PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH
UNIVERSITY OF BISKRA
Faculty of Arts and Languages
Department of English Language and Literature

COURSE: MATHEMATICAL STATISTICS

Quantitative Data Analysis in Applied Linguistics

A course for second year master students specialized in Science of Languages

Prepared by: Moustafa Amrate

ACADEMIC YEAR: 2022/2023

Table of contents

Introduction

Quantitative data analysis has become an essential tool for researchers and scholars in many fields, including Applied Linguistics and Teaching English as a Foreign Language (TEFL). It is the process of using statistical methods and tools to analyze and interpret numerical data to uncover patterns, relationships, and trends (Romeijn, 2014). This coursebook details a statistics course designed for master two students of English who are specialized in Science of Languages. The goal of this course is to provide students with a deep understanding of the principles and methods of quantitative data analysis in the fields of Applied Linguistics and TEFL. This is to equip students with the essential skills and knowledge necessary to design, collect, manage, analyze, and interpret quantitative data manually, as well as using the different computer programs, such as Excel, SPSS and R studio. Upon completion of the course, students will have a solid foundation in quantitative data analysis, and they will be well-equipped to apply their knowledge in their graduation research projects.

The course will be delivered through a lecture format covering several key topics related to quantitative data analysis in Applied Linguistics. The first introductory lectures cover key concepts in statistics such as sampling approaches, data types, and hypothesis testing. This is followed by two lectures introducing descriptive statistics and its different measures, including measures of central tendency, variability, and distribution. Students are then introduced to inferential statistics with comprehensible definitions for advanced statistical testing, parametric assumptions, as well as some statistical guideline for hypothesis testing. More important, this lecture introduces students to the different advanced parametric and non-parametric statistical tests, along with the assumptions required to choose the appropriate test. Finally, the course will teach students how to use data visualization tools to represent quantitative data in a meaningful and compelling way.

To ensure that students have a solid grasp of the course content, both formative and summative assessments will be utilized. Formative assessments will be conducted through a variety of online and in-person activities, designed to enhance students' understanding of statistical methods and provide feedback to refine their skills. These assessments will allow students to practice and apply the concepts they have learned, with the aim of deepening their understanding and building confidence. The summative assessment, on the other hand, will be administered at the end of semester 3 of the master's degree program in Science of Language and will take the form of a 90-minute, face-to-face exam. The exam will assess students' ability to comprehend course concepts and to analyze and interpret real-world data using the methods and principles taught throughout the program.

Prerequisites

Before enrolling in this course, students should have:

A basic level in arithmetic.

A knowledge of basic concepts in Research Methods in Applied Linguistics.

Course objectives

By the end of this course, the student will be able to:

1. Choose the appropriate statistical test for the Master's research project.
2. Use the relevant statistical test for their research manually or by using Excel, SPSS, and R studio.
3. Interpret the results generated through the different statistical tests.

Course card

Module	Statistics
Level:	Master
Year of study:	2nd year
Term/ semester:	3rd semester
Unit:	Discovery
Credit:	1
Coefficient:	1
Format:	Lectures (a hybrid of online and face-to-face lectures)
Evaluation Format:	End of term exam (100%)
Lecturer:	Moustafa Amrate
Email:	*moustafa.amrate@univ-biskra.dz*

Reading list for the course

Dörnyei, Zoltán. (2007). Research methods in applied linguistics. Oxford university press.

Field, Andy. (2013). Discovering statistics using IBM SPSS statistics. sage.

Mackey, A., & Gass, S.M. (2015). Second Language Research: Methodology and Design (2nd ed.). Routledge. https://doi.org/10.4324/9781315750606

McKinley, J., & Rose, H. (Eds.). (2019). The Routledge handbook of research methods in applied linguistics. Routledge.

LECTURE 1: INTRODUCTION AND BASIC CONCEPTS

1.1. Introduction to statistics

Statistics is the discipline that concerns the collection, organization, analysis, interpretation, and presentation of data (Romeijn, 2014). In Statistics, the phenomena being studied should be quantified and therefore are reduced to numbers. Statistics is comprised of two main branches, a) descriptive statistics and b) inferential statistics.

a) “**Descriptive statistics** can help to provide a simple summary or overview of the data, thus allowing researchers to gain a better overall understanding of the data” (Mackey & Gass, 2015, pp. 250-251). This summary of the data is done through measures of frequency, measures of central tendency (i.e., the mode, median, and mean), and measures of spread (i.e., variance and standard deviation).

b) Inferential statistics refer to statistical tests that allow the researchers to infer the results obtained with a sample to the general population (hence the name inferential). Such tests are very important as they ensure the researcher that the difference found between tests or groups is not due to chance but rather to a real relationship (causal or correlational) between the variables being studied.

1.2. Important concepts in statistics

1.2.1. Data

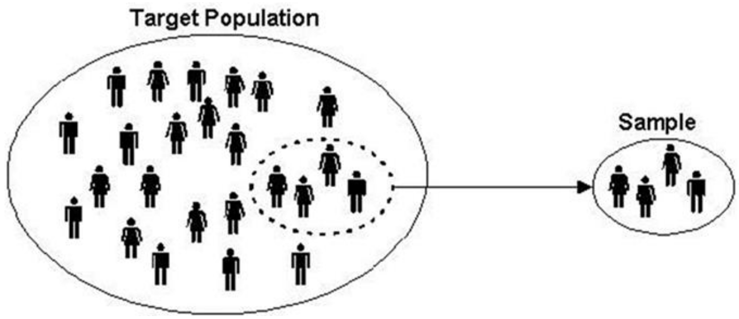
The term data (plural for datum) refers to facts, information, and observations about an object (or objects), a phenomenon (or phenomena), a person (or a group of people). Data can take either a qualitative or a quantitative form.

Qualitative data refers to the type of descriptive information that does not make use of statistical procedures (Mackey & Gass, 2015, pp. 250-251). Such data are often collected using observations, questionnaires, or interviews. Instead of quantification, the collection and analysis of qualitative data involves careful and detailed descriptions of objects, phenomena, or people’s narratives, perceptions, and beliefs about objects or phenomena.

Quantitative data consists of information that is, in some way or the other, quantifiable. In other words, we can put quantitative data into numbers, figures, and graphs, and process it using statistical procedures. When using quantitative analysis, we are usually interested in how much and how many there is/are of whatever we are interested in. The ultimate goal of quantitative data analysis is to generate facts that are generalizable.

1.2.2. A population vs. a sample

A **population** is the entire group that the researcher wants to draw conclusions about. A **sample** is the specific group that the researcher will collect data from. The size of the sample is always less than the total size of the population.



1.2.3. Variables

In research, variables refer to the population, objects, or phenomena that the researcher is trying to measure.

Dependent variable: The variable that is influenced by the independent variable.

Independent variable: The factor that causes the dependent variable to change.

Example: The influence of **English media** on ***EFL learners' listening comprehension.***

1.2.4. Hypotheses

Research hypotheses can be used to express what the researcher expects the results of the investigation to be. The hypotheses are based on observations or on what the literature suggests the answers might be (Mackey & Gass, 2005: 19).

a) Null hypothesis: (H_0) is a neutral statement used as a basis for testing. The null hypothesis states that there is no relationship between items under investigation. In advanced inferential statistical tests, the objective is often to reject the null hypothesis by providing a proof that there is a relationship between variables X and Y.

b) Non-directional hypothesis (two-tailed/ two-way hypothesis): In these types of hypotheses, a difference between variables under investigation is predicted, but the researcher is not interested in the direction of the difference/ relationship., but its direction is not specified.

Examples:

- Text-to-speech technologies will affect Algerian EFL learners' pronunciation awareness.

- There will be a relationship between family income and students' academic performance.

c) Directional hypothesis (one-tailed): In these types of hypotheses, the researcher indicates the predicted direction of the relationship between two or more variables.

Examples:

- Older learners will prefer speaking more than younger learners.
- Novice teachers will give less instructions in the target language than experienced teachers.

1.2.5. Causation vs. correlation

"Correlational research attempts to determine the relationship between or among variables; it does not determine causation." (Mackey & Gass, 2005, p. 284). On the other hand, causal relationships are unveiled by determining the causes and effects of a particular phenomenon. In research, there is a general consensus that correlation does not equal causation. In other words, just because there a relationship between two variables, it does not mean they cause each other. For example, high classroom attendance is often correlated with high academic achievement. However, it is not necessary that high classroom attendance is causing the high academic achievement (or the opposite). Perhaps, there is a third factor (e.g., willingness to study or proficiency) that is directly contributing to academic achievement.

1.2.6. Probability

In statistics, probability refers to the likelihood of an event taking place.[2] For example, the probability of getting heads or tails when tossing a coin is 50%. In statistical tests, the probability value (or p-value) measures the probability that an observed relationship/ difference could have occurred due to random chance. Low p-values ($p < 0.05$) indicate a great chance of real statistical significance (i.e., rejecting the null hypothesis); meanwhile, high p-values indicate a low chance of statistical significance (i.e., accepting the null hypothesis).

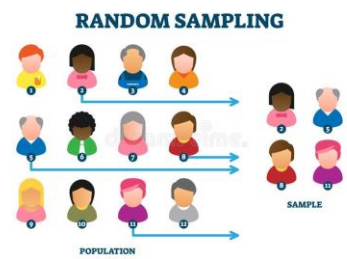
1.2.7. Bias

"Bias takes place when a statistical model/ test over or underestimates a relationship between two or more variables. There are two main types of bias, a) selection bias and b) confirmation bias. **Selection bias** happens when a researcher tries to generalize findings that were collected in a non-random way. **Confirmation bias**, on the other hand, takes place when the researcher starts the data collection/ analysis with a predetermined assumption" (Vickery, 2021).

LECTURE 2: SAMPLING IN RESEARCH

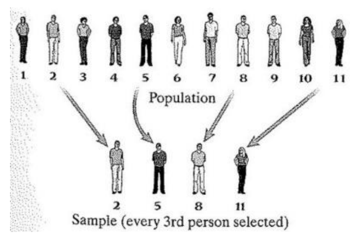
A ***population*** is an entire group that the researcher wants to draw conclusions about. A ***sample*** is a specific group that the researcher will collect data from. ***Sampling*** is the process of selecting individuals to form a sample that will estimate the characteristics of the whole population. The size of the sample is always less than the total size of the population.

2.1. PROBABILITY SAMPLING (REPRESENTATIVE): a type of sampling where all of the participants of the population have an equal chance (probability) of being selected in the sample. Probability sampling includes *random sampling*, *systematic sampling*, *stratified random sampling*, and *cluster sampling*.

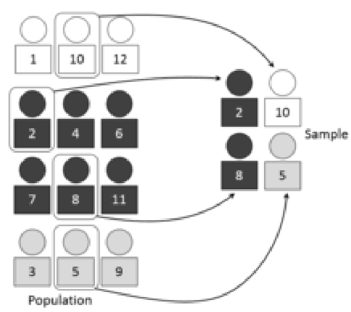


1. **Random sampling:** the key to random sampling is that each unit in the population has an equal probability of being selected in the sample. Using random sampling protects against bias being introduced in the sampling process, and hence, it helps in obtaining a representative sample.

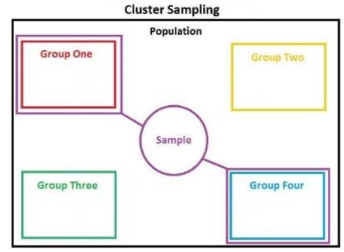
Example**:** Researcher goes to an English school and randomly selects 60 participants to take part in a study.



2. **Systematic sampling** is similar to random sampling, but it is usually slightly easier to conduct. Every member of the population is listed with a number, but instead of randomly selecting participants, individuals are chosen according to some systematic rule – e.g., every fourth unit/student, etc.
Example: Researcher goes to a classroom at an English school and selects every 3rd student.



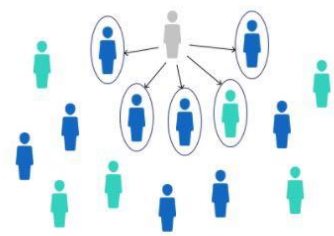
3. **Stratified random sampling:** used when the population has different groups (strata) and the researcher needs to ensure that those groups are fairly represented in the sample. In stratified random sampling, independent samples are randomly drawn from each group.
Example: Researcher goes to an international English school with students from Arabic, French, Italian, and Chinese L1 backgrounds. In stratified sampling, the researcher has to randomly select participants from each L1 subgroup in the English school.



