DATA ANALYSIS EXERCISE
FAMILIES IN SOCIAL CONTEXT
FAMILY CHANGE: 1950 TO 1990

LEARNING OBJECTIVES:

Skill

- Using software to access and analyze census data
- Identifying independent and dependent variables
- Employing control variables
- Quantitative writing
- Learning how to construct, read, and interpret bivariate tables displaying frequencies and percentages
- Creating visual tools representing quantitative data in the form of charts or graphs
- Identifying population trends over time
- Using real world data to enhance and support key course concepts

Substance

- As a part of this examination, we will consider changes in family patterns between 1950 and 2000—a set of changes we will call "the second transformation." In this exercise, you will use census data to trace some of these changes.
    - In the first part of this exercise, you will examine marital status, number of children born to women, and household type in 1950, 1960, 1970, 1980, and 1990. The purposes are to identify the changes that have taken place, to assess their magnitude, and to consider the pace and timing of these changes.
    - In the second part of this exercise, you will examine one of these family variables in 1990. First, you will examine the relationship between race and ethnicity and the family variable. The purpose is to see if racial and ethnic groups differ on this measure of family behavior. Next, you will examine the relationship of family income to the family variable. You will control for "class" to see if racial and ethnic differences remain or diminish. The purpose is to see if the racial/ethnic differences are due primarily to class differences.

DESCRIPTION

In the first week of the course, we will examine the historical development of contemporary families in the United States. As a part of this examination, we will consider changes in family patterns between 1950 and 1990—a set of changes we will call "the second transformation." In this exercise, you will use census data to trace some of these changes.

In the first part of this exercise, you will examine marital status, number of children born to women, and household type in 1950, 1960, 1970, 1980, and 1990. The purposes are to identify the changes that have taken place, to assess their magnitude, and to consider the pace and timing of these changes.

In the second part of this exercise, you will examine one of these family variables in 1990. First, you will examine the relationship between race and ethnicity and the family variable. The purpose is to see if racial and ethnic groups differ on this measure of family behavior. Next, you will examine the relationship of family income to the family variable. You will control for "class" to see if racial and ethnic differences remain or diminish. The purpose is to see if the racial/ethnic differences are due primarily to class differences.

CENSUS DATA

Every ten years, the United States Bureau of the Census surveys the population of the country to determine its size and characteristics. Attempting to locate and collect information on every person living in the United States is a very difficult undertaking, and many individuals are missed and many groups undercounted (homeless people, for example). Nevertheless, the U.S. Census makes great efforts to collect information on everyone, and census data provide an extremely valuable resource for examining the changing size, composition, and characteristics of the population of the United States.

Analysis of census data is difficult because of the large number of individuals surveyed and the extensive information collected on each individual— analysis is time consuming and expensive. The data that you will use was prepared by the Social Science Data Analysis Network (SSDAN) at the Population Studies Center at the University of Michigan. The information is assembled in data sets composed of a small number of variables for more efficient analysis. (A variable is a characteristic of an individual such as marital status, age, or income). The SSDAN staff actually created a few of the data sets that you will use specifically for this class. Each data set that you will use for this exercise and for subsequent data analysis projects in the class contains information on a random a sample of individuals in the United States who are 15 years of age and over surveyed by the U.S. Census.

USING WebCHIP

1. Go to the DataCounts! site: http://www.ssdan.net/datacounts
2. You will be using WebCHIP software to produce crosstabluations that will show changes in family behavior over time and differences in family behavior between ethnic/racial and income groups.

CROSSTABULATIONS AND GRAPHING

You will be using two techniques to examine changes in the family between 1950 and 1990. A crosstabulation shows the relationship between two variables, for example "year" and "marital status" or "race/ethnicity" and "number of children." The crosstabulation below shows the relationship between "year" and "education" for adults.   It shows the educational attainment of individuals in four different census years. In 1950, 65.8% of the population had not completed high school. In 1980, 33.5% of the population had not completed high school. (Notice that the table is percentaged across the rows. The numbers in the last column are the total numbers in each year and the overall sample total.) We can conclude that the educational attainment of individuals in the United States has increased between 1950 and 1980. By calculating the "percentage differences" we can assess the magnitude of these changes. For

example, we can say that 32.3% more people had not completed high school in 1950 than in 1980 (subtracting 33.5% from 65.8%). We can say that 22.2% more people completed high school in 1980 than in the 1950. And we can say that the increase in college attendance occurs at an accelerating rate since 1.4% more people completed college in 1960 than in 1950, 3.0% more people completed college in 1970 than in 1960, and 5.6% more people completed college in 1980 than in 1970.

