МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ

РОССИЙСКОЙ ФЕДЕРАЦИИ

ФГАОУ ВО «СЕВЕРО-КАВКАЗСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

ИНСТИТУТ ИНФОРМАЦИОННЫХ

ТЕХНОЛОГИЙ И ТЕЛЕКОММУНИКАЦИЙ

КАФЕДРА ПРИКЛАДНОЙ ИНФОРМАТИКИ

**Лабораторная работа №2**

по дисциплине

«Искусственный интеллект и машинное обучение»

**Выполнил:**

Сердюков Никита Анатольевич

Студент 2 курса группы \_ПИН-б-о-22-1

Направления подготовки

09.03.03 Прикладная информатика

очной формы обучения

Ставрополь, 2023 г.

Тема: Визуализация данных.

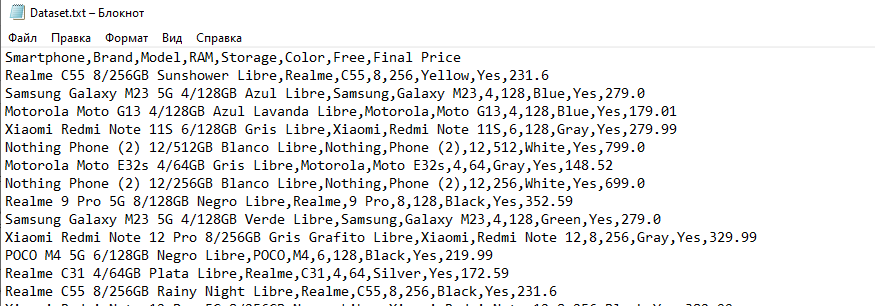
Цель работы: изучение программных средств для визуализации наборов данных.

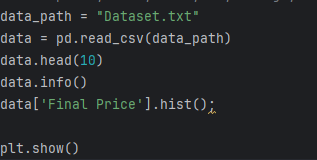
Выполнение работы:

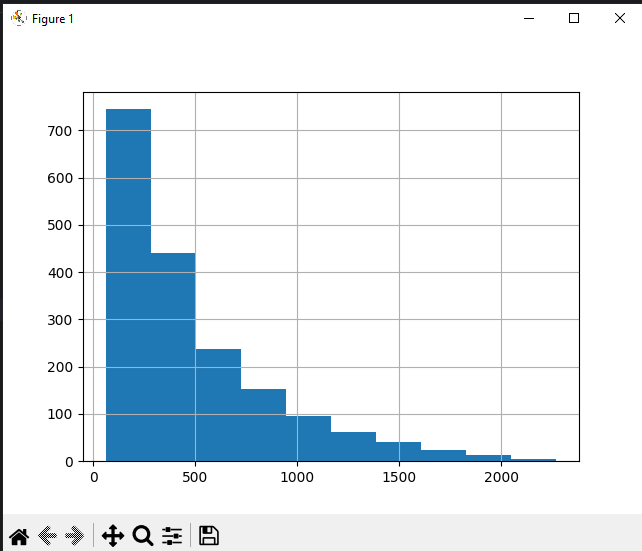
Датасет: Smartphones Price Dataset

Описание датасета: информацию о различных смартфонах, позволяя провести детальный анализ их технических характеристик и цен. Он охватывает широкий спектр смартфонов различных брендов, моделей и конфигураций. Признаки: Смартфон, Бренд, Модель, ОЗУ, Хранилище, Бесплатная сим-карта(по контракту), Финальная цена

