

PSCI 2702B Winter 2024

## ASSIGNMENT TWO

### Instructions

This assignment asks you to analyze two bivariate relationships and one set of control tables using the Canadian Press data we collected earlier in the semester.

You will complete your assignment in an MS Word document. Your assignment should include:

- A title, with your name and student number.
- Variable information: provide the names, coding, and categories of three variables.
- One of these variables should be a dependent variable, and the others should be independent variables (i.e. one dependent variable you believe is caused by the other variables).
- **In each table, at least one of the variables must be nominal.**
- Hypotheses: include a statement about the relationship between each independent variable and the dependent variable. Also include a rationale for each hypothesis (2-4 sentences).
- Crosstabulations and analysis: provide a crosstabulation (the dependent variable and an independent variable), and a brief description of the relationship in each table; include the appropriate measures of association and statistical significance, and briefly analyze these (i.e. What is the strength of the relationship? How certain can we be in rejecting the null hypothesis?).
- Control tables and analysis: provide control tables, treating one of your independent variables as a control variable (pick whichever one of the two independent variables you would like to act as a control). Analyze these tables.
- Include your *Stata* syntax at the end of your assignment.

**The rest of this document provides detailed, step-by-step instructions on how to obtain the information you need to complete this assignment.**

**Do not collaborate with other students in completing your assignment. Two or more students are not permitted to submit the same work for evaluation. Submitting the same assignment as another student will be considered plagiarism.**

# 1 Assignment Elements

Below is an example of how your assignment should look in an MS Word document. Refer to this as you complete your assignment.

**PSCI 2702B Assignment 2**  
**Student Name: Stephen White**  
**Student Number: 0000000**

## Part 1: Hypotheses

*[Offer a statement about the expected relationship between the first independent variable and the dependent variable, and then a statement about the expected relationship between the second independent variable and the dependent variable.*

*Next, provide a rationale for each hypothesis – that is, why you expect to find this relationship. Each rationale should be 2-4 sentences long.]*

## Part 2: Variables

Dependent Variable

Name: Vote

Categories and Coding

Did not vote	0
Voted	1

Independent Variable One

Name: Age

Categories and Coding

Less than 30 years	1
31-45 years	2
More than 45 years	3

Independent Variable Two

Name: Income

Categories and Coding

Less than \$40,000	1
\$40,000-90,000	2
More than \$90,000	3

### Part 3: Cross-tabulations

*[Note that the number of categories in the rows and columns of your tables will depend on the number of categories in your variables. Please see the ASSIGNMENT TWO TABLE TEMPLATES document associated with this assignment. You can use these as templates for your own tables.]*

Table 1 Vote by Age

	Age		
Vote	Less than 30 years	31-45 years	More than 45 years
Did not vote	25%	35%	45%
Voted	75	65	55
Total	100%	100%	100%

*[Paragraph analyzing Table 1 should be placed here, 4-6 sentences. Remember to compare column percentages across rows. Consult course lectures to see how this is done.]*

*[Short paragraph stating and analyzing the appropriate measure of association and statistical significance (including its p-value) for Table 1 should be placed here. 2-4 sentences]*

Table 2 Vote by Income

	Income		
Vote	Less than \$40,000	\$40,000-90,000	\$40,000-90,000
Did not vote	25%	35%	45%
Voted	75	65	55
Total	100%	100%	100%

*[Paragraph analyzing Table 2 should be placed here, 4-6 sentences. Remember to compare column percentages across rows. Consult course lectures to see how this is done.]*

*[Short paragraph stating and analyzing the appropriate measure of association and statistical significance (including its p-value) for Table 2 should be placed here. 2-4 sentences]*

### Part 4: Control Tables

*[Note that the number of control tables will depend on the number of categories in your control variable: if that variable has two categories, then there will only be two control tables]*

Table 3 Vote by Age by Low Income (Less than \$40,000)

	Age		
Vote	Less than 30 years	31-45 years	More than 45 years
Did not vote	25%	35%	45%
Voted	75	65	55
Total	100%	100%	100%

Table 4 Vote by Age by Middle Income (\$40,000-90,000)

	Age		
Vote	Less than 30 years	31-45 years	More than 45 years
Did not vote	25%	35%	45%
Voted	75	65	55
Total	100%	100%	100%

Table 5 Vote by Age by High Income (More than \$90,000)

Vote	Age		
	Less than 30 years	31-45 years	More than 45 years
Did not vote	25%	35%	45%
Voted	75	65	55
Total	100%	100%	100%

