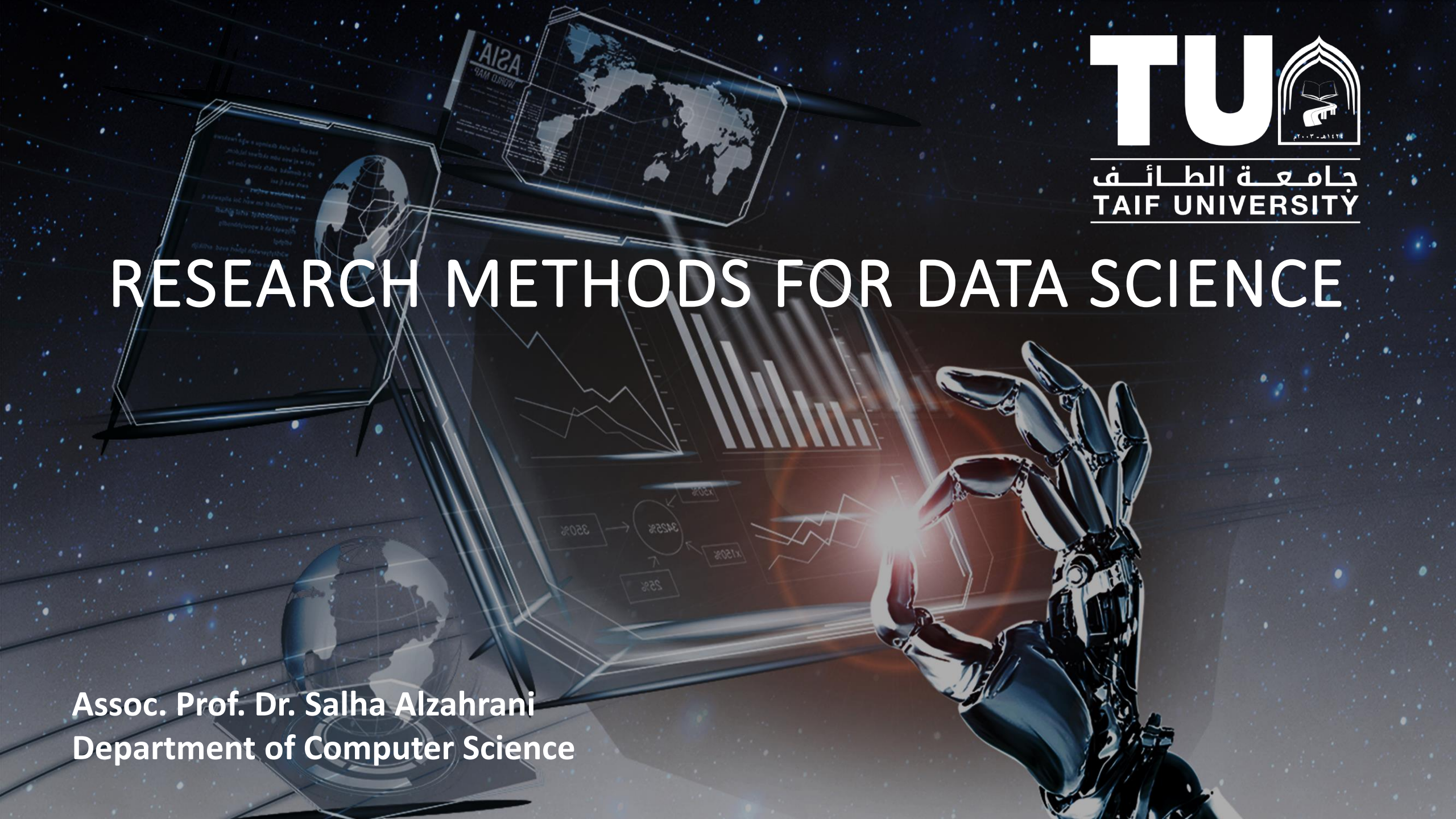


جامعة الطائف
TAIF UNIVERSITY

RESEARCH METHODS FOR DATA SCIENCE

Assoc. Prof. Dr. Salha Alzahrani
Department of Computer Science

STEP 1: Formulating a research problem

Part III

My Notes

~variables

~Hypothesis

What is a variable?

An image, perception or concept that is capable of measurement – hence capable of taking on different values – is called a **variable**. According to Kerlinger (1986: 27), 'A variable is a property that takes on different values. Putting it redundantly, a variable is something that varies ... A variable is a symbol to which numerals or values are attached.' Black and Champion (1976: 34) define variables as 'rational units of analysis that can assume any one of a number of designated sets of values'. A variable, then, is a concept that can be measured on any one of the four types of **measurement scale**, which have varying degrees of precision in measurement (measurement scales are discussed later in this chapter).

