

**SSRIC Teaching Resources**  
**Introduction to the Study of War**  
**Cynthia McMeekin, California State University, Los Angeles**  
**and Emily Acevedo, California State University, Los Angeles**

**Table of Contents**

|  |    |
|--|----|
| 1. Introduction .....  | 2  |
| 2. What is war? .....  | 2  |
| a. Interstate war .....  | 3  |
| b. Civil war .....   | 3  |
| c. Nonstate wars .....   | 4  |
| d. Descriptive statistics .....  | 4  |
| e. Exercise one – Descriptive statistics in SPSS .....   | 8  |
| 3. Levels of Analysis .....  | 16 |
| a. Individual actors and war .....   | 17 |
| b. Groups and war .....  | 17 |
| c. Recoding variables .....  | 17 |
| d. Exercise two – Recoding variables in SPSS .....   | 18 |
| 4. Levels of Analysis .....  | 21 |
| a. States and war .....  | 21 |
| b. The international system and war .....  | 21 |
| c. Creating composite variables .....  | 22 |
| d. Exercise three – Creating composite variables in SPSS .....   | 23 |
| 5. Theories of war .....   | 25 |
| a. Democratic Peace Theory .....   | 25 |
| b. Power Transition Theory .....   | 25 |
| c. Testing theories of war through statistical analyses .....  | 26 |
| d. Testing for association of variables with a T-test, One-way ANOVA,<br>Correlation, or Chi-Square test ..... | 27 |
| e. Exercise four – Testing association of variables.....   | 29 |
| 6. Conclusion.....   | 31 |
| 7. References .....  | 32 |

## **I. Introduction**

War is a dominant subject in the field of international relations. The systematic study of war dates back to Thucydides observations of the Peloponnesian War in the fifth century B.C. Though war is a rare event, when it does occur, it devastates lives, uproots people from their homes, and results in the deaths of combatants and noncombatants. In part, the negative impact of war has driven scholars to continuously study war; because despite the fact that all wars impart death and destruction, war is a consistent feature of human history. Scholars study the general causes of war, strategies of war, the psychology of war, patterns of war, how wars end, the intensity of war, the timing of wars, and a myriad of other topics related to war.

This module is designed to introduce students to the study of war and basic statistics. No prior knowledge of statistics is necessary. This module serves as a primer in using data to analyze war. It uses an integrative approach that allows students to practice analyzing war data as they learn about how war is defined, the levels of analysis used in the study of war, and what major theories of international relations have to say about war. After completing this module, students should have a general understanding of descriptive statistics, including measures of central tendency and variance. They will learn how to create and recode variables in SPSS and be able to create composite variables. It also shows them how to test theories of war to determine which theory is better supported by the data.

## **II. What is war?**

In the year 2001, after the 9/11 terrorist attacks on the World Trade Center and the Pentagon, President George W. Bush gave the Taliban government in Afghanistan an ultimatum. Meet the demands of the U.S. or suffer the same fate as the terrorists. Bush demanded that the government of Afghanistan: 1. Hand over all leaders of al Qaeda, the terrorist network deemed to be responsible for the 9/11 attacks; 2. Release foreign prisoners; 3. Protect journalists and aid workers; 4. Shut down all terrorist training camps and turn over all known terrorists and their supporters to the U.S.; and 5. Allow U.S. access to terrorist camps (Bush, September 20, 2001). The Taliban government refused the demands of President Bush and the United States, and their allies began a bombing campaign against Afghanistan on October 7, 2001. This attack initiated the U.S. War on Afghanistan.

This attack and subsequent actions in Afghanistan are consistently referred to as the Afghanistan War by journalists, politicians, and American citizens alike. American and allied soldiers were sent to fight the Taliban in Afghanistan. Between the years of 2001 and 2016, there have been 110,166 battle related deaths on all sides of the conflict (Lacina and Gleditsch, 2005). Based on U.S. government estimates, The Afghani government that is supported by the U.S. now controls approximately 60% of the territory within the country with the rest either contested or under control of insurgents such as the Taliban. The fighting that began in 2001 continues today. U.S. soldiers began fighting against the Taliban government then against the Taliban insurgents and other insurgent groups. However, the United States Congress has not has not declared war in Afghanistan. For the past 17 years, American troops have been sent to Afghanistan, despite the

absence of a formal declaration of war by the U.S. Congress. This brings up the question, what is war? Clearly, the conflict in Afghanistan meets some definition of war, despite it not having been officially declared a war by Congress. This section will discuss three different classifications of war commonly used in international relations scholarship.

The study of war is made easier with a shared definition of the concept so that empirical studies of conflict can be easily compared with each other. While there is not a universal definition of war, there are definitions that are commonly used in the field of international relations. In particular, the dataset that will be used in this module's exercises presents commonly utilized definitions of interstate wars, civil war (intrastate war), and non-state wars. The definitions presented in this section will pay particular attention to how the Correlates of War (COW) data project defines these distinct classifications of war.

**Interstate war** is often what first comes to mind when posed with the question, what is war? Some of the most devastating and memorable wars in recent history are interstate wars, such as World War I and World War II. An interstate war includes the participation of two or more states recognized in the international system. When states have diplomatic disagreements, there is the possibility that the discord will escalate to physical combat between states, but what level of violence meets the definition of war? In 1962, the United States discovered that the Soviet Union placed missiles on the island of Cuba capable of striking the US. This discovery created a diplomatic crisis for the US and the Soviet Union and brought the two countries to the brink of nuclear war. The crisis lasted for thirteen days and during the heightened tensions, the Soviet Union shot down a U.S. spy plane flying over Cuba, killing one American soldier. Though the crisis escalated to physical combat and death of a combatant related to the crisis, war was averted. A single act of war does not necessarily lead to war. While any specific level of violence indicating war between states may seem arbitrary, Small and Singer (1982) established a now generally accepted number of 1,000 battle related deaths within a twelve-month period to define war between two or more countries. The combatants must engage in sustained combat through organized armed forces. In cases such as World I and World War II, where multiple states are included, an individual member state is classified as a war participant by meeting either of two alternative criteria: a minimum of 100 fatalities or a minimum of 1,000 armed personnel engaged in active combat (Small and Singer, 1982, p. 56). This commonly accepted definition of interstate war in the field of international relations has helped with data collection and has facilitated comparison of research that focuses on interstate war.

**Civil war (intrastate war)** is war between a state government and an organized group within the state's borders. Small and Singer (1982) also develop a standard definition of civil war that is frequently used in research conducted by scholars who study war. Civil war must meet some of the same criteria as an interstate war. There must be 1,000 battle deaths within 12 months through sustained combat committed by organized armed forces. To distinguish civil war from a one-sided mass killing, genocide, severe government repression, or unorganized violent protest, the weaker side must demonstrate effective resistance. Small and Singer describe effective resistance as both sides being organized for violence and having at least 5% of the casualties. For example, if the government kills 1000 members of a group they are "at war" with, the

government must have incurred 50 casualties in order to demonstrate effective resistance of the group fighting the government. El Salvador experienced a civil war from 1979 – 1992 that meets the above criteria.

Civilians in El Salvador suffered significant state repression for most of the twentieth century. From the beginning of the century, the military and the economic elite have strongly influenced who governs El Salvador. The alliance generally protected the status quo and thwarted any attempts to alter the economic policies and political leadership by leftists and moderates alike. Rather than opening political space for opposing views, the military regularly invoked repression to quiet the opposition. Over time, there was a break between hardliners and reformers within the military. A military coup in 1979 sought to implement reform by ousting the current leaders and instituting a reform-minded military junta. The junta promised to address human rights abuses, confront political corruption, and address problems of economic disparity (Mazzei, 2009). Though the coup was successful, the resultant struggle for leadership within the military junta altered the directive of the new government and ended up increasing repression, and despite limited land reform, generally maintained status quo policies that benefited the existing elite. Following the coup, the extreme right-wing sector of the economic elite developed a political-military strategy that entailed the use of paramilitary forces in conjunction with a political party (Montgomery, 1996). As the right ramped up the violence, leftist rebels responded with violent acts of their own until low level violence escalated to civil war.

