

УТВЕРЖДЕНА

Приказом Ректора АНО ВО
«Центральный университет»
Е.В. Ивашкевич
от «26» июня 2025 г. № 0626.32

**Рабочая программа дисциплины (модуля)
«Causal analysis (Причинно-следственный анализ)»
дополнительной профессиональной программы – программы
профессиональной переподготовки «Академия data science»**

Траектория: Продуктовая аналитика

**Москва
2025**

Содержание

1. Краткая характеристика дисциплины (модуля)	3
2. Тематический план	4
3. Содержание дисциплины (модуля)	4
4. Учебно-методическое обеспечение	5
5. Материально-техническое обеспечение	5
6. Методические и оценочные материалы	7

1. Краткая характеристика дисциплины (модуля)

Изучение дисциплины (модуля) «Causal analysis (Причинно-следственный анализ)» критически важно в эпоху больших данных, поскольку позволяет отличать истинные причинно-следственные связи от случайных корреляций, минимизируя риски ошибочных выводов в бизнесе и аналитике. Освоение этих методов обеспечивает специалистам инструменты для точной оценки воздействия вмешательств, что повышает эффективность стратегических решений в маркетинге, экономике и управлении проектами.

Цель изучения дисциплины (модуля): формирование навыков точного установления причинно-следственных связей в данных для обоснованного принятия решений в аналитике и бизнесе.

Задачи изучения дисциплины (модуля):

- научиться определять истинные причины изменений в данных, отличая их от случайных корреляций, для корректного анализа эффектов;
- приобрести умение строить графовые модели зависимостей, чтобы визуализировать и оценивать цепочки причинно-следственных связей;
- сформировать навык применения квазиэкспериментальных методов для оценки влияния факторов в условиях отсутствия рандомизации;
- освоить использование propensity score matching для балансировки сравниваемых групп и минимизации смещений в оценках;
- научиться анализировать гетерогенные эффекты воздействия с помощью regression discontinuity design для выявления пороговых изменений.

В результате освоения дисциплины (модуля) обучающийся должен:

знать:

- разницу между корреляцией и причинно-следственной связью;
- как работают каузальные модели и графы;
- принципы работы "Rubin's Causal Model";
- основные каузальные модели;
- графовые модели установления причинно-следственных связей;
- почему результатам рандомизированных контролируемых экспериментов доверяют;
- в каких случаях и каким образом применять квазиэксперименты;
- как работает, преимущества и ограничения "Propensity score matching";
- как работает, преимущества и ограничения "Difference in differences";
- как работает, преимущества и ограничения "Regression discontinuity design".

уметь:

- воспроизводить каузальные модели;
- воспроизводить графовые модели причинно-следственных связей;
- отличать корреляции и причинно-следственную связь во время принятия решений;
- применять квазиэксперименты в учебных кейсах;
- применять "Propensity score matching" в учебных кейсах;
- применять "Difference in differences" в учебных кейсах;
- применять "Regression discontinuity design" в учебных кейсах.

владеть:

- навыком определения гетерогенных эффектов воздействия;
- навыком нивелирования гетерогенных эффектов воздействия;
- навыками применения на практике квазиэкспериментов;
- навыками применения "Propensity score matching" на практике;
- навыками применения "Difference in differences" на практике;
- навыками применения "Regression discontinuity design" на практике.

2. Тематический план

№ п/п	Наименование раздела дисциплины (модуля)	Трудоемкость, академические часы				ТКУ (текущий контроль успеваемости)
		<i>Очная форма</i>				
		Аудиторная работа		Контроль	Самостояте льная работа	
		Лекции	Семинары (практичес кие занятия)			
1	Корреляция и причинно-следственная связь	1	3		16	Домашние задания
2	Каузальные модели и графы	1	6		16	Домашние задания
3	Rubin's Causal Model	1	4		16	Домашние задания
4	Рандомизированные контролируемые эксперименты	1	4		16	Домашние задания
5	Гетерогенные эффекты воздействия	1	4		16	Контрольная работа
6	Квазиэксперименты	1	2		16	Домашние задания
7	Propensity score matching	1	2		16	Домашние задания
8	Difference in differences	1	2		18	Домашние задания
9	Regression discontinuity design	1	2		18	Домашние задания
	<i>Экзамен</i>			6		
	Итого:	9	29	6	148	
	Объем дисциплины (модуля) (в ак. ч.)	190				

