



АКАДЕМИЯ
искусственного интеллекта
для школьников

#AIAcademy

**План и методические рекомендации по проведению урока
по теме «Статистический анализ данных. Линейный регрессионный анализ».**

Основная идея: привлечение внимания обучающихся к одной из перспективных областей профессиональной деятельности в ИТ-индустрии — **науке о данных**.

Цель урока:

- сформировать у учащихся представление о статистическом анализе данных и его роль в интерпретации полученных результатов;
- научить использовать программу Microsoft Excel для работы с табличными данными.

Задачи урока:

- расширить представление у учащихся о табличных данных и методах работы с ними;
- сформировать навыки статистического анализа данных;

- сформировать навыки использования базовых средств табличного процессора Microsoft Excel для статистического анализа данных (визуализация данных с помощью диаграммы разброса, построение математической модели линейной (парной) зависимости и ее интерпретация);
- сформировать метапредметные результаты, связанные с:
 - умением работать с информацией, анализировать и структурировать полученные знания и синтезировать новые, устанавливать причинно-следственные связи;
 - ставить цель и находить оптимальные способы ее достижения, проводить ситуационную и ретроспективную рефлексию, участвуя в подведении итогов отдельных этапов и урока в целом;
 - умением вступать в диалог и вести его, взаимодействовать в команде.

Планируемые результаты:

личностные:

- готовность и способность к осознанному выбору и построению дальнейшей индивидуальной траектории образования благодаря ориентации в мире профессий и профессиональных предпочтений с учётом устойчивых познавательных интересов;

метапредметные умения и опыт:

- выдвигать версии решения проблемы, формулировать гипотезы, предвосхищать конечный результат;
- определять цель деятельности на основе определенной проблемы и существующих возможностей;
- строить рассуждение на основе сравнения данных, выделяя общие признаки;
- излагать полученную информацию, интерпретируя ее в контексте решаемой задачи;
- выделять информационный аспект задачи, оперировать данными, использовать модель решения задачи и др.

Тема урока «**Статистический анализ данных. Линейный регрессионный анализ**» раскрывается в два академических часа и состоит из вводной теоретической части и практической. Сопровождается

презентацией в рамках содержательной части урока, а также подкрепляется проектами в Microsoft Excel для учителя (с заданиями, комментариями и ответами) и ученика (с заданиями, комментариями, но без ответов) в практической части урока.

Практическая работа включает алгоритм, справочный материал и пошаговые инструкции по выполнению заданий. Содержательное наполнение заданий основано на реальной практической задаче (сюжет и данные, используемые в задаче, вымышленные; все совпадения случайны). Кроме того, учитель имеет доступ к файлам с исходными данными задач, что позволит значительно оптимизировать процесс, не тратя времени на создание массивов данных (ввод информации). Подробное описание действий в инструкциях («выделите, откройте» и т.п.) позволит включить в работу даже школьников, не имеющих опыта работы с табличным процессором.

Практическую работу можно выполнять как во фронтальном (синхронном) режиме, объясняя рекомендуемые приемы работы, так и организуя индивидуальную работу школьников в удобном для них темпе.

Рекомендации по распределению времени: учитель может распределить учебное время следующим образом: теория – 0,5 часа, практика – 1,5 часа или теория и практика по 1 часу, ориентируясь на уровень подготовки учебной группы. Если есть дополнительное время, то на практическую часть можно заложить и больше времени (на усмотрение учителя).

Важно: на уроке можно выполнять не все предложенные из практикума. Все зависит от подготовленности группы. Учитель может самостоятельно формировать набор заданий для урока. Часть разобрать и выполнить на уроке, а другую часть дать учащимся для домашнего выполнения.

Также можно выбрать другую стратегию. Увеличить количество часов для практической части до 2 часов.

Для проведения урока учителю понадобится:

- компьютер с проекционным оборудованием;
- компьютерный класс с выходом в интернет и с установленным пакетом Microsoft Excel;
- опорная презентация;
- загруженный файл с заданиями для учеников на каждый учебный компьютер.

