

МИНИСТЕРСТВО ТРАНСПОРТА РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«РОССИЙСКИЙ УНИВЕРСИТЕТ ТРАНСПОРТА»
(РУТ (МИИТ))



Рабочая программа дисциплины (модуля),
как компонент образовательной программы
базового высшего образования
по направлению подготовки
09.03.01 Информатика и вычислительная техника,
утвержденной первым проректором ФГАОУ ВО
РУТ (МИИТ) Тимониным В.С.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Бизнес-анализ для Data Engineering в транспортной отрасли

Направление подготовки: 09.03.01 Информатика и вычислительная техника

Направленность (профиль): IT-сервисы и технологии обработки данных на транспорте

Форма обучения: Очная

Рабочая программа дисциплины (модуля) в виде
электронного документа выгружена из единой
корпоративной информационной системы управления
университетом и соответствует оригиналу

Простая электронная подпись, выданная ФГАОУ ВО РУТ
(МИИТ)

ID подписи: 937226

Подписал: директор центра Проневич Ольга Борисовна

Дата: 21.10.2026

1. Общие сведения о дисциплине (модуле).

Целью освоения дисциплины «Бизнес-анализ для Data Engineering в транспортной отрасли» является формирование у обучающихся компетенций в области проектирования, разработки и эксплуатации систем сбора, хранения и анализа больших данных, генерируемых объектами транспортной инфраструктуры и подвижным составом. Дисциплина нацелена на подготовку специалистов, способных выстраивать сквозные конвейеры данных – от телематического датчика до витрины ситуационного центра – и принимать инженерные и управленческие решения на основе методов бизнес-анализа в логистике, пассажирских перевозках и управлении техническим состоянием парка транспортных средств.

Задачи: Изучение специфики источников данных на транспорте (бортовые системы, CAN-шины, системы взимания платы, АСУ ТП инфраструктуры), стандартов их передачи и протоколы потокового обмена (MQTT, Kafka); освоение методов и инструментов Data Engineering для построения масштабируемых корпоративных хранилищ и озер данных (Data Lake / Data Vault), ориентированных на геопространственные и временные ряды транспортной телематики; получение практических навыков нормализации/денормализации данных для аналитических задач и разработки ER-диаграмм и схем связей сущностей; развитие системного мышления в области обеспечения достоверности, безопасности и отказоустойчивости данных в соответствии с требованиями транспортной безопасности и цифровой трансформации отрасли.

2. Планируемые результаты обучения по дисциплине (модулю).

Перечень формируемых результатов освоения образовательной программы (компетенций) в результате обучения по дисциплине (модулю):

АД-2 - Способен проектировать, реализовывать и оптимизировать витрины данных для поддержки бизнес-аналитики.

Обучение по дисциплине (модулю) предполагает, что по его результатам обучающийся будет:

Знать:

Знает принципы нормализации/денормализации данных для аналитических задач.

Уметь:

Умеет разрабатывать ER-диаграммы и схемы связей сущностей для витрин.

Владеть:

Способен проектировать, реализовывать и оптимизировать витрины данных для поддержки бизнес-аналитики.

3. Объем дисциплины (модуля).

3.1. Общая трудоемкость дисциплины (модуля).

Общая трудоемкость дисциплины (модуля) составляет 3 з.е. (108 академических часа(ов)).

3.2. Объем дисциплины (модуля) в форме контактной работы обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации образовательной программы на иных условиях, при проведении учебных занятий:

Тип учебных занятий	Количество часов	
	Всего	Семестр №3
Контактная работа при проведении учебных занятий (всего):	48	48
В том числе:		
Занятия лекционного типа	16	16
Занятия семинарского типа	32	32

3.3. Объем дисциплины (модуля) в форме самостоятельной работы обучающихся, а также в форме контактной работы обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации образовательной программы на иных условиях, при проведении промежуточной аттестации составляет 60 академических часа (ов).

3.4. При обучении по индивидуальному учебному плану, в том числе при ускоренном обучении, объем дисциплины (модуля) может быть реализован полностью в форме самостоятельной работы обучающихся, а также в форме контактной работы обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации образовательной программы на иных условиях, при проведении промежуточной аттестации.

4. Содержание дисциплины (модуля).

4.1. Занятия лекционного типа.