3. Содержание дисциплины (модуля)

№п/п	Наименование раздела дисциплины (модуля)	Содержание дисциплины (модуля) по темам
1	Корреляция и причинно-следственная связь	Введение в причинно-следственный анализ, отличие от корреляции. Ошибки интерпретации корреляций и роль скрытых переменных
2	Каузальные модели и графы	Основы каузальных моделей: переменные, эффекты, допущения Directed Acyclic Graphs (DAG): построение и чтение. Скрытые переменные и идентифицируемость причинных эффектов
3	Rubin's Causal Model	Потенциальные исходы и формализация причинного эффекта Средние и индивидуальные эффекты воздействия, роль допущений
4	Рандомизированные контролируемые эксперименты	Дизайн и проведение RCT: валидность, источники смещений Интерпретация и переносимость результатов RCT
5	Гетерогенные эффекты воздействия	Методы выявления гетерогенных эффектов Практика анализа и интерпретации гетерогенных эффектов
6	Квазиэксперименты	Концепция квази-экспериментов, область применения и ограничения
7	Propensity score matching	Логика метода Propensity Score Matching и практическая реализация
8	Difference in differences	Difference-in-Differences: предпосылки, применение и интерпретация
9	Regression discontinuity design	Regression Discontinuity Design: логика, примеры и ограничения

4. Учебно-методическое обеспечение

Университет располагает полным набором лицензионного и свободно распространяемого программного обеспечения, включая продукты отечественного производства.

Каждый слушатель в течение всего периода обучения получает индивидуальный неограниченный доступ к электронно-библиотечной системе и электронной информационно-образовательной среде университета. Эти системы предоставляют возможность доступа к ресурсам из любой точки, где есть подключение к сети Интернет, как на территории университета, так и за его пределами.

Слушателям обеспечен удаленный доступ к современным профессиональным базам данных и информационным справочным системам.

Основная литература:

1. Шифрин, Б. М. Основы теории принятия решений : учебное пособие / Б. М. Шифрин, М. Б. Шифрин. - Москва ; Вологда : Инфра-Инженерия, 2024. - 148 с. - ISBN 978-5-9729-1903-1. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/2171802>.

2. Карпушкин, С. В. Основы теории принятия решений : учебное пособие / С. В. Карпушкин. – Москва ; Вологда : Инфра-Инженерия, 2025. - 128 с. – ISBN 978-5-9729-2664-0. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/2226635>.

Дополнительная литература:

1. Целых, А. Н. Принятие решений на основе методов машинного обучения : учебное пособие по курсам «Модели и методы инженерии знаний», «Методы анализа больших данных» / А. Н. Целых, Н. В. Драгныш, Э. М. Котов ; Южный федеральный университет. - Ростов-на-Дону ; Таганрог : Издательство Южного федерального университета, 2022. - 113 с. - ISBN 978-5-9275-4246-8. - Текст : электронный. - URL: <https://znanium.com/catalog/product/2132258>.

2. Зиятдинов, Н. Н. Основы теории принятия решений : учебно-методическое пособие / Н. Н. Зиятдинов, Т. В. Лаптева, И. В. Логинова ; Минобрнауки России, Казан. нац. исслед. технол. ун-т. - Казань : КНИТУ, 2023. - 104 с. - ISBN 978-5-7882-3352-9. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/2172683>.

5. Материально-техническое обеспечение

Университет располагает материально-технической базой, соответствующей действующим противопожарным правилам и нормам и обеспечивающей проведение всех видов дисциплинарной и междисциплинарной подготовки, практической и научно-исследовательской работ обучающихся, предусмотренных учебным планом.

