ORGANIZATION OF STATISTICAL RESEARCH

Lecturer: ***Ph.D. Khorosh M.V.***
Department of Public Health with MLE

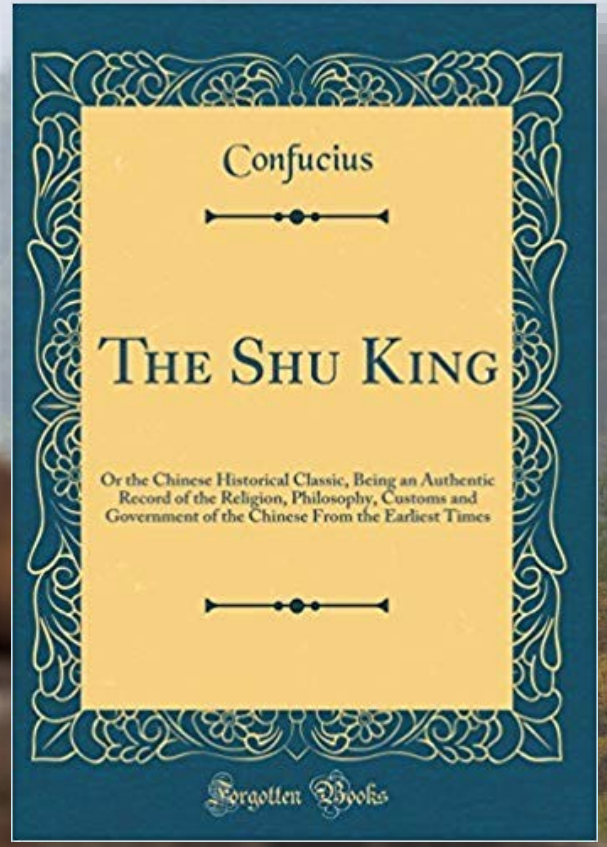
1. History of statistical researches

2. Stages of statistical research

Statistics is a field of knowledge that outlines the general issues of collecting, measuring and analyzing of mass statistical (quantitative or qualitative) data; study of the quantitative aspect of mass social phenomena in numerical form.

The term **“statistics”** was introduced to science by German scientist **Gottfried Achenvall** in 1746, proposing to replace the name of the course “State Studies” to “Statistics”, thus starting the development of statistics as a science and academic discipline.

HISTORY OF STATISTICAL (EPIDEMIOLOGICAL) RESEARCHES

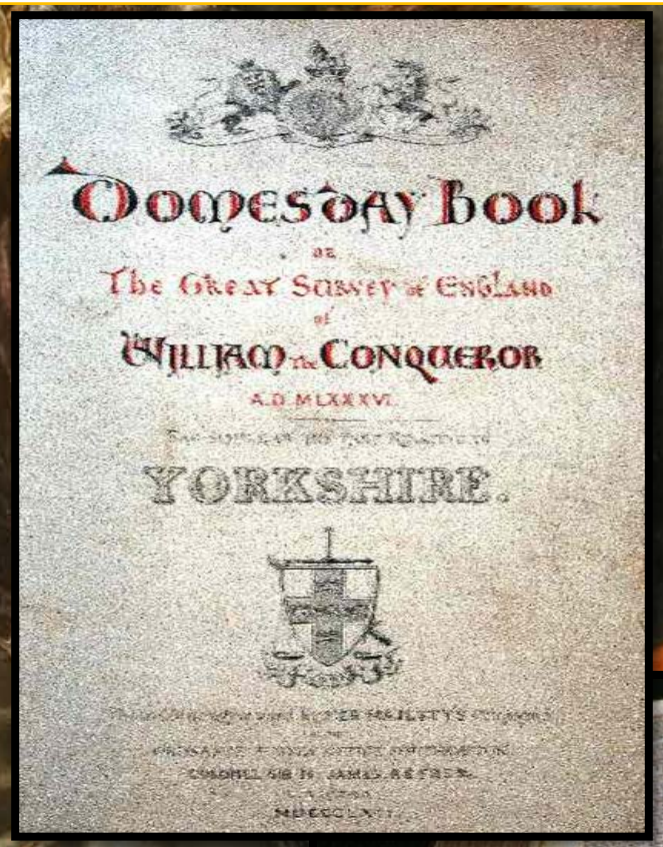


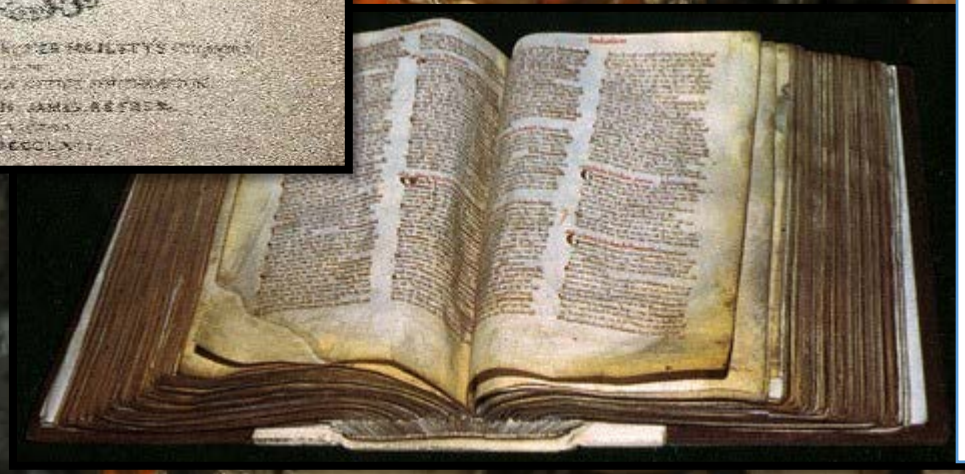
The book "Shu-King" written by Confucius (551-479 BC) mentions a ***census in China in 2238 BC***.

HISTORY OF STATISTICAL (EPIDEMIOLOGICAL) RESEARCHES

In **550 BC in ancient Rome**, a special organization was created — the CENSUS — for conducting censuses of free citizens. Initially, censuses were conducted every 5 years, and later - every 10 years.

HISTORY OF STATISTICAL (EPIDEMIOLOGICAL) RESEARCHES





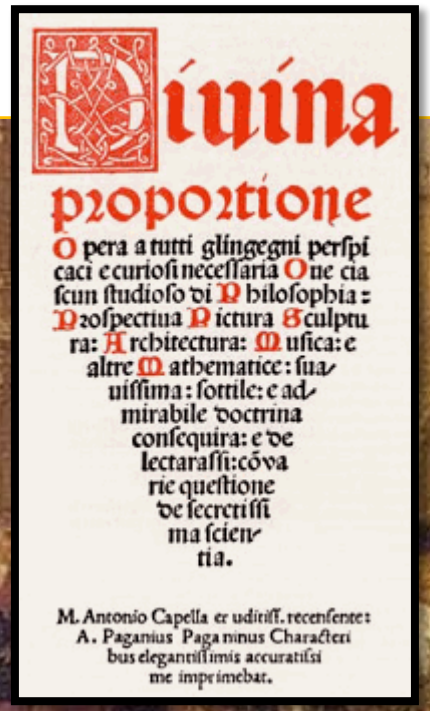
Middle Ages (near 1000 AD).

A unique monument of human activity at that time was the **"Domesday Book"** (Book of the Last Judjement), which is the material of the general land census of all England.

