ЧАСТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ «ВОСТОЧНАЯ ЭКОНОМИКО-ЮРИДИЧЕСКАЯ ГУМАНИТАРНАЯ АКАДЕМИЯ» (Академия ВЭГУ)

ОДОБРЕНА	УТВЕРЖДЕНА
Ученым советом Академии ВЭГУ	приказом ректора Академии ВЭГУ
(протокол от 28 июня 2021 г. , № 4)	от 31.08. 2021 № 71/а

Рабочая программа дисциплины
Методы анализа данных

Кафедра: управления, информатики и общенаучных дисциплин

Основная образовательная программа: 09.03.03 Прикладная информатика, направленности (профиля) «Проектирование и разработка информационных систем»

1. Общая характеристика

1.1 Наименование

Данная учебная дисциплина называется «Методы анализа данных», включена в Реестр автономных дидактических компонентов Академии ВЭГУ и реализуется в рамках ООП Академии ВЭГУ: 09.03.03 Прикладная информатика (профиль проектирование и разработка информационных систем) по заочной форме обучения, в т.ч. с использованием электронного обучения и дистанционных образовательных технологий.

1.2 Цели реализации

1.2.1 В результате освоения данной дисциплины обучающийся должны овладеть знаниями, умениями и навыками в рамках формирования следующих компетенций:

- способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач (УК-1);
- способен проводить обследование организаций, выявлять информационные потребности пользователей, формировать требования к информационной системе (ПК-1);
- способен использовать современные аналитические инструменты при решении задач, связанных с профессиональной деятельностью (ПК-10).

1.2.2 Обучающийся, освоивший данную дисциплину, должен знать:

- принципы сбора, отбора и обобщения информации, методики системного подхода для решения профессиональных задач;
- основные математические и статистические методы анализа больших объемов информации из различных предметных областей;

- современные программные средства анализа больших объемов информации.

1.2.3 Обучающийся, освоивший данную дисциплину, должен уметь:

- анализировать и систематизировать разнородные данные, оценивать эффективность процедур анализа проблем и принятия решений в профессиональной деятельности;

- анализировать и систематизировать входную информацию в больших объемах; проводить системный анализ предметной области, собирать и систематизировать информацию о структуре компании и ее бизнес-процессах;

- выбирать и использовать оптимальные инструментальные средства для анализа данных.

1.2.4 Обучающийся, освоивший данную дисциплину, должен владеть навыками:

- навыками научного поиска и практической работы с информационными источниками; методами принятия решений;

- навыками применения статистических методов для обработки и анализа больших объемов информации;

- навыками поиска информации, необходимой для профессиональной деятельности; методикой работы с современными программными средствами анализа больших объемов информации.

1.3 Место в структуре ООП

1.3.1 Данная дисциплина относится к блоку Б1 части, формируемой участниками образовательных отношений, дисциплина по выбору (Б1.В.ДВ.08.01) и изучается по заочной форме с применением ЭО и ДОТ на 6 семестре (на 3 курсе) обучения.

1.3.2 Логически и содержательно-методически данная дисциплина связана с такими автономными дидактическими компонентами данной ООП как: математика, информатика, программирование, информационные системы и технологии, теория вероятностей и математическая статистика, исследование операций и методы оптимизации, теория систем и системный анализ, базы данных, пакеты прикладных программ, введение в машинное обучение, интеллектуальные информационные системы.

1.3.3 Изучению данной дисциплины должно предшествовать освоение обучающимся программы: математика, информатика, программирование, теория вероятностей и математическая статистика.

1.3.4 Освоение данной дисциплины необходимо как предшествующее для прохождения обучения по программам: информационные технологии в управлении, WEB-программирование, интернет-технологии, интеллектуальные информационные системы, выполнение и защита выпускной квалификационной работы.

1.4 Объем

1.4.1 Общий объем данной дисциплины (трудоемкость учебной нагрузки обучающегося при освоении программы, включающая в себя все виды его учебной деятельности, предусмотренные учебным планом для достижения планируемых результатов обучения) составляет 5 зачетных единиц или 180 академических часов вне зависимости от формы обучения, применяемых образовательных технологий, реализации с использованием сетевой формы, реализации по индивидуальному учебному плану, в том числе при ускоренном обучении.

1.4.2 Объемы учебной нагрузки обучающегося при освоении программы дисциплины по видам учебной деятельности составляют:

Виды учебной деятельности	**Объем, в академических часах**
	По заочной форме с применением ЭО и ДОТ
Занятия лекционного типа	**4**
Занятия семинарского типа	**-**
Проектирование	**-**
Индивидуальная работа с обучающимся	**4**
Самостоятельная работа обучающегося	**142**
Аттестация	**30**
Всего	**180**

2. Структура и содержание

2.1 Содержание разделов и тем

Разделы и темы		**Содержание (дидактические единицы)**	**Учебные занятия**	
			заочная форма с применением ЭО и ДОТ	
№ п/п	**Наименование**		**виды**	**Объем, академических часов**
1	**2**	**3**	**4**	**5**
1	Ведение в анализ больших данных. Обзор источников информац	Основные определения, термины, задачи анализа больших данных. Вопросы безопасности. Понятие Data Mining. Когнитивный анализ данных. Обзор источников информации для Big Data (открытые источники информации: статистические сборники, опубликованные отчеты и результаты исследований; доступ к закрытой информации). Методики сбора данных.	Занятия лекционного типа	1
			Индивидуальная работа с обучающимся	1
		- проработка материала лекции;	Самостоя-	10