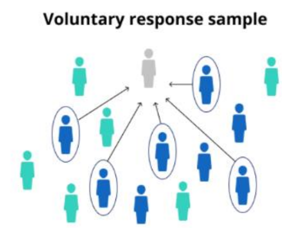
4. **Cluster sampling** is a probability sampling method in which you divide a population into clusters, such as districts, schools or classrooms, and then randomly select some of these clusters (groups) as your sample.
Example: Researcher goes to an English school with eight classrooms. In cluster sampling, the research can randomly select two classrooms to do the study.

2.2. NON-PROBABILITY SAMPLING (NON- REPRESENTATIVE): a type of sampling where the participants do not have an equal chance (probability) of being selected in the sample. Non-probability sampling includes *convenience sampling*, *voluntary sampling*, *snowball sampling*, *quota sampling*, and *purposive sampling*.

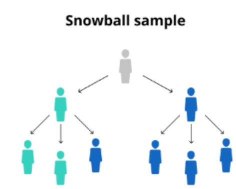
1. Convenience sampling (also known as availability sampling) relies on data collection from population members who are conveniently available to participate in study. In other words, this sampling method involves getting participants who are *not very far* from the researcher wherever *convenient*.
Example: An MA student investigating students' opinions on university support services. After each class, the student asks one of their friends to complete a survey. This is a convenient way to gather data, but as the student only surveyed their friends, the sample is not representative of all university students.



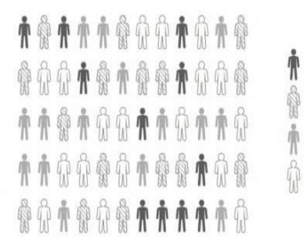
2. Voluntary response sampling similar to a convenience sample, a voluntary response sample is mainly based on ease of access. However, instead of the researcher choosing participants and directly contacting them, people volunteer themselves. Voluntary response samples are always at least somewhat biased, as some people will inherently be more likely to volunteer than others.
***Example*:** A researcher posts an advertisement for an experiment at an English school. Only motivated students join the study.



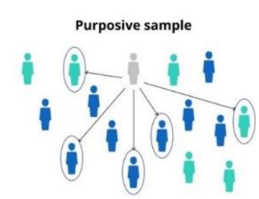
3. Snowball sampling is usually done when there is a very small population size. In this type of sampling, the researcher asks the initial participant to identify another potential participant who also meets the criteria of the research.
Example: A researcher doing a study on disabled EFL teachers in Algerian high schools. After interviewing one participant, and due to the limited number of potential participants, the researcher asks the first participant to identify other disabled EFL teachers who work in high schools.



4. Quota sampling is defined as a non-probability sampling method in which researchers create a sample involving individuals that represent a population with its various subgroups (e.g. age, gender, education, race, or religion). However, unlike *Stratified Random Sampling,* participants are chosen through a non-random sample selection (i.e. only available/ volunteering participants).
Example: In an English school where students are 20% male and 80% female, the researcher chooses to interview a quota of 20 male students and 80 female students to get a proportional balance. However, unlike Stratified Random Sampling, the researcher interviews only the participants who are *conveniently* available.



5. Purposive sampling (judgement sampling): subjects are chosen to be part of the sample with specific purpose in mind. The researcher believes that some subjects are more fit for the research compared to other individuals.
Example: A researcher at an English school chooses only students with fluent English to do interviews.



LECTURE 3: LEVELS OF MEASUREMENT

Quantitative data can be captured/ collected through two main levels of measurement: 1) Categorical data and 2) Numerical data. Categorical data are divided into two distinct subcategories: A) Nominal data and B) Ordinal data. Similarly, numerical data are also divided into two subcategories: A) Interval data and B) Ratio data. It is necessary to understand these different types of measurement as they are necessary for choosing the appropriate statistical analyses. The following table summarizes the different types of data and their qualities:

	CATEGORICAL DATA		*NUMERICAL DATA*	
	Nominal data	**Ordinal data**	**Interval data**	**Ratio data**
Can be named	✓	✓	✓	✓
Can be ordered		✓	✓	✓
Equal & meaningful distance between values			✓	✓
Has an absolute zero value				✓

3.1. Categorical data

Categorical data (also known as qualitative variables) describe qualitative attributes.

A. Nominal data

Nominal data are a type of categorical data that are “labeled” or “named” data which can be divided into various groups that do not overlap. Data are not measured or evaluated in this case; they are just assigned to multiple unique groups. “We may, for example, ask participants to report whether they are native or non-native speakers of English. Here, for statistical analysis (e.g., frequency counts or percentages), we can code this data by assigning 1 to native speakers of English and 2 to non-native speakers of English.” (Phakiti, 2010, p. 40). Other examples of nominal data include gender, country of birth, nationality, eye color …etc.

As demonstrated in the examples, nominal data are alphabetical with qualitative property/ nature. For example, while we can calculate the number of male/female students in a particular context, gender in itself is not something we can measure. Because of this qualitative nature, it is impossible to calculate the mean for nominal data. Furthermore, because nominal data do not overlap, they do not have an intrinsic order. For example, there is no natural order for gender or being a native or a non-native speaker of English.

B. Ordinal data

Ordinal data is a type of categorical data where the variables have a natural order. The following table provides some examples of ordinal data:

Question/ item	Options				
Likert scales: How satisfied are you with the course?	1. Very satisfied ☐	2. Satisfied ☐	3. Indifferent ☐	4. Dissatisfied ☐	5. Very dissatisfied ☐
Educational level:	Kindergarten ☐	Primary school ☐	Middle school ☐	High school ☐	University ☐
Age group:	0 – 18 ☐	19 – 34 ☐	35 – 49 ☐	50 – 64 ☐	65 + ☐

"Ordinal data can tell us that an individual is greater or less than others in a characteristic or aspect being measured, but they do not tell us how much greater or lesser in terms of equal intervals (i.e., the distance between categories is not precise). For example, in an agreement scale ranging from 1 (strongly disagree) to 5 (strongly agree), we can see that one is greater or less than another in terms of levels of agreement, but we cannot say that the distance between the scores of 4 (agree) and 5 (strongly agree) is the same as the distance between the scores of 1 (strongly disagree) and 2 (disagree)" (Phakiti, 2010, p. 40).

3.2. Numerical data

Numerical data (also referred to as continuous variables) describe quantitative attributes.

A. Interval data

Interval data are a type of continuous numerical data which are measured along a scale, in which each point is placed at an equal distance from the other (i.e., equidistant). For example, the distance between 10° and 11° is the same as the distance between 20° and 21° (i.e., 1 degree Celsius). Examples of interval data include time, date, temperature, IQ (Intelligence Quotient) test … etc.

A distinctive feature of interval data is that they often do not have an absolute zero (unlike ratios). In other words, the value 0 in interval data, does not imply the absence of the attribute. For example, the value 0 degree in Celsius refers to the freezing temperature of water. In fact, interval data can go below zero (e.g., -10 degrees Celsius).

B. Ratio data

Ratio data is defined as a type of numerical data that can be categorized, ranked, with an equal and definitive distance between values and an absolute "zero" being treated as a point of origin. Examples of ratio data include age, height, weight, speed, monthly income … etc. A true zero means there is an absence of the variable of interest. In other words, there can be no negative numerical value in ratio data (the value zero mean the lack of something).

ACTIVITY: Indicate the type of data in each of these examples:

Duration of completing a reading task	Nominal	☐	Number of students in a classroom	Nominal	☐	Language proficiency level	Nominal	☐	Number of grammatical errors in an essay	Nominal	☐
	Ordinal	☐		Ordinal	☐		Ordinal	☐		Ordinal	☐
	Interval	☐		Interval	☐		Interval	☐		Interval	☐
	Ratio	☐		Ratio	☐		Ratio	☐		Ratio	☐
Country of residence	Nominal	☐	Self-reported frequency of taking in class notes	Nominal	☐	Language teaching approach	Nominal	☐	Percentage of lectures attended	Nominal	☐
	Ordinal	☐		Ordinal	☐		Ordinal	☐		Ordinal	☐
	Interval	☐		Interval	☐		Interval	☐		Interval	☐
	Ratio	☐		Ratio	☐		Ratio	☐		Ratio	☐
Reading strategy	Nominal	☐	IELTS/ TOEFL score	Nominal	☐	Frequency of speech intonation/ pitch	Nominal	☐	Students' Excel knowledge (beginner, intermediate, advanced)	Nominal	☐
	Ordinal	☐		Ordinal	☐		Ordinal	☐		Ordinal	☐
	Interval	☐		Interval	☐		Interval	☐		Interval	☐
	Ratio	☐		Ratio	☐		Ratio	☐		Ratio	☐

LECTURE 4: INTRODUCTION TO DESCRIPTIVE STATISTICS

"***Descriptive statistics*** can help to provide a simple summary or overview of the data, thus allowing researchers to gain a better overall understanding of the data set" (Mackey & Gass, 2015, p. 292). Because raw data are not in and of themselves revealing, they must be organized and described in order to be informative. This lecture will present an overview of three different types of descriptive statistics: (1) measures of frequency; (2) measures of central tendency; and (3) measures of variability or dispersion.

4.1. Measures of frequency

Measures of frequency are used to indicate how often a particular behavior or phenomenon occurs. For example, in second language studies, researchers might be interested in tallying how often learners make errors in forming the past tense, or how often they engage in a particular classroom behavior. This is usually done through a table format where the researcher presents frequencies and percentages. The following table, for example, presents the frequency of phonemic pronunciation errors made by beginner EFL students in a spontaneous speech task.

Participants	**Gender**	**Vowel errors**	**Consonant errors**	**Totals**
Student 1	Male	17	3	20
Student 2	Female	15	8	23
Student 3	Female	25	9	34
Student 4	Male	18	6	24
Student 5	Female	14	2	16
Totals *(%)*	*n = 5 (2 male, 3 females)*	**89** *(76.07%)*	**28** *(23.93%)*	**117** *(100%)*

Measures of frequency can also be visualized through pie charts, bar graphs, and line charts.

4.2. Measures of central tendency

Second language researchers often use one or more measures of central tendency to provide precise quantitative information about the typical behavior of learners with respect to a particular phenomenon. There are three commonly used measures of central tendency, namely: 1) the *mode*, 2) the *median*, and 3) the *mean*.

A) The mode: is the most frequent score obtained by a particular group of learners. For example, if the ESL proficiency test scores recorded for a group of students were 92, 78, 92, 74, 89, and 80, the mode would be 92 because two students in this sample obtained that score.

B) The median: is the score at the center of the distribution—that is, the score that splits the group in half. For example, in our series of ESL proficiency test scores (92, 78, 92, 74, 89, and 80), we would find the median by first ordering the scores (74, 78, 80, 89, 92, 92) and then finding the score at the center. Since we have an even number of scores in this case (i.e., six), we would take the midpoint between the two middle scores (80 and 89), or 84.5.

C) The mean (average): is derived from adding up all the numbers (*Sum*) and dividing by the total number of observations (e.g., participants). The following table presents the formulas needed for calculating the mean.

	FORMULA	EXPLANATION
Population mean (μ)	$\mu = \frac{(\Sigma\ xi)}{N}$	$Population\ Mean\ = \frac{Sum}{Population\ size}$
Sample mean (x̄)	$\bar{x} = \frac{(\Sigma\ xi)}{n}$	$Sample\ Mean\ = \frac{Sum}{Sample\ size}$

For our scores (92, 78, 92, 74, 89, and 80), the mean would be the sum of all scores divided by the number of observations, or (505 /6 =) 84.17. It should be kept in mind that even though the mean is commonly used, it is sensitive to extreme scores especially if the number of participants is small.

4.3. Measures of spread (dispersion):

"Measures of dispersion describe variability of the numeral data away from the central tendency" (Phakiti, 2010, p. 44)[1]. "Measures of dispersion [particularly standard deviation] can serve as a quality control for measures of central tendency; the smaller the standard deviation, the better the mean captures the behavior of the sample." (Mackey & Gass, 2015, p. 303). In statistics, there are two main measures of dispersion, the *variance* and *standard deviation*.

A) The variance: is the expectation of the squared deviation of a random variable from its population mean or sample mean. The variance is a measure of how far a set of numbers is spread out from their average value. The more spread the data, the larger the variance is in relation to the mean. It is calculated by dividing the sum of squared deviations from the mean by the population or sample size.

	FORMULA	EXPLANATION
Population variance (σ2)	$\sigma^2 = \frac{\Sigma\ (xi - \ \mu)^2}{N}$	$Population\ Vriance\ = \frac{Sum\ of\ (observation\ value - population\ mean)^{Squared}}{Population\ size}$
Sample variance (S^2)	$S^2 = \frac{\Sigma\ (xi - \ \bar{x})^2}{n-1}$	$Sample\ Vriance\ = \frac{Sum\ of\ (observation\ value - sample\ mean)^{Squared}}{Sample\ size - 1}$

[1] Phakiti, A. (2010). Analyzing quantitative data. In B. Paltridge & A. Phakiti (Eds.), *Continuum Companion to Research Methods in Applied Linguistics* (pp. 39-49). London: Continuum.

