# Задание на финальный этап дисциплины ИИ и АД олимпиады Волга-ИТ 2024

Подавляющее большинство людей потребляет воду и электричество. Также обычно люди живут в домах или квартирах, где как раз и потребляют данные ресурсы. Если эти ресурсы не генерируются самим человеком, то за них необходимо заплатить. Но некоторые другие услуги ЖКХ оплачиваются исходя из количества проживающих в доме или квартире, и поэтому нам важно знать это значение.

**Задача**: необходимо определить количество проживающих в доме или квартире на основе объёмов потребляемых ими коммунальных ресурсов.

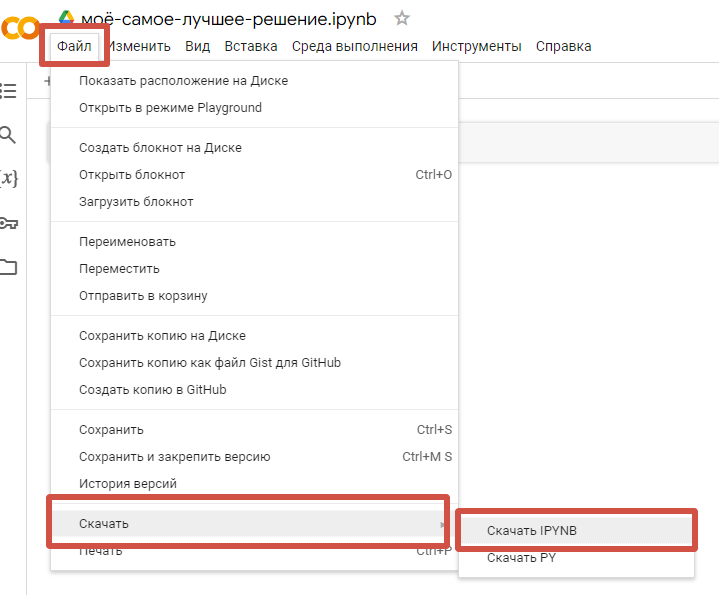
**На вход** предоставляется два файла:

* Обучающая выборка (*training.csv*), где уже указано количество проживающих.
* Часть контрольной выборки (*check.csv*) без указания количества проживающих в квартирах/домах. Служит для проверки модели.

**На выходе** должен быть предоставлен:

* Исходный код на python или jupiter notebook.  
  Если используется google colab или аналог, notebook должен быть оттуда выгружен.
* Файл (*result.csv*) с определённым количеством проживающих для каждой квартиры/дома из контрольной выборки (*check.csv*).

Как скачать notebook из google colab.



#### Особенности входных файлов

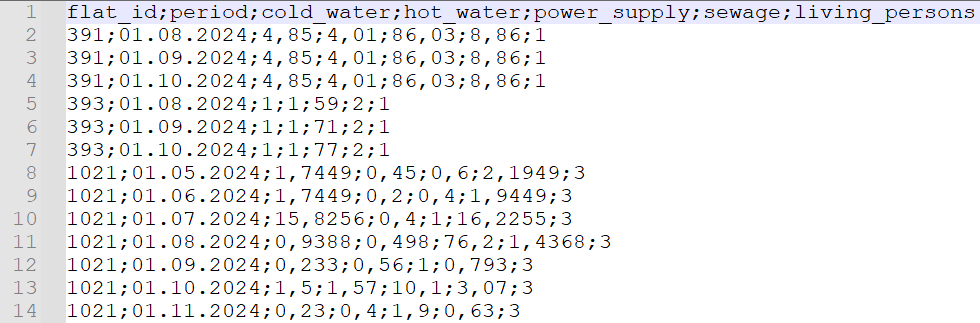
Формат всех файлов - CSV с заголовком.  
Разделители - «;», десятичный разделитель - «,», формат даты - «ДД.ММ.ГГГГ».

**Файл training.csv**Основной файл для обучения модели.

Содержит следующие столбцы:

* flat\_id - идентификатор квартиры (дома);
* period - период, за которой учитывается объём ресурсов (1 число месяца);
* cold\_water - объём потребления холодной воды (за месяц);
* hot\_water - объём потребления горячей воды;
* power\_supply - объём потребления электроэнергии;
* sewage - объём водоотведения (обычно сумма холодной и горячей воды);
* living\_persons - **количество проживающих** по данным системы.

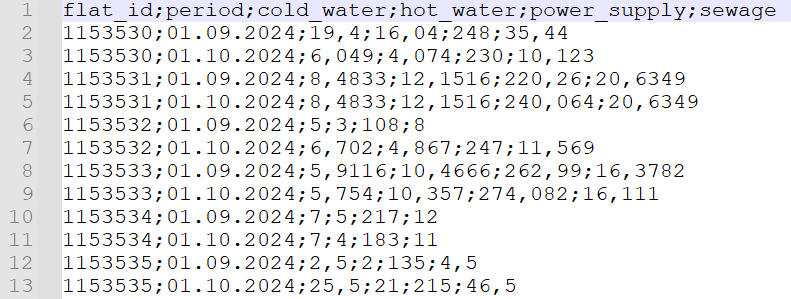
Пример содержимого файла *training.csv*.



**Файл check.csv**  
Содержит данные для проверки работы модели.

Формат - аналогичный *training.csv* кроме наличия столбца *living\_persons*.

Пример содержимого файла *check.csv*.

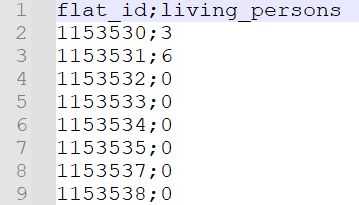


#### Особенности выходных файлов

На выходе решение должно генерировать **файл *result.csv*** со следующими столбцами:

* flat\_id - идентификатор квартиры (дома) из файла *check.csv*;
* living\_persons - **количество проживающих** (целое число!) по данным работы решения.

Пример содержимого файла *result.csv*.



#### Шаблон для google colab

Был разработан блокнот в google colab, где уже реализован корректный импорт и экспорт данных для выполнения задания.

#### Особенности данных

* Объём потребления ресурсов считается за месяц (поле *period*).
* Для каждой квартиры может быть указано несколько периодов.
* Количество проживающих указано и предсказывается на текущий месяц.
* Выгружались только данные с количеством проживающих **от 1 до 7**.
* В файле, отправляемом на проверку, количество проживающих должно быть целым числом.