

TRƯỜNG ĐẠI HỌC VĂN LANG
KHOA CÔNG NGHỆ THÔNG TIN



BÀI TẬP LỚN CÁ NHÂN
Học Phần: Các Hệ Hỗ Trợ Ra Quyết Định

CÔNG CỤ BI VÀ TRỰC QUAN HÓA

Họ và tên:

Đặng Thị Thu Hằng - 2274801040002

Trần Khánh Linh - 2274802010480

GVHD: Nguyễn Thái Anh

TP. Hồ Chí Minh – năm 2024

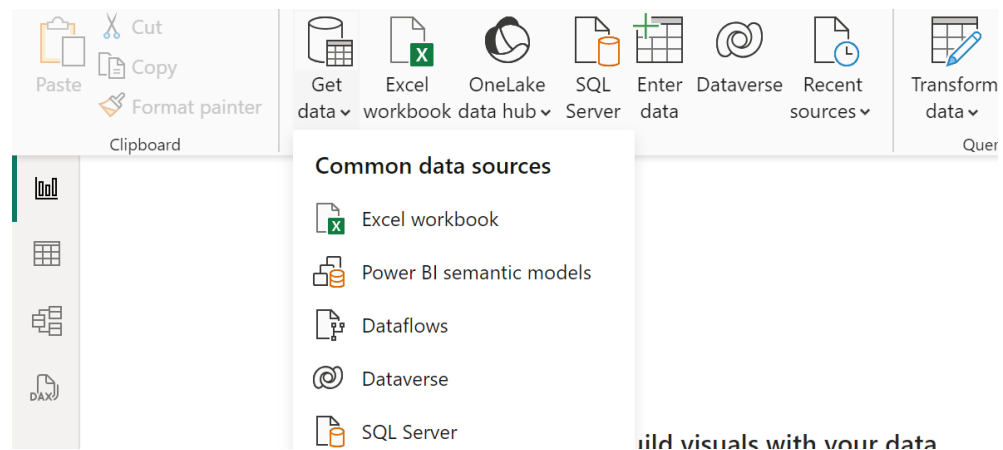
CÂU 1 - TỔNG QUAN VỀ POWER BI

Power BI được Microsoft xây dựng và phát triển từ năm 2013 - là một giải pháp báo cáo toàn diện, bao gồm từ việc chuẩn bị dữ liệu, trực quan hóa, phân phối đến quản lý dữ liệu thông qua các công cụ và nền tảng trực tuyến.

Power BI có thể đáp ứng nhu cầu từ những bài báo cáo đơn giản với một nguồn dữ liệu duy nhất đến những mô hình dữ liệu phức tạp và chủ đề nhất quán. Trải nghiệm công cụ Power BI giúp tạo ra những bài báo cáo sinh động, đẹp mắt, tương tác dễ dàng hỗ trợ người dùng đưa ra quyết định nhanh chóng và hiệu quả. Với những khả năng này, Power BI trở thành công cụ thiết yếu cho việc phân tích và hỗ trợ ra quyết định.

I. TÍNH NĂNG QUAN TRỌNG CỦA POWER BI

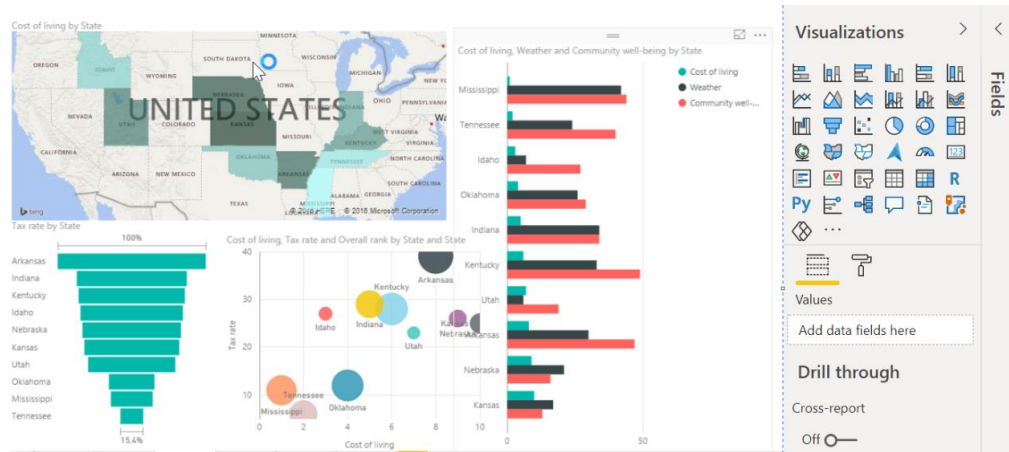
- **Kết nối và thu thập dữ liệu:** Power BI cho phép kết nối với nhiều nguồn dữ liệu khác nhau như cơ sở dữ liệu SQL Server, Excel, Azure, ... và nhiều API khác, thậm chí có thể trích xuất dữ liệu dạng bảng trực tiếp từ URL từ trang web. Người dùng có thể dễ dàng kết hợp dữ liệu từ nhiều nguồn khác nhau để tạo nên một bộ dữ liệu hoàn chỉnh. Không những thế, Power BI còn hỗ trợ dữ liệu đám mây lẫn dữ liệu local, giúp linh hoạt trong việc xử lý và chuẩn bị dữ liệu.



- **Thực hiện các thao tác làm sạch và biến đổi dữ liệu** thông qua ứng dụng Power Query Editor. Đây cũng là một phần của Power BI dùng để xử lý dữ liệu và đưa dữ liệu về dạng bảng biểu tiến hành xây dựng mô hình dữ liệu (data model) theo nhu cầu sử dụng.

	months_at_customer	age	policy_number	policy_state	policy_col_min	policy_col_max
1	328	48	523585	OH		250
2	228	42	242888	IN		250
3	134	29	687098	OH		200
4	256	41	227812	IL		250
5	228	44	367455	IL		500
6	256	39	104394	OH		250
7	137	34	413978	IN		250
8	165	37	429027	IL		100
9	27	33	485645	IL		100
10	212	42	636501	IL		100
11	235	42	543630	OH		100
12	447	61	214618	OH		100
13	60	23	842643	OH		100
14	121	34	626808	OH		100
15	180	38	644082	OH		250
16	473	58	892874	IN		100
17	70	26	558538	OH		500
18	140	31	275565	IN		500
19	160	37	921282	OH		500
20	196	39	149172	IN		500
21	460	42	189430	IN		250
22	217	41	433876	IL		500
23	370	55	285496	IL		100

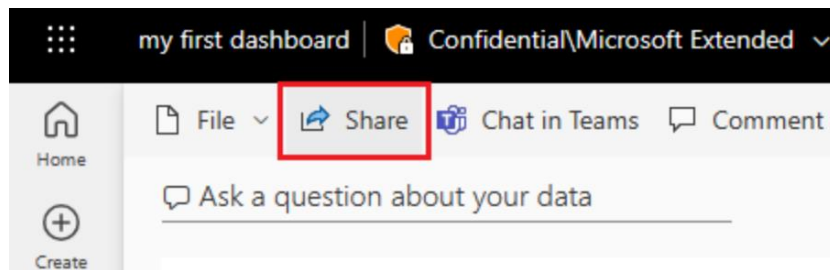
- **Trực quan hóa dữ liệu:** Là một trong những phần cốt lõi và nền tảng cơ bản của Power BI. Công cụ cung cấp hàng loạt biểu đồ, bản đồ, và đồ thị tương tác đa dạng. Người dùng có thể tạo ra các báo cáo trực quan, dễ hiểu từ những trường dữ liệu phức tạp thông qua hình ảnh sinh động. Bên cạnh đó, khả năng tương tác giữa các thành phần trên báo cáo còn giúp người xem có thể lọc và phân tích thông tin chi tiết một cách dễ dàng.



