

PSYCHOLOGICAL RESEARCH

The Whys and Hows of the Scientific Method and Statistics

1A: THE PURPOSE OF STATISTICS

Statistics are tools that we use to understand sets of data. Consider the following fictional data set.

A	60	95	H	65	195
B	78	260	I	64	135
C	62	120	J	71	180
D	72	155	K	70	188
E	71	170	L	70	160
F	70	162	M	76	220
G	64	135	N	74	235

Assume that the data shown here are the heights (in inches) and weights (in pounds) of 14 students in a class. Answer the following questions as best as you can.

1. What does a “typical” student in this class look like in terms of height and weight?
2. Is there anyone in the class who has the same height or weight as another student? Identify any students with the same score on the two measures.
3. What are the highest and lowest weights and heights in the class? What do these values tell you about the group of students in the class?
4. How do the two different sets of measures (known as distributions) differ? Do you notice anything about these two distributions that distinguishes them?

1B: SCIENCE IN THE MEDIA

More than ever before, we are presented with data and statistics in the news. However, most of the time these analyses are reports of reports (secondary sources). In other words, they are interpreted, summarized, and often simplified by reporters. This exercise is intended to demonstrate the importance of interpreting reports as a critically informed consumer.

Part I

Instructions: Find an article in the newspaper (online newspapers are fine) that reports the results of a research finding. (*Hint*: Check the science section.)

For your chosen article, try to identify as many “scientific method details” from the research articles as you can by answering the following questions.

1. What was the title and who were the authors of the original study upon which the statements in the article are based? What was the hypothesis for the research?
2. What methodology was used (e.g., experimental, correlational, case study)? Who were the participants, and how were they recruited?
3. What were the conclusions of the research?
4. What were the limitations of the study? How convinced are you by the study’s results?
5. What questions about the research do you have? What other details were left out that would be useful in evaluating the quality of the study?

Part II

Instructions: Now try to find the original research article on which the news story was based. You may be able to find the article with a search of Google Scholar, or you may need to use a research articles database (e.g., PsycINFO) to search for the original research article. Read the original article, paying close attention to the scientific method details that you summarized earlier. Then answer the following questions.

1. How well do you think the news story conveys the research findings presented in the article? (*Hint*: Read the abstract summary of the article to get a simple summary of the findings of the study.) Were the original findings accurately portrayed in the news story? Why or why not?
2. How do you think the popular press article should be changed to provide a more accurate depiction of the original research article published by the researchers?

1C: UNDERSTANDING YOUR DATA

The following table contains a data set that describes the top 25 salaries for Major League Baseball players as of opening day of the 2016 season:

Player	Team	Position	Age (as of April 28, 2016)	Salary
Kershaw	Dodgers	Pitcher	28	34.57
Greinke	Diamondbacks	Pitcher	32	34
Price	Red Sox	Pitcher	30	30
Verlander	Tigers	Pitcher	33	28
Cabrera	Tigers	First base	33	28
Hernandez	Mariners	Pitcher	30	25.85
Sabathia	Yankees	Pitcher	35	25
Lester	Cubs	Pitcher	32	25
Howard	Phillies	First base	36	25
Pujols	Angels	Designated hitter	36	25
Cano	Mariners	Second base	33	24
Hamels	Rangers	Pitcher	32	23.5
Teixeira	Yankees	First base	36	23.13
Mauer	Twins	First base	33	23
Ramirez	Red Sox	First base	32	22.75
Scherzer	Nationals	Pitcher	31	22.14
Upton	Tigers	Left field	28	22.13
Tanaka	Yankees	Pitcher	27	22
Reyes	Rockies	Shortstop	32	22
Gonzalez	Dodgers	First base	33	21.86
Crawford	Dodgers	Left field	34	21.61
Werth	Nationals	Left field	36	21.57
Ellsbury	Yankees	Center field	32	21.14
Davis	Orioles	First base	30	21.12
Shields	Padres	Pitcher	34	21

1. Who are the individuals in this data set?
2. In addition to the player's name, how many variables does the data set contain? Which of these variables take numerical values?
3. What do you think are the units in which each of the numerical values is expressed? For example, what does it mean when Howard's salary is listed as 25?
4. What is the most common position in the data set? What is the most common salary? Do you think the most common salary will be the same as the average salary? Why or why not?

