

## Data Analysis Assignment 2

### Who are the Poor?

### What Will You Learn By Doing This Assignment?

#### *Skills*

- You will learn why it is important to understand what percentages refer to (percentage of what?!)
- You will gain some experience writing about simple, two-variable tables.

#### *Substance*

- You will learn something about what kinds of households in the United States are at greatest risk of being poor.

### The Problem

This assignment requires you to do a simple analysis of poor households in the United States. It asks you to think about two questions and to think about why these are DIFFERENT questions:

- a. Are certain kinds of households more likely to be poor than others?
- b. What sorts of households make up the largest proportion of the poor?

You will be looking at three characteristics that may be associated with poverty: race, age, and family type.

### Getting the Information You Need

Let's start by looking at differences in earnings. You can access WebCHIP through the SSDAN website. Use these instructions:

1. Go to <http://www.ssdan.net/datacounts/data/>
2. From there, click "Browse" on the left sidebar. Find "**census2000**" in the drop-down box and select it.
3. Scroll down through the list of data sets until you find **fpov2k.dat**. Highlight and click "submit."
4. You can also click [here](#) to launch the dataset in WebCHIP.

This will open the data file you will be using. The data file contains information from the Census about poor households in the US and allows you to categorize the poor by race, age, and family type.

Use the "Marginals" Function to see the breakdown of the population for each of the variables in the data set: poverty, race, age and family type. Now you can begin to analyze which households are likely to be poor.

You are trying to find out what factors affect a household's chances of being poor. So, that means that "poverty" is your dependent variable, i.e., the thing you're trying to explain. Race, Family Type and Age are the "independent variables" you're looking at, i.e., the factors you think might affect a household's chances of being poor.

Select "poverty" as your row variable.

Select one of the other variables in the data set: either race, age, or family type as your column variable.

Now, create a crosstab. What this does is to get the computer to calculate your chances of being poor if you are in certain age groups. At the top of the page, you'll still see the marginals you computed earlier. Below that, you'll see a table showing you each age group and the percentage of households headed by someone in that age group that is poor and non-poor. For example, it shows that 30.4% of households headed by a person under 25 are poor. Notice that if you add the percentages of the poor and non-poor for each age group, it adds up to 100%.

Underneath the percentages, you will see a number in parentheses. This is the number of households in that group. For example, the table shows that there are 788,564 poor households headed by someone under the age of 25.

Copy the information from the table you made (note: you probably won't be able to print the information on the screen. You can hand copy it, or, if you know how, you can do a screen shot and print that. Or, you can ask your computer to "copy" the data in the table paste it into a program like Excel and print that. The easiest thing to do is to copy the data by hand, although that's obviously slow.).

Now repeat the process for the other two variables. Remember to make poverty your "row variable," then pick one of the column variables, then hit "cross-tabs." Copy or print your results each time.

You now have the data you need to do the assignment.

## **Analyzing What You Found**

First, you need to think about the chances that households headed by members of certain age, racial or family-type groups will be poor. You can find information about this by looking at the percentages in the tables you've produced.

Then, you need to think about whether households headed by members of certain age, racial or family-type groups make up the largest proportion of poor households. You can find information about this by looking at the numbers in parentheses under the percentages (Note: it will help you to answer this part of the question if you add up the numbers for all groups among the poor. This will give you a total number of poor households to think about).

## Data Analysis Assignment 2 – Worksheet

**Step 1: Calculate the “marginals” – i.e., find out how the people in the population are distributed among the various categories in each of the variables you are using: race, age, family type and poverty status**

To get this information, use the `marginals` function. Then look at the data that appears. Fill in the information in the spaces below:

RACE: White (non-Latino) \_\_\_\_\_

Black \_\_\_\_\_

Asian \_\_\_\_\_

Latino \_\_\_\_\_

AGE: <25

25-34 \_\_\_\_\_

35-44

45-54

55-64

65+ \_\_\_\_\_

**FAMILY TYPE:**

Married Couple \_\_\_\_\_

Male-Headed

Female-Headed

POVERTY STATUS:

Poor

Non-Poor

**Step 2: Calculate Cross-Tabs for Poverty and Age (i.e., find out what proportion of each age group is poor). BE SURE TO COPY BOTH THE PERCENTAGE AND THE ACTUAL NUMBER INTO THE CELLS BELOW.**

*To do this, set “poverty” as your “row” variable and “age as the “column” variable. Create a crosstab and look at the table that appears (you may have to scroll down to see the entire table). Copy this information into the matrix below.*

|          | <25 | 25-34 | 35-44 | 45-54 | 55-64 | 65+ |
|----------|-----|-------|-------|-------|-------|-----|
| Poor     |     |       |       |       |       |     |
| Non Poor |     |       |       |       |       |     |

**Step 3: Calculate Cross-Tabs for Poverty and Race (i.e., find out what proportion of each racial group is poor). BE SURE TO COPY BOTH THE PERCENTAGE AND THE ACTUAL NUMBER INTO THE CELLS BELOW.**

*To do this, leave “poverty” as your “row” variable and “race” as the “column variable.” Create a crosstab and look at the table that appears.*

|          | White | Black | Asian | Latino |
|----------|-------|-------|-------|--------|
| Poor     |       |       |       |        |
| Non Poor |       |       |       |        |

**Step 4: Calculate Cross-Tabs for Poverty and Family Type (i.e., find out what proportion of each type of family is poor). BE SURE TO COPY BOTH THE PERCENTAGE AND THE ACTUAL NUMBER INTO THE CELLS BELOW.**

*To do this, select “poverty” as your “row” variable and “family type” as the “column variable.” Create a crosstab and look at the table that appears*

|          | Married Couple | Male-Headed | Female-Headed |
|----------|----------------|-------------|---------------|
| Poor     |                |             |               |
| Non Poor |                |             |               |

## Writing Up Your Analysis

Write a report describing what you found. Make sure that you back up what you say by making specific reference to the tables you produced – that means, talk about the numbers!! Your report should be typed and should contain the following paragraphs:

- a. A paragraph discussing the chances that households headed by people of different ages will be poor.
- b. A paragraph discussing the chances that households headed by members of different racial groups will be poor.
- c. A paragraph discussing the chances that households with different family structures will be poor (i.e., single-father, single-mother, married couple).
- d. A paragraph discussing whether most poor households are headed by members of a particular age group or racial group and whether most poor households have a particular family structure. (i.e., single-father, single-mother or married couple).
- e. A paragraph explaining why the two questions posed in this assignment are not the same question (“are certain kinds of households more likely to be poor than others?” and “what sorts of households make up the largest proportion of the poor?”)