



# Methods of Multivariate Data Analysis

## Working program of the academic discipline (Syllabus )

### Details of the academic discipline

Level of higher education	<i>Second (master's)</i>
Branch of knowledge	<i>05 Social and behavioral sciences</i>
Specialty	<i>054 Sociology</i>
Educational program	<i>Conflict resolution and mediation</i>
Discipline status	<i>Selective</i>
Form of education	<i>full-time/part-time/distance/mixed</i>
Year of training, semester	<i>First year, 2nd semester</i>
Scope of the discipline	<i>4 credits</i>
Semester control/ control measures	<i>Test</i>
Lessons schedule	
Language of teaching	<i>Ukrainian</i>
Information about the course leader / teachers	<i>Lecturer: Doctor of Sciences in Education, Professor Liubov Panchenko, +380963352397, <a href="mailto:liubov.felixovna@gmail.com">liubov.felixovna@gmail.com</a> Practical / Seminar: Liubov Panchenko Laboratory: Liubov Panchenko</i>
Placement of the course	<i>Link to remote resource ( Moodle )</i>

### Program of educational discipline

#### 1. Description of the educational discipline, its purpose, subject of study and learning outcomes

Within the framework of the discipline, students will acquire knowledge and skills regarding the use of methods of multidimensional data analysis in sociology, interpretation of such analysis. Within the framework of the course, modern information technologies and means of visualization of the obtained data are studied. The planned types of training sessions include lectures, practical and laboratory classes, modular control work, calculation work.

The purpose of the discipline is to form students' abilities to choose adequate methods of analysis and to use multivariate statistics to analyze sociological data and research data, in particular peace and conflict research.

The main tasks of the academic discipline

- provide students with basic information about the essence of multidimensional statistics, its classification and use in sociology;
- form an understanding of discriminant analysis, multidimensional scaling, longitudinal analysis, survival analysis, structural equation modeling and other multivariate methods;
- teach them to use the capabilities of computer tools for multidimensional analysis of research data in sociology correctly, in particular for peace and conflict research;

- to form in master's students the need for further development of knowledge and skills in the field of application of multidimensional analysis methods and computer tools in future professional activity.

After mastering the academic discipline, master's students must demonstrate the following learning outcomes:

**KNOWLEDGE :**

- the main directions of using multidimensional statistical methods in sociology, the specifics and difficulties of their use
- basic concepts of multivariate statistics, purpose of discriminant, factor and cluster analysis, multidimensional scaling, loglinear analysis, survival analysis, structural equation modeling
- the main stages of carrying out each of the methods and statistics used in them
- capabilities of computer packages for statistical processing of multidimensional sociological information and presentation of its results

**SKILLS :**

- to determine the class of problems in sociology, in particular the analysis of peace and conflicts, for which it is necessary to apply multivariate statistical methods
- use multidimensional methods of analysis
- analyze and interpret empirical research data using these methods;
- choose methods adequate to the tasks;
- use modern packages for statistical processing of information;
- present the results of the research with the help of computer activities (including on the Internet).

**2. Pre-requisites and post-requisites of the discipline (place in the structural and logical scheme of training according to the relevant educational program)**

The discipline relies on students' knowledge of general sociology, methodology and methods of sociological research, mathematical and statistical methods of analyzing sociological information

**Content of the academic discipline**

Allocation of study time for the credit module

Form of education	Credit modules	In total		Distribution of study time by types of classes			Semester certification
		Credits	hours	Lectures	Practical training	SRS	
<i>Full-time</i>	<i>1</i>	<i>4</i>	<i>120</i>	<i>18</i>	<i>36</i>	<i>66</i>	<i>Test</i>
<i>Correspondence</i>	<i>1</i>	<i>4</i>	<i>120</i>	<i>8</i>	<i>4</i>	<i>108</i>	<i>Test</i>

