

Bộ giáo dục và đào tạo
Trường Đại học Ngoại ngữ - Tin học TP.HCM



Đề tài

PHÂN TÍCH DỮ LIỆU THỊ TRƯỜNG BẤT ĐỘNG SẢN

GVHD: Th.s Vũ Đình Ái

HVTH: 21DH114518 Lê Đại Thắng
21DH113372 Nguyễn Tiến Thê
21DH114491 Lê Đức Phú

Tháng 3 Năm 2024

Mục lục

I. Giới thiệu.....	1
II. Nội dung.....	2
III. Phương pháp phân tích	3
IV. Kết quả phân tích	4
IV.1. Diện tích của các căn hộ có ảnh hưởng đến giá không?	4
IV.2. Số tầng của các căn hộ có ảnh hưởng đến giá không?	5
IV.3. Số phòng ngủ có ảnh hưởng đến giá hay không?	6
IV.4. Số phòng tắm có ảnh hưởng đến giá hay không?	7
IV.5. Có bao nhiêu số lượng căn hộ phân bố ở các quận khác nhau?.....	8
IV.6. Số lượng căn hộ trong các quận có giá bán cao là bao nhiêu căn?	9
IV.7. Số lượng căn hộ trong các quận có giá bán thấp nhất là bao nhiêu căn?.....	10
IV.8. So sánh giá căn hộ đã được bàn giao với chưa bàn giao?.....	11
IV.9. Các kiểu nội thất được ưa chuộng trong các căn hộ?	12
IV.10. Các thể loại nhà ở được ưa chuộng ?.....	13
IV.11. Hướng ban công có ảnh hưởng tới giá thuê không?	14
IV.12. So sánh nhà ở có giấy tờ và chưa có thì có ảnh hưởng tới giá trị của bất động sản hay không?	15
IV.13. Số lượng xuất hiện của mỗi loại căn hộ?	16
IV.14. Phân bố giá của các căn hộ?	17
IV.15. Tình trạng các loại nội thất được xuất hiện nhiều trong các loại bất động sản?.....	19
IV.16. Có bao nhiêu bất động sản chưa được bàn giao theo quận?	20
IV.17. Hướng ban công phổ biến nhất được các nhà bán bất động sản xây dựng nhiều nhất	21
IV.18. Các giấy tờ pháp lý thông dụng cho bất động sản ?	21
IV.19. Ma trận tương quan giữa căn hộ với giá của căn hộ.	23
V. Kết luận.....	24
Tài liệu tham khảo (IEEE).....	25

Danh mục hình ảnh

Hình 1. Quy trình phân tích	3
Hình 2. Biểu đồ diện tích ảnh hưởng đến giá của căn hộ.....	4
Hình 3. Biểu đồ số tầng ảnh hưởng đến giá của căn chung cư.....	5
Hình 4. Biểu đồ số phòng ngủ ảnh hưởng đến giá của căn hộ	6
Hình 5. Biểu đồ số phòng tắm ảnh hưởng đến giá của căn hộ	7
Hình 6. Biểu đồ hiển thị số lượng căn hộ phân bố ở các quận	8
Hình 7. Biểu đồ số lượng căn hộ có giá cao nhất ở các quận.....	9
Hình 8. Biểu đồ số lượng căn hộ có giá thấp nhất ở các quận	10
Hình 9. Biểu đồ so sánh giá tình trạng của BĐS	11
Hình 10. Biểu đồ nội thất được ưa chuộng.....	12
Hình 11. Biểu đồ thể loại nhà được ưa chuộng	13
Hình 12. Biểu đồ hướng ban công có ảnh hưởng đến giá không	14
Hình 13. Biểu đồ thể hiện giá qua giấy tờ	15
Hình 14. Biểu đồ số lượng của các loại căn hộ	16
Hình 15. Biểu đồ histogram thể hiện phân bố giá của căn hộ	17
Hình 16. Biểu đồ phân bố giá của căn hộ.....	18
Hình 17. Biểu đồ tình trạng nội thất của các căn hộ.....	19
Hình 18. Biểu đồ tình trạng nội thất của các căn hộ.....	19
Hình 19. Biểu đồ các căn hộ chưa được bàn giao ở các quận	20
Hình 20. Biểu đồ hướng ban công phổ biến nhất.....	21
Hình 21. Biểu đồ các giấy tờ pháp lý thông dụng	22
Hình 22. Biểu đồ giá trị trung bình của các loại căn hộ	23
Hình 23. Ma trận tương quan giữa căn hộ với giá của căn hộ	24

I. Giới thiệu

Phần trình bày sẽ giới thiệu về bộ dữ liệu phân tích.

Dự án phân tích dữ liệu bất động sản của chúng em nhằm mục đích khám phá và hiểu sâu hơn về xu hướng, giá cả, nhu cầu và cung ứng, các yếu tố ảnh hưởng và tiềm năng đầu tư trong lĩnh vực bất động sản.

Trong quá trình nghiên cứu, chúng em đã áp dụng các công cụ phân tích dữ liệu tiên tiến như Pandas và Numpy cho việc xử lý dữ liệu cùng với Matplotlib và Seaborn để trực quan hóa dữ liệu một cách sinh động và dễ hiểu. Môi trường làm việc chính được chọn là Google Colab.

Quá trình thực hiện dự án không chỉ giúp nhóm chúng em nâng cao kỹ năng phân tích dữ liệu, mà còn phát triển khả năng suy luận và tư duy giải quyết vấn đề thông qua việc tương tác với dữ liệu. Thông qua việc đặt ra và trả lời các câu hỏi dựa trên dữ liệu, chúng em đã học được cách tiếp cận và phân tích dữ liệu một cách có hệ thống, đồng thời áp dụng những hiểu biết sâu sắc về bất động sản.

Đề tài này sẽ tiến hành phân tích các số liệu bao gồm giá cả, diện tích, đặc điểm và các yếu tố có liên quan đến bất động sản để có những cái nhìn sâu sắc hơn về giá trị thị trường bất động sản đưa ra những định hướng, quyết định mua bán, đầu tư hoặc cho thuê bất động sản trong khu vực này.

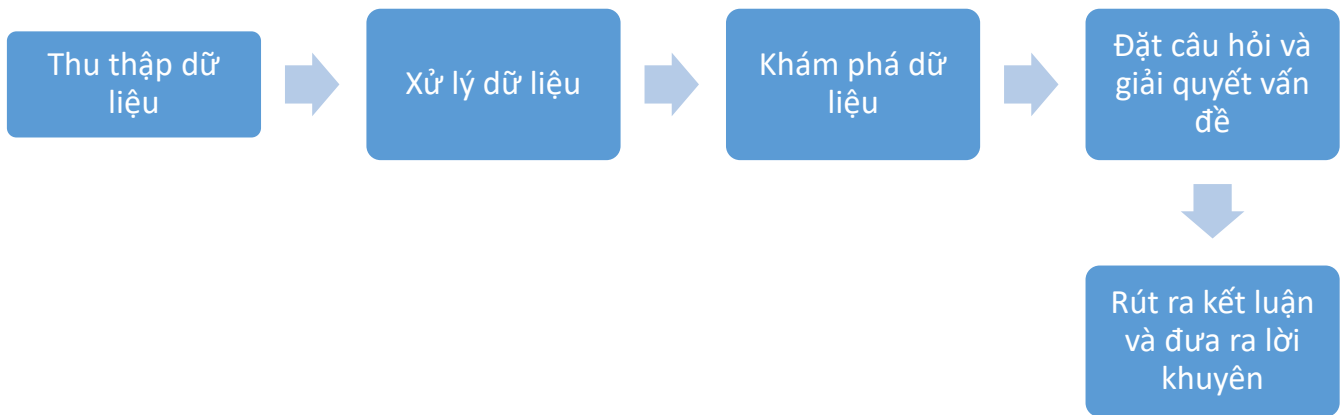
II. Nội dung

Phần trình bày chi tiết các thông tin của bộ dữ liệu

- Tên bộ dữ liệu: Bất động sản thành phố Hồ Chí Minh(2020).
- Nguồn dữ liệu: Github
- Cấu trúc gồm: 15 cột là chi tiết của 20 câu hỏi, 24949 dòng dữ liệu, chi tiết:

Câu hỏi số (Số cột)	Cột tương ứng trên dataset	Nội dung
0	DiaChi	Địa chỉ của chung cư.
1	TinhTrangBDS	Là chung cư này đã bàn giao chưa, hay vẫn còn đang trong quá trình xây dựng.
2	DienTich	Diện tích thực ở(sử dụng) trên sổ hồng, đơn vị: triệu/m2.
3	Gia/m2	Giá bán của căn hộ.
4	PhongNgu	Số lượng phòng ngủ.
5	TenPhanKhu	Căn hộ đó thuộc block nào trong khu chung cư ấy. Vì 1 khu chung cư có rất nhiều block, các block ở vị trí khác nhau sẽ có giá khác nhau.
6	SoTang	Căn hộ nằm ở tầng thứ mấy.
7	PhongTam	Số lượng nhà vệ sinh.
8	Loai	Chung cư hay nhà ở xã hội.
9	GiaiyTo	Giấy tờ pháp lý của căn hộ, có đang tranh chấp hay không, có hợp pháp hay không.
10	MaCanHo	Mã căn hộ (giống như số nhà).
11	TinhTrangNoiThat	Căn hộ đã có nội thất hay chưa(sofa, lò vi sóng, máy lạnh,...).
12	HuongCuaChinh	Hướng cửa chính của căn hộ.
13	HuongBanCong	Hướng ban công của căn hộ.
14	DacDiem	Đặc điểm căn hộ (căn trong góc, hay căn chính giữa,...).
15	Gia	Giá bán của căn hộ.

III. Phương pháp phân tích

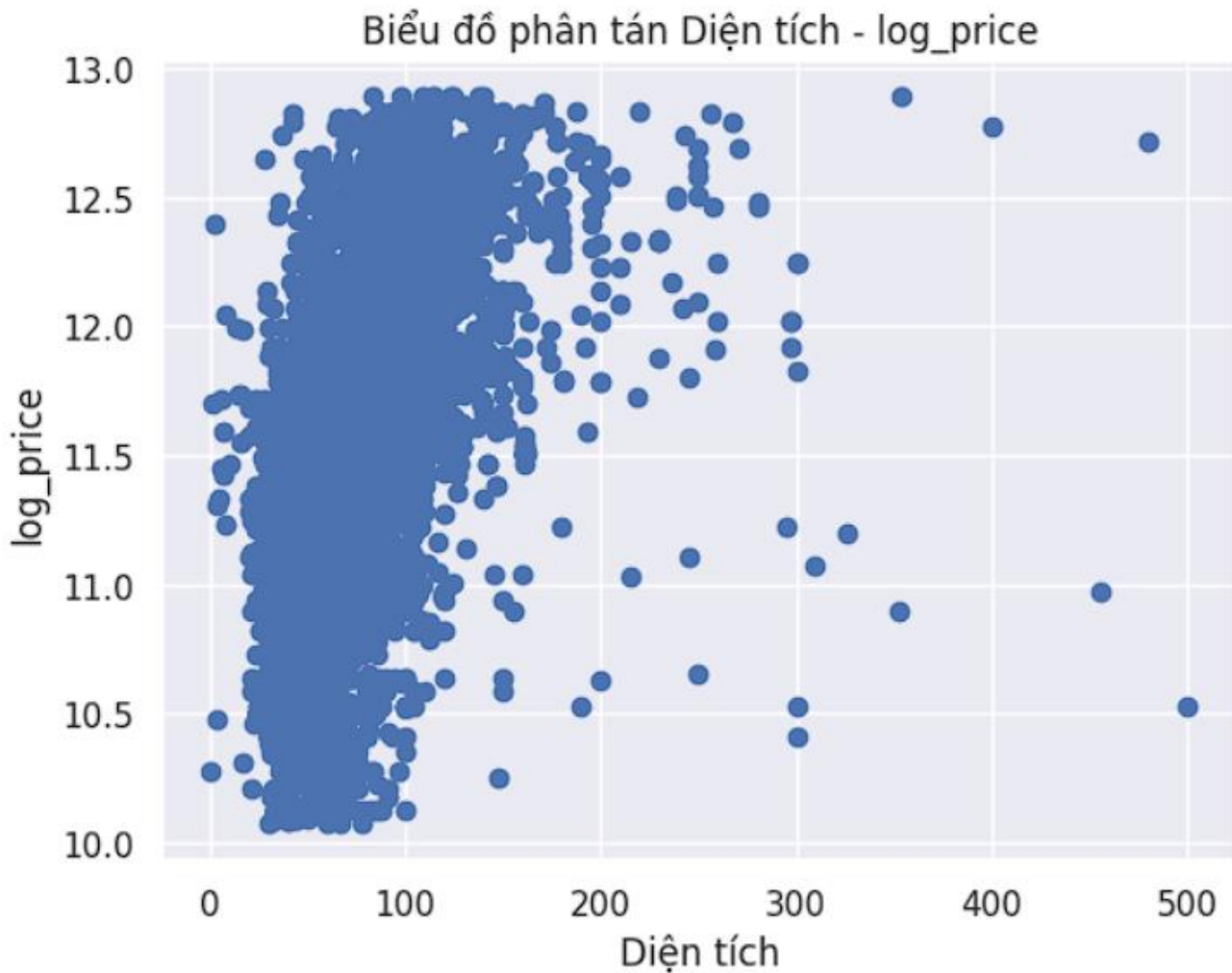


Hình 1. Quy trình phân tích

Mô tả: Nhóm bắt đầu với việc chọn bộ dữ liệu, kiểm tra các thông tin cơ bản của bộ dữ liệu để đặt ra những vấn đề cơ bản và chung nhất, sau đó tiến hành xử lý dữ liệu bằng cách lấy những dữ liệu liên quan đến diện tích, phòng ngủ, phòng tắm, số tầng trong tập dữ liệu và xóa đi các trường dữ liệu như ‘Gia/m2’ bởi vì trường dữ liệu này có thông tin tương đương nhau như vậy. Tiếp đến nhóm sẽ từ thuộc tính ‘DiaChi’ tạo thêm một cột dữ liệu mới đó là ‘Quan’. Tiếp theo trong quá trình vẽ biểu đồ thì nhận các dữ liệu có ‘Sotang > 81’ và ‘Dientich > 500’ thì biểu đồ không thể hiện hết được các thông số và khó có thể đọc được biểu đồ nên nhóm quyết định loại bỏ các dòng dữ liệu đó đi. Kế tiếp nhóm tạo thêm 2 cột gọi là ‘USD’, ‘log_price’ tại vì tiền ở Việt Nam quá nhiều số nên sẽ đổi ra ‘USD’ và ‘log_price’ là giá trị logarit của USD để giảm độ lớn của giá trị và làm cho phân phối dữ liệu chuẩn hơn giúp cho việc đọc các biểu đồ thuận tiện hơn. Cuối cùng nhóm sẽ loại bỏ giá trị ngoại biên (outlier) để không ảnh hưởng đến việc tính toán và hình dạng phân phối của biểu đồ.

