Excel: Conjunctive Analysis of Case Configurations

Summary by Grant Drawve

Conjunctive Analysis of Case Configurations (CACC) is being used more frequently within criminal justice research. In short, CACC relies on categorical data to identify patterns within the data as the patterns relate to a specific outcome. This How To is not intended to provide a background on CACC but provide an example of how CACC could be done in Excel. The additional resources below are meant to assist and provide greater background if needed on CACC.

Additional Resources:

- Article: Miethe, T.D., Hart, T.C., Regoeczi, W.C. (2008). The conjunctive analysis of case configurations: An exploratory method for discrete multivariate analyses of crime data. *Journal of Quantitative Criminology, 24*, 217-241.
- Article: Hart, T.C. (2019). Identifying situational clustering and quantifying its magnitude in dominant case configurations: New methods for conjunctive analysis. Available online first *Crime & Delinquency*. https://doi.org/10.1177/0011128719866123
- Statistical Package: R package for CACC by Esteve, Moneva, and Hart
  - https://zenodo.org/record/3459112#.XuTyykVKiUk
- Webpage: How to construct an automated process for CACC with the Association Rules by Wheeler
  - Andrewpwheeler.com
- Webpage: Spatial walk-through between ArcGIS and SPSS. This has a RTM foundation but a similar process could be used for any polygon file by Barnum
  - https://www.riskterrainmodeling.com/

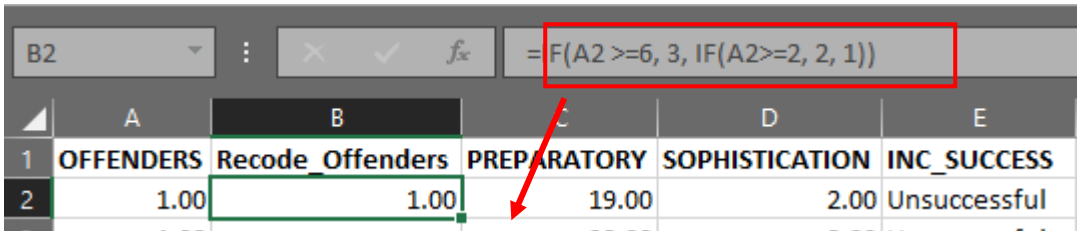
Data for current application:

- American Terrorism Study (ATS) data for Domestic Events (U.S.): $n = 310$
  - Similar data were used in, Gruenewald, J., Drawve, G., & Smith, B.L. (year). The situated contexts of American terrorism: A conjunctive analysis of case configurations. *Criminal Justice & Behavior, 46*(6), 884-901.
- Independent Variables:
  - Number of Terrorist Offenders
  - Number of Preparatory Incidents
  - Level of Weapon Sophistication
- Dependent Variable:
  - Successful Terrorist Attack

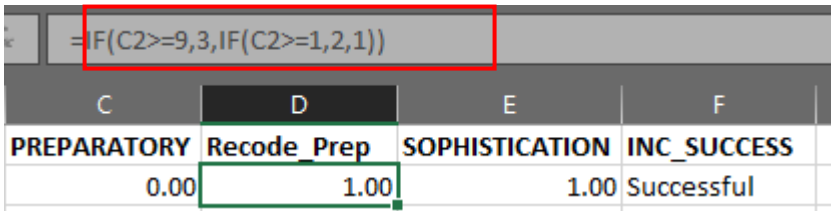
	A	B	C	D
1	OFFENDERS	PREPARATORY	SOPHISTICATION	INC_SUCCESS
2	6.00	62.00	2.00	Unsuccessful
3	4.00	12.00	2.00	Unsuccessful
4	5.00	10.00	2.00	Unsuccessful
5	6.00	31.00	2.00	Successful
6	4.00	19.00	2.00	Unsuccessful
7	4.00	2.00	2.00	Unsuccessful
8	7.00	54.00	1.00	Unsuccessful

