

СИЛЛАБУС
Осенний семестр 2025 – 2026 учебного года
Образовательная программа «6В06102 Информационные системы»

ID и наименование дисциплины	Самостоятельная работа обучающегося (СРО)	Кол-во кредитов			Общее кол-во кредитов	Самостоятельная работа обучающегося под руководством преподавателя (СРОП)	
		Лекции (Л)	Практ. занятия (ПЗ)	Лаб. занятия (ЛЗ)			
102302, Наука о данных	4	1,7	-	3,0	5	4	
АКАДЕМИЧЕСКАЯ ИНФОРМАЦИЯ О ДИСЦИПЛИНЕ							
Формат обучения	Цикл, модуль компонент	Типы лекций	Типы практических занятий		Форма и платформа итогового контроля		
Офлайн	П КпВ	Обзорные лекции – введение в науку о данных, место дисциплины среди других наук	работа с Jupyter Notebook, написание кода на Python, базовые операции с данными		Устный офлайн		
Лектор - (ы)	Байкувеков Мейржан Берикович						
e-mail:	baikuvekov_meirzhan3@live.kaznu.kz						
Телефон:	+7 702 6478122						
Ассистент- (ы)	Ілесова Бахытгүл Ералықызы						
e-mail:	Bakhytgul@live.kaznu.kz						
Телефон:	+7 747 983 3787						
АКАДЕМИЧЕСКАЯ ПРЕЗЕНТАЦИЯ ДИСЦИПЛИНЫ							
Цель дисциплины	Ожидаемые результаты обучения (РО)*				Индикаторы достижения РО (ИД)		
Целью изучения дисциплины «Наука о данных» является формирование у студентов теоретических знаний и практических навыков в области сбора, обработки, анализа и интерпретации данных с использованием современных инструментов и методов. Дисциплина направлена на развитие умений применять методы статистики, машинного обучения и визуализации данных для решения	1. Владеть базовыми понятиями и этапами науки о данных, понимать её роль и задачи в современных информационных системах.				1.1 Студент объясняет ключевые этапы жизненного цикла данных и их взаимосвязь.		
					1.2 Студент описывает области применения науки о данных и её значение в решении прикладных задач.		
	2. Уметь собирать, очищать, преобразовывать и анализировать данные с использованием современных инструментов и библиотек.				2.1 Студент выполняет загрузку и предобработку данных в среде Python (Pandas, NumPy).		
					2.2 Студент применяет методы очистки и преобразования данных для подготовки их к анализу.		
	3. Применять статистические методы и исследовательский анализ данных для выявления закономерностей и тенденций.				3.1 Студент рассчитывает основные статистические показатели и интерпретирует их.		
					3.2 Студент проводит EDA и визуализирует результаты анализа с помощью графических инструментов.		
	4. Проектировать и реализовывать базовые модели машинного обучения для решения задач классификации, регрессии и кластеризации.				4.1 Студент обучает и тестирует модели машинного обучения на реальных данных.		
					4.2 Студент оценивает точность и эффективность моделей с использованием соответствующих метрик.		
	5. Демонстрировать навыки самостоятельного проектирования и реализации аналитического проекта на основе реальных данных.				5.1 Студент разрабатывает мини-проект, включающий полный цикл анализа данных.		

