**Санкт-Петербургский филиал федерального государственного автономного образовательного учреждения высшего образования "Национальный исследовательский университет "Высшая школа экономики"**

Факультет Санкт-Петербургская школа социальных наук и востоковедения Национального исследовательского университета «Высшая школа экономики»

Департамент социологии

# Рабочая программа дисциплины
# Байесовская статистика (преподается на английском языке)

для образовательной программы «Современный социальный анализ»
направления подготовки 39.04.01 "Социология"
уровень магистратура

Разработчики программы
Соколов Б.О., к. полит. н., доцент, bssokolov@hse.ru

Согласована методистом ОСУП
«30» августа 2018 г.

Т.Г. Ефимова _____

Утверждена Академическим советом образовательной программы

«30» августа 2018 г., № протокола ___1____

Академический руководитель образовательной программы

Е.Л. Омельченко _____

Санкт-Петербург, 2018

*Настоящая программа не может быть использована другими подразделениями университета и другими вузами без разрешения кафедры-разработчика программы.*

# Аннотация

| | | | |
|---|---|---|---|
| Название дисциплины | **Байесовская статистика (преподается на английском языке)** | | |
| Образовательная программа | **«Современный социальный анализ»** | | |
| Тип дисциплины | **По выбору** | | |
| Требования к уровню знаний студентов, необходимых для освоения дисциплины (пререквизиты) | Количественный анализ социологических данных | | |
| Объем з.е. | **3** | | |
| Объем в часах | **Аудиторная работа** | **Самостоятельная работа** | **Всего** |
| | **28** | **86** | **114** |
| Краткое описание курса | Данный курс представляет собой введение в основные принципы байесовской статистики. В рамках курса рассматриваются такие темы, как теорема Байеса для дискретных и непрерывных распределений, вычисление параметров по методу Монте-Карло цепей Маркова и диагностика апостериорных выборок, линейные и обобщённые линейные байесовские регрессии, оценка качества подгонки байесовских моделей, байесовское усреднение моделей. Кроме того, обсуждаются различные практические вопросы, связанные с оценкой, визуализацией и представлением байесовских моделей с использованием языка программирования R, в частности пакетов *MCMCpack*, rjags и *rstanarm*. | | |
| Образовательные результаты по дисциплине | Студент в результате освоения дисциплины:<br>• Разбирается в основных принципах байесовской статистики<br>• Умеет применять изученные методы в собственной исследовательской/аналитической деятельности<br>• Использует среду программирования R, в частности пакеты *lavaan*, *semPlot* и *semTools*, для выполнения комплексных статистических расчётов | | |
| Краткое содержание дисциплины | Понятие условной вероятности. Теорема Байеса для дискретных и непрерывных распределений. Правдоподобие модели, априорные и апостериорные распределения параметров. Вычисление параметров по методу Монте-Карло цепей Маркова и диагностика апостериорных выборок. Линейные и обобщённые линейные байесовские регрессии: интерпретация параметров и статистический вывод. Оценка качества подгонки байесовских моделей: апостериорный предиктивный анализ. Байесовские процедуры отбора наилучшей модели: фактор Байеса, информационные критерии, кросс-валидация. Байесовское усреднение моделей | | |
| Образовательные технологии | В рамках данного курса задействуются следующие образовательные технологии: | | |

| | |
|---|---|
| | • Метод проектов<br>• Исследовательский метод (количественный и качественных анализ реальных политологических данных) |
| Формы контроля | Итоговая оценка за курс вычисляется согласно следующей формуле:<br><br>$$\mathbf{O_{итог} = 0{,}65 * O_{накопленная} + 0.35 * O_{проект}\ (2)}$$<br><br>Где $\mathbf{O_{накопленная}}$ – накопленная оценка за курс, а $\mathbf{O_{проект}}$ – оценка за итоговый экзамен (проходящий в форме защиты индивидуального исследовательского проекта). $\mathbf{O_{накопленная}}$ рассчитывается следующим образом:<br><br>$$\mathbf{O_{накопленная} = 0.5 * O_{дз} + 0.1 * O_{класс} + 0.4 * O_{презентация}\ (1)}$$<br><br>В вышеприведённой формуле $\mathbf{O_{дз}}$ – среднее арифметическое оценок за домашние задания (письменные задания по анализу количественных данных), $\mathbf{O_{класс}}$ – оценка за аудиторную работу на протяжении всего курса, и $\mathbf{O_{презентация}}$ – оценка за промежуточную презентацию индивидуального исследовательского проекта. Все оценки в рамках данного курса выставляются по шкале 1-10 баллов. |
| Литература | <u>Основная</u><br>Müller P. ,Andres F. (2015) Bayesian Nonparametric Data Analysis https://link.springer.com/book/10.1007%2F978-3-319-18968-0#about<br><u>Дополнительная</u><br>Kaeding M. (2015) Bayesian Analysis of Failure Time Data Using P-Splines https://link.springer.com/book/10.1007%2F978-3-658-08393-9#about<br>John Kruschke (2011) *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*. Academic Press https://library.books24x7.com/toc.aspx?bookid=56584 |
| Преподаватель | Соколов Б. О, доцент, к. полит. н., bssokolov@hse.ru |

