telmar
the FUTURE of
Media Planning

Factor Analysis

Quick User Guide

Updated 3/16/16

Factor Analysis
Quick User Guide

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

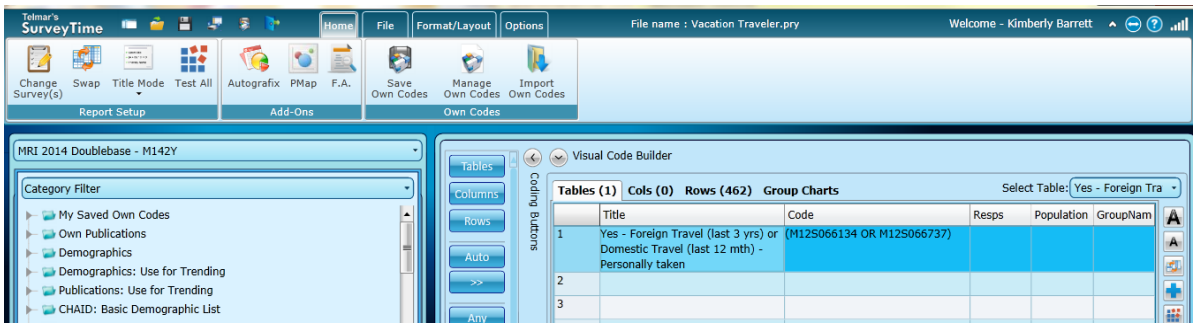
Factor Analysis
Quick User Guide

Steps to building a Factor Analysis

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

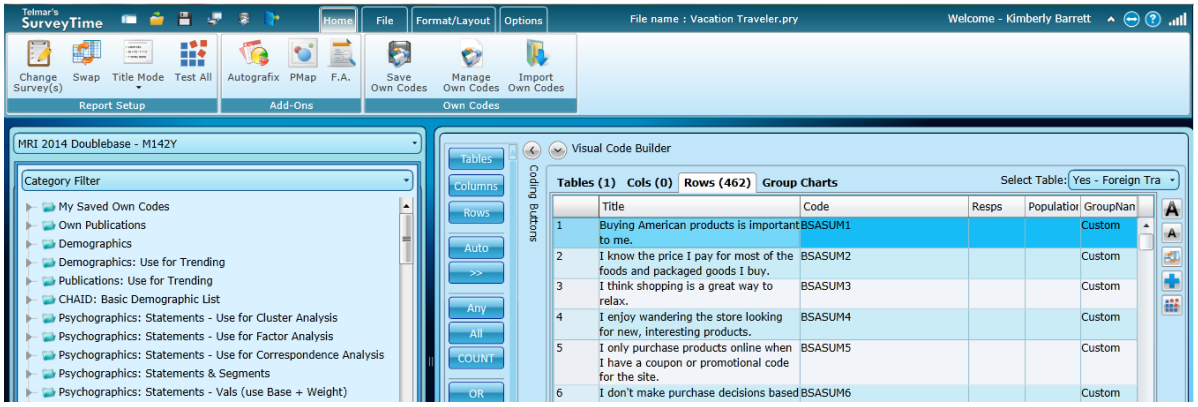
- Select a survey (e.g. MRI 2015 Doublebase)

- 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.



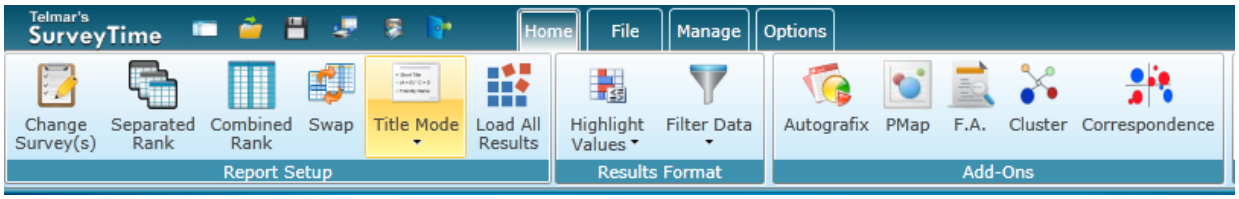
- Input the variables or questions that you are using to explain the target audience for analysis in the **Rows**
  - In this example - 462 psychographic statements from (Use for Crosstab) section were used
  - Variables can come from any section of the codebook

- **Caution:** Do not use calculated values such as <u>mean</u> or <u>median</u> in **Tables** or **Rows**



