



Богатый домовладелец

Итоговая аттестация: Программирование на Python

Шаббир Кафи Ул, М03-3046, 07.06.2024

Московский физико-технический институт

Автор

- **Шаббир Кафи Ул**
- kafiulshabbir@phystech.edu
- 1 курс магистратуры
- Физтех-школа
Аэрокосмических
Технологий (ФАКТ)
- Московский Физико-
Технический Институт



1. Введение

Зачем нужно разбирать данную проблему сейчас?

- В последние годы рынок краткосрочной аренды жилья через платформы, такие как Airbnb, существенно вырос.
- Тем не менее, существует ограниченное понимание того, какие факторы наиболее сильно влияют на цены листингов.
- Это знание может помочь как владельцам жилья, так и арендаторам лучше ориентироваться на рынке.

1.1. airbnb



The screenshot displays the Airbnb homepage. At the top, the navigation bar includes the Airbnb logo, links for 'Stays', 'Experiences', 'Online Experiences', 'Airbnb your home', a globe icon, and a user profile icon. Below this is a search bar with four sections: 'Where' (Search destinations), 'Check in' (Add dates), 'Check out' (Add dates), and 'Who' (Add guests), followed by a red search button with a magnifying glass icon. A horizontal row of category icons follows: 'Caves', 'Icons', 'Earth homes', 'OMG!', 'Mansions', 'Windmills', and a right arrow. To the right of these icons are 'Filters' and a toggle for 'Display total before taxes'. The main content area features three travel recommendations, each with a 'Guest favorite' badge, a heart icon, a photo, and text details.

Destination	Distance	Dates	Price	Rating
Vals, Switzerland	2,203 kilometers away	Jun 13 – 18	99,004P night	★ 5.0
Ravello, Italy	2,371 kilometers away	Sep 13 – 18	19,100P night	★ 4.81
Nar, Turkey	1,914 kilometers away	Jun 15 – 20	12,345P night	★ 4.89

1.2. Актуальность

**Для владельцев жилья
знание факторов,
влияющих на цену, поможет
оптимизировать свои
предложения, а для
арендаторов — выбрать
наиболее выгодные
варианты**



1.3. Отчет проекта (06.05.2024, 18:42)

Проект по анализу данных на Python: Богатый домовладелец

Лебедев Матвей^{1, a} and Шаббир Кафи^{1, b}

¹Московский физико-технический институт, Долгопрудный, 141701

^alebedev.ma@phystech.edu M02-301я

^bkafiulshabbir@phystech.edu M03-304б

5 июня 2024 г.

Аннотация

В последние годы рынок краткосрочной аренды жилья через платформы, такие как Airbnb, существенно вырос. Тем не менее, существует ограниченное понимание того, какие факторы наиболее сильно влияют на цены листингов. Это знание может помочь как владельцам жилья, так и арендаторам лучше ориентироваться на рынке.

Содержание

1	Введение	2
1.1	Проблема	2

<https://github.com/kafiulshabbir/project-data-visual>

1.4. Цель

Исследовательский вопрос:
Какие факторы объясняют
вариации цен на листинги
Airbnb в Амстердаме в будние
дни?

