Exploratory Factor Analysis
Common factor analysis

- This is what people generally mean when they say "factor analysis"
- This family of techniques uses an estimate of common variance among the original variables to generate the factor solution
- It is based on the fundamental assumption that some underlying factors, which are smaller in number than the observed variables, are responsible for the covariation among them
Questions to answer

- How many different factors are needed to explain the pattern of relationships among these variables?
- What is the nature of those factors?
- How well do the hypothesized factors explain the observed data?
- How much purely random or unique variance does each observed variable include?

- As with previous techniques we have discussed, the goal is *dimension reduction*, i.e. to describe a number of variables in a simpler form
Issues

- Factor solutions in the exploratory endeavor will be different depending on the data, algorithm and other researcher choices.
- The goal with exploratory factor analysis is to discover structure, not determine it.
- There are also differences in reporting such that one may see quite different results from sample to sample, study to study.
Factor Analysis

- There are four basic steps:
  - data collection and generation of the correlation matrix
  - extraction of initial factor solution
  - rotation and interpretation (also validation)
  - construction of scales or factor scores to use in further analyses

- A good factor:
  - Makes sense
  - Will be easy to interpret
  - Possesses simple structure
    - Items have low cross-loadings
Factor Analysis

- Factor analysis can be seen as a family of techniques, of which both PCA and EFA are members.
- Factor analysis is a statistical approach that can be used to analyze interrelationships among a large number of variables and to explain these variables in terms of their common underlying dimensions (factors).
- It involves finding a way of condensing the information contained in a number of original variables into a smaller set of dimensions (factors) with a minimum loss of information.
Principal Components Analysis

- Principle components analysis (PCA) is a statistical technique applied to a single set of variables to discover which variables in the set form coherent subsets that are independent of one another
  - Provides a unique solution, so that the original data, the covariance or correlation matrix, can be reconstructed from the results
  - Looks at the total variance among the variables, so the solution generated will include as many factors/components as there are variables, although it is unlikely that they will all meet the criteria for retention
- Variables that are correlated with one another which are also largely independent of other subsets of variables are combined into factors
- Factors are generated which are thought to be representative of the underlying processes that have created the correlations among variables
- The underlying notion of PCA is that the observed variables can be transformed into linear combinations of an underlying set of hypothesized or unobserved components (factors)
- PCA is typically exploratory in nature
Common factor model

- PCA and common factor analysis may utilize a similar method and are conducted with similar goals in mind.
- The difference between PCA and common FA involves the underlying model:
  - The common factor model for factor analysis
    - PCA assumes that all variance is common, which is akin to assuming they are perfectly reliable:
      - All ‘unique factors’, i.e. sources of variability not attributable to a factor, set equal to zero
    - The common factor model on the other hand holds that the observed variance in each measure is attributable to relatively small set of common factors (latent characteristics common to two or more variables), and a single specific factor unrelated to any other underlying factor in the model.
Comparison of underlying models

- **PCA**
  - Extraction is the process of forming PCs as linear combinations of the measured variables as we have done with our other techniques
    - \( PC_1 = b_{11}X_1 + b_{21}X_2 + ... + b_{k1}X_k \)
    - \( PC_2 = b_{12}X_1 + b_{22}X_2 + ... + b_{k2}X_k \)
    - \( PC_f = b_{1f}X_1 + b_{2f}X_2 + ... + b_{kf}X_k \)

- **Common factor analysis**
  - Each measure \( X \) has two contributing sources of variation: the common factor \( \xi \) and the specific or unique factor \( \delta \):
    - \( X_1 = \lambda_1 \xi + \delta_1 \)
    - \( X_2 = \lambda_2 \xi + \delta_2 \)
    - \( X_f = \lambda_f \xi + \delta_f \)
Example

- Consider the following example from Holzinger and Swineford 1939
  - Ol’ skool yeah!
- Variables
  - Paragraph comprehension
  - Sentence completion
  - Word meaning
  - Addition
  - Counting dots
- Each person’s score is a reflection of the weighted combination of the common factor (latent variable) and measurement error (uniqueness, unreliability)
- In these equations, $\lambda$ represents the extent to which each measure reflects the underlying common factor
Factor Analysis

- When standardized the variance can be decomposed as follows:

\[ \text{var}(X_i) = \text{var}(\lambda_i \xi + \delta_i) = \lambda^2 + \text{var}(\delta_i) = 1 \]

- \(\lambda_i\), in unsquared form, is now interpretable as a correlation coefficient and its square the proportion of the variation in X accounted for by the common factor, i.e. the communality.

