Московский Государственный Технический Университет

им. Н.Э. Баумана

Отчет по лабораторной работе №2

по курсу

Технологии Машинного Обучения

# Выполнил:

## Муравьев О.М.

ИУ5-62

\_\_\_\_\_\_\_\_\_\_\_\_\_

# Проверил:

## Гапанюк Ю.Е.

\_\_\_\_\_\_\_\_\_\_\_\_\_

Москва, 2019

# Задание

Часть 1.

Выполните первое демонстрационное задание "demo assignment" под названием "Exploratory data analysis with Pandas" со страницы курса <https://mlcourse.ai/assignments>

Задание заключается в том, чтобы ответить на вопросы, по повожу датасета, используя запросы pandas.

Часть 2.

Выполните следующие запросы с использованием двух различных библиотек - [Pandas](https://pandas.pydata.org/) и [PandaSQL](https://github.com/yhat/pandasql):

* один произвольный запрос на соединение двух наборов данных
* один произвольный запрос на группировку набора данных с использованием функций агрегирования

Сравните время выполнения каждого запроса в Pandas и PandaSQL.

# Код и результаты выполнения

1. Подключаем библиотеки:

import numpy as np

import pandas as pd

pd.set\_option('display.max.columns', 100)

%matplotlib inline

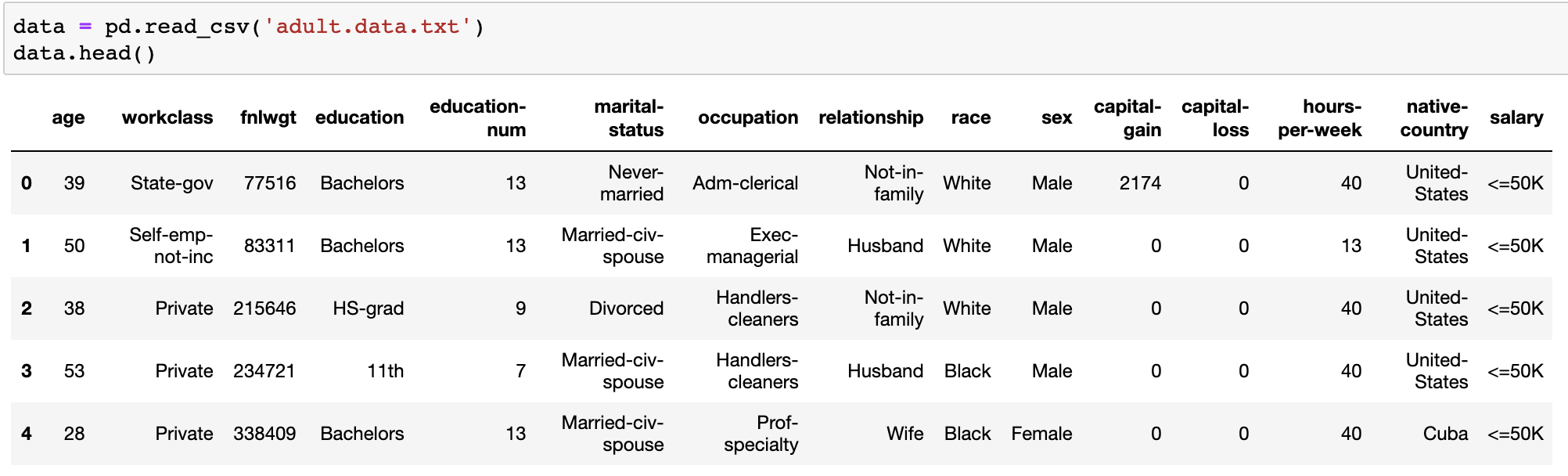
import matplotlib.pyplot as plt

import seaborn as sns

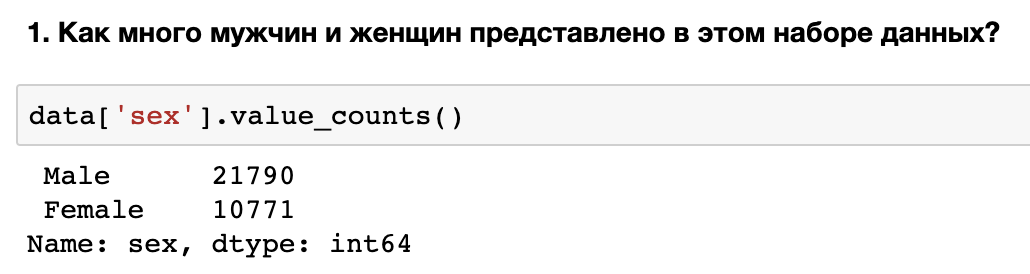
import warnings

warnings.filterwarnings('ignore')

