



Правительство Российской Федерации

**Федеральное государственное автономное образовательное учреждение
высшего профессионального образования
"Национальный исследовательский университет
"Высшая школа экономики"**

Факультет социологии

**Программа дисциплины
"Анализ социологических данных - 1"**

для направления 040200.62 «Социология» подготовки бакалавра

Автор программы:

А.Р.Бессуднов, D.Phil., abessudnov@hse.ru

Одобрена на заседании кафедры методов сбора и анализа социологической информации

«__» _____ 20__ г.

Зав. кафедрой И.М.Козина

Рекомендована секцией УМС «Социология» «__» _____ 20__ г.

Председатель В.Г.Ледяев

Утверждена УС факультета социологии «__» _____ 20__ г.

Ученый секретарь Е.В.Надеждина _____ [подпись]

Москва, 2011

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



Область применения и нормативные ссылки

Настоящая программа учебной дисциплины устанавливает минимальные требования к знаниям и умениям студента и определяет содержание и виды учебных занятий и отчетности.

Программа предназначена для преподавателей, ведущих данную дисциплину, учебных ассистентов и студентов направления подготовки бакалавра 040200.62 «Социология». Программа разработана в соответствии с:

- оригинальным образовательным стандартом НИУ ВШЭ по социологии;
- Образовательной программой 040200.62 «Социология»;
- Рабочим учебным планом университета по направлению подготовки «Социология», утвержденным в 2011 г.

1 Цели освоения дисциплины

Целями освоения дисциплины «Анализ социологических данных – 1» являются:

- изучение и практическое освоение базовых методов статистического анализа данных в социальных науках,
- изучение и практическое освоение компьютерных программ, применяемых для статистического анализа данных (пакета SPSS),
- приобретение понимания специфики работы с количественными данными в социальных науках, понимания типов задач, которые могут быть решены с помощью статистических методов.

2 Компетенции обучающегося, формируемые в результате освоения дисциплины

В результате освоения дисциплины студент должен:

- Знать методы статистического анализа данных в пределах программы курса.
- Уметь ставить и понимать социологические задачи, которые могут быть решены с помощью статистического анализа данных; понимать специфику данных, используемых в статистическом анализе.
- Иметь навыки самостоятельного статистического анализа данных на компьютере в программе SPSS.

В результате освоения дисциплины студент осваивает следующие компетенции:

Компетенция	Код по ФГОС/ НИУ	Дескрипторы – основные признаки освоения (показатели достижения результата)	Формы и методы обучения, способствующие формированию и развитию компетенции
Способность использовать основные законы естественнонаучных дисциплин в профессиональной деятельности, применять методы математического анализа и моделирования, теоретического и экспериментального исследования (формируется частично)	ОК-11	- применяет методы статистического анализа данных для исследования общественных закономерностей - интерпретирует результаты статистического анализа данных (статистический вывод)	Лекции, практические занятия, чтение литературы из списка литературы



Компетенция	Код по ФГОС/ НИУ	Дескрипторы – основные признаки освоения (показатели достижения результата)	Формы и методы обучения, способствующие формированию и развитию компетенции
Владение основными методами, способами и средствами получения, хранения, переработки информации, навыки работы с компьютером как средством управления информацией <i>(формируется частично)</i>	ОК-13	- использует компьютер для статистического анализа данных	Практические занятия, выполнение самостоятельных работ
Владение иностранным языком на уровне, достаточном для разговорного общения, а также для поиска и анализа иностранных источников информации <i>(формируется частично)</i>	ОК-15	- использует англоязычные источники информации для поиска информации о статистическом анализе данных - демонстрирует владение английским языком для поиска информации о статистическом анализе данных	Чтение англоязычной литературы из списка литературы
Способность самостоятельно формулировать цели, ставить конкретные задачи научных исследований в различных областях социологии и решать их с помощью современных исследовательских методов с использованием новейшего отечественного и зарубежного опыта и с применением современной аппаратуры, оборудования, информационных технологий <i>(формируется частично)</i>	ПК-2	- демонстрирует умение решать поставленные социологические задачи с помощью анализа данных статистическими методами (такими как дисперсионный анализ, регрессионный анализ, факторный анализ и пр.) - оценивает возможности решения поставленных социологических задач с помощью статистических методов	Лекции, практические занятия, выполнение самостоятельных работ
Способность и готовность участвовать в составлении и оформлении научно-технической	ПК-3	- демонстрирует умение оформлять результаты статистического анализа данных в письменных работах в соответствии со стандартами,	Выполнение самостоятельных работ

Компетенция	Код по ФГОС/ НИУ	Дескрипторы – основные признаки освоения (показатели достижения результата)	Формы и методы обучения, способствующие формированию и развитию компетенции
документации, научных отчетов, представлять результаты исследовательской работы с учетом особенностей потенциальной аудитории <i>(формируется частично)</i>		принятыми в академических публикациях	
Умение обрабатывать и анализировать данные для подготовки аналитических решений, экспертных заключений и рекомендаций <i>(формируется частично)</i>	ПК-8	- владеет методами статистического анализа данных в пределах программы курса - применяет методы статистического анализа данных для анализа социологических проблем	Практические занятия, выполнение самостоятельных работ
Способность и готовность к планированию и осуществлению проектных работ в области изучения общественного мнения, организации работы маркетинговых служб <i>(формируется частично)</i>	ПК-9	- владеет методами анализа данных опросов общественного мнения - корректно интерпретирует результаты опросов общественного мнения	Практические занятия, выполнение самостоятельных работ

3 Место дисциплины в структуре образовательной программы

Настоящая дисциплина является обязательной дисциплиной профессионального цикла бакалаврской программы направления 040100.62 «Социология».