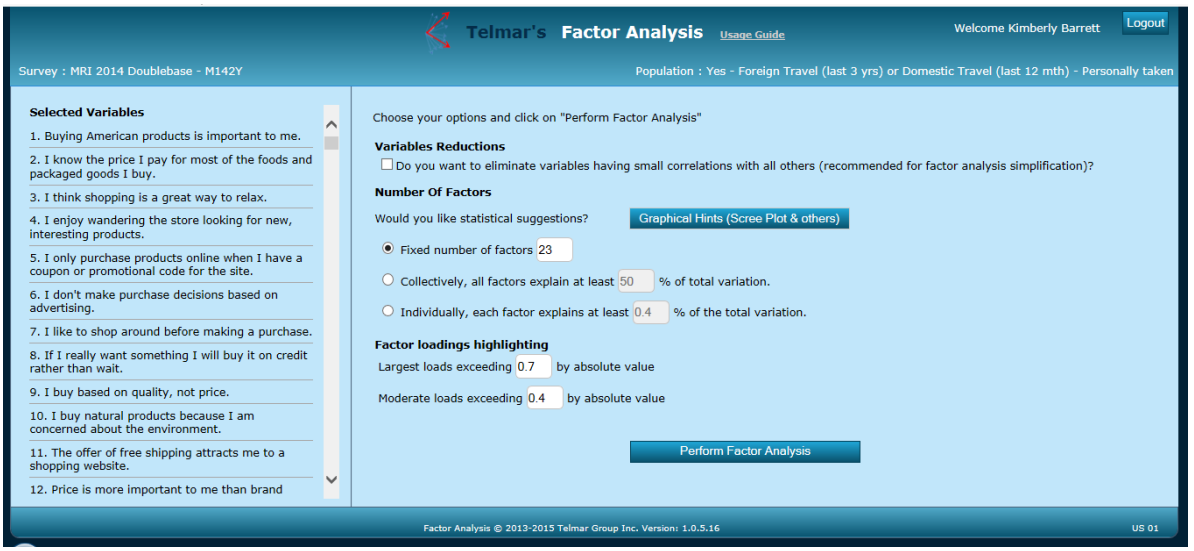
Factor Analysis
Quick User Guide

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

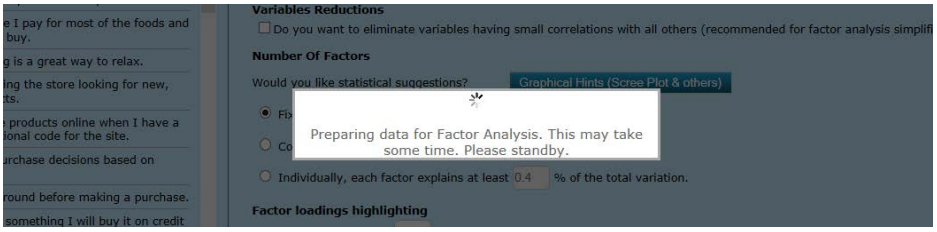


Selecting Factor Analysis Options

- 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**, <u>or</u> the user can adjust this number if so desired
  - 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)

- **Variables Reduction**: Is an option that allows the end user an option to eliminate variables before 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

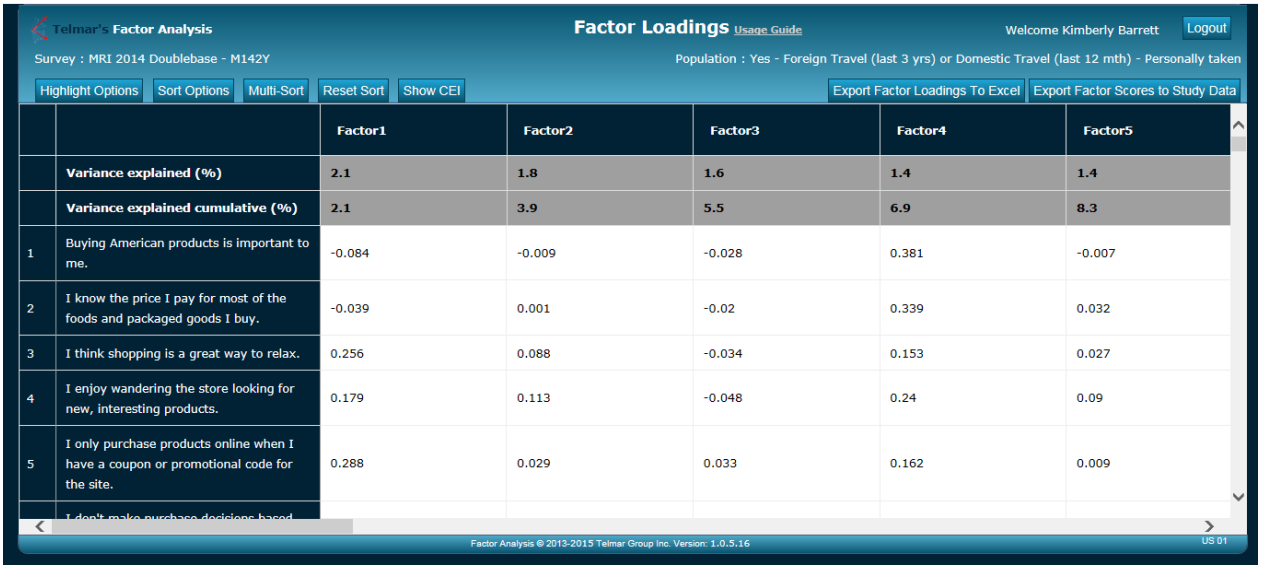
  Individually, each factor explains at least xx.x% of the Total Variation solution: A solution for those who wants all factors to be "significant." One rule of thumb: the value should be around 2 - 2.5 times higher than one divided by number of variables (for instance, 2*(1/10) = 20% for 10 variables.)

- **Factor loadings highlighting:** 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.