	ии.	- чтение специальной литературы; - подготовка к тестированию	тельная работа обучающегося	
2	Технологии хранения и обработки больших данных.	Обзор технологий хранения больших данных. Базы данных. Системы управления базами данных. Модели данных. Подготовка исходных данных для анализа: первичная обработка и визуализация имеющихся данных.	Занятия лекционного типа	0,5
			Индивидуальная работа с обучающимся	0,5
		- проработка материала лекции; - чтение специальной литературы; - подготовка к тестированию	Самостоятельная работа обучающегося	10
3	Статистические методы анализа данных.	Основные понятия математической статистики. Методы анализа данных: дескриптивная статистика, параметрические, непараметрические, номинальные методы (корреляционный, регрессионный, дисперсионный анализы, кластерный, дискриминантный, факторный анализы).	Занятия лекционного типа	0,5
			Индивидуальная работа с обучающимся	0,5
		- проработка материала лекции; - чтение специальной литературы; - подготовка к тестированию	Самостоятельная работа обучающегося	15
4	Современные программные средства анализа больших объемов информации.	Обзор современных популярных программных средства анализа данных: Statistica, SPSS, Excel, R-Studio и другие; их преимущества и недостатки.	Занятия лекционного типа	0,5
			Индивидуальная работа с обучающимся	0,5
		- проработка материала лекции; - чтение специальной литературы; - подготовка к тестированию	Самостоятельная работа обучающегося	15
5	Сбор и хранение больших данных.	Поиск источников информации в сети Интернет: открытые и закрытые источники данных. Портал открытых данных РФ. Сохранение данных в программе Excel. Преобразование и первичная обработка данных.	Занятия лекционного типа	0,5
			Индивидуальная работа с обучающимся	0,5
		- проработка материала лекции;	Самостоя-	25

		- чтение специальной литературы; - подготовка к тестированию; - выполнение письменной аттестационной работы	тельная работа обучающегося	
6	Методы обработки и анализа больших данных	Представление исходных данных в программе R-Studio (векторы, массивы, матрицы, списки, таблицы). Статистическая обработка данных в программах Excel и R-Studio: подсчет описательных статистик, графическое представление данных. Группировка данных, обнаружение значимых корреляций, зависимостей и тенденций в результате анализа имеющейся информации, выявления отношений между данными различного типа. Применение различных методов выделения, извлечения и группировки данных, которые позволяют выявить систематизированные структуры данных и вывести из них правила для принятия решений и прогнозирования их последствий (регрессионный, дисперсионный, кластерный, дискриминантный, факторный анализы).	Занятия лекционного типа	0,5
			Индивидуальная работа с обучающимся	0,5
		- проработка материала лекции; - чтение специальной литературы; - подготовка к тестированию; - выполнение письменной аттестационной работы	Самостоятельная работа обучающегося	42
7	Визуализация исходной информации и аналитических данных.	Возможности графического представления информации в программе R-Studio: графические функции отображения одномерных и многомерных данных, графический вывод с использованием графических параметров	Занятия лекционного типа	0,5
			Индивидуальная работа с обучающимся	0,5
		- проработка материала лекции; - чтение специальной литературы; - подготовка к тестированию; - выполнение письменной аттестационной работы	Самостоятельная работа обучающегося	25

2.2 Перечень обеспечения СРС

Студенту предлагается выполнить следующие виды самостоятельной работы:

– проработка конспекта лекции;

– анализ учебников, учебных пособий, специальной литературы по данной теме (с указанием страниц);

– проведение научных исследований;

– выполнение домашней контрольной работы, письменное или устное решение задач, разбор конкретных ситуаций;

– подготовка к дискуссии по определенной проблеме на базе прочитанной литературы, изучения нормативных актов, практики т.д.;

– подготовка списка литературы (библиографии) и подборка нормативных источников по определенной тематике, их изучение и т.д.).

При выполнении самостоятельной работы, обучающемуся предоставляется следующее учебно-методическое обеспечение:

- информационные ресурсы, перечисленные в разделе 4: Информационные ресурсы данной программы;
- электронные курсы, размещенные в вертикальном меню Кампуса ВЭГУ;
- материалы, размещенные в разделах Диск, Задачи, Обсуждение, Сообщение ПГ Методы анализа данных Кампуса ВЭГУ 24;
- материалы лекционных и практических занятий по дисциплинам бакалавриата 09.03.03 – Прикладная информатика, расположенные по адресу http://cp.insto.ru/extranet/

3. Фонд оценочных средств

3.1 Этапы формирования компетенций

Компетенция		Этапы формирования		
код	**содержание**	**знать**	**уметь**	**владеть навыками**
1	**2**	**3**	**4**	**5**
УК-1	способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	принципы сбора, отбора и обобщения информации, методики системного подхода для решения профессиональных задач	анализировать и систематизировать разнородные данные, оценивать эффективность процедур анализа проблем и принятия решений в профессиональной деятельности	навыками научного поиска и практической работы с информационными источниками; методами принятия решений
	Типовые контрольные задания: - дайте определение понятию «информационные ресурсы»; - объясните, что означает «информационный поиск»; - расскажите, что такое информационно-коммуникационные технологии; - перечислите основные компоненты процесса поиска информации; - определите понятие «информационные системы»; - охарактеризуйте портал открытых данных РФ; - определите сущность понятия «большие данные»; - определите понятие Data Mining; - расскажите, каковы главные проблемы безопасности «больших данных»; - дайте характеристику принципу безопасности «intelligence».			

