

STATISTICS

INTRODUCTION, HISTORY

SUBJECT AND TASK OF STATISTICS,

CENTRAL STATISTICAL OFFICE

- **SZTE Mezőgazdasági Kar,
Hódmezővásárhely, Andrásy út 15.**
- GPS coordinates (according to GOOGLE map): 46.414908, 20.323209

Aim of the subject

Name of the subject:	Statistics
Curriculum codes:	<u>EMA15121 lect</u> , <u>EMA151211 pract</u>
Weekly hours (lecture/seminar):	(2 x 45' lectures + 2 x 45' seminars) / week
Semester closing requirements:	Lecture: exam (2 written); seminar: 2 written
Credit:	Lecture: 2; seminar: 1
Suggested semester:	2nd semester
Pre-study requirements:	–
Fields of training:	For foreign students
<p>Objective: Students learn and utilize basic statistical techniques in their engineering work.</p> <p>The course is designed to acquaint students with the basic knowledge of the rules of probability theory and statistical calculations. The course helps students to be able to apply them in practice. The areas to be acquired: data collection, information compressing, comparison, time series analysis and correlation study, reviewing the overall statistical services, land use, crop production, production statistics, price statistics and the current system of structural business statistics.</p>	

Suggested literature:

- Abonyiné Palotás, J., 1999: Általános statisztika alkalmazása a társadalmi-gazdasági földrajzban. (Use of general statistics in socio-economic geography.) JATEPress, Szeged, 123 p.
- Szűcs, I., 2002: Alkalmazott Statisztika. (Applied statistics.) Agroinform Kiadó, Budapest, 551 p.
- Reiczigel J., Harnos, A., Solymosi, N., 2007: Biostatisztika nem statisztikusoknak. (Biostatistics for non-statisticians.) Pars Kft. Nagykovács
- Rappai, G., 2001: Üzleti statisztika Excellel. (Business statistics with excel.) KSH
- Hunyadi, L., Vita L., 2008: Statisztika I. (Statistics I.) Aula Kiadó, Budapest, 348 p.
- Hunyadi, L., Vita, L., 2008: Statisztika II. (Statistics II.) Aula Kiadó, Budapest, 300 p.
- Hunyadi, L., Vita, L., 2008: Statisztikai képletek és táblázatok (oktatási segédlet). [Statistical formulae and tables (tutorial)]. Aula Kiadó, Budapest, 51 p.
- Kerékgyártó, Gy-né, Balogh, I., Sugár, A., Szarvas, B., 2008: Statisztikai módszerek és alkalmazásuk a gazdasági és társadalmi elemzésekben. (Statistical methods and their application in economic and social analyses.) AULA Kiadó, Budapest, 446 p.

A little introduction

- A theoretical mathematician, an applied mathematician and a statistician appeared for a particular job. On the interview they were asked how much is $1 + 1$?
- **Theoretical mathematician:** I can prove that it exists, but not that there is only one solution.
- **Applied Mathematician:** The answer is approximately 2.00, the estimate of the standard deviation is 0.01.
- **Statistician:** What do you want, what should be?
 - <http://www.ilstu.edu/~gcramsey/Gallery.html>

Bonmots on Statistics

"I believe only in the statistics, which I myself falsify." [Winston Churchill](#)

"Statistics are like a bikini: Show a lot of things, but the gist covers." [Aaron Levenstein](#)

"Hungry people can not be fed by statistics." [Lloyd George](#)

"Small lies, big lies and statistics!" [Benjamin Disraeli](#)

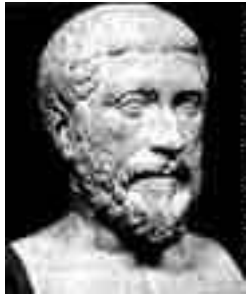
"Statistics do not answer all knowledge."

"Life in truth it is nothing more than statistical error of death."