1. The first step for CACC is to code your variables into numeric categories. Below provides an outline of how these variables are recoded with steps on how to complete this in Excel. If the recode is more than 2 categories, this will require nested **IF** statements

- Number of Offenders – Originally a continuous count measure
  - Range: 1 – 36
  - Mean: 2.76
  - Recode Categories:
    - 1 = 1 (potentially Lone Wolf)
    - 2 – 5 = 2
    - 6 + = 3
  - Insert a column next to Offenders for your Recode. Use a similar **IF** statement as seen below to recode the measure into 3 categories, then Double-Click the green box to autofill the cells below with the same function. Scroll down and double-check your recode to make sure it populated correctly.



- Number of Preparatory Events
  - Range: 0 – 76
  - Mean: 8.13
  - Recode Categories:
    - 0 = 1
    - 1 – 8 = 2
    - 9 + = 3
  - Insert a column next to Preparatory for your Recode. Use a similar **IF** statement as seen below to recode the measure into 3 categories, then Double-Click the green box to autofill the cells below with the same function



- Level of Sophistication
  - Range: 0 – 2
  - Mean: 1.09
  - Since Sophistication is already categorized within ATS, we are going to leave it as it is coded
    - 0 = Low, 1 = Moderate, 2 = High

- Successful Attack
  - Categories: Unsuccessful, Partial Success, and Successful
  - Recode:
    - 0 = Unsuccessful
    - 1 = Partial Success & Successful
  - Copy & Paste the INC_SUCCESS column next to itself. Change the title of the second column to Recode_Success. Here, you can use the **Find and Replace** function in Excel. Highlight *only* the Recode_Success column then use the **Find and Replace** function. You can run Unsuccessful first since "successful" is part of the word "unsuccessful"
    - Find what = Unsuccessful
      - Replace with= 0
    - Find what = Partial Success
      - Replace with= 1
    - Find what = Successful
      - Replace with= 1

F	G
INC_SUCCESS	**Recode_Success**
Successful	1
Successful	1
Successful	1
Successful	1
Successful	1
Successful	1
Unsuccessful	0
Unsuccessful	0
Successful	1
Successful	1
Unsuccessful	0
Unsuccessful	0
Partial success	1

Now you have your independent and dependent variables recoded for CACC.

There would be the potential for 54 different configurations (3*3*3*2).

2. To make a cleaner workspace, **copy and paste** your final variables onto a new sheet (Recode_Offenders, Recode_Prep, Sophistication, and Recode_Success). Once this is completed, insert a column titled, "Count". **Assign the cells within that column to "1"** (For those familiar with Excel, there are multiple ways to do this, but to ease the process, a Count column was created). At this point, if any of your numbers have decimal places, remove the decimal points so there are only whole numbers.

