

# Мидл аналитик данных

## 01 Кому подойдёт курс

Курс рассчитан на аналитиков данных уровня junior и junior+, которые хотят освоить продвинутый SQL, автоматизацию на Python и BI-инструменты, научиться проводить A/B-тесты и использовать ИИ для оптимизации работы.

Он также подойдёт аналитикам уровня middle, которые хотят глубже разобраться в продвинутых методах экспериментов, каузальном анализе и практическом ML для прогнозов, и специалистам смежных ролей, которые планируют систематизировать знания по данным, закрыть пробелы в статистике и Python и перейти в аналитику или усилить решения в своей текущей роли.

## 02 Чему научитесь на курсе: какие знания и навыки освоите

### Реальные задачи из практики аналитика

- Вы погрузитесь в ситуации, с которыми middle-аналитик сталкивается каждый день. Будете искать причины падения метрик, проектировать дашборды, узнаете, как проводить A/B-тесты и интерпретировать результаты экспериментов.
- Все кейсы — реальные задачи, которые влияют на бизнес-решения.

### Освоите востребованный стек технологий

- Вы научитесь писать сложные SQL-запросы, автоматизировать расчёты на Python, строить отчёты в BI-инструментах и применять статистические методы, а ещё работать с ML и ИИ.
- В расширенной программе — работа с DWH, оркестрацией данных, ClickHouse и Data Lake.
- Для максимальной версии — создание RAG-ботов и работа с чувствительными данными. Особый упор на гибкие навыки аналитика.
- Каждый инструмент вы отработаете на примерах, которые встречаются в реальной работе аналитика.

### Станете самостоятельным специалистом

- Главный навык сильного аналитика — умение разбираться в незнакомых задачах без подсказок.
- Вы научитесь критически оценивать результаты и находить ошибки в данных. Освоите ИИ как инструмент: сможете формулировать запросы, проверять ответы и использовать нейросети для ускорения рутины.
- Главное — перестанете теряться перед новой технологией или нестандартной задачей и сможете сами прокладывать путь к решению.

### Научитесь говорить с бизнесом на одном языке

Аналитика бесполезна, если выводы не доходят до стейкхолдеров. Вы научитесь переводить расплывчатые запросы в точные метрики, аргументировать результаты и защищать свои решения перед командой и руководством. Научитесь оформлять выводы так, чтобы они становились основой для действий.

# Мидл аналитик данных

02 Чему научитесь на курсе: какие знания и навыки освоите

## Получите обратную связь от практикующих экспертов

- Ваши решения будут проверять наставники — практикующие аналитики. Они помогут увидеть слабые места, подскажут, как улучшить код и логику, и поделятся реальными кейсами из своего опыта.
- Для максимальной версии предусмотрены индивидуальные консультации.

## Подготовка к собеседованиям и карьерному росту

- Вы разберёте типичные задачи, которые встречаются на собеседованиях для middle-аналитиков — от написания кода на Python до интерпретации результатов экспериментов.
- С наставниками потренируетесь отвечать на вопросы и получите обратную связь, чтобы уверенно проходить собеседования и претендовать на позиции с более высокой зарплатой.
- *Подготовка к собеседованиям недоступна при оплате курса работодателем*

03 Как проходит обучение

## Оптимальный темп

- Курс разбит на спринты — блоки, в которых мы подробно разбираем отдельные темы. В зависимости от версии вы пройдёте от 8 до 15 спринтов. Спринты открываются по очереди, чтобы вы могли спокойно встроить учёбу в рабочий график.
- Вы будете изучать теорию, практиковаться, поучаствуете в воркшопах и выполните индивидуальный проект.

## Теория и практика на платформе

- В каждом спринте — гайды, конспекты, реальные задачи и кейсы с комментариями экспертов. Материалы в формате текста, иллюстраций и небольших задач помогают сразу отрабатывать навыки на платформе и в рабочих инструментах.
- Учиться можно из любой точки мира — нужен только интернет.

## Воркшопы для отработки навыков

- Несколько раз за курс мы соберёмся на воркшопы. Вместе с наставником и другими студентами вы будете решать задачи, приближённые к реальной работе: разбирать сложные кейсы, проектировать решения и прокачивать гибкие навыки.
- Это возможность взглянуть на одну проблему с разных сторон и научиться аргументировать свой подход.

