Forecasting tourism demand in Amsterdam with Google Trends

A research into the forecasting potential of Google Trends for tourism demand in Amsterdam

Elke Rödel

University of Twente

Master Thesis

Author:	E.L. (Elke) Rödel
Student number:	1756419
Education:	Master Business Administration
Course code:	201500102
Supervisor:	A.B.J.M. (Fons) Wijnhoven
Second supervisor:	H.G. (Harry) van der Kaap
Graduation period:	December 2016 – October 2017
Publication date and –place:	19th of October 2017 in Losser

Table of Content

List of figures

List of tables

Abstract

The tourism industry is still growing worldwide and is now responsible for 9 % of the Dutch domestic product, the tourism industry is contributing to the economic growth. Since tourism demand modelling and forecasting has attracted much attention from researchers and progress had been made in this area. This study focusses on the forecasting value of Google Trends for tourism demand by overnight stays in hotels in Amsterdam. The literature indicates that Google Trends has some value for forecasting tourism from which the extent will be measured in this study. The customer journey theory was used to subtract search query terms in a deductive way and the other way around tourist statistics were linked to Google Trends with the use of Google Correlate. The researcher found that data provided by Google Trends can be useful for forecasting night passes in hotels in Amsterdam if the fitting keywords are used. The extent to which the explorative research indicates usefulness is moderate with an average adjusted r-square of 37.4 %. The conclusion for the data driven phase of this research is that there are many correlating search queries, but only very view with possible predictive value or even Google Trends output. Therefore, it was concluded that the data generated from Google Trends in the inductive research has very low to zero usefulness for forecasting night passes in hotels in Amsterdam.

Keywords:

Forecasting; Tourism Demand; Google Trends; Customer Journey in Tourism.

1 Introduction

This chapter presents the problem statement of this study, the central research question and the sub questions. Both the practical- and theoretical relevance will be discussed followed by the scope and outline of this master thesis.

1.1 Problem statement

According to the World Tourism Organization UNWTO (2016) tourism is a still growing industry worldwide, and has been growing strongly for six years in a row now. The Organization for Economic Cooperation and Development, OESO (2016), adds that the tourism industry is now responsible for 9 % of the domestic product and employment worldwide. Also in the Netherlands, the tourism industry is still growing, as is its contribution to the Dutch economy (CBS, Trendsrapport toerisme, recreatie en vrije tijd 2016, 2016). The biggest contribution to the Dutch economy by the tourism industry derives from tourism in Amsterdam (ING Economisch Bureau, 2016). With over thirteen million night passes, Amsterdam is responsible for 30 % of the total number of night passes in the Netherlands and is the biggest touristic hotspot in the Netherlands for both business- and touristic guests. These numbers do even exclude night passes in and around the airport Schiphol, since that might be night passes from people on transit. Since the touristic demand pressures Amsterdam, the city designed a dispersion policy and introduces tourists to other attractive sights and places (ING Economisch Bureau, 2016). With this policy, the city provides other places with the opportunity to gain from tourism and stimulates an even bigger contribution of the tourism industry to the Dutch economy (ING Economisch Bureau, 2016).

Tourism demand modelling and forecasting has attracted much attention of both academics and practitioners (Höpken, Ernesti, Fuchs, Kronenberg, & Lexhagen, 2017). Advances in information technologies have given rise to a massive amount of big data, generated by users. This data includes search query data, social media mentions, and mobile device locations (Mayer-Schönberger & Cukier, 2013). Among the previous years, different variables were used to measure tourism demand, tourist arrival, holiday tourist arrival and business tourist arrival were the most popular measures (Song, Li, Witt, & Fei, 2010) the difference being the nature of the arrivals. Also, tourist expenditure in the destination was often used as the demand variable (Kulendran & Wong, Modeling Seasonality in Tourism Forecasting, 2005). In recent years, there has been an interest in exploiting search query data, which is available through sources, such as Google Trends (www.google.com/trends) to model processes or cycles such as the customer journey (Rivera, 2016). Since the internet is commonly used by tourists for travel planning, Choi & Varian (2012) suggest that Google Trends data about destinations may predict actual tourists' visits to that specific destination.

As discussed above, big data may provide new possibilities for forecasting tourism demand. However, there are some challenges regarding the analysis, capture, search, sharing, storage, transfer, and visualization and information privacy of big data. These challenges require new programs or technologies to uncover hidden values from these large amounts of data (Hashem, et al., 2015). Google

Trends might be one of those programs needed to enlarge the advantage of the use of big data for forecasting tourism demand. This has led to the following research question for this study:

"To what extent is the data provided by Google Trends useful for forecasting night passes in hotels in Amsterdam?"

To be able to answer this central research question the following sub-questions will be answered first:

- *What forecasting techniques are used for forecasting tourism demand now?*
- *What is Google Trends and how does it work?*
- *Is Google Trends a better forecaster than the traditional forecasting method for tourism in Amsterdam?*

1.2 Theoretical- and practical relevance

From a theoretical perspective, this study contributes to literature concerning forecasting tourism demand and extends on Google Trends literature. This thesis analyses the forecasting potential of Google Trends for tourism demand as overnight stays in hotels, which enhances the current knowledge about tourism demand and Google Trends.

The practical relevance of this study lies in the possibility of using Google Trends for forecasting tourism demand in the form of overnights stays in hotels. The use of a free method for forecasting demand based on online actions of actual tourists, provides the hospitality and tourism industry with the opportunity to respond more accurately to the demand, which might lead to better experiences for the tourists and better achievements for the companies. The practical relevance for the city of Amsterdam is the possibility to react more accurately to the demand and maintain their dispersion policy. This might lead to a better spread of tourism in the Netherlands and a larger contribution to the Dutch economy. The total hospitality and tourism industry can benefit from this research since it might provide them with new opportunities for intervening in the customer journey.

1.3 Thesis outline

There are several chapters written for this master thesis. The first chapter, written above, handles the problem indication, the problem statement, research question and the theoretical- and practical relevance. The following chapter gives the theoretical framework for forecasting tourism demand, Google Trends and the customer journey. The third chapter describes the methodology for this study and the fourth chapter presents the results of the data collection and -analysis. The final chapter presents the conclusions of this research, the discussion, the limitations and the indications for further research.

2 Theory

This chapter presents theory of forecasting tourism demand, big data, Google Trends and customer journey in tourism.

2.1 Forecasting tourism demand

Researchers have aimed to forecast tourism demand for years. According to Witt and Witt (1995), whom reviewed the progress made on this subject, the set of techniques used for forecasting tourism demand is limited. Quantitative methods used are econometric models, spatial models or time-series models. In empirical studies, the Delphi method and scenario planning were used. Song and Li (2008) perfected the study and found that tourism demand forecasting heavily relies on secondary data in terms of estimations and can be broadly divided in two categories, namely quantitative- and qualitative methods. From the 121 post-2000 empirical studies reviewed in their paper quantitative forecasting techniques were used in 119 cases. The quantitative methods most used were time-series techniques and econometric models in addition to identify causal relationships. Artificial intelligence was also used as a method, but not as extensive as the other two.

According to these researches quantitative research is the most commonly used category for forecasting tourism demand. From the research of Song and Li (2008) it can be concluded that econometric models and time-series models are mostly used in research into forecasting tourism demand.

2.1.1 Econometric models

Econometric models are models that attempt to replicate the important structures of the real world. These models can include any number of simultaneous multiple regression analysis and linear equations with several interdependent variables (Frechtling, 2011). The major advantage of the econometric approach is the ability to analyze causal relations (Song & Li, 2008). In this case the relation between the dependent variable tourism demand, in the form of touristic night passes, and the possible influencing factors (explanatory variables) such as search query data. This type of analysis "fulfils many useful roles other than just being a device for generating forecasts; for example, such models consolidate existing empirical and theoretical knowledge of how economies function, provide a framework for a progressive research strategy, and help explain their own failures" (Clements & Hendry, 1998, p. 16). For tourism demand, econometric analysis is useful for interpreting the change of tourism demand from an economist's perspective. Since it provides possible recommendations for change in policies or confirms the effectiveness of existing policies (Song & Li, 2008). The most popular measure of tourism demand over the last few years in tourist arrivals. Specifically, this is measured by total tourists arrivals form an origin to a destination. This can later on be decomposed further into holiday tourists arrivals, business tourists arrivals etcetera (Kulendran & Wong, 2005). Other measures used for tourism demand in literature are tourism revenues (Akal, 2004), tourism employment (Witt, Song, & Wanhill, 2004) and tourism import and export (Smeral, Long-Term Forecasts for International Tourism, 2004). Common used econometric models in literature are time varying parameter (TVP) models, the vector autoregressive (VAR) model and the error correction (ECM) model (Song & Li, 2008). With regard to forecasting performances, these models generally predict well although there remains more research to be done to create significant improvements (Song & Li, 2008).

2.1.2 Time-series models

"A time-series model explains a variable with regard to its own past and a random disturbance term" (Song & Li, 2008, p. 210). In other words, a time-series analysis aims to understand reasons for historical patterns in data and to forecast future values (Cryer & Chan, 2008). In contrast to econometric models, time-series models do not explain causal relationships. They however look for time patterns such as trends, cycles and seasonal fluctuations in a single series of historical data. The patterns that are found get modeled mathematically (Andrew, Cranage, & Lee, 1990). The focus lies on exploring the historic trends and patterns of the particular time series based on trends and patterns that were identified by the model. Time-series can be used as a forecasting technique under the assumption that the patterns that were identified in the past, will also occur in the future (Makridakis, Wheelwright, & Hyndman, 2005). Data collection and model estimation for time-series models is less costly since it only requires historical observations of a variable (Song & Li, 2008). Cowpertwait and Metcalfe (2009) add that a typical feature of time-series modelling is that it mostly uses observations that come from a single unit and that are spaced strictly over equal intervals in time. However, these features are typical more than necessary since time-series can be collected form many units and can tolerate deviation across time periods. Time-series models have been widely used for tourism demand forecasting in the past four decades with the dominance of the integrated autoregressive moving-average models (ARIMAs) and were proposed by Box and Jenkins (1970). Depending on the time series simple- or seasonal ARIMAs were used (Song & Li, 2008). Song and Li (2008) state that seasonality is a dominant feature of the tourism industry which makes decisionmakers very interested in the seasonal variation in tourism demand. Cho (2001) and Goh and Law (2002) contradict, they find that simple ARIMAs without seasonality features outperform seasonal ARIMA's (SARIMA's). The performance of their tested ARIMAs was above average of all forecasting models considered. On the other hand, Smeral and Wüger (2005) found that neither the ARIMA or SARIMA model could outperform the Naïve 1 (no-change) model.

2.1.3 Scenario planning studies

Considering the potential effects of crises, disasters and other one-off events, it is important to not only take post events but also possible future events into account. Forecasting risks is of great importance for the tourism industry since it has an impact on tourism demand (Song & Li, 2008). Very little attention has been payed to these issues which lead to Prideaux, Laws and Faulkner (2003) arguing that commonly used forecasting methods have little ability to cope with unexpected events. They found that, although these events occur unexpected they may be associated with some level of certainty. Thus, the effects of these events on tourism demand are, based on appropriate scenario analysis, predictable to some extent. To encounter these possible events tools such as risk assessment, historical research, scenarios and the Delphi approach are suggested (Song & Li, 2008). In forecasting tourism demand very little attention was payed to these methods. Although integration between qualitative and quantitative forecasting approaches to produce a series of scenario forecasts based on several, different assumptions was recommended in literature (Song & Li, 2008).

2.1.4 Artificial intelligence

In addition to the econometric and time series models, artificial intelligence (AI) techniques, have emerged in the tourism demand forecasting literature (Song & Li, 2008). The artificial neural network method (ANN) is a computing technique that imitates the learning process of the human brain (Law, 2000). Kon and Turner (2005) provided an overview of the applications of this in forecasting tourism demand. Empirical evidence shows that ANNs can outperform classic forecasting models ((Burger, Dohnal, Kathrada, & Law, 2001); (Cho, 2003) and (Kon & Turner, 2005)). Despite the satisfactory forecasting performance, AI techniques embody important limitations. Because they lack of theoretical underpinning, and are unable to interpret tourism demand from an economic perspective the techniques provide very little help in tourism policy evaluation. Therefore, the scope of practical applications of AI techniques in tourism demand analysis is restricted (Song & Li, 2008).

2.1.5 Forecasting by the City of Amsterdam

The City of Amsterdam publishes reports concerning the tourism branch in total on yearly bases ((Gemeente Amsterdam, 2014); (Gemeente Amsterdam, 2015); (Gemeente Amsterdam, 2016)). The reports concern trends in tourism, incoming tourism, Dutch tourism, the effect of tourism on the Dutch economy, employment opportunities in the branch and recreation. In these reports the City presents factsheets about the past year and forecasts about the remainder of the year.

The forecasts about the number of overnight stays in hotels are made on the bases of data from a yearly hotelier survey, tourism statistics presented by the CBS and data from the Cities' hotel database (Gemeente Amsterdam, 2016, p. 42). However, the forecasts presented by the City of Amsterdam are not explicit. They concern overnights stays in hotels from all guests, no distinction is being made between nationalities of the tourists, and they forecast only the remaining 6 months of the current year. For example, the forecast of the amount of overnights stays in hotels for the total year 2015 was an increase of 3 % pertaining to 2014 (Gemeente Amsterdam, 2015). However, this forecast was made in July 2015, the data of the first 6 months of 2015 were already taken into account.

The reports provide no insight in the specific forecasting techniques used to design these forecasts. But, while reading the reports the researcher found that it is most likely that the City of Amsterdam used a form of econometric modeling. The reports describe several causal relations which can be analyzed with the use of econometric models. An example of a causal relation described, is the organization of more recreative events in the City which leads to an increase in touristic overnight stays in hotels (Gemeente Amsterdam, 2016).

2.2 Big data

Big data is known as one of the most popular and most frequently used terms to describe the growth and availability of data in the modern age, which is likely to be maintained or even accelerate in the future (Hassani & Silva, 2015). A classification of big data can be found in Figure 2.1.

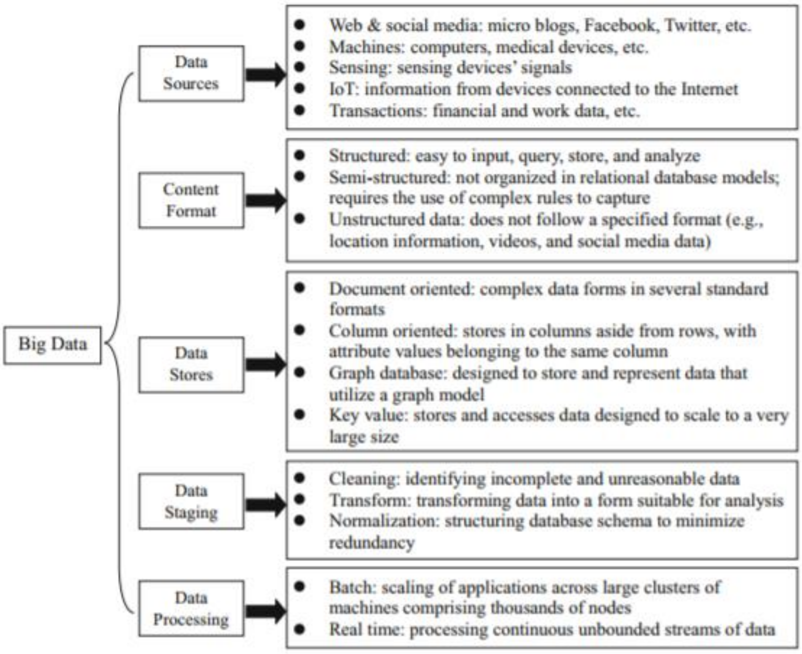


Figure 2.1 Big Data classification (Hashem, et al., 2015, p. 101)

Since the new availability of web-based data sources such as, search engine traffic, customer feedback on review platforms and web traffic, has a natural relation with tourism demand, this big data has been used for tourism demand prediction (Höpken, Ernesti, Fuchs, Kronenberg, & Lexhagen, 2017). Previous studies on tourism are mostly based on surveys or experts' views, this means that they used samples from the total population and do not have real data about all tourists (Song & Liu, 2017). According to Song and Liu (2017) the use of tourism related big data appears to have advantages over traditional methodologies. Firstly, the reliability of the data is higher since it is unprovoked data based on users' real actions and not on samples which allows us to consider all aspects of the information in order to provide accurate results instead of biased conclusions due to data loss because of the usage of sampled data. Secondly, since tourism big data is produced by tourists themselves, it enriches the knowledge of tourism businesses' target markets and is useful for analyzing consumers' demand for touristic products and services (Hendrik & Perdana, 2014). It is a possibility to cross-reference the big data with other data sources. This might lead to determining the balance between the supply and demand of touristic products and services. The last major advantage is the possibility of nowcasting which is the usage of real-time data to describe simultaneous online activities before data sources are made available (Bollier & Firestone, 2010).

On the other side, big data is associated with several issues (Hofacker, Malthouse, & Sultan, 2016). Big data presents observed behavior, but does not provide traditional constructs such as motivation and attitude. Also, having a database does not mean that it can be useful. If the data is of low quality effective database managing is not possible, which means there is nothing to gain from the data (Even, Shankaranarayanan, & Berger, 2010). Therefore, the quality of the big data cannot be assumed. Furthermore, the relevance of the data may vanish in minutes. For example, knowing that a consumer is in proximity to a store is not useful when the consumer has already moved to another location. Normandeau (2013) adds that in this world of real time data you need to determine at what point the data is no longer relevant to the current analysis. Then, the representativeness of the big data forms an issue. How the data was sampled and potential biases creating the sampling procedure. If, for example, some people complain online about a feature of a product, the manufacturer cannot be sure about the nature of the complaint. It could be from an actual customer, but it could also be from a competitor who wants to sabotage the brand (Hu, Bose, Koh, & Liu, 2012). The generalizability of research on the bases of big data forms another issue. While the data might form a complete census from some period, even if it is free from measurement errors, omitted variables and sampling errors, one cannot assume that the results generalize (Hofacker, Malthouse, & Sultan, 2016). The use of big data in correlation research is another issue. The danger with correlation research lies in not understanding the causal relationship between the variables. An alleged cause may be correlated with the outcome of interest, but the correlation could be due to reversed causality or omitted variables (Hofacker, Malthouse, & Sultan, 2016). The last, but certainly not least, issue with the use of big data is privacy. The consumer is often not aware that data is being collected since data sources include online navigation, social media participation and location data form mobile beacons. This is all intimate data which the consumer might not be willing to provide (Hofacker, Malthouse, & Sultan, 2016).

Collaborating, the use of big data has its pros and cons. Since the advantages of using big data in forecasting tourism demand are real, with respect to the disadvantages, the use of big data is worth the risk of continuing in this line.

2.3 Search query data

As specified in the section above a small part of big data is search query data, the keywords that users enter in a search engine. Search query data is valuable for forecasting tourism demand since it provides information about the tourists' interests, intentions and opinions. Tourists use search engines to obtain information about their travel destination, their routes, sights they want to see and other tourists' opinions (Fresenmaier, Xiang, Pan, & Law, 2010). Yang, Pan, Evans and Lv (2015) agree by stating that search query data, including volume and content, captures tourists' attention to travel destinations and is useful in accurately forecasting tourism demand. They found that models that use search query data helped to significantly decrease the forecasting errors of corresponding ARMA models without the search query data input. Varian (2014) argued that real-time data from Google search queries are a good way to nowcast tourists' activities since the correlation analysis of data obtained from Google is a six-week lead on reported values. A clear example of Varians' use of Google search query data is his nowcast of the flu, which identifies possible flu outbreaks one to two weeks earlier than official health reports. However, the literature does point out some challenges in the modelling process of tourism forecasting based on search query data (Li, Pan, Law, & Huang, 2017). The researcher needs to select keywords related to tourism, obtain search query data, select appropriate data and construct econometric models. The most challenging are keyword selection and selecting the appropriate data, since it should be related to tourism which might not always be the case when a specific keyword is used. In their study, Li, Pan, Law and Huang (2017) selected the appropriate keywords by listing all influenza related terms and then eliminating all terms that might indicate something else than the specific information they were trying to find. Further, they delimited their research to bound geographical areas to be certain about the meaning of the queries. With this approach Li et al. (2017) managed to develop a forecasting model that is able to forecast flu outbreak earlier than traditional methods.

