Специальный курс по выбору студента

Кафедры математических и компьютерных методов анализа

**Работа с данными в индустрии (1/2 г.)**

Лекторы: доцент О.В. Попов, приглашенные докладчики.

**Аннотация курса:** Задача спецкурса – дать учащемуся теоретические знания о работе с данными в индустрии, а также обзор инструментария. Базовые навыки работы будут осваиваться студентами в формате практикума или live coding.

**Программа курса**

|  |  |
| --- | --- |
| Тема 1 | Знакомство с SQL.   * Основные термины * Базовые элементы языка. * Синтаксис оператора SELECT. * Внесение изменений в БД. * 5 Запросы на объединение. |
| Тема 2 | Структура баз данных.   * Ключи, историчность * Откуда берутся данные (источники) |
| Тема 3 | Python   * Знакомство с Python на примере jyputernotebook * Основные библиотеки работы с данными(PandasNumPy) |
| Тема 4 | * Библиотеки для визуализация данных в Python |
| Тема 5 | Визуализация данных   * Способы представления информации. * Разновидности графиков, диаграмм * Применение |
| Тема 6 | Визуализация данных   * Визуализация в excel * Визуализация в питон * Практика |
| Тема 7 | Как проводить тесты А-Б, что еще бывает   * Что такое тесты и зачем они нужны * Простые примеры * Математика, статистические критерии, статистическая значимость |
| Тема 8 | Как проводить тесты А-Б, что еще бывает   * Что происходит на практике * Как сделать правильный тест * Множество примеров (хороших и плохих) |
| Тема 9 | Что измерять – какие бывают метрики.   * Как в зависимости от цели выбрать, на что смотреть. * Как делать выводы. * На какие конверсии можно смотреть * Примеры (с демонстрацией, что показательно, а что нет) |
| Тема 10 | Как следить за изменениями   * Мониторинг характеристик, отчеты * Примеры, когда что-то меняется ("что, почему и как засечь") |
| Тема 11 | Модели   * Какие используют аналитики и зачем (с примерами) * Простейшие прогнозирующие * Выбор параметров модели |
| Тема 12 | Модели   * Финансовые модели * Модели роста |
| Тема 13 | Линейная и логистическая регрессии   * Теория * Примеры |
| Тема 14 | Построение логистической регрессии   * Подготовка данных * Выбор признаков * AUC ROC, WoE |

Форма промежуточной аттестации – экзамен.