IV. Kết quả phân tích

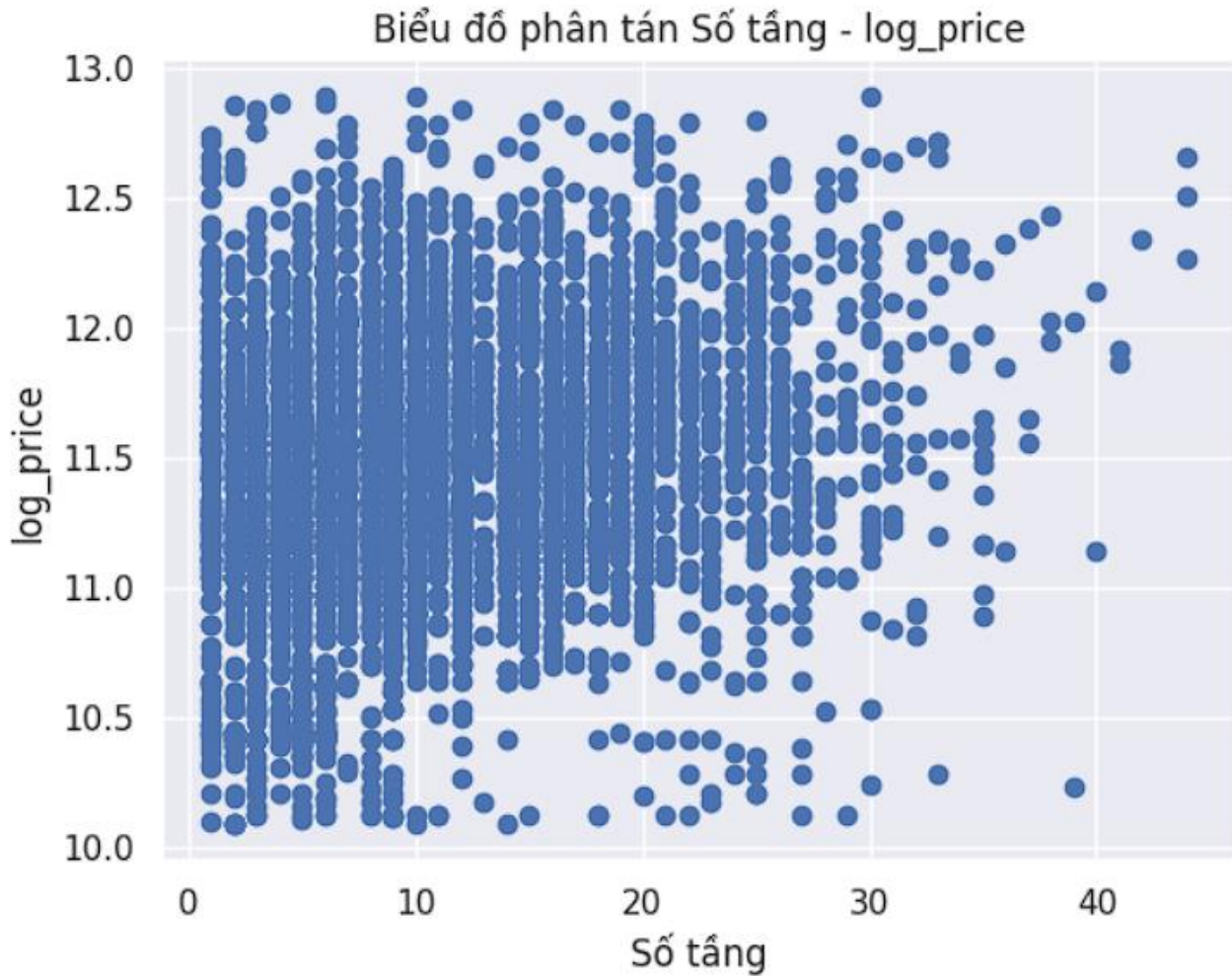
IV.1. Diện tích của các căn hộ có ảnh hưởng đến giá không?



Hình 2. Biểu đồ diện tích ảnh hưởng đến giá của căn hộ

Dựa trên nhận xét từ biểu đồ phân tán Diện tích - log_price, ta có thể kết luận rằng diện tích căn hộ có ảnh hưởng đến giá chung cư. Biểu đồ cho thấy một mối quan hệ tương đối tuyến tính giữa diện tích và log_price, với diện tích lớn hơn thường đi kèm với log_price cao hơn. => Diện tích càng rộng thì giá càng tăng. Có 1 số trường hợp diện tích cao mà giá thấp thì chúng ta phải xem xét rằng căn chung cư đó ở vị trí ở đâu.

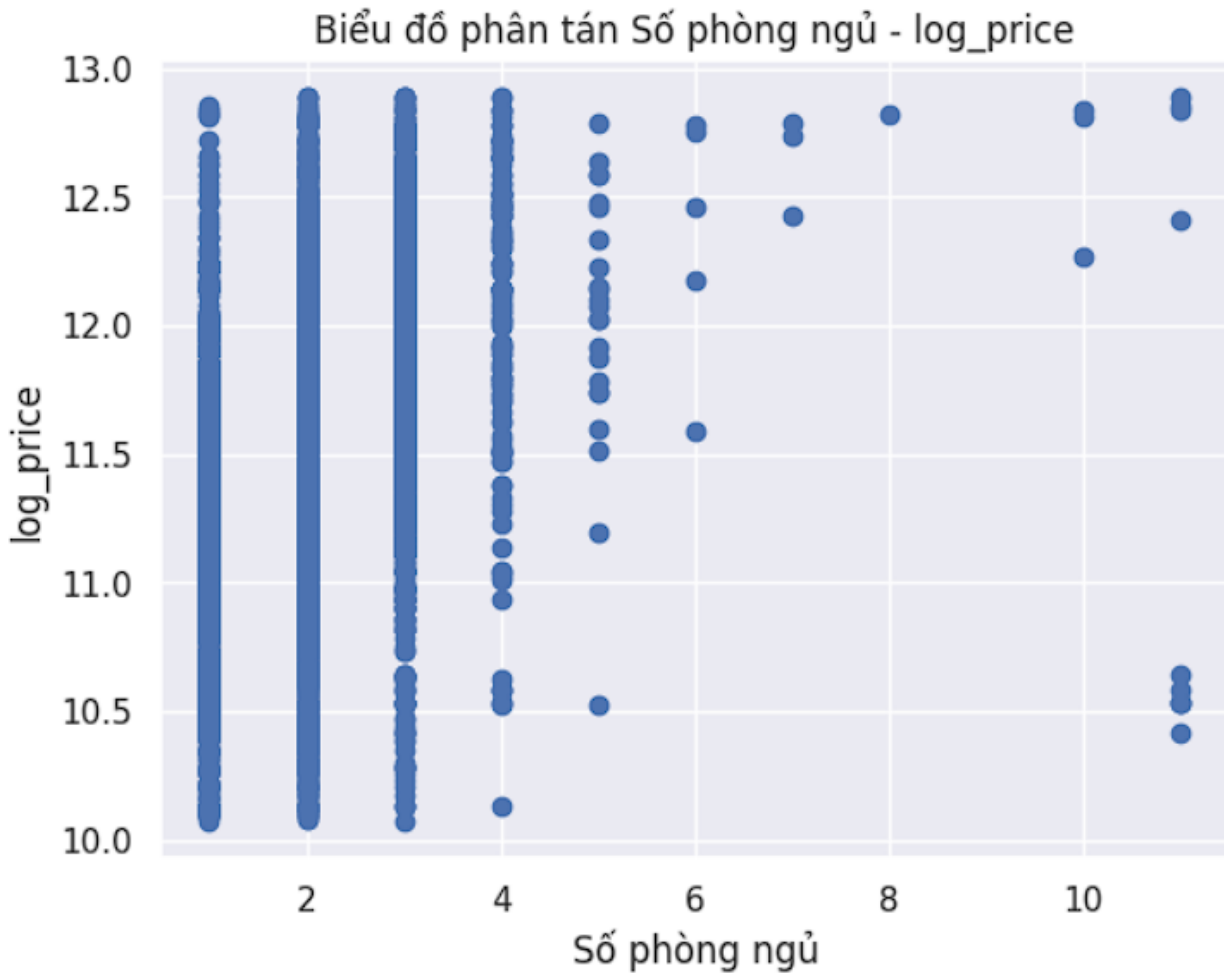
IV.2. Số tầng của các căn hộ có ảnh hưởng đến giá không?



Hình 3. Biểu đồ số tầng ảnh hưởng đến giá của căn chung cư

Dựa trên biểu đồ phân tán Số tầng - log_price, ta có thể nhận thấy các điểm trên biểu đồ phân tán không theo một hình dạng hoặc mô hình đặc trưng nào. Điều này cho thấy rằng số tầng không đơn giản là yếu tố quyết định duy nhất cho giá chung cư và có thể có nhiều yếu tố khác có ảnh hưởng đến giá.

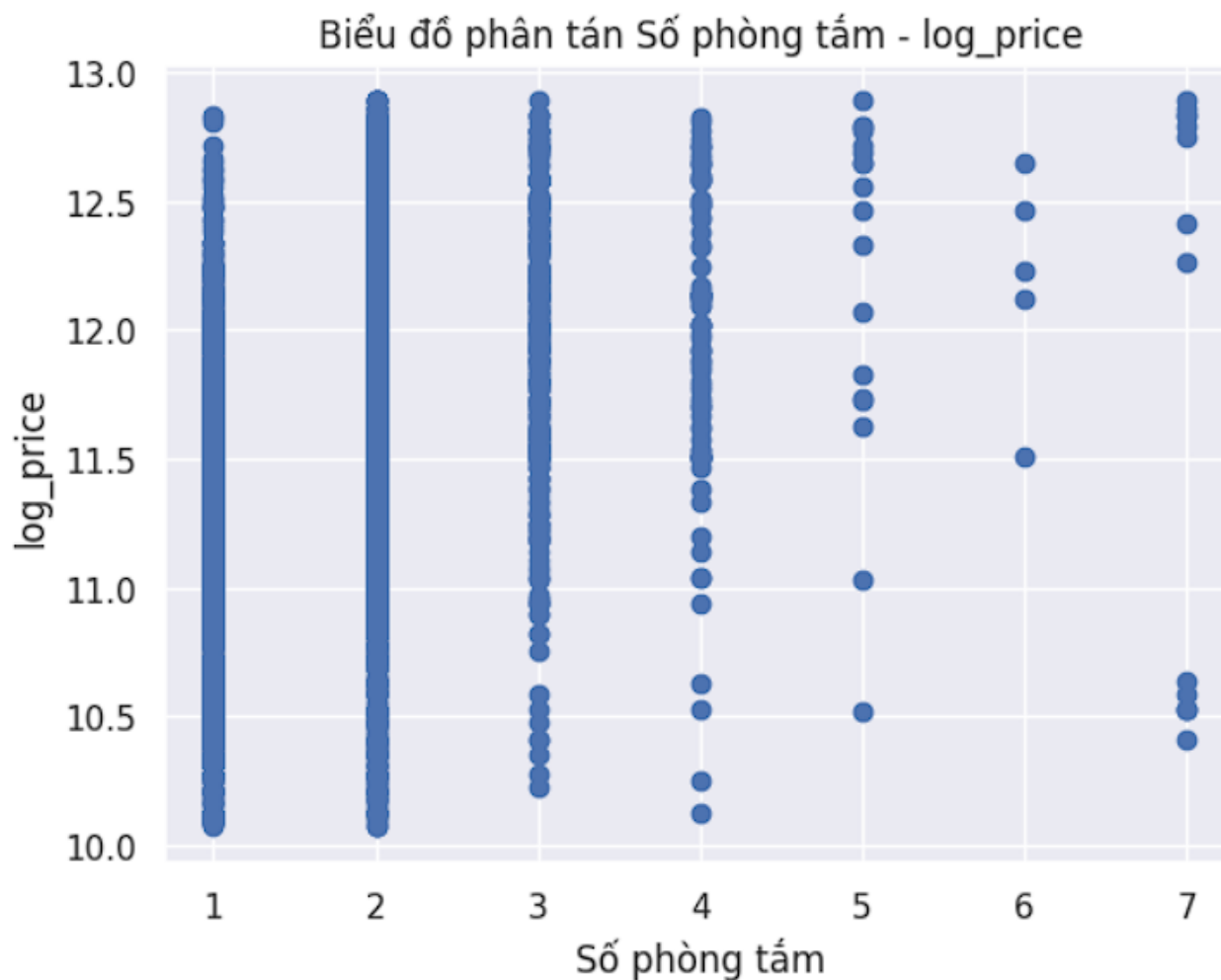
IV.3. Số phòng ngủ có ảnh hưởng đến giá hay không?