Google Trends is a program that uses search query data to detect trends. The program illustrates how often a particular keyword is entered as search term for the total search volume across various countries and in various languages (Choi & Varian, 2012) and is a free service available via www.google.com/trends. The program does not report the raw data about these search queries but presents a query index. This index is made from a query share, which is the total volume of search query terms within a geographic region, divided by the total number of entered queries in that region in the given period of time. Otherwise the places with the highest volume would always be ranked the highest, which would not be accurate. The index is published in values from 0 to 100, the index of 100 indicates the maximum query share for the category determined. Therefore, Google Trends shows Google search engine users interests through time (Google, How Trends Data is Adjusted, 2017).

Researchers have made multiple attempts to use Google Trends as a forecaster in several situations ((Teng, et al., 2017); (Xu & Reed, 2017); (Pollett, et al., 2017)) as did Choi and Varian (2012) in relation to tourism. They found that models that include Google Trends data tend to outperform models that exclude Google Trends data by five to twenty percent and recommend further research into this topic. The researchers designed a linear formula which was estimated for each country and then fitted to the actual visitors data. They found good fits with a R^2 of 73.3 percent.

2.4 Customer journey

Before the customer journey for the tourism sector is defined, it is important to understand the general definition of consumer behavior. The process by which a consumer chooses to purchase or use a product or service is defined as the consumer behavior process (Horner & Swarbrooke, 2011). In addition, tourism research mainly views travel planning as a complex and multi-faced decision making process (Fesemaier & Jeng, 2000) which indicates a complex customer journey. In earlier research, the first models of the customer journey for tourism were developed. These models identified determinants and describe phases in the decision making process (Swarbrook & Horner, 1999).

The customer journey type that is mostly used in the hospitality industry is the process and experience oriented approach (Nenonen, Rasila, Junnonen, & Kärnä, 2008). This is so due to the recognition of the process nature of services and the premises processes are carried out in. The aim of this approach is having a comprehensive description of the clients' process. Literature proposes three types of customer journeys in tourism, the customer journey funnel (Lewis, 1903), service mapping and sequential incident technique (Strauss & Weinlich, 1997) and the visitors journey cycle which has interrelating stages (Lane, 2007).

The basic method to describe and understand the process of a customer is service blueprinting (Koljonen & Reid, 2000). In service blueprinting the processes of services and interactions are visualized as a flowchart. This approach looks at the processes form the companies' view rather than the customer perspective and illustrates actions or events. Lewis (1903) proposed the customer journey funnel as presented in Figure 2.2 Customer journey funnel, this funnel is accepted in literature. Both the awareness and interest stages are stages where the tourist might use the Google search engine to become aware of destinations available and learn more about them.

Figure 2.2 Customer journey funnel (Lewis, 1903)

Other methods within the process and experience oriented approach are service mapping and the sequential incident technique (Strauss & Weinlich, 1997). The first is, comparable to service blueprinting, industry focused whereas sequential incident technique (SIT) focusses more on the customer perspective. The most important part of SIT is the service map which presents the customer path, reflecting the course of a typical customer process (Strauss & Weinlich, 1997). Interactions and transactions are chronologically presented in a flowchart with only a horizontal axis. Figure 2.3 Figure 2.3 SITprovides the flowchart of a holiday transaction, which in the case of Strauss and Weinlich (1997) concerns a holiday in a club resort. Although this method was designed before the active use of internet in the tourism customer journey this method does recognize the information stage. This is the stage where nowadays the Google search engine might be used by the tourists.

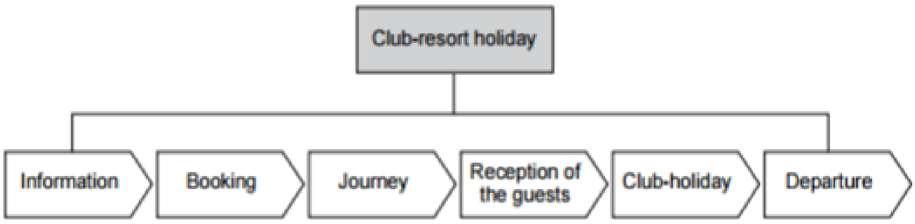


Figure 2.3 SIT (Strauss & Weinlich, 1997, p. 42)

The customer journey is the cycle of interaction between a customer and an organization (Nenonen, Rasila, Junnonen, & Kärnä, 2008). This is a visual process-oriented method for structuring peoples' experiences. It describes the transition from 'never being a customer' to 'always a customer'. Within this journey the value of customers changes and the mental models, the flow of interactions and possible touch points are taken into account. This cycle usually starts when the customer wants or needs a service and continues to the point where it is reclaimed. Different phase classifications are used by different authors. The phases from a customer experience perspective are 'need', 'enquire', 'approach', 'recommendation', 'purchase', 'experience' and 'problem' (Nenonen, Rasila, Junnonen, & Kärnä, 2008). From the process perspective, the phases are 'orientation', 'approach', 'action', 'depart' and 'evaluation' (Nenonen, Rasila, Junnonen, & Kärnä, 2008). However, the phases from the process perspective mentioned by Nenonen et al. (2008) do seem to match the phases described in the customer journey funnel (Lewis, 1903) or in service blueprinting or SIT (Strauss & Weinlich, 1997) which suggested that these process perspective phases might not be to accurate.

The phases from the customer experience perspective that might include the use of the Google search engine are need, 'I am considering a purchase, where do I go?' and enquire, 'I make general enquiries to possible suppliers'. From the process perspective, the orientation phase is the phase where the tourist might use the Google search engine.

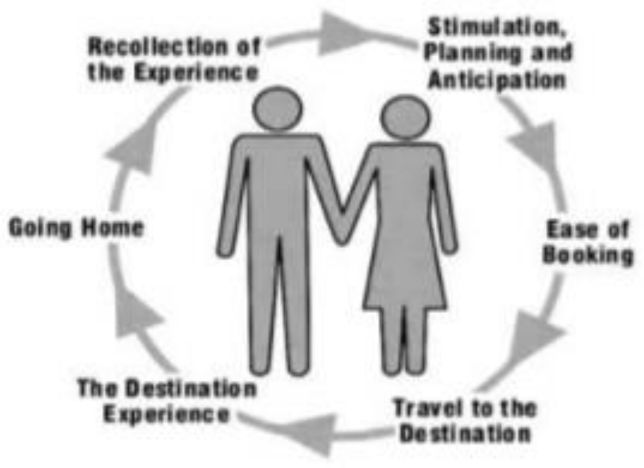


Figure 2.4 Visitors journey (Lane, 2007, p. 252)

In addition, according to Lane (2007) the customer journey in tourism is called the visitors' journey. Within the visitors' journey six interrelated stages have been identified from the tourists' perspective as shown in Figure 2.4 Visitors journey. The stages identified are: 'stimulation, planning and anticipation', 'ease of booking', 'travel to the destination', 'the destination experience', 'going home' and 'recollection of the experience'. This model was designed to understand the tourists and enable the industry to engage in the process tourists go through and add value to it. The first stage, 'stimulation, planning and anticipation' is the stage wherein Google might be used by the tourist. Lane (2007) states that the tourist might look for the destination itself and activities in accessible formats.

2.4.1 Selection of constructs and search terms derived from the customer journey

Concluding, there are different methods for describing the customer journey. Some are presented as flowcharts while others are cycles. All models describe comparable processes with phases that overlap between the several processes. With the aim of finding the corresponding search terms for the data collection the phases of the customer journey which might include the use of the Google search engine are operationalized in Figure 2.5. These stages are a combination of the models from Lewis (1903) and Nenonen, Rasila, Junnonen and Kärnä (2008) which are combined since the researcher found they form an addition to each other, since the customer journey funnel from Lewis (1903) is designed from the organizations perspective and the phases from the customer journey cycle from Nenonen et al. (2008) are designed from a customers perspective. The combination of the models provides the most complete description of the stages.

While operationalizing the customer journey in tourism the phases awareness or need and enquire or planning where chosen. In these phases tourists become aware of their interest and start orienting their vacations. In these stages are defined as two separate phases since in the first phase, tourists can still decide not to visit Amsterdam while in the second phase they already have chosen Amsterdam as their destination. The awareness or need phases is operationalized in the search behaviors ‘search for a type of holiday’ and ‘search for a destination’, since Lane (2007) states that in the first stage of his model (stimulation, planning and anticipation) tourist search for a destination. The anticipation falls within the awareness or need stage. The second search behavior was chosen since Strauss and Weinlich (1997) describe this in the information phase of their model which fits within the operationalized awareness or need phase. From these search behaviors the search queries ‘travel Amsterdam’ and ‘Amsterdam’ are derived which will cover the awareness or need stage.

The second phase, the enquire or planning phase, is divided in the search behaviors ‘search for hotels’, ‘search for ways of traveling’ and ‘search for activities’. Since Lane (2007) describes that in the customer journey planning phase, tourists search for places to stay and activities to do. This leads to the operationalized search queries ‘hotel Amsterdam’, since this study focusses on overnight stays in hotels, and ‘tourist info Amsterdam’. Strauss and Weinlich (1997) find that while planning, tourists take the journey into account, which leads to the search behavior ‘search for ways of traveling’ into the specifications ‘by car’, ‘by plane’ and ‘by train’. The specifications ‘by plane’ and ‘by train’ are operationalized in the queries as ‘flight Amsterdam’ and ‘train Amsterdam’. NBTC Holland Marketing (2016) states that these ways of traveling are most common for foreign tourists to get to Amsterdam with 32 % by plane and 9 % by train, which leads to the decision not to take traveling by car into account.

In the operationalization the search queries ‘city trip Amsterdam’, ‘holiday Amsterdam’ and ‘visit Amsterdam’ are derived directly from the enquire or planning phase. These queries relate directly to the planning of the stay and can be used by tourists to gather information about any one of the other search queries in this phase.

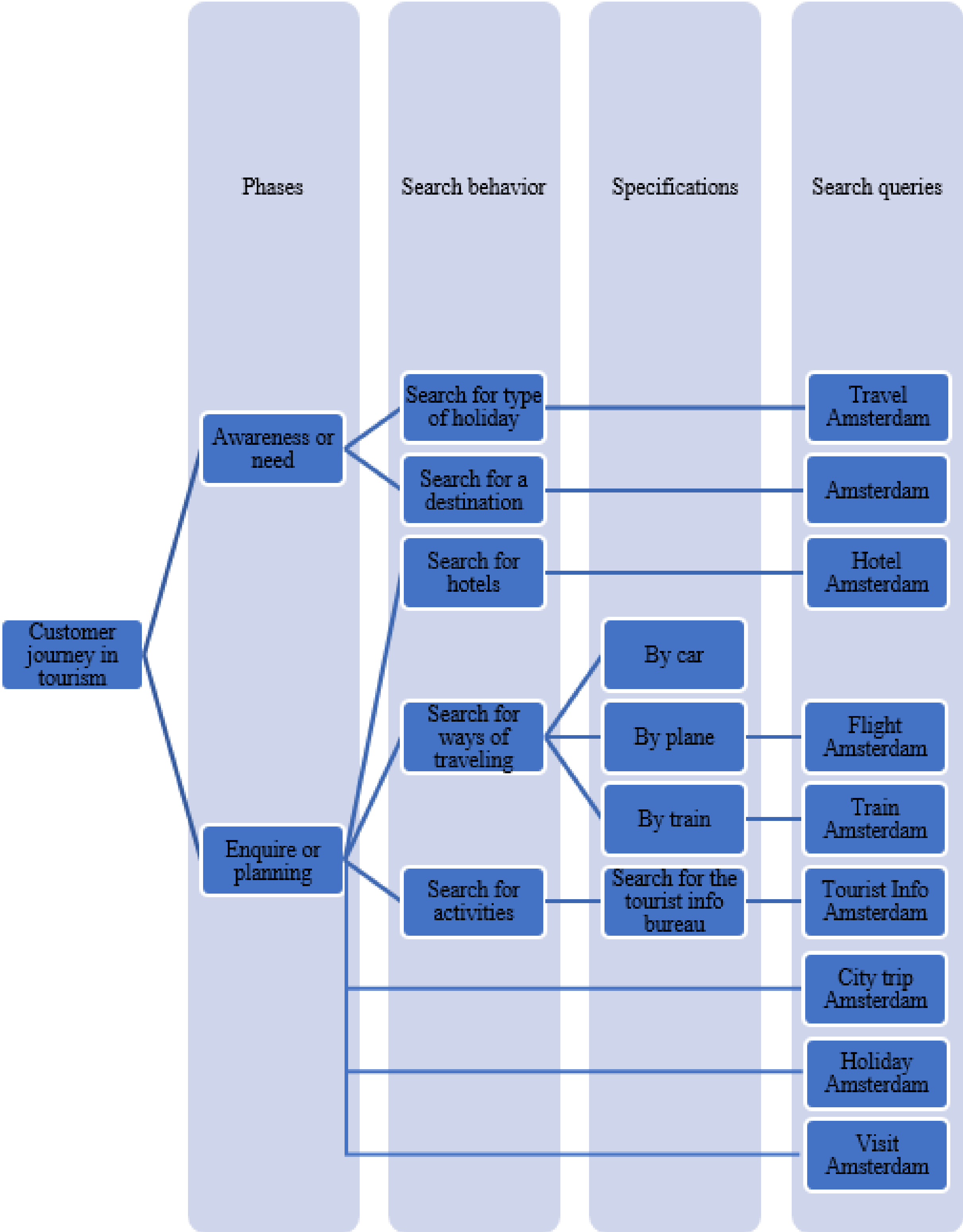


Figure 2.5 Operationalization of the customer journey in tourism

3 Methodology

This chapter describes the methodology of this study. First the chosen research design, followed by the collection- and analysis of data.

3.1 Research design

This study aims at forecasting tourism demand, in the form of touristic night passes in hotels, with Google Trends with the use of an deductive and inductive approach. The quantitative data sources that will be used are Google Trends and the data from CBS StatLine. The study takes the form of time-series modeling since it is based on trends and patterns identified by Google Trends. This form of research is convenient since it only requires historical observations of a variable (Song & Li, 2008).

An important part of this research in forecasting, is finding the relevant Google Trends keywords or search queries. For that, both a deductive and an inductive approach will be used. First, the customer journey theory is used to subtract search query terms. It will be investigated how strong the relation is between these terms and the pattern of specific tourists' statistics of the city of Amsterdam.

Besides that, a more common inductive approach in using Google Trends data will be used. This will be done by linking tourist statistics to Google Trends the other way around with the use of Google Correlate. Based on this, we try to find other keywords and search queries for this field of interest.

With this design, the distinction between explaining and predicting will be made which leads us to the following technique questions:

- *What are relevant keywords for forecasting night passes in hotels in Amsterdam according to the customer journey theory?*
- *What are relevant keywords for forecasting night passes in hotels in Amsterdam according to Google Correlate?*

3.2 Data Collection

Data will be collected from Google Trends, reports from the city of Amsterdam (2017) and the electronic database CBS StatLine (2017). This study will focus on the calendar years 2014, 2015 and 2016 and on tourism demand in the form of night passes in hotels in Amsterdam. This means that the sample consists of thirty-six months.

Data from the reports will be filtered while reading the reports, but for the data that will be collected from Google Trends keyword selection is of major importance. As pointed out in the theory of this study, the researcher needs to use keywords related to tourism, obtain search query data and select appropriate data. The most challenging is keyword selection.

The keywords used for the first part of the research are derived from the customer journey theory. For the awareness or need phase of the customer journey the keywords 'Amsterdam' and 'Travel Amsterdam' are chosen. For the enquire or planning phase the following keywords are chosen, 'Hotel Amsterdam', 'Flight Amsterdam', 'Train to Amsterdam', 'City trip Amsterdam', 'Holiday Amsterdam', 'Visit Amsterdam' 'Tourist info Amsterdam' and 'Visit Amsterdam'. This part of the research is divided in two phases since in the first phase tourists can also decide not to visit Amsterdam while in the second phase Amsterdam is already chosen as the destination.

For the second part of the research keywords derived from the data will be used. These keywords are subtracted from Google Correlate on the bases of data concerning the night passes in hotels in the City of Amsterdam. Google Correlate is an online, automated method for query or keyword selection that does not require such prior knowledge. Instead, given a temporal or spatial pattern of interest (a dependent variable), Google Correlate determines which keywords best mimic the data. Google correlate computes the Pearson Correlation Coefficient, also known as ‘r’ (Google, 2017). However, spurious correlation exists, strong correlations do not always imply causation (Vigen, 2017). Therefore, the correlations found will be filtered on the bases of a possible relation with tourism demand in the City of Amsterdam. For this research the keywords that have a correlation between r = 0.80 and r = 1.0 and could possibly relate to tourism demand in the City of Amsterdam are examined.

When the most accurate keywords are selected and entered the function of the Google Trends program translating these keywords into various languages used by tourists all over the world (Choi & Varian, 2012) makes the data exists out of a national sample per country which represents the entire population without sampling bias (Li, Pan, Law, & Huang, 2017). This enhances the reliability of the research (Song & Liu, 2017). To filter out data that does not relate to tourism or falls outside the scope of this research the Google Trends filters will be used. The geographical area will be set at the United Kingdom and Germany since these countries represent the biggest part of foreign tourists in the City of Amsterdam form Europe, an exact representation of the percentages can be found in appendix I (CBS, 2017). The settings from the time filter will be set to the calendar year 2014 up to and including 2016. The category filter will be set to travelling since the program than only counts the search query data that fits within this category. The last filter will be set at Google Search since that is the part of the search engine tourists use for their travel planning.

3.3 Data Analysis

A benefit of Google Trends as a source of data is that it is a suitable data source of timely information (Xu & Reed, 2017). Google Trends provides the data on weekly bases, but the analysis of this study requires a monthly time series since that is how the report concerning tourism demand in Amsterdam are presented. The research would be more effective if weekly timeseries of the dependent variable would be used but unfortunately this data is unavailable for the researcher. Therefore, the Google Trends data needs to be aggregated. To aggregate a weighted average will be used; the index of each week is weighted according to the share of the week that falls in the month. This means that a week that falls completely in a month gets the weight of seven divided by the total number of days in that month. Weeks that extend across two months are weighted by the amount of days they fall within the one month divided by the number of days in that month and then the residual amount of days get divided by the number of days in the second month. The study is aware of the assumption being made that search behavior of tourists is constant across all days of the week and recognizes this as in flaw in the research. However, there was no better way found of aggregating the data.

When the data is aggregated analysis will be done by the steps of the regression model estimation process (Frechtling, 2011). This process was chosen due to its ability to work without the use of a base model with time lag (Frechtling, 2011) pertaining to the method used by Barreira, Godinho and Melo (2013) which works with seasonality and time lags. An analysis which includes time lags might provide a more accurate representation of the forecasting value of Google Trends but due to the availability of data on monthly bases it is not possible to establish an accurate time lag for the customer journey in tourism. Since this monthly based data is available, the regression model estimation process is most useful (Frechtling, 2011). The correlations among the explanatory variables, the data derived from Google Trends, will be examined to identify any multicollinearity between the independent variables, which will then be removed from the model. The multicollinearity tests outcome needs to fall within the tolerance > 0.2 and the variance inflation factor VIF < 5 not to be excluded from the model (Grande, 2015). Next the expected relationships will be specified and initial models will be identified. Thereafter, the validity of the model is going to be evaluated and the significance assessed.

4 Results

In this chapter, the results of the data analysis will be presented. First data of the actual night passes in hotels in Amsterdam will be analyzed, then the correlations of the explorative research are presented followed by the results of the inductive research.

4.1 Hotel night passes in the City of Amsterdam

The night passes in hotels in the City of Amsterdam with a German or United Kingdom origin represent on average over the years 2014, 2015 and 2016 30.6 percent of the total amount of night passes in hotels in the city. This are 10,017,000 actual night passes in this time period (CBS, 2017). As can been seen in Figure 4.1 Hotel night passes in the City of Amsterdam with German or United Kingdom origin, this happens with seasonality. The seasonality in comparable between Germany and the United Kingdom and experiences peaks and descends in approximately the same periods.