Изучение данной дисциплины базируется на следующих дисциплинах:

- «Теория вероятности и математическая статистика»

Основные положения дисциплины должны быть использованы в дальнейшем при изучении следующих дисциплин:

- «Анализ социологических данных-2»
- «Введение в методы измерения в социологии»
- «Логлинейный анализ в социологии»
- «Методы анализа латентных признаков»
- «Методы классификации в социологии»
- «Методы непараметрической статистики в социологии»
- Научно-исследовательский семинар



4 Тематический план учебной дисциплины

№	Название раздела	Всего часов	Аудиторные часы			Самостоятельная работа
			Лекции	Семинары	Практические занятия	
1	Введение в статистический анализ данных в социологии	18	2	-	6	10
2	Описательная статистика и статистические графики	18	2	-	6	10
3	Статистические гипотезы и их тестирование. Дисперсионный анализ	18	2	-	6	10
4	Таблицы сопряженности и их анализ	18	2	-	6	10
5	Корреляция и парная линейная регрессия	18	2	-	6	10
6	Множественная линейная регрессия	36	4	-	12	20
7	Регрессионные модели для бинарных и категориальных зависимых переменных	36	4	-	12	20
8	Идентификация причинно-следственных связей и регрессионный анализ	18	2	-	6	10
		180	20	-	60	100



5 Формы контроля знаний студентов

Тип контроля	Форма контроля	1 год				Параметры
		1	2	3	4	
Текущий (раз в две недели)	Домашнее задание			5	5	Письменное домашнее задание (раз в две недели)
Промежуточный	Зачет			1	1	Самостоятельная практическая работа на компьютере в аудитории
Итоговый	Экзамен				1	Письменный экзамен

5.1 Критерии оценки знаний, навыков

Оценки по всем формам текущего, промежуточного и итогового контроля выставляются по 10-ти балльной шкале.

Оценки за домашние задания, зачет и экзамен выставляются, исходя из следующих критериев:

- статистическая правильность решения задачи,
- статистическая полнота решения задачи,
- наличие и корректность социологической интерпретации статистического вывода,
- корректность оформления статистического вывода.

6 Содержание дисциплины

Тема 1. Введение в статистический анализ данных в социологии

Для решения каких задач в социальных науках используется статистический анализ данных? Примеры исследований, построенных на статистическом анализе данных. Статистический анализ данных в разных социальных науках (социология, экономика, политология, психология). Источники данных для статистического анализа: опрос, перепись, агрегированная статистика. Основные методы анализа. Обзор содержания курса.

Основы работы в SPSS. Синтаксис SPSS.

Тема 2. Описательная статистика и статистические графики.

Шкалы измерений (интервальная, порядковая, категориальная). Средние величины, медиана, мода. Дисперсия. Распределения и их характеристики.

Статистическая визуализация. График плотности распределения. Диаграммы рассеивания. Гистограммы. Столбиковые диаграммы. Линейные диаграммы. Коробчатые диаграммы. Круговые диаграммы.

Тема 3. Статистические гипотезы и их тестирование. Дисперсионный анализ

Понятие статистической гипотезы. Тестирование гипотез. Тесты на равенство средних и пропорций. Доверительные интервалы и их интерпретация. Ошибки типа I и типа II. Понятие статистической мощности. Корректная интерпретация статистической значимости. Величина статистических эффектов и статистическая значимость. Дисперсионный анализ

Тема 4. Таблицы сопряженности и их анализ.

Таблицы сопряженности. Критерий хи-квадрат. Отношения шансов (преобладания). Логлинейные модели. Применение логлинейных моделей в исследованиях социальной мобильности.

Тема 5. Корреляция и парная линейная регрессия.



Коэффициент корреляции Пирсона как мера связи между двумя интервальными переменными. Парная линейная регрессия. Связь между корреляцией и регрессией. Оценка и интерпретация регрессионных коэффициентов и стандартных ошибок. Статистическая значимость коэффициентов. Регрессии с фиктивными переменными.

Тема 6. Множественная линейная регрессия.

Понятие статистического контроля. Оценка и интерпретация регрессионных коэффициентов в моделях множественной регрессии. Эффекты взаимодействия. Коэффициент детерминации R^2 . F-статистика и F-тест. Допущения моделей множественной регрессии и диагностика моделей. Гетероскедастичность. Нелинейные связи. Статистические выбросы. Мультиколлинеарность. Принципы построения регрессионных моделей.

