Задание 2.

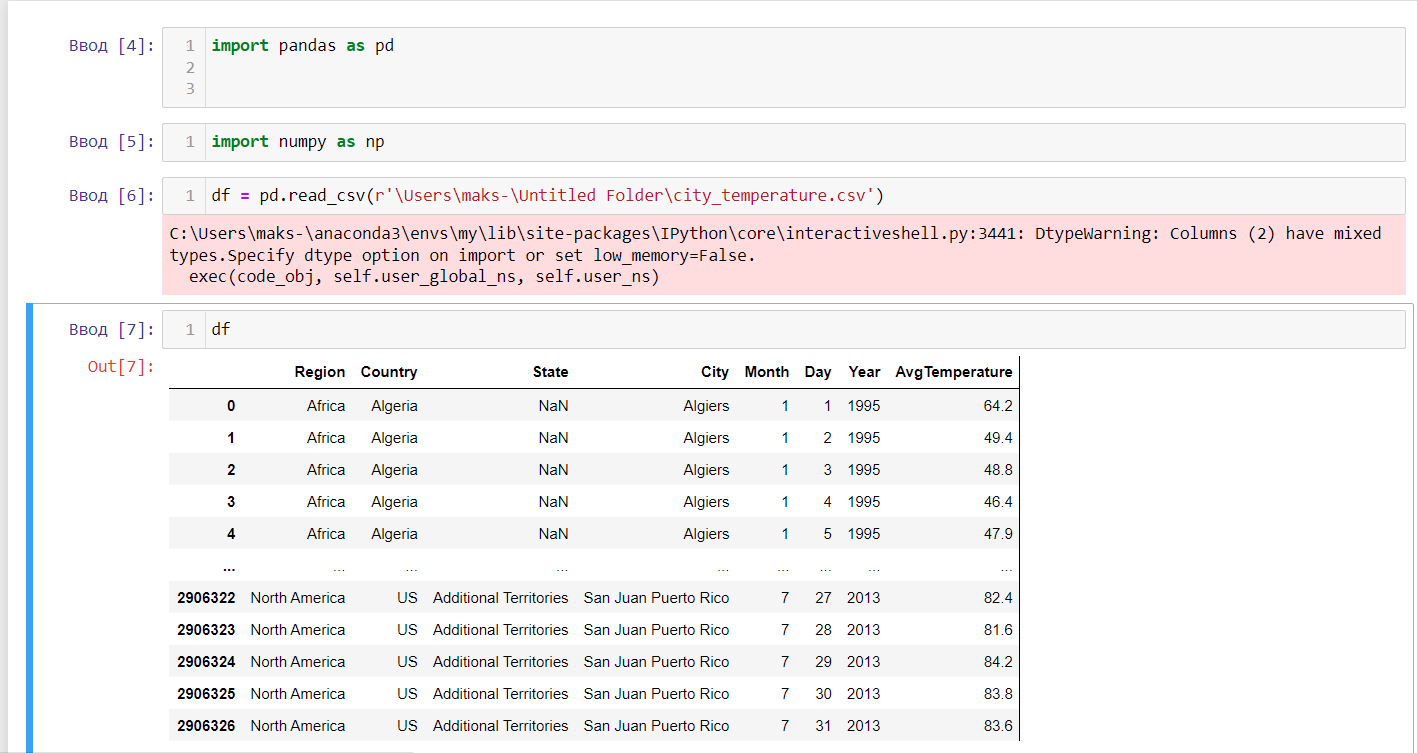
Дата сет взят с сайта:  
 <https://www.kaggle.com/sudalairajkumar/daily-temperature-of-major-cities>

Набор данных содержит информацию о температуре в крупных городах мира на протяжении нескольких лет в разных странах.

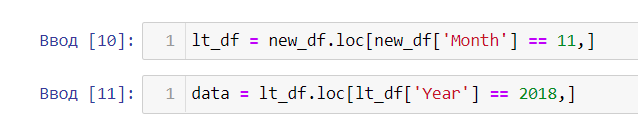
Среднесуточные значения температуры находятся в файле city\_temperature.csv.