«... instead of using words only in a comparative and superlative degree and resorting to speculative arguments, I embarked on the path of expressing my opinions in terms of numbers, weights, and measures»

William Petty, 1623-1687

HISTORY OF STATISTICAL (EPIDEMIOLOGICAL) RESEARCHES


Divina
proportione
Opera a tutti gl'ingegni perspi
caci e curiosi necessaria Ove cia
scun studioso di Philosophia:
Prospectiva Pictura Sculptu
ra: Architectura: Musica: e
altre Mathematice: sua
vissima: sottile: e ad
mirabile doctrina
consequira: e de
lectarassi: cõva
rie questione
de secretissi
ma scien
tia.
M. Antonio Capella eruditiss. recensente:
A. Paganius Paganinus Characteri
bus elegantissimis accuratissi
me imprimebat.

Renaissance. Characterized by the formation of descriptive statistics.

An important event of the time:

Luca Pacioli (Franciscan monk and mathematician) created an encyclopedic work **"The sum of arithmetic, geometry, the theory of proportions and relations" (1494)**. In the section "Treatise on accounts and records" of this work laid the foundation of accounting.

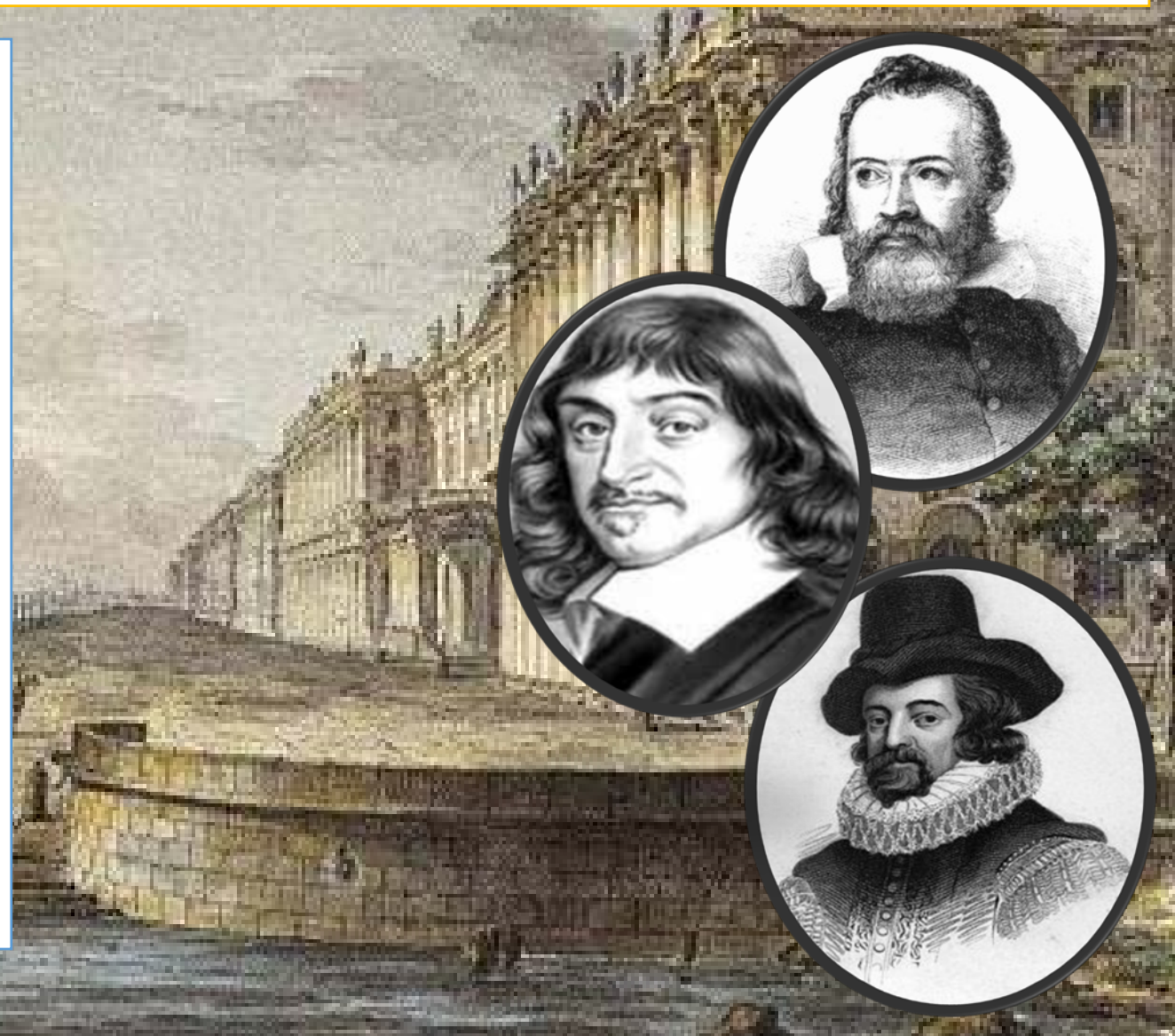
In the 16th – 17th centuries, the development of primary accounting – simple registration of facts and analytical accounting prepared the emergence of statistics.

HISTORY OF STATISTICAL (EPIDEMIOLOGICAL) RESEARCHES

The next stage in the development of statistics refers to the **period of pre-monopoly capitalism** (the end of the 18th-19th centuries).

In the second half of the seventeenth century there was a turn in the history of the development of science.

The works of such scientists as G. Galileo, F. Bacon, R. Descartes, B. Spinoza, I. Newton, and others laid the foundations of modern science.



HISTORY OF STATISTICAL (EPIDEMIOLOGICAL) RESEARCHES

Their work created the prerequisites for the emergence of new techniques:

correlation theory

experiment planning theory

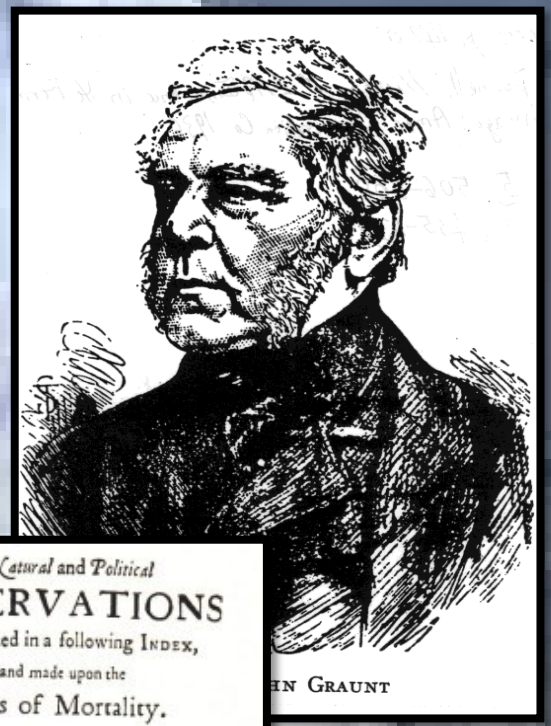
time series theory

theory of component analysis and rank correlation

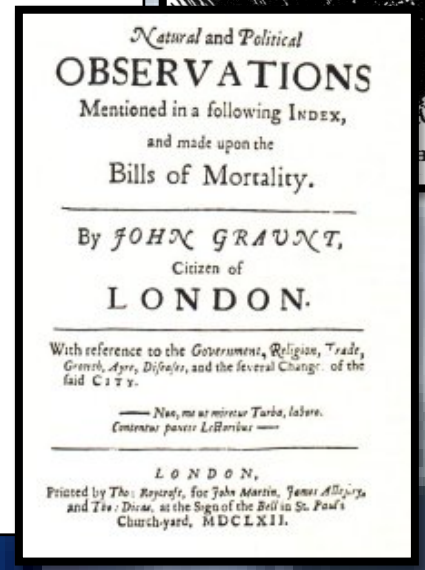
At the end of the nineteenth century statistical method becomes universal.

