ALQUSOUR ACADEMY
اكاديمية القصور

Q.A.J.U.S.T

# BIOSTATISTICS

#لجنة_كلية_الصيدلة

## لطلبة الصيدلة والعلوم الطبية

## Subject:

## First Exam – Part One

ALQUSOUR ACADEMY

خاص
للفصل الدراسي الأول
2016/2015

Prod. Date: **16/10/2015** Pages: **18** Price: **60**

ساعـــات الـــدوام الرسمـــي

السبت - الخميس: 11:00 ظهراً - 12:00 ليلاً
الجمعـــــــة: 2:00 ظهراً - 12:00 ليلاً

**تحذير: محاضراتنا (الملخصات) متوفرة لدى أكاديمية القصور .**

اربد / بجانب اربيلا مول / سلام سنتر / الطابق الرابع

للإستفسار والتسجيل: 0785 70 60 08 / 0795 33 99 34

Like facebook.com/alqusouracademy.page

# Introduction

## What is biostatistics?

**Biostatistics:** is the science of collecting biological (or life) data , analyzing, and interpreting (draw conclusions) data to understand and improve human health.

Example:
We have 600 students mark… this is the data.
We analyze data and found that average marks for females is 90% and for males is 60%.
Can we conclude that females are smarter??
To conclude: We should be careful before concluding that females are smarter than males. It is possible that males and females differ in sample number.

## Why do we need biostatistics?

1) Essential for scientific methods for investigation .
   - Formulate hypothesis:

   Example: My hypothesis is: there is a relation between vitamin D deficiency and miscourage.

   To investigate whether this is true or not we need biostatistics to conduct study.
   - Design a study to test hypothesis.
   - Collect reliable and unbiased data.
   - Evaluate data and draw conclusions.

   Example: From death rates we know the leading cause of death in the country.

2) To understand and critique scientific literature.

   So we can judge studies through looking at:
   - Sources of data.
   - How did they conduct the stud y.
   - How did they collected data. (should be without bias)
   - The sample size.

---

Example 1:

A medical school in Jordan (not JUST) reports that average starting salary of its graduate was over 4000 JD/month.

Are we in the wrong school (in JUST) ?

With a close look to their data we find that only 30% of graduating students returned the salary survey.

Note that students who are successful are more likely to fill out the survey..... this is bias.

Then we conclude that salary estimates are optimistic.

---

Example 2:

In a study we took a sample of patients from clinic and compare their biological markers with a sample of controls from population.

Biomarker: is any measurement from inside your body.

Example: high liver enzyme is a biomarker for liver injury.

The bias here is that: usually those attending clinics are females, educated and affluent thus our comparison with controls will give misleading results.

We can't determine if the biomarker is associated with the disease or with gender, class or education.

---

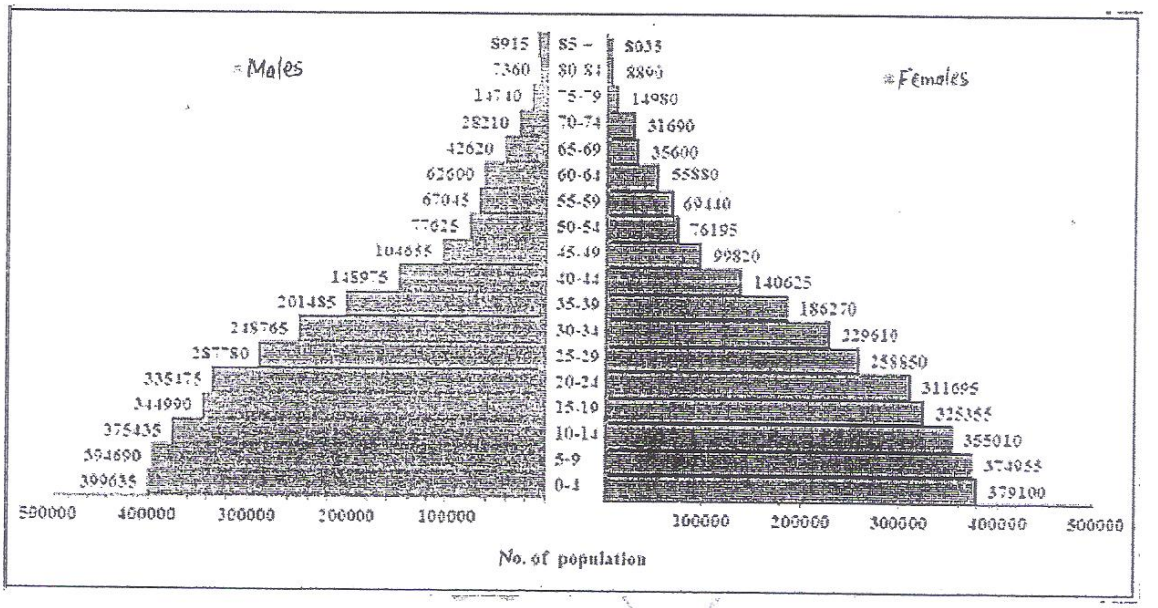
## Some indicators and numbers about Jordan:

| Item | 2012 |
|---|---|
| Area ($Km^2$) | 88778 |
| Total population (million) | 6.4 |
| % Population growth rate | 2.2 |
| Total fertility rate | 3.8 |
| Adult male illiteracy rate (% for > 15 yrs of age) | 3.5 |
| Adult female illiteracy rate (% for >15yrs of age) | 10.1 |
| Crude birth rate (per 1000 pop) | 28.1 |
| Total life expectancy (expected years to live)<br>71.6 males and 74.4 females | 73 |
| Crude death rate (per 1000 pop) | 7 |
| Infant mortality rate (per 1000 live births) | 23 |
| Maternal mortality rate per 100000 live births | 19.1 |
| Unemployment rate (%) | 12.2 |

**Table comments:**

- The growth rate 2.2 this means population will be double next time.
- Maternal mortality rate is high.

# AGE DISTRIBUTION OF JORDANIAN POPULATION (2010)



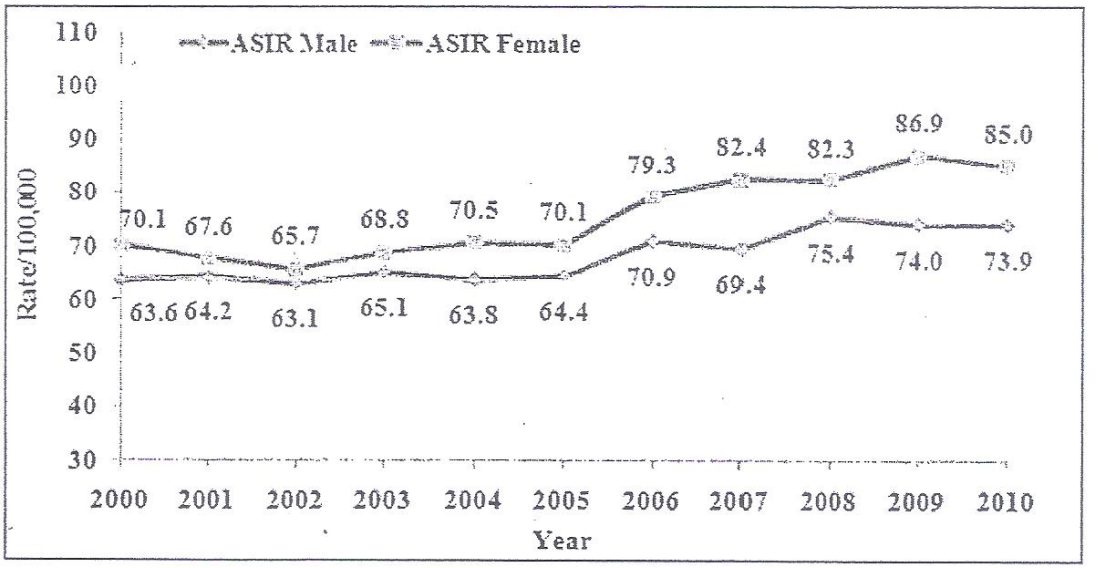

It is a pyramid because more people die with aging.