B) The Standard deviation (*SD*): "is the average point from the mean which indicates on average how much the individual scores spread around the mean." (Phakiti, 2010, p. 44). The standard deviation is "high if the sample is heterogeneous and contains extreme scores, whereas [it is] low in a homogeneous sample with all the scores clustered around the mean." (Dörnyei, 2007, p. 214). The standard deviation is obtained by calculating the square root ($\sqrt{\ }$) of the variance.

	FORMULA	EXPLANATION
Population standard deviation (σ)	$\sigma = \sqrt{\frac{\Sigma\,(xi - \mu)^2}{N}}$	$Population\ SD = \sqrt{\frac{Sum\ of\ (observation\ value - population\ mean)^{Squared}}{Population\ size}}$
Sample standard deviation (S)	$S = \sqrt{\frac{\Sigma\,(xi - \bar{x})^2}{n-1}}$	$Sample\ SD = \sqrt{\frac{Sum\ of\ (observation\ value - sample\ mean)^{Squared}}{Sample\ size - 1}}$

LET'S PRACTICE

Calculate the sample variance and standard deviation for the ESL proficiency scores (i.e., 92, 78, 92, 74, 89, 80).

Sample variance

$$S^2 = \frac{\Sigma\,(xi - \bar{x})^2}{n-1}$$

$$S^2 = \frac{((92-84.17)^2 + (78-84.17)^2 + (92-84.17)^2 + (74-84.17)^2 + (89-84.17)^2 + (80-84.17)^2)}{6-1}$$

$$S^2 = \frac{((7.83)^2 + (-6.17)^2 + (7.83)^2 + (-10.17)^2 + (4.83)^2 + (-4.17)^2)}{5}$$

$$S^2 = \frac{(61.31 + 38.07 + 61.31 + 103.43 + 23.33 + 17.39)}{5}$$

$$S^2 = \frac{304.84}{5}$$

$$S^2 = 60.97$$

Sample standard deviation

$$S = \sqrt{S^2}$$

$$S = \sqrt{60.97}$$

$$S = 7.81$$

LECTURE 5: APPLICATIONS OF DESCRIPTIVE STATISTICS

5.1. Practice

The following dataset represent the scores achieved by university Algerian EFL students in the International English Language Testing System (i.e., IELTS): 6.00, 5.50, 6.00, 7.00, 6.00, 6.50, 5.50, 6.50, 5.00, 6.00.

1. Using these results, calculate the sum, mode, median, sample mean, sample variance, and standard deviation.

Students	**IELTS Scores**	**Differences ($x_i - \bar{x}$)**	**Differences squared $(x_i - \bar{x})^2$**
Student 1	6.00	0.00	0.00
Student 2	5.50	-0.50	0.25
Student 3	6.00	0.00	0.00
Student 4	7.00	1.00	1.00
Student 5	6.00	0.00	0.00
Student 6	6.50	0.50	0.25
Student 7	5.50	-0.50	0.25
Student 8	6.50	0.50	0.25
Student 9	5.00	-1.00	1.00
Student 10	6.00	0.00	0.00
Sum (Σ):	60.00		**$\Sigma(x_i - \bar{x})^2 = 3.00$**
Mode:	6.00		
Median:	6.00		
Sample mean ($\bar{x}$):	6.00		
Sample variance (S^2):	0.33		
Sample standard deviation (S):	0.58		

2. Report the results in the table below.

Table 1. Algerian university EFL students' performance in the IELTS exam

	n	***Min***	***Max***	***$\bar{x}$***	***S^2***	***S***
IELTS scores	10	5	7	6.00	0.33	0.58

Notes. n = sample size, $\bar{x}$ = sample mean, S^2 = sample variance, S = sample standard deviation.

5.2. Generating descriptive statistics in Excel

You can use Excel to generate to generate descriptive statistics.

Statistics	**Steps**			
Sum (Σ):		**IELTS Scores**		**IELTS Scores**
	Student 1	5.00	*Student 1*	5.00
	Student 2	4.50	*Student 2*	4.50
	Student 3	5.00	*Student 3*	5.00
	Student 4	6.00	*Student 4*	6.00
	Student 5	5.00	*Student 5*	5.00
	Student 6	5.50	*Student 6*	5.50
	Student 7	4.50	*Student 7*	4.50
	Student 8	5.50	*Student 8*	5.50
	Student 9	4.00	*Student 9*	4.00
	Student 10	5.00	*Student 10*	5.00
	Sum (Σ):	=sum(B2:B11)	**Sum (Σ):**	50.00

The mean (x̄):

	IELTS Scores
Student 1	5.00
Student 2	4.50
Student 3	5.00
Student 4	6.00
Student 5	5.00
Student 6	5.50
Student 7	4.50
Student 8	5.50
Student 9	4.00
Student 10	5.00
Sum (Σ):	50.00
Sample mean (x̄):	=AVERAGE(B2:B11)

	IELTS Scores
Student 1	5.00
Student 2	4.50
Student 3	5.00
Student 4	6.00
Student 5	5.00
Student 6	5.50
Student 7	4.50
Student 8	5.50
Student 9	4.00
Student 10	5.00
Sum (Σ):	50.00
Sample mean (x̄):	5.00

The variance:

	IELTS Scores	Differences (xi- x̄)	Differences squared (xi- x̄)²
Student 1	5.00	0.00	0.00
Student 2	4.50	-0.50	0.25
Student 3	5.00	0.00	0.00
Student 4	6.00	1.00	1.00
Student 5	5.00	0.00	0.00
Student 6	5.50	0.50	0.25
Student 7	4.50	-0.50	0.25
Student 8	5.50	0.50	0.25
Student 9	4.00	-1.00	1.00
Student 10	5.00	0.00	0.00
Sum (Σ):	50.00	**Σ(xi- x̄)² =**	3.00
Sample mean (x̄):	5.00		
Sample variance (S²):	=D13/9		
Sample standard deviation (S):			

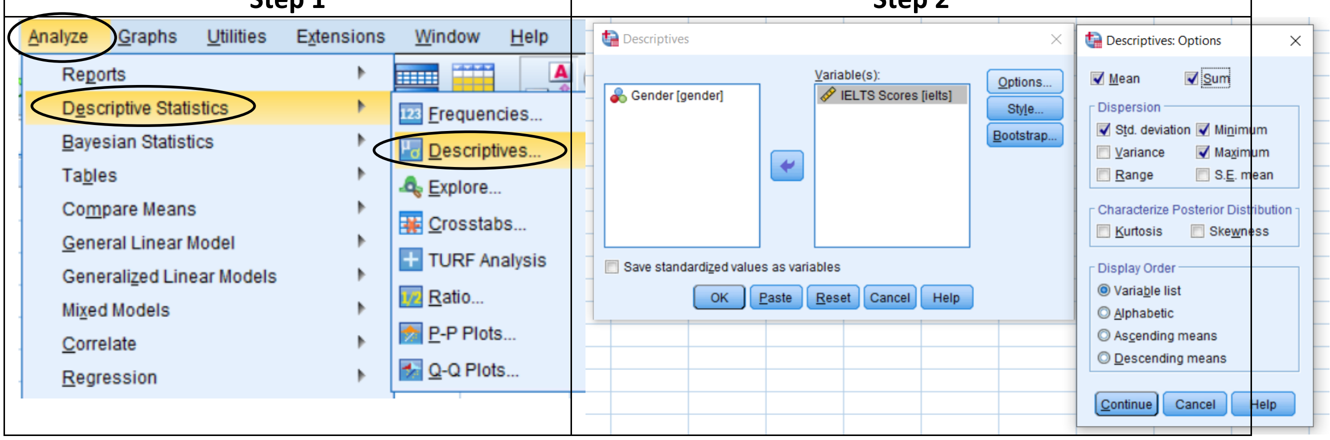
Standard deviation:

	IELTS Scores	Differences (xi- x̄)	Differences squared (xi- x̄)²
Student 1	5.00	0.00	0.00
Student 2	4.50	-0.50	0.25
Student 3	5.00	0.00	0.00
Student 4	6.00	1.00	1.00
Student 5	5.00	0.00	0.00
Student 6	5.50	0.50	0.25
Student 7	4.50	-0.50	0.25
Student 8	5.50	0.50	0.25
Student 9	4.00	-1.00	1.00
Student 10	5.00	0.00	0.00
Sum (Σ):	50.00	**Σ(xi- x̄)² =**	3.00
Sample mean (x̄):	5.00		
Sample variance (S²):	0.33		
Sample standard deviation (S):	=SQRT(B15)		

5.3. Generating Descriptive Statistics in SPSS

Descriptive statistics can also be easily generated using SPSS.

Step 1	Step 2



LECTURE 6: INTRODUCTION TO INFERENTIAL STATISTICS AND HYPOTHESIS TESTING

"Inferential statistics are used to help us look beyond raw data and descriptive statistics. They help us make inferences about population parameters" (Phakiti, 2010, p. 44). "Given that it is impossible to gather data from all members of the population, inferential statistics can allow researchers to generalize findings to other, similar language learners; that is, to make inferences." (Mackey & Gass, 2005, p. 269).

6.1. Assumptions for inferential statistical tests

There are two main categories of inferential statistical tests, namely: **1) parametric tests**, such as the t-test, One-Way ANOVA, and Pearson correlation, and **2) non-parametric tests**, such as the Mann-Whitney test, Kruskal walis test, Spearman correlation. "Although parametric tests are more preferable in quantitative research, non-parametric tests are [also] important for applied linguistics research because some data are not always strongly interval or continuous." (Phakiti, 2010, pp. 45-46).

A number of assumptions (conditions) should be met before the researcher can decide on using the appropriate inferential testing category with their data. "These assumptions (conditions) are not optional and if they are not met, there is a heightened risk of making a false inference." (Phakiti, 2010, p. 45). The conditions of each testing category are summarized in the table below:

ASSUMPTIONS	PARAMETRIC TESTS	NON-PARAMETRIC TESTS
Type of data	Numerical data	Categorical data
Sampling approach	Random sampling	Non-random sampling
Distribution of results	Normal distribution of data	Non-normal distribution of data
Homogeneity of variances	Homogeneous variances	Heterogeneous variances

<u>Assumption 1</u>: Type of data

The type of data is essential in determining the appropriate inferential statistical tests. Usually, numerical data are used in parametric tests; while categorical data are used in non-parametric tests.

<u>Assumption 2</u>: Sampling approach

Because the main aim of inferential statistics is to generalize the results from the sample to the population, it is preferable to have a sample that is randomly recruited from the population. If the sample is randomly recruited, parametric tests can be used; otherwise, non-parametric tests can serve as an alternative.

<u>Assumption 3</u>: Distribution of results

A distribution describes the clustering of scores in a dataset. In a normal distribution, the numbers (e.g., scores on a particular test) cluster evenly around the mean. ***Figure 10.7*** shows that the measures of central tendency coincide at the midpoint. Parametric tests require the data to be normally distributed. However, if the data are not normally distributed, non-parametric tests should be used. The two well-known tests of normality, namely, the Kolmogorov–Smirnov test and the Shapiro–Wilk test are the most widely used methods to test the normality of data.

