



## ORIGINAL ARTICLE


**Citation:** Nagy, B., Gabor, M. R., Bacoş, I. B., Kabil, M., Zhu, K., & Dávid, L. D. (2023). Google and Apple mobility data as predictors for European tourism during the COVID-19 pandemic: A neural network approach. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 18(2), 419–459. doi: 10.24136/eq.2023.013

Contact to corresponding author: Kai Zhu, hizhukai@163.com

Article history: Received: 9.02.2023; Accepted: 1.06.2023; Published online: 30.06.2023


### **Benedek Nagy**

*Sapientia Hungarian University of Transylvania, Romania*

 [orcid.org/0000-0001-7027-0194](https://orcid.org/0000-0001-7027-0194)


### **Manuela Rozalia Gabor**

*George Emil Palade University of Medicine, Pharmacy, Science, and Technology of Targu Mures, Romania*

 [orcid.org/0000-0003-1025-869X](https://orcid.org/0000-0003-1025-869X)


### **Ioan Bogdan Bacoş**

*George Emil Palade University of Medicine, Pharmacy, Science, and Technology of Targu Mures, Romania*

 [orcid.org/0000-0002-4055-6287](https://orcid.org/0000-0002-4055-6287)


### **Moaz Kabil**

*Hungarian University of Agriculture and Life Sciences, Hungary  
Cairo University, Egypt*

 [orcid.org/0000-0002-0658-3871](https://orcid.org/0000-0002-0658-3871)

### **Kai Zhu**


*Hubei University, China*

 [orcid.org/0000-0003-4041-2918](https://orcid.org/0000-0003-4041-2918)

### **Lóránt Dénes Dávid**

*John von Neumann University, Hungary*

*Hungarian University of Agriculture and Life Sciences, Hungary*

 [orcid.org/0000-0001-7880-9860](https://orcid.org/0000-0001-7880-9860)

Copyright © Instytut Badań Gospodarczych / Institute of Economic Research (Poland)

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Google and Apple mobility data as predictors for European tourism during the COVID-19 pandemic: A neural network approach

**JEL Classification:** *C45; R12; F14*

**Keywords:** *mobility; tourism; Google mobility data; Apple mobility data; Europe; COVID-19 pandemic*

### Abstract

**Research background:** The COVID-19 pandemic has caused unprecedented disruptions to the global tourism industry, resulting in significant impacts on both human and economic activities. Travel restrictions, border closures, and quarantine measures have led to a sharp decline in tourism demand, causing businesses to shut down, jobs to be lost, and economies to suffer.

**Purpose of the article:** This study aims to examine the correlation and causal relationship between real-time mobility data and statistical data on tourism, specifically tourism overnights, across eleven European countries during the first 14 months of the pandemic. We analyzed the short longitudinal connections between two dimensions of tourism and related activities.

**Methods:** Our method is to use Google and Apple's observational data to link with tourism statistical data, enabling the development of early predictive models and econometric models for tourism overnights (or other tourism indices). This approach leverages the more timely and more reliable mobility data from Google and Apple, which is published with less delay than tourism statistical data.

**Findings & value added:** Our findings indicate statistically significant correlations between specific mobility dimensions, such as recreation and retail, parks, and tourism statistical data, but poor or insignificant relations with workplace and transit dimensions. We have identified that leisure and recreation have a much stronger influence on tourism than the domestic and routine-named dimensions. Additionally, our neural network analysis revealed that Google Mobility Parks and Google Mobility Retail & Recreation are the best predictors for tourism, while Apple Driving and Apple Walking also show significant correlations with tourism data. The main added value of our research is that it combines observational data with statistical data, demonstrates that Google and Apple location data can be used to model tourism phenomena, and identifies specific methods to determine the extent, direction, and intensity of the relationship between mobility and tourism flows.

### Introduction

The COVID-19 pandemic has had devastating effects on tourism. In April and May 2020, in most European countries, tourism arrivals fell to a 0-close value (Ghorbani *et al.*, 2023; Mousazadeh *et al.*, 2023). Globally, in 2020, tourism arrivals decreased by 73% compared to 2019 (UNWTO Dash-

board). In Romania, the tourist arrival decrease was even more significant: in 2020, the downfall was 83%, or it was at the 17% level of the previous year (2019), according to the UNWTO data (World Tourism Organization, 2022). During the COVID pandemic and ever since, we have had a large-scale open database about the movement and real motion of people in different types of urban and nonurban areas around the world. Since February 2020, Google has continuously published COVID-19 Community Mobility Reports, which contain data about the presence and mobility of persons in six different space categories: parks, transit stations, workplaces, retail & recreation, grocery & pharmacy and residential. At the same time, Apple has also launched a similar database, referring to the movement of people measured in only three categories (A3): transit, driving and walking. Both big data providers – Google and Apple – gather data on the individual level from smartphone devices that have set the location determination function "ON". The data are aggregated and published on the level of cities and states by both providers and on the county/regional level by Google.

Our aim is to analyze the relation between the GM6 & A3 dimensions of movement, measured by Google and Apple, and tourist overnight stays within a period of 12–14 months (according to the mobility dimensions compared) by using both statistical and econometric methods. Our intuitive research hypothesis is a powerful correlation or causal relationship between mobility indices and overnights, but some particularities regarding some dimensions of people mobility and the difference in correlations among some European countries are useful for future research by applying different types of quantitative and/or qualitative methods. For the present research, we compared the state-level indicators in European countries with tourism arrivals, available on the Eurostat website (Eurostat, 2021). For one single case, we compared the daily data of the Apple and Google dimensions, and we found statistically significant correlations in all analyzed countries.

It is important to mention that the present results are complementary to complex research made by the authors (Nagy *et al.*, 2022) for one European country (Romania), but using only Google Mobility dimensions and big data: daily data for all 41 Romanian counties and Bucharest. The statistical method to find the best predictor of overnight stays for Romanian tourism was the structural equation model (Nagy *et al.*, 2022). With the present research, the authors validate whether the relations between mobility data

and tourism indicators during the COVID-19 pandemic for Romania indicate a particular situation for an East European country or not. Therefore, we introduce in the study all European countries with complete data for the abovementioned indicators.

The present paper is structured as follows: first, we present the specific data used in the research and the statistical methods used, our own methodological approach for the research, the research results and discussions, conclusions, and limitations of the research. Our theoretical and practical contribution to the international scientific literature is from a methodological viewpoint and demonstrates that some specific mobility data are a good predictor, statistically significant, of tourism overnight stays for the European country. From a methodological standpoint, we make a theoretical and practical contribution to the international scientific literature by demonstrating that some specific mobility data are a good predictor, statistically significant, of tourism overnight stays for the European countries studied. The main limitations of the research are linked to incomplete data for all European countries and the lack of similar research to construct the research hypothesis. The novelty of the results is that the present results can be used as core hypotheses for future research on other continents and/or at a specific country level.

## **Literature review**

The new availability of these data about the movement and fluctuation of people in different types of locations is very interesting; nonetheless, several articles have been published in the last two years using these data sets. Most of these articles focus on COVID-19 infections and effects, the spread of new cases and the relationship between the mobility of people (modeled with Google mobility data) and the diffusion of the pandemic in different countries and other areas. First, Tamagusko and Ferreira (2020) sought a connection between the  $R_t$  value of the pandemic in Portugal — the scale of contagiousness — and the mobility of people, but they did not show a clear and direct relationship between the two lines of data. On the other hand, Irini *et al.* (2021) pointed out that there is a close connection between new COVID cases and some Google mobility dimensions: daily infections can be linked with lower rates of the residential index (staying at home). There are also other articles dealing with the link between the spread of the

virus or other effects of policies and the movement of people, measured with data from Google Community Reports: Saha *et al.* (2021) about the importance of policies for flattening the curve of the pandemic; Cot *et al.* (2021) about the imprint of social distancing in Europe and the US; Wang *et al.* (2021) about the changes in socioeconomic activity since 2020; Sulyok and Walker (2021) about the mobility and mortality links across Scandinavia; Hakim *et al.* (2021) about the number of cases, the effects of restriction policies and the mobility of people in some Asian countries; Jacobsen and Jacobsen (2020) about the effects of stay-at-home orders and voluntary behavior changes and many others. In addition to the effects of mobility on pandemics, other subjects were explored in the last 2–3 years with the help of Google and, in some cases, Apple data. Munawar *et al.* (2021) analyze how the transport system is impacted because of the policies adopted by the Australian government for the containment of COVID-19. Yang *et al.* (2021a) compare three of the Google mobility dimensions (Retail and Recreation, Parks and Transit Stations) of 9 world cities using tourism and travel data gathered with Mastercard. They conclude that cities administered different policies during the severe periods of the pandemic, resulting in a different change point of mobility, and point out that tourism could predict changes in mobility (which is a reverse logic compared to the one used in the present study).

Another approach relates to tourism during and after the COVID period (post-COVID), which is summarized effectively by Yang *et al.* (2021b). They have gathered 249 articles in five key subject areas on tourism, such as psychological effects and behavior, risk perceptions, well-being and mental health, motivation and behavioral intention and responses, strategies, and resilience: organization and government, some of which also address economic effects. Geng *et al.* (2021) measure park attendance with the help of Google Community data and the impact of this on the social and mental wellbeing of visitors — this would be the only direct use of Google data for tourism-related research purposes, from what we have found thus far. An interesting new study that integrates the statistical data from Eurostat on retail with the retail data from Google Community Reports (Szász *et al.*, 2022) observes the short-term drivers and long-term implications of the pandemic in the online retail sector. They construct a good model that could explain the short-term evolution of the online retail sector during the pandemic.

In the most recent studies, Google mobility dimensions are directly linked to the COVID-19 pandemic (dealing with the spread of the virus) and the social effects of this actual situation (Murray, 2021; Atalay & Solmazer, 2021; Ibarra-Espinosa *et al.*, 2021). The economy and, within that, the tourism relations with the movement of people within a region, country or continent in various types of locations has not yet been researched, and only a very few papers have yet to be researched, with only a few papers published. One of these states that urban park visitations have increased since the beginning of the pandemic (Geng *et al.*, 2020), and the other one searches for the possibility of forecasting tourist arrivals with SARIMA models, but not Google Mobility, which we plan to use; instead, they apply Google Trends data (which reflect the number of searches through Google search engines worldwide) (Bangwayo-Skeete & Skeete, 2015).