*[Paragraph analyzing Tables 3-5 should be placed here, 8-12 sentences. Remember to compare column percentages across rows. Consult course lectures to see how this is done. Remember to interpret the effect of the control variable on the original relationship.]*

*[A paragraph stating and analyzing the appropriate measures of association and statistical significance (including p-values) for Tables 3-5 should be placed here. 3-5 sentences]*

### **Part 5: Stata Syntax**

[Copy and paste all of your Stata commands here]

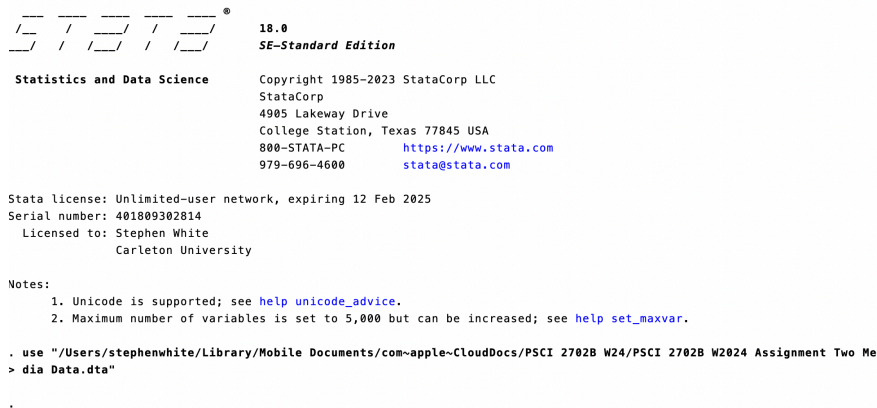
## 2 Using the Media Content Dataset

Once you have identified the variables, you can open the dataset.

From Brightspace, download the file **PSCI 2702B W2024 Assignment Two Media Data.dta** to the desktop

Then, double-click the icon named **PSCI 2702B W2024 Assignment Two Media Data.dta**

This will open the statistical software Stata/SE 18.0 (or version 17.0). In the main part of the screen you should see a “results” window, which should look something like this:



```

      _ _ _ _ _
     / / / / /
    _/_/_/_/_/

18.0
SE-Standard Edition

Statistics and Data Science
Copyright 1985-2023 StataCorp LLC
StataCorp
4905 Lakeway Drive
College Station, Texas 77845 USA
800-STATA-PC      https://www.stata.com
979-696-4600      stata@stata.com

Stata license: Unlimited-user network, expiring 12 Feb 2025
Serial number: 401809302814
Licensed to: Stephen White
            Carleton University

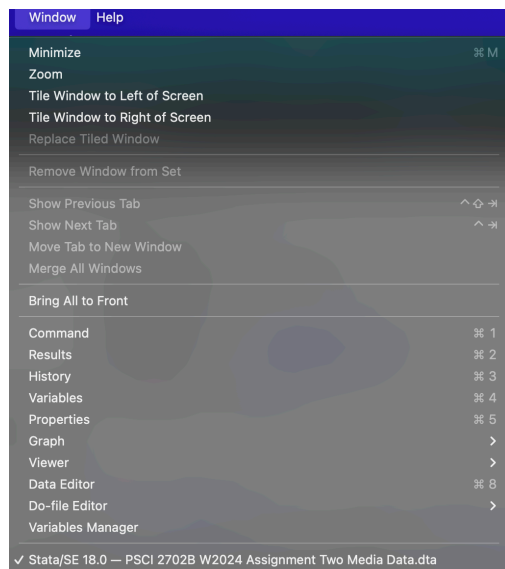
Notes:
  1. Unicode is supported; see help unicode\_advice.
  2. Maximum number of variables is set to 5,000 but can be increased; see help set\_maxvar.

. use "/Users/stephenwhite/Library/Mobile Documents/com~apple~CloudDocs/PSCI 2702B W24/PSCI 2702B W2024 Assignment Two Media Data.dta"
> dia Data.dta"
.
```

In this window, you will see the results of all commands you ask Stata to execute for you.

You might also see other windows to the left, right, and/or below the results window. If you do not, you can make them visible by using the “window” menu at the top of your screen.