HISTORY OF STATISTICAL (EPIDEMIOLOGICAL) RESEARCHES - MEDICINE

	Численность нас., т. ч.	Ср.-год. общий прирост, %	Доля городского нас., %
1751	429,9	1,7[1]	—
1801	848,9	1,5[1]	до 5,0
1851	1657,6	0,6[1]	5,8[2]
1870	1769,0	—8,2[3]	7,5
1900	2655,9[4]	1,0[5]	12,5
1920	3147,6[4]	0,65[5]	16,1
1940	3695,6[4]	0,65[5]	26,8
1950	4029,8[4]	0,25[5]	32,3
1970	4598,3[4]	0,1[3]	50,9
1975	4715,8[4]	0,5[3]	58,0
1980	4787,8[4]	0,3[3]	59,8



HN GRAUNT



Natural and *Political*

OBSERVATIONS

Mentioned in a following INDEX,

and made upon the

Bills of Mortality.

By JOHN GRAUNT,

Citizen of

LONDON.

With reference to the *Government*, *Religion*, *Trade*, *Growth*, *Ayre*, *Diseases*, and the several Changes of the said CITY.

Nec, me ut miretur Turba, laboro,
Contentus paucis Lectoribus.

LONDON,

Printed by *Tho: Roycroft*, for *John Martin*, *James Allestry*, and *Tho: Dicas*, at the Sign of the *Bell* in St. *Paul's* Church-yard, MDCLXII.

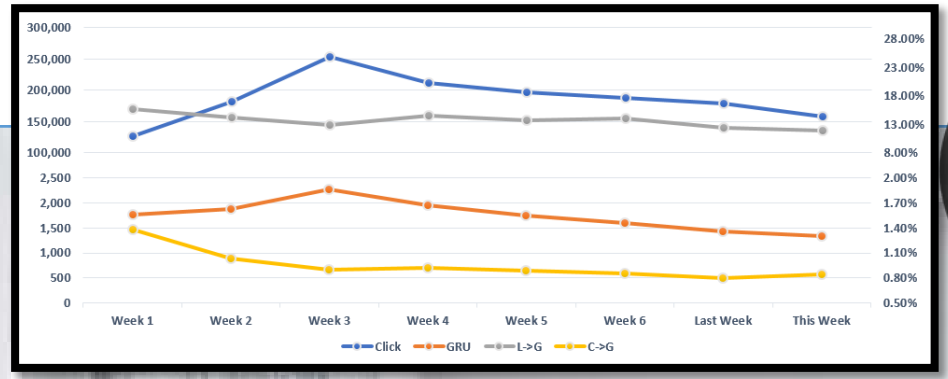
The first attempts to calculate the average life expectancy were undertaken in the **second century AD**.

In the Middle Ages, they were continued in the form of the **“Florentine Table”** (1526).

J. Graunt (1662) based on data on births and deaths in London, he calculated a table of excess life expectancy.

PRACTICE AND STATISTICAL RESEARCH

In practical and research activities, the doctor analyzes the results of his activities not only at the individual level, but also at the group and population levels.



This is necessary for the doctor **to confirm the level of qualification**, as well as for further improvement and professional specialization.

The ability to properly organize and conduct a statistical research is **necessary for all physicians** of different profiles, heads of institutions and health authorities.

Such knowledge and skills contribute to the **improvement of the quality and effectiveness of medical care** for the population through continuous training of personnel (an essential element of resource support) and, thus, the competitiveness of health care institutions of various forms of ownership in a market economy.

PRACTICE AND STATISTICAL RESEARCH

Healthcare managers in operational and prognostic work CONSTANTLY USE statistical data.

!!! ONLY A QUALIFIED ANALYSIS !!!

of statistical data, an assessment of events and relevant conclusions allow us:

- to make the right managerial decision
- contribute to a better organization of work
- more accurate planning and forecasting

PRACTICE AND STATISTICAL RESEARCH

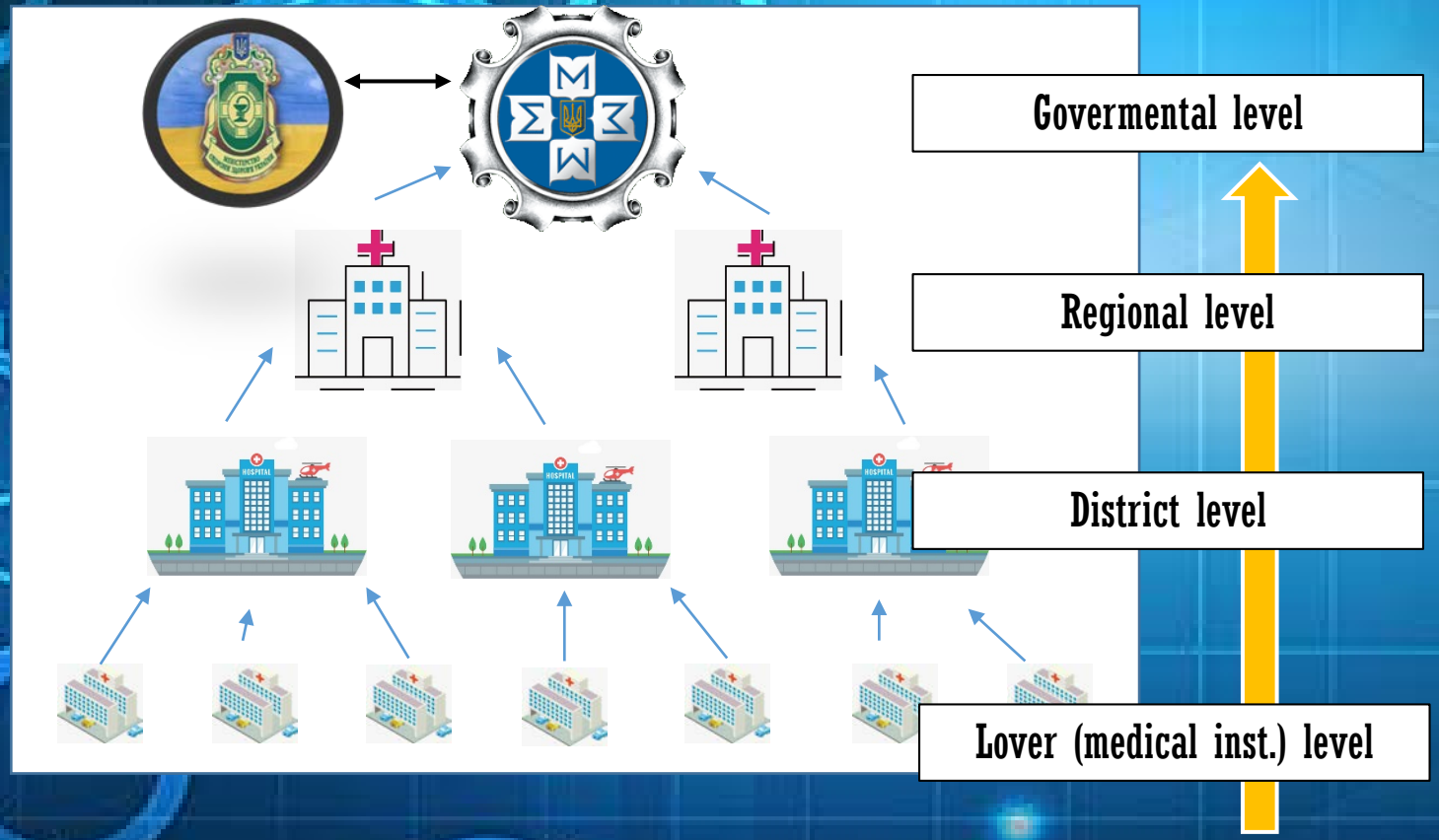
Statistics helps to:

- monitor the activities of the institution
- manage it quickly
- evaluate the quality and effectiveness of treatment and prevention work

The heads of healthcare system in the preparation of current and future work plans should be based on the study and analysis of trends and patterns in the development of both health care and the health status of the population in their district, city, region, etc.