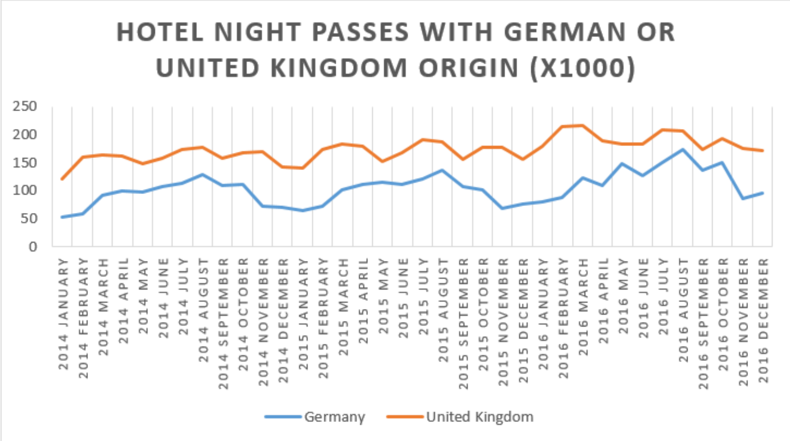


Figure 4.1 Hotel night passes in the City of Amsterdam with German or United Kingdom origin

The hotel night passes with German origin (Figure 4.2 Hotel night passes, German origin) represent on average 11,5 percent of the total number with 3,770,000 actual night passes in the given period. The number of German night passes has increased over the years with 1,112,000 in 2014 to 1,469,000 in 2016 (CBS, 2017). It stands out that the seasonality forms a less fluent line in 2016 where more peaks and descends are noticeable.

On average 19.1 percent of the hotel night passes in 2014 till and up to 2016 have an United Kingdom origin. This are 6,427,000 overnight stays in hotels (CBS, 2017). Figure 4.3 Hotel night passes, United Kingdom origin shows that the seasonality with United Kingdom origin has more peaks and

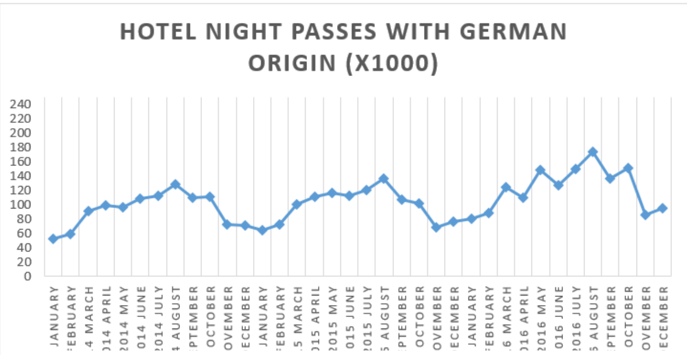


Figure 4.2 Hotel night passes, German origin

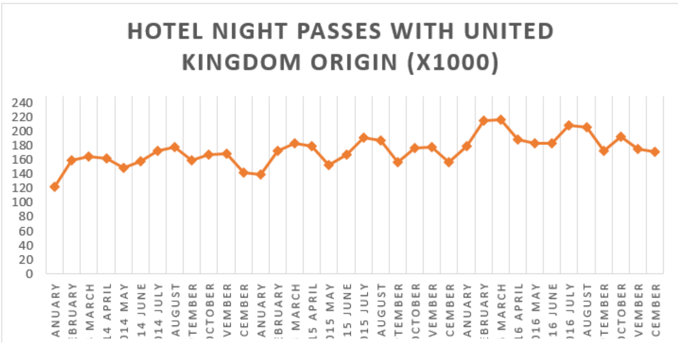


Figure 4.3 Hotel night passes, United Kingdom origin

descends over the years than the German origin. However, this seasonality is comparable over the years with a noticeable increase in the amount of hotel night passes.

When comparing the actual night passes with a German origin with the Google Trends output for the same period in phase 1 (keywords ‘Amsterdam’ and ‘travel Amsterdam’) in Figure 4.7**Fout! Verwijzingsbron niet gevonden.**, it stands out that the keyword ‘travel Amsterdam’ shows no similarities of any kind with the actual night passes. Continuing, the keyword ‘Amsterdam’ shows a curve with possible seasonality which does not fit to the curve of the actual passes but shows some similarities.

The comparison of these variables for the second phase with a German origin (keywords ‘hotel Amsterdam’, ‘flight Amsterdam’, ‘holiday Amsterdam’, ‘train Amsterdam’, ‘tourist info Amsterdam’ and ‘visit Amsterdam’) in Figure 4.6 shows no results for ‘tourist info Amsterdam’, ‘flight Amsterdam’, ‘holiday Amsterdam’ and view results for ‘visit Amsterdam’. For the keywords ‘hotel Amsterdam’ and ‘train Amsterdam’ curves with seasonal effects are drawn which, to some extent, show similarities to the actual night passes.

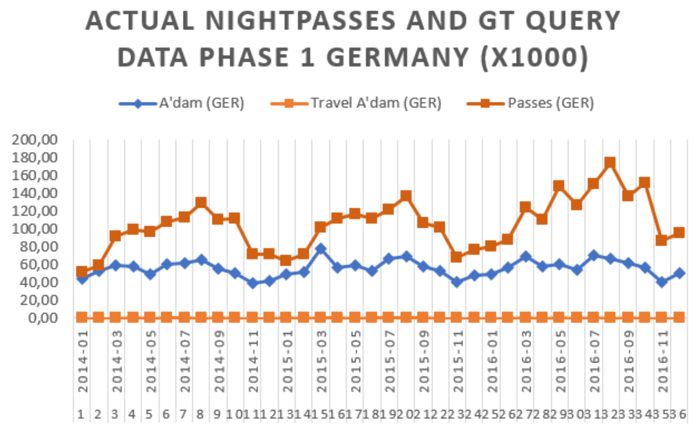


Figure 4.7 Hotel night passes Germany vs GT data phase 1

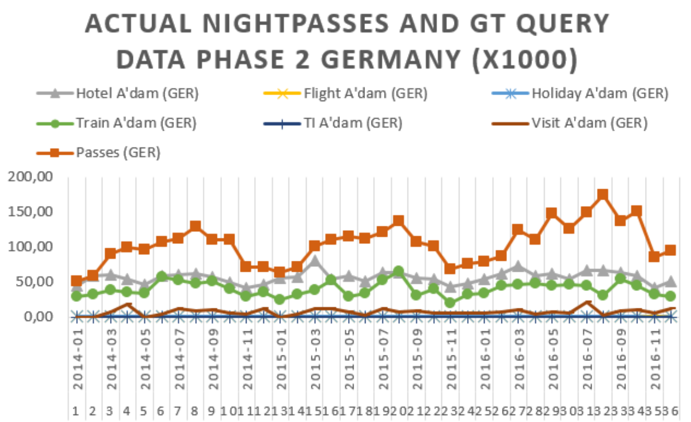


Figure 4.6 Hotel night passes Germany vs GT data phase 2

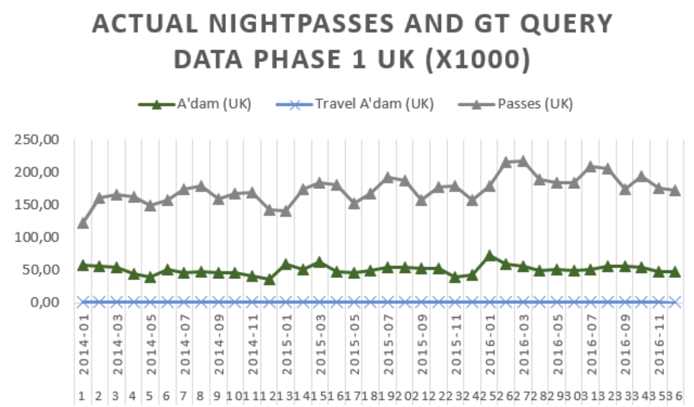


Figure 4.5 Hotel night passes UK vs GT data phase 1

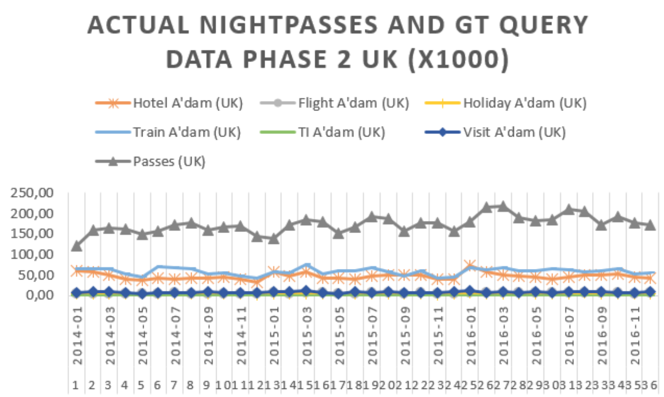


Figure 4.4 Hotel night passes UK vs GT data phase 2

While comparing the same for the tourists with an United Kingdom origin for phase 1 in Figure 4.5 it stand out that, again, the keyword ‘travel Amsterdam’ shows no similarities of any kind with the data of the actual night passes. The keyword 'Amsterdam' does show results but mediate signs of seasonality which the data of the actual night passes does show.

The comparison of the keywords of phase 2 with the actual night passes with an United Kingdom origin is shown in Figure 4.4. It stands out that the keywords 'tourist info Amsterdam', 'flight Amsterdam', 'holiday Amsterdam' and 'visit Amsterdam', again, do not show any signs of similarity in relation to the actual night passes. The keywords 'train Amsterdam' and 'hotel Amsterdam' on the other hand, do show signs of similarity. Although the curves show less peaks and descents than the curve of the actual night passes signs of seasonality are, to some extent, recognizable.

4.2 Explorative research

This section discusses the results for the phases 1 (awareness or need) and 2 (enquire or planning).

4.2.1 Predictive value awareness or need stage

The keywords used in this stage are 'Amsterdam' and 'Travel Amsterdam' for both Germany and the United Kingdom. It stands out that for the keyword 'Travel Amsterdam' Google Trends provides small values (0 up and till 3) while it does provide higher values for the key word 'Amsterdam' (0 up and till 70).

Before the regression analysis could be started the researcher checked for multicollinearity for both Germany and the United Kingdom. The values that were found fall within the > 0.2 for tolerance and < 5.0 for VIF criteria (Grande, 2015) which indicates very low multicollinearity. The SPSS output for this phase can be found in appendix II.

The multiple regression analysis for Germany in phase 1 led to the following regression formula: *Passes (GER) = - 35.862 + 2.586 * A'dam + -40.016 * Travel A'dam.* The corresponding adjusted R-square indicates that 46.5 % of the variation in passes (GER) can be explained by the independent variables together. PSPSS is 0.000 which is smaller than the significance level alpha (α = 0.05) which indicates that at least one of the independent variables predicts Passes (GER). Since the p-value for Travel A'dam is bigger than alpha this variable is not significant and can be left out of the model and therefore a new regression analysis, with only the significant variables, was done. This led to the following formula: ***Passes (GER) = -17.101 + 2.181 * A'dam***. PSPSS = 0.000 for the model which means p < 0.001 which is smaller than alpha (α = 0.05). Also, PSPSS = 0.000 for the variable A'dam which is also smaller than alpha (α = 0.05). This indicates significance for the model. The adjusted r-square states that 45.2 % of the variation in Passes (GER) can be explained by the model.

The regression formula for phase 1 of the United Kingdom is the following: *Passes (UK) = 138.390 + 1.127 * A'dam + -15.860 * Travel A'dam.* Adjusted R-square indicates that 6.9 % of the variation in passes (UK) can be explained by the independent variables together. However, p = 0.115 for the model which is bigger than alpha (α = 0.05) which indicates that none of the independent variables are predictors of passes (UK). Therefore, the presented formula is not significant and has no predictive value for passes (UK).

The multiple regression analysis SPSS output can be found in appendix III for both Germany and the United Kingdom. An overview of these results is presented in Table 1 Overview results phase 1Table 1.

Table 1 Overview results phase 1

Phase 1				
	Regression coefficients	*P-value*	*B*	*Adjusted R^2*
Germany				
A'dam	2.181	0.000		
Model		0.000	-17.101	0.452
United Kingdom				
Model		0.115		

4.2.2 Predictive value enquire or planning stage

For this stage, the used keywords are 'Hotel Amsterdam', 'Flight Amsterdam', 'Train Amsterdam', 'Holiday Amsterdam', 'City trip Amsterdam', 'Visit Amsterdam' and 'Tourist info Amsterdam' for both Germany and the United Kingdom. It stands out that the keyword 'Tourist Info Amsterdam' provides the value 0 over the total timespan of this research for both Germany and the United Kingdom which indicates that this keyword is not very useful for forecasting tourism demand in Amsterdam. The keyword 'City trip Amsterdam' did not provide any data since this keyword did not generate any values. Google Trends was unable to provide any trend data on the bases of this keyword. The system reported that the search term does not provide enough search query data to display within Google Trends. This does not necessarily indicate that these terms are useless in forecasting tourism demand in Amsterdam. It does indicate that Google Trends does not possess enough search query data within the boundaries of this research. For the other keywords used Google Trends did generate data that is useable within this research.

Also for this keywords multicollinearity had to be assessed. All the values found for both Germany and the United Kingdom meet the criteria of tolerance > 0.2 and VIF < 5.0 which indicates that there is very low multicollinearity for both countries in phase 2. The multicollinearity SPSS output of phase 2 can be found in appendix IV.

As mentioned, for the keyword 'Tourist Info Amsterdam' Google Trends generated a constant value of 0 for both the German and the United Kingdom tourist origin. Therefore, we deleted this variable for the multiple regression analysis. For phase 2 with tourists with a German origin the following regression formula was generated: *Passes (GER) = - 15.822 + 1.269 * hotel A'dam + 15.111 * flight A'dam + - 15.999 * holiday A'dam + 1.145 * train A'dam + 0.708 * visit A'dam.* The adjusted R-square indicates that 40.2 % of the variation in passes (GER) can be explained by the independent variables together. The p-value is 0.001 which is smaller than alpha (α = 0,05) which indicates that at least one of the independent variables predicts passes (GER). Since the individual p-values for flight

A’dam (p = 0.458), holiday A’dam (p = 0.205) and visit A’dam (p = 0.432) are bigger than alpha these variables can be left out of the model. A new multiple regression analysis with only the significant variables led to the following formula: ***Passes (GER) = -10.011 + 1.276 * Hotel A’dam + 1.049 * Train A’dam.*** PSPSS of the model = 0.000 which means P < 0.001 which is smaller than alpha (α = 0.05). The p-value for Hotel A’dam = 0.017 and for Train A’dam the p-value = 0.019, this is both smaller than alpha (α = 0.05) which means the model is significant. The adjusted r-square reports that 39.1 % of the variation in passes (GER) can be explained by the model.

The regression formula for tourists with an United Kingdom origin is: Passes (UK) = 100.295 + -0.730 * hotel A’dam + 13.409 * flight A’dam + 2.927 * holiday A’dam + 0.411 * train A’dam + -1.368 * visit A’dam. The adjusted r-square indicates that 26.1 % of the variation in passes (UK) can be explained by the independent variables together. The p-value is 0.014 which is smaller than alpha (α = 0,05) which indicates that at least one of the independent variables predicts passes (UK). Since the individual p-values for hotel A’dam (p = 0.207), holiday A’dam (p = 0.580), Train A’dam (p = 0.414) and visit A’dam (p = 0.523) are bigger than alpha these variables can be left out of the model. A new multiple regression analysis with only the significant variables led to the following formula: **Passes (UK) = 98.244 + 12.276 * Flight A’dam**. The p-value for the model is p = 0.001 which is smaller than alpha (α = 0.05), the p-value for the variable flight A’dam = 0.001 which is also smaller than alpha. This indicates that the model is significant. The adjusted r-square reports that 28.0 % of the variation in passes (UK) can be explained by the model.

The multiple regression analysis SPSS output of phase 2 for both Germany and the United Kingdom can be found in appendix V. An overview of the results is presented in Table 2.

Table 2 Overview results phase 2

Phase 2				
	Regression coefficients	*P-value*	*B*	*Adjusted R^2*
Germany				
Hotel A’dam	1.276	0.017		
Train A’dam	1.049	0.019		
Model		0.000	-10.011	0.391
United Kingdom				
Flight A’dam	12.276	0.001		
Model		0.001	98.244	0.280

4.3 Inductive research

This section presents the results for the third, data driven, phase of this research.

4.3.1 Google Correlate generated keywords

For this inductive part of the research keywords are derived from the data with the use of Google Correlate. The data concerning the night passes in hotels in the City of Amsterdam was uploaded to

Google Correlate which then presented correlating search queries. For tourists with a German origin ‘hop on hop off’ (r = 0.9603), ‘sehenswürdigkeiten in der nahe’ (r = 0.9423) which translates to things to see in the area and ‘one day trip’ (r = 0.9374) were derived. Keywords with a high correlation with night passes in hotels in the City of Amsterdam with an origin in the United Kingdom are ‘cosas que ver’ (r = 0.8394) which translates to ‘things to do’ and ‘cosas que haser’ (r = 0.8376) which translates to ‘what to do’. For the United Kingdom, no more possibly meaningful correlations were found. A representation of all the correlations found for both Germany and the United Kingdom can be found in appendix VI followed by an overview of the keywords that will be used in the different phases in appendix VII.

4.3.2 Predictive value data driven research

For this third, inductive phase, the keywords used are ‘hop on hop off’, ‘sehenswürdigkeiten in der nahe’ and ‘one day trip’ for Germany and ‘cosas que ver’ and ‘cosas que haser’ for the United Kingdom.

For the term ‘hop on hop off’ Google Trends has provided the researcher with search query data to analyze. The terms ‘sehenswürdigkeiten in der nahe’ and ‘one day trip’ on the other hand did not since for both of these keywords all search query values are a constant 0. This indicates that these two keywords are not useful for forecasting tourism demand in Amsterdam.

Although Google Correlate measured very strong correlations between the keywords ‘cosas que ver’ (r = 0.8394) and ‘cosas que haser’ (r = 0.8376) for overnight stays in hotels in the City of Amsterdam from tourist with an United Kingdom origin, Google Trends was unable to provide any trend data on the bases of these keywords. The system reported that the search terms do not provide enough search query data to display within Google Trends. This does not necessarily indicate that these terms are useless in forecasting tourism demand in Amsterdam. It does indicate that Google Trends does not possess enough search query data within the boundaries of this research. Which in this case were our time filter set to 2014 - 2016, location filter set to United Kingdom and the category filters set to travel.

Since Google Correlate searches for correlations without any form of time lag (Google, 2017) the results for the previous correlation did show search queries that are related to leisure or free time, but not necessarily to night passes in Amsterdam. Since tourists spend time planning their vacation it is possible that, when the data is aggregated to form where it seems that there is a time lag, more meaningful correlations will be found. Therefore, the data was aggregated for both countries with a 1 month time lag, a 2 month time lag, a 3 month time lag, a 4 month time lag, a 5 month time lag and a 6 month time lag and then inserted into Google Correlate. These time lags where all tested since the actual time lag between searching or planning a vacation could not be derived from theory. All the correlations Google Correlate found for these time lags for both countries can be found in appendix VIII.

For Germany three search queries that might be meaningful were found. Within the 1 month time lag these queries are ‘wohnung mit garten mieten’ (renting a house with a garden) with a correlation of r = 0.9043 and ‘kurzurlaub niederlande’ (short vacation the Netherlands) with a correlation of r = 0.8910.

Within the 2 and 3 month time lag the possibly meaningful query ‘reispass baby’ (passport baby) was found with a correlation of r = 0.8766. However, when the keyword ‘wohnung mit garten mieten’ was entered into Google Trends the program was unable to provide any data due to too little search query data to display within Google Trends. The keyword ‘kurzurlaub niederlande’ did generate Google Trends, this indicates that it might have predictive value for forecasting tourism demand in the City of Amsterdam. The last generated keyword for Germany, ‘reispass baby’, also provided the researcher with Google Trends output, but that had the constant value of 0 which means it has no predictive value.

The possibly meaningful search query that was found for the United Kingdom falls within the 1 month time lag and is ‘Amsterdam all inclusive’. The keyword has a correlation of r = 0.8565 and does generate Google Trends data within the framework of this research which might indicate predictive value.

5 Conclusion and discussion

This chapter describes the conclusions that are drawn from the results, the discussion, the limitations of the research and the implications for further research.

5.1 Conclusion

For the first two phases of the research it can be concluded that 4 of the keywords derived from the theory have predictive value for forecasting tourism demand in the City of Amsterdam.

Within the first phase, the awareness or need phase, it can be concluded that the keyword 'Amsterdam' has predictive value for tourist with a German origin. With an adjusted r-square of 45.2 % it can be concluded that the predictive value of this keyword is significant. The first phase did not lead to any significant keywords with predictive value for tourists with an United Kingdom origin.

The second phase, the enquire or planning phase, led to more keywords with predictive value. For tourists from Germany it can be concluded that 39.1 % of the variation in the amount of night passes in hotels in the City of Amsterdam can be explained by the variables 'Hotel Amsterdam' and 'Train Amsterdam'. The amount of night passes from tourists with an United Kingdom origin can be forecasted with the use of Google Trends by the keyword 'flight Amsterdam' whereby 28.0 % of the variation can be explained by this keyword.

