Module 1: Population structures and cohorts
Kyle Crowder

Introduction

This module provides a gentle introduction to the use of WebCHIP software and census data to investigate basic population issues. In the first part of this module, you will use data from the 1990 U.S. census to create population pyramids for several racial and ethnic groups. These population pyramids provide the ability to view the age and sex structure of a population. They not only allow us to view the cumulative impacts of past patterns of fertility, mortality, and migration, but also provide hints about what a particular population is likely to look like in the future. In the second part of the module you will continue our exploration of basic U.S. population structures by comparing some of the basic characteristics of a number of distinct birth cohorts.

Learning objectives

Skill

- Develop an understanding of WebCHIP procedures for examining cross-tabulations and introducing controls.
- Develop hypotheses and identify the independent and dependent variable implied in an hypothesis.
- Investigate and describe the relationship between two variables and assess whether the relationship supports or contradicts the hypothesis being tested.

Substance

- Explore racial differences and historical trends in the structure of the U.S. population. Examine differences in socioeconomic differences between birth cohorts.

Data information

For this exercise you will use several datasets drawn from census information from U.S. censuses between 1950 and 1990. Most data files contain information on the entire population of the U.S. as of the year of data collection 1990. Some data files contain information on employment and occupational characteristics of the population and, therefore, contain only population members who are old enough to legally work (16+).

Part I. Racial and ethnic differences in population structures

A. Start your exploration of U.S. population structures by using the dataset *POPUSA9.dat* to create a series of tables showing the distribution of the U.S. population by age and sex for the following racial/ethnic groups.

   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 **"cen1990"** in the drop-down box and select it.
4. Scroll down through the list of data sets until you find **"popusa9.dat"** Highlight and click "submit."
5. You can also click [here](#) to launch the dataset in WebCHIP.

- To see the percentages of males and females falling into each age category you would create a table showing the values of the variable "age" by the values of the variable "gender." For this exercise, we want the percentage of each gender group that falls into each age category. For example, we want to know what percentage of the female population is age 0-4, what percentage of the female population is in the 5-14 age category, etc. In other words, you want the total number in each of the categories of the gender variable to be your denominator in calculating your percentages. Create a Percent Down table with "Age" as the row variable and "Gender" as the column variable.
- If things go as planned, your percentages within the "Female" category should add up to 100 as should the percentages within the "Male" category.

Use your results to complete the following table:

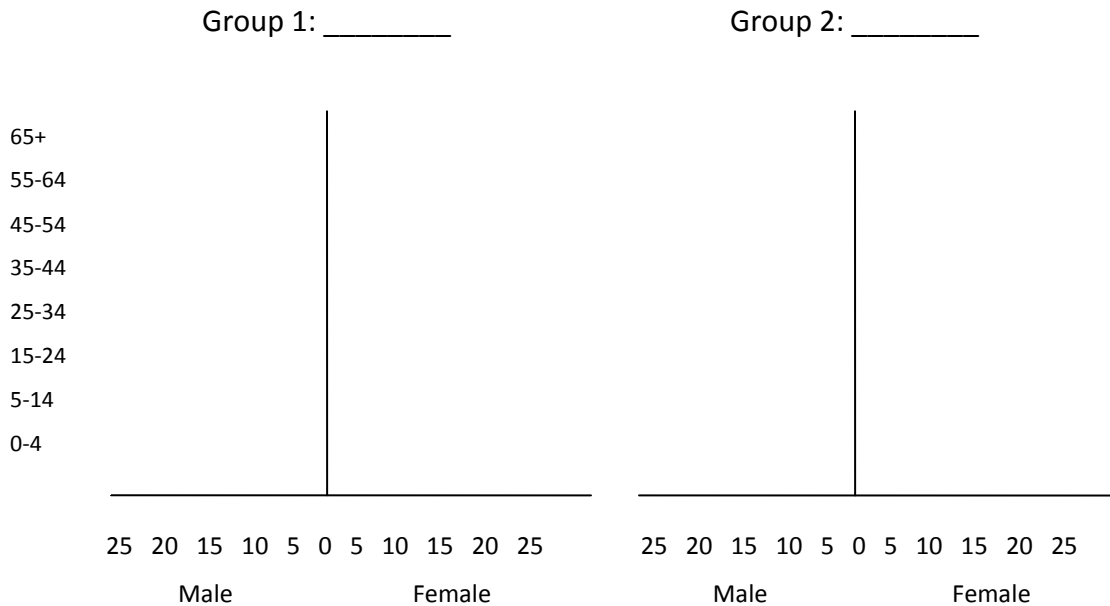| Age	Gender	
	Male	Female
65+		
55-64		
45-54		
35-44		
25-34		
15-24		
5-14		
0-4		
Total	100%	100%