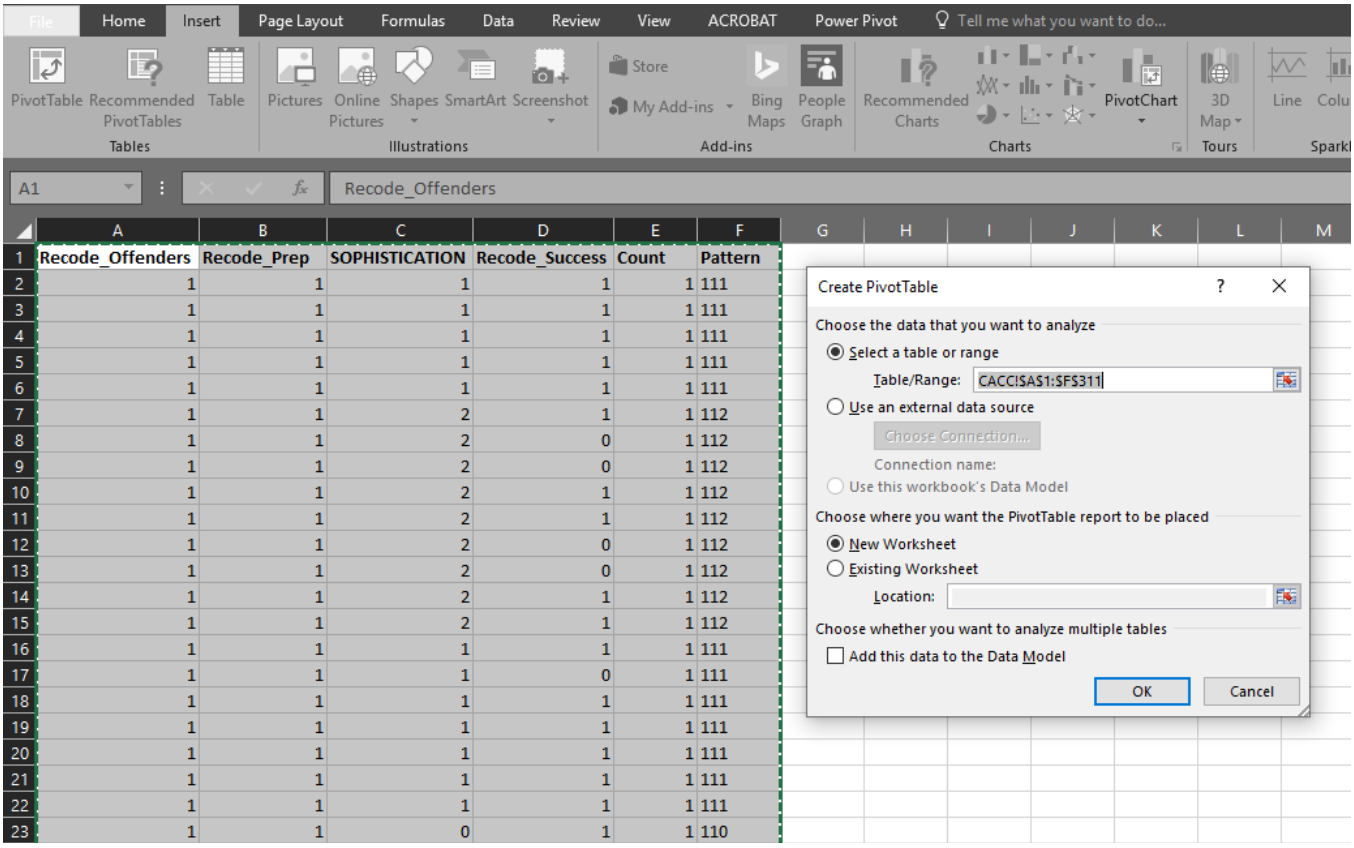
	A	B	C	D	E
1	Recode_Offenders	Recode_Prep	SOPHISTICATION	Recode_Success	Count
2	1	1	1	1	1
3	1	1	1	1	1
4	1	1	1	1	1
5	1	1	1	1	1
6	1	1	1	1	1
7	1	1	2	1	1
8	1	1	2	0	1

3. With the New Worksheet, Insert another column titled, "Pattern". Here, you are going to fill the cell value with only your independent variables with the Concatenate function in Excel (Check out video on YouTube channel if help is needed with concatenate). You only want to do this for your independent variables, not Recode_Success or the Count column. Double-Click the green box to autofill the cells below with the same function. What you will notice is a unique configuration of each variable combined into one field.

