

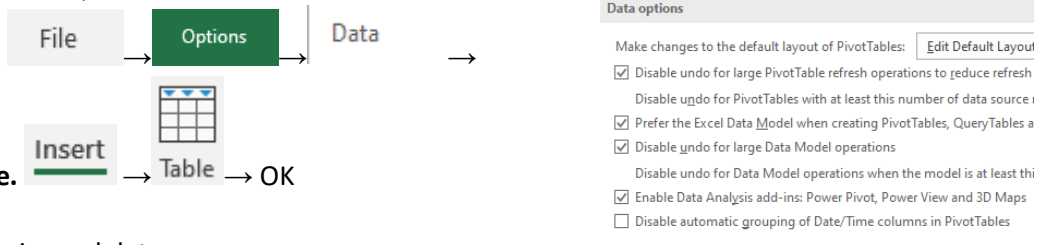


Instructions. Download the Young People Survey data and create the needed variables.

Data: This dataset was collected from 1010 statistics students and their friends at the Comenius University in Bratislava. The survey was administered in Slovak language and later translated into English.

1. Download the dataset and questionnaire from <https://github.com/AMDeLouize/Statistics-Workshops>. Save the dataset to the computer.

2. Turn on Data add ins.

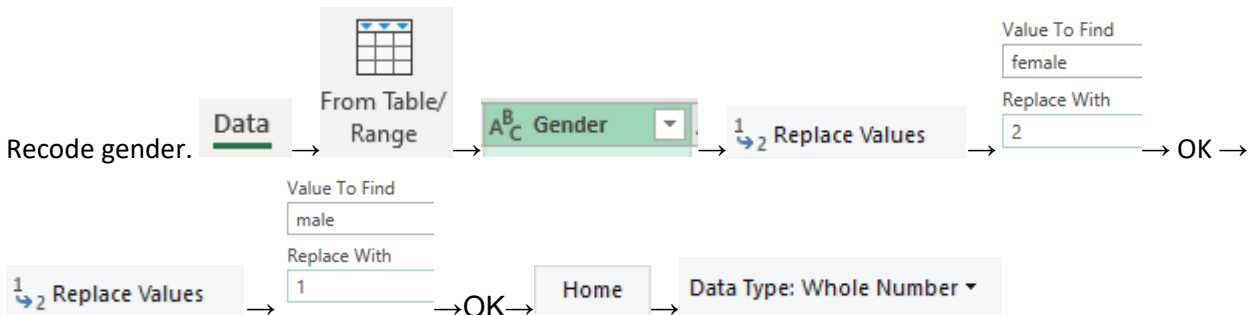


3. Turn data into a table.

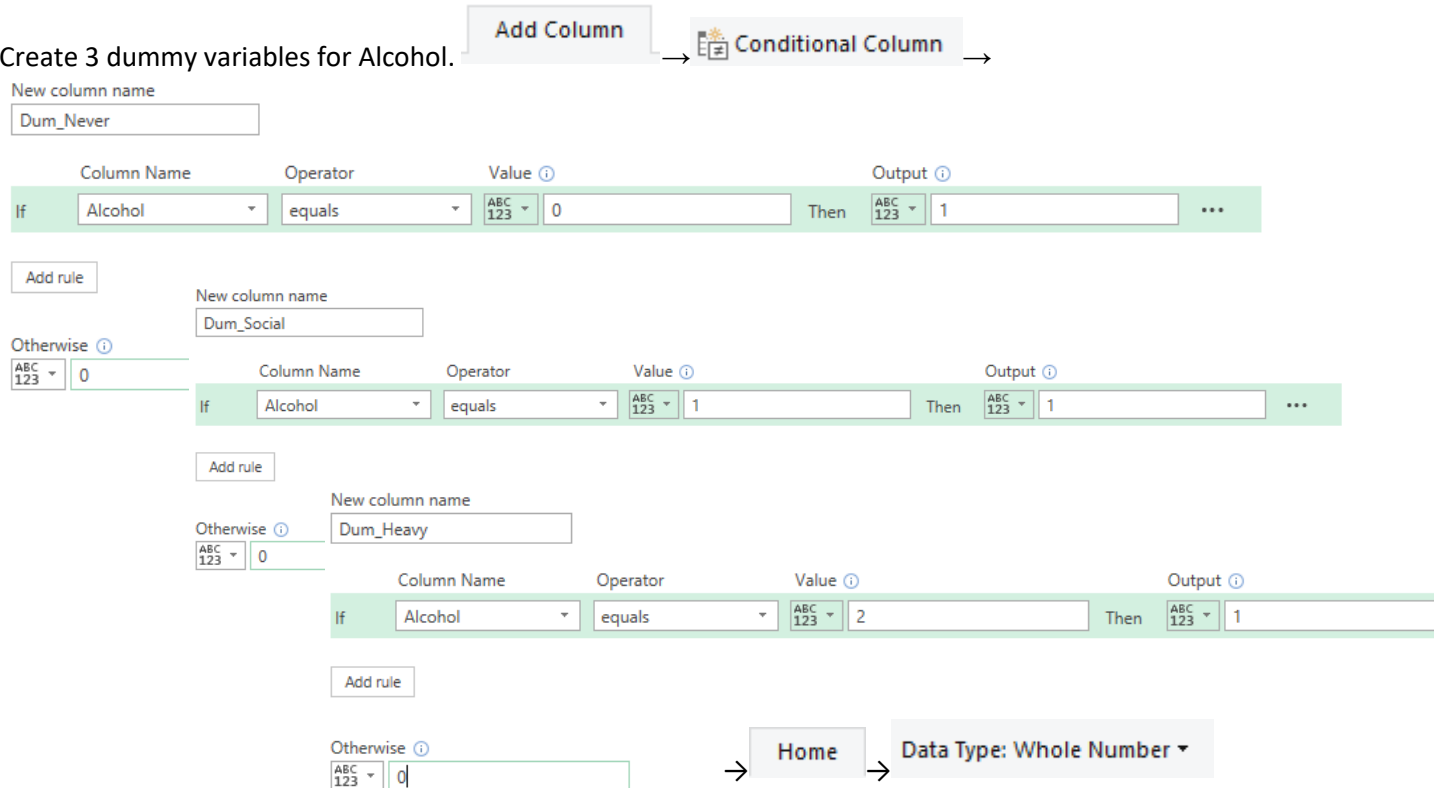


4. Review the questionnaire and data.

5. Recode gender.



6. Create 3 dummy variables for Alcohol.



Instructions. Run descriptive statistics for categorical/nominal variables (Alcohol and Gender) and for continuous/ordinal variables (Music, Movies, Pop, Finances, Age, Regret, Charity, Friends).

1. Create new tab named Descriptives. Hold Ctrl and select the variables listed above. Copy and Paste these columns to the new Descriptives tab.

2. Turn data into a table. → → OK
3. Move Alcohol and gender to the end of the table by copying and pasting or create a different table for them.
4. Descriptives (sample size, mean, median, standard deviation, range, skewness, and kurtosis) for continuous variables.

→ →

Click this up arrow then select the columns Music, Pop, Movies, Regret, Charity, Friends, and Age.

5. Run descriptive statistics (sample size, frequency/percentage, skewness, kurtosis) for categorical/nominal variables (Alcohol and Gender). Create the following below the results for the continuous variables (Note. In my workbook Alcohol is in column J and Gender is in column K, you may need to change these letters if your variables are in a different location):

Alcohol			Gender		
		$n/\text{Sample Size/ Count}$			$n/\text{Sample Size/ Count}$
		Skewness			Skewness
		Kurtosis			Kurtosis
Never	0	Frequency	Man	1	Frequency
		Percentage			Percentage
Social	1	Frequency	Woman	2	Frequency
		Percentage			Percentage
Heavy	2	Frequency			
		Percentage			

Percentage = $\text{Frequency}/n \times 100$

$= \text{COUNTIF}(K:K, 1) / \text{COUNT}(J:J) * 100$

In my worksheet "COUNTIF(K:K, 1)" is in R22 and "=COUNT(J:J)" is in R19



Instructions. Download the Young People Survey data and create descriptive data of the relevant variables. Then, check the assumptions for analysis.