# 1 Normative references

This syllabus outlines minimum requirements, objectives, learning outcomes, and content of the course.

The syllabus is designed for the instructors teaching this discipline, their teaching assistants, and the students of the Master Program “Modern Social Analysis” who have taken the course “Bayesian Statistics (offered in English).”

The syllabus complies with the following documents:

- The Educational Standard of the National Research University “Higher School of Economics” for the master program “Modern Social Analysis”;
- The Educational Program “Sociology” for the master program “Modern Social Analysis”;
- The University Curriculum of the National Research University “Higher School of Economics” for students of the master program “Modern Social Analysis”

# 2 Course overview

Bayesian data analysis is a rapidly developing field of statistics, which has many useful applications in various areas of political science, sociology, and international relations. The goal of this course is to provide a brief and “mostly harmless” (that is, as informal as possible) introduction to the theory and application of Bayesian statistical methods. The course begins with the basic concepts of Bayesian statistics (e.g., Bayes’s rule. priors, likelihood, and posterior distribution). Then we consider various approaches to the estimation and assessment of Bayesian models (with most attention to the MCMC-based methods) in the context of generalized linear models. Next we learn about main Bayesian approaches to model selection, including Bayes factors, DIC, and cross-validation methods. We conclude by discussing Bayesian model averaging (BMA), a powerful Bayesian approach to reducing model specification uncertainty.

Students are assumed to have basic knowledge of statistics and be familiar with several conventional statistical methods, most importantly regression analysis. Knowledge of advanced topics, such as multilevel regression analysis and maximum-likelihood estimation, is helpful, but not critical. In addition, for practical exercises we will use R programming environment, so another major prerequisite is a basic knowledge of R.

# 3 Intended Learning Outcomes

After completing this course, a student will be able to:

- Understand the basic principles of Bayesian analysis, the opportunities which this statistical method offers for social scientists, and its limitations.
- Apply Bayesian methods to the analysis of real data sets.
- Properly report the results of Bayesian analysis in research papers

The course develops the following competences:

| Компетенция | Код по ОС ВШЭ | Уровень формирования компетенции | Дескрипторы – основные признаки освоения (показатели достижения результата) | Формы и методы обучения, способствующие формированию и развитию компетенции | Форма контроля уровня сформированности компетенции |
|---|---|---|---|---|---|
| Способен рефлексировать (оценивать и перерабатывать) освоенные науч- | СК-1 | РБ | Магистрант задает вопросы, сравнивает различные способы анализа, пробует выполнить задание разными методами | Семинары, самостоятельная работа на индивидуальным проектом | Аудиторная работа, домашние задания |

