

**Образовательная автономная некоммерческая организация высшего образования
«МОСКОВСКАЯ ВЫСШАЯ ШКОЛА СОЦИАЛЬНЫХ
И ЭКОНОМИЧЕСКИХ НАУК»**

Факультет социальных наук
(наименование факультета)

УТВЕРЖДЕНО

Ректор ОАНО «МВШСЭН»

Электронная подпись

ОБРАЗОВАТЕЛЬНАЯ ПРОГРАММА ВЫСШЕГО ОБРАЗОВАНИЯ

Современная социальная теория
(наименование образовательной программы)

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
реализуемой без применения электронного (онлайн) курса**

Б1.В.ДВ.06.01. Анализ данных в количественном исследовании
(код и наименование РПД)

Бакалавриат
(уровень образования)

39.03.01 Социология
(код, наименование направления подготовки)

Очная
(форма обучения)

Год набора - 2022

Москва, 2021 г.

Автор–составитель:

Кандидат социологических наук, доцент кафедры теоретической социологии и эпистемологии, Астахова А.С.

(ученое звание, ученая степень, должность) (наименование кафедры) (Ф.И.О.)

ИО заведующего кафедрой теоретической социологии и эпистемологии, кандидат социологических наук, доцент кафедры теоретической социологии и эпистемологии, Астахова А.С.

(наименование кафедры) (ученое звание, ученая степень,) (Ф.И.О.)

Рабочая программа дисциплины «Б1.В.ДВ.06.01. Анализ данных в количественном исследовании» рассмотрена и одобрена на заседании кафедры социологии МВШСЭН

протокол от «17» мая 2022 г. № 5

СОДЕРЖАНИЕ

1. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения образовательной программы.....	4
2. Объем и место дисциплины в структуре ОП ВО	5
3. Содержание и структура дисциплины.....	5
4. Материалы текущего контроля успеваемости обучающихся и фонд оценочных средств промежуточной аттестации по дисциплине.....	7
5. Методические указания для обучающихся по освоению дисциплины.....	12
6. Учебная литература и ресурсы информационно-телекоммуникационной сети "Интернет", включая перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине.....	12
6.1. Основная литература.....	12
6.2. Дополнительная литература.....	12
6.3. Учебно-методическое обеспечение самостоятельной работы.....	12
6.4. Нормативные правовые документы.....	12
6.5. Интернет-ресурсы.....	12
6.6. Иные источники.....	13
7. Материально-техническая база, информационные технологии, программное обеспечение и информационные справочные системы	13

1. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения образовательной программы

1.1. Дисциплина Б1.В.ДВ.06.01. Анализ данных в количественном исследовании обеспечивает овладение следующими компетенциями:

Код компетенции	Наименование компетенции	Индикатор
ПК-3	Способность обрабатывать, анализировать и интерпретировать данные из первичных и вторичных источников	ПК-3.2. Анализирует и интерпретирует информацию в профессиональной деятельности.
ПК-4	Способность отбирать, классифицировать и корректно интерпретировать обобщенную профессиональную информацию при подготовке научных отчетов и презентации результатов исследований с учетом особенностей потенциальной аудитории.	ПК-4.1. Определяет круг релевантных источников для работы над проектом, отбирает и классифицирует профессиональную информацию

1.2 В результате освоения дисциплины у студентов должны быть сформированы:

ОТФ/ТФ/ трудовые или профессиональные действия	Код компонента компетенции	Результаты обучения
Первичная обработка данных с применением программного обеспечения, соответствующего задаче и типу данных.	ПК-3	на уровне знаний: основные принципы планирования сбора данных социологического исследования;
		на уровне умений: осуществлять интерпретацию данных;
		на уровне навыков: обрабатывать и анализировать данные социологического исследования.
Первичная	ПК 4	на уровне знаний:

обработка данных с применением программного обеспечения, соответствующего задаче и типу данных.		основная проблематика и тенденции актуальных современных социологических исследований
		на уровне умений: составлять проектную документацию: программы, инструкции, технические задания;
		на уровне навыков: применения методов социологического исследования в зависимости от тематики и масштаба социологического исследования и требований заказчика; презентации результатов исследования с учетом особенностей аудитории.

