FORECASTING THE AUTO MARKET.
GOOGLE TRENDS AS A FORECASTING
TOOL IN UKRAINE.

by

Sviatoslav Hohol

A thesis submitted in partial fulfillment of the
requirements for the degree of

MA in Business and Financial Economics

Kyiv School of Economics

2024

Thesis Supervisor: ______________________________ Professor Maksym Obrizan

Approved by
Head of the KSE Defense Committee, Professor [Type surname, name]

Date ______________________________________

ACKNOWLEDGMENTS

The author would like to express gratitude to his family for their strong support throughout his master's studies. Their encouragement and understanding were a source of strength during this journey.

A special thank you goes to my supervisor, Professor Maksym Obrizan, whose concise and insightful advice served as a guiding light and a food for thought. His feedback improved the quality of my work a lot.

I also want to express my appreciation to the faculty at the Kyiv School of Economics for their dedication and expertise. Their commitment to education and the wealth of knowledge they shared influenced my academic growth a lot.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

CCF - Cross-correlation function

EV - Electric vehicle

GT – Google Trends

IEA - International Energy Agency

LPG - Liquefied petroleum gas

MIA - Ministry of Internal Affairs of Ukraine

VIN - Vehicle identification number

CHAPTER 1. INTRODUCTION

The automotive industry is one of the most significant sectors in the global economy. It drives economic growth worldwide. The advent of cars in the XVIII century changed society and the economy (Briney, 2020). Today we cannot imagine our life without transport and logistics. Its impact is multifaceted and influences not just economic metrics but also social norms and the environment.

Another important economic driver is access to the Internet. Online search plays a crucial role in consumer behavior because it provides immediate access to product information, reviews, and price comparisons between different marketplaces. It also helps to discover trends and make informed purchasing decisions for consumers.

In the developing automotive industry understanding of market dynamics is important for manufacturers and dealers to optimize inventory management and production. Traditional market research methods often rely on historical data and lag behind real-time market and sentiment changes. The motivation of a study is to review and establish a connection between online users' sentiments and real-world economic outcomes, such as the purchase of a car. Digital platforms store a lot of data on consumer behaviors and preferences. Google Trends, in particular, provides real-time insights about consumers' search interests. The key idea of the paper is to evaluate the potential of Google Trends indices as an indicator of demand in the Ukrainian car market.

Early detection of search behavior changes can help market participants respond on time and efficiently engage with clients. So, car dealers and sellers can smoothly prepare their inventory for possible future demand shifts. Moreover, Google Trends is a free and easily available tool that provides data across various time and geographical regions in contrast to the usually long-term conventional research and forecast tools. Usage of this approach can be interesting also for companies that seek to reduce research-associated costs.

So, in this work, we strive to define a degree of how Google Trends are related to real sales in the Ukrainian car market and the time lag appropriate for the specific brands, models, or price segments. It is worth to note that we do not aim to state about existence of causal relationships between the time series of rising search interest and subsequent increase in sales. Instead, the primary goal is to evaluate the reliability of Google Trends as an auxiliary tool in the automotive decision-making process. Although not a substitute for traditional forecasting models, our approach can serve as valuable indicator observed in a real-time. Therefore, this tool can help car market participants to act proactively rather than reactively and improve overall responsiveness to market conditions.

CHAPTER 2. INDUSTRY OVERVIEW AND RELATED STUDIES

2.1 Ukrainian Car Market Overview.

Nowadays, Ukraine does not take a high place in terms of automobilization level. Although the Ukraine area is the biggest in Europe, the quantity of cars per 1000 people is almost the lowest in the region.

Table 1. Autobilization level of Ukraine and neighbors.

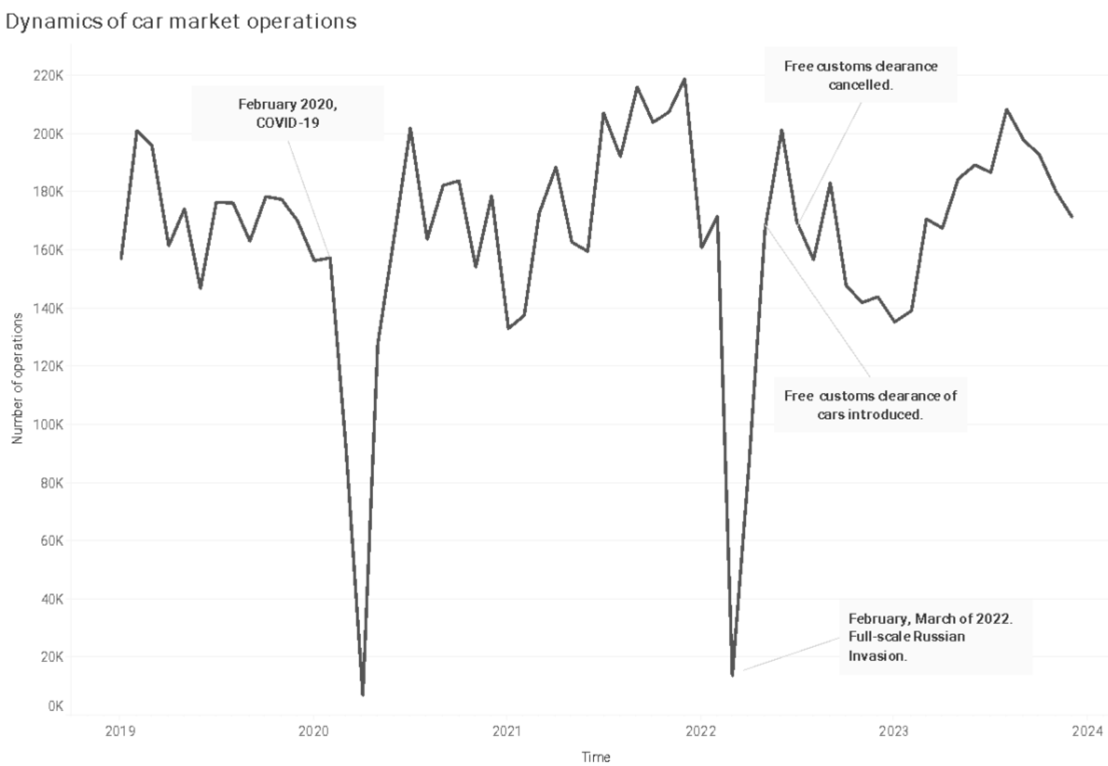
Country	Motor vehicles per 1000 people
United States	872
Poland	817
Germany	633
Romania	465
Hungary	478
Ukraine	204

Source: Countryeconomy online portal (2021).

Table 1 presents some countries and their vehicles per 1000 people index. The United States, with almost 900 motor vehicles per 1000 people, is a world leader. Ukraine's motor vehicle index of 204 cars per 1000 people is significantly lower than in other countries listed in the table. This low level of vehicle ownership can be explained by economic, social, and infrastructural factors. These factors impact the affordability and necessity of private vehicle ownership in Ukraine. It is also significantly influenced by economic policies, including high tax rates on vehicle imports and so-called "protectionist" measures in the past. On the one hand, products of the national automotive industry were not very attractive to consumers in comparison with imported vehicles. On the other hand, import tax rates were relatively higher than in neighboring countries or well-developed economies. These factors made car ownership less accessible for Ukrainians. (Dmitriyev, Shevchenko 2017)

Also, there is no tradition of old car utilization in Ukraine. There is a legislation base targeted for auto recycling programs passed by the Ukrainian Parliament, but recycling culture is in its formative stages. The law, intended to rid our country of outdated automotive relics, is not being effectively implemented, and old car owners do not have an incentive to participate in utilization or recycling programs. (Venzhega et al, 2016) That is why it is not efficient to evaluate market dynamics and volume by only the number of cars. The reason is that the analysis will be overwhelmed by old, non-working vehicles and irrelevant records. So, we suppose that it will be better to evaluate market dynamics by the number of registrations over time.

Figure 1. Dynamics of car market operations.



Source: Ministry of Internal Affairs, register of vehicles 2019-2023.

Figure 1 and Table 2 illustrate the dynamics of car market operations in Ukraine from 2019 through 2023, including the annual growth rates for each year. Each record reflects registration action between new and old car owners.

Usually, number of operations fluctuates from 1.7 to 2.1 million per year. The time series is visually non-stationary due to a few shocks, so there are a few noticeable declines in the graph. The first one is due to the COVID-19 pandemic, which led to a significant drop in yearly market activity at almost 15% of a pre-pandemic level. As most of the people stayed at home and there were quarantine restrictions, demand for cars dropped. The second huge deviation of almost -21% occurred in the year 2022 when the full-scale russian invasion happened. In February-March 2022 number of operations dropped by 90-95% in comparison with the previous months. This indicates that the market was almost inactive. It is also important to note that after each fall, the shock was followed by significant growth and recovery.

Table 2. Dynamics of the Ukrainian Vehicle Market.

Year	2019	2020	2021	2022	2023
# of operations	2079481	1771329	2201307	1745908	2124732
Growth rate		-14,8%	24,3%	-20,7%	21,7%

Source: Ministry of Internal Affairs, register of vehicles 2019-2023.

In April 2022, the government canceled import duties on cars. As a result, there was a significant recovery in the market. The number of transactions and imported cars increased significantly. However, due to a lack of revenues for the road fund, the import duty on cars was returned in July 2022. After that, the market returned to pre-war levels of activity. A trend line shows returns as insignificant (due to the non-stationarity), but it is still a positive result of a 6,77% average growth rate for 2019-2023 years.

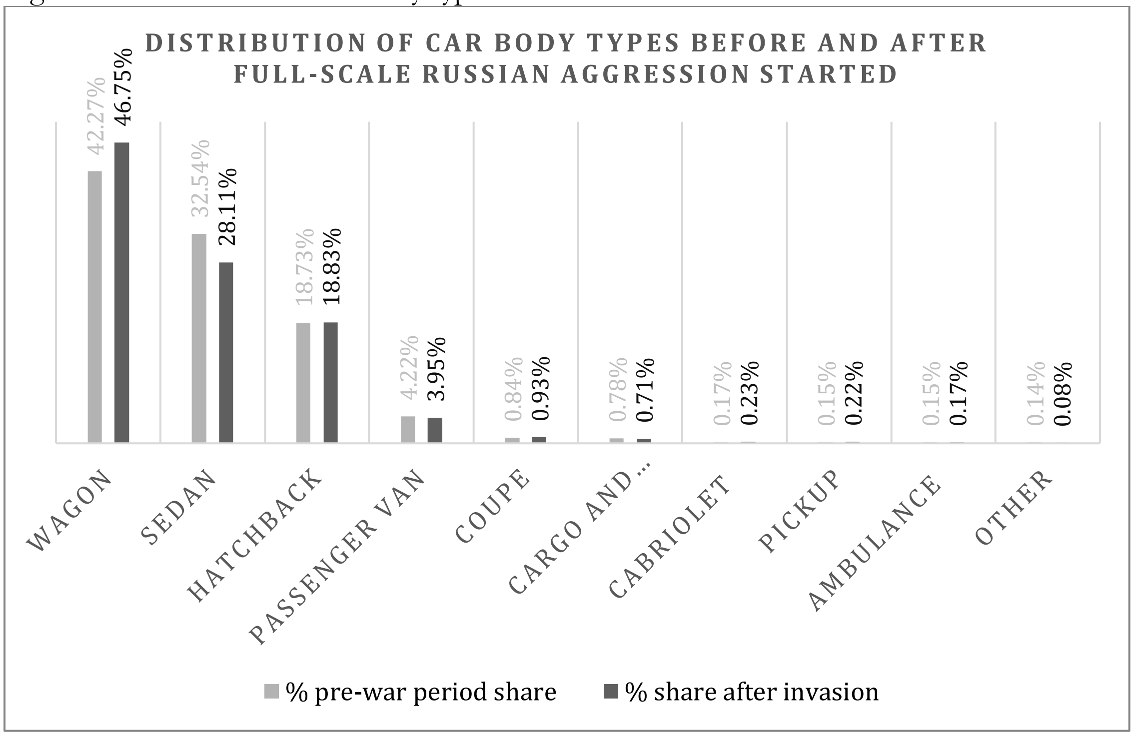
Consumer preferences of Ukrainians regarding cars, reasons for purchasing and sources of information are presented in the paper of Zlatova (2014). According to this paper, the most significant reason for car purchase, mentioned by 68% of participants, is the emergence of new needs. The breakdown of the old car and extra funds are both the reasons for 36% of buyers. Additionally, 46% of respondents attribute their purchase to the moral and physical obsolescence of their previous vehicle. Less influential factors are the release of new car models (18%), maintaining prestige (14%), new car technologies (11%), and fashion trends (7%).