Hình 4. Biểu đồ số phòng ngủ ảnh hưởng đến giá của căn hộ

Dựa trên biểu đồ phân tán Số phòng ngủ - log_price, ta có thể nhận thấy một mối quan hệ tương đối tuyến tính giữa số phòng ngủ và log của giá chung cư. Các điểm dữ liệu trên biểu đồ phân tán có xu hướng tăng dần khi số phòng ngủ tăng, cho thấy số phòng ngủ lớn hơn thường đi kèm với log_price cao hơn.

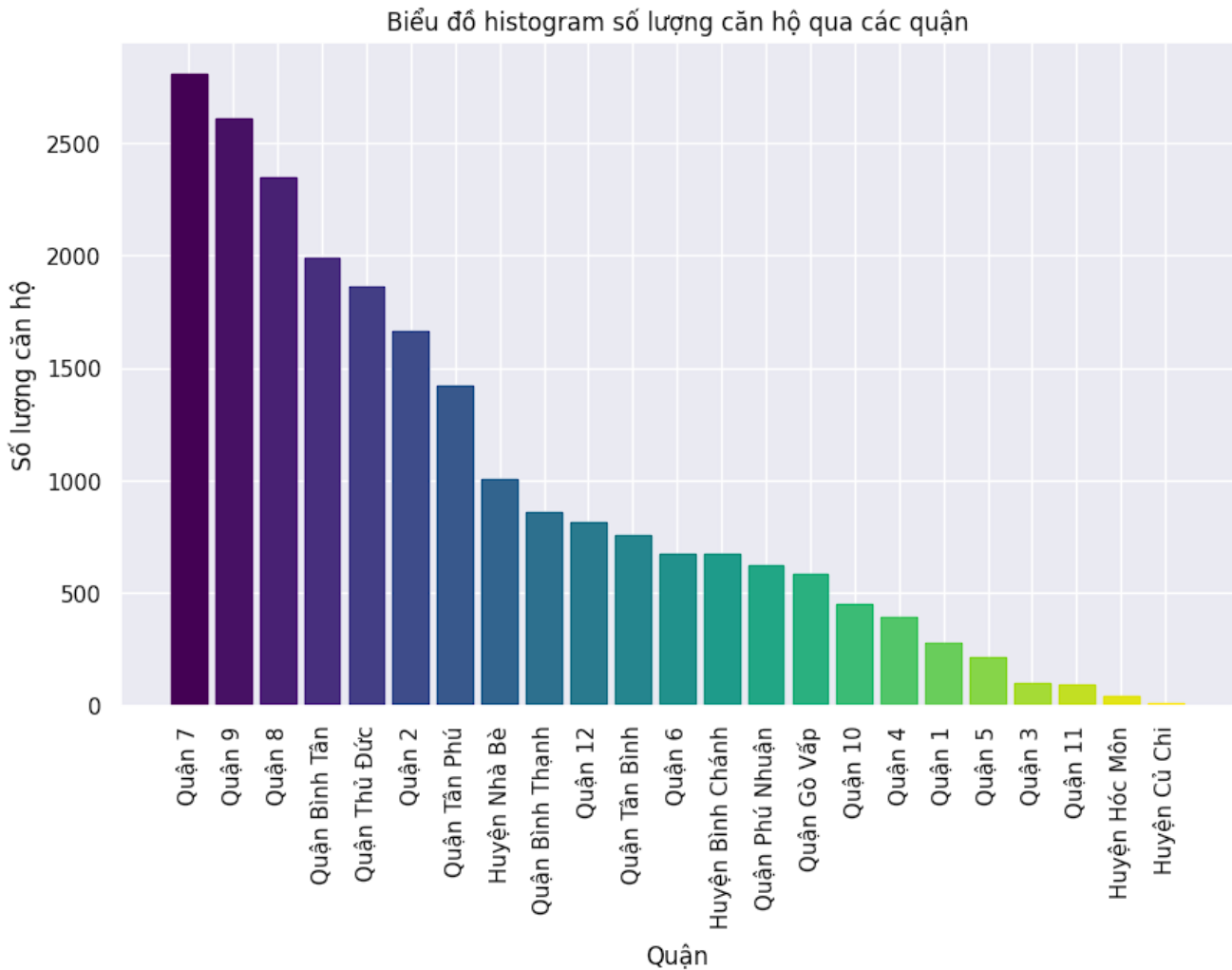
IV.4. Số phòng tắm có ảnh hưởng đến giá hay không?



Hình 5. Biểu đồ số phòng tắm ảnh hưởng đến giá của căn hộ

Dựa trên biểu đồ phân tán Số phòng tắm - log_price, ta có thể nhận các điểm dữ liệu trên biểu đồ phân tán có xu hướng tăng dần khi số phòng tắm tăng, cho thấy số phòng tắm lớn hơn thường đi kèm với log_price cao hơn.

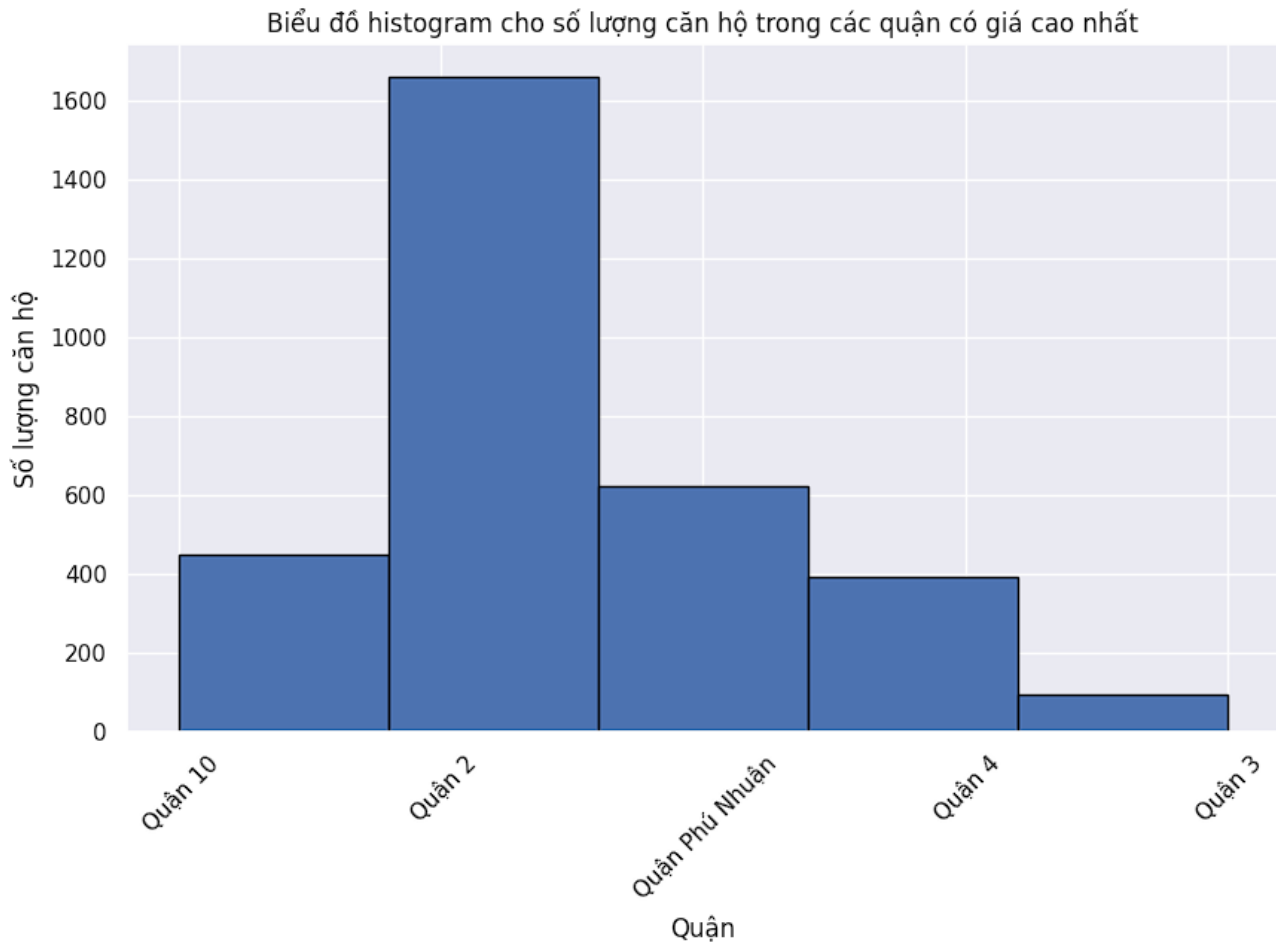
IV.5. Có bao nhiêu số lượng căn hộ phân bố ở các quận khác nhau?



Hình 6. Biểu đồ hiển thị số lượng căn hộ phân bố ở các quận

Dựa vào biểu đồ histogram trên cho ta thấy ‘Quận 7’, ‘Quận 9’, ‘Quận 8’ là nơi được rao bán căn hộ trên trang chợ tốt nhiều nhất. Và huyện Hóc Môn, huyện Củ Chi là nơi có số căn hộ ít được rao bán trên thị trường chợ tốt.

IV.6. Số lượng căn hộ trong các quận có giá bán cao là bao nhiêu căn?



Hình 7. Biểu đồ số lượng căn hộ có giá cao nhất ở các quận

Dựa trên biểu đồ trên ta có thể thấy quận 2 là nơi có số lượng căn hộ nhiều nhất ở TPHCM, tiếp đến lần lượt là quận Phú Nhuận, quận 10, quận 4, quận 3

1. Phân bố số lượng căn hộ:

Biểu đồ thể hiện số lượng căn hộ có giá cao nhất tập trung chủ yếu ở Quận 2 với hơn 1600 căn hộ.

Quận 4, Quận 10 và Quận Phú Nhuận có số lượng căn hộ xếp sau với khoảng 400 và 600 căn hộ tương ứng.

Quận 3 có số lượng căn hộ thấp nhất, chỉ khoảng gần 200 căn hộ.