My Notes

The difference between a variable and a concept

My Notes

The main difference between a **concept** and a variable is measurability. Concepts are mental images or perceptions and therefore their meanings vary markedly from individual to individual, whereas variables are measurable, though, of course, with varying degrees of accuracy depending upon the measurement scale used. A concept as such cannot be measured, whereas a variable can be subjected to measurement by crude/refined or subjective/objective units of measurement.

Concepts are subjective impressions which, if measured as such, would cause problems in comparing responses obtained from different respondents.

According to Young (1966: 18):

Types of variables

A variable can be classified in a number of ways. The classification developed here results from looking at variables in three different ways (see Figure 5.1):

- the causal relationship;
- the study design;
- the unit of measurement.

My Notes

From the viewpoint of causal relationship

My Notes

In research terminology, change variables are called **independent variables**, outcome/effect variables are called **dependent variables**, the unmeasured variables affecting the cause-and-effect relationship are called **extraneous variables**, and the variables that link a cause-and-effect relationship are called **intervening variables**. To give a little more detail:

From the viewpoint of causal relationship

1. **Independent variable** – the cause supposed to be responsible for bringing about change(s) in a phenomenon or situation.
2. **Dependent variable** – the outcome or change(s) brought about by introduction of an independent variable.
3. **Extraneous variable** – several other factors operating in a real-life situation may affect or effect changes in the dependent variable. These factors, not measured in the study, may increase or decrease the magnitude or strength of the relationship between independent and dependent variables.
4. **Intervening variable** – sometimes called the confounding variable (Grinnell 1988: 203), it links the independent and dependent variables. In certain situations the relationship between an independent and a dependent variable cannot be established without the intervention of another variable. The cause, or independent, variable will have the assumed effect only in the presence of an intervening variable.

My Notes

From the viewpoint of the study design

My Notes

A study that examines association or causation may be a controlled/contrived experiment, a quasi-experiment or an *ex post facto* or non-experimental study. In controlled experiments the independent (cause) variable may be introduced or manipulated either by the researcher or by someone else who is providing the service. In these situations there are two sets of variables (see Figure 5.6):

- **Active variables** – those variables that can be manipulated, changed or controlled.
- **Attribute variables** – those variables that cannot be manipulated, changed or controlled, and that reflect the characteristics of the study population, for example age, gender, education and income.

From the viewpoint of the measurement unit

My Notes

From the viewpoint of the unit of measurement, there are two ways of categorising variables:

- whether the unit of measurement is categorical (as in nominal and ordinal scales) or continuous in nature (as in interval and ratio scales);
- whether it is qualitative (as in nominal and ordinal scales) or quantitative in nature (as in interval and ratio scales).

On the whole there is very little difference between categorical and qualitative, and between continuous and quantitative, variables. The slight difference between them is explained below.

From the viewpoint of the measurement unit

My Notes

Categorical variables are measured on nominal or ordinal measurement scales, whereas for **continuous variables** the measurements are made on either an interval or a ratio scale. There are three types of categorical variables:

- **constant variable** – has only one category or value, for example taxi, tree and water;
- **dichotomous variable** – has only two categories, as in male/female, yes/no, good/bad, heads/tails, up/down and rich/poor;
- **polytomous variable** – can be divided into more than two categories, for example religion (Christian, Muslim, Hindu); political parties (Labour, Liberal, Conservative); and attitudes (strongly favourable, favourable, uncertain, unfavourable, strongly unfavourable).

Types of measurement scales

- nominal or classificatory scale;
- ordinal or ranking scale;
- interval scale;
- ratio scale.

My Notes

Constructing hypothesis

My Notes

According to Grinnell (1988: 200):

> A hypothesis is written in such a way that it can be proven or disproven by valid and reliable data – it is in order to obtain these data that we perform our study.

From the above definitions it is apparent that a hypothesis has certain characteristics:

1. It is a tentative proposition.
2. Its validity is unknown.
3. In most cases, it specifies a relationship between two or more variables.

The functions of a hypothesis

While some researchers believe that to conduct a study requires a hypothesis, having a hypothesis is not essential, as already mentioned. However, a hypothesis is important in terms of bringing clarity to the research problem. Specifically, a hypothesis serves the following functions:

- The formulation of a hypothesis forces you to precisely specify what you want to find out about, thus bringing specificity and clarity to your study.
- The specificity and clarity needed to construct a hypothesis ensure you only collect the information you need, thereby providing focus to the study. This also enhances the validity of your study as it ensures you are measuring what you set out to measure.
- As it provides a focus, the construction of a hypothesis enhances objectivity in a study.
- The testing of a hypothesis enables you to specifically conclude what is true or what is false, thus enabling you to contribute towards theory formulation.

