****Московский государственный технический университет им. Н.Э. Баумана  
Кафедра «Системы обработки информации и управления»

Лабораторная работа №1  
по дисциплине  
«Методы машинного обучения»  
на тему

«Разведочный анализ данных. Исследование и  
визуализация данных»

Выполнил:  
студент группы ИУ5-22М  
Вей Пхьоу Ту

Москва — 2021 г.

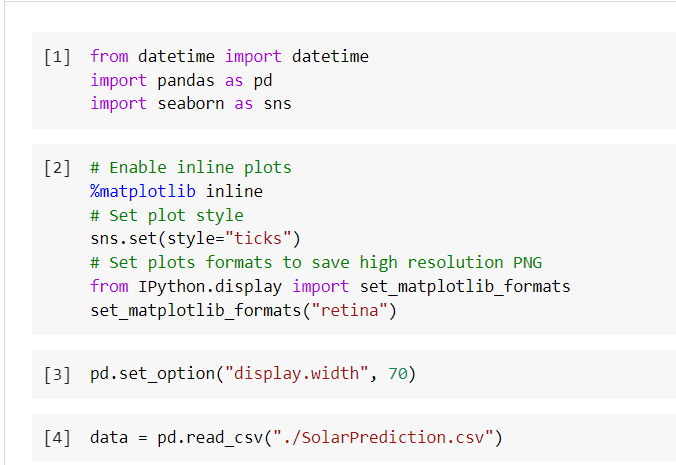
1. **Цель лабораторной работы**Изучить различные методы визуализации данных [1].

**2. Задание**Требуется выполнить следующие действия [1]:  
• Выбрать набор данных (датасет).  
• Создать ноутбук, который содержит следующие разделы:  
 1. Текстовое описание выбранного набора данных.  
 2. Основные характеристики датасета.  
 3. Визуальное исследование датасета.  
 4. Информация о корреляции признаков.  
• Сформировать отчет и разместить его в своем репозитории на GitHub

**3. Ход выполнения работы  
 3.1. Текстовое описание набора данных**

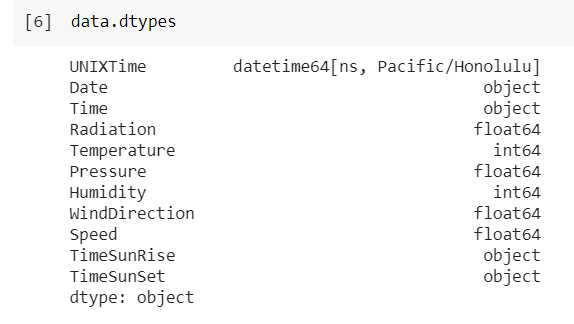
Сердце - это удивительный орган. Он бьется ровно, ровно, примерно от 60 до 100 раз каждую минуту. Это примерно 100 000 раз в день. Иногда твое сердце выходит из ритма. Ваш врач называет нерегулярное или неправильное сердцебиение аритмией. Аритмия (также называемая дисритмией) может вызывать неравномерное сердцебиение или сердцебиение, которое либо слишком медленное, либо слишком быстрое.

**3.2. Основные характеристики набора данных**  
Подключим все необходимые библиотеки [1]

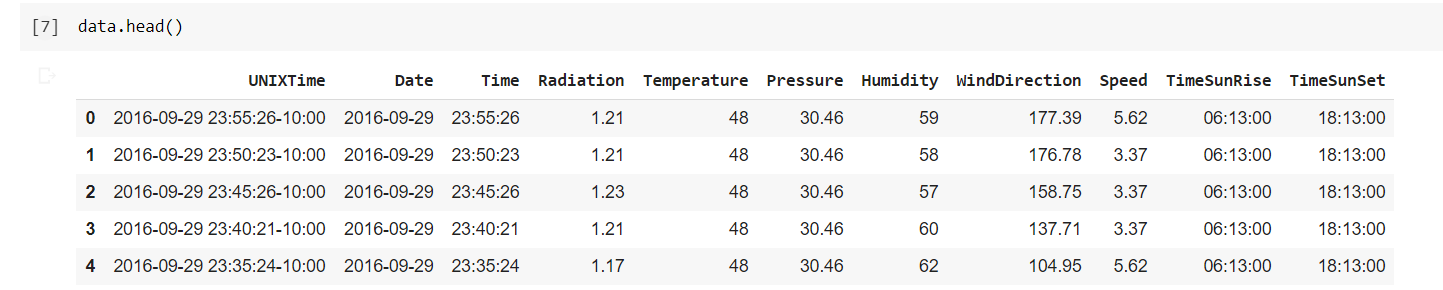


Настроим отображение графиков [3,4]:

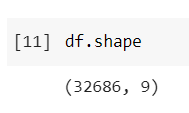
Загрузим непосредственно данные[5]



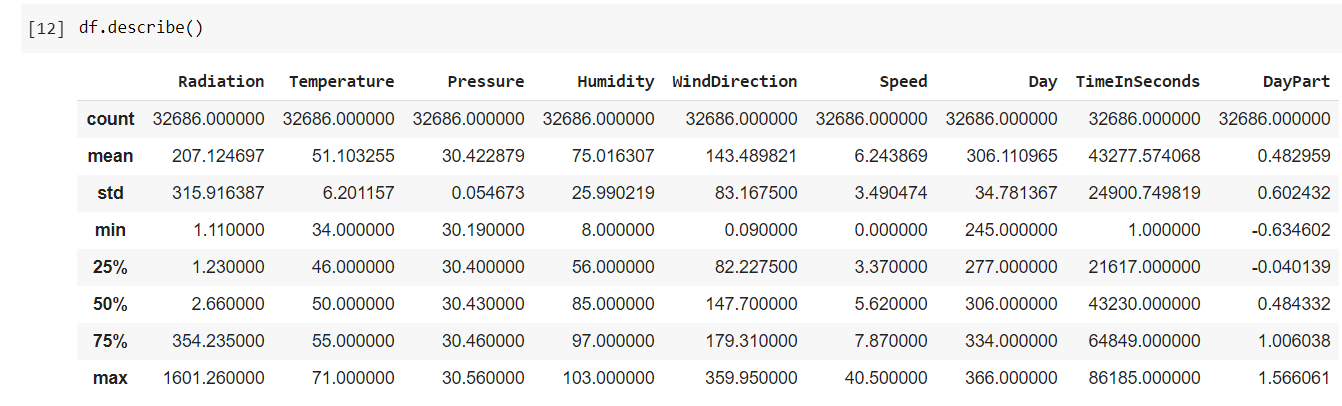
Посмотрим на данные в данном наборе данных:



Проверим размер набора данных:

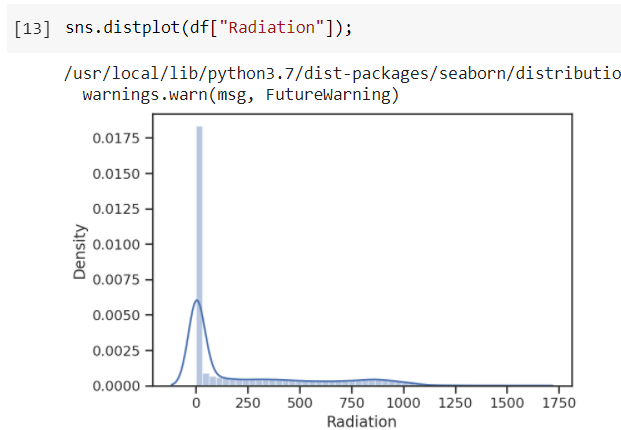


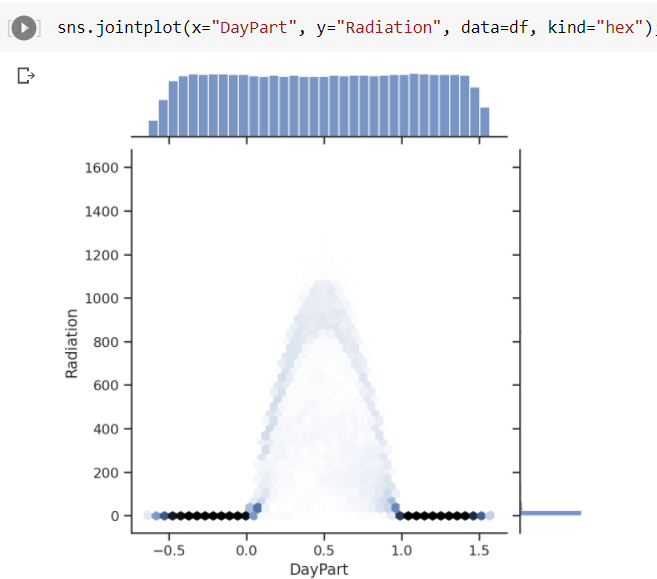
Проверим основные статистические характеристики набора данных:

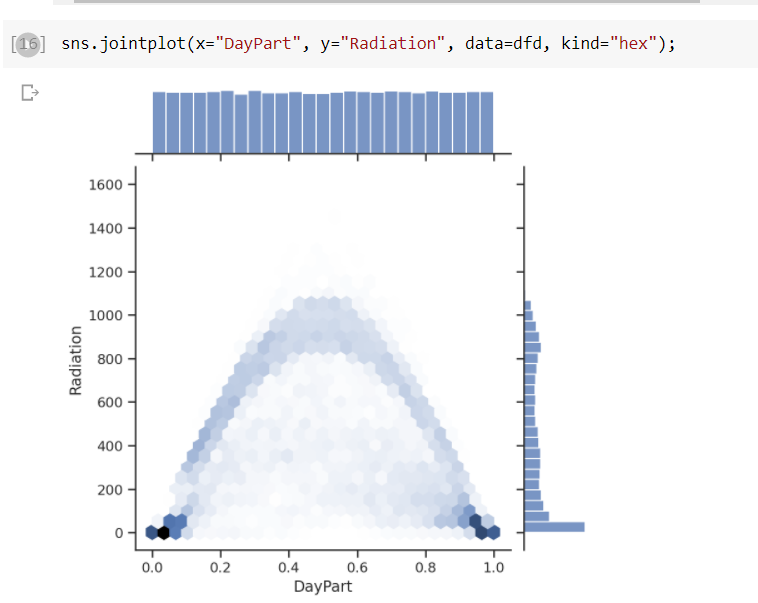
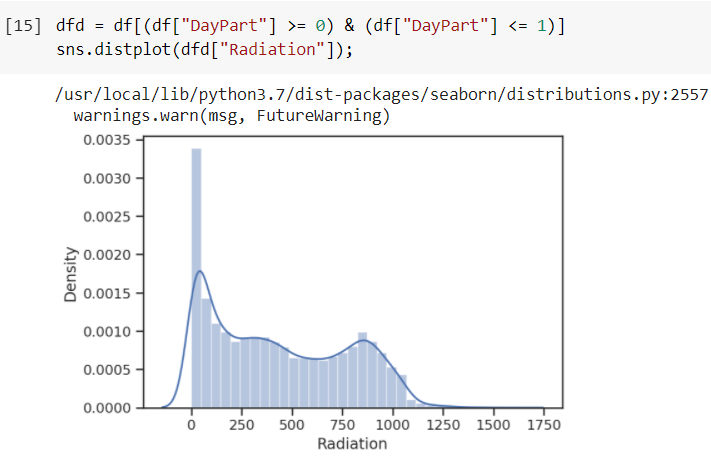


**3.3. Визуальное исследование датасета**

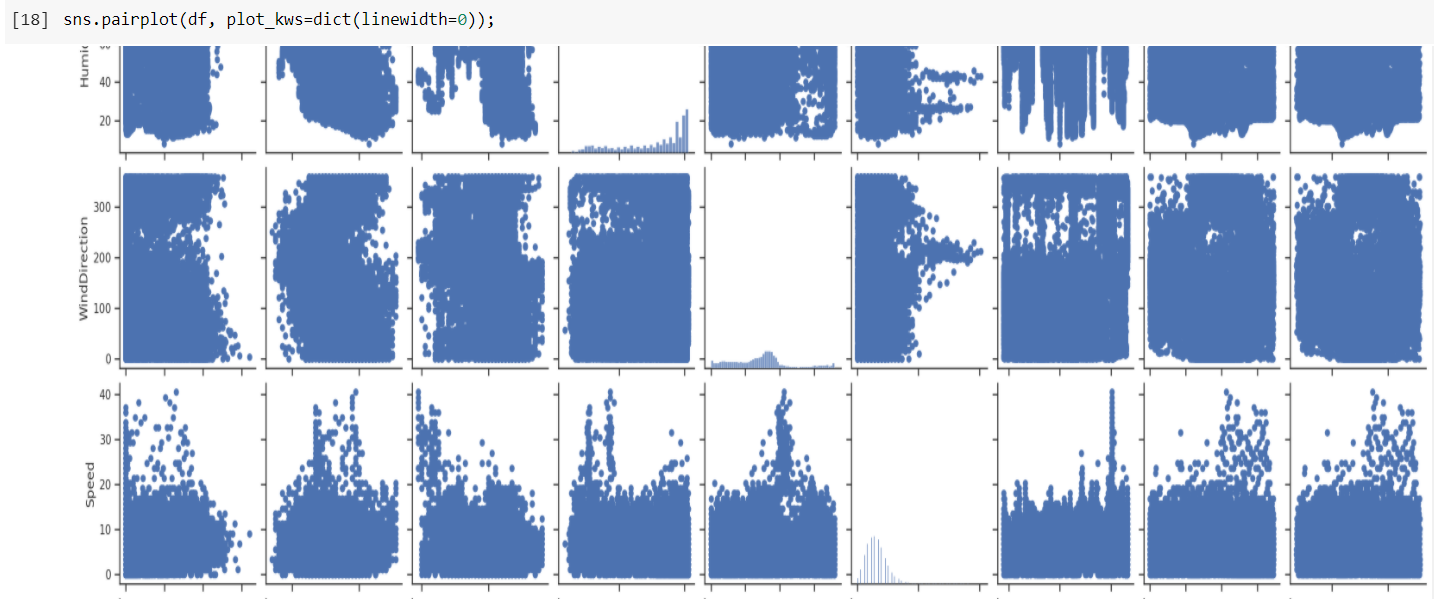
Давайте оценим распределение целевого атрибута - Рейтинг:



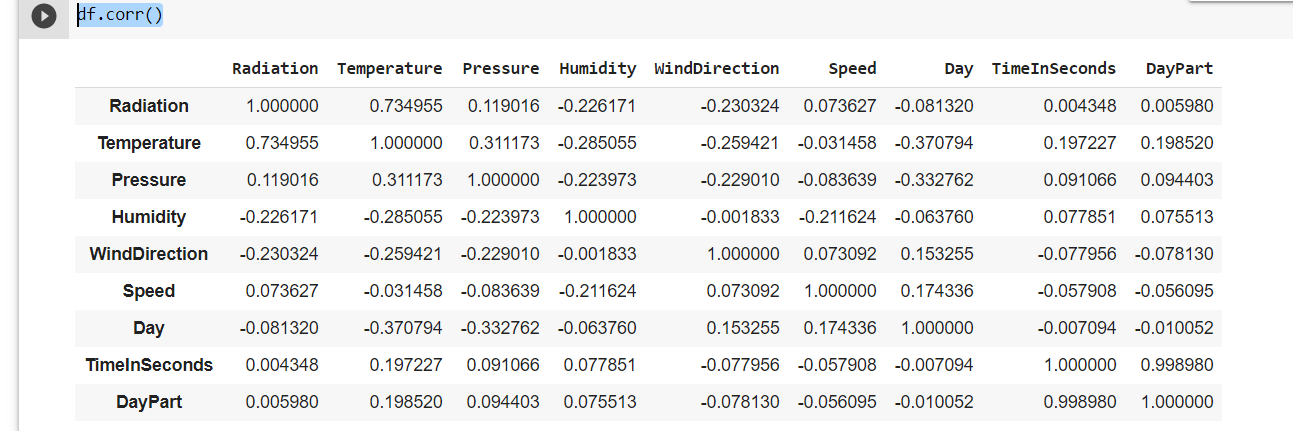




Построим парные диаграммы по всем показателям по исходному набору данных:

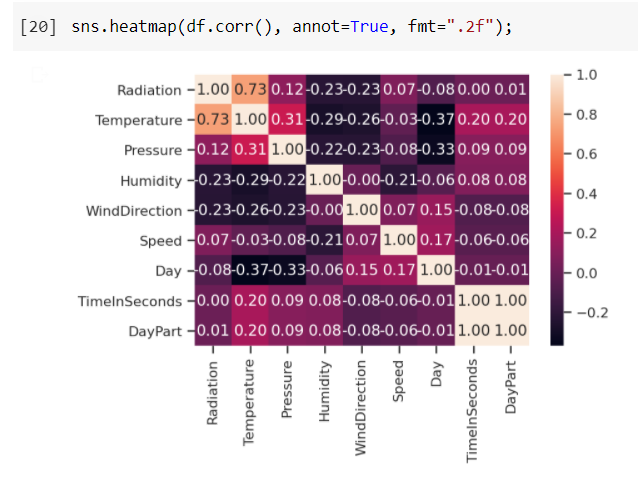


**3.4. Информация о корреляции признаков**

****

Построим корреляционную матрицу по всему набору данных:

Визуализируем корреляционную матрицу с помощью тепловой карты:



**Список литературы**

[1] Гапанюк Ю. Е. Лабораторная работа «Разведочный анализ данных. Исследование  
и визуализация данных» [Электронный ресурс] // GitHub. –– 2019. –– Режим доступа: https://github.com/ugapanyuk/ml\_course/wiki/LAB\_EDA\_VISUALIZATION (дата  
обращения: 13.02.2019)