F2 =CONCATENATE(A2,B2,C2)

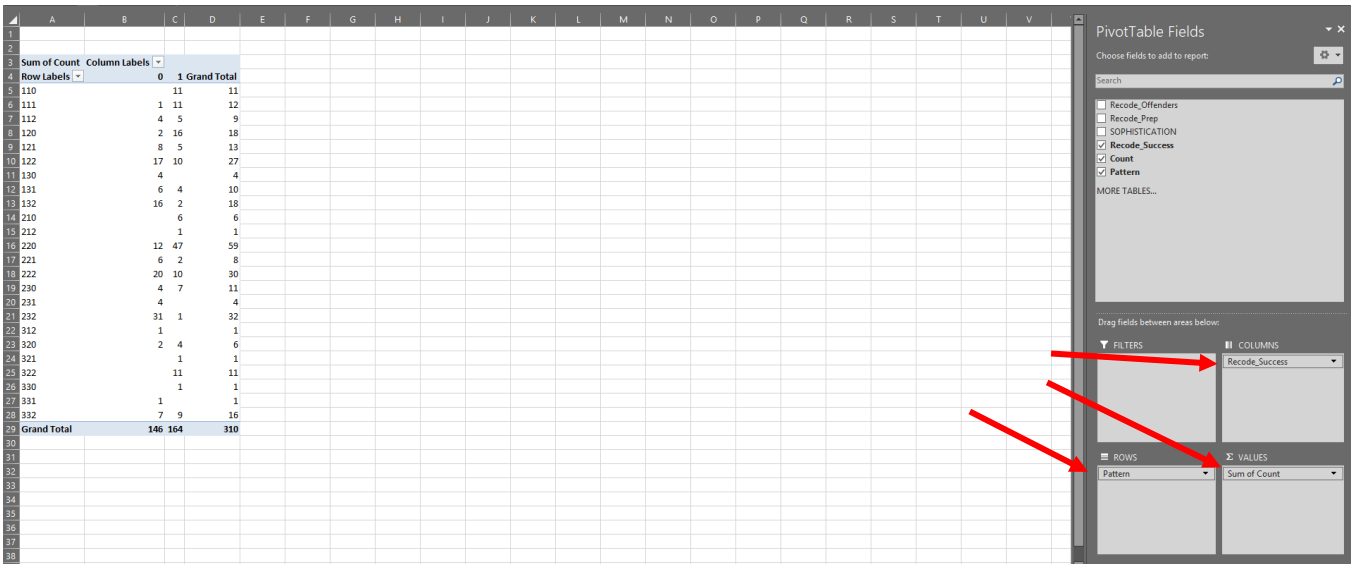
	A	B	C	D	E	F
1	Recode_Offenders	Recode_Prep	SOPHISTICATION	Recode_Success	Count	Pattern
2	1	1	1	1	1	111
3	1	1	1	1	1	111
4	1	1	1	1	1	111
5	1	1	1	1	1	111
6	1	1	1	1	1	111
7	1	1	2	1	1	112
8	1	1	2	0	1	112
9	1	1	2	0	1	112
10	1	1	2	1	1	112
11	1	1	2	1	1	112
12	1	1	2	0	1	112

4. Now, highlight the data table (only the columns and rows with data) and insert a **PivotTable**. Go to/Click the **Insert** tab at the top and click **PivotTable**. Since you have already highlighted your data table, this is generated in the Table/Range. You can leave the "Choose where you want the PivotTable report to be placed" to a New Worksheet.



	A	B	C	D	E	F
1	Recode_Offenders	Recode_Prep	SOPHISTICATION	Recode_Success	Count	Pattern
2	1	1	1	1	1	111
3	1	1	1	1	1	111
4	1	1	1	1	1	111
5	1	1	1	1	1	111
6	1	1	1	1	1	111
7	1	1	2	1	1	112
8	1	1	2	0	1	112
9	1	1	2	0	1	112
10	1	1	2	1	1	112
11	1	1	2	1	1	112
12	1	1	2	0	1	112
13	1	1	2	0	1	112
14	1	1	2	1	1	112
15	1	1	2	1	1	112
16	1	1	1	1	1	111
17	1	1	1	0	1	111
18	1	1	1	1	1	111
19	1	1	1	1	1	111
20	1	1	1	1	1	111
21	1	1	1	1	1	111
22	1	1	1	1	1	111
23	1	1	0	1	1	110