ПК-1	- способен проводить обследование организаций, выявлять информационные потребности пользователей, формировать требования к информационной системе	основные математические и статистические методы анализа больших объемов информации из различных предметных областей	анализировать и систематизировать входную информацию в больших объемах; проводить системный анализ предметной области, собирать и систематизировать информацию о структуре компании и ее бизнес-процессах	навыками применения статистических методов для обработки и анализа больших объемов информации
	Типовые контрольные задания: - перечислите технологии хранения больших данных; - расскажите, какие модели данных вы знаете; - объясните, что включает первичная обработка данных; - определите различия между параметрическими, непараметрическими и номинальными методами; - опишите основную идею корреляционного анализа; - объясните, для чего применяют регрессионный анализ; - раскройте основную идею дисперсионного анализа; - раскройте цели применения факторного анализа; - проведите статистическую обработку данных в программе R-Studio			
ПК -10	способен использовать современные аналитические инструменты при решении задач, связанных с профессиональной деятельностью	современные программные средства анализа больших объемов информации	выбирать и использовать оптимальные инструментальные средства для анализа данных	навыками поиска информации, необходимой для профессиональной деятельности; методикой работы с современными программными средствами анализа больших объемов информации
	Типовые контрольные задания - перечислите программные средства анализа данных; - раскройте преимущества работа с данными в программе R-Studio; - расскажите, каковы возможности представления данных в программе R-Studio; - решите задачу корреляционного анализа, определите зависимости между переменными; - продемонстрируйте возможности графического представления информации в программе R-Studio: графические функции отображения одномерных и многомерных данных, графический вывод с использованием графических параметров			

3.2 Показатели, критерии и шкала оценивания

3.2.1 Для оценивания компетенций обучающегося на этапе их формирования по результатам освоения программы данной дисциплины применяется четырехбалльная шкала оценивания (оценки: «отлично», «хорошо», «удовлетворительно» или «неудовлетворительно»).

3.2.2 При применении технологий, использующих иные шкалы измерения (тестирование, балльно-рейтинговой, рейтинговой и т.д.), они для окончательного оформления переводятся в четырехбалльную шкалу по следующим параметрам: 90 и более процентов максимально-возможной суммы – «отлично», 70-89% - «хорошо», 50-69% - «удовлетворительно», менее 50 % - «неудовлетворительно».

3.2.3 При формировании оценки обучающегося используются следующие показатели и критерии оценивания результатов освоения программы данной дисциплины и соответствующего этапа формирования компетенций обучающегося:

Оценка	**Критерий**	**Индикатор (показатель)**
«отлично»	усвоение программы в полном объеме	задание выполнено без замечаний, полное и логически стройное изложение содержания при ответе или в отчете, тесное увязывание теории вопроса с практикой, отсутствие затруднений с объяснением всех аспектов выполнения задания, хорошее владение умениями и навыками по программе, знание монографической литературы, наличие умений самостоятельно обобщать и излагать материал
«хорошо»	твердое владение материалом в рамках программы	задание выполнено без существенных замечаний, грамотное изложение ответа (отчета), отсутствие существенных неточностей, правильное применение теоретических положений и владение необходимыми навыками при выполнении практических заданий
«удовлетворительно»	владение только основным материалом программы	задание в основном выполнено, допущение неточностей при правильном в основном ответе, нарушение последовательности в его изложении, неусвоение отдельных существенных деталей, наличие затруднений в выполнении практических заданий
«неудовлетворительно»	невладение значительной (и значимой) частью материала программы	задание не выполнено, допуск обучающимся при ответе принципиальных ошибок, большие затруднения при выполнении практических работ, ответ свидетельствует об отсутствии знаний по предмету

3.2.4 Аттестация по данной дисциплине может осуществляться по балльно-рейтинговой системе (далее - БРС), которая представляет собой строго последовательное прохождение обучающимся контрольных (реперных, рубежных) точек (далее – КТ-1, КТ-2 и т.д.) с получением оценки за качество показанных результатов в виде определенной конечной суммы баллов. При этом используются следующие индикаторы начисления баллов.

3.2.4.1 Общее количество баллов БРС распределяется следующим образом:

- за прохождение предварительного этапа – 20 баллов;
- за прохождение первой контрольной точки – до 20 баллов;
- за прохождение второй контрольной точки – до 20 баллов;
- за прохождение третьей контрольной точки – до 20 баллов.

3.2.4.2 Перевод набранной суммы по итогам всей БРС в четырехбалльную шкалу оценивания осуществляется по следующим параметрам: 72 и более баллов – «отлично», 56-71 баллов - «хорошо», 40-55 баллов - «удовлетворительно», менее 40 баллов - «неудовлетворительно».

3.2.4.3 За прохождение предварительного этапа начисляется до 20 баллов пропорционально изученным обучающимся разделам лекционного материала и набранным при тестировании в самом представленном для изучения ресурсе количеству баллов.

3.2.4.4 Начисление баллов по рубежной аттестации по первой контрольной точке осуществляется в зависимости от результативности участия на вебинаре.

3.2.4.4.1 При он-лайн участии на вебинаре баллы начисляются по следующим критериям:

1) 5 баллов, если участие ограничилось только присутствием или одним нерезультативным действием (вопрос или выступление не соответствовали теме);

2) 10 баллов, если были два и более нерезультативных действия;

3) 15 баллов, если среди двух или более произведенных обучающимся действий как минимум одно было результативным (правильное изложение материала, точно заданный вопрос, аргументированная и объективная рецензия);

4) 20 баллов, если все произведенные обучающимся два и более действий были результативными.