First, select the “variables” window:



You should then see a window that lists all of the variables in the dataset, as well as labels describing the variable. These are the data you and your classmates collected. Refer to Assignment One for details on these variables (I have combined some variable categories, and added some other data). The variable window will look something like this:

[illegible]

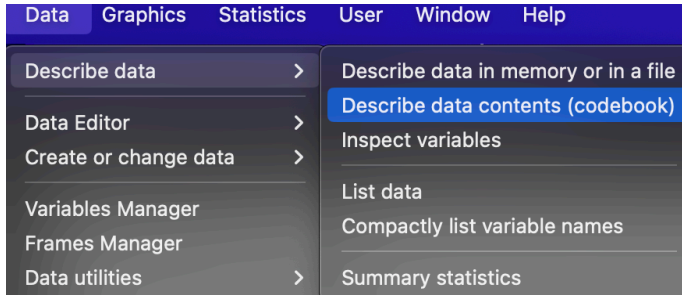
Using the window menu, you can also select the “command” window. You should then see a window in which you will type commands for Stata to execute. This will get you the results you need for your assignment. That window looks something like this:

Command

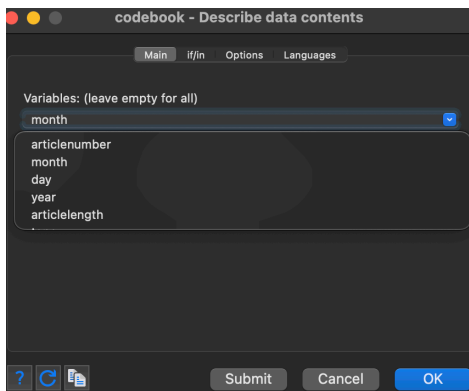
### 3 Obtaining Categories and Codes for Your Variables

To complete part one of the assignment, you will need the categories and codes (i.e. values) for each of the variables you select.

Begin by choosing the “Data” drop-down menu from the top of the screen, and then select “describe data” and “describe data contents (codebook)”:



This will bring up a “codebook” window, where you can select the variables you are interested in examining from the “variables” drop-down menu (you can select multiple variables). Press “ok” when you have selected them:



In the results window, you will see the details of your variable(s). You can use this to choose your variables, and you can record this information in order to complete part one of the assignment:

```
. codebook month
```

---

month	Month
-------	-------

---

```

      Type: Numeric (byte)
      Label: month

      Range: [1,12]
      Unique values: 12
      Examples: 3    March
                6    June
                8    August
                11   November

      Units: 1
      Missing .: 0/359
```

## 4 Saving Your Commands

You should save a record of all of the things you have asked Stata to do for you by selecting and copying all of the commands from the “history” window and pasting them at the end of your assignment:

History		
	Command	_rc
1	use "/Users/steph...	
2	describe taxes	
3	codebook day	
4	codebook tone	
5	codebook tone	
6	by year, sort : tabu...	
7	codebook month	
9	tabulate ht_regtyp...	
10	tabulate fh_status...	

At a later time, if you would like to ask Stata to execute a command again, all you need to do is copy the command from your assignment, paste the command into the command window, and press return:

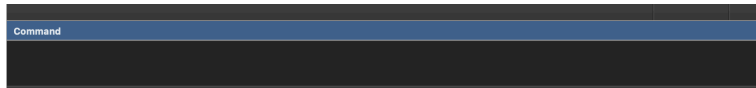
Command
codebook month



## 5 Cross-tabulation

For the assignment you will need to generate *contingency tables* (also known as *cross-tabulations*) and appropriate *measures of association* and *statistical significance*. By examining a cross-tabulation you will be able to see whether a relationship exists between two variables in your data.

Begin by entering the command to generate your table in the command window:



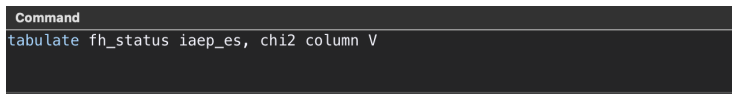
The command should take the following form:

```
tabulate DEPVAR INDVAR, chi2 column V
```

Where DEPVAR is the name of your dependent variable, and INDVAR is the name of your independent variable.

tabulate asks Stata to generate a table for you, while chi2 asks Stata to calculate the Chi-square statistic and its p-value (which tells you the statistical significance), column asks Stata to calculate column percentages, and V asks Stata to calculate the Cramer's V measure of association (NOTE: for 2x2 tables, V is equal to Phi).

Here is an example using different data:



Here, I have asked Stata to generate a table with the dependent variable fh\_status, measuring the level of freedom in different countries, and the independent variable iaep\_es, which measures the type of electoral system.