Previous research on the topic of enhancing the accuracy of COVID-19 prediction by integrating epidemiological and mobility data (García-Cremades *et al.*, 2021) aimed to assess various models that can be used to make early predictions about the progression of the COVID-19 outbreak, with the ultimate goal of developing a decision support system to aid policy-makers. For example, spatiotemporal disease models and graph neural networks were integrated to enhance the forecasting accuracy of weekly COVID-19 cases in Germany (Fritz *et al.*, 2022). This research demonstrated the essentiality of incorporating mobility data while also highlighting the adaptability and comprehensibility of the proposed methodology. Similarly, mobile phone data have been used to guide public health interventions throughout the various stages of the COVID-19 pandemic (Oliver *et al.*, 2020).

Foundational research on human mobility has revealed that combined and identified mobile phone data can aid in modeling the geographic diffusion of epidemics (Finger *et al.*, 2016; Tizzoni *et al.*, 2014; Wesolowski *et al.*, 2012; Bengtsson *et al.*, 2015; Wesolowski *et al.*, 2015).

Consequently, partnerships have emerged between researchers, governments, and private enterprises, particularly mobile network operators and location intelligence firms, to gauge the efficacy of mitigation measures in several nations, such as Austria, Belgium, Chile, China, Germany, France, Italy, Spain, the United Kingdom, and the United States (Oliver *et al.*, 2020; Shortall *et al.*, 2020; Kraemer *et al.*, 2020; Lai *et al.*, 2020; Lyons, 2020; Brodeur *et al.*, 2021).

## Research methods

Due to the lack of similar comparable research to construct the research hypothesis, our article/research main hypothesis is that tourism statistical data can be linked with different statistical methods using observational data from Google and Apple, published in early 2020 about the mobility of different levels of communities in six different dimensions by Google and three by Apple. Thus, if verified, early predictive models, other econometric models (based on causal relationship between variables) and machine learning methods can be constructed based on Google and Apple mobility data for tourism arrivals (or other tourism indices) since the tourism statistical data are published with a 2–4-month delay. Because tourism statistical data are published with a 2–4-month delay, early predictive models and other econometric models based on Google and Apple mobility data can be constructed for tourism arrivals (or other tourism indices).

The variables used as inputs in the present research were collected as follows:

- Daily mobility data were collected from Google Mobility Community Reports (6 mobility indices) and Apple Mobility (3 mobility indices) reports from March 2020 to April 2021 (Table 1) for 11 European countries with completed data for all inputs in the study: Belgium, Finland, Norway, the Czech Republic, Germany, Estonia, Spain, Italy, Luxembourg, the Netherlands, and Slovakia. From daily data, the authors transformed it into monthly data by calculating the monthly mean average for each European country and every month, both for Google Mobility and for Apple Mobility. We chose these countries if their overnight data were available for the months, we analysed in the Eurostat database.
- The number of overnight stays was collected. The Eurostat database was used to collect monthly data on the number of nights spent in tourist accommodations by tourists in every chosen European country. from the Eurostat database. For better and more comparable data, the authors also transformed the tourism indicators into relative indicators, choosing a fixed base for February 2020 because the mobility indicators are being presented in Google and Apple reports.

For a better understanding of the measurements made by the Google and Apple mobility reports (Apple mobility reports, 2021), we present a short description of these indicators according to the initial reports. The authors ensured the representativeness of the data by introducing in the

study only 11 European countries with complete data, and it can be observed that these countries have a good distribution from a geographical perspective. We have a relatively good spatial representativeness of the research sample, and therefore, the present results can be inferred in the general population at the European level.

Based on the abovementioned particularities, the *research flowchart* together with statistical/econometrical/machine learning methods/analysis are presented in Figure 1.

As shown in Figure 1, a complex statistical methods were applied to analyze the relationship, causality, and association between the tourism indicator and the mobility of people measured by Apple and Google mobility dimensions:

1. Descriptive statistics (Gabor, 2013; 2016) (presented as the mean  $\pm$  standard deviation) for both Google and Apple mobility dimensions and overnight as a fixed base index for each of the 11 European countries (presented in Figure 2 from the next section); to analyze if they exist or not, along the 11 European countries, positive and/or negative linear standard deviations for the analyzed period and to establish what statistical methods to use.
2. The one-sample Kolmogorov–Smirnov test (Lambin, 1990; Aaker *et al.*, 1998; Jolibert & Jourdan, 2006; Fenneteau & Bialés, 1993; Vendrine, 1991; Churchill, 2001) was used to test the normal distribution of the data and to choose adequate statistical methods and tests according to these results.
3. Spearman’s rank correlation coefficient (Giannelloni & Vernet, 2003; Evrard *et al.*, 2003; Pupion & Pupion, 1998; Saporta, 1990; Gabor, 2016) was used to analyze the sense, intensity, and statistical significance of the association between GM6 and A3 mobility dimensions and overnight stays for the monthly data of the 11 European countries. Heatmaps were used to make better comparisons of the correlations (see Figure 3 in the next section).
4. Pearson correlation coefficient (Giannelloni & Vernet, 2003; Evrard, Pras & Roux, 2003; Pupion & Pupion, 1998; Saporta, 1990; Gabor, 2016) to analyze the sense, intensity, and statistical significance of the link between all daily data of the GM6 and A3 mobility indices for four European countries: Belgium, the Czech Republic, Germany, and Italy. Heatmaps were used to make better comparisons of the correlations (see Figure 4 in the next section).



5. One-way ANOVA (Gauthy *et al.*, 2005; Giannelloni & Vernetto, 2003; Gabor, 2016; Hayes, 1998) to test if there is statistically significant variability between the 11 European countries according to the average value of the mobility data.
6. The Kruskal–Wallis (independent samples) test (Gauthy *et al.*, 2005; d’Astous, 2005) is used to test if there are the same distributions along all 11 European countries for overnights and mobility indices and to determine whether the distributions of overnights and mobility indices are consistent across all 11 European countries.
7. An overall multilinear regression model (with collinearity diagnostic) (Baron *et al.*, 1996; Giannelloni & Vernetto, 2003; Malhorta, 2004; Vendrine, 1991; Jolibert & Jourdan, 2006) for all 11 European countries, with GM6 and A3 dimensions as independent variables and overnight stay as a fixed base index as the dependent variable, was used to identify the best and most statistically significant predictors of European tourism during the COVID-19 pandemic.
8. A set of methods using GM6 and A3 dimensions is as follows:
  - Factor analysis (PCA – Principal Component Analysis with Varimax rotation) (Malhorta, 2004; Jolibert & Jourdan, 2006; Fenneteau & Bialès, 1993) or regrouping the 9 mobility dimensions (GM6 and A3) into new factors (principal components) that will explain an important significant percentage of the total variance (at least 70%) and to determine, for the pandemic time, how mobility data are grouped into groups and which of them better explained the total variance, overall, for all 11 European countries from the research studied;
  - A proper regression model (Baron *et al.*, 1996; Giannelloni & Vernetto, 2001; Malhorta, 2004; Vendrine, 1991; Jolibert & Jourdan, 2006) starting with the principal components resulting from the factorial analysis (PCA) as independent variables and overnight stays as the fixed base index as the dependent variable is needed to validate (or not) the overall multilinear regression model from point 7.
9. The group of methods using only Google mobility data is as follows:
  - A factor analysis (PCA – principal component analysis with varimax rotation) (Malhorta, 2004; Jolibert & Jourdan, 2006; Fenneteau & Bialès, 1993) to regroup the GM6 mobility data into new factors that will explain at least 70% of the total variance (and to detect, for the COVID-19 pandemic time, in what way the mobility data are grouped, to detect how the mobility data are grouped for the COVID-19 pandemic time

and to identify GM6 and/or A3 data that better explained the total variance, overall, for all 11 European countries;

- A proper regression model (Baron *et al.*, 1996; Giannelloni & Vernet, 2001; Malhorta, 2004; Vendrine, 1991; Jolibert & Jourdan, 2006) using the principal components from PCA as independent variables and overnight stays as a fixed base index as a dependent variable.
10. Neural networks (Kolková & Ključnikov, 2021; Barclay *et al.*, 1995; McCormick & Salcedo, 2017) with the multilayer perceptron (MLP) algorithm were used to find the best mobility data predictor overnight stays for all 11 European countries separately for Google Mobility data and Apple Mobility data. Since regression analysis, in many circumstances, is not sufficient to explain the association between the predictors and outcomes and is not sufficient to accommodate all the ways in which the values of one predictor may affect the impact of other predictors, we also applied neural network analysis by using SPSS software.

The statistical methods from points 8 and 9 were used to analyze which of the two mobility data sets, Apple or Google, are more determinant and good predictor for European tourism during the COVID-19 pandemic.

## Results

The starting point for the complex statistical analysis, using inferential statistics and descriptive and explicative methods for data analysis, is the results of the descriptive statistics indicators, the mean and standard deviation of each variable for all 11 European countries.

In Tables 2–5, there are the results for descriptive statistics for all variables from the study as linear mean deviation in Table 2–4 and mean+ SD in Table 5). In Tables 2–4, positive linear mean deviations are marked with blue and the negative one with red for each country. According to the results from Tables 2–4, there are many differences between the 11 countries. For example, from Figure 2, overnight (number), Germany has the largest positive deviation from the mean of all countries, but at the same time, it has a negative deviation for overnights as a fixed base index (with February 2020 as the fixed base). Only Italy has both positive linear deviations for overnights as a number and for overnights as a fixed base index. Additionally, the Netherlands has a positive linear mean deviation for overnights as a number and the largest positive linear mean deviation for over-

nights as a fixed base index (+38.46%), followed by Italy (+29.70%) and Luxembourg (+21.54%). All these differences need to be analyzed with complex statistical methods to determine whether there are interdependence and/or determination relationships.

According to the results from Table 5, there are many differences between the 11 countries. Regarding the Google Mobility (GM) indices, there are two atypical countries:

- For the GM Workplace index, it can be seen from Table 4 that Spain has the largest negative linear mean deviation.
- For the GM residential index, Finland has the largest positive deviation compared with the rest of the countries.
- For the indices GM Retail & Recreation, GM Grocery & Pharmacy, and GM Parks, we can observe a positive group of deviations formed by Finland, Norway, and Estonia and a negative group formed by Spain, Italy, and Luxembourg. For the GM Transit Stations index, the extremes are Estonia (+12.64) and the Netherlands (-11.56).