3.2.4.4.2 При просмотре вебинара в записи и предоставлении обучающимся письменного отзыва о нем, определяются следующие степени и суммы баллов:

1) 5 баллов, если отзыв написан формально, малосодержательно, но свидетельствует о просмотре всего вебинара;

2) 10 баллов, если в отзыве достаточно аргументировано выделены позитивная (что понравилась, было понятно, интересно) и (или) негативная (что не понравилась, было непонятным, неинтересным) стороны вебинара;

3) 15 баллов, если в отзыве есть обоснованное, логичное сопоставление позитивных и негативных итогов занятия;

4) 20 баллов, если отзыв в дополнение к третьему уровню содержит существенные предложения по улучшению организации вебинара или аргументировано описывает проблему, сформировавшуюся по итогам вебинара.

3.2.4.5 При тестировании по второй контрольной точке, если тестовые задания имеют разные степени сложности, каждые 5% максимально-возможной суммы правильных ответов приравнивается одному баллу БРС.

При равной сложности всех тестовых заданий каждый правильный ответ приносит 1 балл БРС.

3.2.4.6 Творческая аттестационная работа оценивается по шестибалльной шкале (0 баллов – отсутствует, 1 балл – имеется, но абсолютно не соответствует заданию и (или) дисциплине, 2 балла – в большей части не соответствует заданию (дисциплине), хотя есть определенное приближение к сути задания (дисциплины), 3 балла – суть задания выявлена, но неполно, 4 балла – имеются только отдельные неточности, 5 – нет претензий к исполнению) по следующим направлениям:

- определение и фиксация проблемы;
- формулирование ответа (рабочей гипотезы);
- аргументы и иллюстрации в пользу ответа (рабочей гипотезы);
- использование концептуального и понятийного аппарата дисциплины.

Сумма набранных баллов за все 4 направления является количеством баллов БРС, начисляемых в общий рейтинг за третью контрольную точку.

3.3 Типовые контрольные задания

3.3.1 При подготовке обучающегося к аттестации и при оценивании результатов освоения программы данной дисциплины (для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы) используются типовые контрольные задания по:

- истории развития той отрасли знания, изучению которой посвящена данная дисциплина (модуль); объекту, предмету и методам (инструментам), применяемым в данной дисциплине (модуле); месту, значению данной дисциплины (модуля) в деятельности человека и ее связям с другими дисциплинами (модулями); проблемам и задачам, решаемым в рамках данной дисциплины (модуля) и т.д.;
- существу теорий, концепций, систем описания и объяснения, гипотез, выдвигаемых в рамках данной дисциплины, а также категориям и понятиям (терминам), являющимся существенными для данной дисциплины;
- применению компетенций, сформированных в ходе освоения программы, для решения конкретной задачи, объяснения конкретного факта (явления), разрешения конкретной ситуации и т.д.

3.3.2 Данные типовые задания при проведении конкретных аттестационных испытаний переформатируются в вопросы аттестационных билетов или тестовые задания в соответствии с правилами, установленными для Фонда оценочных средств Академии ВЭГУ. По предметам вопросов типовых заданий могут быть сформулированы несколько различающихся по форме и аспекту рассмотрения вопросов аттестационных билетов, тестовых заданий или тем письменных

работ. Комплекты вопросов для аттестационных билетов, тестовых заданий и тем в виде соответствующих баз хранятся в Центре аттестации Академии ВЭГУ в режиме конфиденциальности и предъявляются в виде набора аттестационных билетов, тестов или отдельной темы, формируемых по установленным в Академии ВЭГУ правилам, на каждую конкретную аттестацию.

3.3.3 Типовые контрольные вопросы и задания для подготовки и проведения промежуточной аттестации:

– Определите сущность понятия «большие данные».
– Опишите методики анализа больших данных.
– Процесс аналитики анализа больших данных.
– Дайте характеристику Big Data на мировом рынке.
– Охарактеризуйте Big Data в России.
– Определите понятие Data Mining.
– Вопросы безопасности больших данных.
– В чем состоит когнитивный анализ данных.
– Какие модели данных вы знаете?
– Основные описательные статистики.
– Определите различия между параметрическими, непараметрическими и номинальными методами.
– Опишите основную идею корреляционного анализа.
– Регрессионный анализ.
– Основная идея дисперсионного анализа.
– Сущность кластерного анализа.
– Дискриминантный анализ: модель и общая процедура выполнения.
– Цели факторного анализа.
– Программные средства анализа данных: Statistica, SPSS, Excel; их преимущества и недостатки.
– Преимущества работа с данными в программе R-Studio.
– Представление исходных данных в программе R-Studio.
– Выполнение анализа данных в R-Studio.

3.3.4 Типовые (примерные) темы для письменных работ:

– Сбор данных. Информационные ресурсы. Портал открытых данных РФ. Сбор информации из различных источников в сети Интернет.
– Хранение и анализ данных в программе Excel.
– Знакомство с программой R-Studio. Синтаксис.
– Представление исходных данных в программе R-Studio.
– Преобразование исходных данных в программе R-Studio (векторы, массивы, матрицы, списки, таблицы).
– Статистическая обработка данных в программе R-Studio: подсчет описательных статистик.

– Графическое представление данных в R-Studio.

– Законы распределения вероятностей, реализованные в R.

– Корреляционный анализ. Определение зависимостей между переменными.