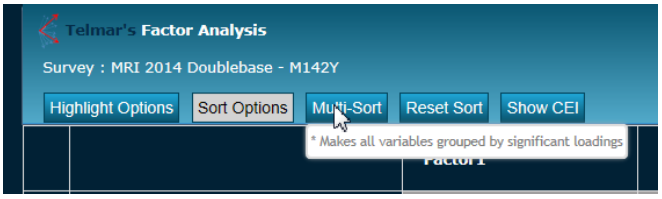
Factor Analysis
Quick User Guide

Reviewing Factor Analysis Output and Sorting Results

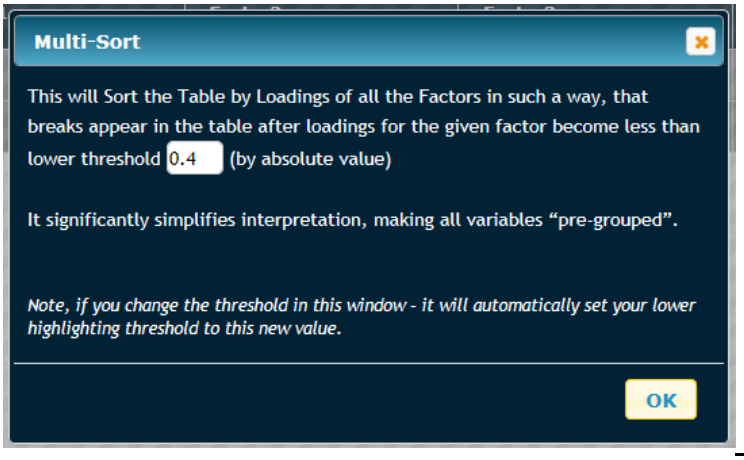
- A new Web browser window will open with a **Factor Loadings** matrix. The order of variables (rows) is as carried over from SurveyTime (as seen in the left-hand column). At this point, the user has a few options to customize the view of the matrix to facilitate the interpretation of the results.



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



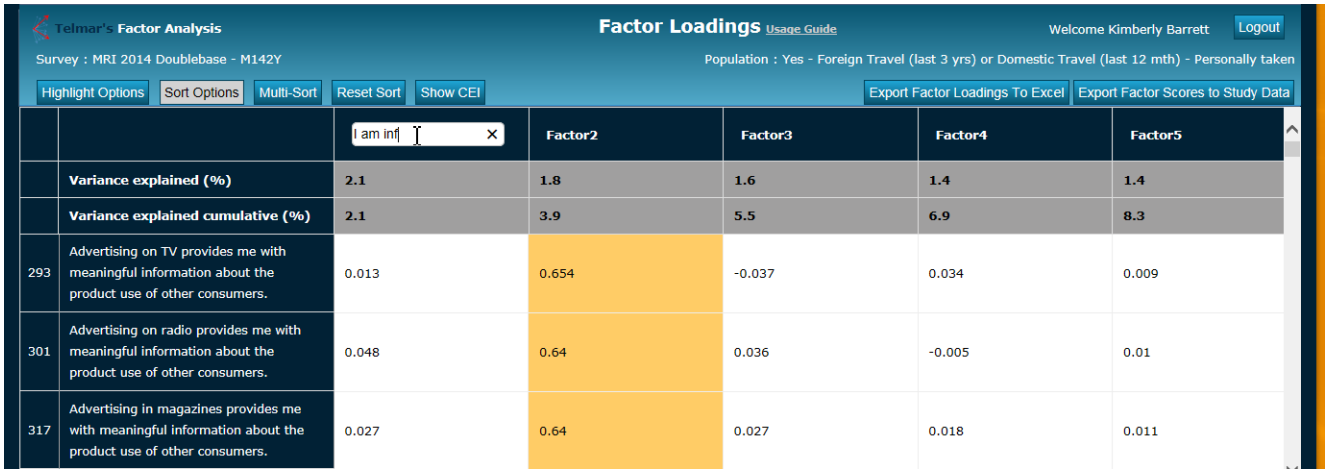
- The **Multi-Sort** window opens to display the default lower threshold. To select the default value of 0.4, click **OK.**



Factor Analysis
Quick User Guide

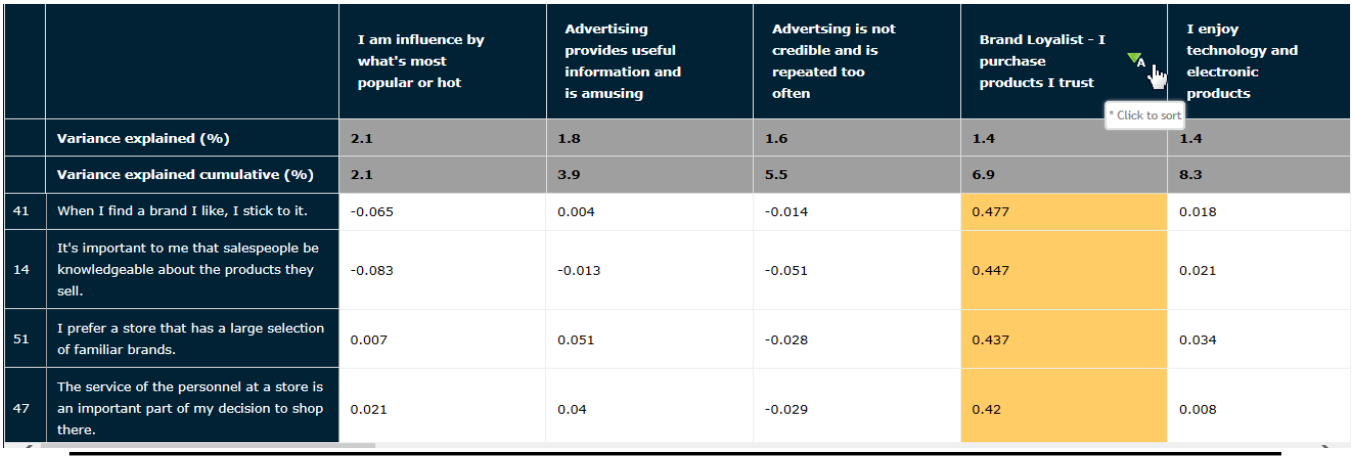
Reviewing Factor Analysis Results

- 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 header 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.*

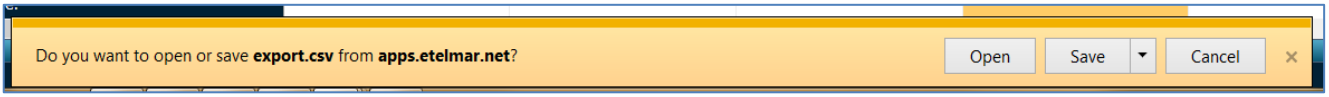
- If the factors are NOT clear, the user can change the **Factor Selection Options** and re-run the report.
  - Click **Logout** to exit Factor Analysis application. From SurveyTime, click **"F.A."**, to re-run Factor Analysis, from the **Home** tab in the **Add-Ons** section.



