//Set a log file

log using "Lab 1 Notes", replace text

use "anes\_timeseries\_cdf"

set more off

\*\*\* Describe will give you basic information about the variable and how it was coded\*\*

describe

\*\*\*Many variables are coded with strings of numbers and letters that can be difficult to interpret\*\*

\*\*\*These can be resolved for ease of use with the rename command\*\*\*

\*\*\*Syntax is rename varname new varname\*\*\*

rename VCF0004 Year

rename VCF0208 Southerners\_ft

rename VCF0148 SocialClass

rename VCF0606 WasteTax

\*\*\*The tab command gives you an idea of what the responses are and how many///

\*\*\*respondents selected each category\*\*\*

\*\*\* Adding ,m to the command lists the missing values\*\*\*

tab Year, m

tab Southerners\_ft, m

tab SocialClass, m

tab WasteTax, m

\*\*\*tab1 is an spost command that gives you a series of one way tabs\*\*\*

tab1 SocialClass WasteTax, m

\*\*\*tabstat will give you a table of statistics that you ask it to give\*\*\*

\*\*\*mean is the average value of the data, median is the middle value\*\*\*

\*\*\*p25 is the value at the 25th percentile, p75 is the 75th\*\*\*

\*\*\*min is the minimum value, max is the maximum\*\*\*

\*\*\*sd is standard deviation, variance is variance\*\*\*

tabstat Southerners\_ft SocialClass, statistics( mean median p25 p75 min max sd variance ) ///

columns(variables)

\*\*\*You can use /// to continue the line of code on the next line\*\*\*

\*\*\*codebook varname, compact gives you the number of observations, unique obs///

\*\*\* mean, min, max, and the label\*\*\*

codebook Southerners\_ft SocialClass, compact

\*\*\*You can also use the list command to list what responses were given\*\*\*

list Southerners\_ft SocialClass in 1/5

\*\*\*1/5 refers to observations 1 through 5\*\*\*

\*\*\*NOTE: THE ANES HAS SEVERAL THOUSAND RESPONSES, so try not to use list unless you///

\*\*\*need to\*\*\*

\*\*\*If you need to look at continuous variables use the list and not tab command\*\*\*

\*\*\*you can also glance at your data by simply typing edit\*\*\*

edit

\*\*\*A histogram is a good way to display the frequency with which data appears\*\*\*

hist SocialClass

hist SocialClass, discrete

\*\*\*\*Change the number of bins or size of bins through hist varname, bin(#) or

\*\*\*hist varname, width(#)\*\*\*\*

hist SocialClass, bin(9)

hist SocialClass, width(2)

\*\*\*You can also add titles to your histograms through title("Title")

\*\*\*to include the frequency along the x-axis add the freq option

\*\*\*to add the norm plot add the normal option

hist SocialClass, frequency normal title("Social Class Distribution") xtitle("Frequency of Responses") ytitle("Social Class")

graph save Graph "/Users/Patrick/Desktop/histogram.png"

\*\*\*A scatter plot can show the relationship between data\*\*\*

scatter SocialClass WasteTax, jitter(4)

//all commands have options you can use to make different specifications.

//the options follow the comma after the main part of the command

//to know what the options are use the help command

\*\*\*\*You can also use graph editor

help scatter

graph save Graph "/Users/Patrick/Desktop/Graph1.gph"

//QQ plots, these are used to compare the distributions of two variables

qqplot SocialClass WasteTax

//the graph will fit a line, we can see this is not a good fit

qnorm WasteTax

//check the distribution of a variable

save lab1.dta, replace