Regarding the Apple Mobility indices (Table 3), Germany and Estonia consistently have a positive deviation for all 3 Apple indices, with mention that for Apple Walking, Finland (+20.56) and Norway (+25.99) have the highest level of positive linear mean deviation. Italy and the Czech Republic are the only countries with a negative linear mean deviation for all 3 Apple indices, with mention that the Czech Republic has a low level of negative linear mean deviation (-60.34) for the Apple Transit index and Norway for apple driving (-23.16).

Regarding the Apple Mobility indices, Germany and Estonia consistently have a positive linear mean deviation for all three Apple indices, with Finland (+20.56) and Norway (+25.99) having the highest level of positive deviation for Apple Walking. Italy and the Czech Republic are the only countries with a negative deviation for all three Apple indices, with the Czech Republic having the lowest level of negative linear mean deviation (-60.34) for the index Apple Transit and Norway having the highest level of negative linear mean deviation (-23.16).

Due to the abovementioned positive and negative differences between the 11 countries for all 9 mobility indices and overnights, we test the normal distribution of data by using the one-sample Kolmogorov–Smirnov test. The results indicate that only for these variables do the data have a normal distribution: GM Retail & Recreation, GM Transit Stations, GM

Residential, and Apple Transit. Therefore, for statistical analysis, we predominantly used nonparametric methods, including correlations.

To determine the direction and intensity of the correlation between Google Mobility and Apple indices, we applied nonparametric Spearman correlations. At the level of all 11 European countries, the correlation matrices are presented below (Table 6) using heatmap diagrams (the red color for inverse correlations and the blue color for direct correlations), and the statistically significant correlations are also marked. At the level of the 11 European countries, all correlations are statistically significant either of low intensity or of high intensity. All correlations are statistically significant, either low or high intensity, at the level of the 11 European countries. It is also observed that the GM residential indicator negatively correlates with all other indicators, thus:

- Strong intensity negative correlation with X1 (-0.914), X4 (-0.818), overnights as fixed base index (Feb. 2020) (-0.739).
- The fixed base index (Feb. 2020) has a strong intensity negative correlation with X1 (-0.914), X4 (-0.818), and overnights.
- Medium to a strong negative correlation with X3 (-0.712), X2 (-0.653), and X5 (-0.698).

There are also positive correlations:

- Strong intensity correlation between Y and X1 (0.823), X7 (0.741), and X9 (0.758).
- There are also some strong correlations between X1 and X3, X4, X7, and X9, which would suggest the higher necessity of mobility for shopping, recreation, and visiting parks.

We have also chosen four European countries that have daily data for the mobility indicators for both Google and Apple: Belgium, the Czech Republic, Germany, and Italy (432 days between 2/15/2020 and 4/21/2021). We have checked the correlations between each Google and Apple mobility dimension. The synoptic table in the form of a heatmap of these values is presented in Table 7. Generally, these daily data have better correlations (and are more accurate due to the higher number of cases) than the monthly data. In general, daily data have better correlations (and are more accurate due to the higher number of cases) than monthly data. It can be said that in these four countries, there are good correlations between Apple Driving and GM Retail & Recreation, GM Parks and GM Transit stations, Apple Transit with GM retail and transit, Apple walking with GM Retail & Recreation, GM Parks, and GM Transit. GM Grocery & Pharmacy and GM

Workplace do not correlate too strongly with the Apple mobility dimensions, but they show a relatively stronger correlation with Apple transit. GM Residential has strong negative correlations with all three Apple mobility dimensions. GM Grocery & Pharmacy and GM Workplace have a weak correlation with the Apple mobility dimensions but a stronger correlation with Apple Transit. GM Residential has strong negative correlations with all three Apple mobility dimensions.

GM Retail & Recreation, which also contains a strong recreation-oriented dimension, along with GM Parks and possibly transit, are the "fun" parts, the recreation dimensions, and the new places where people use their phone more for orientation. The grocery (and pharmacy) and workplace are the routine dimensions, where there is no need for location determination or route-finding applications.

We note that the negative correlation of the GM Residential index with all Apple mobility indices for each of the four countries is confirmed, pointing out, once again, that the Residential dimensions show the "not-mobility" or the home staying trend of people. The negative correlation of the GM Residential index with all Apple mobility indices for each of the four countries is confirmed, emphasizing once again that the Residential dimensions reflect people's "not-mobility" or home-staying trend.

Because the descriptive statistics in Tables 2–4 show positive and negative deviations of significantly different values, we tested with nonparametric and parametric statistical tests if there were statistically significant differences in the values or averages of the indicators between countries. One-way ANOVA indicates results — except for the variable overnight stay -fixed base indicator (February 2020) — for all variables, there are or are not statistically significant differences in the variations from the average of the data depending on the country. Additionally, p values > 0.05 but below 0.1 were recorded for the Apple Driving (0.063) and GM Workplace (0.063) indicators.

The Kruskal–Wallis test (independent samples) was applied to test whether the distributions of Google Mobility and Apple indices are the same in the 11 countries from the study. The results indicate that for the following variables, there are statistically significant differences between countries: overnight stays — number (0.000), Google Mobility Retail (0.008), Google Mobility Grocery (0.000), Google Mobility Parks (0.001), Google Mobility Transit (0.001), Google Mobility Residential (0.003), and for all 3 indicators Apple: Apple Driving (0.038), Apple Transit (0.000), and

Apple Walking (0.000). There are no differences between countries regarding only the distribution of overnight stays in the FB index ( $p$  value = 0.083) and for Google Mobility Workplace ( $p$  value = 0.066).

Because only 4/11 variables had normally distributed data (according to the results of the Kolmogorov–Smirnov test), we applied the median test to all variables. The results indicate that there are statistically significant differences between countries for all variables except Google Mobility Workplace (0.160) and Google Mobility Residential (0.249), which are practically for the most important indicators of the pandemic time for the entire world, not only for European countries.

To determine whether there is a causal relationship between Google and Apple mobility indices and overnight stays (point (7) from Methodology), we applied the multilinear regression model with overnight stays as dependent variables and all 9 mobility indices as independent variables for the 11 countries and for 14 months. First, we applied the method both overnight stay as a number and as a fixed base index as a dependent variable. Since for the model with an overnight stay as the dependent variable, the  $R^2$  coefficient was low, we decided to continue the analysis only overnight stays as a fixed base index. Due to the powerful and statistically significant correlations of each mobility index with overnights as a fixed base, multicollinearity comes with the analysis. The results are presented in the tables below, after multicollinearity test we excluded GM Residential (X6) dimensions due to the collinearity of this variable (VIF = 19.320). For analysis, linear regression with the Enter method was used

The determinant coefficients  $R^2$  is 0.615 meaning that over 60% of total variance is explained by the GM and Apply data. Additionally, ANOVA test results indicate statistical significance of the model ( $p$  value = 0.000).

For the final model the results for the multicollinearity test indicate that all VIF values are between 1 and 10 (Table 8).

According to the Beta standardized coefficient values (Table 8), the following independent variables have a statistically significant contribution and explain the variance of overnight stays as a fixed base index across all 11 European countries: GM Retail & Recreation ( $p$  value = 0.000; SC Beta = +0.573); Apple Driving ( $p$  value = 0.003; SC Beta = +0.370); and Google Mobility Workplace ( $p$  value = 0.018; SC Beta = +0.370):

- for increasing the overnights by one unit of GM Retail & Recreation as the fixed base index increases by 0.573.

- as the fixed base index rises by 0.370 points.
- the overnights as the fixed base index decrease by 0.244 for every unit of GM increase.

According to the results of the regression model, only 3 of the 8 independent variables have a statistically significant contribution to the explanation of overnight stays for the countries included in the study. Therefore, we continued the analysis, and we applied the factorial analysis using the 8 independent variables to find new factors (point (8) from Research Methods section). Therefore, we applied PCA with Varimax rotation and Kaiser normalization using the criteria of initial eigenvalues  $> 1$ . The results show that two principal components explain a significant percentage of the total variance, 79.73 percent.

According to the results of the rotated component matrix (Table 9), the 8 variables are grouped as follows (see also Figure 2):

- The principal component PC1 is formed by the following: GM Parks, Apple Walking, Apple Driving, GM Retail & Recreation and Apple Transit, practically all the mobility indices that indicate the leisure mobility of the people. PC1 explains 66.203% of the total variance and can be named "Leisure and recreation activities".
- The principal component PC2 is formed by the following: GM\_Workplace, GM\_Grocery & Pharmacy, GM\_Transit, practically all the indices that indicate domestic or routine activities. PC2 explains 13,52% of the total variance and can be named "Domestic and routine activities".

To analyze which component resulting from PCA is a good predictor of overnight stays, we applied regression analysis (point (8) from Research Methods section), but used these two factors as independent variables and overnight stay as a dependent variable. The results show a good value for the  $R^2$  determinant coefficient (0.556) and a statistical significance for the model (ANOVA p value = 0.000).

The regression coefficients (Table 10) indicate that both principal components are good predictors for overnight with a fixed base index (p value  $< 0.05$ ).

According to the results from Table 10, PC1 – "Leisure and recreation activities" – has an important and direct contribution for overnights with a fixed base for all 11 European countries (0.729), followed by the second factor PC2 – "Domestic and routine activities" (0.143). Therefore, with this statistical approach, we discover that not only GM Retail & Recreation and

Apple Driving but also GM Parks, Apple Walking and Apple Transit are good predictors of overnight stays for these countries. Below, we present the histogram (Figure 3) and normal P-P plot (Figure 4) for regression model with factors from PCA by using all mobility data (GM and Apple).

In the Research Methods section, last paragraph, point (9), we mention that we applied the same combination of methods (regression and factorial analysis) but only using GM data as independent variables. In the next paragraphs, we present these results. This part is dedicated to the validation analysis (or not) of the previous results obtained. For the first model of linear regression analysis with only five Google Mobility indices as independent variables (except Google Mobility Residential with collinearity) and overnights as a dependent variable, we obtain a good value of determinant coefficient of the model  $R^2$  (0.578), the model statistically significant ( $p$  value = 0.000) according to the ANOVA results. In this model, there are no collinearity statistics, and all five VIF values are between 1 and 10. The results are presented in Table 11.

As shown in Table 11, 3 from 5 of the GM data are statistically significant ( $p$  value < 0.05): GM Retail & recreation, GM Grocery & Pharmacy, and GM Parks. However, if we apply the statistical regression rule, which accepts all variables with a  $p$  value of 0.1 as statistically significant, we find that all five GM data are statistically significant.

