

Updated 3/16/16



About Factor Analysis

- Telmar's Factor Analysis summarizes the correlation in a set of data and converts this into interpretable patterns.
- It simplifies the structure of a set of data.
- It reduces a "large" number of correlated variables (rows) e.g. lifestyle questions to a smaller set of uncorrelated variables or themes called factors. It does this by grouping survey questions (rows or variables) together (based on their correlation).
- It is used to understand a single target audience. Factor Analysis compares the row (e.g. 350 + lifestyle) questions for one target, unlike correspondence which compare and group similar brands (in columns) based on the row (e.g. 350 + lifestyle) questions.
- Unlike Cluster Analysis which groups respondents together, Factor Analysis groups questions (rows) which are answered in a similar way.

Features

- New direct launch from SurveyTime or TNT+ 5.0 application
- Easily interpret sorted results displayed within Factor Analysis module
- Export results to an Excel spreadsheet
- Export of factors as people back to crosstab as audiences. Carry over results to other Telmar applications such as SurveyTime, TNT+ 5.0, and Cluster.

What can it be used for?

- It is an excellent aid for questionnaire design, as it will identify questions answered in a similar way (due to their high correlation with each other)
- Understand media and their relationship with content e.g. word of mouth, favorite TV programs, topic of interest in magazines, etc.
- As a part of the media planning process to identify appropriate media for lifestyle driven target audiences
- Factor Analysis can be used as a selection procedure for running a Cluster Analysis



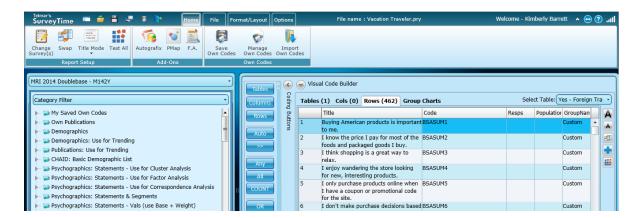
Steps to building a Factor Analysis

Preparing the Target, Variables (rows) and Launching Factor Analysis in SurveyTime

- Select a survey (e.g. MRI 2015 Doublebase)
- Input the target audience to be analyzed in the Tables (not in Columns)
 - In this example Foreign Travel (last 3 years) or Domestic Travel (last 12 month)
 Personally taken.



- Input the variables or questions that you are using to explain the target audience for analysis in the **Rows**
 - In this example 462 psychographic statements from (Use for Crosstab) section were used
 - Variables can come from any section of the codebook
- Caution: Do not use calculated values such as mean or median in Tables or Rows



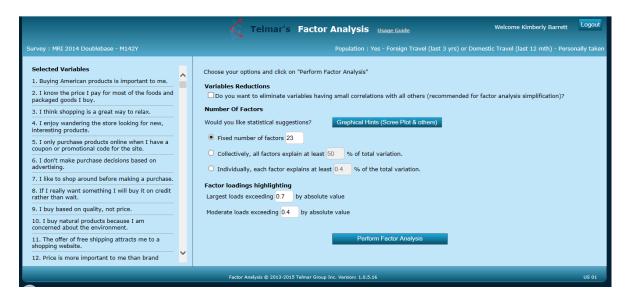


• Click "F.A.", to start Factor Analysis, from the Home tab in the Add-Ons section

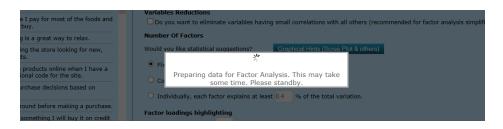


Selecting Factor Analysis Options

- Web browser will open a <u>new</u> window, displaying the Factor Analysis Options
- The target audience will display as the **Population** versus **Tables**
 - In this example Foreign Travel (last 3 years) or Domestic Travel (last 12 month)
 Personally taken
- The Selected Variables section appear in the left-hand column
 - In this example 462 psychographic statements placed in Rows
- The system will generate a default Number of Factors, or the user can adjust this number if so desired
 - o In this example the system default choose 23 factors
- After selecting Factor Analysis options, then click



• The next message to appear will be the processing message. Please do not close this window.





Selecting Factor Analysis Options (Advanced Options)

<u>Variables Reduction</u>: Is an option that allows the end user an option to eliminate variables <u>before</u> running the Factor Analysis. Checking the box opens the option to select/change the typical default absolute value of "0.1". In this case, the Variable Reduction is unchecked (default setting)

Helpful hint: If one variable has low correlations with all others - it should be eliminated from the FA solution. Low correlated variables create unnecessary noise in the data (e.g. may inflate the necessary number of factors, etc.)