# Мидл аналитик данных

03 Как проходит обучение

## Практические работы

После каждого спринта — проект на реальных данных. Вы спроектируете дашборд, проведёте A/B-тест, настроите пайплайн или постройте прогнозную модель. Опытные ревьюеры проверят код, логику и выводы, дадут развёрнутую обратную связь. Вы сможете её учесть и добавить удачные работы в портфолио.

## Общение и поддержка

Куратор поможет с организационными вопросами, а наставник объяснит материал. Другие участники курса поддержат и помогут справиться с заданиями.

## Что вас ждёт на курсе

Диплом  
о профессиональной  
переподготовке  
или сертификат

Практика, основанная  
на решении реальных  
рабочих задач

Обучение от экспертов  
из Яндекса и других крупных  
компаний

# Мидл аналитик данных

## Сравнение версий

	Базовая версия	Расширенная версия	Максимальная версия
Длительность обучения	4 месяца	6 месяцев	7 месяцев
В каких случаях подойдёт	<p>Если хотите освоить полный стек middle-аналитика: продвинутый SQL и Python, продуктовые метрики, A/B-тесты, ML и автоматизацию с ИИ</p> <p>Если нужно научиться самостоятельно решать задачи — от поиска причин падения метрик до внедрения прогнозных моделей</p>	<p>Если планируете работать с инфраструктурой данных: проектировать DWH, настраивать ETL-пайплайны, оркестрацию и работать с ClickHouse</p> <p>Если хотите не просто считать метрики, а контролировать качество данных и выстраивать надёжные процессы — от источника до витрины</p>	<p>Если готовы выйти на экспертный уровень: создавать RAG-ботов, внедрять LLM-автоматизацию и разбираться в Privacy &amp; Compliance</p> <p>Если хотите прокачать гибкие навыки, получить индивидуальные консультации с наставником и подготовиться к собеседованиям на позиции middle+</p> <p>Подготовка к собеседованиям недоступна при оплате курса работодателем</p>
Практические работы	8 проектов и 5 ревью	12 проектов и 9 ревью	14 проектов и 9 ревью
Воркшопы с экспертами	7 воркшопов с экспертами	11 воркшопов с экспертами	11 воркшопов с экспертами
Модуль ключевых компетенций мидл аналитика	<ul style="list-style-type: none"> <li>Освоите продвинутый SQL и Python для сложных запросов, очистки данных и автоматизации расчётов</li> <li>Научитесь строить продуктовые метрики, анализировать воронки, проводить A/B-тесты и каузальный анализ</li> <li>Примените машинное обучение для прогнозирования и ИИ для ускорения рутинных задач</li> <li>Будете учиться доносить свои идеи и уверенно защищать свои решения</li> </ul>		
Модуль по DE	✗	<ul style="list-style-type: none"> <li>Научитесь проектировать хранилища данных, витрины и работать с историчностью</li> <li>Освоите оркестрацию пайплайнов, контроль качества данных и мониторинг в продакшене</li> <li>Спроектируете и создадите таблицы в ClickHouse с оптимальными партициями и ключами сортировки под задачи аналитики.</li> </ul>	
Модуль по гибким навыкам аналитика и Privacy and compliance	✗	✗	<ul style="list-style-type: none"> <li>Освоите ответственное обращение с данными и аудит витрин на PII</li> <li>Научитесь аргументировать методологию, работать с конфликтами и защищать решения перед бизнесом</li> </ul>

# Мидл аналитик данных

## Сравнение версий

	Базовая версия	Расширенная версия	Максимальная версия
Модуль «RAG-боты и автоматизация LLM»	×	×	<ul style="list-style-type: none"> <li>Разберётесь в архитектуре RAG и создадите собственного ассистента</li> <li>Настроите семантический поиск и интегрируете бота в Telegram</li> </ul>
Факультативный модуль «Подготовка к собеседованиям» для самостоятельного обучения	×	×	<ul style="list-style-type: none"> <li>Потренируетесь решать задачи по SQL, Python/pandas, статистике и теории вероятностей</li> <li>Разберёте бизнес-кейсы и узнаете, на что обращают внимание при найме middle-аналитиков</li> </ul>
Консультации с наставником по разбору вашего кейса	×	×	2 индивидуальных встречи по 45 минут с опытным аналитиком данных

