**ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH**

**TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN**

**KHOA CÔNG NGHỆ THÔNG TIN**

□□&□□Logo

Description automatically generated

**BÁO CÁO ĐỒ ÁN CUỐI KỲ**

**MÔN HỌC: TRỰC QUAN HÓA DỮ LIỆU**

**PHÂN TÍCH GIÁ NHÀ Ở HÀ NỘI**

Giảng viên: TS. Bùi Tiến Lên

Nhóm thực hiện: Nhóm 7

Lớp: 21\_21

TP Hồ Chí Minh, tháng 7 năm 2024

NỘI DUNG

[I. Thông tin nhóm 1](#_Toc171241300)

[II. Mức độ hoàn thành của từng thành viên 1](#_Toc171241301)

[III. Nội dung báo cáo 2](#_Toc171241302)

[a. Mục tiêu của đề tài 2](#_Toc171241303)

[b. Phạm vi đề tài 2](#_Toc171241304)

[c. Kết quả nghiên cứu 2](#_Toc171241305)

[1. Dashboard 2](#_Toc171241306)

[2. Câu hỏi 1: Sheet 1 … 2](#_Toc171241307)

[3. Câu hỏi 2: 2](#_Toc171241308)

[d. Các tiêu chí đánh giá và mức độ hoàn thành 2](#_Toc171241309)

[1. Kết hợp nguồn dữ liệu đáng tin cậy 2](#_Toc171241310)

[2. Phù hợp với mục đích 2](#_Toc171241311)

[3. Rõ ràng và dễ hiểu 2](#_Toc171241312)

[4. Sự tích hợp và liên kết 2](#_Toc171241313)

[5. Tương tác và điều hướng 2](#_Toc171241314)

[6. Thiết kế hấp dẫn 2](#_Toc171241315)

[7. Phân tích dữ liệu 2](#_Toc171241316)

[8. Tích hợp AI 2](#_Toc171241317)

[IV. Tài liệu tham khảo 3](#_Toc171241318)

# Thông tin nhóm

Nhóm 7

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Họ và tên | MSSV | Mức độ hoàn thành nhiệm vụ |
| 1 | Trương Hoàng Nhân | 21120402 | 100% |
| 2 | Nguyễn Hoàng Quân | 21120403 | 100% |
| 3 | Lưu Kiến Đạt | 21120429 | 100% |
| 4 | Hoàng Thu Thủy | 20120382 | 100% |
| 5 | Lê Bảo Hiếu | 21120451 | 100% |

# Mức độ hoàn thành của từng thành viên

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Họ và tên | Công việc | Mức độ hoàn thành |
| 1 | Trương Hoàng Nhân (Lead) | * Phân công, quản lý tiến trình * Viết báo cáo * Vẽ biểu đồ Tableau * Làm slide | 100% |
| 2 | Nguyễn Hoàng Quân | * Vẽ biểu đồ Tableau * Chọn chủ đề và dataset * Tìm hiểu GeoJSON | 100% |
| 3 | Lưu Kiến Đạt | * Vẽ biểu đồ Tableau * Viết báo cáo phần dashboard | 100% |
| 4 | Hoàng Thu Thủy | * Vẽ biểu đồ Tableau | 100% |
| 5 | Lê Bảo Hiếu | * Vẽ biểu đồ Tableau * Viết báo cáo | 100% |

