

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ СЕМЕНА КУЗНЕЦЯ

"ЗАТВЕРДЖУЮ"
Заступник керівника
(проректор з науково-педагогічної роботи)



М. В. Афанасьєв

ОСНОВИ АНАЛІЗУ ДАНИХ
робоча програма навчальної дисципліни

Галузь знань	12 Інформаційні технології
Спеціальність	122 Комп'ютерні науки
Освітній рівень	другий (магістерський)
Освітня програма	Бізнес-аналітика та інформаційні системи у підприємстві

Вид дисципліни
Мова викладання, навчання та оцінювання

вибіркова
англійська

Завідувач кафедри економічної теорії,
статистики та прогнозування

О.В. Раєвцева

Харків
ХНЕУ ім. С. Кузнеця
2018

APPROVED

at a meeting of the Department of Economic Theory, Statistics and Forecasting
Protocol № 2 on 1.09.2018

Compiled by:

Rayevneva O., doctor of economic sciences, professor of the Department of Economic Theory, Statistics and Forecasting;

Grynevych L., candidate of economic sciences, associate professor of the Department of Economic Theory, Statistics and Forecasting;

Brovko O., candidate of economic sciences, associate professor of the Department of Economic Theory, Statistics and Forecasting.

**Letter of renewal and re-approval
of the academic discipline**

Academic year	Date of the meeting of the department - developer of syllabus	Protocol number	The signature of the head of the department

1. Introduction

Annotation of the discipline:

Data analysis is a science that studies important information and arrays of a large amount of data. It combines statistical methods for analyzing data using a number of IT technologies. The discipline offers a broad approach to data analysis and data searching methods and their application in practice. It combines cutting-edge research and practice in related fields and provides students with the knowledge and ability to initiate and conduct data analysis that can turn company data into commercially valuable information.

Study of the discipline "Fundamentals of data analysis" belongs to the cycle of selective disciplines of educational degree "Master" for students of specialty 122 "Computer Science" of the educational and professional program "Business analytics and information systems in entrepreneurship" and aims at the formation of students with basic competencies about the ability to analyze, to objectively evaluate phenomena and processes, to establish the dependence of the effective indicator on a set of factors, to identify the trends of changes in the phenomena under study and to predict their solution current and interpretation of the survey results and more.

The object of the discipline is social and economic processes and components of social, cultural, economic, information and business systems.

The subject of the discipline is the fundamental statistical methods of data research.

The purpose of the discipline: the formation of theoretical knowledge, applied skills and skills in the organization of statistical observations, methods of statistical analysis and forecasting socio-economic phenomena and processes.

The main tasks of studying the discipline "Fundamentals of data analysis" are:

definition of the principles of observation;

mastering of data processing skills;

studying the methods of constructing indices and their research;

formation of a sample and its analysis from the standpoint of representativeness;

construction of data research indicators and their analysis.

Course	1M	
Semester	1	
Number of ECTS credits	5	
Audit lessons	lectures	20
	laboratory	20
Independent work	110	
Form of final control	Examination	

Structural-logical scheme of studying the discipline:

Previous disciplines	The following disciplines
Statistics	Visualization of data and visual analytics
Computer Science	Statistical thinking for the science of data
Business Economics	Writing course papers, consulting projects

2. Competence and outcomes of studying in a discipline:

Competence	Learning outcomes
The ability to identify, interpret and analyze the needs of stakeholders of economic relations.	The ability to determine the object, subject of the study of socio-economic phenomena and processes. The ability to form the information base of the study.
The ability to apply systemic thinking to understand complex systemic behavior, including interaction between constituent and other systems (social, cultural, legislative, environmental, business, etc.).	The ability to calculate the average and marginal error of selective observation of socio-economic systems under different selection schemes. The ability to determine the sample size based on the use of application packages.
The ability to establish the regularity of the behavior of economic objects on the basis of the formation of homogeneous groups	The ability to conduct primary and secondary grouping of statistical information. The ability to develop layouts of statistical tables. The ability to group statistics using the MS Excel Add-in "Data Analysis".
The ability to identify quantitative and qualitative characteristics of the course of economic phenomena	The ability to perform calculations of average and relative indicators in the study of socio-economic systems. The ability to analyze the intensity of dynamics, identify the presence of trends in dynamic ranks and predict socio-economic phenomena based on the method of extrapolation. The ability to distinguish seasonal characteristics in the studied series of dynamics
The ability to form innovative management decisions regarding the development of economic processes and objects	The ability to construct various forms of expression of indices and assess their interconnection.