Помещения, которые представляют собой учебные аудитории для проведения занятий лекционного типа, занятий семинарского (практического) типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также помещения для самостоятельной работы и помещения для хранения и профилактического обслуживания учебного оборудования. Помещения укомплектованы специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Изучение дисциплины (модуля) обеспечивается в учебных аудиториях, оснащенных:

- столами и стульями;
- компьютерной техникой;
- механическими калькуляторами;
- специализированным оборудованием, включая демонстрационное оборудование.

Помещения для самостоятельной работы обучающихся, в том числе приспособленные для использования инвалидами и лицами с ограниченными возможностями здоровья,

оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду Университета.

Обучающимся предоставляется доступ (в том числе удаленный) к ресурсам информационно-телекоммуникационной сети «Интернет», электронным ресурсам (в том числе электронным библиотечным системам, современным профессиональным базам данных и информационным справочным системам):

№	Наименование портала (издания, курса, документа)	Ссылка
1.	Научная электронная библиотека elibrary.ru библиотека	https://elibrary.ru/defaultx.asp
2.	База данных для IT-специалистов	https://habr.com
3.	База данных ScienceDirect	https://www.sciencedirect.com
4.	Официальный сайт Министерства науки и высшего образования Российской Федерации	https://minobrnauki.gov.ru/
5.	Федеральный портал «Российское образование»	https://www.edu.ru/
6.	Информационная система "Единое окно доступа к образовательным ресурсам"	http://window.edu.ru/
7.	Единая коллекция цифровых образовательных ресурсов	http://school-collection.edu.ru/
8.	Федеральный центр информационно - образовательных ресурсов	http://fcior.edu.ru/

Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), в том числе комплект лицензионного программного обеспечения, современные профессиональные базы данных и информационные справочные системы:

Наименование ПО	Производство	Лицензионное / свободно распространяемое
Операционные системы:		
Microsoft Imagine (Windows Client, Server)	зарубежное	лицензионное
Браузеры:		
Яндекс.Браузер	отечественное	свободно распространяемое
Google Chrome	зарубежное	свободно распространяемое
Офисные приложения:		
Microsoft Imagine (Visio, OneNote)	зарубежное	лицензионное
TeXstudio	зарубежное	свободно распространяемое
Adobe Acrobat Reader	зарубежное	свободно распространяемое
Программное обеспечение для планирования и учета времени:		
Toggle app	зарубежное	свободно распространяемое
Системы управления проектами:		
Microsoft Imagine (Project)	зарубежное	лицензионное
Системы управления базами данных:		
Microsoft Imagine (SQL Server)	зарубежное	лицензионное
Системы резервного копирования (backup):		
Acronis Backup Advanced for HyperV	зарубежное	лицензионное
Справочно-правовые системы:		
КонсультантПлюс: справочно-правовая система	отечественное	лицензионное
Средства антивирусной защиты:		

Kaspersky Endpoint Security для бизнеса Стандартный Russian Edition	отечественное	лицензионное
Среды разработки:		
Visual Studio Code	зарубежное	свободно распространяемое
Bash (Unix shell)	зарубежное	свободно распространяемое
Anaconda	зарубежное	свободно распространяемое
Robotic Operating System	зарубежное	свободно распространяемое
CopelliaSim	зарубежное	свободно распространяемое
Google Colaboratory	зарубежное	свободно распространяемое
Пакеты программных средств и библиотек:		
AutoPsy	зарубежное	свободно распространяемое
Interactive Disassembler (IDA)	зарубежное	свободно распространяемое
Системы управления библиографической информацией:		
Zotero	зарубежное	свободно распространяемое
Сервисы и службы:		
Bind	зарубежное	свободно распространяемое
Docker	зарубежное	свободно распространяемое

6. Методические и оценочные материалы

Методические указания для обучающихся по освоению дисциплины (модуля)

В процессе изучения дисциплины (модуля) «Causal analysis (Причинно-следственный анализ)» в рамках текущего контроля успеваемости используются такие виды учебной работы, как лекции, семинары, контрольная работа, домашние задания, а также различные виды самостоятельной работы обучающихся по заданию преподавателя, направленные на развитие навыков профессиональной лексики, закрепление практических профессиональных компетенций, поощрение инициатив.