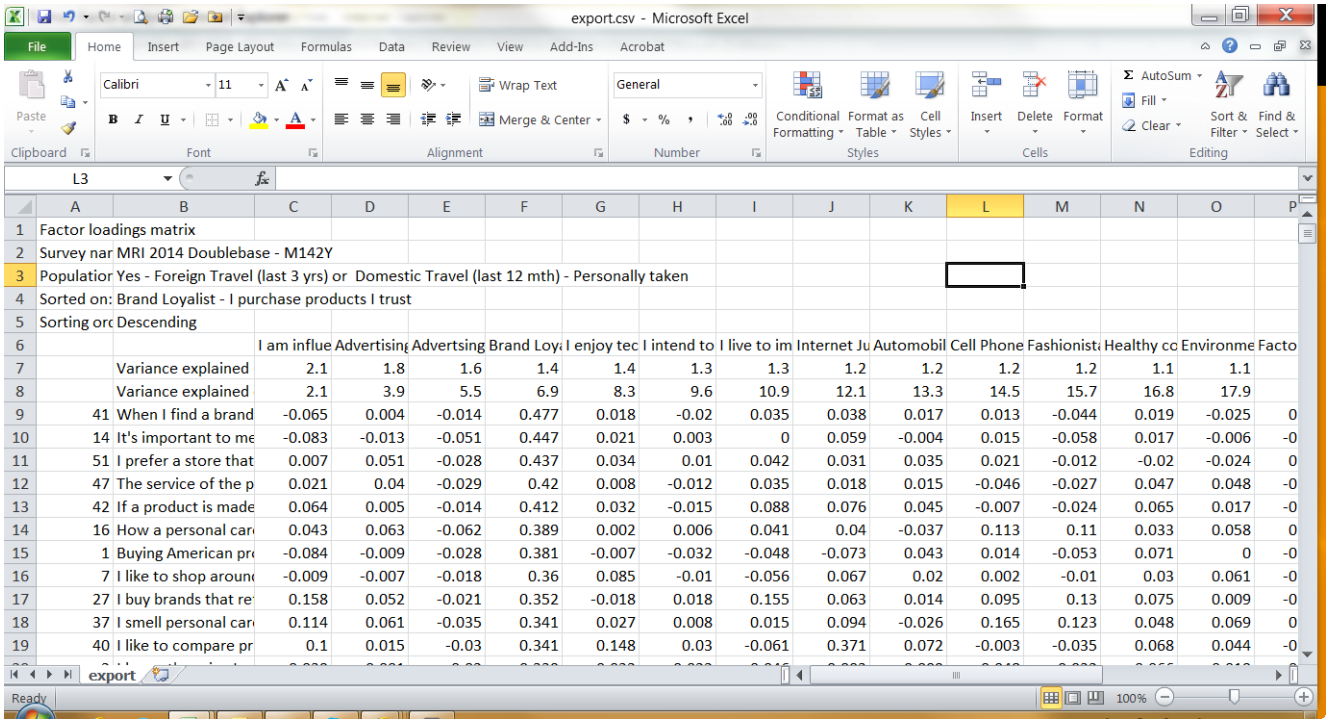
Factor Analysis
Quick User Guide

Factor Analysis - Export Factor Loading to Excel

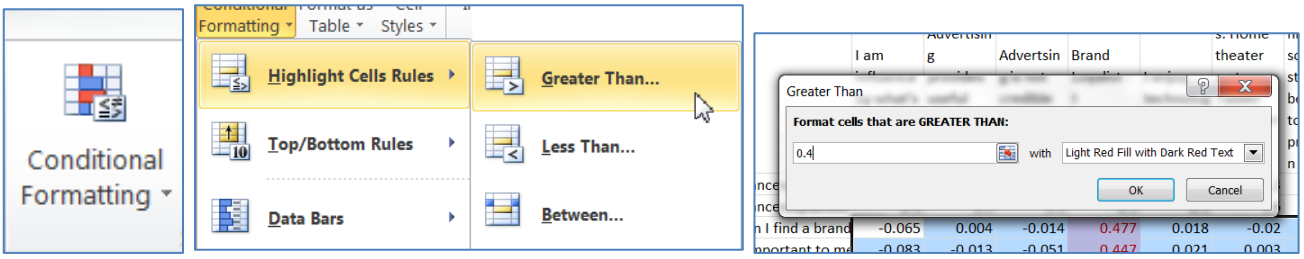
- Next, click **Export Factor Loadings To Excel** to format results in Excel into relevant groupings.
- An Excel message will appear at the bottom of web browser, Click **Open**



- The .csv file open and it looks like this



- Use **Conditional Formatting** to highlight values like shown in Factor Analysis program
  - 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**