– Регрессионный анализ (линейная зависимость). Построение линейной модели. Проверка адекватности построенной модели.

– Регрессионный анализ (нелинейная зависимость). Определение вида зависимости. Построение модели. Проверка адекватности построенной модели.

– Однофакторный дисперсионный анализ.

– Многофакторный дисперсионный анализ.

– Факторный анализ.

– Кластерный анализ.

– Возможности графического представления информации в программе R-Studio: графические функции отображения одномерных и многомерных данных, графический вывод с использованием графических параметров.

3.4 Методические материалы по процедурам оценивания

3.4.1 Методика (в том числе технологические и организационные аспекты), определяющая процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций, установлены положениями об организации образовательной деятельности по программам высшего образования, о текущем контроле успеваемости и промежуточной аттестации, о балльно-рейтинговой системе аттестации, о фонде оценочных средств и изданными в соответствии с ними другими локальными нормативными актами, в т.ч. внутривузовскими стандартами, Академии ВЭГУ.

3.4.2 Обеспечение аттестации в Академии ВЭГУ должно удовлетворять, в первую очередь, требованиям нацеленности на результат обучения, системности, унификации и объективности, что на практике означает следующие:

- применение единых критериев оценивания по всем образовательным программам;

- сопряженность всех видов контрольно-измерительных материалов с конечными компетенциями и друг с другом (высокая степень валидности) и их репрезентативность с содержанием программы;

- выведение итоговых оценок по результатам проверки знаний, умений и навыков по всем дидактическим единицам оцениваемой программы;

- применение единых контрольно-измерительных материалов и процедур аттестации к обучающимся всех форм и технологий обучения;

- привлечение к осуществлению аттестации не задействованных в подпроцессе обучения обучающих, в т.ч. практикующих специалистов необразовательных организаций;

- широкое применение инструментальной среды;

- обязательная экспертиза текстов письменных работ (и других видов оформления результатов проектирования) на авторство (отсутствие заимствований, плагиата, копирования);

- ежегодное обновление тем письменных работ.

Порядок разработки, использования и хранения контрольных измерительных материалов (включая требования к режиму их защиты, порядку и условиям размещения информации, содержащейся в контрольных измерительных материалах, в сети «Интернет») устанавливается в соответствии с федеральными требованиями положением Академии ВЭГУ о фонде оценочных средств.

3.4.3 При организации аттестации обучающихся в Академии ВЭГУ обязательно исполнение следующих положений:

- аттестации подлежат только образовательные программы, назначенные для данного аттестуемого (правилами приема, учебными планами и другими соответствующими документами);

- аттестацию может проводить только тот обучающий (эксперт), который закреплен для данной процедуры по данной группе аттестуемых;

- аттестация проводится только с использованием утвержденных контрольно-измерительных материалов и по утвержденной форме и процедуре;

- аттестация проводится только для тех обучающихся, которые имеют соответствующий допуск;

- результаты аттестации должны быть зафиксированы в утвержденных для этого документах.

4. Информационные ресурсы

4.1 Основная литература

№ п/п	Выходные данные основной учебной литературы	Адрес доступа к полнотекстовому варианту в Электронно-библиотечной системе Академии ВЭГУ
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.	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.	http://www.iprbookshop.ru/26445.html

4.2 Дополнительная литература

№ п/п	Выходные данные основной учебной литературы	Адрес доступа к полнотекстовому варианту в Электронно-библиотечной системе Академии ВЭГУ

1	Илышев, А. М. Общая теория статистики : учебник / А. М. Илышев. — Москва : ЮНИТИ-ДАНА, 2012. — 536 с.	http://www.iprbookshop.ru/10504.html
2	Далингер, В. А. Теория вероятностей и математическая статистика с применением Mathcad : учебник и практикум для вузов / В. А. Далингер, С. Д. Симонженков, Б. С. Галюкшов. — 2-е изд., испр. и доп. — Москва : Издательство Юрайт, 2020. — 145 с.	https://urait.ru/bcode/452017
3	Мельниченко, А. С. Математическая статистика и анализ данных : учебное пособие / А. С. Мельниченко. — Москва : Издательский Дом МИСиС, 2018. — 45 с.	http://www.iprbookshop.ru/78563.html
4	Маккинли, Уэс Python и анализ данных / Уэс Маккинли ; перевод А. Слинкина. — 2-е изд. — Саратов : Профобразование, 2019. — 482 с.	http://www.iprbookshop.ru/88752.html
5	Воронов, В. И. Data Mining - технологии обработки больших данных : учебное пособие / В. И. Воронов, Л. И. Воронова, В. А. Усачев. — Москва : Московский технический университет связи и информатики, 2018. — 47 с.	http://www.iprbookshop.ru/81324.html
6	Программные системы статистического анализа. Обнаружение закономерностей в данных с использованием системы R и языка Python : учебное пособие / В. М. Волкова, М. А. Семёнова, Е. С. Четвертакова, С. С. Вожов. — Новосибирск : Новосибирский государственный технический университет, 2017. — 74 с.	http://www.iprbookshop.ru/91682.html
7	Синева, И. С. Анализ данных в среде R. Ч. 1 : учебное пособие / И. С. Синева. — Москва : Московский технический университет связи и информатики, 2018. — 32 с.	http://www.iprbookshop.ru/92422.html
8	Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход : монография / Б. Ю. Лемешко, С. Б. Лемешко, С. Н. Постовалов, Е. В. Чимитова. — Новосибирск : Новосибирский государственный технический университет, 2011. — 888 с.	http://www.iprbookshop.ru/47719.html
9	Александровская, Ю. П. Многомерный статистический анализ в экономике : учебное пособие / Ю. П. Александровская. — Казань : Казанский национальный исследовательский технологический универси-	http://www.iprbookshop.ru/79330.html