Also, in this work, a survey was conducted in Ukraine on the sources of information before buying a car. So, the most popular source is the Internet, with 61% of respondents relying on it. Friends and acquaintances are also significant and mentioned by 36% of participants. Other sources include specialists (29% share), magazines (14% share), and car dealerships (11% share). Traditional media sources like television, billboards, and radio are almost insignificant, with a 0-1% share in the survey.

Other preferences of Ukrainian consumers about passenger cars can be revealed by analyzing the database of vehicle records of the Ministry of Internal Affairs for 2019-2023 years. The dataset consists of records about all vehicle registrations. Figure 2 below represents the percentage share of various vehicle body types. The most popular vehicle body type in Ukraine are wagons. They have around 44% of the market share. Wagons are so popular due to their versatility, big cargo space, and practicality. It makes them one of the best options for families. Sedans choose 31% share of consumers, and hatchbacks account for 19% share. Passenger vans represent 4.1%, and coupes have a very low share of up to 0.99% of the market. Other vehicle types, such as cargo and passenger vans, cabriolets, pickups, and ambulances, have minimal shares, each below 1%.

We intentionally split the study period into two subgroups - before and after russia's full-scale invasion in Ukraine in February 2022 to evaluate potential consumer preferences shift due to the military demand. The most interesting thing is that due to the army demand, the share of pickups and ambulance cars increased. Around 50% increase for pickups and 12% increase for ambulance cars. It is worth noting that our dataset includes only operations made by individuals (volunteers) but does not include military or government imports. Moreover, it does not include vehicles imported as humanitarian aid and transferred to the balance sheet of military units because this information is not publicly available.

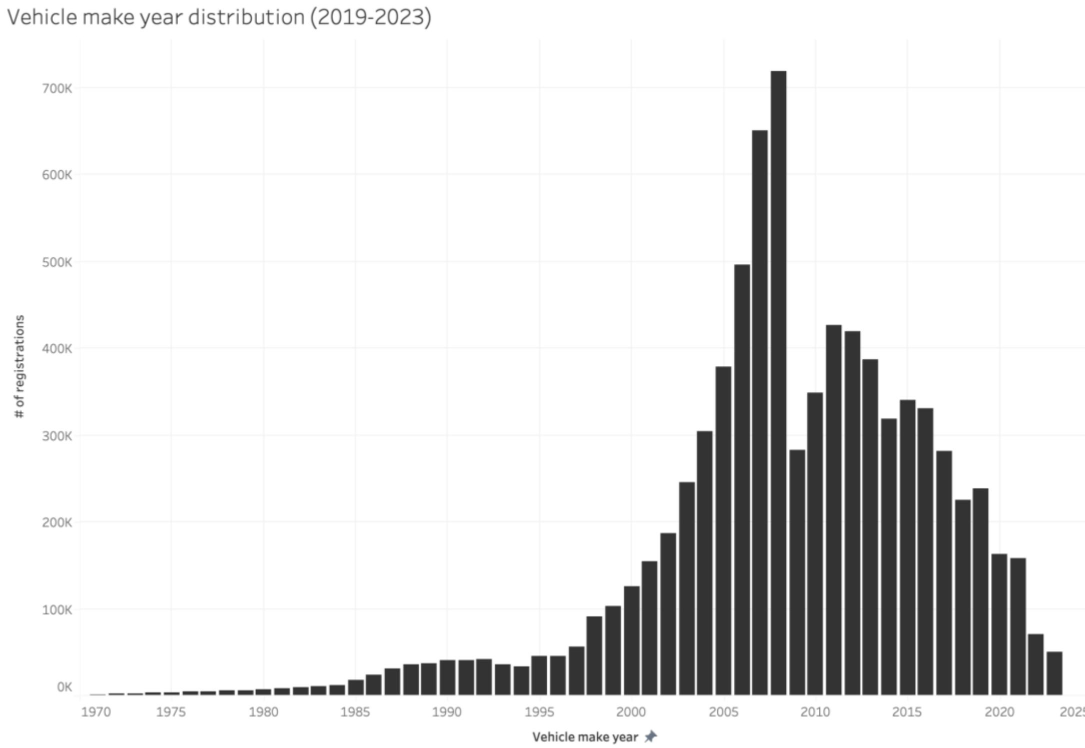
Figure. 2 Distribution of car body types 2019-2023 in Ukraine.



Source: Ministry of Internal Affairs, register of vehicles 2019-2023.

Another indicator that can characterize the features of the Ukrainian car market is the average age of a car. Figure 3 below displays the distribution of vehicle production years for cars registered between 2019 and 2023 and offers a depiction of which production years dominate the Ukrainian car market. The most popular cars in the study period are those produced in 2007-2008. Intuitively, vehicles of this age are the most popular because of their cheaper price. By the way, cars that are over a decade old are unlikely to be very reliable. The decline in registrations for newer models could be connected to the economic challenges faced by the country that made newer vehicles less accessible to the population. So, the presence of older cars may also imply a lower rate of new car purchases that are driven by financial constraints or a preference for imported used vehicles over new ones.

Figure. 3 Vehicle make year distribution 2019-2023 in Ukraine.

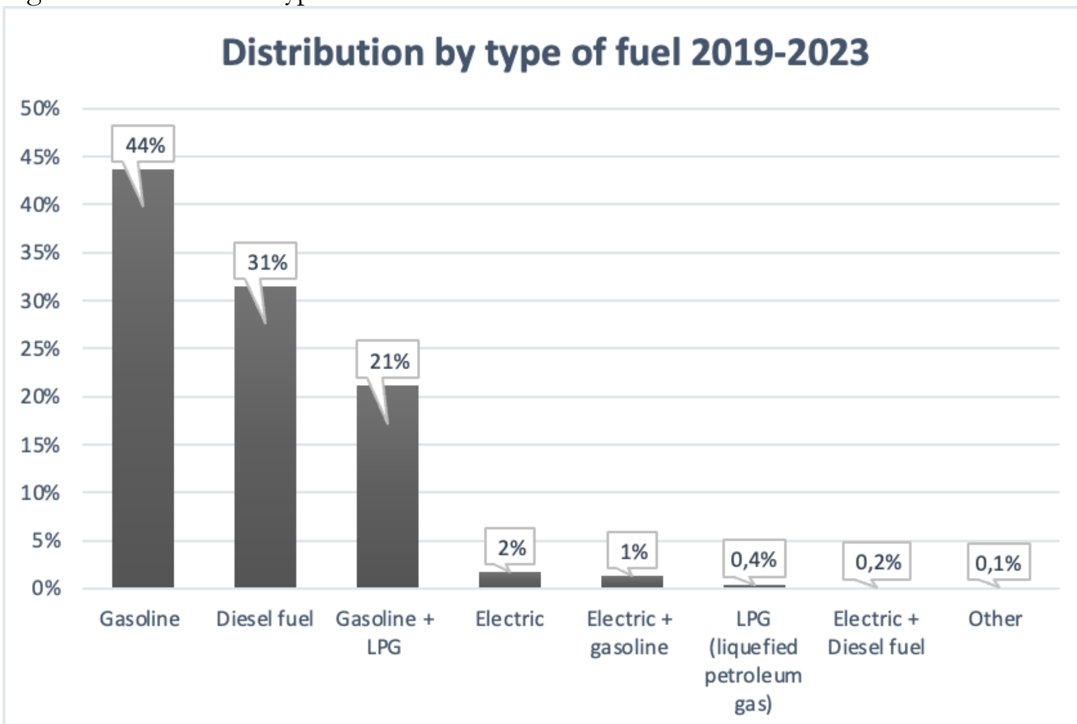


Source: Ministry of Internal Affairs, register of vehicles 2019-2023.

In conclusion, the Ukrainian car market with sufficiently old cars differs significantly from the markets of developed economies such as the US and Germany, where new cars almost completely dominate over older ones. The reasons for this are the peculiarities of the Ukrainian economy, low level of income, high import duties and the almost complete absence of favorable financing instruments for car purchases: loans, installments, leasing, etc.

Preferences about cars of Ukrainian consumers can also be revealed through the choice of fuel type of vehicle. This choice is influenced by factors like fuel availability, cost, environmental regulations, and consumer preferences. Figure 4 presents the distribution of vehicle fuel types based on registration data from 2019 to 2023.

Figure. 4 Vehicle fuel type distribution 2019-2023 in Ukraine.

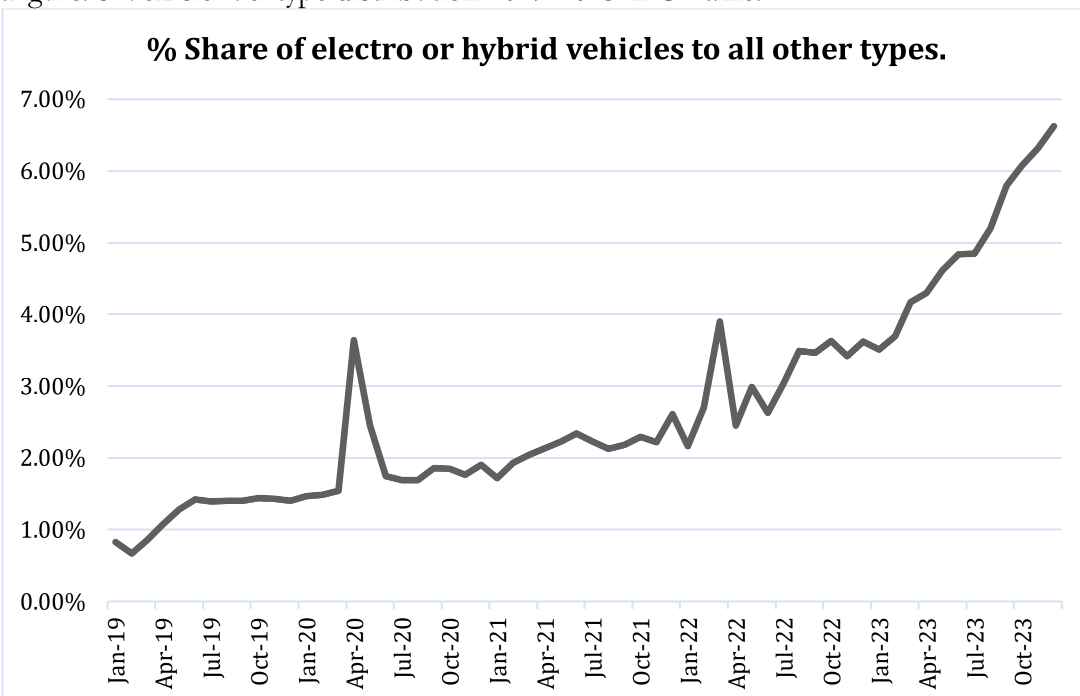


Source: Ministry of Internal Affairs, register of vehicles 2019-2023.