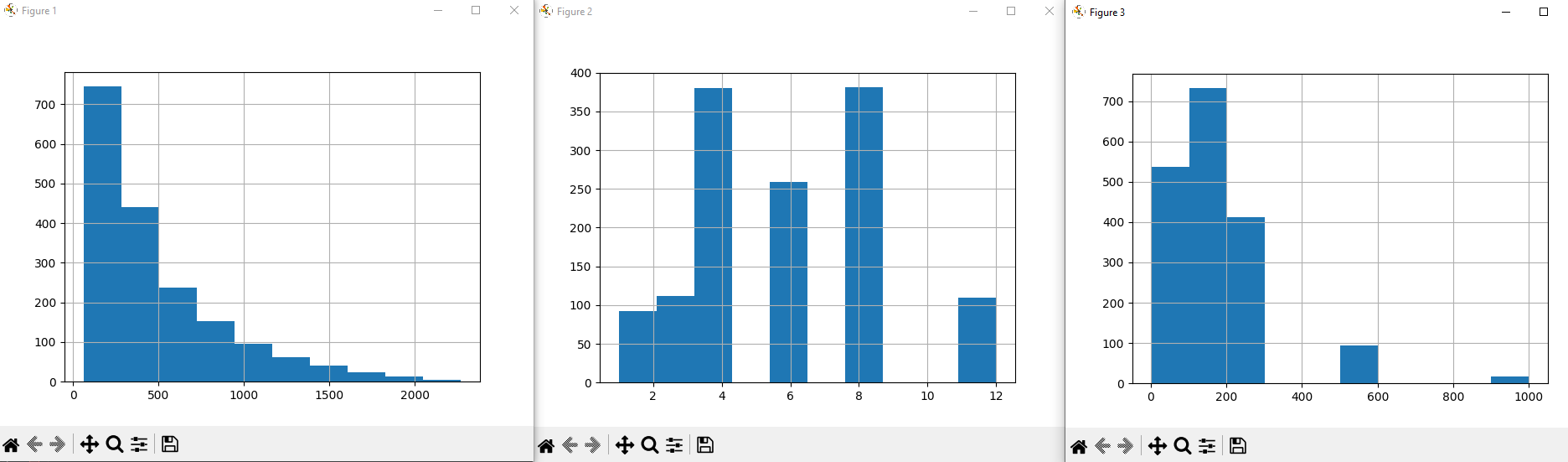
1. . Визуализация количественных признаков:





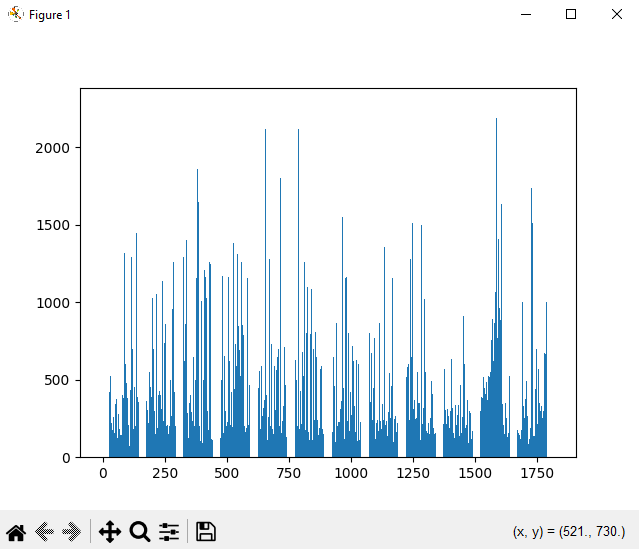


1. Применение метода hist() для визуализации распределения нескольких признаков



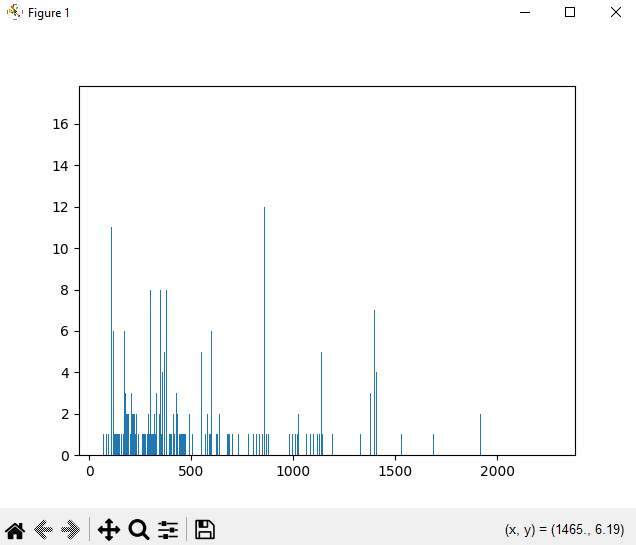
1. Построение гистограммы с использованием matplotlib