# Nội dung báo cáo

## Mục tiêu của đề tài

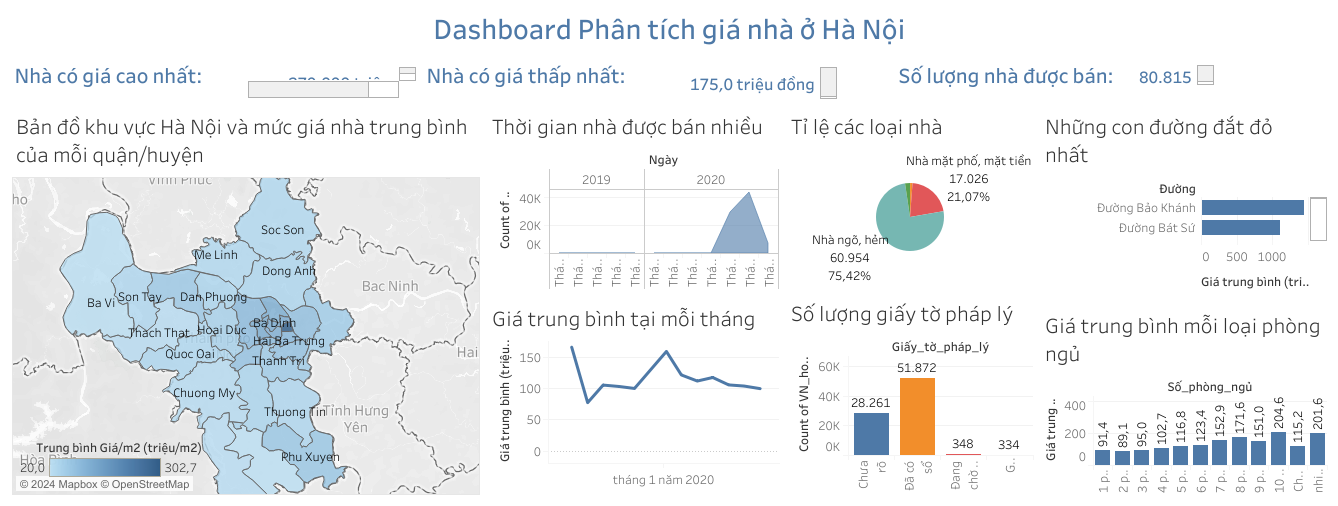
* Hiểu về thị trường nhà đất tại Hà Nội.
* Trực quan được sự phân bố giá nhà tại từng quận huyện.
* Cung cấp thêm thông tin hỗ trợ cho việc mua bán nhà như thông tin về giấy tờ pháp lý, loại nhà, mật độ nhà được bán và giá bình quân, từ đó rút ra được quận trọng tâm của Hà Nội.
* Quan sát được sự ảnh hưởng của bối cảnh thực tế của giai đoạn đó lên thị trường nhà đất của Hà Nội.
* Tạo ra chatbot giải đáp thắc mắc của người dùng khi xem bảng phân tích.

## Phạm vi đề tài

* Dựa trên bộ dữ liệu VN housing datasets (Hanoi), ta xác định được phạm vi địa lý của đề tài này nằm tại thành phố Hà Nội. Phạm vi nội dung bao gồm địa chỉ, thông tin giấy tờ pháp lý và giá bán của những căn nhà được rao bán trong thành phố Hà Nội trong vòng 1 năm từ tháng 8 năm 2019 đến tháng 8 năm 2020.