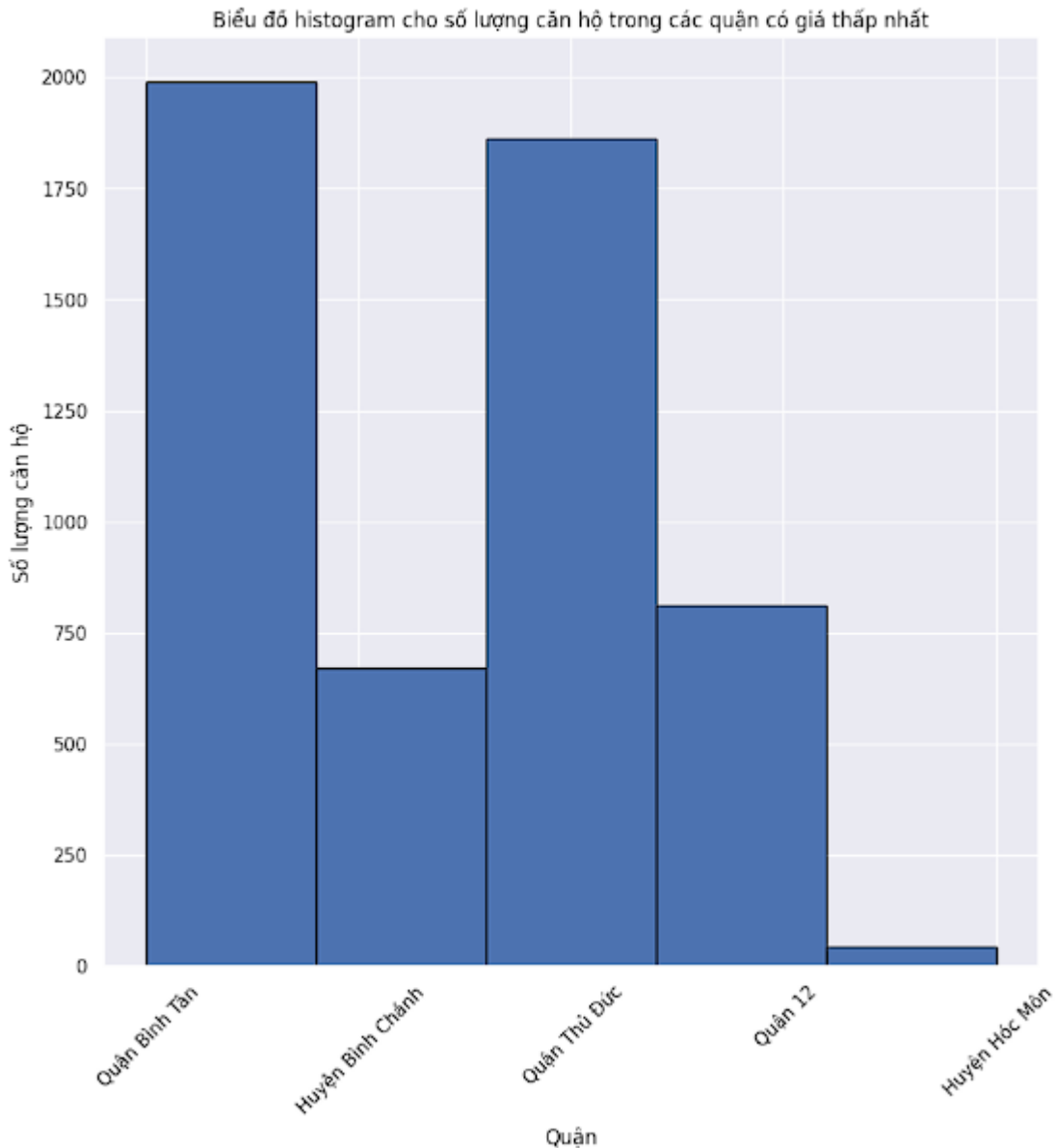
3. Nhận xét chung:

Sự chênh lệch số lượng căn hộ giữa các quận: Biểu đồ cho thấy sự chênh lệch khá lớn về số lượng căn hộ có giá cao nhất giữa các quận.

Quận 10 là khu vực có nhiều căn hộ có giá cao nhất.

Quận 3 là khu vực có số lượng căn hộ có giá thấp nhất.

IV.7. Số lượng căn hộ trong các quận có giá bán thấp nhất là bao nhiêu căn?



Hình 8. Biểu đồ số lượng căn hộ có giá thấp nhất ở các quận

Dựa trên biểu đồ trên ta có thể thấy số lượng căn hộ có giá thấp đa phần tập trung ở các huyện Bình Chánh, Quận 12, Huyện Hóc Môn. Đặc biệt là quận Bình Tân và Thủ Đức có gần 2000 căn hộ giá cao nhất.

Biểu đồ thể hiện số lượng căn hộ có giá bán thấp nhất tại các quận ở Thành phố Hồ Chí Minh:

Quận Bình Tân có số lượng căn hộ giá thấp nhất với hơn 2000 căn.

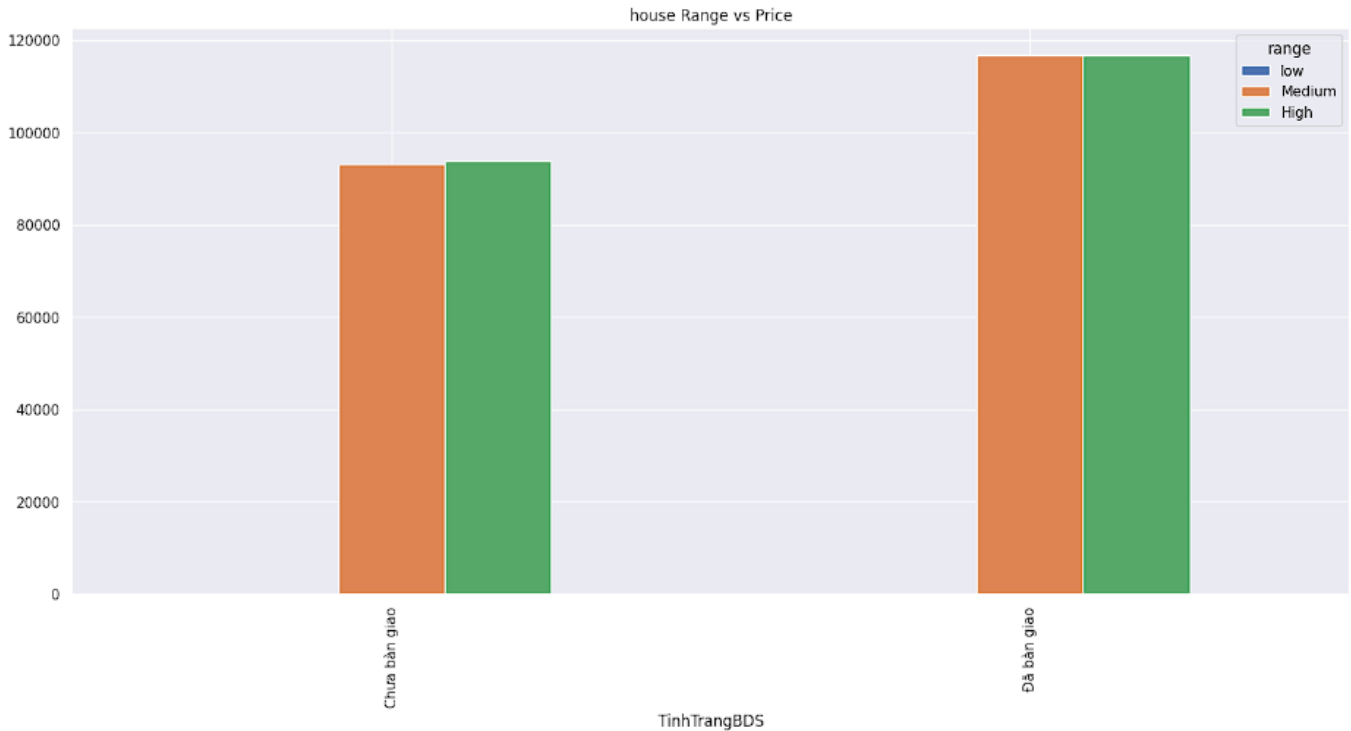
Quận Thủ Đức có số lượng căn hộ giá thấp thứ hai với hơn 1750 căn.

Quận 12 có số lượng căn hộ giá thấp thứ ba với hơn 800 căn.

Quận Bình Chánh có số lượng căn hộ giá thấp thứ tư với khoảng 650 căn.

Huyện Hóc Môn có số lượng căn hộ thấp nhất chỉ khoảng hơn 100 căn.

IV.8. So sánh giá căn hộ đã được bàn giao với chưa bàn giao?



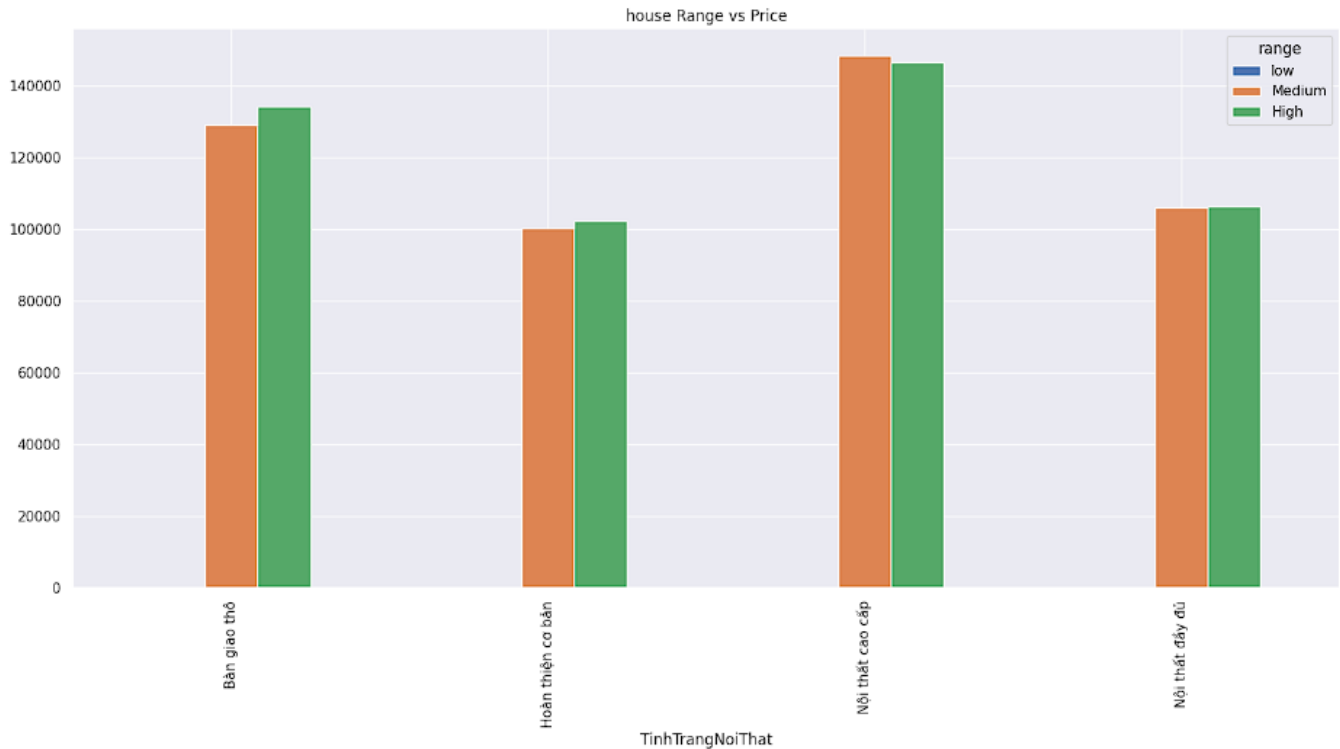
Hình 9. Biểu đồ so sánh giá tình trạng của BDS

Dựa trên biểu đồ trên ta có thể thấy những căn chung cư đã được bàn giao có giá cao hơn so với những căn chưa được bàn giao.

Giá căn hộ chưa bàn giao có xu hướng tăng ở tất cả các phân khúc.

Giá căn hộ đã bàn giao có xu hướng giảm ở phân khúc thấp và trung bình, nhưng tăng ở phân khúc cao cấp.

IV.9. Các kiểu nội thất được ưa chuộng trong các căn hộ?



Hình 10. Biểu đồ nội thất được ưa chuộng

Dựa trên biểu đồ trên ta có thể thấy nội thất đã được bàn giao thô và nội thất cao cấp được ưa chuộng hơn trong những khu căn hộ có giá cao.

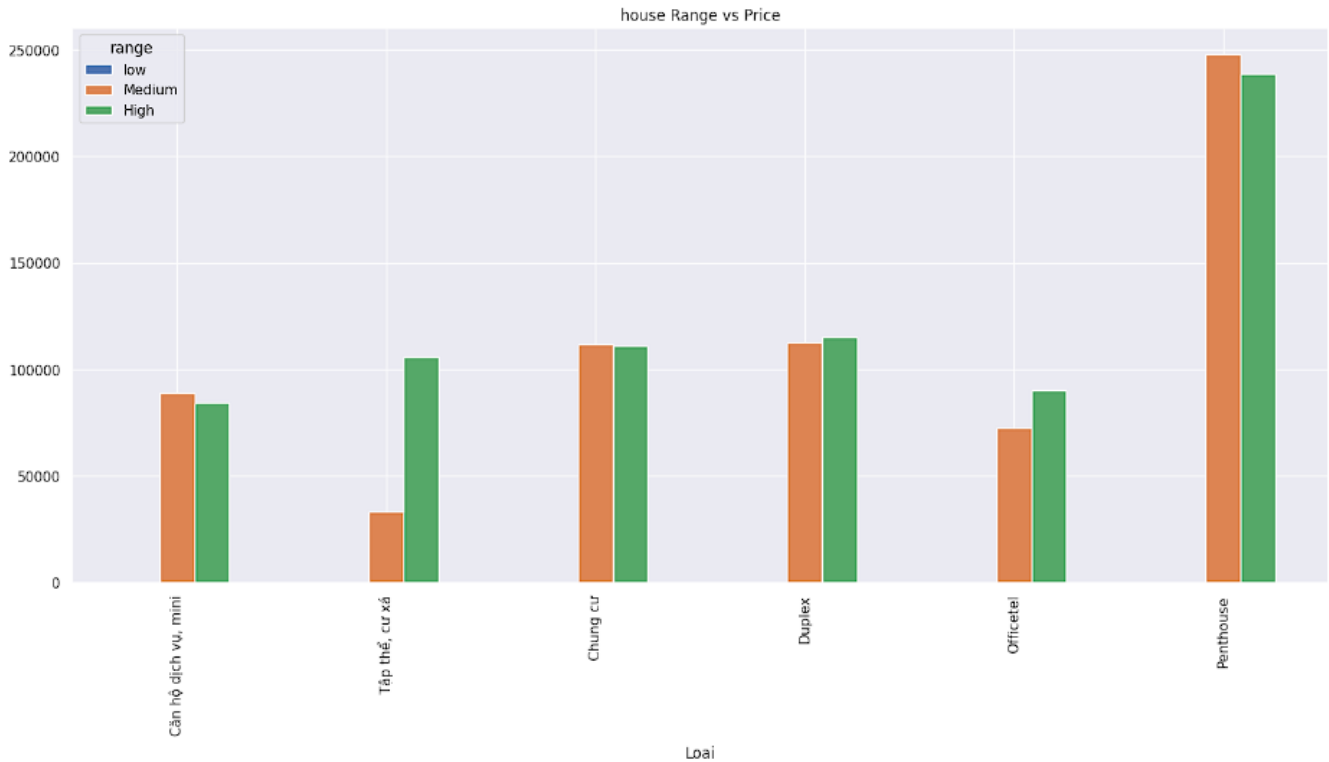
Nội thất cao cấp: Đây là kiểu nội thất được ưa chuộng nhất trong các căn hộ, với mức độ ưa chuộng cao ở hai phân khúc trung bình và cao.