Names of sections, topics	Hours in general	Distribution of study time by types of classes			
		Lectures	Practical (Seminar classes)	Laboratory	SRS
Topic 1. Fundamentals of multivariate statistics and their use in social research and peace and conflict research. Computer tools. R as a tool for multivariate statistics.	13	2	4	0	7
Topic 2. Data preparation for multivariate analysis.	13	2	4	0	7
Topic 3. Multivariate variance analysis and its use for social research data analysis.	13	2	4		7
Topic 4. Discriminant analysis and its use for the analysis of social research data.	13	2	4	0	7
Topic 5. Loglinear analysis in social research	13	2	4	0	7
Topic 6. Survival analysis and time series analysis	13	2	4	0	7
Topic 7. Multidimensional scaling in sociological and marketing research, peace and conflict research	13	2	4	0	7
Topic 8. Multidimensional methods in social network research (multidimensional correlations and regressions, correspondence analysis, multidimensional scaling , cluster analysis).	16	2	4	0	10
Topic 9. Structural equation modeling in sociological research	13	2	4	0	7
<b>In total</b>	120	18	36	0	66

### 3. Educational materials and resources

#### Recommended Books

1. Tabachnik B., Fidell L. Using multivariate statistics. 6<sup>th</sup> edition. - Pearson, 2013.
2. Bühl A., Tsöfel P. SPSS: art processing information. Analysis statistical data and recovery hidden regularities. – St. Petersburg : DiaSoft. YUP LLC. - 2002. - 608 p. (Translation from German).
3. Kramer D. Mathematical processing data in social sciences: modern methods. - M.: Publishing Center " Akademiya ", 2007.
4. Paniotto V.I., Maksimenko V.S., Kharchenko N.M. Statistical analysis of sociological data. - K.: KM Academy, 2004. - 270 p.
5. Panchenko L. F. Data analysis practicum - Luhansk, Publishing House of State University "Taras Shevchenko LNU ", 2013. - 269 p.
6. Panchenko L.F. Mathematical and statistical methods of sociological information analysis. K., KPI named after Igor Sikorskyi, 2018.

#### Additional literature

1. V. Paniotto , A. Hrushetskyi . Is modeling dead yet? The history of social modeling in Ukraine and an agent-oriented approach based on the example of forecasting the language situation in Ukraine. —