Gasoline is the most common fuel type, with 44% of registered cars during the study period. Diesel fuel follows at 31% share. It is worth noting that regular diesel and petrol engines are the most commonly used worldwide. Their popularity is explained by their relative technological simplicity and maintainability. Hybrid and electric engines, despite their increased efficiency, are more expensive to maintain, so fewer people are willing to own cars with such engines. Also, a relatively high share of private vehicles with a combination of gasoline and LPG reflects a shift towards more cost-effective and potentially cleaner alternatives. Fully electric cars, despite growth in popularity, represent a small share of 2%. While there is a significant trend and interest in greener technologies, the transition in Ukraine is very gradual. Other fuel types, including electric-gasoline hybrids and LPG alone, account for less than 1% each, so it highlights the dominance of traditional fuels despite alternatives.

Nevertheless, according to the IEA report in 2023, nearly one in five cars sold was electric, with sales reaching almost 14 million. 95% of these sales were in China, Europe, and the U.S. In Ukraine, the electric vehicle market is still in its early stages, and shares of EV and hybrid vehicles is much smaller. However, the adoption of electric and hybrid vehicles was steadily increasing each year, as shown in figure 5 below.

Figure. 5 Vehicle fuel type distribution 2019-2023 in Ukraine.

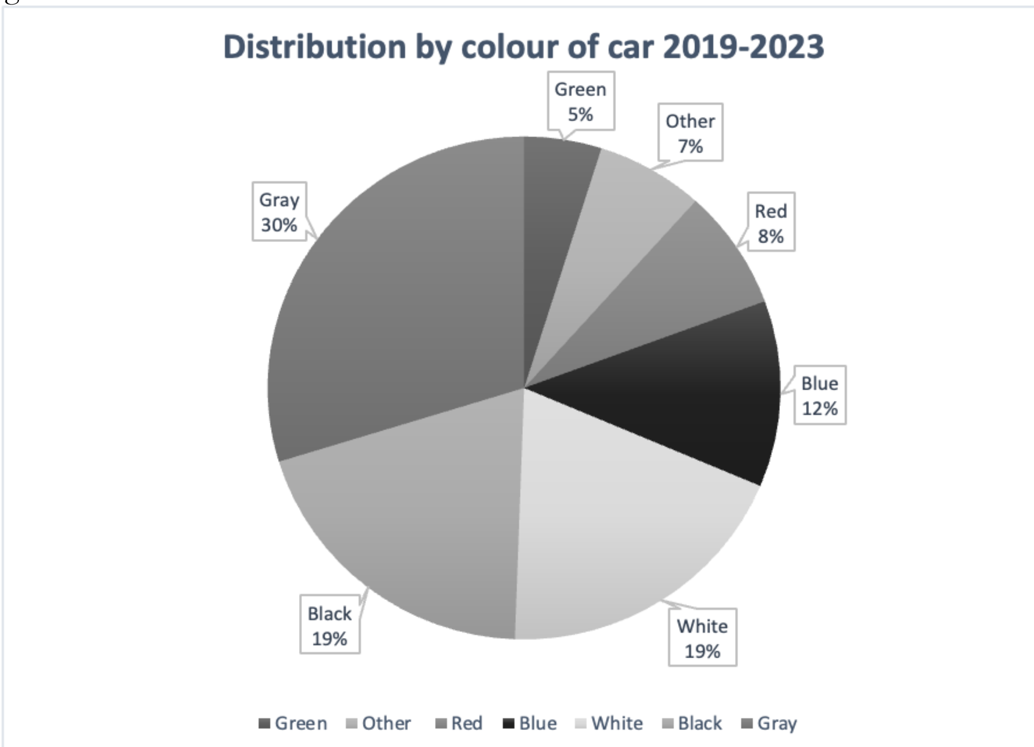


Source: Ministry of Internal Affairs, register of vehicles 2019-2023.

By the end of 2023, electric and hybrid vehicles reached almost 7% share to all vehicle sales in the country. It confirms that, despite all the challenges, there is a growing interest in EV and hybrid vehicles in Ukraine.

Another interesting preference that could be revealed is the choice of the color of the car. It can be influenced by cultural preferences, resale value, and sometimes perceptions of safety. Some people would like to avoid black or dark cars because they heat up faster, and this body is more difficult to maintain and keep clean.

Figure. 6 Vehicle color distribution 2019-2023 in Ukraine.



Source: Ministry of Internal Affairs, register of vehicles 2019-2023.

Figure 6 reveals that gray is the most preferred car color and has around 30% share of all registrations. Then goes black and white, each with a 19% share.

2.2 Google Trends validity

Google Trends is a tool that provides a snapshot of the search queries submitted to Google by users. It is anonymous and presented in an aggregated way. It allows all users to track interest in specific topics globally, at a region or city level. Indexes cover search activities in real-time and include information since 2004 and up to 72 hours before the query.

Indices presented in Google Trends are normalized and scaled from 0 to 100. Normalization is done by dividing each data point by the number of total searches in a given region and timeframe to ensure that variations in search volume don't skew the results. This method allows for comparisons across different periods and regions, even if the absolute search volumes can vary significantly (Google support, n.d.).

The validity and prediction power of Google Trends are already discovered in a variety of papers. Varian and Choi (2009) found that Google Trends can be considered as a valid indicator for predicting demand shifts. The main focus of the work was partially different from ours and insisted on the improvement of the forecast by using this tool. Methodologically, Google Trends indices were included as an additional variable to the conventional seasonal AR and fixed effects models. As a result, some models performed in a better way. While the improvement in prediction accuracy varies, it was quite significant in some cases. For example, the inclusion of additional GT variable led to an 18% increase in the accuracy of predictions of models about automotive examples and a 12% enhancement for housing demand changes.

Another work by Kinski (2016) focuses on the analysis of the relationship between internet searches for car models and new vehicle sales data across Germany and the United States. The author`s approach was firstly to use cross-correlation analysis between two time-series data of actual sales and Google Trends indices. The purpose was to identify the appropriate time lag between Google Trends changes and the demand shifts for particular car models. Then, with adjustment for optimal lag simple OLS was done. The main findings proved a significant relationship between GT and actual sales. However, the explanatory power varies for different models from totally insignificant to 10-20%. So, only a small portion of the changes in new car model sales could be explained by fluctuations in the Google Search query index in their work.

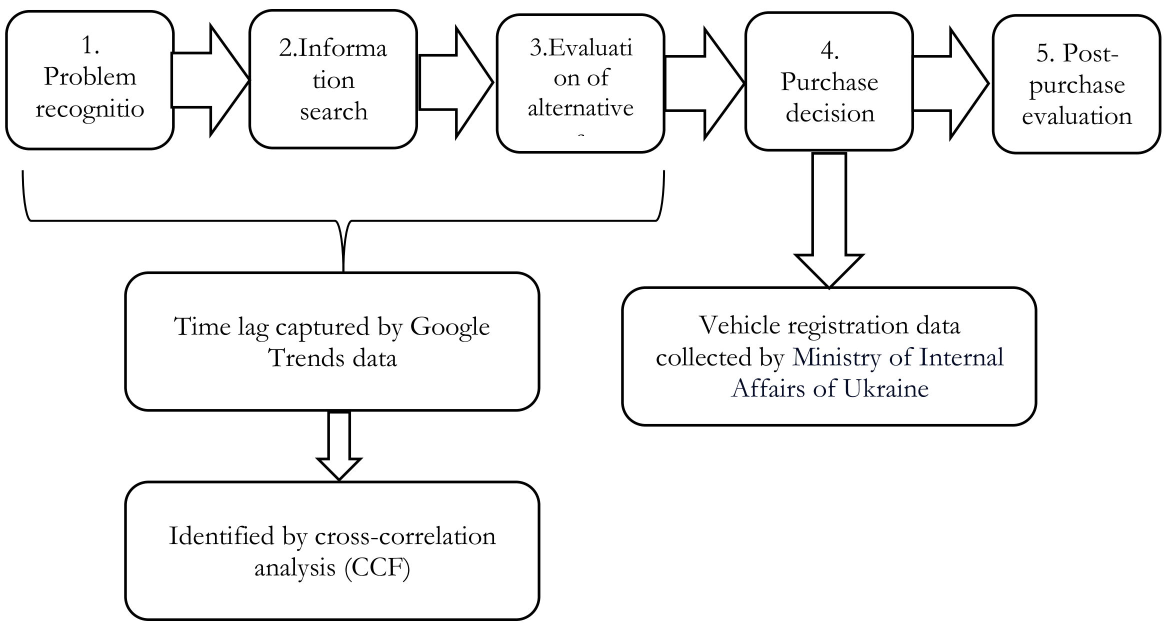
Almost a similar approach was used in Wijnhoven, Fons & Plant, Olivia's (2017) work where the Netherland automotive market was investigated. Except for Google Trends indices in their models, authors also took into account other measures of online sentiment. Metrics like the Positivity to Negativity (PN) ratio, negative and positive mentions number about particular car models were web-scrapped and added to the models. The obtained results also proved a significant connection between online sentiment and new car sales. Moreover, there were no differences in the strength of the correlations between low and high-priced models.

This work can contribute to both academic and practical fields by demonstrating the utility of Google Trends data in predicting economic variables like car sales. From an academic point of view, research steps aside from conventional ways of demand prediction and offers an alternative approach. Existing literature does not cover the Ukrainian automotive market, so it also gives an opportunity for analysis across different countries.

From a practical point of view, the obtained results can be interesting for automotive market participants, such as dealers and auto consultants. It can help dealerships adjust inventory management, marketing, and sales strategies in response to consumer interest trends.

According to a Google survey (2023), 92% of car buyers all over the world conduct their research and compare alternatives online before purchasing. The main reason is that it allows them to spend less time and effort on visiting dealerships. It also gives them an opportunity to consult themselves more easily.

Figure 7. Data identification on the stages of the purchase process based on Kotler.

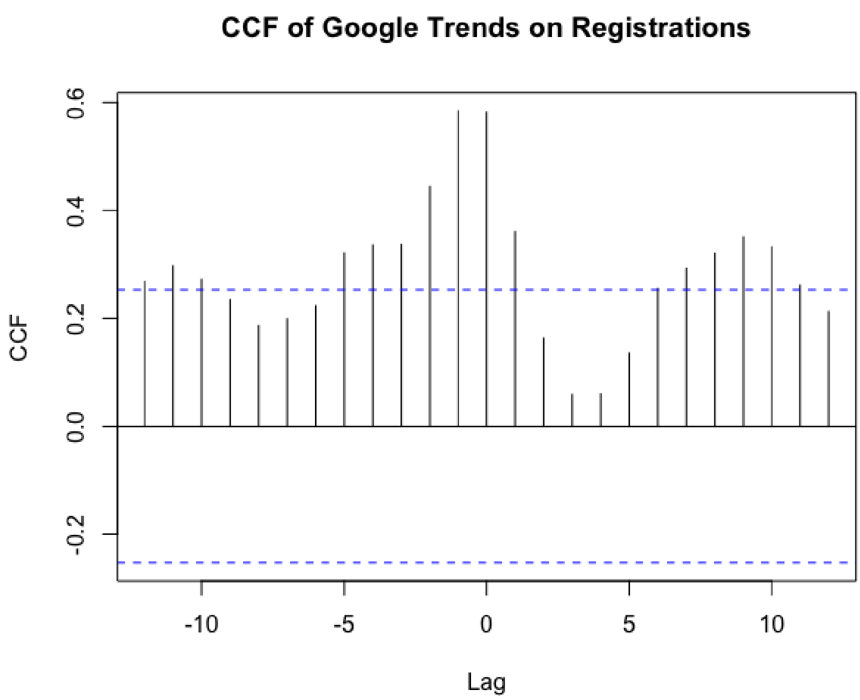


Source: Kotler (2012), author.

