

# Quiz #1

*Dustin Leatherman*

*April 22, 2019*

## 1

A researcher obtained five variables measuring job characteristics (feedback, task significance, task variety, task identity, autonomy) and seven variables measuring job satisfaction (supervisor satisfaction, career-future satisfaction, financial satisfaction, workload satisfaction, company identification, kind-of-work satisfaction, general satisfaction) from 784 executives in the corporate branch of a large retail merchandizing corporation. The researcher had two questions: (a) Are measures of job satisfaction associated with measures of job characteristics? (b) If so, what features of job satisfaction are related to what features of job characteristics?

**What analysis procedure(s) do you recommend and why?**

When looking to determine the relationship between one or more sets of variables, utilizing canonical correlation analysis is the best choice. In this case, canonical correlation analysis would be able to help the researcher ascertain the correlations between one or more job characteristics and one or more job satisfaction measures.

## 2

Researchers constructed a questionnaire containing 200 questions about job satisfaction in the nursing profession. To reduce its length, they administered it to a sample of 500 current and retired nurses and used PCA to identify 15 main factors accounting for 98.3% of the variation in all 200 questions. Subsequently, they reduced the questionnaire to a manageable 30 questions specifically designed to measure those 15 factors. Now the researchers want to use discriminant function analysis to identify factors that distinguish between nurses who will stay in the profession and nurses who will leave the profession for work elsewhere.

**Should the researchers be concerned that the reduction from 200 to 30 questions might have overlooked key factors that could distinguish these two groups?**

The aim of PCA is to reduce dimensions within a dataset without losing variance explained. Assuming that the 15 main factors explain an appropriate amount of variance within the dataset, they should not be concerned.

## 3

A researcher is interested in characterizing different aspects of the American Protestant and Catholic churches' membership. These variables are the average percentage of weeks that individuals attended a church meeting, the average number of secular organizations to which members belong, the average percentage of members that describe themselves as being strong church members, and the average annual income of members. Principal components analysis on the standardized variables was used to try to reduce the dimensions.

**(a) Use the following table to decide how many principal components should be retained. Justify your reasons.**

PC	$\lambda_i$	Cumulative %
1	2.9284544	73.21
2	0.6403080	89.22

PC	$\lambda_i$	Cumulative %
3	0.3271641	97.40
4	0.1040735	100

There are two rules that can be used to select which principle components to retain: the 80% rule, and any eigenvalue greater than the average eigenvalue rule. Since the variables are standardized, the average eigenvalue is 1. By the second rule, PC1 would be the only one chosen. By the 80% rule, both PC1 and PC2 would be chosen. PC1 and PC2 should be chosen in this instance because even if PC2 is not as significant in our calculations, the extra variance explained in the dataset is nice to have to ensure that the original variance in the dataset is as intact as possible.

**(b) Below is a scatterplot of the first two principal components. Use this scatterplot and the equation below to interpret the first principal component.**

The first principle component contains two positive and two negative coefficients which are roughly the same magnitude. The positive coefficients are associated with Church Attendance and the average percent of members that describe themselves as strong church members. The negative coefficients are associated with the average number of secular organizations to which members belong and the average annual income of members. This principle component represents the difference between average church attendance plus self-described strong church members and average number of secular organizations plus average annual income of members. I would expect that points on the extreme right end of the x axis would be churches with strong members, large attendance, and low involvement elsewhere. It is no surprise to find Jehovah Witness and Mormon in that category.