

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



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

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)**

**Прикладной анализ данных**

Направление подготовки: 09.03.01 Информатика и вычислительная техника

Направленность (профиль): Цифровая инженерия транспортных процессов

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 РУТ (МИИТ)  
ID подписи: 937226  
Подписал: руководитель образовательной программы  
Проневич Ольга Борисовна  
Дата: 30.04.2026

## 1. Общие сведения о дисциплине (модуле).

Целью освоения дисциплины являются: формирование компетенций, позволяющих использовать соответствующий математический аппарат и инструментальные средства для анализа данных и извлечения информации, закономерностей, знаний и шаблонов из больших объемов необработанных данных.

Задачами освоения дисциплины являются:

- изучение методов сбора, подготовки и измерения данных;
- освоение статистических методов анализа (дескриптивный анализ, анализ связей, многомерный анализ, анализ временных рядов);
- применение методов кластеризации и классификации;
- развитие навыков постановки задачи анализа данных и интерпретации полученных результатов.

## 2. Планируемые результаты обучения по дисциплине (модулю).

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

**ПК-1** - Способен анализировать большие данные с использованием существующей в организации методологической и технологической инфраструктуры.

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

### **Знать:**

- типы шкал измерений (номинальная, порядковая, интервальная, относительная, дихотомическая);
- этапы решения задачи анализа данных;
- базовые методы статистического анализа (корреляционный, регрессионный, дисперсионный, факторный);
- постановки задач кластерного анализа и классификации;
- метрики расстояний и меры близости объектов.

### **Уметь:**

- применять дескриптивный анализ для описания исходных данных;
- проверять гипотезы о стационарности, нормальности, независимости, однородности;
- использовать методы регрессионного и корреляционного анализа для выявления связей;

- строить деревья решений, применять метод ближайшего соседа и метод k-средних.

**Владеть:**

- навыками работы с Google Colab, Jupyter Notebook;
- методами оценки качества кластеризации и классификации;
- интерпретацией коэффициентов линейной регрессии;
- базовыми подходами к анализу временных рядов.

3. Объем дисциплины (модуля).

3.1. Общая трудоемкость дисциплины (модуля).

Общая трудоемкость дисциплины (модуля) составляет 2 з.е. (72 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):	32	32
В том числе:		
Занятия лекционного типа	16	16
Занятия семинарского типа	16	16

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 40 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>Корреляционный анализ</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>- Сущность корреляции, ковариация. Коэффициенты Пирсона, Спирмена, Кендалла.</li> <li>- Проверка значимости, ограничения.</li> </ul>
2	<p>Регрессионный анализ</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>- Парная и множественная линейная регрессия. Метод наименьших квадратов.</li> <li>- Интерпретация коэффициентов. Коэффициент детерминации R<sup>2</sup>, стандартная ошибка.</li> </ul>
3	<p>Анализ временных рядов</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>- Тренд, сезонность, случайная компонента. Скользящее среднее, экспоненциальное сглаживание.</li> <li>- Индексы сезонности, коэффициенты неравномерности.</li> </ul>
4	<p>Сравнение средних: t-критерий Стьюдента и А/В тестирование</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>- Проверка гипотез, t-критерий для зависимых и независимых выборок.</li> <li>- А/В тестирование: метрики, размер выборки, мощность теста. Доверительные интервалы.</li> </ul>
5	<p>Дисперсионный анализ (ANOVA)</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>- Однофакторный ANOVA, сравнение трёх и более групп. Проверка условий применимости.</li> <li>- Транспортные кейсы (сравнение времени рейса по разным маршрутам).</li> </ul>
6	<p>ABC-анализ и Парето (правило 80/20)</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>- Принцип Парето, построение диаграммы. ABC-анализ: категории А, В, С.</li> <li>- Применение на транспорте: приоритизация причин задержек, управление запасами.</li> </ul>
7	<p>Процентильный анализ и оценка рисков (количественная)</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>- Процентили (90-й, 95-й), интерпретация.</li> <li>- Расчёт показателя RPN (Severity ? Occurrence ? Detection) в FMEA. Построение матрицы рисков.</li> </ul>
8	<p>Методы теории надёжности</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>- Интенсивность отказов, наработка на отказ (MTBF). Вероятность безотказной работы.</li> <li>- Применение к подвижному составу и инфраструктуре.</li> </ul>
9	<p>Диаграмма Исикавы («рыбья кость») и метод «5 почему»</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>- Структура причинно-следственной диаграммы. Последовательное выявление корневой причины.</li> <li>- Транспортные примеры (опоздания, аварии, поломки).</li> </ul>
10	<p>Анализ видов и последствий отказов (FMEA / FMECA)</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>- Цели, этапы, таблица FMEA. Определение Severity, Occurrence, Detection.</li> <li>- Ранжирование рисков по RPN.</li> </ul>
11	<p>Анализ дерева неисправностей (FTA) и дерева событий (ETA)</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>- Дедуктивный метод (FTA): построение дерева отказов.</li> <li>- Индуктивный метод (ETA): развитие сценариев от исходного события. Расчёт вероятностей (И, ИЛИ).</li> </ul>