Этап урока	Номер слайда	Комментарии для учителя	Дополнительные материалы
Мотивация к учебной деятельности	---	<p>Учителю следует создать условия, чтобы ученики внутренне собрались, подготовились и нацелились на «покорение новых вершин».</p> <p>Вступительное слово.</p> <p><i>На прошлом уроке мы научились определять связь между величинами с помощью корреляции и интерпретировать полученные результаты. Но как выявленная связь между величинами может нам помочь в дальнейшем анализе данных? Давайте разбираться.</i></p>	<p>Рекомендация (не обязательно):</p> <p>Корреляция \neq Каузальность.</p> <p>AsapSCIENCE [Электронный ресурс]</p> <p>– Режим доступа:</p> <p>https://www.youtube.com/watch?v=un0OkHP8IxE</p>
Актуализация знаний	1	После вводной лекции учителя следует обратить внимание на то, что корреляционный анализ позволяет выявить линейную связь между	<p>Рекомендация для учителя:</p> <ul style="list-style-type: none"> ● рекомендуемая презентация:

		<p>признаками, это позволяет нам построить математическую модель линейной зависимости.</p> <p>Учитель может задать вопросы для обсуждения:</p> <p><i>Вспомним прошлое занятие.</i></p> <p><i>Как выглядела связь между величинами? (точки собирались в линию, очень похожую на прямую)</i></p> <p><i>Какие параметры мы получали? (коэффициент при X и свободных коэффициент)</i></p> <p><i>Подумайте, а можем ли мы, зная данные параметры, рассчитать что-нибудь наперед? (конечно можем, подставить интересующий нас X и рассчитать Y)</i></p> <p><i>Это и есть прогнозирование, то есть узнать, что будет в будущем, и это позволяет делать линейная регрессия.</i></p> <p><i>Как мы говорили ранее, в школьных предметах также можно увидеть корреляцию, а следовательно и построить математическую модель линейной зависимости. Давайте вспомним зависимости, которые вы разобрали на прошлом</i></p>	<p><i>7_Статистический_анализ_данных_Линейный_регрессионный.pptx;</i></p> <ul style="list-style-type: none"> ● организовать групповую работу в классе и предоставить возможность самостоятельно изучить поставленные вопросы для обсуждения (у учащихся должен быть компьютер с выходом в Интернет).
--	--	--	---

		<p>уроке из школьных предметов математики и физики.</p> <p>математика:</p> <ul style="list-style-type: none"> ● путь — время (график зависимости этих двух величин есть скорость); ● ширина — длина (график зависимости этих двух величин есть площадь); <p>физика:</p> <ul style="list-style-type: none"> ● сила тока в участке цепи — сопротивление участка (график зависимости двух величин есть напряжение на этом участке); <p>и т.п.</p> <p>Подумайте, если бы вам предложили на координатной плоскости начертить график этих зависимостей. Как бы он выглядел? Напрямую, то есть линейная функция.</p> <p>Можно дать подсказку: вспомнить диаграмму разброса; если бы эти «облака» были бы более вытянутыми, на что они станут похожи.</p> <p>Вспомните общую формулу линейной функции? $y=ax+b$</p>	
--	--	---	--

		<i>Эта формула отражает линейную зависимость между двумя признаками. Поэтому ее называют «парная линейная регрессия».</i>	
Целеполагание, постановка проблемы	2-7	<p>После обсуждения учителю следует обратить внимание на то, что ошибки визуализации данных возникают при неправильном масштабировании. Из-за этого несущественные различия кажутся существенными. Это и вводит в заблуждение. Учитель формирует вместе с учащимися основную цель урока:</p> <p><i>«Сегодня нам предстоит разобраться, как линейный регрессионный анализ при анализе данных помогает правильно интерпретировать результаты. Понять, в каких случаях линейный регрессионный анализ не помогает, а даже мешает делать правильные выводы и как корреляционный анализ помогает линейному регрессионному».</i></p> <p>Учитель раскрывает тему урока. Транслирует новый материал.</p> <p><i>Статистическое использование слова «регрессия» исходит из явления, известного как регрессия к</i></p>	<p>Рекомендация для учителя:</p> <ul style="list-style-type: none"> ● рекомендуемая презентация: <i>7_Статистический_анализ_данных_Линейный_регрессионный.pptx.</i> <p>Подробнее:</p> <p>Основы линейной регрессии. StatSoft. Портал Знаний: глобальный интеллектуальный ресурс.[Электронный ресурс] – Режим доступа:</p>