Figure 7 presents a simplified consumer decision-making process model developed by Cotler (2012), which outlines the typical stages a consumer goes through before and after making a purchase. Stages 1-3 are related to the research about needed products and, in our case, are captured by Google Trends data. This linear process starts with the problem recognition stage, where the consumer identifies a need for a vehicle. Then goes an information search stage, where the consumer looks for information about potential options. The third stage is the evaluation of alternatives, in which different options are compared to determine the best choice. The first three stages create a time lag between the research process and the purchase decision. Then it leads to a purchase decision, the stage at which the consumer decides to buy a particular car.

The optimal time lag for each particular model can be defined using cross-correlation analysis. By definition, “cross-correlation is a measure of similarity of two waveforms as a function of a time lag applied to one of them. This is also known as a sliding dot product or inner product. It is commonly used to search a long-duration signal for a shorter, known feature.” Rabi (2024). In practice, it is used to measure the degree to which two series (time series) are correlated at different lags. Essentially, it quantifies how well the movements in one series can predict or are related to the movements in another series over different time intervals. Results of this stage of analysis can prove the hypothesis that there exists a measurable time lag between rising search interest in a specific vehicle model and the corresponding increase in its sales.

Figure 8. Example of cross-correlation function usage for Volkswagen Golf.



Source: Google Trends, Ministry of Internal Affairs, register of vehicles 2019-2023.

Figure 8 above presents the example of the result of CCF analysis between the Google Trends index and actual sales of one of the most popular vehicles in Ukraine - the Volkswagen Golf. Y-Axis (CCF) represents the cross-correlation coefficient at different lags. The values can range from -1 to 1, where 1 confirms a perfect positive correlation, -1 means a perfect negative correlation, and 0 confirms that there is no correlation at all. X-Axis (Lag) indicates the lag between both time series. A lag of 0 means the two series are aligned in the same period. Positive lags indicate the number of time units that Google Trends data leads registrations. Negative lags indicate the number of time units that registrations lead to Google Trends data.

It has common sense to take into consideration only negative lags for this example, as online interest usually occurs before the purchase decision. Moreover, positive lags are mostly insignificant. From the given analysis, we can state that the best option for further analysis is to adjust the GT index for one month back.

The second hypothesis can be presented as an increase in Google Trends search index for specific vehicle brand and model correlates with the subsequent rise in actual sales. So, after matching 2 time series of Google Trends and actual sales of a particular model, linear regression can be used to define the strength of the relationship between Google Trends and actual car sales of particular car models.

Suggested model is presented below as equation 1.

$$Sales_t = Google\ Trends_t + sales_{(t-1)} + \ sales_{(t-12)} + invasion + free\ trade + \in \tag{1}$$

Where:

$Sales_{(t-1)}$ - is sales of a particular vehicle model adjusted for lag 1.

$Sales_{(t-12)}$ - sales of a particular vehicle model adjusted for 12 months back lag in order to adjust the model for seasonality.

$Google\ trends_{(t-n)}$ - GT index adjusted for optimal lag (if lag is found).

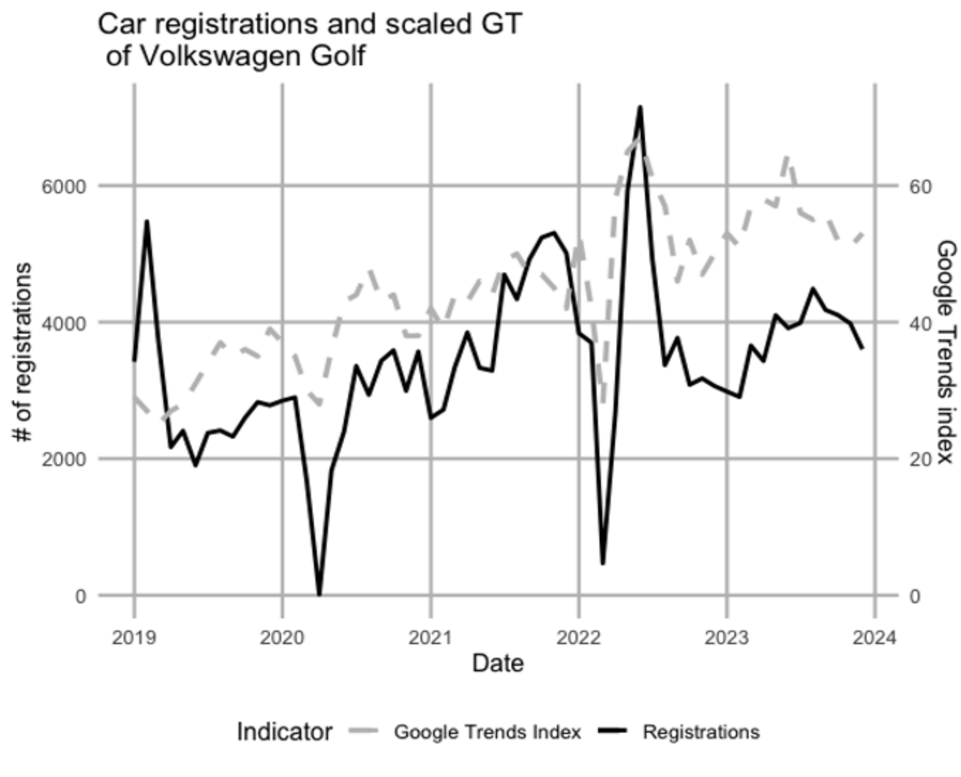
$Invasion$ - dummy variable to incorporate the effect of russian invasion of Ukraine in February 2022. It equals 0 for periods before February 2022 and 1 otherwise.

$Free\,trade$ - dummy variable that equals 1 for the months when the import duties were canceled.

So, the regression analysis can help in predicting future car sales based on the trends observed in Google search data in real-time. We can assess how sensitive car sales are to changes in search frequency when analyzing the regression coefficients. A positive coefficient on GT will mean that an increase in searches correlates with an increase in sales. It also suggests that search activity could serve as an indicator of market demand. Moreover, a statistically significant coefficient can show that changes in search volume can reliably predict changes in sales.

To inspect visually the relationships between online sentiment and subsequent purchase decisions Figure 9 was developed.

Figure 9. Time Series of Volkswagen Golf Registrations and Google Trends index.



Source: Google Trends, Ministry of Internal Affairs, register of vehicles 2019-2023.

The Black Line on the graph represents the monthly count of Volkswagen Golf registrations. The Gray dashed line measures the relative search interest in Volkswagen Golf on Google over time. The data has been scaled (multiplied by 100) to better compare it visually with the registration numbers.

Both lines exhibit some seasonality with peaks and troughs at various points throughout the year. It indicates seasonal interest, usually peaks are observed in the second half of the year. A possible explanation is that people start to think more about owning a car when the colder months approach. Visually, at several points, Google Trends data precede similar movements in registrations. For example, a spike at the beginning of 2021 in Google Trends is followed by a rise in registrations.

The significant drop around early 2020 in both registrations and Google Trends is related to an external shock of the COVID-19 pandemic. It had a significant impact on car sales and consumer interest. A similar situation was observed in February-March of 2022 when a full-scale russian invasion began. At this time, car market activity dropped at 90-95% overall. The same thing happened with the GT index, consumers did not create search queries about vehicles. By the way, after the sharp drop, both lines show recovery, with registrations recovering faster than Google Trends. It indicates a stronger return to purchasing despite lingering lower search interest.

Seeing the impact of external shocks on the automotive market, it is important to account at least for the effect of the military invasion and the cancelation of import duties in our model. The invasion variable in the model captures the effect of the russian full-scale invasion of Ukraine in February 2022, which led to abrupt changes in consumer behavior. The same day it started, there was a sharp decline in car sales which was followed by several weeks of lull in the car market. Therefore, in order to revive market activity, Verkhovna Rada passed a bill №7190 canceling customs duties, value-added tax, and excise duty on imported vehicles for all individuals on the first of April 2022. So, the free trade variable accounts for the periods when import duties on cars were canceled as it substantially increased the number of imported cars. However, later, there a lack of revenues to the Road Fund occurred. The maintenance of transport infrastructure should be a sustainable process, so the preferential customs clearance was canceled on 1 July 2022. After cancellation all tariffs returned to pre-war norms. It is also important to note that the free trade variable will not be added to the models with EVs because import duties on them were canceled before the investigated period had started.

As we work with the time-series data, it is also crucial to check for autocorrelation in regression residuals. When autocorrelation is present - it violates the key assumption of OLS regression. The estimated coefficients may be biased and not efficient. It is important in the context of our analysis because car sales are influenced by cycles and market conditions that evolve over time. Ignoring autocorrelation could lead to over-optimistic conclusions about the relationships between the variables in our model, like Google Trends data, sales, and the impact of external events like the war or import duties cancelation. Therefore, Durbin-Watson and Ljung-Box tests were applied to verify the robustness of the results.

By its nature, the Durbin-Watson test focuses on first-order autocorrelation. It calculates a DW test statistic that ranges from 0 to 4. Values around 2 indicate no autocorrelation, and values around 0 and 4 confirm positive and negative correlation accordingly. (CFI,2024). The Ljung-Box test, on the other hand, is a more general test for autocorrelation. It is used to check autocorrelation at a group of lags collectively (Statology, 2020). As a result, the test provides a p-value, and if it is below the common threshold of 0.05, we can state that there is autocorrelation at some lag in the residuals.

CHAPTER 4. DATA

The data for this work will consist of 2 main parts: Google Trends indices and actual car registrations collected in the database called "Information about vehicles and their owners" collected and published by service centers of the Ministry of Internal Affairs of Ukraine.

Google Trends data is available online by request. First, we submit a request for a specific search query. After that, we can refine our query in more detail and choose a region, time period, and other parameters. The index value ranges from 0 to 100. Where 0 means a complete lack of online interest in the query or a lack of relevant data. A mark of 100 signifies that the query is at the pick of popularity. By the way, the score is normalized and the value does not provide information about the number of searches that were made. Later, we can adjust our request by following parameters:

- Timeframe. It can be chosen from last hour, month, year, and up to 2004
- Geography. Adjustments for a country, region, and city level can be done.
- Category. It allows one to filter and obtain results in a specific category if the word has a few meanings
- Type of a search. We can choose among searches on the internet, YouTube, Google Shopping, and picture search.

After all adjustments, the Google Trend index can be downloaded in an Excel format file and treated as time series data.

The second source is a dataset that contains information on vehicles and their owners, published by the Ministry of Internal Affairs of Ukraine. The unit of observation is every registration action with cars related to the change of car owner carried out in the service center of the Ministry of Internal Affairs. Service centers are the only entity in Ukraine that can certify the sale of a car and change of ownership. The dataset is published on the government portal on a yearly basis and includes registration of new, used, and imported vehicles. It has 1.5-2.5 million observations each year and includes the following variables:

- Type of person and operation: a physician or legal entity performing a registration
- The registration address of the new owner is in KOATUU format.
- Date of operation.
- Department code: helps geographically distinguish service centers.
- Vehicle characteristics: brand, model, vehicle identification number (VIN), make year, color, kind, body type, fuel type, capacity, weight, and others.
- New registration certificate number, assigned car plate number to the new owner.

The next challenge was to form a representative sample among dozens of car brands and thousands of models. Exhaustive research will need analysis for each particular vehicle brand and model. But our approach is to cover at least the most popular vehicle models in a few segments. In this case, the selection includes the top 10 car models of 2023, the top 5 models of new cars for 2023, the top 5 electric vehicles for 2023, and a random selection of executive-class cars presented in regular yearly reports of the Automotive Market Research Institute of Ukraine.

Table 3. Sample of selected car models.

