

**Федеральное государственное бюджетное образовательное
учреждение высшего образования
«РОССИЙСКАЯ АКАДЕМИЯ НАРОДНОГО ХОЗЯЙСТВА
И ГОСУДАРСТВЕННОЙ СЛУЖБЫ
ПРИ ПРЕЗИДЕНТЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»**

Институт общественных наук

(наименование института)

Кафедра теоретической социологии и эпистемологии

(наименование кафедры)

УТВЕРЖДЕНА

кафедрой теоретической
социологии и эпистемологии

Протокол от «10» июня 2020 г.

№5

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Б1.В.ДВ.01.23.08 Многомерный анализ данных

(индекс и наименование дисциплины)

Minor "Современные методы анализа данных"

38.03.02 Менеджмент

(код, наименование направления подготовки)

Стратегическое управление компанией (Liberal Arts)

направленность (профиль)

бакалавр

+- (квалификация)

очная

(форма обучения)

Год набора - 2021

Москва, 2020 г.

Автор–составитель:

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

Заведующий кафедрой теоретической социологии и эпистемологии, декан философско-социологического факультета, к.с.н. Вахштайн В.С.

Содержание

1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения программы	4
2. Объем и место дисциплины в структуре ОП ВО	4
3. Содержание и структура дисциплины	5
4. Материалы текущего контроля и фонд оценочных средств промежуточной аттестации по дисциплине	6
5. Методические указания для обучающихся по освоению дисциплины	9
6. Учебная литература и ресурсы информационно-телекоммуникационной сети "Интернет", включая перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине	9
6.1. Основная литература.	9
6.2. Дополнительная литература.	10
6.3. Учебно-методическое обеспечение самостоятельной работы.	10
6.4 Нормативные правовые документы.	10
6.5. Интернет-ресурсы, справочные системы.	10
6.6. Иные рекомендуемые источники.	11
7. Материально-техническая база, информационные технологии, программное обеспечение и информационные справочные системы	11

1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения программы

1.1. Дисциплина Б1.В.ДВ.01.23.08 «Многомерный анализ данных» обеспечивает овладение следующими компетенциями с учетом индикатора:

Код компетенции	Наименование компетенции	Код индикатора компетенции	Наименование индикатора компетенций
СК ОС LA - 25	Способен осуществлять подготовку, анализ, интерпретацию данных из первичных и вторичных источников, а также визуализировать и представлять результаты исследования	СК ОС LA-25.1.	Представляет результаты проведенного анализа данных, может описать и объяснить происхождение полученных значений; демонстрирует владение специализированными программами для работы с данными;

1.2. В результате освоения дисциплины у студентов должны быть сформированы:

Код индикатора компетенции	Результаты обучения
СК ОС LA-25.1	на уровне знаний: <ul style="list-style-type: none"> • основных принципов сбора, подготовки, хранения данных; • основных принципов визуализации данных; основных принципов интерпретации данных при подготовке отчетов, аналитических заключений;
	на уровне умений: <ul style="list-style-type: none"> • осуществлять ввод и «чистку» данных; осуществлять интерпретацию данных;
	на уровне навыков: <ul style="list-style-type: none"> • пользоваться справочными информационными системами для поиска дополнительной информации; • описания и базовой интерпретации первичных и вторичных данных; использования инструментов визуализации данных.

2. Объем и место дисциплины в структуре ОП ВО

- Общая трудоемкость дисциплины: 108 ак./81 астр. часов (3 ЗЕТ);
- 30 ак./22,54 астр. часов выделено на контактную работу с преподавателем и 42 ак./31,5 астрон. часа на самостоятельную работу обучающихся;
- Б1.В.ДВ.01.23.08. Многомерный анализ данных осваивается на 4 курсе обучения (7 семестр).
- дисциплина реализуется после изучения: Б1.О.10 Основы математики, Б1.О.11 Информационные технологии, Б1.В.ДВ.01.23.04 Анализ данных: теория и практика.

3. Содержание и структура дисциплины

Таблица 1.