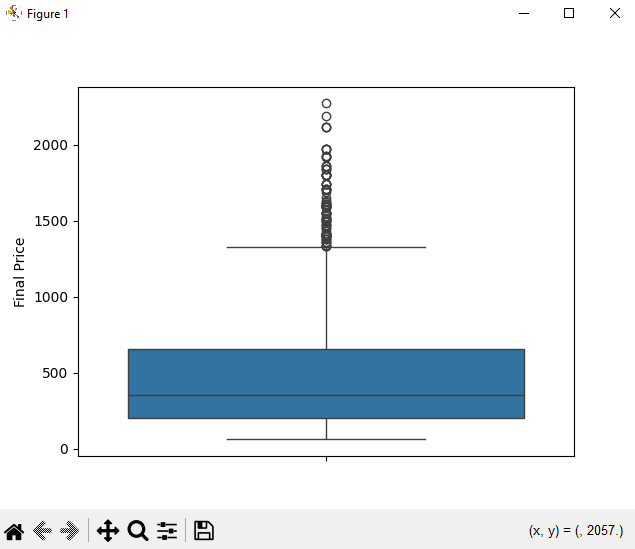
1. Использование matplotlib для представления распределения значений признака



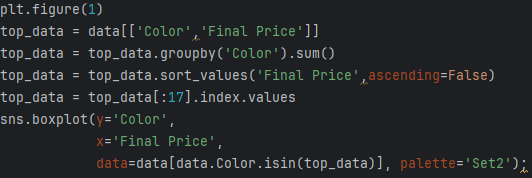


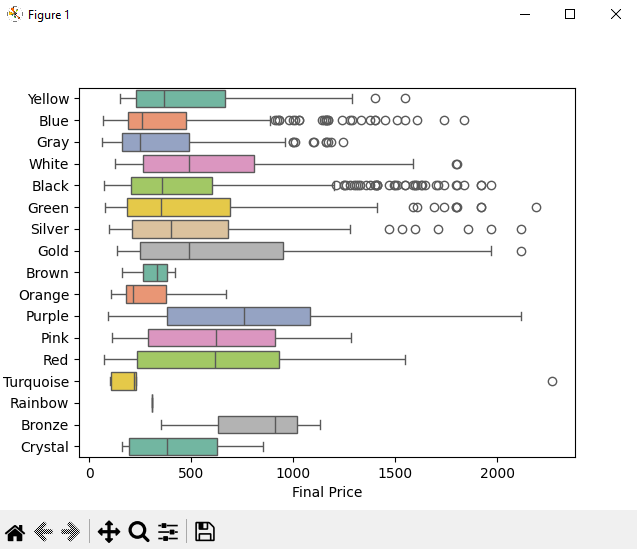
1. График «ящик с усами» для отдельного признака м визуализацию данных

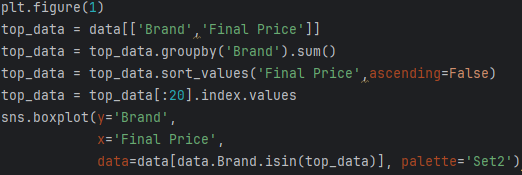


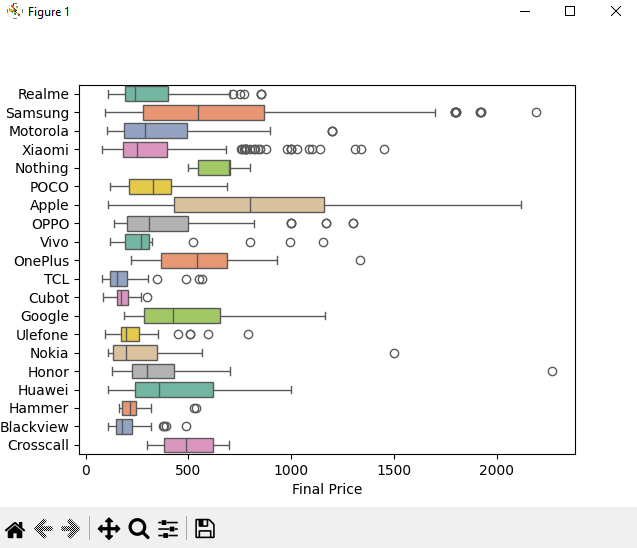


1. Использование boxplot для анализа признака цвета телефона и его бренда от цены