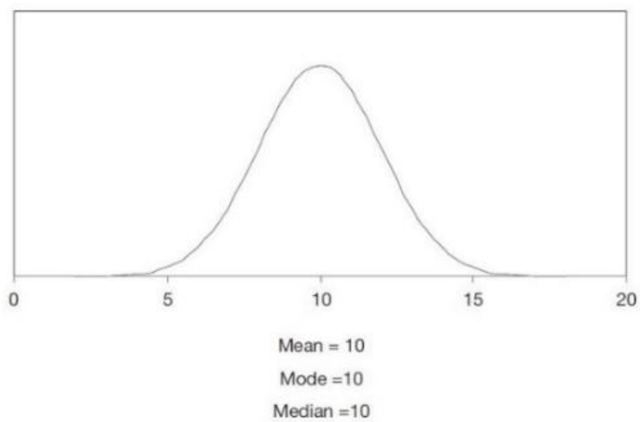


FIGURE 10.7 Normal distribution

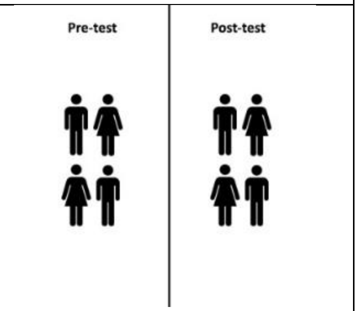
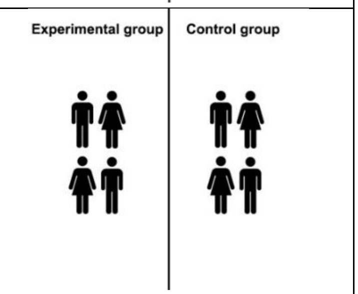
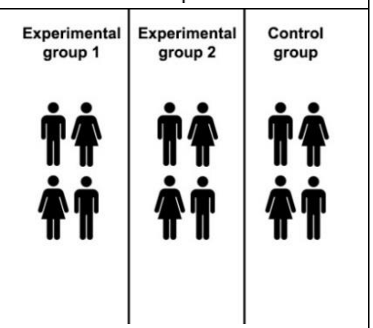
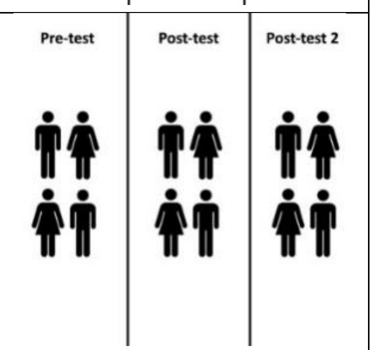
<u>Assumption 4</u>: Homogeneity of variance

Parametric tests require the groups being tested to have similar variances (i.e., homogenous variances). There are a number of tests that examine homogeneity of variances, but the most robust test is the Levene's test.

IMPORTANT NOTE. *Non-parametric tests should be used if any of the above assumptions are not met.*

6.2. Classification of inferential statistical tests

The following table classifies the most widely used inferential statistical tests and summarizes their purposes.

PARAMETRIC TESTS	EQUIVALENT NON-PARAMETRIC TESTS	PURPOSE
Paired t-test	***Wilcoxon Rank sum test***	To test difference between the means of two related samples (e.g., testing students vocabulary scores before and after a treatment [pre-test and post-test designs]) 
Independent t-test	***Mann-Whitney U test***	To test difference between the means of two independent samples (e.g., experimental group vs. control group) 
ANOVA (One way analysis of variance)	***Kruskal Wallis Test***	To test the difference between the means of more than two independent samples (e.g., control group vs. experimental group1 vs. experimental group 2) 
Repeated measures ANOVA	***Friedman test***	To test the difference between the means of more than two related samples (e.g., testing the pronunciation of EFL students in the first, second, and third semesters) 
Pearson correlation r	***Spearman correlation***	To measure correlation between two sets of data (e.g., testing the relationship between classroom attendance and Oral Expression exam results)

N.B. *On the electronic version of this file, the test names above contain links that will direct you to YouTube tutorials demonstrating how you can conduct each test on the software package for statistics SPSS.*

LECTURE 7: IMPORTANT CONCEPTS IN HYPOTHESIS TESTING

While it is possible to conduct the different inferential statistical tests manually, it is easier and more efficient to use a software package for statistical analysis. Currently, IBM SPSS is among the most widely used software packages for statistical analysis. **N.B.** The electronic copy of **Lecture 6** contains direct links to YouTube tutorials demonstrating the necessary steps for conducting the different inferential statistical tests on SPSS.

7.1. Null and alternative hypotheses

Before conducting inferential statistical tests on SPSS, the researcher should consider two hypotheses. They are called the null hypothesis (H_0) and the alternative hypothesis (H_a). These hypotheses contain opposing viewpoints. ***The null hypothesis (H_0)*** is a hypothesis that says there is no statistical significance between the two variables. ***The alternative hypothesis (H_α)***, on the other hand, indicates that there is a statistical significance between the two variables being measured.

Example

In a study investigating the effects of English audiobooks on EFL learners' listening comprehension, the null hypothesis would be: $\mathbf{H_0}$ = English audiobooks have no significant effect on EFL learners' listening comprehension. Meanwhile, the alternative hypothesis would be: $\mathbf{H_a}$ = English audiobooks have a significant effect on EFL learners' listening comprehension.

7.2. Interpreting the significance level (alpha value)

After conducting the inferential test in SPSS, the software generates the significance level. This is also known as the *Alpha level (α)*. The p-value is the probability that the null hypothesis is true. In other words, the p-value is the probability of making the wrong decision or that the results are due to random chance.

The p-value generated by the statistical tests in SPSS (i.e., $p< 0.05$) is a fractional number. To better understand this number, we should multiply it by 100 to get 5%. This number means that there is a 5% (or lower) that the null hypothesis is false. Usually:

- If $P \leq 0.05$, the results are statistically significant. This means that there is a 5% chance or less that the null hypothesis is correct.

- If $P > 0.05$, the results are not statistically significant. This means that there is more than a 5% chance that the null hypothesis is correct. Therefore, it is more likely that the results are due to chance.

The table below demonstrates the necessary decisions in each scenario.

Result	**Decision**
$P \leq 0.05$	Reject the null hypothesis (i.e., the results are statistically significant)
$P > 0.05$	Accept the null hypothesis (i.e., the results are not statistically significant)

7.3. Bad practices in advanced statistical testing

There are number of bad practices that the researchers should avoid when conducting inferential statistical tests. In this lecture, we focus on type 1 and type 2 errors, p-hacking, and publication bias.

I. Type I and Type II errors

"In statistical hypothesis testing, a type I error is the mistaken rejection of an actually true null hypothesis (also known as a "false positive" finding or conclusion; example: "*an innocent person is convicted*"), while a type II error is the mistaken acceptance of an actually false null hypothesis (also known as a "false negative" finding or conclusion; example: "*a guilty person is not convicted*")" (Shuttleworth & Wilson, n.d.).

	True Null Hypothesis	**False Null Hypothesis**
Reject H_0	**Type 1 error (False positive)** [*finding significant results when there are no significant results*]	Correct decision
Accept H_a	Correct decision	**Type 2 error (False negative)** [*finding no significant results when there are significant results*]

To avoid these errors, the researcher should choose the appropriate inferential tests (i.e., parametric or non-parametric test) based on the type of data, sample size, and distribution of results.

II. P-hacking

"P-hacking (also known as *significance chasing* or *selective inferencing*) is the misuse of data analysis to find patterns in data that can be presented as statistically significant" (Davey Smith & Ebrahim, 2002). This can also be done by only presenting the part of data that can generate significant results. In addition to being an unethical practice, p-hacking increases the chance of committing a type 1 (i.e., finding significant results when there are no significant results).

III. Publication bias

Publication bias is a type of bias that occurs when the outcome of a study influences the researcher's decision on publishing or withholding the findings.

PRACTICE

Indicate which advanced statistical test should be used in the following scenarios.

SCENARIO	TEST
1. To investigate the effectiveness of teaching English vocabulary with songs, you compared between an experimental group listening to songs with subtitles, a group listening to songs without subtitles, and group receiving traditional teaching. The study was conducted with a large sample and the data were normally distributed.	
2. You want to compare the grammatical accuracy of 60 adult Algerian EFL learners and 60 adult Tunisian EFL learners. The data is not normally distributed.	
3. You want to understand the relationship between the number of years living abroad and the pronunciation intelligibility of 8 ESL learners. The data are normally distributed.	
4. You want to compare the pronunciation accuracy of an Experimental group involving 48 ESL learners who have been practicing with computer-assisted language learning technologies and that of a control group involving 40 EFL learners who have been practicing in a traditional classroom setting. The data are normally distributed.	
5. To investigate the effects of English podcasts on the listening comprehension of 7 EFL learners, you conduct a pre-test that is followed by 6 weeks of Podcasts listening sessions and end the training with a post-test.	
6. You tested the lexical complexity in the speech of 12 EFL learners during 4 occasions. The data was normally distributed.	

LECTURE 8: CORRELATION

So far, we have looked at the statistical tests examining causal relationships and differences between variables. In this lecture, we cover statistical tests examining the relationship between variables. In statistics, correlation refers to statistical tests that investigate the degree to which two or more variables are linked to each other. For example, a correlational study in Applied Linguistics would look at the relationship between the number of years living abroad and ESL/ EFL learners' pronunciation accuracy. "The statistical procedure to achieve [this purpose] is called '*correlation analysis*' and it allows us to look at two variables and evaluate the strength and direction of their relationship." (Dörnyei, 2007, p. 223). "Correlation research attempts to determine the relationship between or among variables; it does not determine causation" (Mackey & Gass, 2015, p. 284). In other words, the fact that two variables are linearly linked to each other does not mean that one is causing the other. For example, just because British people spend more money in shops during the winter, it does not mean that the cold temperature is causing people in the UK to spend more money. It just happened that the winter season coincides with the British holiday season, a period that is known for increased spending.

8.1. Detecting a correlation

Generally, correlations can be detected or noticed when summarizing and describing data. For example, the table below reports the percentage of public library members in relation to their social class.

Percentage of public library members by their social class origin

PUBLIC LIBRARY MEMBERSHIP	**SOCIAL CLASS STATUS**	
	Middle class	***Working class***
Member	86%	37%
Non-member	14%	63%
Total	100%	100%

Adapted from (Cohen et al., 2002, p. 529)

However, for a more accurate detection and measurement of correlation, two main statistical tests are used, namely: the 1) **Pearson product-moment correlation** (or the Pearson correlation) for interval and ratio data and 2) **Spearman rank order correlation** (or the Spearman's correlation) for ordinal data. Correlational tests are designed to answer the following questions:

1) *Is there a relationship between two (or more) variables?*

2) *What is the direction of the relationship?*
3) *What is the strength of the relationship?*

1. Pearson correlation

"Pearson product-moment correlation is the standard type, computed between two continuous variables. When we talk about 'correlation' in general, this is what we usually mean. The Pearson product-moment correlation coefficient is symbolized by the lower-case letter '**r**'". (Dornyei, 2007, p.224). This test is used when the data of the variables being correlated meet the conditions for parametric tests (i.e., 1) **numerical data**, 2) **large sample size**, and 3) **normal distribution of data**).

2. The Spearman rank order correlation

The Spearman correlation is the non-parametric equivalent for the Pearson correlation. "As the name suggests, the correlation is based on the ranks of the data (in a rank order) rather than the actual [numerical] values" (Dornyei, 2007, p. 230). In other words, this test is used when the data of the variables being correlated do not meet the conditions for parametric tests. The Spearman correlation test is appropriate when the data are ordinal (rather than numerical), the sample size is very small, and the data are not normally distributed. The Spearman rank order correlation is represented with the Greek letter '**ρ**'.

8.2. Interpreting correlation test results

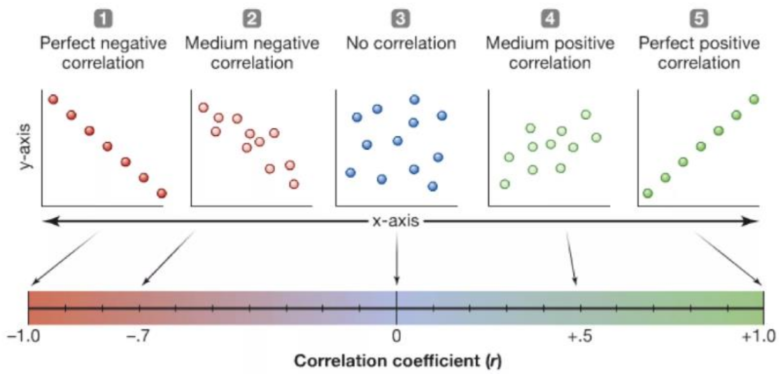
When investigating correlation using the Pearson or the Spearman correlation tests, a coefficient is generated. This coefficient indicates the direction as well as the strength of a relationship between two variables.

1. The direction of a correlation

A correlation can either be positive or negative. Where the two variables (or sets of data) fluctuate in the same direction, i.e., as one increases so does the other, or as one decreases so does the other, a **positive correlation** (i.e., a positive relationship) is said to exist. Correlations reflecting this pattern are prefaced with a plus sign (**+**) to indicate the positive nature of the relationship. Thus, **+1.0** would indicate a perfect positive correlation.

On the other hand, a **negative correlation** is to be found when an increase in one variable is accompanied by a decrease in the other variable. Negative correlations are prefaced with a minus sign (-). Thus, **-1.0** would represent perfect negative correlation.

2. The strength of a correlation



Researchers are often interested in the strength of a correlation rather than its direction. Theoretically speaking, correlation tests can give us *no correlation*, *low correlation*, *medium correlation*, *strong correlation*, or a *perfect correction*. The following figure provide a visualization of the possible correlation magnitudes.

However, "perfect correlations of +1.00 or -1.00 [or even 0.00] are rarely found and most coefficients of correlation in social research are around +0.50 or less". (Cohen, Manion, & Morrison, 2002, p. 530).