- Modern methods of data collection and analysis in sociology / Scientific editor. E. I. Golov akha and T. Ya. Lyubiva . - K.: Institute of Sociology of the National Academy of Sciences of Ukraine, 2013. - 140 p.
2. Panchenko L.F. On the issue of the use of quantitative methods in conflict studies // Bulletin of the National Technical University of Ukraine "Kyiv Polytechnic Institute". Politology. Sociology. Right: Coll. of science works – Kyiv, 2018. – No. 3 (39). – P.21–27.
  3. Panchenko L.F. Preparation of future sociologists for computer analysis of demographic processes and structures // Information Technologies and Learning Tools . - 2018. - Vol. 65(3). – C.166–183.
  4. The R environment (R– project ) <http://www.r-project.org/about.html>.
  5. Shrodt F. Sem mortals sins modern quantitative analysis in political science/ Per s Eng. A.A. Shirokanova // Sociology: methodology, method, mathematics modeling. – 2017. – No. 43. – C.154–210.
  6. Eberwein WD The Quantitative Study of International Conflict : Quantity and Quality ? An Assessment of Empirical Research / Wolf-Dieter Eberwein // Journal of Peace Research. - 1981. - Vol . 18, No. 1. – P. 19–38.
  7. Smith RP Quantitative Methods in Peace Research / Ron P. Smith // Journal of Peace Research. - 1998. - Vol . 35, No. 4. – P. 419–427.
  8. Beck N. Improving Quantitative Studies of International Conflict: A Conjecture / Nathaniel Beck , Gary King , and Langche Zeng // American Political Science Review . - 2000. - No. 94. - P. 21–36. <https://gking.harvard.edu/files/gking/files/improv.pdf> \_
  9. Druckman D. Doing Research : Methods of Inquiry for Conflict Analysis / Daniel Druckman / SAGE Publication . - 2005. - 387 p.
  10. Gagliardone I. Public Opinion Research in a Conflict Zone : Grassroots Diplomacy in Darfur / Gagliardone Iginio and Stremlau Nicole / 2008. – [Electronic resource]. – Mode of access: [http://repository.upenn.edu/cgcs\\_publications/13](http://repository.upenn.edu/cgcs_publications/13)
  11. Mitchell C. Conflict , Social Change and Conflict Resolution . An Enquiry . The Berghof Handbook Dialogue . -2005. – #5. – 25 years [Electronic resource]. – Access mode: [https://www.berghof-foundation.org/fileadmin/redaktion/Publications/Handbook/Dialogue\\_Chapters/dialogue5\\_mitchell\\_lead-1.pdf](https://www.berghof-foundation.org/fileadmin/redaktion/Publications/Handbook/Dialogue_Chapters/dialogue5_mitchell_lead-1.pdf)
  12. Coleman PT The Handbook of Conflict Resolution : Theory and Practice / Peter T. Coleman , Morton Deutsch, Eric C. Marcus / [Electronic resource] . – Access mode: <https://books.google.com.ua/books?id=ay74AgAAQBAJ&printsec=frontcover&hl=ru#v=onepage&q&f=false>
  13. Caplan R. Studying Conflict And Practicing Peacebuilding / Richard Caplan // E-International Relations . - July 20, 2017. - [Electronic resource]. – Access mode: <http://www.e-ir.info/2017/07/20/studying-conflict-and-practicing-peacebuilding/>
  14. Schrodt PA. Prediction of Interstate Conflict Outcomes Using a Neural Network / Philip A. Schrodt // Social Science Computer Review . - 1991. - Vol . 9, No. 3. – R. 359–380.
  15. Habtemariam E. Artificial Intelligence for Conflict Management / E. Habtemariam , Tshildzi Marwala, Monica Lagazio / 2007. – [Electronic resource]. – Access mode: [https://www.researchgate.net/publication/220487240\\_Artificial\\_Intelligence\\_for\\_Conflict\\_Management](https://www.researchgate.net/publication/220487240_Artificial_Intelligence_for_Conflict_Management)
  16. Agresti , A. 1996. An Introduction that Categorical Data Analysis . John Wiley & Sons , Inc. New York , New York , USA.
  17. Field A. Discovering Statistics Using SPSS Second Edition.
  18. Loglinear Models. Angela Jeansonne <http://lib.slon.pp.ru/ComputerScience/Compression/Log%20Linear%20Models.pdf>
  19. Upton H. Analysis conjugation tables. M. Finances and statistics. 1982. (The Analysis of Cross-tabulated Data. /Graham JG Upton. Lecturer in Mathematics University of Essex).
  20. Agresti Alan. Categorical Data Analysis.Second Edition.University of Florida.2002 John Wiley & Sons , Inc. , Hoboken , New Jersey .

#### 4. Methods of mastering an educational discipline (educational component)

Educational classes in the discipline "Methods of multivariate data analysis" are conducted in the form of lectures, practical and laboratory classes. In the lectures, the teacher introduces master's students to key statistical concepts, basic statistical methods used to analyze social research data. Lectures take place in a dialogic form using multimedia presentations, questions and discussion. Practical and laboratory classes are held using computer devices, where students perform the tasks of a computer practicum on data analysis in the Excel, SPSS (PSPP), R environments. All classes are accompanied by materials and tasks in the Moodle distance learning system .

##### Topics and content of lectures

**Lecture 1. Introduction to multivariate statistics.** Purpose and essence of multivariate analysis. Dependent and independent variables. Experimental and non-experimental research. Continuous, discrete and dichotomous data. Samples and general populations. Descriptive statistics and hypothesis testing. Orthogonality (independence of variables): standard and sequential analysis. A linear combination of variables. The number and nature of variables included in the analysis. Statistical power. Data in multivariate analysis. Choosing a method of multivariate analysis. Purpose and objectives of the course. Multivariate analysis software: Microsoft Excel, SPSS (PSPP), R, Amos , etc. *R* as a tool for multivariate statistics (Advantages of R. R Packages. Basic Data Types in R. Vectors, Factors, Matrices, Lists, Data Tables. Data Export and Import. RStudio Shell. R Multivariate Capabilities. Packages for Sociologists).