PRACTICE AND STATISTICAL RESEARCH

The traditional statistical system in health care is based on obtaining data in the form of reports that are compiled in lower level institutions and then summed up at intermediate and higher levels.



The reporting system has not only advantages (a single program, ensuring comparability, indicators of workload and use of resources, simplicity and low cost of collecting materials), but also certain disadvantages (low efficiency, rigidity, non-flexible program, limited information, uncontrolled accounting errors, etc.).

PRACTICE AND STATISTICAL RESEARCH

Analysis of the work done should be carried out by doctors, not only on the basis of existing reporting documentation, but also through **specially conducted selective statistical studies**.

IMPORTANT

METHODOLOGY OF STATISTICAL RESEARCH

***Statistics* is a social science that studies the quantitative side of mass social phenomena in close connection with their qualitative features..**

The analysis of quantitative indicators allows to identify the most important patterns of various processes in society and thus learn the qualitative nature of phenomena in their relationship.

Medical (sanitary) statistics examines the state of public health and healthcare system, establishes the reliability of the results of research conducted in clinics and laboratories, the effectiveness of new drugs, methods of diagnosis and prevention, the quality of the doctor’s work.

The following sections of medical statistics are distinguished:

1) public health statistics;
2) healthcare statistics;
3) research statistics, or theoretical medical statistics.

Statistical research is a scientifically organized collection, compilation and analysis of data (facts) about socio-economic, demographic and other phenomena and processes of public life in the state with the registration of their most essential features in the accounting documentation.

THE SPECIFICITY OF THE STATISTICAL STUDY ARE:

- **PURPOSEFULNESS**
- **ORGANIZATION**
- **MASS CHARACTER**
- **CONSISTENCY (COMPLEXITY)**
- **COMPARABILITY**
- **DOCUMENTATION**
- **CONTROLLABILITY**
- **PRACTICALITY**

MAIN STAGES OF STATISTICAL RESEARCH:

1. Preparatory (organizational) stage: drawing up a plan and research program of statistical research

2. Collection of statistical material

3. Developing material and analysis of the data

4. Conclusions, suggestions.

STATISTICAL RESEARCH SHOULD BE:

- Have a socially useful goal and universal (state) significance
- To relate to the subject of statistics in the specific conditions of its place and time
- Express the statistical type of accounting (not bookkeeping and not operational)
- Conducted according to a previously developed program with its scientifically based methodological and other software;
- Collect mass data (facts), which reflect the entire set of causal and other factors that diversify the phenomenon

STATISTICAL RESEARCH SHOULD BE:

- Register in the form of accounting documents of the established sample.
- Guarantee the absence of observation errors or reduce them to the minimum possible.
- To provide for certain quality criteria and methods for monitoring the collected data, ensuring their reliability, completeness and content
- Focus on cost-effective data acquisition and processing technology
- To be a reliable information base for all subsequent stages of statistical research and all users of statistical information.

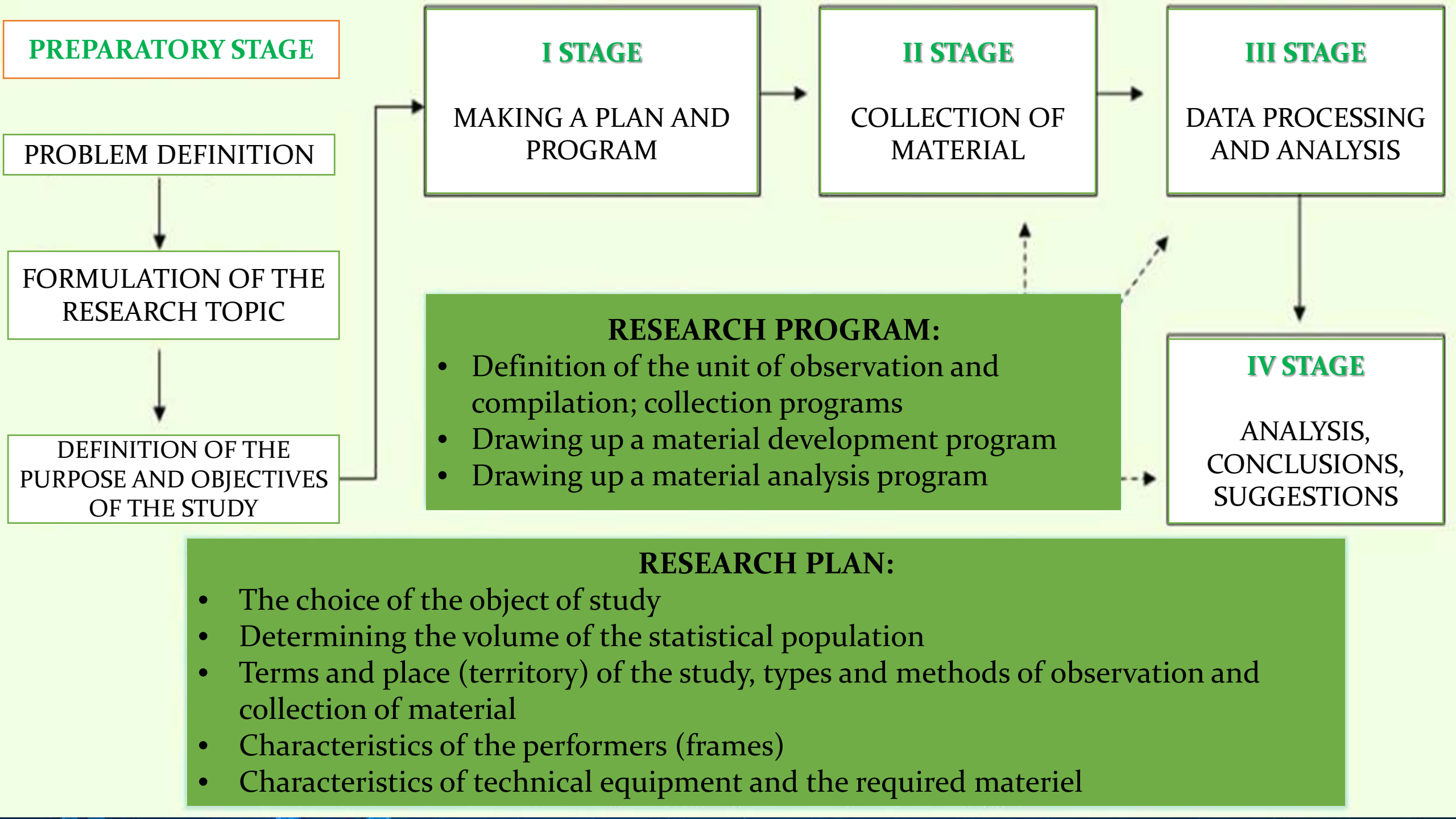
RESEARCH NOT SATISFACTING THESE REQUIREMENTS, ARE NOT STATISTICAL

NON-STATISTICAL STUDIES

- mother for the playing child (personal question);
- the audience for theatrical production (no record for the spectacle);
- scientist for physico-chemical experiments with their measurements, calculations and documentary registration (not mass-public data);
- doctor for patients with keeping medical cards (operational records);
- bookkeeper for the movement of funds in the bank account of the company (bookkeeper);
- journalists for public and personal life of government officials or other celebrities (not the subject of statistics).