For the application of factor analysis (PCA with Varimax rotation and Kaiser normalization) accordingly with point (9) from research Methods section, we consider the rule of a minimum of six variables for PCA, and therefore, we rolled the method with all GM indices. Additionally, the number of factors was manually chosen, with the total variance explained being 88.23%. The rotated component matrix is presented in Table 12.

By comparing these results with the previous results of factorial analysis (with all mobility data, GM and Apple indices), we conclude that the groupings are the same (except for the Apple Mobility indices), but there are changes in the order of factors. PC1 from the previous analysis is now the second (PC2 in Table 12), formed by GM Parks and GM Retail & Recreation, and it explains 38.53% of the total variance explained. This principal component refers to the dynamics or movement of people for recreation and leisure activities specific to tourism. The second factor from the previous analysis is, in these conditions, the first one (PC1 in Table 12) with GM Residential. This principal component refers to the domestic/static activities specific to lockdown time during a pandemic. PC1 from this analysis ex-



plains 49.70% of the total variance explained. In Figure 5, we present the component plot in the rotated space.

As in the previous analysis, we continue by applying the linear regression model with the principal components as independent variables and the overnights as dependent variables. The value of determinant coefficient  $R^2$  is 0.533 for this model, and the ANOVA indicates a statistically significant model (p value = 0.000). The coefficient table is presented in Table 13.

The results indicate that both factors are statistically significant (0.007 and 0.000) and, most importantly, they validate the previous results. More specifically, the PC2 (GM) formed by GM Parks and GM Retail & Recreation is a good predictor alone (without Apple Mobility indices such as in the first analysis) of overnight stays as a fixed base index. For each one-unit increase in PC2 (GM), the overnight stays increase by 0.711 units.

Because regression analysis is, in many circumstances, not sufficient to explain the association between the predictors and outcomes and is not sufficient to accommodate all the ways in which the values of one predictor may affect the impact of other predictors, we also applied neural network analysis with the multilayer perceptron (MLP) algorithm by using SPSS software. We chose MLP because of its flexibility and lack of distribution assumptions for the data we were analyzing.

The results of the neural network for GM data indicate good predictors overnight stays as a fixed base indicator and are presented in Figure 6 and Table 14.

The input layers of the MLP neural network had all the GM data as a continuous variable. As a fixed base (February 2020), the output layer was the overnight stay. We choose one hidden layer and one node as the model architecture, and we also show the synaptic weights. The blue color of the synaptic weights in Figure 6 represents weight 0, while the gray color represents weight  $> 0$ . The difference between the contribution and prediction of GM Parks to European tourism (positive, gray synaptic weight) and, for example, GM Workplace (negative, blue synaptic weight) can be seen.

The model returns four hidden layer nodes: H (1:1), H (1:2), H (1:3), and H (1:4), but the synaptic weights are different, opposite for hidden layer H (1:4) compared with the other three hidden layers in the network. In Table 15, the parameter estimates are presented. The results for GM data show that hidden layer H (1:4), composed of GM Parks in collaboration with GM Retail & Recreation, is the best predictor of tourism during the COVID-19

pandemic for the 14 months included in the research, with a direct, positive contribution (0.991).

The indirect contribution as a predictor for tourism is given by the hidden layer H (1:1), composed of the opposite of GM Transit (-0.589) and GM Residential (0.595)

In Figure 8, the normalized importance of independent variables for GM data indicates that GM Parks and GM Residential have a 100% percentage of importance opposite of GM Transit, which has more importance than other variables for the outcome, overnight stays. According to the findings, GM Parks and GM Residential are extremely important in all eleven European countries.

The results of the neural network for Apple Mobility data indicate good predictors overnight stays as a fixed base indicator and are presented in Figure 7 and Table 15.

For Apple Mobility data, the MLP – neural network returns two hidden layers, H (1:1) and H (1:2), but only one of them has a significant contribution as a predictor for tourism. The hidden layer H (1:2) (1.098) combines Apple driving (0.595) and Apple Walking (0.515), opposite to Apple Transit (-0.121).

In Figure 9, the normalized importance of independent variables for Apple Mobility data confirmed these results with 100% importance of Apple Driving for overnights as a predictor.

Therefore, we find that the best predictors for European tourism during the COVID-19 pandemic are as follows:

- for Google Mobility data: the Google Mobility Parks together with Google Mobility Retail & Recreation;
- for Apple Mobility data: the Apple Driving together with Apple Walking.

## **Discussions**

In our exploratory research, we connected the observational daily data of As we presented before, Google collects mobility data in six dimensions: Retail and Receptions, Parks, Grocery and Pharmacy, Transit, Residential and Workplace, and Apple collects data in three different dimensions: Walking, Driving and Transit.

We also found strong parametric and nonparametric correlations between GM Retail&Recreation and GM Parks, GM Retail & Recreation and GM Transit stations, GM Retail &Recreation with Apple Walking and Apple Driving, and GM Parks and Apple Walking. These links made us believe that these dimensions can reflect the recreation time and activities of the populations. There were also some weak relations between GM Parks and GM Workplace. GM Workplace and Apple Walking, signaling that some other dimensions reflect the other type, not so much the leisure but routine-related activities of people. There was GM Residential, which correlates negatively and strongly with everything else, signaling that when people stay at home, they are not outdoors, so the GM Residential data are also reliable.

The same results and connections were also revealed from the daily data analysis. Here, we compared Apple data with Google data, which showed an even higher correlation, approximately 0.75 – 0.85, between the following data pairs: Apple Driving with GM Retail & Recreation, GM Parks, GM Transit stations, Apple Transit with GM Retail & Recreation, GM Transit stations, and Apple walking with GM Retail & Recreation and GM Parks. GM Residential & Recreation correlates negatively and strongly with all three Apple dimensions at the level of daily data as well. These results show that people movement is reflected well and alike by the two big tech companies and their data, but some dimensions tend to “go together,” such as Retail and Recreation with Walking, Driving, Parks and Walking, while Parks does not match with Transit, Workplace, etc.

However, the main question was whether these observational data match the statistical data of tourism, analyzed with a regression model. Our first regression model, which used all the Apple and Google dimensions (except the Residential), was below our expectations. Most of the dimensions were weak predictors and/or with no significance. However, when we reduced the number of factors with principal component analysis, we obtained two excellent predictors, one containing GM Parks, Apple Walking, Apple Driving, GM Retail & Recreations and Apple Transit, named PC1 – Leisure and recreation activities, and another containing GM Workplace, GM Grocery & Pharmacy, GM Transit, named PC2 – Domestic and routine activities. The first one (PC1 – Leisure and recreation activities) explains more than 66% of the change in tourism arrivals, and the second (PC2 – Domestic and routine activities) explains only 13,5%. After all, we can conclude that some dimensions of mobility are strongly

related to tourism activities. In this way, these dimensions, especially GM Retail and Recreations and GM Parks, can predict (type nowcasting) tourism arrivals in most countries, since the GM data can be seen just days after the data were collected, but tourism arrivals are released 2–4 months later

We discovered strong parametric and nonparametric correlations between GM Retail & Recreation and GM Parks, GM Retail & Recreation and GM Transit stations but also between GM Retail & Recreation and Apple Walking, Apple Driving, and GM Parks. These connections led us to believe that these dimensions could reflect people's leisure time and activities. There were also some strained relationships between GM Parks and GM Workplace. GM Workplace and Apple Walking, signaling that some other dimensions reflect other types of activities, not so much leisure but routine-related ones. There was GM Residential, which strongly correlates negatively with everything else, signaling that when people stay at home, they are not outdoors, so the GM Residential data are also reliable.

The same results and connections were also revealed from the daily data analysis. Here, we compared Apple data with Google data, and they showed an even higher correlation, approximately 0.75 to 0.85, between the following data pairs: Apple Driving with GM Retail & Recreation, GM Parks, and GM Transit Stations; Apple Transit with GM Retail & Recreation, GM Transit Stations; and Apple Walking with GM Retail & Recreation and GM Parks To mention, GM Residential & Recreation correlates negatively and strongly with all three Apple dimensions at the level of daily data as well. These results show that people's movement is reflected well and alike by the two big tech companies and their data, but some dimensions tend to "go together," such as retail and recreation with walking, driving, parks, and walking, while parks do not match with transit, workplace, etc.

## **Conclusions**

In the research, we connected the observational daily data from Google and Apple, resulting from location tracking technologies on smartphones, from 11 European countries during the COVID-19 pandemic's months in 2020 and 2021 with other tourism statistical data from the same countries. The study's goal was to determine whether there was a strong, medium, or consistent link between these data from various countries and seasons. Our

results show that there is a reliable connection between the movement of people in certain countries (Beckers *et al.*, 2021), reflected in their real-time smartphone tracking technologies and statistical data from tourism service providers, food, hotels, and/or other accommodations (Hall *et al.*, 2022). Nonetheless, these connections are different, not so much determined by the source (provider) as by the dimensions of observations. As we presented before, Google collects mobility data in six dimensions: retail and recreation, parks, grocery and pharmacy, transit, residential, and workplace, while Apple collects data in three different dimensions: walking, driving, and transit. It appears that there is a very strong link between tourism flows (overnight stays) and GM retail and recreation by all methods, as well as a strong link with Apple Walking, Driving, and GM parks. These categories best describe tourism activity (represented here by overnight stays) and appear to be the leisure and recreation categories of mobility. The two transit dimensions of Apple and Google only partially describe tourism flows, possibly because public transport transit also includes other types of movement. Additionally, there is a negative and strong relationship with the GM Residential dimension, which also seems logical: if people spend more time at home, they are not traveling. Finally, there is a weak link with the GM Grocery and Pharmacy and Workplace dimensions. These could be the routine dimensions of mobility, with little or no relation to tourism activities.

In terms of recommendations for practice, it can be interesting to download data for each locality — since the Google data is available at the level of towns — and compare it with different tourism data (overnights, arrivals, but also with tourism resources and supply, such as accommodation services, food and beverage locations, cultural productions). There is a possible connection with these statistical data, and if one can construct a similar model, it can be used for predicting a more precise short-term tourism evolution, since it will be linked only to the local mobility data.

Our model also allows to predict on short term the evolution of statistical data for overnights. Especially GM Retail and Recreation and GM Parks can predict (type nowcast) the tourism overnights in most of the countries, since the GM data can be seen just days after the data were collected, but the tourism arrivals are released 2–4 months later.