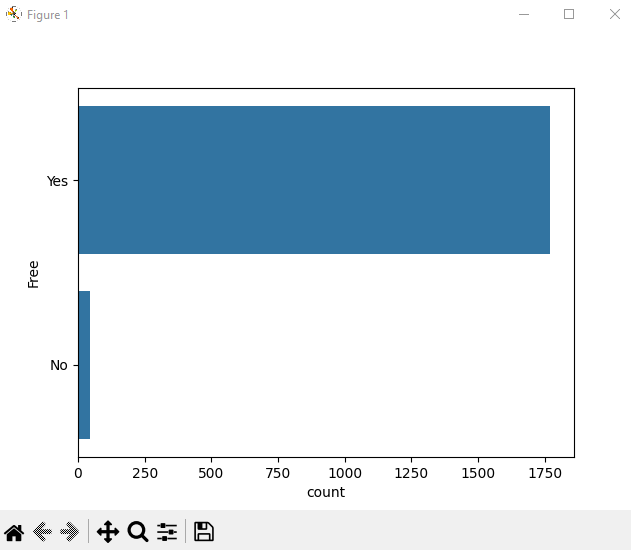




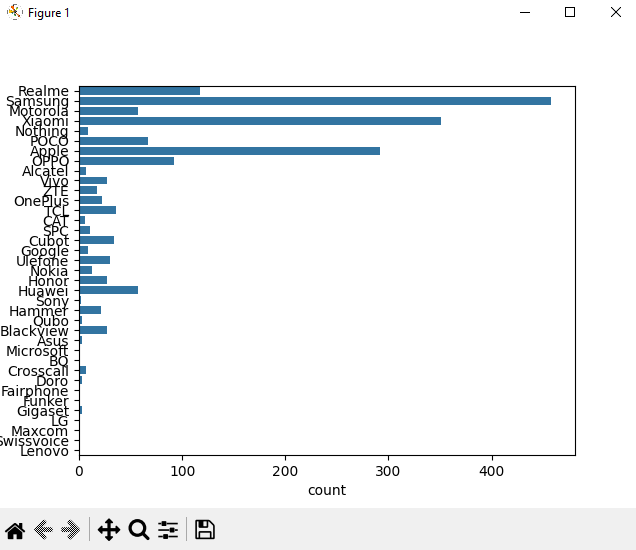


1. Сделаем график по категориальным признакам. Первый по наличию бесплатной сим-карты, второй по брендам.