"A man's death is a tragedy. Millions' death are statistics." [Sztálin](#)

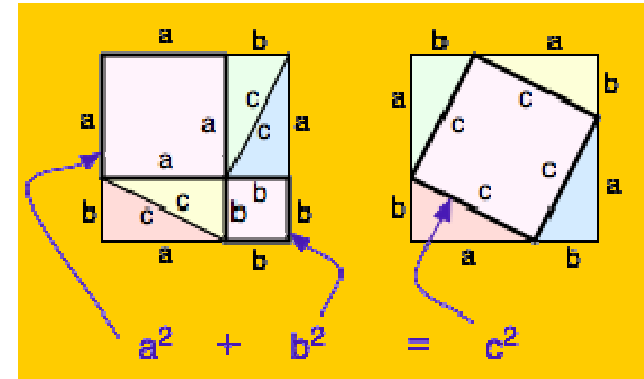
"I am a great friend of statistics; not that I think it just as much prove as many people say, but because since support of every allegation by statistical data has become fashionable, setting up fake items involves a little more difficulty, and advocates of scientific paradoxes provide their stupidity in at least a right system."
[Eötvös József](#)

▪ MATHEMATICS IN ANCIENT TIMES, HELLAS



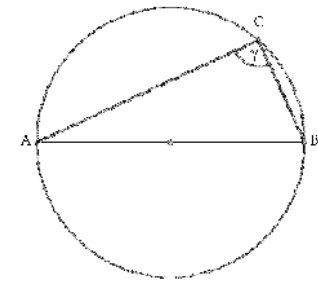
Pythagoras (i.e. 569 – i.e. 475)

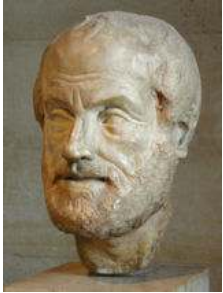
➤ $a^2 + b^2 = c^2$



Thales (i.e. 624 – i.e. 546)

- **Thales theorem:** If the end points A and B of the diameter of a circle are connected with an arbitrary point C of the arc from A to B, then C angle in the triangle ABC will be a right angle.





- Aristotle (i.e. 384 – i.e. 322)

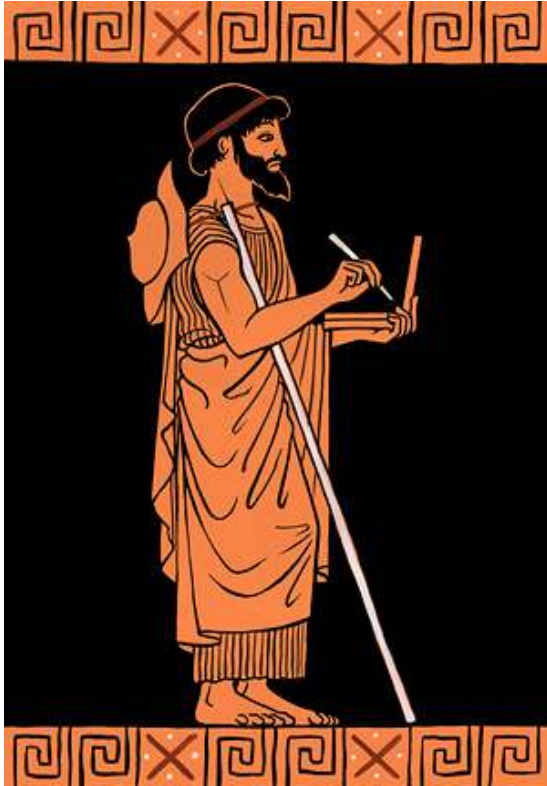
- ✓ **Classification** is an organizing methodology, which dates back to ancient Greek times. Aristotle was the first who formed groups and attempted to comprehend the essence of social subgroups. Having observed that dolphins have placenta, he concluded that they are mammals, not fish. Almost two millennia this conclusion was the subject of ridicule among scientists.



