Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
How did COVID-19 change what people buy: Evidence from a supermarket chain

Danas Zuokas*, Evren Gul, Alvin Lim

Retailer Products Research & Development, NielsenIQ, Chicago, IL, 60606, USA

ARTICLE INFO

Keywords:
Consumer behavior
COVID-19 pandemic
Panic buying
Retail
Sales forecast
Time series clustering

ABSTRACT

This research takes a retrospective view of the COVID-19 pandemic and attempts to accurately measure its impact on sales of different product categories in grocery retail. In total 150 product categories were analyzed using the data of a major supermarket chain in the Netherlands. We propose to measure the pandemic impact by excess sales—the difference of actual and expected sales. We show that the pandemic impact is twofold: (1) There was a large but brief growth at 30.6% in excess sales associated with panic buying across most product categories within a two-week period; and (2) People spending most of their time at home due to imposed restrictions resulted in an estimated 5.4% increase in total sales lasting as long as the restrictions were active. The pandemic impact on different product categories varies in magnitudes and timing. Using time series clustering, we identified eight clusters of categories with similar pandemic impacts. Using clustering results, we project that product categories used for cooking, baking or meal preparation in general will have elevated sales even after the pandemic.

1. Introduction

It is hard to overestimate the COVID-19 pandemic impact on the retail industry. The data provided by the statistical office of the European Union (Eurostat) shows that in the 27 European Union countries year-over-year decrease in retail trade (except those of motor vehicles and motorcycles) was 6.9% in March, 17.5% in April and 2.2% in May of 2020. This drop was mostly driven by non-food products (including fuel). While the retail sales of food, beverages and tobacco increased by 8.3% in March, 0.8% in April and 4.7% in May, respectively, as compared to the same periods in 2019. This increase stayed at 3.0% on average from March 2020 to February 2021. These were mainly caused by various governmental measures to control the pandemic and the change in people’s shopping behavior.

At the very start of the pandemic, with the rise of the number of daily new confirmed cases and deceased, governments responded by banning public gatherings, closing educational institutions, and canceling international flights. People reacted by rushing to the stores. Probably the most iconic image of such behavior is a shopping cart piled up with packs of toilet paper. Panic buying received a huge interest among researchers in various scientific disciplines such as consumer research, retail management, psychology, social sciences and humanities, among others (Billore and Anisimova, 2021). Using evidence from several countries, Islam et al. (2021) showed, for example, that perceived limited quantity and time scarcity incite people to buy in an impulsive and obsessive manner.

With strict measures (lockdowns) imposed, people immediately confronted a new reality—staying at home for most of their time. This brought many challenges and problems: working from home (Kaushik and Guleria, 2020), home schooling (Letzel et al., 2020), social isolation (Hwang et al., 2020), among others. With all these changes in everyday life, food consumption habits had to change too. People had to think more on how they get their next meal, relying on food delivery services (Tandon et al., 2021) or preparing their own meals (Bennett et al., 2021). To avoid crowds, consumers chose to order goods online (Baarsma and Groenewegen, 2021) or taking various precautionary measures as they visited brick-and-mortar stores (Wang et al., 2020).

All this led to the change in consumption of various product categories. With this research we are trying to estimate this change. First, time series models are fitted with sales data ending just before the start of the pandemic. These models are used to estimate expected future sales. We subtract these values from actual future sales to get the excess sales.
sales due to the impact of the pandemic. This generalizes the main methodological contribution of our paper – to provide a framework to measure the pandemic impact without pandemic data acting as exogenous information. This contrasts with the approach where various pandemic metrics (the number of new daily cases, dummy variables indicating lockdown period, etc.) act as exogenous variables in a regression model for sales. Studies using similar and alternative approaches are reviewed, and arguments supporting our research strategy are presented in Section 2. Section 3 provides the details on the time series models and presents the estimated impact of the pandemic for 150 product categories of the supermarket chain. To summarize the results of the pandemic impact on all product categories, we group them using data clustering and provide explanations on the clustering results. Clustering methods and distinct clusters themselves are presented in Section 4. Thus, we contribute to the existing literature by providing an extensive reference for the researchers who are interested in the consumer behavior during the COVID-19 pandemic. Finally, in Section 5, we review the benefits of our research for business practitioners. We also assess the potential limitations in our strategy and consider some future directions of research to extend the current work.

2. Literature review

A wide range of research articles are dedicated to measure pandemic impact on consumer behavior, household expenditure, and retail sales. No matter how diverse, they could be grouped with regards to data sets and methods used. One class of research, e.g., Laato et al. (2020), Omar et al. (2021), Eger et al. (2021), use respondent survey data and structural equation modeling (a well established framework in social sciences with rigorous procedures) to investigate how consumers’ latent (not measured directly), inner (psychological) factors influence their purchasing behavior. In contrast, our study does not look for the motives behind the changed purchasing behavior but detects the behavioral change itself using historical purchase data.