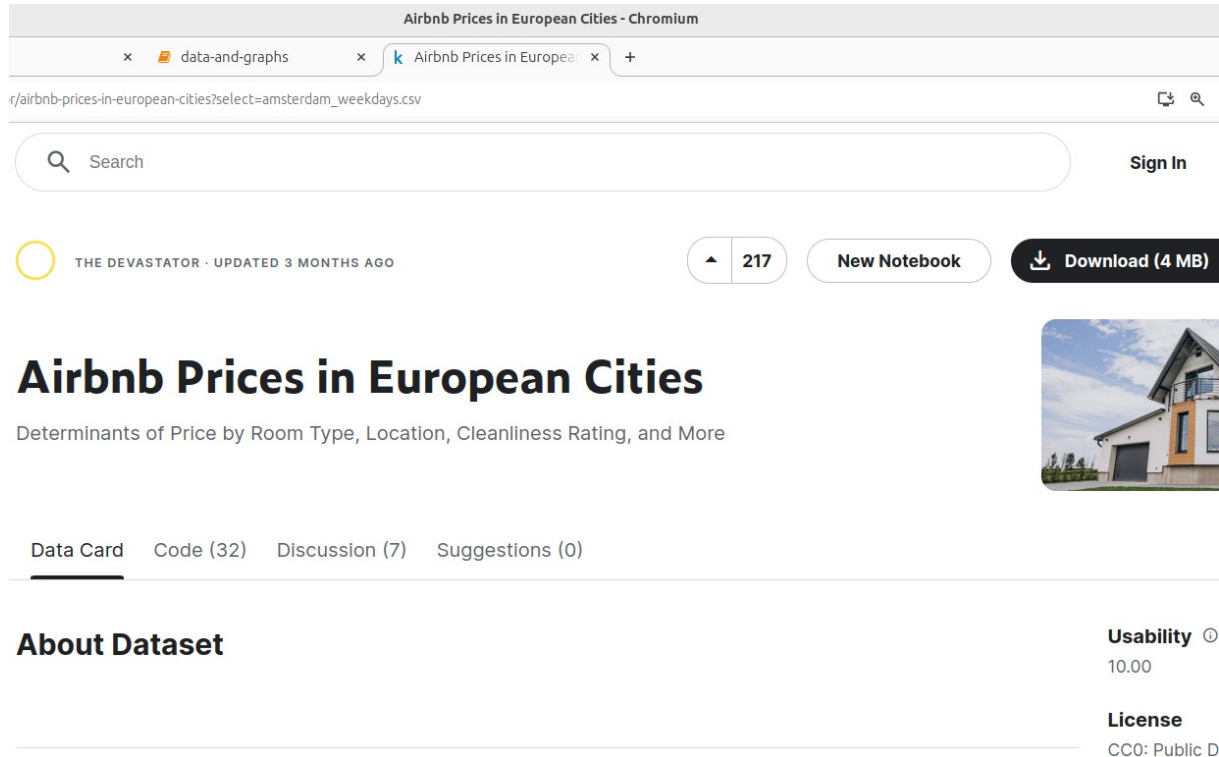


2. Методы, используемые для анализа

- **Данные собраны и предоставлены в виде CSV-файла. Источник данных - платформа Airbnb**
- Данные содержат информацию о краткосрочной аренде жилья в Амстердаме в будние дни. Каждый объект недвижимости представлен множеством характеристик, включая тип комнаты, вместимость, оценки чистоты и удовлетворенности гостей, а также различные индексы привлекательности и доступности ресторанов.
- Данные могут использоваться для анализа факторов, влияющих на цены и популярность аренды, а также для оценки качества обслуживания.

2.1. Data source: Методы сбора данных [1]

https://www.kaggle.com/datasets/thedevastator/airbnb-prices-in-european-cities?select=amsterdam_weekdays.csv



The screenshot shows a web browser window with the title "Airbnb Prices in European Cities - Chromium". The address bar shows the URL "r/airbnb-prices-in-european-cities?select=amsterdam_weekdays.csv". The page features a search bar, a "Sign In" button, and a "Download (4 MB)" button. The dataset title "Airbnb Prices in European Cities" is prominently displayed, followed by a subtitle "Determinants of Price by Room Type, Location, Cleanliness Rating, and More". Below this, there are tabs for "Data Card", "Code (32)", "Discussion (7)", and "Suggestions (0)". The "About Dataset" section is visible, and the "Usability" is rated 10.00. The "License" is listed as "CC0: Public D".