Nội thất đầy đủ: Đây là kiểu nội thất phổ biến nhất trong các căn hộ, với mức độ ưa chuộng cao ở cả hai phân khúc giá nhà (trung bình và cao).

Hoàn thiện cơ bản: Kiểu nội thất này cũng khá phổ biến, đặc biệt là ở phân khúc nhà giá trung bình và cao.

Bàn giao thô: Đây là kiểu nội thất ít phổ biến nhất, chỉ được ưa chuộng ở phân khúc nhà giá trung bình và cao.

IV.10. Các thể loại nhà ở được ưa chuộng ?



Hình 11. Biểu đồ thể loại nhà được ưa chuộng

Dựa trên biểu đồ trên ta có thể thấy được những căn có giá trung bình đến cao thường là Penthouse.

Penthouse: Loại nhà được ưa chuộng nhất, với mức độ ưa chuộng cao nhất trong tất cả các phân khúc giá (cao, trung bình).

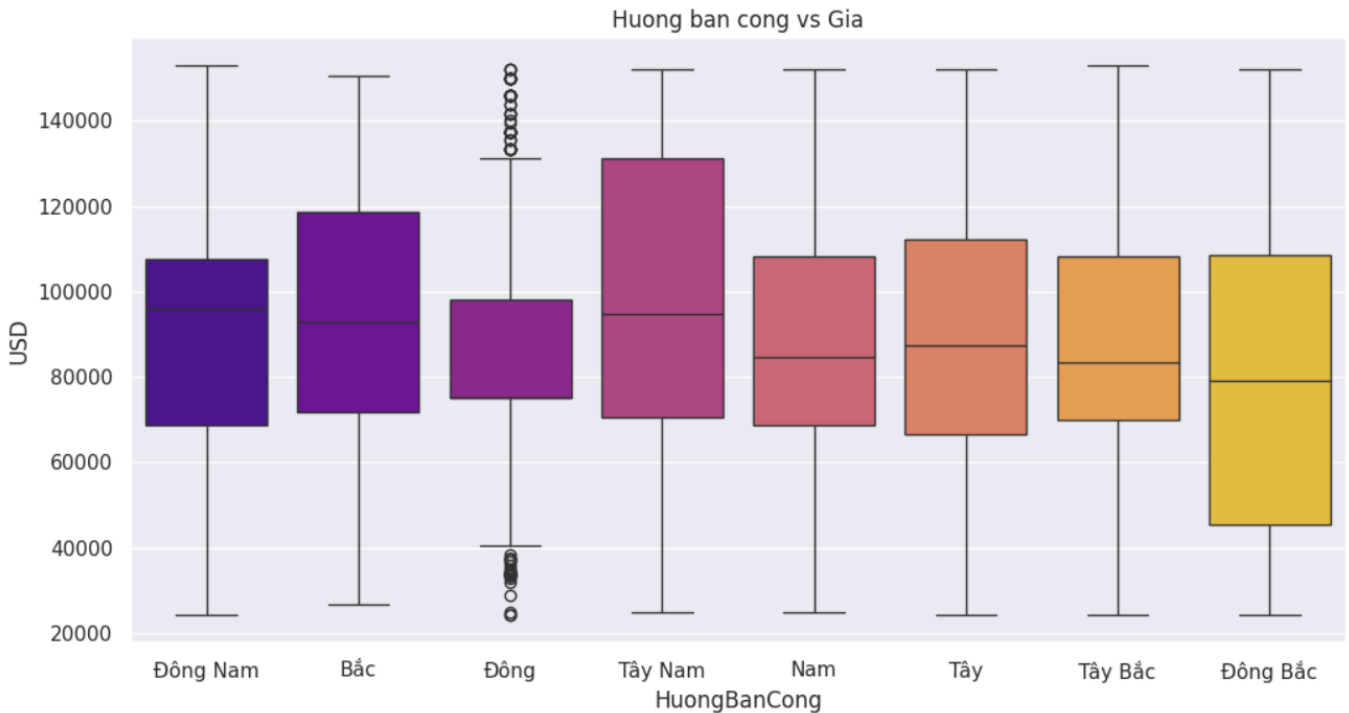
Căn hộ trung cư, Duplex: Loại nhà phổ biến, với mức độ ưa chuộng cao đều ở hai phân khúc trung bình và cao.

Căn hộ dịch vụ, mini: Cũng được ưa chuộng, đặc biệt ở phân khúc giá trung bình và cao.

Officetel: Ít phổ biến hơn so với chung cư và Duplex, nhưng vẫn có nhu cầu nhất định ở phân khúc giá trung bình và cao.

Căn hộ dịch vụ, mini: Ít phổ biến nhất, chỉ được ưa chuộng ở phân khúc giá cao.

IV.11. Hướng ban công có ảnh hưởng tới giá thuê không?



Hình 12. Biểu đồ hướng ban công có ảnh hưởng đến giá không

```
Đông Nam 1552
Đông 823
Tây Nam 767
Đông Bắc 707
Bắc 679
Tây 623
Tây Bắc 575
Nam 529
Name: HuongBanCong, dtype: int64
```

Dựa vào số lượng căn hộ theo từng hướng ban công:

Hướng Đông Nam (Đông Nam): Là hướng phổ biến nhất với 1552 căn hộ.

Hướng Đông (Đông): Có 823 căn hộ hướng Đông.

Hướng Tây Nam (Tây Nam): Có 767 căn hộ hướng Tây Nam.

Hướng Đông Bắc (Đông Bắc): Có 707 căn hộ hướng Đông Bắc.

Hướng Bắc (Bắc): Có 679 căn hộ hướng Bắc.

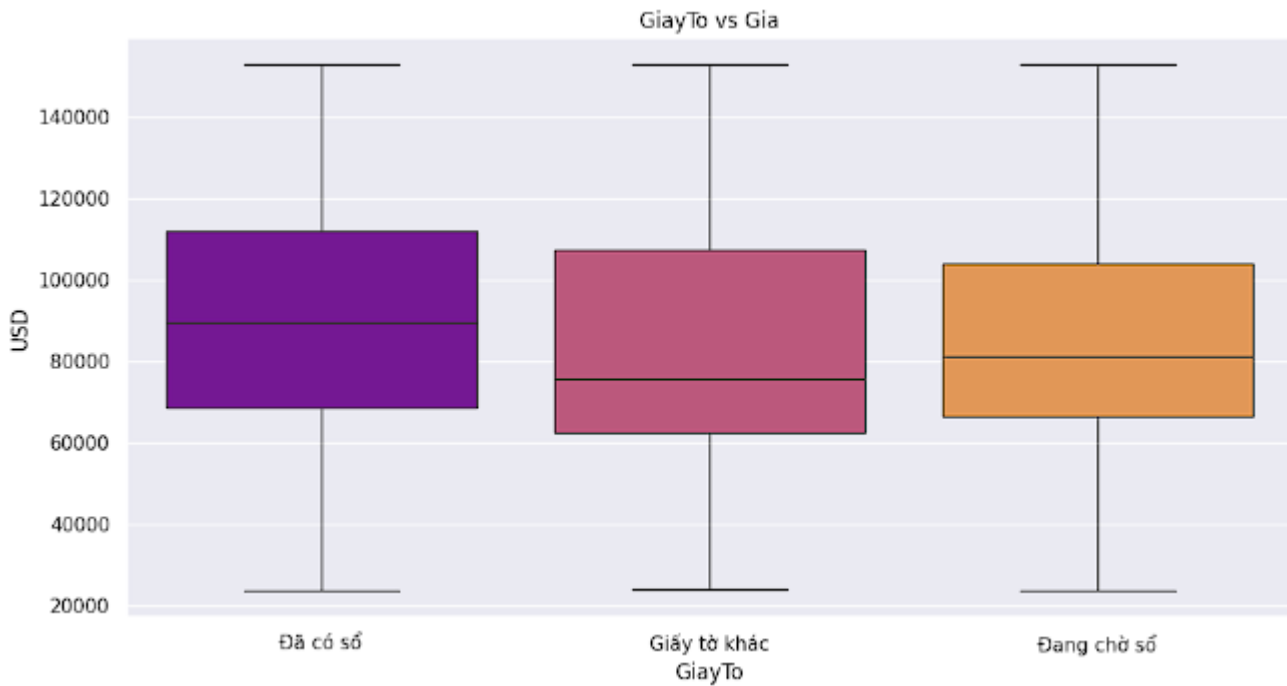
Hướng Tây (Tây): Có 623 căn hộ hướng Tây.

Hướng Tây Bắc (Tây Bắc): Có 575 căn hộ hướng Tây Bắc.

Hướng Nam (Nam): Có 529 căn hộ hướng Nam.

Từ phân tích trên, có thể thấy hướng Đông Nam (Đông Nam) là hướng được ưa chuộng nhất với số lượng căn hộ nhiều nhất, trong khi hướng Nam (Nam) là hướng ít được lựa chọn nhất. Điều này có thể phản ánh sở thích và tiện ích của cư dân đối với ánh sáng, gió và hướng nhìn.

IV.12. So sánh nhà ở có giấy tờ và chưa có thì có ảnh hưởng tới giá trị của bất động sản hay không?



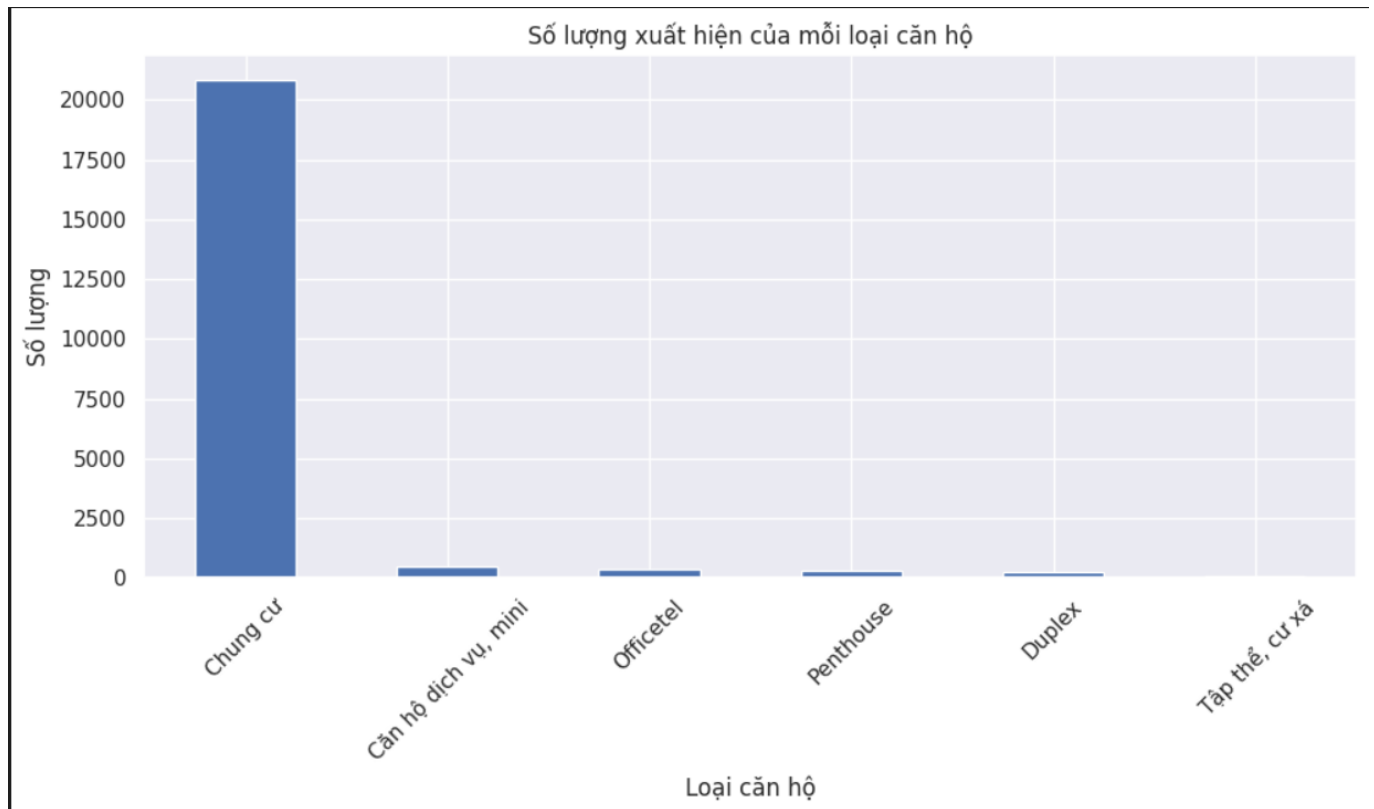
Hình 13. Biểu đồ thể hiện giá qua giấy tờ

Có thể thấy trường dữ liệu giấy tờ không ảnh hưởng quá nhiều tới giá nhà.