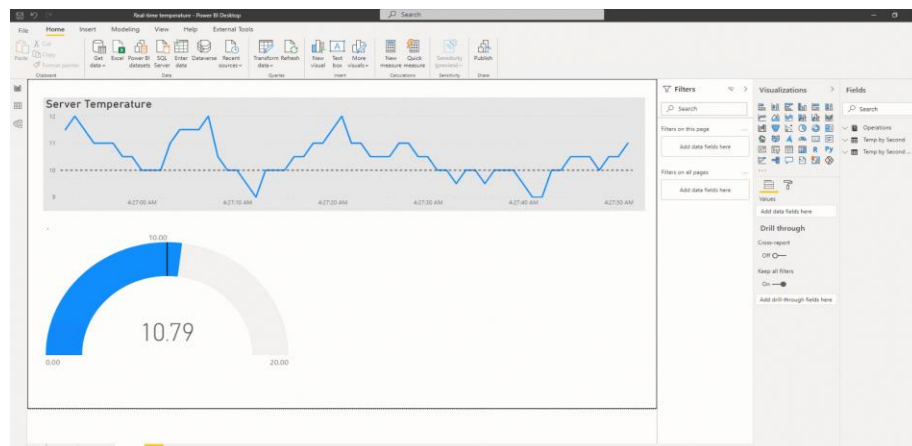
- **Phân tích nâng cao:** Power BI hỗ trợ các tính năng phân tích dữ liệu nâng cao, bao gồm khả năng sử dụng ngôn ngữ DAX (Data Analysis Expressions) để tạo ra các công thức tính toán phức tạp. Ngoài ra, Power BI tích hợp với các công cụ phân tích dữ liệu khác như Python và R, cho phép thực hiện các phân tích thống kê và dự báo chính xác hơn. Tính năng này giúp người dùng không chỉ tạo ra báo cáo mà còn phân tích chuyên sâu về xu hướng và mối quan hệ trong dữ liệu.



- **Chia sẻ và cộng tác:** hỗ trợ việc chia sẻ báo cáo dễ dàng qua Power BI Service, giúp người dùng đăng tải và chia sẻ báo cáo với các thành viên trong tổ chức hoặc đối tác bên ngoài. Người dùng có thể thiết lập quyền truy cập, cùng cộng tác trực tiếp trên các báo cáo, và theo dõi thông tin mới nhất nhờ khả năng đồng bộ hóa dữ liệu tự động. Điều này giúp tạo ra một môi trường làm việc hợp tác, hiệu quả cho các nhóm dựa trên dữ liệu chung.

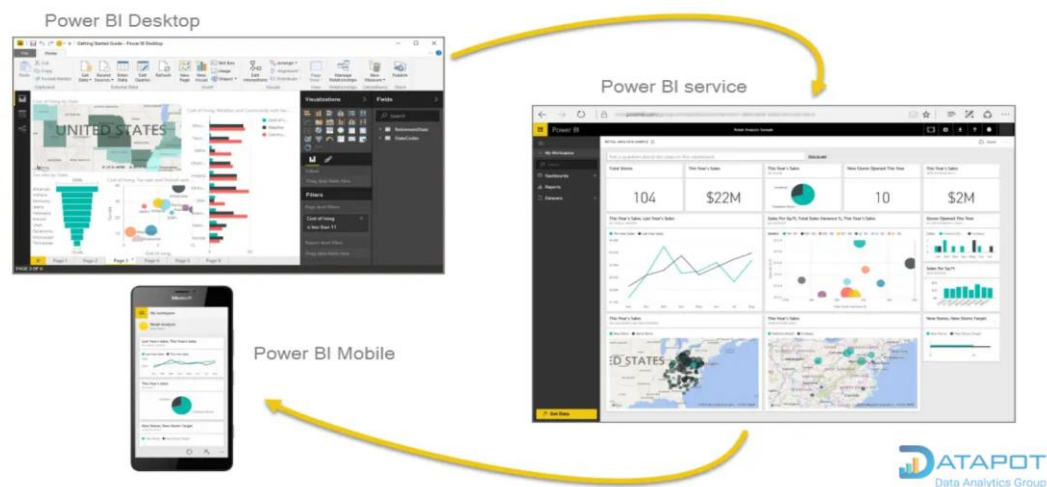


- **Real-time data:** Power BI cho phép người dùng tích hợp và theo dõi dữ liệu thời gian thực từ nhiều nguồn khác nhau như cảm biến, thiết bị IoT, và các nền tảng dữ liệu trực tuyến. Các bảng điều khiển (dashboards) có thể hiển thị thông tin mới nhất ngay khi có thay đổi, giúp tổ chức đưa ra các quyết định nhanh chóng và kịp thời trong những tình huống cần sự phản ứng nhanh như giám sát hoạt động kinh doanh, quản lý hệ thống hoặc theo dõi sản xuất.



- **Bảo mật dữ liệu:** Power BI tích hợp các tính năng bảo mật dữ liệu mạnh mẽ, bao gồm mã hóa, xác thực đa yếu tố (MFA) và phân quyền truy cập dựa trên vai trò (role-based access). Điều này đảm bảo rằng chỉ những người có thẩm quyền mới có thể truy cập và chỉnh sửa báo cáo hoặc mô hình dữ liệu.