**Metabolic Syndrome in Jordan:**

| Disease | Total % | Males % | Females % |
|---|---|---|---|
| Hypertension | 44.6 | 42.2 | 46.0 |
| Diabetes | 19.5 | 19.7 | 19.8 |
| High total blood cholesterol | 36.0 | 35.6 | 36.8 |
| Cholesterol HDLP | 60 | 61.5 | 59.2 |
| Triglyceride | 38.6 | 48.1 | 34.2 |
| Overweight | 30.2 | 35.5 | 22.5 |
| Obesity BMI | 44.1 | 12.3 | 60.7 |
| Metabolic Syndrome (having at least 3 diseases from above) | 36.3 | 28.7 | 40.9 |

# CANCER IN JORDAN 2000-2010

*Figure (5) Trend of Crude incidence rate of cancer by sex, 2000-2010 , Jordan*

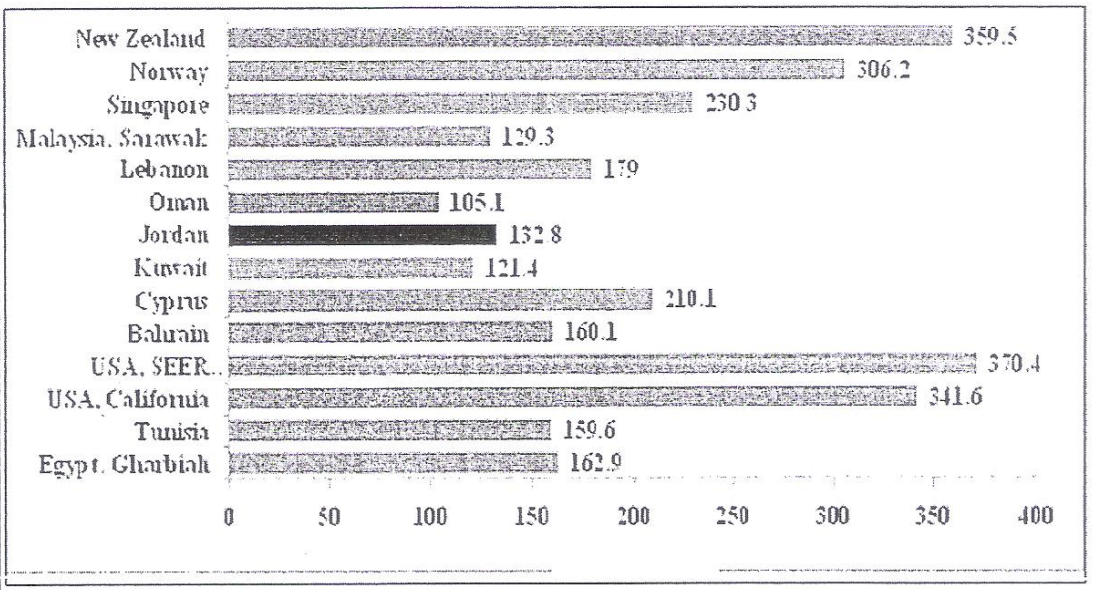





- Cancer increased for both males and females with years.
- From the graph we conclude it is more in females. (so graphs make it easier to conclude)