		<p><i>среднему, приписываемого сэру Френсису Гальтону (1889).</i></p> <p><i>Он показал, что, хотя высокие отцы имеют тенденцию иметь высоких сыновей, средний рост сыновей меньше, чем у их высоких отцов. Средний рост сыновей «регрессировал» и «двигался вспять» к среднему росту всех отцов в популяции. Таким образом, в среднем высокие отцы имеют более низких (но всё-таки высоких) сыновей, а низкие отцы имеют сыновей более высоких (но всё-таки довольно низких).</i></p>	<p>http://statistica.ru/theory/osnovy-lineynoy-regressii/</p> <p>Линейная регрессия: создавайте прогнозы с помощью простой для понимания математической формулы. IBM. [Электронный ресурс] – Режим доступа: https://www.ibm.com/ru-ru/analytics/learn/linear-regression</p>
	8-9	<p>Учитель может продемонстрировать встроенные функции MS Excel и построение диаграммы разброса с линией тренда.</p>	
Поиск путей решения проблемы	---	<p><i>Ребята, вы уже знаете, что для аналитика важно уметь работать с данными (собирать, сохранять, извлекать, обрабатывать их), также получать новые расчетные данные для дальнейших выводов и прогнозирования.</i></p> <p><i>Но очень важно, чтобы данные были корректны и анализ подтверждался статистическими методами.</i></p>	<p>Рекомендация для учителя:</p> <ul style="list-style-type: none"> ● рекомендуемый раздаточный материал для учителя: 7_Статистический_анализ_данных_Линейный_регрессионный (для учителя).xlsx; ● рекомендуемый раздаточный материал для ученика:

		<p><i>Это позволит избежать ошибок в исследовании и подтасовки данных.</i></p> <p><i>Мы с вами будем работать в приложении Microsoft Excel. Будем данные визуализировать с помощью диаграммы разброса и определять наличие связи между признаками. А также подтверждать наличие связи с помощью коэффициента Пирсона. Для этого мы будем использовать встроенные функции. Переходим к практической части.</i></p> <p>Консультирует. Демонстрирует работу в приложении Microsoft Excel.</p> <p>Учитель повторяет с учащимся интерфейс приложения Microsoft Excel (книга, лист, столбцы, строки, ячейка, адресация (имя ячейки), поле имени, строка формул). Для получения новой информации (обогащения данных) требуется не только использовать формулы и относительную адресацию, но и встроенные функции Microsoft Excel.</p> <p>Учитель должен сделать акцент на том, что в рамках урока потребуются команды, находящиеся на ленте «Главная» (это упростит поиск необходимых команд для решения поставленной</p>	<p><i>7_Статистический_анализ_данных_Линейный_регрессионный.xlsx</i></p> <p><i>Учащиеся могут работать одновременно с учителем. Знакомиться с интерфейсом. Если учащиеся знакомы с электронной таблицей, то данный этап на уроке можно сократить как по времени, так и по содержанию.</i></p> <p>Планирование путей достижения намеченной цели. Осуществление учебных действий по плану. Решение практических задач.</p>
--	--	---	--

		задачи). При работе с диаграммами будет доступна вспомогательная панель « Конструктор ».	
Решение проблемы	---	<p><i>Вам предстоит самостоятельно познакомиться с коэффициентами линейной регрессии и разобраться с интерпретацией. Решить несколько кейс-задач. Давайте познакомимся с содержанием практической работы.</i></p> <p><i>У каждого этапа есть инструкция и справочный материал, которые помогут решить поставленную задачу. Также вы можете использовать и внешние источники информации (интернет).</i></p> <p><i>Напоминаю, что все команды, которые потребуются для решения задачи, находится на ленте «Главная». Можете пользоваться справочным материалом Microsoft Excel.</i></p> <p><i>При работе с диаграммами появляется дополнительная панель «Конструктор».</i></p> <p>Задача 1. Исследование зависимости пути от времени.</p> <p>Задача 2. Исследование рисков от курения.</p>	<p>Рекомендации для учителя:</p> <p><i>В зависимости от уровня базовой подготовки учащихся можно организовать фронтальную работу, осуществлять контроль и объяснять методы решения поставленной задачи или индивидуальную, в удобном для каждого темпе с внешним контролем.</i></p> <p>Чтобы учащиеся могли самостоятельно осваивать материал, необходимо им показать, как использовать встроенный справочный материал Microsoft Excel (при вызове встроенной функции можно раскрыть справку, также при наведение курсора на команду появляются подсказки).</p> <p>Подробнее:</p>