2. Объем и место дисциплины в структуре ОП ВО

Объем дисциплины

- Общая трудоемкость дисциплины 144/108 ак./астр. часов (4 ЗЕ);
- 52/39 ак./астр. часов выделено на контактную работу с преподавателем и 56/42 ак./астр. часов на самостоятельную работу обучающихся;

Место дисциплины в структуре ОП ВО

- Б1.В.ДВ.06.01. Анализ данных в количественном исследовании осваивается на 4 курсе обучения (7 семестр);
- дисциплина реализуется после изучения: Б1.О.07 Основы математики, Б1.О.14 Основы социологии

3. Содержание и структура дисциплины

№ п/п	Наименование тем (разделов),	Объем дисциплины (модуля), ак. час.						Форма текущего контроля успеваемости **, промежуточной аттестации
		Всего	Контактная работа обучающихся с преподавателем по видам учебных занятий				СР	
			Л	ЛР	ПЗ	КСР		
Тема 1	Статистический анализ данных в социологии. Статистические гипотезы и их тестирование. Дисперсионный анализ	18/13, 5	2/1,5		6/4,5		10/7,5	П, ОПЗ, Д
Тема 2	Таблицы сопряженности и их анализ и корреляция	18/13, 5	2/1,5		6/4,5	4/3	6/4,5	П, ОПЗ, Д
Тема 3	Регрессионный анализ	18/13, 5	2/1,5		6/4,5		10/7,5	П, ОПЗ, Д
Тема 4	Факторный анализ.	18/13, 5	2/1,5		4/3	4/3	8/6	П, ОПЗ, Д
Тема 5	Кластерный анализ	18/13, 5	4/3		4/3		10/7,5	П, ОПЗ, Д
Тема 6	Модели деревьев решений	16/12	2/1,5		2/1,5		12/9	П, ОПЗ, Д

№ п/п	Наименование тем (разделов),	Объем дисциплины (модуля), ак. час.						Форма текущего контроля успеваемости **, промежуточной аттестации
		Всего	Контактная работа обучающихся с преподавателем по видам учебных занятий				СР	
			Л	ЛР	ПЗ	КСР		
Консультация		2/1,5						
Промежуточная аттестация		36/27						Экзамен
	Всего:	144/108	14/10, 5		28/21	8/6	56/42	

Примечание: ** – формы текущего контроля успеваемости: посещаемость (П), ответ на практическом занятии (ОПЗ), дискуссия (Д)

Содержание дисциплины

№ п/п	Наименование тем (разделов)	Содержание тем (разделов)
1	Статистический анализ данных в социологии. Статистические гипотезы и их тестирование. Дисперсионный анализ	Источники данных для статистического анализа: опрос, перепись, агрегированная статистика. Основные методы анализа. Обзор содержания курса. Основы работы в STATA. Синтаксис STATA. Описательная статистика: основные процедуры. 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
2	Таблицы сопряженности и их анализ и корреляция	Таблицы сопряженности. Критерий хи-квадрат. Отношения шансов (преобладания). Логлинейные модели. Применение логлинейных моделей в исследованиях социальной мобильности. Коэффициент корреляции Пирсона как мера связи между двумя интервальными переменными. Другие коэффициенты корреляции в зависимости от используемых шкал.
3	Регрессионный анализ	Парная линейная регрессия. Связь между корреляцией и регрессией. Оценка и интерпретация регрессионных коэффициентов и стандартных ошибок. Статистическая значимость коэффициентов. Регрессии с фиктивными переменными. 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. 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. 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
4	Факторный анализ.	Модель факторного анализа как модель латентных переменных. Различные подходы к определению числа факторов. Процент объясненной дисперсии как показатель качества факторной модели. Индивидуальные значения факторов. Сохранение факторов как новых переменных. Вращение матрицы факторных нагрузок. Ортогональные и неортогональные методы вращения.
5	Кластерный анализ	Иерархический агломеративный кластерный анализ. Кластерный анализ методом k-средних. Проблемы выбора меры расстояния и формы кластера. Проблема устойчивости кластеризации. Методы оценки устойчивости. Описание и интерпретация результатов кластеризации.
6	Модели деревьев решений	Особенности работы алгоритмов CHAID и CRT. Ошибки классификации, определение понятия риска. Цена ошибки классификации (Misclassification Cost). Возможность задавать различные значения цены ошибки классификации для разных типов ошибок, интерпретация соответствующих моделей. Проверка качества модели, способы решения проблемы излишней подгонки дерева: кросс-проверка, V-кратная кросс-проверка, отсечение ветвей.