прикладных задач в науке, бизнесе и промышленности, а также на подготовку студентов к самостоятельной аналитической работе и исследовательской деятельности в области Data Science.		5.2 Студент презентует результаты проекта и обосновывает выбор методов и инструментов.
Пререквизиты	Алгоритмы, структуры данных и программирование; Объектно-ориентированное программирование: Системный анализ и проектирование ИС	
Постреквизиты	Модели и методы интеллектуальных информационных систем: Основы машинного обучения; Программирование на языке Java	
Учебные ресурсы	Литература: основная, дополнительная. - Провост Ф., Фоусет Т. <i>Data Science для бизнеса. Что нужно знать о данных и аналитике</i> – М.: Манн, Иванов и Фербер, 2020. Классический курс по применению методов анализа данных в бизнес-задачах. - Герасименко В. П., Орлов А. И. <i>Наука о данных: основы и методы</i> – СПб.: Питер, 2021. – Системное изложение основ науки о данных, включая обработку, анализ и визуализацию. - Андреев А. А. <i>Машинное обучение и анализ данных</i> – М.: БХВ-Петербург, 2022. – Учебное пособие по ключевым методам машинного обучения и их реализации на Python. - Стародубцев А. А. <i>Python и анализ данных: практическое руководство</i> – СПб.: Питер, 2023. – Практическое руководство по обработке и визуализации данных средствами Python. - Соколов А. В. <i>Машинное обучение и наука о данных. Практика применения алгоритмов</i> – М.: ДМК Пресс, 2022. – Подробное описание алгоритмов и инструментов машинного обучения с примерами кода. Исследовательская инфраструктура - Лаборатория анализа данных и машинного обучения – специализированная учебно-исследовательская среда для проведения практических и проектных работ, экспериментов по обработке и моделированию данных. - Облачные вычислительные ресурсы и онлайн-платформы – Google Colab, Kaggle, GitHub и другие сервисы, предоставляющие доступ к вычислительным мощностям и реальным наборам данных. - Электронные образовательные ресурсы и научные базы данных – доступ к современным научным публикациям, библиотекам и открытым наборам данных для проведения исследовательских проектов. - Учебные аудитории и коворкинг-пространства университета – для проведения семинаров, групповых дискуссий, презентаций проектов и защиты итоговых работ. Профессиональные научные базы данных . - IEEE Xplore Сайт: IEEE Xplore - ACM Digital Library а\$т: ACM Digital Library . - SpringerLink Car: SpringerLink - ScienceDirect C a r : Science Direct - Web of Science Caùt: Web of Science . Интернет-ресурсы (не менее 3-5) https://www.kaggle.com – платформа для соревнований по анализу данных, публикаций и обмена датасетами.	

2. <https://datasetsearch.research.google.com> – поисковик открытых наборов данных от Google.
 3. <https://data.gov> – портал открытых государственных данных США.
 4. <https://data.world> – онлайн-сообщество для обмена наборами данных и совместной работы над проектами.
- Программное обеспечение.**

1. Python
2. Anaconda Distribution
3. Jupyter Notebook / JupyterLab

Академическая политика дисциплины

Академическая политика дисциплины определяется Академической политикой и Политикой академической честности КазНУ имени аль-Фараби.

Документы доступны на главной странице ИС Univer.

Интеграция науки и образования. Научно-исследовательская работа студентов, магистрантов и докторантов – это углубление учебного процесса. Она организуется непосредственно на кафедрах, в лабораториях, научных и проектных подразделениях университета, в студенческих научно-технических объединениях. Самостоятельная работа обучающихся на всех уровнях образования направлена на развитие исследовательских навыков и компетенций на основе получения нового знания с применением современных научно-исследовательских и информационных технологий. Преподаватель исследовательского университета интегрирует результаты научной деятельности в тематику лекций и семинарских (практических) занятий, лабораторных занятий и в задания СРОП, СРО, которые отражаются в силлабусе и отвечают за актуальность тематик учебных занятий и заданий.

Посещаемость. Дедлайн каждого задания указан в календаре (графике) реализации содержания дисциплины. Несоблюдение дедлайнов приводит к потере баллов.

Академическая честность. Практические/лабораторные занятия, СРО развивают у обучающегося самостоятельность, критическое мышление, креативность. Недопустимы плагиат, подлог, использование шпаргалок, списывание на всех этапах выполнения заданий.

Соблюдение академической честности в период теоретического обучения и на экзаменах помимо основных политик регламентируют «Правила проведения итогового контроля», «Инструкции для проведения итогового контроля осеннего/весеннего семестра текущего учебного года», «Положение о проверке текстовых документов обучающихся на наличие заимствований».

Документы доступны на главной странице ИС Univer.

Основные принципы инклюзивного образования. Образовательная среда университета задумана как безопасное место, где всегда присутствуют поддержка и равное отношение со стороны преподавателя ко всем обучающимся и обучающимся друг к другу независимо от гендерной, расовой/ этнической принадлежности, религиозных убеждений, социально-экономического статуса, физического здоровья студента и др. Все люди нуждаются в поддержке и дружбе ровесников и сокурсников. Для всех студентов достижение прогресса скорее в том, что они могут делать, чем в том, что не могут. Разнообразие усиливает все стороны жизни.