**Nonstate war** involves fighting between two or more non-state armed groups, that are not only organized for war, but can demonstrate effective resistance, and commit at least 100 troops to the war or suffer 25 battle-related deaths (Small and Singer 1982). Nonstate wars can be fought within a state or across state boundaries. A nonstate actor (NSA) can be deemed a war participant if it either commits 100 armed personnel to the war or suffers 25 battle-related deaths. What distinguishes nonstate wars from interstate and civil wars is that this conflict is waged between nonstate actors. Civil wars can also include NSAs in a conflict such as the current Syrian civil war that includes other NSAs such as The Islamic State of Iraq and the Levant, also known as the Islamic State of Iraq and Syria (ISIS).

An example of a nonstate war is the war between Jewish and Arab communities in Palestine from 1947-1948. The war began in response to the adoption of the United Nations General Assembly Resolution calling for the creation of “Independent Arab and Jewish States and the Special International Regime for the City of Jerusalem” (A/RES/1811 (II)). According to Morris (1982), the fighting was initiated by Arab communities targeting Jewish populations that eventually spiraled into militarized campaigns by both communities.

### **Descriptive Statistics**

Statistics provides a useful tool with which to study war. Descriptive statistics summarize the information, central tendencies and variance, on a single variable. This section will describe measures of central tendency and variance. Examining the central tendencies and variance of individual variables helps to reveal features of data that are not obvious from looking at an entire dataset. The next section asks the student to use SPSS, the statistical software package, to run

descriptive statistics to analyze interstate war data. The Correlates of War Project has collected and stored information on interstate wars between the years of 1816 and 2007. The interstate war dataset consists of a number of variables. For example, one variable, "WhereFought," provides information on what region or regions a war was fought. The regions include: The Western Hemisphere, Europe, Africa, Middle East, Asia, Oceania, or some combination thereof. Each region and combination of regions is represented by a unique number. For example, The Western Hemisphere is represented by a 1, Europe is represented by a 2, etc.

Each variable within the dataset can be labeled a nominal variable, an ordinal variable, or an interval variable. **Nominal variables** are categorical and have no numeric order to them. For example, the region in which a war is fought is a nominal variable. There is no numeric or rank order to the regions. They simply categorize where the war took place. **Ordinal variables**, on the other hand, can be rank ordered. To demonstrate, perhaps a study is examining whether wars are shorter in length when a neutral third party intervenes to help negotiate a settlement between warring parties. A dataset might include a variable regarding third party interventions in the settlement of wars. The variable might be coded 0 for no intervention, 1 for an ineffective intervention, 2 for a partially effective intervention, and 3 for an effective intervention by a neutral third party. The categories of the variable can clearly be ranked from no intervention to an effective intervention. The distance between the variables, however, is not uniform. For instance, the distance between no intervention and an ineffective intervention is not necessarily the same as the distance between an ineffective intervention and a partially effective intervention. The distance between these categories cannot be accurately assessed. **Interval variables** can be rank ordered and have values of equal intervals that have some meaning. You can think of interval variables as "count variables." In the interstate war data that will be examined in Exercise One, the variable "BattleDeaths" is a record of all battle related deaths that occurred during the war. Each number represents one death, and the distinction between 5 battle deaths and 6 battle deaths is meaningful.

Knowing the difference between nominal, ordinal, and interval variables is important when using descriptive statistics because the type of variable will indicate which measures of central tendency are most appropriate. There are three measures of central tendency: mode, median, and mean.

**Mode** is a measure of central tendency that identifies the category with the most observations. The mode can be assessed for nominal variables, ordinal variables, and interval variables. For example, consider the variable that records the region in which wars were fought. The mode for this variable would be the region that has experience the most wars. The mode for this nominal variable can be assessed.

**Median** is another measure of central tendency. It represents the middle value. Half of the observations are above that value and half are below. The median can be found for ordinal and interval variables. If the median value of battle deaths for wars between 1816 and 2007 were 100,000 battle deaths, half of the wars would have battle deaths fewer than 100,000 and half of the wars would have more than 100,000 battle deaths.

Finally, the **mean** value measures the average of all values for a variable and can only be calculated for interval variables. The mean is calculated by summing the values of several observations and dividing by the number of observations. If the mean number of battle deaths for wars between 1816 – 2007 were 4,500, that would be the average number of deaths in each war.

Measures of central tendency and variance work together to provide a complete picture of any variable. Variance measures the dispersion of the variable, or the distance of the observations from a measure of central tendency. For example, consider again the variable that describes the region in which each war was fought. The mode identifies the region with the most wars fought. A basic measure of dispersion is the **percentage of the observed values** that are in the modal category. The smaller the percentage, the greater the dispersion of the variable across categories. If the modal category has all the observations, then this variable has the least amount of dispersion; and if the observations are spread out equally among categories, the variable would have the greatest amount of dispersion. If all wars were fought exclusively in Europe, the variable is not dispersed, and if all regions had an equal number of wars, then this variable is greatly dispersed. This measure of variance is the only measure of variance for nominal variables.

Another measure of variance is the **range** of a variable. The range measures the distance between the value of the highest observation and the lowest observation. If the lowest number of battle deaths for any war is 1,000 and the highest number of deaths is 500,000, then the range of this variable is  $500,000 - 1,000 = 499,000$ . Range can be assessed for ordinal and interval variables. F

Finally, the **standard deviation** measures the distance each observation, on average, differs from the mean. In other words, this measure describes how spread out the observations are from the average observation. Figure 1 below shows three distributions of different variables. The blue variable has the greatest distribution and therefore the highest standard deviation from the mean. The observations for the red variable are much more closely centered around the mean and represents the lowest measure of standard deviation in the figure. This measure of dispersion is appropriate for interval level variables.

### Graphical Representations of Data Spread

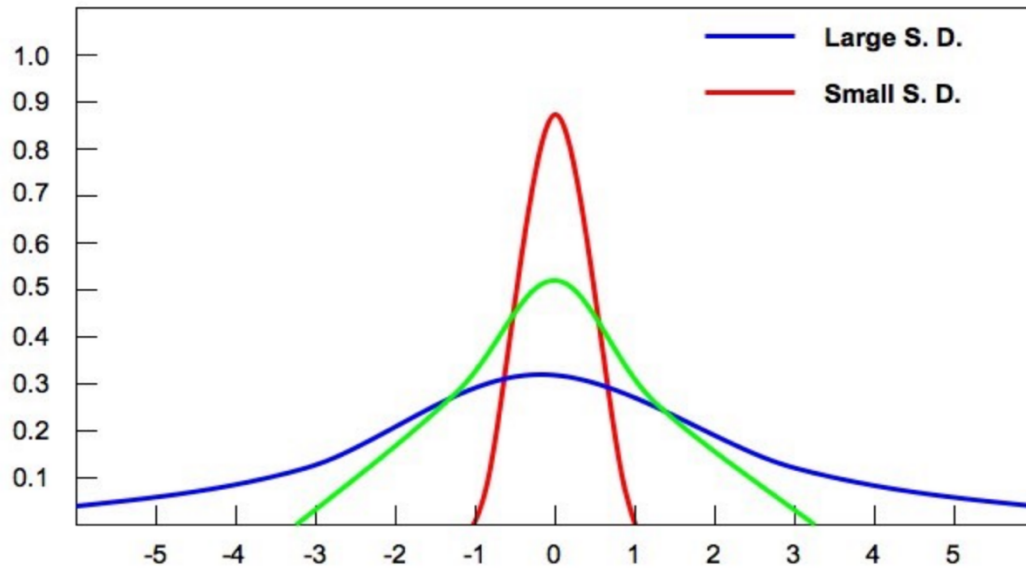


Figure 1

Source: Bioninja: Your one stop biology resource. <http://www.old-ib.bioninja.com.au/standard-level/topic-1-statistical-analysis/11-statistical-analysis.html>

Measures of central tendency and variance help to reveal the characteristics of variables within a dataset. Exercise One, below, provides instructions on generating descriptive statistics in SPSS. In particular, the instructions indicate how to produce frequency tables, histograms, and bar charts. These tables and graphs offer information regarding the measures of central tendency and variance discussed above. A frequency table offers information about the number of observations and percent on each value. A bar chart is a series of bars in which each bar represents a number or percentage of observations in a particular category. A histogram is a type of bar chart where the height of the bars is proportional to the frequency of an observations and appropriate for use with interval variables. Please follow the instructions and answer the questions in the exercise based on your reading of this module, the codebook for the dataset on Interstate War, and the tables and graphs generated through SPSS.