4. Материалы текущего контроля успеваемости обучающихся и фонд оценочных средств промежуточной аттестации по дисциплине

В ходе реализации дисциплины Б1.В.ДВ.06.01. Анализ данных в количественном исследовании используются следующие методы текущего контроля и успеваемости обучающихся:

- при проведении занятий лекционного типа:
 - контроль посещаемости;
- при проведении практических занятий:
 - ответ на практическом занятии;
- при контроле результатов самостоятельной работы студентов:
 - дискуссия;

Экзамен проводится с применением следующих методов (средств): устный экзамен, состоящий из ответа на 2 вопроса.

Материалы текущего контроля успеваемости обучающихся.

Посещаемость лекционных и практических занятий, ответы на практических занятиях. Обучающиеся готовят ответы на вопросы к практическим занятиям, исходя из списка

вопросов ниже и списка литературы.

Типовые оценочные материалы по темам.

ВОПРОСЫ К ПРАКТИЧЕСКИМ ЗАНЯТИЯМ

Тема 1

1. Методология количественных методов исследования.
2. Стадии статистического исследования.
3. Виды статистических группировок.

Тема 2

1. Этапы построения группировки.
2. Принципы построения статистических группировок.
3. Классификация рядов распределения.

Тема 3

1. Атрибутивные ряды распределения.
2. Вариационные ряды распределения.
3. Классификация статистических таблиц.

Тема 4

1. Формы средней величины.
2. Виды средних величин.
3. Структурные средние величины.

Тема 5

1. Дисперсия и способы ее расчета.
2. Виды дисперсий.
3. Вариация качественных признаков.
4. Дисперсия альтернативного признака.

Тема 6

1. Классификация связей.
2. Графический метод в анализе взаимосвязей явлений и процессов.
3. Аналитические группировки в анализе взаимосвязей явлений и процессов.
4. Метод приведения параллельных данных в анализе взаимосвязей явлений и процессов.

Оценочные средства для промежуточной аттестации.

Код компетенции	Компетенция	Индикатор оценивания	Критерий оценивания
------------------------	--------------------	-----------------------------	----------------------------

ПК-3	Способен обрабатывать, анализировать и интерпретировать данные из первичных и вторичных источников.	ПК-3.2. Анализирует и интерпретирует информацию в профессиональной деятельности.	<p>Базовый уровень: знает критерии оценки качества источника данных; как составляется адекватное описание данных.</p> <p>Повышенный уровень: может формулировать и описывать предварительные этапы работы с данными в программе исследования; осуществлять сбор первичных или отбор вторичных данных, соответствующих задачам исследования; осуществлять развернутый анализ данных в зависимости от задач исследования; предлагать интерпретацию данных;- на основе интерпретации данных подготовить аналитическое заключение и сделать вывод о предложенных рекомендациях.</p>
ПК-4	Способен отбирать, классифицировать и корректно интерпретировать обобщенную профессиональную информацию при подготовке научных отчетов и презентации результатов исследований с учетом особенностей потенциальной аудитории	ПК-4.1. Определяет круг релевантных источников для работы над проектом, отбирает и классифицирует профессиональную информацию	<p>Базовый уровень: знает критерии отбора релевантных источников и принципы написания обзора, отличия обзора литературы для исследовательских целей от реферата;</p> <p>Повышенный уровень: может составить обзор источников, критически охарактеризовать и классифицировать источники.</p>

Типовые оценочные средства

Вопросы к экзамену¹:

1. Линейный коэффициент корреляции.
2. Парная регрессия.
3. Интерпретация уравнения регрессии.
4. Коэффициенты ассоциации и контингенции.
5. Коэффициенты взаимной сопряженности Пирсона и Чупрова.
6. Коэффициенты ранговой корреляции.
7. Требования к исходной информационной базе.
8. Программно-методологические вопросы статистического наблюдения.
9. Временные ряды. Классификация временных рядов.
10. Аналитические показатели ряда динамики.
11. Средние показатели ряда динамики.

