**Week 2 Quiz**

**Question 1**

Parte superior do formulário

**Fill in the missing words. Choose from: levels, groups, entities, numbers, measurements.**

Numeric variables are those like Age (in years) that we can think of as **measurements**. A categorical variable, like AgeDecade, defines **groups** that our  **entities** belong to.

Parte inferior do formulário

**Question 2**

Parte superior do formulário

**Which of the following statements concerning categorical variables are TRUE?**

Select all the answers you think are correct.

**Pie charts and stacked bar charts are not very good at showing changes and differences.**

The order categories are placed in by default in bar charts is generally the order they first appear in the relevant column of the data.

**Summarising of categorical variables is done well by using the percentages of individuals falling into each category.**

If a categorical variable is nominal, any order used in a bar chart is equally useful.

**An ordinal variable is a set of groupings that has a natural order.**

**Nominal variables contain groups that have no natural order and can often usefully be plotted in the order of their frequencies.**

**Correct**

Parte inferior do formulário

The first statement is **TRUE**. Pie charts are very difficult to use if you want to assess group differences, unless there are only 2 groups. Stacked bar charts make comparisons between different groups difficult.

The second statement is **FALSE**. The (default) order for categorical variables is normally alpha-numeric (numbers and then letters alphabetically).

The third statement is **TRUE**. Using percentages gives a clear impression of the order of the groups by size as well as giving a consistent indication of the fraction of all the entities that belong to that group.

The fourth statement is **FALSE**. The categories should be ordered from largest to smallest (or smallest to largest) so that comparisons can be made by simply looking across the plot (and the group labels), or in some order (e.g. alphabetic) that lets us find particular groups quickly.

The fifth statement is **TRUE**. **AgeGroup** has a natural order whereas **MaritalStatus** does not.

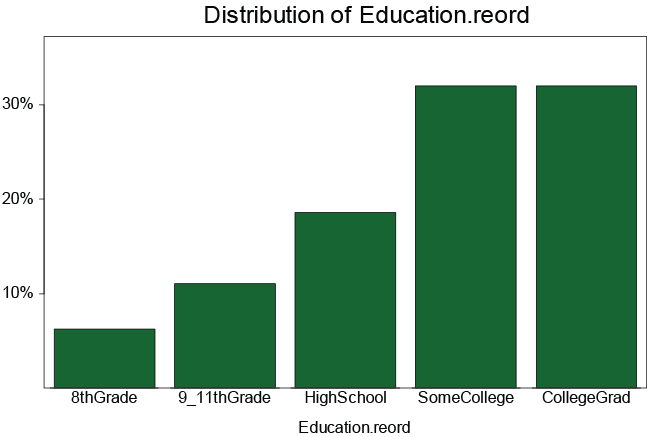
The sixth statement is **TRUE**. By plotting in a bar chart by frequencies, comparisons can be easily made.

**Question 3**

Parte superior do formulário

**The following statements refer to the plot of Education from the NHanes2009-2012 data set.**

Education records the *highest level* of Education the person has attained.



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

*About one third* of people are in the **SomeCollege** category.

The percentage of people who have **SomeCollege** education is *quite a lot bigger* than the percentage whose highest level is graduating from high school (**HighSchool**).

***Most people* are in the SomeCollege category.**

**(**This statement is **FALSE** – About 30% of the people were in the **SomeCollege** category. 70% of people are outside of this category.)

The *majority* of people have obtained *at least some* college education.

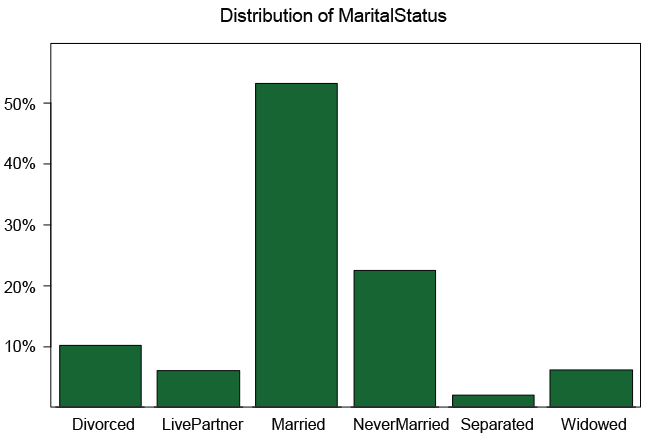
The sum of the bar heights should be 100%.

Parte inferior do formulário

**Question 4**

Parte superior do formulário

**The following statements refer to this plot of MaritalStatus from the NHanes2009-2012 data set.**



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

*More than twice as many* people are **Married** as **NeverMarried**.

*More* people are **Divorced** than **Widowed**.

**About 1 in 3 marriages end in divorce.**

**MaritalStatus** is a *categorical* variable.

**Correct**

This statement is **FALSE** - This type of information is not visible on the graph.

Parte inferior do formulário