# Мидл аналитик данных

4, 6 или 7 месяцев

продолжительность курса

8, 12 или 14 проектов

в портфолио

2 ЧАСА

00

Бесплатная часть.  
Тест-драйв обучения

2 НЕДЕЛИ

01

Продвинутый SQL

2 НЕДЕЛИ

02

Продвинутый Python  
для аналитики

2 НЕДЕЛИ

03

Стратегическое  
понимание продукта

2 НЕДЕЛИ

04

BI-инструменты

2 НЕДЕЛИ

05

Статистика  
и эксперименты:  
A/B-тесты end-to-end

2 НЕДЕЛИ

06

Causal Inference

2 НЕДЕЛИ

07

Machine Learning для  
аналитических задач

2 НЕДЕЛИ

08

Использование ИИ  
в аналитике

2 НЕДЕЛИ

09

Хранилища  
и DWH-архитектура

Расширенная версия Максимальная версия

2 НЕДЕЛИ

10

Оркестрация  
для аналитиков

Расширенная версия Максимальная версия

2 НЕДЕЛИ

11

ClickHouse

Расширенная версия Максимальная версия

2 НЕДЕЛИ

12

Data Lake: архитектура  
и обработка данных

Расширенная версия Максимальная версия

2 НЕДЕЛИ

13

RAG-боты  
и автоматизация LLM

Максимальная версия

1 НЕДЕЛЯ

14

Privacy & Compliance

Максимальная версия

1 НЕДЕЛЯ

15

Soft Skills аналитика

Максимальная версия

2 НЕДЕЛИ



Бонусный модуль  
«Подготовка  
к собеседованиям»

Подготовка к собеседованиям  
недоступна при оплате курса  
работодателем

Максимальная версия

# Бесплатная часть. Тест-драйв обучения

00

2 часа

Сначала вы познакомитесь с программой и форматом обучения: узнаете, как устроены спринты, какие навыки будете отрабатывать на практике и кто будет сопровождать вас во время учёбы.

Затем пройдёте несколько уроков. Они помогут погрузиться в процесс и разобрать конкретную задачу аналитика: перевод бизнес-вопроса в чёткую структуру анализа. Вы увидите, как мы выстраиваем подачу материала, и как строится работа над реальными кейсами.

Всё это поможет понять, насколько вам подходит курс, и выбрать оптимальную версию.

## Содержание

Темы

1. О чём этот курс и как проходит обучение
2. Тест-драйв обучения. Метрики и гипотезы

# Продвинутый SQL

01

2 недели

Научитесь писать сложные запросы с разными типами JOIN, использовать CTE и вложенные запросы для структурирования кода, освоите оконные функции и базовые принципы оптимизации.

Создадите аналитическую витрину на реальных данных: рассчитаете ключевые продуктовые метрики — продажи по категориям, активные пользователи, скользящие суммы, примените ранжирование пользователей и проанализируете план выполнения запроса, чтобы убедиться в его оптимальности. На выходе — готовое решение с обоснованием принятых проектных решений.

## Содержание

Темы

1. Сложные JOIN
2. CTE и вложенные запросы
3. Оконные функции
4. Оптимизация запросов

Проект

На основе данных рассчитаете метрики, необходимые для анализа бизнеса, и предложите оптимальную структуру запросов.

Инструменты  
и технологии

- SQL (сложные JOIN, CTE, оконные функции)
- PostgreSQL
- EXPLAIN / EXPLAIN ANALYZE

# Продвинутый Python для аналитики

02

2 недели

Освоите pandas для полного цикла анализа: от исследовательского анализа и очистки до автоматизации расчётов и упаковки результатов. Научитесь выявлять аномалии, визуализировать распределения и писать воспроизводимый код.

Реализуете сквозной пайплайн: изучите структуру данных, проведёте EDA, обработаете пропуски и выбросы, рассчитаете ключевые метрики. Вынесете повторяющуюся логику в функции, добавите автоматические проверки качества и оформите код так, чтобы анализ можно было воспроизвести на новых данных.

## Содержание

Темы	<ol style="list-style-type: none"><li>1. Работа с pandas</li><li>2. Исследовательский анализ данных (EDA) и очистка данных</li><li>3. Автоматизация аналитических расчётов</li><li>4. Упаковка результатов</li></ol>
------	--

Проект	На основе данных проведёте полный анализ: от изучения структуры до расчёта продуктовых метрик и визуализации. Упакуете код в воспроизводимый пайплайн с автоматическими проверками.
--------	---

Инструменты и технологии	<ul style="list-style-type: none"><li>• Python: pandas, numpy, matplotlib, seaborn</li><li>• Jupyter Notebook</li></ul>
--------------------------	---

# Стратегическое понимание продукта

03

2 недели

Научитесь переводить бизнес-задачи в измеримые метрики: строить дерево метрик, анализировать пользовательские воронки, выявлять точки роста и формулировать проверяемые продуктовые гипотезы.