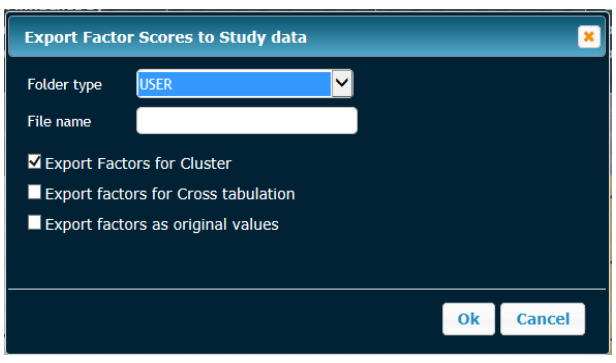
- At this stage, the report will be colored coded. Additional report formats can include:
  - Factor title row – set to **Wrap Text** for compacting the title
  - Expand variable column (B) to read the full questions of each row
  - If user did not elect to rename Factors in software, it is *recommended*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

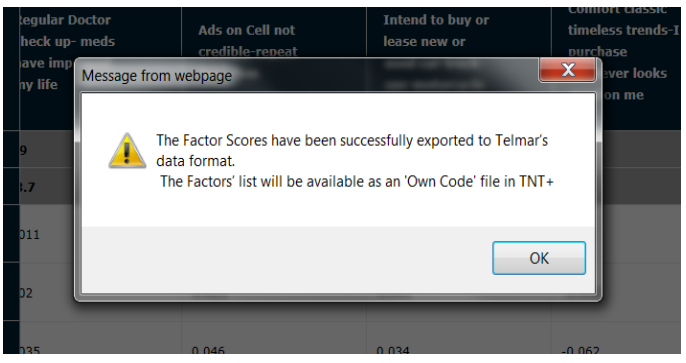
- Save formatted file in latest version of Excel to your personal computer for later review.

A	B	C	D	E	F	G	H	I	J
		I am influence by what's most popular or hot	Advertising provides useful information and is amusing	Advertising is not credible and is repeated too often	Brand Loyalist - I purchase products I trust	I enjoy new technology and electronic products	Buy electronics: Home theater system-Tablet-Smartphone-DVD Player	I live to impress others - achieve a high social status - be at the top of profession	Junkie - very useful for information and research and entertainme
	Variance explained (%)	2.1	1.8	1.6	1.4	1.4	1.3	1.3	1.2
	Variance explained cumulative (%)	2.1	3.9	5.5	6.9	8.3	9.6	10.9	12.1
286	Lifestyle (you personally): Have a child graduate from colle	0.012	0.025	-0.02	-0.017	-0.001	0.027	-0.01	-0.021
287	Lifestyle (you personally): Have a child get married	-0.031	0.022	-0.024	0.006	0.016	0.017	-0.019	-0.058
288	Lifestyle (you personally): Retire from full-time work	0.015	0.02	0.025	-0.039	0.007	0.069	0.023	-0.094
289	Lifestyle (you personally): Collect lump-sum from pensionfi	0.037	0.009	-0.001	-0.033	0.019	0.063	0.015	-0.08
285	Lifestyle (you personally): Have a child go away to college	-0.002	-0.004	0.002	0.004	-0.001	0.03	-0.001	0.004
284	Lifestyle (you personally): Become a grandparent	-0.044	0.034	-0.026	-0.02	0.002	0.016	-0.03	-0.122
60	I like to connect with brands through social-networking sites	0.565	0.035	-0.006	0.1	0.064	0.035	0.059	0.043
19	A celebrity endorsement may influence me to consider or b	0.542	0.037	0.018	0.089	0.019	0.024	0.064	-0.057
59	I prefer to buy things my friends or neighbors would approv	0.507	0.01	0.006	0.151	0.027	0.01	0.089	0.03
43	I like to change brands often for the sake of variety and nov	0.463	0.031	0.034	0.091	0.013	0.009	0.063	-0.048
50	I'm always one of the first of my friends to try new products	0.455	0.026	0.006	0.134	0.134	0.045	0.138	-0.009
17	I like to share my opinions about products and services by p	0.454	0.053	0.02	0.125	0.092	0.029	0.044	0.061
15	I am influenced by what's hot and what's not.	0.437	0.047	-0.004	0.118	0.038	0.025	0.12	0.016
58	Brand name is the best indication of quality.	0.409	0.031	0.016	0.171	-0.003	-0.008	0.103	-0.044
293	Advertising on TV provides me with meaningful informatio	0.013	0.654	-0.037	0.034	0.009	0.017	0.06	-0.057
301	Advertising on radio provides me with meaningful informat	0.048	0.64	0.036	-0.005	0.01	0.011	0.065	-0.022
317	Advertising in magazines provides me with meaningful infi	0.027	0.64	0.027	0.018	0.011	0.013	0.017	-0.027
302	Advertising on radio provides me with useful information at	-0.029	0.637	0.024	0.045	0.023	0.009	0.054	0.011
292	Advertising on TV provides me with useful information abo	-0.018	0.636	-0.062	0.05	0.016	0.022	0.048	-0.038
294	Advertising on TV provides me with useful information abo	-0.105	0.634	-0.062	0.12	0.013	0.017	0.049	-0.017
300	Advertising on radio provides me with useful information at	-0.021	0.633	-0.017	0.031	0.025	0.006	0.048	-0.001
316	Advertising in magazines provides me with useful informat	0.023	0.625	0.019	0.009	0.005	0.017	0	-0.035
318	Advertising in magazines provides me with useful informat	-0.065	0.623	0.033	0.079	0.001	0.015	0.006	0.009

