

Домашнее задание №2. Введение в Pandas и визуализацию данных

1. Структура данных. Даны три таблицы `people.csv`, `cars.csv`, `appartments.csv` со следующей структурой:

Таблица	Поле	Тип	Описание
people.csv	ID	число	Идентификатор
people.csv	Name	строка	Имя
people.csv	Height	число	Рост
people.csv	Weight	число	Вес
people.csv	Age	число	Возраст
people.csv	Sex	строка	Пол
people.csv	Profession	строка	Профессия
people.csv	Salary	Число	Зарплата
cars.csv	ID	число	Идентификатор
cars.csv	ID_person	число	Идентификатор владельца
cars.csv	Model	строка	Модель автомобиля
cars.csv	Power	число	Мощность
cars.csv	Price	число	Цена автомобиля
appartments.csv	ID	число	Идентификатор
appartments.csv	ID_person	число	Идентификатор владельца
appartments.csv	Square	число	Площадь квартиры
appartments.csv	Rooms	число	Количество комнат
appartments.csv	Price	число	Цена квартиры

0. Используя методы из `numpy.random`, случайно сгенерируйте по два пункта для каждого последующего задания, которые вам предстоит выполнить. В качестве сида используйте номер строки в гугл-доке с оценками.
1. Обработать пропущенные значения, используя одну из следующих стратегий:

0 a Удалить строки, содержащие пропуски;

1 b Заменить минимальным значением;

2 c Заменить наиболее часто встречающимся значением (модой);

3 **d Заменить медианой;**

4 e Заменить средним значением.

Добавить в таблицу столбцы-индикаторы, сигнализирующие, где была произведена замена пропущенного значения. Имя столбца-индикатора сделать путем прибавления к имени заменяемого столбца суффикса `'_indicator'`. Сохранить полученные `DataFrame`-ы в csv-файлы без индексов, но с заголовками столбцов с префиксом `'_stage1'`, то есть `people_stage1.csv`, `cars_stage1.csv`, `appartments_stage1.csv`.

2. В таблицу people добавить один из следующих столбцов:
 - a Бинарный признак того, что у человека есть хотя бы одна машина
 - b Числовой признак с количеством машин
 - c Бинарный признак того, что у человека есть хотя бы одна квартира или хотя бы одна машина
 - d Бинарный признак того, что у человека есть хотя бы одна машина
 - e Числовой признак с количеством квартир
3. Отсортировать таблицу cars, полученную в результате предыдущего пункта, следующим образом:
 - a По весу владельца, деленному на рост владельца по возрастанию;
 - b По возрасту владельца по возрастанию;
 - c По стоимости автомобиля, деленной на мощность по убыванию;
 - d По доходу владельца по убыванию;
 - e По стоимости по убыванию и после по мощности по возрастанию.

Сохранить данные с индексами и без заголовков в cars_stage3.csv
4. Посчитать среднее количество
 - a стоимости имущества (машин и квартир)
 - ✓ b квартир
 - c стоимости машин
 - d стоимости квартир
 - e машин

для людей, удовлетворяющим следующим критериям

 - a пенсионеров (мужчин старше 65 и женщин старше 60);
 - b студентов (но не студенток);
 - c женщин;
 - d мужчин;
 - ✓ e молодых (до 35 лет) преподавателей.
5. Выполнить следующее задание:
 - a Сохранить таблицу максимальных зарплат в зависимости от пола, профессии и марки машины.
 - b Сохранить таблицу средних зарплат в зависимости от пола, профессии и количества квартир.
 - c Сохранить таблицу минимальных зарплат в зависимости от пола, профессии и количества машин.

- d Сохранить таблицу минимальных и максиксамальных зарплат в зависимости от пола, профессии и количества машин.
 - e Сохранить таблицу средних, минимальных и максимальных зарплат в зависимости от пола, профессии и количества квартир.
6. Построить гистограмму зарплат для людей, удовлетворяющим следующим критериям:
- a пенсионеров (мужчин старше 65 и женщин старше 60);
 - b студентов (но не студенток);
 - c женщин;
 - d мужчин;
 - e молодых (до 35 лет) преподавателей.
- Сделать визуализацию несколькими способами (на одном графике, на разных, с разным уровнем детализации и т.д.)
7. Визуализировать следующую зависимость:
- a заплата в зависимости от возраста;
 - b количество квартир в зависимости от возраста;
 - c заплата в зависимости от профессии;
 - d пол в зависимости от роста;
 - e количество машин в зависимости от количества квартир.
8. (опционально) Предложите интересное задание задание по pandas и визуализации