	тет, 2017. — 96 с.	
10	Яковлев, В. Б. Статистика. Расчеты в Microsoft Excel : учебное пособие для вузов / В. Б. Яковлев. — 2-е изд., испр. и доп. — Москва : Издательство Юрайт, 2020. — 353 с.	https://urait.ru/bcode/453051

4.3 Интернет-ресурсы

№ п/п	Наименование ресурса	Адрес доступа к полнотекстовому варианту (в ЭБС Академии ВЭГУ или других ресурсах в сети «Интернет»)
Современные профессиональные базы		
1	Образовательный математический сайт Exponenta.ru	http://old.exponenta.ru/
2	Вся математика в одном месте	http://www.allmath.ru/highermath.htm
3	Отделение математических наук РАН	http://math.ras.ru/
Информационные справочные системы		
1	Теория вероятностей и математическая статистика	http://zyurvas.narod.ru/resursy.html
2	Правовая система КонсультантПлюс	http://www.consultant.ru/
3	Информационно-правовой портал Гарант	http://www.garant.ru/
4	Официальный сайт Росстата	http://www.gks.ru
5	Общероссийский математический портал	http://www.mathnet.ru/
Иные ресурсы Интернет		
1	Математический форум	https://dxdy.ru/
2	Научный журнал «Фундаментальная и прикладная математика»	http://mech.math.msu.su/~fpm/rus/fpmosn.htm
3	Анализ и визуализация данных. Мастицкий С.Э., Шитиков В.К.	https://r-analytics.blogspot.com/p/blog-page.html
4	Статьи по Big Data	https://m7876.wiki.zoho.com/Introduction-to-R.html

4.4 Информационные технологии

4.4.1 АСО Академии ВЭГУ

Образовательный процесс по данной дисциплине в Академии ВЭГУ ведется с широким использованием Автоматизированной системы обучения. В частности, применяются следующие составные части (модули) АСО:

№ п/п	Полное наименование	Область применения в образовательной деятельности
1.	«1С-Битрикс: внутренний портал учебного заведения»	Платформа для интеграции всех сервисов и создания виртуальных рабочих кабинетов участников образовательного процесса. 1. Создание ЭИОС для обучающегося: - редактирование индивидуального учебного плана, обучающегося; - изучение материалов по доступным дисциплина (модулям) (электронный курс, материалы для самоконтроля и прохождение аттестации, расписание трансляций лекций, очных заня-

		тий и вебинаров, просмотр видео материалов); - средства электронных коммуникаций (форумы, комментарии, чат) для общения с участниками процесса обучения (в учебной и предметных группах); - просмотр электронной зачётной книжки; - получение информации о набранных кредитах (оценках); - автоматическое зачисление в предметные и учебные группы; - формирование портфолио обучающегося (данные по IMS ePortfolio Specification (http://www.imsglobal.org/ep/); - биллинговая система (on-line оплата обучения, просмотр истории оплаты). 2. Создание ЭИОС для обучающего: - получение информации по нагрузке; - планирование и проведение вебинаров; - разработка и экспертиза контрольно-измерительных материалов (КИМ); - проверка эссе обучающихся; - средства электронных коммуникаций (форумы, комментарии, чат) для общения с участниками процесса обучения (в учебных и предметных группах); - публикация мультимедийного обучающего контента; - формирование персонального портфолио. 3. Для организаторов образовательного процесса: - система мониторинга (получение информации об активности пользователей; организация опросов пользователей); - участие в группах (учебных, предметных, общих); - поддержка основных элементов коммуникаций (форумы, комментарии); - оценивание и организация объектов социальной сети; - отслеживание рейтингов и достижений; - публичное портфолио пользователя; - работа с заявками (отправка, отслеживание выполнения).
2.	«Компас-В» на платформе «1С: Предприятие 8.2»	Автоматизирует работу: - приёмной комиссии (ведение базы абитуриентов, зачисление на обучение, финансовые и маркетинговые отчёты), - деканата (документооборот приказов по движению обучающихся, репозиторий документов обучающихся, оповещение обучающихся по электронной почте, SMS), - учебно-методического управления (ведение базы преподавателей, формирование и учёт учебной нагрузки).
3.	Программный комплекс «Автоматизированная среда аттестации АСА»	Автоматизирует процедуры: - разработки, экспертизы и публичной экспертизы контрольно-измерительных материалов; - формирования и использования фонда оценочных средств по конкретным программам, дисциплинам (модулям, предметам, видам учебной работы); - проведения обучения с использованием балльно - рейтинговой системы (БРС); - допуска, проведения приема экзаменов, зачетов, письменных

		аттестационных работ; - видеопротоколирования процесса аттестации; - оформления документации по процессам аттестации (аттестационных ведомостей, заявлений на оплату выполненной обучающими работы, отчетов); - контроля успеваемости обучающихся; - мониторинга удовлетворенности обучающимися качеством контрольно-измерительных материалов и процедурами аттестации.
4.	Система программных продуктов LMS Moodle	Используется для: - публикации электронных курсов; - просмотра результата прохождения электронного курса и последующего его учёта в рамках БРС.
5.	Система автоматизации библиотек ИРБИС64	Используется для: организации доступа к электронной библиотеке для: просмотра полнотекстовых вариантов основной и дополнительной литературы; просмотра учебно-методических комплексов дисциплин.
6.	Программный продукт Автоматизированная информационная система библиотеки «Электронная библиотека»	Используется для организации процесса разработки и обновления полнотекстовых электронных версий учебных материалов и рабочих программ дисциплин.
7.	Прикладное программное обеспечение "Мираполис"	On-line сервис интегрированный в АСО, используется для: - организации и проведения вебинаров, интерактивных занятий, on-line консультаций и лекций; - просмотра записей вебинаров.