- Plato (i.e. 427 – i.e. 347)

- ✓ „Vile man is who does not know that the diagonal of the square is not comparable to its side.”
(That is, there is no unity, which is an integer multiple of the square diagonal and side too.)
- ✓ „Mathematics is like a game: suitable for young people, not too heavy, entertaining, and it does not pose a threat to the state.”

Right-thinking #1



Herodotus, c.500 BC

“A decision is wise, even if it led to disastrous consequences if it was the best decision based on the information available; and a decision is wrong, even if it is the best possible consequences if it was unreasonable to expect the consequences.”

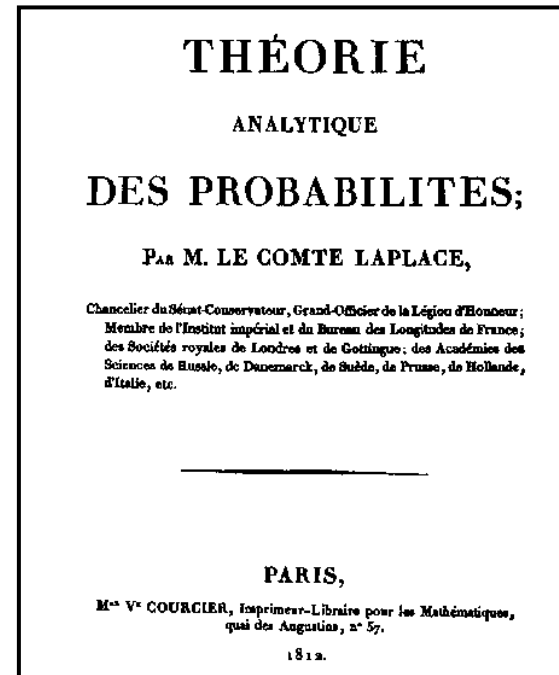
The best should be done with what you have, not what you want.

Right-thinking #2



Pierre-Simon Laplace
(1749 - 1827)

“Probability theory is not more than
common sense - limited for calculation”



Right-thinking #3



William of Occam
(1288 - 1348 AD)

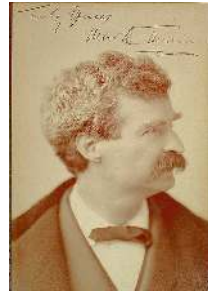
Occam's Razor

“Frustra fit per plura, quod fieri potest per pauciora.”

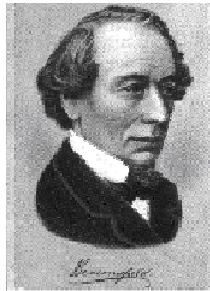
“There is no point to do more for something, if it can be reached with less.”

In order to solve problems, models are highlighted, that are simple.

At first statistics had bad rap...



Mark Twain



Benjamin Disraeli

“There are three kinds of lies:
lies, shameless lies and
statistics”

Mark Twain

This was frequently justified:

The Economist

June 3, 2004

... at University of Girona in Spain, two researchers found that 38% of papers in journal Nature contained one or more statistical errors ... ,

The methods used for data processing is often seen as easy recipes ...

- Correlated 'residues'
- Incorrect assumptions



Systematic errors and/or erroneous results

... owever, in analysis, as in life, the recipes do not always work the way you want.





"The problem is that we (statisticians) call modern statistics, was developed under the strong pressure of biologists."

Jerzy Neyman

Origin of statistics

Originally, State Statistics was developed. (The word "statistics" itself evolved from word "state" meaning in Latin "status".) Statistic, since antiquity, informed the leaders of the states, what taxes to levy their subjects and how many soldiers can expect a coming war.

The statistic has become a real science only after the bourgeois revolutions. Pioneers were John Graunt (1620-1674) and William Petty (1623-1687). In capitalism not only the leaders of the states, but the capitalist entrepreneurs also were interested in statistical surveys, more and more serious mathematical tools were used up for data processing, increasing benefits, such as insurance.

A good insurance is based on accurate assessment and proper mathematical reasoning. Since the 17th centuries mathematical statistics gradually developed into an independent arm of mathematics with the following main aims: to win more reliable information on assessment, monitoring and measurement data: the statistical sample.