## Exercise One: Descriptive Statistics in SPSS

For these examples, use the Correlates of War database, COW War Data, 1816 - 2007 (v4.0): Inter-state War.

**1. Read the codebook for the Inter-State War** data available on the COW website: <http://cow.dss.ucdavis.edu/data-sets/COW-war>.

- a. Download the data on interstate war from the website above and save it as an Excel file.
- b. Open SPSS. **File** > **Import Data** > **Excel**
- c. Choose your saved Excel File and the dataset will open in the data editor.
- d. Below you will find several ways to practice, using this dataset, with descriptive statistics output from SPSS.

### 2. Frequency Distribution: Table Form

In this example, you enter four variables at once to generate a frequency table. In general, for generating descriptive statistics it is efficient to run several variables at once. However, you probably will not use all the descriptive statistics generated in such a procedure.

**Analyze** > **Descriptive Statistics** > **Frequencies**

Select, Initiator, Outcome, WhereFought, BatDeath and move them over in the “Variables” box. Click **Statistics**.

Select “Mean”, “Std. Deviation”, and “Range”. **Click Continue** > **OK**.

You should get the following output in your Output window.



# Frequencies

## Statistics

|                |         | Initiator | Outcome | WhereFought | BatDeath   |
|----------------|---------|-----------|---------|-------------|------------|
| N              | Valid   | 337       | 337     | 337         | 337        |
|                | Missing | 0         | 0       | 0           | 0          |
| Mean           |         | 1.68      | 2.09    | 4.93        | 95195.55   |
| Std. Deviation |         | .468      | 1.528   | 3.476       | 501199.696 |
| Range          |         | 1         | 7       | 18          | 7500009    |

- Having read the codebook for the interstate war dataset, are the Initiator, Outcome, WhereFought, and BatDeath variables nominal, ordinal, or interval variables?
  - Initiator \_\_\_\_\_
  - Outcome \_\_\_\_\_
  - WhereFought \_\_\_\_\_
  - BatDeaths \_\_\_\_\_
- For which of these variables is the mean an appropriate measure of central tendency?
 

\_\_\_\_\_
- Examine the standard deviation for Battle Deaths. Would you say the data are very dispersed or centered closely around the mean?
 

\_\_\_\_\_
- Explain the meaning of the number 95,195.55 in the table above.
 

\_\_\_\_\_

\_\_\_\_\_

## Frequency Table

|       |       | Initiator |         |               |                    |
|-------|-------|-----------|---------|---------------|--------------------|
|       |       | Frequency | Percent | Valid Percent | Cumulative Percent |
| Valid | 1     | 109       | 32.3    | 32.3          | 32.3               |
|       | 2     | 228       | 67.7    | 67.7          | 100.0              |
|       | Total | 337       | 100.0   | 100.0         |                    |

|       |       | Outcome   |         |               |                    |
|-------|-------|-----------|---------|---------------|--------------------|
|       |       | Frequency | Percent | Valid Percent | Cumulative Percent |
| Valid | 1     | 155       | 46.0    | 46.0          | 46.0               |
|       | 2     | 119       | 35.3    | 35.3          | 81.3               |
|       | 3     | 4         | 1.2     | 1.2           | 82.5               |
|       | 4     | 28        | 8.3     | 8.3           | 90.8               |
|       | 6     | 30        | 8.9     | 8.9           | 99.7               |
|       | 8     | 1         | .3      | .3            | 100.0              |
|       | Total | 337       | 100.0   | 100.0         |                    |

5. Look at the frequency table for the variable “Outcome.” Using the interstate war codebook, identify what the numbers 1, 2, 3, 4, 6, and 8 signify.

a. 1: \_\_\_\_\_

b. 2: \_\_\_\_\_

c. 3: \_\_\_\_\_

d. 4: \_\_\_\_\_

e. 6: \_\_\_\_\_

f. 8: \_\_\_\_\_

6. Identify the mode for the “Outcome” variable.

\_\_\_\_\_

7. Identify the mode for the “WhereFought” variable. Be sure to identify the region and not just the number that represents the region.

\_\_\_\_\_

8. Examine your frequency table for the BatDeaths variable. Identify the median number of battle deaths in your output.

\_\_\_\_\_

### WhereFought

|       |       | Frequency | Percent | Valid Percent | Cumulative<br>Percent |
|-------|-------|-----------|---------|---------------|-----------------------|
| Valid | 1     | 35        | 10.4    | 10.4          | 10.4                  |
|       | 2     | 110       | 32.6    | 32.6          | 43.0                  |
|       | 4     | 16        | 4.7     | 4.7           | 47.8                  |
|       | 6     | 62        | 18.4    | 18.4          | 66.2                  |
|       | 7     | 88        | 26.1    | 26.1          | 92.3                  |
|       | 11    | 10        | 3.0     | 3.0           | 95.3                  |
|       | 12    | 1         | .3      | .3            | 95.5                  |
|       | 13    | 2         | .6      | .6            | 96.1                  |
|       | 14    | 4         | 1.2     | 1.2           | 97.3                  |
|       | 15    | 3         | .9      | .9            | 98.2                  |
|       | 16    | 2         | .6      | .6            | 98.8                  |
|       | 17    | 1         | .3      | .3            | 99.1                  |
|       | 18    | 2         | .6      | .6            | 99.7                  |
|       | 19    | 1         | .3      | .3            | 100.0                 |
|       | Total | 337       | 100.0   | 100.0         |                       |