Nhìn chung, giá bất động sản Đã có sổ cao hơn so với các loại hình khác. Giá trị trung vị của nhóm này nằm ở mức 120.000 USD, cao hơn đáng kể so với nhóm Đang chờ sổ (80.000 USD) và nhóm giấy tờ khác (60.000 USD).

Biểu đồ cũng cho thấy sự phân bố giá trị rộng hơn cho nhóm nhà Đã có sổ, thể hiện sự đa dạng trong phân khúc thị trường này.

IV.13. Số lượng xuất hiện của mỗi loại căn hộ?



Hình 14. Biểu đồ số lượng của các loại căn hộ

Dựa trên biểu đồ trên ta có thể thấy số lượng căn hộ được xuất hiện nhiều nhất trên chợ tốt là những loại căn hộ chung cư.

Số lượng căn hộ chung cư là cao nhất, tiếp theo là căn hộ dịch vụ, mini, officetel, penthouse, duplex và tập thể, cư xá không ưa chuộng nên ít xuất hiện trên chợ.

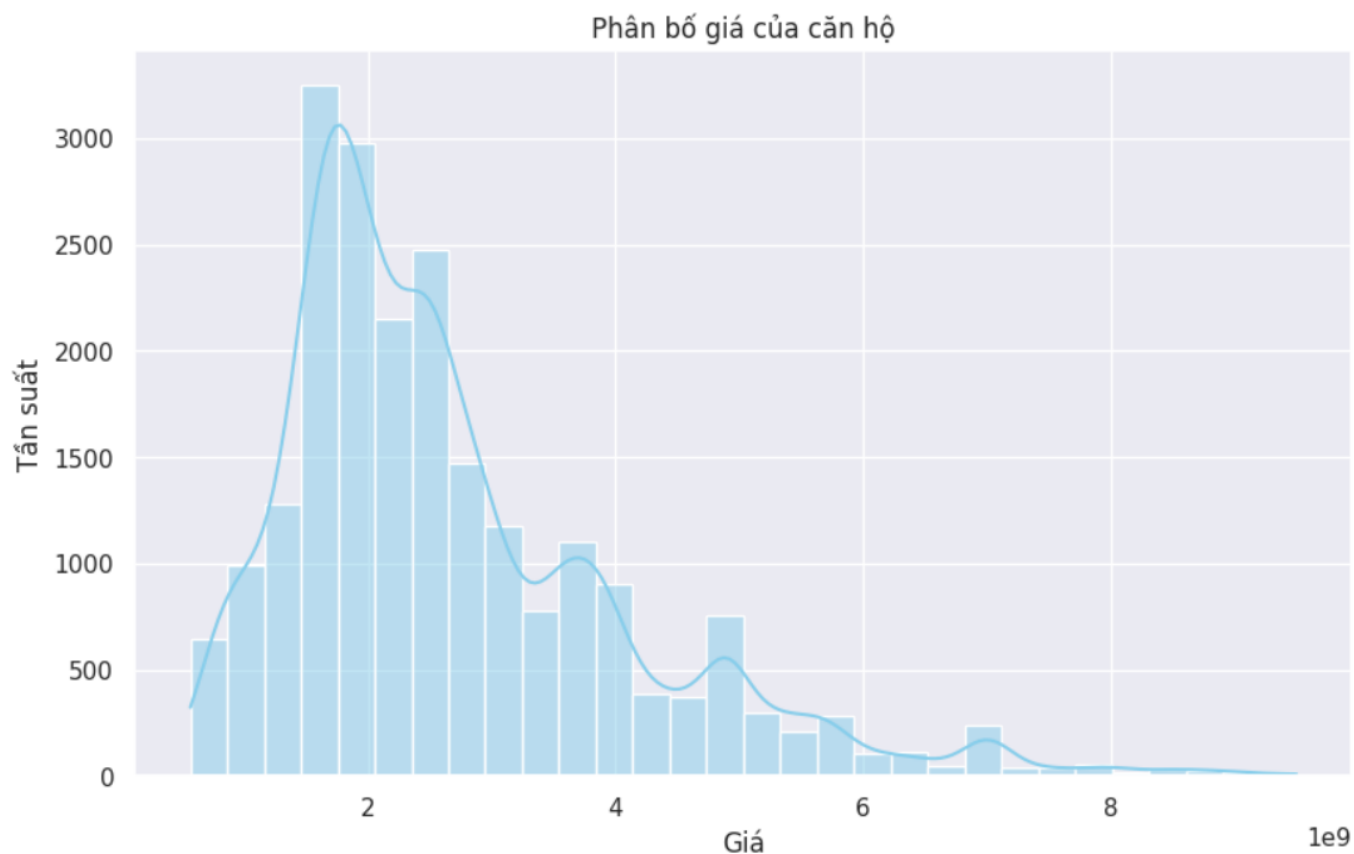
Sự gia tăng số lượng căn hộ có thể do nhiều nguyên nhân:

Nhu cầu nhà ở ngày càng tăng do tốc độ đô thị hóa nhanh.

Mức sống của người dân ngày càng được cải thiện, có khả năng mua nhà ở tốt hơn.

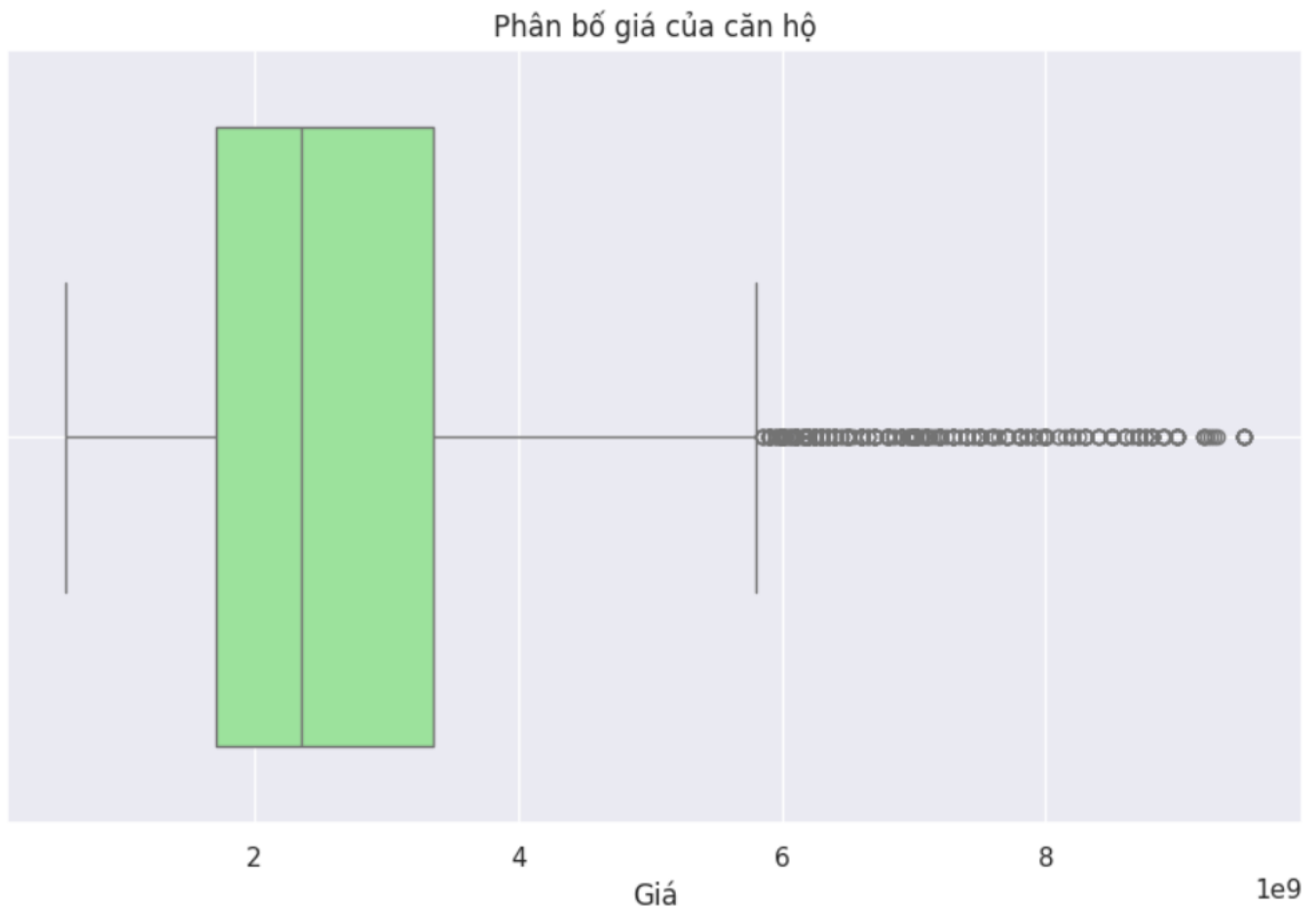
Các nhà đầu tư đổ vốn vào thị trường bất động sản, xây dựng nhiều dự án căn hộ mới.

IV.14. Phân bố giá của các căn hộ?



Hình 15. Biểu đồ histogram thể hiện phân bố giá của căn hộ

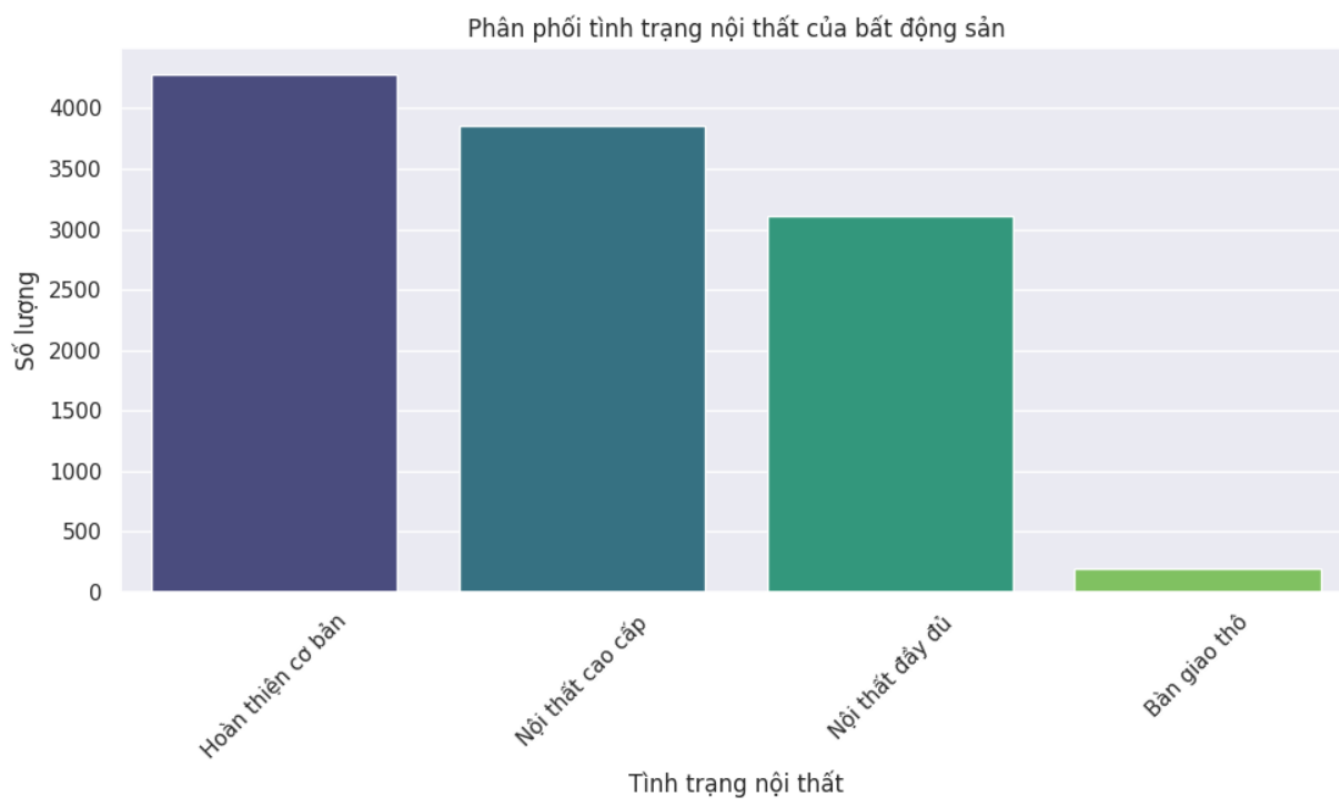
Histogram: Phân phối giá của căn hộ không đồng đều, có thể có sự tập trung mạnh ở một số khoảng giá cụ thể. Dữ liệu có thể bị lệch (skewed), với một số giá trị cao tập trung ở phần đuôi bên phải của phân phối.