PRACTICE

Calculate the Pearson correlation with the following dataset.

Participants	**Attendance** (x)	**Marks** (y)
Student 1	4	12
Student 2	7	16
Student 3	2	8
Student 4	5	12
Student 5	3	10
Student 6	6	14
	Σ =	**Σ** =
	x̄ =	**ȳ** =
	***r* =**	

LECTURE 9: REPORTING RESULTS AND DATA VISUALIZATION

9.1. Data visualization

Data visualization is a process that involves converting raw data into graphical representations that highlight key trends and patterns in the results. Quantitative data can be visualized through tables, charts, graphs, infographics or geospatial figures. The process of data visualization is as important in reporting data as descriptive and advanced statistics (Larson-Hall & Plonsky, 2015). In Applied Linguistics, data visualization follows the APA formatting style.

Tables

Tables are often the first step when introducing datasets or summarizing data/results in rows and columns. The following table is an example of the table format in APA style 7th edition.

Adapted from (Publication Manual of the American Psychological Association 2020: The Official Guide to APA Style (7th Ed.), 2020, p. 60)

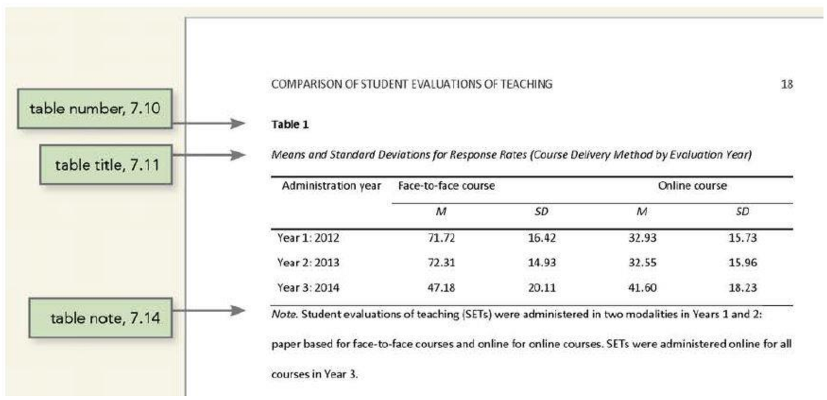


Table 1

Means and Standard Deviations for Response Rates (Course Delivery Method by Evaluation Year)

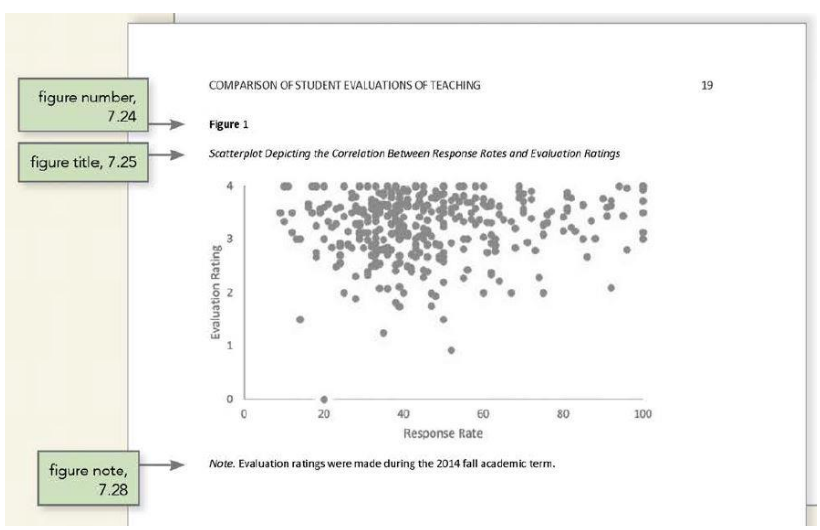
Administration year	Face-to-face course		Online course	
	M	*SD*	*M*	*SD*
Year 1: 2012	71.72	16.42	32.93	15.73
Year 2: 2013	72.31	14.93	32.55	15.96
Year 3: 2014	47.18	20.11	41.60	18.23

Note. Student evaluations of teaching (SETs) were administered in two modalities in Years 1 and 2: paper based for face-to-face courses and online for online courses. SETs were administered online for all courses in Year 3.

Figures

Figures are graphics that visualize data in different shapes that make the interpretation of results easy and accessible to the reader. The following figure demonstrate how figures should be reported according to the APA styles 7th edition.

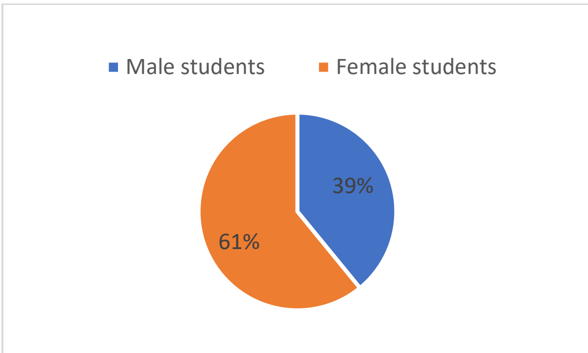
Adapted from (Publication Manual of the American Psychological Association 2020: The Official Guide to APA Style (7th Ed.), 2020, p. 60)



9.2. Data visualization techniques

9.2.1. Charts and graphs

Charts are graphical representations that summarize datasets in a concise and intuitive way. Charts can take the form of line charts, pie charts, or bar charts. Graphs, on the other hand, visualize trends over time and relationships between variables. The following figure shows the difference between charts and graphs:



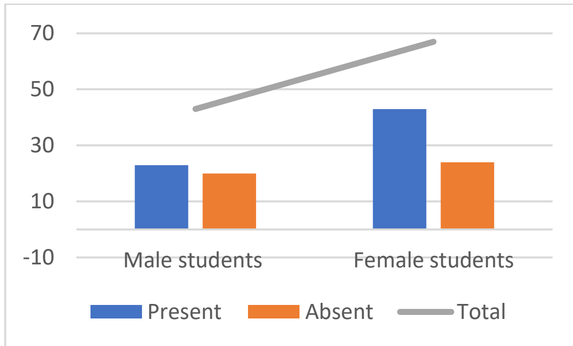
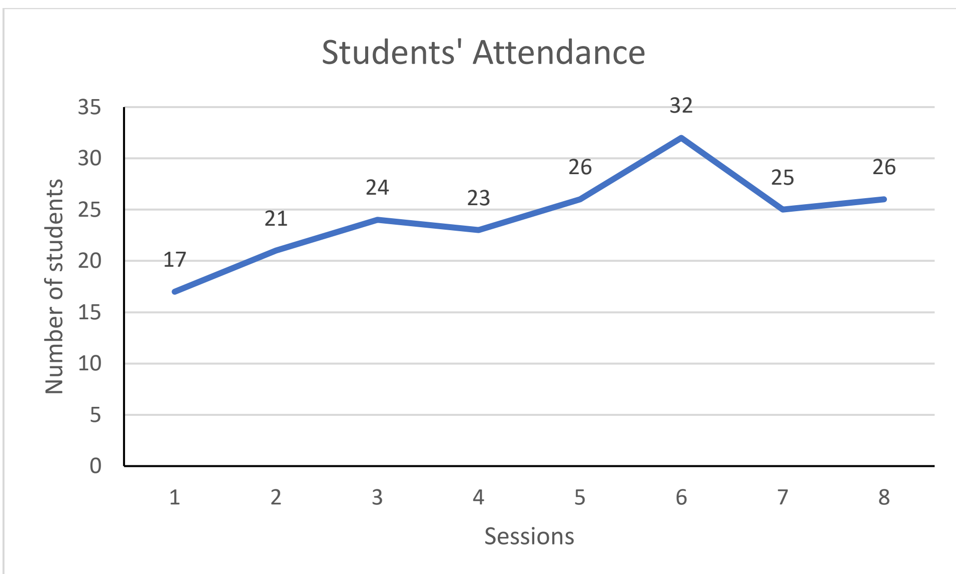


Chart **Graph**

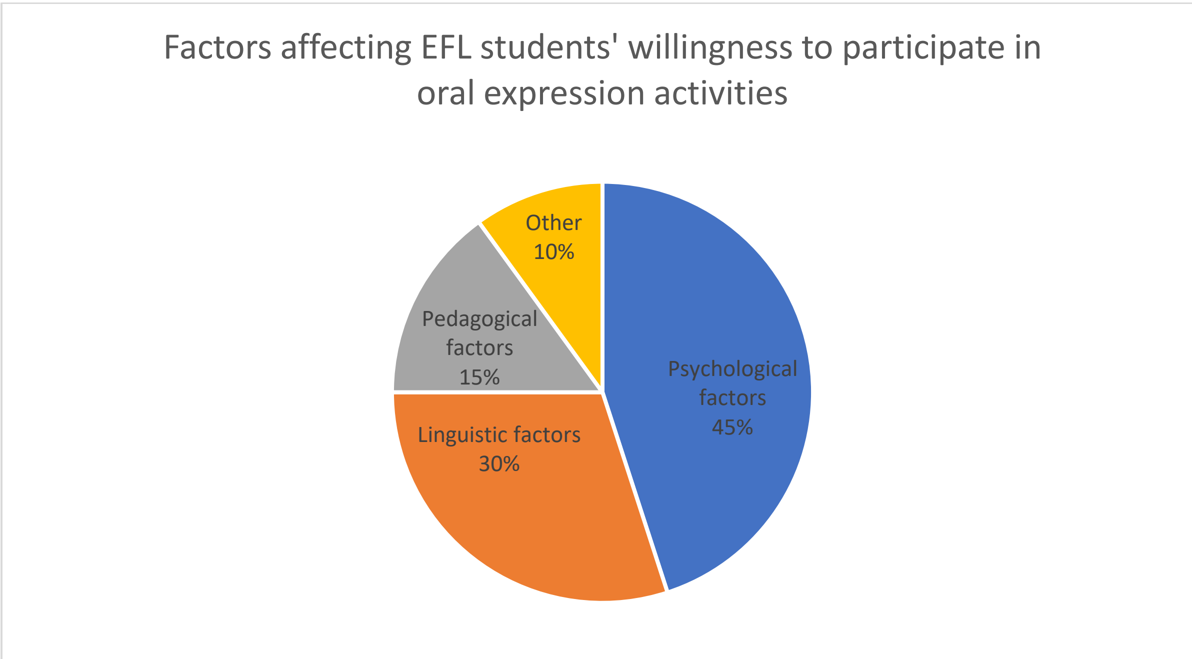
9.2.2. Line charts

A line chart visualizes data through a series of points or markers in relation to specific x and y axises. Line charts are well known for connecting markers with a straight line.



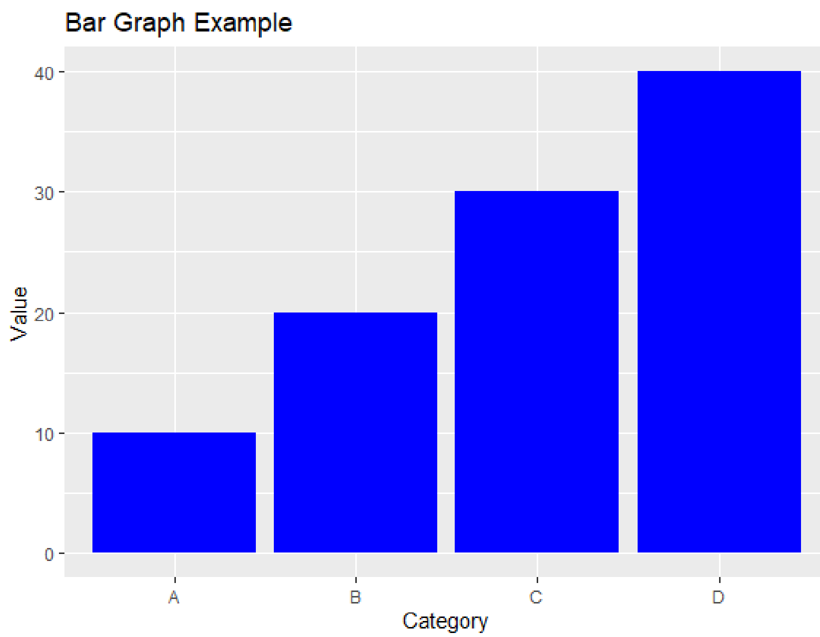
9.2.3. Pie chart

A pie chart is a circular visual representation often using percentages to illustrate proportions of particular variables in a whole dataset.