Airbnb Prices in European Cities - Chromium

data-and-graphs x Airbnb Prices in Europeal x +

r/airbnb-prices-in-european-cities?select=amsterdam_weekdays.csv

Search Sign In

THE DEVASTATOR · UPDATED 3 MONTHS AGO 217 New Notebook Download (4 MB)

Airbnb Prices in European Cities

Determinants of Price by Room Type, Location, Cleanliness Rating, and More

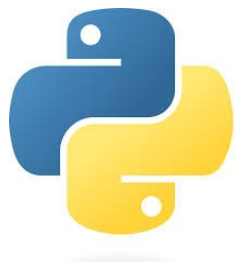
Data Card Code (32) Discussion (7) Suggestions (0)

About Dataset

Usability 10.00

License CC0: Public D

2.3. Методы, используемые для анализа



seaborn



pandas

matplotlib

Импорт библиотек

[10]:

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from scipy.stats import ttest_ind
%matplotlib inline
```

Считываем данные и преобразуем

[2]:

```
file_path = 'amsterdam_weekdays.csv'
df = pd.read_csv(file_path)
df.head()
```

[2]:

	Unnamed: 0	realSum	room_type	room_shared	room_private	person_capacity	host_is_superhost	m
0	0	194.033698	Private room	False	True	2.0	False	
1	1	344.245776	Private room	False	True	4.0	False	
2	2	264.101422	Private room	False	True	2.0	False	

2.4. Аналитические методы:

1. **Визуализация данных:** Использование гистограмм, scatter plot, box plot и heatmap для анализа распределения данных и зависимостей между переменными.
2. **Описательная статистика:** Средние значения, медианы, стандартные отклонения и другие описательные статистики для ключевых переменных.
3. **Корреляционный анализ:** Определение корреляций между переменными с помощью коэффициента корреляции Пирсона.
4. **T-тест:** Сравнение средних значений между двумя группами.

3.1. Результаты

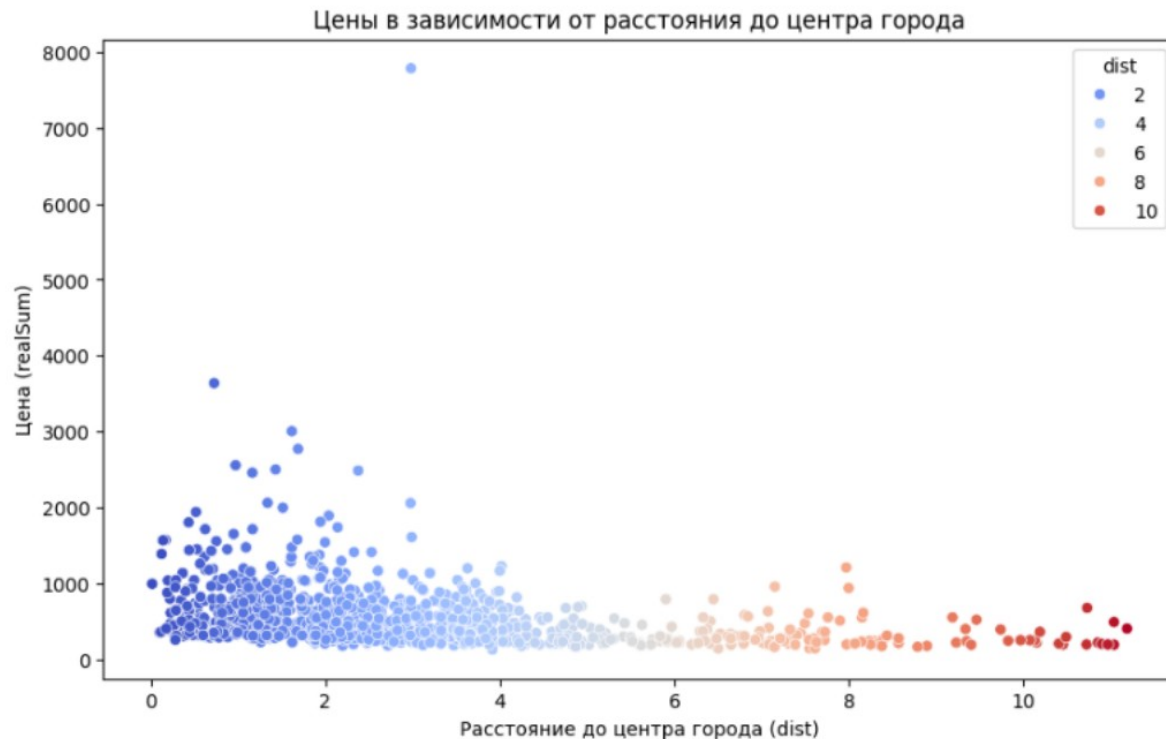
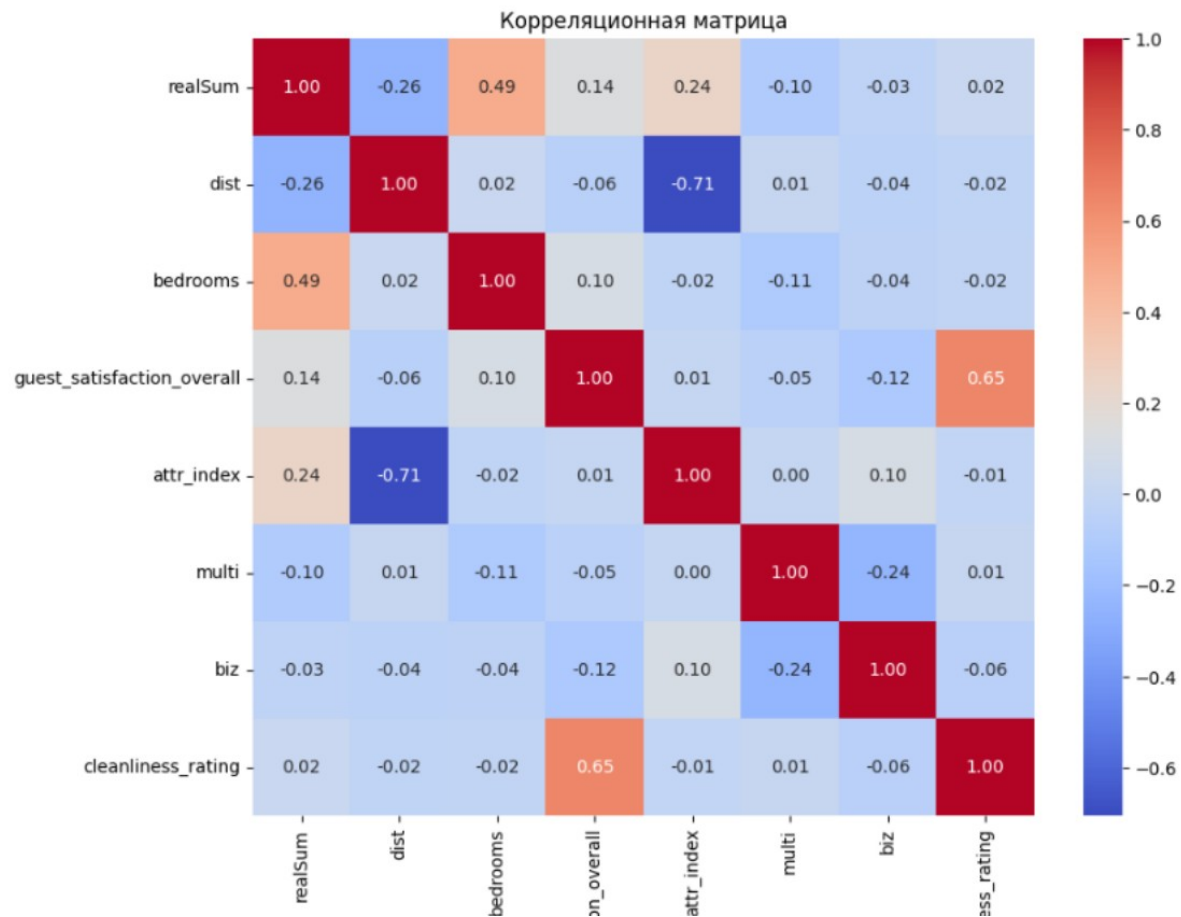
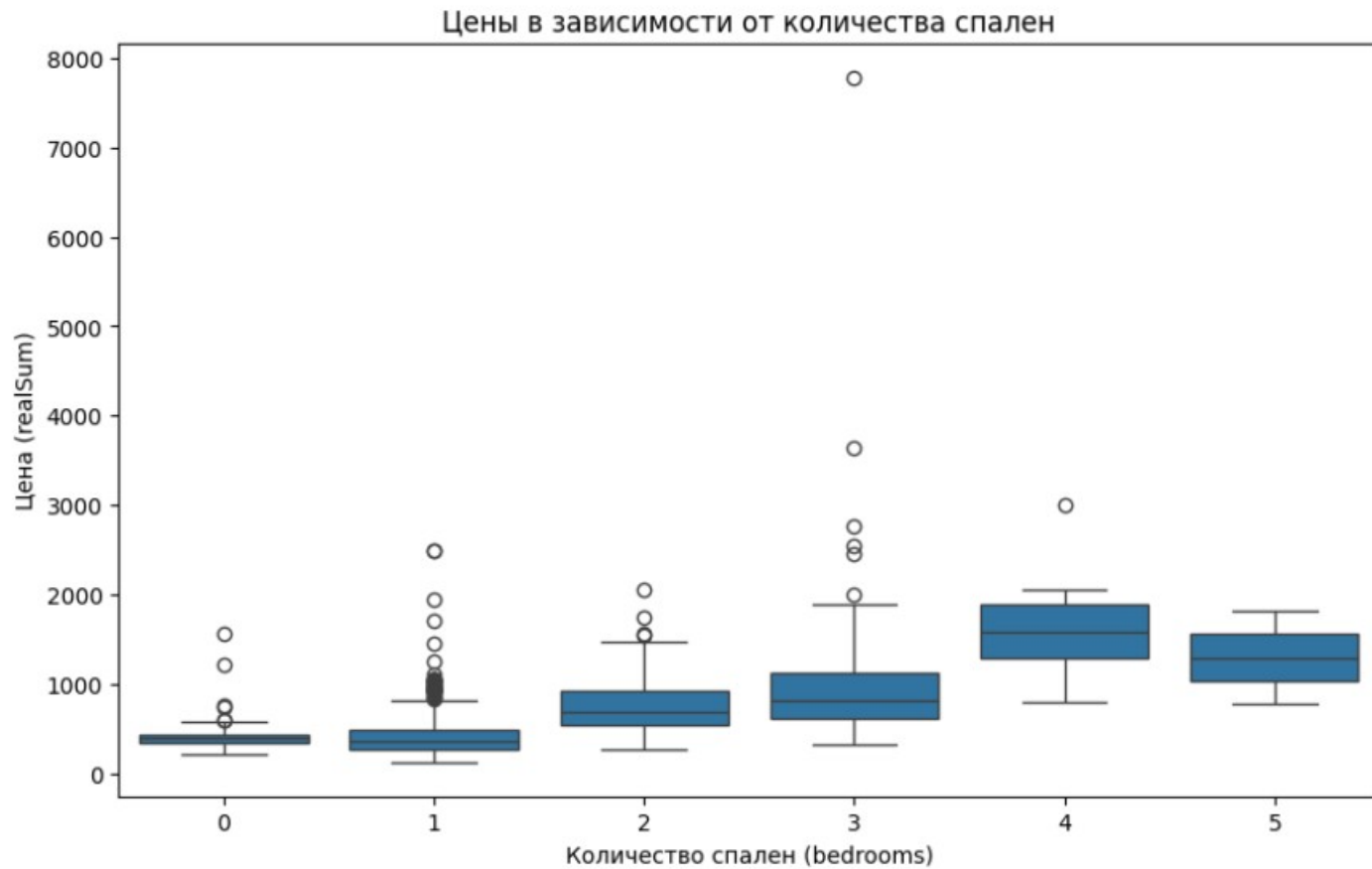


Рис. 3: для наших гипотез, чтобы подкрепить наши выводы из корреляционного анализа и описательной статистики. Листинги, расположенные ближе к центру города, имеют более высокие цены.