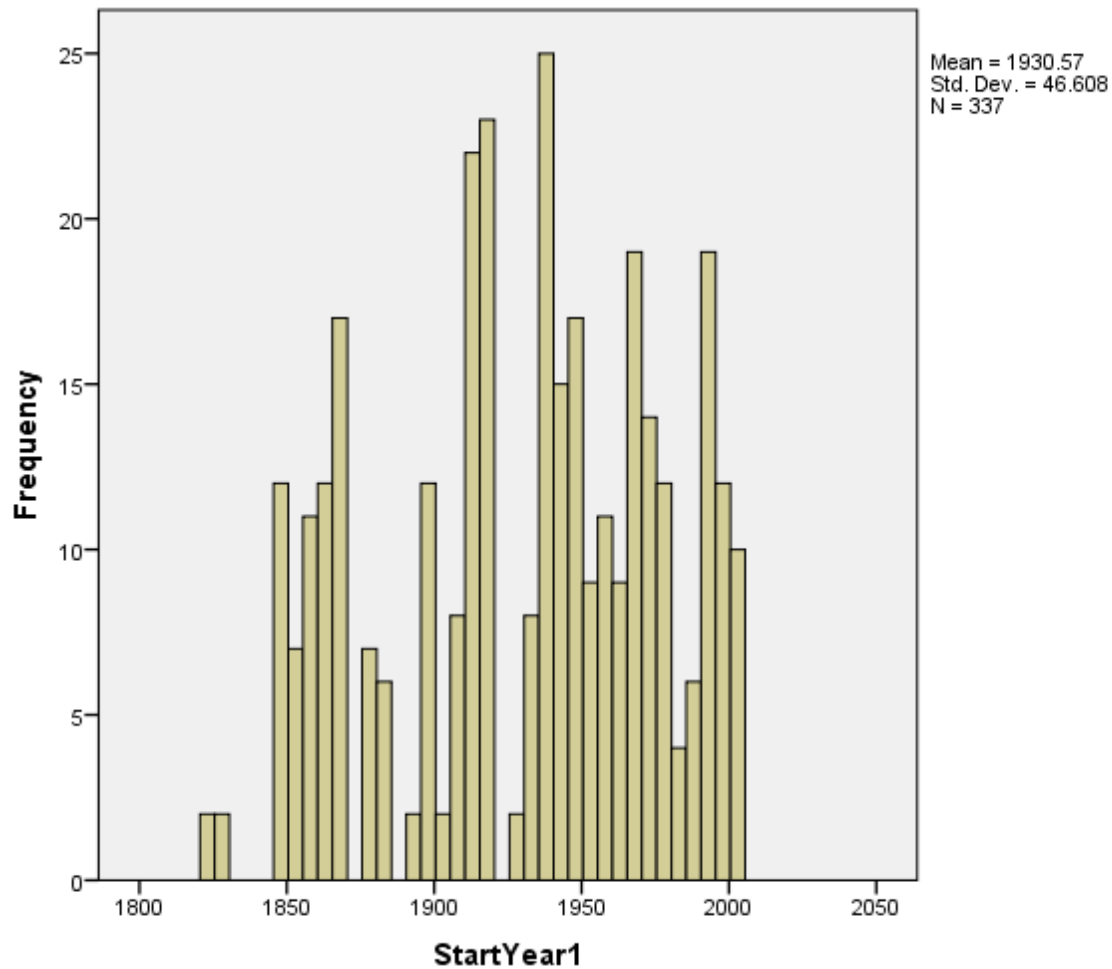
### BatDeath

|       |     | Frequency | Percent | Valid Percent | Cumulative Percent |
|-------|-----|-----------|---------|---------------|--------------------|
| Valid | -9  | 1         | .3      | .3            | .3                 |
|       | 0   | 18        | 5.3     | 5.3           | 5.6                |
|       | 2   | 3         | .9      | .9            | 6.5                |
|       | 6   | 1         | .3      | .3            | 6.8                |
|       | 10  | 1         | .3      | .3            | 7.1                |
|       | 14  | 1         | .3      | .3            | 7.4                |
|       | 21  | 1         | .3      | .3            | 7.7                |
|       | 22  | 1         | .3      | .3            | 8.0                |
|       | 23  | 1         | .3      | .3            | 8.3                |
|       | 24  | 2         | .6      | .6            | 8.9                |
|       | 33  | 1         | .3      | .3            | 9.2                |
|       | 34  | 1         | .3      | .3            | 9.5                |
|       | 44  | 1         | .3      | .3            | 9.8                |
|       | 79  | 1         | .3      | .3            | 10.1               |
|       | 92  | 1         | .3      | .3            | 10.4               |
|       | 97  | 1         | .3      | .3            | 10.7               |
|       | 100 | 14        | 4.2     | 4.2           | 14.8               |
|       | 111 | 1         | .3      | .3            | 15.1               |
|       | 114 | 1         | .3      | .3            | 15.4               |
|       | 120 | 1         | .3      | .3            | 15.7               |
|       | 122 | 1         | .3      | .3            | 16.0               |
|       | 125 | 1         | .3      | .3            | 16.3               |
|       | 140 | 2         | .6      | .6            | 16.9               |
|       | 150 | 1         | .3      | .3            | 17.2               |
|       | 169 | 1         | .3      | .3            | 17.5               |

### 3. Graphical Form

#### Histogram

**Graphs** > **Legacy Dialogs** > **Histogram**. Move **StartYear1** into “Variable” box, and click **OK**. You should get a histogram like the one below in you Output window.



9. Between 1800 and 2012, identify the most peaceful 50 years based on the data provided in the histogram you have created.

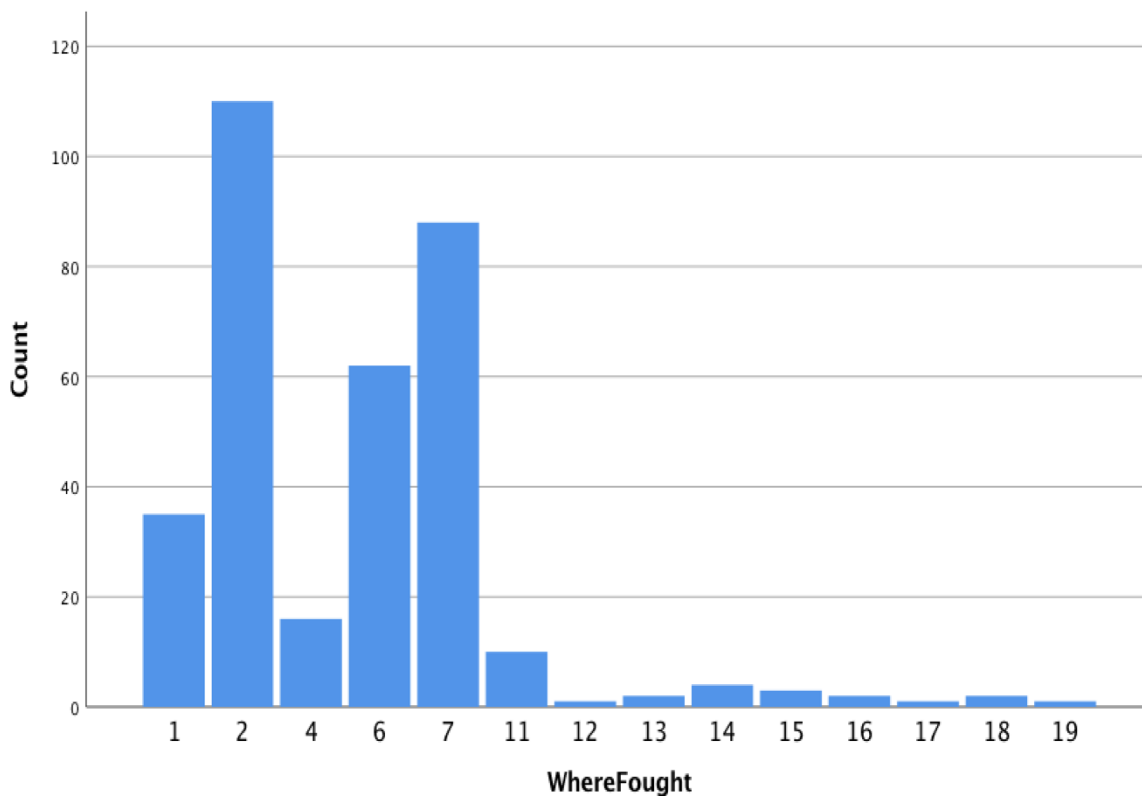
---

#### 4. Bar Graph

**Graphs** > **Legacy Dialogs** > **Bar**. The default is “Simple.” Click **Define**. Move WhereFought over to “Category Axis.” Under “Bars Represent,” choose either “N of cases” for a frequency or “% of Cases” for percent.

#### Graph

[DataSet1] /Users/cynthiamcmeekin/Documents/Data Grant/COW Data/Inter-StateWarData.sav



10. Look at the bar graph. Identify the three regions that have experienced the most war between 1816 and 2007.

- a. \_\_\_\_\_
- b. \_\_\_\_\_
- c. \_\_\_\_\_

You have many other options for graphs, too. Experiment.

### III. Levels of Analysis: Individuals and Groups

The previous section defined different types of wars. This section will explore different levels of analysis through which scholars view war. A researcher must determine whether they will assess the international system as a whole, relationships between parts of the international system, or parts of the international system in particular (Gebhard, 2017). As mentioned above, researchers study many different aspects of war. As we consider different levels of analysis, we will focus on the general question, what causes war? This section will closely consider four different levels of analysis: the individual, group, state, and the international system.

In an attempt to interpret world politics and explain phenomena, IR scholars typically focus on three levels of analysis – Individual, State, Systemic. Kenneth Waltz (1954) in *Man, State and War*, explores what causes war through the exploration of three “images,” noting that most scholars attributed war to either individuals, domestic structures, or the international system itself. However, one level of analysis that is often excluded by IR scholars is the group level of analysis. Groups are important and should be acknowledged in light of the increased focus on non-state actors that are able to operate across state borders and affect states, such as drug trafficking organizations, human trafficking, or terrorism.

