

DA6823

Exercise #3

Name: Austin Thrash - FFK221

This third exercise is to give you practice at creating and extracting a single construct (abstract constructs) from multiple variables.

This means you are going to repeat the steps for the five point scales for the eight questions that form the two abstract constructs in the principal components analysis. **That is, when you are done and ready to do the PCA, there should be eight happy 5 point scale questions that represent the two abstract constructs you have chosen.**

You can copy and edit code from exercise 2 that created the variables for the market segmentation project and just take the new code for the eight (8) five point questions for the PCA and tack it onto the end of your exercise #2 code.

1. Run your PCA analysis on the eight or more variables that make up the abstract construct variables that you named in exercise #1 and have coded here in this exercise. Then answer the following questions:
2. Decide which extraction technique to use and tell me why

Since our dependent variables are interval and normal, based on the IDRE chart we are to perform factor analysis.

3. Decide which rotation method you are going to use and tell me why

We will use Kaiser-Varimax rotation as it eliminates the middle ground when attempting to identify more specific factors.

4. Run the PCA analysis and answer the additional questions below:
 - a. What is the criteria for determining that a factor was extracted?

If the eigenvalues are greater than 1 then it will be a factor that we should consider.

- b. How many factors were extracted – hopefully just two factors? Was it the number you expected? If not, you may have to replace a variable or two that was not working and may have caused the excess number of factors and rerun it. **You should end up with only two extracted factors that make the Kaiser eigenvalue cutoff point.** Cut and paste the eigenvalue table.

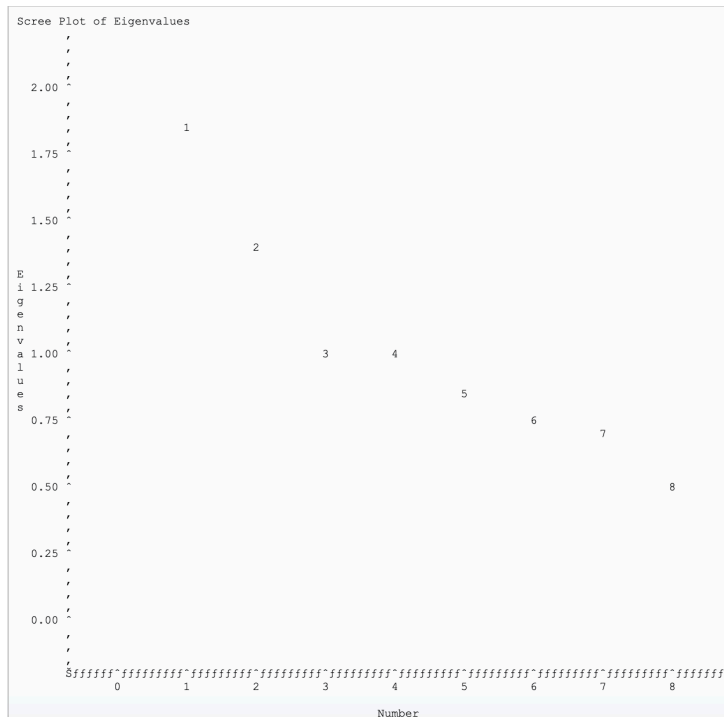
The number of factors extracted was two. This is the number of factors we need, however I did have to replace some variables in order to achieve only two extracted factors.

Eigenvalues of the Correlation Matrix: Total = 8 Average = 1				
	Eigenvalue	Difference	Proportion	Cumulative
1	1.89775347	0.22397932	0.2372	0.2372
2	1.67377415	0.69936123	0.2092	0.4464
3	0.97441292	0.17216911	0.1218	0.5682
4	0.80224381	0.01551496	0.1003	0.6685
5	0.78672885	0.04483683	0.0983	0.7669
6	0.74189202	0.11078159	0.0927	0.8596
7	0.63111044	0.13902609	0.0789	0.9385
8	0.49208435		0.0615	1.0000

c. What percentage of the variance was explained by the extracted factors?

The percentage of variance explained is 0.4464 or 44.64%. This value is not acceptable as we are looking to obtain a percentage over or near 0.5. Therefore one of our variables may be causing issues between distinguishing the two factors.

d. Make a not so pretty scree plot in SAS. How do you interpret it? Cut and paste it.



We can notice there seems to be a bit of a pile up at 3 and 4. The scree plot should similar to an exponential curve, so we should probably look at removing or changing these variables to help explain more variance in our factors.

- e. Find the correct Factor Pattern matrix and cut it and paste it. Circle or highlight which variables go with which factor. **Don't grab the wrong Factor Pattern matrix!!!**

Rotated Factor Pattern			
		Factor1	Factor2
careful_money	I'M CAREFUL WITH MY MONEY	-0.00866	0.81623
manage_money	I AM VERY GOOD AT MANAGING MONEY	0.03636	0.78385
debt	DON'T LIKE THE IDEA OF BEING IN DEBT	-0.02637	0.62699
hold_out	I HOLD OUT ON BUYING THINGS UNTIL SALE	0.07484	0.30897
best_brands	PRFR SHPPNG SPC STRS BC TND CRRY BST BR	0.65085	0.01102
fav_brand	I ALWAYS LOOK FOR FAVORITE BRANDS FIRST	0.65062	0.04749
toiletries	IN TOILETRIES, BRANDS BUY IS IMPORTANT	0.62378	0.07390
despite	DESPITE PRICE IMPT TO HAVE BRAND I LIKE	0.74212	-0.00072

The blue group goes with factor 2 and the orange group goes with factor 1. Its worth noting that by removing either the 'hold_out' or 'best_brands' variables drastically increases the amount of variance explained, so these variables may not be the best to use.

- f. Interpret the two factors extracted – that is, what does each of the two factors represent.

The first factor is Brand Loyalty and the second factor is Cost Consciousness.

- g. Turn in your report and the code that created the output.