1. Подключаем набор данных

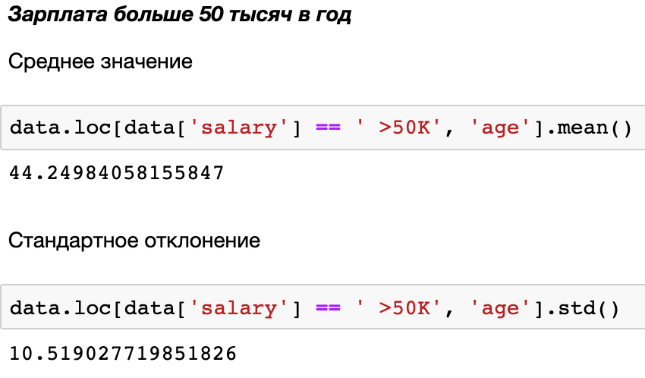


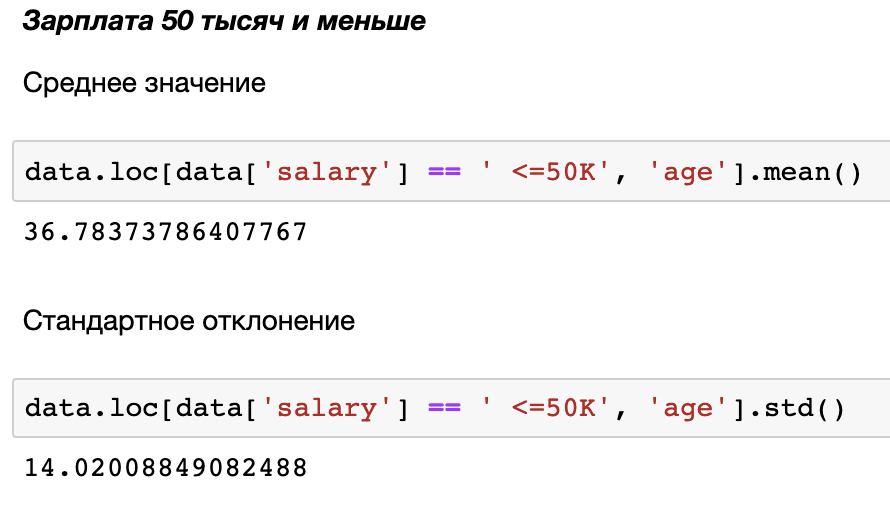
1. Запросы

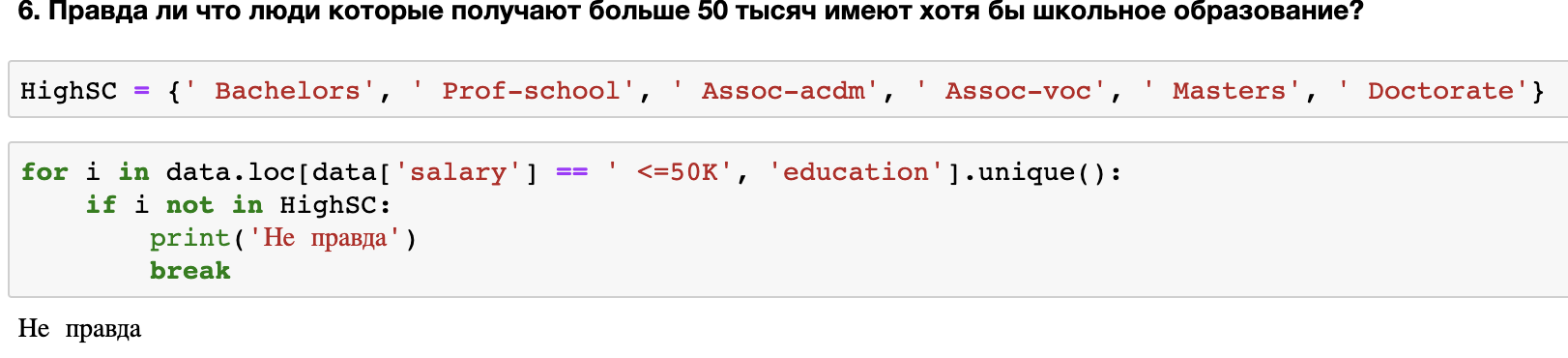


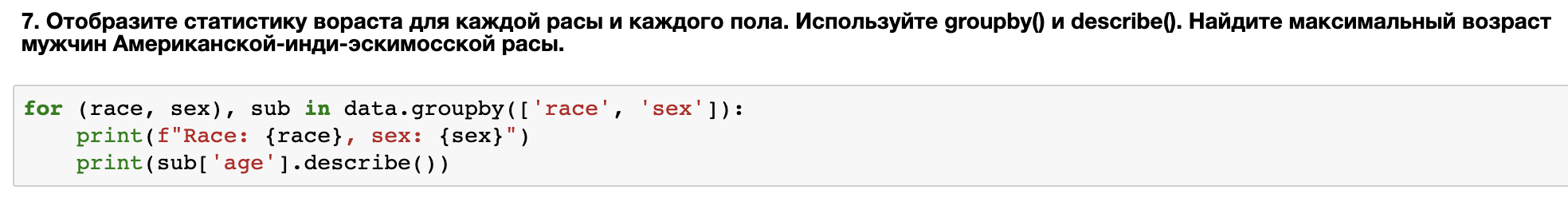