| Компетенция | Код по ОС ВШЭ | Уровень формирования компетенции | Дескрипторы – основные признаки освоения (показатели достижения результата) | Формы и методы обучения, способствующие формированию и развитию компетенции | Форма контроля уровня сформированности компетенции |
|---|---|---|---|---|---|
| ные методы и способы деятельности | | | | | |
| Способен предлагать концепции, модели, изобретать и апробировать способы и инструменты профессиональной деятельности | СК-2 | РБ | Магистрат интерпретирует эмпирические данные, применяет разработанные концептуальные модели и предлагает свои объяснительные схемы. | Семинары, самостоятельная работа на индивидуальным проектом | Аудиторная работа, домашние задания |
| Способен к самостоятельному освоению новых методов исследования, изменению научного и научно-производственного профиля своей деятельности | СК-3 | РБ | Магистрант сначала изучает разнообразные исследовательские дизайны на примере работ, опубликованных в ведущих социологических журналах, а затем предлагает свои способы операционализации теоретических проблем и конструирует адекватные концептуальным задачам аналитические модели | Семинары, самостоятельная работа на индивидуальным проектом | Аудиторная работа, домашние задания |
| Способен принимать управленческие решения, оценивать их возможные последствия и нести за них ответственность. | СК-5 | РБ | Магистрант демонстрирует способность принимать управленческие решения, оценивать их возможные последствия и нести за них ответственность. | Семинары, самостоятельная работа на индивидуальным проектом | Аудиторная работа, домашние задания |
| Способен вести профессиональную, в том числе научно-исследовательскую деятельность в международной среде | СК-8 | РБ | Магистрант использует иностранную литературу и иностранный опыт при проведении исследований | Семинары, самостоятельная работа на индивидуальным проектом | Аудиторная работа, домашние задания, экзамен |
| Способен свободно общаться на иностранных языках для целей профессионального и научного общения | ПК-8 | РБ | Магистрант свободно читает обязательную и дополнительную литературу на английском языке; способен поддержать дискуссии на английском языке в рамках индивидуальной и групповой работы на семинарских занятиях. | Участие в дебатах на семинарских занятиях, самостоятельная работа с литературой, подготовка устных докладов на семинарских занятиях и письменных работ | Аудиторная работа, домашние задания, экзамен |
| Способен использовать в профессиональной деятельности основные требования информационной безопасности, в том числе в об- | ПК-10 | РБ | Магистрант демонстрирует способность использовать в профессиональной деятельности основные требования информационной безопасности, в том числе в области защиты государственной безопасности. | Семинары, самостоятельная работа на индивидуальным проектом | Аудиторная работа, домашние задания |

| Компетенция | Код по ОС ВШЭ | Уровень формирования компетенции | Дескрипторы – основные признаки освоения (показатели достижения результата) | Формы и методы обучения, способствующие формированию и развитию компетенции | Форма контроля уровня сформированности компетенции |
|---|---|---|---|---|---|
| ласти защиты государственной безопасности | | | | | |

## 4 Prerequisites

This course is an elective course for students of the master program 39.04.01 “Comparative Social Research”. Taking this course requires the successful prior completion of the following disciplines:

- Quantitative data analysis

To succeed in this course, students are expected to possess the following knowledge and competences:

- The ability to develop new concepts and models and apply them in professional work (GC-M2);
- The ability to work with information and existing literature (PC-11/IC-M4);
- The ability to speak a foreign (English) language at a level sufficient for informal communication as well as for the search and analysis of foreign sources of information (PC-8/IC-M2).

The course is strongly related and complementary to other compulsory and elective courses and provides crucial prerequisites for later courses and research projects as well as skills necessary for the master thesis. The course takes place in the first module of the second year of the program, giving students the important skills in designing and conducting their own research as well as assessing the quality of research projects published in a series of working papers and peer-reviewed social science journals.

## 5 Content and Structure of the Course

This course is worth 3 credit units.

| **№** | **Topic** | **Total Hours** | **Directed Study** | | **Self-Directed Study** |
|---|---|---|---|---|---|
| | | | *Lectures* | *Tutorials* | |
| 1 | Introduction. Basic concepts of Bayesian analysis | 18 | 0 | 6 | 14 |
| 2 | General principles of Bayesian inference. Priors and likelihood | 24 | 0 | 6 | 16 |
| 3 | Bayesian model estimation: Gibbs sampling and Hamiltonian Monte Carlo | 18 | 0 | 4 | 14 |
| 4 | Bayesian model evaluation. Posterior predictive checks. | 18 | 0 | 4 | 14 |
| 5 | Bayesian model comparison | 18 | 0 | 4 | 14 |
| 6 | Bayesian model averaging | 18 | 0 | 4 | 14 |
| **Total volume of hours** | | **114** | **0** | **28** | **86** |