3. Program of the academic discipline

Content module 1. Methodological bases of data analysis

Theme 1. Nature of observation of socio-economic objects

1.1. The concept of statistical observation. Formation of information base of statistical observation.

The main requirements of the formation of the information base of the study in accordance with the set task. Ways to obtain information for the research. The essence of statistical observation. Types of observations. Stages of statistical observation.

1.2. The method of analytic grouping.

Methods of study of interconnections between indicators. The essence of the method of analytical grouping of the information space. Stages of analytical grouping. Checking the relevance of the relationship between the surveyed indicators.

1.3. Provision of statistical data: tables, charts, maps.

Types of tables. Rules for constructing tables, charts, maps. Graphic representation of information space and research results. Cartograms and rules for their construction.

Laboratory work on the theme 1. Formation of a representative sample.

Theme 2. Sampling: its construction and verification of its representativeness

2.1. The essence of selective observation, the causes and conditions of its application.

The essence of selective observation. The main tasks of selective observation. Basic concepts of sampling method. Advantages of selective observation.

2.2. Theoretical foundations of selective observation. Types and methods of selecting units in a sample population that ensure the representativeness of the sample.

Classification of selective observation. Methods of selection in a sample collection: repeated, non-repetitive and mechanical selection. Methods of selecting aggregate units in the formation of a sample population. Basic parameters of the general and sample aggregates.

2.3. Classification of sampling errors. The order of determination of sampling errors (representativeness) at different methods of selection.

The essence of errors in the formation of the population. Types of sample errors and methods for calculating them. Methods of checking the quality of statistical evaluation.

2.4. Ways of distributing the results of selective observation to the general population.

Methods of propagation of selective observation to the general population. The essence of direct reckoning. The essence of the method of coefficients.

2.5. Determination of the required sample size.

The main factors affecting the sample size. Methods and methods for calculating the required sample size.

Laboratory task on the theme 2. Tools for visualization of statistical data

Theme 3. Grouping data and studying groups

3.1. Grouping as the basis of the science of data

The essence of grouping data in solving the problem. Basic requirements for data grouping. Types of grouping in the scientific processing of data. The essence of classifications of groupings. The main tasks facing the grouping. Stages of grouping.

3.2. Methods of visualization of grouped data

Visualization, what is the essence of understanding data. Functions of visualization of data in the study of various systems (economic, informational, legal, demographic). Data visualization tasks. Classification of existing data visualization methods. Ways of applying visualization methods. Graphical methods of express analysis. Visualization methods for complex systems. Instruments and techniques for rendering grouped data.

Laboratory work on the topic. 3. Grouping of data and research of groups. 3.1. Grouping as the basis of the science of data

The essence of grouping data in solving the problem. Basic requirements for data grouping. Types of grouping in the scientific processing of data. The essence of classifications of groupings. The main tasks facing the grouping. Stages of grouping.

3.2. Methods of visualization of grouped data

Visualization, what is the essence of understanding data. Functions of visualization of data in the study of various systems (economic, informational, legal, demographic). Data visualization tasks. Classification of existing data visualization methods. Ways of

applying visualization methods. Graphical methods of express analysis. Visualization methods for complex systems. Instruments and techniques for rendering grouped data. Laboratory work on the theme. 3. Grouping of data and research of groups.

Content module 2. Methods of research and analysis of the components of socio-economic systems

Theme 4. Indicators of data research

4.1. Essence of research indicators.

The essence of the indicators of the study of socio-economic phenomena. Classification of existing research indicators. The main functions and objectives of the indicators. Methodological bases for the use of data research indicators.

4.2. Types and forms of indicators (indicators) when analyzing data.

Characteristics of types of indicators. The essence and importance of generalizing indicators of research data of different systems. Classification of research indicators by the method of their use. The main groups of indicators of complex analysis in the study of socio-economic processes.

4.3. Methods and methods of calculating the indicators of research.

Characteristics and nature of the indicators of the study by the method of calculation, by the nature of the phenomena under investigation, by the way of reflection and by the degree of aggregation. Terms of use of relative and absolute indicators.

Laboratory work on the theme 4. Statistical tools for studying the dynamics of phenomena and processes.

Theme 5. World indexes and their research

5.1. Indices and their classification.

The essence of the index analysis. Main tasks of index analysis. Classification of world indexes. Methodological principles of constructing indices. Methods of calculating the main economic indexes.

5.2. General Indices of Quantitative Indicators.

The essence of general and individual quantitative indices. Basic requirements for the construction of the index of quantitative indicators. The general rule of constructing the index of quantitative indicators. Stages of constructing quantitative indices.

5.3. General indexes of quality indicators.

Basic requirements for the construction of the index of quality indicators. List of requirements for the construction of the index of qualitative indicators. Methodology for determining the indices of qualitative indicators and their economic essence.

