

Data Practice with writeup

One important goal of this course is to provide you with instructions in theory and experiences in applying survey software and methods to analyzing sample survey data. This assignment requires you to practice a set of survey data skills. Specifically:

- 1) importing the data into SPSS;
- 2) preparing data for analysis through labeling variables, coding data, handling missing data, and screening data;
- 3) testing the reliability of survey items; and
- 4) exploring the measurement structure through exploratory factor analysis (EFA).

Your report should explain the procedures, steps, and judgement calls you take to get ready for data analysis. It should also include the results of your reliability and EFA tests and your interpretations and evaluations of these results. The best way to complete your report is to refer to the Method section of the journal articles you have reviewed in this course.

Your write up should be double-spaced, in 12-point Times New Roman font with 1” margins around.

Data file and codebook

The data files can be found in the google folder “DBA-Survey-2020S” in Subfolder titled “Data Practice” (see the link below). Both SPSS and Excel files are provided, with the name of “**GUAO_Student assignment**”. For the analysis practice, you can use either data file – ensure you select the correct files for this assignment.

The Codebook titled “Survey_Codebook_Practice_Students” is also provided in the good drive. You will need the Codebook to rename variables, update the variable labels and gain some understanding of the variables in the data file.

https://drive.google.com/drive/u/0/folders/1YyjZr1QRV7zAq9fu65wDr1yJv1gQ_yDj

Preparing data for analysis

You will focus on the constructs of *job alternatives*, *career satisfaction*, and *organizational commitment* in this assignment. In the data file,

Job alternatives (JA) is measured by three variables - *T1_Job_Offer_1*, *T1_Job_Offer_2*, and *T1_Job_Offer_3*. Please rename these three JA variables into three new variables as *JA_3*, *JA_4*, *JA_5* according to the codebook.

Career satisfaction (Car_Sat) is measured by five variables -namely, *T1_Car_Sat_1*, *T1_Car_Sat_3*, *T1_Car_Sat_4*, *T1_Car_Sat_5*, *T1_Car_Sat_6*.

Organizational commitment (Org_Commit) is measured by twelve variables, named as *T1_Org_Commit_1*, *T1_Org_Commit_2*,... and *T1_Org_Commit_12*. Moreover, please complete the labels for the variables *T1_Org_Commit_9*, *T1_Org_Commit_10*, *T1_Org_Commit_11* and *T1_Org_Commit_12*, according to the Codebook.

You should screen data and identify any “fishy” patterns. Describe how you “screen” the data including the missing responses. Describe how you handle the missing data and the rationale for your strategy.

Testing the reliability of survey items

Calculate the Cronbach’s Alphas for *job alternatives*, *career satisfaction*, and *organizational commitment* measures respectively. What patterns do you observe for each set of measures? Do you obtain satisfying cronbach’s coefficients? If no, what is the issue? How do you address it to ensure the reliability of the scale?

Exploratory Factor Analysis (EFA)

Conduct the EFA to better understand the measurement structure of *Organizational commitment* (Org_Commit), using the variables of *T1_Org_Commit_1*, *T1_Org_Commit_2*,... and *T1_Org_Commit_12*.

Your analysis and report should address the following:

- Provide a correlation matrix of all variables analyzed. What information can you gather from the correlation matrix?
- Report the measures of sampling adequacy and the appropriate Chi-square statistics. What do the test statistics tell you?
- Conduct a factor analysis using the principal components (PCA) method extracting the maximum number of factors. Explain the utility and meaning (in your own words) of the eigenvalues and factor loadings.
- Run a factor analysis using the maximum likelihood (ML) method using the eigenvalue criterion as an initial guide for the number of factors to be extracted. What are your initial conclusions?
- Discuss the two rotations (varimax and promax) in your factor models tested above. Which one gives you a better fit? Why?
- Briefly describe the differences and similarities between PCA and factor analysis. How do these two approaches provide different information in the analysis you have conducted?

- What is the final factor model do you select for this data? What are some of your final remarks/observations regarding your analysis?

In completing this assignment, you are expected to be specific and analytical in your responses. Simply producing tables or output is not sufficient and if not accompanied by any explanation or discussion it will not be given any credit.

Your typed assignment answers should include the relevant outputs of your computations (estimates, tables and/or charts). You may also append the full syntax you used to compute your answers.

Resources and tips

- Ruel et al., Chapter 9, 11, 12, 13
- Fabrigar et al., (1999)
- Revisit the assigned reading below and see how the authors report their reliability and EFA results
 - Edmondson (1999)
 - Mitchell et al. (2001)
 - Rybowski et al. (1999)
- The website below provides step by step explanations and helpful syntax for analyses

<https://stats.idre.ucla.edu/spss/seminars/efa-spss/>