№ п/п	Тематика лекционных занятий / краткое содержание
1	<p>Тема 1. Введение в бизнес-анализ и Data Engineering в цифровой экосистеме транспорта.</p> <p>Рассматриваются вопросы:</p> <p>Роль бизнес-аналитика и Data-инженера на стыке транспортной отрасли и IT.</p> <p>Цифровая зрелость транспортных компаний: от АСУ ТП до предиктивной аналитики.</p> <p>Обзор MaaS (Mobility as a Service) и цифровых логистических платформ.</p> <p>Потоки данных в транспортном хабе: телематика, пассажиропоток, транзакции, погода, дорожная обстановка.</p>
2	<p>Тема 2. Источники данных на транспорте и стандарты их передачи.</p> <p>Рассматриваются вопросы:</p> <p>Телематические протоколы: CAN-шина, FMS-стандарт, Wialon IPS.</p> <p>Данные систем взимания платы (ETC, Free Flow), валидация билетов (AFC).</p> <p>AIS (Автоматическая идентификационная система) на водном транспорте и ADS-B в авиации.</p> <p>Работа с сырыми (raw) логами: стандартизация временных рядов с метриками скорости, координат и расхода топлива.</p>
3	<p>Тема 3. Специфика сбора и потоковой обработки телематических данных. Часть 2.</p> <p>Рассматриваются вопросы:</p> <p>Архитектура IoT-платформ для транспортных средств (MQTT, брокеры сообщений).</p> <p>Потоковая обработка (Stream Processing): Apache Kafka и Apache Flink для детекции отклонений от маршрута (геозоны).</p> <p>Протоколы MTConnect и OPC-UA для железнодорожной инфраструктуры и депо.</p> <p>Обработка событий комплексов фотовидеофиксации (распознавание номеров, типов ТС) в режиме реального времени.</p>
4	<p>Тема 4. Проектирование хранилищ данных для транспортной аналитики.</p> <p>Рассматриваются вопросы:</p> <p>Архитектура Data Lake и Data Vault 2.0 для логистического оператора.</p> <p>Моделирование медленно меняющихся измерений (SCD) для справочников транспортных средств и водителей.</p> <p>Проектирование витрин данных: анализ средней скорости доставки, коэффициент выпуска парка на линию.</p> <p>Специфика хранения гео-треков и маршрутных графов (PostGIS, H3-индексы).</p>
5	<p>Тема 5. Управление пропускной способностью и расписанием.</p> <p>Рассматриваются вопросы:</p> <p>Анализ «Узких мест»: сбор данных с индуктивных петель и датчиков веса в дорожном покрытии.</p> <p>Предиктивная аналитика пассажиропотока на основе исторических данных транзакций (АСОП).</p> <p>Имитационное моделирование (AnyLogic) и оптимизация графика движения на основе больших данных.</p> <p>Визуализация композитного показателя эффективности (KPI) расписания.</p>
6	<p>Тема 6. Надежность, техническое обслуживание и ремонт (ТОиР) на основе данных.</p> <p>Рассматриваются вопросы:</p> <p>Концепция Prescriptive Maintenance для локомотивов и воздушных судов.</p> <p>Анализ временных рядов вибрации и пробега, прогнозирование остаточного ресурса (RUL).</p> <p>Создание Feature Store для инженерных признаков износа узлов и агрегатов.</p> <p>Цифровой двойник инфраструктуры: мониторинг состояния пути (рельсовая дефектоскопия) в BI-системах.</p>
7	<p>Тема 7. Экономика и безопасность транспортных данных. Часть 3.</p> <p>Рассматриваются вопросы:</p> <p>Юнит-экономика рейса: расчет маржинальности на основе машинных данных (факт vs план).</p> <p>Выявление аномалий и фрода: скрученный пробег, несанкционированные сливы топлива, подмены</p>

№ п/п	Тематика лекционных занятий / краткое содержание
	<p>карт тахографа.</p> <p>Аналитика «Карт риска» на основе открытых данных ДТП и геоданных погоды.</p> <p>Требования транспортной безопасности (ФЗ-16 «О транспортной безопасности») к архитектуре хранения данных и обезличиванию.</p>
8	<p>Тема 8. MLOps и искусственный интеллект в производственных процессах транспорта.</p> <p>Рассматриваются вопросы:</p> <p>Жизненный цикл ML-модели в транспортной задаче (от планирования до мониторинга дрейфа данных).</p> <p>Применение компьютерного зрения (CV) в задачах контроля безопасности (усталость водителя, занятость полосы).</p> <p>Построение рекомендательных систем интервалов движения для общественного транспорта.</p> <p>Внедрение федеративного обучения для кооперации перевозчиков без нарушения приватности.</p>

4.2. Занятия семинарского типа.