The individual level of analysis is concerned with human nature and finds the source of war is rooted in the aggressive nature of man (Hobbes, 1651). The key actors examined at this level of analysis are those key leaders who exhibit influence and power and can influence international events. Once these key individuals are identified, it is then necessary to assess factors that affect their goals, thought processes, and decision-making processes.

The group level of analysis involves the study of group dynamics, such as intra- and intergroup conflict and cohesion, leadership, power, and norms. This group level of analysis could include nonstate actors, social movements, or interest groups, examining “how they relate to the state level and where they position themselves with respect to the global dimension of the issues they are dealing with” (Gebhard, 2017: 35).

The state level of analysis examines the nature of the state and finds the cause of war is rooted in the internal organization of domestic structures and politics of individual states (Wright, 1942). These domestic variables – government type, economy, population, resources, and stability - affect states' behavior at the international level. Many different theories within IR view the state as a central unit of analysis when examining state behavior, although each holds different interpretations for why states engage in war.

The systemic level of analysis focuses on the international system and sees the source of conflict stemming from the nature of the international system. The structure of the international system is seen as anarchic since there is no world government to restrict states. This system is made up of independent states.



## **Individual Actors and War**

The individual level of analysis traces the cause of war to human nature. The emphasis on individuals attempts to explore how individuals matter and how much difference one leader make in foreign policy decisions. IR scholars tend to study specific elements of human nature such as personality, perceptions, choices and activities of decision makers. These factors help one understand not only who the person is, but how that person makes decisions.

## **Groups and War**

Groups are especially important in the study of war. The group level of analysis is not focused exclusively on violent groups. When examining the decisions of a state (either foreign or domestic policy), using a group level of analysis may shed light on why a state chose a particular policy. For instance, if one is exploring the United States decision to invade Iraq in 2003, a group level analysis may illustrate that powerful interest groups, such as American oil companies, lobbied the state to ensure that their interests are protected.

However, in the Correlates of War project, non-state actors are those engaged in conflict, either with one another or with a state. The Islamic State of Iraq and the Levant, also known as the Islamic State of Iraq and Syria (ISIS), is a central actor in several conflicts throughout the Middle East; yet because it is not a recognized state, it is hard to explore this actor using the individual or state level of analysis. In contrast, a group level analysis of ISIS in Syria might provide a critical understanding of the issues relevant to ISIS. It might also shed light on how this group's actions influence a state's decision to either engage or not engage in the Syrian conflict.

## **Recoding variables**

Sometimes the numbers in an original dataset need to be changed in order to conduct your analysis. You can modify data values by recoding them. This is often useful when you are trying to collapse or merge categories. For example, when looking at a measure of Gross Domestic Product (GDP) for states, the dataset may include ranges of values that you would like to collapse into one category. If two ranges in the dataset are 1 :  $\leq 10$  billion dollars and 2 : 10 billion - 20 billion dollars, you may want to recode this so that 1 :  $\leq 20$  billion.

There are two ways that you can recode variables in SPSS. The first way is to recode the values within an existing variable, and the other way is to create a new variable based on the recoded values of an existing variable. It is highly recommended that you recode into different variables, which leaves your original variable unchanged. It is always best to leave your original data intact, in the event that you make an error in recoding or if you decide against recoding the variable.

In the following exercise, you will learn to use the COW dataset on nonstate wars to recode variables.

## Exercise Two: Recoding Variables in SPSS

For these practice exercises, please use the Correlates of War database, COW War Data, 1816 - 2007 (v4.0): Non-State Wars.

### 1. Read the codebook for the Non-State War data available on the COW website:

<http://cow.dss.ucdavis.edu/data-sets/COW-war>.

- Download the data on Non-State Wars from the website above and save it as an Excel file.
- Open SPSS. **File** > **Import Data** > **Excel**
- Choose your saved Excel File and the dataset will open in the data editor.
- Below you will find several ways to practice, using this dataset, with descriptive statistics output from SPSS.

### 2. Recoding Variables in SPSS

The original variable “WhereFought” consists of six different regions indicating what region a war was fought. For the purpose of this exercise, you will recode “WhereFought” so that only two regions will be reflected: West or NonWest. The original values of 1 (Western Hemisphere) and 2 (Europe) under “WhereFought” will be recoded to 1 indicating West, while the remaining values of 4 (Africa), 6 (Middle East), 7 (Asia) and 9 (Oceania) will be recoded to 2 indicating NonWest.

Follow these instructions to recode the variable “WhereFought”

- Click on **Transform** > **Recode into Different Variables**
- Select “WhereFought” and move it to the Input Variable -> Output Variable box
- In the Output Variable area, give your recoded variable the name “WestNonWest” then click Change
- Click the **Old and New Values** button.
- Click on Range under Old Values on the left side and enter 1 through 2, then under New Values on the right side, enter 1, then click on “Add”
- Click on All Other Values under Old Values on the left side, then under New Values on the right side, enter 2, then click on “Add” **Click Continue** > **OK**.

Your newly recoded variable, **WestNonWest**, should now appear in the last column in data view and the last row in variable view.

Now check to see that your recode was successful by comparing the frequencies on your new and old variables.

**Analyze** > **Descriptive Statistics** > **Frequencies**. Move both your old and newly recoded variables over into the “Variables” section, click **OK**.

In your SPSS Output window, you should see output like that below.

## Frequencies

|                |         | Statistics  |             |
|----------------|---------|-------------|-------------|
|                |         | WhereFought | WestNonWest |
| N              | Valid   | 62          | 62          |
|                | Missing | 0           | 0           |
| Mean           |         | 4.68        | 1.7581      |
| Std. Deviation |         | 2.553       | .43175      |
| Range          |         | 8           | 1.00        |

## Frequency Table

|       |       | WhereFought |         |               | Cumulative |
|-------|-------|-------------|---------|---------------|------------|
|       |       | Frequency   | Percent | Valid Percent | Percent    |
| Valid | 1     | 15          | 24.2    | 24.2          | 24.2       |
|       | 4     | 19          | 30.6    | 30.6          | 54.8       |
|       | 6     | 5           | 8.1     | 8.1           | 62.9       |
|       | 7     | 19          | 30.6    | 30.6          | 93.5       |
|       | 9     | 4           | 6.5     | 6.5           | 100.0      |
|       | Total | 62          | 100.0   | 100.0         |            |

|       |       | WestNonWest |         |               | Cumulative |
|-------|-------|-------------|---------|---------------|------------|
|       |       | Frequency   | Percent | Valid Percent | Percent    |
| Valid | 1.00  | 15          | 24.2    | 24.2          | 24.2       |
|       | 2.00  | 47          | 75.8    | 75.8          | 100.0      |
|       | Total | 62          | 100.0   | 100.0         |            |

Note that the frequencies of “WestNonWest” are the same as “WhereFought” (62 cases). Your recode was successful! In your old variable, “WhereFought,” the old values of 1 and 2 totaled 15 cases and the remaining values for 4, 6, 7 and 9 add up to 47 cases totaling 62 cases. If these numbers do not match up, something went wrong.

You will now recode another variable to see how comfortable you are with this process.

The original variable “Outcome” consists of seven different outcomes of a nonstate war. For the purpose of this exercise, you will recode “Outcome” so that only two outcomes are reflected: End or Continuation. The original values of 1 (Side A wins), 2 (Side B wins), 3 (Compromise), and 6 (Stalemate) will be recoded to 1 indicating an end to the war, while the remaining values of 4 (War transformed), 5 (War ongoing), and 7 (Conflict persists below war level) will be recoded to 2 indicating continuation of war.

Follow these instructions to recode the variable “Outcome”

- i. **Transform** > **Recode into Different Variables**
- ii. Select “Outcome” and move it to the Input Variable -> Output Variable box
- iii. In the Output Variable area, give your recoded variable the name “EndCont” then click Change
- iv. Click the **Old and New Values** button
- v. Click on Range under Old Values on the left side and enter 1 through 3, then under New Values on the right side, enter 1, then click on “Add”
- vi. Click on Value under Old Values on the left side and enter 6, then under New Values on the right side, enter 1, then click on “Add”
- vii. Click on **All other values** under Old Values on the left side, then under New Values on the right side, enter 2, then click on “Add” **Click Continue** > **OK**.