Székely J. Gábor „Paradoxes of random mathematics“

Development of statistics (1)

■ Founding fathers

- ❑ Hermann Conring (1606-1681)
forerunner of modern German government statistics
- ❑ Gottfried Achenwall (1719-1772)
German philosopher and statistician
- ❑ Johann Peter Süßmilch (1707-1767)
German priest, statistician and demographer
- ❑ John Graunt (1620-1674)
English, founder of modern demography
1662 - Observations on the Bills of Mortality
- ❑ Sir William Petty (1623-1687)
British economist and demographer

Development of statistics (2)

■ Development of mathematics

- Blaise Pascal (1623-1662)
French mathematician, physicist and philosopher on religion
- Jacob Bernoulli (1654-1705)
Swiss mathematician
- Pierre-Simon de Laplace (1749-1827)
French mathematician and astronomer
- Johann Carl Friedrich Gauss (1777-1855)
German mathematician, physicist and astronomer
- Siméon-Denis Poisson (1781-1840)
French mathematician and physicist
- Pafnutyij Lvovics Csebisev (1821- 1894)
Russian mathematician Russian mathematician
- Andrej Andrejevics Markov (1856-1922)
Russian mathematician

Development of statistics (3)

- Development of computer technology

- Development of the performance of computers
- Artificial intelligence research
- OLAP technology

Online Analytical Processing (OLAP) is an interactive multi-dimensional reporting production application. It is an essential tool of business decision that allows quick access to data and online queries.

- Spread of business management systems (SAP, Baan)
- Spread of Decision Support Systems (DSS)

An interactive information system that relies on user-friendly hardware and software tools to form and display information so that to supports the leadership in the decision-making process.

- Appearance of statistical software packages

SAS, SPSS, Matlab, Statistica, S-plus, R, WinIDAMS, OpenStat, XLStat, etc.

Development of statistical activities

- Middle Ages: religious demographic events
 - 1563: Trident synod
 - 1611: Nagyszombat synod
- Until 17th century: describing statistics
- Since mid-17th century: political arithmetic
- 18-19th century: development of mathematics, probability theory
- 19th century: development of National Statistical Services (civil registration of births)
- 20th century: development of computer technology

Central Statistical Office (CSO)

- **President: Dr. Gabriella Vukovich**
- **Characteristics of CSO:**
 - prestigious organization,
 - It is under the direct supervision of the Government,
 - professionally independent,
 - a nationwide organization.
- **Its main tasks are:**
 - designing surveys,
 - recording data,
 - data processing,
 - storage of data,
 - data analysis and publication,
 - protection of individual data.

Central Statistical Office



Chronology of the development of the CSO

- 1867: start of the operation of the Statistical Service.
- Mid-20th century: developing coordinated economic statistical systems.
- 1968: Application of CMEA and UN recommendations.
- Since 1990: Development of relations with the EU.
- 1992: Geneva conference – statistical principles.

Reporting activities of CSO

When informing, the agency will apply the following principles:

- objectivity, professionalism,
- full disclosure (while protecting individual data),
- simultaneous information.

Law XLVI, 1993

„The task and aim of statistics is providing a realistic and objective picture on the status and evolution of the society, economy, ownership and the environment for the bodies of power and administration, as well as society organizations and members.”

Law XLVI, 1993

„ ... individual data collected for statistical purposes can only be used for other purposes, or else disclose (to pass), if the reporting has consented in writing, except social and fiscal bodies performing public interest tasks: these data can be disclosed without consent.”