We can also add to the implications the possibility of identifying statistical data collection problems at the local level. If there is a good model between mobility and tourism data at the national level, but there is a sig-

nificant gap between the two dimensions at the local level, it may indicate a data collection problem. This is because mobility data is more reliable as they are exhaustive and not declarative (but observational).

The limitations of the present research are as follows: (1) Google and Apple mobility data includes other dimensions in such as shopping, online retail, accommodation, urban transit travel, etc.; (2) there are missing values for many European countries (members and non-members of the European Union), this way we could not run the comparisons in the case of other countries; (3) lack of similar research results to compare the present research results; and (4) lack of a similar situation worldwide for objective comparisons of the present results; (5) We had chronological data for only 14 months (in addition to the base period), which is relatively short; (6) Different countries usually represent different cultures, which can create different types of tourism and different patterns of mobility (and different locations), which could be a challenge in comparing and measuring the same flows).

Similar research has been conducted in European countries, but with online retail as the dependent variable, high-frequency data on GPS-based population mobility and government stringency during the COVID-19 pandemic, and only one dimension from Google Mobility Reports, GM Residential (Szász *et al.*, 2022). Therefore, we consider the present research an important contribution to the existing theoretical and practical knowledge of the relationship between mobility data (Apple and Google) and tourism indicators (overnight stays).

For future research, the authors intend to conduct a detailed analysis at the national level for Romania, with data being available at the county level, to apply other quantitative methods (such as gray relational analysis) and to include other important available dimensions for the tourism industry. Other further research is necessary to compare the results across different time periods and geographical locations to establish a short-term forecasting model for tourism. Another important question remains on what influences mobility? We have successfully pointed out the extent to which mobility can affect tourism flows, but the question remains: what defines mobility as a social characteristic? What are the dimensions of more movement in some countries and how are these determined by other economic, social and cultural factors?

This is because Google and Apple mobility data appear immediately, while tourism statistical data are typically published with a delay of two to three months.

## References

- Aaker, D. A., Kumar, V., & Day, G. S. (1998). *Marketing research*. Indianapolis: John Wiley & Sons.
- Apple mobility database (2021). Apple makes mobility data available to aid COVID-19 efforts. Retrieved from <https://www.apple.com/newsroom/2020/04/apple-makes-mobility-data-available-to-aid-covid-19-efforts> (31.05.2021).
- Apple mobility reports (2021). COVID-19 - Mobility Trends Reports. Retrieved from <https://covid19.apple.com/mobility> (31.05.2021).
- Atalay, S., & Solmazer, G. (2021). The relationship between cultural value orientations and the changes in mobility during the Covid-19 pandemic: a national-level analysis. *Frontiers in Psychology*, 12, 578190. doi: 10.3389/fpsyg.2021.578190.
- Bangwayo-Skeete, P. F., & Skeete, R. W. (2015). Can Google data improve the forecasting performance of tourist arrivals? Mixed-data sampling approach. *Tourism Management*, 46, 454–464. doi: 10.1016/j.tourman.2014.07.014.
- Barclay, D., Higgins, C., & Thompson, R. (1995). The partial least squares (PLS) approach to causal modeling: Personal computer use as an illustration. *Technological Studies*, 2(2), 285–309.
- Baron, T., Biji, E., Tövisi, L., Wagner, P., Isaic-Maniu, Al., Korca, M., & Porojan, D. (1996). *Theoretical and economics statistics*. Bucharest: Didactică și Pedagogică.
- Beckers, J., Weekx, S., Beutels, P., & Verhetsel, A. (2021). COVID-19 and retail: The catalyst for e-commerce in Belgium? *Journal of Retailing and Consumer Services*, 62, 102645. doi: 10.1016/j.jretconser.2021.102645.
- Bengtsson, L., Gaudart, J., Lu, X., Moore, S., Wetter, E., Sallah, K., Rebaudet, S., & Piarroux, R. (2015). Using mobile phone data to predict the spatial spread of cholera. *Scientific reports*, 5(1), 8923. doi: 10.1038/srep08923.
- Brodeur, A., Gray, D., Islam, A., & Bhuiyan, S. (2021). A literature review of the economics of COVID-19. *Journal of Economic Surveys*, 35(4), 1007–1044. doi: 10.1111/joes.12423.
- Churchill, G. A. (2001). *Basic marketing research*. The Dryden Press.
- Cot, C., Cacciapaglia, G., & Sannino, F. (2021). Mining Google and Apple mobility data: Temporal anatomy for COVID-19 social distancing. *Scientific Reports*, 11(1), 4150. doi: 10.1038/s41598-021-83441-4.
- d’Astous, A. (2005). *Design of marketing research*. Montréal: Chenelière Éducation.
- Eurostat (2021). Nights spent at tourist accommodation establishments by residents/non-residents. Retrieved from <https://ec.europa.eu/eurostat/databrowser/view/tin00171/default/table?lang=en> (15.06.2021).

- Evrard, Y., Pras, B., & Roux, E. (2003). *Marketing- studies and researches in marketing*. Paris: Dunod.
- Fenneteau, H., & Bialès, C. (1993). *Da statistical analysis – applications and study cases for marketing*. Paris: Ellipses.
- Finger, F., Genolet, T., Mari, L., de Magny, G. C., Manga, N. M., Rinaldo, A., & Bertuzzo, E. (2016). Mobile phone data highlights the role of mass gatherings in the spreading of cholera outbreaks. *Proceedings of the National Academy of Sciences*, 113(23), 6421–6426. doi: 10.1073/pnas.1522305113.
- Fritz, C., Dorigatti, E., & Rügamer, D. (2022). Combining graph neural networks and spatio-temporal disease models to improve the prediction of weekly COVID-19 cases in Germany. *Scientific Reports*, 12(1), 3930. doi: 10.1038/s41598-022-07757-5.
- Gabor, M. R. (2013). *Market prospecting through statistical methods*. Bucharest: CH Beck.
- Gabor, M.R. (2016). *Analysis and inference of statistical data*. Bucharest: CH Beck.
- García-Cremades, S., Morales-García, J., Hernández-Sanjaime, R., Martínez-España, R., Bueno-Crespo, A., Hernández-Orallo, E., López-Espín, J. J., & Cecilia, J. M. (2021). Improving prediction of COVID-19 evolution by fusing epidemiological and mobility data. *Scientific Reports*, 11(1), 15173. doi: 10.1038/s41598-021-94696-2.
- Gauthy Sinéchal, M., & Vandercammen, M. (2005). *Marketing studies – methods and tools*. Buxelles: De Boeck & Larcier.
- Geng, D., Innes, J., Wu, W., & Wang, G. (2021). Impacts of COVID-19 pandemic on urban park visitation: A global analysis. *Journal of Forestry Research*, 32, 553–567. doi: 10.1007/s11676-020-01249-w.
- Ghorbani, A., Mousazadeh, H., Akbarzadeh Almani, F., Lajevardi, M., Hamidizadeh, M. R., Orouei, M., Zhu, K., & Dávid, L. D (2023). Reconceptualizing customer perceived value in hotel management in turbulent times: A case study of Isfahan metropolis five-star hotels during the COVID-19 Pandemic. *Sustainability*, 15(8), 7022. doi: 10.3390/su15087022.
- Giannelloni, J. L., & Vernet, E. (2003). *Marketing studies*. Paris: Vuibert.
- Google Mobility database (2021). Mobility Report CSV Documentation. Retrieved from [https://www.google.com/covid19/mobility/data\\_documentation.html?hl=en](https://www.google.com/covid19/mobility/data_documentation.html?hl=en) (31.05.2021).
- Google mobility reports (2022). Community mobility reports. Retrieved from <https://www.google.com/covid19/mobility/> (20.12.2022).
- Hakim, A. J., Victory, K. R., Chevinsky, J. R., Hast, M. A., Weikum, D., Kazazian, L., Bhatkoti, R., Schmitz, M. M., Lynch, M., & Marston, B. J. (2021). Mitigation policies, community mobility, and COVID-19 case counts in Australia, Japan, Hong Kong, and Singapore. *Public Health*, 194, 238–244. doi: 10.1016/j.puhe.2021.02.001.



- Hall, M. C., Prayag, G., Fieger, P., & Dyason, D. (2020). Beyond panic buying: Consumption displacement and COVID-19. *Journal of Service Management*, 32(1), 113–128. doi: 10.1108/JOSM-05-2020-0151.
- Hayes, B. E. (1998). *Measuring customer satisfaction – survey design, use and statistical analysis methods*. Wisconsin: ASQ Quality Press.
- Ibarra-Espinosa, S., de Freitas, E. D., Ropkins, K., Dominici, F., & Rehbein, A. (2021). Association between COVID-19, mobility and environment in São Paulo, Brazil. *medRxiv*. doi: 10.1101/2021.02.08.21250113.
- Irini, F., Kia, A. N., Shannon, D., Jannusch, T., Murphy, F., & Sheehan, B. (2021). Associations between mobility patterns and COVID-19 deaths during the pandemic: A network structure and rank propagation modelling approach. *Array*, 11, 100075. doi: 10.1016/j.array.2021.100075.
- Jacobsen, G. D., & Jacobsen, K. H. (2020). Statewide COVID-19 stay-at-home orders and population mobility in the United States. *World Medical & Health Policy*, 12(4), 347–356. doi: 10.1002/wmh3.350.
- Jolibert, A., & Jourdan, Ph. (2006). *Marketing research – research methods and marketing studies*. Paris: Dunod.
- Kolková, A., & Ključnikov, A. (2021). Demand forecasting: An alternative approach based on technical indicator Pbands. *Oeconomia Copernicana*, 12(4), 1063–1094. doi: 10.24136/oc.2021.035.
- Kraemer, M. U., Yang, C. H., Gutierrez, B., Wu, C. H., Klein, B., Pigott, D. M., Open COVID-19 Data Working Group, Faria, N. R., Li, R., Hanage, W. P., Brownstein, J. S., Layan, M., Vespignani, A., Tian, H., Dye, C., Pybus, O. G., & Scarpino, S. V. (2020). The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science*, 368(6490), 493–497. doi: 10.1126/science.abb4218.
- Lai, S., Ruktanonchai, N. W., Zhou, L., Prosper, O., Luo, W., Floyd, J. R., Wesolowski, A., Santillana, M., Zhang, C., Du, X., Yu, H., & Tatem, A. J. (2020). Effect of non-pharmaceutical interventions to contain COVID-19 in China. *Nature*, 585(7825), 410–413. doi: 10.1038/s41586-020-2293-x.
- Lambin, J. J. (1990). *Marketing research- analysis, measure, forecasting*. Paris: McGraw.
- Lyons, K. (2020). Governments around the world are increasingly using location data to manage the coronavirus. Retrieved from <https://theverge.com/2020/3/23/21190700/eu-mobile-carriers-customer-data-coronavirus-south-korea-taiwan-privacy> (31.05.2021).
- Malhorta, N. (2004). *Marketing studies with SPSS*. Paris: Pearson Education France.
- McCormick, K., & Salcedo, J. (2017). *SPSS statistics for data analysis and visualization*. Indianapolis: John Wiley & Sons.
- Mousazadeh, H., Ghorbani, A., Azadi, H., Almani, F. A., Zangiabadi, A., Zhu, K., & Dávid, L. D. (2023). Developing sustainable behaviors for underground heritage tourism management: The case of Persian Qanats, a UNESCO World Heritage Property. *Land*, 12(4), 808. doi: 10.3390/land12040808.