4.4.2 Специальные информационные технологии

Информационные технологии, используемые при осуществлении образовательного процесса по данной дисциплине, включает следующий перечень программного обеспечения и информационных справочных систем:

Программный продукт	Договор
1) Microsoft Office 2007 2) 7-Zip 3) Acrobat Reader 4) Антивирус Kaspersky Endpoint Security 10 for Windows 5) Mozilla Firefox 6) Google Chrome 7) VLC 2.2.6 8) Microsoft Windows 7 Pro-	1) Лицензия Microsoft Open License №43509314 дата выдачи настоящей Лицензии 19.02.2008 «Microsoft® Office 2007 Russian Academic. Microsoft® Windows® Vista Business Russian Upgrade Academic» 2) лицензия GNU LGPL (Свободное ПО) 3) лицензионное соглашение с компанией Adobe (Свободное ПО) 4) лицензия № 1FB6-170208-101930-190-411 5) лицензия GNU LGPL (Свободное ПО)

fessional Academic Open License 9) Microsoft Windows Server 2008	6) лицензия LGPL (Свободное ПО) 7) лицензия GNU LGPL (Свободное ПО) 8) лицензия № 62875440 9) лицензия № 47623222

4.5 Материально-техническая база

4.5.1 В процессе обучения данной дисциплины используется специальный виртуальный учебный кабинет – предметная группа модуля «Кампус ВЭГУ 24» (далее – ПГ).

4.5.2 Информационно-предметная среда ПГ размещается в разделах меню ПГ, а так же в разделах и подразделах меню «Кампус ВЭГУ 24».

Размещение материалов (или обеспечение прямого доступа к информационным ресурсам) в разделах меню «Кампус ВЭГУ 24» осуществляется в централизованном порядке структурными подразделениями Академии ВЭГУ (материалы Электронно-библиотечной системы, локальные нормативные акты Академии ВЭГУ, записи проведенных учебных занятий и т.п.).

В разделах меню ПГ размещение материалов организовывается преподавателем, являющимся владельцем данной ПГ (т.е. закрепленным за реализацией данной дисциплины).

4.5.3 Взаимодействие обучающихся в ПГ осуществляется на принципах тьюторинга и интерактива.

Тьюторинг осуществляется посредством постановки преподавателем задач перед обучающимся, консультирования, проверки и оценивания исполнения задач.

Интерактивный метод обучения обеспечивается организацией преподавателем в ПГ взаимодействия обучающихся друг с другом в сфере освоения программы дисциплины в разных формах и способах, в т.ч.:

- обязательных в рамках этапа (контрольных точек) БРС и факультативных занятий в форме дискуссий, групповой (командной) творческой работы, группового проектирования, кейс-метода, «мозгового штурма» или игры;
- широким использованием ПГ для общего обмена мнениями при постоянном контроле со стороны преподавателя и корректировки оценок и поощрения лучших сообщений;
- вовлечения обучающихся в формирование базы знаний.

4.5.4 Для ведения образовательной деятельности в учебных корпусах Академии ВЭГУ имеются учебные аудитории для проведения занятий лекционного типа, занятий семинарского типа, курсового проектирования (выполнения курсовых работ), групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также помещения для самостоятельной работы и помещения для хранения и профилактического обслуживания учебного оборудования. Они укомплектованы специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории. Для проведения занятий лекционного типа предлагаются

наборы демонстрационного оборудования и учебно-наглядных пособий, обеспечивающие тематические иллюстрации. Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети Интернет и обеспечением доступа в электронно-информационную среду Академии ВЭГУ.

Полный перечень материально-технического обеспечения по дисциплине указан в Справке о материально-техническом обеспечении основной профессиональной образовательной программы высшего образования.

5. Методические указания обучающемуся

Общая трудоемкость дисциплины составляет 5 зачетных единиц или 180 часов.

Обработка и хранение больших данных является одной из наиболее важных задач, возникающих при создании сложных аналитических систем ввиду того, что скорость предобработки и доступа к данным прямо влияет на итоговую скорость анализа данных. Под термином «большие данные» (*Big data)* подразумеваются наборы данных, хранение и обработка которых с помощью традиционно используемых на данных момент компьютерных систем представляет трудности. На данный момент для хранения и обработки данных больше 50 терабайт невозможно использовать средства реляционных баз данных, традиционно используемых для данных меньших объемов. В то же время необходимость в обработке и хранении больших данных возникает во множестве предметных областей - анализ финансовых рынков, анализ социальных сетей, медицинской информатики, обработка больших коллекций статей. Основой для обработки больших данных является хранилище данных. Разработка систем хранения больших данных является одним из приоритетных направлений таких компаний как Яндекс, Microsoft, IBM, Google.

Целью освоения дисциплины «Методы анализа данных» является теоретическая и практическая подготовка студентов к работе с большими данными. Знания, полученные в результате освоения дисциплины, помогут при сборе и анализе огромных объемов структурированной или неструктурированной информации, при разработке моделей данных и получении новых знаний. Все это необходимо выпускнику, освоившему программу бакалавриата, для решения различных задач практической и научно-исследовательской деятельности.