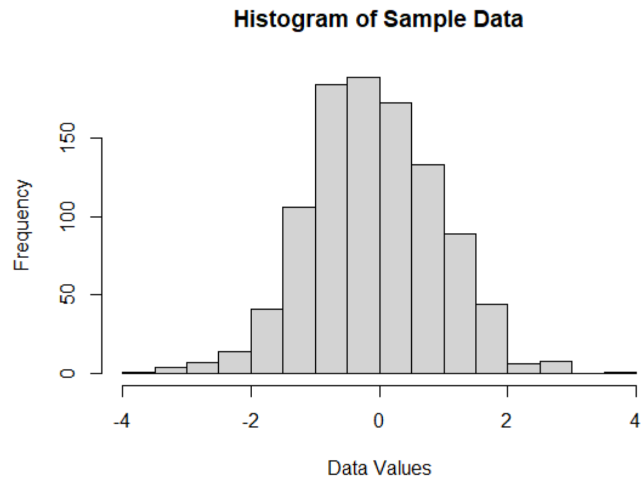
9.2.4. Bar chart

A bar chart is a chart that represents categorical data (nominal or ordinal) through bars that differ in height depending on the value they visualize.



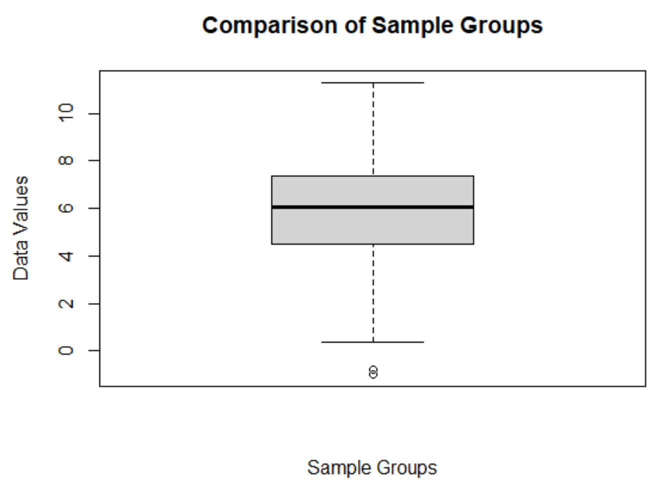
9.2.5. Histogram

A histogram is a type of a bar charts that plot the frequency of numerical data. Although the terms bar charts and histograms are sometimes used interchangeably, the chart types are different. While bar charts visualize categorical data, histograms visualize numerical data showing the distribution.



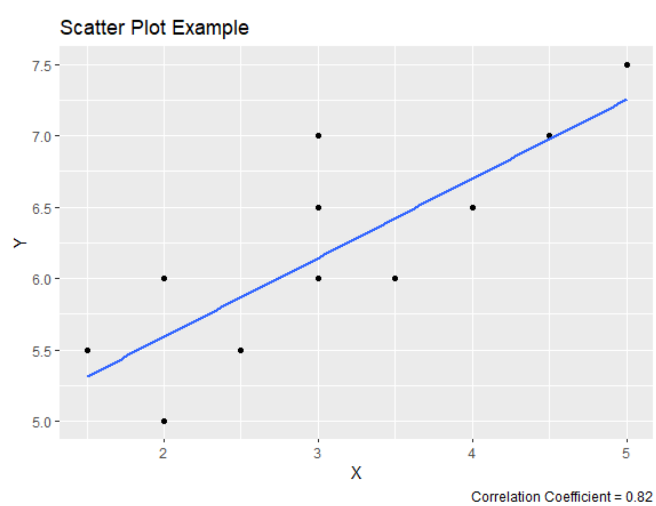
9.2.6. Boxplots

Boxplots are used to visually summarize the descriptive statistics of one or more groups of participants. This summary includes minimum and maxim values through the whiskers, the spread of data around the mean through the lower and upper quartiles (the box), and the media with the median line.



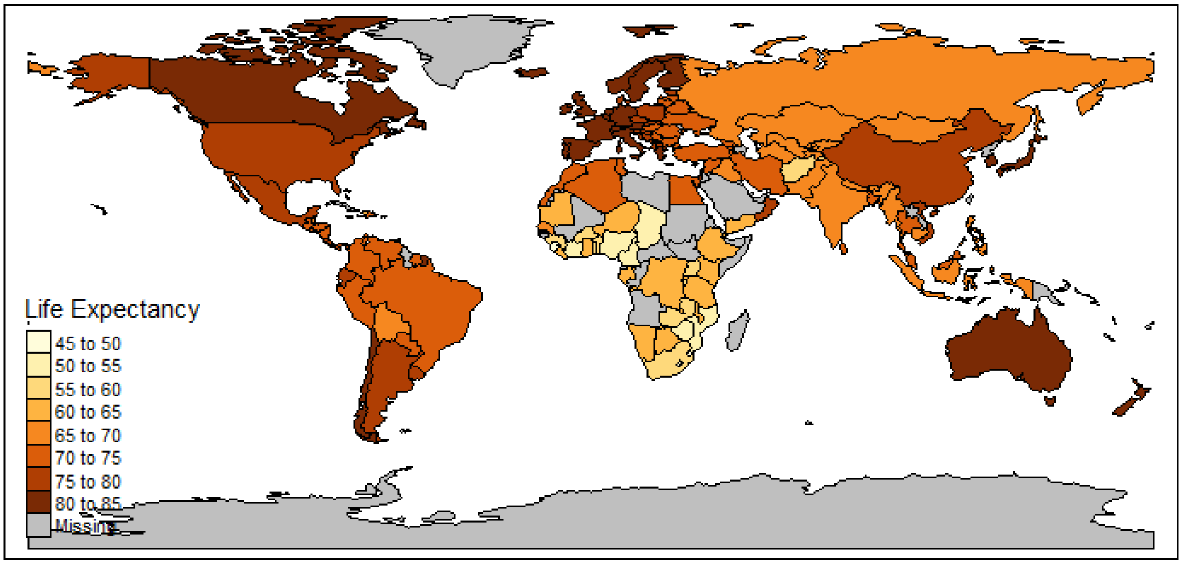
9.2.7. Scatter plots

Scatter plots are diagrams that visualize the relationship between two (or more) variables through dots plotted in relation to horizontal and vertical axes.



9.2.8. Geospatial

Geospatial refers to graphs visualizing statistics on geographical maps. The following example of a geospatial reports the number of Berber speakers in different regions of Algeria.



9.2.9. Infographics

Infographics are a visualization tool to represent the data in a creative and visually appealing way.

(29 nfographic Statistics You Need to Know in 2022, 2022)

9.3. Data visualization tools

Statistics analysis tools are essential for processing, analyzing, and visualizing data accurately. Three main statistics analysis tools used by researchers are Excel, SPSS, and R Studio. Excel is a widely used spreadsheet software that includes basic statistical functions, making it an easy-to-use tool for visualizing data. SPSS is a specialized software that is designed specifically for statistical analysis. It has advanced features for descriptive and inferential statistics and is commonly used in social science research. R Studio, on the other hand, is an open-source software that is highly customizable and provides a wide range of statistical and graphical tools for data analysis. It requires some programming skills, but its flexibility and power make it a popular choice for data scientists and researchers. The table bellow compares between these three tools in terms of use, functionality, scalability, cost, cusomizability, and reproducibility.

Feature	Excel	SPSS	R Studio
Ease of Use	Widely available and easy to learn	Easy to use automated data analysis options	Requires programming skills and lacks user-friendly interface
Statistical Functionality	Basic functions only	Comprehensive statistical features including advanced analyses and graphical capabilities	Customizable and programmable analyses with extensive statistical capabilities and visualization tools
Scalability	Poor scalability for larger datasets	Limited by hardware, but can handle moderate-sized datasets	Can handle large datasets and big data with ease
Cost	Affordable	Expensive	Open-source and free
Customizability and Flexibility	Customizable analyses and graphing	Limited customizability, but provides automation for repetitive tasks	Highly customizable with the ability to create own functions and packages
Reproducibility and Collaboration	Limited ability to reproduce analyses and results	Limited ability to reproduce analyses and results, but provides collaboration options	Highly reproducible analyses and results with collaboration options

LECTURE 10: REVISION AND PRACTICE

Calculate the sum, mean, variance, and standard deviation for the IELTS scores achieved by 8 EFL students over the period of three semesters.

Participants	**Time 1**			**Time 2**			**Time 3**		
student 1	2			3			4		
student 2	4			4			5		
student 3	3			5			5		
student 4	6			6			6		
student 5	4			4			5		
student 6	5			6			6		
student 7	3			4			4		
student 8	4			4			4		
Sum (Σ)									
Mean (x̄)									
Variance (S^2)									
SD (S)									

- Which inferential statistical test should be used with the dataset above?

ANSWERS

Participants	**Time 1**	*(xi - x̄)*	*(xi - x̄)^2*	**Time 2**	*(xi - x̄)*	*(xi - x̄)^2*	**Time 3**	*(xi - x̄)*	*(xi - x̄)^2*
student 1	2	*-1.88*	*3.53*	3	*-1.50*	*2.25*	4	*-0.88*	*0.77*
student 2	4	*0.12*	*0.01*	4	*-0.50*	*0.25*	5	*0.12*	*0.01*
student 3	3	*-0.88*	*0.77*	5	*0.50*	*0.25*	5	*0.12*	*0.01*
student 4	6	*2.12*	*4.49*	6	*1.50*	*2.25*	6	*1.12*	*1.25*
student 5	4	*0.12*	*0.01*	4	*-0.50*	*0.25*	5	*0.12*	*0.01*
student 6	5	*1.12*	*1.25*	6	*1.50*	*2.25*	6	*1.12*	*1.25*
student 7	3	*-0.88*	*0.77*	4	*-0.50*	*0.25*	4	*-0.88*	*0.77*
student 8	4	*0.12*	*0.01*	4	*-0.50*	*0.25*	4	*-0.88*	*0.77*
SUM	**31.00**		*10.88*	**36.00**		*8.00*	**39.00**		*4.88*
MEAN	**3.88**			**4.50**			**4.88**		
VAR	**1.55**			**1.14**			**0.70**		
SD	**1.25**			**1.07**			**0.83**		

- The inferential test that should be used is the non-parametric Friedman test.

References

Cohen, L., Manion, L., & Morrison, K. (2002). Research Methods in Education. *Research Methods in Education*. https://doi.org/10.4324/9780203224342

Davey Smith, G., & Ebrahim, S. (2002). Data dredging, bias, or confounding. *British Medical Journal*, *325*(7378), 1437–1438. https://doi.org/10.1136/bmj.325.7378.1437

Dörnyei, Z. (2007). *Research Methods in Applied Linguistics*. Oxford university press.

Larson-Hall, J., & Plonsky, L. (2015). Reporting and Interpreting Quantitative Research Findings: What Gets Reported and Recommendations for the Field. *Language Learning*, *65*(S1), 127–159. https://doi.org/10.1111/LANG.12115

Mackey, A., & Gass, S. M. (2015). Second language research: Methodology and design, second edition. In *Second Language Research: Methodology and Design, Second Edition*. Taylor and Francis. https://doi.org/10.4324/9781315750606/SECOND-LANGUAGE-RESEARCH-ALISON-MACKEY-SUSAN-GASS

Phakiti, A. (2010). Analyzing Quantitative Data. In B. Paltridge & P. Aek (Eds.), *Continuum companion to research methods in applied linguistics* (pp. 39–49). A&C Black.

Publication manual of the American Psychological Association 2020: the official guide to APA style (7th ed.) (7th ed.). (2020). American Psychological Association.

Romeijn, J.-W. (2014). *Philosophy of Statistics*. https://plato.stanford.edu/entries/statistics/

Shuttleworth, M., & Wilson, L. (n.d.). *Type I Error and Type II Error - Experimental Errors in Research*. Explorable.Com. Retrieved February 22, 2023, from https://explorable.com/type-i-error

Vickery, R. (2021, January 30). *8 Fundamental Statistical Concepts for Data Science*. Towardsdatascience.Com. https://towardsdatascience.com/8-fundamental-statistical-concepts-for-data-science-9b4e8a0c6f1c

Appendices

Appendix 1. Exam 1 (January, 2023) – Questions

ACTIVITY ONE (10 points)

- Tick (✓) the correct answer.