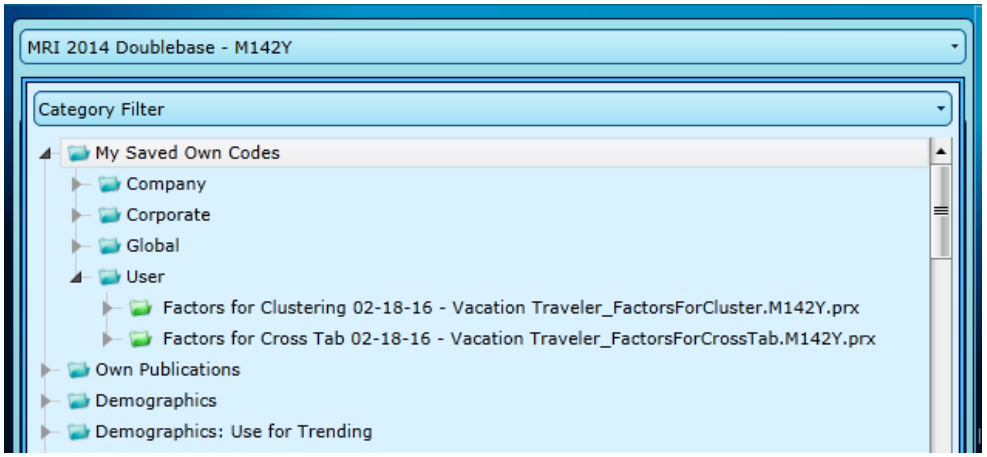
Factor Analysis - Export Factor Scores to Study Data

- 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**

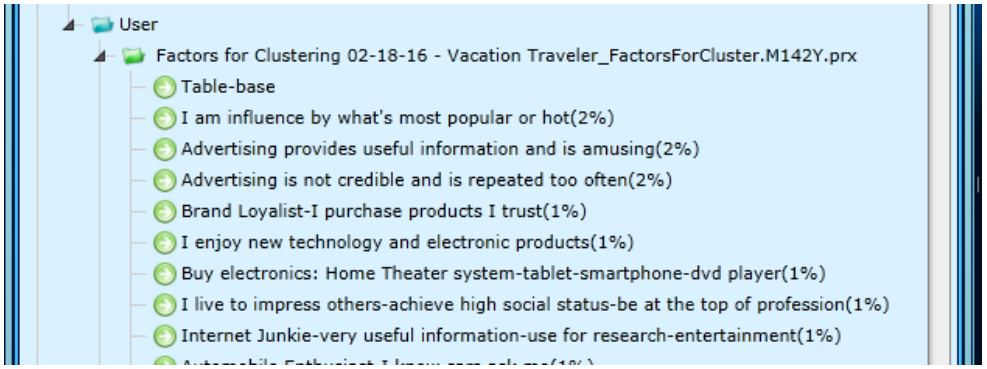




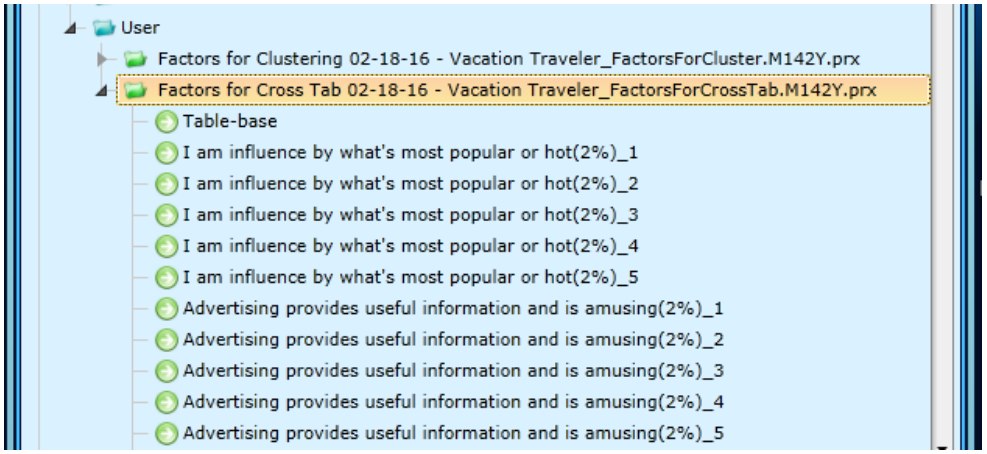
- The exported factors will appear under **My Saved Own Codes.**



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



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

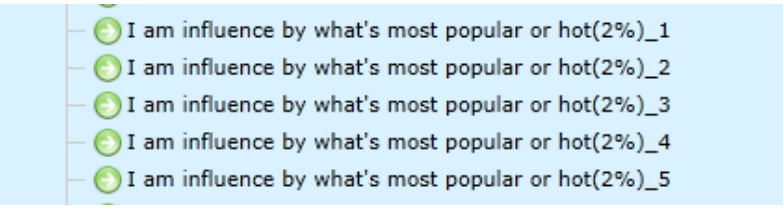


How to create a Cross tab report using Factor Analysis Scores

- 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.

