SYLLABUS

INTERNATIONAL EUROPEAN UNIVERSITY

EUROPEAN SCHOOL OF BUSINESS

Empirical methods of software engineering
Educational program «Software engineering»

2024

SYLLABUS

1	Name of the course and educational program

Empirical methods of software engineering
Educational program "Software engineering"

2	Course description

The study of the academic discipline "Empirical Methods of Software Engineering" is aimed at mastering and using knowledge about the mathematical foundations of the practical application of empirical methods and tools of statistical analysis in software engineering. For this purpose, a set of knowledge and skills is formed for students to master the methods of constructing mathematical models using statistical methods, develop logical and algorithmic thinking of students, familiarize themselves with the possibilities of applying statistical methods for processing and analyzing empirical software information, and form in students an understanding of the basic principles that underlie the use of computing capabilities in processing and analyzing experimental data.

3	Study prerequisites

The academic discipline is related to the disciplines "Higher and Applied Mathematics", "Probability Theory and Mathematical Statistics", "Object-Oriented Programming", "Computer Discrete Mathematics".

4	Amount of credits/hours

4 ECTS credits/ 120 hours

5	Training format

Blended learning

6	Classroom location

Audience 405. https://dist.ieu.edu.ua/enrol/index.php?id=598

7	Information about the teacher

Ivan Kazachkov, Dr. technical Sciences, professor

8	Department

Department of Information Technologies

9	Office location

Kyiv, Akademika Glushkova Ave., 42 B, room 505

10	Schedule of counseling

Every Monday from 12:00 to 16:00 with prior appointment via corporate mail

11	E-mail of the teacher

ivankazachkov@ieu.edu.ua

12	Course objectives

The aim of the course is to provide students with knowledge of scientific concepts and principles of practical application of empirical methods and tools of statistical analysis, relevant in the aspect of software, providing an appropriate set of skills and abilities necessary to increase the efficiency of the professional and scientific activities of future software engineering specialists.

13	The role of academic discipline in achieving program results

PR03. Know the basic processes, phases, and iterations of the software life cycle.

PR05. Know and apply relevant mathematical concepts, methods of domain, system and object-oriented analysis and mathematical modeling for software development.

PR06. Ability to choose and use a software development methodology appropriate to the task.

PR07. Know and apply in practice the fundamental concepts, paradigms and basic principles of the functioning of language, instrumental and computational tools of software engineering.

PR11. Select input data for design, guided by formal methods of requirements description and modeling

14	Learning outcomes

Know:
- basic statistical methods y,
- the possibilities of their application for processing and analyzing empirical software
- information.

Be able:
- use graphing tools, diagrams, histograms, etc. based on the obtained experimental data;
- use computer capabilities when processing and analyzing experimental data;
- apply statistical procedures to solving problems and functions of special programs, including spreadsheets.

15	Course content

Section 1. PRELIMINARY KNOWLEDGE OF PROFESSIONAL AND APPLIED ASPECTS OF SOFTWARE ENGINEERING AND EMPIRICAL RESEARCH METHODS

Topic 1. Introduction to methods of experimental research of software engineering problems.

Topic 2. Background information on probability theory and descriptive statistics.

Topic 3. Random variables. Important discrete and continuous istributions.

Section 2. STATISTICAL ESTIMATION OF DISTRIBUTION PARAMETERS AND TESTING STATISTICAL HYPOTHESES

Topic 4. Fundamentals of the theory of statistical estimation.

Topic 5. Fundamentals of statistical hypothesis testing.

Topic 6. Consent Criteria.

Topic 7. Testing hypotheses related to the normal distribution.