- The remaining is that which is accounted for by the specific factor or uniqueness (individual differences, measurement error, some other known factor e.g. intelligence)

\[ \text{var}(\delta) = \theta_{ii} \]

\[ 1 - \theta_{ii}^2 = \text{communality} \]
Measurement error

- The principal axis extraction method of EFA can be distinguished from PCA in terms of simply having communalities on the diagonal of the correlation matrix instead of 1s
- What does this mean?
- Having 1s assumes each item/variable is perfectly reliable, i.e. there is no measurement error in its ability to distinguish among cases/individuals
Measurement error

- The fact that our estimates in psych are not perfectly reliable suggests that we use methods that take this into account.
- This is a reason to use EFA over PCA in cases involving measurement scales or items, as the communalities can be seen as the lower bound (i.e. conservative) estimate of the variables’ reliability.
- Note however, that low communalities are not interpreted as evidence of poor fit so much as evidence that the variables analyzed have little in common with one another, and thus are not ‘reliable’ measures of a proposed factor solution.
Two-factor solution

- From our example before, we could have selected a two factor solution
  - Quantitative vs. Verbal reasoning
- Now we have three sources of variation observed in test scores
  - Two common factors and one unique factor
- As before, each $\lambda$ reflects the extent to which each common factor contributes to the variance of each test score
- The communality for each variable is now $\lambda^2 + \lambda^2$. 

\[
X_1 = \lambda_{11} \xi_1 + \lambda_{12} \xi_2 + \delta_1 \\
X_2 = \lambda_{21} \xi_1 + \lambda_{22} \xi_2 + \delta_2 \\
X_3 = \lambda_{31} \xi_1 + \lambda_{32} \xi_2 + \delta_3 \\
X_4 = \lambda_{41} \xi_1 + \lambda_{42} \xi_2 + \delta_4 \\
X_5 = \lambda_{51} \xi_1 + \lambda_{52} \xi_2 + \delta_5 
\]
Analysis: the Correlation Matrices

- Observed Correlation Matrix
  - Note that the manner in which missing values are dealt with will determine the observed correlation matrix
  - In typical FA settings one may have quite a few, and so casewise deletion would not be appropriate
  - Best would be to produce a correlation matrix resulting from some missing values analysis (e.g. EM algorithm)

- Reproduced Correlation Matrix
  - That which is produced by the factor solution
    - Recall that in PCA it is identical to the observed
    - It is the product of the matrix of loadings and the transpose of the pattern matrix (partial loadings)*

- Residual Correlation Matrix
  - The difference between the two

\[ \bar{R} = AA' \]

\[ R_{res} = R - \bar{R} \]
Analysis: is the data worth reducing?

- **The Kaiser-Meyer-Olkin Measure of Sampling Adequacy**
  - A statistic that indicates the proportion of variance in your variables that might be caused by common underlying factors
  - An index for comparing the magnitudes of the observed correlation coefficients to the magnitudes of the partial correlation coefficients
    - If two variables share a common factor with other variables, their partial correlation will be small once the factor is taken into account
    - High values (close to 1.0) generally indicate that a factor analysis may be useful with your data.
      - If the value is less than 0.50, the results of the factor analysis probably won't be very useful.

- **Bartlett's test of sphericity**
  - Tests the hypothesis that your correlation matrix is an identity matrix (1s on the diagonal, 0s off-diagonals), which would indicate that your variables are unrelated and therefore unsuitable for structure detection
Analysis: Extraction Methods

- **Principal (Axis) Factors**
  - Estimates of communalities (SMC) are in the diagonal; used as starting values for the communality estimation (iterative)
  - Removes unique and error variance
  - Solution depends on quality of the initial communality estimates
- **Maximum Likelihood**
  - Computationally intensive method for estimating loadings that maximize the likelihood of sampling the observed correlation matrix from a population
- **Unweighted least squares**
  - Minimize off diagonal residuals between reproduced and original R matrix
- **Generalized (weighted) least squares**
  - Also minimizes the off diagonal residuals
  - Variables with larger communalities are given more weight in the analysis
- **Alpha factoring**
  - Maximizes the reliability of the factors
- **Image factoring**
  - Minimizes ‘unique’ factors consisting of essentially one measured variable
Analysis: Rotation Methods