Questions:

1. Which age category contains the largest percentage of the male and female populations? Which cohort do the people in this age category represent (refer back to the cohort names listed on the "Introduction and Key Definitions" page)?
2. For which age category is there the greatest sex imbalance (the greatest disparity between the relative percentages of men and women? What factors might lead to such an imbalance in this age category?

B. Now take a look at differences in the structure of the population of specific racial and ethnic groups. You can do this by following the same set-up you used in part A above, but now introduce a control for "RaceLat," the variable indicating which race/ethnicity group a person belongs to. Recall that controlling for a variable means that you are looking at some distribution or relationship within specific categories of the control

variable. By controlling for race/ethnicity, we will be looking at the distribution of "Age" by "Gender" within specific categories of race/ethnicity.

If all goes as planned, you should have separate tables (much like the one in part A) for each of the following race/ethnicity groups: Blacks, Asians, Non-Latino whites, American Indians, Latinos, and Non-Latino others.

Choose TWO of these race/ethnicity groups and construct a population pyramid for each group that you choose.

Group 1: _____   Group 2: _____

	65+
	55-64
	45-54
	35-44
	25-34
	15-24
	5-14
	0-4

25 20 15 10 5 0 5 10 15 20 25   25 20 15 10 5 0 5 10 15 20 25

Male   Female   Male   Female

Questions:

3. For each group, what does the population pyramid indicate about the dynamics of the population (e.g., does the population of the group appear to be growing, declining, stable, or what)? Explain.
4. Based on these population pyramids, what can you tell about group differences in levels of mortality and fertility? Make sure to point to specific differences in the age and gender composition of the two groups to illustrate your points.
5. What role do you think migration plays in creating the different population structures for the two groups you selected?
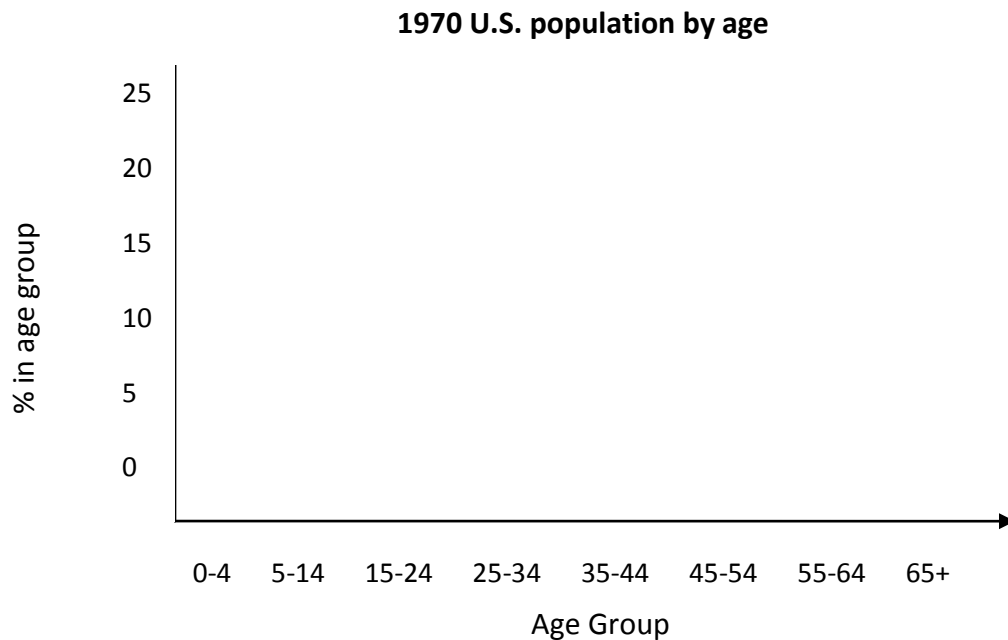
Part II: Cohort sizes, cohort differences, and potential economic repercussions