SYLLABUS

Section 3. BASIC TYPES OF STATISTICAL ANALYSIS
Topic 8. Analysis of variance
Topic 9. Correlation and regression analysis
Topic 10. Multivariate analysis and other statistical methods
Topic 11. Empirical methods for assessing software reliability

16 — Course materials and requirements

1. Martynenko M.A., O.M. Neschadym, V.M. Safonov. Probability theory and mathematical statistics: Textbook. Part 1 Kyiv: CP "KOMPRINT", 2012. 288.
2. Martynenko M.A., O.M. Neschadym, V.M. Safonov. Probability Theory and Mathematical Statistics: Textbook. Part II Kyiv: Publishing House "KOMPRINT", 2013. 278 p.
3. Sulima I.M., I.I. Kovtun et al. Applied Mathematics: Probability Theory. Mathematical Statistics. Kyiv: NAU, 2005. 148 p.
4. Bidyuk P.I., Tkach B.P., Harington Tom. Mathematical Statistics. Kyiv: Personal. 2017. 255 p.
5. Sheftel Z. G. Theory Probability – 2nd ed., revised and supplemented . Kyiv : Higher School, 1994. 192 p.
6. Hollander M., Wolfe D. Non-parametric methods of statistics. Trans. With English M. : Edition House "Williams", 2003 p

17 — Technical requirements for working on the course

In order to access the course materials, you will need regular access to a computer and the Internet. In order to successfully study and pass the exam from the training course, it is necessary to constantly familiarize yourself with the materials posted on the university's remote platform (Moodle) in the course "Empirical methods of software engineering". You also need to create reporting documents for the performance of practical work and upload them to the platform (the platform can only be used from a corporate email account).

In the case of problems with access to the distance learning platform, it is necessary to notify the dean's office or the headmaster, or the course teacher directly.

18 — Learning process

The process of studying the course "Empirical methods of software engineering" includes lectures and practical sessions.

During the lectures, such teaching methods as lecture, lecture-conversation, discussion, discussion of problematic issues, demonstration, and analysis of various situations will be used according to the topic of the lectures.

During practical classes, such teaching methods as surveys, testing, performance of individual tasks, performance of analytical and calculation works, solved specific problems and situations will be used).

SYLLABUS

19	Signs of discipline

Term of teaching	Semester	International Disciplinary integration	Course study	Cycles: general training/ professional training/ free choice
1 semester	8th semester	No	4ty course	Cycle of professional training

20	Evaluation policies

You will have different ways to demonstrate your knowledge and skills throughout the semester. This includes how you attend class, how and what you contribute to topic discussions, how you complete and complete lab assignments and tests on time, how you complete independent work assignments, and the ability to present your work. In addition, it is possible to perform tasks that are performed individually or in a small group in the form of a student scientific work.

Activities during the semester	Maximum number of points during the semester
Current work (attendance, supervision at lectures, completion of practical work)	20
Tests (8)	16
Practical work (8 pieces)	24
Together	**60**
Exam	**40**
TOTAL	**100**

21	Rating scale

The grade for the discipline is defined as the sum of the points scored for the current activity in the semester. Each module includes an assessment score for the student's current work. Module control activities are carried out upon completion of the study of the taught material of this module. The minimum number of points for the current educational activity, which allows the discipline to be counted as completed, must be at least 60. The maximum point for the discipline is 100.
The total grade for studying the discipline is set according to the national and European scale (EKTS).
The overall final grade in points, according to the national scale and according to the ECTS scale, is entered in the student's assessment and examination information, study card and student's assessment book.

Rating scale: national and ECTS

The sum of points for all types of educational activities	Evaluation ECTS	Evaluation on a national scale	
		for an exam, course project (work), practice	for credit
90-100	A	perfectly	Enrolled
82-89	B	good	
74-81	C	good	
66-73	D	satisfactorily	
60-65	E	satisfactorily	
30-59	FX	unsatisfactory with possibility reassembly	not counted with the possibility of retaking
1-29	F	unsatisfactory with mandatory repeated sstudy of the discipline	not enrolled with mandatory repeated study of the discipline

22	How to find out your score

To check your assignment grades and read the teacher's comments, you need to check the relevant tabs on the distance learning platform (Moodle) in this course.

You can also get information about the received grades in the joint chat of the subject group (Viber or Telegram) or directly from the course instructor via corporate mail, messengers or by appointment on the days of consultations.

23	Course policies

For the productive educational and cognitive activity of the applicants when studying the discipline, thematic lectures are held and practical classes are conducted in the form of laboratory works.