Recapitulating, when answering the research question *"To what extent is the data provided by Google Trends useful for forecasting night passes in hotels in Amsterdam?"* on the bases of the theory driven part of this research, the answer is that data provided by Google Trends can be useful for forecasting night passes in hotels in Amsterdam if the fitting keywords are used. The extent to which the explorative research indicates usefulness is on average 37.4 %.

The conclusion for the third phase of this research is that there are many correlating search queries, but only 2 with possible predictive value or even Google Trends output. The data generated keywords 'kurzurlaub niederlande' for Germany and 'Amsterdam all inclusive' for the United Kingdom have predictive value. From this part of the research the answer to the research question *"To what extent is the data provided by Google Trends useful for forecasting night passes in hotels in Amsterdam?"* is that the data generated from Google Trends has very low to zero usefulness for forecasting night passes in hotels in Amsterdam.

Concluding, in comparison with the current method of forecasting used by the City of Amsterdam these conclusions are useful in practice. Since the forecasts made by the City are not explicit and do not take nationalities into account the use of forecasting with Google Trends would increase the usability of the forecasts both for the City of- and the hospitality industry in Amsterdam.

5.2 Discussion

Shmueli (2010) states that blind predictions, without theory, are less likely to be effective. In the first two phases of this study theory on the customer journey in tourism is used. This enhances the validity of the finding of this study. However, the theories on customer journey in tourism are broad and lack of specific steps tourist take within this journey. The search terms derived from the theory are created by the researcher which, although they are based on the theory that is available, minimizes the effectiveness of the models created in this research. In the third phase, the data collection was data driven which led to few results. This might not have been a blind prediction, but it was also not a theory driven prediction. The data driven phase was, as Shmeili (2010) already suggested, less effective than the theory driven predictions.

5.3 Limitations and further research

Within this research, the comparison of Google Trends analysis is limited to Germany and the United Kingdom and narrowed down to 10 or 11 search terms per country for hotel night passes in the City of Amsterdam. The comparison of other countries, search terms and other destinations is left as endeavors for future researchers. Furthermore, the researcher did not segment the tourist in categories such as short or long stays or leisure or business visits since the data did not allow it. But since forecasting potential of Google Trends is confirmed in this research, future research might investigate these segments in more detail which can lead to more useful outcomes for the cities and hoteliers to act upon.

Future research can be perfected by not only using Google Trends data as the independent variable, but adding other factors that influence the customer journey in tourism such as economic state and political stability from the destination or origin country of the tourist. The consideration of other factors will enhance the generalizability of the research since Google Trends only provides absolute volume data. The actual queries remain unreported which affects the generalizability of this study because it remains unclear which number of users used the search queries.

The lack of meaningful Google Correlate output for this research in the start of phase 3 led to shifts in the data to fake a time lag. This generated some results which could not be analyzed further within the timeframe of this research. Future research might analyze this results in more detail which provide the theory with more data driven results, which also contribute to the customer journey theory for tourism.

Due to the lack of data on weekly bases of the depend variable, hotel night passes in the City of Amsterdam, the Google Trends data had to be aggregated. The aggregation was made under the assumption that search behavior of tourists is constant across all days of the week. Which is a flaw in the research but no better way was found, this is left endeavors for future researchers. The mandatory aggregation to monthly bases results in the researcher being unable to distinguish a time lag into the forecasting model. When more specific data would have been accessible the results of this research might be different and the model more effective for practical usage. However, this research did confirm the forecasting potential of Google Trends for forecasting tourism demand in Amsterdam. If future research pursues on this finding and analyses the forecasting potential on weekly or daily bases with an established time lag the new models will improve the ability of the City of Amsterdam to respond to tourism demand and maintain their dispersion policy. Furthermore, not only the City of Amsterdam can provide from these models, this research can also be done for other cities or even large hotel chains. This can lead to the ability of cities and hoteliers to respond adequately to tourism demand and more effective intervening in the customer journey of tourists.

References

Akal, M. (2004). Forecasting Turkey's revenues by ARMAX model. *Tourism Management, 25*(5), 565-580. doi:https://doi.org/10.1016/j.tourman.2003.08.001

Andrew, W., Cranage, D., & Lee, C. (1990). Forecasting Hotel Occupancy Rates with Time Series Models: An Empirical Analyses. *Journal of Hospitality & Tourism Research, 14*(2), 173-182. doi:10.1177/109634809001400219

Barreira, N., Godinho, P., & Melo, P. (2013). Nowcasting unemployment rate and new car sales in south-western Europe with Google Trends. *Netnomics, 14*, 129-165. doi:DOI 10.1007/s11066-013-9082-8

Bollier, D., & Firestone, C. (2010). *The promise and peril of big data.* Washington DC: The Aspen Institute.

Box, G., & Jenkins, G. (1970). *Time series analysis, forecasting and control.* San Fransisco: Holden Day.

Burger, C., Dohnal, M., Kathrada, M., & Law, R. (2001). A practitioners guide to time-series methods for tourism demand forecasting — a case study of Durban, South Africa. *Tourism Management, 22*(4), 403-409. doi:https://doi.org/10.1016/S0261-5177(00)00068-6

CBS. (2016). *Trendsrapport toerisme, recreatie en vrije tijd 2016.* Retrieved January 25, 2017, from file:///C:/Users/Gebruiker/Downloads/Trendsrapport%20toerisme%20recreatie%20en%20vrij%20tijd%202016.pdf

CBS. (2017, 07 10). Hotels; Gasten, Overnachtingen, Woonland, Regio. Retrieved from http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=82061NED&D1=1&D2=0-2,4,8,11,16,35,42&D3=23&D4=50,67,84&HDR=G3,G2,T&STB=G1&VW=T

CBS. (2017, 06 05). *StatLine*. Retrieved from CBS StatLine: http://statline.cbs.nl/statweb/

Cho, V. (2001). Tourism Forecasting and its Relationships with Leading Economic Indicators. *Journal of Hospitality & Tourism Research, 25*(4), 399-420. doi:https://doi.org/10.1177/109634800102500404

Cho, V. (2003). A comparison of three different approaches to tourist arrival forecasting. *Tourism Management, 24*(3), 323-330. doi:https://doi.org/10.1016/S0261-5177(02)00068-7

Choi, H., & Varian, H. (2012). Predicting the present with Google Trends. *Economic Record, 88*(s1), 2-9. doi:10.1111/j.1475-4932.2012.00809.x

City of Amsterdam. (2017). *Toerisme in Amsterdam.* Retrieved January 25, 2017, from http://www.ois.amsterdam.nl/toerisme

Clements, M., & Hendry, D. (1998). *Forecasting economic time series.* Cambridge: Cambridge University Press.

Cowpertwait, P., & Metcalfe, A. (2009). *Introductory time series with R.* New York: NY Springer-Verlag.

Cryer, J., & Chan, K. (2008). *Time Series Analysis. With Applications to R. Design.* New York: Springer.

Even, A., Shankaranarayanan, G., & Berger, P. (2010). Managing the Quality of Marketing Data: Cost/benefit Tradeoffs and Optimal Configuration. *Jounal of Interactive Marketing, 24*(3), 209-221. doi:https://doi.org/10.1016/j.intmar.2010.01.001

Fesemaier, D., & Jeng, J. (2000). Assessing structure in the pleasure trip planning process. *Tourism Analysis, 5*(1), 13-27.

Frechtling, D. (2011). *Forecasting Tourism Demand: Methods and Strategies.* Oxford: Butterworth Heinemann.

Fresenmaier, D., Xiang, Z., Pan, B., & Law, R. (2010). A Framework of Search Engine Use for Travel Planning. *Journal of Travel Research, 50*(6), 587-601. doi:10.1177/0047287510385466

Gemeente Amsterdam. (2014). *Toerisme in Amsterdam 2013-2014.* Retrieved 10 15, 2017, from OIS Amsterdam: https://www.ois.amsterdam.nl/pdf/2014_toerisme.pdf

Gemeente Amsterdam. (2015). *Toerisme in MRA 2014-2015.* Retrieved 10 15, 2017, from OIS Amsterdam: https://www.ois.amsterdam.nl/pdf/2016_toerisme%20mra%202014%202015.pdf

Gemeente Amsterdam. (2016). *Toerisme in MRA 2015-2016.* Retrieved 10 15, 2017, from OIS Amsterdam: https://www.ois.amsterdam.nl/pdf/2017_toerisme%20mra%202015%202016.pdf

Goh, C., & Law, R. (2002). Modeling and forecasting tourism demand for arrivals with stochastic nogstationary seasonality and intervention. *Tourism Management*, *23*(5), 499-510. doi:https://doi.org/10.1016/S0261-5177(02)00009-2

Google. (2017, 07 15). *Google Correlate Tutorial*. Retrieved 07 15, 2017, from Google Correlate: https://www.google.com/trends/correlate/tutorial

Google. (2017). *How Trends Data is Adjusted*. Retrieved January 24, 2017, from https://support.google.com/trends/answer/4365533?hl=en{&}ref{_}topic=4365599

Grande, T. (2015, 12 1). *Understanding and identifying multicolinearity in regression using SPSS.* Retrieved 23 07, 2017, from YouTube: https://www.youtube.com/watch?v=pZsOn6wnGSo

Hashem, I., Yaqoob, I., Anuar, N., Mokhtar, S., Gani, A., & Khan, S. (2015). The rise of "big data" on cloud computing: Review and open research issues. *Information Systems, 47*(1), 98-115. doi:10.1016/j.is.2014.07.006

Hassani, H., & Silva, E. (2015). Forcasting with Big Data: A Review. *Annals of Data Science, 2*(1), 5-19. doi:10.1007/s40745-015-0029-9

Hendrik, H., & Perdana, D. (2014). Trip guidance: A linked data based mobile tourists guide. *Advanced Science Letters, 20*(1), 75-79. doi:10.1166/asl.2014.5285

Hofacker, C., Malthouse, E., & Sultan, F. (2016). Big Data and consumer behavior: imminent opportunities. *Journal of Consumer Marketing, 33*(2), 89-97. doi:0.1108/JCM-04-2015-1399

Höpken, W., Ernesti, D., Fuchs, M., Kronenberg, K., & Lexhagen, M. (2017). Big Data as Input for Predicting Tourist Arrivals. In R. Schegg, & B. Stangl, *Information and Communication Technologies in Tourism 2017* (pp. 187-199). Cham: Springer International Publishing AG 2017.

Horner, S., & Swarbrooke, J. (2011). *Consumer behaviour in Tourism.* Amsterdam: Elsevier Ltd.

Hu, N., Bose, I., Koh, N., & Liu, L. (2012). Manipulation of online reviews: An analysis of ratings, readability and sentiment. *Decison Support Systems, 52*(3), 674-684. doi:https://doi.org/10.1016/j.dss.2011.11.002

ING Economisch Bureau. (2016, March). *Toerisme groeit in steeds meer provincies.* Retrieved February 2, 2017, from https://www.ing.nl/media/ING%20-%20Toerisme%20groeit%20in%20meeste%20provincies%20-%20maart%202016_tcm162-101300.pdf

Koljonen, E., & Reid, R. (2000). Walk-through audit provides focus for service improvements for Hong Kong Law Firm. *Managing Service Quality, 10*(1), 32-46.

Kon, S., & Turner, L. (2005). Neural Network Forecasting of Tourism Demand. *Tourism Economics, 11*(3), 301-328. doi:https://doi.org/10.5367/000000005774353006

Kulendran, N., & Wong, K. (2005). Modeling Seasonality in Tourism Forecasting. *Journal of Travel Research, 44*(2), 163-170. doi:10.1177/0047287505276605

Kulendran, N., & Wong, K. (2005). Modeling Seasonality in Tourism Forecasting. *Journal of Travel Research, 44*(2), 163-170. doi:https://doi.org/10.1177/0047287505276605

Lane, M. (2007). The Visitor Journey: the new road to success. *International Journal of Contemporary Hospitality Management, 19*(3), 248-254. doi:http://dx.doi.org/10.1108/09596110710739949

Law, R. (2000). Back-propagation learning in improving the accuracy of neural network-based tourism demand forecasting. *Tourism Management, 21*(4), 331-340. doi:https://doi.org/10.1016/S0261-5177(99)00067-9

Lewis, E. (1903). *Advertising department: catch-line and argument.* Toronto: The Book-Keeper.

Li, X., Pan, B., Law, R., & Huang, X. (2017). Forecasting tourism demand with composite search index. *Tourism Management, 59*, 57-66. doi:10.1016/j.tourman.2016.07.005

Makridakis, S., Wheelwright, S., & Hyndman, R. (2005). *Forecasting: methods and applications.* Hoboken, New Jersey: Wiley.

Mayer-Schönberger, V., & Cukier, K. (2013). *Big Data: A Revolution That Will Transform How We Live, Work and Think.* Boston New York: Houghton Mifflin Hartcourt.

Mohebbi, M., Vanderkam, D., Kodysh, J., Schonberger, R., Choi, H., & Kumar, S. (2011, 06 09). *Google Correlate Whitepaper.* Retrieved from http://anokacountyhistory.org/wp-content/uploads/whitepaper.pdf

NBTC Holland Marketing. (2016). *Kerncijfers 2016 Gastvrijheidseconomie.* Retrieved February 2, 2017, from http://kerncijfers.nbtc.nl/nl/magazine/category.html

NBTC Holland Marketing. (2016). *Kerncijfers gastvrijheidseconomie 2016.* Retrieved 07 21, 2017, from http://kerncijfers.nbtc.nl/nl/magazine/11909/821275/reisproces_-_vervoer.html

NBTC Holland Marketing. (2017). *Home page*. Retrieved January 25, 2017, from http://www.nbtc.nl/nl/homepage/hollandmarketing.htm

Nenonen, S., Rasila, H., Junnonen, J., & Kärnä, S. (2008). Customer Journey - a method to investigate user experience. *Proceedings of the Euro FM Conference Manchester*, (pp. 54-63). Manchester.

Normandeau, K. (2013). Beyond volume, variety and velocity is the isue of big data veracity. *Inside Big Data*.

OESO. (2016). *The Contours of World Development*. Retrieved January 25, 2017, from http://www.theworldeconomy.org/index.htm

Pollett, S., Boscardin, W., Azziz-Baumgartner, E., Tinoco, Y., Soto, G., Romero, C., & Rutherford, G. (2017). Evaluating Google Flu Trends in Latin America: important lessons for the next phase of digital disease detection. *Clincal Infectious Diseases, 64*(1), 34-41. doi:10.1093/cid/ciw657

Prideaux, B., Laws, E., & Faulkner, B. (2003). Events in Indonesia: exploring the limits to formal tourism trends forecasting methods in complex crisis situations. *Tourism Management, 24*(4), 475-487. doi:https://doi.org/10.1016/S0261-5177(02)00115-2

Rezaei, S., Ali, F., Amin, M., & Jayashree, S. (2016). Online impulse buying of tourism products. *Journal of Hospitality and Tourism Technology, 7*(1), 60-83. doi:10.1108/JHTT-03-2015-0018

Rivera, R. (2016). A dynamic linear model to forecast hotel registrations in Puerto Rico using Google Trends data. *Tourism Management, 57*, 12-20. doi:10.2016/j.tourman.2016.04.008

Shmueli, G. (2010). To explian or to predict? *Statistical Science, 25*(3), 289-310. doi:10.1214/10-STS330

Smeral, E. (2004). Long-Term Forecasts for International Tourism. *Tourism Economics, 10*(2), 145-166. doi:https://doi.org/10.5367/000000004323142461

Smeral, E., & Wüger, M. (2005). Does Complexity Matter? Methods for Improving Forecasting Accuracy in Tourism: The Case of Austria. *Journal of Travel Research, 44*(1), 100-110. doi:https://doi.org/10.1177/0047287505276596

Song, H., & Li, G. (2008). Tourism demand modelling and forecasting - A review of recent years. *Tourism Management, 29*, 203-220. doi:10.1016/j.tourman.2007.07.016

Song, H., & Liu, H. (2017). Predicting Tourist Demand Using Big Data. In Z. Xiang, & D. Fresenmaier, *Analystics in Smart Tourism Design* (pp. 13-29). Cham: Springer International Publishing Switzerland 2017.

Song, H., Li, G., Witt, S., & Fei, B. (2010). Tourism Demand Modeling and Forecasting: How Should Demand Be Measured? *Tourism Economics*(16), 63-81. doi:10.5367/000000010790872213

Statstutor. (2017, 07 15). *Pearson's correlation.* Retrieved from Statstutor: http://www.statstutor.ac.uk/resources/uploaded/pearsons.pdf

Steinbauer, A., & Werthner, H. (2007). Consumer Behaviour in e-Tourism. In M. Sigala, L. Mich, & J. Murphy, *Information and Communication Technologies in Tourism 2007* (pp. 65-76). Vienna: Springer. doi:10.1007/978-3-211-69566-1_7

Strauss, B., & Weinlich, B. (1997). Process-oriented measurement of service quality. *European Journal of Marketing, 31*(1), 33-55. doi:http://dx.doi.org/10.1108/03090569710157025

Swarbrook, J., & Horner, S. (1999). *Consumer Behaviour in Tourism.* Oxford: Butterworth and Heinemann.

Teng, Y., Bi, D., Xie, G., Jin, Y., Huang, Y., Lin, B., . . . Tong, Y. (2017). Dynamic Forecasting of Zika Epidemics Using Google Trends. *Plos One, 12*(1). doi:10.1371/journal.pone.0165085

University of Twente. (2016). *Master of Science in Business Administration; MSc BA Thesis Manual 2016-2017.* Retrieved December 12, 2016, from https://www.utwente.nl/ba/master/thesis-and-graduation/

UNWTO. (2016). *Close to one billion international tourists in the first nine months of 2016*. Retrieved January 25, 2017, from http://media.unwto.org/press-release/2016-11-07/close-one-billion-international-tourists-first-nine-months-2016

Varian, H. (2014). Big Data: New Tricks for Econometrics. *Journal of Economic Perspective, 28*(2), 3-28. doi:10.1257/jep.28.2.3

Verhoeven, N. (2007). *Wat is onderzoek? Praktijk methoden en technieken voor het hoger onderwijs.* . Den Haag: Boom Onderwijs.

Vigen, T. (2017, 07 15). *Spurious Correlations*. Retrieved from tylervigen.com: http://www.tylervigen.com/about

Witt, S., & Witt, C. (1995). Forecasting tourism demand: A review of empirical research. *International Journal of Forecasting, 11*, 447-475. doi:10.1016/0169-2070(95)00591-7

Witt, S., Song, H., & Wanhill, S. (2004). Forecasting Tourism-Generated Employment: The Case of Denmark. *Tourism Economics, 10*(2), 167-176. doi:https://doi.org/10.5367/000000004323142407

Xu, X., & Reed, M. (2017). Perceived pollution and inbound tourism in China. *Tourism Management Perspectives, 21*, 109-112. doi:10.1016/j.tmp.2016.12.006

Yang, X., Pan, B., Evans, J., & Lv, B. (2015). Forecasting Chinese tourist volume with search engine data. *Tourism Management, 46*, 386-397. doi:10.1016/j.tourman.2014.07.019

Appendices

Appendix I

Hotel night passes in the City of Amsterdam (x1000)	2014	% off foreign countries total	2015	% off foreign countries total	2016	% off foreign countries total
All countries	12.538		12,899		13,983	
The Netherlands	2,341		2,238		2,192	
Foreign countries total	10.197		10,661		11,790	
Belgium	335	3%	308	3%	355	3%
Germany	1,112	11%	1,189	11%	1,469	12%
France	700	7%	674	6%	741	6%
Italy	689	7%	722	7%	733	6%
United Kingdom	1,905	19%	2,046	19%	2,296	19%
United States of America	1,130	11%	1,147	11%	1,354	11%

Appendix II

SPSS output multicollinearity phase 1 for Germany:

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	Travel A'dam (GER), A'dam (GER)[b]	.	Enter

a. Dependent Variable: Passes (GER)

b. All requested variables entered.

Coefficients[a]

Model		Collinearity Statistics	
		Tolerance	VIF
1	A'dam (GER)	,524	1,909
	Travel A'dam (GER)	,524	1,909

a. Dependent Variable: Passes (GER)

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions		
				(Constant)	A'dam (GER)	Travel A'dam (GER)
1	1	2,513	1,000	,00	,00	,04
	2	,480	2,289	,01	,00	,52
	3	,007	18,868	,99	1,00	,44

a. Dependent Variable: Passes (GER)

SPSS output multicollinearity phase 1 for the United Kingdom:

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	Travel A'dam (UK), A'dam (UK)[b]	.	Enter

a. Dependent Variable: Passes (UK)

b. All requested variables entered.