Your newly recoded variable, **EndCont**, should now appear in the last column in data view and the last row in variable view.

Now check to see that your recode was successful by comparing the frequencies on your new and old variables and answer the following questions:

1. Explain why collapsing the variable “Outcome” may be beneficial for determining the outcome of a non-state war.

---

---

2. Write down the mean for the original variable “Outcome” and the recoded variable “EndCont.”

---

3. How do you know if your recode was successful? Compare the old values for “Outcome” and the new values for EndCont.”

---

#### **IV. Levels of Analysis: States and the International System**

##### **States and War**

The state level of analysis assumes states are independent actors with their own interests. Studying war at the state level examines how characteristics such as the state's political system, its economy, and historical ties and experiences with other countries affect when states will engage in war. Interactions between states may lead to cooperation or conflict. Because there is no international organization with authority over all states, and no international police force, states must resolve their own conflicts. Disagreements between states may be settled diplomatically, and in some cases unresolved disputes may lead to war. Examining how state interests diverge and conflict between two or more states leads to war is to examine the occurrence of war at the state level of analysis.

The democratic peace theory asserts that well established democracies do not go to war with each other (Ray, 1993). Democracies go to war with non-democracies and non-democracies fight each other, newly democratized states may also engage in war with democracies. Once a country has fully established a democratic government, scholarly research demonstrates that it is not likely to go to war with another well-established democracy. This suggests that the relationship between democracies has some set of unique interests or characteristics that help to restrain two or more democratic countries from fighting each other. Studying these interests and characteristics requires examining the occurrence of war at the state level of analysis. At this level of analysis, state level characteristics and interests affect state behavior and the decision to go to war.

##### **The International System and War**

Studying war at the international system level involves examining how the system affects state behavior. Realists (discussed in more detail below) often utilize this level of analysis. While states remain the main actors at this level of analysis, the balance of power within the system helps to drive state behavior and influence when states will go to war with each other. The distribution of power (military and material capabilities) among the most powerful states within the international system determines the polarity of the international system. A multipolar system has three or more powerful states with relatively equal economic and military influence that exist within the system. A bipolar system contains just two major powers of roughly equal economic and military power. Finally, a unipolar system has one major superpower that dominates the international system.

While IR scholars disagree on what polarity is the most stable and least likely to erupt into warfare between great powers, there is general agreement that the distribution of power can influence international relations. Waltz (1979) argued that because the international system is anarchic and has no central authority that can enforce rules on states, all states exist in a self-help system. The ultimate goal of each state is survival, according to Waltz. In order to survive, states seek to acquire more power than other states to ensure survival if a dispute should devolve into a war.

One way to avoid war and to improve the chances of survival for all states is to establish an international balance of power. For example, a bipolar system wherein two major powers are close to equal in military and material capabilities can prevent war between these major powers. States of roughly equal power hesitate to go to war because it is not clear who would win. States that are uncertain of their chances of survival in a war will seek to avoid war.

### **Composite Variables**

In order to study the distribution of power in the international system, researchers must find a way to measure the power of states. The Correlates of War Project has created a dataset that measures power with seven different variables. Based on the studies of many IR scholars, military expenditure, number of military personnel, total population, urban production, iron and steel production, energy consumption, and a **composite variable** that combines all other variables to create a single unique score for each state. This variable is referred to as CINC in the dataset you will use for the exercise below. Composite variables are created by combining the information from two or more variables. The CINC variable takes information from the other six variables mentioned above to give each state a single score each year that can be easily compared to the scores of other countries.

But what if your particular theoretical approach to the causes of war suggested that the military expenditure as a percentage of total population was an important measure? Or, perhaps you would like to know the mean of all military expenditure over time rather than considering unique yearly scores? Luckily, we are not limited by the way that the data were collected. We can create our own composite variables.

In the following exercise, you will learn to use the COW dataset on military and national capabilities to compute unique variables.

### Exercise Three: Creating composite variables in SPSS

For these practice exercises, please use the Correlates of War database, COW War Data, 1816 - 2007 (v4.0): National Material Capabilities (v5.0).

1. **Read the codebook for the National Material Capabilities** data available on the COW website: <http://cow.dss.ucdavis.edu/data-sets/national-material-capabilities>.
  - a. Download the data on National Material Capabilities from the website above and save it as an Excel file.
  - b. Open SPSS. **File** > **Import Data** > **Excel**
  - c. Choose your saved Excel File and the dataset will open in the data editor.

### 2. Creating a Composite Variables in SPSS

There are several ways you can choose to compute new variables – from very simple additive techniques to complex mathematical expressions. Below are directions for two ways to compute new variables. The first shows you how to compute a new variable using subtraction. The National Material Capabilities dataset has a variable representing the total state population and urban population for each observation. You will create a variable denoting the non-urban population by subtracting the urban population from the total state population. The second variable you will create is a measure of military expenditures per person, per year, in each state.

### 3. Computing Non-urban Population Variable

- a. From the menu bar, select “Transform”
- b. Select “Compute Variable”
- c. In the box named Target Variable type the name of your new variable: nonurbanpop
- d. In the “Numeric Expression” box type the expression “tpop – upop.” Click OK.

A new variable named nonurbanpop will appear in your dataset.

1. Identify the total population, urban population, and non-urban population for Canada in the year 2002, Mexico for the year 2011, and Portugal for the 1835.

Canada 2002

Total population \_\_\_\_\_

Urban population \_\_\_\_\_

Non-urban population \_\_\_\_\_

Mexico 2011

Total population \_\_\_\_\_

Urban population \_\_\_\_\_

Non-urban population \_\_\_\_\_

Portugal 1835

Total population \_\_\_\_\_

Urban population \_\_\_\_\_

Non-urban population \_\_\_\_\_

#### 4. Computing Military Expenditures Per Person Variable

- a. From the menu bar, select “Transform”
- b. Select “Compute Variable”
- c. In the box named Target Variable, type the name of your new variable: milexpp
- d. In the “Numeric Expression” box, type the expression “milex/tpop” Click OK.

A new variable named “milexpp” will appear in your dataset.

2. Identify the per person military expenditures for the U.S. in 2012, Guatemala in 1984, and Uruguay in 1994.

US 2012 \_\_\_\_\_

Guatemala 1984 \_\_\_\_\_

Uruguay 1994 \_\_\_\_\_



## **V. Theories of War**

Two major theories that dominate the study of IR are realism and liberalism. The origins of realism date back to Thucydides studies of the Peloponnesian wars of 5<sup>th</sup> Century B.C. The foundations of IR liberalism are often associated with Kantian ideas of the 18<sup>th</sup> century. Theories of realism and liberalism come in many different shapes and sizes. Liberal theories of international relations tend to focus on how states can cooperate despite an anarchic international system. Cooperation can lead to avoiding war between states. Realist theories, in contrast, are apt to concentrate on how anarchy influences the behavior of states and drives them to compete for relative power. In order to survive, states seek to acquire more power than other states. This section will consider two theories of war that feature prominently in the liberal and realist literature.

### **Democratic Peace Theory**

The democratic peace theory is grounded in liberal assumptions. Liberalism traces back to the Enlightenment period in the 18<sup>th</sup> century where individuals were viewed as rational and capable of improving their condition. Immanuel Kant argued that states could overcome anarchy through collective action, once states shared three features: a representative democracy, adherence to international law and organizations, and advanced commercial integration (Russett and Oneal, 2000).

The democratic peace theory is based on Immanuel Kant's argument and stresses three inter-related concepts: democracy, economic interdependence, and membership in international organizations. This is known as the Kantian peace triangle. This theory suggests that democracies do not fight other democracies, and in general, are more peaceful than nondemocracies. Furthermore, states that are economically interdependent with another state tend to avoid conflict as it may threaten their economic growth. Lastly, states who join international organizations are less likely to fight against other members.

### **Power Transition Theory**

Power transition theory is grounded in realist assumptions. It presupposes that states are the most relevant actors in the international system; the international system is anarchic and as such states must help themselves; and relative power in the international system drives state behavior. Power transition theorists suggest that war in the international system is more likely to occur when a rising power (a state that was weaker than the hegemon but is now approaching parity with the most powerful state in the system) is unhappy with the status quo.

