

Fundamentals of Statistics

Intellectus

The image features a white background with two teal-colored geometric shapes. On the left, there is a large teal trapezoid that tapers towards the right. On the right side, there is a smaller teal triangle that tapers towards the left. The word "Introduction" is centered between these two shapes.

Introduction

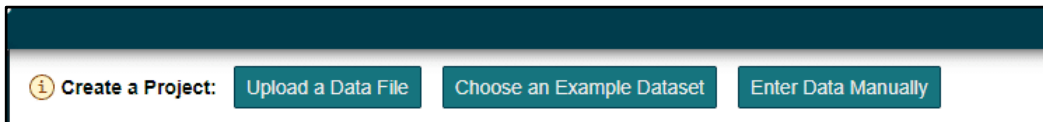
Opening Example

Choosing an example dataset

Throughout this module, you will be asked to open and run analyses on various example datasets supplied by Intellectus. If you are unsure of how to find an example dataset, follow the steps listed below.

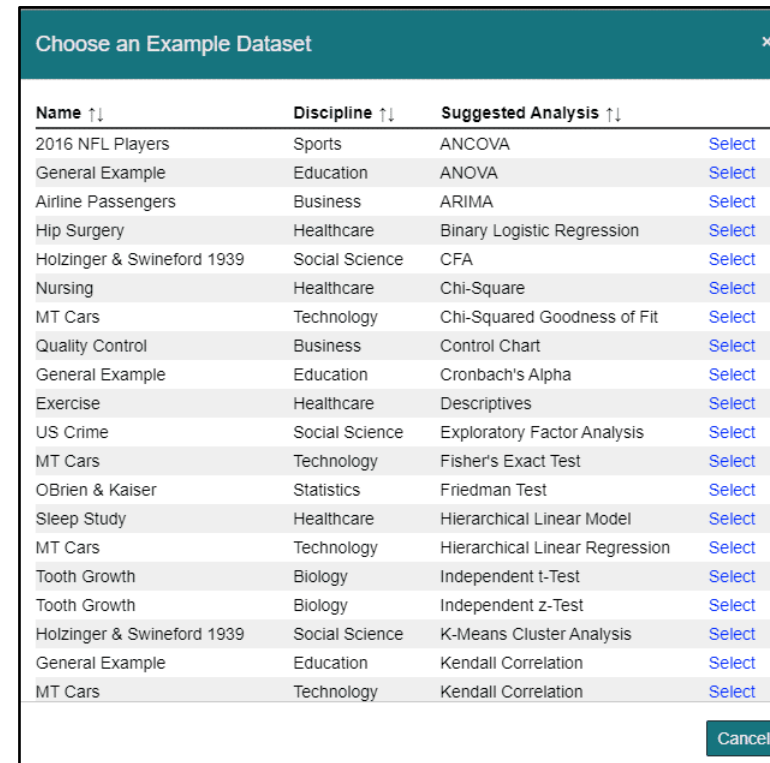
Step 1

Select “Choose and Example Dataset”



Step 2

Selecting the dataset



A screenshot of a dialog box titled 'Choose an Example Dataset'. It contains a table with three columns: 'Name', 'Discipline', and 'Suggested Analysis'. Each row represents a dataset with its corresponding discipline and a suggested analysis method. A 'Cancel' button is located at the bottom right of the dialog.

Name ↑↓	Discipline ↑↓	Suggested Analysis ↑↓	
2016 NFL Players	Sports	ANCOVA	Select
General Example	Education	ANOVA	Select
Airline Passengers	Business	ARIMA	Select
Hip Surgery	Healthcare	Binary Logistic Regression	Select
Holzinger & Swineford 1939	Social Science	CFA	Select
Nursing	Healthcare	Chi-Square	Select
MT Cars	Technology	Chi-Squared Goodness of Fit	Select
Quality Control	Business	Control Chart	Select
General Example	Education	Cronbach's Alpha	Select
Exercise	Healthcare	Descriptives	Select
US Crime	Social Science	Exploratory Factor Analysis	Select
MT Cars	Technology	Fisher's Exact Test	Select
OBrien & Kaiser	Statistics	Friedman Test	Select
Sleep Study	Healthcare	Hierarchical Linear Model	Select
MT Cars	Technology	Hierarchical Linear Regression	Select
Tooth Growth	Biology	Independent t-Test	Select
Tooth Growth	Biology	Independent z-Test	Select
Holzinger & Swineford 1939	Social Science	K-Means Cluster Analysis	Select
General Example	Education	Kendall Correlation	Select
MT Cars	Technology	Kendall Correlation	Select

There are three columns: the name of the dataset, discipline, and suggested analysis. For the most part, we will be looking at the suggested analysis column in order to determine which dataset to open.

Opening Example

Step 3

Details of the dataset

Data Description

General Example Dataset

This dataset consists of randomly generated data for testing various analyses. It consists of 300 rows and 25 variables that can be recoded to conduct most analyses within Intellectus Statistics.

Suggested Analysis

MANCOVA

- *MathScores* and *ScienceScores* as Dependent Variables
- *Group* as the Independent Variable
- *Reading1* as a Covariate

[Go Back](#) [Create Project with this Dataset](#)

After opening the dataset, you will be presented with a description of the dataset and details of the suggested analysis. Select "Create a Project with this Dataset."