Branches of the statistical activity

- population statistics
- economic Statistics
- sector statistics
- corporate and operating statistics
- social statistics
- justice statistics

Development of statistics in Hungary (1)

In the Middle Ages

- ✓ decimal note (ninth, tenth)
- ✓ urbaria (since 1530)
 - regulating mutual rights and obligations of the landlord and the serf. It was also a collection of written regulations, including important information about the management retaining a census among others the serfs livestock, tools, size of the site and its other parameters, as well
- ✓ serfdom census (in 1700s)
- ✓ censuses (in 1800s)

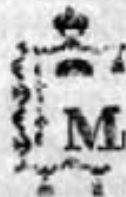
Cover paper of the urbarium, Pozsonypüspöki borough, 1782

URBARIOMA

*Püspöky Mező Városnak
a mellynek,
Az Esztergami Ersek
Földes Ura.*

ELSŐ PUNCTOM

Jobbágy Ház-Helynek Mivoltáról.



§. 1.

Mivel az Jobbágyoknak kötelességet, 's minden Adozást az Jobbágy Helyeknek voltához, és Minőségéhez kölletik szabni, ezek pedig nem mindenőt egyarányofak; hanem az Határoknak Mivoltához, és azoknak egyéb hástón Vételeihez, vagy is Fogynokozáshoz szükséges azokat alkalmatuzozni: azért ezen —
Mező Városnak — egy egyfű Jobbágy Ház-Helynek belső Fundusa, az az Ház-Helye, annak Udvara, szűrés, és veteményes kertve oly nagy, és téres Helyt foglallyon, a' minémű két PofonyiMérő Mag alá kívántaok; a' mennyivel pedig azon belső Fundus kissebbnek találtaok, azat a' költő szűrés földökben, vagy Rétekben kölletik helyre hozni, ellenben a' mennyire két PofonyiMérőnél nagyobb leszen, tehát az a' külső Appertinentiákban szamláltaok; hanem hatsak oly kevéssel köllik halladná, hogy egy ferrály Mérőnél többet nem tenne, az olyan Tekélystge, nem unelő Tekemethen venni.

Development of statistics in Hungary (2)

Major exponents of scientific statistics

Mátyás Bél (1684-1749)

István Hatvani (1718-1786)

Márton Schwartner (1759-1823)

Elek Fényes (1807-1876)

Károly Keleti (1833-1893)

Alfréd Rényi (1921-1969)

Major statistical periodicals

Magyar Statisztikai Évkönyv ([Hungarian Statistical Yearbook](#))

Magyar Statisztikai Zsebkönyv ([Hungarian Statistical Handbook](#))

Gazdaságstatisztikai Évkönyv (1990-től) [[Statistical Yearbook of Economics\(1990 onwards\)](#)]

Élelmiszeripari Statisztikai Zsebkönyv ([Statistical Handbook of Food Industry](#))

Külkereskedelmi Statisztikai Évkönyv ([Statistics Yearbook of Foreign Trade](#))

Magyarország Demográfiai Évkönyve ([Demographic Yearbook of Hungary](#))

Ipari és építőipari statisztikai évkönyv ([Statistical Yearbook of Industry and construction](#))

Mezőgazdasági statisztikai évkönyv ([Statistical Yearbook of Agriculture](#))

Területi statisztikai évkönyv ([Regional Statistical Yearbook](#))

Nemzetközi Statisztikai Évkönyv ([International Statistical Yearbook](#))

[FAO Trade Yearbook](#)

[FAO Production Yearbook](#)

[EUROSTAT publications](#)

www.ksh.hu

Monthly or quarterly

- ✓ latest figures
- ✓ Statistical reports of CSO
- ✓ Monthly Statistical releases
- ✓ Statistical Review
- ✓ Economics and Statistics
- ✓ Regional Statistics
- ✓ Demography
- ✓ methodological papers

Data - Information - Knowledge

- Concept of „data” used by statistics is always a result of experiment, observation, investigation, and often appears as a figure - for the real world
- **data** mean observations or facts that gathered, organized and evaluated become into information and then into **knowledge**.

The concept of statistics

- **practical activities**, for data collecting, organizing, processing and analysis;
- statistical activity is **scientific methodology with modeling purposes**;
- **Information set**, the totality of the data collected and systematized
- The subject of the statistics: **multitude** - elements of the population are the same according to given characteristics, while based on other parameters they are different

Statistics - as a practical activity – is collecting, processing and analyzing information on individuals of mass phenomena, in other words a concise quantitative characterization on the analyzed phenomena.