Выстройте иерархию показателей, исследуете путь пользователя и найдёте узкие места. Оцените влияние потерь на каждом этапе на итоговый результат и рассчитаете потенциал улучшений.

На основе этих данных разработаете приоритизированный список гипотез и оформите рекомендации для бизнеса.

## Содержание

Темы	<ol style="list-style-type: none"><li>1. Построение дерева метрик</li><li>2. Анализ пользовательской воронки</li><li>3. Генерация и приоритизация продуктовых гипотез</li></ol>
------	---

# Содержание

---

**Проект** Для заданного продукта и бизнес-запроса постройте дерево метрик, проведёте анализ воронки, сформулируете и приоритезируете гипотезы, а также подготовите выводы на языке бизнеса.

---

**Инструменты и технологии**

- Дерево метрик
- Funnel Analysis
- ICE/RICE
- Python

---

## VI-инструменты

# 04

**2 недели** Освоите работу с Apache Superset: научитесь работать с датасетами, строить визуализации и проектировать дашборды под бизнес-задачи. Настройте фильтры и интерактивность, а главное — превратите набор графиков в историю с понятными выводами и рекомендациями.

На практике пройдёте полный цикл: от загрузки сырых данных до публикации готового дашборда. Заполните Dashboard Canvas, разработаете макет, соберёте визуализации в Superset и настройте взаимодействие между ними. Затем отберёте ключевые инсайты, подготовите презентацию и сформулируете рекомендации, понятные стейкхолдерам.

# Содержание

---

**Темы**

1. Первые шаги в Superset
2. Разработка дашборда в Superset
3. Проектирование Dashboard Canvas
4. Презентация аналитических выводов

---

**Проект** Для заданного бизнес-кейса на основе данных создадите интерактивный дашборд в Superset и подготовите короткую презентацию с выводами и рекомендациями.

---

**Инструменты и технологии**

- Apache Superset
- Dashboard Canvas
- SQL
- Jinja

---

# Статистика и эксперименты: A/B-тесты end-to-end

05

2 недели

Освоите полный цикл проведения A/B-тестов: от планирования и валидации до оценки результатов и принятия бизнес-решений.

На практике сначала спроектируете дизайн-эксперимент под конкретную бизнес-гипотезу: определите метрики, единицу рандомизации и длительность, оцените риски. Затем проанализируете результаты реального теста — проверите корректность его проведения, оцените эффект и примете аргументированное решение, которое ляжет в основу рекомендаций для бизнеса.

## Содержание

Темы

1. Планирование эксперимента
2. Валидация эксперимента
3. Оценка эффекта: классические методы
4. Оценка эффекта: продвинутые методы
5. Принятие решения и коммуникация

Проект

Разработаете дизайн A/B-теста и проанализируете результаты готового эксперимента с выводом для бизнеса.

Инструменты  
и технологии

- Статистические методы: t-тест, бутстрап, доверительные интервалы
- Продвинутые техники: CUPED, стратификация, дельта-метод
- Расчёт размера выборки и MDE
- Python: pandas, numpy, matplotlib, seaborn, scipy, statsmodels
- Jupyter Notebook

# Causal Inference

06

2 недели

Освоите методы каузального анализа для ситуаций, когда A/B-тест невозможен. Научитесь применять Diff-in-Diff, Causal Impact и Propensity Score Matching, проверять допущения параллельных трендов и корректно интерпретировать результаты.

Поработаете с реальными данными и оцените эффект от постепенного редизайна, который внедрялся без A/B-теста. Сформируете корректные группы, примените Diff-in-Diff для расчёта causal-эффекта на ключевые метрики и проверите выполнение допущений. Сравните результаты с наивным подходом «до/после», увидите, как сильно может искажаться оценка без учёта каузальности, и подготовите бизнес-отчёт с интерпретацией и ограничениями.