1. Copy table of continuous variables to the Assumptions tab.
2. Normality. Can be evaluated with the Skewness (want values between -2 and 2) and Kurtosis (want values between -7 and 7) values obtained during descriptives or by visually evaluating histograms. Create histograms using the following:

- a. → → → OK →

The Histogram dialog box shows: Input Range: \$A:\$A, Bin Range: (empty), Labels: checked, Output options: Output Range: \$J:\$J, New Worksheet Ply: (empty), New Workbook: (empty), Pareto (sorted histogram): unchecked, Cumulative Percentage: unchecked, Chart Output: unchecked.
- b. Select column with frequency → → → 2D Clustered Column (First Option)
- c. Which variable is not normally distributed? _____
- d. Transform this variable so that the categories are “Likes Music A Lot” (Strongly Agree – 5) and “Does Not Like Music a Lot” (everything else – 1 to 4)

- i. Go to data worksheet → Select table () → → → →

→

New column name
Music_Dichotomized

	Column Name	Operator	Value	Output
If	Music	equals	1	1
Else If	Music	equals	2	1
Else If	Music	equals	3	1
Else If	Music	equals	4	1
Else If	Music	equals	5	2

Add rule

- e. Re-run skewness/kurtosis or create a histogram with this variable to make sure it is sufficiently normal now.
3. Independence – Review the dataset. Is every participant only in it once?
 4. Homoscedasticity (Equality of Variances) and Linearity
 - a. Select both Finances and Pop by holding down the Ctrl button → → → Scatter
 5. Multicollinearity – Run correlations between all x and y variables (See next activity), all should be < .9





Instructions. Use the Young People Survey data to perform a correlation in Excel.


Data. This dataset was collected from 1010 statistics students and their friends at the Comenius University in Bratislava. The survey was administered in Slovak language and later translated into English.


Objective. To investigate the relatedness of music and movie preferences.

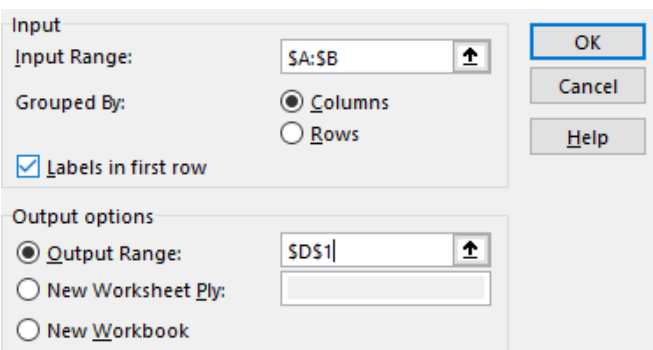
Hypothesis 1. People who enjoy music will also tend to enjoy movies.

Hypothesis 2. _____

- Optional: Look at the questionnaire and variables and come up with a second hypothesis of the relatedness of two ordinal or continuous variables (e.g., Could see if a certain type of music is correlated with a certain type of movie preference).
- Copy Music_Dichotamized and Movies to the correlation tab.