#	Group	Brand	Model	# of operations (2019-2023)
1	Top 10 2023 year	Volkswagen	Passat	298 505
2		Daewoo	Lanos	127 605
3		Skoda	Octavia	258 823
4		Volkswagen	Golf	225 395
5		Renault	Megane	224 424
6		Ford	Focus	254 357
7		BMW	5 series	82 878
8		Skoda	Fabia	112 460

#		Brand	Model	# of operations (2019-2023)
9		Chevrolet	Aveo	81 412
10		Opel	Astra	112 240
11	Top 5 2023 of new cars	Renault	Duster	39 098
12		Toyota	RAV4	76 976
13		Mazda	CX-5	36 810
14		Hyundai	Tucson	63 555
15		Toyota	Land Cruiser Prado	26 296
16	Top 5 2023 of electric cars	Volkswagen	ID.4	8 741
17		Volkswagen	E-Golf	7 115
18		Nissan	Leaf	34 603
19		Tesla	Model 3	11 609
20		Renault	Zoe	5 470
21	Executive class car	Audi	A7	3 628
22		BMW	X5	69 014
23		Porsche	Panamera	2 647
24		Volkswagen	Arteon	1 174
25		Audi	Q7	32 739

Source: Automotive Market Research Institute of Ukraine (2024), author.

So, the overall sample consists of 25 vehicle models in 3 different groups (table 3 above):

- The first group - top 10 car models on the secondary and new market of 2023 due to the annual report of Automotive Market Research Institute of Ukraine. (2024) This category represents the most popular and widely sold vehicles in

Ukraine. It includes models with different body types (SUVs, sedans, hatchbacks), types of fuel, and years of production.

- The second group - is a top 5 car model by the number of new vehicle registrations without mileage. Despite the fact that the market for new cars in Ukraine is relatively small, we consider it necessary to include the most popular new models in the sample.
- The third selection includes the most popular EV vehicles. Electric vehicles represent a segment that transforms the automotive industry, so it is important to include them in the research.
- The fourth group consists of executive-class vehicles. This selection provides an overview of the luxury segment of the market. This group has the lowest number of registrations during the investigation period. However, revealing preferences from this segment is important for a more comprehensive analysis as it offers some contrast to the most popular segments.

CHAPTER 5. RESULTS

In this chapter, we will provide the results of the calculations made following the approach described in Chapter 3. So, the first task was to identify possible optimal lags using cross-correlation analysis for further adjustment. As it can be concluded from table 4 below, adjusting for a time lag is not so effective as it was in related papers. In most instances, the highest CCF values were observed at lag 0, or the results were insignificant at all. This outcome can be explained by the features of the Ukrainian automotive market. A huge portion of cars in Ukraine are imported from the EU or the US salvage market. Unlike the markets of other countries, where customers can purchase vehicles and register them same day, the process for imported cars involves ordering, waiting for delivery, and sometimes repairing the vehicle before registration. Since our dataset only contains information on registrations, this difference creates a mismatch in the time cycle of the purchase-registration process. So, Hypothesis 2 about measurable time lag is rejected, it is worth to check for lags, but mostly no sense to adjust. Results of CCF analysis are presented in appendix.

Table 4. Results of the OLS estimation. [1]

Brand, model	Optimal lag	R-squared	Intercept	GT	Sales(t-1)	Sales(t-12)	Invasion	Free trade	Note
Volkswagen Passat	n/a	0.71	-3748** (1143)	82.69*** (14.57)	0.51*** (0.09)	0.04 (0.07)	-1250*** (301.90)	2331*** (553.50)	-
Daewoo Lanos	n/a	0.7	-1492* (580.60)	181.1*** (28.72)	0.34** (0.11)	-0.01 (0.08)	-693.9*** (135.3)	148.4 (265.6)	-
Skoda Octavia	0	0.67	-2678** (935.1)	102.0*** (15.76)	0.30** (0.10)	0.04 (0.08)	-1262*** (259.0)	928.7 (470.5)	-
Renault Megane	0	0.71	-1177 (623.6)	63.19*** (10.16)	0.33** (0.10)	-0.01 (0.09)	-1247*** (262.3)	1597*** (422.6)	-
Ford Focus	0	0.65	-1409* (555.5)	71.16*** (12.14)	0.27* (0.11)	0.09 (0.10)	-729.0*** (157.8)	651.3* (295.8)	-
Volkswagen Golf	1	0.75	-1966** (669.0)	88.92*** (16.62)	0.46*** (0.09)	0.02 (0.08)	-1069*** (277.0)	1198* (456.8)	-
Toyota Camry	0	0.54	-546,5 (444.4)	55.97*** (13.88)	0.35* (0.14)	-0.02 (0.11)	159.4 (110.2)	-172.9 (252.8)	-
Skoda Fabia	0	0.6	-954,9 (493.9)	52.78*** (9.91)	0.37** (0.11)	-0.02 (0.10)	-442.7*** (116.0)	420.1 (231.4)	-

[1] Standard errors are presented in parenthesis. *p<0.1; **p<0.05; ***p<0.01 indicate statistical significance levels.

Brand, model	Optimal lag	R-squared	Intercept	GT	Sales(t-1)	Sales(t-12)	Invasion	Free trade	Note
Chevrolet Aveo	1	0.6	-858,6 (436.9)	25.27*** (5.13)	0.42** (0.13)	-0.09 (0.09)	-126.9 (86.43)	-11.90 (193.6)	-
Opel Astra	n/a	0.65	-500,2 (382.6)	23.53*** (5.70)	0.47*** (0.10)	0.06 (0.09)	-904.9*** (199.0)	886.5*** (237.9)	-
Average for group of significant results	**0.29**	**0.65**	-	-	-	-	-	-	-
Renault Duster	n/a	0.47	502.4** (167.7)	6.58** (2.25)	0.18 (0.15)	-0.25 (0.13)	-158.4* (66.35)	-260.1 (136.0)	-
Toyota RAV4	0	0.51	-396.2 (310.8)	22.28*** (3.91)	-0.08 (0.15)	0.02 (0.11)	89.3 (77.96)	-223.5 (186.2)	-
Mazda CX-5	0	0.77	-178.0 (95.12)	13.15*** (1.69)	0.03 (0.12)	0.00 (0.10)	19.1 (39.23)	-153.4 (88.25)	-
Hyundai Tucson	0	0.51	-432.8 (276.4)	17.23*** (3.22)	0.13 (0.13)	0.11 (0.11)	38.82 (64.42)	123.9 (141.2)	-
Toyota LC Prado	0	0.53	-191.5 (131.4)	11.48** (3.27)	0.25 (0.14)	0.11 (0.10)	91.79* (36.39)	-116.7 (80.97)	-
Average for group of significant results	**0**	**0.56**	-	-	-	-	-	-	-
Volkswagen ID4	0	0.98	-5.59 (7.92)	2.55*** (0.65)	0.69*** (0.08)	-0.03 (0.12)	23.50 (22.51)	-	-
Volkswagen e-Golf	0	0.97	-35.47*** (9.58)	2.32*** (0.44)	0.62*** (0.06)	0.60* (0.23)	-10.87 (12.91)	-	-
Nissan Leaf	0	0.78	-150.5 (87.64)	6.98*** (1.89)	0.66*** (0.10)	0.13 (0.15)	-0.45 (57.79)	-	-
Tesla Model 3	0	0.95	-67.97** (19.96)	5.10*** (1.08)	0.67*** (0.09)	0.01 (0.14)	3.48 (19.05)	-	-
Renault Zoe	0	0.94	-26.61** (8.41)	1.10*** (0.30)	0.64*** (0.08)	0.48** (0.16)	0.55 (11.56)	-	-
Average for group of significant results	**0**	**0.924**	-	-	-	-	-	-	-
Audi A7	0	0.68	-17.60 (14.07)	0.90*** (0.25)	0.48** (0.15)	-0.27 (0.15)	5.56 (6.29)	-5.89 (11.90)	-
BMW X5	0	0.64	-579.65* (266.64)	38.49*** (8.55)	0.31* (0.13)	0.03 (0.11)	-108.33 (95.88)	436.57** (151.94)	-
Porsche Panamera	n/a	0.29	21.08* (9.65)	0.07 (0.15)	0.50** (0.15)	0.02 (0.14)	-0.09 (4.77)	-3.15 (10.02)	not significant
Volkswagen Arteon	0	0.84	-10.98* (4.87)	0.32** (0.11)	0.69*** (0.11)	0.02 (0.17)	2.34 (4.39)	2.24 (5.69)	-
Audi Q7	0	0.69	-243.27* (98.59)	10.86*** (1.63)	0.15 (0.12)	0.01 (0.12)	-16.90 (36.09)	121.74 (65.79)	-
Average for group of significant results	**0**	**0.71**	-	-	-	-	-	-	-

Hypothesis 1 about the positive effect of GT on subsequent sales is failed to reject. For 24 out of 25 sample models, the rise of the GT index shows a significant and positive effect on sales, the more online interest is observed – the more cars will be sold in the further period. The magnitude varies across models and lies from 2 to 180 registrations per 1 point of the Google Trends index. This suggests that GT is a good predictor of market interest and sales performance and, more importantly - could serve as an auxiliary tool for business decisions.

The lagged sales variables helped to cope with the issue of autocorrelation in residuals. The sales from the previous month variable is significant for almost all models, so it confirms the autocorrelation presence in car sales data. It also suggests that there is some inertia in car sales. So, if a model sells well in one month, it will likely continue selling well in the following month. At the same time, the lagged sales at 12 months are generally less significant compared to the (t-1) lag. However, when testing the model specification, it was found that the presence of this variable, although statistically insignificant, increases the R-squared and confirms that the model explains a larger portion of the variance in data.

The dummy variable for the invasion, as expected, shows a negative effect on sales for most car models. It is highly significant in 10 out of 25 selected models. This confirms that the Russian invasion of Ukraine had a detrimental impact on car sales that lasts up to now. In addition to the immediate danger from hostilities, the war also affects consumer hesitation. War creates economic uncertainty that reduces the activity of the automotive market.

The free trade dummy is significant for some of the models, especially for mostly imported and pre-owned cars from EU markets: Volkswagen Passat, Renault Megane, Ford Focus, Opel Astra, etc. This effect confirms that the removal of import duties had a positive impact on car sales and made vehicles more affordable for Ukrainian consumers.

The R-squared values across models range from 0.47 to 0.98. It indicates that the models generally fit the data well. The highest values of R-squared with an average of

0,92 is observed in the sample group of electric vehicles. We have no clear explanation for this feature, and this topic will need further research.

The next task was to check whether there is no autocorrelation in the residuals of each model. Autocorrelation in residuals can indicate that the model does not fully capture the time-dependence structure of the data, which can lead to inefficient estimates. In order to test for the autocorrelation in residuals, the Ljung-Box test and the Durbin-Watson tests were applied due to our methodological approach. Results are presented in Table 5 below.

Table 5. Results of autocorrelation tests for residuals.