# 6 Indicative Assessment Methods and Strategy

| Type of Control | Form of Control | Detailed explanation |
|---|---|---|
| Weekly | Home assignments | Students are expected to complete written home assignments outside of class time. Some assignments are conceptual, but some of them require using R and empirical data, therefore students should plan to spend time in the computer lab outside of regular class hours. All assignments are due at the beginning of the following class time |
| | Class work | Students are expected to participate in all class activities, such as group discussions, responses to the instructor’s questions, and completion of simple conceptual and programming exercises. |
| | Progress reports | Students are expected to write a short research paper (see below). Throughout the course they will have multiple opportunities to present their progress. It is not mandatory to present each class day (except on mid-term presentation day), but frequent presentations may improve personal cumulative grade (conditional on their overall quality) and will definitely help in the preparation of the final paper. |
| Final | Exam | All course participants must write a short research paper (5-15 pages) in which they will try to apply Bayesian methods to the topic in modern social research that they are interested in. This paper must follow the model of English-language published scientific journal articles. The most important aspects of the paper to be graded are the creativity of the research idea, the operationalization and proper statement of hypotheses, and the appropriate use of Bayesian statistics. See detailed final paper requirements below, in Section 10.<br><br>Co-authored projects are encouraged, but, with a co-authored project, I will have higher expectations on the paper and presentation. Final versions should be submitted not later than noon of October 20nd, 2018. Then all papers will be evaluated by the instructor and preliminary grades will be announced. Then those course participants who disagree with their preliminary grades will have an opportunity to defend their project papers (by publicly presenting them in class). Presentations must be given using PowerPoint or LaTeX software. |

# 7 Assessment criteria

### 7.1 Class work:

You may expect an excellent grade (8-10 on a 0-10 scale) for class work if you actively participate in in-class discussions and other activities, frequently present your homework and progress reports, and demonstrate a good performance when responding to the instructor’s conceptual questions and performing simple in-class programming exercises. The general rule is the lower your activity the lower your grade, though it is worth noting that the overall quality of your in-class performance also matters in this respect. If you make mistakes regularly when responding to my questions, performing exercises, or presenting your own work, it will negatively affect your grade.

Please notice that attendance is not obligatory, so you can safely miss a few classes. However, if you miss all classes, your grade for class work will be zero. Also keep in mind that if you miss class you are still responsible for everything covered in class, including announcements. 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.

| Grades | Assessment Criteria |
|---|---|
| «Excellent» (8-10) | A critical analysis which demonstrates original thinking and shows strong evidence of preparatory research and broad background knowledge. |
| «Good» (6-7) | Shows strong evidence of preparatory research and broad background knowledge. Excellent oral expression. |
| «Satisfactory» (4-5) | Satisfactory overall, showing a fair knowledge of the topic, a reasonable standard of expression. Some hesitation in answering follow-up questions and/or gives incomplete or partly irrelevant answers. |
| «Fail» (0-3) | Limited evidence of relevant knowledge and an attempt to address the topic. Unable to offer relevant information or opinion in answer to follow-up questions. |

**7.2 Homework assignments;**

Half of your written home assignments will be conceptual assignments and the other half will be related to preparation of your final paper.

You may expect an excellent grade (8-10 on a 0-10 scale) for conceptual assignments if you are able to (a) give correct answers to the stated questions, (b) write (if necessary) correct R code, and (c) interpret properly the results of Bayesian analyses done by other researchers, and also (d) demonstrate a proper understanding (and usage) of relevant Bayesian terminology. As to conceptual assignments, I encourage you to work in groups on the homework, but you always need to write your own solutions including your computer code. Also, it is hugely beneficial to attempt the problems sets on your own before working in groups.

| Grades | Assessment Criteria |
|---|---|
| «Excellent» (8-10) | Has a clear argument, which addresses the topic and responds effectively to all aspects of the task. Fully satisfies all the requirements of the task; rare minor errors occur; |
| «Good» (6-7) | Responds to most aspects of the topic with a clear, explicit argument. Covers the requirements of the task; may produce occasional errors. |
| «Satisfactory» (4-5) | Generally addresses the task; the format may be inappropriate in places; display little evidence of (depending on the assignment): independent thought and critical judgement include a partial superficial coverage of |

| | |
|---|---|
| | the key issues, lack critical analysis, may make frequent errors. |
| «Fail» (0-3) | Fails to demonstrate any appropriate knowledge. |

**7.3 Progress reports and mid-term presentation**

Progress reports are home assignments related to the preparation of your final paper. After each lecture you are expected to apply learned new methods to your own data and statistical model and written a short report describing your results.