• Number of Factors:

Fixed Number of Factors solution: Recommended if one has an idea about the potential data structure (e.g., they expect to see three distinct behavioral patterns), or if a user knows the "limit" of the number of the factors.

OR

Collectively, all Factors explain at least xx% of Total Variation solution: The solution is preferred if one has a minimum expectation of total variance to be explained. One option is to try between 40 and 60% and see how interpretable the solution is.

OR

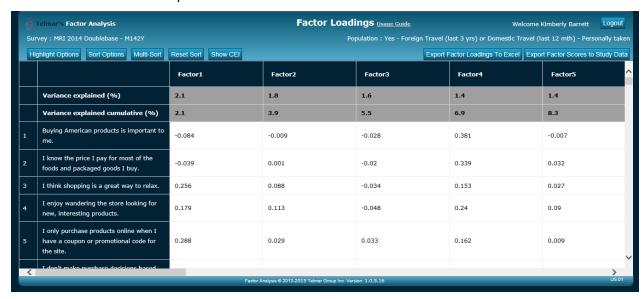
Individually, each factor explains at least xx.x% of the Total Variation solution: A solution for those who wants all factors to be "significant." One rule of thumb: the value should be around 2 - 2.5 times higher than one divided by number of variables (for instance, 2*(1/10) = 20% for 10 variables.)

<u>Factor loadings highlighting</u>: Highlighting is made for simplification of interpretation. The recommended thresholds follow the logic: 0.7 (which squared gives about 50% variance of the variable explained by the factor); 0.4 squared - (which gives about 10%.) One could follow their own logic and change the low and high values.



Reviewing Factor Analysis Output and Sorting Results

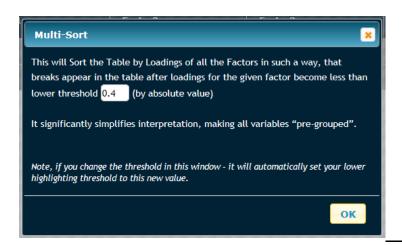
A <u>new</u> Web browser window will open with a **Factor Loadings** matrix. The order of variables (rows) is as carried over from SurveyTime (as seen in the left-hand column). At this point, the user has a few options to customize the view of the matrix to facilitate the interpretation of the results.



• Click on Multi-Sort to aid interpretation of the factors



• The **Multi-Sort** window opens to display the default lower threshold. To select the default value of 0.4, click **OK**.





Reviewing Factor Analysis Results

- At this stage, a report like the one below will appear with Highlights and each column will be sorted.
 - The user should try to identify what the variables have in common (for each factor).
- To view tops scores for each factor, click the <u>header</u> of each factor column
- Having viewed each factor, the user should see commons patterns
 - In this example the tops scores for the Factor 2 appear to be logically consistent with each other as they are about the level of usefulness of Advertising across multiple media platforms.



- Renaming Factors: If the factor is clear, the user can give it a descriptive name based on what the questions have in common. Double-click the factor title (e.g. Factor1) to rename the factor. If renamed, these will appear in all Exports of data.
- If the factors are NOT clear, the user can change the Factor Selection Options and rerun the report.
 - Click Logout to exit Factor Analysis application. From SurveyTime, click "F.A.", to re-run Factor Analysis, from the Home tab in the Add-Ons section.

		I am influence by what's most popular or hot	Advertising provides useful information and is amusing	Advertsing is not credible and is repeated too often	Brand Loyalist - I purchase products I trust	I enjoy technology and electronic products
	Variance explained (%)	2.1	1.8	1.6	1.4	1.4
	Variance explained cumulative (%)	2.1	3.9	5.5	6.9	8.3
41	When I find a brand I like, I stick to it.	-0.065	0.004	-0.014	0.477	0.018
14	It's important to me that salespeople be knowledgeable about the products they sell.	-0.083	-0.013	-0.051	0.447	0.021
51	I prefer a store that has a large selection of familiar brands.	0.007	0.051	-0.028	0.437	0.034
47	The service of the personnel at a store is an important part of my decision to shop there.	0.021	0.04	-0.029	0.42	0.008