My Notes

Types of hypothesis

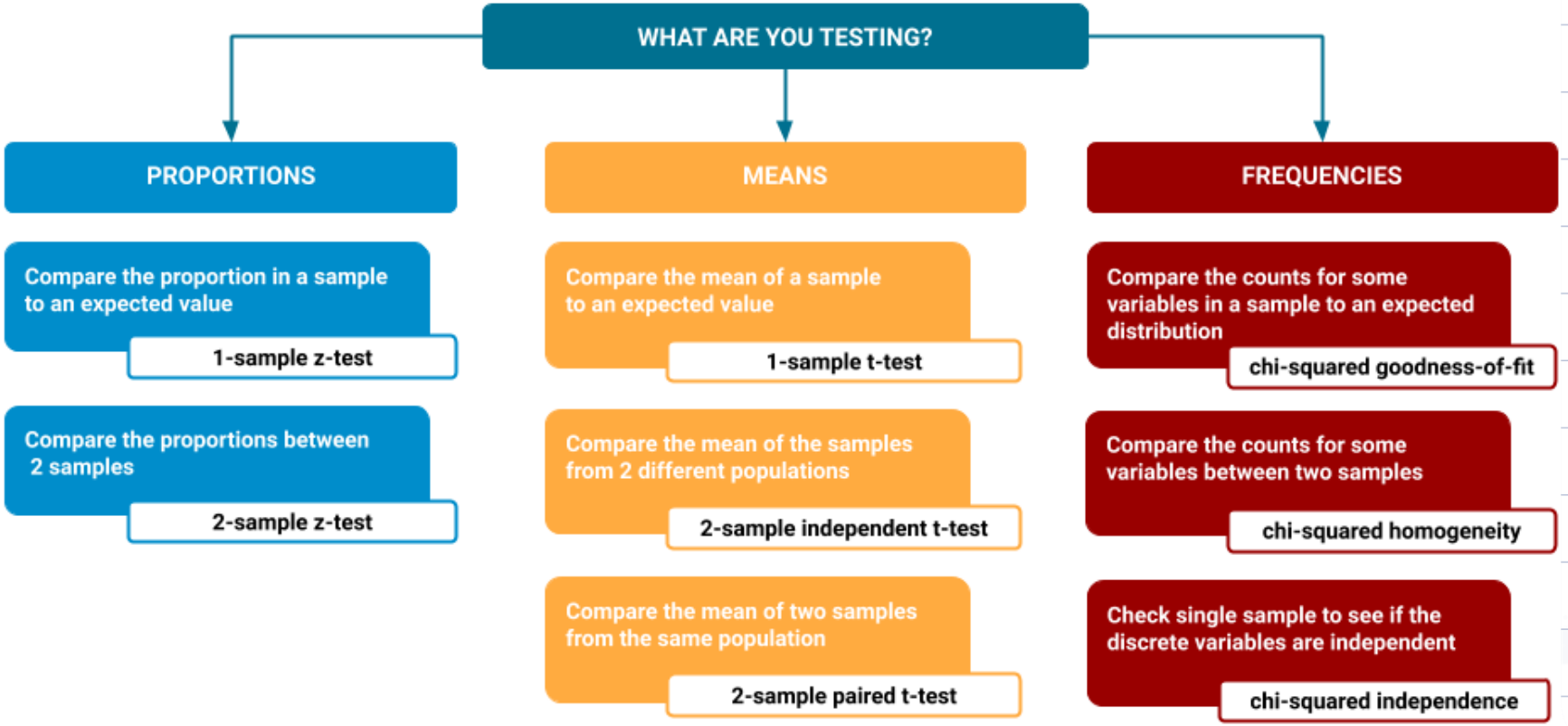
Theoretically there should be only one type of hypothesis, that is the research hypothesis – the basis of your investigation. However, because of the conventions in scientific enquiry and because of the wording used in the construction of a hypothesis, hypotheses can be classified into several types. Broadly, there are two categories of hypothesis:

1. research hypotheses;
2. alternative hypotheses.

Alternative hypothesis: The formulation of an alternative hypothesis is a convention in scientific circles. Its main function is to specify explicitly the relationship that will be considered as true in case the research hypothesis proves to be wrong. In a way, an alternative hypothesis is the opposite of the research hypothesis.