1. Non-parametric test is used when the researcher:

◯ **A.** Recruits volunteers
◯ **B.** Recruits a sample size of more than 30 participants
◯ **C.** Seeks to reflect the population's characteristics
◯ **D.** Selects every member of the population

2. 73 students rated library services for 5 years (*1 = not effective, 10 = very effective*). The test used to assess the difference between the yearly ratings is:

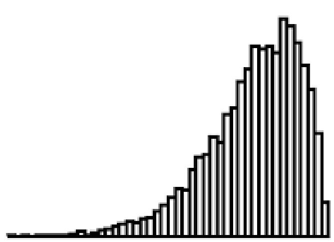
◯ **A.** Kruskal-Wallis test
◯ **B.** Friedman test
◯ **C.** One-way ANOVA test
◯ **D.** Repeated Measures ANOVA test

3. Age group (e.g, 0 – 18, 19 – 30 …) is considered:

◯ **A.** Nominal data
◯ **B.** Ordinal data
◯ **C.** Interval data
◯ **D.** Ratio data

4. A Type 2 error happens when we:

◯ **A.** Accept a true null hypothesis
◯ **B.** Reject a true null hypothesis
◯ **C.** Accept a false null hypothesis
◯ **D.** Reject a false null hypothesis



5. The histogram above shows:

◯ **A.** A normal distribution
◯ **B.** A negatively skewed distribution
◯ **C.** A positively skewed distribution
◯ **D.** None of the above

6. P = 0.03 in *Kolmogorov–Smirnov* test means:

◯ **A.** Data are normally distributed
◯ **B.** Data are not normally distributed
◯ **C.** Data are significant
◯ **D.** Data are not significant

7. A correlation between EFL learners' confidence level and their reported willingness to speak in the classroom requires:

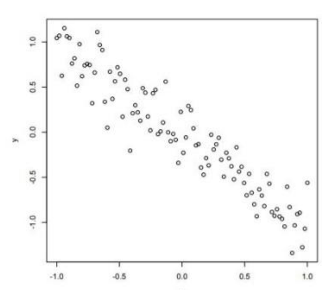
◯ **A.** Pearson correlation
◯ **B.** Spearman correlation
◯ **C.** Multiple regression
◯ **D.** Point-Biserial correlation

8. Parametric statistical tests are used when:

◯ **A.** Participants are randomly recruited
◯ **B.** The data are interval
◯ **C.** The data are normally distributed
◯ **D.** All of the above

9. The null hypothesis (H_0) is:

◯ **A.** Positive
◯ **B.** Negative
◯ **C.** Directional
◯ **D.** Neutral



10. The plot above shows a correlation close to:

◯ **A.** -1
◯ **B.** -0.5
◯ **C.** +1.0
◯ **D.** +0.5

Turn the page …

ACTIVITY TWO (5 points)

- Indicate the sampling approach described in each of the following statements.

9. The researcher thinks that some subjects are more fit for the research compared to other individuals — **1.**________________________

10. Used when the population has different groups and those groups need to be fairly represented in the sample — **2.**________________________

11. Each unit in the population has an equal probability of being selected in the sample — **3.**________________________

12. The researcher asks the initial subject to identify potential participants — **4.** ______________________

13. Involves taking samples according to some specified rule e.g., every fourth unit — **5.** ______________________

ACTIVITY THREE (5 points)

- Calculate the sum, the sample mean, the sample variance, and the sample standard deviation for the grammatical accuracy results and report them in the table below.

Algerian EFL learners	Grammatical errors	Chinese EFL learners	Grammatical errors
Student 1	18	*Student 1*	29
Student 2	42	*Student 2*	51
Student 3	34	*Student 3*	35
Student 4	11	*Student 4*	21
Student 5	27	*Student 5*	42

RESULTS

Groups	n	Σ	$\bar{X}$	S^2	S
Algerian EFL learners					
Chinese EFL learners					

Assuming the data are not normally distributed, what is the most appropriate inferential statistical test to assess the difference between Algerian and Chinese EFL learners' grammatical errors?

____

BEST OF LUCK!

Appendix 2. Exam 1 (January, 2023) – Answers and marking rubric

ACTIVITY ONE (10 points) **(1 point for each correct answer)**

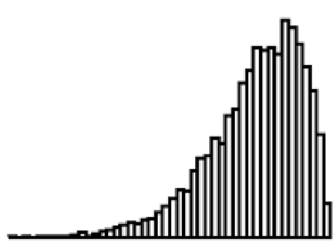
- Tick (✓) the correct answer.

1. Non-parametric test is used when the researcher:
◯ **A.** Recruits volunteers
◯ **B.** Recruits a sample size of more than 30 participants
◯ **C.** Seeks to reflect the population's characteristics
◯ **D.** Selects every member of the population

2. 73 students rated library services for 5 years (*1 = not effective, 10 = very effective*). The test used to assess the difference between the yearly ratings is:
◯ **A.** Kruskal-Wallis test
◯ **B.** Friedman test
◯ **C.** One-way ANOVA test
◯ **D.** Repeated Measures ANOVA test

3. Age group (e.g, 0 – 18, 19 – 30 …) is considered:
◯ **A.** Nominal data
◯ **B.** Ordinal data
◯ **C.** Interval data
◯ **D.** Ratio data

4. A Type 2 error happens when we:
◯ **A.** Accept a true null hypothesis
◯ **B.** Reject a true null hypothesis
◯ **C.** Accept a false null hypothesis
◯ **D.** Reject a false null hypothesis



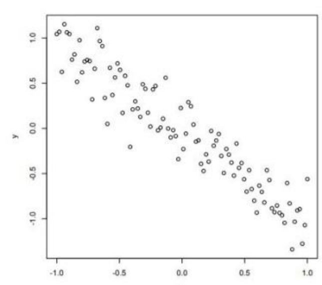
5. The histogram above shows:
◯ **A.** A normal distribution
◯ **B.** A negatively skewed distribution
◯ **C.** A positively skewed distribution
◯ **D.** None of the above

6. *P* = 0.03 in *Kolmogorov–Smirnov* test means:
◯ **A.** Data are normally distributed
◯ **B.** Data are not normally distributed
◯ **C.** Data are significant
◯ **D.** Data are not significant

7. A correlation between EFL learners' confidence level and their reported willingness to speak in the classroom requires:
◯ **A.** Pearson correlation
◯ **B.** Spearman correlation
◯ **C.** Multiple regression
◯ **D.** Point-Biserial correlation

8. Parametric statistical tests are used when:
◯ **A.** Participants are randomly recruited
◯ **B.** The data are interval
◯ **C.** The data are normally distributed
◯ **D.** All of the above

9. The null hypothesis (H_0) is:
◯ **A.** Positive
◯ **B.** Negative
◯ **C.** Directional
◯ **D.** Neutral



10. The plot above shows a correlation close to:
◯ **A.** -1
◯ **B.** -0.5
◯ **C.** +1.0
◯ **D.** +0.5

Turn the page …

ACTIVITY TWO (5 points)

- Indicate the sampling approach described in each of the following statements.

The researcher thinks that some subjects are more fit for the research compared to other individuals	**1. PURPOSIVE SAMPLING (1 pts)**
Used when the population has different groups and those groups need to be fairly represented in the sample	**2. STRATIFIED RANDOM SAMPLING (1 pts)**
Each unit in the population has an equal probability of being selected in the sample	**3.RANDOM SAMPLING (1 pts)**
The researcher asks the initial subject to identify potential participants	**4.SNOWBALL SAMPLING (1 pts)**
Involves taking samples according to some specified rule e.g., every fourth unit	**5. SYSTEMATIC SAMPLING (1 pts)**

ACTIVITY THREE (5 points)

- Calculate the sum, the sample mean, the sample variance, and the sample standard deviation for the grammatical accuracy results and report them in the table below.

Algerian EFL learners	Grammatical errors	Chinese EFL learners	Grammatical errors
Student 1	18	*Student 1*	29
Student 2	42	*Student 2*	51
Student 3	34	*Student 3*	35
Student 4	11	*Student 4*	21
Student 5	27	*Student 5*	42

RESULTS

Groups	n (0,5 pts)	Σ (0,5 pts)	$\bar{X}$ (0,5 pts)	S^2 (1,5 pts)	S (1,5 pts)
Algerian EFL learners	**5**	**132**	**26.40**	**152.30**	**12.34**
Chinese EFL learners	**5**	**178**	**35.60**	**133.80**	**11.57**

Assuming the data are not normally distributed, what is the most appropriate inferential statistical test to assess the difference between Algerian and Chinese EFL learners' grammatical errors?

The inferential statistical test is: **THE MANN–WHITNEY U TEST** **(0,5 pts)**

Appendix 3. Exam 2 (January, 2022) – Exam 2

ACTIVITY ONE (10 points)

Tick (✓) the correct answer.

- The median in [114, 70, 74, 80, 72, 76] is:

◯ 74
◯ 74.5
◯ 75
◯ 75.5

- To investigate the relationship between a single dependent variable and several independent variables, we use:

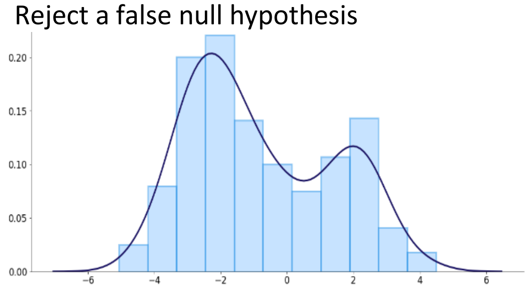
◯ Pearson correlation
◯ Spearman correlation
◯ Multiple regression
◯ Point-Biserial correlation

- The average time for completing a reading task is:

◯ Nominal data
◯ Ordinal data
◯ Interval data
◯ Ratio data

- A Type 1 error happens when we:

◯ Accept a true null hypothesis
◯ Reject a true null hypothesis
◯ Accept a false null hypothesis
◯ Reject a false null hypothesis



- The histogram above shows:

◯ A normal distribution
◯ A non-normal distribution
◯ A skewed distribution
◯ A normal distribution that is skewed

- *P* = 0.76 in *Kolmogorov–Smirnov* test means:

◯ Data are normally distributed
◯ Data are not normally distributed
◯ Data are significant
◯ Data are not significant

- A correlation between EFL learners' confidence level and their reported willingness to speak in the classroom requires:

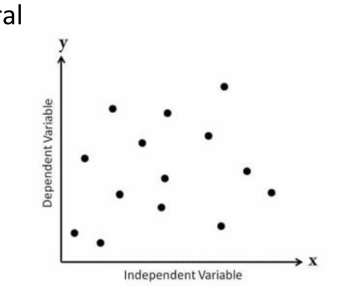
◯ Pearson correlation
◯ Spearman correlation
◯ Multiple regression
◯ Point-Biserial correlation

- Non-parametric statistical tests are used when:

◯ Participants are randomly recruited
◯ The data are interval
◯ All of the above
◯ None of the above

- The null hypothesis (H_0) is:

◯ Positive
◯ Negative
◯ Directional
◯ Neutral



- The plot above shows a correlation close to:

◯ **A.** r = -1
◯ **B.** r = -0.5
◯ **C.** r = 0.0
◯ **D.** r = +0.5

Turn the page …

ACTIVITY TWO (5 points)

- Consider the following scenarios and write the appropriate statistical test.

SCENARIO	ASSUMPTIONS	TEST
A researcher investigated the relationship between out-of-class reading time and grammatical accuracy	• Non-normal distribution of data • The sample consisted of 54 randomly recruited participants	
A researcher tested the development of EFL learners' use of simple word stress during three occasions of the academic year	• Normal distribution of data • The sample consisted of 120 randomly recruited participants	
A researcher tested ESL students' oral fluency before and after using games during the Oral Expression sessions	• Normal distribution of data • The sample consisted of 82 randomly recruited participants	
A researcher used Likert scale questions (***1 = Not useful, 5 = Very useful***) to compare between male and female students' opinions about library services	• Normal distribution of data • The sample consisted of 60 randomly recruited participants	
A researcher compared between the pronunciation intelligibility of EFL learners from Algeria, China, and Thailand	• Normal distribution of data • The sample consisted of 24 participants (8 per-group)	

ACTIVITY THREE (5 points)

Calculate the sum, the sample mean, the sample variance, and the sample standard deviation for the grammatical accuracy results and report them in the table below.

Algerian EFL learners	Grammatical errors	Chinese EFL learners	Grammatical errors
Student 1	29	*Student 1*	39
Student 2	46	*Student 2*	17
Student 3	37	*Student 3*	22
Student 4	55	*Student 4*	27
Student 5	42	*Student 5*	30

RESULTS

Groups	n	Σ	$\bar{X}$	S^2	S
Algerian EFL learners					
Chinese EFL learners					

Assuming the data are normally distributed, what is the most appropriate inferential statistical test to assess the difference between Algerian and Chinese EFL learners' grammatical errors?

………………………………………………………………….

BEST OF LUCK!

Appendix 4. Exam 2 (January, 2022) – Answers and marking rubric

ACTIVITY ONE (10 points) **(1 point for each correct answer)**

Tick (✓) the correct answer.