□ Grouping of statistical methodology I.

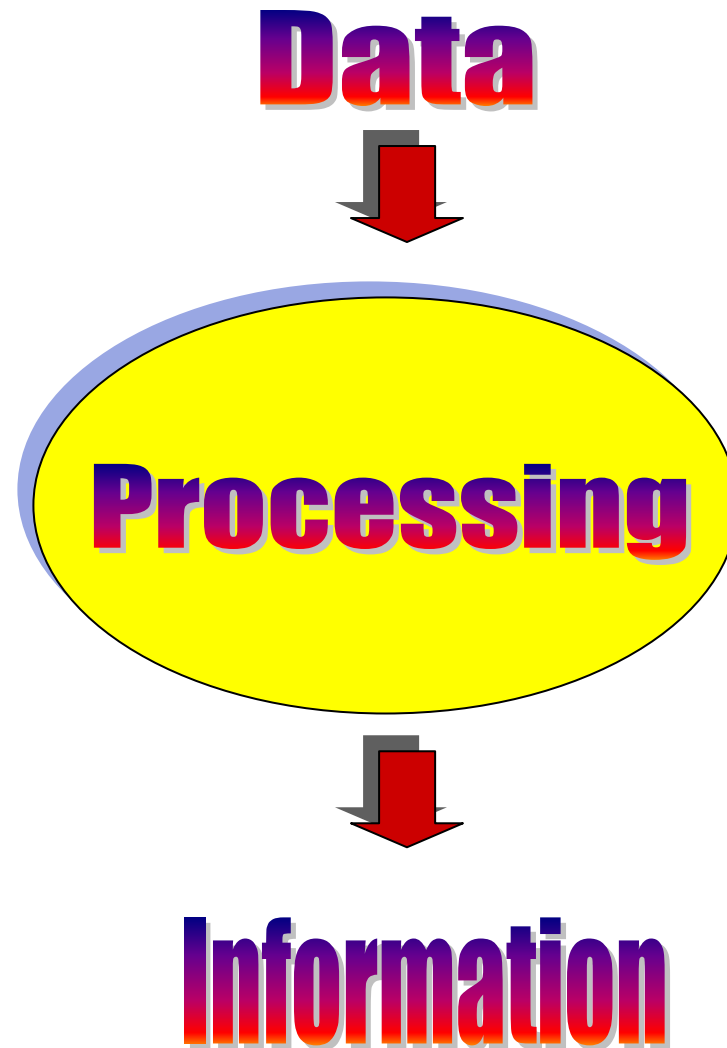
- ***descriptive statistics***: collecting and summarizing information, and methods (data collection, representation, grouping, simple arithmetic operations) of quantification and characterization of information
- ***Statistical inference (mathematical method)***: is used when full observation of individuals of mass phenomena is not possible. In this case, a smaller group is observed and based on the results of its analysis conclude to the mass phenomenon as a whole, and its characteristics.
→ **research statistics**

□ Grouping of statistical methodology II.

- General statistics
- trade statistics (e.g. population statistics).

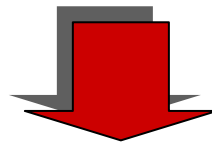
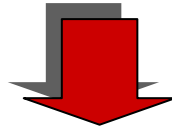
How will be information from data?

information is data organized and processed for a given aim that has been used directly in making particular end use.



How will be knowledge from information?

Information



Knowledge

How will be knowledge from information?

Information



Knowledge

Phases of the statistical work

1. Statistical monitoring (data collection)

1.1. Preparation of the statistical program: setting the task and the plan made to organize the monitoring.

1.1.1. Job setting

1. Statistical observation

1.1.1. Job setting

Records:

- ✓ purpose of the observation,
- ✓ object of the observation,
- ✓ unit of the observation,
- ✓ range of the observation :
 - comprehensive,
 - partial,
 - representative
 - non-representative
- ✓ method of observation
- ✓ frequency of observation

1.1.2. Communication boards, preparation of questionnaires:

According to the observation plan, questionnaires comprise various criteria data (must be accompanied by a duty of implementing instructions and data recording).