## Kết quả nghiên cứu

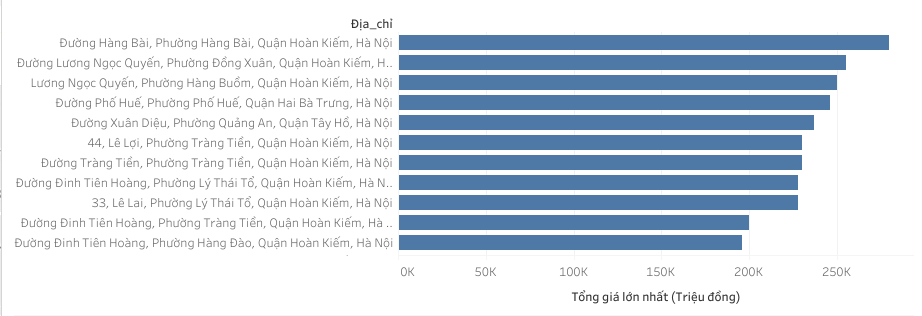
### Dashboard

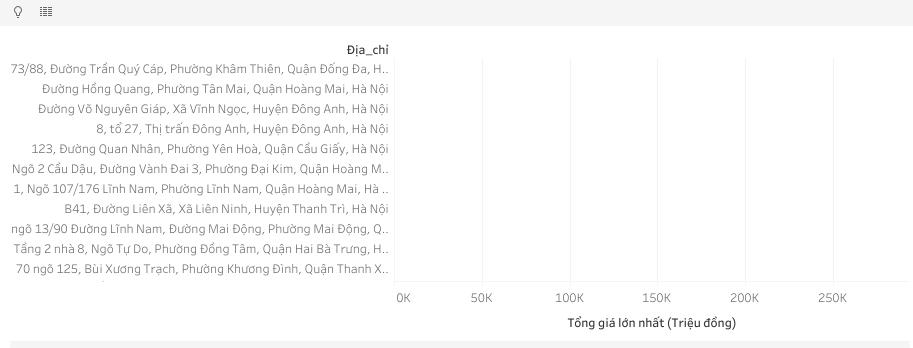


* Tổng thể, dashboard này cung cấp một cái nhìn toàn diện về thị trường nhà ở Hà Nội với các biểu đồ rõ ràng và được tổ chức. Sử dụng các loại biểu đồ khác nhau để trình bày các khía cạnh khác nhau của dữ liệu, làm cho nó trở thành một công cụ hữu ích cho việc phân tích.
* **Bản đồ**:
  + Bản đồ cung cấp cái nhìn trực quan về giá nhà theo các quận, huyện, rất hữu ích cho phân tích địa lý.
  + Màu sắc gradient giúp dễ dàng nhận diện các khu vực có giá cao hoặc thấp hơn.
  + Mỗi khi chọn vào một quận/huyện, các biểu đồ còn lại tự động thay đổi vùng dữ liệu từ Hà Nội thành quận/huyện được chọn.

### Giá nhà cao nhất/thấp nhất là bao nhiêu?

* Ý nghĩa:
  + ***Đối với người bán:***
    - **Định giá bất động sản:** Chủ sở hữu nhà và các công ty bất động sản có thể sử dụng biểu đồ này để định giá bất động sản của mình phù hợp với mức giá thị trường, tránh định giá quá cao hoặc quá thấp.
    - **Quyết định bán nhà:** Chủ sở hữu nhà có thể quyết định thời điểm và giá bán nhà dựa trên các mức giá cao nhất và thấp nhất trong thị trường.
  + ***Đối với người mua:***
    - **Phân tích khả năng mua:** Biểu đồ giúp người mua đánh giá khả năng mua nhà của mình dựa trên mức giá cao nhất và thấp nhất trong thị trường, từ đó quyết định xem có nên tiếp tục tìm kiếm trong khu vực đó hay chuyển sang khu vực khác có giá phù hợp hơn.
    - **Ưu tiên lựa chọn:** Người mua có thể ưu tiên các khu vực có giá nhà thấp hơn nhưng vẫn đáp ứng được nhu cầu về tiện ích, vị trí và môi trường sống.
  + ***Đối với nhà đầu tư***
    - Sử dụng biểu đồ này để **xác định những khu vực tiềm năng để đầu tư.** Nếu khoảng cách giữa giá cao nhất và thấp nhất lớn, có thể có cơ hội đầu tư vào các bất động sản giá rẻ và chờ đợi tăng giá.
    - **Đánh giá rủi ro:** Biểu đồ giúp nhà đầu tư đánh giá rủi ro liên quan đến việc đầu tư vào các bất động sản có giá cao hoặc thấp.

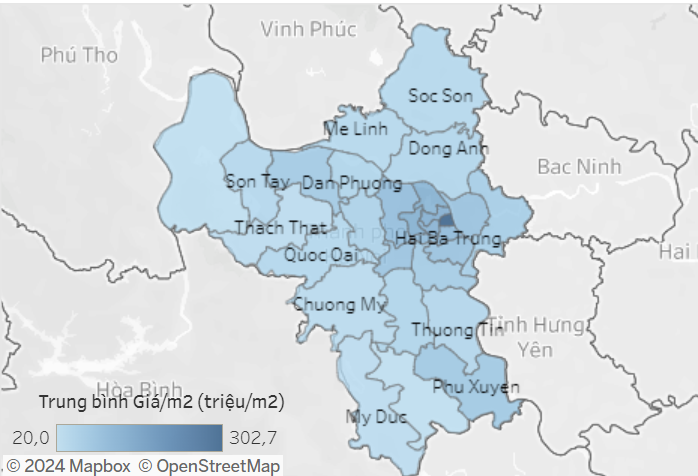




* + ***Kết luận:*** *Đường Hàng Bài, phường Hàng Bài, quận Hoàn Kiếm là khu vực có giá nhà cao nhất khoảng 279.998 (triệu đồng). Khu vực 73/88 đường Trần Quý Cáp, phường Khâm Thiêm, quận Đống Đa có giá nhà thấp nhất (175 triệu).*