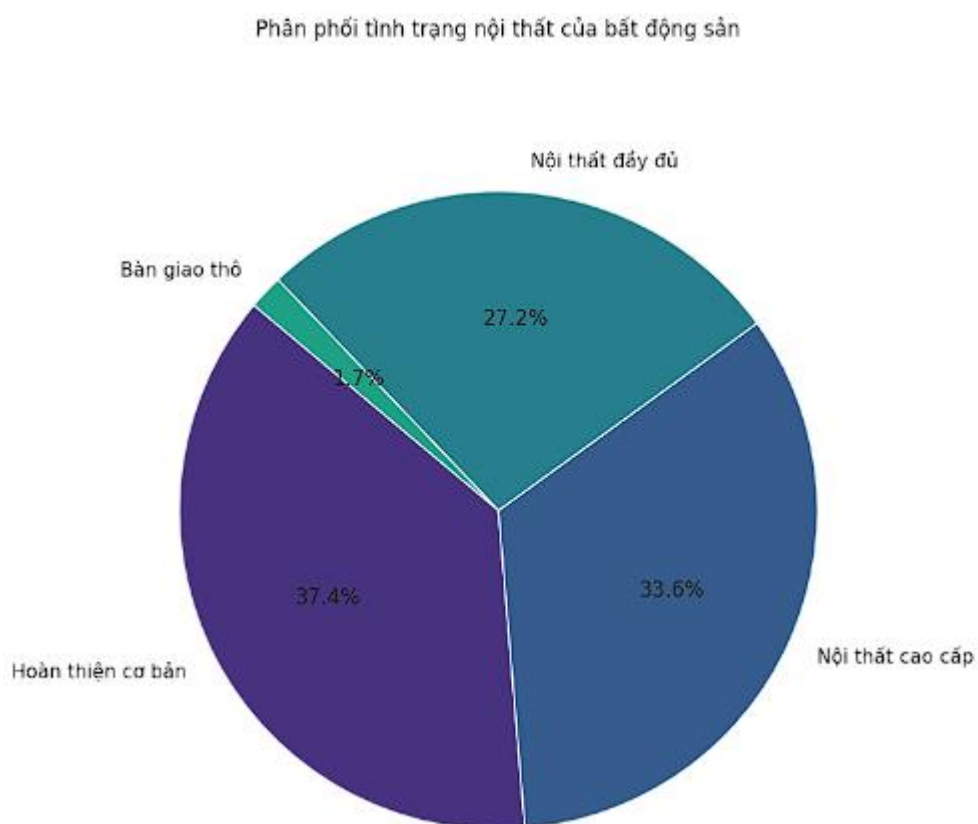
Hình 16. Biểu đồ phân bố giá của căn hộ

Boxplot: Boxplot cho thấy phân phối của giá căn hộ tập trung chủ yếu vào một khoảng giá có giá trị trung vị (median). Có một số giá trị ngoại biên (outliers) ở phía trên của biểu đồ, cho thấy sự biến động lớn trong giá của một số căn hộ. Nhìn chung, dựa trên hai biểu đồ này, có thể kết luận rằng phân bố giá của căn hộ có sự không đồng đều và có sự biến động lớn, đồng thời cũng có sự tập trung ở một số khoảng giá cụ thể. Điều này có thể phản ánh sự đa dạng trong thị trường bất động sản và yếu tố khác nhau ảnh hưởng đến giá của các căn hộ..

IV.15. Tình trạng các loại nội thất được xuất hiện nhiều trong các loại bất động sản?



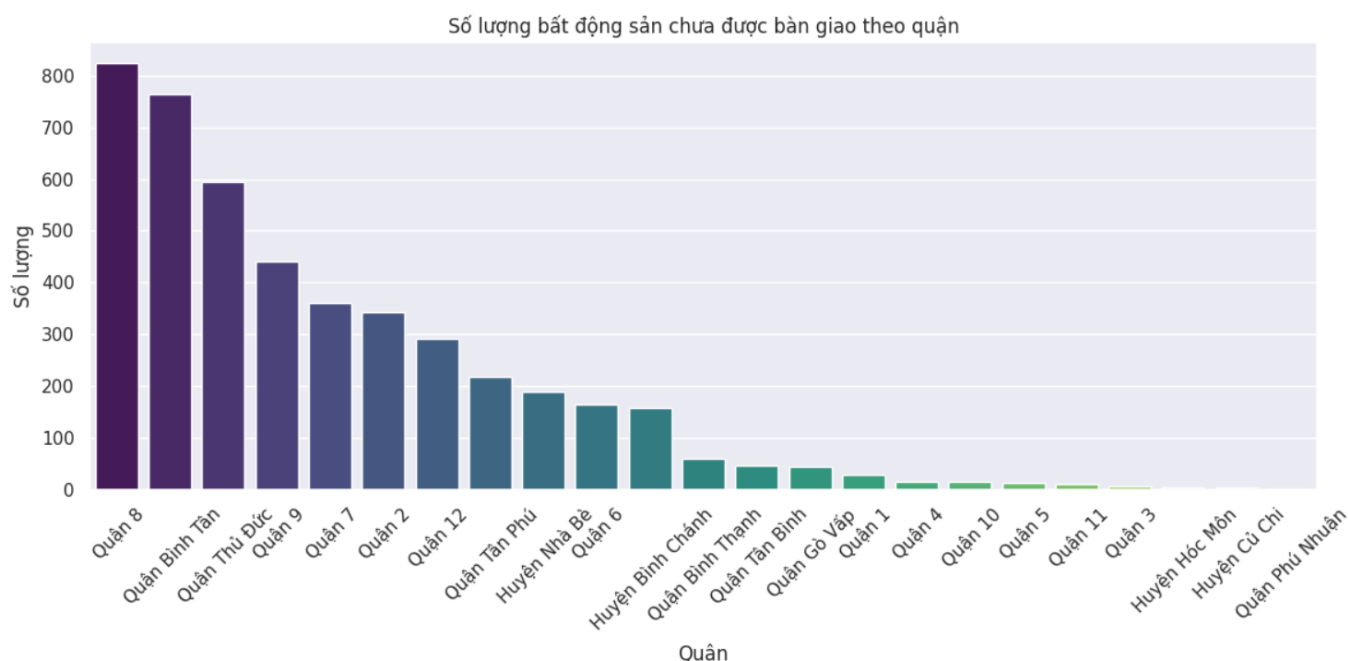
Hình 17. Biểu đồ tình trạng nội thất của các căn hộ



Hình 18. Biểu đồ tình trạng nội thất của các căn hộ

Dựa trên biểu đồ ta có thể thấy nội thất cao cấp và nội thất được hoàn thiện cơ bản được các nhà đầu tư lựa chọn cao nhất trong việc bán căn hộ.

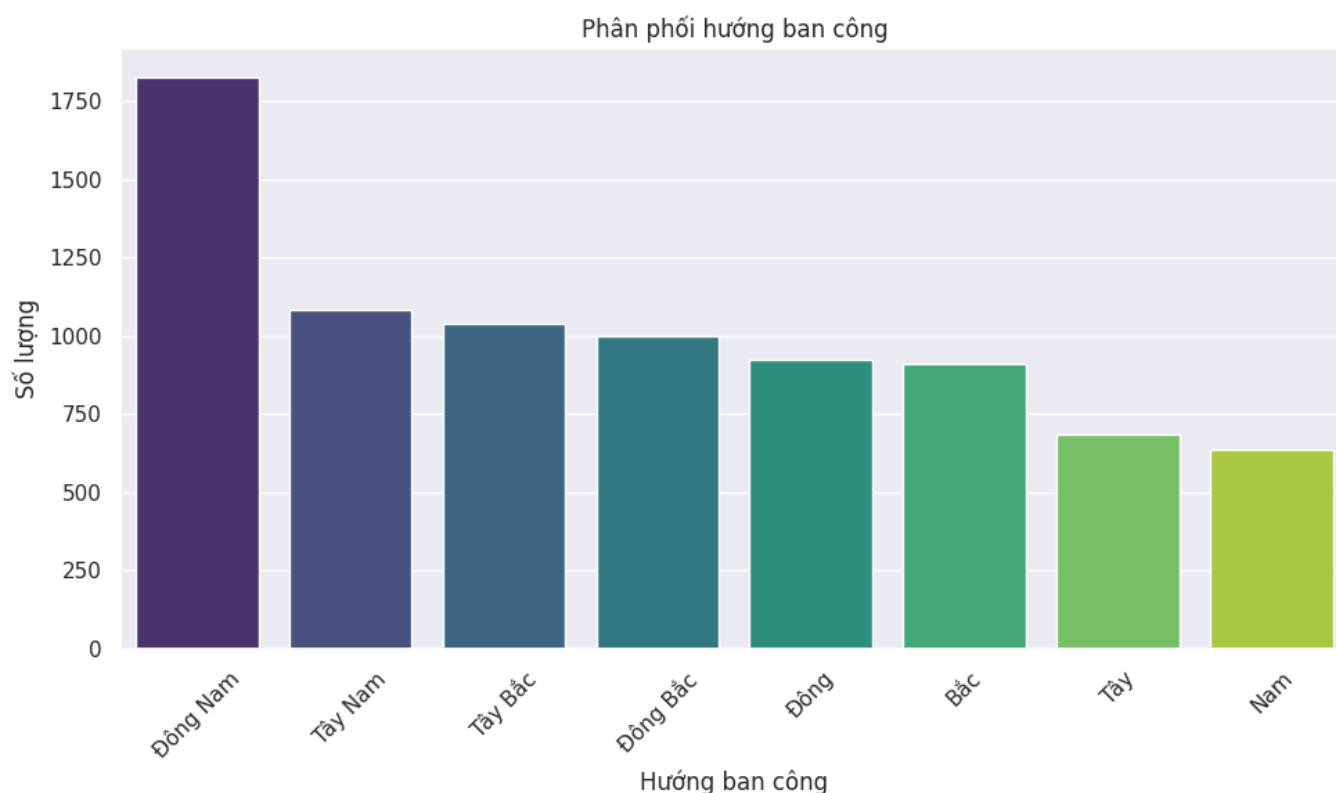
IV.16. Có bao nhiêu bất động sản chưa được bàn giao theo quận?



Hình 19. Biểu đồ các căn hộ chưa được bàn giao ở các quận

Dựa trên biểu đồ ta có thể thấy quận 8, quận Bình Tân, Quận Thủ Đức có nhiều căn hộ chưa được bán trên thị trường nên người mua có thể suy nghĩ về việc lựa chọn nơi sinh sống và giá cả phù hợp cho bản thân mình.

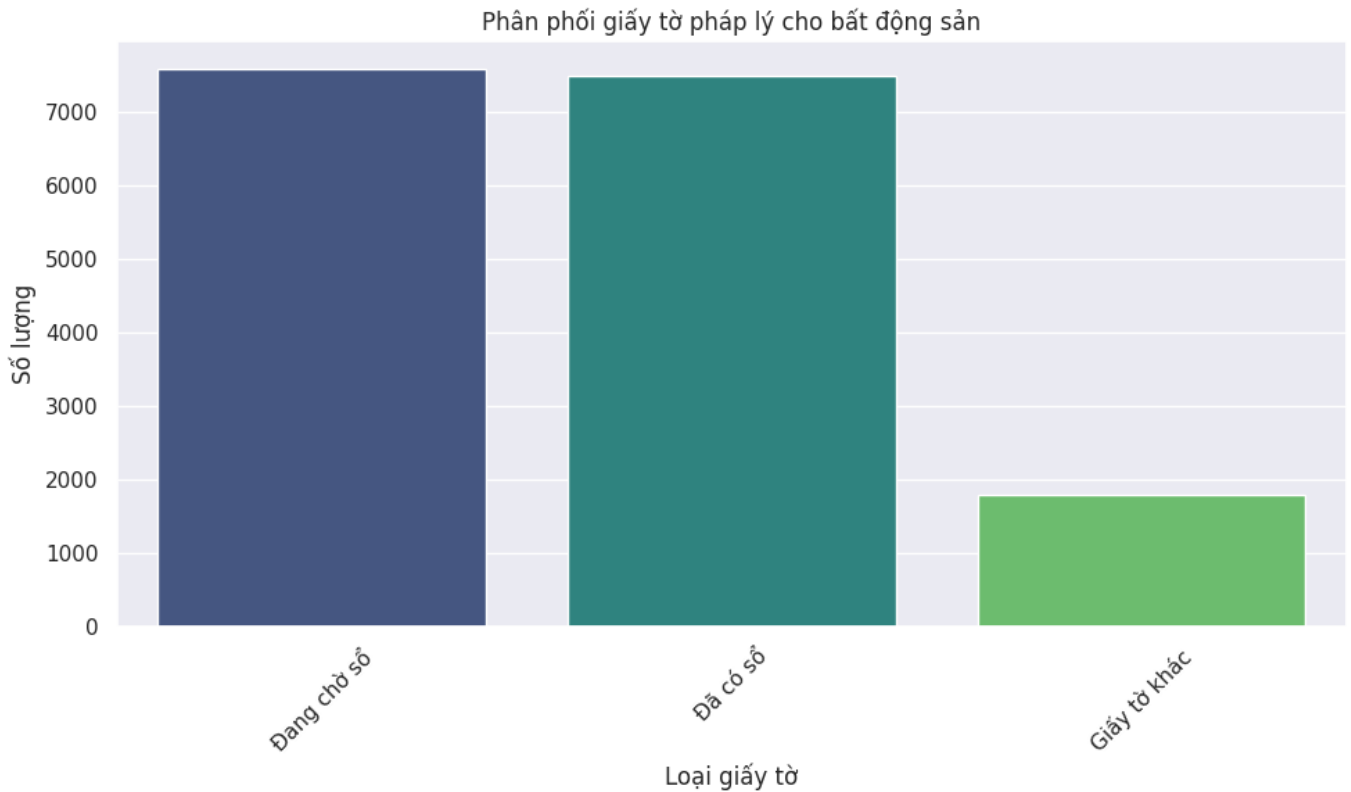
IV.17. Hướng ban công phổ biến nhất được các nhà bán bất động sản xây dựng nhiều nhất