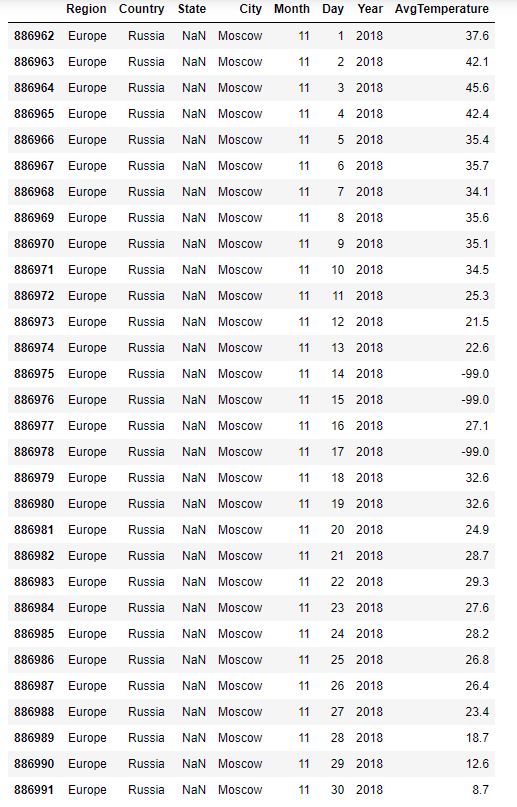
Для анализа была выбрана температура в городе Москва за ноябрь 2018 года.

Загрузка дата сета:

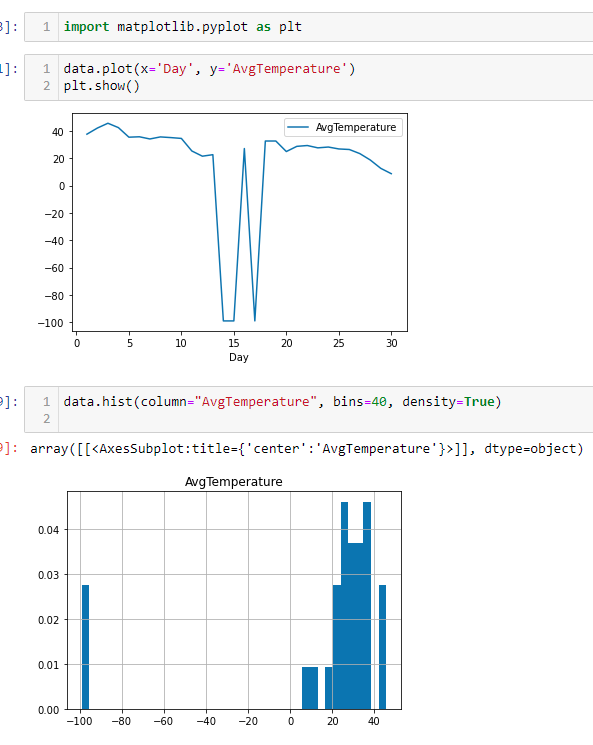


Определение выборки:





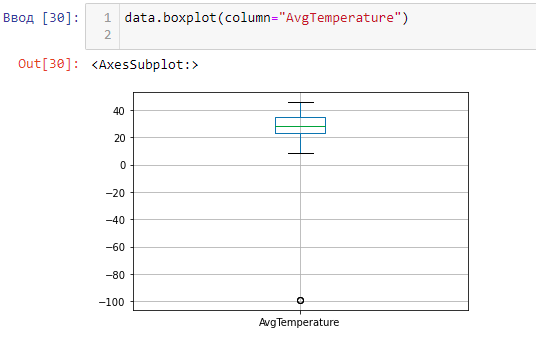
Построим график зависимости дня от температуры на протяжении ноября месяца, а также построим гистограмму распределения:



Можно заметить, что в дата сете присутствуют выбросы.

-99 градусов по фаренгейту встречается в наборе данных дважды и не является корректным и реальным значением температуры.

Построим boxplot:



После анализа данных можно сделать вывод о нормальном распределении исследуемой величины.

Рассчитав Percentile и IQR в excel подверждаем наличие выброса в наших данных