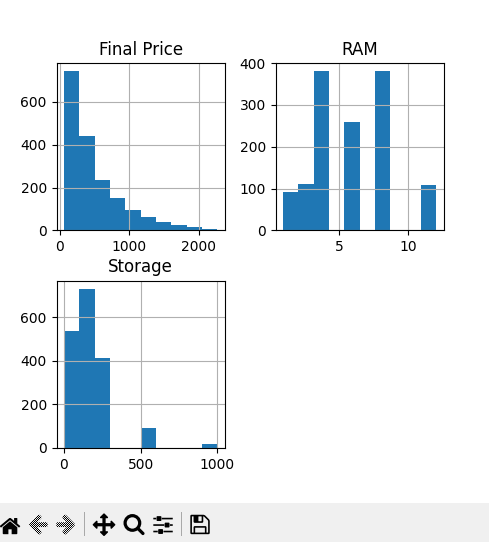




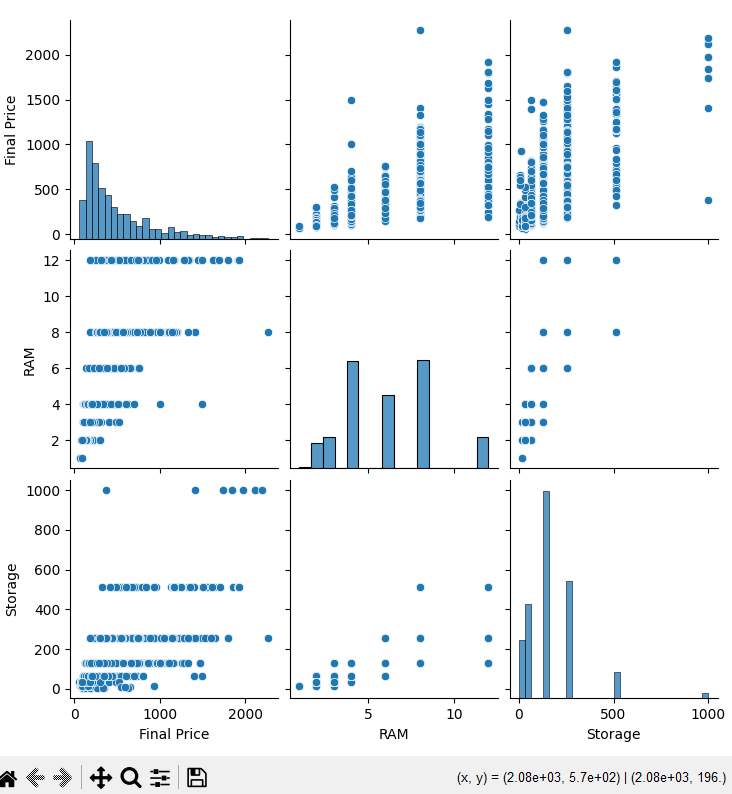


1. Создадим графики количественных признаков

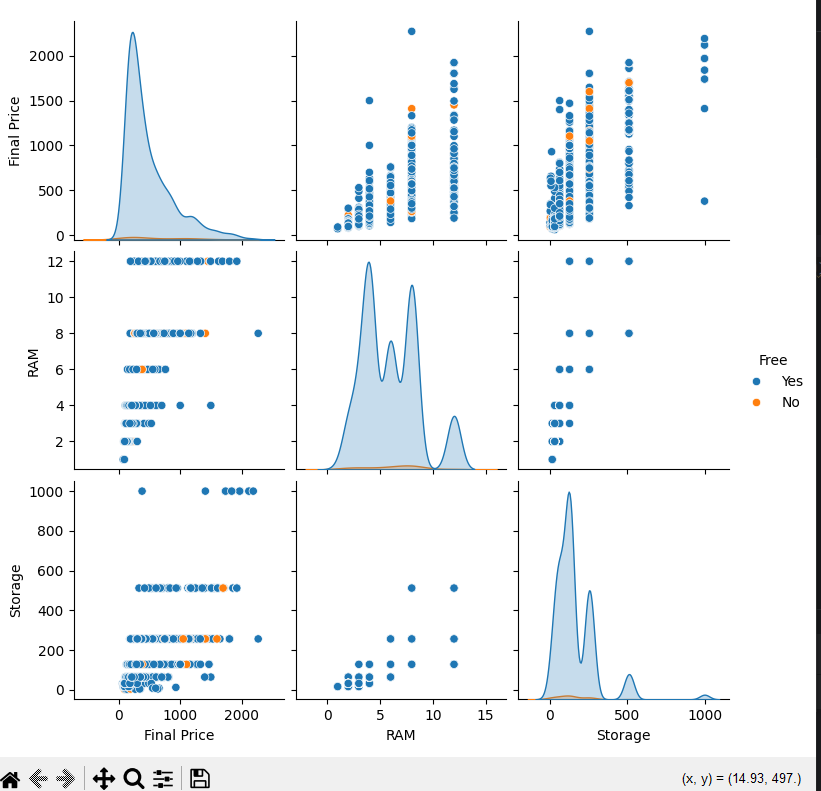




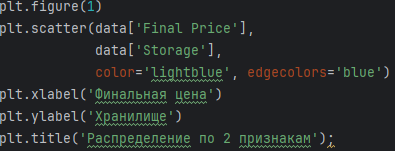


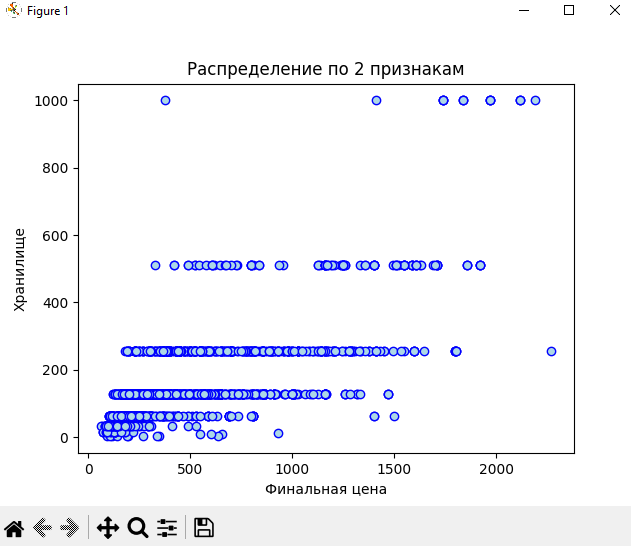




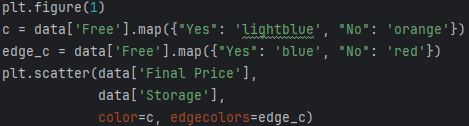


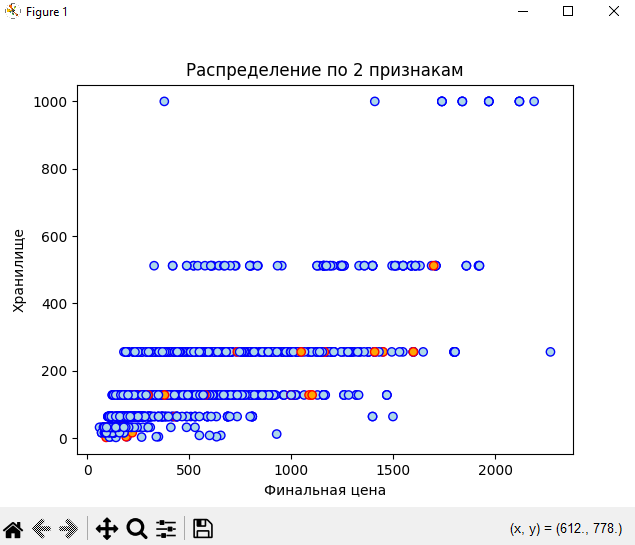
1. Создадим графики распределения по 2 признакам: Цене и объему памяти





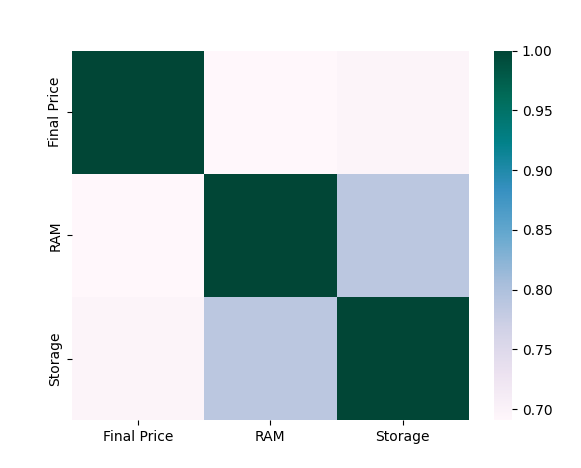
Добавим окрашивание по наличию сим-карты





1. Создадим визуальное представление корреляции количественных значений





Листинг:

import numpy as np  
import pandas as pd  
from matplotlib import pyplot as plt  
import seaborn as sns  
  
# Smartphone,Brand,Model,RAM,Storage,Color,Free,Final Price  
data\_path = "Dataset.txt"  
data = pd.read\_csv(data\_path)  
data.head(10)  
data.info()  
plt.figure(1)  
feats = ['Final Price','RAM','Storage']  
sns.heatmap(data[feats].corr(), cmap=plt.cm.PuBuGn);  