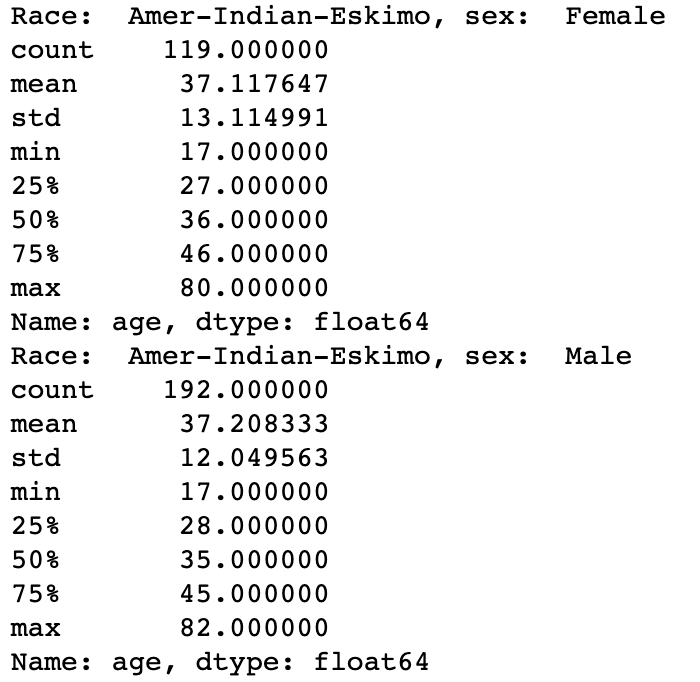
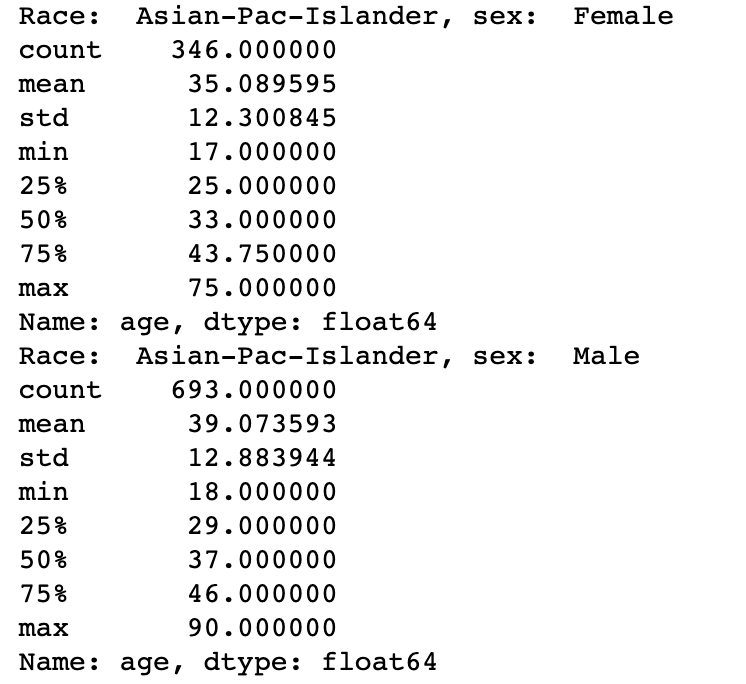
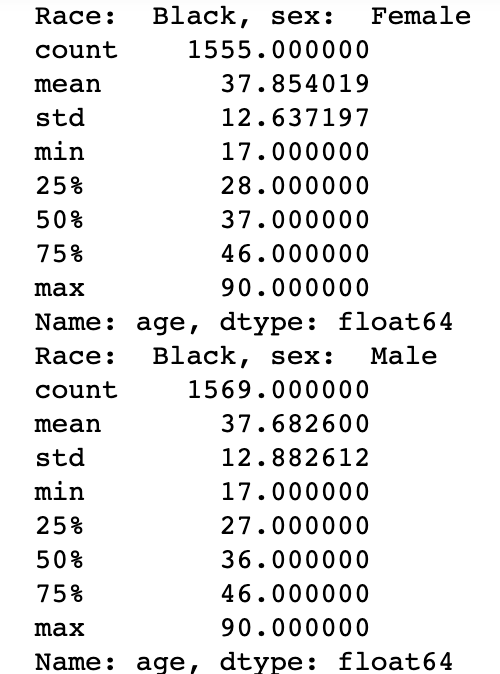
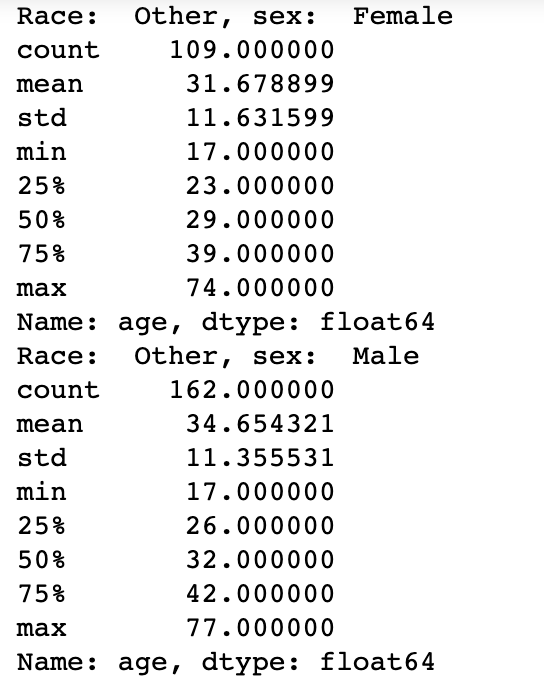
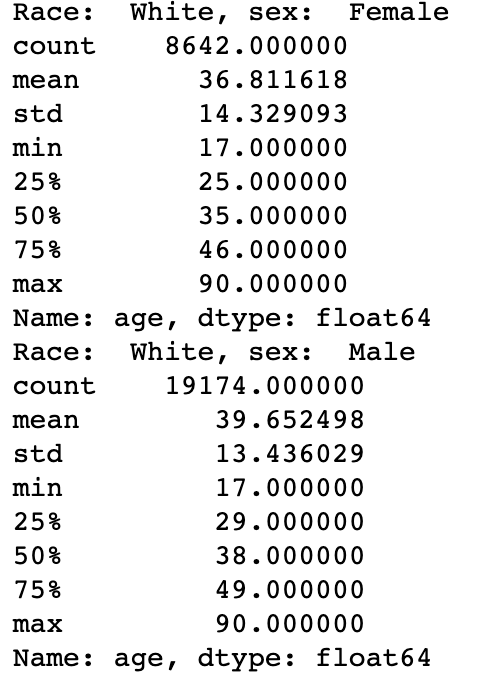