Step 4

Levels of measurement

Levels of Measurement

Variable Name	Level of Measurement
Reading1	Scale
Reading2	Scale
Reading3	Scale
Reading4	Scale
ReadingPost1	Scale
Sex	Nominal
FavoriteColor	Nominal
Group	Nominal
MathScores	Scale
ScienceScores	Scale
TypicalLikert1	Scale
TypicalLikert2	Scale
TypicalLikert3	Scale

[Help Page](#) [Video Tutorial](#) [Continue](#)

Here you can manage the levels of measurement for each variable. For the examples presented in this module, these settings should be left alone.

Overview

What is Statistics?

Statistics is a branch of mathematics dealing with the collection, analysis, interpretation, and presentation of numerical data.



Data is present nearly everywhere in our daily lives.

- Daily commute time
- Student test scores
- The heights of a given population
- The number of steps taken by individuals
- The monetary values of all the cars on the street



Data Collection and Measurement

Data Collection and Measurement

Why do we measure and analyze data?



Consider the following scenario:

Imagine you are a scientist attempting to determine if milk is good or bad for humans. To determine this, you give it to your child Mike, who grows up big and strong with healthy bones. Milk must be good for you - just look how Mike turned out. But then you hear about your friend's child Johnny, who gets stomach aches and feels ill every time he drinks milk. Now it seems milk might not be good for humans after all. You don't really have a way of seeing exactly what the milk is doing in each child, so how can you make a definitive statement for all humans?

The answer is statistics.

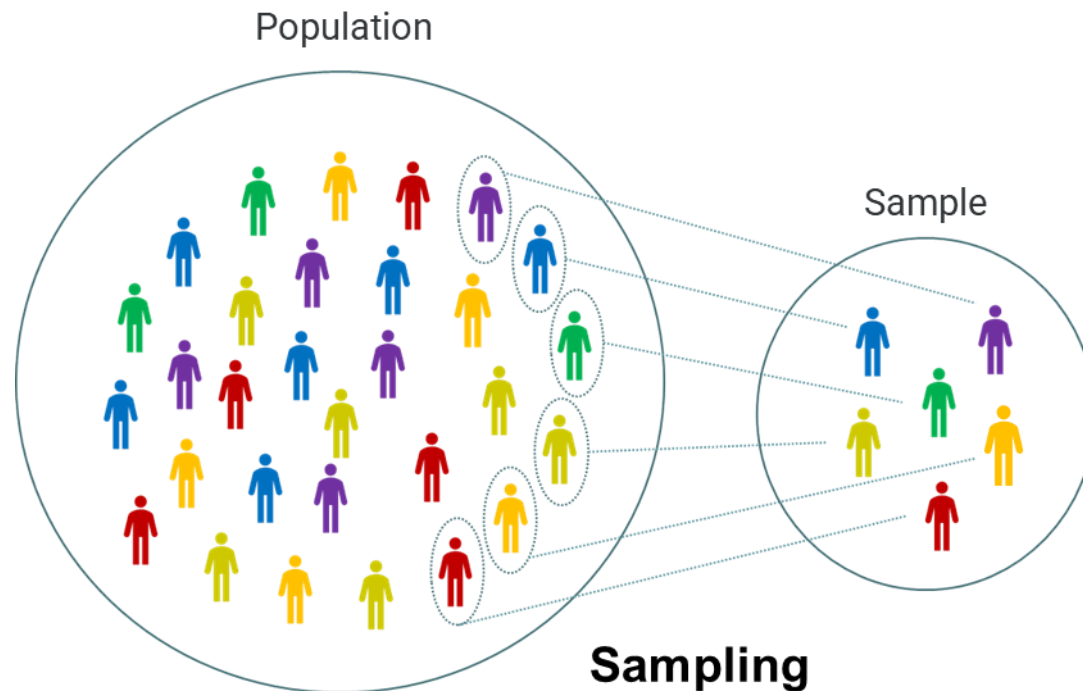
Milk will not affect everyone the same, so the key to determining if it should be considered healthy is to **measure** the effects across a large number of humans, and then **analyze** these results to see if there are any trends or differences.

What to Measure

It is impossible to measure everything.

Being the determined scientist that you are, you continue with your research on milk. You figure that if you can record the effects of milk on everyone, you will know for sure the answer to whether or not milk is healthy.

But this is not possible. You only have so much time and only know so many people. So instead, you observe the effects of milk on all the people you are able to (your friends and family, and anyone else willing to let you watch them drink milk). You have just engaged in something called **sampling**.



Sampling



Definition:

Sampling is the act of measuring and recording a pre-determined number of observations from a population or process.

In our milk example, the sample is anyone you were able to observe the effects of milk on. Below are some terms and facts associated with sampling.

- A single recorded measurement is called an **observation**.
- The number of observations you have is called the **sample size**.
- Generally, the larger the sample size, the more that will be revealed about the population as a whole. A good way to think about this is to imagine a painting. Do you know more about what the painting depicts if you can only see a tiny corner, or if you can see all but the edges?

Your measurement of the effect milk on Mike is an observation. If you manage to observe the effects of milk on 99 other people, you have a sample size of 100. The more people you can observe the effects of milk on, the more you will learn about whether it should be considered healthy.

Review

Summary

Term	Definition
Sampling	The act of measuring and recording a predetermined number of observations from a population or process.
Observation	Single recorded measurement.
Sample size	The number of observations.