### Khu vực có giá nhà đắt đỏ nhất là ở đâu?

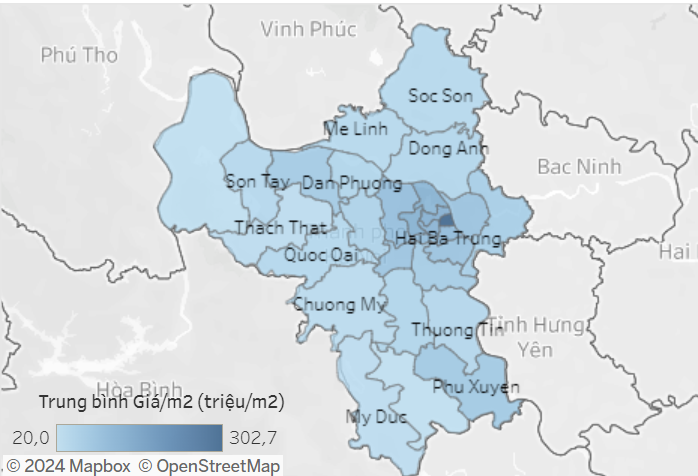
* Ý nghĩa:
* ***Nhà đầu tư bất động sản***
* **Xác định khu vực đầu tư tiềm năng:** Nhà đầu tư có thể xác định những khu vực có giá nhà cao để đầu tư, kỳ vọng vào sự ổn định và tiềm năng sinh lời cao trong tương lai.
* **Phân tích rủi ro và lợi nhuận:** Hiểu được khu vực nào có giá nhà cao nhất giúp nhà đầu tư đánh giá rủi ro và tiềm năng lợi nhuận khi đầu tư vào các khu vực đó.
* ***Người mua nhà***
* **Lập kế hoạch tài chính:** Người mua nhà có thể dựa vào thông tin này để lập kế hoạch tài chính chi tiết, xác định liệu họ có đủ khả năng tài chính để mua nhà ở những khu vực đắt đỏ hay không.
* **Lựa chọn khu vực sinh sống:** Giúp người mua hiểu rõ giá trị bất động sản ở các khu vực khác nhau, từ đó đưa ra quyết định mua nhà phù hợp với nhu cầu và ngân sách của mình.
* ***Chủ sở hữu nhà đất***
* **Định giá bất động sản:** Chủ sở hữu có thể sử dụng thông tin này để định giá tài sản của mình sao cho phù hợp với thị trường, tránh định giá quá cao hoặc quá thấp.
* **Quyết định bán nhà:** Nếu giá nhà ở khu vực của họ nằm trong nhóm các khu vực có giá nhà cao nhất, họ có thể cân nhắc bán nhà để thu lợi nhuận cao hơn.



* ***Kết luận:*** Theo biểu đồ, khu vực có giá nhà đắt đỏ nhất là quận Hoàn Kiếm, 302.7 triệu đồng/m2. Các Quận nội thành lân cân với Hoàn Kiếm như Hai Bà Trưng, Ba Đình... giá nhà cũng rất cao, >100 triệu đồng/m2.

### Khu vực nào có giá nhà rẻ nhất?

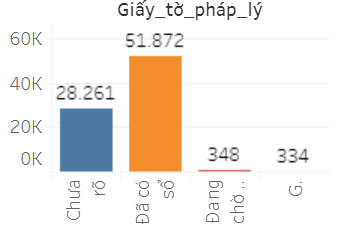
* Ý nghĩa:
* ***Nhà đầu tư bất động sản***
* **Cơ hội đầu tư giá rẻ:** Nhà đầu tư có thể tìm thấy các cơ hội đầu tư vào những khu vực có giá nhà rẻ, từ đó kỳ vọng vào tiềm năng tăng giá trong tương lai khi khu vực phát triển.
* **Chiến lược đa dạng hóa đầu tư:** Đầu tư vào các khu vực có giá nhà rẻ có thể giúp nhà đầu tư đa dạng hóa danh mục đầu tư, giảm rủi ro và tận dụng các cơ hội tăng trưởng dài hạn.
* ***Người mua nhà***
* **Lựa chọn phù hợp với ngân sách:** Những người có ngân sách hạn chế có thể tìm thấy các lựa chọn nhà ở phù hợp với khả năng tài chính của họ ở các khu vực có giá nhà rẻ.
* **Tận dụng tiềm năng tăng giá:** Mua nhà ở các khu vực có giá nhà rẻ có thể mang lại lợi ích tài chính lâu dài nếu khu vực đó phát triển và giá nhà tăng.
* ***Chủ sở hữu nhà đất***
* **Định giá tài sản hợp lý:** Sử dụng thông tin về các khu vực có giá nhà rẻ để định giá tài sản của mình một cách hợp lý, tránh việc định giá quá cao so với thị trường.
* **Quyết định cải thiện nhà ở:** Nâng cấp hoặc cải thiện bất động sản để tăng giá trị tài sản nếu giá nhà ở khu vực của họ thuộc nhóm thấp nhất.