5.4. Two-factor system of interconnected indices.

The essence of the system of interconnected indices. Determination of the influence of individual factors. The essence of the two-factor system of the index. Formation of analytical conclusions on the received calculations of index analysis.

5.5 Index method of analysis of the middle-level dynamics.

Qualitative indicators that characterize economic phenomena. Classification of the indices of dynamics of averages. Method of calculating the index of constant, variable and structural changes. The existence of interconnection between the middle-level dynamics indexes.

5.6. Chain and Basic Indices.

Index system to characterize the dynamics of a complex phenomenon. Distribution on chain and basis indices depending on the comparison base. Characteristics of a system of indices with a variable base of comparison. Characteristics of a system of

indices with a permanent base.

Laboratory work on the theme 5. Indicators of the study of world trends.

4. The order of evaluation of the results of training

The system of evaluation of the developed competencies of students takes into account the types of occupations, which, according to the curriculum program, include lectures, laboratory classes, and independent work.

Assessment of the developed competencies in students is carried out using a 100-point accumulation system. In accordance with the Provisional Regulations "On the Procedure for Assessing the Results of Students' Learning Based on the Accumulated Bulletin-Rating System" KhNEU them. S. Kuznets, control measures include:

current control during the semester during lectures and laboratory classes and is estimated by the sum of the points scored (the maximum amount is 60 points; the minimum amount that allows the student to take the exam - 35 points);

final / semester control, conducted in the form of a semester exam, according to the schedule of the educational process.

The system for assessing students' knowledge, skills and abilities requires rating of all forms of conducting classes. Assessment of student's knowledge is carried out on a 100-point system of accumulation.

Final control is carried out on the basis of the exam. Examination ticket consists of solving five settlement and analytical tasks (2 stereotypical, 2 diagnostic, 1 heuristic).

Criteria for evaluating a stereotyped task:

1 point - solving the problem contains only calculation formulas that do not correspond to the situation being considered;

2 points - the task was solved incorrectly, but contains elements of the decision concerning the situation;

3 points - the answer contains only correctly written calculation formula;

4 points - the answer contains only correctly written calculation formula and explanation to it;

5 points - the answer contains only correctly written calculation formula, explanation to it and justification of use of this formula;

6 points - the answer contains only correctly written calculation formula, an explanation to it, justification of the use of this formula and substitution into the formula;

7 points - the answer contains only correctly written calculation formula, an explanation to it, justification of the use of this formula, substitution into the formula and correct calculation with indication of units of measurement of the result;

8 points - the answer contains only correctly written calculation formula, an explanation to it, justification of using this particular formula, substitution into the formula, correct calculation and conclusion based on the calculation result.

Criteria for evaluating diagnostic tasks:

1 point - solving the problem contains only calculation formulas that do not correspond to the situation being considered;

2 points - solving the problem contains only the calculation formulas or some elements from the separate discipline topics relating to the situation;

3 points - solving the problem contains calculation formulas and explanations to them, the decision has only begun;

4 points - the problem is solved incorrectly, but some calculations correspond to the

situation under consideration;

5 points - the problem is solved completely, but there is a gross error that affects the final result of the calculations, the conclusions are of a superficial nature, not all formulas are explained;

6 points - the task is done correctly in half, the expediency of using analytical tools in the analysis of the proposed situation, the conclusions are superficial;

7 points - the task is done correctly in full, qualitatively executed, the expediency of using analytical tools in the analysis of the proposed situation, but based on the results of calculations made not complete analytical findings and there is inaccuracy in the calculations;

8 points - the task is completed in full, qualitatively executed, the expediency of using analytical tools in the analysis of the proposed situation, but based on the results of calculations are not complete analytical conclusions;

9 points - the task was executed impeccably, without any error, qualitatively executed, the expediency of using analytical tools in the analysis of the proposed situation, based on the results of calculations made reasoned analytical conclusions and generalizations.

Criteria for assessing the heuristic problem:

1 point - the problem solving contains only the formulas of calculations or some elements from separate subjects of discipline;

2 points - the problem is solved incorrectly, but some calculations correspond to the situation under consideration;

3 points - the problem is solved methodically correctly half, the conclusions are superficial;

4 points - the problem is solved methodically correctly, but there are minor mistakes that do not affect the final result, the conclusions are not sufficiently reasoned; or the task is solved correctly, without errors, but there are no complete conclusions and explanations for the calculations;

5 points - the problem is solved methodically correctly, all necessary analytical tools are used in the decision, explanations for the calculations, reasoned conclusions are made, but some inaccuracies in the reasoning of the answer or conclusions are made;

6 points - the task was executed impeccably, qualitatively executed, the expediency of using one or another analytical toolkit in the analysis of the proposed situation, based on the results of calculations made an argumentated analytical conclusion and generalization.