# Содержание

---

Темы	<ol style="list-style-type: none"><li>1. Основы каузального анализа</li><li>2. Проектирование квази-эксперимента</li><li>3. Расчёт эффекта</li><li>4. Проверка допущения параллельных трендов</li><li>5. Интерпретация результатов</li></ol>
Проект	Оцените эффект воздействия методом Diff-in-Diff и сравните с наивной оценкой «до/после».
Инструменты и технологии	<ul style="list-style-type: none"><li>• Diff-in-Diff</li><li>• Python: pandas, numpy, statsmodels, linearmodels</li><li>• Jupyter Notebook</li></ul>

---

## Machine Learning для аналитических задач

07

2 недели	<p>Научитесь применять ML для решения бизнес-задач: классификации, регрессии и прогнозирования. Освоите полный цикл работы — постановку задачи и выбор метрик, построение базовых и продвинутых моделей, оценку качества и интерпретацию результатов в терминах бизнеса.</p> <p>На практике пройдёте все этапы ML-проекта. Определите тип задачи под конкретную бизнес-цель, подготовьте данные, построите baseline и сложные модели, сравните их и выберете лучшую. Затем проинтерпретируете влияние признаков и подготовите краткий отчёт с выводами и рекомендациями.</p>
----------	--

# Содержание

---

Темы	<ol style="list-style-type: none"><li>1. Постановка ML-задачи и выбор метрик</li><li>2. Построение базовых моделей</li><li>3. Построение продвинутых моделей</li><li>4. Оценка и сравнение моделей</li><li>5. Интерпретация модели и бизнес-выводы</li></ol>
Проект	На исторических данных о пользователях решите бизнес-задачу. Постройте и сравните несколько моделей, интерпретируете результаты и подготовите выводы для бизнеса.
Инструменты и технологии	<ul style="list-style-type: none"><li>• Линейная регрессия</li><li>• Логистическая регрессия</li><li>• Random Forest</li><li>• Gradient Boosting</li><li>• Python: pandas, numpy, matplotlib, seaborn, scikit-learn</li><li>• Jupyter Notebook</li></ul>

---

2 недели

Освоите применение больших языковых моделей для решения аналитических задач: сбор и обработка данных, генерация и отладка кода, создание визуализаций и формулировка выводов. Научитесь критически оценивать результаты работы ИИ, отделять факты от галлюцинаций и понимать, где автоматизация действительно уместна. Соберёте данные с помощью LLM, обработаете их, построите визуализации, проверите и доработаете сгенерированный код, а затем оформите результаты в виде таблиц, графиков и текстовых выводов.

## Содержание

Темы

1. Сбор данных через ИИ
2. Анализ и обработка данных
3. Визуализация данных
4. Валидация и доработка сгенерированного кода
5. Аналитические выводы с ИИ

Проект

Выполните исследовательский проект с опорой на ИИ по одной из предложенных тем. Готовую работу можно добавить в портфолио.

Инструменты и технологии

- LLM
- Промптинг, генерация кода
- Jupyter Notebook
- Python: pandas, numpy, matplotlib, seaborn, plotly, requests, BeautifulSoup, Selenium

Расширенная версия

Максимальная версия

2 недели

Освоите принципы проектирования хранилищ данных: познакомитесь с подходом «факты и измерения», научитесь различать OLTP и OLAP-системы, проектировать схемы «звезда» и работать с историчностью данных через SCD Type 2. Поймёте, как устроено слоирование данных и где проходит граница ответственности аналитика и дата-инженера.

Спроектируете аналитическую витрину под бизнес-метрики.

Определите факты и измерения с гранулярностью, выберете измерения с историчностью. Напишите SQL для фактовой таблицы, измерения с SCD Type 2 и представления для агрегаций. Получите готовую схему витрины и код для реальных проектов.

## Содержание

Темы	<ol style="list-style-type: none"><li>1. Хранилище и аналитическое моделирование данных</li><li>2. Историчность данных и Slowly Changing Dimensions</li><li>3. Слоирование и организация хранилища</li><li>4. Проектирование витрин данных</li></ol>
------	--

Проект	На основе сырых данных из Postgres спроектируете витрину по схеме «звезда» с учётом историчности, напишете SQL для ключевых таблиц и представления.
--------	---

Инструменты и технологии	<ul style="list-style-type: none"><li>• PostgreSQL</li><li>• Моделирование данных (факты, измерения, гранулярность)</li><li>• Slowly Changing Dimensions (SCD Type 2)</li><li>• Слоирование: raw, staging, mart</li></ul>
--------------------------	---

# Оркестрация для аналитиков

# 10

Расширенная версия

Максимальная версия

2 недели

Научитесь управлять вычислительными процессами с Apache Airflow: проектировать направленные ациклические графы, выстраивать ETL- и ELT-пайплайны, внедрять контроль качества данных и отлаживать задачи в боевом режиме. Разберётесь, как обеспечить стабильность расчётов и в каких случаях выбирать тот или иной подход к трансформациям.