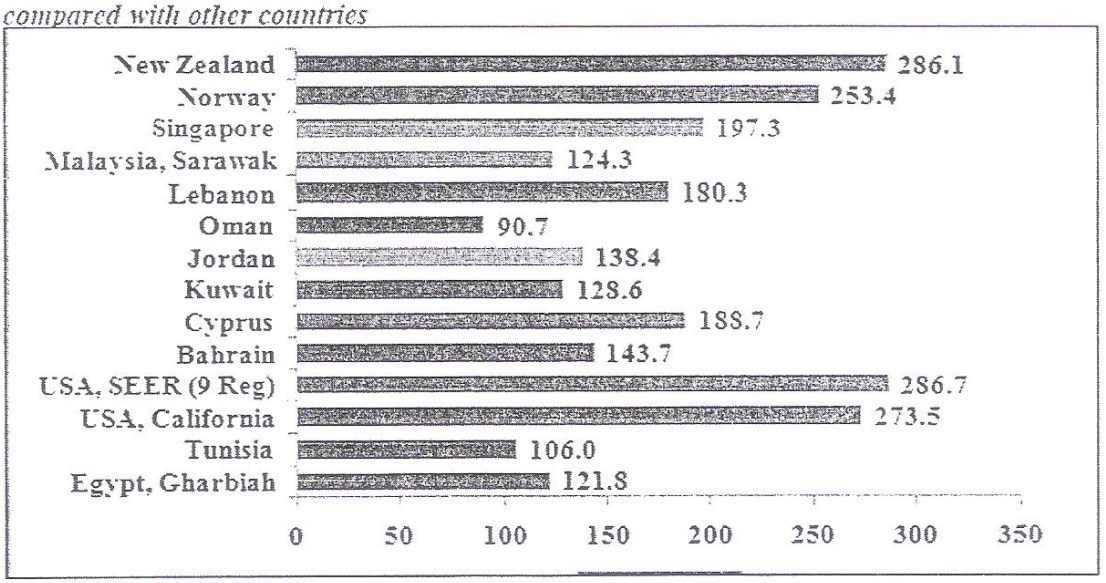
# CANCER RATE

*Figure (10) Age-Standardized Incidence Rate for Male cancers ( all sites) compared with other countries.*




# CANCER RATE

*Figure (11) Age-Standardized Incidence Rate for Female cancers ( all sites) compared with other countries*



**Cancer in Jordan:** Our situation not bad compared to other countries..

- We have lower cancer rate compared to Egypt and New Zealand.
- We have high cancer rate compared to Kuwait. Possible explanations for this maybe:

a) Good health care system in Kuwait.

b) Kuwait is small country so sure will have less cases, and services will be distributed to less people number.

c) A problem in reporting system so less cases were registered in Kuwait.

This is cancer order in 2010:

## Common Cancers

*Table (14) Ten most common cancers among Jordanian Males, 2010-Jordan*

| Rank | Cancer | No | % |
|---|---|---|---|
| 1 | Colorectal | 332 | 14.2 |
| 2 | Lung | 311 | 13.3 |
| 3 | Prostate | 218 | 9.4 |
| 4 | U. Bladder | 186 | 8.0 |
| 5 | Leukemia | 127 | 5.5 |
| 6 | Non-Hodgkin lymphoma | 120 | 5.2 |
| 7 | Brain, Nervous system | 106 | 4.5 |
| 8 | Stomach | 90 | 3.9 |
| 9 | Larynx | 74 | 3.2 |
| 10 | Hodgkin disease | 65 | 2.8 |

*N.B: The total of the top Ten male cancers accounted 1629 ( 70%)*

*Table (15) Ten most common cancers among Jordanian Females , 2010*

| Rank | Cancer | No | % |
|---|---|---|---|
| 1 | Breast | 941 | 37.4 |
| 2 | Colo-rectal | 226 | 9.0 |
| 3 | Thyroid | 136 | 5.4 |
| 4 | Non-Hodgkin lymphoma | 130 | 5.2 |
| 5 | Uterus | 113 | 4.5 |
| 6 | Leukemia | 91 | 3.6 |
| 7 | Ovary | 84 | 3.3 |
| 8 | Lung | 68 | 2.7 |
| 9 | Hodgkin disease | 67 | 2.7 |
| 10 | Stomach | 62 | 2.5 |

*N.B: The total of the top ten female cancers accounted 1918( 76.1%)*

# Definitions

**Population (N):** the largest collection of entities for which there is an interest at a particular time.

- Contains all possible observations.