- The median in [114, 70, 74, 80, 72, 76] is:

◯ 74
◯ 74.5
◯ 75
◯ 75.5

- To investigate the relationship between a single dependent variable and several independent variables, we use:

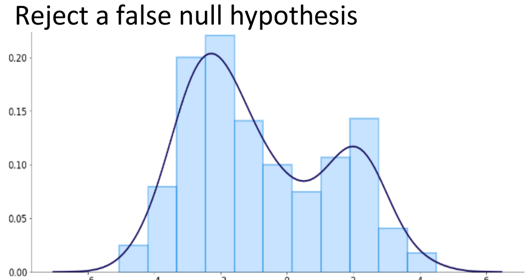
◯ Pearson correlation
◯ Spearman correlation
◯ Multiple regression
◯ Point-Biserial correlation

- The average time for completing a reading task is:

◯ Nominal data
◯ Ordinal data
◯ Interval data
◯ Ratio data

- A Type 1 error happens when we:

◯ Accept a true null hypothesis
◯ Reject a true null hypothesis
◯ Accept a false null hypothesis
◯ Reject a false null hypothesis



- The histogram above shows:

◯ A normal distribution
◯ A non-normal distribution
◯ A skewed distribution
◯ A normal distribution that is skewed

- *P* = 0.76 in *Kolmogorov–Smirnov* test means:

◯ Data are normally distributed
◯ Data are not normally distributed
◯ Data are significant
◯ Data are not significant

- A correlation between EFL learners' confidence level and their reported willingness to speak in the classroom requires:

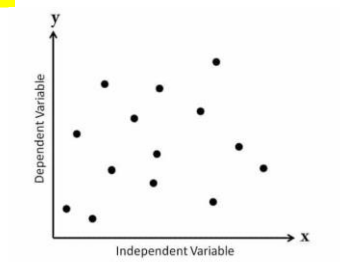
◯ Pearson correlation
◯ Spearman correlation
◯ Multiple regression
◯ Point-Biserial correlation

- Non-parametric statistical tests are used when:

◯ Participants are randomly recruited
◯ The data are interval
◯ All of the above
◯ None of the above

- The null hypothesis (H_0) is:

◯ Positive
◯ Negative
◯ Directional
◯ Neutral



- The plot above shows a correlation close to:

◯ **A.** r = -1
◯ **B.** r = -0.5
◯ **C.** r = 0.0
◯ **D.** r = +0.5

Turn the page …

ACTIVITY TWO (5 points) **(1 point for each correct answer)**

- Consider the following scenarios and write the appropriate statistical test.

SCENARIO	ASSUMPTIONS	TEST
A researcher investigated the relationship between out-of-class reading time and grammatical accuracy	• Non-normal distribution of data • The sample consisted of 54 randomly recruited participants	**Spearman correlation**
A researcher tested the development of EFL learners' use of simple word stress during three occasions of the academic year	• Normal distribution of data • The sample consisted of 120 randomly recruited participants	**Repeated measures ANOVA**
A researcher tested ESL students' oral fluency before and after using games during the Oral Expression sessions	• Normal distribution of data • The sample consisted of 82 randomly recruited participants	**Paired t-test**
A researcher used Likert scale questions (***1 = Not useful, 5 = Very useful***) to compare between male and female students' opinions about library services	• Normal distribution of data • The sample consisted of 60 randomly recruited participants	**Mann-Whitney U test**
A researcher compared between the pronunciation intelligibility of EFL learners from Algeria, China, and Thailand	• Normal distribution of data • The sample consisted of 24 participants (8 per-group)	**Kruskal Wallis Test**

ACTIVITY THREE (5 points)

Calculate the sum, the sample mean, the sample variance, and the sample standard deviation for the grammatical accuracy results and report them in the table below.

Algerian EFL learners	Grammatical errors	Chinese EFL learners	Grammatical errors
Student 1	29	*Student 1*	39
Student 2	46	*Student 2*	17
Student 3	37	*Student 3*	22
Student 4	55	*Student 4*	27
Student 5	42	*Student 5*	30

RESULTS

Groups	n	Σ	$\bar{X}$	S^2	S
Algerian EFL learners	5 **(0.25 pts)**	209 **(0.25 pts)**	41.8 **(0.25 pts)**	94.70 **(0.75 pts)**	9.73 **(0.75 pts)**
Chinese EFL learners	5 **(0.25 pts)**	135 **(0.25 pts)**	27 **(0.25 pts)**	69.50 **(0.75 pts)**	8.34 **(0.75 pts)**

Assuming the data are normally distributed, what is the most appropriate inferential statistical test to assess the difference between Algerian and Chinese EFL learners' grammatical errors?

Independent t-test **(0.5 pts)**

Appendix 5. Exam 3 (January, 2023) – Questions

ACTIVITY ONE (10 points)

- Tick (✓) the correct answer (CHOOSE ONLY **ONE** ANSWER).

- A representative sample:
  - ◯ Is recruited using non-probability sampling
  - ◯ Guarantees generalizability
  - ◯ Seeks to reflect the population's characteristics
  - ◯ Selects every member of the population

- A teacher assessed his students' pronunciation on a scale from ***1 = Incomprehensible*** to ***5 = Native-like***. The data are:
  - ◯ Nominal data
  - ◯ Ordinal data
  - ◯ Interval data
  - ◯ Ratio data

- In the *Shapiro wilk* test, P = 0.53, means the:
  - ◯ Data are normally distributed
  - ◯ Data are not normally distributed
  - ◯ Data are significant
  - ◯ Data are not significant

- The mode in [63, 62, 66, 67, 63, 70, 67, 68, 61] is:
  - ◯ 66 and 67
  - ◯ 63 and 67
  - ◯ 66
  - ◯ 70

- The median in [1, 2, 3, 4, 5, 6, 8, 9] is:
  - ◯ 4
  - ◯ 5
  - ◯ 4 and 5
  - ◯ 4.5

- If $\sigma^2 = 13.77$, then:
  - ◯ $\sigma = 1.69$
  - ◯ $\sigma = 2.70$
  - ◯ $\sigma = 3.71$
  - ◯ $\sigma = 4.72$

- To compare pretest and posttest vocabulary test results of a randomly recruited sample, the researcher should use:
  - ◯ Repeated measures ANOVA
  - ◯ Independent t-test
  - ◯ One way ANOVA
  - ◯ Paired t-test

- Non-Parametric statistical tests are used when:
  - ◯ The sample is randomly recruited
  - ◯ The data are interval
  - ◯ All of the above
  - ◯ None of the above

- The null hypothesis (H_0) is:
  - ◯ Assumed to be true until proven otherwise
  - ◯ Assumed to be false until proven otherwise
  - ◯ A hypothesis that cannot be proven
  - ◯ A hypothesis that cannot be disproven

- A correlation can be:
  - ◯ Positive
  - ◯ Negative
  - ◯ All of the above
  - ◯ None of the above

Turn the paper …

ACTIVITY TWO (5 points)

The following table represent a sheet used by an English teacher to keep track of students' attendance in the Oral Expression module (**X** = present, **O** = Absent). Calculate the frequency of attendance and percentage of sessions attended for each student.

	Student 1	***Student 2***	***Student 3***	***Student 4***	***Student 5***	***Student 6***
Session 1	X	O	O	X	X	X
Session 2	O	X	X	X	O	X
Session 3	X	X	O	X	O	X
Session 4	X	X	O	X	X	X
Session 5	O	X	O	X	O	O
Session 6	O	X	O	X	O	O

	Student 1	*Student 2*	*Student 3*	*Student 4*	*Student 5*	*Student 6*
Frequency of attended sessions	***E.g.*** 3	....	....	....	....	....
Percentage of sessions attended	***E.g.*** 50.00 %	.... %	.... %	.... %	.... %	.... %

ACTIVITY THREE (5 points)

- Calculate the sum, the sample mean, the sample variance, and the sample standard deviation for the two tests in the data set below.

PARTICIPANTS		PRETEST	POSTTEST
Student 1		29	39
Student 2		46	17
Student 3		37	22
Student 4		55	27
Student 5		42	30
(0.5 pts)	n	....	....
(0.5 pts)	Σ	....	....
(0.5 pts)	$\bar{X}$	....	....
(1.5 pts)	S^2	....	....
(1.5 pts)	S	....	....

Assuming the data are normally distributed, what the most appropriate inferential statistical test with the above data? *(0.5 pts)*

Answer: ________________________________

BEST OF LUCK!

Appendix 6. Exam 3 (January, 2023) – Answers and marking rubric

ACTIVITY ONE (10 points) **(1 point for each correct answer)**

- Tick (✓) the correct answer (CHOOSE ONLY **ONE** ANSWER).

- A representative sample:
  - ◯ Is recruited using non-probability sampling
  - ◯ Guarantees generalizability
  - ◯ Seeks to reflect the population's characteristics
  - ◯ Selects every member of the population

- A teacher assessed his students' pronunciation on a scale from ***1 = Incomprehensible*** to ***5 = Native-like***. The data are:
  - ◯ Nominal data
  - ◯ Ordinal data
  - ◯ Interval data
  - ◯ Ratio data

- In the *Shapiro wilk* test, *P* = 0.53, means the:
  - ◯ Data are normally distributed
  - ◯ Data are not normally distributed
  - ◯ Data are significant
  - ◯ Data are not significant

- The mode in [63, 62, 66, 67, 63, 70, 67, 68, 61] is:
  - ◯ 66 and 67
  - ◯ 63 and 67
  - ◯ 66
  - ◯ 70

- The median in [1, 2, 3, 4, 5, 6, 8, 9] is:
  - ◯ 4
  - ◯ 5
  - ◯ 4 and 5
  - ◯ 4.5

- If $\sigma^2 = 13.77$, then:
  - ◯ $\sigma = 1.69$
  - ◯ $\sigma = 2.70$
  - ◯ $\sigma = 3.71$
  - ◯ $\sigma = 4.72$

- To compare pretest and posttest vocabulary test results of a randomly recruited sample, the researcher should use:
  - ◯ Repeated measures ANOVA
  - ◯ Independent t-test
  - ◯ One way ANOVA
  - ◯ Paired t-test

- Non-Parametric statistical tests are used when:
  - ◯ The sample is randomly recruited
  - ◯ The data are interval
  - ◯ All of the above
  - ◯ None of the above

- The null hypothesis (H_0) is:
  - ◯ Assumed to be true until proven otherwise
  - ◯ Assumed to be false until proven otherwise
  - ◯ A hypothesis that cannot be proven
  - ◯ A hypothesis that cannot be disproven

- A correlation can be:
  - ◯ Positive
  - ◯ Negative
  - ◯ All of the above
  - ◯ None of the above

Turn the paper …

ACTIVITY TWO (5 points)

The following table represent a sheet used by an English teacher to keep track of students' attendance in the Oral Expression module (**X** = present, **O** = Absent). Calculate the frequency of attendance and percentage of sessions attended for each student.

	Student 1	***Student 2***	***Student 3***	***Student 4***	***Student 5***	***Student 6***
Session 1	X	O	O	X	X	X
Session 2	O	X	X	X	O	X
Session 3	X	X	O	X	O	X
Session 4	X	X	O	X	X	X
Session 5	O	X	O	X	O	O
Session 6	O	X	O	X	O	O

	Student 1	*Student 2*	*Student 3*	*Student 4*	*Student 5*	*Student 6*
Frequency of attended sessions	***E.g.*** 3	**5** **(0.5 pts)**	**1** **(0.5 pts)**	**6** **(0.5 pts)**	**2** **(0.5 pts)**	**4** **(0.5 pts)**
Percentage of sessions attended	***E.g.*** 50.00 %	**83.33%** **(0.5 pts)**	**16.66%** **(0.5 pts)**	**100%** **(0.5 pts)**	**33.33%** **(0.5 pts)**	**66.66%** **(0.5 pts)**

ACTIVITY THREE (5 points)

- Calculate the sum, the sample mean, the sample variance, and the sample standard deviation for the two tests in the data set below.

PARTICIPANTS		**PRETEST**	**POSTTEST**
Student 1		29	39
Student 2		46	17
Student 3		37	22
Student 4		55	27
Student 5		42	30
(0.5 pts)	***n***	5 **(0.25 pts)**	5 **(0.25 pts)**
(0.5 pts)	Σ	209 **(0.25 pts)**	135 **(0.25 pts)**
(0.5 pts)	$\bar{X}$	41.8 **(0.25 pts)**	27 **(0.25 pts)**
(1.5 pts)	S^2	94.70 **(0.75 pts)**	69.50 **(0.75 pts)**
(1.5 pts)	S	9.73 **(0.75 pts)**	8.34 **(0.75 pts)**

Assuming the data are normally distributed, what the most appropriate inferential statistical test with the above data? *(0.5 pts)*

Answer: **Paired t-test** **(0.5 pts)**

BEST OF LUCK!