3.2. Результаты



3.3. Результаты



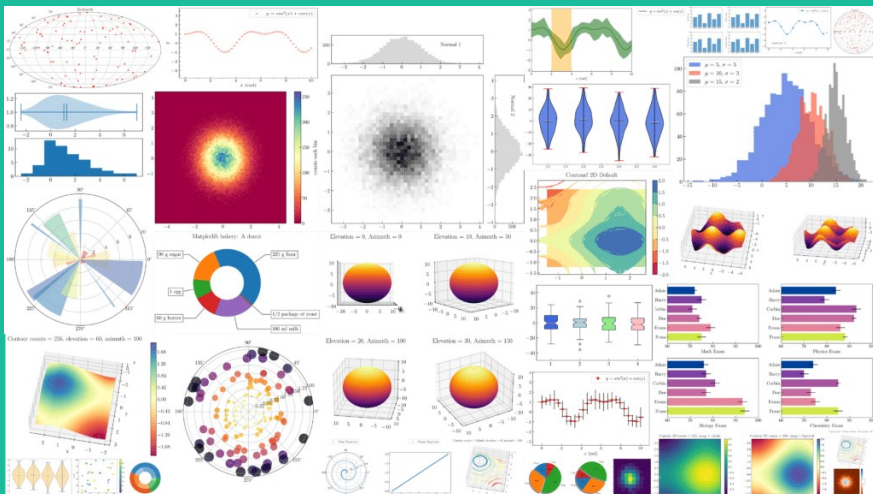
4. Выводы

На цену листинга в основном влияют следующие факторы:

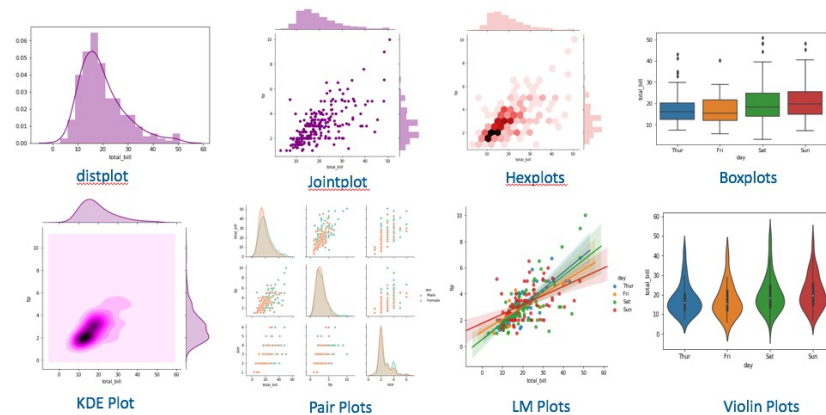
1. **Близость к центру города:** Чем ближе листинг к центру, тем выше цена.
2. **Количество спален:** Листинги с большим количеством спален имеют более высокие цены.
3. **Тип жилья:** Целые дома или квартиры стоят дороже, чем отдельные комнаты.
4. **Доступность для бронирования на несколько комнат:** Листинги, доступные для бронирования на несколько комнат, имеют более высокие цены.
5. **Рейтинг удовлетворенности гостей:** чем больше рейтинг, тем дороже листинг.

Список литературы

1. Airbnb Prices in European Cities — kaggle.com.
https://www.kaggle.com/datasets/thedevastator/airbnb-prices-in-european-cities?select=amsterdam_weekdays.csv. [Accessed 05-06-2024].
2. Israel Cohen, Yiteng Huang, Jingdong Chen, Jacob Benesty, Jacob Benesty, Jingdong Chen, Yiteng Huang, and Israel Cohen. Pearson correlation coefficient. Noise reduction in speech processing, pages 1–4, 2009.
3. Daniel Guttentag. Progress on airbnb: a literature review. *Journal of Hospitality and Tourism Technology*, 10(4):814–844, 2019.
4. Jonathan Halket, Lars Nesheim, and Florian Oswald. The housing stock, housing prices, and user costs: The roles of location, structure, and unobserved quality. *International Economic Review*, 61(4):1777–1814, 2020.
5. Tae Kyun Kim. T test as a parametric statistic. *Korean journal of anesthesiology*, 68(6):540, 2015.
6. Prabhaker Mishra, Chandra M Pandey, Uttam Singh, Anshul Gupta, Chinmoy Sahu, and Amit Keshri. Descriptive statistics and normality tests for statistical data. *Annals of cardiac anaesthesia*, 22(1):67–72, 2019.



Seaborn Plots



Спасибо!