МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

федеральное государственное бюджетное образовательное учреждение

высшего образования

''УЛЬЯНОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ''

Кафедра «Информационно-вычислительные комплексы»

«Методы искусственного интеллекта»

Отчет по лабораторной работе №3

**Выполнила**:

студентка гр. ИСТбд-41

Кашицына Н. А.

**Проверил**:

преподаватель

Шишкин В. В

Ульяновск, 2022

**Исследование библиотек CSV, pandas**

**Задание:**

1. Создать симулированный набор данных и записать его на диск в виде csv файла со следующими параметрами:

- количество строк не менее 1000 (задается случайным образом);

- структура набора:

- табельный номер;

- Фамилия И.О.;

- пол;

- год рождения;

- год начала работы в компании;

- подразделение;

- должность;

- оклад;

- количество выполненных проектов

2. Прочитать сгенерированный набор данных в виде списков и получить с помощью программирования и методов библиотеки numpy для разных по типу признаков столбцов (не менее 3) основные статистические характеристики (например, для порядкового типа: минимум, максимум, среднее, дисперсия, стандартное отклонение, медиана, мода).

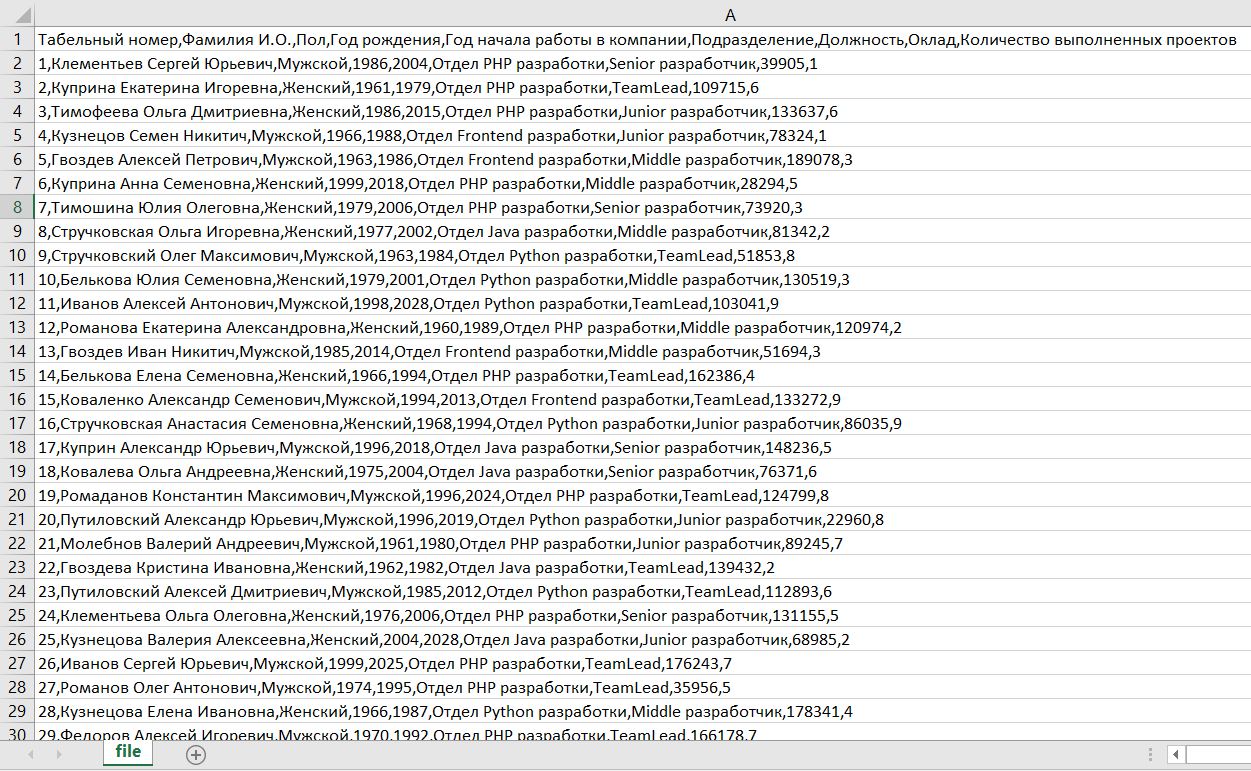
3. Прочитать сгенерированный набор данных в виде датафрейма и получить с помощью методов библиотеки pandas для тех же столбцов те же статистические характеристики. Продемонстрировать применение не менее 3 методов библиотеки pandas.

4. Построить не менее 3 разнотипных графиков.

5. Оценить возможности библиотек csv, numpy, pandas в форме отчета по лабораторной работе.

**Результат:**

1) Набор данных в файле csv:



2) Статистические характеристики:

Numpy

Количество сотрудников: 1119

Опыт работы

Минимальный стаж: -12

Максимальный стаж: 44

Средний стаж: 16.29

Заработная плата

Минимальная зарплата, в рублях: 20039

Максимальная зарплата, в рублях: 199811

Средняя зарплата, в рублях: 111621.94

Медианное значение зарплаты, в рублях: 111903.5

Дисперсия зарплаты, в рублях: 2586449827.15

Стандартное отклонение зарплаты, в рублях: 50857.15

Проекты

Максимальное количество проектов на одного сотрудника: 10

Минимальное количество проектов на одного сотрудника: 1

Среднее количество проектов на одного сотрудника: 5.45

Количество выполненных проектов: 6248

Pandas

Количество сотрудников: 1146

Проекты

Максимальное количество проектов на одного сотрудника: 10

Минимальное количество проектов на одного сотрудника: 1

Среднее количество проектов на одного сотрудника: 5.45

Количество выполненных проектов: 6248

Зарплата

Максимальная зарплата, в рублях: 199811

Минимальная зарплата, в рублях: 20039

Средняя зарплата, в рублях: 111621.94

Медианное значение зарплаты, в рублях: 111903.5

Дисперсия зарплаты, в рублях: 2588708735.3

Стандартное отклонение зарплаты, в рублях: 50879.35

Отделы

Количество отделов: 4

Количество сотрудников в отделе Java разработки: 286

Количество сотрудников в отделе Python разработки: 301

Количество сотрудников в отделе Frontend разработки: 0

Количество сотрудников в отделе PHP разработки: 291