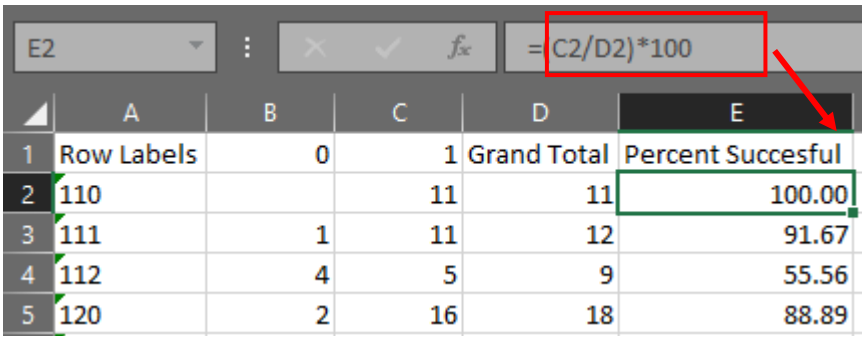
5. In the PivotTable sheet, move the Pattern to rows, Recode_Success to columns, and Count to values (make sure it is set to Sum). What this displays is the configuration of IVs with the DV category separated on 0 and 1. The Grand Total column indicates how many cases had that overall configuration. For instance, 11 cases had 110, which is 1 Offender, 0 Preparatory Events, and Low level of sophistication. Out of those 11 cases, all of them were successful.



6. With the generated PivotTable, highlight from "Row Labels" across and down to the "16". **Copy and Paste** this into a New Worksheet. When Pasting, Paste only the "Values".

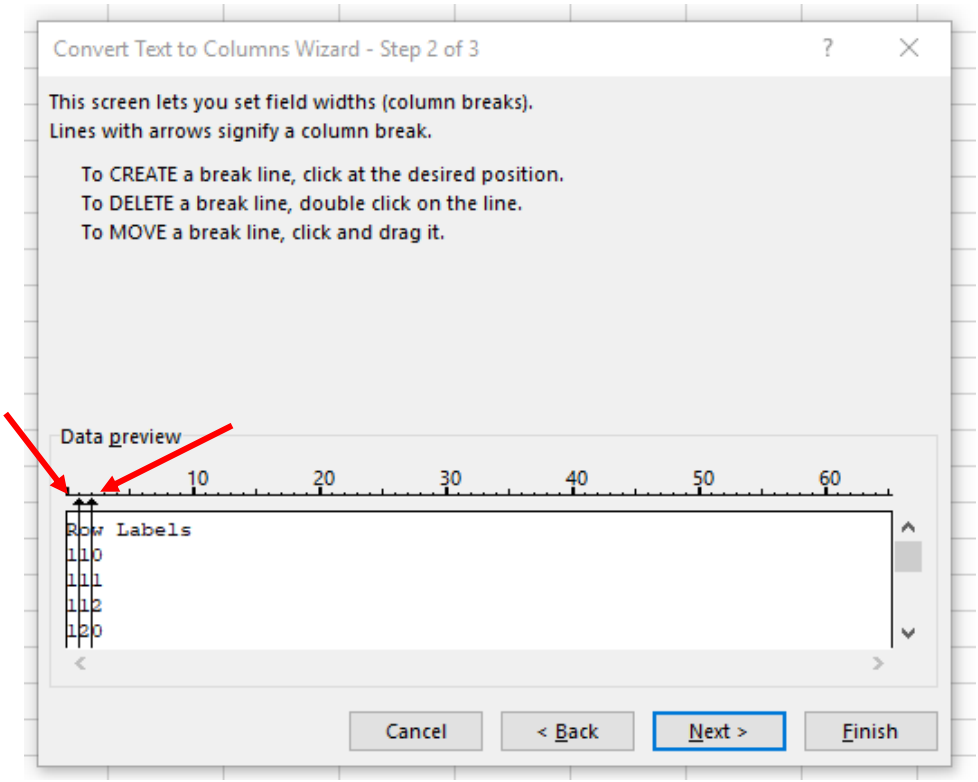
	A	B	C	D
1	Row Labels	0	1	Grand Total
2	110		11	11
3	111	1	11	12
4	112	4	5	9
5	120	2	16	18
6	121	8	5	13
7	122	17	10	27
8	130	4		4
9	131	6	4	10
10	132	16	2	18
11	210		6	6
12	212		1	1
13	220	12	47	59
14	221	6	2	8
15	222	20	10	30
16	230	4	7	11
17	231	4		4
18	232	31	1	32
19	312	1		1
20	320	2	4	6
21	321		1	1
22	322		11	11
23	330		1	1
24	331	1		1
25	332	7	9	16