¹ Студентам по билетам задается 2 вопроса.

12. Компоненты временного ряда.
13. Модели. Классификация моделей.
14. Этапы построения статистических моделей.
15. Прогностика как метод научного познания.
16. Прогноз и предсказание. Классификация методов прогноза.
17. Этапы построения моделей статистического прогнозирования.
18. Понятие основной тенденции развития.
19. Методы выявления тенденции в целом в ряду динамики.
20. Методы выявления видов тенденции.
21. Модели тенденции средней и дисперсии.
22. Методы определения типа тенденции явлений и процессов.
23. Метод аналитического выравнивания.
24. Методы выбора модели тенденции различных процессов и явлений.
25. Дисперсионный метод анализа.
26. Понятие периодической компоненты.
27. Методы выявления периодической компоненты.
28. Автокорреляция: причины возникновения.
29. Автокорреляция: методы выявления.
30. Коэффициент автокорреляции.
31. Критерий Дарбина-Уотсона.
32. Модели авторегрессионных преобразований.
33. Мультиколлинеарность: методы выявления.
34. Мультиколлинеарность: методы устранения.
35. Интерпретация статистических моделей регрессии.
36. Прогноз методом экстраполяции тренда.

Шкала оценивания.

Форма промежуточной аттестации	Критерии оценивания	Оценка (баллы)
Устный ответ на экзамене ²	<ul style="list-style-type: none"> • Студент подробно излагает содержание вопроса: исчерпывающе, последовательно, четко и аргументированно излагает материал. • Демонстрирует критическую оценку возможностей и ограничений представленного подхода/метода: рассматривает возможную критику, условия, в которых данные концепции или методы неприменимы. • Демонстрирует способность проследить и реконструировать аргументацию авторов по теме, ссылается в ответе на первоисточники или актуальные исследования. • Поясняет утверждение на уместных примерах. • Уверенно отвечает на дополнительные вопросы, свободно 	Отлично (81-100)

² Преподавателям предлагается оценить ответ по каждому блоку по 5 критериям, выставляя за каждый критерий до 20 баллов, баллы выставляются по двум вопросам отдельно.

Форма промежуточной аттестации	Критерии оценивания	Оценка (баллы)
	ориентируется в теме.	
	<ul style="list-style-type: none"> • Студент подробно излагает содержание билета, но упускает некоторые аспекты рассматриваемого подхода/метода. • Демонстрирует критическую оценку возможностей представленного подхода/метода, но не всегда способен проследить его ограничения. • Демонстрирует способность частично реконструировать аргументацию авторов, на которых ссылается. • Поясняет утверждение на уместных примерах. • Отвечает на дополнительные вопросы по теме с небольшими паузами в разговоре. 	Хорошо (61-80)
	<ul style="list-style-type: none"> • Студент излагает содержание билета поверхностно; • Демонстрирует критическую оценку возможностей представленного подхода/метода на уровне «здравого смысла». • Демонстрирует способность частично реконструировать аргументацию авторов, на которых ссылается, но допускает ошибки. • Поясняет утверждение на примерах, но не всегда сразу может обосновать их уместность. • Отвечает на дополнительные вопросы по теме с паузами в разговоре и ошибками. 	Удовлетворительно (41-60)
	<ul style="list-style-type: none"> • Студент излагает содержание билета сжато, не отражая сути вопроса; • Не демонстрирует критическую оценку возможностей представленного подхода/метода. • Не способен реконструировать аргументацию авторов, допускает серьезные ошибки. • Не использует примеры, либо предложенные примеры не отражают суть вопроса. • Не отвечает на дополнительные вопросы по теме. 	Неудовлетворительно (0-40)

4.4. Методические материалы

Текущий контроль осуществляется с целью мониторинга актуальной ситуации подготовки студентов к этапу промежуточного контроля. Текущий контроль осуществляется в формах: контроля посещаемости, ответов на практических занятиях, дискуссиях. Преподаватель отслеживает текущую ситуацию по указанным формам текущего контроля и выносит студентам рекомендации для улучшения качества подготовки к промежуточному контролю.

Промежуточный контроль по курсу осуществляется в форме экзамена.