Coefficients[a]

Model		Collinearity Statistics	
		Tolerance	VIF
1	A'dam (UK)	,802	1,248
	Travel A'dam (UK)	,802	1,248

a. Dependent Variable: Passes (UK)

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions		
				(Constant)	A'dam (UK)	Travel A'dam (UK)
1	1	2,971	1,000	,00	,00	,00
	2	,020	12,262	,21	,07	,96
	3	,009	18,087	,79	,93	,04

a. Dependent Variable: Passes (UK)

Appendix III

Multiple regression SPSS output phase 1 for Germany, initial model:

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	Travel A'dam (GER), A'dam (GER)[b]	.	Enter

a. Dependent Variable: Passes (GER)

b. All requested variables entered.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,704[a]	,495	,465	20,92908

a. Predictors: (Constant), Travel A'dam (GER), A'dam (GER)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	14176,353	2	7088,176	16,182	,000[b]
	Residual	14454,870	33	438,026		
	Total	28631,222	35			

a. Dependent Variable: Passes (GER)

b. Predictors: (Constant), Travel A'dam (GER), A'dam (GER)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-35,862	28,043		-1,279	,210
	A'dam (GER)	2,586	,540	,819	4,792	,000
	Travel A'dam (GER)	-40,016	36,840	-,186	-1,086	,285

a. Dependent Variable: Passes (GER)

Multiple regression SPSS output phase 1 for Germany:

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	A'dam (GER)[b]	.	Enter

a. Dependent Variable: Passes (GER)

b. All requested variables entered.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,691[a]	,477	,462	20,98436

a. Predictors: (Constant), A'dam (GER)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	13659,554	1	13659,554	31,020	,000[b]
	Residual	14971,668	34	440,343		
	Total	28631,222	35			

a. Dependent Variable: Passes (GER)

b. Predictors: (Constant), A'dam (GER)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-17,101	22,151		-,772	,445
	A'dam (GER)	2,181	,392	,691	5,570	,000

a. Dependent Variable: Passes (GER)

Multiple regression SPSS output phase 1 for the United Kingdom, initial model:

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	Travel A'dam (UK), A'dam (UK)[b]	.	Enter

a. Dependent Variable: Passes (UK)

b. All requested variables entered.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,350[a]	,123	,069	19,79119

a. Predictors: (Constant), Travel A'dam (UK), A'dam (UK)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1807,194	2	903,597	2,307	,115[b]
	Residual	12925,806	33	391,691		
	Total	14733,000	35			

a. Dependent Variable: Passes (UK)

b. Predictors: (Constant), Travel A'dam (UK), A'dam (UK)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	138,390	25,298		5,470	,000
	A'dam (UK)	1,127	,528	,389	2,134	,040
	Travel A'dam (UK)	-15,860	13,576	-,213	-1,168	,251

a. Dependent Variable: Passes (UK)

Appendix IV

SPSS output multicollinearity phase 2 for Germany:

Warnings

For models with dependent variable Passes (GER), the following variables are constants or have missing correlations: TI A'dam (GER). They will be deleted from the analysis.

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	Visit A'dam (GER), Holiday A'dam (GER), Flight A'dam (GER), Train A'dam (GER), Hotel A'dam (GER)[b]	.	Enter

a. Dependent Variable: Passes (GER)

b. All requested variables entered.

Coefficients[a]

Model		Collinearity Statistics	
		Tolerance	VIF
1	Hotel A'dam (GER)	,561	1,783
	Flight A'dam (GER)	,747	1,338
	Holiday A'dam (GER)	,778	1,285
	Train A'dam (GER)	,615	1,626
	Visit A'dam (GER)	,744	1,344

a. Dependent Variable: Passes (GER)

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions					
				(Constant)	Hotel A'dam (GER)	Flight A'dam (GER)	Holiday A'dam (GER)	Train A'dam (GER)	Visit A'dam (GER)
1	1	5,662	1,000	,00	,00	,00	,00	,00	,01
	2	,216	5,123	,00	,00	,00	,01	,00	,82
	3	,058	9,903	,01	,00	,22	,18	,17	,00
	4	,044	11,319	,01	,00	,00	,70	,31	,03
	5	,012	22,091	,43	,16	,77	,04	,48	,11
	6	,009	24,818	,55	,83	,00	,06	,03	,04

a. Dependent Variable: Passes (GER)

SPSS output multicollinearity phase 2 for the United Kingdom:

Warnings

For models with dependent variable Passes (UK), the following variables are constants or have missing correlations: TI A'dam (UK). They will be deleted from the analysis.

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	Visit A'dam (UK), Train A'dam (UK), Flight A'dam (UK), Hotel A'dam (UK), Holiday A'dam (UK)[b]	.	Enter

a. Dependent Variable: Passes (UK)

b. All requested variables entered.

Coefficients[a]

Model		Collinearity Statistics	
		Tolerance	VIF
1	Hotel A'dam (UK)	,425	2,352
	Flight A'dam (UK)	,559	1,788
	Holiday A'dam (UK)	,388	2,579
	Train A'dam (UK)	,550	1,817
	Visit A'dam (UK)	,621	1,609

a. Dependent Variable: Passes (UK)

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions					
				(Constant)	Hotel A'dam (UK)	Flight A'dam (UK)	Holiday A'dam (UK)	Train A'dam (UK)	Visit A'dam (UK)
1	1	5,927	1,000	,00	,00	,00	,00	,00	,00
	2	,026	15,195	,13	,00	,01	,02	,04	,62
	3	,021	16,649	,12	,03	,01	,40	,00	,26
	4	,010	23,893	,08	,46	,57	,05	,00	,00
	5	,009	26,293	,04	,46	,40	,15	,28	,12
	6	,007	29,023	,63	,05	,01	,39	,67	,00

a. Dependent Variable: Passes (UK)

Appendix V

SPSS output multiple regression analysis phase 2 for Germany, initial model:

Warnings

For models with dependent variable Passes (GER), the following variables are constants or have missing correlations: TI A'dam (GER). They will be deleted from the analysis.

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	Visit A'dam (GER), Holiday A'dam (GER), Flight A'dam (GER), Train A'dam (GER), Hotel A'dam (GER)[b]	.	Enter

a. Dependent Variable: Passes (GER)

b. All requested variables entered.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,698[a]	,487	,402	22,11777

a. Predictors: (Constant), Visit A'dam (GER), Holiday A'dam (GER), Flight A'dam (GER), Train A'dam (GER), Hotel A'dam (GER)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	13955,353	5	2791,071	5,705	,001[b]
	Residual	14675,870	30	489,196		
	Total	28631,222	35			

a. Dependent Variable: Passes (GER)

b. Predictors: (Constant), Visit A'dam (GER), Holiday A'dam (GER), Flight A'dam (GER), Train A'dam (GER), Hotel A'dam (GER)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-15,822	30,655		-,516	,610
	Hotel A'dam (GER)	1,269	,583	,380	2,176	,038
	Flight A'dam (GER)	15,111	20,092	,114	,752	,458
	Holiday A'dam (GER)	-15,999	12,352	-,192	-1,295	,205
	Train A'dam (GER)	1,145	,468	,408	2,449	,020
	Visit A'dam (GER)	,708	,889	,121	,796	,432

a. Dependent Variable: Passes (GER)

SPSS output multiple regression analysis phase 2 for Germany:

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	Train A'dam (GER), Hotel A'dam (GER)[b]	.	Enter

a. Dependent Variable: Passes (GER)

b. All requested variables entered.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,652[a]	,426	,391	22,32108

a. Predictors: (Constant), Train A'dam (GER), Hotel A'dam (GER)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	12189,615	2	6094,807	12,233	,000[b]
	Residual	16441,607	33	498,231		
	Total	28631,222	35			

a. Dependent Variable: Passes (GER)

b. Predictors: (Constant), Train A'dam (GER), Hotel A'dam (GER)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-10,011	25,508		-,392	,697
	Hotel A'dam (GER)	1,276	,506	,382	2,519	,017
	Train A'dam (GER)	1,049	,425	,374	2,468	,019

a. Dependent Variable: Passes (GER)

SPSS output multiple regression analysis phase 2 for the United Kingdom, initial model:

Warnings

For models with dependent variable Passes (UK), the following variables are constants or have missing correlations: TI A'dam (UK). They will be deleted from the analysis.

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	Visit A'dam (UK), Train A'dam (UK), Flight A'dam (UK), Hotel A'dam (UK), Holiday A'dam (UK)[b]	.	Enter

a. Dependent Variable: Passes (UK)

b. All requested variables entered.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,605[a]	,366	,261	17,63964

a. Predictors: (Constant), Visit A'dam (UK), Train A'dam (UK), Flight A'dam (UK), Hotel A'dam (UK), Holiday A'dam (UK)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5398,291	5	1079,658	3,470	,014[b]
	Residual	9334,709	30	311,157		
	Total	14733,000	35			

a. Dependent Variable: Passes (UK)

b. Predictors: (Constant), Visit A'dam (UK), Train A'dam (UK), Flight A'dam (UK), Hotel A'dam (UK), Holiday A'dam (UK)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	100,295	24,952		4,019	,000
	Hotel A'dam (UK)	-,730	,566	-,287	-1,289	,207
	Flight A'dam (UK)	13,409	4,354	,599	3,080	,004
	Holiday A'dam (UK)	2,927	5,233	,131	,559	,580
	Train A'dam (UK)	,411	,496	,162	,829	,414
	Visit A'dam (UK)	-1,368	2,120	-,119	-,646	,523

a. Dependent Variable: Passes (UK)

SPSS output multiple regression analysis phase 2 for the United Kingdom:

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	Flight A'dam (UK)[b]	.	Enter

a. Dependent Variable: Passes (UK)

b. All requested variables entered.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,548[a]	,300	,280	17,41302

a. Predictors: (Constant), Flight A'dam (UK)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4423,744	1	4423,744	14,590	,001[b]
	Residual	10309,256	34	303,213		
	Total	14733,000	35			

a. Dependent Variable: Passes (UK)

b. Predictors: (Constant), Flight A'dam (UK)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	98,244	19,915		4,933	,000
	Flight A'dam (UK)	12,276	3,214	,548	3,820	,001

a. Dependent Variable: Passes (UK)

Appendix VI

Correlations found for Germany:

Compare US states
Compare weekly time series
Compare monthly time series

Shift series 0 months
Country:
Germany

Documentation
Comic Book
FAQ
Tutorial
Whitepaper
Correlate Algorithm

Correlate Labs
Search by Drawing

Correlated with **Duitsland**

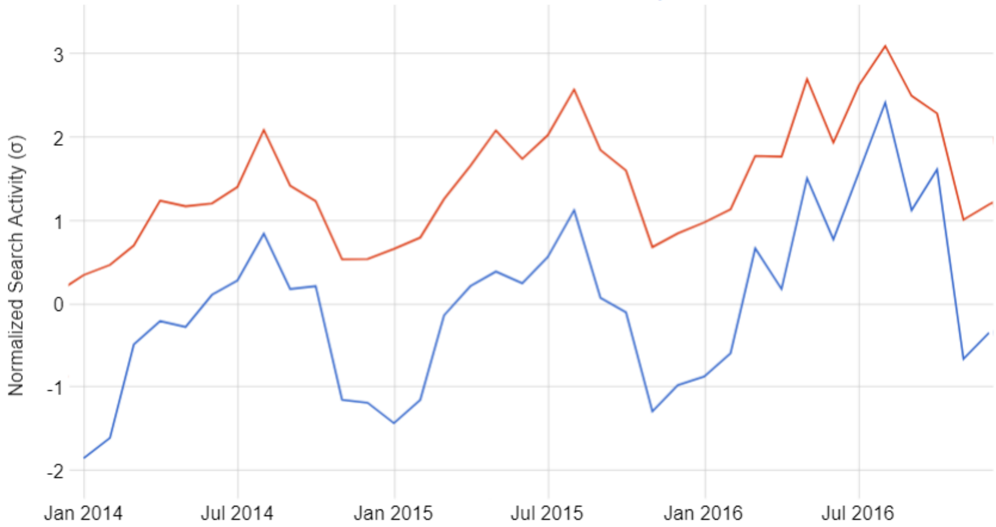
0.9603 hop on hop off
0.9597 hop off
0.9593 hop on
0.9591 cafe amsterdam
0.9486 fischbude
0.9471 parken wohnmobil
0.9461 stellplatz
0.9454 prerow leuchtturm
0.9453 kennzeichen mk
0.9438 parking
0.9429 rock cafe amsterdam
0.9423 sehenswürdigkeiten in der nähe
0.9417 hard rock cafe amsterdam
0.9416 kassel sehenswürdigkeiten
0.9407 hop on hop off hamburg
0.9394 wetter bratislava
0.9389 outdoor active
0.9386 wohnmobilstellplatz
0.9380 hamburg hop on hop off
0.9376 strandhalle
0.9374 one day trip
0.9369 gemünd
0.9357 eltz
0.9349 europa park
0.9345 idar oberstein
0.9343 meran sehenswürdigkeiten
0.9342 den haag shopping
0.9342 wohnmobil parken
0.9332 öffentliche toiletten
0.9327 oberstein
0.9327 city sightseeing
0.9321 fussweg
0.9315 fisch klette
0.9314 algarve shopping
0.9309 bikepark hahnenklee
0.9304 parken
0.9301 karden
0.9300 touri
0.9298 markt palma
0.9293 idar
0.9287 corte ingles palma
0.9285 kennzeichen st
0.9285 rhein sehenswürdigkeiten
0.9284 reichsburg cochem
0.9282 rundweg
0.9280 bus hop
0.9279 ankunft flug
0.9278 mallorca uhrzeit
0.9277 womo stellplatz
0.9277 markt mallorca
0.9274 outdooractive
0.9274 p+r
0.9271 parken kostenlos
0.9271 guell
0.9270 giftbude norderney
0.9268 hrensko
0.9264 coves
0.9261 abflug flughafen
0.9255 womo
0.9250 einchecken
0.9249 rasender roland binz
0.9249 kennzeichen
0.9247 public transport
0.9245 reichsburg
0.9243 hop on hop off bus
0.9240 handgepäck
0.9233 schifffahrt cochem
0.9233 veste oberhaus
0.9231 dhv wetter
0.9229 giftbude
0.9226 saxon switzerland
0.9223 bad schandau
0.9221 park guell
0.9220 leiden holland
0.9219 waterbus
0.9218 elektromobil
0.9216 palma öffnungszeiten
0.9216 sightseeing
0.9215 airport bus
0.9215 zeit mallorca
0.9211 bus tour
0.9211 maastricht wetter
0.9210 schloss schwerin
0.9210 schandau
0.9210 burg eltz
0.9208 what to see
0.9205 brauhaus am kreuzberg
0.9203 bogenparcour
0.9202 ulm sehenswürdigkeiten
0.9201 sächsische schweiz wetter

User uploaded activity for **Duitsland** and Germany Web Search activity for **hop on hop off** (r=0.9603)

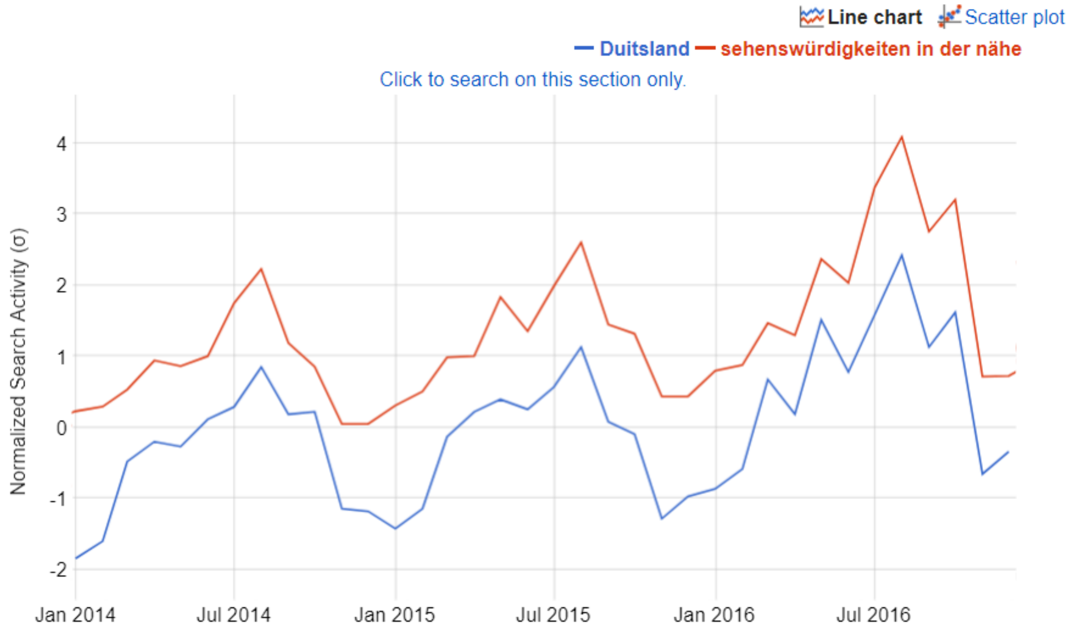
Line chart Scatter plot

— Duitsland — hop on hop off

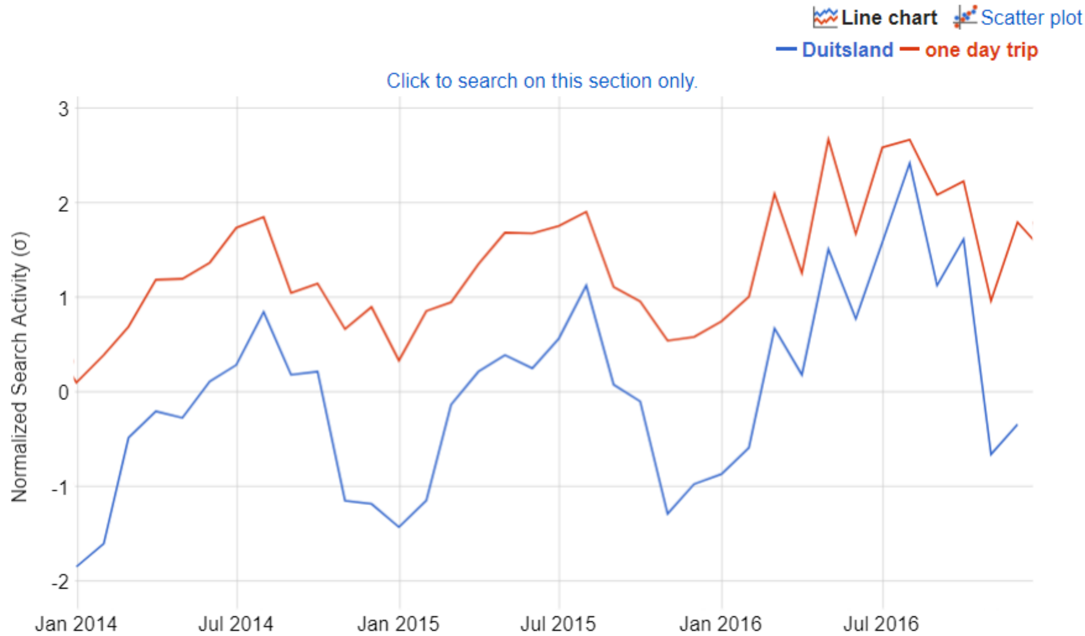
Click to search on this section only.