Statistical aggregate is a set of units with mass character, typicalness, qualitative homogeneity, and the presence of variation.

The statistical aggregate consists of materially existing objects (Workers, enterprises, countries, regions), is the object of statistical research.

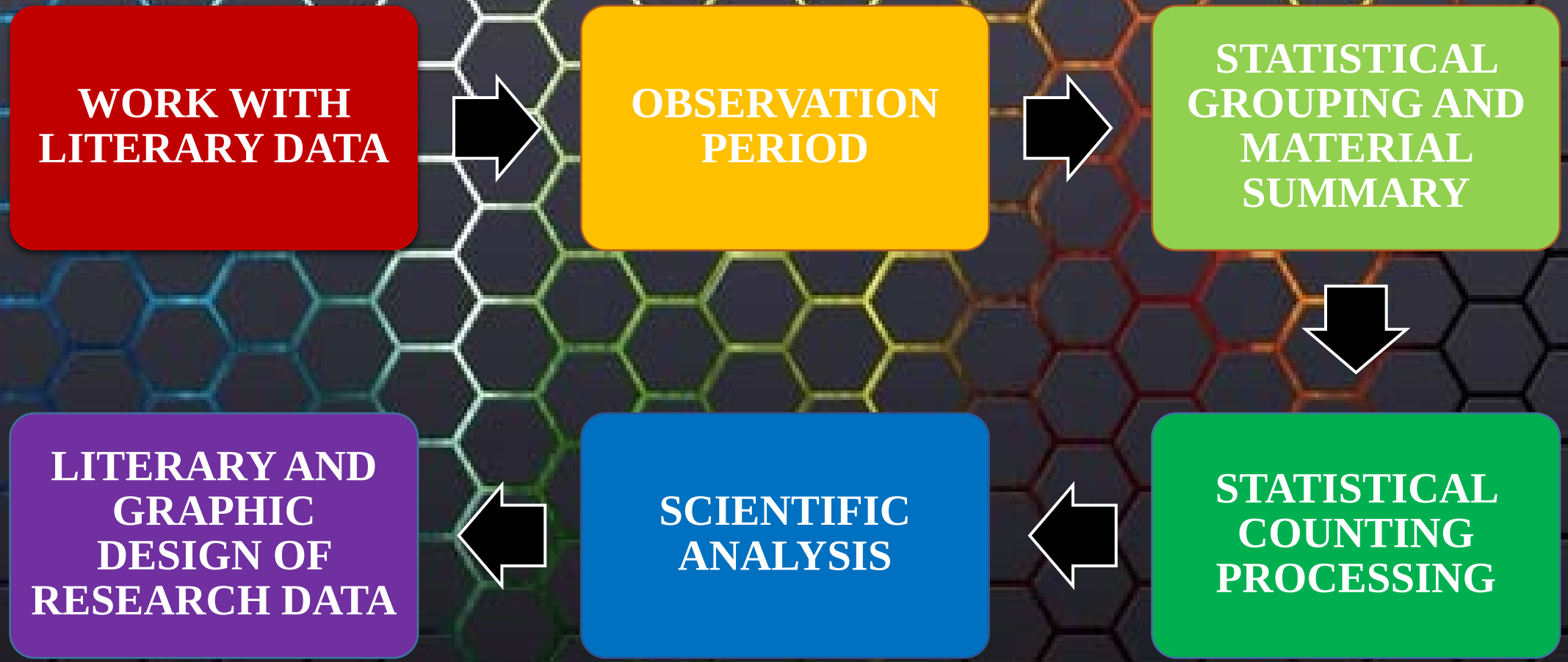

PREPARATORY STAGE
PROBLEM DEFINITION
FORMULATION OF THE RESEARCH TOPIC
DEFINITION OF THE PURPOSE AND OBJECTIVES OF THE STUDY
I STAGE
MAKING A PLAN AND PROGRAM
II STAGE
COLLECTION OF MATERIAL
III STAGE
DATA PROCESSING AND ANALYSIS
IV STAGE
ANALYSIS, CONCLUSIONS, SUGGESTIONS
RESEARCH PROGRAM:
• Definition of the unit of observation and compilation; collection programs
• Drawing up a material development program
• Drawing up a material analysis program
RESEARCH PLAN:
• The choice of the object of study
• Determining the volume of the statistical population
• Terms and place (territory) of the study, types and methods of observation and collection of material
• Characteristics of the performers (frames)
• Characteristics of technical equipment and the required materiel

THE PLAN OF STATISTICAL RESEARCH IS MADE IN ACCORDANCE WITH THE PLANNED PROGRAM.

The main issues of the plan are:

- ✓definition of the purpose of the study
- ✓object definition
- ✓determining the duration of the work at all stages
- ✓indication of the type of statistical observation and method
- ✓determining where the observations will be made
- ✓finding out by what forces and under whose methodological and organizational leadership research will be conducted.

ORGANIZATION OF STATISTICAL RESEARCH IS SHARED ON SEVERAL STEPS:



I stage of statistical research
PROGRAMMING AND RESEARCH PLAN

The program of statistical research provides the solution of the following questions:

- determining the unit of observation and drawing up a material collection program
- drawing up a material development program
- drawing up a program for analyzing the collected material.
- identification of logistical and organizational needs