Экзамен проводится в форме устного ответа включает ответ на 2 вопроса по билету. Преподавателям предлагается оценить ответ по 5-ти критериям (см.п.4.3.2), выставя за каждый критерий до 20 баллов. Экзамен считается сданным, если студент прошел порог 41 балл, что соответствует оценке «удовлетворительно» и переносится в зачетную книжку и ведомость.

5. Методические указания для обучающихся по освоению дисциплины

Общие положения

Дисциплина подразумевает работу над двумя взаимодополняющими видами деятельности: освоение содержания дисциплины в рамках лекционных и практических занятий (а также самостоятельной работы).

Технология организации самостоятельной работы обучающихся включает использование информационных и материально-технических ресурсов образовательного учреждения.

Перед выполнением обучающимися внеаудиторной самостоятельной работы преподаватель проводит консультирование по выполнению задания, который включает цель задания, его содержания, сроки выполнения, ориентировочный объем работы, основные требования к результатам работы, критерии оценки. Во время выполнения обучающимися внеаудиторной самостоятельной работы и при необходимости преподаватель может проводить индивидуальные и групповые консультации.

Самостоятельная работа может осуществляться индивидуально или группами обучающихся в зависимости от цели, объекта, конкретной тематики самостоятельной работы, уровня сложности, уровня умений обучающихся.

6. Учебная литература и ресурсы информационно-телекоммуникационной сети "Интернет", включая перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

6.1. Основная литература.

1. Стружкин, Н. П. Базы данных: проектирование : учебник для академического бакалавриата / Н. П. Стружкин, В. В. Годин. — М. : Издательство Юрайт, 2017. — 477 с. Режим доступа: <https://www.biblio-online.ru/book/BF8DDE6E-054D-4BB4-A6FA-2E9898529E96>.

6.2. Дополнительная литература.

1. Миркин, Б. Г. Введение в анализ данных : учебник и практикум / Б. Г. Миркин. — М. : Издательство Юрайт, 2016. — 174 с. Режим доступа: <https://www.biblio-online.ru/book/E486A3B0-1AE9-4179-8D48-FA24B626C7C9>

6.3. Учебно-методическое обеспечение самостоятельной работы.

1. Положение об организации самостоятельной работы студентов образовательной автономной некоммерческой организации высшего образования «Московская высшая школа социальных и экономических наук», утверждено приказом ОАНО «МВШСЭН» №112/1 от 31 августа 2017 г. <https://www.msses.ru/sveden/document/>

6.4. Нормативные правовые документы.

Нормативные правовые документы в процессе освоения дисциплины не используются.

6.5. Интернет-ресурсы.

1. ЭБС IPRBooks. Режим доступа: www.iprbookshop.ru .
2. Федеральный информационный портал «Экономика. Социология. Менеджмент» (Разделы: Книги и статьи; Учебные программы; Журнальный зал) Режим доступа: <http://ecsocman.hse.ru> .
3. Общедоступные электронные библиотеки гуманитарной и общественно-научной литературы (Библиотека «Гумер», «Ихтика», «Якова Кротова», «Philosophy.ru» и др.).

4. Базы данных научной периодики и книг (НЭБ/Elibrary.ru, JSTORE, ProQuest, EBSCO и др.).

6.6. Иные источники.

1. Толстова Ю.Н. Математико-статистические модели в социологии. М.: ГУ ВШЭ, 2008.

7. Материально-техническая база, информационные технологии, программное обеспечение и информационные справочные системы

Для преподавания дисциплины требуются аудитории для проведения занятий лекционного и практического типов, вместимостью от 10 до 20 человек, укомплектованные партами, стульями, настенной маркерной или электронной доской, флипчартом. Для ведения занятий лекционного типа требуется оборудование для демонстрации визуальных методических материалов (презентаций в PowerPoint, видео и т.д.): ноутбук, мобильный или стационарный проектор.

Для обеспечения самостоятельной работы обучающихся в рамках освоения дисциплины требуется помещение, укомплектованное партами, стульями, компьютерами с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду организации.

Для обеспечения самостоятельной работы обучающихся требуется следующее лицензионное ПО:

- пакет Microsoft Office, включающий программы Word и PowerPoint.

Для обеспечения преподавания дисциплины требуется помещение для хранения демонстрационного оборудования (ноутбуков, проекторов, пультов для проекторов).