Все обучающиеся, особенно с ограниченными возможностями, могут получать консультативную помощь по телефону/ e-mail baikuvekov_meirzhan3@live.kaznu.kz либо посредством видеосвязи в MS Teams <https://us04web.teams.us/j/73674837609?pwd=AID37HhIShnBYtJOgU7qqmDQ9oK4na.1>

Интеграция МООС (massive open online course). В случае интеграции МООС в дисциплину, всем обучающимся необходимо зарегистрироваться на МООС. Сроки прохождения модулей МООС должны неукоснительно соблюдаться в соответствии с графиком изучения дисциплины.

ВНИМАНИЕ! Дедлайн каждого задания указан в календаре (графике) реализации содержания дисциплины, а также в МООС. Несоблюдение дедлайнов приводит к потере баллов.

ИНФОРМАЦИЯ О ПРЕПОДАВАНИИ, ОБУЧЕНИИ И ОЦЕНИВАНИИ

Балльно-рейтинговая буквенная система оценки учета учебных достижений				Методы оценивания
Оценка	Цифровой эквивалент баллов	Баллы, % содержание	Оценка по традиционной системе	Критериальное оценивание – процесс соотнесения реально достигнутых результатов обучения с ожидаемыми результатами обучения на основе четко выработанных критериев. Основано на формативном и суммативном оценивании. Формативное оценивание – вид оценивания, который проводится в ходе повседневной учебной деятельности. Является текущим показателем успеваемости. Обеспечивает оперативную взаимосвязь между обучающимся и преподавателем. Позволяет определить возможности обучающегося, выявить трудности, помочь в достижении наилучших результатов, своевременно корректировать преподавателю образовательный процесс. Оценивается выполнение заданий, активность работы в аудитории во время лекций, семинаров, практических занятий (дискуссии, викторины, дебаты, круглые столы, лабораторные работы и т. д.). Оцениваются приобретенные знания и компетенции.
A	4,0	95-100	Отлично	
A-	3,67	90-94		
B+	3,33	85-89	Хорошо	

			Суммативное оценивание – вид оценивания, который проводится по завершению изучения раздела в соответствии с программой дисциплины. Проводится 4 раза за семестр при выполнении СРО. Это оценивание освоения ожидаемых результатов обучения в соответствии с дескрипторами. Позволяет определять и фиксировать уровень освоения дисциплины за определенный период. Оцениваются результаты обучения. <table><tr><th>Формативное и суммативное оценивание</th><th>Баллы % содержание</th></tr><tr><td>Активность на лекциях</td><td>0</td></tr><tr><td>Работа на практических занятиях</td><td>25</td></tr><tr><td>Самостоятельная работа</td><td>25</td></tr><tr><td>Проектная и творческая деятельность</td><td>10</td></tr><tr><td>Итоговый контроль (экзамен)</td><td>40</td></tr><tr><td>ИТОГО</td><td>100</td></tr></table>	Формативное и суммативное оценивание	Баллы % содержание	Активность на лекциях	0	Работа на практических занятиях	25	Самостоятельная работа	25	Проектная и творческая деятельность	10	Итоговый контроль (экзамен)	40	ИТОГО	100
Формативное и суммативное оценивание	Баллы % содержание																
Активность на лекциях	0																
Работа на практических занятиях	25																
Самостоятельная работа	25																
Проектная и творческая деятельность	10																
Итоговый контроль (экзамен)	40																
ИТОГО	100																
B	3,0	80-84															
B-	2,67	75-79															
C+	2,33	70-74															
C	2,0	65-69															
C-	1,67	60-64															
D+	1,33	55-59															
D	1,0	50-54															
FX	0,5	25-49															
F	0	0-24															

Календарь (график) реализации

Календарь (график) реализации содержания дисциплины. Методы преподавания и обучения.