* ***Kết luận:*** Quận Ba Vì có giá nhà bình quân thấp nhất (20 triệu đồng/m2). Các quận ngoại thành như Sơn Tây, Thạch Thất, Quốc Oai, Chương Mỹ, Mỹ Đức, Mê Linh, Sóc Sơn cũng có giá nhà rẻ, <40 triệu đồng / m2.

### Tổng số nhà đã được bán ra

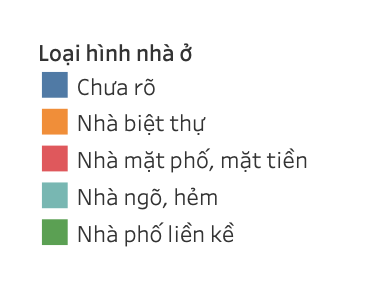
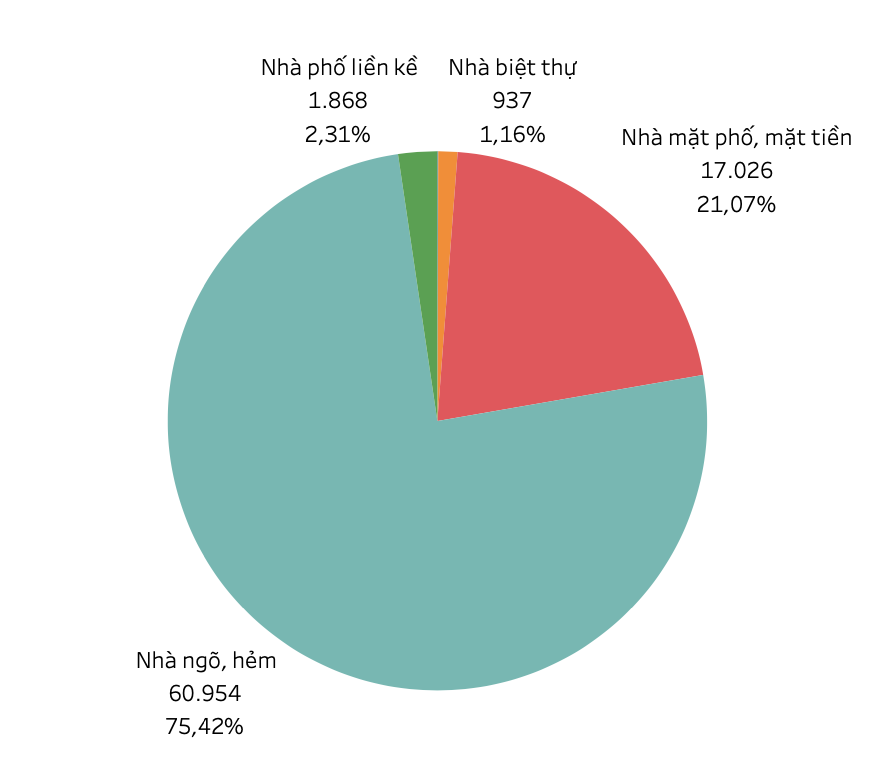
* Ý nghĩa:
* ***Nhà đầu tư bất động sản***
* **Đánh giá nhu cầu thị trường:** Số lượng nhà được bán ra phản ánh nhu cầu của thị trường bất động sản. Nếu số lượng nhà bán ra cao, điều này cho thấy nhu cầu mua nhà lớn, từ đó giúp nhà đầu tư đánh giá tiềm năng sinh lời của việc đầu tư vào khu vực này.
* **Xác định chiến lược đầu tư:** Thông tin về số lượng nhà bán ra giúp nhà đầu tư xác định các chiến lược đầu tư phù hợp, chẳng hạn như mở rộng danh mục đầu tư hoặc tìm kiếm các cơ hội đầu tư mới.
* ***Người mua nhà***
* **Hiểu rõ thị trường:** Người mua nhà có thể dựa vào thông tin này để hiểu rõ hơn về tình hình thị trường bất động sản, từ đó đưa ra quyết định mua nhà một cách thông minh và có kế hoạch.
* **Đánh giá tính cạnh tranh:** Nếu số lượng nhà bán ra lớn, điều này có thể cho thấy thị trường cạnh tranh cao, người mua cần chuẩn bị kỹ càng và hành động nhanh chóng để nắm bắt cơ hội.
* ***Chủ sở hữu nhà đất***
* **Đánh giá thời điểm bán nhà:** Thông tin về số lượng nhà bán ra giúp chủ sở hữu đánh giá thời điểm thích hợp để bán nhà. Nếu thị trường đang có nhiều giao dịch, họ có thể cân nhắc bán nhà để thu lợi nhuận tốt hơn.
* **Hiểu biết về nhu cầu thị trường:** Chủ sở hữu có thể sử dụng thông tin này để hiểu rõ nhu cầu của người mua và điều chỉnh giá bán hoặc cải thiện chất lượng nhà ở để thu hút người mua.