EDUCATIONAL ATTAINMENT BY YEAR: 1950-1980

	NO HIGH SCHOOL DEGREE	HIGH SCHOOL DEGREE	COLLEGE DEGREE	TOTAL
1980	33.5%	50.2%	16.2%	13274
1970	47.6	41.8	10.6	10941
1960	59.0	33.4	7.6	9955
1950	65.8	28.0	6.2	8711
ALL	49.6	39.6	10.8	42881

In this crosstabluation, "year" is the independent variable and "education" is the dependent variable. "Year" is called the independent variable because it is the causal variable. Change in year (i.e., the passage of time and associated developments) is "causing" increases in education. Changes in education are not causing changes in year.

Graphing presents a picture of the data in a crosstabulation that represents the relationship between independent and dependent variables. For this exercise, you will be constructing line graphs. For latter projects you may choose to construct other types of graphs. (I will pass out a handout on this.)

Line graphs allow you to represent changes over time as you will do in this exercise (or to compare two or more groups on the same graph as you may do in subsequent data analysis projects).

For the above table, you would draw 3 line graphs. One representing each category of the dependent variable. Notice that the independent variable (year) is on the horizontal axis and the dependent variable is on the vertical axis. Below is an example of a line graph for the category NO HIGH SCHOOL DEGREE.

PERCENT WITHOUT HIGH SCHOOL DEGREES
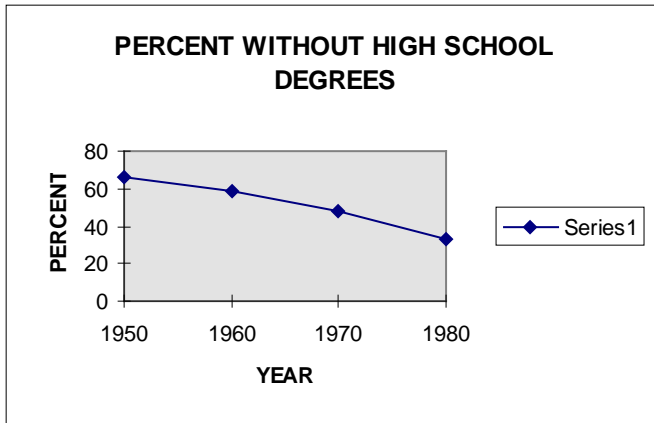
Construct your line graphs on lined paper using 2 lines to represent increments of 10%.

EXERCISE INSTRUCTIONS

A. Marital status, childbearing and household type: 1950-1990

First, examine information on the changes in marital status that occurred between 1950 and 2000. To open the dataset:

1. Go to http://www.ssdan.net/datacounts
2. Click on the "Data" in the menu bar
3. From there, click "Browse" on the left sidebar. Find **"centrend"** in the drop-down box and select it.
4. Scroll down through the list of data sets until you find **"marr502k.dat"** Highlight and click "submit."
5. You can also click here to launch the dataset in WebCHIP.
6. Create a Percent Across table with "Year" as the row variable and "Marital" as the column variable (the marital status of individuals). Remember that the independent variable should always be the row variable. Always enter the row variable first when specifying variables using the crosstabulation command
7. Record the percentages in the crosstabulation in the spaces provided on the answer sheet (combining the categories as shown—simply add the divorced and separated row percentages together). Construct line graphs for each category of dependent variable--omitting widowed--showing the changes by year (year on the horizontal axis and percent on the vertical axis).