- After extraction, initial interpretation may be difficult
- Rotation is used to improve interpretability and utility
  - Refer back to the PCA mechanics handout for the geometric interpretation of rotation
  - By placing a variable in the n-dimensional space specified by the factors involved, factor loadings are the cosine of the angle formed by a vector from the origin to that coordinate and the factor axis
    - Note how PCA would be distinguished from multiple regression
    - PCA minimizes the squared distances to the axis, with each point mapping on to the axis forming a right angle (as opposed to ‘dropping straight down’ in MR)
    - MR is inclined to account for variance in the DV, where PCA will tilt more to whichever variable exhibits the most variance
Analysis: Rotation Methods

- So factors are the axes
  - Orthogonal Factors are at right angles
  - Oblique rotation allows for other angles
    - Often achieve simpler structure, though at the cost that you must also consider the factor inter-correlations when interpreting results
- Repositioning the axes changes the loadings on the factor but keeps the relative positioning of the points the same
- Length of the line from the origin to the variable coordinates is equal to the communality for that variable
Example: Rotation

![Diagram showing unrotated and rotated solutions for pain relievers.](image)

**Table 5.6** Factor loadings for pain relievers

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Unrotated Solution</th>
<th>Rotated Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>No upset stomach</td>
<td>0.579 -0.452</td>
<td>0.139 0.721</td>
</tr>
<tr>
<td>No bad side effects</td>
<td>0.522 -0.572</td>
<td>0.017 0.774</td>
</tr>
<tr>
<td>Stops the pain</td>
<td>0.645 0.436</td>
<td>0.772 0.097</td>
</tr>
<tr>
<td>Works quickly</td>
<td>0.542 0.542</td>
<td>0.764 -0.051</td>
</tr>
<tr>
<td>Keeps me awake</td>
<td>-0.476 0.596</td>
<td>-0.034 -0.762</td>
</tr>
<tr>
<td>Limited relief</td>
<td>-0.613 -0.439</td>
<td>-0.750 -0.074</td>
</tr>
</tbody>
</table>

Note that the variance of the two factors for the original and rotated solutions sum to the same amount.
Analysis: Rotation Methods

- **Orthogonal rotation** keeps factors uncorrelated while increasing the meaning of the factors
  - **Varimax** – most popular
    - ‘Cleans up the factors’
    - Makes large loadings larger and small loadings smaller
  - **Quartimax**
    - ‘Cleans up the variables’
    - Each variable loads mainly on one factor
    - Varimax works on the columns of the loading matrix; Quartimax works on the rows
    - Not used as often; simplifying variables is not usually a goal
  - **Equamax**
    - Hybrid of the two that tries to simultaneously simplify factors and variables
    - Not that popular either
Analysis: Rotation Methods

- Oblique Rotation Techniques
  - Direct Oblimin
    - Begins with an unrotated solution
    - Has a parameter $\gamma$ that allows the user to define the amount of correlation acceptable; gamma values near $-4 \rightarrow$ orthogonal, $0$ leads to mild correlations (also direct quartimin) and close to $1$ highly correlated
  - Promax$^2$
    - Solution is orthogonally rotated initially (varimax)
    - This is followed by oblique rotation
    - Orthogonal loadings are raised to powers in order to drive down small to moderate loadings
Analysis: Orthogonal vs. Oblique output

- **Orthogonal Rotation**
  - **Factor matrix**
    - Correlation between observed variable and factor for the unrotated solution
  - **Pattern vs. structure matrix**
    - **Structure matrix**
      - Loadings, i.e. structure coefficients
      - Correlation between observed variable and factor
    - **Pattern matrix**
      - Standardized, partialled coefficients (weights, loadings)
    - These structure and pattern matrices are the same in orthogonal solutions and so will not be distinguished
  - **Factor Score Coefficient matrix**
    - Coefficients used to calculate factor scores (like regression coefficients) from original variables (standardized)
Analysis: Orthogonal vs. Oblique output

- **Oblique Rotation**
  - Factor matrix
    - Correlation between observed variable and factor for the unrotated solution
  - Structure Matrix
    - Simple correlation between factors and variables
    - Factor loading matrix
  - Pattern Matrix
    - Unique relationship between each factor and variable that takes into account the correlation between the factors
    - The standardized regression coefficient from the common factor model
    - The more factors, the lower the pattern coefficients as a rule since there will be more common contributions to variance explained
  - For oblique rotation, the researcher looks at both the structure and pattern coefficients when attributing a label to a factor

- Factor Score Coefficient matrix
  - Again used to derive factors scores from the original variables
- Factor Correlation Matrix
  - correlation between the factors
Analysis: Factor scores