Лекция – систематическое, последовательное, монологическое изложение преподавателем учебного материала, как правило, теоретического характера.

В процессе лекций рекомендуется вести конспект лекций: кратко и схематично фиксировать основные идеи, выводы и обобщения лекции; выделять важные мысли, ключевые слова и термины. Необходимо отметить вопросы или материалы, которые вызывают затруднения, и попытаться найти ответы в рекомендованной литературе. Если разобраться в материале не удастся, следует сформулировать вопрос и задать его преподавателю на консультации или во время семинарского (практического) занятия.

Семинар – это форма учебной деятельности, проводимая в учебном заведении под руководством преподавателя, где слушатели активно участвуют в обсуждениях, практических заданиях и других формах взаимодействия.

Для успешной подготовки к семинару рекомендуется заранее ознакомиться с темой занятия и основными материалами, чтобы иметь возможность активно участвовать в обсуждении. Также полезно подготовить вопросы и идеи для обсуждения, что поможет глубже понять материал и продемонстрировать заинтересованность.

Контрольная работа – письменная работа с набором задач, которые нужно решить за ограниченное время.

Цель контрольной работы – получить специальные знания по одной или нескольким темам дисциплины и продемонстрировать навыки их практического применения.

Домашнее задание – набор задач по темам недели.

При работе над домашними заданиями важно внимательно ознакомиться с требованиями и сроками выполнения. Рекомендуется разбивать задания на этапы, чтобы

избежать перегрузки и лучше усвоить материал. Использовать различные источники информации, включая учебники и онлайн-ресурсы, для более глубокого понимания темы.

Самостоятельная работа – работа слушателей, направленная на углубленное изучение отдельных тем и вопросов учебной дисциплины (модуля).

В процессе самостоятельной работы слушатели взаимодействуют с рекомендованными материалами при минимальном участии преподавателя. Задачи слушателя включают работу с конспектами лекций (обработка текста), повторное изучение учебных материалов планов и тезисов ответов, изучение дополнительных тем, выполнение учебно-исследовательских заданий и другое.

Система оценивания результатов обучения по дисциплине (модулю)

Оценивание уровня учебных достижений обучающихся по дисциплине (модулю) осуществляется в виде текущего контроля успеваемости и промежуточной аттестации.

Промежуточная аттестация по дисциплине (модулю) осуществляется в форме **экзамена**.

Для оценивания текущего контроля успеваемости и промежуточной аттестации используется десятибалльная шкала оценивания, которая соотносится с традиционной пятибалльной шкалой следующим образом:

Десятибалльная оценка	Пятибалльная оценка	Оценка за зачет	Общая характеристика результата обучения по дисциплине (модулю)
10	Отлично	Зачтено	Слушатель полностью владеет знаниями, изложенными в рабочей программе, и глубоко осмысляет дисциплину (модуль). Он самостоятельно и логически последовательно отвечает на все вопросы, акцентируя внимание на наиболее важном. Умеет анализировать, сравнивать, классифицировать, обобщать, конкретизировать и систематизировать изученный материал, выделяя ключевые моменты и устанавливая причинно-следственные связи. Четко формулирует ответы, уверенно интерпретирует результаты анализов и других исследований, а также решает сложные задачи. Слушатель хорошо знаком с методами исследования, необходимыми для практической деятельности, и умеет связывать теоретические аспекты дисциплины (модуля) с практическими задачами.
9	Отлично	Зачтено	
8	Отлично	Зачтено	
7	Хорошо	Зачтено	Слушатель обладает знаниями предмета почти в полном объеме рабочей программы и самостоятельно, логически последовательно и всесторонне отвечает на все вопросы, акцентируя внимание на наиболее значимых моментах. Он умеет
6	Хорошо	Зачтено	