Тема 7. Регрессионные модели для бинарных и категориальных зависимых переменных.

Модель линейной вероятности. Логистическая регрессия. Интерпретация коэффициентов логистической регрессии. Шансы и отношения шансов. Предсказанные вероятности. Мультиномиальная логистическая регрессия. Презентация эффектов в линейной и логистической регрессии.

Тема 8. Идентификация причинно-следственных связей и регрессионный анализ.

Идентификация причинно-следственных связей и регрессионный анализ. Контрфактуальная модель Неймана-Рубина. Проблема пропущенных переменных. Проблема обратной причинно-следственной зависимости. Возможные решения: регрессии с фиксированными эффектами, инструментальные переменные, мэтчинг.

7 Образовательные технологии

Занятия проводятся в форме лекций и практических занятий в компьютерном классе.

8 Оценочные средства для текущего контроля и аттестации студента

8.1 Тематика заданий текущего контроля

Каждые две недели студентам предлагается письменное домашнее задание, которые включает 3-4 задачи для решения в SPSS. Задачи обычно включают в себя самостоятельный анализ данных, социологическую интерпретацию результатов анализа и письменное представление результатов.

9 Порядок формирования оценок по дисциплине

Оценка по курсу складывается из оценок за домашние задания (40%), два письменных зачета в конце каждого модуля (20% каждый) и письменный экзамен в конце четвертого модуля (20%). Оценка за экзамен является блокирующей, т.е. студент, неудовлетворительно сдавший экзамен, не может получить удовлетворительную итоговую оценку за курс в целом.

Письменные домашние работы сдаются каждые две недели. Домашние работы не принимаются с опозданием, за исключением случаев, когда они не могли быть выполнены в срок по уважительной причине (эти случаи отдельно обговариваются с преподавателем).

На пересдаче студенту не предоставляется возможность получить дополнительный балл для компенсации оценки за текущий контроль.

В диплом ставится оценка за итоговый контроль, которая является результирующей оценкой по учебной дисциплине.

10 Учебно-методическое и информационное обеспечение дисциплины

10.1 Базовые учебники

- Тюрин Ю.Н., Макаров А.А. Анализ данных на компьютере. 4-е изд. М.: Форум, 2008.



- Крыштановский А.О. Анализ социологических данных с помощью пакета SPSS. М.: ГУ ВШЭ, 2006.

- Толстова Ю.Н. Анализ социологических данных. М.: Научный мир, 2000.

- Agresti, A. and Finley, B. Statistical Methods for the Social Sciences. 3rd ed. Pearson, 1997.

10.2 Основная литература

Полные библиографические описания основной литературы см. выше в разделе «Базовые учебники».

Тема 1

- Agresti & Finley, ch.1

Тема 2

- Agresti & Finley, ch.3-4

- Крыштановский, гл.1

Тема 3

- Agresti & Finley, ch.5-7, 12

- Тюрин и Макаров, гл.3

- Крыштановский, гл.3

- Толстова, ч.2, гл.1

Тема 4

- Agresti & Finley, ch.8

- Крыштановский, гл.2

- Толстова, ч.2, гл.2.1-2.3

Тема 5

- Agresti & Finley, ch.9

- Крыштановский, гл.4.1-4.3, 4.5

- Тюрин и Макаров, гл.8

Тема 6

- Agresti & Finley, ch.10-11, 14

- Крыштановский, гл.4.4

- Тюрин и Макаров, гл.8

Тема 7

- Agresti & Finley, ch.15

- Крыштановский, гл.4.6

Тема 8

- Gangl, M. "Causal inference in sociological research". Annual Review of Sociology 2010, vol.36, pp.21-47.

10.3 Дополнительная литература

- Gelman, A. and Hill, J. Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge: Cambridge University Press, 2007.
- Long, J.S. Regression Models for Categorical and Limited Dependent Variables. Sage, 1997.
- Long, J.S. and Freese, J. Regression Models for Categorical Dependent Variables Using Stata. Stata Press, 2006.
- Fox, J. Applied Regression Analysis and Generalized Linear Models. Sage, 2008. 2nd ed.
- Holland, P.W. "Statistics and causal inference". Journal of the American Statistical Association 1986, vol.81, p.945-960.
- Abbott, A. "The causal devolution". Sociological Methods and Research 1998, vol. 27(2), p.148-181.



10.4 Программные средства

Для успешного освоения дисциплины, студент использует следующие программные средства:

- SPSS

10.5 Дистанционная поддержка дисциплины

Материалы к курсу, включающие в себя программу, слайды к лекциям и практическим занятиям, домашние задания и базы данных, доступны в Интернете по адресу <http://sites.google.com/site/bessudnov> .

11. Материально-техническое обеспечение дисциплины

На всех лекциях и практических занятиях используется проектор. Практические занятия проводятся в компьютерном классе с установленной на компьютерах программой SPSS.