II. ĐỊNH NGHĨA CÁC THÀNH PHẦN CHÍNH CỦA POWER BI

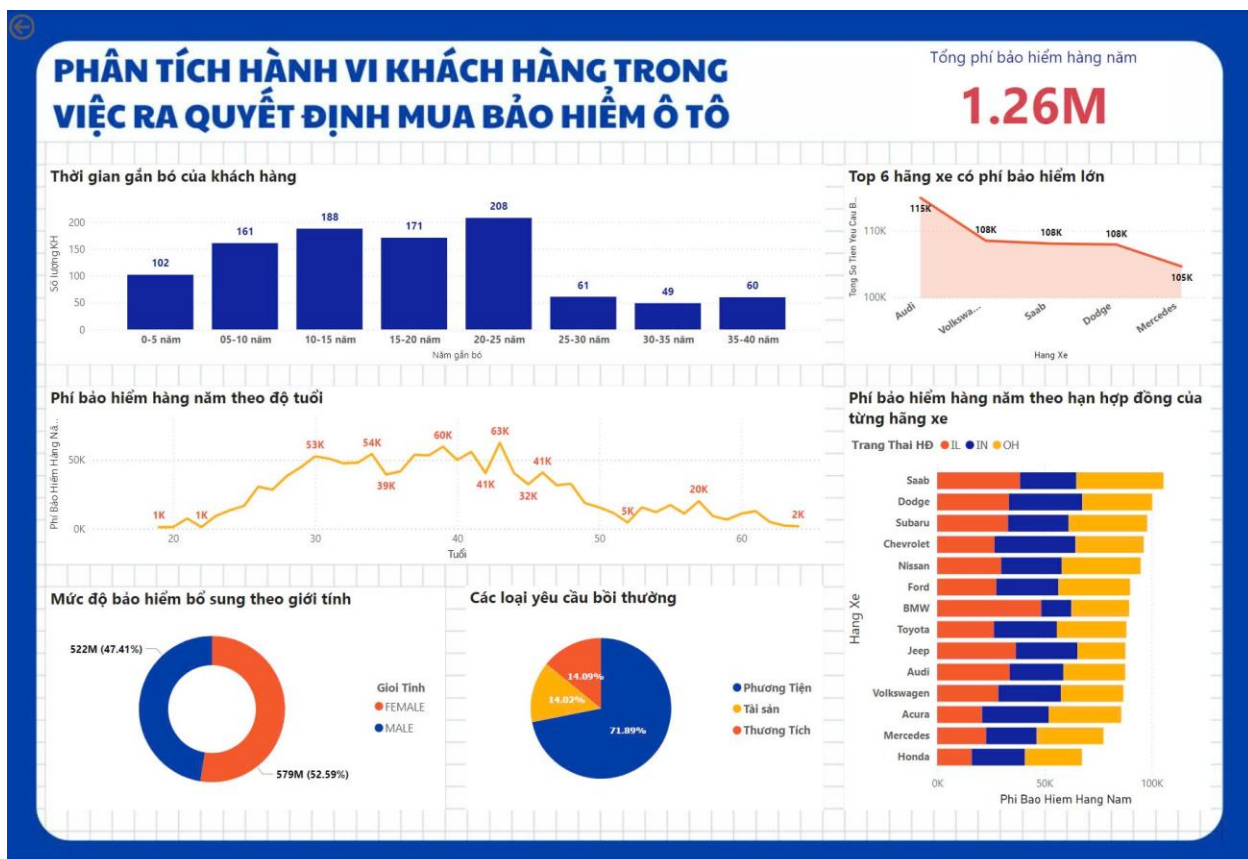


- **Power BI Desktop** là ứng dụng miễn phí được cài đặt trên máy tính với hệ điều hành Windows hoặc Linux. Ứng dụng cung cấp môi trường phát triển báo cáo mạnh mẽ. Người dùng có thể kết nối với nhiều nguồn dữ liệu, thực hiện các thao tác chuẩn bị và xử lý dữ liệu, tạo các biểu đồ trực quan và báo cáo tương tác. Đây là công cụ chính để các nhà phân tích và chuyên gia dữ liệu phát triển và thiết kế báo cáo trước khi chia sẻ chúng qua các dịch vụ Power BI khác.
- **Power BI Service** là nền tảng trực tuyến trên đám mây (SaaS) của Microsoft, nơi người dùng có thể tải lên, chia sẻ và phân phối các báo cáo được tạo từ Power BI Desktop. Power BI Service hỗ trợ cộng tác trực tuyến, cho phép các thành viên trong nhóm làm việc cùng nhau trên các báo cáo, thiết lập quyền truy cập và theo dõi thông tin từ bất kỳ đâu, bất kỳ lúc nào.
- **Power BI Mobile** là ứng dụng dành cho thiết bị di động (Android, iOS và Windows) cho phép người dùng truy cập, xem và tương tác với các báo cáo và bảng điều khiển (dashboards) trên điện thoại hoặc máy tính bảng. Với mục đích xem nhanh báo cáo trên giao diện điện thoại, Power BI Mobile hỗ trợ làm việc linh hoạt, giúp người dùng có thể cập nhật và theo dõi dữ liệu mọi lúc mọi nơi.

CÂU 2 – THIẾT KẾ, TẠO MỘT DASHBOARD VÀ GIẢI THÍCH

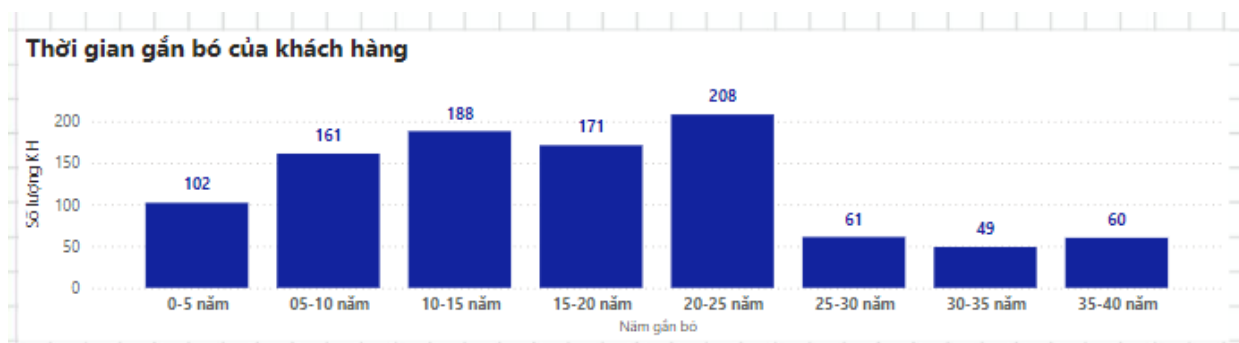
months_as_customer	Tháng Làm Khách Hàng	incident_severity	Mức Độ Sự Cố
age	Tuổi	incident_state	Bang Sự Cố
policy_number	Số Hợp Đồng	incident_city	Thành Phố Sự Cố
policy_state	Bang Hợp Đồng	incident_hour_of_the_day	Giờ Trong Ngày Của Sự Cố
policy_csl	Giới Hạn CSL Hợp Đồng	number_of_vehicles_involved	Số Lượng Xe Liên Quan
policy_deductable	Khấu Trừ Hợp Đồng	property_damage	Thiệt Hại Tài Sản
policy_annual_premium	Phí Bảo Hiểm Hàng Năm	bodily_injuries	Số Lượng Thương Tật
umbrella_limit	Giới Hạn Bảo Hiểm Phụ	police_report_available	Báo Cáo Cảnh Sát Có Sẵn
insured_sex	Giới Tính Được Bảo Hiểm	total_claim_amount	Tổng Số Tiền Yêu Cầu Bồi Thường
insured_education_level	Trình Độ Học Vấn Người Được Bảo Hiểm	injury_claim	Yêu Cầu Bồi Thường Thương Tật
insured_occupation	Nghề Nghiệp Người Được Bảo Hiểm	property_claim	Yêu Cầu Bồi Thường Tài Sản
insured_hobbies	Sở Thích Người Được Bảo Hiểm	vehicle_claim	Yêu Cầu Bồi Thường Phương Tiện
insured_relationship	Mối Quan Hệ Người Được Bảo Hiểm	auto_make	Hãng Xe
incident_date	Ngày Sự Cố	auto_model	Mẫu Xe
incident_type	Loại Sự Cố	auto_year	Năm Sản Xuất Xe
collision_type	Loại Va Chạm	fraud_reported	Báo Cáo Gian Lận

Hình ảnh minh họa các tên cột trong data được dịch sang tiếng việt.



