

Лабораторная работа №3

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.

Оценка возможностей библиотек `csv`, `numpy`, `pandas`.

CSV

CSV - формат представления табличных данных, где каждая строка файла - строка таблицы, разделитель может быть любым. Разработан для работы с CSV файлами, созданными в Excel, легко настраивается для работы с любыми разновидностями этого формата. CSV содержит объекты и методы

2. Вывод статистических данных по прочитанным спискам при помощи numpy

```
-----Основные статистические характеристики с использованием Numpy-----
Количество сотрудников: 1203 чел.
Количество мужчин: 587 чел.
Количество женщин: 616 чел.

Год рождения самого возрастного сотрудника: 2000 год
Год рождения самого молодого сотрудника: 1960 год

Количество должностей: 11

Среднее арифметическое значение оклада: 57798.26 руб.
Размах оклада: 84853 руб.
Медианное значение оклада: 57008 руб.
Стандартное отклонение оклада: 24230.86 руб.
Дисперсия оклада: 587134597.26 руб.

Суммарное количество выполненных проектов: 29789 шт.
Среднее количество выполненных проектов сотрудника: 25 шт.
Мода количества выполненных проектов: 5 шт.
```

3. Вывод статистических данных по прочитанным датафреймам при помощи pandas

```
-----Основные статистические характеристики с использованием Pandas-----
Количество сотрудников: 1203 чел.
Количество мужчин: 587 чел.
Количество женщин: 616 чел.

Год рождения самого возрастного сотрудника: 2000 год
Год рождения самого молодого сотрудника: 1960 год

Количество должностей: 11

Среднее арифметическое значение оклада: 57798.26 руб.
Размах оклада: 84853 руб.
Медианное значение оклада: 57008 руб.
Стандартное отклонение оклада: 24240.94 руб.
Дисперсия оклада: 587623061.98 руб.

Суммарное выполненных количество проектов: 29789 шт.
Среднее количество выполненных проектов на одного сотрудника: 25 шт.
Мода количества выполненных проектов: 5 шт.
```

4. Построенные разнотипные графики

График с использованием функции plot()

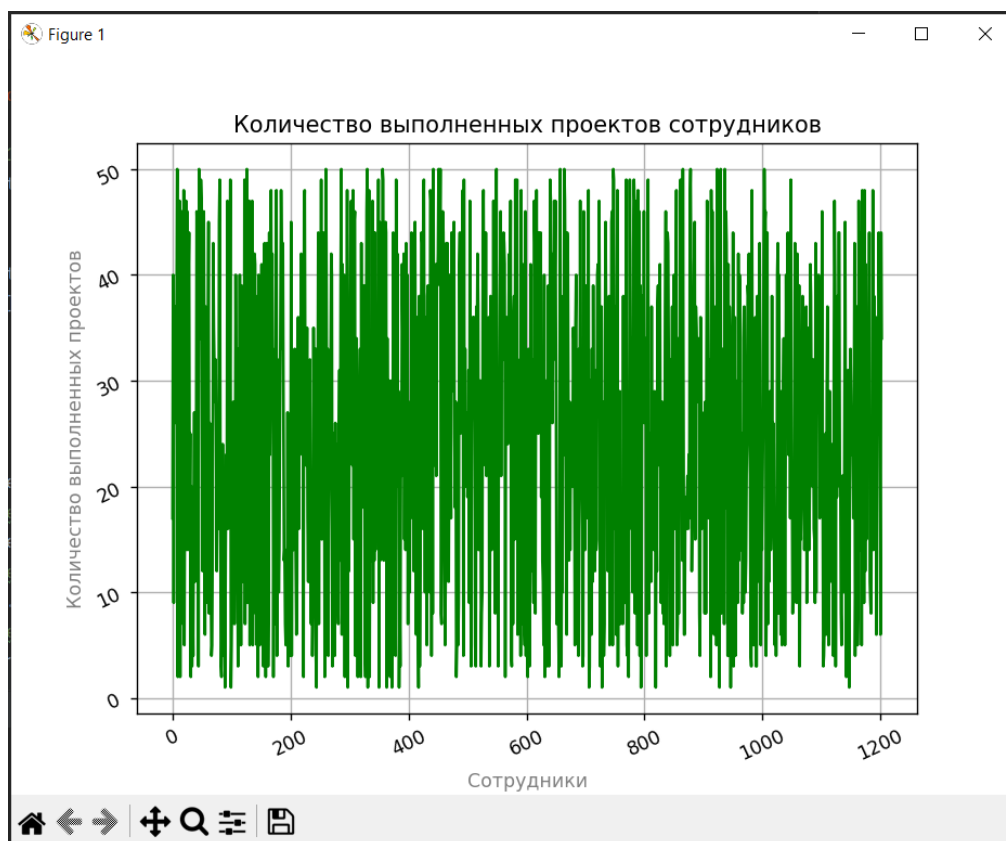


График с использованием функции hist()

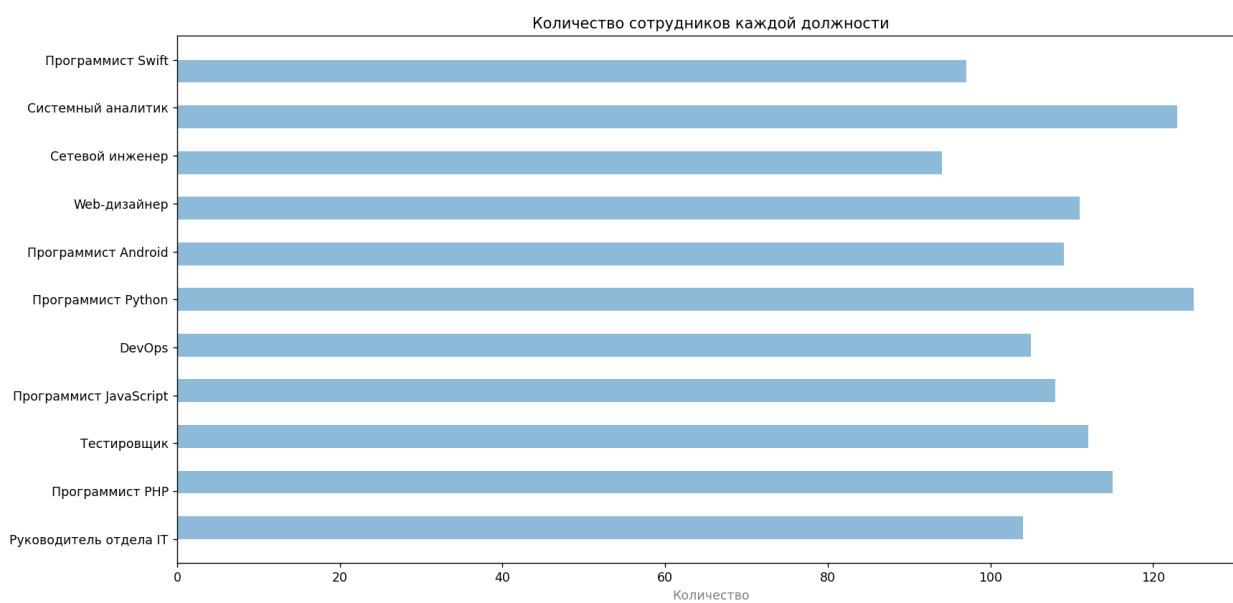


График с использованием функции barh()