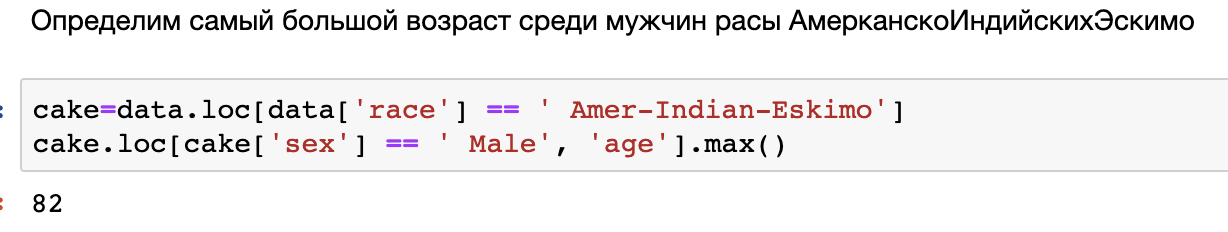








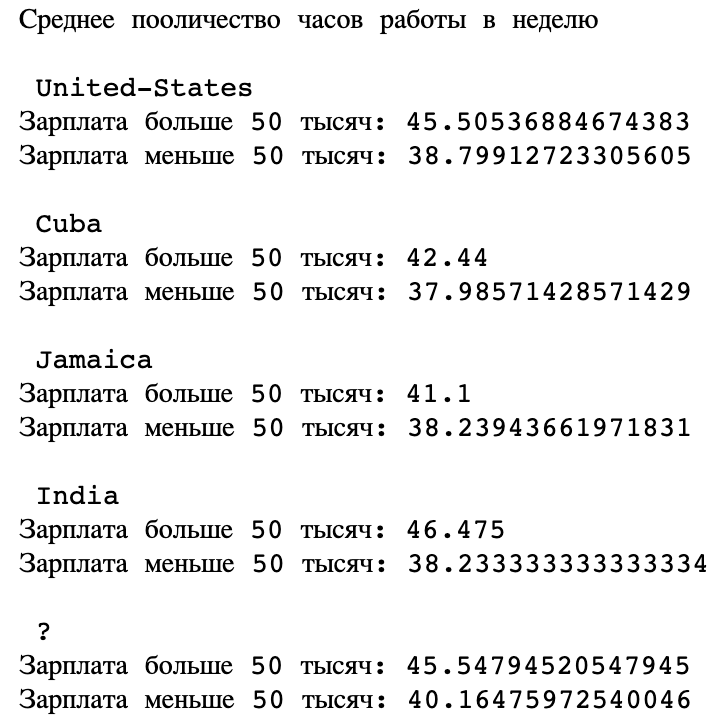
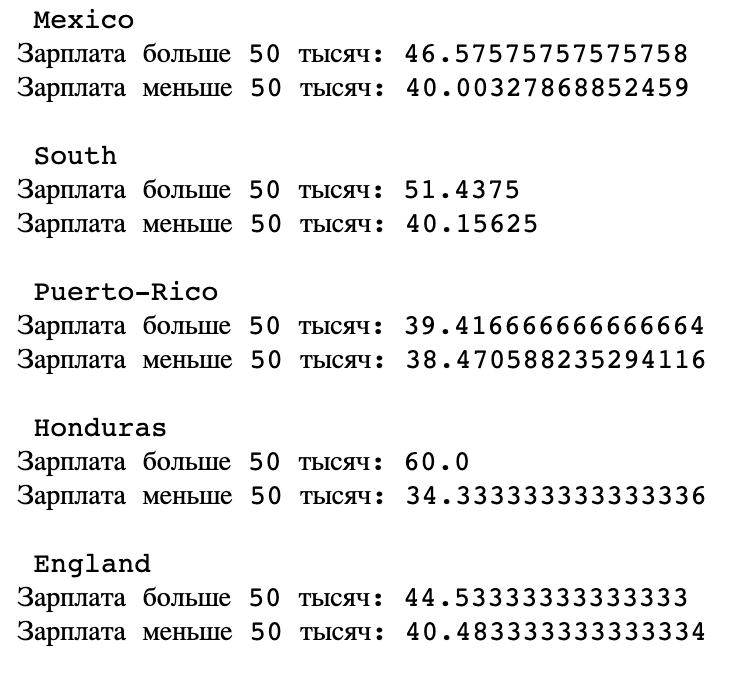