The most important aspects of progress reports to be graded are: (a) correct application of Bayesian methods to your data and model (0-3 points), (b) specific Bayesian terminology (0-3 points), as well as (c) the ability to interpret the results of your analyses correctly (0-3 points).

Notice that (f) style and formatting issues (e.g. correct formatting of tables, figures, in-text citations, and references) also affect the final grade (0-1 points). I do not expect that your style and grammar will be perfect, but I should be able, at least, to understand from the text of your assignment what exactly you have done.

On Wednesday, September 26$^{th}$, all students must make a [mid-term] presentation of their progresses. At the presentation, you will be asked to publicly present (with main focus on research design, operationalization, and the substantive interpretation of the results). You will have about 10 minutes to tell the audience about your progress (in English). You will get a higher grade for that presentation if you are able to successfully demonstrate abilities to (a) formulate an original research question, develop a theoretical model, and properly operationalize it; (b) correctly apply Bayesian methods to your data and model; and (c) interpret the results properly, as well as (d) good presentation skills.

| Grades | Assessment Criteria |
|---|---|
| «Excellent» (8-10) | A well-structured, analytical presentation of project work. Shows strong evidence and broad background knowledge. In a group presentation all members contribute equally and each contribution builds on the previous one clearly; Answers to follow-up questions reveal a good range and depth of knowledge beyond that covered in the presentation and show confidence in discussion. |
| «Good» (6-7) | Clearly organized analysis, showing evidence of a good overall knowledge of the topic. The presenter of the project work highlights key points and responds to follow up questions appropriately. In group presentations there is evidence that the group has met to discuss the topic and is presenting the results of that discussion, in an order previously agreed. |
| «Satisfactory» (4-5) | Takes a very basic approach to the topic, using broadly appropriate material but lacking focus. The presentation of project work is largely unstructured, and some points are irrelevant to the topic. Knowledge of the topic is limited and there may be evidence of basic misunderstanding. In a group presentation, most of the work is done by one or two students and the individual contributions do not add up. |
| «Fail» (0-3) | Fails to demonstrate any appropriate knowledge. |

**7.4 Final paper**

See detailed info below, in Section 10.5 of this syllabus.

# 8 Detailed Course Content

**<u>Day 1. Introduction (Wednesday, Sept 5, room 210, 55/2 Sedova str.).</u>**

***Content***: frequentist vs. epistemic concepts of probability, Bayes’s rule, prior and posterior distributions, likelihood, discrete probability examples, simple continuous distributions examples, popular R packages for Bayesian analysis.

**Day 2. General principles of Bayesian inference (Sept 12, room 210, 55/2 Sedova str.).**

***Content***: Bayesian linear regression, choice of priors, interpretation of model parameters, Bayesian inference, credibility intervals, Bayesian generalized regression modeling, examples in *MCMCpack* and *rstanarm*.

**Day 3. MCMC estimation (Sept 19, room 210, 55/2 Sedova str.).**

***Content***: MCMC estimation, Gibbs sampling, Hamiltonian Monte-Carlo, main convergence diagnostics, INLA, variational inference.

**Day 4. Mid-term presentations of research projects (Sept 26, room 210, 55/2 Sedova str.)**

We will devote this day to the detailed discussion of your progress with your research projects.

**Day 5. Bayesian model evaluation (Oct 3, room 210, 55/2 Sedova str.).**

***Content***: Posterior predictive distribution, posterior predictive checks, posterior predictive P-value, visual checks.

**Day 6. Model selection and comparison (Oct 10, room 210, 55/2 Sedova str.).**

***Content***: Bayes factors, Bayesian Information Criterion (BIC) and Deviance Information Criterion (DIC), WAIC, leave-one-out cross-validation.

**Day 7. Bayesian model averaging (Oct 17, room 210, 55/2 Sedova str.).**

***Content***: What is BMA, why it can be useful for social scientists, model priors selection, most popular BMA algorithms and their implementation in R.

# 9 Educational techniques

The goal of this course above and beyond teaching specific methods is to enable students to use the methods covered in the course on a stand-alone basis whenever they need this in the future. Therefore, every reasonable effort should be made to make the material understandable and comprehensible, depending on the level of the student.

It is also crucially important to encourage those students who have already understood new material to share their understanding with the others has demonstrated rewarding results.

Regular Q&A sessions are needed at the end of each session, both on lectures and on practical exercises.