**Lecture 2. Data preparation for multivariate analysis.** Screening algorithm ungrouped data. Checking for inputs, missing data, distributions, and univariate outliers. Linearity and homoscedasticity . Transformations of variables. Finding multidimensional outliers. Variables that cause emissions. Multicollinearity.

**Lecture 3. Multidimensional variance analysis.** Purpose of multivariate variance analysis and its use in sociology. Examples of data. Step-by-step calculation algorithms. Presentation of results: descriptive statistics, Box's test , multivariate tests, homogeneity of variance test. Interpretation of results.

**Lecture 4. Discriminant analysis and its use in sociology.** Purpose of discriminant analysis and its use in sociology. Examples. Discriminant function. Stages of execution. Presentation of results: group statistics, test of identity of group means, correlation matrix, eigenvalues, Wilks' lambda , canonical coefficients of the discriminant function, classification matrix. Interpretation of results.

**Lecture 5. Loglinear analysis in social research.** The purpose of the analysis. Comparison of variance and loglinear analysis. The maximum likelihood ratio chi-square test for the main variables and their interactions. Testing k-order and higher-order interactions . Estimation of model parameters. Stepwise elimination method. Interpretation of results.

**Lecture 6. Survival analysis. Time series analysis.** General purpose and description of survival analysis. Examples of research. Types of research questions. Limitations of survival analysis. Fundamental equations for survival analysis. Tables of life. Graphs of life tables. Test for group differences.

Purpose of *time series analysis*. Examples of research. Types of research questions. Samples of autocorrelation. Seasonal cycles and trends. Prognostication. The effect of the intervention. Limitations of time series analysis. Basic equations of ARIMA models. Interpretation of results.

**Lecture 7. Multidimensional scaling and its use in sociology and marketing.** Purpose of multidimensional scaling. Examples of use in sociology and marketing. Metric and non-metric methods.

Spatial maps. Stages of implementation. Statistics used: similarity score, preference ranks, stress, rotation, stimulus coordinates. Interpretation of results.

**Lecture 8. Multidimensional methods in social network research.** Multivariate correlations and regressions, correspondence analysis, multivariate scaling, cluster analysis. Social network analysis software. Features of NodeXL, Gephi, R.

**Lecture 9. Fundamentals of structural equation modeling.** Purpose of the SEM method. The use of SEM in sociology and behavioral sciences. Basic graphical elements: manifest variable, latent variable, directional relationship, covariance (bidirectional relationship). Structural models of correlation, multiple regression, factor analysis. Stages. Models: path model, confirmatory factor analysis, structural model. Evaluation and correction of the model. SEM software. Amos SPSS.

### Weekly planning of practical and laboratory work

No z/p	Learning outcomes	Control evaluation measures	Deadline, one week
1	<p>Knowledge:</p> <ul style="list-style-type: none"> <li>– the main directions of using mathematical and statistical methods, the specifics and difficulties of their use in social science research</li> <li>– the essence of multidimensional statistics and directions of its use in sociology</li> </ul>	Implementation and report on the work of the computer workshop	Week 1-18
2	Knowledge of data screening , experience with SPSS modules for data verification, analysis of their distributions, missing values, transformations, etc.	Implementation and report on the work of the computer workshop	1-2
2	Knowledge: purpose and essence of multivariate variance analysis. Proof of construction of variance analysis models in Excel, SPSS, R environments	Implementation and report on the work of the computer workshop.	Week 3-4
3	Knowledge: purpose and essence of discriminant analysis. Proof of construction of variance analysis models in SPSS, R environments	Implementation and report on the work of the computer workshop	Week 5-6
4	Knowledge: purpose and essence of longitudinal analysis. Proof of construction of longitudinal analysis models in SPSS, R environments.	Implementation and report on the work of the computer workshop	Week 7-8
5	Understanding the essence and purpose of survival analysis and time series analysis Evidence of building models of survival analysis and time series analysis in SPSS, R environments	Implementation and report on the work of the computer workshop,	Week 9-10
6	Understanding the essence and purpose of multidimensional scaling in SPSS, R environments. Experience building multidimensional scaling models in SPSS, R environments	Implementation and report on the work of the computer	Week 11-12

		workshop	
7	Understanding the essence and purpose of multivariate analysis methods for social network analysis. Experience building multivariate correlation and regression models	Implementation and report on the work of the computer workshop, RGR	Week 13-14
8	Knowledge of the possibilities and essence of modeling with structural equations, proof of construction of models using AMOS SPSS	Implementation and report on the work of the computer workshop of the MKR	Week 15-16
9	Test		17

