**Техническое задание на прохождение стажировки Python Developer\Data engineer**

Данное техническое задание предназначена для оценки возможностей стажера и выявления сильных и слабых сторон. Стажеру желательно, но необязательно строго следовать инструкциям в ТЗ, в случае отклонения от инструкции, будет достаточно объяснения. **Также важно осознать, что мы не требуем выполнения всех пунктов в приложенном ТЗ, нам важно оценить процесс подхода к решению проблемы.**

**Задача №1 Добыча и подготовка данных:**

* 1. Собрать данные по актуальным объявлениям по продаже авто в городе Алматы. Минимальный набор полей, который необходимо получить (минимальное кол-во записей +10 тысяч):
* Марка
* Модель
* Год выпуска авто
* Цена авто

**Если пункт выполнить не удалось, можете воспользоваться прикрепленным массивом данных – данные.csv**

* 1. Используя возможности языка Python, обработайте полученный массив данных (из пункта 1.1):
     1. Удалите записи с неполным набором полей
     2. Почистите\унифицируйте поля марка, модель (напр. «LandRover» -> «Land Rover»)
     3. Приведите поля год выпуска и цена авто к числовому типу, для цены в валюте тенге (напр. «100 000 тг.» -> «100000»)
  2. Используя возможности языка Python, нормализуйте данные на несколько справочников
  3. Загрузите полученные данные (из пункта 1.3) в базу данных (любой SQL СУБД):
     1. Создайте таблицы и заполните их, используя полученные данные
     2. Настройте зависимости и индексы между таблицами

**Задача №2 Подготовка SQL функции**

2.1 Напишите SQL запрос, который выведет данный набор информации в одном OUTPUT-е:

* Марка + Модель (поле должно быть уникальным, те. не дублироваться в результате)
* Средняя цена по срезу (Марка + Модель)
* Максимальная цена по срезу
* Минимальная цена по срезу
* Кол-во записей по срезу
* Кол-во уникальных годов авто по срезу
* Самый популярный год авто в срезе (год авто, у которого самое большое кол-во записей)

2.2 Оберните запрос (из пункта 2.1) в функцию SQL, с двумя, необязательными аргументами на вход:

* Марка
* Модель

Которые будут фильтровать результат запроса по полю Марка + Модель. В случае если ничего не передается, фильтрация не применяется, а возвращается полный результат запроса.

На выходе функция возвращает результат выполнения запроса (из пункта 2.1) применёнными условиями

**Задача №3 Разработка рассылки в Telegram канал на Python**

3.1 Необходимо создать Telegram канал, в который будет отправляться рассылка в виде Excel файла, внутри которого будет храниться результат выполнения функции (из пункта 2.2). Для выполнения будет достаточно написать Python скрипт, при запуске которого будет отправляться сообщение в Telegram канал

3.2\* Настройте бота в Telegramканале, который сможет отправлять файл по запросу в чате. Например, пользователь обращается к боту, запрашивает, вывести информацию по конкретной марке модели, и бот отправляет в канал, файлик с результатом запроса по данной марке и модели авто

**Задача №4 Отправить результат выполнения ТЗ**

4.1 Распишите в одном Word файле весь ход работ, чем детальнее, тем лучше. Желательный формат:

Пункт 1.1 – подробное объяснение

**Даже в случае невыполнения одного (всех) пунктов, обязательно опишите ваши мысли, как бы вы их сделали, и какие пробелы обнаружили в своих знаниях при выполнении задания**

4.2 Соберите все исходники, материалы и файлы (вкл. Пункт 4.1) в один архив

Базу можете выгрузить dump и тоже приложить.

4.3 Отправить архив в What’s Up на номер 8 707 555 8210 до истечения отведенного времени