* ***Kết luận:*** Có khoảng 80.815 ngôi nhà được bán ra.

### Tỉ lệ giữa các loại nhà khác nhau

* Ý nghĩa:
* ***Nhà đầu tư bất động sản***
  + **Tìm ra được mô hình kinh doanh đem lại lợi nhuận cao:** Các nhà phố, mặt tiền, các căn hộ chung cư cao cấp, sẽ được đầu tư để mang lại lợi nhuận cao hơn cho nhà đầu tư.
* ***Người mua nhà***
  + **Chọn được loai nhà phù hợp với nhu cầu và khả năng:** Những loại nhà sẽ được chọn tùy vào nhu cầu của người mua nhà, một chỗ ở nhỏ gọn, giá cả phải chăng để dễ dàng trang trải cuộc sống hay một ngồi nhà mặt phố có đủ các tiện ích xã hội.
* ***Chủ sở hữu nhà đất:***
  + **Chọn được loại nhà phù hợp để xây dựng:** Những loại nhà sẽ được chủ nhà đất chọn để xây dựng như biệt thự rồi chuyển sang cho các nhà đầu tư bất động sản lấy lời nhuận



* ***Kết luận:*** Các căn nhà ở trong ngõ, hẻm chiếm tỉ lệ cao nhất với số lượng và tỉ lệ chiếm lần lượt là ***60,9%*** và ***75,42%***, có thể vì giá cả ở Hà Nội khá đắt đỏ, nên người thuê nhà có xu hướng chọn những căn nhà trong ngõ, hẻm để tiết kiệm chi phí.

### Thời gian nhà được bán nhiều

* Ý nghĩa:
* ***Nhà đầu tư bất động sản***
* **Cơ hội đầu tư:** Có thể lựa chọn thời điểm bán nhà để đạt được mức thỏa thuận tốt nhất hoặc thời điểm mua nhà vào thời điểm giá rẻ.
* ***Người mua nhà***
* **Lựa chọn thời điểm mua nhà:** Người mua có thể lựa chọn thời điểm thị trường không khan hiếm để mua được nhà với giá phải chăng.
* ***Chủ sở hữu nhà đất***
* **Lựa chọn thời điểm bán:** Tương tự như nhà đầu tư, người sở hữu cũng có thể lựa chọn thời điểm bán nhà hợp lý, có lợi cho bản thân.