My Notes

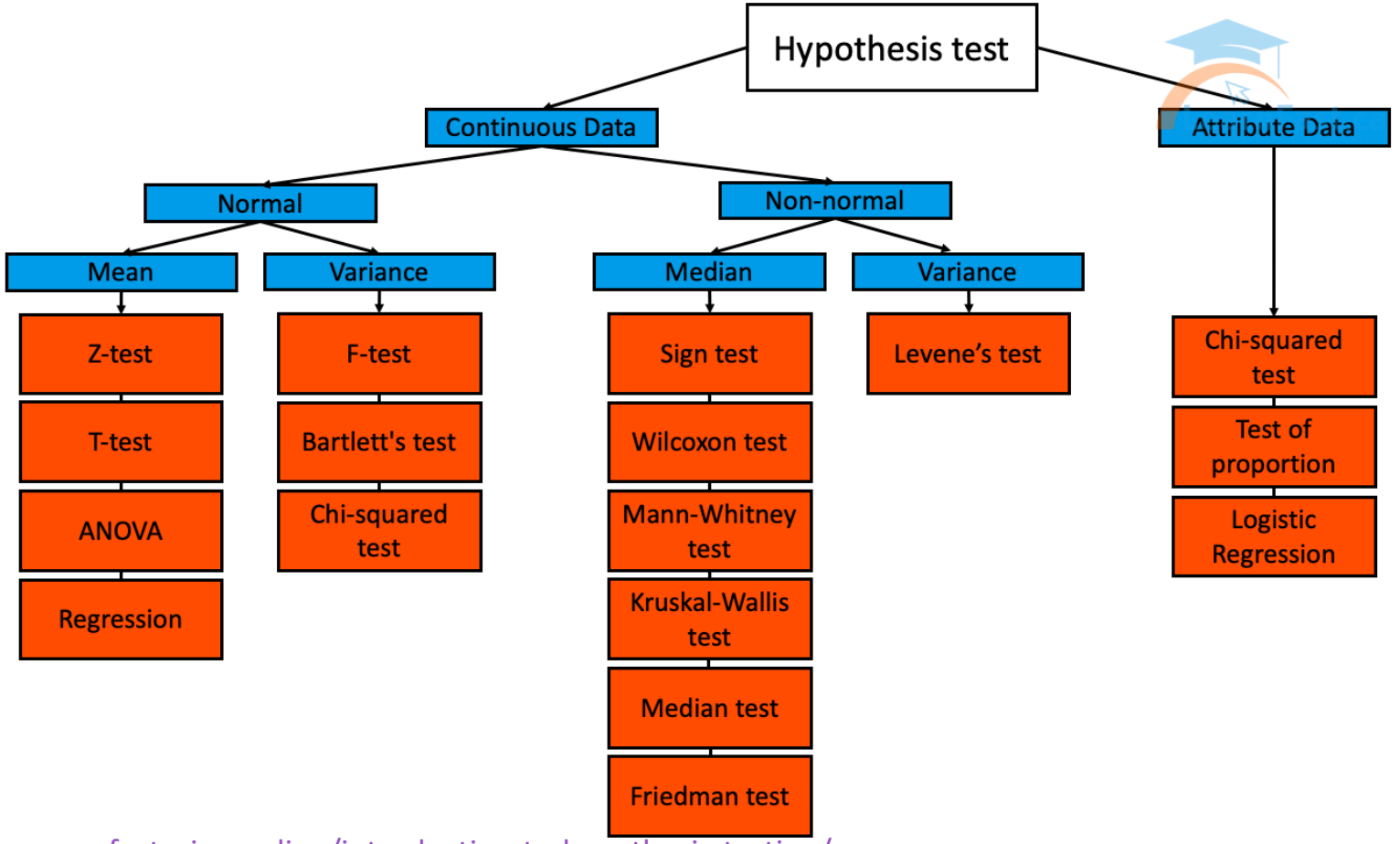
Testing of hypothesis



https://sonalake.com/latest/hypothesis-testing-of-frequency-based-samples/

My Notes

Testing of hypothesis



https://leanmanufacturing.online/introduction-to-hypothesis-testing/

My Notes

Errors in testing a hypothesis

As already mentioned, a hypothesis is an assumption that may prove to be either correct or incorrect. It is possible to arrive at an incorrect conclusion about a hypothesis for a variety of reasons. Incorrect conclusions about the validity of a hypothesis may be drawn if:

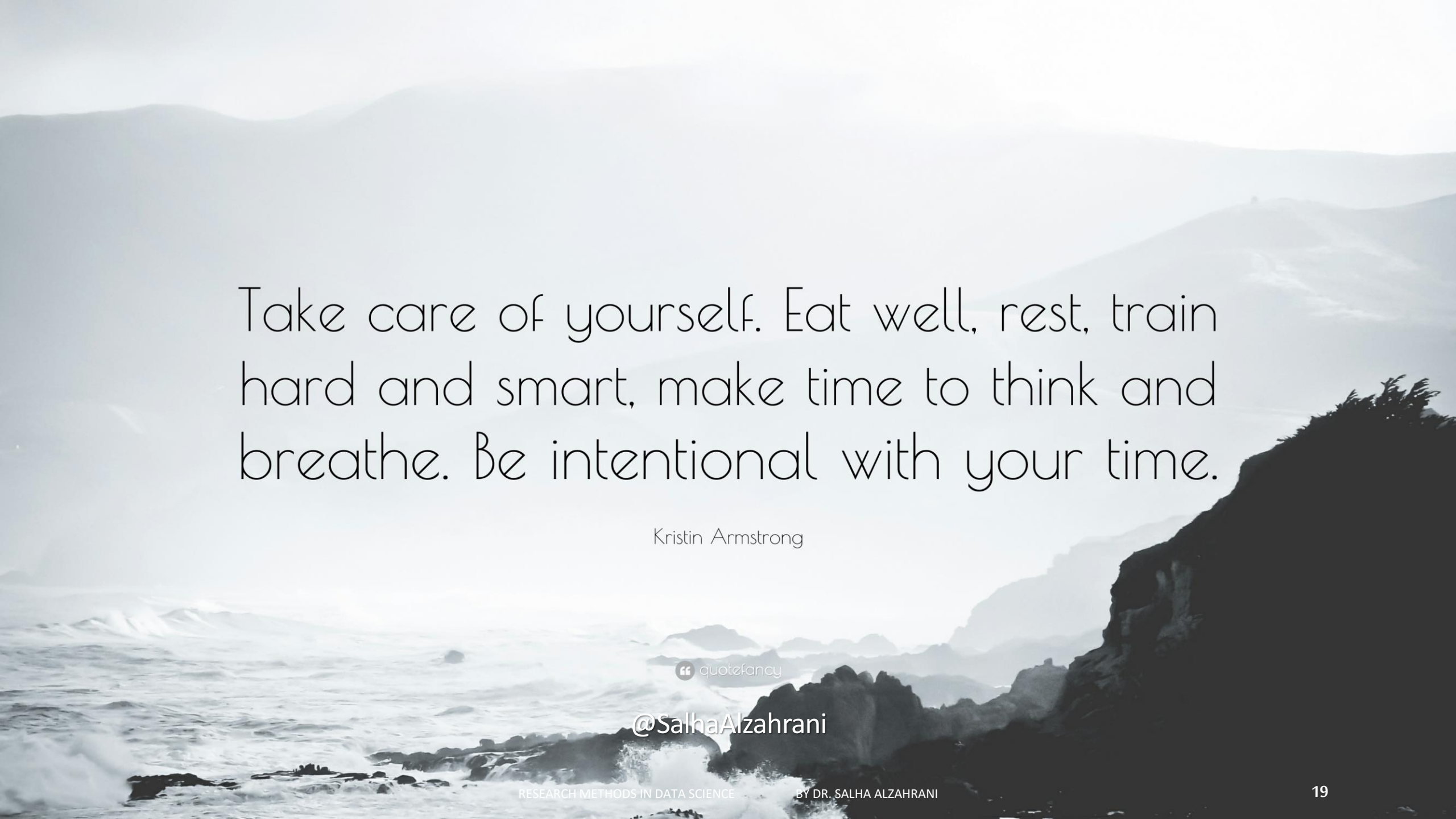
- the study design selected is faulty;
- the sampling procedure adopted is faulty;
- the method of data collection is inaccurate;
- the analysis is wrong;
- the statistical procedures applied are inappropriate; or
- the conclusions drawn are incorrect.

Hence, in drawing conclusions about a hypothesis, two types of error can occur:

- *Rejection* of a null hypothesis when it is true. This is known as a **Type I error.**
- *Acceptance* of a null hypothesis when it is false. This is known as a **Type II error.**

My Notes

Watch this video: Hypotheses vs. Research Questions


Take care of yourself. Eat well, rest, train hard and smart, make time to think and breathe. Be intentional with your time.
Kristin Armstrong
quotefancy
@SalhaAlzahrani