



HIGHER SCHOOL OF ECONOMICS

NATIONAL RESEARCH UNIVERSITY

The Federal Government Autonomous Institution of Higher Education
National Research University Higher School of Economics

International Laboratory for Applied Network Research

Master's Program: Applied Statistics with Network Analysis

Contemporary Data Analysis: Methodology and Methods of Interdisciplinary Research

Room TBD

Author: Valentina Kuskova
Office: Myasnitskaya 20-231
Phone: +7-916-974-9287
E-mail: vkuskova@hse.ru; all student submissions – masna@hse.ru
Office Hours: Thursdays 15:00-18:00 and by appointment
Course website: Link given in class

I. Required readings

This is an introductory course, and it has a lot of required readings. All the readings are listed in a separate document (“Course Schedule”) and will be provided in advance via class web site.

II. Course Summary

This course is a required foundational course for masters' students in “Applied Statistics with Network Analysis” program, designed to familiarize them with the most recent developments in interdisciplinary statistical methods. The students will get an overview of data and approaches to analyzing them (remember, “data” is always plural!), including complex models. The course will also emphasize problem formulation at the intersection of mathematics and social sciences, integrate the most important concepts from probability theory, and overall, is designed as a “gateway” to graduate work in statistics, where the mathematical concepts are bridged with applied concepts and research design, depending on the discipline.

III. Field of application and normative references

The program of the course builds up the minimum requirements to the knowledge and skills of a student and determines the content and types of classes and assessment. The program is designed for instructors of this discipline, teaching assistants and students of the Master's program “Applied Statistics with Network Analysis.” The program is developed according to:

- Educational standards of NRU HSE for the master's programs in “Applied mathematics and informatics” field, approved by HSE Scientific Council in December of 2017
- University Academic Plan of NRU HSE for Master level education, developed in 2014.

IV. Course Description

This course is about interdisciplinary use of statistics, and unlike other statistical courses, which focus on specific methods, this course will focus on the broader areas within statistics. There are four major areas within the field of statistics that it will focus on: data management, statistical inference, statistical prediction, and statistical reporting. The course will cover several major areas that apply universally to any field of study: all about data (sources, including surveys – and methodology of survey design; sampling; approaches to missing data; issues with small and large samples, including the overview of “big data,” and bootstrapping); problem formulation in applied disciplines; the role of probability theory (examining social problems from probabilistic perspective; different distributions and their applications; mathematical expectation, its application and parameter estimation; hypothesis testing and the notion of statistical significance and confidence); parametric vs. non-parametric data analysis; social mathematical models; complex models, including network models; and current issues in forecasting and planning.

V. Course Goals and Outcomes

- a. The general objectives of the course are:
- To ensure that you the most advanced recent approaches to working with data, including approaches to incomplete data, different data sizes, and network data.
 - To provide you with an understanding of the role that probability plays in statistical analysis, including the concepts of statistical significance and confidence.
 - To provide you with an overview of the most recent advances in network science and applied statistical methods, complex statistical modeling, analysis, and forecasting
 - To familiarize you with the requirements and guidelines of scientific publishing both in Russia and abroad.
 - To help you continue developing your written and oral communication skills.

VI. Students' Competencies to be Developed by the Course

The Course develops the following competencies

Competencies	NC/NRU-HSE Code	Descriptors - the learning outcomes (the indicators of achievement)	Teaching forms and methods of that contribute to the development of a competence
<u>Systemic Competencies</u>			
1. Ability to reflect (evaluate and reprocess) studied scientific methods and techniques of professional activity.	CK-1	Can evaluate and reprocess methods and techniques of contemporary data analysis for a given problem.	Lectures, readings, in-class exercises, data analysis projects
2. Ability to create new theories, invent new techniques and tools of professional activity.	CK-2	Can use their knowledge in contemporary data analysis to create new theories, invent new techniques and tools of professional activity.	Lectures, readings, in-class exercises, data analysis projects

Competencies	NC/NRU-HSE Code	Descriptors - the learning outcomes (the indicators of achievement)	Teaching forms and methods of that contribute to the development of a competence
3. Ability to independently learn new research methods, change the scientific and production profile of their activity.	CK-3	Can use new methods and techniques of contemporary data analysis, additional packages and tools, without direct supervision.	Lectures, readings, in-class exercises, data analysis projects
4. Ability to analyze, verify, evaluate the completeness of information in the process of their professional activities, to replenish and synthesize missing information, if necessary.	CK-6	Able to analyze, verify, evaluate the completeness of information, can integrate information found from various sources and compensate for lack of data by adjusting models.	Lectures; independent work.
5. Is able to conduct a professional, including scientific and research activities, in the international environment.	CK-8	Can conduct professional, including scientific and research activities, in the international environment. Concisely and precisely expresses research ideas in English in written and oral communication.	Lectures, readings, in-class exercises, data analysis projects
<u>Social and Personal Competencies</u>			
6. Ability to transmit the norms of a healthy lifestyle, to lead by their example.	ΠΚ-1	Transmit the norms of a healthy lifestyle, and are capable to lead by their example.	In-class exercises
7. Ability to use social and multicultural differences to solve problems in professional and social activities.	ΠΚ-3	Can effectively solve problems in professional and social activities.	Lectures, in-class exercises, data analysis projects
8. Ability to determine and transmit common goals in professional and social activities.	ΠΚ-4	Can effectively determine and transmit common goals in professional and social activities.	In-class exercises, data analysis projects
9. Ability to build professional activity and business, make choices, guided by the principles of social responsibility.	ΠΚ-7	Can build professional activity, business, and make choices on the principles of social responsibility.	Lectures, readings, in-class exercises, data analysis projects
10. Ability to generate	ΠΚ-8	Can effectively generate	Lectures, readings, in-

