Exploring Two Variables: Explanatory and Response

Learning Objective: Classify a data analysis situation (involving two variables) according to the "role-type classification," and state the appropriate display and/or numerical measures that should be used in order to summarize the data.

While it is fundamentally important to know how to describe the distribution of a single variable, most studies pose research questions that involve exploring the relationship between **two** variables using the collected data.

Here are a few examples of such research questions with the two variables highlighted:

Examples

1. Is there a relationship between **gender** and **test scores** on a particular standardized test?

   Other ways of phrasing the same research question:

   - Is performance on the test related to gender?
   - Is there a gender effect on test scores?
   - Are there differences in test scores between males and females?

2. How is the **number of calories** in a hot dog related to (or affected by) the **type of hot dog** (beef, meat or poultry)? In other words, are there differences in the number of calories among the three types of hot dogs?
3. Is there a relationship between the **type of light** a baby sleeps with (no light, night-light, lamp) and whether or not the child develops **nearsightedness**?
4. Are the **smoking habits** of a person (yes, no) related to the person's **gender**?
5. How well can we predict a student's freshman year **GPA** from his/her **SAT score**?

6. What is the relationship between driver's **age** and sign legibility **distance** (the maximum distance at which the driver can read a sign)?
7. Is there a relationship between the **time** a person has practiced driving while having a learner's permit, and **whether or not this person passed the driving test**?
8. Can you predict a person's **favorite type of music** (classical, rock, jazz) based on his/her **IQ level**?

In most studies involving two variables, each of the variables has a role. We distinguish between:

- the **explanatory** variable (also commonly referred to as the **independent variable**)-—the variable that claims to explain, predict or affect the response; and
- the **response** variable (also commonly referred to as the **dependent variable**)-—the outcome of the study.

Typically the **explanatory** (or independent) variable is denoted by **X,** while the **response** (or dependent) variable is denoted by **Y**.

Explanatory and Response Variables

In this course, we will use the terms **explanatory** and **response** variables, instead of **independent** and **dependent** variables.

Now let's go back to some of the examples and classify the two relevant variables according to their roles in the study:

Example: 1

We want to explore whether the outcome of the study—the score on a test—is affected by the test-taker's gender.

Therefore:

- **Gender** is the **explanatory**
- **Test score** is the **response**

Example: 3

In this study we explore whether the nearsightedness of a person can be explained by the type of light that person slept with as a baby.

Therefore:

- **Light type** is the **explanatory**
- **Nearsightedness** is the **response**

Example: 5

Here we are examining whether a student's SAT score is a good predictor for the student's GPA freshman year.

Therefore:

- **SAT score** is the **explanatory**
- **GPA of freshman year** is the **response**

Example: 7

Here we are examining whether a person's outcome on the driving test (pass/fail) can be explained by the length of time this person has practiced driving prior to the test.

Therefore:

- **Time** is the **explanatory**
- **Driving test outcome** is the **response**

Scenario: Example Research Questions with Two Variables

Now, using the same reasoning, the following exercise will help you to classify the two variables in the other examples. Refer to the eight examples listed below when answering the questions.

Examples

1. Is there a relationship between **gender** and **test scores** on a particular standardized test?

   Other ways of phrasing the same research question:

   - Is performance on the test related to gender?
   - Is there a gender effect on test scores?
   - Are there differences in test scores between males and females?

2. How is the **number of calories** in a hot dog related to (or affected by) the **type of hot dog** (beef, meat or poultry)? In other words, are there differences in the number of calories among the three

types of hot dogs?

3. Is there a relationship between the **type of light** a baby sleeps with (no light, night-light, lamp) and whether or not the child develops **nearsightedness**?
4. Are the **smoking habits** of a person (yes, no) related to the person's **gender**?
5. How well can we predict a student's freshman year **GPA** from his/her **SAT score**?
6. What is the relationship between driver's **age** and sign legibility **distance** (the maximum distance at which the driver can read a sign)?
7. Is there a relationship between the **time** a person has practiced driving while having a learner's permit, and **whether or not this person passed the driving test**?
8. Can you predict a person's **favorite type of music** (classical, rock, jazz) based on his/her **IQ level**?

Learn By Doing

1/1 point (graded)

In example 2, is the number of calories the explanatory or response variable?

- ○ explanatory
- ◉ response ✔

Answer

Correct:

In this study we are exploring whether the type of hot dog affects the number of calories it has. Thus, the number of calories is the response variable.

Submit

Learn By Doing

1/1 point (graded)

In example 2, is the type of hot dog the explanatory or response variable?

- ◉ explanatory ✔
- ○ response

Answer

Correct:
In this study we are exploring whether the type of hot dog affects the number of calories it has. Thus, the type of hot dog is the explanatory variable.

Submit

Learn By Doing

1/1 point (graded)
In example 4, is gender the explanatory or response variable?

- (•) explanatory ✔
- ( ) response

Answer
Correct:
Note that in most studies involving gender, it will be the explanatory variable. We are not trying to see how a person's smoking habits affect the person's gender, but rather explore whether gender is one of the factors that can explain whether the person is a smoker or not.

Submit

Learn By Doing

1/1 point (graded)
In example 4, is smoking habit the explanatory or response variable?

- ( ) explanatory
- (•) response ✔

Answer
Correct:
If a person's gender is one of the factors that can explain whether the person is a smoker or not, then smoking habits is the response variable.

Submit

Learn By Doing

1/1 point (graded)

In example 6, is age the explanatory or response variable?

- (•) explanatory ✔
- ( ) response

Answer

Correct:

In this study we are examining how legibility distance changes with age. Thus, we can say that age affects legibility distance, making it the explanatory variable.

Submit

Learn By Doing

1/1 point (graded)

In example 6, is legibility distance the explanatory or response variable?

- ( ) explanatory
- (•) response ✔

Answer

Correct:

Since we are examining how age affects legibility distance, legibility distance is the response variable.

Submit

Learn By Doing

1/1 point (graded)

In example 8, is favorite type of music the explanatory or response variable?

- ( ) explanatory

response ✔

Answer
Correct:
Since the purpose of this study was to explore whether a person's IQ level can explain or help predict the person's favorite type of music, type of music is the response variable.

Submit

Learn By Doing

1/1 point (graded)
In example 8, is IQ level the explanatory or response variable?

explanatory ✔

response

Answer
Correct:
Since the purpose of this study was to explore whether a person's IQ level can explain or help predict the person's favorite type of music, IQ level is the explanatory variable.

Submit

Many Students Wonder ...

Question: Is the role classification of variables always clear? In other words, is it always clear which of the variables is the explanatory and which is the response?

Answer: No. There are studies in which the role classification is not really clear. This mainly happens in cases when both variables are categorical or both are quantitative. An example is a study that explores the relationship between students' SAT Math and SAT Verbal scores. In cases like this, any classification choice would be fine (as long as it is consistent throughout the analysis).