In classes and during his stay at the university, the student must treat teachers, staff and other students with respect, attend classes according to the schedule, come on time and not leave the classroom without the teacher's permission. It is necessary to complete all academic tasks and their work within the specified time.

The teacher, in turn, must constantly raise his professional level, pedagogical skill, and general culture, provide conditions for students to master educational programs at the level of mandatory requirements for the content, level and scope of education, and promote comprehensive professional development of students. It is mandatory to follow the educational and thematic plan, not to be late for classes, not to allow any manifestations of corruption, discrimination, bullying, harassment and oppression of the rights of those seeking education.

Education is based on the application of active learning methods. Active participation is expected and the norm. Attendance and active participation make up 80% of the grade. A student who, for good reasons, documented, was not subject to current control has the right to undergo current control within a two-week period after returning to studies.

A student who was absent from classes without valid reasons, did not participate in current control activities, did not liquidate academic debt, is not allowed to take the final semester control of knowledge in this discipline, and on the day of the exam in the examination information by a scientific and pedagogical employee the grade "not admitted" is issued. Retaking the exam in the discipline is prescribed on the condition that all types of educational, independent (individual) work provided for in the work curriculum of the discipline are performed, and is carried out in accordance with the liquidation schedule approved by the directorate.

The academic integrity of any institution of higher education requires integrity in teaching and research, so academic integrity is required of all MEU students. Academic dishonesty is prohibited in all programs at our university. All participants in the educational process are guided by the principles of academic integrity.

24	Completing the task late, correcting grades, working out

Assignment reports must be uploaded to Moodle by the due dates specified in the course schedule. Best practice would be to complete assignments as soon as possible after receipt to allow enough time to actively participate in class. If more time is needed to complete the task, flexible deadlines are available. Completed assignments are accepted for full credit until the last class in the discipline on the schedule, after which 40% partial credit based on the grade received will be awarded within a week of the last day of class. Assignments that were not submitted at all will receive 0.

If classes are missed for more than one week due to illness or other reasons, it is necessary to contact the teacher to agree on alternative options for completing tasks. Deadlines work both ways, and meeting them ensures that your instructor provides timely feedback on your assignments to ensure you stay on course.

25	Teacher's response time (about checking assignments)

Via corporate mail (within 24 hours), via messengers (within 1-2 hours).

26	Effective communication

Effective communication is essential to success in this course, we recommend using the following channels:
Forum of questions and answers: for general course questions, you need to check the FAQ section in Moodle and then post your question in the Q&A forum to ask your colleagues or the instructor (guaranteed to receive a notification by e-mail every time a new publication or an answer to a question appears);
E-mail: have a personal question related to studying the course, write to the teacher directly;
Social networks, messengers: personal communication with classmates, teacher;
Face-to-face meeting: communication with classmates during classes and with the teacher on consultation days.

27	Policy of publication and distribution of course materials

Students may not post, publish, sell, or otherwise publicly distribute course materials without written Permission the teacher. Such materials include: lecture notes, slides (presentations) of lectures, video or audio recordings, tasks, problem sets, tests, other students' works and answers, etc. Students who sell, post, publish, or distribute course materials without written permission or otherwise may be subject to disciplinary action, up to and including withdrawal.
The use of generative AI is permitted subject to adherence to the principles of academic integrity.

SYLLABUS

28	Expected workload and involvement of students

Approximately 2-3 hours per week should be allocated to work in this course. If circumstances arise that force you to spend more time on one of the tasks, you must inform the teacher by e-mail (messenger).
An extension of the submission deadline is possible only under the condition that the teacher is informed in advance that it is impossible to submit the assignment by the specified time. Students are expected to have a backup plan in case of computer malfunctions or Internet outages.