Brand	Ljung-Box P-value	DW test value	Note	Brand	Ljung-Box P-value	DW test value	Note
Volkswagen Passat	0,88	2,03	No autocorrelation in residuals	**Hyundai Tucson**	0,6	2,15	No autocorrelation in residuals
Daewoo Lanos	0,69	2,11	No autocorrelation in residuals	**Toyota LC Prado**	0,39	2,24	No autocorrelation in residuals
Skoda Octavia	0,18	2,31	No autocorrelation in residuals	**Volkswagen ID4**	0,12	2,43	No autocorrelation in residuals
Renault Megane	0,21	2,33	No autocorrelation in residuals	**Volkswagen e-Golf**	0,01	1,22	Positive autocorrelation detected
Ford Focus	0,56	2,12	No autocorrelation in residuals	**Nissan Leaf**	0,32	2,24	No autocorrelation in residuals
Volkswagen Golf	0,24	2,33	No autocorrelation in residuals	**Tesla Model 3**	0,14	2,23	No autocorrelation in residuals
Toyota Camry	0,3	2,28	No autocorrelation in residuals	**Renault Zoe**	0,94	2,01	No autocorrelation in residuals
Skoda Fabia	0,36	2,22	No autocorrelation in residuals	**Audi A7**	0,66	1,94	No autocorrelation in residuals
Chevrolet Aveo	0,49	2,18	No autocorrelation in residuals	**BMW X5**	0,08	2,48	No autocorrelation in residuals
Opel Astra	0,08	1,51	No autocorrelation in residuals	**Porsche Panamera**	0,56	2,16	No autocorrelation in residuals
Renault Duster	0,6	2,14	No autocorrelation in residuals	**Volkswagen Arteon**	0,02	2,57	Positive autocorrelation detected

Brand	Ljung-Box P-value	DW test value	Note	Brand	Ljung-Box P-value	DW test value	Note
Toyota RAV4	0,15	2,4	No autocorrelation in residuals	**Audi Q7**	0,12	2,47	No autocorrelation in residuals
Mazda CX-5	0,03	2,58	Positive autocorrelation detected	-	-	-	-

Most of the models (22 out of 25) have no significant autocorrelation in residuals. It is indicated by the high p-values from the Ljung-Box test and Durbin- Watson values are close to 2. It means that the model residuals do not show any time-related patterns and that the model specification has been chosen properly. However, there are a few models, like the Volkswagen e-Golf, Opel Astra, Mazda CX-5, and Volkswagen Arteon, where the tests confirm possible autocorrelation. The LB test p-values for these models are lower than 0,05, and DW test results have a significant deviation from 2. These cases might need further investigation, it could be adding of additional variables or usage of another model specification (ARIMA or GLS) to address this issue.

CHAPTER 6. CONCLUSIONS AND RECOMMENDATIONS

6.1 Conclusions

The results of the analysis confirm that Google Trends can be a useful tool for predicting sales trends in the Ukrainian automotive market. One of the hypotheses, that there is a positive relationship between the Google Trends index and car sales, failed to reject for almost all models in the dataset. However, the second hypothesis, about the measurable time lag between GT changes and actual sales, is rejected due to the peculiarities of the Ukrainian automotive market.

By the way, we cannot state that there is a causal relationship between the rising of online interest and the subsequent increase of car sales. The aim of the work is rather to identify the degree and significance of that time series of online interest and registration actions moving in parallel. Therefore, Google Trends can serve as an auxiliary tool for making business decisions in the automotive sector. It should complement other traditional methods, but not replace them.

Moreover, it is crucial to understand that sales will not occur in isolation and just because of the GT index rise. They are the results of productive efforts of workers, business management, and economic or social circumstances. External factors like economic stability, consumer purchasing power, and market conditions of course influence sales outcomes a lot.

The findings from this research may have practical implications for automotive market participants - car dealers, importers, and auto consultants. Google Trends could offer a low-cost alternative for businesses to monitor market interest in real time and, as a result, adjust inventory and marketing strategies. By exploring the rise of online interest, dealers can react faster to demand shifts than they would with more traditional methods of market analysis. For instance, if the GT index for the Volkswagen Golf increases by 10 points in the current month, due to our results, there will be sold almost 900 hundred more vehicles in the whole market in the next month. So dealers could proactively import more cars to serve excess demand effectively. Moreover, they can also use this information to place better emphasis on marketing strategies or incentives during periods of high interest. Dealers can also reduce the risk of overstocking or understocking because sometimes it is crucial for the buyer to receive their vehicle right after purchase and not wait for delivery or import process.

6.2 Limitations and Further Research

There are several limitations to this research that should be addressed in future studies. The first one is related to the impact of the economic shocks. During the study period, there were 2 of them - COVID-19 and full-scale russian invasion. These events caused drastic drops in sales but had a more gradual impact on the Google Trends indices. The war has a lasting effect on consumer behavior and economic conditions that cannot be measured precisely.

The second limitation is related to the sample size. In our work, we considered 25 of the most popular vehicle models, but the market consists of thousands of different models. Further research may include a broader sample in order to ensure that the findings are generalizable across the entire market.

The third limitation is connected with data availability. It may be helpful to incorporate prices for completed deals. Information about prices could be web-scrapped from online platforms, but it is not a very reliable source. On the secondary car market, prices may differ significantly from those stated in the advertisements. The information specified in the purchase agreements between individuals is not publicly available. Moreover, while GT only reflects search interest, it may also be good to incorporate data about the traffic on other online advertisement platforms, such as Autoria or OLX.

The Ukrainian car market is unique due to the high proportion of imported used and salvage vehicles, mainly from the EU and the US. The time between purchase, delivery, repair, and registration creates discrepancies that make it difficult to establish a clear time lag between search interest and sales. Future research may also focus on adjusting the methodology in order to account for these market peculiarities.

REFERENCES

Briney, A. (2020, August 26). Who invented the car? *ThoughtCo.* https://www.thoughtco.com/who-invented-the-car-4059932

Corporate Finance Institute. (n.d.). Durbin-Watson statistic. *Corporate Finance Institute.* https://corporatefinanceinstitute.com/resources/data-science/durbin-watson-statistic/

Dmitriyev, I., & Shevchenko, I. (2017). Problems and prospects of development of the automotive industry in Ukraine. *Scientific Journal of Polonia University*, 20(1). http://pnap.ap.edu.pl/index.php/pnap/article/view/115

International Energy Agency. (2024). Global EV outlook 2024: Trends in electric cars. *International Energy Agency.* https://www.iea.org/reports/global-ev-outlook-2024/trends-in-electric-cars

Kinski, A. (2016). Google trends as a complementary tool for new car sales forecasting: A cross-country comparison along the customer journey. *University of Twente.* https://essay.utwente.nl/70462/

Kotler, P. and Keller, K.L. (2012) *Marketing Management.* 14th Edition, Pearson Education.

Motor vehicle in use. (2021). *Countryeconomy*. https://countryeconomy.com/business/motor-vehicle-in-use

Rabi. (2024). Crosscorelation. *MATLAB Central File Exchange.* https://www.mathworks.com/matlabcentral/fileexchange/30510-crosscorelation

Statology. (n.d.). Ljung-Box test. *Statology.* https://www.statology.org/ljung-box-test/

Ukrainian market of new passenger cars: Results of 2023. (2024). *Automotive Market Research Institute.* https://eauto.org.ua/news/482-ukrajinskiy-rinok-novih-legkovih-avto-pidsumki-2023-roku

Ukrainian market of used cars: Results of 2023. (2024). *Automotive Market Research Institute.* https://eauto.org.ua/news/480-ukrajinskiy-rinok-legkovih-avto-z-probigom-pidsumki-2023-roku

Understanding the car buyer’s online journey. (2023). *ACV Auctions.* https://www.acvmax.com/blog/understanding-the-car-buyers-online-journey#:~:text=According%20to%20a%20recent%20Google,research%20vehicles%20online%20before%20purchasing%C2%B9

Varian, H. R., & Choi, H. (2009). Predicting the Present with Google Trends. *SSRN*. https://ssrn.com/abstract=1659302 or http://dx.doi.org/10.2139/ssrn.1659302

Venzhega, V., Rudyk, A., & Pasov, G. (2016). Features of car recycling in accordance with the Law of Ukraine "On Vehicle Recycling". *Technical Sciences and Technologies*. http://nbuv.gov.ua/UJRN/tnt_2016_3_8

Wijnhoven, F., & Plant, O. (2017). Sentiment analysis and Google Trends data for predicting car sales. *ResearchGate*. https://www.researchgate.net/publication/330104910_Sentiment_Analysis_and_Google_Trends_Data_for_Predicting_Car_Sales

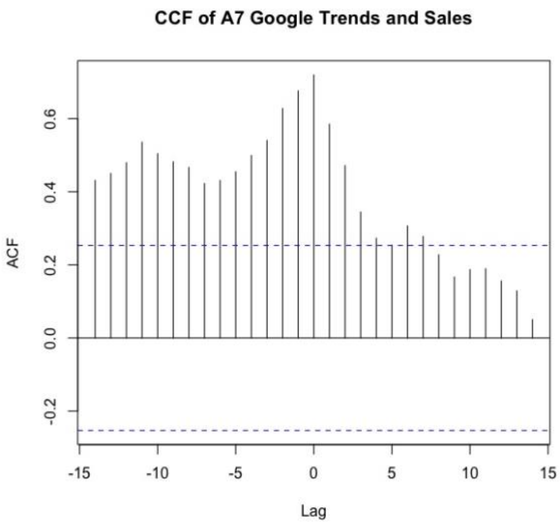
Zlatova, I. O. (2014). Research of consumer preferences in the market of new cars of Ukraine in the conditions of crisis. *Ekonomika: realii vremeni*, 5(15), 160-166. http://economics.opu.ua/files/archive/2014/n5.html

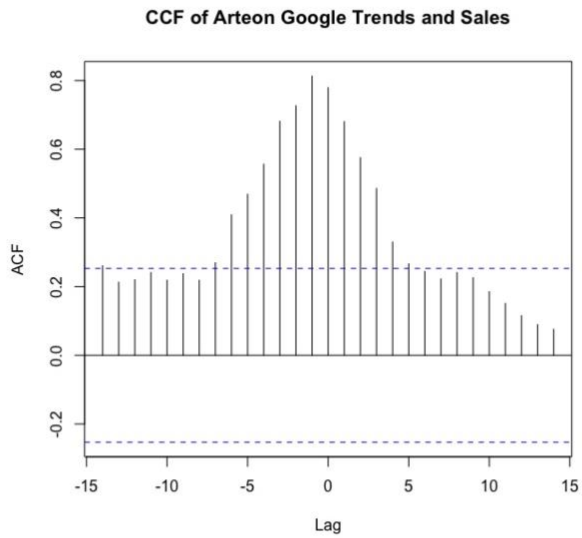
FAQ about Google Trends data. (n.d.). *Google Support*. https://support.google.com/trends/answer/4365533?hl=en

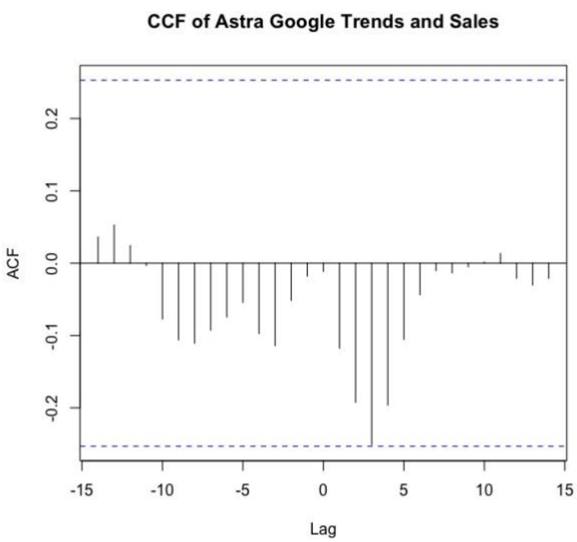
Top 10 for 10 years. What brands of electric vehicles are popular in the Ukrainian market? (2024). *Automotive Market Research Institute*. https://eauto.org.ua/news/571-top-10-za-10-rokiv-yaki-marki-elektromobiliv-populyarni-na-ukrajinskomu-rinku