		<p><i>Задача 3. Исследование профилактическим мер вакцинации.</i></p> <p><i>Задача 4. Исследования здоровья и благополучия подростков.</i></p> <p><i>Не спешите. Прочитайте внимательно описание данных. В задачах предложены реальные данные, которые ранее были использованы в исследованиях. В каждой задаче есть содержание исследования и описание ожидаемых результатов. Необходимо подтвердить или опровергнуть наличие тех или иных связей. Сделать итоговый вывод.</i></p> <p><i>Выполняя задание шаг за шагом, вы увидите, как анализ становится более глубоким и информативным. Подумайте, почему так получается?</i></p> <p><i>Вам потребуется переходить на другую вкладку «Данные». Будьте внимательны.</i></p> <p><i>Желаю вам удачи. Если у вас что-то не получается, я вам обязательно помогу.</i></p>	<p>R^2 — это коэффициент корреляции в квадрате, его можно увидеть на диаграмме разброса при добавлении линии тренда.</p>
--	--	---	---

		<p>Учитель должен сделать акцент на том, что задания нужно выполнять последовательно, это позволит учащимся разобраться с встроенными функциями, которые агрегируют данные по одному или нескольким параметрам и после рассчитывают статистические показатели.</p> <p>Надо обратить внимание учащихся, что данные по исследованию постепенно дополняются. Это дает возможность получать новые показатели и делать дополнительный анализ.</p> <p>Консультирует.</p> <p>В предложенном практическом задании исходные данные произвольные.</p>	
Самостоятельная работа с использованием полученных знаний	---	Направляет, советует, консультирует.	<p>Рекомендации для учителя:</p> <p><i>организовать самостоятельную работу таким образом, что если ребенок не знает как выполнить задание, чтобы он смог найти информацию или пример решения подобной задачи (в книге, справке, интернете и т.п.) Это позволит ему</i></p>

			получить навык поиска необходимой информации в различных источниках.
Систематизация знаний	---	<p>Консультирует, направляет.</p> <p><i>При проведении анализа данных необходимо четко формулировать цели исследования и использовать различные статистические методы, чтобы не получить абсурдных результатов.</i></p> <p>Просмотр видео, в котором показывают абсурдные отношения между несвязанными признаками и как это в дальнейшем влияет на конечный результат исследования.</p>	<p>Рекомендации для учителя:</p> <p><i>создать условия учащимся, для работы по выявлению связи изученной на уроке темы с изученным ранее материалом, связи с жизнью.</i></p>
Объяснение домашнего задания	---	<p>Домашнее задание после первого урока:</p> <p><i>Доклад на тему: «Статистические методы анализа данных»;</i></p> <p><i>Ошибки в статистическом анализе данных».</i></p> <p>Домашнее задание после второго урока:</p> <p><i>найти примеры зависимостей из школьных предметов; собрать данные, рассчитать</i></p>	<p>У учеников должна быть возможность выбора домашнего задания в соответствии со своими предпочтениями.</p> <p>Необходимо наличие заданий разного уровня сложности</p>

		<i>коэффициенты линейной регрессии (не более 10 записей) и построить диаграмму разброса. Также можно дать задание на самостоятельный поиск исходных данных (или получение их), с последующим анализом.</i>	
Оценивание	---	Консультирует, обосновывает оценки.	Учащиеся самостоятельно оценивают работу (самооценка, взаимооценивание результатов работы одноклассников).
Рефлексия учебной деятельности	---	Благодарит учеников за урок.	Учащиеся называют тему урока, его этапы, перечисляют виды деятельности на каждом этапе, определяют предметное содержание. Делятся мнением о своей работе на уроке.