1.2. Execution of the survey:

can be occurred by:

- ✓ self-counting
- ✓ The help of writers

1.3. Data checking:

It covers:

- ✓ the completeness of the data (on-site verification of the data collection),
 - ✓ accuracy of the information
 - ✓ mechanical
 - ✓ logic
 - ✓ verification of the statistical system
- } ***control***

2. Data storage and processing (1)

✓ 2.1. Survey:

- ✓ - data bank or database - organized mass data storage and quick access.

✓ 2.2. Media:

- for recording and preserving data,
- questionnaires, statistical publications, microfilms,
- Automatic storage media (floppy disk, magnetic tape, CD, DVD, USB)

2. Data storage and processing (1)

✓ 2.3. Classification system:

classification or grouping is a basic tool of processing and analysis.

✓ 2.3.1. Traditional Classification:

grouping of the population is performed on the basis of a characteristic feature or grouping criteria. Correct choice of the criteria is important. Due to the classification, the main population to be tested is disconnected according to the multitude of different sub-populations.

2. Requirements for statistics

- Revealing reality
- It should be solid and transparent
- Fast report of the results

- fulness
- free from overlapping
- homogeneity

Grouping

- ✓ Easy grouping: according to an individual criterion.
- ✓ Advanced grouping: grouping by two or more criteria.

Advanced grouping

- ✓ **Parallel-group breakdown:**

classification of the same population into different groups according to independent characteristics (e.g. according to active employees, payment categories and age).

- ✓ **Combinative grouping:**

further grouping of data that have already been classified previously by another criterion (e.g. distribution of active employees according to gender and salary scales).

✓ 2.3.2 Automatic classification:

using computer equipments, automatic classification through combinative grouping, according to different characteristics of the population elements (e.g. classification of companies according to production type, text cataloging of library search engines ...).

✓ 2.3.3. Hierarchical classification:

classification is divided into different levels and each level complies with the triple requirements of the classification methodology (completeness, free from overlapping, homogeneity).

plant ⇒ main sector ⇒ industry group ⇒ sector

TEAOR, FEOR, BTO, új SZJ, Customs Tariff.....

The official Hungarian statistical service

management activities
decision making



information

The only information is generally not sufficient to make the correct decision, but many logically interrelated information, information system is needed.

Production, processing, storage and distribution of information plays an important role in economic development and in society.

The official Hungarian statistical service

The statistical information system

operating side:

operations and procedures, which make possible convert and move the components of statistical information in the system for generation, capture, register and store.