### 5. Independent work of a student/graduate student

Types of independent work: preparation for practical and laboratory classes (review of the materials of the current lecture), performance of tasks for independent work from the manual, performance of calculation work. All deadlines are given in the Moodle system on a weekly basis.

## Politics of discipline

### 6. Policy of academic discipline (educational component)

#### Violation of deadlines and incentive points

Incentive points	
Criterion	Weight score
Participation in the conference / publication of theses	5 points
Publication of a scientific article	10 points

#### Attending classes

Attending lectures, practical and laboratory classes is mandatory.

#### Missed evaluation control measures

A missed modular test can be written again, but before the intermediate certification is issued, taking into account the time needed by the teacher to check the work.

#### The procedure for contesting the results of assessment control measures

Students have the opportunity to raise any issue relating to the screening process and expect it to be dealt with in accordance with pre-defined procedures. In order to appeal the control measures, the student must submit an application, in which the reason for the appeal should be indicated, and the facts of the teacher's bias should be given. The teacher must discuss this application with the student personally at the consultation. The commission may decide to conduct the control measure again or reject the application. The commission's decision is final and cannot be appealed.

## Calendar border control

At the first certification, the student receives "certified" if his current rating at the time of certification is 20 or more points.

At the second certification, the student receives "certified" if his current rating at the time of certification is 40 or more points. If this indicator does not meet the requirements, "not certified" is issued.

Work-out of "non-certification" is carried out during consultations with the teacher by oral answers to questions about material not mastered by the student and demonstration of completed practical and laboratory work.

Criterion		The first certification	The second attestation
Term of attestation		Week 8	Week 14
Conditions for obtaining attestation	Current rating	≥ 20 points	≥ 40 points
	...	+	+
	...	...	...
	...	...	...
	...	...	...

## Academic integrity

The policy and principles of academic integrity are defined in Chapter 3 of the Code of Honor of the National Technical University of Ukraine "Ihor Sikorsky Kyiv Polytechnic Institute". More details: <https://kpi.ua/code>.

## Norms of ethical behavior

Standards of ethical behavior of students and employees are defined in Chapter 2 of the Code of Honor of the National Technical University of Ukraine "Ihor Sikorskyi Kyiv Polytechnic Institute". More details: <https://kpi.ua/code>.

## 7. Types of control and rating system for evaluating learning outcomes (RSO)

No z/p	Assessment control measure	%	Weight score	How many	In total
1.	Laboratory classes	60	3	20	60
2.	Computational graphic work - a data analysis case (RGR)	20	20	1	20
3.	Modular control work (MCR)	15	15	1	15
4	Incentive points	5	5	1	5
	In total				100

Semester control: [assessment](#)

Conditions for admission to semester control: [semester rating of more than 40 points](#).

Table of correspondence of rating points to grades on the university scale:



<i>Scores</i>	<i>Rating</i>
100-95	Perfectly
94-85	Very good
84-75	Fine
74-65	Satisfactorily
64-60	Enough
Less than 60	Unsatisfactorily
Admission conditions not met	Not allowed

#### **8. Additional information on the discipline (educational component)**

**The list of questions for assessment is provided in the Moodle system .**

**The working program of the academic discipline ( syllabus ):**

**Compiled** by professor of the department, doctor of pedagogical sciences, professor L.F. Panchenko.

**Approved by the Department of Sociology (protocol No. 10 dated April 8, 2023).**

**Agreed by the Methodical Council of KPI named after Igor Sikorskyi (protocol No. 8 dated June 2, 2023).**