Easterlin and other theorists have argued that the size of the cohort a person is born into has important repercussions on their life chances. Specifically, members of larger cohorts face more competition for such things as jobs, housing, and even marriage partners. In the following exercises you will get the opportunity to develop and test some hypotheses based on this general idea.
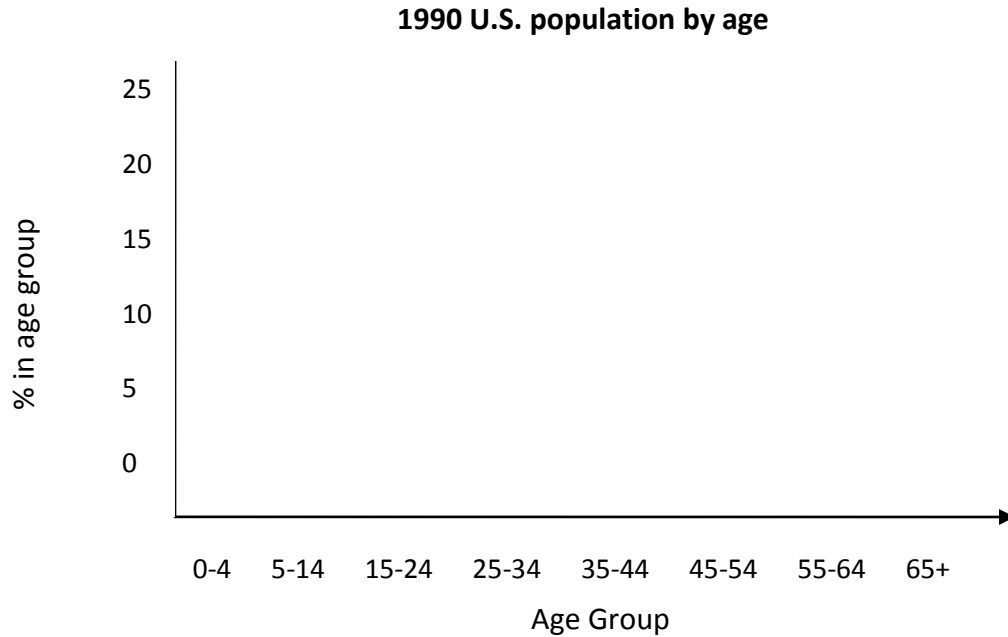
A. Let's start by looking at some changes in the age distribution of the U.S. population. Use the dataset *popstruct.dat* to examine the age distribution of the population by census year (your cross-tab will contain the variable "Age" by the variable "Period").

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 **"popstruc.dat"** Highlight and click "submit."
5. You can also click here to launch the dataset in WebCHIP.

Create a bar chart showing the percentage of the population in 1970 within the various age categories.

1970 U.S. population by age

Do the same thing for the population as of 1990.

1990 U.S. population by age

Chart with y-axis labeled "% in age group" (0, 5, 10, 15, 20, 25) and x-axis labeled "Age Group" with categories: 0-4, 5-14, 15-24, 25-34, 35-44, 45-54, 55-64, 65+

Questions:
1. Describe the key differences between the age distribution in 1970 and the age distribution in 1990? What are the largest age categories in the two years?
2. Use these charts to identify the Generation X cohort. What percentage of the population did this cohort make up in 1970? What percentage did it make up in 1990?
3. What do these line charts tell you about the relative size of the cohorts identified under "Key Concepts" on the introduction page? Which cohort appears to be largest and which is smallest?
4. What do these differences in cohort size tell us about changes in patterns of fertility during the 20th century?