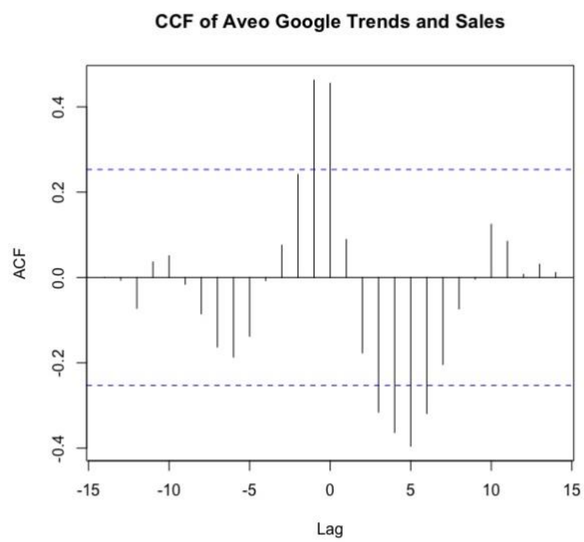
APPENDIX

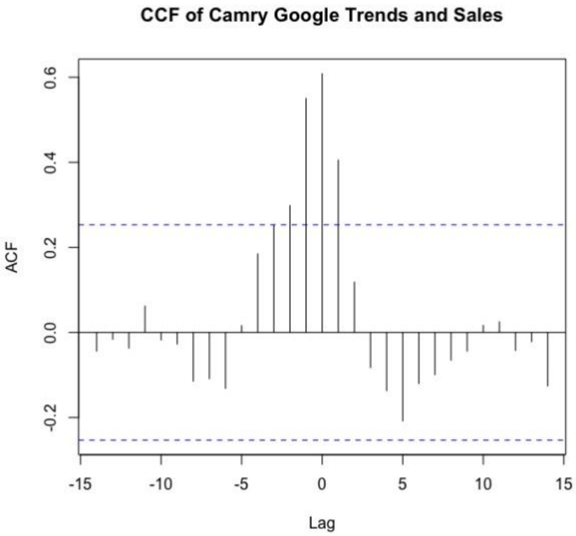
CCF GRAPHS FOR MODELS FROM SAMPLE

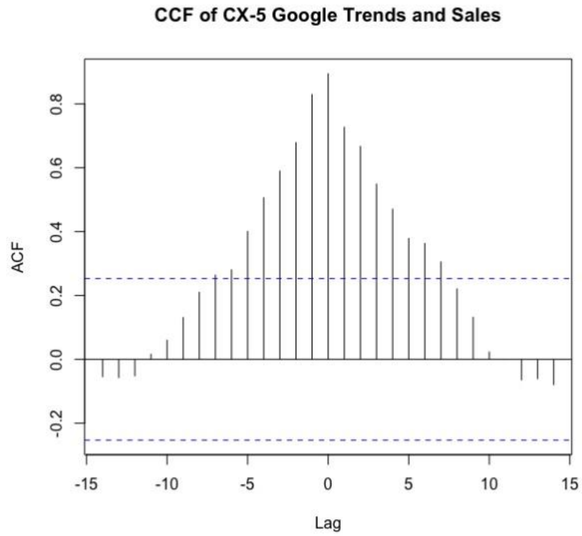


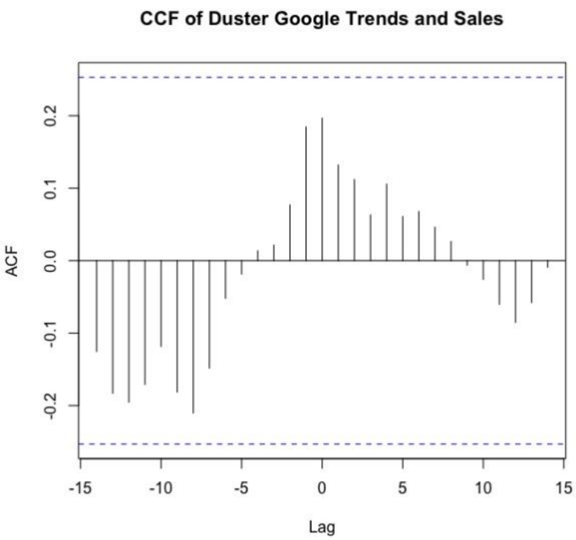


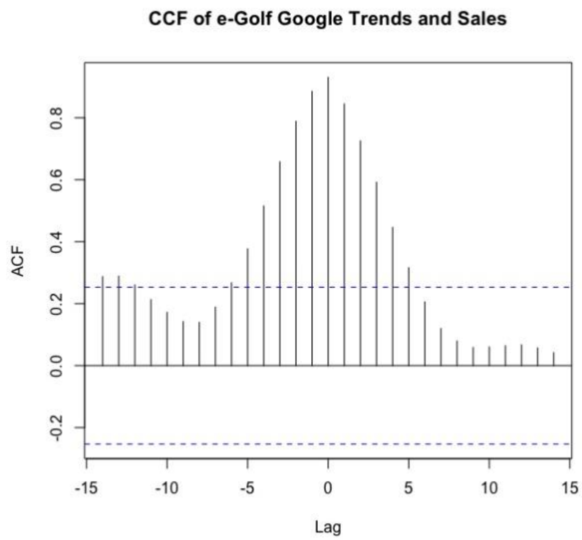


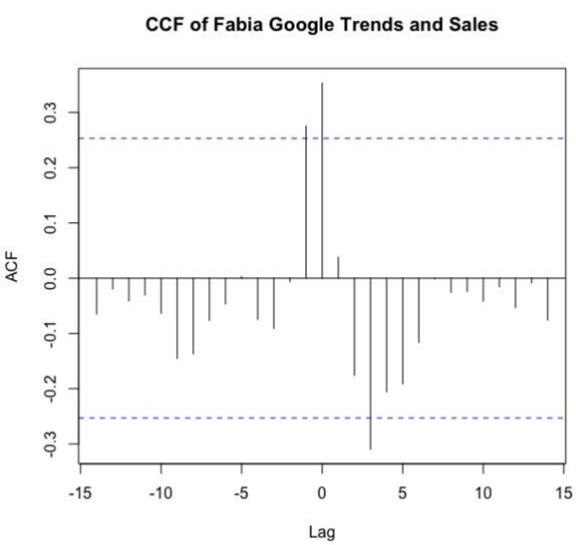


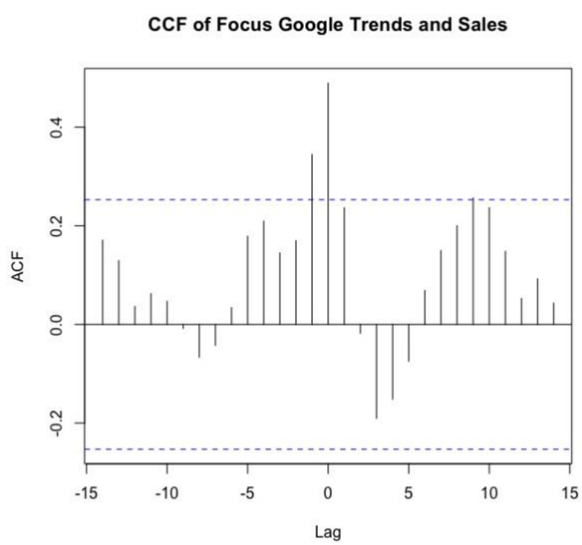

CCF of Camry Google Trends and Sales
ACF
Lag

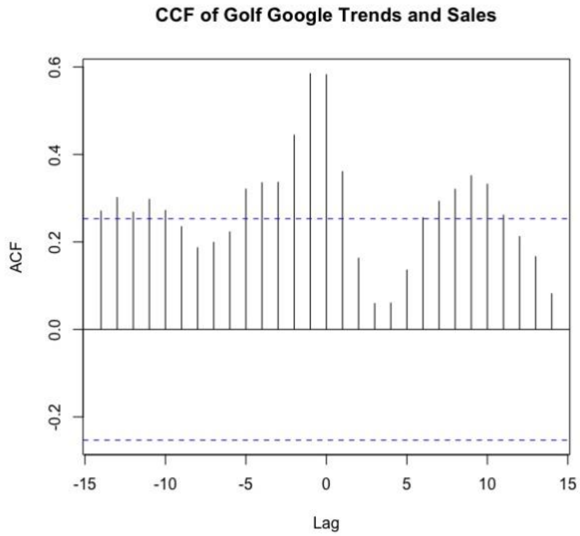

CCF of CX-5 Google Trends and Sales
ACF
Lag

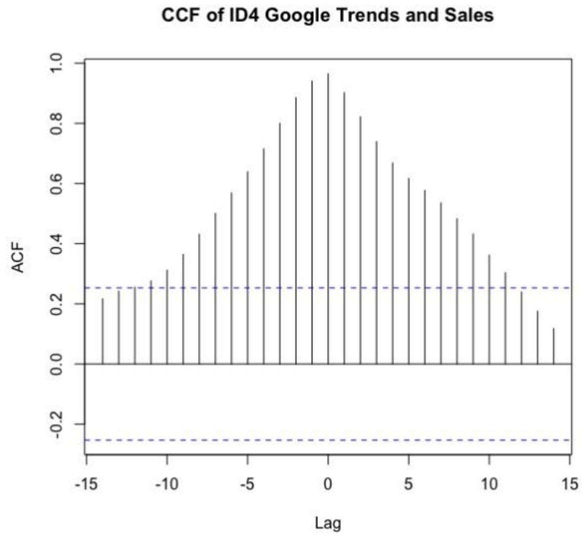

CCF of Duster Google Trends and Sales
ACF
Lag

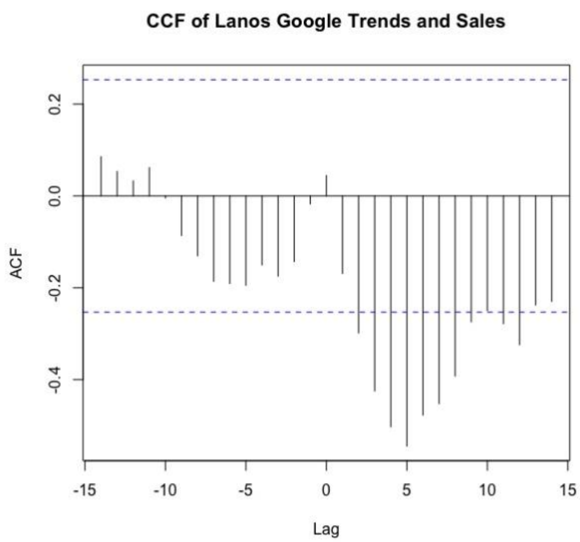

CCF of e-Golf Google Trends and Sales
ACF
Lag

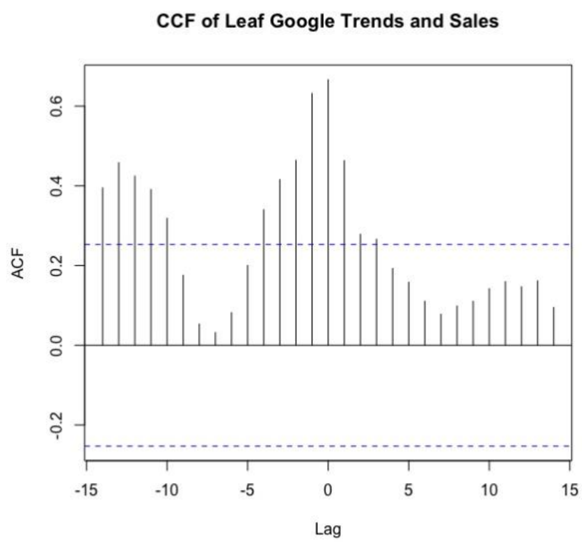

CCF of Fabia Google Trends and Sales
ACF
Lag

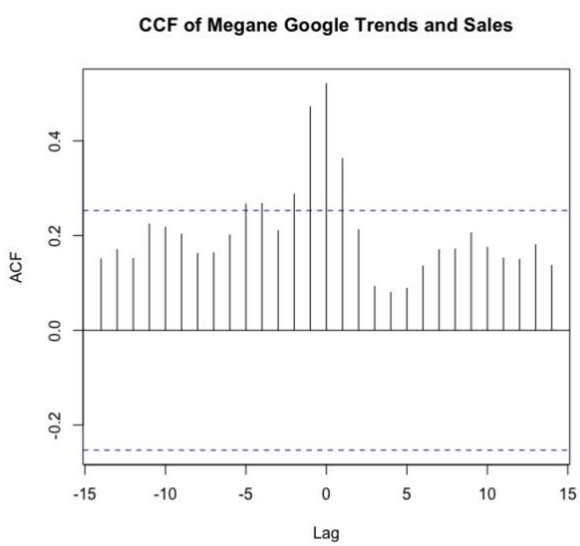

CCF of Focus Google Trends and Sales
ACF
Lag

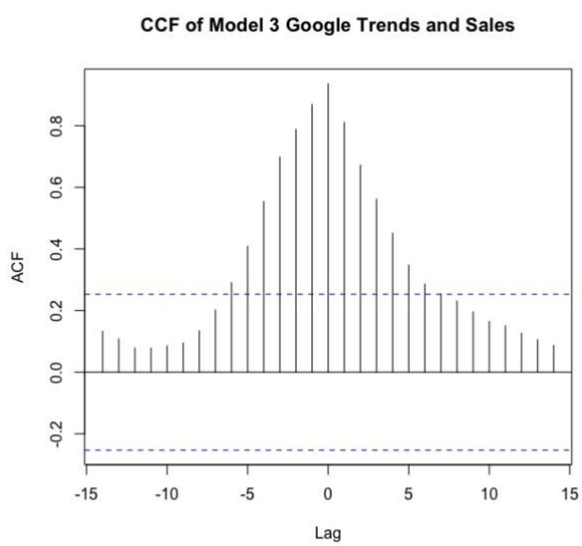

CCF of Golf Google Trends and Sales
ACF
Lag


CCF of ID4 Google Trends and Sales
ACF
Lag

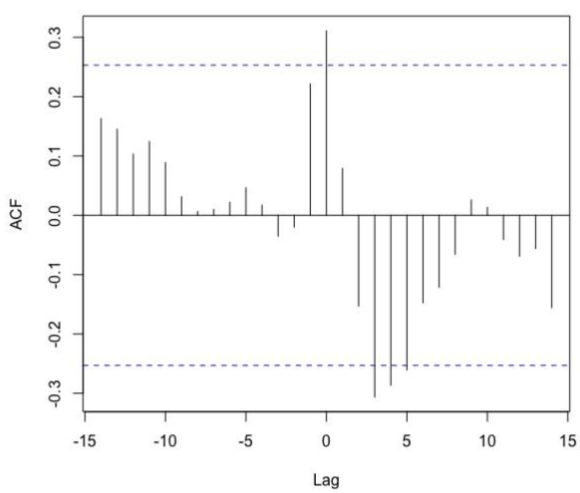