# c = data['Free'].map({"Yes": 'lightblue', "No": 'orange'})  
# edge\_c = data['Free'].map({"Yes": 'blue', "No": 'red'})  
# plt.scatter(data['Final Price'],  
# data['Storage'],  
# color=c, edgecolors=edge\_c)  
# plt.xlabel('Финальная цена')  
# plt.ylabel('Хранилище')  
# plt.title('Распределение по 2 признакам');  
#data[feats].hist(figsize=(5,5));  
#sns.pairplot(data[feats + ['Free']], hue='Free');  
#sns.countplot(data[data['Brand'].isin(data['Brand'].value\_counts().tail(40).index)]['Brand']);  
#sns.countplot(data['Free']);  
# top\_data = data[['Brand','Final Price']]D  
# top\_data = top\_data.groupby('Brand').sum()  
# top\_data = top\_data.sort\_values('Final Price',ascending=False)  
# top\_data = top\_data[:20].index.values  
# sns.boxplot(y='Brand',  
# x='Final Price',  
# data=data[data.Brand.isin(top\_data)], palette='Set2');  
# sns.boxplot(data['Final Price']);  
# hist = data['Final Price'].value\_counts()  
# plt.bar(hist.index, hist);  
# plt.bar(data.index, data['Final Price'])  
# data['Final Price'].hist();  
# plt.figure(2)  
# data['RAM'].hist();  
# plt.figure(3)  
# data['Storage'].hist();  
  
plt.show()

Вывод: изучил программные средства для визуализации наборов данных.