№ п/п	Наименование тем (разделов)	Объем дисциплины, час.						Форма текущего контроля успеваемост и ⁴ , промежуточ ной аттестации
		Всего	Контактная работа обучающихся с преподавателем по видам учебных занятий				СР	
			Л	ЛР	ПЗ	КС Р		
1	Факторный и кластерный анализ.	22/16,5			8/6		14/10,5	ПЗ
2	Модели деревьев решений	24/18			10/7,5		14/10,5	ПЗ
3	Методы анализа качественных данных. Интерпретация данных.	24/18			10/7,5		14/10,5	О, ПЗ
Консультация		2/1,5						
Промежуточная аттестация		36/27						Экзамен
Всего:		108/81			28/21		42/31,5	

Примечание: ** – формы текущего контроля успеваемости: опрос (О), практические задания (ПЗ)

Содержание дисциплины

№ п/п	Наименование тем (разделов)	Содержание тем (разделов)
Тема 1	Факторный анализ. Кластерный 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. 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.
Тема 2	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.
Тема 3	Методы анализа качественных данных. Интерпретация данных.	Описание и интерпретация данных: различие в количественных и качественных исследованиях. Виды данных в качественном исследовании: текстовые данные, иконическая информация. Подходы к работе с

№ п/п	Наименование тем (разделов)	Содержание тем (разделов)
		текстовыми и визуальными данными: кодирование, категоризация, конденсация смыслов, интерпретация. Дискурс-анализ как тактика и способ качественного анализа данных: преимущества и ограничения. Виды дискурс-анализа (критический, исторический и др.).

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

4.1 Формы и методы текущего контроля успеваемости обучающихся и промежуточной аттестации.

4.1.1. В ходе реализации дисциплины Б1.В.ДВ.01.23.08. Многомерный анализ данных используются следующие методы текущего контроля и успеваемости обучающихся:

– при проведении практических занятий:

- Опрос
- Практические задания

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

4.2. Материалы текущего контроля успеваемости.

Примерные задачи для подготовки к практическим занятиям по темам:

Тема 1. Факторный и кластерный анализ.

1. Произвести кластеризацию целевых групп рекламной информации исходя из статистических данных по потребителям.

Тема 2. Модели деревьев решений

1. Решить методом дерева решений, какую сумму вы потратите на продукты в определённый день.

Тема 3. Методы анализа качественных данных. Интерпретация данных.

1. Требуется изучить наличие/отсутствие связи между рядом показателей по городам России и количеством жителей в данных городах. Какие коэффициенты можно применять в данном случае? Обоснуйте ответ и произведите расчет и описание данных.

4.3 Оценочные средства для промежуточной аттестации.

4.3.1. Перечень компетенций образовательной программы. Индикаторы и критерии оценивания компетенций

Код компетенции	Наименование компетенции	Код индикатора компетенции	Наименование индикатора компетенций

СК ОС LA-25	Способен осуществлять подготовку, анализ, интерпретацию данных из первичных и вторичных источников, а также визуализировать и представлять результаты исследования	СК ОС LA-25.1	Решает задачи, связанные с коммуникацией компании с потребителями в сети интернет, в сфере её маркетинговой деятельности, используя теоретические знания и практические навыки
-------------	--	---------------	--

Индикатор оценивания	Критерии Оценивания
СК ОС LA-25.1 Решает задачи, связанные с коммуникацией компании с потребителями в сети интернет, в сфере её маркетинговой деятельности, используя теоретические знания и практические навыки	Базовый уровень - знает принципы сбора и анализа данных, преимущества и ограничения первичных и вторичных данных, базовые операции, применимые к различным типам данных. Повышенный уровень – может применять различные методы анализа данных и формулировать модели объяснения и интерпретации данных, критически описывать противоречия в данных.

4.3.2 Типовые оценочные средства

Выбор вопросов для ответа на экзамене осуществляется случайным образом методом выбора студентом двух случайных номеров из общего списка. Каждому номеру соответствует определенный вопрос. Выбранные вопросы могут относиться к любой из изученных тем.

Примерные вопросы к экзамену

1. Факторный анализ
2. Кластерный анализ
3. Модели деревьев решений
4. Анализ данных в качественном исследовании: основные подходы
5. Описание и интерпретация данных: исходные требования.

Шкала оценивания

Форма промежуточной аттестации	Критерии оценивания	Оценка (баллы)
Устный ответ на	<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 	Отлично (41-60)

Форма промежуточной аттестации	Критерии оценивания	Оценка (баллы)
экзамен ¹	<p>критику, условия, в которых данные концепции или методы неприменимы.</p> <ul style="list-style-type: none"> • 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 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 (21-41)
	<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 (1-21)
	<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)

**Итоговая семестровая оценка выставляется, исходя из формулы:
40% текущий контроль 60% экзамен.**

¹ Преподавателям предлагается оценить ответ по 5 критериям, выставляя за каждый критерий до 20 баллов, баллы выставляются по двум вопросам отдельно.