На практике настройте среду с Airflow и Postgres, разработаете граф для движения данных от сырья до витрины. Напишите трансформации для промежуточных слоёв, добавьте проверки на целостность и уникальность, запустите процесс и разберёте логи.

В результате получите работающий конвейер с автоматическим пересчётом метрик и встроенным контролем.

# Содержание

---

Темы	<ol style="list-style-type: none"><li>1. Запуск и настройка окружения</li><li>2. Проектирование DAG</li><li>3. Реализация ETL/ELT-трансформаций</li><li>4. Встраивание проверок качества</li><li>5. Мониторинг, отладка и продакшн-практики</li></ol>
------	---

---

Проект	В подготовленном окружении создадите DAG, который последовательно очищает сырые данные, формирует витрину и проверяет качество на ключевых этапах.
--------	--

---

Инструменты и технологии	<ul style="list-style-type: none"><li>• Apache Airflow</li><li>• PostgreSQL</li><li>• ETL/ELT-пайплайны</li><li>• Мониторинг и отладка</li><li>• SQL</li><li>• Python</li></ul>
--------------------------	---

---

## ClickHouse

11

Расширенная версия

Максимальная версия

2 недели

Изучите особенности колоночной СУБД ClickHouse, ориентированной на аналитические запросы. Освойте проектирование таблиц с эффективным партиционированием и первичным ключом, научитесь создавать материализованные представления для ускорения отчётов и оценивать влияние структуры данных на производительность.

Погрузитесь в ClickHouse. Создадите агрегирующее материализованное представление, выполните несколько запросов и сравните скорость выполнения. По итогам оформите обоснование принятых архитектурных решений.

## Содержание

---

Темы	<ul style="list-style-type: none"><li>• Подготовка среды ClickHouse</li><li>• Проектирование таблиц</li><li>• Оптимизация производительности</li></ul>
------	--

---

Проект	На данных спроектируете аналитический слой в ClickHouse: настроите партиционирование, создадите материализованное представление и проверите его эффективность.
--------	--

---

Инструменты и технологии	<ul style="list-style-type: none"><li>• ClickHouse (MergeTree, Materialized Views)</li><li>• Оптимизация аналитических запросов</li><li>• SQL</li><li>• CSV</li></ul>
--------------------------	---

---

Расширенная версия

Максимальная версия

## 2 недели

Познакомитесь с концепцией Data Lake и современными подходами к хранению больших данных. Изучите форматы Parquet, Delta Lake, принципы работы Apache Spark и научитесь проектировать слоистую архитектуру (raw/staging/curated). Поймёте, как выбор формата и партиционирования влияет на производительность обработки.

Спроектируете структуру Data Lake под задачи продуктовой аналитики. Загрузите сырые данные событий мобильного приложения в raw-слой, выполните очистку и объединение в Spark, рассчитаете ключевые метрики — DAU, выручку и конверсию.

## Содержание

### Темы

- Проектирование Data Lake
- Очистка и объединение в Spark
- Расчёт продуктовых метрик
- Parquet и партиционирование

### Проект

На данных спроектируете Data Lake, обработаете данные в Spark и рассчитаете продуктовые метрики с сохранением в Parquet.

### Инструменты и технологии

- Apache Spark
- Parquet, Delta Lake
- Data Lake архитектура (raw/staging/curated)
- Оптимизация хранения и партиционирование
- CSV

Максимальная версия

2 недели

Погрузитесь в архитектуру RAG (Retrieval-Augmented Generation) и научитесь создавать ассистентов, которые отвечают на вопросы по документам. Освойте семантический поиск, работу с эмбедингами и векторными базами данных, а также интеграцию с внешними источниками через MCP-серверы.

Соберёте RAG-ассистента на no-code платформе: подготовите документы, настройте семантический поиск и проверите качество ответов. Научитесь подключать бота, чтобы ассистент стал доступен в интерактивном формате.