- Munawar, H. S., Khan, S. I., Qadir, Z., Kouzani, A. Z., & Mahmud, M. P. (2021). Insight into the impact of COVID-19 on Australian transportation sector: An economic and community-based perspective. *Sustainability*, 13(3), 1276. doi: 10.3390/su13031276.
- Murray, T. (2021). Stay-at-home orders, mobility patterns, and spread of COVID-19. *American Journal of Public Health*, 111(6), 1149–1156. doi: 10.2105/AJPH.2021.306209.
- Nagy, B., Gabor, M. R., & Bacoş, I. B. (2022). Google mobility data as a predictor for tourism in Romania during the COVID-19 pandemic—A structural equation modeling approach for big data. *Electronics*, 11(15), 2317. doi: 10.3390/electronics11152317.
- Oliver, N., Lepri, B., Sterly, H., Lambiotte, R., Deletaille, S., De Nadai, M., Letouze, E., Ali Salah, A., Benjamins, R., Cattuto, C., Colizza, V., Cordes, N., Fraiberger, S. P., Koebe, T., Lehmann, S., Murillo, J., Pentland, A., Pham, P. N., Pivetta, F., Saramaki, J., Scarpino, S. V., Tizzoni, M., Verhulst, S., & Vinck, P. (2020). Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle. *Science Advances*, 6(23), eabc0764. doi: 10.1126/sciadv.abc0764.
- Pupion, G., & Pupion, P.C. (1998). *Non - parametric tests with economics applications and administration*. Paris: Economica.
- Saha, J., Mondal, S., & Chouhan, P. (2021). Spatial-temporal variations in community mobility during lockdown, unlock, and the second wave of COVID-19 in India: A data-based analysis using Google's community mobility reports. *Spatial and Spatio-temporal Epidemiology*, 39, 100442. doi: 10.1016/j.sste.2021.100442.
- Saporta, G. (1990). *Data probabilistic analysis and statistics*. Paris: Technip.
- Shortall, R., Mouter, N., & Van Wee, B. (2022). COVID-19 passenger transport measures and their impacts. *Transport Reviews*, 42(4), 441–466. doi: 10.1080/01441647.2021.1976307.
- Sulyok, M., & Walker, M. D. (2021). Mobility and COVID-19 mortality across Scandinavia: A modeling study. *Travel Medicine and Infectious Disease*, 41, 102039. doi: 10.1016/j.tmaid.2021.102039.
- Szász, L., Bálint, C., Csíki, O., Nagy, B. Z., Rácz, B. G., Csala, D., & Harris, L. C. (2022). The impact of COVID-19 on the evolution of online retail: The pandemic as a window of opportunity. *Journal of Retailing and Consumer Services*, 69, 103089. doi: 10.1016/j.jretconser.2022.103089.
- Tamagusko, T., & Ferreira, A. (2020). Data-driven approach to understand the mobility patterns of the Portuguese population during the COVID-19 pandemic. *Sustainability*, 12(22), 9775. doi: 10.3390/su12229775.
- Tizzoni, M., Bajardi, P., Decuyper, A., Kon Kam King, G., Schneider, C. M., Blondel, V., Smoreda, Z., González, M. C., & Colizza, V. (2014). On the use of human mobility proxies for modeling epidemics. *PLoS Computational Biology*, 10(7), e1003716. doi: 10.1371/journal.pcbi.1003716.
- Vendrine, J. P. (1991). *Data processing in marketing - in 10 questions, 13 applications, 27 examples and commented exercises*. Paris: Les Éditions d'Organisation.

- Wang, S., Tong, Y., Fan, Y., Liu, H., Wu, J., Wang, Z., & Fang, C. (2021). Observing the silent world under COVID-19 with a comprehensive impact analysis based on human mobility. *Scientific Reports*, 11(1), 14691. doi: 10.1038/s41598-021-94060-4.
- Wesolowski, A., Eagle, N., Tatem, A. J., Smith, D. L., Noor, A. M., Snow, R. W., & Buckee, C. O. (2012). Quantifying the impact of human mobility on malaria. *Science*, 338(6104), 267–270. doi: 10.1126/science.1223467.
- Wesolowski, A., Qureshi, T., Boni, M. F., Sundsøy, P. R., Johansson, M. A., Rasheed, S. B., Engø-Monsen, K., & Buckee, C. O. (2015). Impact of human mobility on the emergence of dengue epidemics in Pakistan. *Proceedings of the National Academy of Sciences*, 112(38), 11887–11892. doi: 10.1073/pnas.1504964112.
- World Tourism Organization (2022). Tourism Data Dashboard - International Tourism and Covid-19. Retrieved from <https://www.unwto.org/tourism-data/unwto-tourism-dashboard> (27.02.2022).
- Yang, A., Yang, J., Yang, D., Xu, R., He, Y., Aragon, A., & Qiu, H. (2021a). Human mobility to parks under the COVID-19 pandemic and wildfire seasons in the Western and Central United States. *GeoHealth*, 5(12), e2021GH000494. doi: 10.1029/2021GH000494.
- Yang, Y., Zhang, C. X., & Rickly, J. M. (2021b). A review of early COVID-19 research in tourism: Launching the annals of tourism research's curated collection on coronavirus and tourism. *Annals of Tourism Research*, 91, 103313. doi: 10.1016/j.annals.2021.103313.



Ministry of Education and Science  
Republic of Poland

---

The journal is co-financed in the years 2022–2024 by the Ministry of Education and Science of the Republic of Poland in the framework of the ministerial programme “Development of Scientific Journals” (RCN) on the basis of contract no. RCN/SN/0129/2021/1 concluded on 29 September 2022 and being in force until 28 September 2024.

## Annex

**Table 1.** Google and Apple mobility data description

Google Mobility (GM) Indices		Apple Mobility Indices	
Indice name	Short description, according to Google	Indice name	Short description, according to Apple
GM Retail& Recreation	Mobility trends for places like restaurants, cafes, shopping centers, theme parks, museums, libraries, and movie theaters.	Apple Driving	Aggregated was data collected from Apple Maps, by counting the number of requests made to Apple Maps for directions. The data sets are then compared to reflect a change in the volume of people driving, walking, or taking public transit around the world. Data availability in a particular city, country, or region is subject to many factors, including minimum thresholds for direction requests made per day.
GM Grocery& Pharmacy	Mobility trends for places like grocery markets, food warehouses, farmers markets, specialty food shops, drug stores, and pharmacies.		
GM Parks	Mobility trends for places like local parks, national parks, public beaches, marinas, dog parks, plazas, and public gardens.	Apple Transit	
GM Transit Stations	Mobility trends for places like public transport hubs such as subway, bus, and train stations.		
GM Workplace	Mobility trends for places of work.	Apple Walking	
GM Residential	Mobility trends for places of residence.		

Source: Nagy *et al.* (2022).

**Table 2.** Linear mean deviation for overnights

	Overnights (number)		Overnights-fixed base Febr. 2020	
	mean	st. dev.	mean	st. dev.
<b>Belgium</b>	984172.33	2623500.90	92.61	1.54
<b>Finland</b>	934741.73	2672931.50	75.23	-15.84
<b>Norway</b>	1478335.47	2129337.77	97.31	6.24
<b>Czechia</b>	1544172.87	2063500.37	77.92	-13.15
<b>Germany</b>	15594744.71	11987071.48	77.92	-13.15
<b>Estonia</b>	166871.13	3440802.10	99.15	8.08
<b>Spain</b>	5656858.93	2049185.70	68.53	-22.54
<b>Italy</b>	9166706.53	5559033.30	120.77	29.70
<b>Luxembourg</b>	22075.73	3585597.50	112.61	21.54
<b>Netherlands</b>	4482188.47	874515.24	129.53	38.46
<b>Slovakia</b>	452675.73	3154997.50	50.08	-40.99
<b>Mean EU 11</b>	<b>3607673.232</b>		<b>91.0629</b>	

**Table 3.** Linear mean deviation for Apple Mobility indices

	Apple Driving		Apple Transit		Apple Walking	
	mean	st. dev.	mean	st. dev.	mean	st. dev.
<b>Belgium</b>	85.77	-5.61	64.20	-0.33	68.22	-6.50
<b>Finland</b>	89.55	-1.83	67.56	3.03	95.28	20.56
<b>Norway</b>	68.22	-23.16	57.58	-6.95	100.71	25.99
<b>Czechia</b>	90.22	-1.16	4.19	-60.34	48.49	-26.23
<b>Germany</b>	105.19	13.81	82.31	17.78	87.93	13.21
<b>Estonia</b>	110.30	18.92	80.40	15.87	77.45	2.73
<b>Spain</b>	92.10	0.72	78.43	13.90	47.28	-27.44
<b>Italy</b>	82.07	-9.31	51.97	-12.56	58.27	-16.45
<b>Luxembourg</b>	84.91	-6.47	66.51	1.98	83.62	8.90
<b>Netherlands</b>	86.69	-4.69	42.14	-22.39	73.74	-0.98
<b>Slovakia</b>	110.15	18.77	69.50	4.97	80.93	6.21
<b>Mean EU 11</b>	<b>91.3776</b>		<b>64.5257</b>		<b>74.7199</b>	