Неделя	Название темы			Кол-во часов	Макс. балл
МОДУЛЬ 1. Основы науки о данных и обработка данных.					
1	Л 1. Введение в науку о данных: понятия, этапы и области применения			1	0
	ЛЗ 1. Работа с датасетами: загрузка и обзор данных в Python (Pandas)			2	0
2	Л 2. Источники данных и форматы. Сбор и загрузка данных			1	0
	ЛЗ 2. Очистка и нормализация данных на примере реального набора			2	0
3	Л 3. Предобработка данных: очистка, фильтрация, преобразование			1	0
	ЛЗ 3. Анализ пропущенных и выбросных значений			2	0
	СРОП 1. Консультации по выполнению СРО1 на тему «Обзор современных инструментов и библиотек для науки о данных»				0
4	Л 4. Исследовательский анализ данных (EDA): методы и инструменты			1	0
	ЛЗ 4. Проведение EDA: статистическое описание и визуализация данных			2	0
5	Л 5. Визуализация данных: принципы и библиотеки (Matplotlib, Seaborn)			1	0
	ЛЗ 5. Создание аналитического отчета с использованием графиков			2	15
	СРО 1. Прием СРС1 в форме оцениваться по отчёту			0	20
МОДУЛЬ 2. Статистический анализ и машинное обучение.					
6	Л 6. Основы статистики и вероятностей в науке о данных			1	0
	ЛЗ 6. Вычисление основных статистических показателей и их интерпретация			2	15
	СРОП 2. Консультации по выполнению СРО 2 «Анализ примеров применения Data Science в бизнесе и науке»				0
7	Л 7. Корреляция и причинно-следственные связи			1	0
	ЛЗ 7. Корреляционный анализ данных			2	15
	СРО 2. Прием СРС2 в форме оцениваться по отчёту				20
8	Л 8. Введение в машинное обучение: задачи и методы			1	0
	ЛЗ 8. Построение и обучение модели классификации (логистическая регрессия)			2	15
Рубежный контроль 1					100
9	Л 9. Модели классификации: логистическая регрессия, kNN			1	0
	ЛЗ 9. Реализация алгоритма kNN для задач классификации			2	10
	СРОП 3. Консультация по выполнению СРО 3. «Обзор открытых наборов данных и их особенности»				0
10	Л 10. Модели регрессии: линейная и полиномиальная регрессия			1	0
	ЛЗ 10. Применение моделей регрессии к реальному набору данных			2	10
	СРО 3. Прием СРС3 в форме оцениваться по отчёту				15
МОДУЛЬ 3 Продвинутое методы и практические проекты					
11	Л 11. Методы кластеризации: k-means и иерархическая кластеризация			1	0
	ЛЗ 11. Реализация алгоритма кластеризации (k-means)			2	10
12	Л12. Снижение размерности: PCA и t-SNE			1	0
	ЛЗ 12. Снижение размерности данных с помощью PCA			2	10
13	Л 13. Валидация и оценка качества моделей			1	0
	ЛЗ 13. Оценка и сравнение моделей с использованием метрик точности			2	10
	СРОП4. Консультация по выполнению СРО 4. «Написание аналитического отчета по мини-исследованию данных»				0
14	Л 14. Работа с временными рядами			1	0
	ЛЗ 14. Анализ временных рядов: тренды и сезонность			2	10
15	Л 15. Проектирование и реализация проекта по анализу данных			1	0

	ЛЗ 15. Итоговый мини-проект: анализ и визуализация данных по выбранной теме	2	10
	СРО 4. Прием СРС4 в форме оцениваться по отчёту		15
Рубежный контроль 2			100
Итоговый контроль (экзамен)			100
ИТОГО за дисциплину			100