Десятибалльная оценка	Пятибалльная оценка	Оценка за зачет	Общая характеристика результата обучения по дисциплине (модулю)
			анализировать, сравнивать, классифицировать, обобщать, конкретизировать и систематизировать изученный материал, выделяя его ключевые аспекты и устанавливая причинно-следственные связи. Формулирует свои ответы, уверенно интерпретирует результаты анализов и других исследований, а также решает сложные ситуационные задачи. Слушатель хорошо знаком с методами исследования, необходимыми для практической деятельности, и умеет связывать теоретические аспекты предмета с практическими задачами.
5	Удовлетворительно	Зачтено	Слушатель обладает базовыми знаниями по дисциплине (модулю), но испытывает трудности при самостоятельных ответах и использует неточные формулировки. В ходе ответов он допускает ошибки, касающиеся сути вопросов. Слушатель способен решать только самые простые задачи и владеет лишь минимальным набором методов исследования.
4	Удовлетворительно	Зачтено	
3	Не сдан	Не зачтено	Слушатель не овладел обязательным минимумом знаний по предмету и не может ответить на вопросы, даже если преподаватель задает дополнительные наводящие вопросы.
2	Не сдан	Не зачтено	
1	Не сдан	Не зачтено	

Дисциплина (модуль) «Causal analysis (Причинно-следственный анализ)» оценивается следующим образом:

Активность	Вес	Описание
Домашние задания	60%	Набор задач по темам недели
Контрольная работа	20%	Письменная работа с набором задач, которые нужно решить за ограниченное время
Экзамен	20%	Письменная или устная работа над заданием, направленным на проверку полученных знаний и навыков по дисциплине (модулю)

Формула расчёта итоговой оценки по дисциплине (модулю) «Causal analysis (Причинно-следственный анализ)»: « $0,6 \times$ среднее за домашние задания + $0,2 \times$ контрольная работа + $0,2 \times$ экзамен».

Текущий контроль успеваемости обучающихся по дисциплине (модулю)

Примерные домашние задания

Домашнее задание 1.

1. Проанализируйте предоставленный датасет с экономическими показателями и определите, какие корреляции могут быть ошибочно приняты за причинно-следственные связи, объяснив роль скрытых переменных.
2. Опишите реальный пример из бизнеса, где сильная корреляция не указывает на причинно-следственную связь, и предложите метод для проверки истинной причины.
3. Постройте простую модель на основе данных о продажах и рекламе, иллюстрируя, как скрытые переменные (например, сезонность) могут исказить интерпретацию корреляции.
4. Разработайте сценарий кейс-стади, где ошибки в интерпретации корреляций привели к неудачному бизнес-решению, и предложите корректный подход к анализу причинно-следственных связей.
5. Используя данные о здоровье и питании, сравните корреляционный и причинно-следственный анализ, объяснив, почему второй более надежен для выводов.

Домашнее задание 2.

1. Постройте DAG для гипотетического сценария влияния образования на доход, включив скрытые переменные, и объясните, как это помогает идентифицировать причинные эффекты.
2. Проанализируйте существующий DAG из литературы по маркетингу и предложите изменения для учета дополнительных допущений о переменных и эффектах.
3. Создайте каузальную модель на основе данных о влиянии маркетинговых кампаний на продажи, определив ключевые переменные, эффекты и допущения.
4. Опишите, как скрытые переменные в DAG могут препятствовать идентифицируемости причинных эффектов, и предложите способы их выявления на примере экономических данных.
5. Разработайте DAG для анализа причинно-следственных связей в цепочке поставок и объясните, как чтение графа помогает в принятии решений.

Домашнее задание 3.

1. Примените концепцию потенциальных исходов к данным рандомизированного эксперимента по тестированию лекарства, рассчитав средний эффект воздействия.
2. Опишите индивидуальные эффекты воздействия в Rubin's модель на примере A/B-теста маркетинговой кампании, обсудив роль допущений.
3. Проанализируйте датасет без рандомизации и используйте Rubin's модель для оценки причинного эффекта, объяснив ограничения допущений.
4. Разработайте план применения Rubin's Causal Model к кейсу оценки влияния обучения на производительность сотрудников, включая расчет средних эффектов.
5. Сравните индивидуальные и средние эффекты воздействия в модели Rubin на основе экономических данных, предложив интерпретацию для бизнес-решения.