Báo cáo phân tích hành vi khách hàng về bảo hiểm ô tô cho thấy tổng phí bảo hiểm hàng năm đạt 1.260 triệu USD, phản ánh sự quan tâm lớn của khách hàng. Khách hàng thường duy trì bảo hiểm trong thời gian dài (5-25 năm), đặc biệt là trong độ tuổi 35-40, khi họ có xu hướng bảo vệ tài sản ô tô nhiều hơn. Các hãng xe cao cấp như Audi và Subaru có phí

bảo hiểm cao hơn do liên quan đến giá trị xe. Việc mua bảo hiểm bổ sung giữa nam và nữ khá cân bằng. Yêu cầu bồi thường phổ biến nhất là bảo vệ phương tiện (71.89%), cho thấy khách hàng chú trọng vào việc sửa chữa xe khi gặp sự cố.

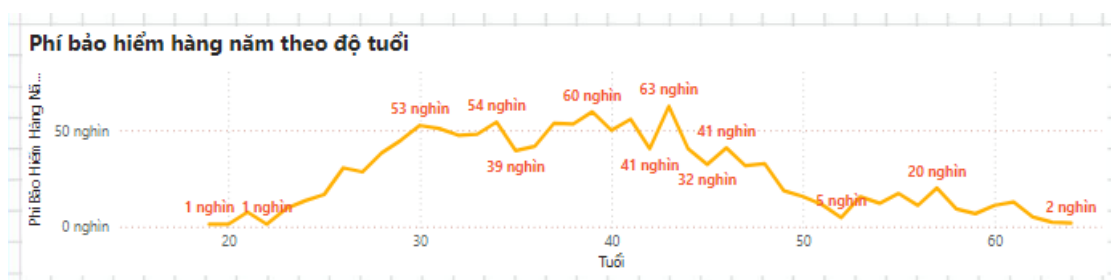


Hình ảnh minh họa biểu đồ “Thời gian gắn bó của khách hàng với bảo hiểm của công ty bảo hiểm”.

Biểu đồ "Thời gian gắn bó của Khách Hàng" thể hiện số lượng khách hàng theo từng khoảng thời gian từ 0 đến 40 năm. Nhóm khách hàng gắn bó từ 0-5 năm có khoảng 102 người, cho thấy công ty thu hút nhiều khách hàng mới. Tuy nhiên, việc duy trì và gia tăng sự gắn bó của nhóm này rất quan trọng vì họ dễ rời bỏ nếu không được đáp ứng tốt về dịch vụ và chính sách.

Nhóm khách hàng gắn bó từ 5 đến 25 năm có số lượng lớn nhất, với đỉnh cao ở nhóm 20-25 năm (208 người), cho thấy công ty có khả năng giữ chân khách hàng tốt trong giai đoạn này. Các nhóm 5-10 năm có 161 người và 10-15 năm có 188 người, cho thấy sự gia tăng liên tục của khách hàng mới. Tuy nhiên, từ 15-20 năm (171 người) có sự giảm nhẹ, nhưng sau đó số lượng lại tăng, cho thấy khách hàng có xu hướng duy trì bảo hiểm lâu dài sau 5 năm gắn bó.

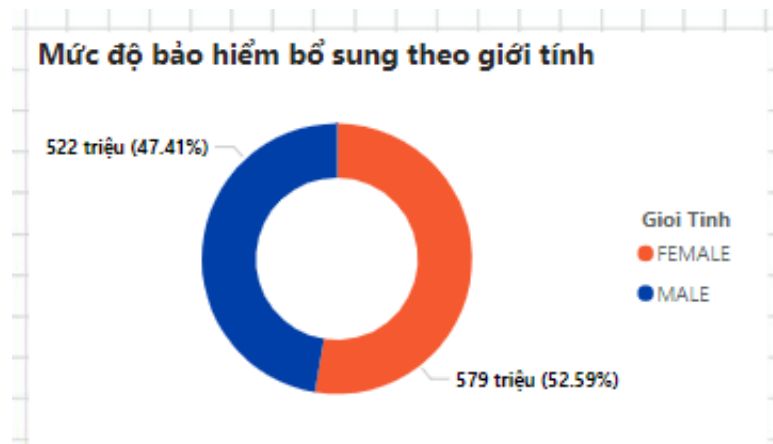
Sau 25 năm, số lượng khách hàng giảm mạnh, như nhóm 25-30 năm (61 người), 30-35 năm (49 người), và 35-40 năm (60 người). Điều này có thể do yếu tố tuổi tác, thay đổi nhu cầu cá nhân, hoặc tài chính. Khách hàng lâu năm cũng có thể muốn chuyển sang các gói bảo hiểm khác. Để giữ chân nhóm khách hàng này, công ty có thể phát triển các gói bảo hiểm tùy chỉnh phù hợp với người lớn tuổi và cung cấp ưu đãi đặc biệt cho khách hàng lâu năm.



Hình ảnh minh họa cho biểu đồ “Phí bảo hiểm hàng năm theo độ tuổi”

Biểu đồ “Phí bảo hiểm hàng năm theo độ tuổi” cho thấy mối quan hệ giữa phí bảo hiểm và độ tuổi khách hàng. Ở nhóm dưới 20 tuổi, phí bảo hiểm rất thấp, chỉ khoảng 1 nghìn USD, phản ánh rủi ro thấp vì người trẻ thường ít sở hữu xe riêng và tham gia giao thông ít. Từ tuổi 20, phí bảo hiểm tăng mạnh và đạt 53 nghìn USD ở tuổi 30, cho thấy rủi ro cao hơn khi khách hàng bắt đầu sở hữu xe riêng và tham gia giao thông thường xuyên hơn. Độ tuổi 40 đạt đỉnh phí bảo hiểm (63 nghìn USD) do tần suất và mức độ rủi ro giao thông cao nhất. Sau 40 tuổi, phí giảm dần, từ 41 nghìn USD xuống 32 nghìn USD, khi người lái xe có kinh nghiệm hơn và cẩn trọng hơn. Từ 50 tuổi trở đi, phí bảo hiểm giảm mạnh, còn 5 nghìn USD ở tuổi 50 và 2 nghìn USD ở tuổi 60, cho thấy người lớn tuổi lái xe ít và chọn phương tiện an toàn hơn.