The general recommendation is to put emphasis on training the skills to perform the same types of analysis autonomously; therefore, the more time students get to practice their data analysis skills on different data sets, the more reliable the success of the course.

## 9.1 Recommendations to the Instructor

Try using as many ways of approaching students as possible. Since the knowledge and readiness to learn in English is non-equal among students, be always prepared to stratify exercises as well as theory for different levels. Find and use the youtube.com tutorials on R and lavaan. Small groups are always encouraged whenever possible and reasonable.

## 9.2 Recommendations to the Students

The key goal of this course is to equip you with a toolkit of state-of-the-art methods of statistical analysis especially relevant for the purposes of comparative social research. You are offered an array of methods of data analysis which you are likely to use while staying in the social sciences and beyond.

To learn these methods successfully, do not hesitate to pose your questions to the instructor. Additionally, try keeping a vocabulary on each topic covered, with the most important terms explained with examples. This will help you at the exam and in the future as your personal reference book. In your free time, read and watch as much about the topic as possible. Having read the same topic from several textbooks and also websites (including Bayesian methods-related Wikipedia articles!) as a rule improves your understanding substantially.

Watch and learn extra beyond the classes. Look through the pages of the recent issues of such journals as the *Sociological Methodology*, *Sociological Methods & Research*, *Political Analysis*, *American Sociological Review*, *American Political Science Review*, *Survey Research and Methods*, *Political Science: Research and Methods*, etc., to learn how to report the results of Bayesian analysis. Complete online Bayesian courses on Coursera.org, DataCamp.com, edX.org, or any other online learning platform to recap on basic Bayesian concepts and methods if you feel you could benefit from it.

Do not wait until the end of the course to see me regarding problems with course materials or your performance (it will be too late to address deficiencies at the end of the semester). If you are aware that you must achieve a particular grade in this course, please see me during the first week of the course. This will allow me to alert you of deficiencies in your performance. There is nothing that either of us can do at the end of the course

## 10 Final Paper Requirements

Note: (This list of recommendations has been built upon a framework developed by A. Almakaeva, M. Tawat, and S. Spencer: see Section 12 of this file: https://www.hse.ru/data/2017/04/27/1168759732/program-1711306534-5jHlOiLKAq.pdf).

All course participants also must write a short research paper (5-15 pages) in which they will try to apply Bayesian methods to the topic in cross-cultural social research that they are interested in. The most important aspects of the paper to be graded are the creativity of the research idea, the operationalization and proper statement of hypotheses, and the appropriate use of Bayesian statistics.

### 10.1 Format

MS Word or LaTeX, 5-15 pages long, typed, double-spaced, standard margins. Notice that page count does not include abstract, references, and appendices, as well as figures and tables.

### 10.2 Purpose

This paper is an academic research paper presenting the results of your secondary data analysis. This paper should demonstrate your 1) logical reasoning, 2) your skill at interpreting the results of structural equation modeling, and 3) your ability to communicate in academic writing style.

### 10.3 Content

This paper, like all academic research papers, must contain the following sections:

1. Abstract
2. Introduction
3. Literature review
4. Data and Methods
5. Findings

6. Discussion/conclusion
7. References.
8. *Appendix (-ces) (If necessary)*

ABSTRACT: A one-paragraph summary of the research question and (only) main findings (on a separate page, not counted in the page count).

INTRODUCTION: In this section you formulate the research question and establish its scientific relevance (i.e. explain “why it is important to study this topic”, may also include social or policy relevance).

LITERATURE REVIEW: This section examines your research question in terms of the theory that generated it, and reviews existing sociological research addressing the question, including research that may be only partially related. If your specific topic appears to be understudied, this section should address what is available on related topics. The literature review generally includes a mention of how the current research replicates previous research, contradicts previous research, or somehow modifies or extends previous research. At the end of this section, you must clearly state the hypothesis or hypotheses to be tested in subsequent empirical analysis. It should be obvious how the hypotheses are related to the theoretical background outlined in the literature review.

For the length of this paper, your literature review must make reference to at least five (5) recent academic peer-reviewed journal (WoS Q1) articles on your topic.

DATA & METHODS: This section briefly describes the dataset and analytical methods that you use. This section should (a) explain how the research question is operationalized into testable hypotheses; (b) clearly state the concepts to be tested in the hypotheses, as well as label the independent variables, the dependent variable, and any intervening or control variables that are included; (c) describe which variables are measuring which concepts and from which sources data on these variables come; and (d) identify which statistical methods are used to analyze data. Notice that the list (table) of descriptive statistics on all variables has to be reported in the Appendix.