Here are the results:

```
. tabulate fh_status iaep_es, chi2 column V
```

Key
<i>frequency</i> <i>column percentage</i>

Freedom Status	Electoral System				Total
	Plurality	Majority	Proportio	Mixed sys	
1	229 27.89	315 21.52	888 58.00	539 49.40	1,971 40.17
2	326 39.71	473 32.31	477 31.16	412 37.76	1,688 34.40
3	266 32.40	676 46.17	166 10.84	140 12.83	1,248 25.43
Total	821 100.00	1,464 100.00	1,531 100.00	1,091 100.00	4,907 100.00

Pearson chi2(6) = 778.4146   Pr = 0.000  
Cramér's V = 0.2816

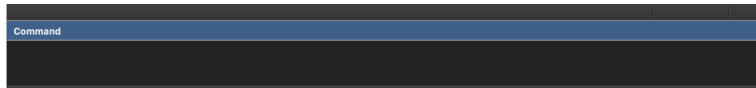
The table reports both the frequency and percentage for each cell. Below the table, the Chi-square and its statistical significance (the p-value, called Pr in Stata, which is 0.000) and Cramer's V (0.2816) are reported.

Repeat this set of commands for both independent variables you have selected. Record the table data in your assignment.

## 6 Control Tables

For the assignment you will also need to generate *control tables* and appropriate *measures of association* and *statistical significance*. By examining a these you will be able to see whether a relationship exists between two variables in your data after controlling for one of the independent variables.

Begin by entering the command to generate your table in the command window:



The command should take the following form:

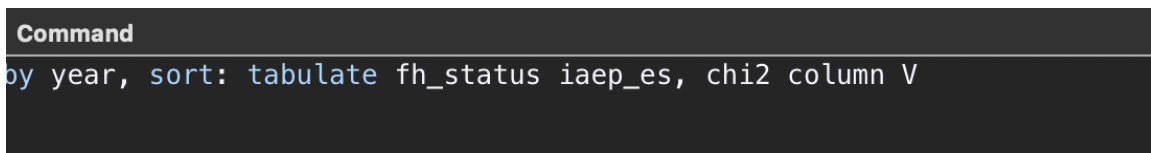
by CONTROL, sort: tabulate DEPVAR INDVAR, chi2 column V

Where DEPVAR is the name of your dependent variable, and INDVAR is the name of your independent variable, and CONTROL is the name of the other independent variable you have decided to treat as your control variable.

tabulate asks Stata to generate a table for you, while chi2 asks Stata to calculate the Chi-square statistic and its p-value (which tells you the statistical significance), column asks Stata to calculate column percentages, and V asks Stata to calculate the Cramer's V measure of association (NOTE: for 2x2 tables, V is equal to Phi).

by CONTROL, sort: asks Stata to generate separate tables for each category of your control variable.

Here is an example using different data:



Here, I have asked Stata to generate a table with the dependent variable fh\_status, measuring the level of freedom in different countries, and the independent variable iaep\_es, which measures the type of electoral system. I am controlling a third variable, year.

Here are the results:

-> year = 2011

Key
<i>frequency</i>
<i>column percentage</i>

Freedom Status	Electoral System				Total
	Plurality	Majority	Proportio	Mixed sys	
1	6 33.33	6 17.65	31 54.39	17 39.53	60 39.47
2	6 33.33	11 32.35	17 29.82	19 44.19	53 34.87
3	6 33.33	17 50.00	9 15.79	7 16.28	39 25.66
Total	18 100.00	34 100.00	57 100.00	43 100.00	152 100.00

Pearson chi2(6) = 20.9497 Pr = 0.002  
Cramér's V = 0.2625

-> year = 2012

Key
<i>frequency</i>
<i>column percentage</i>

Freedom Status	Electoral System				Total
	Plurality	Majority	Proportio	Mixed sys	
1	6 33.33	6 17.65	31 55.36	17 43.59	60 40.82
2	6 33.33	12 35.29	15 26.79	15 38.46	48 32.65
3	6 33.33	16 47.06	10 17.86	7 17.95	39 26.53
Total	18 100.00	34 100.00	56 100.00	39 100.00	147 100.00

Pearson chi2(6) = 17.1466 Pr = 0.009  
Cramér's V = 0.2415

There is a table for each category of the control variable. In this case, there are only two tables. Each table reports both the frequency and percentage for each cell. Below each table, the Chi-square and its statistical significance and Cramer's V are reported.

Record the table data in your assignment.

## 7 Reminder: Saving Your Commands

You should save a record of all of the things you have asked Stata to do for you by selecting and copying all of the commands from the “history” window and pasting them at the end of your assignment:

History		
	Command	_rc
1	use "/Users/steph...	
2	describe taxes	
3	codebook day	
4	codebook tone	
5	codebook tone	
6	by year, sort : tabu...	
7	codebook month	
9	tabulate ht_regtyp...	
10	tabulate fh_status...	

At a later time, if you would like to ask Stata to execute a command again, all you need to do is copy the command from your assignment, paste the command into the command window, and press return:

Command
codebook month