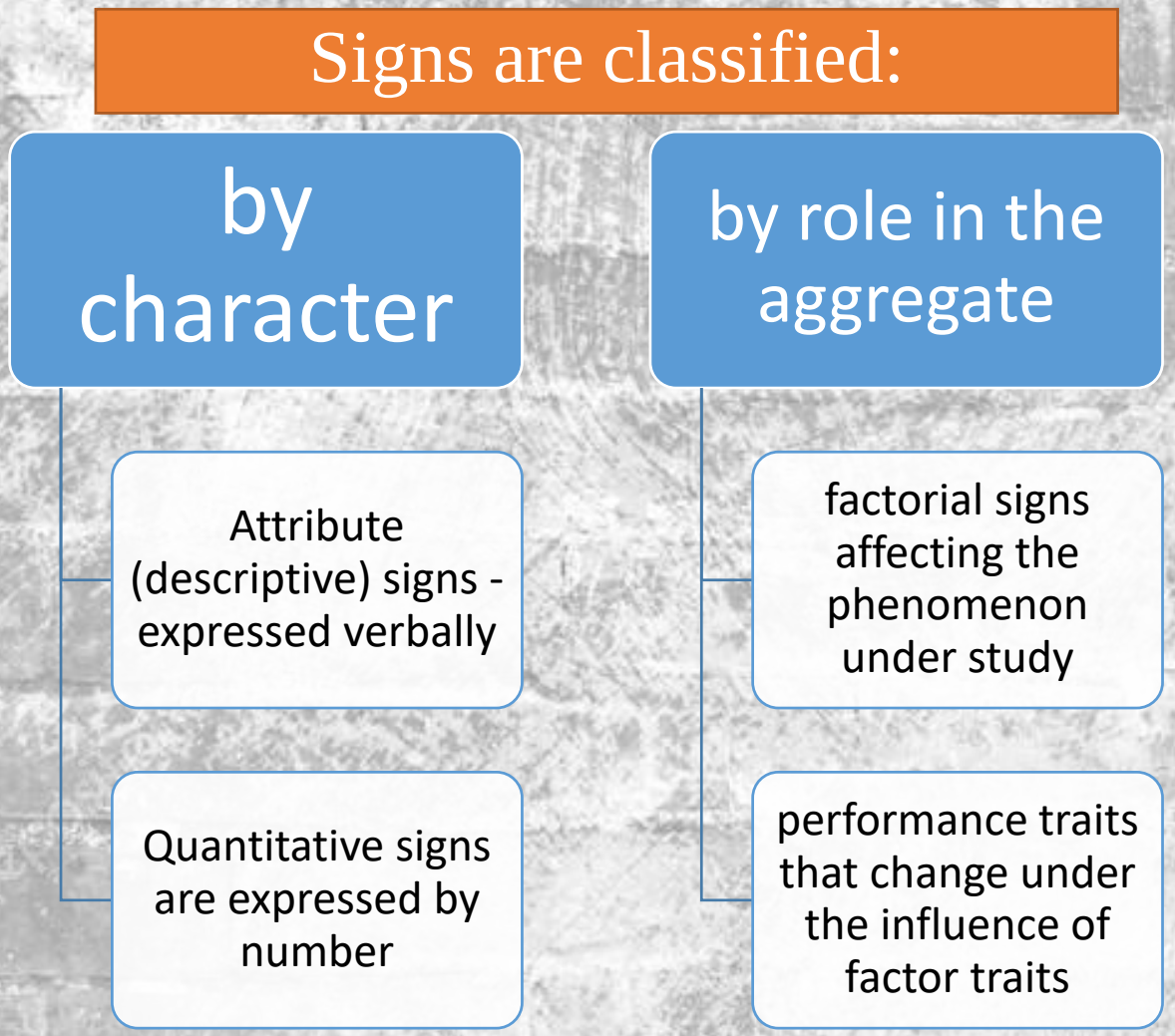
a) Determining the unit of observation and drawing up a material collection program

***The unit of observation** is each primary element of the statistical population.*

The unit of observation is endowed with signs of similarities and differences that are subject to accounting and further observation, therefore, these signs are called counted (accounting).

***Signs taken into account** - signs on which the elements of the observation unit in the statistical population differ.*

a) Determining the unit of observation and drawing up a material collection program



Example: in the study of the prevalence of smoking among students, the observation unit is a student who has been studying in this medical school for all years.

by character

Attribute (descriptive) signs - gender, bad habits, state of health

Quantitative signs are age, number of cigarettes smoked, duration of illness, smoking history

by role in the aggregate

Factorial signs: the presence of bad habits and smoking experience

Performance traits: health status, presence of disease

The material collection program is a consistent presentation of the signs taken into account - questions that need to be answered when conducting this study.

This can be a **questionnaire, questionnaire card or map** specially compiled by the researcher.

!!!The document must have a clear title and approved by experts!!!

Questions (taken into account) should be clear, concise, consistent with the purpose and objectives of the study; for each closed question should provide answers.

The grouping of signs is carried out with the aim of identifying homogeneous groups to study certain patterns of the phenomenon being studied.

Grouping answers by attribute attributes is called **typological**, by quantitative characteristics - **variation**.

Example of a typological grouping:

Group of students by gender:

- man,
- female;

Grouping students by the absence of bad habits:

- smoking students
- non-smokers.

Example of variation grouping:

Grouping students by the number of cigarettes smoked per day:

- 10 or less;
- more than 20


All about ME
NAME
PETS
HOBBIES
favorite CANDY
favorite MOVIE / TV show
FAMILY: how many?
Brothers
Sisters
Are you the oldest, youngest or middle
favorite PLACE
favorite COLOR
favorite MUSIC
your BIRTHDAY
favorite BOOKS to read
favorite VACATION
favorite SUBJECT in school
favorite FOOD

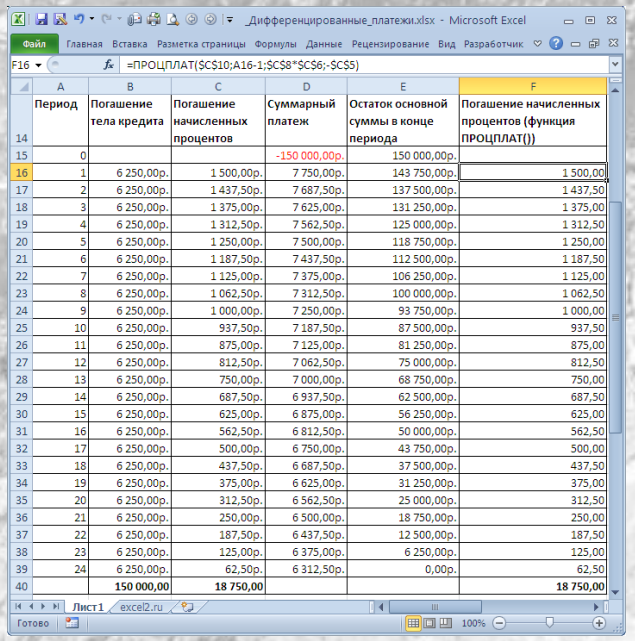

All about ME
NAME
Sam
PETS
dog named Termite
HOBBIES
playing Soccer
taking pictures
LEGOs
favourite CANDY
Snickers
favourite MOVIE / TV show
The LEGO movie
FAMILY: how many?
Brothers 2
Sisters 1
Are you the oldest, youngest or middle
middle
favourite PLACE
my room
favourite COLOUR
blue
favourite MUSIC
Owl City
your BIRTHDAY
April 7th
favourite BOOKS to read
Harry Potter books
favourite VACATION
beach house
favourite SUBJECT in school
Science!
favourite FOOD
pizza

TITLE - Map for studying the prevalence of smoking among medical students

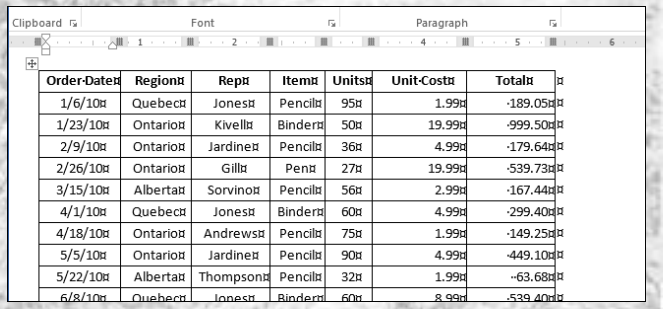
- Full name of the student _____ (write in full)
- Course: I, II, III, IV, V, VI
- Faculty: Medical, Dental, Pharmaceutical,
- Age: up to 20 years, 20, 21, 22, 23, 24, 25 and more
- Gender: Male / Female
- Do you admit that smoking is harmful to health? Yes, No, I do not know
- Who smokes from the people living with you: father, mother, brother, sister, husband, wife, comrade, nobody smokes
- Do you smoke? Yes, No
- Age when you smoked your first cigarette: up to 15 years old, 16-18 years old, over 18 years old
- How many cigarettes (cigarettes) do you smoke per day? 5-10, 11-20, more than 20
- What prompted you to smoke for the first time: an example of parents, an example of teachers, the influence of comrades, the desire to appear adults, the desire to lose weight, curiosity, the desire to keep up with fashion?

b) Drawing up a material development program

The program for the development of the obtained data provides for the compilation of layouts of statistical tables with regard to groupings.