№ п/п	Тематика лекционных занятий / краткое содержание
12	<p>Метод HAZOP и анализ барьеров безопасности (Bow-tie)</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>- HAZOP: ключевые слова (НЕТ, БОЛЬШЕ, МЕНЬШЕ). Анализ отклонений параметров.</li> <li>- Bow-tie: объединение дерева событий и дерева отказов. Барьеры (предотвращающие, смягчающие).</li> </ul>
13	<p>Экспертные методы: Дельфи, парные сравнения, SWIFT</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>- Итеративное анонимное анкетирование. Коэффициент конкордации Кендалла. Матрица парных сравнений.</li> <li>- SWIFT («Что, если?»): структурированное выявление рисков, интеграция с матрицей рисков.</li> </ul>
14	<p>Анализ «затраты-выгода» (Cost-Benefit Analysis)</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>- Расчёт суммарных затрат и выгод. Дисконтирование, чистая приведённая стоимость (NPV).</li> <li>- Принятие решений на транспорте (внедрение барьеров, замена парка).</li> </ul>
15	<p>Транспортно-социологические исследования и исследование безопасности дорожного движения</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>- Методы социологических исследований на транспорте: опросы пассажиров, анкетирование водителей.</li> <li>- Анализ безопасности ДД: очаги аварийности, коэффициенты тяжести и частоты ДТП.</li> </ul>
16	<p>Исследование параметров дорожного движения и технических средств организации движения</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>- Параметры ДД: интенсивность, скорость, плотность, задержки. Пропускная способность.</li> <li>- Эффективность ТСОДД: светофоры, знаки, разметка. Анализ конфликтных точек.</li> </ul>

## 4.2. Занятия семинарского типа.

### Практические занятия

№ п/п	Тематика практических занятий/краткое содержание
1	<p>Корреляционный анализ в Excel</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>- Расчёт коэффициента Пирсона (CORREL, пакет анализа). Расчёт коэффициента Спирмена.</li> <li>- Построение точечной диаграммы и тепловой карты.</li> </ul>
2	<p>Регрессионный анализ</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>- Парная регрессия (тренд на графике, функция ЛИНЕЙН). Множественная регрессия.</li> <li>- Интерпретация коэффициентов, R<sup>2</sup>, p-value, прогнозирование.</li> </ul>
3	<p>Анализ временных рядов</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>- Расчёт скользящего среднего для сглаживания пассажиропотока.</li> <li>- Вычисление индексов сезонности. Построение тренда и прогноза.</li> </ul>
4	<p>t-критерий и А/В тестирование</p> <p>Рассматриваемые вопросы:</p> <ul style="list-style-type: none"> <li>- Расчёт t-статистики (двухвыборочный t-тест). Сравнение времени в пути до/после изменений.</li> <li>- Расчёт размера выборки и доверительных интервалов.</li> </ul>