РУБРИКАТОР СУММАТИВНОГО ОЦЕНИВАНИЯ

КРИТЕРИИ ОЦЕНИВАНИЯ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ

СРС 1. Обзор современных инструментов и библиотек для науки о данных (20% от 100% РК1)				
Критерий	«Отлично» 15-20 %	«Хорошо» 9-14%	«Удовлетворительно» 5-8%	«Неудовлетворительно» 2-4%
Полнота и глубина обзора	Представлен полный анализ инструментов и библиотек (Pandas, NumPy, Scikit-learn, TensorFlow и др.) с их преимуществами и сферами применения.	Анализ содержит большинство ключевых инструментов, но с недостаточной детализацией.	Обзор ограничен небольшим количеством инструментов и кратким описанием.	Обзор не содержит существенной информации или отсутствует.
Структурированность и логика материала	Материал структурирован, логика изложения чёткая и последовательная.	Логика в целом соблюдена, но возможны небольшие несоответствия.	Структура и логика изложения слабые или частично нарушены.	Материал не структурирован и логика отсутствует.
Аналитическое осмысление	Приведено сравнение инструментов, аргументированы их преимущества и недостатки.	Присутствует частичный анализ и сравнение.	Аналитический подход слабый, в основном перечисление.	Анализ отсутствует.

СРС 2. Анализ примеров применения Data Science в бизнесе и науке (20% от 100% РК1)				
Критерий	«Отлично» 15-20 %	«Хорошо» 9-14%	«Удовлетворительно» 5-8%	«Неудовлетворительно» 2-4%
Качество и глубина анализа кейсов	Рассмотрены разнообразные реальные кейсы из бизнеса и науки с глубоким анализом методов и результатов.	Приведены релевантные примеры, но анализ не полностью детализирован.	Примеры поверхностные, без глубокого анализа методов.	Примеры отсутствуют или не соответствуют теме.
Интерпретация и выводы	Выводы логичны, аргументированы и демонстрируют понимание роли Data Science.	Выводы в целом корректны, но частично поверхностны.	Выводы слабо связаны с приведёнными примерами.	Выводы отсутствуют или не соответствуют анализу.
Структурированность работы	Работа логична и хорошо структурирована, есть введение, анализ и выводы.	Структура в целом есть, но с незначительными недочётами.	Структура слабая или не полностью выдержана.	Структура отсутствует.

СРС 3. Обзор открытых наборов данных и их особенности (15% от 100% РК2)

Критерий	«Отлично» 10-15 %	«Хорошо» 6-9%	«Удовлетворительно» 3-5%	«Неудовлетворительно» 0-2%
Полнота и актуальность обзора	Рассмотрены несколько разнообразных и актуальных открытых наборов данных с описанием их структуры и особенностей.	Приведены актуальные наборы данных, но без полного анализа их особенностей.	Примеры наборов ограничены и описаны поверхностно.	Примеры не приведены или не соответствуют теме.
Анализ структуры и качества данных	Дана подробная характеристика структуры, типов данных, качества и ограничений.	Характеристика в целом дана, но с неполной детализацией.	Анализ частично проведён или содержит ошибки.	Анализ отсутствует.
Практическая значимость	Показано, как данные могут быть использованы в проектах Data Science.	Практическая значимость упомянута, но без конкретных примеров.	Практическая ценность раскрыта слабо.	Практическая значимость не рассмотрена.

СРС 4. Написание аналитического отчёта по мини-исследованию данных (15% от 100% РК2)

Критерий	«Отлично» 10-15 %	«Хорошо» 6-9%	«Удовлетворительно» 3-5%	«Неудовлетворительно» 0-2%
Сбор и подготовка данных	Данные собраны корректно, проведена очистка, нормализация и визуальный обзор.	Данные собраны и подготовлены, но есть недочёты в обработке.	Подготовка данных частичная или с ошибками.	Подготовка данных отсутствует или некорректна.
Аналитический анализ и интерпретация	Применены соответствующие методы анализа, результаты интерпретированы и обоснованы.	Методы анализа применены корректно, но интерпретация частично поверхностна.	Анализ ограничен, интерпретация слабая.	Анализ не проведён или некорректен.
Оформление и структура отчёта	Отчёт логично структурирован (введение, методы, результаты, выводы), оформлен грамотно и академично.	Структура отчёта в целом собрана, но есть мелкие недочёты.	Структура слабая или неполная.	Структура отсутствует или не соответствует требованиям.

Декан  Т.С. Иманкулов
 Председатель Академического комитета
 по качеству преподавания и обучения  Ж.А. Бурибаев
 Заведующий кафедрой  А.Н. Шормакова
 Лектор  М.Б. Байкувков