Практические занятия

№ п/п	Тематика практических занятий/краткое содержание
1	<p>Тема 1. Проектирование архитектуры данных транспортного предприятия.</p> <p>Рассматриваются вопросы:</p> <p>Анализ функциональных требований оператора грузоперевозок.</p> <p>Построение концептуальной схемы потоков данных (Context Diagram, DFD).</p> <p>Выбор целевого стека: хранение, потоковая обработка, BI.</p>
2	<p>Тема 2. Парсинг и валидация телеметрических пакетов (FMS/CAN).</p> <p>Рассматриваются вопросы:</p> <p>Чтение бинарного лога рейса транспортного средства.</p> <p>Декодирование параметров: обороты двигателя, уровень топлива, положение педали газа.</p> <p>Очистка выбросов и интерполяция потерь GPS-сигнала в туннелях.</p>
3	<p>Тема 3. Работа с геоданными: построение и анализ треков.</p> <p>Рассматриваются вопросы:</p> <p>Загрузка GPX/CSV треков в гео-расширение PostgreSQL (PostGIS).</p> <p>Пространственные SQL-запросы: поиск остановок в геозоне, расчет пройденного расстояния.</p> <p>Визуализация слоев маршрутов на карте (QGIS / Kepler.gl).</p>
4	<p>Тема 4. Потоковая обработка в реальном времени: детекция заезда в геозону.</p> <p>Рассматриваются вопросы:</p> <p>Развертывание связки Kafka + Spark Streaming (или Flink).</p> <p>Написание обработчика для расчета задержки рейса.</p> <p>Генерация алертов при превышении времени стоянки в неполюженном месте.</p>
5	<p>Тема 5. Моделирование Хранилища данных (DWH) для анализа логистики.</p> <p>Рассматриваются вопросы:</p> <p>Проектирование схемы «Звезда» для факта «Рейс».</p> <p>Создание таблиц в Greenplum или ClickHouse.</p> <p>Реализация загрузки исторического слоя (batch ingestion) с контролем качества (Great Expectations).</p>
6	<p>Тема 6. Анализ нарушения температурного режима (Cold Chain).</p> <p>Рассматриваются вопросы:</p> <p>Анализ временных рядов с датчиков температуры рефрижератора.</p> <p>Расчет триггеров нарушения условий хранения при перевозке скоропорта.</p> <p>Построение дашборда целостности холодовой цепи.</p>

№ п/п	Тематика практических занятий/краткое содержание
7	<p>Тема 7. Расчет KPI парка транспортных средств.</p> <p>Рассматриваются вопросы:</p> <p>Расчет технической готовности (Kcoef. Technical Readiness).</p> <p>Анализ пробега и коэффициент использования (KPI Utilization).</p> <p>Визуализация в Apache Superset / Yandex DataLens.</p>
8	<p>Тема 8. Предиктивная аналитика пассажиропотока на остановке.</p> <p>Рассматриваются вопросы:</p> <p>Подготовка данных валидаторов общественного транспорта.</p> <p>Построение модели временного ряда (Prophet / SARIMA) в Python.</p> <p>Оценка прогноза на часы пик и расчет потребного подвижного состава.</p>
9	<p>Тема 9. Построение цифрового двойника маршрута (AnyLogic / Simulation).</p> <p>Рассматриваются вопросы:</p> <p>Импорт OSM-карт для железнодорожной ветки.</p> <p>Моделирование интервального движения поездов.</p> <p>Сбор симуляционных данных для оценки загруженности станций.</p>
10	<p>Тема 10. Анализ стиля вождения и безопасности (Driver Scorecard).</p> <p>Рассматриваются вопросы:</p> <p>Выделение событий резкого торможения/ускорения (feature engineering).</p> <p>Расчет агрегированного скорингового балла водителя.</p> <p>Кластеризация водителей по стилю управления (K-Means, DBSCAN).</p>
11	<p>Тема 11. Выявление аномалий в данных топливных карт.</p> <p>Рассматриваются вопросы:</p> <p>Сопоставление транзакций заправки и данных уровня топлива в баке.</p> <p>Алгоритм поиска сливов (Isolation Forest).</p> <p>Формирование отчета по расследованию инцидентов (Data Storytelling).</p>
12	<p>Тема 12. Оптимизация маршрутов на основе данных о трафике (Vehicle Routing Problem).</p> <p>Рассматриваются вопросы:</p> <p>Работа с матрицами расстояний (OSRM / Valhalla API).</p> <p>Решение классической задачи коммивояжера с временными окнами (OR-Tools).</p> <p>Сравнение плановой и фактической эффективности нового маршрута.</p>
13	<p>Тема 13. Компьютерное зрение для транспортных узлов.</p> <p>Рассматриваются вопросы:</p> <p>Подключение к RTSP-потoku камеры (или обработка готового датасета).</p> <p>Детекция транспортных средств (YOLO) и подсчет трафика.</p> <p>Классификация типа ТС (легковой, грузовой) и фиксация нарушения ПДД (заезд на полосу).</p>
14	<p>Тема 14. Интеграция с внешними данными: погода и дорожные события.</p> <p>Рассматриваются вопросы:</p> <p>Подключение к API гидрометцентра или сервисов прогноза погоды.</p> <p>Парсинг RSS-лент дорожных служб.</p> <p>Корреляция опозданий рейсов с погодными факторами и коэффициентом сцепления.</p>
15	<p>Тема 15. Конвейер данных (Data Pipeline) для учета топлива.</p> <p>Рассматриваются вопросы:</p> <p>Разработка ETL-процесса в Apache Airflow.</p> <p>Слияние данных CAN-шины, датчиков уровня топлива и транзакций АЗС.</p> <p>Автоматическое формирование ведомости ГСМ.</p>
16	<p>Тема 16. Проектирование витрины для Ситуационного центра транспортной компании.</p> <p>Рассматриваются вопросы:</p>

