excellent ideas.	-0.371	0.055	-0.074	0.559	0.115
O6	I do not have a good imagination.	0.249	-0.135	0.085	-0.480	0.119
O7	I am quick to understand things.	-0.317	-0.075	0.011	0.507	0.046
O8	I use difficult words.	-0.036	0.101	-0.145	0.610	0.003
O9	I spend time reflecting on things.	0.012	0.196	0.230	0.403	-0.110
O10	I am full of ideas.	-0.332	0.156	-0.113	0.617	0.014

These factors are not as easy to interpret as we might like.

Solution? Rotation.

Rotations

Two most common rotations for factor analysis:

Varimax: The goal of *varimax* rotation is to simplify the *columns* of the factor matrix, so that **each factor has only a few variables with large loadings.**

Quartimax: The goal of *quartimax* rotation is to simply the *rows* of the factor matrix, so that **each variable will only load on a few factors.**

Rotations

```
proc factor data=public.ipip
            nfactors=5
            rotate=varimax
            out=ipipScores;
var e1--o10;
run;
```

Rotated Factor Pattern

		Factor1	Factor2	Factor3	Factor4	Factor5
E1	I am the life of the party.	0.712	-0.049	0.056	0.011	0.032
E2	I dont talk a lot.	-0.716	-0.003	-0.121	0.029	-0.035
E3	I feel comfortable around people.	0.669	-0.260	0.264	0.132	-0.017
E4	I keep in the background.	-0.733	0.153	-0.053	-0.023	-0.001
E5	I start conversations.	0.745	-0.076	0.219	0.101	0.071
E6	I have little to say.	-0.602	0.088	-0.158	-0.028	-0.232
E7	I talk to a lot of different people at parties.	0.753	-0.100	0.162	0.046	0.029
E8	I dont like to draw attention to myself.	-0.623	0.024	0.058	0.067	-0.021
E9	I dont mind being the center of attention.	0.670	-0.032	-0.027	-0.008	0.115
E10	I am quiet around strangers.	-0.680	0.189	-0.056	-0.022	-0.017
N1	I get stressed out easily.	-0.107	0.725	0.068	-0.011	-0.073
N2	I am relaxed most of the time.	0.111	-0.553	0.017	-0.054	0.073
N3	I worry about things.	-0.139	0.662	0.176	0.059	-0.008
N4	I seldom feel blue.	0.157	-0.361	-0.040	0.102	-0.071
N5	I am easily disturbed.	-0.036	0.588	-0.024	-0.108	-0.137
N6	I get upset easily.	-0.057	0.768	0.029	-0.082	-0.096
N7	I change my mood a lot.	0.001	0.733	-0.082	-0.152	-0.004
N8	I have frequent mood swings.	-0.014	0.758	-0.091	-0.158	-0.019
N9	I get irritated easily.	-0.040	0.738	-0.187	-0.041	-0.032
N10	I often feel blue.	-0.252	0.648	-0.038	-0.165	0.058

Table Continued From Previous Slide

		Factor1	Factor2	Factor3	Factor4	Factor5
A1	I feel little concern for others.	0.003	0.082	-0.502	-0.007	-0.094
A2	I am interested in people.	0.356	-0.054	0.571	-0.005	0.091
A3	I insult people.	0.128	0.274	-0.462	-0.198	0.092
A4	I sympathize with others feelings.	0.039	0.064	0.800	0.041	0.017
A5	I am not interested in other peoples problems.	-0.136	0.022	-0.704	0.013	-0.024
A6	I have a soft heart.	-0.007	0.164	0.647	0.034	-0.078
A7	I am not really interested in others.	-0.316	0.103	-0.662	-0.001	-0.047
A8	I take time out for others.	0.120	-0.018	0.638	0.091	0.042
A9	I feel others emotions.	0.117	0.120	0.726	0.071	0.065
A10	I make people feel at ease.	0.356	-0.125	0.417	0.159	0.092
C1	I am always prepared.	0.047	-0.100	0.004	0.653	0.122
C2	I leave my belongings around.	0.056	0.106	0.052	-0.585	0.144
C3	I pay attention to details.	-0.040	0.025	0.089	0.465	0.294
C4	I make a mess of things.	-0.060	0.381	-0.046	-0.586	0.020
C5	I get chores done right away.	0.092	-0.082	0.062	0.672	-0.098
C6	I often forget to put things back in their proper	0.001	0.174	0.011	-0.636	0.073
C7	I like order.	-0.045	0.096	0.025	0.608	0.038
C8	I shirk my duties.	-0.059	0.232	-0.171	-0.519	-0.033
C9	I follow a schedule.	0.064	0.037	0.091	0.680	-0.046
C10	I am exacting in my work.	0.037	0.005	0.057	0.535	0.250