Примерные задания для контрольной работы

1. Объясните ключевые отличия между корреляцией и причинно-следственной связью, приведя пример, где сильная корреляция не указывает на причинность.
2. Опишите роль скрытых переменных в ошибках интерпретации корреляций и предложите способ их выявления на примере экономических данных.

3. Постройте простой Directed Acyclic Graph (DAG) для сценария влияния рекламы на продажи, включая скрытые переменные, и объясните, как он помогает в идентифицируемости причинных эффектов.

4. Проанализируйте каузальную модель с переменными, эффектами и допущениями на примере маркетинговой кампании, указав, как скрытые переменные могут нарушить идентифицируемость.

5. Используя концепцию потенциальных исходов в Rubin's Causal Model, рассчитайте средний эффект воздействия для гипотетического эксперимента по тестированию новой стратегии ценообразования.

6. Объясните различия между индивидуальными и средними эффектами воздействия в Rubin's модели, обсудив роль допущений на примере анализа влияния обучения на производительность.

7. Опишите дизайн рандомизированного контролируемого эксперимента (RCT) для оценки эффективности новой маркетинговой стратегии, указав источники потенциальных смещений и способы их минимизации.

8. Проанализируйте интерпретацию результатов RCT, включая переносимость на другие контексты, на примере эксперимента по влиянию диеты на здоровье.

9. Примените метод выявления гетерогенных эффектов воздействия к данным А/В-теста, разделив выборку по демографическим признакам и интерпретируя различия.

10. Разработайте практический план анализа гетерогенных эффектов для кейса оценки влияния скидок на лояльность клиентов, объяснив методы и возможные выводы.

Задания для промежуточной аттестации по дисциплине (модулю)

№ п/п	Задание	Ответ
1.	Назовите основное отличие корреляции от причинно-следственной связи.	не указывает на причинность
2.	Укажите формат публичного обсуждения применения квазиэкспериментов.	дискуссия
3.	Укажите рекомендацию по интерпретации корреляции как причинно-следственной связи.	контролирование скрытых переменных
4.	Укажите формат демонстрации логики Propensity Score Matching.	семинар
5.	Какую роль играют скрытые переменные в ошибках интерпретации корреляций?	искажают связь
6.	Укажите формат отчета по выявлению гетерогенных эффектов воздействия.	график
7.	Укажите рекомендацию по использованию DAG для идентифицируемости эффектов.	построение графа
8.	Укажите метод определения приоритетов в применении каузальных моделей.	приоритизация гипотез/матрица
9.	Оцените, почему введение в причинно-следственный анализ важно для самооценки профессиональных навыков.	помогает совершенствовать подходы
10.	Укажите способ представления интерпретации результатов RCT.	визуализация
11.	Укажите рекомендацию по расчёту средних эффектов воздействия в Rubin's Model.	потенциальные исходы
12.	Укажите рекомендацию по дизайну RCT для минимизации смещений.	рандомизация
13.	Как скрытые переменные влияют на вашу способность определять приоритеты в анализе данных?	затрудняют идентификацию
14.	Укажите формат демонстрации построения и чтения DAG.	демо
15.	Укажите рекомендацию по анализу гетерогенных эффектов в продуктах.	сегментация данных

16.	Укажите метод публичного разбора потенциальных исходов в Rubin's Causal Model.	кейс-стади
17.	Укажите рекомендацию по применению Difference-in-Differences.	предпосылки
18.	Назовите способ совершенствования собственной деятельности на основе понимания ошибок интерпретации корреляций.	самооценка допущений
19.	Укажите способ публичного представления результатов по отличию причинно-следственной связи от корреляции.	презентация
20.	Укажите рекомендацию по использованию Regression Discontinuity Design.	порог