



SYLLABUS OF THE ACADEMIC DISCIPLINE
Applied Statistical Analysis using Python
Educational program component – mandatory (3.0 credits)

| | |
|---|---|
| Educational and Professional Program | Information technology and project management |
| Specialty | 122 – Computer Science |
| Field of knowledge | 12 – Information technology |
| Level of higher education | first (bachelor's) |
| Language of teaching | Ukrainian |
| Teacher profile | Pasichnyk Halyna Saveliivna Head of the Department of Mathematical Modeling, Candidate of Physical and Mathematical Sciences, Associate Professor https://mathmod.chnu.edu.ua/pro-nas/spivrobotnyky/pasichnyk-halyna-saveliivna/ |
| Contact phone number | +38037-258-48-25 |
| E-mail: | h.pasichnyk@chnu.edu.ua |
| Course page in Moodle | https://moodle.chnu.edu.ua/course/view.php?id=4851 |
| Consultations | Wednesday, 13:00 – 14:00 |

ANNOTATION OF THE ACADEMIC DISCIPLINE

This course is designed to position you for success by diving into the real-world of statistics and data science. Dive into the world of Python-based statistical analysis with this comprehensive track designed to equip you with the essential skills to make data-driven decisions. With its user-friendly syntax and robust libraries, Python has become a favorite tool for data scientists and statisticians. Whether it's improving business strategies with A/B testing, drawing conclusions from companies' data, or developing probabilistic models in Bayesian frameworks, you will be well-prepared to tackle challenges in any industry. Start this track today to elevate your capabilities and become a critical contributor to any data-centric team.

Applied Statistics includes planning for the collection of data, managing data, analyzing, interpreting and drawing conclusions from data, and identifying problems, solutions and opportunities using the analysis. This course helps the learners to build critical thinking and problem solving skills in data analysis and empirical research. Learners will learn where data come from, what types of data can be collected, study data design, data management, and how to effectively carry out data exploration and visualization. By leveraging the pandas and matplotlib python libraries, the learners will gain exposure to tools used in data analysis, visualization, and data science using Python programming Language. This course specifically explores inferential statistics - the science of applying statistical techniques to quantify and answer real-world data analysis questions.

EDUCATIONAL CONTENT OF THE EDUCATIONAL DISCIPLINE

| MODULE 1. Important Concepts in Statistics | |
|---|--|
| Topic 1 | Structured data elements |
| Topic 2 | Data and sample distributions |
| Topic 3 | Statistical Experiments and Significance Testing |
| Topic 4 | Regression and Prediction |
| MODULE 2. Using Python libraries | |
| Topic 5 | Python language basics Python language grammar. Data storage structures, control structures. Creating functions and classes. |
| Topic 6 | NumPy library analytical capabilities NumPy library functionality related to the implementation of mathematical calculations, in particular, statistical analysis and generation of pseudo-random variables. |
| Topic 7 | Using the Pandas library for data storage and analysis pandas.Series class. Attributes and methods. Pandas. DataFrame class. Attributes and methods. Methods for combining and grouping data. Data visualization in pandas. |
| Topic 8 | Main data visualization principles in the matplotlib library |
| Topic 9 | Using the Statsmodels library for statistical data analysis Statistics stats, Nonparametric Methods, Empirical Likelihood, Distributions, Graphics |

FORMS, METHODS AND EDUCATIONAL TECHNOLOGIES OF TEACHING

Learning and teaching methods: lectures, laboratory classes, e-learning using the Moodle system, testing, completing INDS tasks.

FORMS AND METHODS OF CONTROL AND EVALUATION

Types and forms of control

1. Current (oral questioning, solving problems)
2. Modular (tests, laboratory work).

Assessment tools: - tests; team projects; analytical reports on the performance of laboratory work, individual tasks and independent work.

Final control - exam.

CRITERIA FOR ASSESSING LEARNING RESULTS

The system for assessing the level of educational achievements is based on the principles of ECTS and is cumulative. During the semester, students complete two tests and 3 laboratory works. Each test is evaluated with a maximum of 5 points, and laboratory works are evaluated with a maximum of 15 points. The final control in the discipline is an oral exam (40 points).

POLICY ON ACADEMIC INTEGRITY

Adherence to the policy on academic integrity by participants in the educational process when studying an academic discipline is regulated by the following documents:

- ✓ «Code of Ethics of Yuriy Fedkovych Chernivtsi National University»
<https://www.chnu.edu.ua/media/jxdfs0zb/etychnyi-kodeks-chernivetskoho-natsionalnoho-universytetu.pdf>
- ✓ «Regulations on the detection and prevention of academic plagiarism at Yuriy Fedkovych Chernivtsi National University»
<https://www.chnu.edu.ua/media/n5nbzwgb/polozhennia-chnu-pro-plahiat-2023-plusdodatky-31102023.pdf>

INFORMATION RESOURCES

- 1 Bengfort B. Applied Text Analysis with Python: Enabling Language Aware Data Products with Machine Learning – O'Reilly – 2018.
2. Wes McKinney Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython 2nd Edition – O'Reilly – 2017.
3. Peter Bruce, Andrew Bruce, Peter Gedeck Practical Statistics for Data Scientists: 50+ Essential Concepts Using R Python and 2nd Edition – O'Reilly Media – 2020.