Table Continued From Previous Slide

		Factor1	Factor2	Factor3	Factor4	Factor5
O1	I have a rich vocabulary.	0.033	-0.043	-0.041	0.048	0.652
O2	I have difficulty understanding abstract ideas.	-0.003	0.232	-0.031	0.006	-0.613
O3	I have a vivid imagination.	0.034	0.118	0.065	-0.091	0.588
O4	I am not interested in abstract ideas.	0.033	0.135	-0.133	0.085	-0.541
O5	I have excellent ideas.	0.225	-0.059	-0.026	0.179	0.621
O6	I do not have a good imagination.	-0.102	0.052	-0.091	0.043	-0.556
O7	I am quick to understand things.	0.082	-0.138	-0.013	0.215	0.542
O8	I use difficult words.	-0.000	0.090	-0.130	-0.048	0.614
O9	I spend time reflecting on things.	-0.163	0.191	0.207	0.052	0.396
O10	I am full of ideas.	0.203	-0.014	0.024	0.052	0.696

Now I have 5 NEW variables,

$$\textbf{Factor1} = \textbf{Extroversion}$$
$$\textbf{Factor2} = \textbf{Neuroticism}$$
$$\textbf{Factor3} = \textbf{Agreeableness}$$
$$\textbf{Factor4} = \textbf{Conscientiousness}$$
$$\textbf{Factor5} = \textbf{Openness}$$

I can learn about my observations by observing their scores on these factors, or by using these factors in a model!

Exploring Gender and Handedness

```
data ipipscores;
set ipipscores;
female=0;
lefty=0;
if gender=2 then female=1;
if hand=2 then lefty=1;
run;
```

Are Females more Agreeable?

```
proc logistic data=ipipscores;
model female(event='1') = factor3;
run;
```

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	0.4548	0.0150	920.3256	<.0001
Factor3	1	0.4699	0.0154	927.1113	<.0001

(yup.)

What else can we say about Females vs. NonFemales?

```
proc logistic data=ipipscores;
model female(event='1') = factor1--factor5;
run;
```

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	0.4753	0.0154	946.8186	<.0001
Factor1(E)	1	0.1221	0.0155	62.2893	<.0001
Factor2(N)	1	0.3955	0.0158	625.5999	<.0001
Factor3(A)	1	0.4975	0.0160	964.5456	<.0001
Factor4(C)	1	0.1046	0.0155	45.7394	<.0001
Factor5(O)	1	-0.2637	0.0158	280.2033	<.0001

Are Lefties more Neurotic?

```
proc logistic data=ipipscores;
model female(event='1') = factor3;
run;
```

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-2.3457	0.0252	8653.0098	<.0001
Factor2	1	0.0223	0.0252	0.7808	0.3769

(nah.)

Differences for Handedness?

```
proc logistic data=ipipscores;
model lefty(event='1') = factor1--factor5;
run;
```

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-2.3485	0.0253	8629.0374	<.0001
Factor1(E)	1	-0.0309	0.0251	1.5138	0.2186
Factor2(N)	1	0.0221	0.0252	0.7689	0.3806
Factor3(A)	1	-0.0151	0.0250	0.3678	0.5442
Factor4(C)	1	-0.0202	0.0252	0.6449	0.4219
Factor5(O)	1	0.0724	0.0255	8.0747	0.0045