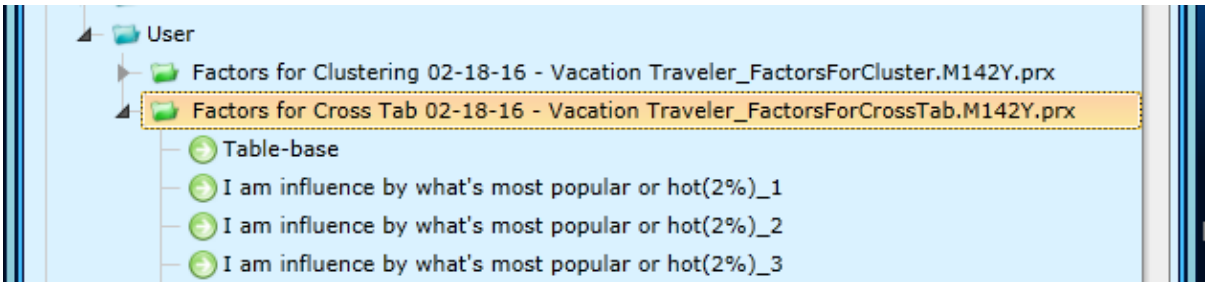


- When creating a cross tab, the user can use the scores separately or merge them in any combination (e.g. combine _4 and _5 to see the positive scores).

Steps to build a Cross tab report using Factor Analysis Scores

- Select a Survey (e.g. MRI 2015 Doublebase)

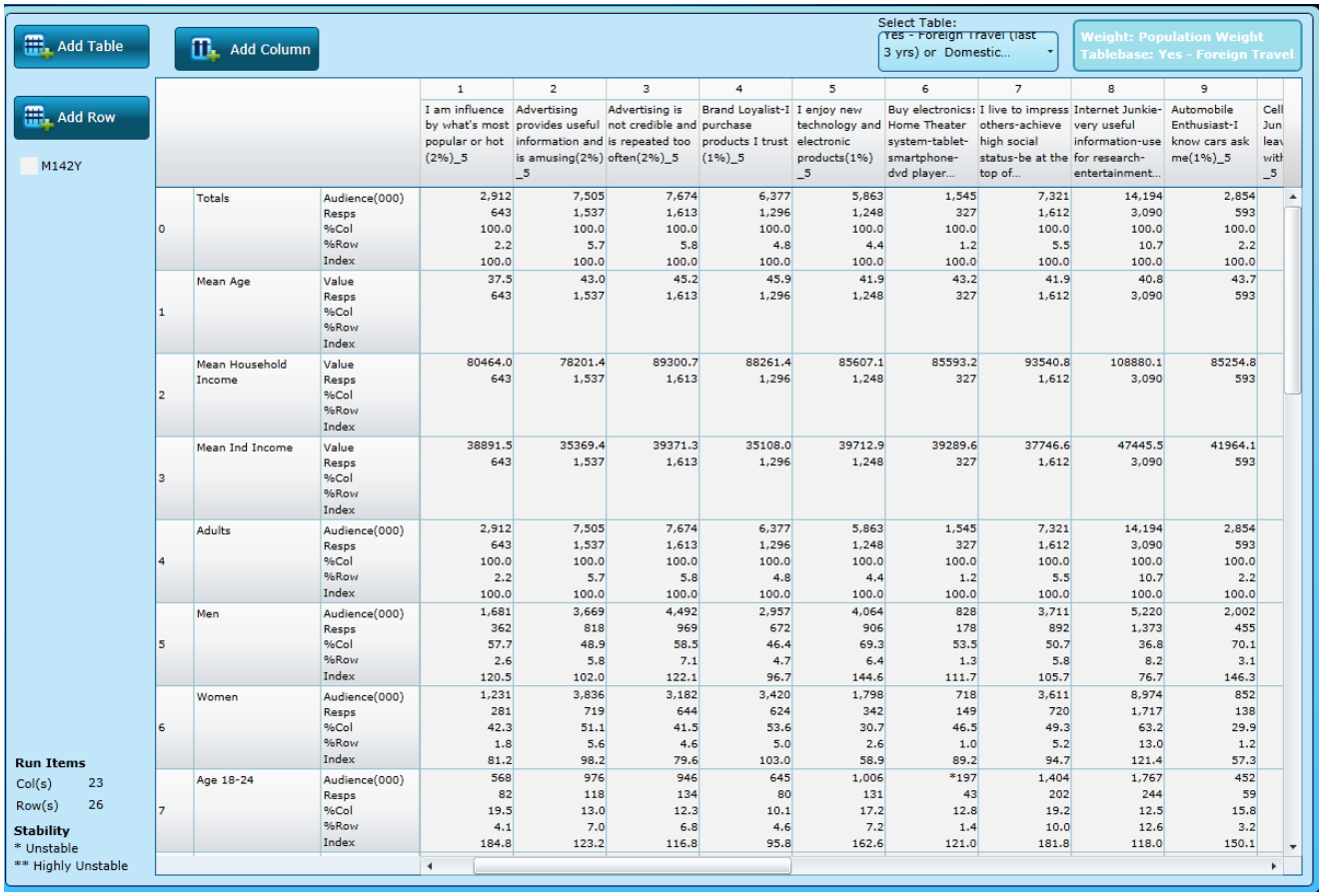
- Go to **Factor for Cross Tab** in **My Saved Own Codes**, then input your **Table–base** to be analyzed in the **Tables**
  - In this example – The *Vacation Traveler: Table – Base* is Foreign Travel (last 3 years) or Domestic Travel (last 12 month) – Personally taken.