1D: DISPLAYING DISTRIBUTIONS

1. Create a frequency table including the range of responses, frequency, proportion, percentage, cumulative proportion, and cumulative frequency for the following data illustrating the number of correct responses on a quiz:

1, 4, 3, 2, 3, 4, 5, 2, 3, 5, 5, 3, 2, 1, 4, 3, 2, 3, 1, 3, 4, 3, 2, 4

2. What percentage of students scored a 3 or lower on the quiz in Question 1?

3. Create a frequency distribution graph of the data from Question 1. What is the shape of this distribution?

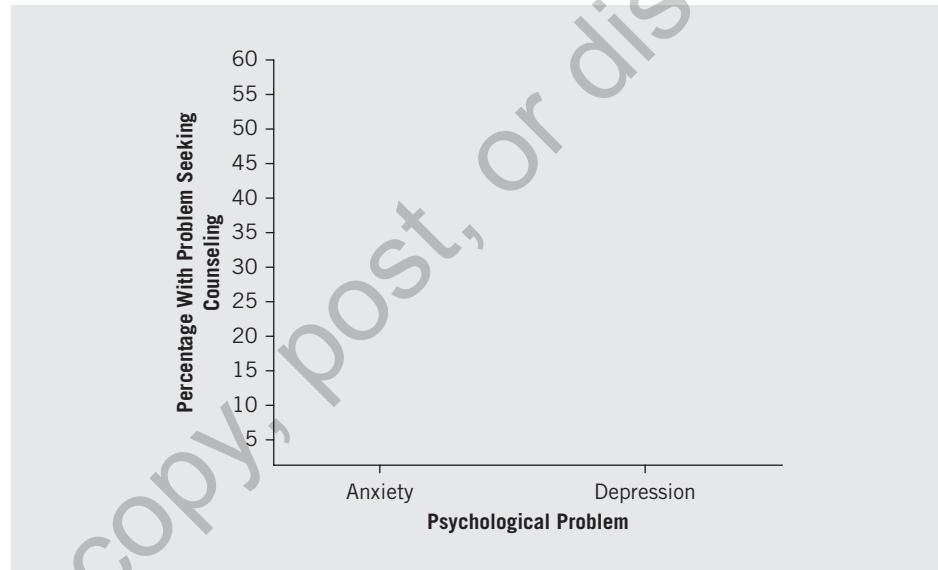
4. What is a typical score in the distribution? What do you know about this score in the distribution?

1E: MAKING AND INTERPRETING GRAPHS

A study has been conducted to compare men and women on the likelihood of seeking counseling for a psychological problem. A survey was completed by 1,000 men and 1,000 women to determine the number of each group suffering from anxiety or depression. The survey also asked if the respondent had sought counseling for his or her anxiety or depression. The following mean values indicate the percentage of those who reported one of the psychological problems and also sought counseling.

	Anxiety	Depression
Men	35%	15%
Women	20%	55%

1. Complete the following bar graph by including a point in the graph for each mean value given previously. Be sure to connect lines for each gender.



2. Re-create the graph as a bar graph using a software package such as Excel. If using Excel, type in the means and variable levels as given previously into a new worksheet, highlight what you have typed, and choose Insert Chart. Under Chart Type, you can choose a bar graph. Chart Options allow you to label the axes and adjust axis scales and fonts.
3. Describe in your own words the results displayed in the graphs.

1F: SETTING UP YOUR DATA IN SPSS: CREATING A DATA FILE

Create a data file in SPSS based on the following data, and then answer the following questions.

ID#: 568 Name: Joe Hart Age: 25 Gender: Male Income: \$23,000 IQ: 105	ID#: 276 Name: Mary Swanson Age: 37 Gender: Female Income: \$41,000 IQ: 115	ID#: 384 Name: Sam Lewis Age: 61 Gender: Male Income: \$56,000 IQ: 125
ID#: 866 Name: Chin Lee Age: 32 Gender: Male Income: \$36,000 IQ: 140	ID#: 231 Name: Al Walton Age: 39 Gender: Male Income: \$29,000 IQ: 95	ID#: 476 Name: Sara Smith Age: 27 Gender: Female Income: \$18,000 IQ: 90
ID#: 07 Name: David Dodge Age: 34 Gender: Male Income: \$29,000 IQ: 115	ID#: 647 Name: Michelle Fried Age: 38 Gender: Female Income: \$22,000 IQ: 105	ID#: 261 Name: Tom Hunt Age: 54 Gender: Male Income: \$38,000 IQ: 115

1. How many individuals are in your data set? Does this match the number of rows in your SPSS data window (it should)?
2. How many variables are in your data set? Does this match the number of columns in your SPSS data window (it should)?
3. Which of your variables are categorical variables? Which are continuous variables?
4. For each variable, what measure of central tendency would you report?

1G: DISPLAYING DISTRIBUTIONS IN SPSS

Part I

Use the Datafile_1.sav data file at <http://edge.sagepub.com/mcbridermstats1e> to create a frequency distribution table for the Quiz 1 variable using SPSS. Then answer the following questions using your table.

1. What percentage of the scores is at or below a score of 7?
2. Where does it appear that most of the scores are located?
3. What does your answer to Question 2 tell you about the difficulty of the quiz?

Part II

Now create a frequency distribution table using SPSS for the second quiz variable. Compare this table to the one you created in Part I to answer the following questions.

1. For which quiz do the scores appear to be more evenly distributed across the scale?
2. Which quiz appeared to be harder? How do you know this?