B. In the U.S., people have typically begun entering the labor force between the ages of 15 and 24. Create a line chart showing changes in the number of persons age 25-34 between 1950 and 1990.

U.S. population age 15-24, 1950 to 1990
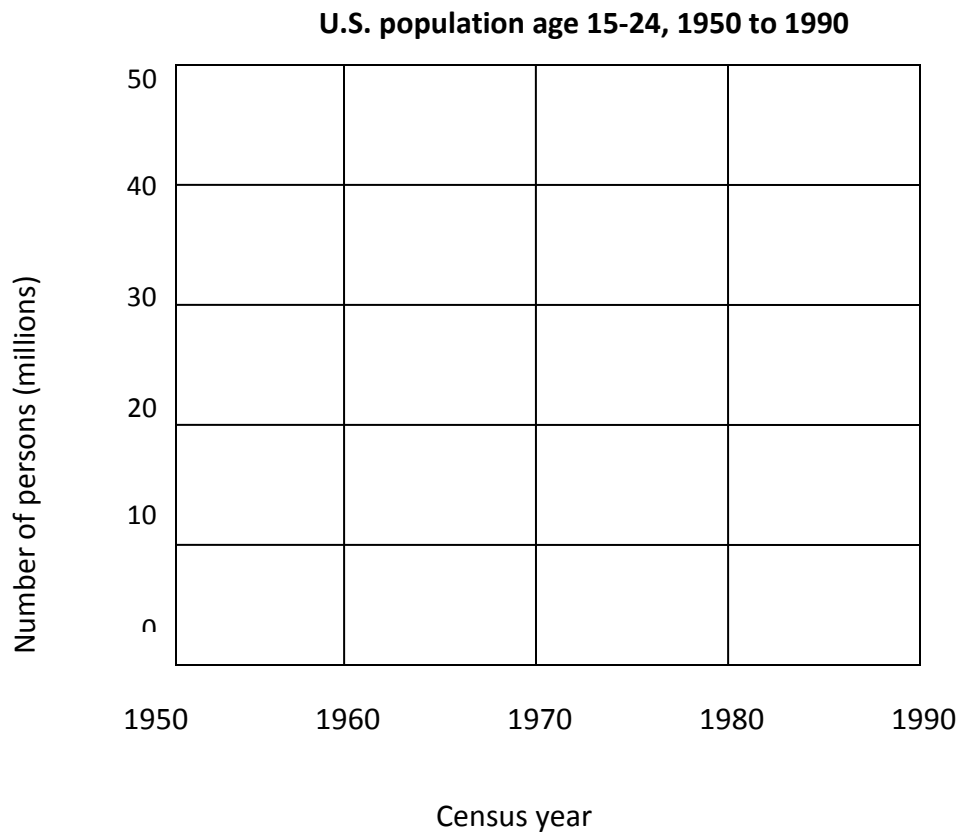
Questions:

5. How did the size of the work-entry population change between 1950 and 1990? In which year was this population the largest? Which cohort is represented in the 15-24 population in that year?
6. How might these differences in cohort size affect the relative economic opportunities of members of different cohorts? Explain your reasoning.
7. Based on your response to question 6, develop a hypothesis that predicts differences in the level of unemployment experienced by the various cohorts at the time they entered the work force. For example, your hypothesis might go something like this: "Because of differences in the size of cohorts, the level of unemployment experienced by young adults should have been highest in (which decade?) when the (which cohort?) reached working ages.
8. What is the independent variable implied in your hypothesis? What is the dependent variable?