Tóm lại, phí bảo hiểm cao nhất ở độ tuổi 30-40 do rủi ro cao, sau đó giảm dần sau 50 tuổi, phản ánh mức độ lái xe và sự cẩn trọng của người lớn tuổi.

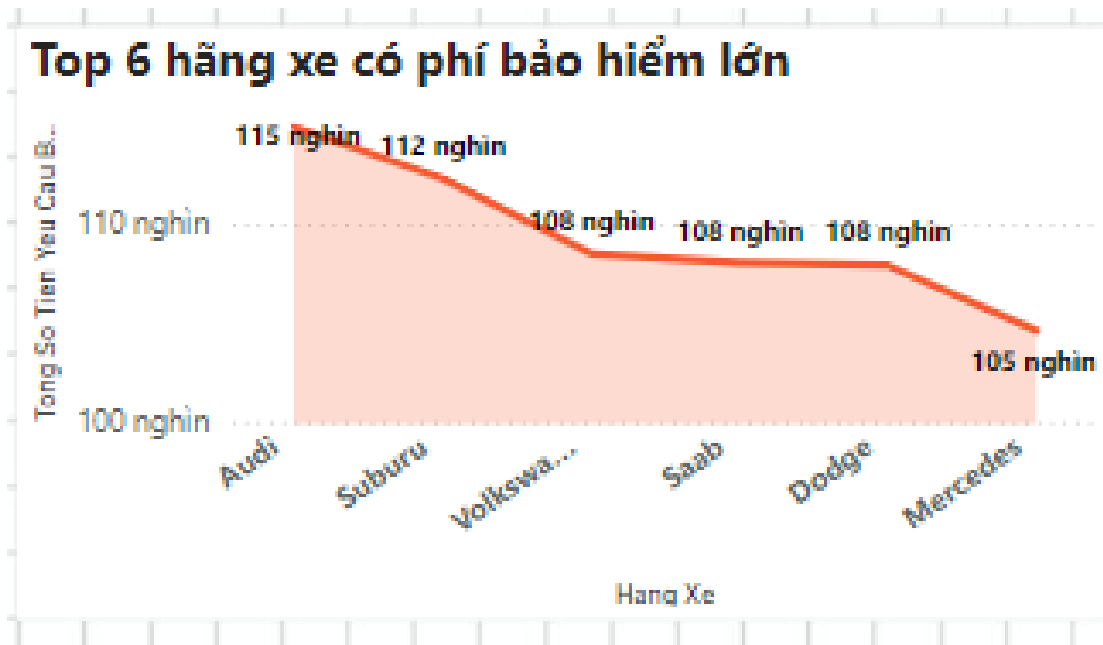


Hình ảnh minh họa cho biểu đồ “mức độ bảo hiểm bổ sung theo giới tính”

Biểu đồ “Mức độ bảo hiểm bổ sung theo giới tính” sử dụng biểu đồ tròn để thể hiện sự chênh lệch về mức bảo hiểm bổ sung giữa nam và nữ. Nữ giới chiếm tổng mức bảo hiểm bổ sung cao hơn

nam giới khoảng 57 triệu USD, cho thấy họ có xu hướng ưu tiên sự an toàn tài chính cao hơn trong các trường hợp rủi ro, như tai nạn hoặc sự cố bất ngờ khi lái xe. Nữ giới có thể chọn các gói bảo hiểm chi trả cao hơn để cảm thấy an tâm hơn trong quá trình lái xe.

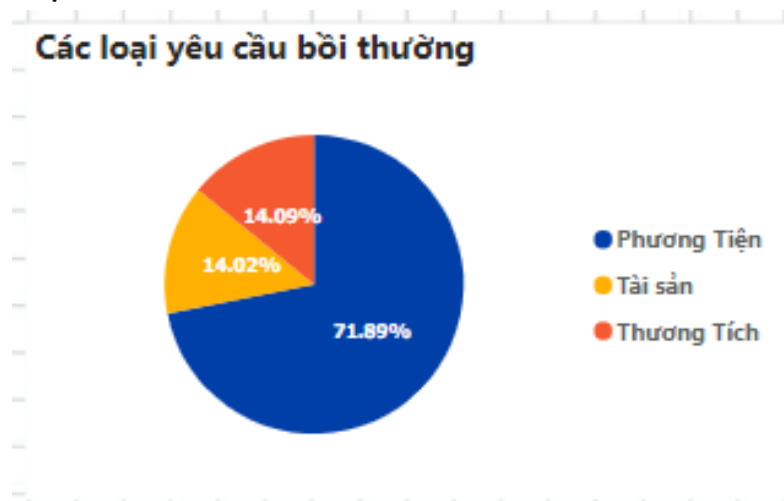
Ngược lại, nam giới có mức bảo hiểm bổ sung thấp hơn, có thể do họ tự tin hơn vào kỹ năng lái xe của mình và thường đánh giá rủi ro thấp hơn, dẫn đến việc chọn các gói bảo hiểm có mức chi trả thấp. Điều này cho thấy sự khác biệt trong cách tiếp cận về rủi ro giữa hai giới, với nữ giới thiên về sự bảo vệ tài chính nhiều hơn, trong khi nam giới chọn mức bảo vệ thấp hơn do tự tin vào khả năng kiểm soát tình huống khi lái xe.



Hình ảnh minh họa cho biểu đồ “Top 6 hãng xe có phí bảo hiểm lớn”

Biểu đồ “Top 6 hãng xe có phí bảo hiểm lớn” được trình bày bằng biểu đồ thanh ngang, hiển thị mức phí bảo hiểm của các hãng xe theo thứ tự giảm dần. Các hãng xe nằm trong danh sách này bao gồm Audi, Subaru, Volkswagen, Saab, Dodge, và Mercedes. Audi có mức phí bảo hiểm cao nhất, đạt 115 nghìn USD, tiếp theo là Subaru với 112 nghìn USD. Các hãng như Volkswagen, Saab, và Dodge có mức phí gần bằng nhau, khoảng 108 nghìn USD, trong khi Mercedes xếp cuối cùng trong Top 6 với 105 nghìn USD.