FINDINGS: This is the section in which you present your findings and explain the results of your statistical tests. Name the test or procedure used to obtain each result (i.e., whether it is an outcome of a linear regression model or some other method). This section should also address whether the analysis of data confirms your hypotheses.

DISCUSSION/CONCLUSION: This section should *briefly* summarize the findings. It should also explain how your findings contribute to the literature reviewed at the beginning of the paper. Discuss whether your findings support or contradict previous research. This section should also briefly discuss limitations of your empirical analysis (especially methodological and data-related ones): for instance, evaluate whether [some of] your concepts could have been measured differently or different tests could be run to answer your substantive research questions. This section may include suggestions for future research or implications for policy or both, but offering future research or policy implications is not a mandatory task.

REFERENCES: All references cited in the text must be listed in the bibliography according to either ASA format, or APA format, or Chicago Manual of Style. You paper should refer to at least 10 peer-reviewed articles published in the top social sciences journals but may also refer to additional sources, such as prominent sociological books, electronic resources, data sets, contributions to edited volumes, etc. For all in-text citations use the author-date format, e.g. (Welzel 2013) or (Inglehart, Ponarin, and Inglehart 2017).

APPENDIX: This section contains all tables and figures referenced in the main text, as well as describe (if necessary) all additional tests and procedures conducted to check the robustness of your main findings under different model assumptions. Notice that replication data (i.e. your data set, as well as an R script used to produce all figures, tables and other results reported in the main text and the appendix) are also worth to be submitted along with the final paper.

***Note**: appendix is not counted in the page count.*

### 10.4 Style and presentation

Prepare tables and figures carefully. Look through the pages of the recent issues of such journals as *Sociological Methodology*, *Political Analysis*, *Sociological Methods & Research*, etc., to learn how to report the results of Bayesian analysis. Remember to include replication code in an appendix. The paper should be accompanied by a Power Point (or equivalent, preferably LaTeX) presentation of the paper. At the **defense**, you will be asked to publicly present the design and the key results of your research. You will have about 10 minutes to tell the audience about your research (in English). You will get a higher grade if you are able to successfully demonstrate (a) a deep understanding of the method, (b) proficient skills in using Bayesian methods to analyze real data and ability to interpret the results of Bayesian analysis correctly, as well as (c) good presentation skills.

### 10.5 Grading Criteria

The most important aspects of your final papers to be graded are: (a) ability to formulate an original research question which at the same time is related to ongoing theoretical debates (0-2 points); (b) ability to operationalize your theoretical/conceptual model using relevant data (0-2 points); (c) appropriate use of Bayesian methods (0-2 points) and (d) specific Bayesian terminology (0-1 points), as well as (e) the ability to interpret the results of your analyses correctly (0-2 points).

Notice that (f) style and formatting issues (e.g. correct formatting of tables, figures, in-text citations, and references) also affect the final grade (0-1 points). I do not expect that your style and grammar will be perfect, but I should be able, at least, to understand from the text of your assignment what exactly you have done,

Final project paper can be written alone or in collaboration with another student (three-author papers are not allowed). If you are working together, please indicate clearly in your reports and presentations, who did what part of the work. Notice that unless a group project deserves an excellent grade (8-10), it will get a final grade one point lower than an individual project of equivalent quality. If you plagiarize, you will fail. The deadline for submitting final papers is noon of October 20$^{th}$, 2018. Late submissions will be graded down (one point on a 1-10 scale per day of delay; papers submitted with a delay of three days or more will be penalized by 3 points, irrespective of the length of delay)

| Grades | Assessment Criteria |
|---|---|
| «Excellent» (8-10) | A well-structured, analytical presentation of project work. Shows strong evidence and broad background knowledge. In a group presentation all members contribute equally and each contribution builds on the previous one clearly; Answers to follow-up questions reveal a good range and depth of knowledge beyond that covered in the presentation and show confidence in discussion. |
| «Good» (6-7) | Clearly organized analysis, showing evidence of a good overall knowledge of the topic. The presenter of the project work highlights key points and responds to follow up questions appropriately. In group presentations there is evidence that the group has met to discuss the topic and is presenting the results of that discussion, in an order previously agreed. |
| «Satisfactory» (4-5) | Takes a very basic approach to the topic, using broadly appropriate material but lacking focus. The presentation of project work is largely unstructured, |

| | |
|---|---|
| | and some points are irrelevant to the topic. Knowledge of the topic is limited and there may be evidence of basic misunderstanding. In a group presentation, most of the work is done by one or two students and the individual contributions do not add up. |
| «Fail» (0-3) | Fails to demonstrate any appropriate knowledge. |