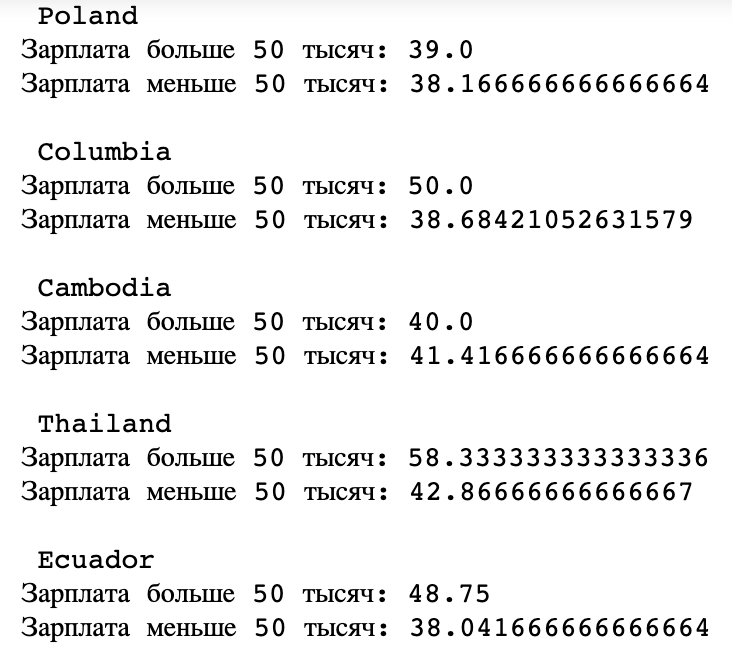
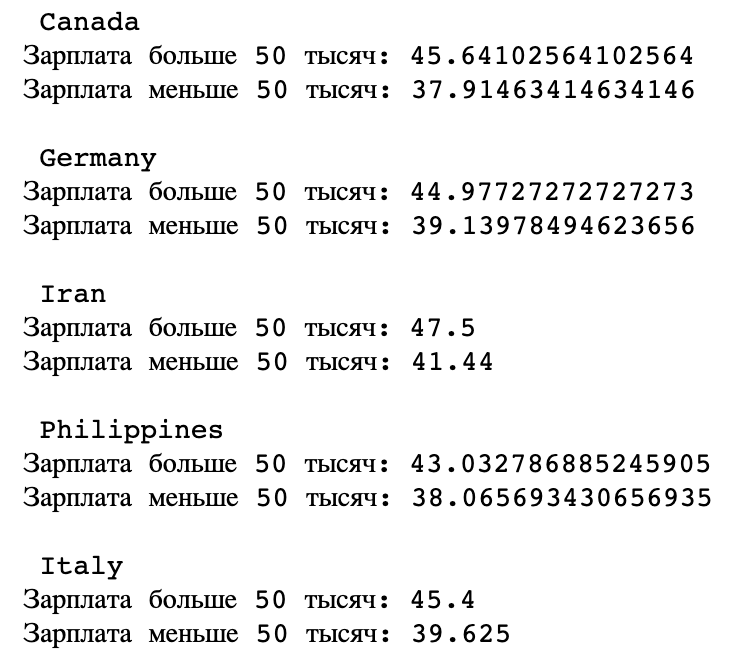
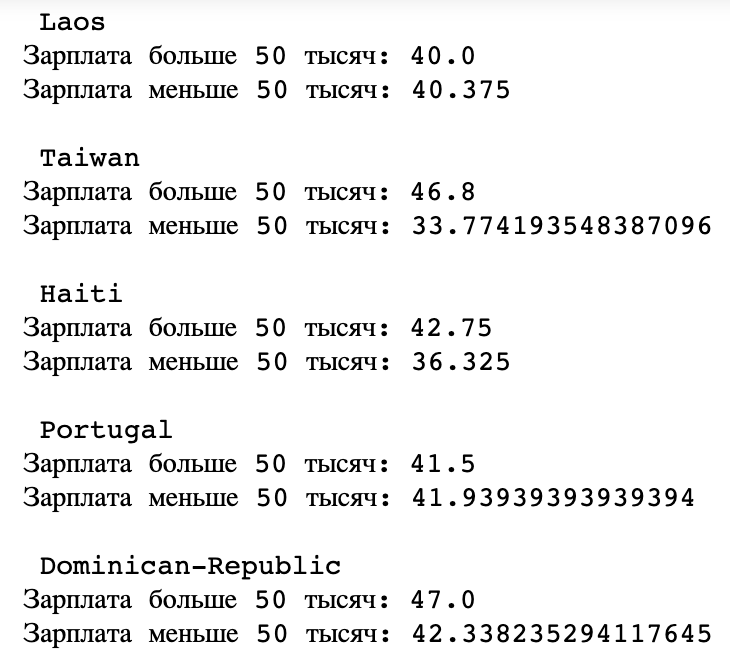
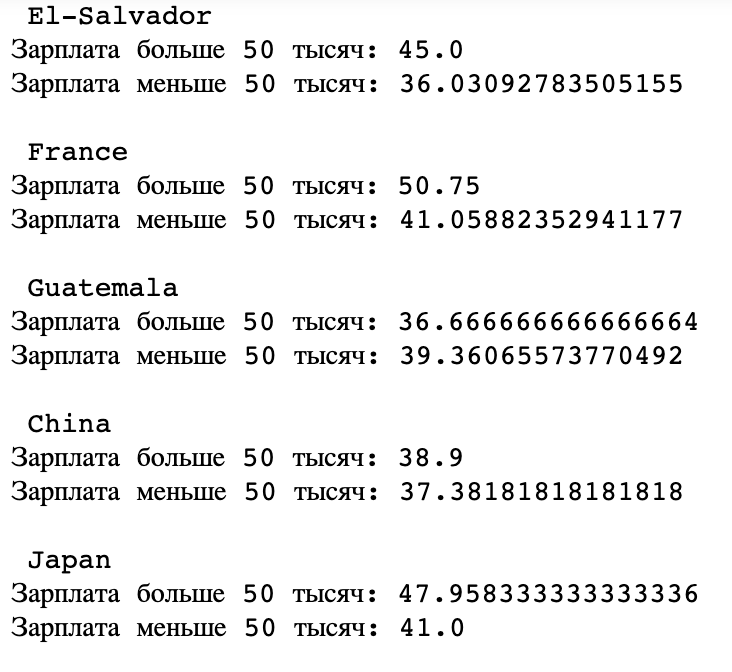
    