Competencies	NC/NRU- HSE Code	Descriptors - the learning outcomes (the indicators of achievement)	Teaching forms and methods of that contribute to the development of a competence
fundamentally new ideas and products, to hold creativity and initiative.		new ideas and products in their professional activities.	class exercises, data analysis projects

VII. Format and Procedures

This course will emphasize preparation for each class period and will involve a high level of class participation. Often, experiential exercises, simulations, and video segments (methodological course webcasts) will be used to illustrate key concepts. Throughout the course, students will be required to complete several practical assignments, utilizing a variety of statistical instruments such as R, Stata, or SAS.

Course progression

Most of the course topics are already set, but the rate of the course progression will depend on your level of comprehension and will be adjusted as necessary.

Messages and Memos for Me

If you have any messages or specific requests for me, please submit them by e-mail to vkuskova@hse.ru ONLY. I will not respond to emails sent to any other address. Ensure that your message includes your name, a complete description of your concern, and a recommendation for resolution.

Stay Informed about Class Schedules & Policies

It is the student's responsibility to stay informed about class schedules and policies. The information you need is included in the MASNA student resources. In addition, announcements will be made regularly in class and on website, and it is your responsibility to keep up with that information. If you are unclear about any policies or other information, please ask promptly. Don't wait and get an unpleasant surprise later.

Participation Ground Rules

In an effort to provide a classroom environment as conducive to learning as possible, the following ground rules should be observed:

1. *Confidentiality.* Concepts and ideas can be taken from the class and discussed freely. However, personal stories or issues raised by individuals are to be kept confidential and as the property of the class.
2. *Respectful Listening.* When differing with another participant's point of view, listen first before raising questions. When another participant raises a point we disagree with or find offensive, it is important to remember that the human being behind that question or comment deserves respect. Please freely utilize the concepts we'll learn in the second week of class.
3. *Participation.* Participants who tend to be quieter are encouraged to contribute to enhancing the learning process by sharing their perspectives and experiences. Those who are aware they are prone to monopolizing discussions are encouraged to self-monitor their behavior and make room for quieter students.

4. *No Zaps*. In keeping with the notion of respectful listening, "putting-down" others in class is discouraged. "Zapping" another person often serves to discourage open and honest exchange of ideas among the whole group.

Homeworks

In this class, homeworks (designed as projects) are essential for learning. As Arnold Schwarzenegger once put it, "Nobody ever got muscles watching me lift the weights!" In a similar manner, simply reading good articles is not enough – you also have to practice your writing skills, and improve on your style by getting feedback for your writing. Projects will be assigned as needed, and specific project requirements and deadlines will be provided with the project descriptions.

Homeworks and Projects must be turned in by midnight of the due date. All work is due by 23:59 pm; work submitted at 12:01 am (1 minute past due date) or later is considered late. Due dates will be clearly posted, and can be handed on paper in class or by midnight electronically.

Copyright Notice

All handouts in this course are copyrighted, including all materials delivered electronically. "Handouts" refers to all materials generated for this class, which include but are not limited to the syllabus, class notes, quizzes, exams, lab problems, in-class materials, review sheets, and additional problem sets. You have the right to download materials from the course website for your own use during this class; however, because these materials are copyrighted, you do not have the right to copy the handouts for other purposes unless the instructor expressly grants permission.

Class preparation

Considering that class preparation is a personal matter and that there is no one formula, the following are some generally recommended guidelines for most cases:

1. Read the assigned material quickly, noting the major issues and a general sense of the layout. Read to get a sense of what the reading is about. Ask yourself how you can relate to the materials covered, and whether all the new terms make sense.
2. Reread the material carefully, annotating, highlighting and distinguishing important information, omissions, and questions raised by the reading.
3. Decide what the most important issues are. Write down questions you don't understand.
4. Discuss the reading with others, before class if possible, to test out your ideas and further your understanding of the issues.
5. Prepare notes to guide your class participation, including: answering assigned questions, summary of the main issue(s), further questions raised by the reading, assumptions made by the chapter, your personal experiences, and possible approaches or solutions to any problems assigned with the material.
6. Always do the assigned homework.