## 11 Grading Components

Your final grade is composed of four components: (a) home assignments (unweighted average), (b) class work, (c) midterm presentation, and (d) final project. The first three components (taken with different weights) define your accumulated grade (which comprises 65% of the overall grade) which is then combined with the score for the final paper (35% of the overall grade).
Formally, your cumulative grade for the course is calculated in the following way:

$$\mathbf{G_{cumulative} = 0.5 * G_{home\ assignment} + 0.1 * G_{class\ work} + 0.4 * G_{mid\text{-}term\ presentation}} \quad \mathbf{(1)}$$

where $G_{home\ assignment}$ is the mean [unweighted] grade for home assignments prepared throughout the course; $G_{class\ assignment}$ is the [single, *post hoc*, on a 1-10 scale] assessment of your overall participation in class activities; and $G_{mid\text{-}term\ presentation}$ is the mean [unweighted] grade for final paper progress reports.
The final (overall) grade for the course (the one you will have in your course records) is calculated in the following way:

$$\mathbf{G_{final} = 0{,}65 * G_{cumulative} + 0.35 * G_{final\ paper}} \quad \mathbf{(2)}$$

where $G_{cumulative}$ is the accumulated grade for the course calculated using the formula above and $G_{final\ paper}$ is the grade for the final research paper.

Both the cumulative grade and the final grade are rounded according to the rules of algebra

All late submissions (both home assignments, progress reports, and the final paper) will be graded down (the maximum grade for a one-day delay is 7 points; this quantity decreases by one point per extra day of delay; all works submitted with a delay of three and more days are evaluated on a 0-4 scale). If you plagiarize, you will fail. However, it is fine to combine your final research project with other class papers (e.g., term paper).

**The minimal passing grade for this course is 4.**

***Re-examination policy***: Students who fail the course are entitled to retake a final exam (precisely speaking, they will have to present and defend an updated and improved version of their final paper). This reexamination option will count only toward the students’ final paper grade (35% of the overall grade for this course). Students’ reexamination grade will not substitute for poor grades earned prior to the final exam in other aspects of the course (home assignments or class work). Therefore, students seeking a minimally passing grade must also turn in home assignments, activity in class, and the paper.

## 12 Readings

Mandatory

Müller P. ,Andres F. (2015) Bayesian Nonparametric Data Analysis
https://link.springer.com/book/10.1007%2F978-3-319-18968-0#about

Optional

Kaeding M. (2015) Bayesian Analysis of Failure Time Data Using P-Splines https://link.springer.com/book/10.1007%2F978-3-658-08393-9#about
John Kruschke (2011) *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*. Academic Press https://library.books24x7.com/toc.aspx?bookid=56584

## 13 Facilities, Equipment and Software

- It is better to have your own personal laptop with R (and maybe Rstudio) software installed (though it is also possible to use university computers for completing home assignments and performing in-class labs). You can read detailed R installation instructions here or simply download the latest version here.

## 14 Special conditions for organization of learning process for students with special needs

The following types of comprehension of learning information (including e-learning and distance learning) can be offered to students with disabilities (by their written request) in accordance with their individual psychophysical characteristics:

1) *for persons with vision disorders:* a printed text in enlarged font; an electronic document; audios(transferring of learning materials into the audio); an individual advising with an assistance of a sign language interpreter; individual assignments and advising.
2) *for persons with hearing disorders: a* printed text; an electronic document; video materials with subtitles; an individual advising with an assistance of a sign language interpreter; individual assignments and advising.
3) *for persons with muscle-skeleton disorders: a* printed text; an electronic document; audios; individual assignments and advising.

If you have questions about the course material, computational issues, or other course-related issues please do not hesitate to contact me via e-mail: bssokolov@gmail.com. Note however that I have a very busy inbox so it may take one or two days for me to respond.

**Good luck and have fun!**