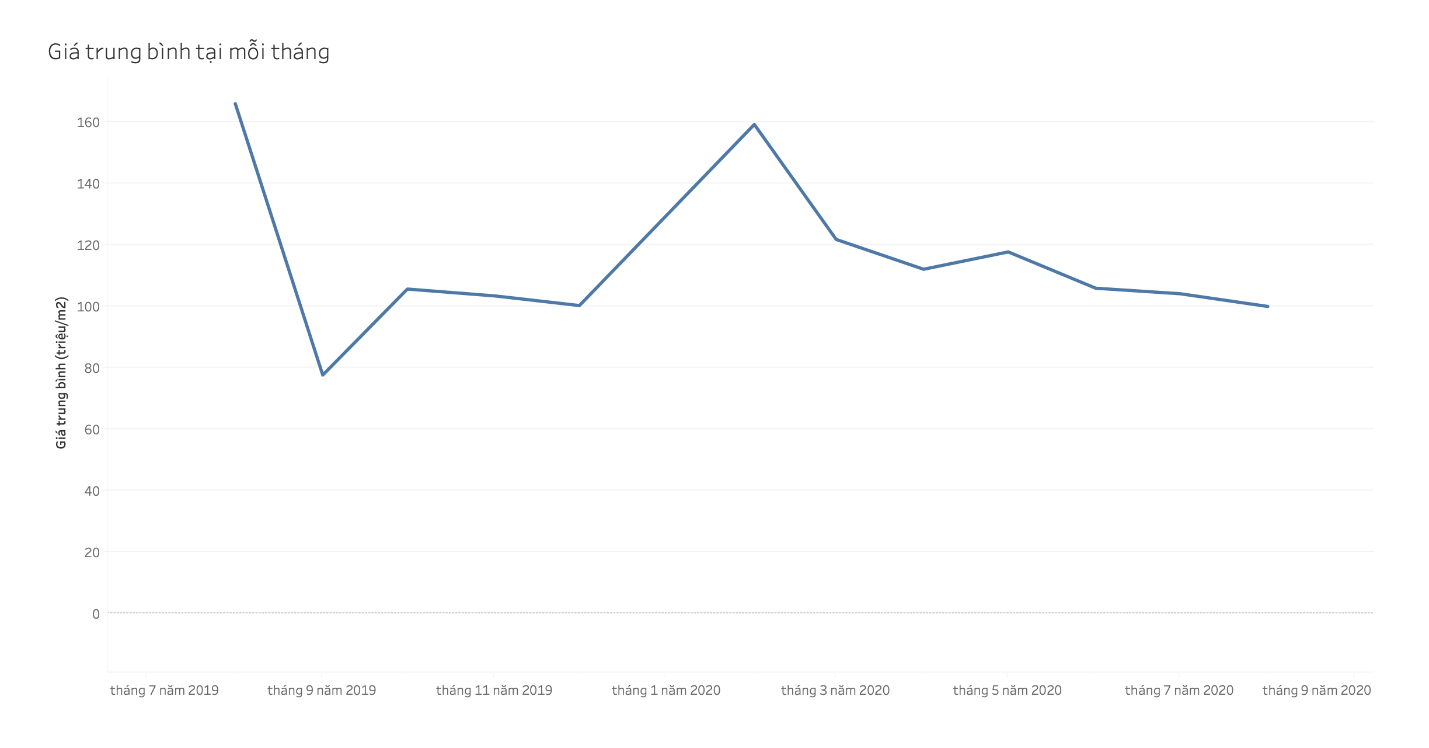
A graph with a blue line

Description automatically generated with medium confidence

* ***Kết luận:*** Số lượng nhà bán ra cao nhất tại thời điểm tháng 7 năm 2020 là 43.885.

### Giá trị trung bình mỗi tháng

* Ý nghĩa:
* ***Nhà đầu tư bất động sản:***
  + **Đánh giá thị trường:** Giá trung bình hàng tháng cho thấy tình hình thị trường bất động sản, liệu nó đang tăng trưởng hay suy giảm.
  + **Quyết định đầu tư:** Dựa vào xu hướng giá, nhà đầu tư có thể quyết định thời điểm mua vào hoặc bán ra để tối đa hóa lợi nhuận.
  + **Lợi nhuận tiềm năng:** Giá trung bình bán nhà cung cấp thông tin về lợi nhuận tiềm năng khi cho thuê hoặc bán lại bất động sản.
  + **So sánh giá trị:** So sánh giá trung bình với giá trị tài sản của mình để xác định chiến lược đầu tư.
* ***Người mua nhà:***
  + **Lập kế hoạch tài chính:** Giá trung bình bán nhà mỗi tháng giúp người mua lập kế hoạch tài chính cụ thể và quyết định số tiền cần dành cho việc mua nhà.
  + **Thời điểm mua nhà:** Theo dõi giá trung bình hàng tháng giúp người mua xác định thời điểm thích hợp để mua nhà với giá tốt nhất.
  + **So sánh giá cả:** Giá trung bình giúp người mua so sánh giá cả giữa các khu vực và loại nhà khác nhau để lựa chọn phù hợp nhất.
* ***Chủ sở hữu nhà đất:***
  + **Đánh giá giá trị tài sản:** Giá trung bình bán nhà mỗi tháng giúp chủ sở hữu đánh giá lại giá trị hiện tại của tài sản bất động sản của họ.
  + **Quyết định bán nhà:** Theo dõi giá trung bình giúp chủ sở hữu quyết định thời điểm bán nhà để thu được lợi nhuận cao nhất.
  + **Lợi nhuận từ cho thuê:** Giá trung bình cũng giúp chủ sở hữu xác định giá thuê hợp lý nếu họ muốn cho thuê tài sản của mình.