CCF of Lanos Google Trends and Sales
ACF
Lag


CCF of Leaf Google Trends and Sales
ACF
Lag

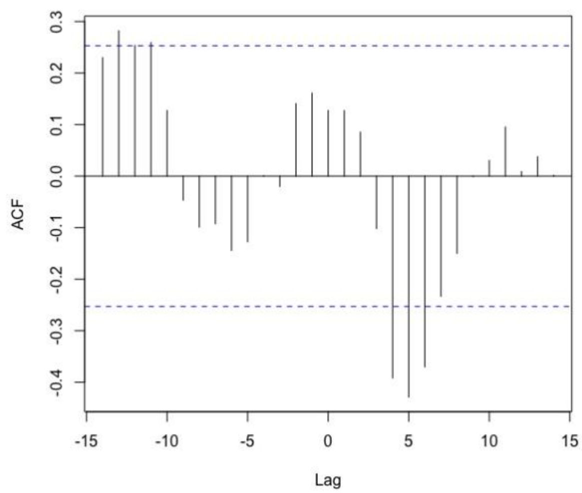

CCF of Megane Google Trends and Sales
ACF
Lag


CCF of Model 3 Google Trends and Sales
ACF
Lag

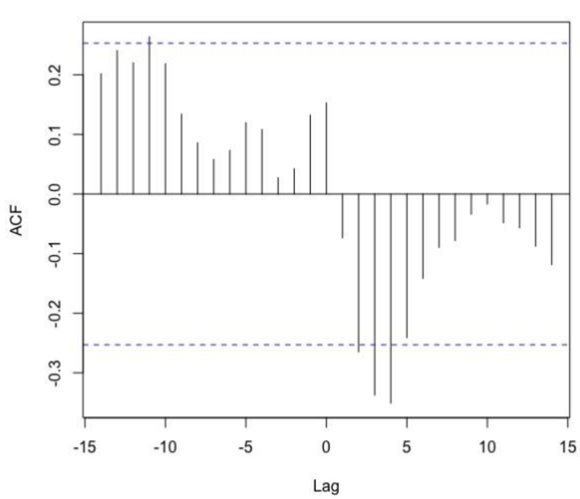
CCF of Octavia Google Trends and Sales



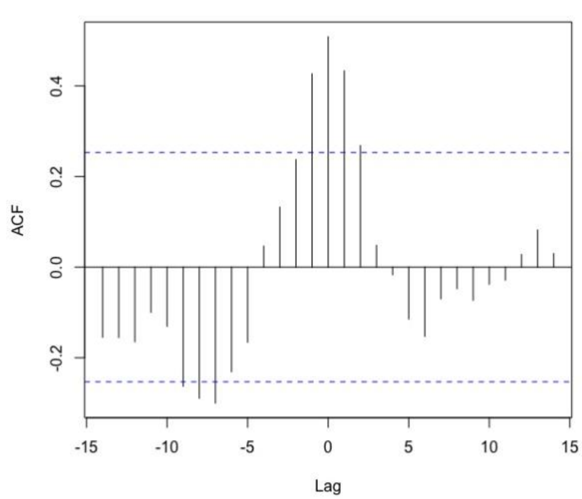
CCF of Panamera Google Trends and Sales



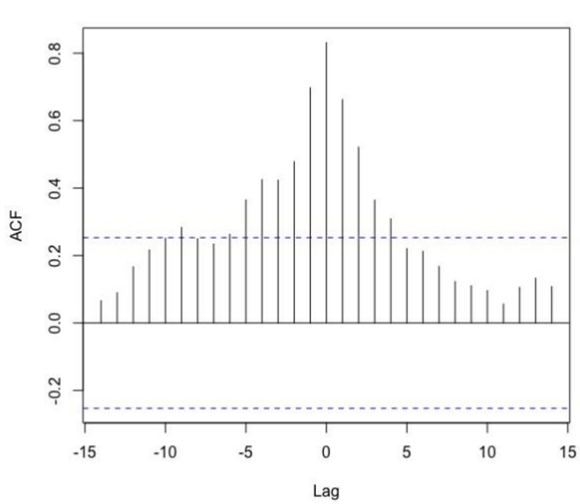
CCF of Passat Google Trends and Sales



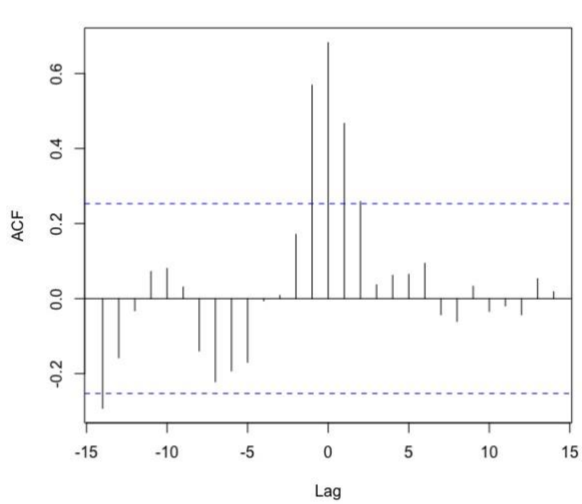
CCF of Prado Google Trends and Sales

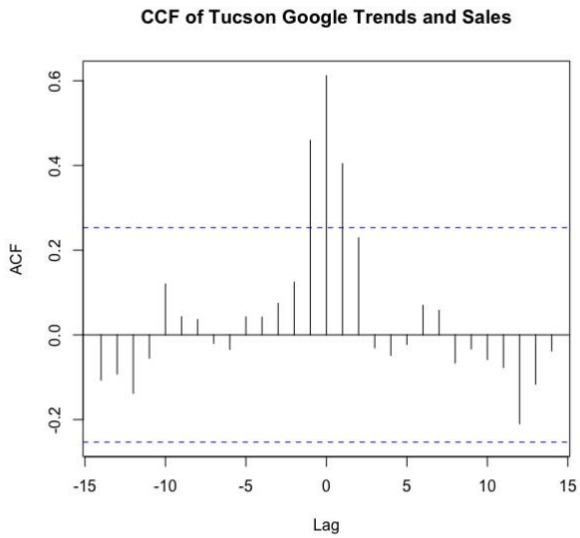


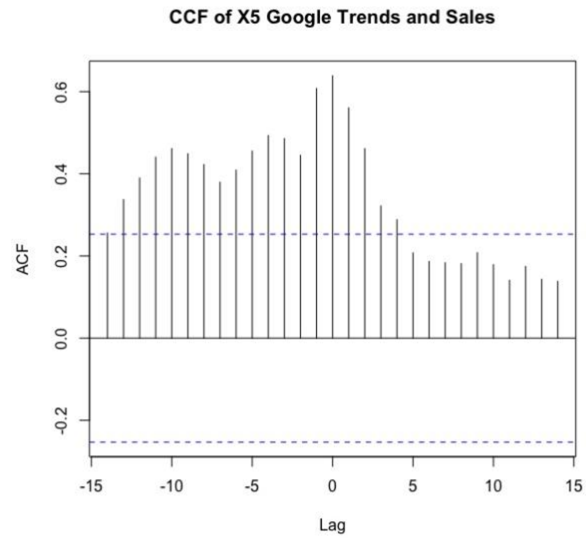
CCF of Q7 Google Trends and Sales

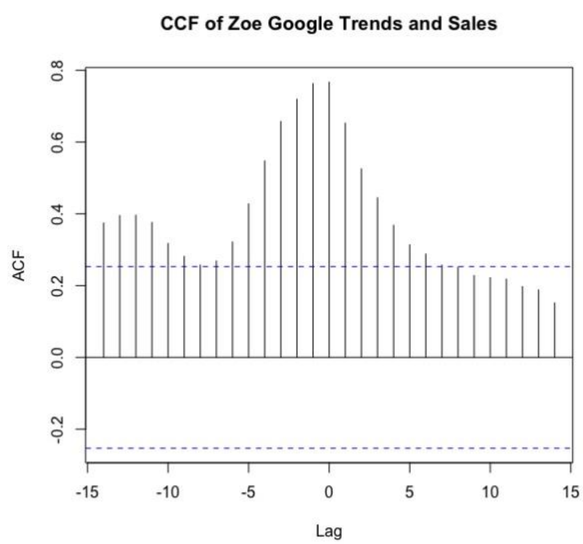


CCF of RAV4 Google Trends and Sales




CCF of Tucson Google Trends and Sales
ACF
Lag
0.6
0.4
0.2
0.0
-0.2
-15
-10
-5
0
5
10
15


CCF of X5 Google Trends and Sales
ACF
Lag
0.6
0.4
0.2
0.0
-0.2
-15
-10
-5
0
5
10
15


CCF of Zoe Google Trends and Sales
ACF
Lag
0.8
0.6
0.4
0.2
0.0
-0.2
-15
-10
-5
0
5
10
15