## Содержание

Темы

- Архитектура и работа RAG
- Создание RAG-ассистента
- Настройка семантического поиска
- Разработка логики RAG-ассистента
- RAG-ассистент в Telegram

Проект

Создадите Telegram-бота с RAG-архитектурой, способного искать информацию по документам и отвечать на вопросы пользователей.

Инструменты и технологии

- RAG-архитектура (retrieval, embeddings, генерация)
- Векторные базы данных
- No-code платформы
- MCP-серверы для внешних источников
- Семантический поиск
- LLM
- Промпт-инжиниринг
- Telegram Bot API
- python-telegram-bot/aiogram

# Privacy & Compliance

Максимальная версия

1 неделя

Разберётесь в основах работы с чувствительными данными: научитесь определять прямые и косвенные идентификаторы, оценивать риски деанонимизации и применять принципы Privacy by Design при проектировании витрин. Поймёте, как соблюдать требования GDPR и 152-ФЗ, не жертвуя аналитической ценностью данных.

Научитесь проводить аудит готовой витрины пользователей: выявите поля, которые могут привести к утечке или деанонимизации, оцените риски и предложите меры по их снижению. Затем спроектируете новую версию витрины — с исключением лишних полей, агрегацией или псевдонимизацией.

# Содержание

---

Темы	<ul style="list-style-type: none"><li>• Privacy &amp; Compliance в аналитике</li><li>• Ответственная работа с данными</li><li>• Privacy by Design: проектирование с учётом приватности</li><li>• Аудит аналитических витрин на PII</li></ul>
------	--

---

Проект	Для витрины данных проведёте privacy-аудит, переработаете структуру с учётом конфиденциальности и подготовите рекомендации для бизнеса.
--------	---

---

Инструменты и технологии	<ul style="list-style-type: none"><li>• Идентификация PII и квази-идентификаторов</li><li>• Принципы GDPR и 152-ФЗ</li><li>• Методы анонимизации и псевдонимизации</li><li>• Privacy by Design</li><li>• GDPR</li><li>• Python: pandas</li><li>• SQL</li></ul>
--------------------------	--

---

## Soft Skills аналитика

15

Максимальная версия

---

1 неделя	<p>Научитесь работать с неопределённостью в данных и доносить выводы с учётом ограничений. Освоите аргументацию выбранных методов, научитесь защищать свою позицию перед бизнесом и конструктивно вести диалог в конфликтных ситуациях.</p> <p>В рамках спринта сможете подготовить защиту одного из своих проектов (или учебного кейса) и потренироваться формулировать цель, обосновывать выбор методов, представлять результаты с указанием уровня неопределённости. При желании проведёте встречу с наставником, чтобы получить обратную связь по качеству выводов, обоснованности подходов и коммуникации.</p>
----------	---

# Содержание

---

Темы	<ol style="list-style-type: none"><li>1. Работа с данными и неопределённостью</li><li>2. Защита методологии и аргументация</li><li>3. Работа с бизнесом и конфликтные ситуации</li></ol>
------	--

---

Инструменты и технологии	<ul style="list-style-type: none"><li>• Аргументация решений</li><li>• Коммуникация с бизнесом</li><li>• Управление неопределённостью</li><li>• Презентационные навыки</li><li>• Конфликт-менеджмент</li><li>• Рефлексия и самоанализ</li></ul>
--------------------------	---

---

# Бонусный модуль «Подготовка к собеседованиям»

Максимальная версия

Подготовка к собеседованиям недоступна при оплате курса работодателем

2 недели

Спринт-задачник для отработки навыков, необходимых на интервью для позиции Middle-аналитика. Все материалы приближены к реальным заданиям, которые встречаются на собеседованиях.

Вы потренируетесь решать SQL-задачи средней сложности с оконными функциями и CTE, выполните алгоритмические задания на Python и задачи на pandas. Разберёте ключевые статистические концепции и проанализируете бизнес-кейсы из разных доменов. На практике решите серию типовых задач по каждому направлению, при желании проведёте mock-интервью с наставником, а по итогам составите индивидуальный чек-лист подготовки и список тем для повторения.

## Содержание

Темы

1. SQL-секции
2. Python / pandas-задачи
3. Статистика и теория вероятностей
4. Разбор бизнес-кейсов
5. На что обращают внимание на собеседовании

Инструменты  
и подходы

- SQL (оконные функции, CTE, оптимизация)
- Python, pandas
- Статистика и теория вероятностей
- Разбор бизнес-кейсов
- Mock-интервью и самопрезентация