

# Making Connections: Descriptive Statistics

## Overview

The purpose of this sheet is to help you make connections between graphical methods (Chapter 2) and numerical summary measures (Chapter 3).

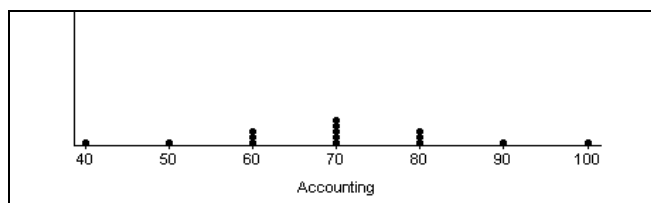
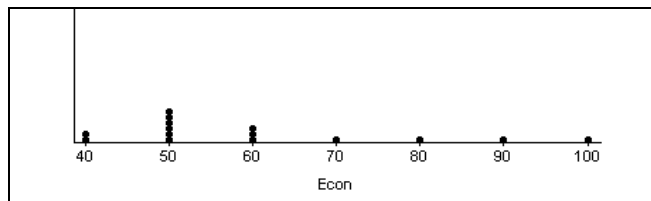
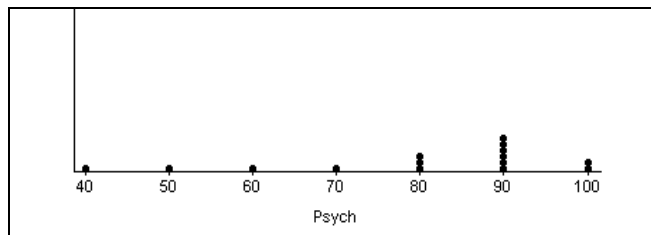
## Scenario

Given below are the scores for the first exam in each of three courses. Fortunately each set of scores has been sorted, which reveals that each exam had scores ranging from 100 down to 40.

<i>Psych</i>	<i>Econ</i>	<i>Accounting</i>
100	100	100
100	90	90
90	80	80
90	70	80
90	60	80
90	60	70
90	60	70
90	50	70
80	50	70
80	50	70
80	50	60
70	50	60
60	50	60
50	40	50
40	40	40

## Graphical Portrayal

Presented below is a *dotplot* of each course's exam scores. In such a plot, each data point is represented by a dot. When a value occurs more than once its dots are stacked, which helps convey its *frequency* of occurrence—just as a histogram's height indicates the frequency of a class.



Take a moment to look each course's dotplot and compare it with the column of exam scores it portrays. Does the shape of each dotplot make sense? How would you describe the *skew* of each distribution? Which of the distributions has the *least* amount of dispersion? To which distribution would the Empirical rule *best* apply?

---

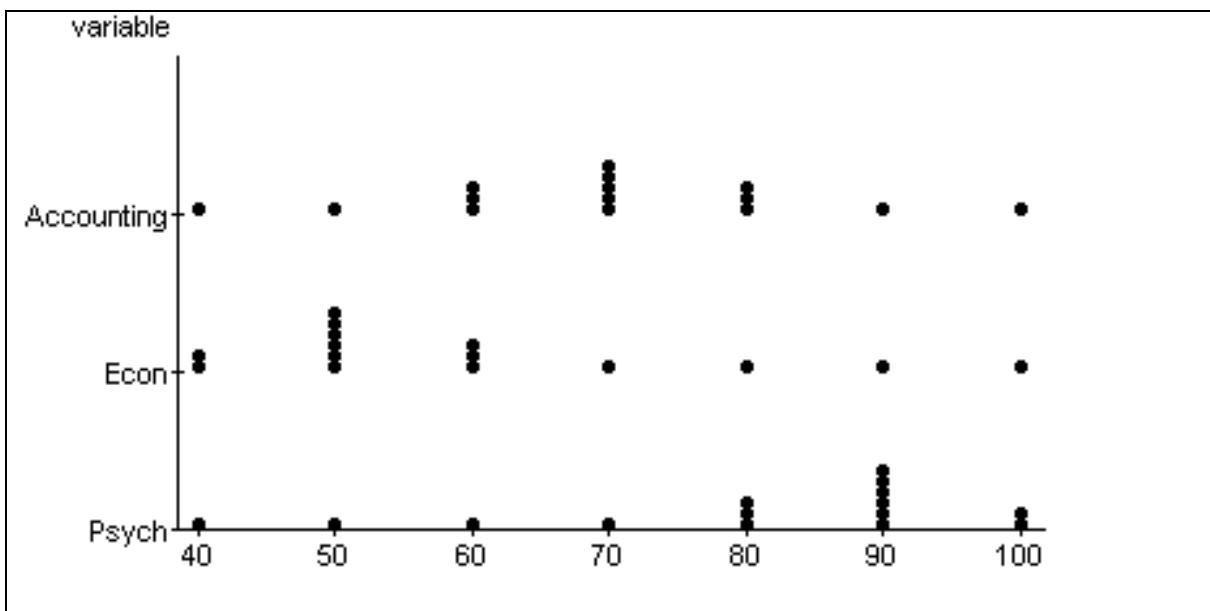
## Numerical Summary

The table below provides a numerical summary of each course's exam scores. Unfortunately the computer program I used to generate these numerical summaries does not print the value of the *mode*. Take a moment to go back to the three columns of scores and determine the *mode* for each class and pencil them in below.

Variable	n	Mean	Variance	Std. Dev.	Median	Range	Min	Max	Q1	Q3
Psych	15	80	314.2857	17.728106	90	60	40	100	70	90
Econ	15	60	314.2857	17.728106	50	60	40	100	50	70
Accounting	15	70	228.57143	15.118579	70	60	40	100	60	80

**Comparing measures of Central Tendency.** I've combined all three dotplots below. On the Psych dotplot, label the value of the median (90), the mean (80), and the mode. Can you see why the *mean* is less than the *median* for the Psych test? Repeat this labeling of median, mean and mode, and the median vs. mean comparison for the Econ and Accounting test score distributions.

**Getting a feel for standard deviation.** Recall that *standard deviation* measures the amount of dispersion of the data values from the mean. Look at the value of each class' standard deviation. Look at their respective dotplots. The Accounting scores have the *smallest* standard deviation (15.12). Note that the standard deviations for the Psych and the Econ tests are the same (17.73). Based on the dotplots, does this make sense? Can you understand why their standard deviations are the same value?



---

## Additional Practice

1. Select one of the three data sets and *manually* calculate the mean, median, standard deviation,  $Q_1$  and  $Q_3$ . Use the values above as check figures.
  2. What proportion of Psych students scored within 2 standard deviations of the mean? Does your finding support Chebyshev's theorem?
  3. What proportion of Accounting students scored within 1 standard deviation of the mean? Does your finding roughly agree with the Empirical rule?
  4. Joe Zin got one of the 80's in the Accounting test. What is the z-score for his Accounting test?
  5. Joe Zin also got (the only) 80 on the Econ test. What is the z-score for his Econ test?
  6. Compare Joe's two z-scores from the previous 2 problems. In which class did Joe have a better *relative* performance compared with his peers?
-