**ТЕСТОВОЕ ЗАДАНИЕ**

Есть два файла, **src\_survey** и **src\_OnOff.**

В таблице **src\_survey** результат опроса респондентов - участников исследования, целью которого является проверка работы оборудования и регистрации респондентов на пиплметре (измерительное устройство для фиксации телесмотрения).

В рамках каждой процедуры в одном домохозяйстве (ДХ) проверяется, кто из респондентов находится в комнате с включенным телевизором. В ДХ может быть несколько приборов (телевизоров), могут проживать несколько респондентов. В опросе содержится информация по каждому телевизору и каждому респонденту. Далее, результат опроса сравнивается с данными регистрации респондента на пиплметре во время опроса, результат сравнения фиксируется в таблице (см. описание полей)

**Нужно проанализировать**:

1. Количество ситуаций (абс, %), когда прибор фактически выключен (данные в таблице **src\_OnOff**), а респондент утверждает, что включен
2. Количество ситуаций (абс, %), когда прибор фактически включен (данные в таблице **src\_OnOff**), а респондент утверждает, что выключен
3. Выделить ДХ с размером более 1 человека, среди них ситуации, телевизор фактически включен, респондент по данным опроса находится в комнате с включенным телевизором, регистрации нет, а у других респондентов в этом ДХ регистрация на этом телевизоре есть. Необходимо сгруппировать результат в зависимости от размера ДХ.

Описание полей:

* статус телевизора по утверждению респондента: *tv state*
* размер ДХ: *Кол-во человек в семье 4+*
* наличие фактической регистрации на пиплметре: *see*
* наличие респондента в комнате с телевизором согласно опросу: *display*
* идентификатор ДХ: *passport* (*PassportNumber* в таблице **src\_OnOff**)
* фактический статус телевизора: DeviceState (ON включен, OFF выключен)
* идентификатор респондента: *button*
* идентификатор телевизора: *unit* (*UnitNumber* в таблице **src\_OnOff**)

Мы ждем:

1. Скрипт Python
2. Результат анализа (данные + небольшой аналитический комментарий)