

Задание 5: Анализ данных средствами pandas (часть 1)

Результатом выполнения задания является работающий исходный код на языке Python. Титульный лист и особый подход к оформлению не требуется. Допускается выполнение задания в интерактивных средах (Jupyter Notebook, Google Colaboratory).

Допускается выполнение задания на компьютере в одной из средств разработки (JetBrains PyCharm, Visual Studio Code).

Задание состоит из последовательности действий, пронумерованных числами. Предполагается после выполнения каждого пункта выводить результат (любым удобным для вас методом, но, чтобы результат каждого пункта был виден).

Ход работы

1. Создайте датафрейм с следующими данными:

Name	Age	Sex
Strom, Mrs. Wilhelm (Elna Matilda Persson)	29	female
Navratil, Mr. Michel ("Louis M Hoffman")	36,5	male
Minahan, Miss. Daisy E	33	female

2. Загрузите данные из файла:
https://drive.google.com/file/d/16gikf9vvMLyldkdHtoDYrPDbqTHP1BxG/view?usp=s_hare_link
3. Загрузите напрямую данные по ссылке:
<https://gist.githubusercontent.com/zaryanezrya/8b4ef51c707cb16d5e88a44dc00a1bb2/raw/41230f49c6268e072dbf102672f670be256922ab/gistfile1.txt> (read_csv).
4. Объедините датафреймы из п. 2, 3 в один датафрейм, удалите дубликаты строк.
5. Задайте в качестве индекса датафрейма id пассажира. Отсортируйте по индексу (sort_values).
6. Отобразите информацию и базовую статистику по датафрейму (info, describe)
7. Поменяйте местами нулевую по порядку строку с элементом с индексом 2 (loc, iloc).
8. В столбце Sex замените female/male на f/m (map).

9. Найдите билеты по которым плыло 6 и более человек, отобразите список людей (groupby, count).
10. Найдите людей, которые плыли в одной каюте с людьми из датафрейма из п1.
11. Добавьте колонку "BirthYear" с годом рождения каждого пассажира.
12. Добавьте колонку "Companion" с количеством попутчиков в одной каюте с этим пассажиром.
13. Поменяйте местами нулевую по порядку строку с элементом с индексом 1.
14. Запишите получившийся датафрейм в csv файл (to_csv).
15. Отобразите топ-10 людей, которые больше остальных заплатили за билет.
16. Отобразите сколько женщин и мужчин выжило/не выжило.
17. Отобразите сколько в каком классе пассажиров выжило/не выжило.

В помощь: <https://colab.research.google.com/drive/1lqhz7ZWfm2zcvYvqjEXJ2j-JlpY2YIZX?usp=sharing>