Example: Jordan.

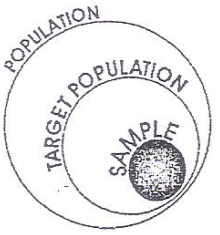


**Target Population:** Large group about which an investigator wishes to draw a conclusion.

- consists of all people (group) that we are interested to study.

Example: Irbid

**Why we can't study whole population?**
It is not practical, costly and time consuming.




**Sample (n):** is simply a part of population.
Should be representative to population.
(We estimate from sample then generalize for population)
**Study Population:** is the population that we can get access to.
**Sample Population:** A subset of individuals that end up participating in your study?

Example:

**Hypothesis: Daily consumption of energy drinks can cause irregular heartbeat among college students**

| | | |
|---|---|---|
| Who do you want to generalize to? | All college students who consume energy drinks in Jordan | Target population |
| What population can you get access to? | College students at JUST who consume energy drinks | Study population |
| Who is in your study? | Students who buy energy drinks from café and willing to participate in the study | Sample population |

**Parameter:** A descriptive measure computed from the data of a population.
**Statistic:** A descriptive measure computed from the data of a sample.

**Descriptive Statistics**: is the methods for organizing, summarizing and presentation of data. (population)
The most common methods of description:

1) Measures of Location (center): Mean, Median, Mode.
2) Measures of spread (Variation): Interquartile Range (IQR), Variance/ Standard Deviation.

Note: These measures will be discussed later on.

**Inferential Statistics:** is part of biostatistics that deals with data analysis and the methods for drawing and reaching decisions about a large body of data based on information obtained from a samples. (sample)

**Important:** Usually our conclusion from sample should be generalized on population but if our sample is not representative or incorrect then take care before generalizing conclusions.

---

**So in brief:**

| **Parameter (denoted by lower case Greek letters such as: μ σ β)** | **Statistic (denoted by lower case Latin letters such as: x, s, b, p, r)** |
|---|---|
| Descriptive statistics | Inferential statistics |
| Less sampling bias | Cheaper, easier, more practical |

---

**Exercises:**

**1) I want to study a chronic disease in Jordan. I live in Irbid. Define target population, study population and sample population.**

Answer:

Target population: is whole Jordanian population. (I want to generalize my results on Jordan)

Study Population: I live in Irbid and have access to Irbid people so my study population is Irbid.

Sample Population: Those who I contact them and agreed to be in my study.

**2) I want to study hypertension in JUST. Define target population, study population and sample population.**

Answer:

Target population: is whole JUST students. (I want to generalize my results on JUST)

Study Population: Suppose I have access to second year medical students so my study population is second year medical students.

Sample Population: I asked your class who wants to participate in the study. Half students agreed to be included so they are my sample population.

**3) Census (التعداد السكاني) is:**

a. Parameter.
b. Statistic

Answer: a

**4) Studies on diseases in Amman (while Jordan is the population) is:**

a. Descriptive statistics.
b. Inferential statistics.

Answer: b (because it is on sample)

**5)** I did census survey. My target population was all live births in Jordan.
I have all birth certificates from 2009 – 2014 and I calculated some parameters and did some analysis like mean and median.
**What type of biostatistics is this?**
a. Descriptive statistics.
b. Inferential statistics.

Answer: a (because we are talking about all population).

**6)** Taking the mean (average) of students in the lecture to study JUST students…
**What is the measurement used here?**
a. Parameter.
b. Statistic

Answer: b (because we studied a sample).

For the same question, **What type of biostatistics is this?**
a. Descriptive statistics.
b. Inferential statistics.

Answer: b (because we studied a sample).