- Factor scores can be derived in a variety of ways, some of which are presented here:
  - Regression
    - Regression factor scores have a mean of 0 and variance equal to the squared multiple correlation between the estimated factor scores and the true factor values. They can be correlated even when factors are assumed to be orthogonal. The sum of squared residuals between true and estimated factors over individuals is minimized.
  - Least squares
    - Minimizes squared residuals of scores and true factor scores
    - Same approach as above but uses the reproduced R matrix instead of the original
  - Bartlett
    - Minimizes the effect of unique factors (consisting of single variables)
  - Anderson-Rubin
    - Same as Bartlett’s but produces orthogonal factor scores
- Once obtained one can use factor scores in other analyses
  - Recall PC regression
Other stuff: An iterative process

- To get an initial estimate of the communalities, we can simply start with the squared multiple correlation coefficient ($R^2$) for each item regressed on the other remaining items
  - There are other approaches, but with this one we can see that if a measure was completely unreliable its $R^2$ value would be zero
- We run the EFA and come to a solution, however now we can estimate the communalities as the sum of the squared loadings for an item across the factors
- We now use these new estimates as communalities and rerun
- This is done until successive iterations are essentially identical
  - Convergence is achieved
- If convergence is not obtained, another method of factor extraction, e.g. PCA, must be utilized, or if possible, sample size increased
Measurement error vs. sampling error

- Measurement error is the variance not attributable to the factor an observed variable purportedly represents \((1 - \text{Reliability})\)
- Sampling error is the variability in estimates seen as we move from one sample to the next
- Just like every person is different, every sample taken would be
- Note that with successive iterations, we increase the likelihood that we are capitalizing on the unique sampling error associated with a given dataset, thus making our results less generalizable
- As one might expect, with larger samples (and fewer factors to consider) we have less to worry about regarding sampling error, and so might allow for more iterations
- Unfortunately there are no hard and fast rules regarding the limitation of iterations, however you should be aware of the trade off
Other stuff: More on sample size

- How big?
- Assume you’ll need lots
- From some simulation studies
  - 1. 4 or more variables per factor with large structure coefficients (e.g. greater than .6) may work with even small samples
  - 2. 10 variables or more per factor, loadings .4 → N ≥ 150
  - 3. Sample size ≥ 300
- The larger the communalites (i.e. the more reliable), the better off you are
Other stuff: How many factors?

- Refer back to PCA notes, there are many ways to determine this
- But recall that we are doing *exploratory* factor analysis
- As such, just as we suggested with cluster analysis, go with the solution that makes the most sense
- Also how you interpret them is, as it was with previous analyses, entirely subjective
Other stuff: Exploratory vs. Confirmatory

- **Exploratory FA**
  - Summarizing data by grouping correlated variables
  - Investigating sets of measured variables related to theoretical constructs
  - Usually done near the onset of research
  - The type of FA and PCA we are talking about

- **Confirmatory FA**
  - More advanced technique
  - When factor structure is known or at least theorized
  - Testing generalization of factor structure to new data, etc.
  - This is tested through SEM methods
Concrete example

- Beer data we did with PCA
  - See appendix for code if you want to go along
- What influences a consumer’s choice behavior when shopping for beer?
- 200 consumers are asked to rate on a scale of 0-100 how important they consider each of seven qualities when deciding whether or not to buy the six pack:
  - COST of the six pack,
  - SIZE of the bottle (volume)
  - Percentage of ALCOHOL in the beer
  - REPUTATION of the brand
  - COLOR of the beer
  - AROMA of the beer
  - TASTE of the beer
- First perform a PCA with varimax rotation
- In descriptives check correlation coefficients and KMO test of sphericity
- Make sure to select that the number of factors to be extracted equals the number variables (7)
- For easier reading you may want to suppress loadings less than .3 in the options dialog box
First we’ll run a PCA and compare the results