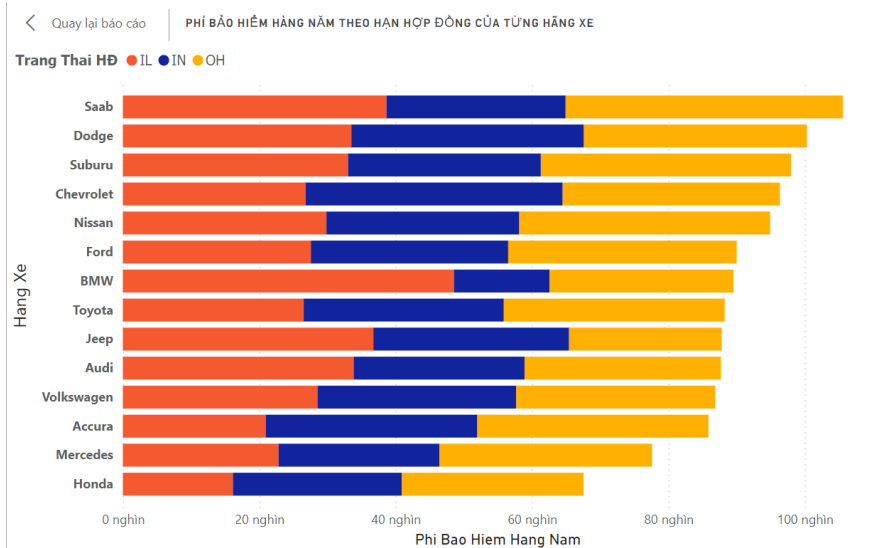
Điều này cho thấy rằng những hãng xe cao cấp thường có chi phí bảo hiểm cao hơn do các chi phí thay thế phụ tùng và sửa chữa đắt đỏ. Mức phí cao nhằm đảm bảo khả năng bù đắp chi phí sửa chữa nếu xảy ra tai nạn hay hỏng hóc, đồng thời bảo vệ giá trị của các dòng xe cao cấp. Khách hàng sử dụng các hãng xe này thường có khả năng chi trả cao hơn và sẵn sàng đầu tư vào bảo hiểm để bảo vệ tài sản của mình.



Hình ảnh minh họa cho biểu đồ “Các loại yêu cầu bồi thường”

Biểu đồ “Các loại yêu cầu bồi thường” sử dụng biểu đồ tròn cho thấy rõ tỷ lệ các loại yêu cầu bồi thường. Với những tỷ lệ trên có thể nhận thấy rằng phương tiện là loại tài sản dễ

bị tổn thất trong quá trình di chuyển và dễ dàng gặp rủi ro hàng ngày như va chạm, trầy xước, hỏng hóc, do chiếm tỷ lệ yêu cầu bồi thường cao nhất chiếm 71.89%. Tài sản và thương tích thường ít gặp hơn vì chỉ liên quan đến các tai nạn lớn và nghiêm trọng nên tỷ lệ của hai loại yêu cầu này chỉ chênh nhau 0.07%. Điều này phản ánh rằng đa số khách hàng đều ưu tiên bảo vệ phương tiện của mình, trong khi bảo hiểm tài sản và thương tích vẫn quan trọng nhưng ít yêu cầu hơn.



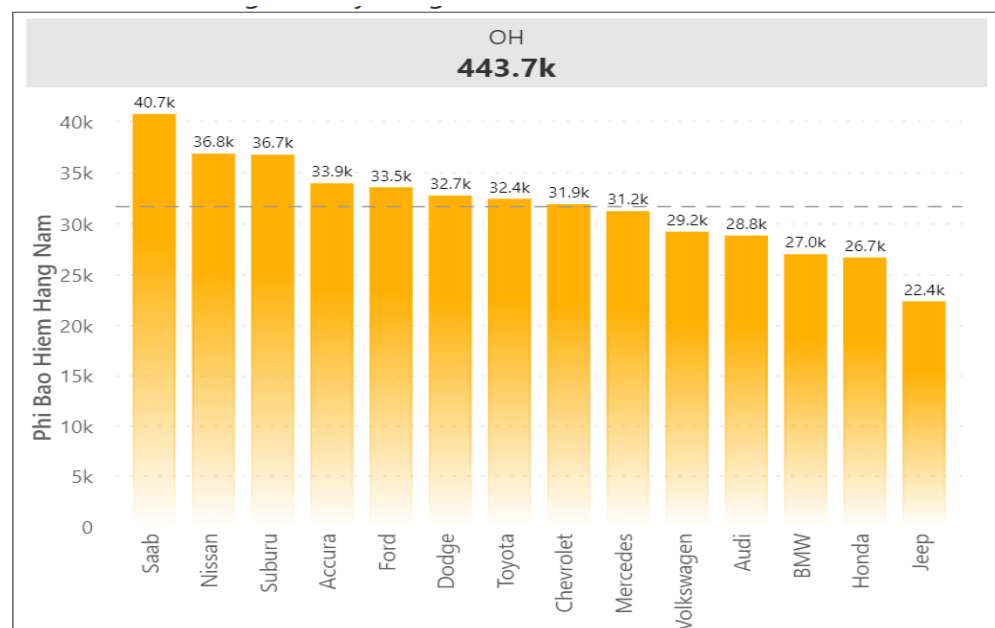
Hình ảnh minh họa cho biểu đồ “Phí bảo hiểm hàng năm theo hạn hợp đồng của từng hãng xe”

Biểu đồ “Phí bảo hiểm hàng năm theo hạn hợp đồng của từng hãng xe” được thể hiện qua biểu đồ chồng cho thấy phí bảo hiểm hàng năm

được phân chia theo các hạn hợp đồng (IL, IN, OH) của từng hãng. Trong đó, IL - Inactive/Lapsed là loại hợp đồng không còn hiệu lực, có thể do đã hết hạn hoặc không được gia hạn, IN - In Force là loại hợp đồng đang có hiệu lực và vẫn được duy trì và OH - On Hold là loại hợp đồng bảo hiểm đang tạm dừng hoặc chưa được xử lý tiếp. Để xem rõ được trạng thái hoạt động của các loại hợp đồng sẽ sử dụng biểu đồ cột.

Hình ảnh minh họa cho biểu đồ “Phí bảo hiểm hàng năm theo hạn hợp đồng theo trạng thái OH - On Hold”