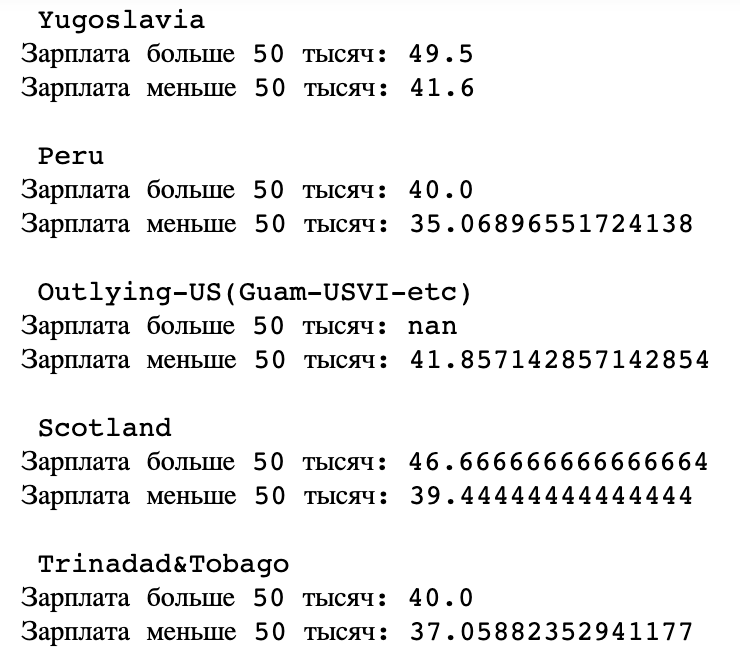
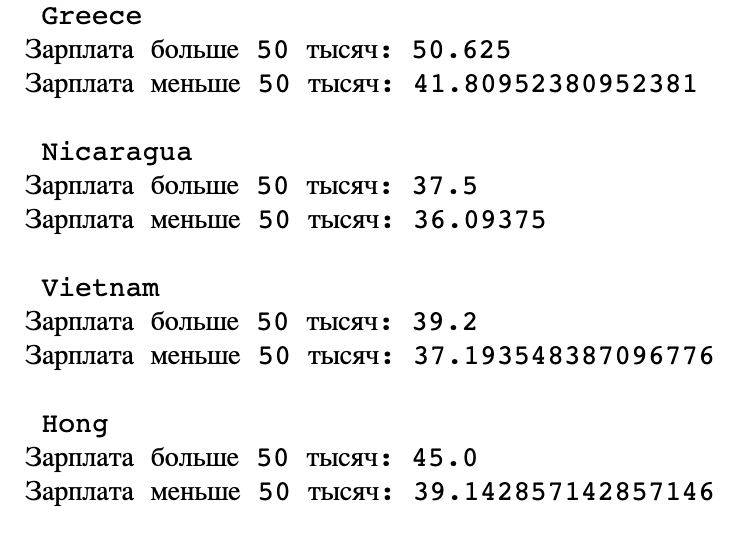
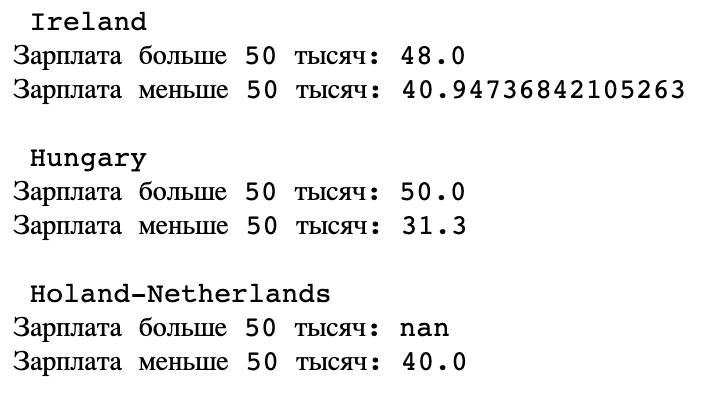