- As always first get to know your correlation matrix
- You should be aware of the simple relationships quite well, and one can already guess the factor structure that may be found
  - The first and last 3 correlate well within their group, reputation, correlates moderately negatively with all the others
- Our tests here indicate we’ll be okay for further analysis

```
Correlation Matrix

<table>
<thead>
<tr>
<th></th>
<th>cost</th>
<th>size</th>
<th>alcohol</th>
<th>reputat</th>
<th>color</th>
<th>aroma</th>
<th>taste</th>
</tr>
</thead>
<tbody>
<tr>
<td>cost</td>
<td>1.00</td>
<td>.832</td>
<td>-.767</td>
<td>-.406</td>
<td>.018</td>
<td>-.046</td>
<td>-.064</td>
</tr>
<tr>
<td>size</td>
<td>.832</td>
<td>1.00</td>
<td>.904</td>
<td>-.392</td>
<td>.179</td>
<td>.098</td>
<td>.026</td>
</tr>
<tr>
<td>alcohol</td>
<td>.767</td>
<td>.904</td>
<td>1.00</td>
<td>-.463</td>
<td>.072</td>
<td>.044</td>
<td>.012</td>
</tr>
<tr>
<td>reputat</td>
<td>-.406</td>
<td>-.392</td>
<td>-.463</td>
<td>1.000</td>
<td>-.443</td>
<td>-.443</td>
<td></td>
</tr>
<tr>
<td>color</td>
<td>.018</td>
<td>.179</td>
<td>.072</td>
<td>-.372</td>
<td>1.000</td>
<td>.999</td>
<td>.903</td>
</tr>
<tr>
<td>aroma</td>
<td>-.046</td>
<td>.098</td>
<td>.044</td>
<td>-.443</td>
<td>.909</td>
<td>1.000</td>
<td>.870</td>
</tr>
<tr>
<td>taste</td>
<td>-.064</td>
<td>.026</td>
<td>.012</td>
<td>-.443</td>
<td>.903</td>
<td>.870</td>
<td>1.000</td>
</tr>
</tbody>
</table>

KMO and Bartlett’s Test

<table>
<thead>
<tr>
<th>KMO and Bartlett’s Test</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Kaiser-Meyer-Olkin Measure of Sampling Adequacy</td>
<td>.665</td>
</tr>
<tr>
<td>Bartlett's Test of Sphericity</td>
<td>Approx. Chi-Square</td>
</tr>
<tr>
<td>df</td>
<td>Sig.</td>
</tr>
</tbody>
</table>
```
PCA

- Recall with PCA we extract *all* the variance
- At this point it looks like we’ll stick with two components/factors which account for almost 85% of the variance
PCA

- We’ve got some strong loadings here, but it’s not easily interpretable
- Perhaps a rotation is in order
- We’ll do varimax and see what we come up with
PCA

- Just going by eigenvalues > 1, it looks like now there maybe a third factor worth considering
- Here, loadings < .3 have been suppressed
- Ah, much nicer, and perhaps we’ll go with a 3 factor interpretation
- One factor related to practical concerns (how cheaply can I get drunk?)
- Another to aesthetic concerns (is it a good beer?)
- One factor is simply reputation (will I look cool drinking it?)

Rotation Sums of Squared Loadings

<table>
<thead>
<tr>
<th>Total</th>
<th>% of Variance</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.808</td>
<td>40.117</td>
<td>40.117</td>
</tr>
<tr>
<td>2.597</td>
<td>37.094</td>
<td>77.211</td>
</tr>
<tr>
<td>1.001</td>
<td>14.295</td>
<td>91.506</td>
</tr>
<tr>
<td>.322</td>
<td>4.593</td>
<td>96.100</td>
</tr>
<tr>
<td>.106</td>
<td>1.520</td>
<td>97.620</td>
</tr>
<tr>
<td>.089</td>
<td>1.270</td>
<td>98.890</td>
</tr>
<tr>
<td>.078</td>
<td>1.110</td>
<td>100.000</td>
</tr>
</tbody>
</table>

Rotated Component Matrix

<table>
<thead>
<tr>
<th>Component</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>cost</td>
<td></td>
<td>.809</td>
<td></td>
<td></td>
<td>.556</td>
<td></td>
<td></td>
</tr>
<tr>
<td>size</td>
<td>.972</td>
<td></td>
<td></td>
<td>.556</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>alcohol</td>
<td>.952</td>
<td></td>
<td>.912</td>
<td></td>
<td>.556</td>
<td></td>
<td></td>
</tr>
<tr>
<td>reput</td>
<td></td>
<td>.977</td>
<td></td>
<td></td>
<td></td>
<td>.939</td>
<td></td>
</tr>
<tr>
<td>color</td>
<td></td>
<td>.977</td>
<td></td>
<td>.939</td>
<td></td>
<td>.939</td>
<td></td>
</tr>
<tr>
<td>aroma</td>
<td>.937</td>
<td></td>
<td>.939</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>taste</td>
<td></td>
<td></td>
<td>.939</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