Power transition theory asserts that the most powerful state in the world sets the rules for diplomatic, economic, military, and cultural interactions between states. For example, the current international system was designed post-WWII. International organizations focus on a market-oriented economy and stress the importance of democratic values (Lemke, 1997). The United States, arguably the world's current hegemon, played a large role in setting up the rules of the international system after the war. If a rising power is dissatisfied with how the current

rules distribute goods in the international system (Kugler and Organski, 1989), this may lead to war. The rising power seeks to change the rules of distribution after winning the war.

### **Testing Theories of War through Statistical Analysis**

Democratic peace theory and power transition theory offer two different explanations as to why states might choose to go to war. While each provides a logically consistent and sound explanation regarding the study of war, how do we determine if they have any basis in fact? Once scholars have created theories based on previous research, the theories can be tested through empirical analysis. There are many tools of empirical research available to IR scholars. One such tool is statistical analysis.

The descriptive methods above helped to summarize and evaluate important properties of single variables. For example, in Exercise One using SPSS, we were able to determine the mean number of battle deaths across all wars in our dataset. In Exercise Four, we will examine the relationship between two variables. A relationship exists between two variables if the values for one variable are connected to the observations of the other variable (Johnson and Reynolds, 2005). An independent variable can cause a change in the dependent variable. For example, if states that initiate war generally have more power than defending states, then the value of the independent variable, power, has a relationship with the dependent variable, initiation of war. It is important to note that relationships between variables can be strong or weak and causal or spurious.

The strength of the relationship between an independent variable and a dependent variable is based on the values of these two variables. The strength of the relationship measures the degree to which changes in the independent variable are linked to changes in the dependent variable. For example, consider this research question: Do more powerful states initiate war more often than less powerful states? Let's hypothetically consider all wars fought between only two states within the years 1900 and 2016. In each war, one state is identified as the more powerful state and the other is identified as the less powerful state. The independent variable is power and the dependent variable is war initiation. A strong relationship between these two variables would be indicated if the most powerful state involved in the war initiated the war most of the time. If the most powerful state initiated the war every time, this would be a perfect relationship. This finding would indicate a seemingly positive response to our research question.

A weak relationship is one where the value of the dependent variable does not differ based on the value of the independent variable. For example, if we look at the wars fought between two states between 1900 and 2016 only to find half of the wars were initiated by the most powerful state and half were initiated by the weakest state, this would indicate a weak relationship; and would suggest that the answer to our research question above is a resounding no.

A strong relationship between two variables indicates that they are significantly associated. If we were to find a strong relationship between power and the initiation of war, we might predict that in the future when two states experience conflict, it is more likely that the stronger state will initiate war. This theory seems plausible and logical. The association between the two variables

is empirical evidence that shows some support for this theory. However, there are instances in which variables can be associated but there is no causal relationship. A common example is the association between people eating ice cream and drownings. Taken alone, these two variables may have a strong relationship. As more people eat ice cream, more people drown. However, there is an omitted variable. More people eat ice when the weather is warm, and more people swim when the weather is warm. Though ice cream eating and drowning may be variables with a strong association, ice cream eating does not cause drowning. This is a spurious relationship.

Democratic peace theory and power transition theory offer two logically consistent and plausible theories about when states will choose to avoid war and when they will choose to go to war. These theories can be tested empirically. In Exercise Four, we will test a much simpler theory. We will begin with the question: Do states that initiate war win more wars than states that do not initiate war? It would seem that states that start wars have some expectation that they can win the war. If a state anticipated losing, the government might make more of an effort at settling a conflict diplomatically and would be less likely to start a war. This makes sense, but we do not have to rely on logic. We can test the theory: States that initiate war win wars more frequently than the non-initiating state. We will test the association between the variables “Initiator” and “Win or Other Outcome.” First, however, we will consider four different tests of association used depending on whether the independent and dependent variables are nominal, ordinal, or interval.

### **Testing for Association of Variables**

There are a variety of ways to test for the relationship of variables using SPSS. For the purpose of this exercise, you will be introduced to four specific tests: T-test, One-Way ANOVA, Pearson’s Correlation, and Chi-Square Test.

#### **T-test**

The t-test compares the means of a variable from TWO groups. An independent samples t-test compares the means of two different groups (e.g., reaction times on a task for women vs. men), whereas a paired t-test compares the means of the same group at two different times (e.g., reaction times for the same people on a task before or after a training period).

When you perform a t-test, you typically are looking for evidence of a significant difference between sample means (2-sample t) or between the sample mean and a hypothesized value (1-sample t). In other words, the t-value measures the size of the difference relative to the variation in your sample. The greater the magnitude of t (this value can be either positive or negative), the greater the evidence *against* the null hypothesis that there is no significant difference. The closer t is to 0, the more likely that there is no significant difference.

#### **One-way ANOVA**

The One-way Analysis of Variance (ANOVA) is very similar to an independent samples t-test, because it is also a comparison of means. You use an ANOVA to determine whether the mean values across three or more groups are significantly different.

For instance, let's say your IV is country (China, Mexico, U.S.) and your DV is number of wars initiated during the period 1800 – 2015. If the U.S. initiated an average of 7.6 wars in that time, China initiated 5.0 wars, and Mexico initiated 1.2 wars, you would use the ANOVA test to determine whether the average number of wars initiated were significantly different from each other.

### **Correlation**

The (Pearson) correlation coefficient is a measure of the strength of the linear relationship between two interval variables. You can also test for correlations between ordinal two variables using the Spearman's rank correlation coefficient. Correlations can be positive or negative, and range from -1 (perfect negative correlation) to 1 (perfect positive correlation). Thus, most correlation coefficients will have decimal values (e.g., .04, .62, .89). A value of 0 indicates no correlation.

### **Chi-Square Test**

The chi-square test is one way to determine whether there is a significant relationship between two nominal variables. The chi-squared statistic is a single number that tells you how much difference exists between your observed counts and the counts you would expect if there were no relationship in the population. There are different chi-square tests. In the example below, we will use Pearson's Chi-Square test.

#### *P-value*

The p-value is not one of the statistical tests. It is a measure used to determine whether an association between two variables is statistically significant. Typically, researchers set a p-value of .05 as the cutoff for significance. Values below this are considered significant. For example, a p-value of .03 would indicate a significant association (using the language above, we would refer to this as a "strong" relationship). Contrarily, a p-value of .07 would indicate a nonsignificant relationship.

To summarize, the type of test you use is determined by the level of measurement of the independent and dependent variables.

- **T-test or One-way ANOVA** if you have a nominal independent variable and an interval dependent variable. Use the t-test if the independent variable has two groups and the ANOVA if the independent variables has three or more groups.
- **Chi Square Test** if you have nominal independent and dependent variables.
- **Pearson's Correlation** if your independent and dependent variables are both interval variables.

In the following exercise, you will use the COW dataset to test theories of war by testing the relationship between two variables.

### Exercise Four: Testing Association of Variables

For these practice exercises, please use the Correlates of War database, COW War Data, 1816 - 2007 (v4.0): National Material Capabilities (v5.0).

We will test the theory that initiators of war are more likely to win wars. In order to test this theory, we will recode the "Outcome" variable to indicate winning the war versus all other outcomes. This will be our nominal dependent variable. The independent variable for this exercise will be the "Initiator" of the war.

**1. Read the codebook for the Interstate War data** available on the COW website: <http://cow.dss.ucdavis.edu/data-sets/national-material-capabilities>.

- a. Download the data on Interstate War from the website above and save it as an Excel file.
- b. Open SPSS. **File** > **Import Data** > **Excel**
- c. Choose your saved Excel File and the dataset will open in the data editor.