7. With the new table, calculate the percent of cases that were successful. Insert a column next to "Grand Total" titled, "Percent Successful". To **calculate** the percentage of cases that were successful, you will want to divide the "1" column, remember it is/was coded as successful, by the "Grand Total" column. In the formula, you can multiple by 100 to obtain the percentage if desired.



	A	B	C	D	E
1	Row Labels	0	1	Grand Total	Percent Succesful
2	110		11	11	100.00
3	111	1	11	12	91.67
4	112	4	5	9	55.56
5	120	2	16	18	88.89

8. Next, let us come back to the independent variables. Insert 2 blank columns to the right of "Row Labels". Highlight your "Row Labels" column and go to the Data tab at the top and click "**Text to Columns**". This column is "**Fixed width**" since each configuration has the same number of characters. Click Next. Create 2 breaklines between that to split the configuration into 3 columns. This separates the "110" to "1", "1", "0". See below for an example of where the break lines should be placed.



Click **Next** and then Click **Finish**. Your table should now look like the one below. Rename the column Headings: A1 = Number of Offenders; B1 = Number of Preparatory Events; C1 = Sophistication

	A	B	C	D	E	F	G
1	R	o	w Labels	0	1	Grand Total	Percent Succesful
2	1	1	0		11	11	100.00
3	1	1	1	1	11	12	91.67
4	1	1	2	4	5	9	55.56
5	1	2	0	2	16	18	88.89
6	1	2	1	8	5	13	38.46
7	1	2	2	17	10	27	37.04
8	1	3	0	4		4	0.00
9	1	3	1	6	4	10	40.00
10	1	3	2	16	2	18	11.11
11	2	1	0		6	6	100.00
12	2	1	2		1	1	100.00
13	2	2	0	12	47	59	79.66
14	2	2	1	6	2	8	25.00
15	2	2	2	20	10	30	33.33
16	2	3	0	4	7	11	63.64
17	2	3	1	4		4	0.00
18	2	3	2	31	1	32	3.13
19	3	1	2	1		1	0.00
20	3	2	0	2	4	6	66.67
21	3	2	1		1	1	100.00
22	3	2	2		11	11	100.00
23	3	3	0		1	1	100.00
24	3	3	1	1		1	0.00
25	3	3	2	7	9	16	56.25

9. For conjunctive analysis, when there are over 100 cases, the minimum number of cases needed per configuration is 10 (dominant case configuration). **Sort** the "Grand Total" column from Largest to Smallest. Then copy and paste "values" into a new sheet of only the rows that have 10 + cases. By pasting values, this also removes the Percent Successful calculation and leaves the percentage (change to 2 decimal places). Also, now delete the columns "0" and "1".

	A	B	C	D	E
1	Number of Offenders	Number of Preparatory Events	Sophistication	Grand Total	Percent Succesful
2	2	2	0	59	79.66
3	2	3	2	32	3.13
4	2	2	2	30	33.33
5	1	2	2	27	37.04
6	1	2	0	18	88.89
7	1	3	2	18	11.11
8	3	3	2	16	56.25
9	1	2	1	13	38.46
10	1	1	1	12	91.67
11	1	1	0	11	100.00
12	2	3	0	11	63.64
13	3	2	2	11	100.00
14	1	3	1	10	40.00

10. Here you might think, well this is difficult to make sense of given the numbers. This is when you should **Find & Replace** the number categories with descriptive text of what that value pertains to in relation to the outcome event of a successful terrorist attack. To do so, highlight the respective column and make the changes similar to as seen below.