**Table 4.** Linear mean deviation for Google Mobility indices

	GM_Retail		GM_Grocery		GM_Parks		GM_Transit		GM_Workplace		GM_Residential	
	mean	st. dev.	mean	st. dev.	mean	st. dev.	mean	st. dev.	mean	st. dev.	mean	st. dev.
Belgium	-35.02	-6.59	-6.28	-1.16	29.26	8.30	-35.30	4.65	-30.52	-3.35	12.01	3.14
Finland	-19.80	8.63	-1.99	3.03	78.76	41.20	-38.67	5.02	-25.47	1.70	65.65	56.78
Norway	-13.46	14.97	2.01	7.03	59.27	21.71	-34.25	-4.60	-26.77	0.40	6.59	-2.28
Czechia	-31.24	-2.81	1.35	6.37	23.49	14.07	-24.69	8.96	-21.93	5.24	7.46	-1.41
Germany	-29.99	-1.56	-3.21	1.81	48.18	10.62	-30.99	2.86	-22.98	4.19	8.15	-0.72
Estonia	-16.38	12.05	0.58	5.60	48.96	11.40	-21.01	12.64	-24.80	2.37	5.84	-3.03
Spain	-36.13	-7.70	-10.42	-5.40	-5.25	-42.81	-37.68	-4.03	-299.30	-272.13	10.14	1.27
Italy	-34.02	-5.59	-13.37	-8.35	6.66	-30.90	-38.94	5.29	-30.61	-3.44	10.64	1.77
Luxembourg	-35.47	-7.04	-11.28	-6.26	26.78	10.78	-29.87	3.78	-33.00	-5.83	12.33	3.46
Netherlands	-27.51	0.92	-4.42	0.60	60.04	22.48	-45.21	-1.56	-26.94	0.23	9.59	0.72
Slovakia	-33.71	-5.28	-8.23	-3.11	36.99	-0.58	-33.55	0.10	-25.94	1.23	8.24	-0.63
Mean EU11	-28.433		-5.0232		37.5583		-33.65		-27.1714		8.8679	

**Table 5.** Descriptive statistics

	Overnights - FB (Febr. 2020)	GM_ Retail& Rerecation	GM _Grocery & Pharmacy	GM _Parks	GM _Transit stations	GM _Workplace	GM _Residential	Apple Driving	Apple Transit	Apple Walking
N	Valid	154	154	154	154	154	154	143	143	143
	Missing	11	11	11	11	11	11	22	22	22
	Mean	91.06	-5.02	37.55	-33.65	-27.17	8.86	91.37	64.52	74.71
	Std. Deviation	92.15	10.52	55.98	14.77	9.86	5.63	35.04	23.86	29.83

**Table 6.** Heatmap for correlation coefficients of monthly data

Overall for all 11 European countries	Y	X1	X2	X3	X4	X5	X6	X7	X8	X9
Y	1	.823**	.478**	.657**	.670**	.408**	-.739**	.741**	.603**	.758**
X1		1	.635**	.742**	.755**	.565**	-.914**	.716**	.586**	.780**
X2			1	.502**	.600**	.552**	-.653**	.423**	.411**	.458**
X3				1	.549**	.284**	-.712**	.674**	.390**	.782**
X4					1	.614**	-.818**	.762**	.683**	.569**
X5						1	-.698**	.446**	.486**	.395**
X6							1	-.793**	-.679**	-.719**
X7								1	.747**	.630**
X8									1	.596**
X9										1

Notes: \*\*. The correlation is significant at the 0.01 level (2-tailed).

Y = overnight FB, X1 = GM retail & recreation, X2 = GM Grocery & Pharmacy, X3 = GM Parks, X4 = Gm Transit stations, X5 = GM Workplace, X6= GM Residential, X7 = Apple Driving, X8= Apple Transit, X9 = Apple Walking.



**Table 7.** Heatmap for correlation coefficients of daily mobility data

Correlation coefficients for Google and Apple daily data	Belgium	Czech Republic	Germany	Italy	Mean
X7 & X1	0.892	0.870	0.826	0.853	0.860
X7 & X2	0.381	0.396	0.335	0.490	0.401
X7 & X3	0.677	0.818	0.747	0.936	0.795
X7 & X4	0.756	0.802	0.794	0.850	0.801
X7 & X75	0.405	0.328	0.374	0.551	0.415
X7 & X6	-0.741	-0.747	-0.719	-0.866	-0.768
X8 & X1	0.793	0.901	0.770	0.860	0.831
X8 * X2	0.496	0.391	0.310	0.620	0.454
X8 & X3	0.374	0.585	0.501	0.736	0.549
X8 & X4	0.835	0.861	0.814	0.904	0.854
X8 & X5	0.522	0.443	0.445	0.692	0.526
X8 & X6	-0.754	-0.721	-0.707	-0.860	-0.761
X9 & X1	0.827	0.873	0.792	0.846	0.835
X9 & X2	0.356	0.382	0.274	0.505	0.379
X9 & X3	0.611	0.716	0.752	0.888	0.742
X9 & X4	0.815	0.832	0.794	0.868	0.827
X9 & X5	0.467	0.391	0.378	0.537	0.443
X9 & X6	-0.757	-0.730	-0.713	-0.836	-0.759
Mean	0.637	0.655	0.614	0.761	

Notes: X1 = GM retail & recreation, X2 = GM Grocery & Pharmacy, X3 = GM Parks, X4 = Gm Transit stations, X5 = GM Workplace, X6= GM Residential, X7 = Apple Driving, X8= Apple Transit, X9 = Apple Walking.

**Table 8.** The regression coefficients

Independent variables	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations	Collinearity Statistics
	B	Std. Error	Beta				
(Constant)	14.248	50.837		.280	.780		
X1	2.640	.600	.573	4.397	.000	.185	5.419
X2	-1.041	.958	-.116	1.086	.279	.274	3.644
X3	.061	.227	.038	.271	.787	.161	6.193
X4	.195	.717	.032	.272	.786	.228	4.386
X5	-2.241	.938	-.244	2.389	.018	.300	3.328
X7	.969	.321	.370	3.021	.003	.209	4.791
X8	-.341	.397	-.088	-.859	.392	.298	3.357
X9	.308	.397	.099	.776	.439	.192	5.218

Notes: Dependent Variable: Overnights as fixed base index (Febr. 2020).

X1 = GM retail & recreation, X2 = GM Grocery & Pharmacy, X3 = GM Parks, X4 = Gm Transit stations, X5 = GM Workplace, X6= GM Residential, X7 = Apple Driving, X8= Apple Transit, X9 = Apple Walking.

**Table 9.** Rotated component matrix

	Component		
	1	2	
	Google Mobility_Parks	.917	.122
	Apple Walking	.879	.266
<b>PC1 - Leisure and recreation activities</b>	Apple Driving	.809	.344
	Google Mobility Retail & Recreation	.720	.567
	Apple Transit	.577	.515
	Google Mobility_Workplace	.103	.941
<b>PC2 - Domestic and routine activities</b>	Google Mobility_Grocery & Pharmacy	.322	.805
	Google Mobility_Transit stations	.545	.724

Notes: Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 3 iterations.

**Table 10.** Regression coefficients

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	93.407	5.558		16.807	.000
3 PC1 (GM & A) – "Leisure and recreation activities"	67.283	5.419	.729	12.416	.000
PC2 (GM & A) – "Domestic and routine activities"	13.195	5.432	.143	2.429	.017

Note: Dependent Variable: Overnights as fixed base index (Febr. 2020).

**Table 11.** Coefficients

Independent variables	Unstandardized Coefficients		Standardized Coefficients	t	Six.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
(Constant)	130.132	21.188		6.142	.000		
X1	2.768	.562	.611	4.924	.000	.200	5.001
X2	-2.277	.873	-.262	-2.610	.010	.305	3.274
X3	.482	.153	.302	3.154	.002	.335	2.983
X4	1.066	.630	.176	1.692	.093	.285	3.506
X5	-1.591	.896	-.176	-1.775	.078	.314	3.183

Notes: Dependent Variable: Overnights as fixed base index (Febr. 2020).

Notes: X1 = GM retail & recreation, X2 = GM Grocery & Pharmacy, X3 = GM Parks, X4 = Gm Transit stations, X5 = GM Workplace.

**Table 12.** Rotated component matrix

	Component	
	PC1	PC2
Google Mobility_Workplace	.963	
Google Mobility_Grocery & Pharmacy	.776	
Google Mobility_Residential	-.755	
Google Mobility_Transit stations	.726	
Google Mobility_Parks		.963
Google Mobility_Retail & Recreation		.741

Notes: Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

Rotation converged in 3 iterations.

**Table 13.** Regression coefficient

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	90.443	5.304		17.050	.000
PC1 (GM)	14.150	5.178	.158	2.733	.007
PC2 (GM)	63.563	5.160	.711	12.317	.000

Note: Dependent Variable: Overnights as a fixed base index (Febr. 2020).

**Table 14.** Parameter estimates for Google Mobility data

Predictor	Predicted				
	Hidden Layer 1				Output Layer
	H(1:1)	H(1:2)	H(1:3)	H(1:4)	Overnights
(Bias)	-.445	.195	-.134	-.669	
GM_Retail_Recreation	-.058	-.550	-.602	.317	
GM_Grocery_Pharmacy	.187	.080	.140	-.337	
Input Layer	GM_Parks	.189	.327	.087	.724
	GM_Transit	-.578	-.422	.604	-.077
	GM_Workplace	.068	.386	.211	-.261
	GM_Residential	.595	-.103	-.219	-.629
	(Bias)				.288
Hidden Layer 1	H(1:1)				-.304
	H(1:2)				-.160
	H(1:3)				-.287
	H(1:4)				.991

**Table 15.** Parameter estimates for Google Mobility data

Predictor	Parameter Estimates			
	Hidden Layer 1		Predicted	
	H(1:1)	H(1:2)	Output Layer overnights_BF_febr20	
(Bias)	-.341	-.235		
Input Layer	Apple_Driving	-.068	.595	
	Apple_Transit	-.622	-.121	
	Apple_Walking	-.392	.515	
Hidden Layer 1	(Bias)			.228
	H(1:1)			.005
	H(1:2)			1.098