№ п/п	Тематика практических занятий/краткое содержание
5	<b>Дисперсионный анализ (ANOVA)</b> Рассматриваемые вопросы: - Однофакторный ANOVA в Excel. Сравнение трёх типов подвижного состава. - Интерпретация F-статистики и p-value.
6	<b>ABC-анализ и диаграмма Парето</b> Рассматриваемые вопросы: - Сортировка данных о причинах задержек, расчёт кумулятивного процента. - Построение диаграммы Парето. ABC-анализ для маршрутов или типов грузов.
7	<b>Процентили и матрица рисков</b> Рассматриваемые вопросы: - Расчёт 90-го и 95-го перцентиля времени задержки (ПЕРСЕНТИЛЬ.ВКЛ). - Построение матрицы рисков (вероятность ? тяжесть), раскраска ячеек.
8	<b>Расчёт показателей надёжности</b> Рассматриваемые вопросы: - Расчёт интенсивности отказов ?. Вычисление наработки на отказ MTBF. - Построение графика вероятности безотказной работы.
9	<b>Диаграмма Исикавы, метод «5 почему» и FMEA</b> Рассматриваемые вопросы: - Построение «рыбьей кости» для транспортной проблемы. Упражнение «5 почему». - Заполнение таблицы FMEA, расчёт RPN, ранжирование рисков.
10	<b>Дерево неисправностей (FTA) и HAZOP</b> Рассматриваемые вопросы: - Построение FTA с элементами И/ИЛИ. Расчёт вероятности вершинного события. - Заполнение таблицы HAZOP: отклонение > причина > последствия > меры.
11	<b>Диаграмма «галстук-бабочка» (Bow-tie)</b> Рассматриваемые вопросы: - Построение Bow-tie: угрозы > событие > последствия, барьеры. - Работа в PowerPoint или Excel с фигурами.
12	<b>Экспертные методы: Дельфи, парные сравнения, SWIFT</b> Рассматриваемые вопросы: - Мини-Дельфи (три тура), расчёт коэффициента конкордации Кендалла. - Матрица парных сравнений. Сессия «Что, если?» с оценкой рисков.
13	<b>Анализ «затраты-выгода» (NPV)</b> Рассматриваемые вопросы: - Таблица затрат и выгод для альтернатив. Расчёт ЧПС (функция ЧПС в Excel). - Выбор экономически эффективного решения.
14	<b>Исследования параметров дорожного движения и безопасности ДД</b> Рассматриваемые вопросы: - Обработка данных об интенсивности, скорости, задержках. Расчёт пропускной способности. - Выявление очагов аварийности, расчёт коэффициентов тяжести и частоты ДТП.
15	<b>Исследования параметров автомобильных дорог, путей сообщения и контактной сети</b> Рассматриваемые вопросы: - Анализ геометрических параметров дорог, расчёт коэффициента аварийности участка. - Оценка износа рельсов и контактного провода, прогнозирование отказов.
16	<b>Исследования парковок, экологии и пассажирских перевозок</b> Рассматриваемые вопросы: - Анализ загрузки парковок, расчёт коэффициента загрузки и оборачиваемости. - Расчёт выбросов CO?. Анализ пассажиропотока и наполнения подвижного состава.

#### 4.3. Самостоятельная работа обучающихся.

№ п/п	Вид самостоятельной работы
1	Подготовка к практическим занятиям
2	Подготовка к промежуточной аттестации
3	Подготовка к текущему контролю
4	Подготовка к промежуточной аттестации.
5	Подготовка к текущему контролю.

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

№ п/п	Библиографическое описание	Место доступа
1	Каган, Е. С. Прикладной статистический анализ данных : учебное пособие / Е. С. Каган. — Кемерово : КеМГУ, 2018. — 235 с. — ISBN 978-5-8353-2413-2.	<a href="https://e.lanbook.com/book/134318">https://e.lanbook.com/book/134318</a>
2	Лабковская, Р. Я. Анализ больших данных : учебное пособие / Р. Я. Лабковская, П. В. Косов. — Санкт-Петербург : СПбГУТ им. М.А. Бонч-Бруевича, 2025. — 152 с. — ISBN 978-5-89160-366-0	<a href="https://e.lanbook.com/book/508654">https://e.lanbook.com/book/508654</a>
3	Крутиков, В. И. Анализ данных : учебное пособие / В. И. Крутиков, В. В. Мешечкин. — Кемерово : КеМГУ, 2014. — 138 с. — ISBN 978-5-8353-1770-7	<a href="https://e.lanbook.com/book/61396">https://e.lanbook.com/book/61396</a>

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

<http://library.miiit.ru/> - электронно-библиотечная система Научно-технической библиотеки МИИТ

<https://e.lanbook.com/> - электронно-библиотечная система Лань

<https://proglib.io/p/raspoznovanie-obektov-s-pomoshchyu-yolo-v3-na-tensorflow-2-0-2020-11-08> - профессиональная библиотека программистов

[https://yandex.cloud/ru/blog/posts/2022/12/andrey-berger-and-yandex-cloud?utm\\_referrer=https%3A%2F%2Fyandex.ru%2F](https://yandex.cloud/ru/blog/posts/2022/12/andrey-berger-and-yandex-cloud?utm_referrer=https%3A%2F%2Fyandex.ru%2F) — библиотека профессиональных статей разработчиков Яндекс

<https://yandex.cloud/ru/blog> - библиотека профессиональных статей разработчиков Яндекс

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

Операционная система Microsoft Windows или аналог

Пакет программ Microsoft Office или аналог

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

Компьютер преподавателя

Компьютеры студентов

экран для проектора, маркерная доска,

Проектор

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

Руководитель образовательной  
программы

О.Б. Проневич

Председатель учебно-методической  
комиссии

Д.В. Паринов