- Number of Offenders:
  - 1 = Lone Assailant
  - 2 = 2 – 5 Offenders
  - 3 = 6 + Offenders
- Number of Preparatory Events:
  - 1 = No Known Prep. Events
  - 2 = 1 – 8 Prep. Events
  - 3 = 9 + Prep. Events
- Weapon Sophistication:
  - 0 = Low
  - 1 = Moderate
  - 2 = High

	A	B	C	D	E
1	Number of Offenders	Number of Preparatory Events	Sophistication	Grand Total	Percent Succesful
2	2-5 Offenders	1-8 Prep. Events	Low	59	79.66
3	2-5 Offenders	9+ Prep. Events	High	32	3.13
4	2-5 Offenders	1-8 Prep. Events	High	30	33.33
5	Lone Assailant	1-8 Prep. Events	High	27	37.04
6	Lone Assailant	1-8 Prep. Events	Low	18	88.89
7	Lone Assailant	9+ Prep. Events	High	18	11.11
8	6+ Offenders	9+ Prep. Events	High	16	56.25
9	Lone Assailant	1-8 Prep. Events	Moderate	13	38.46
10	Lone Assailant	No Known Prep. Events	Moderate	12	91.67
11	Lone Assailant	No Known Prep. Events	Low	11	100.00
12	2-5 Offenders	9+ Prep. Events	Low	11	63.64
13	6+ Offenders	1-8 Prep. Events	High	11	100.00
14	Lone Assailant	9+ Prep. Events	Moderate	10	40.00

11. Now, it is recommended to sort the dominant case configurations (10 or more cases in a configuration) from largest to smallest based on the DV, Percent Successful. If you sum the Grand Total, you would find that 268 of the 310 cases (86 percent) are within the dominant case configurations. The 13 dominant case configurations also represent only 24 percent of the total potential configurations possible.

Row ID	Number of Offenders	Number of Preparatory Events	Sophistication	Grand Total	Percent Succesful
1	Lone Assailant	No Known Prep. Events	Low	11	100.00
2	6+ Offenders	1-8 Prep. Events	High	11	100.00
3	Lone Assailant	No Known Prep. Events	Moderate	12	91.67
4	Lone Assailant	1-8 Prep. Events	Low	18	88.89
5	2-5 Offenders	1-8 Prep. Events	Low	59	79.66
6	2-5 Offenders	9+ Prep. Events	Low	11	63.64
7	6+ Offenders	9+ Prep. Events	High	16	56.25
8	Lone Assailant	9+ Prep. Events	Moderate	10	40.00
9	Lone Assailant	1-8 Prep. Events	Moderate	13	38.46
10	Lone Assailant	1-8 Prep. Events	High	27	37.04
11	2-5 Offenders	1-8 Prep. Events	High	30	33.33
12	Lone Assailant	9+ Prep. Events	High	18	11.11
13	2-5 Offenders	9+ Prep. Events	High	32	3.13

A Row ID column and color were added above to ease in some of the comparisons. Here, we focused on Lone Assailants. As indicated by the CACC results, lone assailants with no known preparatory events and low level of weapon sophistication in their attack were 100 percent successful in their attacks (Row ID #1). Once the level of weapon sophistication increases to moderate in Row ID #3, the likelihood of success does reduce to about 92 percent. Then, if we move down to Row ID #9, we see if lone assailants have been 1-8 preparatory events (below the average) and a moderate level of weapon sophistication, the likelihood of success decreases to about 38 percent.

Note: The example used throughout this How To is a reduced variable list. Variable inclusion and categorization should be completed in respect to extant literature and the theoretical framework being used in the study.