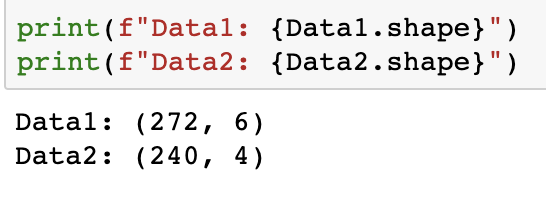
    

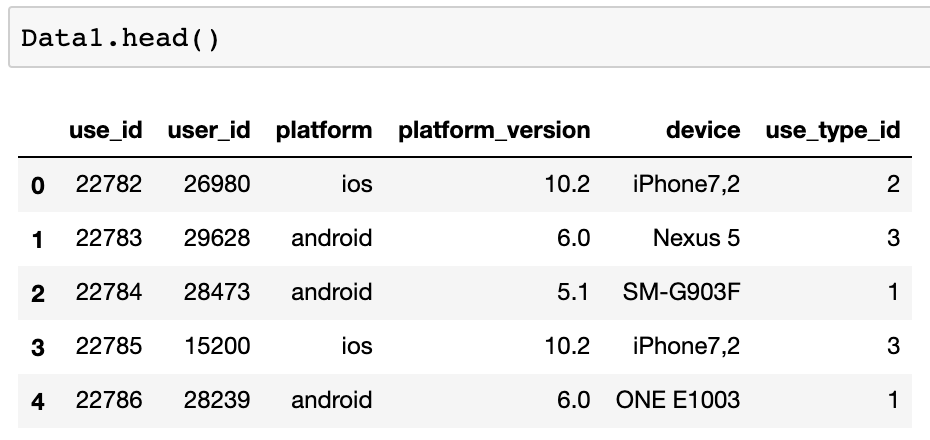
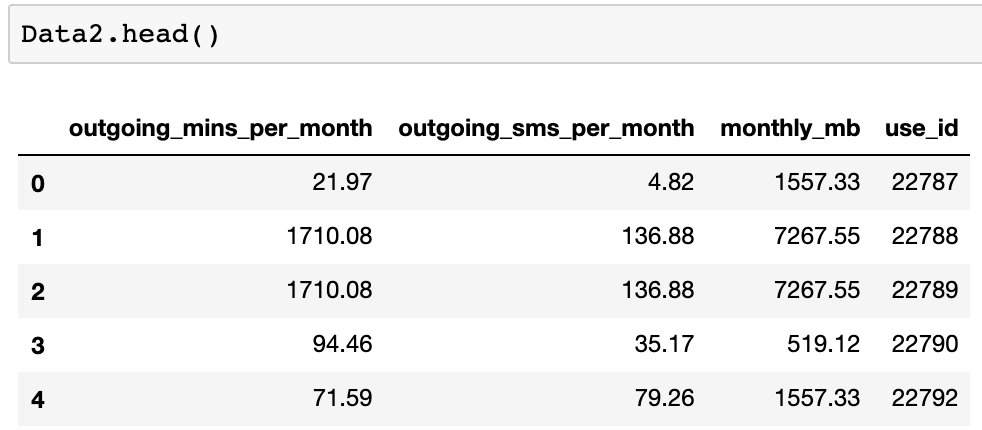
  