QUESTIONS:
1. What changes have taken place in marital status between 1950 and 2000?
2. Was the pace of change more rapid in some decades than in others or has change taken place at a relative constant rate?
3. Use percentage differences to summarize the magnitude of changes that have taken place over the four decades

Second, examine the changes in child bearing that occurred between 1950 and 1990.
1. Go to http://www.ssdan.net/datacounts
2. Click on the "Data" in the menu bar
3. From there, click "Browse" on the left sidebar. Find **"centrend"** in the drop-down box and select it.
4. Scroll down through the list of data sets until you find **"born5090.dat"** Highlight and click "submit."
5. You can also click here to launch the dataset in WebCHIP.
6. Create a Percent Across table with "Year" as the row variable and "Child" as the column variable (the number of children ever born to women).   Record the percentages on the answer sheet (combining categories as shown). Construct line graphs for each category of the dependent variable showing the changes by year.

QUESTIONS:
1. What changes have taken place in the number of children born to women between 1950 and 1990?
2. Was the pace of change more rapid in some decades than in others or has change taken place at a relative constant rate?

Finally, examine changes the prevalence of specific household types that occurred between 1950 and 1990.

1. Go to http://www.ssdan.net/datacounts
2. Click on the "Data" in the menu bar
3. From there, click "Browse" on the left sidebar. Find **"centrend"** in the drop-down box and select it.
4. Scroll down through the list of data sets until you find **"hh5090.dat"** Highlight and click "submit."
5. You can also click here to launch the dataset in WebCHIP.
6. Create a Percent Across table with "Year" as the row variable and "HHtype " as the column variable.   Record the percentages on the answer sheet (combining categories as shown). Record the results on the answer sheet (combining the categories as shown). Construct line graphs for each category of the dependent variable showing the changes by year.

QUESTIONS:
1. What changes have taken place in the distribution of household types between 1950 and 1990?
2. Was the pace of change more rapid in some decades than in others or has change taken place at a relative constant rate?

B. Race/ethnic and class differences in families in 1990

Although many changes have taken place in families since the 1950s, the families of some groups may differ from one another. For this part of the exercise choose either marital status or childbearing.

1. Go to http://www.ssdan.net/datacounts
2. Click on the "Data" in the menu bar
3. From there, click "Browse" on the left sidebar. Find **"custom"** in the drop-down box and select it.
4. Scroll down through the list of data sets until you find **"marrpov9.dat"** Highlight and click "submit."
5. You can also click here to launch the dataset in WebCHIP.
6. Create a Percent Across table with "RaceLat" as the row variable and "Marital" or "Child" as the column variable. Record the percentages on the answer sheet (combining categories as shown). Record the results on the answer sheet.

QUESTIONS:
1. What differences exist between race/ethnic groups on this measure of family behavior? [NOTE: The data in these files are for household heads only and are not directly comparable to the data you have analyzed in Part A.]

Second, create a crosstabulation of "pov4" and "marital' or of "pov4" and "child." (Use the dependent variable you chose above.) Record the row percentages on the answer sheet. variable.

QUESTIONS:
1. Is income related to this measure of family behavior?
2. How do the income groups compare to the patterns you discovered for race/ethnic groups?

Finally, create crosstabulations of "racelat" and "marital" controlling for "pov4" or a crosstabulation of "racelat" and "child" controlling for "pov4." Record the row percentages for all the tables on the answer sheet.

QUESTIONS:
1. What happens to the relationship between race/ethnic group and the measure of family behavior within categories of income?
2. Do the differences between race/ethnic groups that you discovered above remain, diminish in magnitude, or disappear among the four income groups?
3. Which variable is most important in explaining differences in behavior concerning this family variable. (Make sure to look at differences within tables for this part of the exercise. For example, look at the differences between ethnic groups within each category of income. If the differences disappear or are reduced greatly, you can conclude that income and not race/ethnicity is the main source of difference in the dependent variable.)

SUMMARIZING YOUR RESULTS AND DRAWING CONCLUSIONS

Finally, write a paper summarizing the results of your investigation of family change between 1950 and 1990.   Your summary should be based on your answers to the questions posed at each step of the analysis. In your paper, refer to the tables and graphs you have constructed. End your paper with any general conclusions you have reached and some additional questions these data have raised for you. Your paper will be evaluated on organization, comprehensiveness, clarity, and mechanics (grammar, punctuation, and spelling). Turn in your answers sheets and graphs with your paper.

The paper should not exceed four double-spaced, typed pages in length (approximately 1000 words and an absolute maximum).