Для успешного освоения материала дисциплины студентам необходимо наличие компетенций, сформированных на предыдущем уровне образования и непосредственно связанных с дисциплинами общепрофессионального цикла: «Теория вероятностей и математическая статистика», «Информатика», «Программирование». Также понадобятся некоторые навыки работы с пакетами прикладных программ, таких как Excel или R.

По каждой теме предусмотрено выполнение студентами различных видов самостоятельной практической работы:

- самостоятельная работа под контролем преподавателя в форме плановых

консультаций и творческих контактов;

- внеаудиторная самостоятельная работа при выполнении студентом домашних заданий учебного и творческого характера.

По итогам освоения дисциплины предусмотрена рубежная и промежуточная аттестация. Рубежная аттестация проходит в форме балльно-рейтинговой системы в виде контрольных точек, в которые входит изучение электронного курса дисциплины, просмотр вебинаров по различным темам, контрольное тестирование и выполнение письменной аттестационной работы . Промежуточная аттестация предусмотрена в форме экзамена.

Для подготовки к аттестации рекомендуется:

- изучить лекционный и практический материал;
- изучить материалы, представленные по данной дисциплине в библиотеке Академии ВЭГУ или воспользоваться электронной библиотекой;
- использовать самообучающие программы;
- контролировать уровень своих знаний тестами-тренингами.

6. Особенности освоения дисциплины (модуля) для инвалидов и лиц с ограниченными возможностями здоровья.

6.1 Выбор методов обучения лиц с ограниченными возможностями здоровья и инвалидов определяются исходя из их доступности для данной категории обучающихся, определяется содержанием обучения, исходным уровнем имеющихся знаний, уровнем профессиональной подготовки педагогов, особенностями восприятия информации обучающимися. В образовательном процессе предполагается использование социально-активных и рефлексивных методов обучения, технологий социальной реабилитации с целью оказания помощи в установлении полноценных межличностных отношений с другими студентами, создании комфортного психологического климата.

6.2 В целях освоения учебной программы дисциплины инвалидами и лицами с ограниченными возможностями здоровья кафедра обеспечивает:

1) для инвалидов и лиц с ограниченными возможностями здоровья по зрению:

– размещение в доступных для обучающихся, являющихся слепыми или слабовидящими, местах и в адаптированной форме справочной информации о расписании учебных занятий;

– присутствие ассистента, оказывающего обучающемуся необходимую помощь; выпуск альтернативных форматов методических материалов (крупный шрифт или аудиофайлы);

2) для инвалидов и лиц с ограниченными возможностями здоровья по слуху: – надлежащими звуковыми средствами воспроизведения информации;

3) для инвалидов и лиц с ограниченными возможностями здоровья, имеющих нарушения опорно-двигательного аппарата:

– возможность беспрепятственного доступа обучающихся в учебные помещения, туалетные комнаты и другие помещения, а также пребывание в указанных помещениях.

6.3 Образование обучающихся с ограниченными возможностями здоровья может быть организовано как совместно с другими обучающимися, так и в отдельных группах или в отдельных организациях.

6.4 Учебно-методические материалы для самостоятельной работы обучающихся из числа инвалидов и лиц с ограниченными возможностями здоровья предоставляются в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Категории обучающихся	Формы предоставления
С нарушением слуха	в печатной форме; в форме электронного документа
С нарушением зрения	в печатной форме увеличенным шрифтом; в форме электронного документа; в форме аудиофайла;
С нарушением опорно-двигательного аппарата	в печатной форме; в форме электронного документа; в форме аудиофайла;

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

6.5 Процедура промежуточной аттестации для обучающихся с ограниченными возможностями здоровья и инвалидов устанавливается с учетом индивидуальных психофизических особенностей. При необходимости предусматривается увеличение времени на подготовку к промежуточной аттестации, а также предоставляется дополнительное время для подготовки ответа в ходе контроля текущей успеваемости и промежуточной аттестации.

При проведении процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья по дисциплине (модулю) обеспечивается выполнение следующих дополнительных требований в зависимости от индивидуальных особенностей обучающихся:

1. инструкция по порядку проведения процедуры оценивания предоставляется в доступной форме (устно, в письменной форме, устно с использованием услуг сурдопереводчика);

2. доступная форма предоставления заданий оценочных средств (в печатной форме, в печатной форме увеличенным шрифтом, в форме электронного документа, задания зачитываются ассистентом, задания предоставляются с использованием сурдоперевода);

3. доступная форма предоставления ответов на задания (письменно на бумаге, набор ответов на компьютере, с использованием услуг ассистента, устно).

При необходимости для обучающихся с ограниченными возможностями здоровья и инвалидов процедура оценивания результатов обучения по дисциплине (модулю) может проводиться в несколько этапов. Проведение процедуры

оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья допускается с использованием дистанционных образовательных технологий.

6.6 Для освоения дисциплины инвалидами и лицами с ограниченными возможностями здоровья предоставляются основная и дополнительная учебная литература в виде электронного документа в фонде библиотеки и / или в электронно-библиотечных системах, а также предоставляются бесплатно специальные учебники и учебные пособия, иная учебная литература и специальные технические средства обучения коллективного и индивидуального пользования, а также услуги сурдопереводчиков и тифлосурдопереводчиков.

Проректор по учебно-научной и
воспитательной работе А.О. Целищев