* ***Kết luận:*** Giá trung bình bán nhà mỗi tháng là một chỉ số quan trọng giúp đưa ra các quyết định kinh tế hợp lý và tối ưu hóa lợi nhuận hoặc chi phí của mình.

### Số lượng giấy tờ pháp lý

### Bản đồ của khu vực Hà Nội

### Những con đường đắt đỏ nhất Hà Nôi

## Các tiêu chí đánh giá và mức độ hoàn thành

### Kết hợp nguồn dữ liệu đáng tin cậy

* Nhóm đã chọn Kaggle làm nguồn dữ liệu chính vì độ tin cậy và đa dạng của dữ liệu. Nhóm cũng đã thực hiện các bước tiền xử lý để đảm bảo không có lỗi hoặc thiếu sót trong quá trình sử dụng bộ dữ liệu.

### Phù hợp với mục đích

* Việc trực quan hóa dữ liệu được thực hiện với mục tiêu rõ ràng, đảm bảo phản ánh đúng nhu cầu thông tin của các đối tượng khác nhau.

### Rõ ràng và dễ hiểu

* Các biểu đồ và đồ thị được thiết kế để người xem có thể dễ dàng hiểu và nắm bắt thông tin, với mục tiêu truyền đạt thông điệp một cách minh bạch và dễ tiếp cận.

### Sự tích hợp và liên kết

* Đối với các trực quan hóa bao gồm nhiều biểu đồ hoặc đồ thị, nhóm đã chú trọng đến việc liên kết và tích hợp chúng một cách có ý nghĩa, mặc dù chưa làm rõ hơn mối quan hệ giữa các phần được.

### Tương tác và điều hướng

* Dashboard được thiết kế để tương tác một cách mượt mà, với các bộ lọc giúp người dùng dễ dàng điều hướng và quan sát dữ liệu theo từng quận, huyện, ...

### Thiết kế hấp dẫn

* Màu sắc chủ đạo là xanh lam, tạo ra một thiết kế đơn sắc nhưng hấp dẫn, tránh được việc quá tải về màu sắc và giúp người dùng tập trung vào thông tin quan trọng.

### Phân tích dữ liệu

* Phần trực quan hóa đã thể hiện được sự thay đổi theo thời gian của các giá trị, giữa các biến luôn có mối quan hệ rõ ràng, tuy nhiên nhóm chưa áp dụng dự đoán xu hướng trong tương lai.
* Dựa trên những dữ liệu đã trực quan, ta có thể phân tích bối cảnh đằng sau dữ liệu được thể hiện.
* Ví dụ:

A screenshot of a computer

Description automatically generated

* Trên đây là dashboard phân tích giá nhà tại quận Cầu Giấy, ta có thể thấy số lượng nhà bán ra giai đoạn từ tháng 8 năm 2019 đến tháng 5 năm 2020 không đáng kể, có thể nhận định là do tình hình dịch bệnh Covid-19 bùng phát làm cho người dân không thể sinh hoạt xã hội bình thường, dẫn tới tình trạng thị trường bị đình trệ.

### Tích hợp AI

* Do OpenAI đã bắt đầu tính phí cho việc sử dụng API của ChatGPT và nhóm không có cách để mua nên đã chuyển sang sử dụng API miễn phí của Gemini do Google cung cấp. Tuy nhiên, Gemini bị giới hạn hơn so với ChatGPT nên nhóm không thể cài đặt hết toàn bộ chức năng trong khoản thời gian cho phép.

# Tài liệu tham khảo