7) Taking the mean (average) of students in Jordan (population)…**What is the measurement used here?**
a. Parameter.
b. Statistic

Answer: a

I want to study blood pressure in females ages 40 – 50 years in Irbid. I measured blood pressure for those visiting King Abdullah hospital, and I calculated the average.(answer questions 8 to 10)
**8) The average is:**
a. Parameter.
b. Statistic.

Answer: b

**9) The study type is:**
a. Descriptive statistics.
b. Inferential statistics.

Answer: b

**10) If I measured blood pressure in Jordan(Jordan is my population), the study type will be:**
a. Descriptive statistics.
b. Inferential statistics.

Answer: a

---

**Variable:** is a specific characteristic that takes different values (varies) in different persons, places, or things.( age, smoking, gender..)

Random Variable: A variable able to assume different values, determined by chance.
Random variables are denoted by capital letters (example: X, Y).

## There are two kinds of variables:

**1) Quantitative variables (numerical values):**
It has two types:

**a.** Discrete Variables: is a variables that can take only countable or integer values (fixed values). For example, the number of family members can take only integer values, say 1, or 2, or 3, ... etc., but we cannot say we have 3.5 brothers.

**b.** Continuous Variables: are the variables that can take any value.
It may include decimals. For example, age, height and weight. And no gap between the readings.

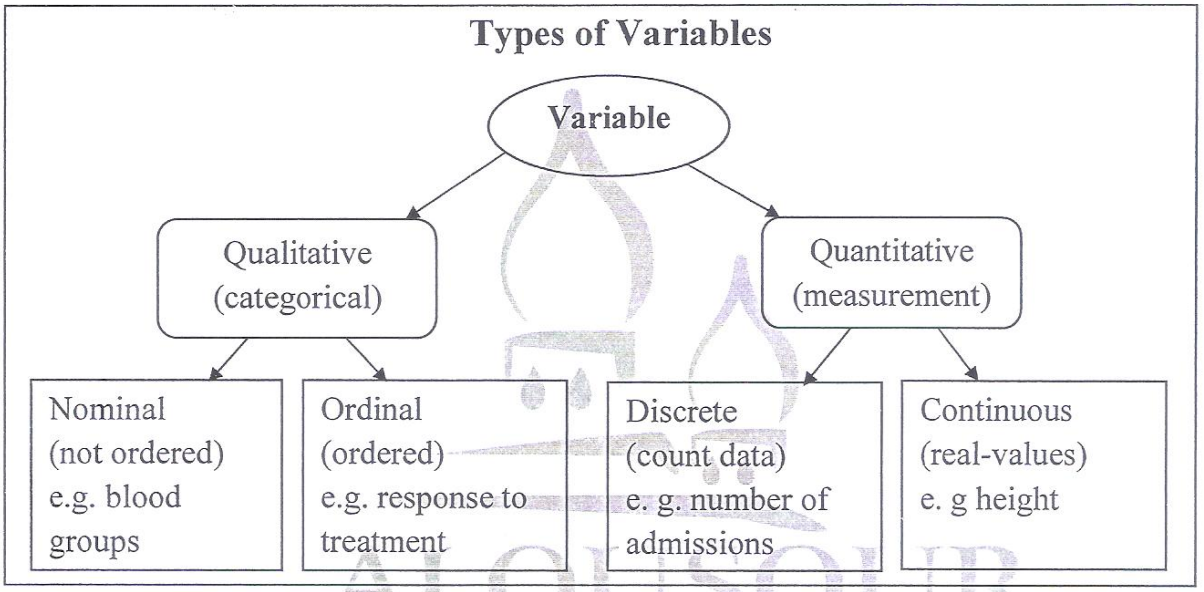
**2) Qualitative Variables (categorical variables):** are the variables that take non-numerical values, such as, blood type or gender.
Categories must not overlap
It has 2 types:

**a**. Nominal : are the variables that can not be ordered or compared by greater or less than. For example: blood type, eye color, gender and marital status.

**b**. Ordinal : same as nominal, but there is some sort of ordering. (differences between categories not considered equal).