C. Now test your hypothesis by looking at cohort differences in employment rates during the prime work-entry ages. Specifically, use the dataset ***emp5090.dat*** to look at changes in the distribution of labor force status of people age 16-24 in between 1950 and 1990. You can do this by looking at a cross-tabulation of the variable "Emp" by the variable "Year," controlling for age ("WkAge"). Use your results to complete the following table.

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 **"emp5090.dat"** Highlight and click "submit."
5. You can also click here to launch the dataset in WebCHIP.

Labor force status of people age 16-24, 1950 to 1990

	1950	1960	1970	1980	1990
% Not in the labor force					
% Unemployed					
% Employed					
Total	100%	100%	100%	100%	100%

Questions:

9. Describe the differences in labor force status across time. Which category underwent the greatest change between 1950 and 1990?
10. Are the trends in the unemployment rate among the young adult population consistent with your hypothesis about the effect of cohort size on economic opportunities? Explain your answer.
11. What other factors may have affected the labor force patterns of young adults during these various decades.

D. Many people argue that as today's younger cohorts age, they will take on many of the same social and economic attributes of older generations. We can test the validity of this assumption by using the ***edoc5090.dat*** dataset to look at cohort differences in educational attainment. To do this, create a stacked bar chart showing the educational distribution of people age 25-34 in each decade from 1950 to 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 **"edoc5090.dat"** Highlight and click "submit."
5. You can also click here to launch the dataset in WebCHIP.

Educational attainment of people age 25-34, 1950 to 1990

Key
College Grad
Some College
H.S. Grad
< H.S.

1950   1960   1970   1980   1990
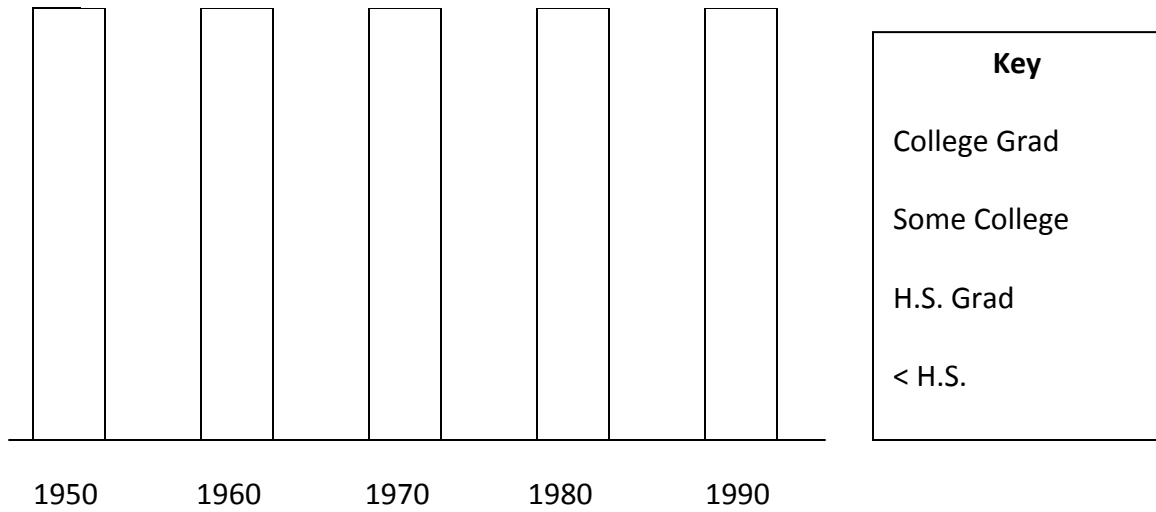
Questions:

12. Describe the basic changes in the educational attainment of young people across the decades. In what year was the educational attainment of this age group the highest? Which cohort did the 25-34 year-olds in that year represent? In which year was educational attainment the lowest and which cohort is represented in that year?
13. What societal changes might help to account for these cohort differences in educational attainment?
14. How might these changes in educational attainment affect the cohort differences in employment observed in part C? Given these patterns, can changes in employment patterns be attributed completely to differences in cohort size?