3. Turn data into a table.  → Table → OK

4.  → Data Analysis → Correlation →



The dialog box shows the following settings:

- Input Range: \$A:\$B
- Grouped By: ☒ Columns
- ☒ Labels in first row
- Output options:
 - ☒ Output Range: \$D\$1
 - ☐ New Worksheet Ply:
 - ☐ New Workbook

Buttons: OK, Cancel, Help

- Create the following table below your correlation results but use the square your correlation results are in for the first equation (instead of E3) and the table your t-value is in for the second equation (instead of E7).

t-value conversion:	$=E3*SQRT(1009)/SQRT(1-E3^2)$
p-value:	$=T.DIST.RT(E7, 1010)$

- Now repeat with your hypothesis 2 variables.
- Write a sentence about the conclusion for each hypothesis:

Hypothesis 1: _____

Hypothesis 2: _____



Instructions. Use the Young People Survey data to perform a *t*-test in Excel.

Data. This dataset was collected from 1010 statistics students and their friends at the Comenius University in Bratislava. The survey was administered in Slovak language and later translated into English.

Objective. To investigate whether men and women differ in their movie and music preferences.

Hypothesis 1. Women will enjoy romantic movies more than men.

Hypothesis 2. _____

- Optional. Look at the questionnaire and variables and come up with a second hypothesis of the relatedness of one categorical and one ordinal or continuous variable (e.g., Could see if women like classical music more).

- Copy Romantic and Gender to a worksheet labeled t-test.

- Turn data into a table. → Table → OK

- Select Gender. → → Insert Slicer → → OK

- Select 1 on the slicer and copy the romantic column to a new column not adjacent to the current table. Now select 2 on the slicer and copy the romantic column to the column next to the newly copied romantic column. Label the second romantic column "Romantic2". This will represent romantic movie preference for females. After you will have two new columns:

Romantic	Gender		Romantic	Romantic2
4	2		2	4
3	2		2	3
2	2		3	2

- Data Analysis → **t-Test: Two-Sample Assuming Equal Variances** →

Input

Variable 1 Range:

Variable 2 Range:

Hypothesized Mean Difference:

☒ Labels

Alpha:

Output options

☒ Output Range:

☐ New Worksheet Ply:

☐ New Workbook

OK Cancel Help

- Now repeat with your hypothesis 2 variables.
- Write a sentence about the conclusion for each hypothesis:

Hypothesis 1: _____

Hypothesis 2: _____

Instructions. Use the Young People Survey data to perform an ANOVA in Excel



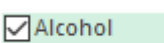

Objective. To investigate the relatedness of drinking alcohol on affinity for music.




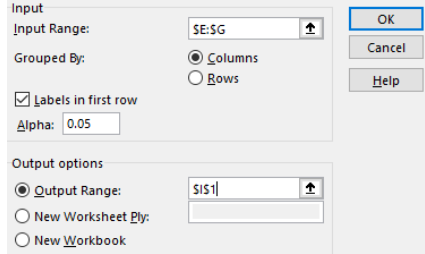
Hypothesis 1. People who drink a lot of alcohol will like pop music more than social drinkers, and social drinkers will like pop music more than non-drinkers.

Hypothesis 2. _____

- Optional. Look at the questionnaire and variables and come up with a second hypothesis of the relatedness of one nominal (categorical) and one continuous variable (e.g., Could see if people who spend more time online have certain music or movie tastes).
- Copy Alcohol and Pop to the One-Way ANOVA tab.

- Turn data into a table.  →  → OK

-  →  →  → Select 0 in slicer and copy and paste pop into new table.
Repeat for 1 and 2. Name Pop0, Pop1, Pop2. Select unfilter on the slicer ().

-  →  →  → 

- Now repeat with your hypothesis 2 variables.
- Write a sentence about the conclusion for each hypothesis:

Hypothesis 1: _____

Hypothesis 2: _____



Instructions. Use the Young People Survey data to perform an ANOVA in Excel.

Objective. To investigate the relatedness of alcohol and gender on entertainment spending.