29	Support services

Electronic schedule: https://rozklad.ieu.edu.ua
Online library: https://onlinelibrary.ieu.edu.ua
Repository: https://sed.ieu.edu.ua/index.php/sed/index
Educational Ombudsman: https://ieu.edu.ua/pro-mieu/ombudsmen

30	Course schedule

Topic name	Content of practical class
Topic 1. INTRODUCTION TO METHODS OF EXPERIMENTAL RESEARCH OF SOFTWARE ENGINEERING PROBLEMS. 1. The place and importance of the academic discipline as a mathematical component of professional training analytical) and its role in the mathematical analysis of experimental ata.lectures; 2. Empirical distribution function, its properties. 2. Typical tasks related to software research 4. General characteristics of empirical methods. 5. Scales of measurement of traits. Relationship between measurement scales 6. Basic methods of processing experimental studies	1. Oral survey 2. Practical Work 1. Software metrics 3. Questions for self-control. Automation of statistical processing of spreadsheets data. 4. Tests
Topic 2. PREVIOUS KNOWLEDGE OF PROBABILITY THEORY AND DESCRIPTIVE STATISTICS. 1. Random selection method. Formation of sample populations from the general population. Representativeness 2. Standard presentation of statistical material (tabular, graphical, analytical) and its role in the mathematical analysis of experimental data 3. Empirical distribution function, its properties 4. Numerical characteristics of a statistical series and their properties, methods finding. 5. Automation of statistical processing of tabular data	1. Oral survey 2. Practical Work 2. Samples and their representation. 3. Questions for self-control. Distributions related to the normal distribution: chi-square, Student (t-distribution), Fisher-Snedecor. 4. Tests

Topic 3. RANDOM VARIABLES. IMPORTANT DISCRETE AND CONTINUOUS DISTRIBUTIONS. 1. Discrete and continuous random variables on R1. 2. Theoretical distribution function. Differential and integral functions distribution of continuous random variables 3. Numerical characteristics of random variables and their properties 4. Multidimensional random variables. Systems of two continuous random variables 5. Laws of distribution of discrete random variables: uniform on finite set, binomial, geometric, hypergeometric distributions. Poisson distribution and others. 6. Laws of distribution of continuous random variables: uniform on interval, exponential, normal, Laplace distribution and others. 7. Distributions related to the normal distribution: chi-square, Student's (tdistribution), Fisher-Snedecor	1. Oral survey 2. Practical Work 3. Numeric characteristics of the statistical sample distribution. 3. Questions for self-control. Application of packages in statistical finding distribution estimates. 4. Tests
Topic 4. BASIS OF THEORY STATISTICAL EVALUATION. 1. General concepts about estimating unknown parameters of general populations and the main problems of the theory of statistical estimation 2. Conclusions from the law of large numbers. Main statistical requirements parameter estimates 3. Estimates for the mathematical expectation and variance of the results observations. 4. Methods of obtaining statistical estimates: method of moments (method of K. Pearson), maximum likelihood method (R. Fisher method). 5. Interval estimates of distribution parameters. Confidence (reliable) interval. Level of significance 6. Interval estimation of the probability of an event through Frequency. 7. The problem of constructing confidence limits for the mathematical expectation when an arbitrary distribution law of the general population 8. Confidence limits for estimating the mathematical expectation of a normal distribution with known and unknown variance (mean squared deviation). Examples. 9. The problem of constructing an interval estimate for the mean square (standard) deviation in a normal population distribution 10. Accuracy of sample estimation. Finding gross errors 11. Application of statistical packages in finding distribution estimates	1. Oral survey 2. Questions for self-control. Analysis repeated paired observations using sign ranks. 4. Tests
Topic 5. BASIC VERIFICATION OF STATISTICAL HYPOTHESES. 1. Statistical models. The idea of random selection. Pragmatic rule 2. Testing statistical hypotheses (general provisions). Simple and complex Hypotheses. Alternatives. Significance level. Statistical criteria. Errors of the first and second kind. Power of the criterion. Algorithm statistical hypothesis testing. 3. Examples of statistical models and hypotheses. 4. Testing statistical hypotheses (applied problems). Test scheme Bernoulli. Sign criterion for one sample 5. Hypothesis testing in two-sample problems. 6. Paired observations. Sign criterion for the analysis of paired repeated observations. Analysis of repeated paired observations using signed ranks (Wilcoxon signed rank sum test).	1. Oral survey 2. Practical Work 4. Testing parametric statistical hypotheses. 3. Questions for self-control. Using statistical packages to analyze the nature of the distribution.. 4. Tests