VIII. Assumptions

We expect you to have a solid command of the English languages, enough to read and comprehend academic literature and follow the writer's style. We also expect you to have solid writing skills at a level sufficient for writing complex academic papers. It will be very difficult to succeed in this class otherwise.

IX. Course Requirements

- a. **Class attendance and participation policy**

You are responsible for attending class. If you miss class, you are still responsible for everything covered in class, including announcements. Absences excuse you (the body) **NOT ANY WORK THAT IS DUE**, even if excuse is documented. Failure to turn in assignments on time will result in a loss of participation/responsibility points, and a zero on the assignment. Similarly, being absent does not excuse you from obtaining handouts and assignments that you may have missed. It is your responsibility to find out what you have missed and to make arrangements to obtain any handouts, assignments, etc. If you are going to be absent or late, then make sure you e-mail or have someone deliver your homework. All work is due before class begins. Exception: religious observance absences as stated below.

When will an excuse be counted as documented? Documented excuses must include: your name, the date(s) of your absence, the reason for the absence, the (legible) name of the person authorizing the excuse, and his/her phone number for verification. If the excuse is for an illness, the date you are allowed to return to school should be included. The excuse must cover the date(s) of your absence. I reserve the right to check excuses for authenticity. Attempts to use forged or invalid excuses will be treated as a case of scholastic dishonesty (See Academic Honesty, below). Please provide documentation of any excused absences within two class periods of returning to class. Please do not wait until the end of the semester to do so; I will reserve the right to refuse any documented excuses submitted to me outside the four-class window.

Why should you bother to attend every class? Most of the new material in this course is introduced and explained **IN CLASS**. You can't learn if you are not there.

Make-up policy

I allow you to drop one lowest grade on homeworks, so no make-ups will be allowed. Exception: religious observances absences as stated below.

Religious Observance

Accommodations will be made for observance of religious holidays, if they are not scheduled as university-wide holidays. I require that you request accommodations in advance, in writing, clearly stating the date(s) of absence, or tasks that you are unable to complete because of religions restrictions (e.g., inability to use technology). Requests for accommodations must be submitted to me no later than Monday of the 2nd week of classes at class time.

b. Course readings:

Please note: whenever possible without violating copyright restrictions, reading materials will be provided in printed or electronic format. There is no textbook for the course; there are many textbooks and I will provide them for you in one form or the other.

Grading Procedures

c. Course assignments and projects:

Assignments	Percent of Grade
#1: Projects	80%
#2: Readings summaries	20%

X. Academic Integrity:

- a. Each student in this course is expected to abide by the Higher School of Economics' Academic Honesty Policy. Any work submitted by a student in this course for academic credit will be the student's own work. As this course requires lots of writing, it is **mandatory** that no part of your work be copied from other sources without being taken in quotes and properly documented. Documentation of the cited literature is an essential skill learned in this course.
- b. You are encouraged to study together and to discuss information and concepts covered in lecture and the sections with other students. You can give "consulting" help to or receive "consulting" help from such students. However, this permissible cooperation should never involve one student having possession of a copy of all or part of work done by someone else, in the form of an e-mail, an e-mail attachment file, a diskette, or a hard copy. Should copying occur, both the student who copied work from another student and the student who gave material to be copied will both automatically receive a zero for the assignment. Penalty for violation of this Policy can also be extended to include failure of the course and University disciplinary action.

XI. Accommodations for Students with Disabilities

NRU Higher School of Economics is committed to ensuring equal academic opportunities and inclusion for students with disabilities based on the principles of independent living, accessible universal design and diversity. I am available to discuss appropriate academic accommodations that may be required for student with disabilities. Requests for academic accommodations are to be made during the first three weeks of the semester, except for unusual circumstances. Students are encouraged to register with Disability Services Center to verify their eligibility for appropriate accommodation

XII. Course Schedule

The topics to be covered in course are listed below. Assigned readings and corresponding assignments, are provided in a separate document and are subject to change throughout the semester. Please check your seminar folder frequently. The dates are not fixed, but it is expected that we will be moving at the rate of a topic (or one part for two-part sequences) per week. In the event that we are progressing faster or slower, you will be notified of all the changes in advance.

Topics covered:

1. Introductions, syllabus & course requirements. Lecture: Overview of the field and all about data
2. Collecting social data: projecting real life into numbers. Scaling procedures: Issues and applications
3. Summated rating scale construction: an Overview of Lisrel for covariance structure analysis
4. Scaling procedures: Issues and applications (continued)
5. Missing data and other data issues
6. Bootstrap, fake data, and other things that go bump in the night
7. All about data: conclusion
8. Introduction to causal inference
9. Causal inference: instrumental variables

10. Spatial Data Analysis
11. Prediction
12. Conclusion: overview of the field