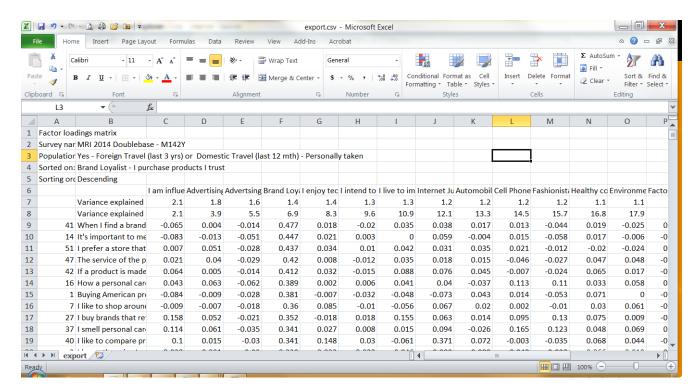
Factor Analysis - Export Factor Loading to Excel

Export Factor Loadings To Excel

- Next, click **Export Factor Loadings To Excel** to format results in Excel into relevant groupings.
- An Excel message will appear at the bottom of web browser, Click Open



The .csv file open and it looks like this

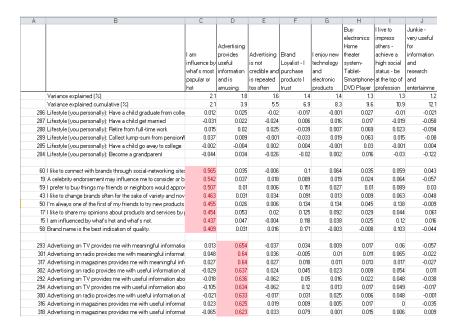


- Use Conditional Formatting to highlight values like shown in Factor Analysis program
 - o Highlight all numbers in the table excluding the variance explained % rows
 - Select Conditional Formatting, then Highlight Cell Rules, then Greater Than
 - Enter 0.4 (threshold) in the Conditional Formatting window, click OK





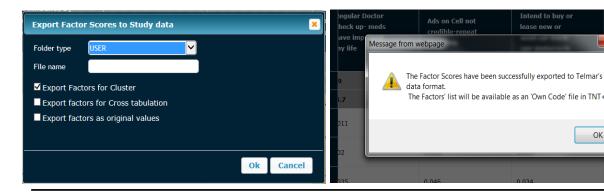
- At this stage, the report will be colored coded. Additional report formats can include:
 - Factor title row set to **Wrap Text** for compacting the title
 - Expand variable column (B) to read the full questions of each row
 - If user did not elect to rename Factors in software, it is recommended to rename factors and save in Excel at this time.
- Save formatted file in latest version of Excel to your personal computer for later



Factor Analysis - Export Factor Scores to Study Data

Export Factor Scores to Study Data

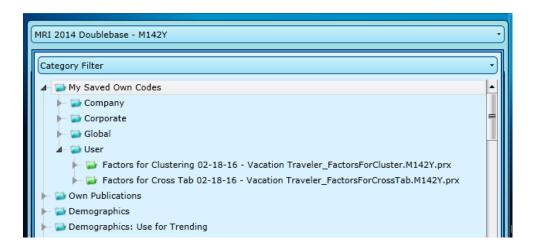
- Next, click Export Factor Scores to Study Data
 - The system provides two export options for further analysis (check both boxes), Export Factors for Cluster and Export Factors for Cross tabulation
 - Select the eTelmar cloud drive to save the exported scores under Folder type User, Company or Corporate clouds
 - Enter File name, then click **OK**



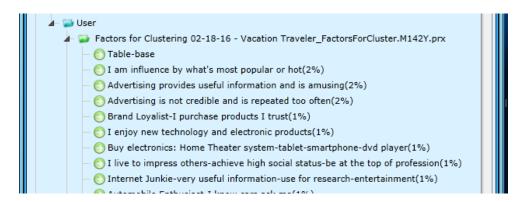
OK



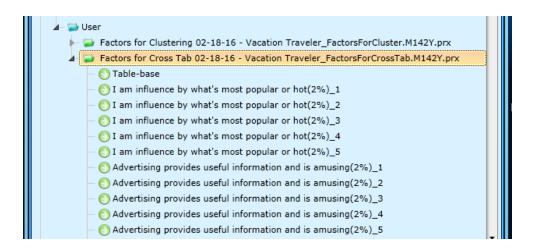
The exported factors will appear under My Saved Own Codes.



Factors for Clustering allow the results to be used in Telmar's Cluster Analysis
program and these data should <u>only</u> be used for creating a cluster solution as
variables for clustering.