Topic 6. CONSENT CRITERIA. 1. Kolmogorov's agreement criteria in the case of a simple hypothesis. 2. K. Pearson's chi-square agreement criterion for a simple hypothesis. 3. Criteria for agreement for a complex hypothesis 4. Fisher's chi-square agreement criterion for a complex hypothesis 5. Agreement criterion for the Poisson distribution. Other agreement criteria	1. Oral survey 2. <u>Practical Work 5</u>. Testing parametric statistical hypotheses; 3. <u>Questions for self-control</u>. Poisson distribution and other agreement criteria. 4. Tests
Topic 7. TESTING HYPOTHESES RELATED TO THE NORMAL DISTRIBUTION. 1. Study of normal samples 2. Testing the hypothesis of a normally distributed mathematical expectation random variable from one sample 3. Testing the hypothesis of a normally distributed mathematical expectation random variable by two ndependent samples. 4. Testing the hypothesis about the variance of a normally distributed general populations with one sample. 5. Testing the hypothesis about the variance of a normally distributed general populations by two independent samples 6. Using statistical packages to analyze the nature of the distribution	1. Oral survey 2. <u>Practical Work 6</u>. Testing the hypothesis about the mathematical expectation of a normally distributed random variable; 3. <u>Questions for self-control</u>. Analysis of repeated paired observations using sign ranks. 4. Tests
Topic 8. DISPERSION ANALYSIS. 1. General analysis of idea variance. 2. Statement of the problem of one-way analysis of variance. Table of observations of one-way analysis 3. General, factor and residual variance 4. Analysis of the interaction of factors. Comparison of the means of the variance analysis method. 5. Using statistical packages when conducting univariate analysis of Two-factor 6. Analysis variance. 7. Relationship between two- and one-factor analysis problems	1. Oral survey 2. <u>Questions for self-control</u>. The connection between tasks and singlefactor analysis.. 4. Tests
Topic 9. CORRELATION AND REGRESSION ANALYSIS. 1. Relationships of features in quantitative measurement scales. Functional and stochastic dependencies. 2. The essence and significance of correlation analysis. Correlation field. Correlation table. Conditional mean. Sample regression equation. 3. Simple linear correlation Correlation coefficient and its properties 4. Applying the least squares method to find the parameters regression 5. Nonlinear correlation. Selective correlation relation 6. Relationship of features measured on a scale of orders 7. Rank correlation. Spearman's rank correlation coefficient. Kendall's rank correlation coefficient	1. Oral survey 2. <u>Practical Work 7</u>. Correlation and regression analysis 3. <u>Questions for self-control</u>. Rank correlation. 4. Tests
Topic 10. MULTIVARIABLE ANALYSIS AND OTHER STATISTICAL METHODS. 1. Factor analysis. 2. Discriminant analysis 3. Cluster analysis 4. Quality control methods	1. Oral survey 2. <u>Questions for self-control</u>. Quality control methods. 4. Tests

SYLLABUS

Topic 11. EMPIRICAL METHODS OF RELIABILITY SOFTWARE ASSESSMENT. 1. Concepts and principles of testing. Types of testing. Baseline testing. 2. Methods of testing conditions. Testing cycles 3. Software settings 4. Tools and methods for detecting software errors . Categories of errors in software .	1. Oral survey 2. Practical Work 8. Processing the results generalization of Experiments. 3. Questions for self-control. Resources methods detection of software errors. 4. Tests

31	Tips for successful study

The goal is unique to everyone, but its correct setting greatly affects the result, as well as the learning process. For example,
- processing the materials of the theoretical component (lectures) of the discipline will provide insight and knowledge about the
- development process and the architecture of the OS itself, and the implementation of the practical component - the acquisition
- of practical skills in the use of methods and tools for creating system software. After all, any training that follows a clear plan and
- with a serious attitude to the material will always be successful.

   So, if you want to successfully master this subject, you must be:
   - persistent, attentive and inquisitive;
   - creative and cheerful, open to communication and discussions
   - ready to receive information and knowledge on the subject not only during lectures, but also during extracurricular hours

See you soon!