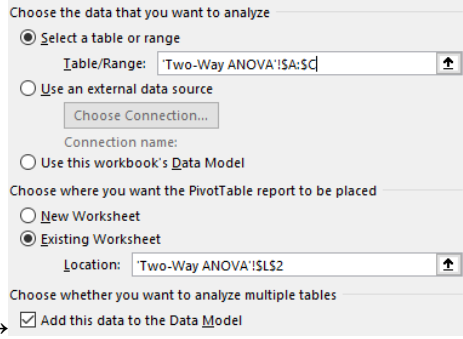
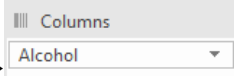
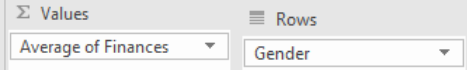
Hypothesis 1a. People who drink a lot of alcohol will spend more money on entertainment than social drinkers, and social drinkers will spend more on entertainment than non-drinkers.

Hypothesis 1b. Men will spend more money on entertainment than women.




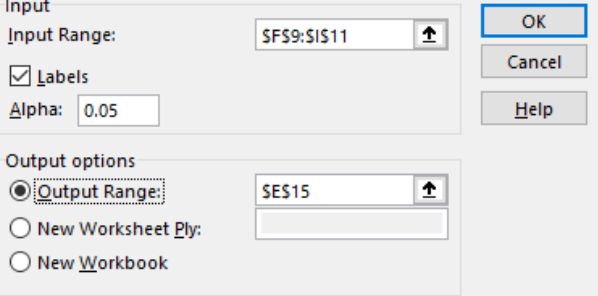
Hypothesis 1c. Drinking in men will lead to bigger differences in entertainment spending than it will for women such that men who drink more will spend significantly more on alcohol than women.

1. Copy Gender, Alcohol, and Finances to the Two-Way ANOVA tab.

2. Turn data into a table.  →  → OK

3.  →  →  →  →  →

	Never	Social	Heavy
Male	3.163636364	3.088983051	2.52991453
Female	3.388059701	3.141486811	2.692307692

4.  →  →  → 

5. Write a sentence about the conclusion for each hypothesis:

Hypothesis 1a: _____

Hypothesis 1b: _____

Hypothesis 1c: _____








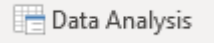
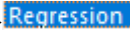
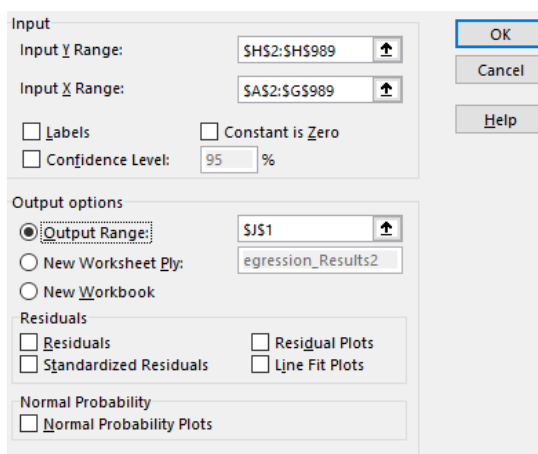
Instructions. Use the Young People Survey data to perform a multiple regression in Excel.

Objective. To investigate whether people who drink alcohol are less likely to save money than people who do not drink.

Hypothesis 1. People who drink a lot and drink socially will save less money than people who do not drink alcohol when controlling for age, gender, the propensity to give to charity, the number of friends, and the propensity for regret.

1. Copy Finances, Age, Gender, Charity, Friends, Regret, Dum Social, and Dum Heavy to a worksheet labeled Multiple Regression. Cut and paste Finances to the last column of the table.

2. **Turn data into a table.**  →  → OK
3. Select "Blanks" and de-select all other values in the column filters and delete rows with missing values.

4.  →  →  → 

5. Write a sentence about the conclusion for the hypothesis:

Hypothesis 1: _____

6. Fill in the following APA table with the multiple regression results

Predictor	<i>B</i>	<i>SE B</i>	β
Male			
Age			
Charity			
Regret			
Friends			
Alcohol (Never Drinker vs.)			
Social Drinker			
Heavy Drinker			
$p < .05$. ** $p < .01$. *** $p < .001$.			