№ п/п	Тематика практических занятий/краткое содержание
	Сбор требований от ЛПР (лиц, принимающих решения). Разработка макета дашборда: карта движения в реальном времени, статус выполнения плана перевозок, инциденты. Защита итогового проекта с обоснованием архитектуры данных.

4.3. Самостоятельная работа обучающихся.

№ п/п	Вид самостоятельной работы
1	Изучение литературы по проектированию витрин данных.
2	Изучение кейсов проектирования витрин в транспортных компаниях.
3	Разработка документации (Data Dictionary) для своей витрины.
4	Написание SQL-скриптов создания объектов витрины.
5	Подготовка к промежуточной аттестации.
6	Подготовка к текущему контролю.

5. Перечень изданий, которые рекомендуется использовать при освоении дисциплины (модуля).

№ п/п	Библиографическое описание	Место доступа
1	Анализ бизнес-процессов : учебное пособие / составители Т. В. Галанина, М. И. Баумгартэн. — Кемерово : КузГТУ имени Т.Ф. Горбачева, 2024. — 238 с. — ISBN 978-5-00137-479-4.	https://e.lanbook.com/book/498959
2	Моделирование и анализ бизнес-процессов : учебное пособие / составители Т. В. Галанина, М. И. Баумгартэн. — Кемерово : КузГТУ имени Т.Ф. Горбачева, 2023. — 164 с. — ISBN 978-5-00137-431-2.	https://e.lanbook.com/book/399725

6. Перечень современных профессиональных баз данных и информационных справочных систем, которые могут использоваться при освоении дисциплины (модуля).

MS Office (Excel, Word)

Notepad++

Браузер Chrome

7. Перечень лицензионного и свободно распространяемого программного обеспечения, в том числе отечественного производства, необходимого для освоения дисциплины (модуля).

Python 3.11/3.12 и JupyterLab / Jupyter Notebook
pandas, NumPy, SciPy, statsmodels и Pingouin
Matplotlib, Seaborn и Plotly
PostgreSQL и DBeaver Community Edition
Microsoft Excel и Power BI Desktop / LibreOffice Calc
Yandex DataLens, Metabase или Apache Superset
Git и GitHub / GitLab

8. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю).

Компьютер преподавателя с доступом к сети Интернет

Компьютеры студентов с установленным Python-окружением и табличным процессором

Доступ к учебной базе событийных данных и экспериментальным выборкам

Доступ к PostgreSQL и BI-системе для визуализации результатов

Доступ к облачной виртуальной машине или Docker-контейнерам для воспроизводимого анализа

Проектор и экран для проектора

Маркерная доска

9. Форма промежуточной аттестации:

Зачет в 3 семестре.

10. Оценочные материалы.

Оценочные материалы, применяемые при проведении промежуточной аттестации, разрабатываются в соответствии с локальным нормативным актом РУТ (МИИТ).

Авторы:

руководитель образовательной
программы

О.Б. Проневич

Согласовано:

Директор

Д.В. Паринов

Директор ЦС IT - программ ВИШ

О.Б. Проневич

Председатель учебно-методической
комиссии

Д.В. Паринов