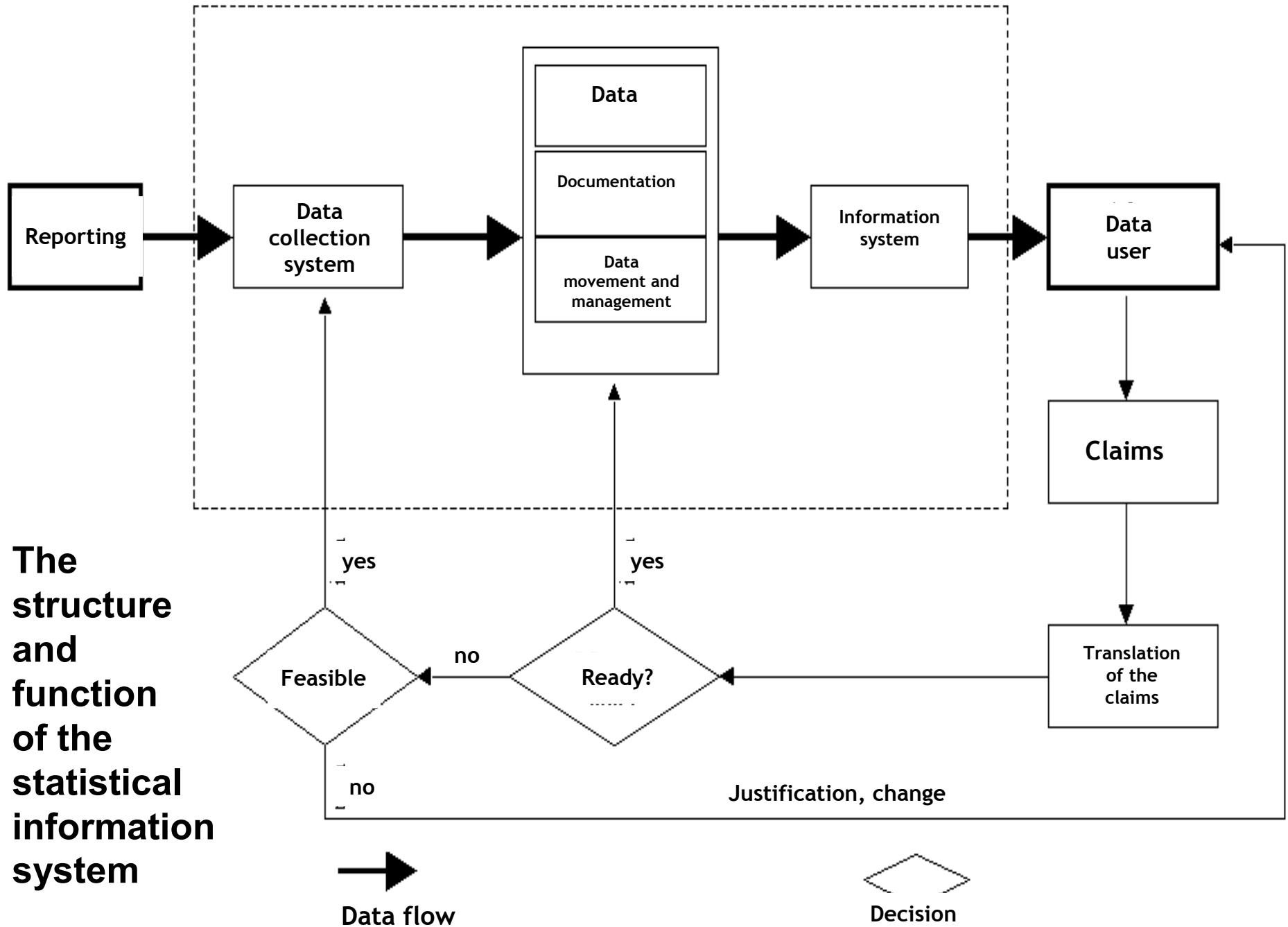
organizational aspects:

organizational unit enabling operation of the information system, or the totality of institutions and their relationships.

The official Hungarian statistical service

- The system includes statistical information gathered in the framework of the official statistical service, which are suitable to take into account the socio-economic processes in a reliable manner, evaluate and provide information about them.
- Consists of:
 - reporting
 - data collection system
 - database
 - information system
 - data user





The structure and function of the statistical information system

The official Hungarian statistical service

Need for statistical information:

government,

other national-level management bodies

This information need is met by a specialized official statistical service in most countries. Its tasks, operational framework and conditions are regulated by separate laws.

Business decision-makers also need information, but their statistical information needs are partly obtained by the companies themselves through surveys carried out for themselves, or partly specialized companies or research institutions (public opinion research, market research firms).



Always look on the bright side
of things!

We finished for today, goodbye!

ямарваа нэг зүйлийн гэгээлэг
талыг нь үргэлж олж харцгаая
өнөөдөртөө ингээд дуусгацгаая, баяртай

让我们总是从光明的一面来看待事物吧！

今天的课程到此结束，谢谢！

دعونا ننظر دائما إلى الجانب المشرق من
الأشياء!

انتهينا لهذا اليوم، وداعا!