POLI 300 PROBLEM SET #5C due 10/20/10

FREQUENCY BAR GRAPHS AND HISTOGRAMS

NOTE: PUT YOUR ANSWERS TO ALL QUESTIONS OTHER THAN #4 and #5 DIRECTLY ON THESE PAGES

NAME _____

1. Refer to the enclosed figure (Attachment A) on the inserted page which shows two (merged) bar graphs, one (light-shaded bars) showing the distribution by age of Berkeley faculty members in 1969-70 and the other (dark-shaded bars) showing the distribution by age of Berkeley faculty members in 1988-89.

   (a) What percent of Berkeley faculty members were 51-55 years old in 1969-70?

   _____

   (b) What percent of Berkeley faculty members were 51-55 years old in 1988-89?

   _____

   (c) What percent of Berkeley faculty members were 40 years old or younger in 1969-70?

   _____

   (d) What percent of Berkeley faculty members were 40 years old or younger in 1988-89?

   _____

   (e) What percent of Berkeley faculty members were 61 years old or older in 1969-70?

   _____

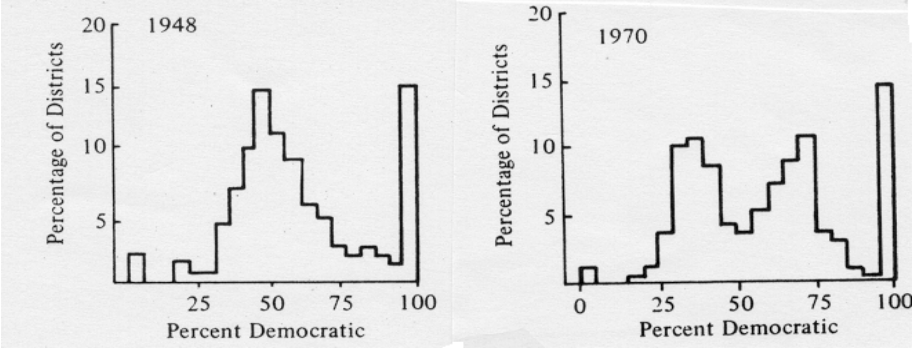
   (f) What percent of Berkeley faculty members were 61 years old or older in 1988-89?

   _____

   (g) In general, how would you describe the change in the age distribution of Berkeley faculty members from 1969-70 to 1988-89?

2. Below are two of the most famous *histograms* in political science.

Figure 1. Congressional Vote in Districts with Incumbents Running, 1948–72



SOURCE: Morris Fiorina, *Congress: Keystone of the Washington Establishment*, pp. 10-11. Fiorina in turn drew this data from David Mayhew, "Congressional Elections: The Case of the Vanishing Marginals," *Polity*, 6 (1974), pp. 295-317, who originally constructed such histogram.

For each, the *units of analysis* (cases) are Congressional districts, and the *population* is all districts with an incumbent running for re-election in the specified year. The *variable* is PERCENT OF THE TWO-PARTY VOTE WON BY THE DEMOCRATIC CONGRESSIONAL CANDIDATE. The *values* of this variable are *equal class intervals* five percentage points wide (e.g., 0-5%, 5-10%, . . . , 95-100%) created for this *continuous* variable. The Republican candidate won in each district in which the value of the variable is 45-50% or less, and the Democratic candidate won in each district in which the value of the variable is 50-55%% or greater. The first graph shows the frequency distribution resulting from the 1948 election and the second graph shows the frequency distribution resulting from the 1970 election. (Note that, in order to reduce clutter in the graphs, the sides of adjacent bars have not been drawn.)

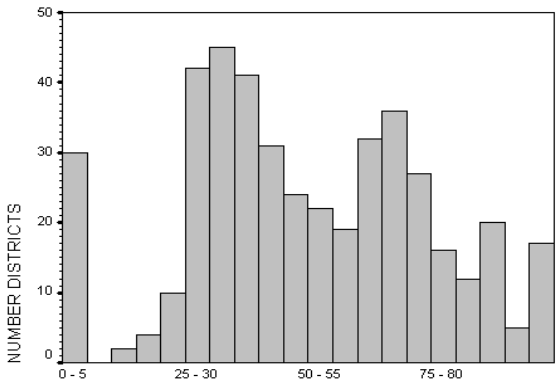
A. In what percent of the districts did the Democratic candidate win 50-55% of the vote in 1948? In 1970?

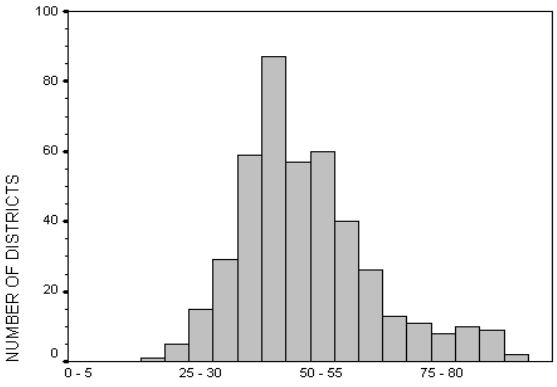
1948: _____ 1970: _____

B. Which party won the greater number of House seats (from among those districts in which an incumbent was running) in 1948? In 1970? How can you determine answers to these questions from the histograms?

C. How would you characterize the change in the shape of the frequency distribution from 1948 to 1970? What do you think this change in shape means in substantive political terms?

