

Национальная олимпиада по анализу данных DANO

Заключительный этап

Задание

Базы данных распределяются между командами на основе их предпочтений согласно специальному алгоритму:

Шаг 0. Команды ранжируют все базы данных от наиболее предпочтительной к наименее.

Шаг 1. Для каждой базы данных выбираются те команды, которые указали ее первым приоритетом.

Шаг 2. Отобранные команды для базы ранжируются по медианному баллу ее участников и записываются все команды, для которых хватило мест (с каждой базой данных может работать определенное количество команд).

Шаг 3. Команды, которые были записаны в результате шага 1 выбывают из распределения и закрепляются за этой базой. Также выбывают базы, на которых закончились места.

Шаг 4. Процедура повторяется для второго, третьего и т.д. предпочтений.

Скачайте базу данных со специальной страницы (<https://dano.hse.ru/data>), изучите ее более внимательно: посмотрите на переменные, посчитайте описательные статистики, постройте предварительные графики и таблицы, обратите внимание на выбросы.

Продумайте все детали: исследовательский вопрос, гипотезу, механизм и пр.

Разработайте дизайн исследования:

- что вам нужно для того, чтобы ответить на исследовательский вопрос
- что нужно чтобы проверить ту или иную гипотезу
- какие таблицы и графики вам понадобятся
- какую информацию из них можно извлечь
- как интерпретировать получаемые результаты
- помогает ли это в вашем исследовании
- несет ли полезную информацию
- действительно ли эти построенные таблицы и графики необходимы и продвигают ваш проект или может быть необходимы другие
- какие методы и модели вам нужны
- в какой последовательности выполнять все расчеты и построения

Обращайте особое внимание на то, что все эти процедуры должны быть оправданы и продвигать вас в направлении поиска ответа на исследовательский вопрос.

Распределите задачи между членами команды. Установите сроки. Придерживайтесь взятых на себя обязательств и данных друг другу обещаний – делайте все в срок. Обсуждайте между собой полученные результаты, ищите наиболее удачный способ проверить ваши гипотезы, наиболее удачные графики и таблицы. Советуйтесь с ментором, обращайтесь к нему за

помощью – его задача помочь вам отобрать правильные идеи и подсказать как их технически реализовать.

Заведите общее облачное пространство, где будут храниться все ваши результаты. Структурируйте, создавайте необходимые папки, называйте папки и документы говорящими именами, оставляйте комментарии.

Как пройдет защита проектов?

Защита проектов пройдет в очном режиме. Ваша команда приедет в Москву и будет защищать проект перед экспертной комиссией.

Соберите все части проекта в единую презентацию. Презентация – основной результат, то, чем вы будете отчитываться по проекту. Будьте готовы показать и рабочие материалы: расчеты в том программном обеспечении, которым вы пользовались, скрипты с кодами, если вы использовали для работы Python или R .

Проверьте, что презентация имеет все необходимое: титульный слайд, на котором указано название проекта и ФИО исполнителей, слайд с описанием использованной базы данных и расшифровкой переменных. Возможно, содержательные части презентации следует разбить на разделы, подумайте на какие. Обязательно наличие слайда с заключением, где кратко представлены результаты и выводы, возможные рекомендации для социально-экономической политики на основе результатов вашего исследования.

Проверьте, что содержательная часть презентации включает в себя все необходимые графики и таблицы, что все грамотно структурировано на слайдах, даны краткие пояснения, выдержан единый стиль, подобраны читаемые шрифты и не режущая глаз цветовая гамма. Избегайте украшательств – это научный стиль, не должно быть ничего лишнего. Советуйтесь с ментором о логике построения презентации и о ее визуальной форме.

Подготовьте рассказ о вашем проекте. Распределите роли, кто что будет рассказывать, но помните, что каждый из членов команды должен разбираться во всем проекте, уметь объяснить любую его часть, уметь прокомментировать любой график или таблицу, любой слайд. Проверьте тайминг: вы должны уложиться в своем рассказе за 15 минут. Потренируйтесь друг перед другом, расскажите ваш проект друзьям или другим финалистам (устройте видеоконференцию), получите от них советы: что было непонятно, где нужно подробнее, где нужно сократить. Проведите свою презентацию ментору, прислушайтесь к его рекомендациям.