Requirements for tables. Layouts of statistical tables should have a clear and concise title corresponding to their content.



b) Drawing up a material development program

In the table distinguish the subject and predicate.

Statistical subject - is what the table says. The tabular subject contains the main features that are the subject of the study, and is usually placed on the left side of the table vertically.

Statistical predicate - this is what characterizes the subject and is placed horizontally.

	Statistical predicate	
Statistical subject		

TYPES OF TABLES

Статьи и Обороты суммарно по статьям.

<-- Показать/скрыть пояснения щёлкнуть +/-

Сортировать по: Статья + Статья -

ядке возрастания
дке уменьшения

N	Статья	Наличные						Счет 5		Суммарно	
		Приход	Расход	При				ход	Расход	Приход	Расход
1	СемРасх	0.00	-8 000.00					0.00	0.00	0.00	-8 000.
2	Подарки	0.00	0.00					0.00	0.00	0.00	-7 020.
3	КоммунПл	0.00	0.00					0.00	0.00	0.00	-3 000.
4	Реклама	0.00	0.00	0.00	0.00	0.00	-700.00	0.00	0.00	0.00	-700.
5	Другие	0.00	-600.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-600.
6	Детям	0.00	-500.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-500.
7	Книги	0.00	-500.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-500.
8	Продукты	0.00	-435.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-435.
9	Аптека	0.00	-255.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-255.
10	Бензин	0.00	-200						0.00	0.00	-200.
11	Накладные	0.00	0						0.00	0.00	-155.
12	КартНал#	11 000.00	0						0.00	11 000.00	-11 000.
13	Налоги	0.00	0						0.00	0.00	0.
14	Проезд	0.00	0						0.00	0.00	0.
15	РСКарт#	0.00	0						0.00	15 000.00	-15 000.
16	Тел Инт ТВ	0.00	0						0.00	0.00	0.
17	Статья 1	0.00	0						0.00	0.00	0.
18	Статья 2	0.00	0						0.00	0.00	0.
19	Статья 3	0.00	0						0.00	0.00	0.
20	Статья 4	0.00	0						0.00	0.00	0.
21	Статья 5	0.00	0						0.00	0.00	0.
22	Другие Пр	2 500.00	0						0.00	2 500.00	0.
23	Товар1	0.00	0.00	0.00	0.00	7 000.00	0.00	0.00	0.00	7 000.00	0.
24	Товар2	0.00	0.00	0.00	0.00	12 000.00	0.00	0.00	0.00	12 000.00	0.
25	Служебная	СЛУЖЕБНАЯ строка. Не заполнять и не удалять.									
		13 500.00	-10 490.00	15 000.00	-18 020.00	19 000.00	-18 855.00	0.00	0.00	47 500.00	-47 365.

Statistical tables are divided into:

- SIMPLE
- GROUP
- COMBINATIONAL

A simple table is a table that allows you to analyze the data, grouped by only one attribute (subject).

Faculty	Total amount of students	
	Absolut number	in %
1. Medical		
2. Dental		
3. Pharmaceutical		
TOTAL:		100,0

A group table is a table in which a connection is established between individual characteristics, i.e. besides the subject, there is a predicate represented by one or more groupings, which are connected (in pairs) with the groups of the subject, but are not related to each other.

Faculty	SEX		Age when was smoked first cigarette			TOTAL
	M	F	before 15	15 - 18	after 18	
1. Medical						
2. Dental						
3. Pharmaceutical						
TOTAL:						

A combination table is a table in which there are two or more predicates that are associated not only with the subject, but also among themselves (one to one).

Faculty	Average amount of say`s smoked cigarret									TOTAL		
	Less than 10			11 - 20			More than 20					
1. Medical	M	F	M+F	M	F	M+F	M	F	M+F	M	F	M+F
2. Dental												
3. Pharmaceutical												
TOTAL:												
Faculty												

C) Drawing up a program for analyzing the collected material.

The analysis program provides a list of statistical techniques necessary to identify patterns of the phenomenon being studied.

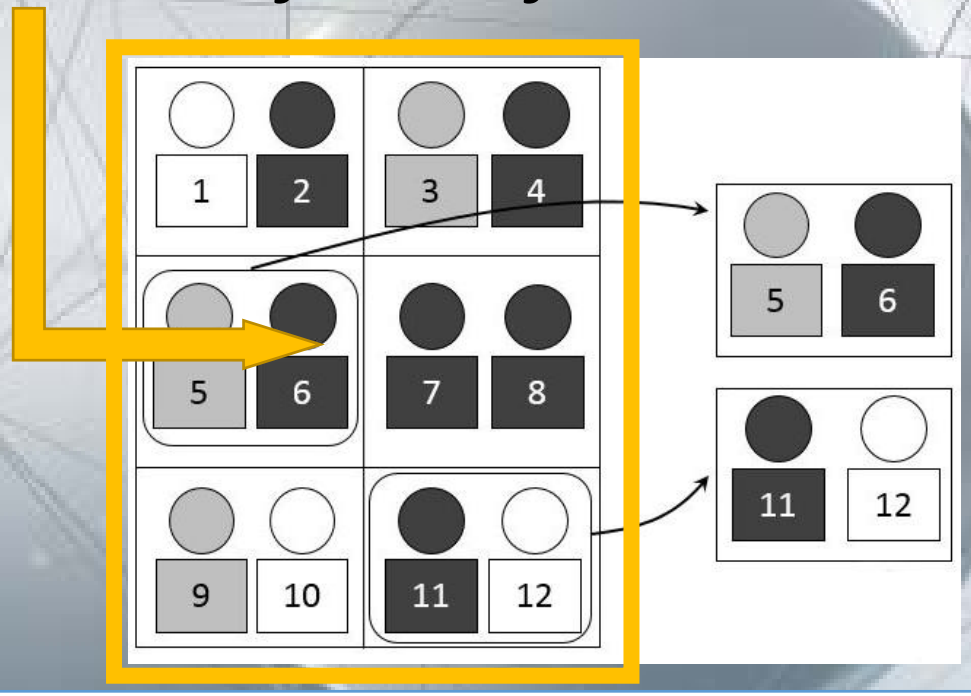
- The use of **generalizing indicators:** absolute, relative, average values and index systems.
- The study of dynamics is usually differential or integral. The fixation of the states of the process forms an integral **dynamic series**, which is investigated on the basis of generalizing analytical indicators, special processing techniques and modeling of the dynamics series.
- Patterns of cause-effect relationships: **correlation and regression analysis**, as well as methods of multivariate statistical analysis.
- The interrelationships of phenomena are also studied using **statistical groupings, comparing parallel series**, building systems of interrelated indices, etc.