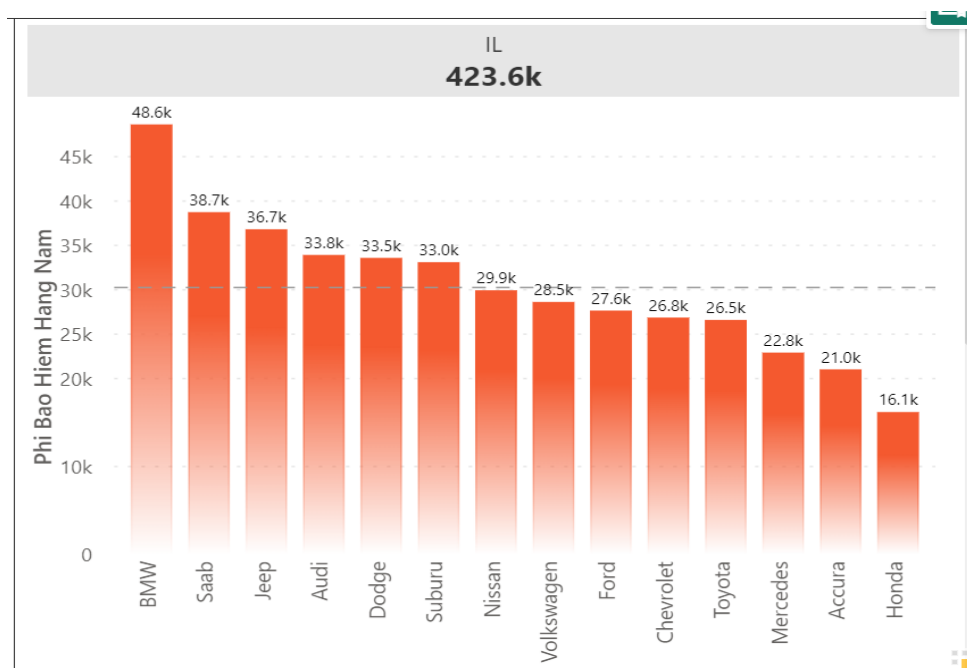
Biểu đồ “Phí bảo hiểm hàng năm theo hạn



hợp đồng theo trạng thái OH - On Hold” thể hiện mức phí bảo hiểm của các hãng xe với hợp đồng bảo hiểm đang tạm dừng hoặc chưa được xử lý tiếp, được tính bằng nghìn USD. Tổng phí bảo hiểm hàng năm cho loại hợp đồng OH là 443,7 nghìn USD. Trong đó, hãng Saab có mức phí cao nhất ở 40,7 nghìn USD, phản ánh chi phí sửa chữa và rủi ro cao liên quan đến loại xe này. Tiếp đến là Nissan và Subaru, lần lượt ở mức 36,8 và 36,7 nghìn USD, chỉ chênh lệch nhau 0,1 nghìn USD và thấp hơn Saab khoảng 4,0 nghìn USD. Điều này có thể phản ánh mức độ phổ biến và chi phí bảo trì của các hãng này.

Ở mức trung bình là các hãng như Acura, Ford, Dodge, Toyota, Chevrolet, và Mercedes, với phí bảo hiểm từ 31,2 đến 33,9 nghìn USD, thể hiện mức rủi ro và chi phí sửa chữa vừa phải. Các hãng Volkswagen, Audi, BMW, Honda, và Jeep có mức phí bảo hiểm thấp nhất, từ 22,4 đến 29,2 nghìn USD. Jeep có mức phí thấp nhất ở 22,4 nghìn USD, cho thấy rủi ro hoặc chi phí sửa chữa ít hơn khi hợp đồng mới mở.

Sự khác biệt trong mức phí bảo hiểm giữa các hãng có thể xuất phát từ nhiều yếu tố như tỷ lệ tai nạn, chi phí linh kiện, độ an toàn, và mức độ phổ biến của xe. Ngoài ra, mức độ mong muốn bảo vệ tài sản của khách hàng cũng ảnh hưởng đến phí bảo hiểm của các hãng xe này trong loại hợp đồng OH.



Hình ảnh minh họa cho biểu đồ “Phí bảo hiểm hàng năm theo hạn hợp đồng theo trạng thái IL- Inactive/Lapsed”

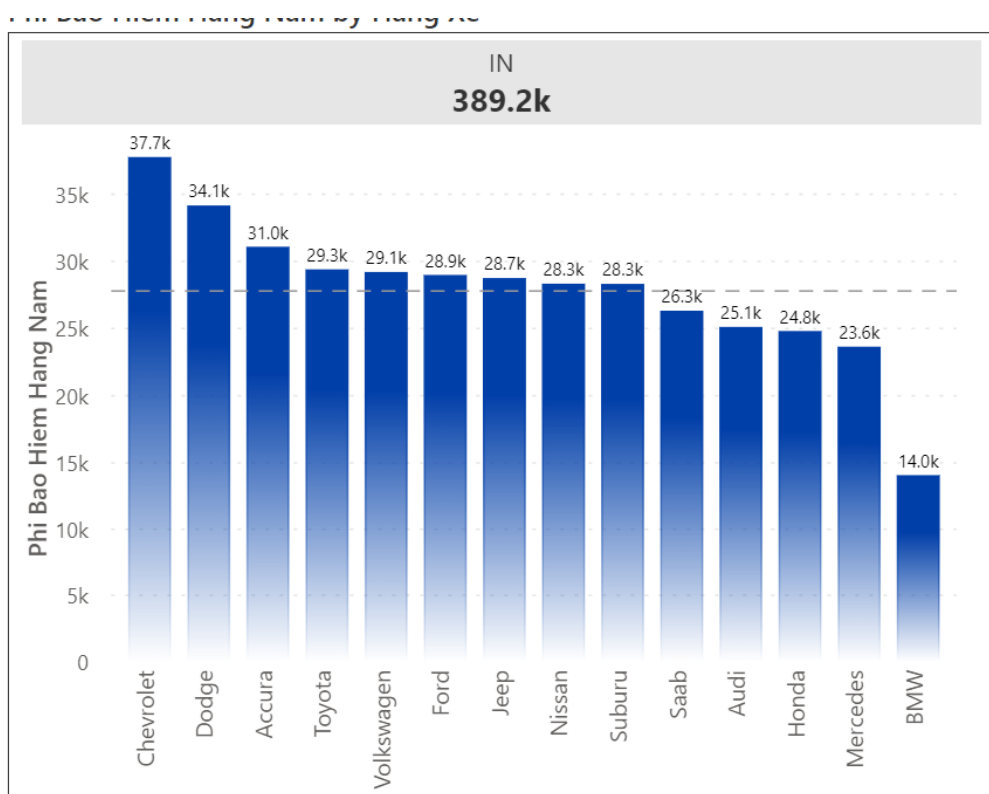
Biểu đồ “Phí bảo hiểm hàng năm theo hạn hợp

đồng theo trạng thái IL - Inactive/Lapsed” cho thấy mức phí bảo hiểm của các hãng xe với tổng phí là 423.6 nghìn USD, dành cho hợp đồng không còn hiệu lực. Hãng BMW có phí bảo hiểm cao nhất là 48.6 nghìn USD, cho thấy mức độ rủi ro và chi phí sửa chữa cao. Saab và Jeep cũng có phí cao lần lượt là 38.7 và 36.7 nghìn USD. Các hãng có mức phí trung bình gồm Audi, Dodge, Subaru, và Volkswagen, từ 33 đến 33.8 nghìn USD. Các hãng có phí thấp gồm Ford, Chevrolet, Toyota, Mercedes, Acura, và Honda (từ 16.1 đến 27.6 nghìn

USD). Những hãng có phí bảo hiểm cao nhưng không được gia hạn lại như BMW và Saab có thể do chi phí bảo hiểm quá cao nên khách hàng không muốn gia hạn lại, trong khi các hãng có phí thấp hơn như Honda có thể không cần bảo hiểm thêm vì xe ít hư hỏng.