Hình 20. Biểu đồ hướng ban công phổ biến nhất

Tuy lựa chọn hướng ban công phụ thuộc vào sở thích cá nhân, nhu cầu sử dụng và yếu tố phong thủy. Nhưng dựa vào biểu đồ trên thì xu hướng xây hướng ban công theo hướng Đông Nam, Tây Nam, Tây Bắc có nhiều sự lựa chọn nhất của các nhà đầu tư bất động sản.

IV.18. Các giấy tờ pháp lý thông dụng cho bất động sản ?



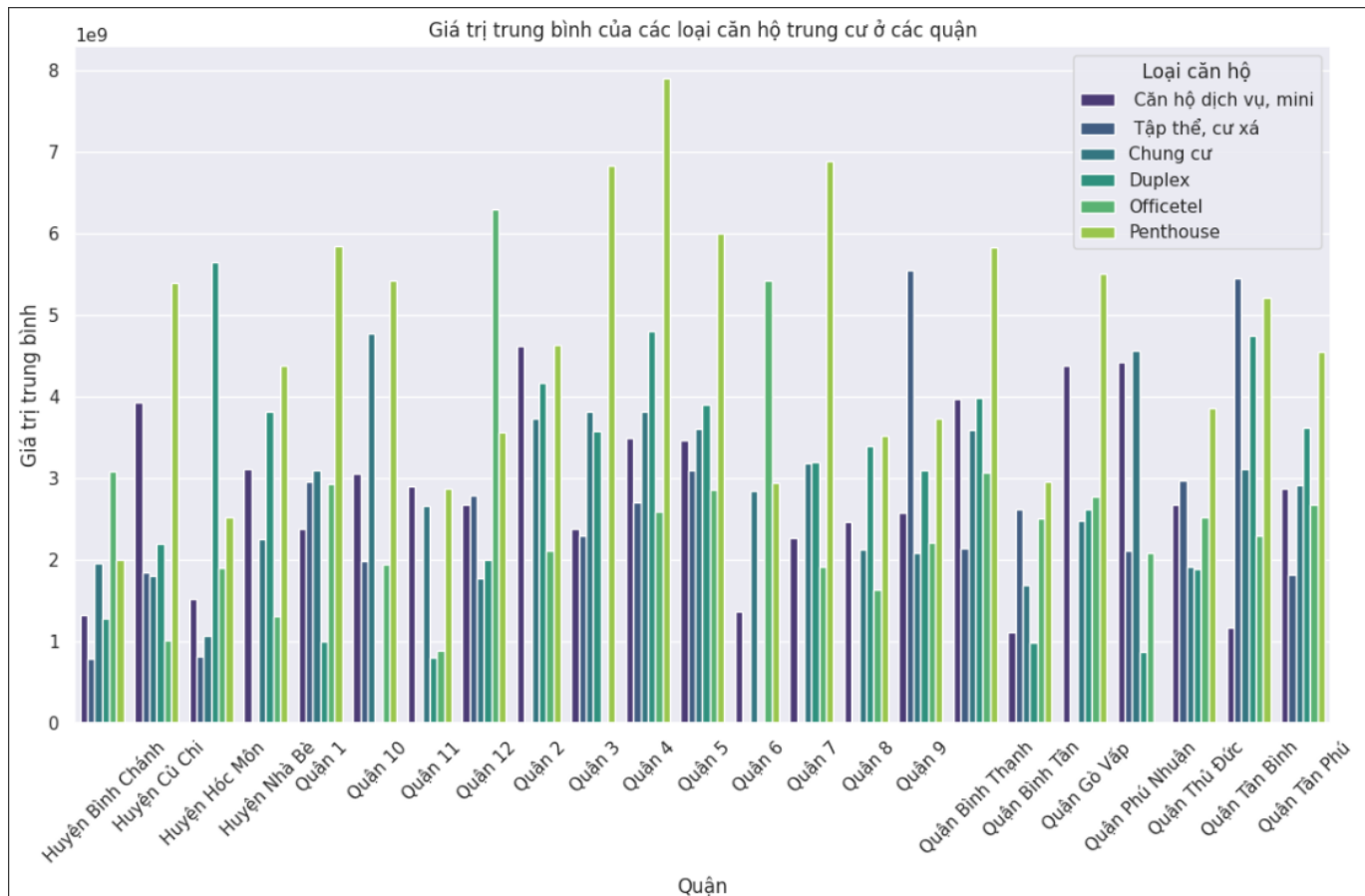
Hình 21. Biểu đồ các giấy tờ pháp lý thông dụng

Giấy tờ khác: cho thấy đây là loại giấy tờ phổ biến nhất trong lĩnh vực bất động sản. Nhưng lại chiếm số lượng rất ít

Đã có sổ: cho thấy đây là loại giấy tờ quan trọng và cần thiết trong các giao dịch bất động sản.

Đang chờ sổ: cho thấy đây là loại giấy tờ ít phổ biến hơn và có thể liên quan đến các thủ tục hành chính đang được thực hiện. Nhưng lại chiếm số lượng cao nhất

IV.19. Giá trị trung bình của các loại căn hộ qua các quận?



Hình 22. Biểu đồ giá trị trung bình của các loại căn hộ

- **Mức giá:**

Giá bán trung bình: Biểu đồ cho thấy mức giá bán trung bình nhà ở dao động từ 1,5 tỷ đến 5 tỷ đồng, với sự khác biệt đáng kể giữa các quận.

Quận có giá cao: Quận 1, 2, 3 và 7 có mức giá bán trung bình cao nhất, trên 4 tỷ đồng.

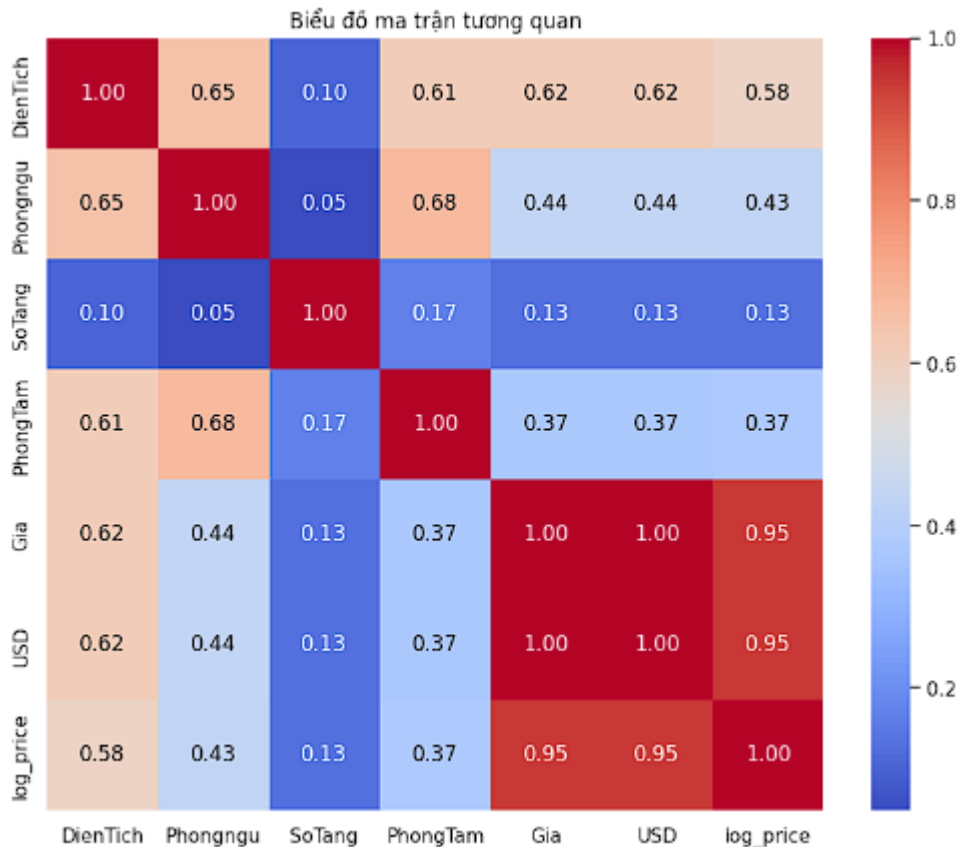
Quận có giá thấp: Các quận 12, Bình Chánh, Hóc Môn và Củ Chi có mức giá bán trung bình thấp nhất, dưới 2 tỷ đồng.

- **Xu hướng:**

Giá bán tăng cao: Nhìn chung, giá bán nhà ở có xu hướng tăng cao ở hầu hết các quận.

Sự chênh lệch giá: Mức độ chênh lệch giá bán giữa các quận khá lớn, có thể ảnh hưởng đến quyết định mua nhà của người dân.

IV.20. Ma trận tương quan giữa căn hộ với giá của căn hộ.



Hình 23. Ma trận tương quan giữa căn hộ với giá của căn hộ

Dựa trên biểu đồ trực quan ta có thể thấy: Diện tích, số tầng, phòng ngủ, phòng tắm của các căn hộ càng nhiều thì giá của các căn hộ càng tăng. Ngoài ra có thể suy đoán rằng, diện tích của căn hộ càng lớn thì trong căn hộ đó có thể có nhiều phòng tắm hoặc phòng ngủ.

V. Kết luận

Trong suốt quá trình làm việc, nhóm đã hoàn tất việc phân tích trên tất cả các thuộc tính của bộ dữ liệu, tìm hiểu và phân tích sâu hơn nhiều vấn đề được đặt ra trên bộ dữ liệu. Bằng cách sử dụng một số công cụ hỗ trợ xử lý dữ liệu như Numpy, Pandas; công cụ trực quan như Seaborn, Matplotlib để tìm hiểu và phân tích những vấn đề liên quan đến giá các căn hộ thu thập được từ tập dữ liệu. Trên đây là những kết quả được chọn lọc từ những phân tích của nhóm, toàn bộ các phân tích sẽ được trình bày trong source code.

Thông qua quá trình phân tích nhóm đã trả lời được rất nhiều câu hỏi liên quan về lĩnh vực Machine Learning & Data Science như Diện tích các căn chung cư có ảnh hưởng đến giá hay không? Số tầng có ảnh hưởng đến giá hay không? Số phòng ngủ có ảnh hưởng đến giá hay không? Số phòng tắm có ảnh hưởng đến giá hay không? Nó có thể giúp các cho người mua có những lựa chọn về bất động sản phù hợp với bản thân mình và giúp họ có lựa chọn đúng đắn về giá cả, không gây những tổn thất cho họ, phù hợp với điều kiện của họ. Bên cạnh đó việc phân tích này còn giúp những người kinh doanh về bất động sản có thể xây dựng những căn hộ, bán những căn hộ theo xu hướng thị trường không gây tổn thất và giúp họ bán được nhiều căn hộ trong tương lai.

Nhu cầu mua bán bất động sản trong tương lai ngày càng nhiều và phổ biến trên đất nước. Nên việc phân tích bất động sản ở TPHCM có thể giúp người mua lựa chọn đúng những căn hộ họ ưng ý trong tầm giá có thể... Đây là những giá trị quan trọng mà bộ dữ liệu mang lại.

Tài liệu tham khảo (IEEE)

- [1] <https://github.com/HungTrinhIT/FinalProject-Datascience/blob/main/Chotot/rawdata.csv>
(Lần cuối truy cập 12/3/2024)
- [2] <https://svcuong.github.io/post/remove-outliers/>
(Lần cuối truy cập 12/3/2024)
- [3] <https://200lab.io/blog/danh-muc-cac-loai-bieu-do-trong-data-visualization/>
(Lần cuối truy cập 12/3/2024)