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), выставляя 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

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

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. Миркин Б. Г. Введение в анализ данных : учебник и практикум / Б. Г. Миркин. — Москва : Издательство Юрайт, 2020. — 174 с. — (Высшее образование). — ISBN 978-5-9916-5009-0. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <http://biblio-online.ru/bcode/450262>
2. Мхитарян В. С. Анализ данных : учебник для вузов / В. С. Мхитарян [и др.] ; под редакцией В. С. Мхитаряна. — Москва : Издательство Юрайт, 2020. — 490 с. — (Высшее образование). — ISBN 978-5-534-00616-2. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <http://biblio-online.ru/bcode/450166>
3. Сидняев Н. И. Теория планирования эксперимента и анализ статистических данных : учебник и практикум для вузов / Н. И. Сидняев. — 2-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2020. — 495 с. — (Высшее образование). — ISBN

978-5-534-05070-7. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <http://biblio-online.ru/bcode/449686>

4. Кремер Н. Ш. Эконометрика : учебник и практикум для вузов / Н. Ш. Кремер, Б. А. Путко ; под редакцией Н. Ш. Кремера. — 4-е изд., испр. и доп. — Москва : Издательство Юрайт, 2020. — 308 с. — (Высшее образование). — ISBN 978-5-534-08710-9. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <http://biblio-online.ru/bcode/449750>

6.2. Дополнительная литература.

1. Федин Ф. О. Анализ данных. Часть 1. Подготовка данных к анализу : учебное пособие / Ф. О. Федин, Ф. Ф. Федин. — Москва : Московский городской педагогический университет, 2012. — 204 с. — ISBN 2227-8397. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/26444.html>
2. Федин Ф. О. Анализ данных. Часть 2. Инструменты Data Mining : учебное пособие / Ф. О. Федин, Ф. Ф. Федин. — Москва : Московский городской педагогический университет, 2012. — 308 с. — ISBN 2227-8397. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/26445.html>
3. Истомина А. П. Анализ данных качественных исследований : практикум / составители А. П. Истомина. — Ставрополь : Северо-Кавказский федеральный университет, 2016. — 94 с. — ISBN 2227-8397. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/66014.html>

6.3. Учебно-методическое обеспечение самостоятельной работы.

1. Положение об организации самостоятельной работы студентов федерального государственного бюджетного образовательного учреждения высшего образования «Российская академия народного хозяйства и государственной службы при Президенте Российской Федерации» (в ред. приказа РАНХиГС от 11.05.2016 г. № 01-2211). http://www.ranepa.ru/images/docs/prikazy-ranhigs/Pologenie_o_samostoyatelnoi_rabote.pdf

6.4 Нормативные правовые документы.

1. Этический кодекс социолога: https://www.ssa-rss.ru/index.php?page_id=84
2. Профессиональный кодекс социолога: https://www.ssa-rss.ru/index.php?page_id=73

6.5. Интернет-ресурсы, справочные системы.

1. <https://www.python.org/>
2. <https://stepik.org/course/67/promo>
3. <https://pythonworld.ru/samouchitel-python>
4. <https://www.r-project.org/>
5. <https://www.sql.ru/>

6.6. Иные рекомендуемые источники.

1. Husson, F., Lê, S., & Pagès, J. (2017). Exploratory Multivariate Analysis by Example Using R (Vol. Second edition). Boca Raton: Chapman and Hall/CRC.
2. Mailund, T. (2017). Beginning Data Science in R : Data Analysis, Visualization, and Modelling for the Data Scientist. New York: Apress.
3. Shmueli, G., Bruce, P. C., Gedeck, P., & Patel, N. R. (2020). Data Mining for Business Analytics : Concepts, Techniques and Applications in Python. Newark: Wiley.
4. Shmueli, G., Bruce, P. C., Yahav, I., Patel, N. R., & Lichtendahl, K. C. (2017). Data Mining for Business Analytics : Concepts, Techniques, and Applications in R. Hoboken, New Jersey: Wiley.

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

Требования к аудиториям (помещениям) для проведения занятий:

Компьютерный класс, оснащенный компьютерами, объединенными в локальную вычислительную сеть с выходом в Интернет через шлюз. Презентационное оборудование (мультимедийный проектор, интерактивные доски и т.д.)

Требования к программному обеспечению общего пользования:

Установленное программное обеспечение MS Excel, R и R Studio.