**Figure 1.** The research flowchart and methods.



Figure 2. The component plot in rotated space for Google and Apple mobility data

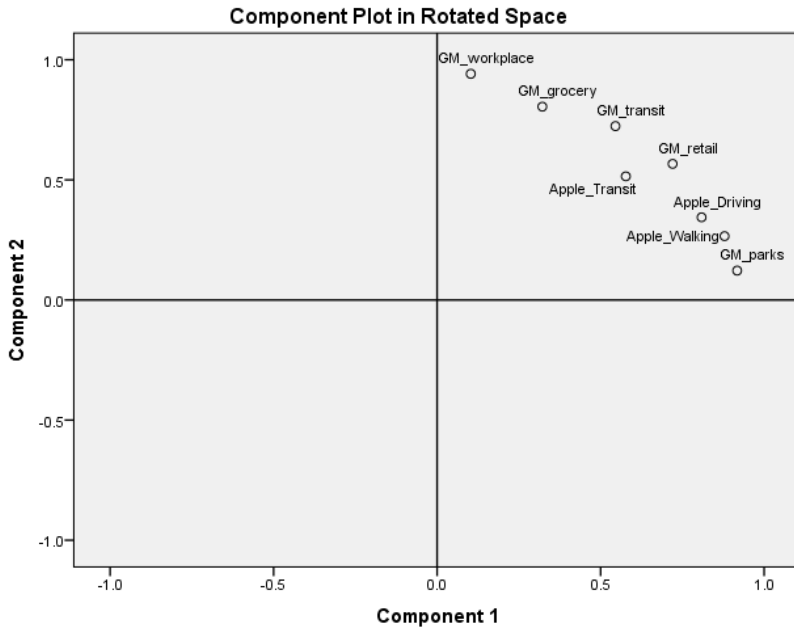


Figure 3. Histogram

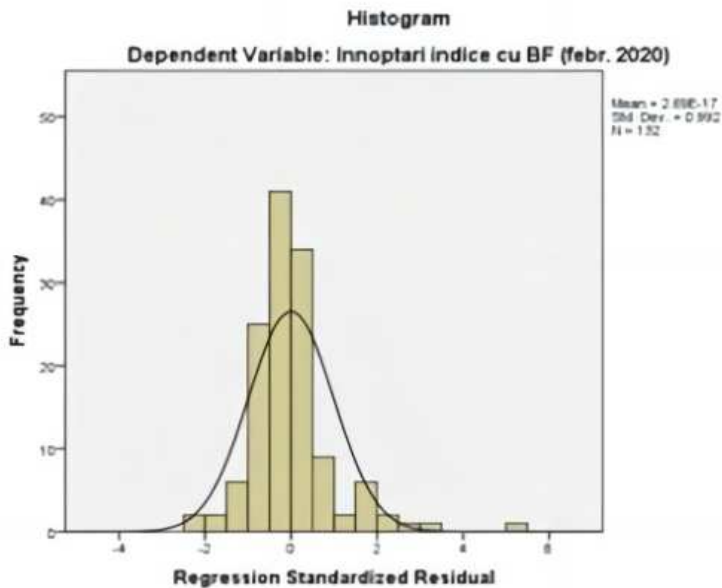


Figure 4. Normal P-P plot for regression

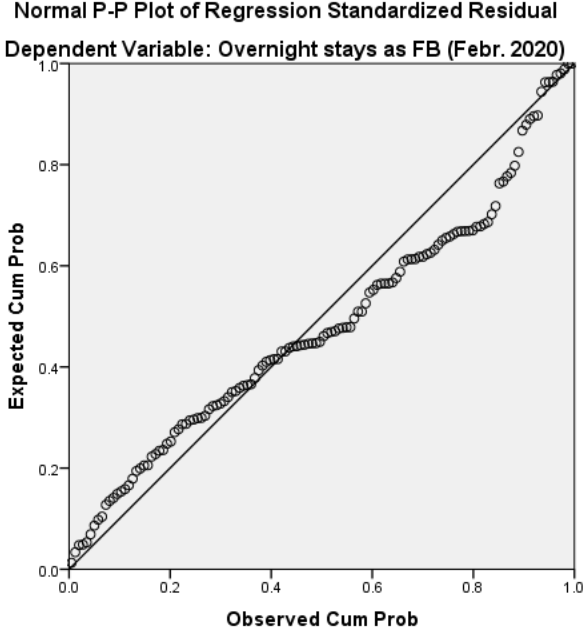


Figure 5. Component plot in rotated space

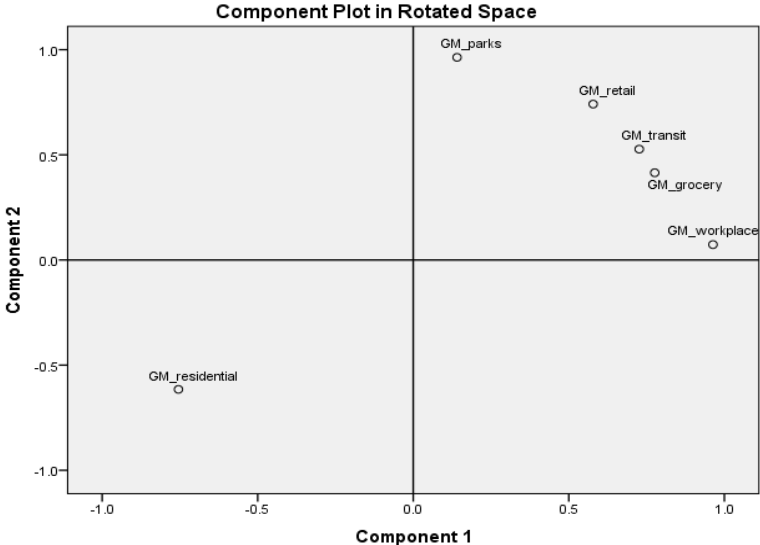


Figure 6. Network diagram for Google Mobility data

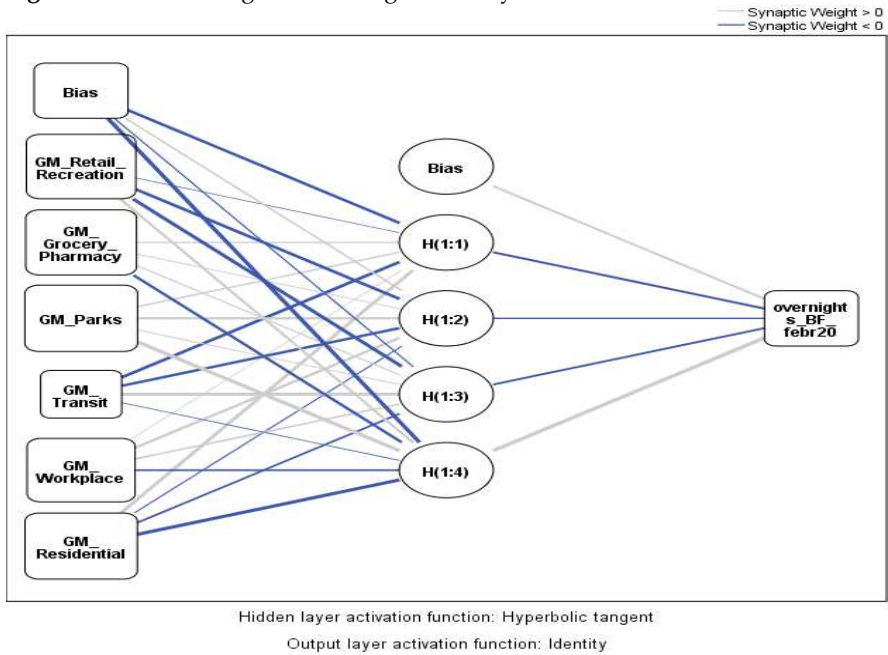
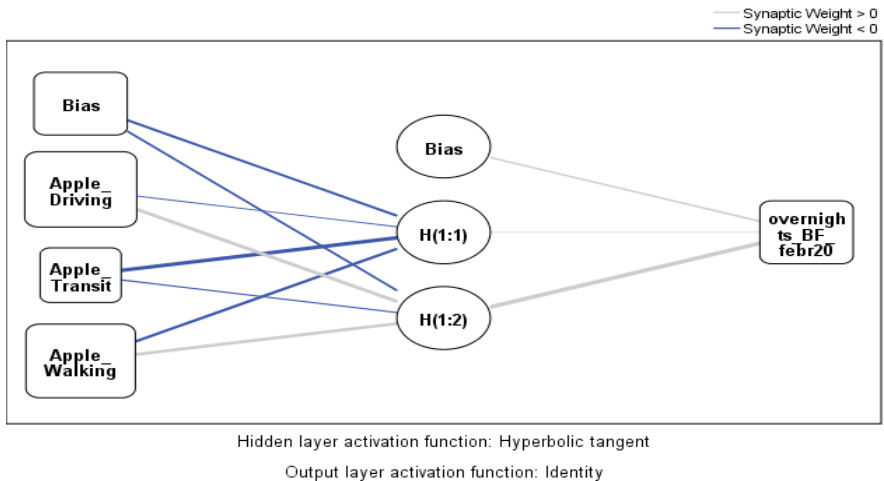
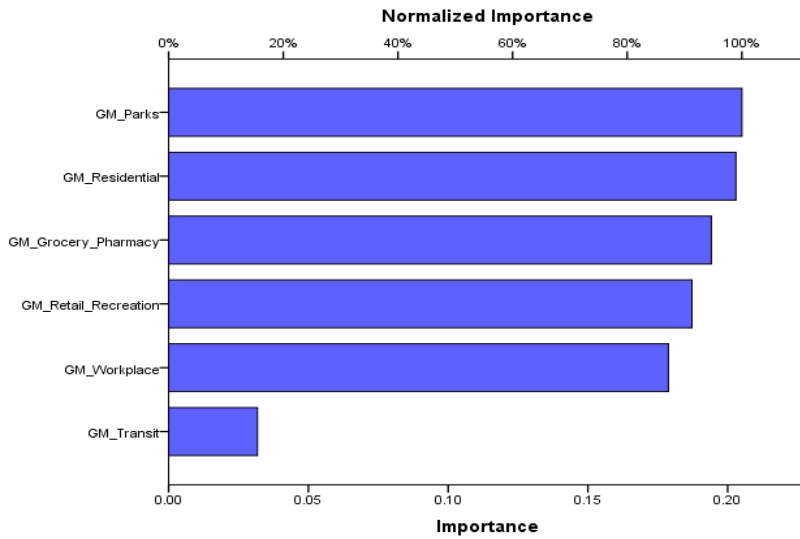


Figure 7. Network diagram for Apple Mobility data





**Figure 8.** The normalized importance of independent variables for Google Mobility data



**Figure 9.** The normalized importance of independent variables for Apple Mobility data