User uploaded activity for **Duitsland** and Germany Web Search activity for **sehenswürdigkeiten in der nähe** (r=0.9423)



User uploaded activity for **Duitsland** and Germany Web Search activity for **one day trip** (r=0.9374)



Correlations found for the United Kingdom:

Compare US states
Compare weekly time series
Compare monthly time series

Shift series 0 months
Country:
United Kingdom

Documentation
Comic Book
FAQ
Tutorial
Whitepaper
Correlate Algorithm

Correlate Labs
Search by Drawing

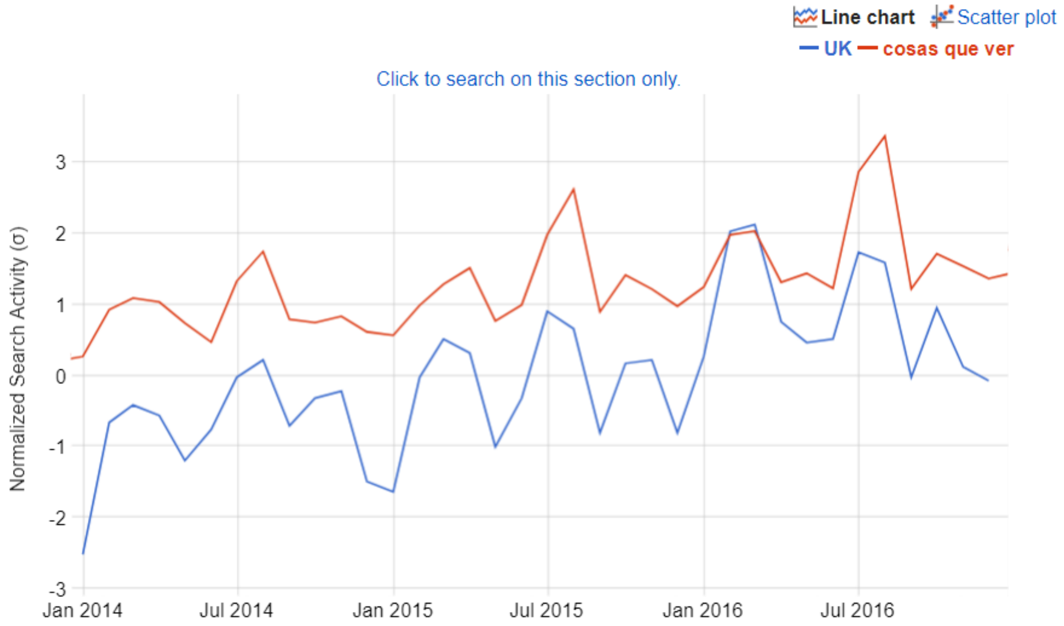
Correlated with **UK**

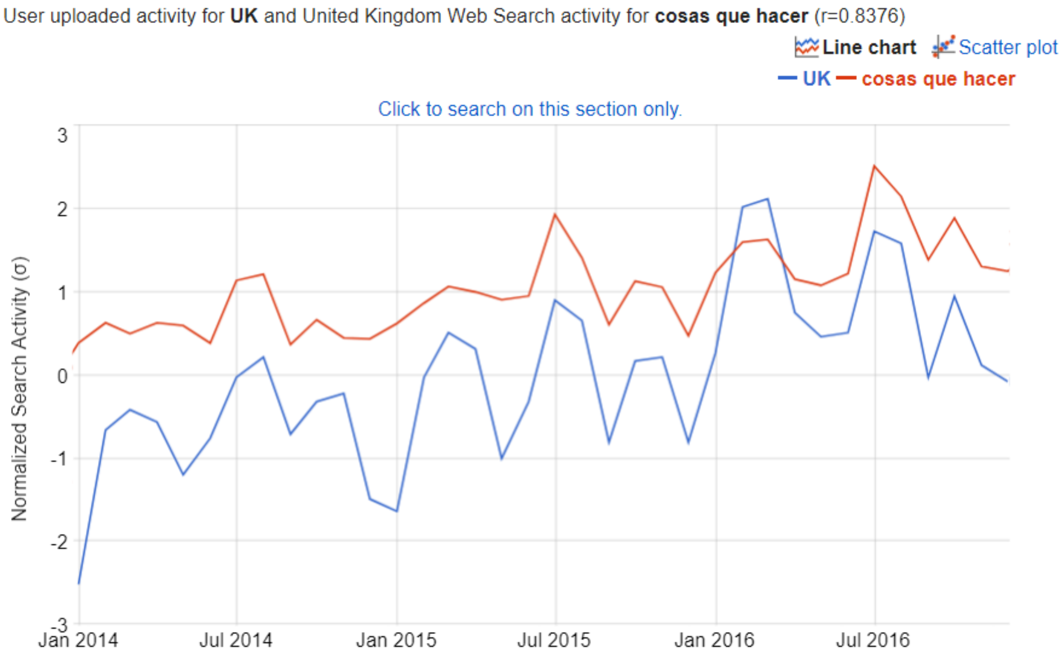
0.8574 tile cleaner
0.8546 places for lunch
0.8525 cosas que
0.8457 gluten free afternoon tea
0.8448 curtains white
0.8423 dry paint
0.8418 ifor williams horse trailer
0.8407 paint mix
0.8394 cosas que ver
0.8390 chalfont drive
0.8384 utsuro
0.8376 cosas que hacer
0.8350 car cleaners
0.8341 tea room
0.8340 mint macaroon
0.8335 white emulsion
0.8335 satinwood
0.8332 cocktail happy hour
0.8319 home of colour
0.8318 sklep budowlany
0.8289 bob with highlights
0.8280 cooper s r53
0.8278 afternoon tea for kids
0.8274 fixings
0.8262 bus tour dublin
0.8253 offside to offside
0.8247 bear crawls
0.8246 handyman services
0.8244 acrylic eggshell
0.8227 saturday
0.8225 wall primer

0.8225 adhesive tiles
0.8223 adhesive remover
0.8217 mini cooper s r53
0.8215 masonry screws
0.8215 polyfilla
0.8214 vintage tea room
0.8211 s3 coupe
0.8211 the running fox
0.8209 orhidee
0.8208 matt paint
0.8204 silk paint
0.8202 protection film
0.8202 contraflow
0.8199 plymouth to newton ab
0.8186 satinwood paint
0.8184 harris paint
0.8184 the pantry
0.8182 single carriageway
0.8179 soft sheen
0.8175 virgin train tickets
0.8174 tea room menu
0.8161 buty nike
0.8161 hair donation
0.8160 the kinmel
0.8156 edging trim
0.8155 healing time
0.8152 brilliant white
0.8152 sheet sander
0.8147 dulux diamond matt
0.8144 gluten free bakery
0.8139 r object
0.8133 dulux white emulsion

0.8133 matt emulsion paint
0.8132 grey yellow
0.8127 vinyl matt
0.8123 budget car hire belfast
0.8123 nc750
0.8122 cake tier stand
0.8118 nationwide select credit card
0.8115 places to eat cambridge
0.8114 contraflow bus
0.8113 gluten free restaurants
0.8112 tea places
0.8112 bars belfast
0.8111 afternoon
0.8110 abc electrification
0.8109 little princess trust
0.8102 what is a clearway
0.8101 plug sockets
0.8097 16000 after tax
0.8088 rooftop restaurant stratford
0.8083 to eat lunch
0.8081 turkish hammam
0.8081 pub lunches
0.8081 remove wall plugs
0.8079 how long is driving theory test
0.8078 golin
0.8076 blue bowl
0.8076 earth cafe

User uploaded activity for **UK** and United Kingdom Web Search activity for **cosas que ver** (r=0.8394)





User uploaded activity for **UK** and United Kingdom Web Search activity for **cosas que hacer** (r=0.8376)
Line chart
Scatter plot
UK
cosas que hacer
Click to search on this section only.
Normalized Search Activity (σ)
3
2
1
0
-1
-2
-3
Jan 2014
Jul 2014
Jan 2015
Jul 2015
Jan 2016
Jul 2016

Appendix VII

	Phase 1 Awareness or need	Phase 2 Enquire or planning	Phase 3 Google correlate
Germany	Amsterdam Travel Amsterdam	Hotel Amsterdam Flight Amsterdam Holiday Amsterdam Train Amsterdam Tourist info Amsterdam Visit Amsterdam City Trip Amsterdam	Hop on hop off Sehenswürdigkeiten in der nahe One day trip
United Kingdom	Amsterdam Travel Amsterdam	Hotel Amsterdam Flight Amsterdam Holiday Amsterdam Train Amsterdam Tourist info Amsterdam Visit Amsterdam City Trip Amsterdam	Cosas que ver Cosas que haser

Appendix VIII

Germany, 1-month delay:

Compare US states

Compare weekly time series

Compare monthly time series

Shift series 0 months

Country:

Germany ▼

Documentation

Comic Book

FAQ

Tutorial

Whitepaper

Correlate Algorithm

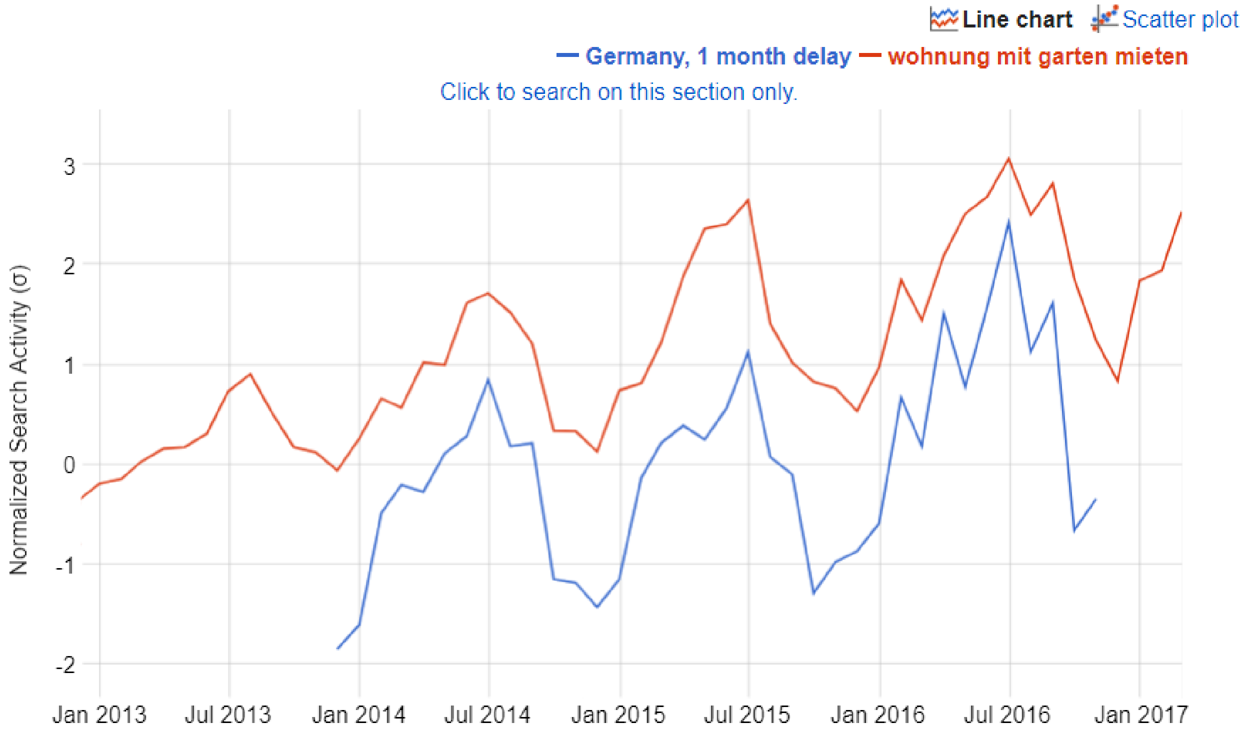
Correlate Labs

Search by Drawing

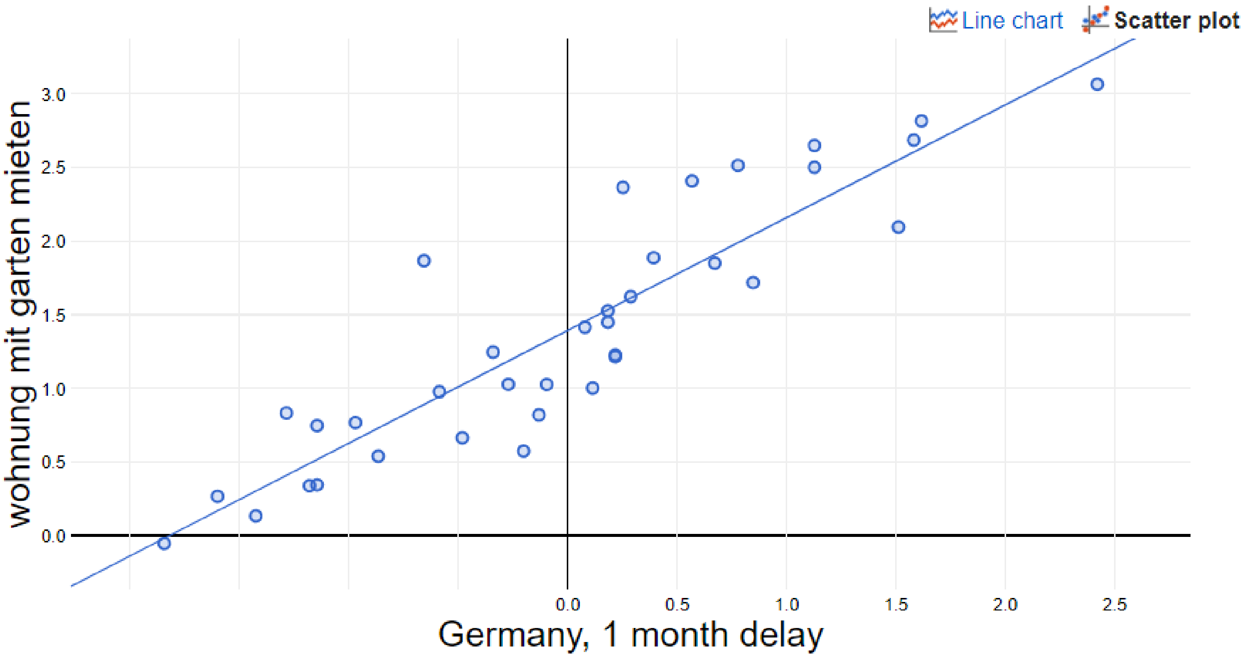
Correlated with **Germany, 1 month**

0.9093 schiebefenster
0.9083 haus mit garten
0.9083 tattoo knöchel
0.9082 lammsbräu
0.9079 luftballons mit helium
0.9043 wohnung mit garten mieten
0.9024 fensterheber motor
0.9021 bulls e-bike
0.9019 müde schlapp
0.9004 batrafen
0.8993 bali impfungen
0.8992 wetter feuerbach
0.8991 kleid streifen
0.8985 keder
0.8980 guttaplast
0.8977 impfung hund
0.8976 führerschein am
0.8968 peugeot tweet
0.8967 flüssigkunststoff
0.8947 eingangstreppe
0.8946 skuter
0.8937 bh ohne bügel
0.8931 36 volt
0.8930 barletta
0.8929 einblutungen
0.8923 bleierne müdigkeit
0.8916 us car werkstatt
0.8914 hortensien kaufen
0.8914 führerschein
0.8912 málna
0.8911 wundheilung
0.8910 kurzurlaub niederlande
0.8908 powergel
0.8908 malinowe
0.8907 klasse l
0.8906 cabrio gebraucht kaufen
0.8903 dachexpert
0.8902 überdachung
0.8899 sukienki na wesele
0.8897 mppt
0.8892 bluterguss
0.8886 weiße flecken
0.8886 zitronen buttermilch
0.8885 alubox
0.8882 mokasyny
0.8881 zadar croatia
0.8879 700x25c
0.8876 kawasaki ninja zx6r
0.8873 ultralight
0.8869 mieten mannheim
0.8868 corvette c3
0.8865 auto mieten mallorca
0.8865 ben jerry
0.8862 schleierkraut
0.8858 kawasaki 800
0.8857 cee stecker
0.8856 fahrrad rahmenschloss
0.8855 terasz
0.8855 jeep cabrio
0.8852 stevens strada 900

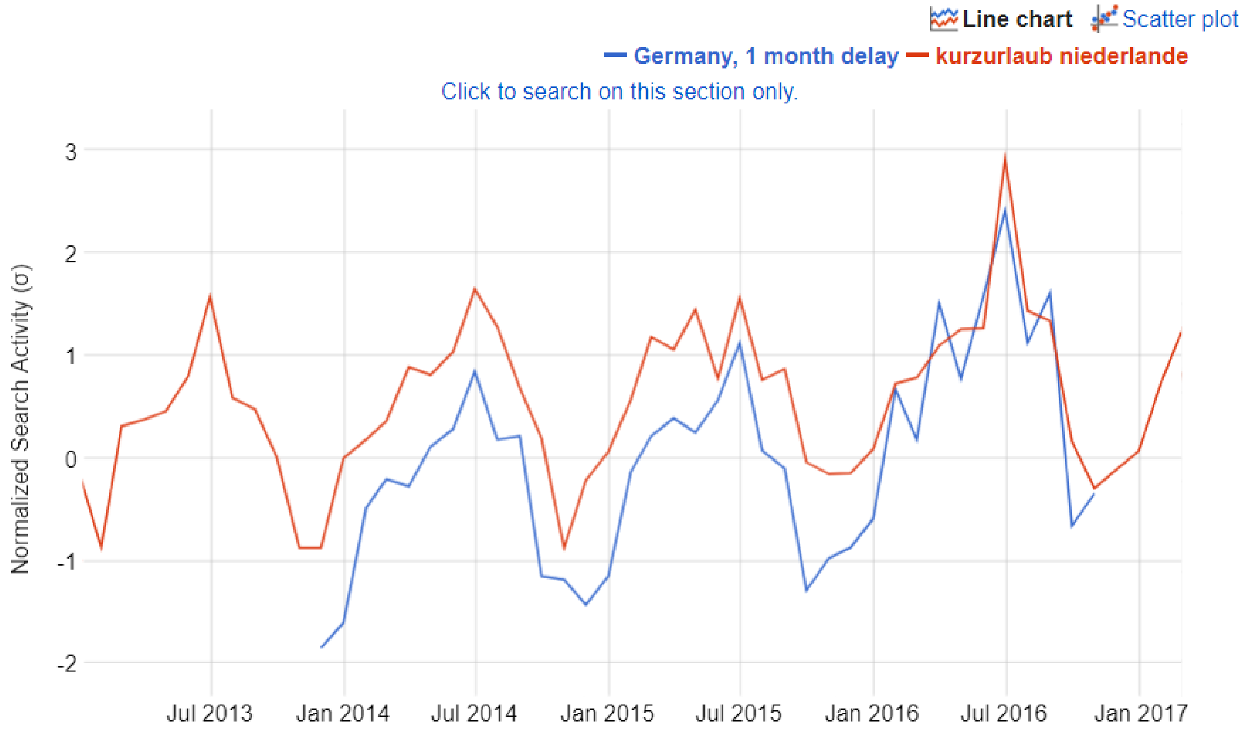
User uploaded activity for **Germany, 1 month delay** and Germany Web Search activity for **wohnung mit garten mieten** (r=0.9043)



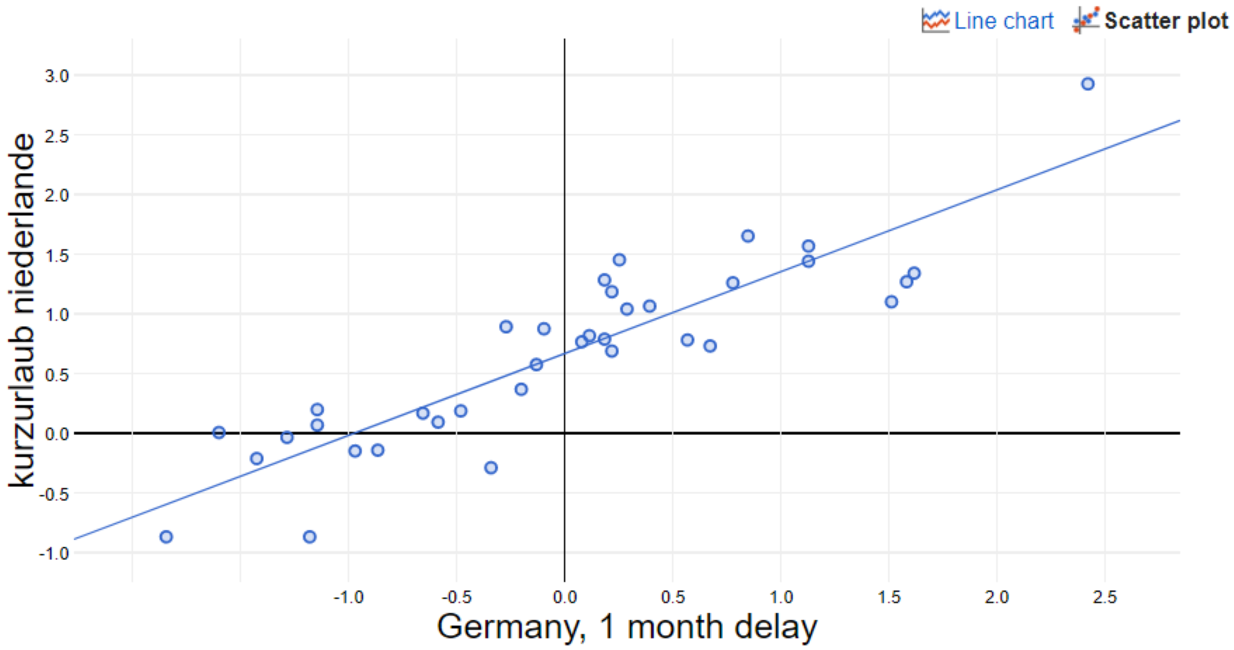
User uploaded activity for **Germany, 1 month delay** and Germany Web Search activity for **wohnung mit garten mieten** (r=0.9043)