*a* Rotation converged in 7 iterations.
Exploratory Factor Analysis

- Now we’ll try the EFA
  - Principal Axis factoring
  - Varimax rotation
- We’ll now be taking into account measurement error, so the communalities will be different
- If we just take the eigenvalues greater than 1 approach, we have 2 factors accounting for 80% of the total variance

<table>
<thead>
<tr>
<th>Trait</th>
<th>Cost</th>
<th>Size</th>
<th>Alcohol</th>
<th>Reputat</th>
<th>Color</th>
<th>Aroma</th>
<th>Taste</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communalities</td>
<td>.738</td>
<td>.912</td>
<td>.866</td>
<td>.499</td>
<td>.922</td>
<td>.857</td>
<td>.881</td>
</tr>
<tr>
<td>Extraction</td>
<td>.745</td>
<td>.914</td>
<td>.866</td>
<td>.385</td>
<td>.892</td>
<td>.896</td>
<td>.902</td>
</tr>
</tbody>
</table>

**Communalities**

<table>
<thead>
<tr>
<th>Factor</th>
<th>Initial Eigenvalues</th>
<th>Extraction Sums of Squared Loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>% of Variance</td>
<td>Cumulative %</td>
</tr>
<tr>
<td>Total</td>
<td>% of Variance</td>
<td>Cumulative %</td>
</tr>
<tr>
<td>1</td>
<td>3.313</td>
<td>47.327</td>
</tr>
<tr>
<td>2</td>
<td>2.616</td>
<td>37.369</td>
</tr>
<tr>
<td>3</td>
<td>.575</td>
<td>8.209</td>
</tr>
<tr>
<td>4</td>
<td>.240</td>
<td>3.427</td>
</tr>
<tr>
<td>5</td>
<td>.134</td>
<td>1.921</td>
</tr>
<tr>
<td>6</td>
<td>.085</td>
<td>1.221</td>
</tr>
<tr>
<td>7</td>
<td>.037</td>
<td>5.27</td>
</tr>
</tbody>
</table>

**Total Variance Explained**

<table>
<thead>
<tr>
<th>Factor</th>
<th>Initial Eigenvalues</th>
<th>Extraction Sums of Squared Loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>% of Variance</td>
<td>Cumulative %</td>
</tr>
<tr>
<td>Total</td>
<td>% of Variance</td>
<td>Cumulative %</td>
</tr>
<tr>
<td>1</td>
<td>3.123</td>
<td>47.327</td>
</tr>
<tr>
<td>2</td>
<td>2.478</td>
<td>35.396</td>
</tr>
<tr>
<td>3</td>
<td>2.478</td>
<td>35.396</td>
</tr>
<tr>
<td>4</td>
<td>2.478</td>
<td>35.396</td>
</tr>
<tr>
<td>5</td>
<td>2.478</td>
<td>35.396</td>
</tr>
<tr>
<td>6</td>
<td>2.478</td>
<td>35.396</td>
</tr>
<tr>
<td>7</td>
<td>2.478</td>
<td>35.396</td>
</tr>
</tbody>
</table>

Extraction Method: Principal Axis Factoring.
EFA

- Here are our initial structure coefficients before rotation
- Similar to before, not so well interpreted
- How about the rotated solution?

<table>
<thead>
<tr>
<th>Factor Matrixa</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>cost</td>
</tr>
<tr>
<td>size</td>
</tr>
<tr>
<td>alcohol</td>
</tr>
<tr>
<td>reputat</td>
</tr>
<tr>
<td>color</td>
</tr>
<tr>
<td>aroma</td>
</tr>
<tr>
<td>taste</td>
</tr>
</tbody>
</table>

Extraction Method: Principal Axis Factoring.

a. 2 factors extracted. 7 iterations required.
EFA

- Much better once again
- But note now we have the reputation variable loading on both, and negatively
- As this might be difficult to incorporate into our interpretation we may just stick to those that are loading highly\(^1\)
- However, this is a good example of how you may end up with different results whether you do PCA or EFA
EFA vs. PCA

- Again, the reason to use other methods of EFA rather than PCA is to take into account measurement error.
- In psych, this would be the route one would typically want to take:
  - Because of the lack of measurement error, physical sciences typically do PCA.
- However, in many cases the interpretation will not change for the most part, more so as more variables are involved:
  - The communalities make up less of the total values in the correlation matrix as we add variables:
    - Ex. 5 variables 10 correlations, 5 communalities
    - 10 variables 45 corr 10 communalities
- Some of the other EFA methods will not be viable with some datasets.
- Gist: in many situations you’ll be fine with either, but perhaps you should have an initial preference for other methods besides PCA.