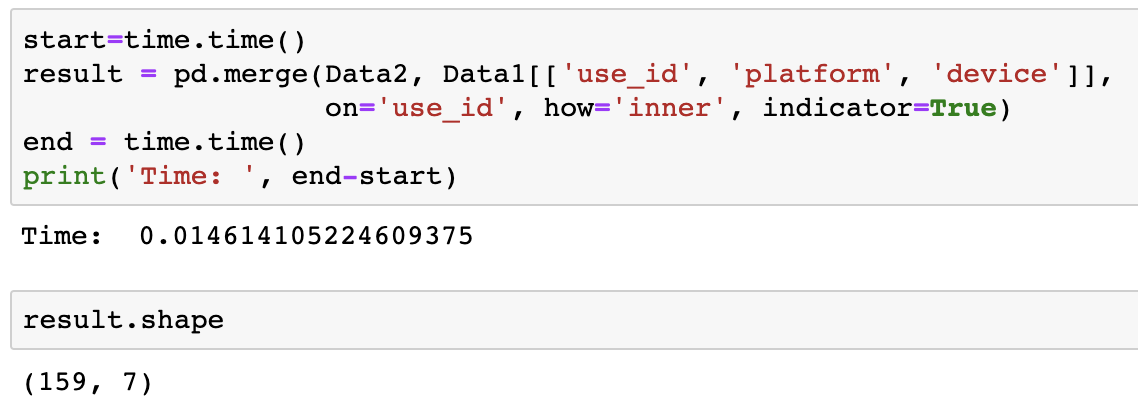
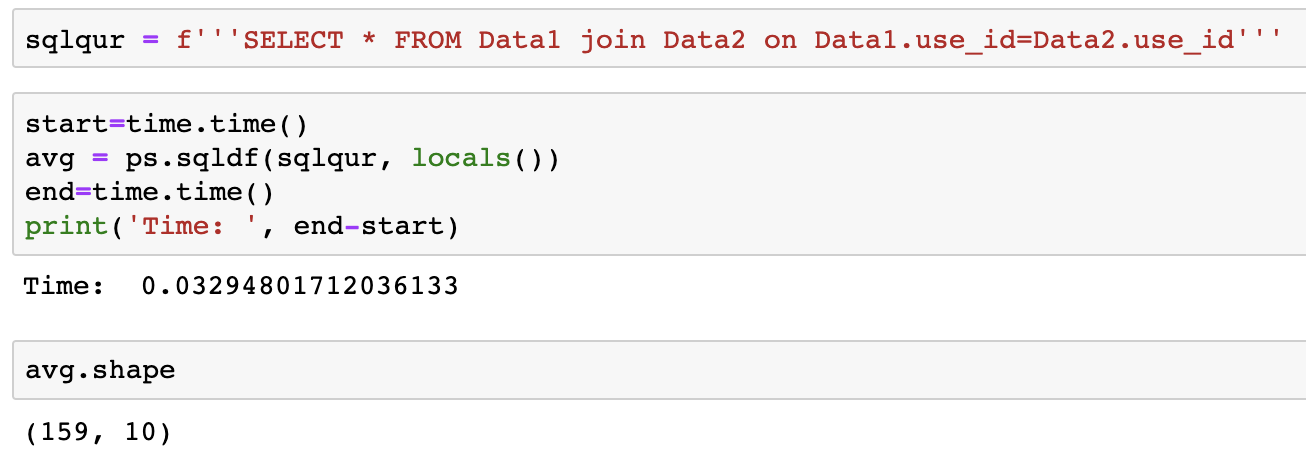
1. Часть 2

Импортируем библиотеки и данные

Объединение таблиц

Группировка