The research plan provides for the solution of the following organizational issues:

- The choice of the object of study
- Determining the volume of the statistical population
- Terms and place (territory) of the study, types and methods of observation and collection of material

- Characteristics of the performers (frames)
- Characteristics of technical equipment and the required materiel
- The object of statistical research is the aggregate from which the necessary information will be collected. This may be the population, students, the sick, hospitalized in hospitals, etc.

Statistical aggregate ***is a group consisting of relatively homogeneous elements, taken together within the known boundaries of time and space in accordance with the goal. The structure of the statistical population: the statistical population consists of units of observation.***



ON THE EXAMPLE OF OUR STUDY - statistical aggregate - these are students studying at a given university for the entire period of study.

THERE ARE TWO TYPES OF AGGREGATE - GENERAL AND SELECTIVE.

- ***The general aggregate (population)*** ***is a group consisting of all relatively homogeneous elements in accordance with the goal.***

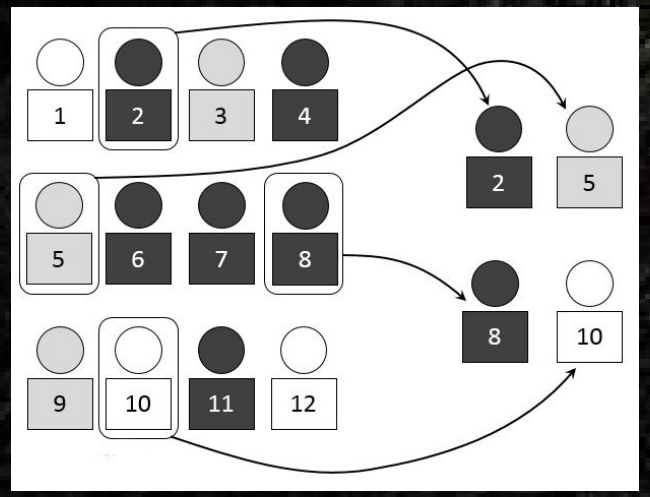
- ***Selective aggregate (set) -*** ***selected for the study part of the general set and designed to characterize the entire general set. It must be representative (representative) in quantity and quality in relation to the general population.***

!!! REPRESENTATIVITY !!!

- ***Quantitative representativeness*** ***based on the law of large numbers and means a sufficient number of elements of the sample, calculated by special formulas and tables.***

$$n = \frac{t^2\sigma^2 N}{t^2\sigma + \Delta^2 N}$$

- ***Qualitative representativeness*** ***is based on the law of probability and means conformity (uniformity) of features characterizing the elements of the sample in relation to the general one.***



In our example, the general population is all medical students; selective aggregate - part of the students of each course and faculty of the university.

- ***The volume of the statistical population** is the number of elements of the population taken for the study.*

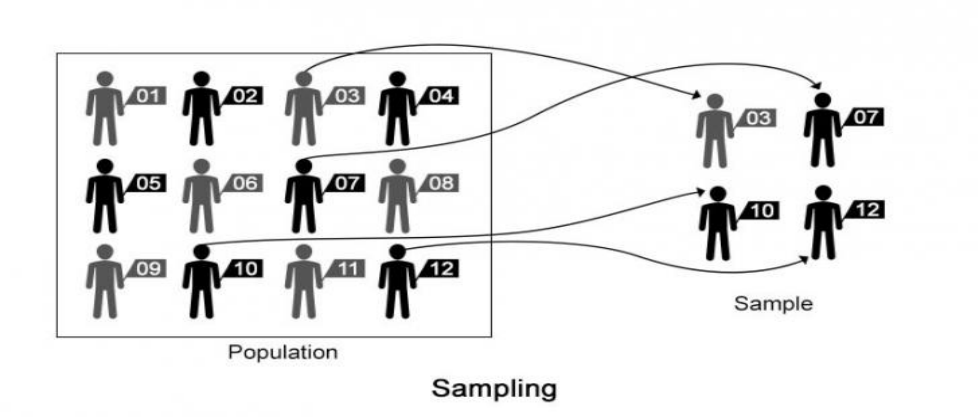
- ***The dates and place (territory) of the study -** is the preparation of a calendar plan for the implementation of this study at this stage in a particular territory.*

- ***Types of observation:***
  - ***current (or constant) observation -** when registration is carried out continuously as the units of observation appear.*
  - ***Example: every case of birth, death, treatment in medical institutions.***
  - ***one-time observation -** when the phenomena being studied are recorded for a specific moment (hour, day of the week, date).*
  - ***Example: population census, the composition of hospital bed capacity.***

WAYS OF CONDUCTING RESEARCH.

For the researcher it is important to determine the method of conducting the study: continuous observation or non-continuous (selective).

- Complete (continuous) observation is the registration of all observation units constituting the entire population.
- Non-continuous (selective) observation - the study of only part of the totality to characterize the whole.



Methods of forming a sample

- **Random selection** is a selection carried out by lot (by the initial letter of the surname or by birthday, etc.).
- **Mechanical selection** is the selection when every fifth (20%) or tenth (10%) observation unit is mechanically selected from the entire population.
- **Serial selection -** when from the general population are selected not individual units, but nests (series), which are selected by random or mechanical sampling. **Example:** to study the incidence of the rural population of the M region, the incidence of the rural population of one, the most typical item is studied. The results apply to the entire rural population of the region.
- **Directional selection** is the selection, when from the general population in order to identify certain patterns are selected only those units of observation, which will reveal the influence of unknown factors while eliminating the influence of known ones. **Example:** when studying the influence of the experience of workers on injuries, workers of one profession, one age, one workshop, one educational level are selected.
- **Typological selection** is the selection of units from previously grouped, same-type qualitative groups. **Example:** when studying the pattern of mortality among the urban population, the studied cities should be grouped according to the population in them.

Characteristics of the performers (frames). *How many people and what qualifications are doing the research. Example: a study on the study of the sanitary and hygienic regime of students in the senior classes of secondary schools of the district is carried out by two doctors and two assistants to the sanitary doctor of the center of hygiene and epidemiology of this administrative district.*

Characteristics of technical equipment and the required material resources:

- *laboratory equipment and devices corresponding to the purpose of the study;*
- *stationery (paper, letterheads);*
- *without additional allocations.*

Stage 2 of statistical research. COLLECTION of MATERIAL

- **It consists in registering individual cases of the phenomenon being studied and characterizing them in the registration forms.**

- ***Control of the collected material*** is the verification of the collected material in order to select records that have defects for their subsequent correction, addition or exclusion from the study. For example, the questionnaire does not indicate gender, age, or no answers to other questions posed.
- ***Encryption*** is the use of symbols for distinguishing features. During manual processing of material, ciphers may be numeric, alphabetic; with machine - only digital.

Stage 3 of statistical research
DEVELOPING MATERIAL AND ANALYSIS OF THE DATA

The responsible stage of the study, at which statistical indicators are calculated (frequency, structure, average size of the phenomenon under study), is given their graphic representation. Dynamics, tendencies are studied, connections between phenomena are established. Predictions are given, etc.

Stage 4 of statistical research
CONCLUSIONS

The forecasts are given, etc. The analysis involves the interpretation of the data obtained, the assessment of the reliability of the research results. In conclusion, conclusions are drawn.

ADDITIONAL STAGES OF STATISTICAL RESEARCH

Stage 5. Literary processing and presentation of the results — involves the final registration of the results of statistical research.

Stage 6. Making a management decision.


THANK YOU FOR
ATTENTION