Hình ảnh minh họa cho biểu đồ “Phí bảo hiểm hàng năm theo hạn hợp đồng theo trạng thái IN - In Force”

Biểu đồ “Phí bảo hiểm hàng năm theo hạn hợp đồng theo trạng thái IN - In Force” cho thấy tổng phí bảo

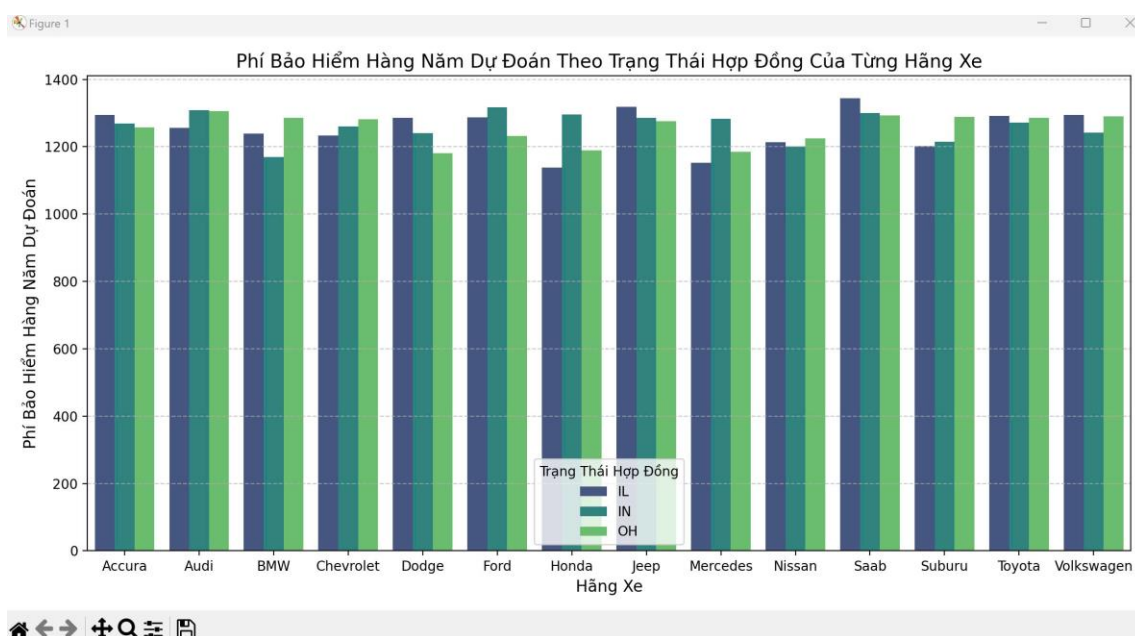


hiểm hàng năm là 389.2 nghìn USD, cho dạng hợp đồng đang thực hiện. Nổi bật nhất là Chevrolet có mức phí cao nhất (37.7 nghìn USD), cho thấy hãng này có nhiều khách hàng sẵn sàng chi trả để bảo vệ xe. Dodge đứng thứ hai với 34.1 nghìn USD. Các hãng như Volkswagen, Ford, Jeep, Nissan và Subaru có mức phí ổn định từ 28.9 đến 28.3 nghìn USD. Saab, Audi, Honda và Mercedes có mức phí từ 26.3 đến 23.6 nghìn USD, trong đó Honda và Audi nổi tiếng với độ bền. BMW có mức phí thấp nhất (14 nghìn USD), có thể do tính năng an toàn hoặc độ tin cậy. Tóm lại, các hãng có mức phí cao thường có độ rủi ro cao, trong khi nhóm trung bình có độ tin cậy vừa phải, và nhóm thấp nhất là những hãng xe bền bỉ, an toàn và tiết kiệm chi phí.

Hợp đồng OH (*On Hold*) có tổng phí bảo hiểm cao nhất, đạt 443.7k USD, cho thấy mức rủi ro cao hơn do xe ít được sử dụng hoặc bảo trì không đều. Hợp đồng IL (*Inactive/Lapsed*) với tổng phí 423.6k USD, phản ánh chi phí cao để bảo vệ các xe không còn hiệu lực hoặc không được gia hạn. Trong khi đó, hợp đồng IN (*In-Force*) có tổng phí bảo hiểm thấp nhất, 389.2k USD, cho thấy các hợp đồng này được duy trì và quản lý tốt, từ đó giảm thiểu rủi ro.

Công ty đối mặt với rủi ro như khó giữ chân khách hàng lâu dài, đặc biệt là sau 40 tuổi, khiến nhu cầu bảo hiểm giảm. Hơn nữa, các hãng xe có phí bảo hiểm cao có thể khiến khách hàng ngại ký hợp đồng dài hạn. Nếu tỷ lệ yêu cầu bồi thường quá cao, công ty sẽ gặp khó khăn trong việc cân bằng chi phí và lợi nhuận. Để khắc phục, công ty nên xây dựng chương trình tri ân hoặc ưu đãi cho khách hàng lâu năm, nhằm tăng sự hài lòng và khuyến khích họ tiếp tục sử dụng dịch vụ. Đồng thời, cung cấp các phương thức thanh toán linh hoạt để giảm áp lực tài chính cho khách hàng.

CÂU 3 - SỬ DỤNG MỘT MÔ HÌNH AI



Hình ảnh minh họa cho biểu đồ thông qua mô hình AI - Random Forest

Sử dụng mô hình AI Random Forest. Random Forest là một mô hình học máy mạnh mẽ, hiệu quả cho cả bài toán phân loại và hồi quy, đặc biệt trong các bài toán có tính phi tuyến cao. Mô hình này thích hợp hơn khi phải xử lý với nguồn dữ liệu lớn. Nó có thể xử lý tốt dữ liệu phân loại như các cột 'Hãng Xe' và 'Trạng Thái Hợp Đồng' vì mỗi nhánh sẽ tự đưa ra quyết định dựa trên các đặc trưng của dữ liệu mà không cần phải xử lý quá nhiều dữ liệu. Phương pháp ensemble giúp kết hợp kết quả của nhiều cây dữ liệu và được huấn luyện trên các tập con dữ liệu ngẫu nhiên. Nhờ vậy mà nó có khả năng giảm thiểu hiện tượng overfitting. Ngoài ra, còn sử dụng GridSearchCV để tối ưu hóa các siêu tham số của Random Forest - cải thiện hiệu suất mô hình. Thêm một điểm cộng dành cho Random Forest là có khả năng hoạt động tốt ngay cả khi dữ liệu có một số giá trị thiếu hoặc có outliers, giúp cho mô hình dự đoán ổn định và đáng tin cậy hơn. Với những ưu điểm như vậy, Random Forest là một lựa chọn hợp lý cho bài toán thể hiện Phí bảo hiểm hàng năm theo trạng thái hợp đồng của từng hãng xe thông qua biểu đồ.