# QUIZ – Week 3

## Question 1

Parte superior do formulário

**Fill in the missing words. Choose from: predictor, identifier, outcome.**

The variable of primary interest that we wish to be able to predict is known as the **outcome** variable. The variable, or variables that we use to try and achieve this is/are known as the **predictor** variable(s).

**Correct**

Parte inferior do formulário

## Question 2

Parte superior do formulário

**When we are looking at relationships between categorical variables, we …..**

**Select all the answers you think are correct.**

**….. can either use separate bar charts or side by side bar charts.**

**….. use separate bar charts to try and determine if there are differences in overall shape for the different groups that make up the predictor variable.**

**….. use side-by-side bar charts to highlight differences between corresponding categories.**

….. can only work with two categorical variables at a time.

**Correct**

The first statement is **TRUE**. It is best to look at both options since they often give us different insights.

The second statement is **TRUE**. If the overall shapes are the same, the predictor variable has no real effect.

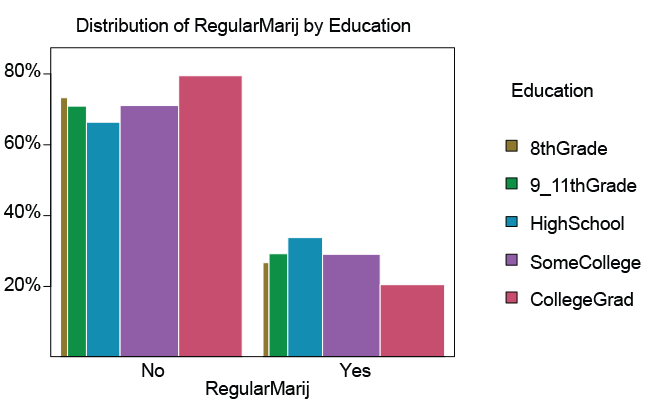
The third statement is **TRUE**. When we detect differences in overall shape in the separate bar charts, side-by-side bar charts allow us to focus our attention on the individual group differences.

Parte inferior do formulário

## Question 3

Parte superior do formulário

**The following statements refer to this plot of responses to the regular-marijuana use question by the highest level of education attainment from the NHANES1000 data set.**



**Participants were asked whether they had been or are a regular user of marijuana. Regular use: at least once a month for a year.**

One of the statements is false. Select the **FALSE** statement.

The 8thGrade bars are very thin because not many people were in this category.

**RegularMarij** is the outcome variable and **Education** is the predictor variable.

**There is no relationship between regular marijuana use and the level of education.**

Those who had graduated from college were the least likely to have said “yes” to the question about using marijuana regularly.

The group whose highest level of education was High School had the highest percentage saying “yes” to using marijuana regularly.

**Correct**

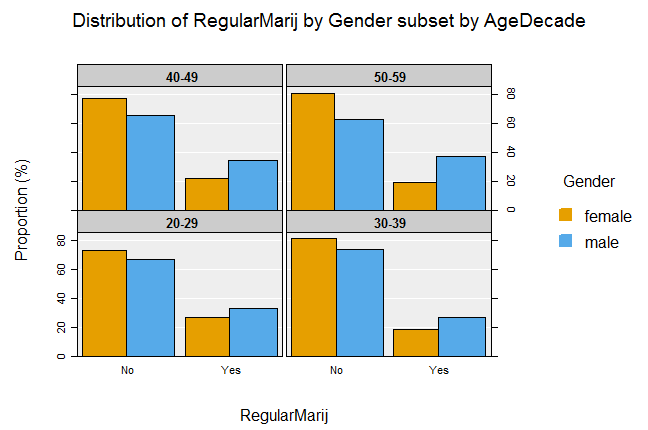
This statement is **FALSE** – we would expect the heights of the “yes” bars for each level of education to be the almost the same if this statement was true.

Parte inferior do formulário

## Question 4

Parte superior do formulário

**The following statements refer to this plot of regular marijuana use by gender & age decade from the NHanes 2009-2012 data set.**



One of the statements is false. Select the **FALSE** statement.

For all age groups, the males were more likely to say “yes“ to the marijuana question than the females.

**RegularMarij is the predictor variable and AgeDecade is the outcome variable.**

The age group where people were least likely to say “yes” to the marijuana question was the 30-39 age group.

Males in the 50-59 age group were almost twice as likely as the females to say “yes” to the marijuana question.

The age group where the responses for the males and females are most similar is the 20-29 age group.

The greatest difference in regular marijuana use between genders is in the 50-59 age group.

**Correct**

This answer is **FALSE** - **RegularMarij** is the outcome variable. It is the distribution of **RegularMarij** that is being plotted each time

Parte inferior do formulário