3. As you may have read, no more than about 30 House races were regarded as potentially "competitive" in 2004. (A "competitive" [or "marginal'] House district is conventionally defined as one in which the winning candidate gets less than 55% of the two-party vote.) The lack of competition in House races is commonly blamed on the practice of drawing House districts so that they are overwhelmingly safe for one or other party (so-called "bipartisan gerrymandering"; see, for example, the Center for Voting & Democracy, "Monopoly Politics 2004: Even Less Competition Than in 2002" (http://www.fairvote.org/2004/index.html). Here are two histograms (produced by SPSS) set up in a manner parallel to the histograms in Question 2 but which pertain to the 2000 election. The first histogram is directly comparable to the previous histograms, except that all House races are shown (not just races in which the incumbent is running for re-election) and the vertical scale shows absolute rather than relative frequencies. The second shows the distribution of House districts not with respect to the percent of the vote received by the Democratic House candidate but rather the percent of the vote received by Democratic Presidential candidate Albert Gore Jr.





Compare and contrast these two histograms and comment on what they imply for the hypothesis that lack of competition for House seats is due to of gerrymandering.

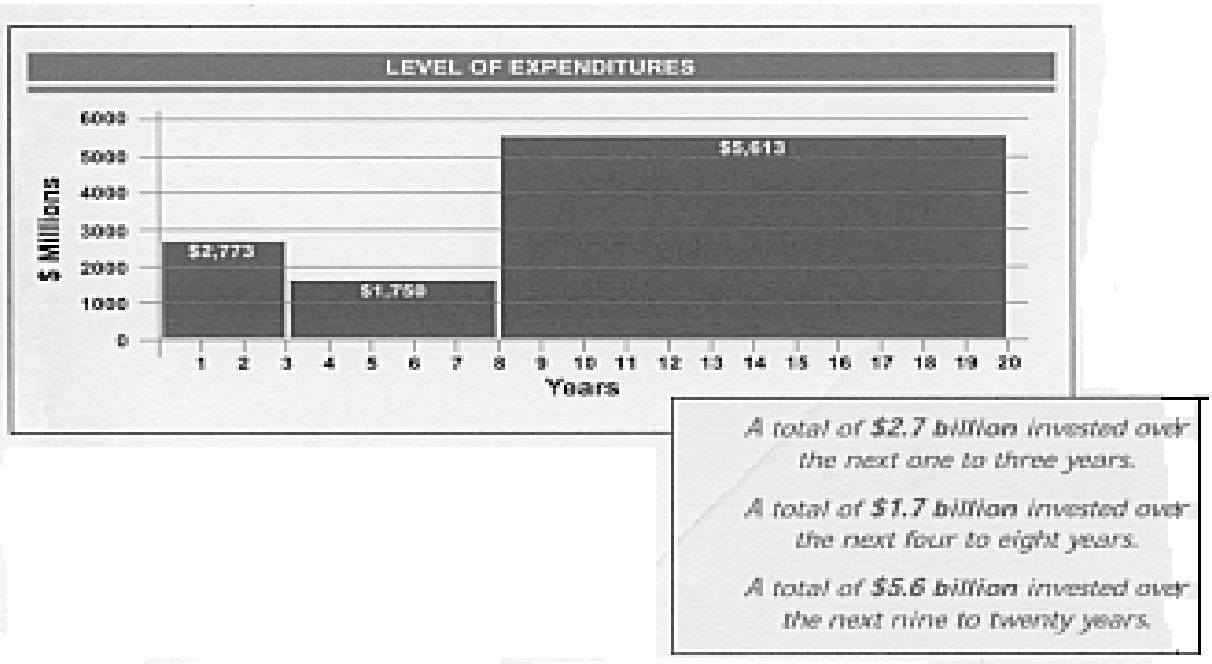
4. Here are the *absolute frequencies* for the variable V60 (AGE) in the SETUPS/NES Codebook for the year 2008 only. (You can get these absolute frequencies using SPSS by running SELECT CASES with V01 = 2008 and then by running FREQUENCIES for V60):

Age Category	Frequency
17-24	259
25-34	428
35-44	368
45-54	501
55-64	358
65-85	391
NA	20

Note. I have changed the "width" of the sixth age category from 65-99 to 65-85. Since very few NES respondents are older than 85, this change produces a more realistically shaped histogram. You should also review the footnote on the bottom of page 9 in Handout #5 on *Frequency Tables, Bar Graphs, and Histograms*.

Showing all necessary calculations along the way, construct a *histogram* of this age distribution. Draw your histogram as carefully and precisely as possible, preferably on graph paper.

5. Help Amtrak (it needs all the help it can get) by properly constructing a chart that depicts its twenty-year investment plan.



6. The density curve histogram in Attachment A shows the (continuous) frequency distribution of Annual Income in Thousands of Dollars among all U.S. households. (Source: hypothetical but half-way plausible for 1990 dollars.)

   a. The number of households that have incomes of about $10,000 is approximately what fraction of the number of households that have incomes of about $20,000?

   _____

   b. The number of households that have incomes of about $60,000 is approximately what fraction of the number of households that have incomes of about $30,000?

   _____

   c. The percent of all households that have incomes of $15,000 or less appears to be approximately

   _____

   d. The percent of all households that have incomes of $30,000 or more (cumulative frequency) appears to be approximately

   _____

   e. The percent of all households that have income in the range from $10,000 to $40,000 appears to be approximately

   _____

   f. The income level such that only about 10% of households have more than that level appears to be approximately

   _____

   g. The income level such that about half the households have less and the other half have more appears to be approximately

   _____