User uploaded activity for **Germany, 1 month delay** and Germany Web Search activity for **kurzurlaub niederlande** (r=0.8910)



User uploaded activity for **Germany, 1 month delay** and Germany Web Search activity for **kurzurlaub niederlande** (r=0.8910)



Germany, 2 months delay:

Compare US states
Compare weekly time series
Compare monthly time series

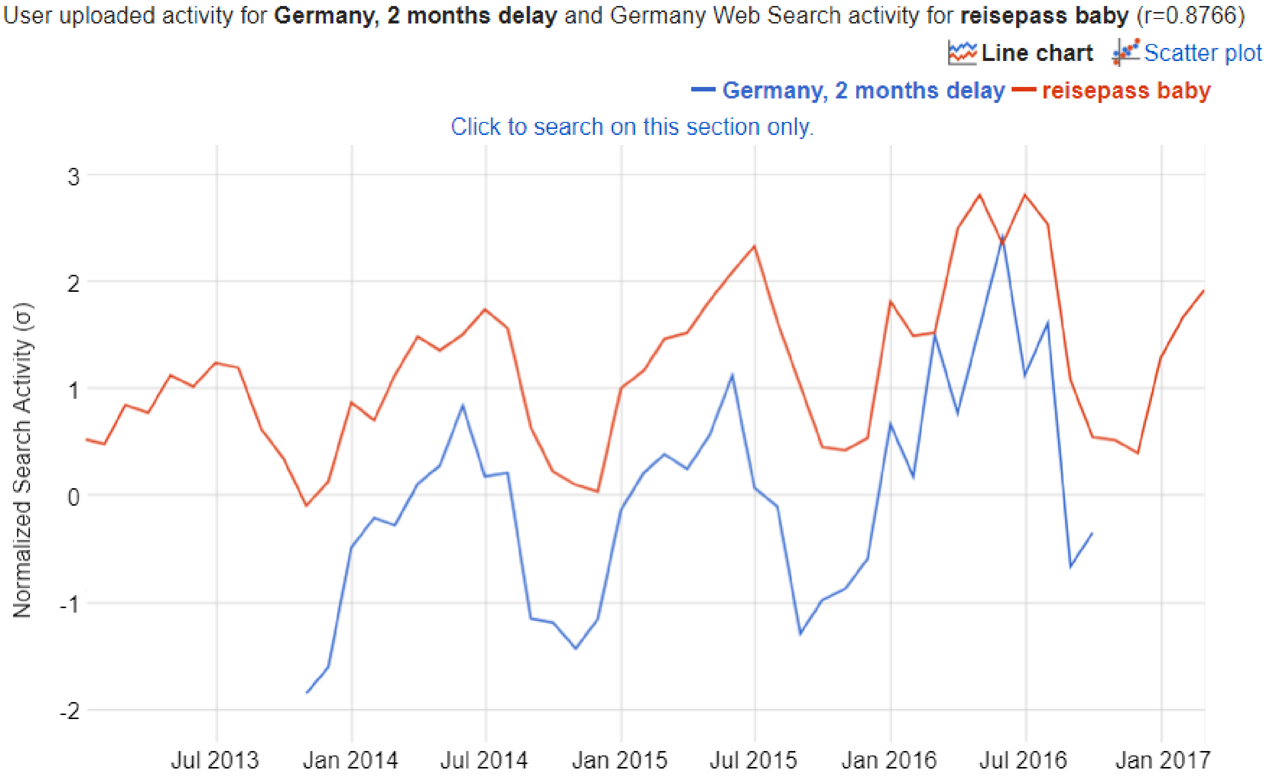
Shift series 0 months
Country:
Germany ▼

Documentation
Comic Book
FAQ
Tutorial
Whitepaper
Correlate Algorithm

Correlate Labs
Search by Drawing

Correlated with **Germany, 2 months delay**

0.9287 rosa kleid
0.9255 nike bikini
0.9197 personalausweis kind
0.9163 freudentränen
0.9144 kleid rosa
0.9077 kleid lang
0.9075 pizzabrot
0.9068 rathaus troisdorf
0.9043 boho hippie
0.9040 schuhe hochzeit
0.9039 für freudentränen
0.8984 gestreift
0.8973 shirt pink
0.8972 langes kleid
0.8951 kleider lang
0.8930 shaping badeanzug
0.8914 rezept joghurt
0.8900 kurzes top
0.8893 hellblaues kleid
0.8881 striped
0.8877 aufkleber hochzeit
0.8873 shirt gestreift
0.8871 t-shirt damen
0.8871 gartenhochzeit
0.8858 große brust
0.8854 brautstrauß rot
0.8839 jga
0.8833 reisebuggy
0.8831 festival job
0.8830 fladenbrot
0.8828 bikini asos
0.8827 partyboot düsseldorf
0.8821 baby shirt
0.8818 hose blau
0.8815 große oberweite
0.8814 flieder kleid
0.8804 badeanzug schwarz
0.8801 kleid hellblau
0.8787 ballermann urlaub
0.8766 reisepass baby
0.8760 cala ratjada urlaub
0.8758 zusammengeklappt
0.8757 blau weiß gestreift
0.8756 repayment
0.8753 zalando badeanzug
0.8752 medat
0.8749 mailand urlaub
0.8748 michael kors ballerina
0.8746 haarband hochzeit
0.8745 tan lotion
0.8741 bbq catering
0.8738 abendkleid blau
0.8734 baby passfoto
0.8733 top bikini
0.8731 taschentücher freudentränen
0.8729 hotel ballermann
0.8729 schwarzer bikini
0.8729 brautjungfer
0.8728 asos bikini
0.8726 hochzeit blumen


User uploaded activity for **Germany, 2 months delay** and Germany Web Search activity for **reisepass baby** (r=0.8766)
Line chart
Scatter plot
Germany, 2 months delay
reisepass baby
Click to search on this section only.
Normalized Search Activity (σ)
3
2
1
0
-1
-2
Jul 2013
Jan 2014
Jul 2014
Jan 2015
Jul 2015
Jan 2016
Jul 2016
Jan 2017


User uploaded activity for **Germany, 2 months delay** and Germany Web Search activity for **reisepass baby** (r=0.8766)
Line chart
Scatter plot
reisepass baby
3.0
2.5
2.0
1.5
1.0
0.5
0.0
0.0
0.5
1.0
1.5
2.0
2.5
Germany, 2 months delay

Germany, 3 months delay:

Compare US states

Compare weekly time series

Compare monthly time series

Shift series 0 months

Country:

Germany ▼

Documentation

Comic Book

FAQ

Tutorial

Whitepaper

Correlate Algorithm

Correlate Labs

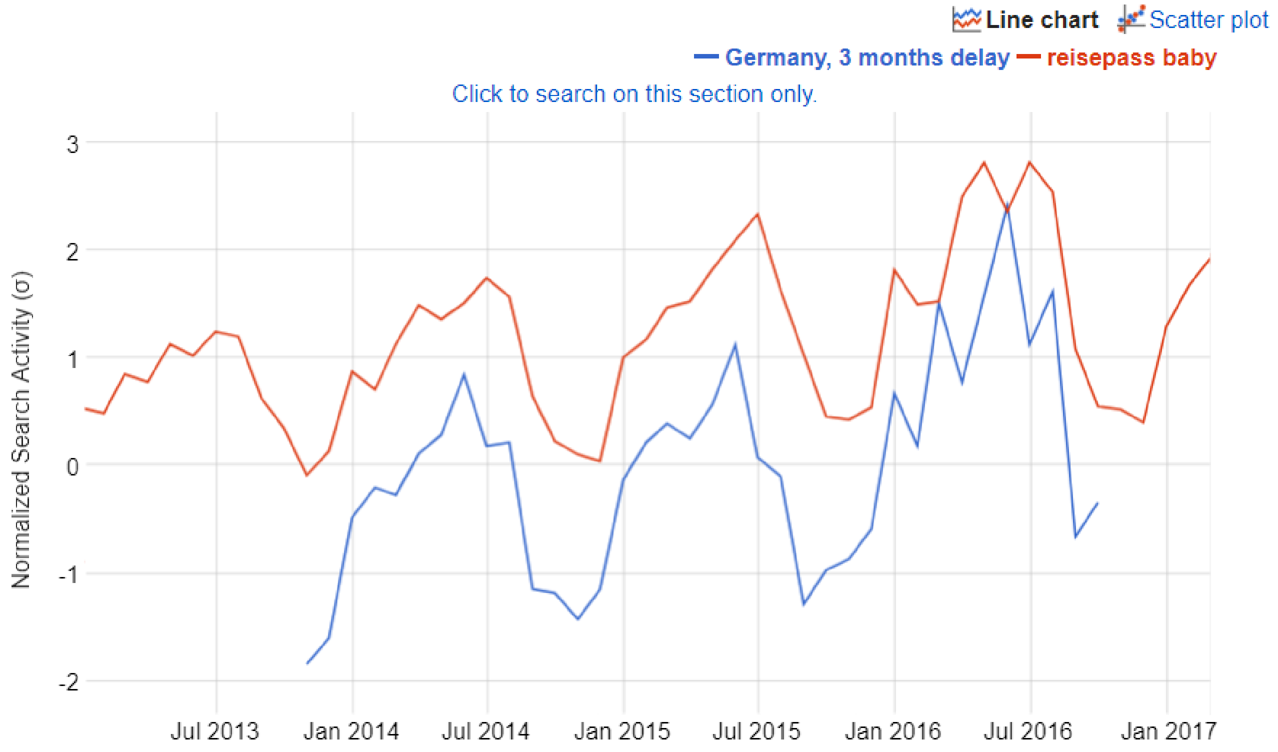
Search by Drawing

Correlated with **Germany, 3 months delay**

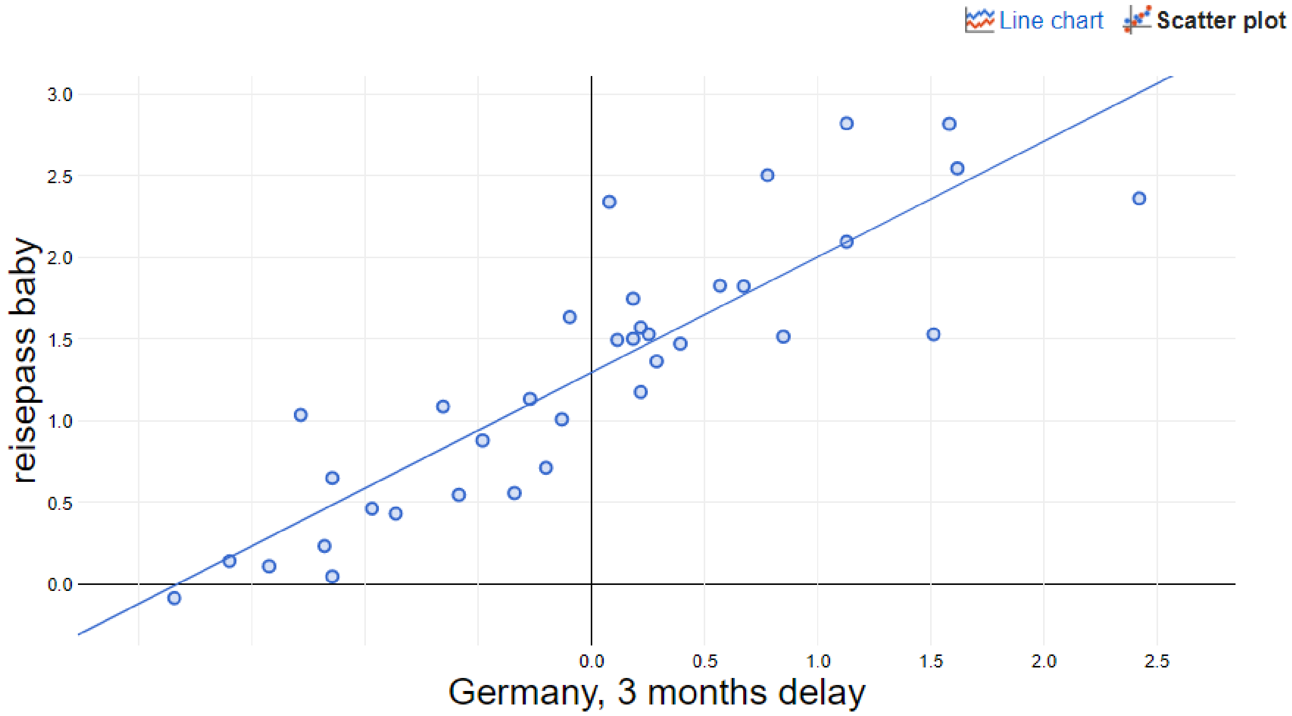
0.9287 rosa kleid
0.9255 nike bikini
0.9197 personalausweis kind
0.9163 freudentränen
0.9144 kleid rosa
0.9077 kleid lang
0.9075 pizzabrot
0.9068 rathaus troisdorf
0.9043 boho hippie
0.9040 schuhe hochzeit
0.9039 für freudentränen
0.8984 gestreift
0.8973 shirt pink
0.8972 langes kleid
0.8951 kleider lang
0.8930 shaping badeanzug
0.8914 rezept joghurt
0.8900 kurzes top
0.8893 hellblaues kleid
0.8881 striped
0.8877 aufkleber hochzeit
0.8873 shirt gestreift
0.8871 t-shirt damen
0.8871 gartenhochzeit
0.8858 große brust
0.8854 brautstrauß rot
0.8839 jga
0.8833 reisebuggy
0.8831 festival job
0.8830 fladenbrot
0.8828 bikini asos

0.8827 partyboot düsseldorf
0.8821 baby shirt
0.8818 hose blau
0.8815 große oberweite
0.8814 flieder kleid
0.8804 badeanzug schwarz
0.8801 kleid hellblau
0.8787 ballermann urlaub
0.8766 reisepass baby
0.8760 cala ratjada urlaub
0.8758 zusammengeklappt
0.8757 blau weiß gestreift
0.8756 repayment
0.8753 zalando badeanzug
0.8752 medat
0.8749 mailand urlaub
0.8748 michael kors ballerina
0.8746 haarband hochzeit
0.8745 tan lotion
0.8741 bbq catering
0.8738 abendkleid blau
0.8734 baby passfoto
0.8733 top bikini
0.8731 taschentücher freudentränen
0.8729 hotel ballermann
0.8729 schwarzer bikini
0.8729 brautjungfer
0.8728 asos bikini
0.8726 hochzeit blumen

User uploaded activity for **Germany, 3 months delay** and Germany Web Search activity for **reisepass baby** (r=0.8766)



User uploaded activity for **Germany, 3 months delay** and Germany Web Search activity for **reisepass baby** (r=0.8766)



Germany, 4 months delay:

Compare US states
Compare weekly time series
Compare monthly time series

Shift series 0 months
Country:
Germany

Documentation
Comic Book
FAQ
Tutorial
Whitepaper
Correlate Algorithm

Correlate Labs
Search by Drawing

Correlated with **Germany, 4 months delay**

0.8896 rotkehlchen
0.8891 glycerin
0.8863 neuroderm
0.8800 yiyecekler
0.8768 tin suresi
0.8731 zayıflatırmı
0.8703 kalk chemie
0.8691 empfindliche
0.8678 diskussion englisch
0.8669 feuchtigkeitscreme
0.8660 beste creme
0.8631 дата выхода серий
0.8628 aktie dividende
0.8627 utdelning
0.8611 kalori
0.8608 research paper
0.8608 tohumu
0.8606 baby neurodermitis
0.8595 nba scores
0.8591 hautrötungen
0.8590 duha
0.8589 draft prospects
0.8575 kartoffeln keimen
0.8569 yumurta
0.8561 cocktailkleid schwarz
0.8553 abituraufgaben mit lösungen
0.8548 qualitativer angebotsvergleich
0.8545 immer halsschmerzen
0.8543 budiair
0.8538 fiil
0.8529 tedavi
0.8519 längen umwandeln
0.8519 rückfettende creme
0.8510 neurodermitis baby
0.8510 sınava
0.8510 keimende kartoffeln
0.8505 haut creme
0.8498 osmanische reich
0.8484 thailand im juli
0.8478 hakkında
0.8465 innenpolitik bismarck
0.8463 das osmanische reich
0.8458 trainingsjacke damen
0.8434 medgurus
0.8430 ballkleid
0.8428 gekeimte kartoffeln
0.8424 nba score
0.8422 abschlussballkleider lang
0.8421 gesichtscreme
0.8416 umwelt englisch
0.8414 doc a
0.8411 was ist ein rechteck
0.8411 haşlanmış yumurta
0.8409 yulaf
0.8408 abi kleid
0.8406 gesicht brennt
0.8405 buchfink weibchen
0.8401 tedavisi
0.8398 komik doğum günü
0.8398 stress definition

Germany, 5 months delay:

Compare US states
Compare weekly time series
Compare monthly time series

Shift series 0 months
Country:
Germany

Documentation
Comic Book
FAQ
Tutorial
Whitepaper
Correlate Algorithm

Correlate Labs
Search by Drawing

Correlated with **Germany, 5 months delay**

0.9154 extreme halsschmerzen
0.9026 frankfurter kranz
0.9006 boyun
0.8981 schuppige kopfhaut
0.8937 crema pasticcera
0.8930 etkileri
0.8922 pasticcera
0.8879 mandelentzündung
0.8871 lachs auflauf
0.8844 süsleme
0.8833 augencreme
0.8826 wetter corralejo
0.8813 gegen trockene kopfhaut
0.8802 gesichtscreme
0.8796 reife haut
0.8792 adiyat suresi
0.8787 dr eckstein
0.8785 yan etkileri
0.8777 spinaci
0.8776 starke halsschmerzen
0.8752 welche foundation
0.8746 foundation für trockene haut
0.8739 şiir
0.8722 bu hafta
0.8722 mandarinen kuchen
0.8711 grüner schleim
0.8706 trockene kopfhaut
0.8697 magen darm kind
0.8695 steuererklärung rückwirkend
0.8690 extremer reizhusten
0.8677 mandarinen torte
0.8676 journee bruchsal
0.8674 brunch frankfurt
0.8672 ständig krank
0.8662 rückwirkend steuererklärung
0.8644 dr spiller
0.8639 zararları
0.8636 nach krampfader op
0.8636 spinat mit lachs
0.8629 creme für trockene haut
0.8623 crema de zahar
0.8622 çayı
0.8617 mobilbagger
0.8611 crema de zahar ars
0.8604 auge zuckt
0.8600 kitabı
0.8600 zahar ars
0.8599 husten hört nicht auf
0.8594 häufig erkältet
0.8590 adiyat
0.8589 deckendes make up
0.8581 blush
0.8570 destanı
0.8569 rachenentzündung
0.8569 chefkoch spinat
0.8564 immer wieder krank
0.8562 pasta sfoglia
0.8562 hustensaft prospan
0.8561 beyaz show konukları
0.8561 ıspanak
0.8557 schoko sahne
0.8554 scarlatina
0.8553 yumurta

Germany, 6 months delay:

Compare US states
Compare weekly time series
Compare monthly time series

Shift series 0 months
Country:
Germany ▼

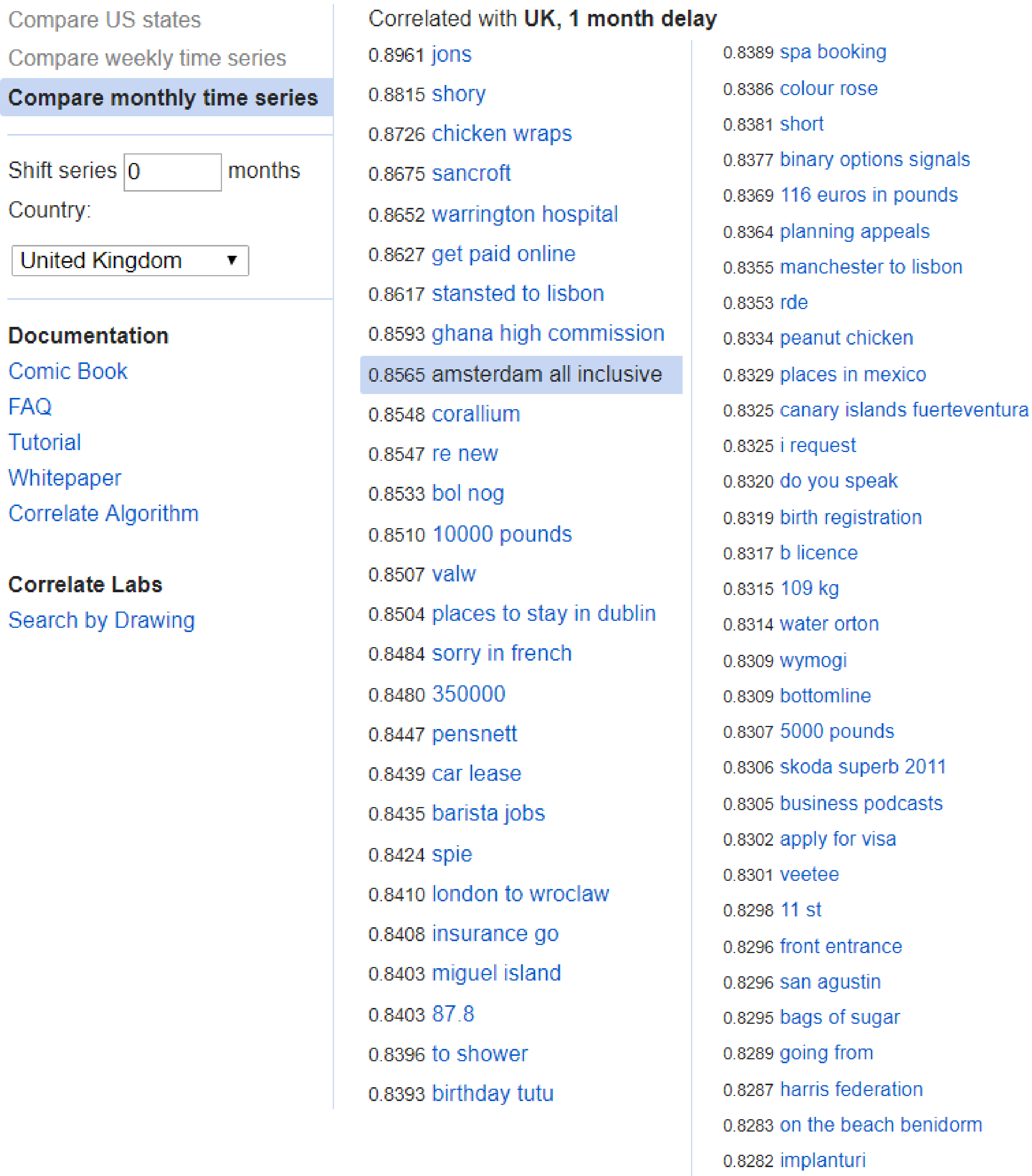
Documentation
Comic Book
FAQ
Tutorial
Whitepaper
Correlate Algorithm

Correlate Labs
Search by Drawing

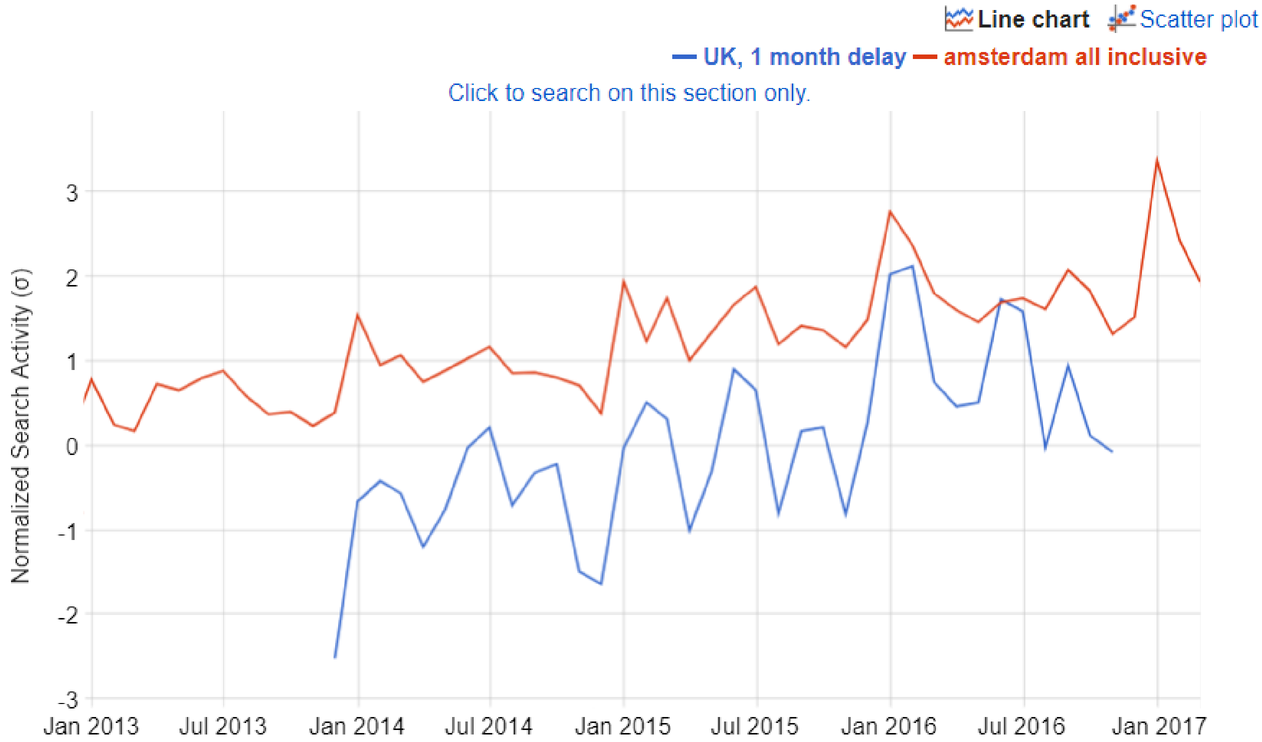
Correlated with **Germany, 6 months delay**

0.9134	rotbäckchen
0.9078	rahmgemüse
0.9073	wetter in madeira
0.9065	rezept vegetarisch
0.8994	plattensäge
0.8994	late entry
0.8991	alkapida
0.8980	ms 261 c-m
0.8980	carne
0.8968	göbel hotel
0.8960	geschnetzeltes
0.8960	cordon
0.8945	fräsmaschine
0.8932	einhell oberfräse
0.8931	imbir
0.8928	multi sanostol
0.8927	cordon bleu
0.8911	lötzinn
0.8907	absauganlage
0.8899	sanostol
0.8894	abstrakte bilder
0.8891	reflex si
0.8888	na zatoki
0.8886	rote linsensuppe
0.8874	netzwerkdose
0.8873	deckenlampe led
0.8872	pastinaken roh
0.8865	stihl ms 261 c-m
0.8864	topfband
0.8864	gratin
0.8862	backenfutter
0.8861	blut im
0.8859	pflanzliche arzneimittel
0.8854	wetter auf madeira
0.8848	almstadl plattling
0.8838	schnittschutzhandschuhe
0.8833	laif
0.8830	starke blähungen
0.8827	graupen
0.8823	kreuzschalter
0.8823	papierschneider
0.8815	bessey
0.8812	schraubstock
0.8810	φασολαδα
0.8808	prossimo turno
0.8805	hühnerfrikassee
0.8796	xenon birne
0.8794	global village dubai
0.8793	mulgatol
0.8790	linsen einweichen
0.8785	сауну
0.8784	rohes sauerkraut
0.8783	einstellungstest bundespolizei
0.8783	hoher blutdruck
0.8782	eukal
0.8782	weitsichtig
0.8780	käsespätzle
0.8777	ist ingwer gesund
0.8777	con carne
0.8777	ingwer gut für
0.8770	schmelzkäse
0.8764	schmerzen brustkorb
0.8757	salzgrotte wolfsburg

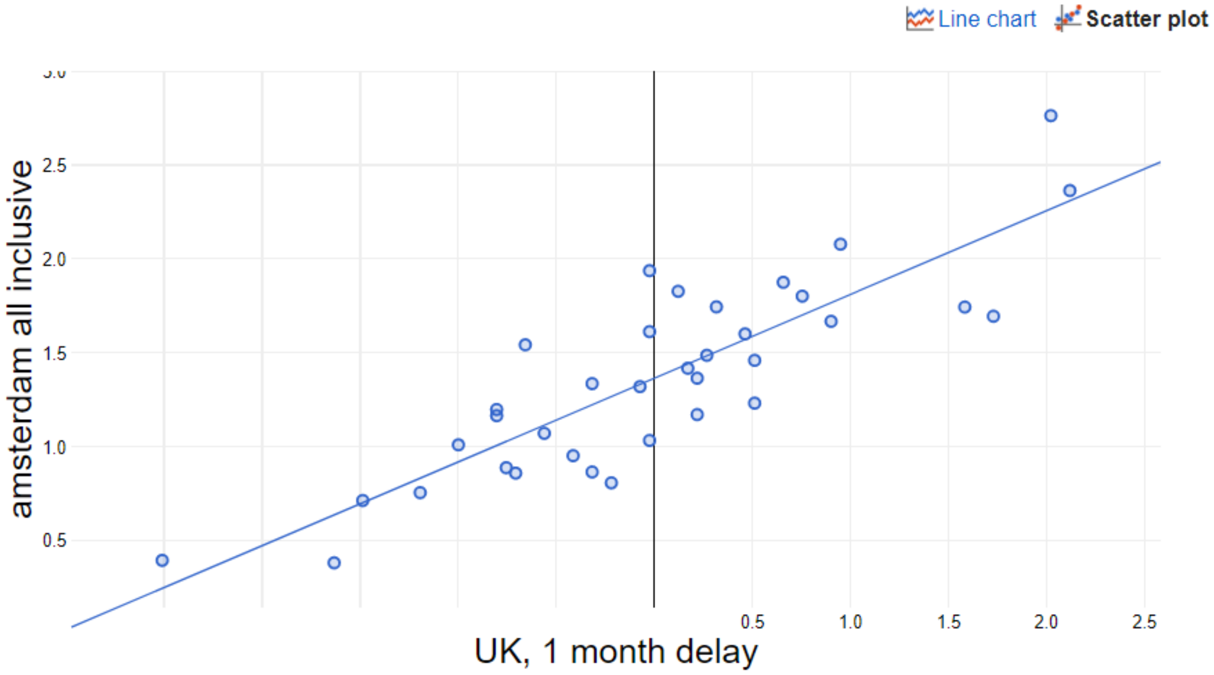
United Kingdom, 1-month delay:



User uploaded activity for **UK, 1 month delay** and United Kingdom Web Search activity for **amsterdam all inclusive** (r=0.8565)



User uploaded activity for **UK, 1 month delay** and United Kingdom Web Search activity for **amsterdam all inclusive** (r=0.8565)



United Kingdom, 2 months delay:

Compare US states

Compare weekly time series

Compare monthly time series

Shift series 0 months

Country:

United Kingdom

Documentation

Comic Book

FAQ

Tutorial

Whitepaper

Correlate Algorithm

Correlate Labs

Search by Drawing

Correlated with **UK, 2 months delay**

0.8461	using fitbit
0.8441	garmin heart rate
0.8406	what is the fitbit
0.8358	push up bra
0.8341	apps for movies
0.8326	best heart
0.8305	does fitbit work
0.8302	fitbit cheapest
0.8289	how to use fitbit
0.8264	king prawns
0.8255	i sent
0.8225	nike dri fit
0.8224	flysky fs-t6
0.8221	up bra
0.8217	cabbage salad
0.8217	best heart rate
0.8217	buy fitbit
0.8197	monitor heart rate
0.8194	raw salmon
0.8190	fitbit windows
0.8188	fitbit heart
0.8181	set up fitbit
0.8179	fitbit heart rate
0.8174	best fitness
0.8170	nike dri
0.8170	what should my heartrate be
0.8167	runni
0.8167	exercise tracker
0.8165	fitbit zip
0.8161	fit it
0.8153	how does fitbit work
0.8150	fitbit tracker
0.8146	fitbit how to
0.8146	how many glasses
0.8146	how to set up fitbit
0.8142	charge fitbit
0.8140	hr review
0.8138	dri fit
0.8128	8.7 stone in kg
0.8126	fitbit iphone
0.8119	fitbit monitor
0.8116	lace bra
0.8116	games steam
0.8109	hr fitbit
0.8106	www.fitbit.com/setup
0.8100	chive dip
0.8100	cheapest fitbit
0.8095	fitbit.com
0.8089	trophies ps4
0.8084	up fit
0.8083	fitbit wrist
0.8082	how does a fitbit work
0.8081	fitbit compatible
0.8079	fit charge hr
0.8071	tumblr backgrounds
0.8069	fitbit/setup
0.8063	the fitbit
0.8055	charge hr fitbit
0.8044	how long to charge fitbit
0.8040	fitbit
0.8040	cooked salmon
0.8035	where can i buy a fitbit
0.8033	rate watch

United Kingdom, 3 months delay:

Compare US states
Compare weekly time series
Compare monthly time series

Shift series 0 months
Country:
United Kingdom ▼

Documentation
Comic Book
FAQ
Tutorial
Whitepaper
Correlate Algorithm

Correlate Labs
Search by Drawing

Correlated with **UK, 3 months delay**

0.8812 ipl philips
0.8518 masażer
0.8411 m&s mens jeans
0.8400 womens size 5
0.8400 plunge bra
0.8372 uk size 8
0.8366 what size
0.8359 mac mineralise
0.8338 uk size 10
0.8334 wwe euroshop
0.8322 cotton stretch
0.8290 yate sort it centre
0.8281 nars illuminator
0.8258 no7 cream
0.8256 rice bowl beechwood
0.8251 types of bras
0.8251 earring gold
0.8248 bronzer brush
0.8241 a 12
0.8232 floral dressing gown
0.8206 nike discount
0.8204 popular trainers
0.8203 piosenkarz
0.8190 food station salford
0.8182 amazon running
0.8177 lace top
0.8177 my psu
0.8176 αγιος
0.8173 now tv box tesco
0.8160 40% of 100
0.8148 skinny
0.8144 mens air max 95
0.8135 shops in debenhams
0.8133 nike discount code
0.8131 heely
0.8127 roller shoes
0.8121 london silver
0.8116 benefit hoola bronzer
0.8113 deep plunge bra
0.8109 radio bluetooth
0.8107 pirelli trofeo
0.8103 baby gift set
0.8102 in order
0.8086 uk 8
0.8086 applique dress
0.8083 the new movie
0.8083 benefit hoola
0.8081 moisturizing gel
0.8080 school romance
0.8074 aldi old colwyn
0.8074 water for constipation
0.8068 star kebab
0.8066 gold rim
0.8053 now tv hd
0.8051 b and m uk
0.8050 trainers
0.8047 redpost equestrian
0.8041 tactical elite
0.8040 30 usd
0.8037 victoria secret promo
0.8029 spade
0.8029 track top
0.8022 adidas zip up

United Kingdom, 4 months delay:

Compare US states
Compare weekly time series
Compare monthly time series

Shift series 0 months
Country:
United Kingdom ▼

Documentation
Comic Book
FAQ
Tutorial
Whitepaper
Correlate Algorithm

Correlate Labs
Search by Drawing

Correlated with **UK, 4 months delay**

0.8686 unhappy synonym
0.8586 part synonym
0.8569 table matlab
0.8545 further synonym
0.8495 ebsco login
0.8410 taxi industry
0.8405 understanding synonym
0.8395 impact synonym
0.8367 cartof
0.8346 role synonym
0.8292 rick grimes actor
0.8291 colouring images
0.8285 din 4
0.8281 drive synonym
0.8267 clearly synonym
0.8261 time.
0.8261 imshow
0.8260 focus synonym
0.8248 activities synonym
0.8248 demonstrate synonym
0.8240 access synonym
0.8232 bcs login
0.8228 mancaruri
0.8223 history synonym
0.8221 establish synonym
0.8220 study synonym
0.8219 for loop r
0.8204 srand
0.8198 lot synonym
0.8196 knowledge synonym
0.8196 time in new
0.8195 delayed concussion
0.8193 introduce synonym
0.8190 but synonym
0.8189 sonequa martin
0.8187 feel synonym
0.8175 mancare de
0.8175 comparison synonym
0.8173 process synonym
0.8171 important synonym
0.8166 list python
0.8163 cdaq
0.8162 random number between
0.8157 developed synonym
0.8156 as synonym
0.8154 huge synonym
0.8144 addition synonym
0.8142 gain synonym
0.8141 peste
0.8140 comprehensive synonym
0.8140 engage synonym
0.8140 walking dead online subtitrat
0.8135 prekladac
0.8134 三 五
0.8132 tv online romanesti
0.8131 promote synonym
0.8130 quiz puns
0.8121 understand synonym
0.8120 while synonym
0.8116 success synonym

United Kingdom, 5 months delay:

Compare US states
Compare weekly time series
Compare monthly time series

Shift series 0 months
Country:
United Kingdom ▼

Documentation
Comic Book
FAQ
Tutorial
Whitepaper
Correlate Algorithm

Correlate Labs
Search by Drawing

Correlated with **UK, 5 months delay**

0.8792	student minds
0.8777	office 365 education
0.8761	breaches
0.8728	henrys
0.8717	pzz
0.8708	dress up adults
0.8702	uob
0.8700	to draw in
0.8699	bible definition
0.8696	drawings
0.8694	soul mean
0.8690	warwick uni email
0.8685	slavery definition
0.8677	relaxation music
0.8672	the noun
0.8662	commercial definition
0.8656	mail rugby
0.8654	daily mail rugby
0.8651	the witches book
0.8649	school hub
0.8645	class dojo points
0.8639	being a parent
0.8638	intr
0.8636	teech
0.8629	to e
0.8628	future learn
0.8624	exeter email
0.8622	identity theft definition
0.8615	synonyms for strong
0.8604	suddenly synonym
0.8601	exclamation mark
0.8597	oxford road
0.8594	top maths
0.8592	behavioural science
0.8591	inside of a
0.8590	uon
0.8584	business quote
0.8580	helping
0.8575	black dance leotard
0.8567	old fire
0.8566	technology meaning
0.8565	line drawing
0.8554	face cartoon
0.8551	drawings of
0.8549	have the
0.8546	telegraph rugby
0.8542	interactive learning
0.8542	a line
0.8535	lycée international
0.8534	manager definition
0.8530	education login
0.8528	to draw
0.8527	ls6 cafe
0.8525	energy transition
0.8521	migrated
0.8516	guardian rugby
0.8512	after concussion
0.8509	teesside uni
0.8507	what is a group of
0.8505	you draw
0.8499	youtube relaxing music
0.8498	old fire station
0.8497	what does a

United Kingdom, 6 months delay:

Compare US states
Compare weekly time series
Compare monthly time series

Shift series 0 months
Country:
United Kingdom ▼

Documentation
Comic Book
FAQ
Tutorial
Whitepaper
Correlate Algorithm

Correlate Labs
Search by Drawing

Correlated with **UK, 6 months delay**

0.8839 saag aloo
0.8781 gluten free bread
0.8779 new york things
0.8732 deep fitted sheet
0.8715 ikea office
0.8693 ikea mirrors
0.8684 best mattress topper
0.8670 laundry basket
0.8607 lodge break
0.8602 theatre new york
0.8589 rustic kitchen
0.8581 tall storage unit
0.8577 drawer storage
0.8575 desk mat
0.8573 new york things to do
0.8568 dundee street
0.8564 new york basketball
0.8564 ikea bookcase
0.8563 queen size bed
0.8556 ekby
0.8555 next bedding
0.8555 pearl hotel new york
0.8539 top restaurants
0.8536 yomvi
0.8528 york basketball
0.8528 small double duvet
0.8522 ikea bedding
0.8511 double bunk
0.8511 ikea ekby
0.8510 queen size
0.8509 wooden breakfast bar
0.8498 west lothian school holidays
0.8498 ikea desk
0.8494 drawer storage unit
0.8492 double bunk beds
0.8491 kids bed
0.8488 long sleeve lace wedding dress
0.8478 door hooks
0.8475 broadway tickets
0.8470 connect printer
0.8470 white corner
0.8468 small double
0.8466 taxi from jfk
0.8465 super king beds
0.8464 muji stationery
0.8459 desk shelf
0.8455 kitchen shelves
0.8454 ikea lack shelf
0.8448 board wall
0.8443 oak
0.8443 best pillows
0.8433 tall storage
0.8430 led grow
0.8428 places to eat in new york
0.8427 bed frame
0.8427 ikea cabin bed
0.8427 double bed
0.8425 ikea office chair
0.8421 double mattress memory foam
0.8420 walnut
0.8419 double sofa bed
0.8414 file box
0.8405 divan base