Factors for Cross Tab allow the factor scores to be analyzed in SurveyTime





How to create a Cross tab report using Factor Analysis Scores

- When factors are created in the Factor Analysis program, their values can be assigned to each respondent and exported to SurveyTime to make the analysis convenient
- The program calculates values for each factor for each respondent, called factor scores.
- The scores variables are made for each factor and they are split into 5 groups
- The group labelled _5 has the highest positive score for that particular factor and the label _1 is the lowest.

```
I am influence by what's most popular or hot(2%)_1
I am influence by what's most popular or hot(2%)_2
I am influence by what's most popular or hot(2%)_3
I am influence by what's most popular or hot(2%)_4
I am influence by what's most popular or hot(2%)_5
```

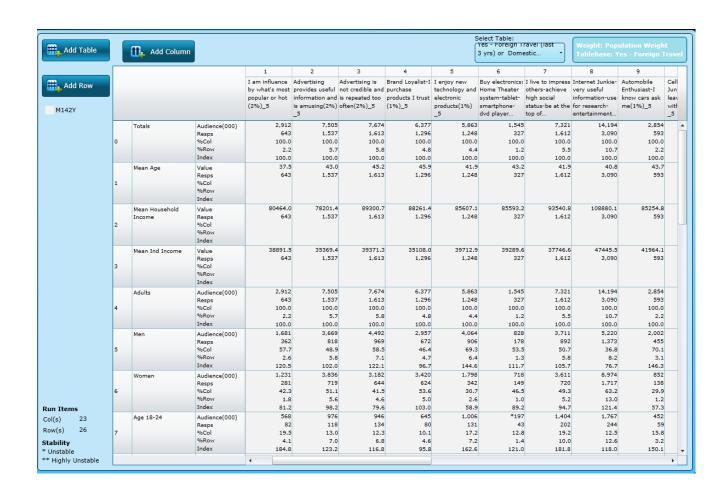
• When creating a cross tab, the user can use the scores separately or merge them in any combination (e.g. combine 4 and 5 to see the positive scores).

Steps to build a Cross tab report using Factor Analysis Scores

- Select a Survey (e.g. MRI 2015 Doublebase)
- Go to Factor for Cross Tab in My Saved Own Codes, then input your Table—base to be analyzed in the Tables
 - In this example The Vacation Traveler: Table Base is Foreign Travel (last 3 years) or Domestic Travel (last 12 month) Personally taken.

- In the **Columns** add the highest factor scores 5 of each factor
- In the **Rows** the user can have any items in the codebook, choose those that help interpret the columns (the factors)
 - o In this example Demographics were selected





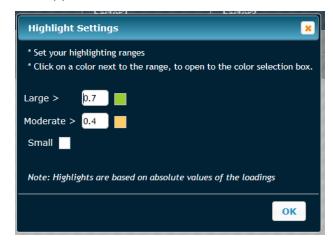


Sorting Results (Additional Options)

• Each Factor (i.e. each column of results) may be sorted by clicking on the right-hand side of each Factor cell. Note, the default sort order is Absolute value.



- Highlighting and Sorting:
 - Highlight Options allows users to change the default values for highlighting in the matrix for Large and Moderate Loadings. Selecting "Small" will cause all other remaining values to be highlighted in the preferred color when clicked as a color palette will appear.



- Sorting Options allow users to change the sorting to Original sort order (or non-Absolute value.) Sorting by absolute values (most often used) puts highly correlated variables first, thus allowing factor interpretation regardless of the sign of the correlations. Sorting by original values separates positive and negative correlations.
- Multi Sort Options allow users to sort the reports an aid in interpretation of the factors. Displays the values highlighted (green or yellow) of each factor.
- o **Reset Sort** removes any type of sorting and reverts matrix to the original order.



- Show CEI (Correlations Exploratory Indicators) attempts to explain why certain correlations between the factor and variable have appeared. If one clicks this button, the clickable CEI icons appear for each factor.
 - CEI signals appear only against the significant loadings as highlighted in yellow (moderate >4) or green (high >7).
 - Dark green cells mean "business as usual", i.e. correlations are high, because respondents answered in more or less similar way, like for all variables of the Factor1 (i.e. all the respondents gave similar answers in both "Disagree" and "Agree" on a 5 grade scale).



Red cells means high correlations between these variables and the factor are most likely to be explained by only part of the respondent pool, like Factor23 (Collect lump sum from pension/IRA/401K) – namely, by those who have very large values of the variable that drive a high correlation for the entire respondent pool (i.e. by respondents, who mainly answered "Agree" on a 5 grade scale).