**2. Follow these instructions to recode the variable "Outcome"**

- a) Click on **Transform** > **Recode into Different Variables**
- b) Select "Outcome" and move it to the Input Variable -> Output Variable box
- c) In the Output Variable area, give your recoded variable the name "WinOther" and label your variable "Win or Other Outcome" then click Change
- d) Click the **Old and New Values** button. Input Old Value "1" and New Value "1" then click on "Add"
- e) Click on Range under Old Values on the left side and enter 2 through 8, then under New Values on the right side, enter 2, then click on "Add" **Click Continue** > **OK**.
- f) Go to the variable view of your dataset so you can see all of your variables.
- g) Find the WinOther row and click on the **Values** box where there are three dots. (The value is currently NONE.) You are going to add the values for this variable.
- h) In the Value box insert "1". In the Label box write "Won the War." Click on "Add"
- i) In the Value box insert "2". In the Label box write "Other Outcome." Click on "Add" then Click on "Ok".
- j) Go to the column labeled "Measure" and ensure that the level of measurement is selected as nominal.

**3. Label "Initiator" Variable**

- a) Go to the variable view of your dataset so you can see all of your variables.
- b) Find the Initiator row and click on the **Values** box where there are three dots. (The value is currently NONE. You are going to add the values for this variable).
- c) In the Value box insert "1". In the Label box write "Initiated the War." Click on "Add"
- d) In the Value box insert "2". In the Label box write "Did Not Initiate the War." Click on "Add" then Click on "Ok".
- e) Go to the column labeled "Measure" and ensure that the level of measurement is selected as nominal.

Now we will test for an association between “Initiator” (independent variable) and the dependent variable “WinOther.”

1. Both our independent and dependent variables are nominal. What statistical test of association should we use for these two variables?

#### 4. Test for Association of Variables

Analyze > Descriptive Statistics > Crosstabs. Put your IV, **Initiator**, in the “Rows” box and your DV, **WinOther**, in the “Columns” box. Click Statistics and select “Chi-square”. Click Continue > Cells, and under “Percentages”, select “Rows”. “Observed” will already be checked; leave it as is. Click Continue > OK.

**Initiator \* Win or Other Outcome Crosstabulation**

|           |                      |                    | Win or Other Outcome |               | Total  |
|-----------|----------------------|--------------------|----------------------|---------------|--------|
|           |                      |                    | Won the war          | Other Outcome |        |
| Initiator | Initiated War        | Count              | 61                   | 48            | 109    |
|           |                      | % within Initiator | 56.0%                | 44.0%         | 100.0% |
|           | Did Not Initiate War | Count              | 94                   | 134           | 228    |
|           |                      | % within Initiator | 41.2%                | 58.8%         | 100.0% |
| Total     | Count                |                    | 155                  | 182           | 337    |
|           | % within Initiator   |                    | 46.0%                | 54.0%         | 100.0% |

**Chi-Square Tests**

|                                    | Value              | df | Asymptotic Significance (2-sided) | Exact Sig. (2-sided) | Exact Sig. (1-sided) |
|------------------------------------|--------------------|----|-----------------------------------|----------------------|----------------------|
| Pearson Chi-Square                 | 6.446 <sup>a</sup> | 1  | .011                              |                      |                      |
| Continuity Correction <sup>b</sup> | 5.867              | 1  | .015                              |                      |                      |
| Likelihood Ratio                   | 6.443              | 1  | .011                              |                      |                      |
| Fisher's Exact Test                |                    |    |                                   | .014                 | .008                 |
| Linear-by-Linear Association       | 6.427              | 1  | .011                              |                      |                      |
| N of Valid Cases                   | 337                |    |                                   |                      |                      |

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 50.13.

b. Computed only for a 2x2 table

First, look at the “Chi-Square Tests” table and note the Pearson chi-square and its p-value (Asymp. Sig [2-sided]). In this case, **Initiator** and **WinOther** are significantly related. Second, look at how the percentages differ between Initiator and “WinOther” in the “Crosstabs” table.

Look at note “a” here. In this case, there is no problem. However, if >20% of the cells had expected cell counts < 5, your Chi-square would be considered somewhat unstable, and you should comment on this. In such a case, if >20% of the cells have expected cell counts < 5 and some cells have an observed count of 0; it limits how much confidence we have in the results of the chi-square analysis.

2. Which group, states that initiated war or states that did not initiate war, won a higher percentage of wars? \_\_\_\_\_
3. Is our theory that war initiators win more wars supported by these results? Explain your reasoning.

---



---



---

### 3. Create a Pie Chart

- a. **Graphs** > **Legacy Dialogs** > **Pie**. Leave default setting". Click **Define**.
- b. Under "Slices Represent" click % of cases.
- c. Move WinOther over to "Define Slices By".
- d. Move Initiator to "Columns" and click "OK."

#### Graph



4. How does the pie graph visually represent the association between our independent and dependent variables that was tested in the Chi-Square above?

---



---



---

### Conclusion

These exercises were intended to introduce students to basic data analysis using the statistical package, SPSS, to analyze data associated with war. We hope that you found these exercises helpful and encourage you to try practicing your skills using other publicly available databases

such as Polity IV through the Center for Systemic Peace available at <http://www.systemicpeace.org/inscrdata.html>.

## References

Babbie, Earl. 2002. *The Basics of Social Research*. Stamford, CT: Wadsworth Publishing Company.

Barbieri, Katherine and Omar M. G. Omar Keshk. 2016. *Correlates of War Project Trade Data Set Codebook*, Version 4.0. Online: <http://correlatesofwar.org>.

Bush, George W. 2001. "Selected Speeches of George W. Bush 2001 – 2008." Retrieved from: [https://georgewbushwhitehouse.archives.gov/infocus/bushrecord/documents/Selected\\_Speeches\\_George\\_W\\_Bush.pdf](https://georgewbushwhitehouse.archives.gov/infocus/bushrecord/documents/Selected_Speeches_George_W_Bush.pdf)

Frude, Neil. 1987. *A Guide to SPSS/PC+*. London, UK: Palgrave MacMillan Education.

Gebhard, Carmen. 2017. "One World, Many Actors." Stephen McGlinchy (Ed.) *International Relations* pp. 32 -45. Bristol, England: E-International Relations Publishing.

Johnson, Janet Buttolph and H.T. Reynolds. 2005. *Political Science Research Methods*. Washington D.C: CQ Press.

Kugler, Jacek & A. F. K. Organski, 1989. 'The Power Transition: A Retrospective and Prospective Evaluation', pp. 171-194 in Manus Midlarsky, ed., *The Handbook of War Studies*. Boston, MA: Unwin Hyman.

Lacina, Bethany & Nils Petter Gleditsch. 2005. Monitoring trends in global combat: A new dataset of battle deaths. *European Journal of Population* 21(2–3): 145–166.

Lemke, Douglas. 1997. "The Continuation of History: Power Transition Theory and the End of the Cold War." *Journal of Peace Research*, Vol. 34, No. 1 (Feb., 1997), pp. 23-36.

Mazzei, Julie. 2009. *Death Squads or Self-Defense Forces?* Chapel Hill, NC: University of North Carolina Press.

Montgomery, Tommie Sue. 1996. *Revolution in El Salvador*. 2nd ed. Boulder, CO: Westview Press.

Morris, Benny. 1948. *A History of the First Arab–Israeli War*. Yale, CT: Yale University Press.

Pevehouse, Jon C., Timothy Nordstrom, and Kevin Warnke. 2004. "The COW-2 International Organizations Dataset Version 2.0," *Conflict Management and Peace Science* 21:101-119.

Pollock, Philip. 2005. *An SPSS Companion to Political Analysis*, Second Edition. Washington DC: CQ Press.



Ray, J. L. 1993. "Wars" between democracies: Rare or non-existent?" *International Interactions*. Vol. 18: 251 – 76.

Russett, Bruce and John Oneal. 2000. *Triangulating Peace: Democracy, Interdependence, and International Organizations (The Norton Series in World Politics)*. New York, NY: Norton Press.

Sarkees, Meredith Reid and Frank Wayman. 2010. *Resort to War: 1816 - 2007*. Washington DC: CQ Press.

Singer, David. 1961. "The Level-of-Analysis Problem in International Relations." *World Politics*, Vol. 14, No. 1, pp. 77-92.

Small, Melvin and J. David Singer. 1982. *Resort to Arms: International and Civil War, 1816-1980*. Beverly Hills, Calif.: Sage.