Example: Why do some people vote, while others do not?
Hypothesis: People’s gender makes a difference. Females are more likely to vote than males.
Measure the relationship between gender and voting in elections.
Both gender and voting are nominal-level variables. → use cross-tabulation analysis.
How do we proceed to test the hypothesis?

1. **Separate cases into groups based on their values for the IV**

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Voted</strong></td>
<td>400</td>
<td>400</td>
</tr>
<tr>
<td><strong>Did not vote</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>400</td>
<td>400</td>
</tr>
<tr>
<td></td>
<td>(100%)</td>
<td>(100%)</td>
</tr>
<tr>
<td></td>
<td>N=800</td>
<td></td>
</tr>
</tbody>
</table>
2. Compare the values of the dependent variable for those groups
3. Decide whether the values for the DV are different for the different groups

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voted</td>
<td>100</td>
<td>320</td>
</tr>
<tr>
<td>(25%)</td>
<td>(80%)</td>
<td></td>
</tr>
<tr>
<td>Did not vote</td>
<td>300</td>
<td>80</td>
</tr>
<tr>
<td>(75%)</td>
<td>(20%)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>400</td>
<td>400</td>
</tr>
<tr>
<td>(100%)</td>
<td>(100%)</td>
<td></td>
</tr>
<tr>
<td>N=800</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4. Are females more likely to vote than males?
Example: Interest in the presidential election and voter turnout.

Hypothesis: “Those who care about the presidential election are more likely to go to the polls than those who do not.”

The independent variables is interest in the presidential election.

The dependent variable is voter turnout in the 2000 election.

Measure the relationship between *interest in the election* and *voter turnout*.

Open the ANES 2008 data
Use v085001b to measure interest in the election and v085036x to measure voter turnout in the 2000 election.

First, let’s see a frequency distribution of these two variables.

Syntax commands
freq var v085001b v085036x.
A Cross-Tabulation Table

- Syntax commands for a cross-tabulation table:
  crosstabs tables v085036x by v085001b.
- The order of the variable is important: “the dependent variable explained by the independent variable.”
- That is, the dependent variable is followed by the independent variable.
A Cross-Tabulation Table

- We use the percentages for interpreting the table.
- Add percentages to the table.
- Syntax commands
  
  \texttt{crosstabs tables v085036x by v085001b}
  
  \texttt{/cells count column.}

- Do the data support hypothesis?
Example: marital status and voter turnout.

Hypothesis: “Married people are more likely to go to the polls than unmarried people.”

The independent variables is marital status.

The dependent variable is voter turnout in the 2008 election.

Measure the relationship between marital status and voter turnout.
Produce a cross-tabulation table.

crosstabs tables v085036x by v083216x
/cells count column.

Do the data support the hypothesis?
Example: Income and preference on government service and spending.

Hypothesis: High income people are more likely to support in a decrease in government service and spending than low income people. Low income people are more likely to support an increase in government service and spending than high income people.

The independent variables is household income.

The dependent variable is support for government spending.
Use v083249 to measure household income. and v083108x to measure support for government service and spending.

Make a cross-tabulation table.

Too many rows and columns!

Use the command recode and reduce the number of categories in the variables.
Syntax commands to produce new variables with fewer categories

recode v083108x (1,2,3=1)(4=2)(5,6,7=3)(-9, -8, -1=sysmis) into service.
execute.
recode v083249 (1 thru 12=1)(13,14,15,16,17=2)(18 thru 25 =3) into income.
execute.

Make a cross-tabulation table with new variables

crosstabs tables service by income
/cells count column.
Give new value labels

value labels income 1 "Low" 2 "Middle" 3 "High".
value labels service 1 "More" 2 "Same" 3 "Less".

Rerun the code for a cross-tabulation table.
crosstabs tables pid by income /cells count column.
Assignment

- Why did people choose Barack Obama in the 2008 presidential election? Why did people vote for John McCain?
- Formulate a hypothesis to answer this question and test it using a cross-tabulation table and the ANES 2008 data set.