Examples:

- students rating: Excellent, very good, good, ...etc..
- Pain severity.
- Levels of satisfaction.




**Example:**

Classify the following variables as discrete or continuous:

1. Number of students in a class room.
2. Time.
3. Number of bacterial count on a plate.
4. Number of missing teeth.
5. PH of a liquid.
6. RBC count.
7. Number of admissions to hospital.
8. A baby head circumference.
9. A student weight.
10. An incubation period of a disease.

**Answers:**
Discrete variables: 1, 3, 4, 6, 7.

---

## Another variables classification:

**1) Dependent Variables (outcome):** are the variables that used to describe or measure the main problem (or disease) under study.

**2) Independent Variables (explanatory):** are the variables that used to describe or measure the factors that are assumed to cause or at least to influence the problem (risk factors).

Example 1: body mass index depends on the weight.
Then body mass index (dependent), weight (independent).

Example 2: differences in weight explain differences in blood pressure, so blood pressure is called an *outcome* variable, while we call weight an *explanatory* variable.

---

## Exercises on variables:

1) **Is your health poor, reasonable, or good. What is the variable type here?**
a. Quantitative.
b. Qualitative.

Answer: b

**2) The above variable is:**
a. discrete.
b. continuous.
c. ordinal.
d. nominal.

Answer: c

**3) Number of cars is:**
a. discrete.
b. continuous.
c. ordinal.
d. nominal.

Answer: a

**4) Faculties names in JUST is considered:**
a. Quantitative.
b. Qualitative.

Answer: b

**5) Faculties names in JUST is:**
a. discrete.
b. continuous.
c. ordinal.
d. nominal.

Answer: d

**6) (Time spent studying) and (test score)... which is the independent and which is the dependent variable?**

Answer:
Scores depends on studying hours so it is dependent.
Studying hours causes (affects) scores so it is independent.

**7) Hair color is :**
. discrete.
b. continuous.
c. ordinal.
d. nominal.

Answer: d

**Why it is important to know the classification of variables?**

Because variables are tabled, summarized, graphed and analyzed in very different ways.

This means statistical tests for continuous variables are different from tests used for nominal data.

---

**Questions:**

1) A qualitative variable:
   a. usually measurements values.
   b. is not numeric.
   c. has only two possible outcomes.
   d. all of the above are correct.

2) A discrete variable is:
   a. an example of a qualitative variable.
   b. usually measurements values.
   c. called categorical variable.
   d. can assume only whole number values.

3) Which of the following is a quantitative variable?
   a. a person's weight.
   b. a person's gender.
   c. fuel efficiency of a car.
   d. whether a person is an University graduate or not.

4) Which one of the following variables is not categorical?
   a. Age of a person.
   b. Gender of a person.
   c. Choice on a test item: true or false.
   d. Marital status of a person .

5) _____ is a set of elements taken from a larger population according to certain rules.
   a. Sample
   b. Population
   c. Statistic
   d. Element

## Answers Key

| Q | Answer |
|---|---|
| 1 | b. is not numeric |
| 2 | d. can assume … |
| 3 | a. a person's weight |
| 4 | a. age of a person. |
| 5 | a. Sample |

---

ALQUSOUR ACADEMY
أكاديمية القصور

### لتقديم الإقتراحات والملاحظات و الشكاوى

* الخط المباشر مع المدير العام : الأستاذ إبراهيم الشواهين الإتصال 0795747445
* في حال عدم الرد إرسال SMS لرقم 0795747445

(الاسم) ----------
(الملاحظة) ---------- لدي ملاحظة ----------

رسالة إرشادية : أعزائي الطلبة هدفنا التفوق معاً . و لترتقي بكم لأعلى الدرجات لابد من إعلامي بأي إقتراح أو ملاحظة أو شكوى في الوقت المناسب و عدم إعلامي بها متأخراً ليتسنى لي حلها و أخذها بعين الإعتبار.

المدير العام
أ. ابراهيم الشواهين



نعتز بثقتكم