- 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)
  - In this example – Demographics were selected

Factor Analysis
Quick User Guide



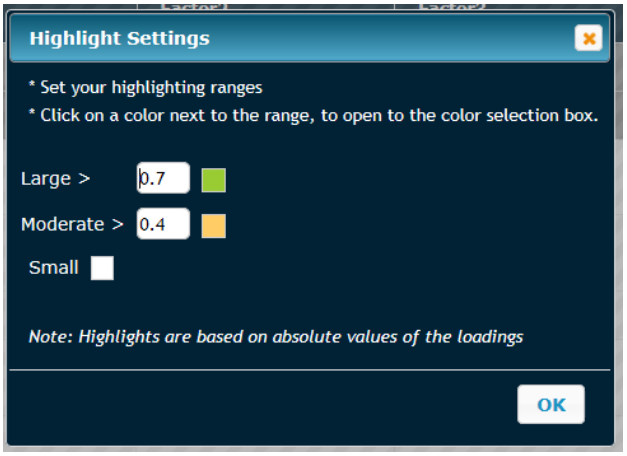
			1	2	3	4	5	6	7	8	9	
			I am influence by what's most popular or hot (2%)_5	Advertising provides useful information and is amusing(2%)_5	Advertising is not credible and is repeated too often(2%)_5	Brand Loyalist-I purchase products I trust (1%)_5	I enjoy new technology and electronic products(1%)_5	Buy electronics: Home Theater system-tablet-smartphone-dvd player...	I live to impress others-achieve high social status-be at the top of...	Internet Junkie-very useful information-use for research-entertainment...	Automobile Enthusiast-I know cars ask me(1%)_5	Cell Jun leav with _5
0	Totals	Audience(000)	2,912	7,505	7,674	6,377	5,863	1,545	7,321	14,194	2,854	
		Resps	643	1,537	1,613	1,296	1,248	327	1,612	3,090	593	
		%Col	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	
		%Row	2.2	5.7	5.8	4.8	4.4	1.2	5.5	10.7	2.2	
		Index	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	
1	Mean Age	Value	37.5	43.0	45.2	45.9	41.9	43.2	41.9	40.8	43.7	
		Resps	643	1,537	1,613	1,296	1,248	327	1,612	3,090	593	
		%Col										
		%Row										
		Index										
2	Mean Household Income	Value	80464.0	78201.4	89300.7	88261.4	85607.1	85593.2	93540.8	108880.1	85254.8	
		Resps	643	1,537	1,613	1,296	1,248	327	1,612	3,090	593	
		%Col										
		%Row										
		Index										
3	Mean Ind Income	Value	38891.5	35369.4	39371.3	35108.0	39712.9	39289.6	37746.6	47445.5	41964.1	
		Resps	643	1,537	1,613	1,296	1,248	327	1,612	3,090	593	
		%Col										
		%Row										
		Index										
4	Adults	Audience(000)	2,912	7,505	7,674	6,377	5,863	1,545	7,321	14,194	2,854	
		Resps	643	1,537	1,613	1,296	1,248	327	1,612	3,090	593	
		%Col	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	
		%Row	2.2	5.7	5.8	4.8	4.4	1.2	5.5	10.7	2.2	
		Index	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	
5	Men	Audience(000)	1,681	3,669	4,492	2,957	4,064	828	3,711	5,220	2,002	
		Resps	362	818	969	672	906	178	892	1,373	455	
		%Col	57.7	48.9	58.5	46.4	69.3	53.5	50.7	36.8	70.1	
		%Row	2.6	5.8	7.1	4.7	6.4	1.3	5.8	8.2	3.1	
		Index	120.5	102.0	122.1	96.7	144.6	111.7	105.7	76.7	146.3	
6	Women	Audience(000)	1,231	3,836	3,182	3,420	1,798	718	3,611	8,974	852	
		Resps	281	719	644	624	342	149	720	1,717	138	
		%Col	42.3	51.1	41.5	53.6	30.7	46.5	49.3	63.2	29.9	
		%Row	1.8	5.6	4.6	5.0	2.6	1.0	5.2	13.0	1.2	
		Index	81.2	98.2	79.6	103.0	58.9	89.2	94.7	121.4	57.3	
7	Age 18-24	Audience(000)	568	976	946	645	1,006	*197	1,404	1,767	452	
		Resps	82	118	134	80	131	43	202	244	59	
		%Col	19.5	13.0	12.3	10.1	17.2	12.8	19.2	12.5	15.8	
		%Row	4.1	7.0	6.8	4.6	7.2	1.4	10.0	12.6	3.2	
		Index	184.8	123.2	116.8	95.8	162.6	121.0	181.8	118.0	150.1	

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.

  - **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).

		Factor1	CEI	Factor2	CEI	Factor3	CEI	Factor4	CEI
	Variance explained (%)	2.1		1.8		1.6		1.4	
	Variance explained cumulative (%)	2.1		3.9		5.5		6.9	
60	I like to connect with brands through social-networking sites.	0.565		0.035		-0.006		0.1	
19	A celebrity endorsement may influence me to consider or buy a product.	0.542		0.037		0.018		0.089	
59	I prefer to buy things my friends or neighbors would approve of.	0.507		0.01		0.006		0.151	
43	I like to change brands often for the sake of variety and novelty.	0.463		0.031		0.034		0.091	
50	I'm always one of the first of my friends to try new products or services.	0.455		0.026		0.006		0.134	
	I like to share my opinions about								

    - **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).

		Factor20	CEI	Factor21	CEI	Factor22	CEI	Factor23	CEI
	Variance explained (%)	0.8		0.8		0.8		0.7	
	Variance explained cumulative (%)	24.5		25.3		26.1		26.8	
	graduate from college								
287	Lifestyle (you personally): Have a child get married	0.023		0.011		-0.052		0.467	
288	Lifestyle (you personally): Retire from full-time work	0.046		0.034		-0.062		0.448	
289	Lifestyle (you personally): Collect lump-sum from pension/IRA/401k	0.027		0.04		-0.057		0.434	
285	Lifestyle (you personally): Have a child go away to college	-0.005		0.02		-0.028		0.417	
284	Lifestyle (you personally): Become a grandparent	0.002		0.007		-0.035		0.41	