The result of the semester exam is evaluated in points (the maximum number is 40 points, the minimum number is scored - 25 points) and is placed in the corresponding column of the examination "Accountancy record of success".

The student should be **considered certified if the sum of the points earned on the results of the final / semester test of success is** equal to or exceeds 60. The minimum number of points for the current and modular control during the semester is 35 and the minimum number of points scored on the exam is 25.

Table 4.1

Distribution of points for weeks

Themes of the content module			Lectures	Laboratory classes	Test tasks	Control work	Total
Content module 1. Methodological bases of data analysis	Theme 1. Nature of observation of socio-economic objects	1 week	1	-	-	-	1
		2 week	-	-	-	-	-
	Theme 2. Sampling: its construction and verification of its representativeness	3 week	1	2	-	-	3
		4 week	-	-	1	-	1
		5 week	1	4	-	-	5
		6 week	-	-	-	-	-
	Theme 3. Grouping data and studying groups	7 week	1	4	-	-	5
		8 week	-	-	1	-	1
		9 week	1	-	-	10	11
Content module 2. Methods of research and analysis of the components of socio-economic systems	Theme 4. Indicators of data research	10 week	-	-	1	-	1
		11 week	1	4	-	-	5
		12 week	-	-	1	-	1
		13 week	1	4	-	-	5
	Theme 5. World indexes and their research	14 week	-	-	1	-	1
		15 week	1	4	-	-	5
		16 week 17 week	2	3	-	10	15
Examination							40
Total			10	25	5	20	100

Grating scale: national and ECTS

The amount of points for all types of educational activities	Rating ECTS	Score on a national scale	
		for exam , course project (work), practice	for the offset
90 – 100	A	excellent	enrolled
82 – 89	B	good	
74 – 81	C		
64 – 73	D	satisfactory	
60 – 63	E		
35 – 59	FX	unsatisfactory	unenrolled
1 – 34	F		

5. References

5.1. Main

1. Андрієнко В.Ю. Статистичні індекси в економічних дослідженнях / В. Ю. Андрієнко. – К. : 2004. – 536 с.
2. Ефимова М.Р., Петрова Е.В., Румянцева В.Н. Общая теория статистики: Учебник. – 2-е изд., испр. И доп. – М.: ИНФРА-М, 2007. – 417 с.
3. Лугін О.Є. Статистика. Підручник. 2-е видання, перероблене та доповнене. – К.: Центр учбової літератури, 2007. – 608 с.
4. Сизова Т.М. Статистика. Учебное пособие. – Изд-во «ИТМО», Санкт-Петербург, 2005. – 190 с.
5. Статистика: Підручник / С. С. Герасименко, А. В. Головач, А. М. Єріна та ін.; За наук. ред. д-ра екон. наук С. С. Герасименка. — 2-ге вид., перероб. і доп. — К.: КНЕУ, 2000. — 467 с.
6. Статистика. Навчальний посібник / Під ред. д.е.н., професора Раєвнєвої О.В. – Харків: Вид. ХНЕУ, 2010. – 520 с.
7. Статистическое моделирование и прогнозирование / Учеб. пособие / Под ред. А.Г. Гранберга. – М.: Финансы и статистика, 2000, – 384 с.
8. Marcelo Fernandes. Statistics for Business and Economic. 2009. - P.150 - http://russell-davidson.arts.mcgill.ca/e257/Marcelo_Fernandes.pdf

5.2. Additional

9. Дягтерев Д.А. Количественные методы в международных исследованиях. Международные процессы, Том 13, № 2, с. 35-54
10. Єріна А. М. Статистичне моделювання та прогнозування: навч. посіб. / А. М. Єріна. – Київ : КНЕУ, 2001. – 170 с.
11. Халафян А.А. STATISTICA 6. Статистический анализ данных. – М.: ООО «Бином-Пресс», 2008. – 512с.
12. G. Scott Owen. History of Visualization. – 2010 –P. 97

5.3. Internet resources

13. Офіційний сайт Всесвітнього Банку [Електронний ресурс]. – Режим доступу: <http://www.worldbank.org/>
14. Офіційний сайт Міждержавного статистичного комітету об'єднання незалежних держав [Електронний ресурс]. – Режим доступу: <http://www.cisstat.com/>
15. Офіційний сайт Європейського центрального банку [Електронний ресурс]. – Режим доступу: <https://www.ecb.europa.eu/stats/html/index.en.html>
16. Офіційний сайт департаменту статистики Організації Об'єднаних Націй [Електронний ресурс]. – Режим доступу: <http://unstats.un.org/unsd/default.htm>