Another set of research analyzes consumer spending using purchase transaction data (basically payment by card, cash withdrawals and incoming money flows) across different expenditure categories, e.g., travel, groceries, restaurants, etc. Transaction data of 760 thousand Danish households were used in Andersen et al. (2020) to compare pre-pandemic year-over-year increase in spending to those after the onset of the pandemic to estimate the pandemic’s impact. Similar pandemic impact estimation is used in Carvalho et al. (2020) wherein 2.1 billion transactions from 1.6 million distinct points of sale in Spain were examined. Daily transaction data in 214 Chinese cities were examined in Chen et al. (2020) to estimate the impact of the pandemic by fitting a linear regression model with a dummy variable indicating the pandemic time period.

Two articles analyzed data that is similar to that of our research and had similar research objectives – to measure the pandemic impact on various grocery product categories. Valls Castelló and López Casasnovas (2021) investigated the sales data of a Catalan supermarket chain with 11% market share and reported sales changes in 12 food product categories. To capture the effect of newly confirmed COVID-19 cases on sales, they included it as a variable in the linear regression model. O’Connell et al. (2021) analyzed purchase data on 138 categories of fast-moving consumer goods from 17 thousand UK households and grouped the categories into four groups. As it is not stated in the paper, we suppose that the authors manually assigned each category to a group. In contrast, we use a data-driven approach to group product categories.

Product categories were assigned to three groups in Jung and Sung (2017). Monthly sales of electronic goods (home appliances, communication devices and computers), semi-luxury goods (clothing, shoes and accessories) and groceries (food, beverages and cosmetics) were analyzed for the impact of the Middle East Respiratory Syndrome (MERS – coronavirus species) outbreak in Korea. To capture the MERS effect, the authors used SARIMAX (seasonal ARIMA with external variables) model with a dummy variable indicating pandemic time period.

Previous studies show that to measure the pandemic impact on sales, researchers take two directions. The first is (auto)regression modeling with pandemic explanatory variables. Including a pandemic time period indicator allows the capture of change in average sales level. Such models might not be capable of inferring other (more dynamic) forms of pandemic impact. Another common model specification is to introduce the number of confirmed cases which captures the pandemic impact with a multiplier effect. This strategy is limiting in explaining sales dynamics of various product categories beyond a scale on the same data.

The second approach is by comparing pre- and post-pandemic situation; expected and actual. For example, 2019 and 2020 daily sales values are compared in O’Connell et al. (2021). In Andersen et al. (2020), excess spending from March 11 to May 3 of 2021 (actual) is compared to that of from January 1 to February 15 (normal). Excess spending is computed as year-over-year change in daily spend. A consistent approach is taken in Panzone et al. (2021) where 12 seasonal ARIMA models are fitted on sales data. Then, the difference of actual and expected sales values are computed with expected values coming from forecasts given by the best model. Our method is similar to this approach, but we improve on it by using three more time series modeling alternatives besides ARIMA and combining forecast results of the models into one final forecast.

Such an approach has the advantage over regression modeling in that a researcher does not have to explicitly specify how the pandemic affects sales. Consequently, the only concern left is accurate time series forecasting, and the methods for this task are well-developed and established. On the condition that forecasts are accurate, this allows us to obtain pandemic impact in its pure form. On the other hand, this does not allow the capture of the effects of different factors contributing to the total pandemic impact. Also, it can be applied to a relatively short time period because forecast accuracy decreases.

3. Sales forecast and pandemic impact

Forecasting is essential in conventional retail analytics. It finds its uses in product pricing, promotion planning, marketing, assortment modeling and other operational areas. Comprehensive and conceptual overview of the research literature on forecasting in retail can be found in Fildes et al. (2019). It shows that for a broad range of different tasks and time series data, various statistical or machine learning methods are often applied.

Before presenting methods that we applied in our study, let us first briefly describe the data that we used. We had access to four years of sales data of a supermarket chain in the Netherlands representing over 20% market share in 2020. 150 categories (most of them food and drinks but also home, body, baby, and pet care) that constitute 98.9% of 2020 total sales were analyzed. Time series were aggregated to four-week periods, and the logarithmic transformation was used – a common practice in econometric modeling.

3.1. Models

Following the outline of Hyndman and Athanasopoulos (2021), we note several major approaches for time series modeling:

- Time series decomposition;
- Time series regression models;
- Exponential smoothing; and
- ARIMA models.
They all are suitable for our analysis. First, they account for trend and seasonality, which are included in most sales time series. Second, they are easy to implement, do not require a heavy computational infrastructure and can select the best model in an autonomous manner if needed. This is important because we are dealing with 150 sales time series and do not have the resources to fine-tune each model separately.

Next, we describe how we apply each of the listed classical methods in our case.

**Time series decomposition.** Separately for each category, using additive decomposition model

\[ y_t = T_t + S_t + R_t, \]

the value of the logarithm of sales, \( y_t \), is split into a trend \( T_t \), a seasonality \( S_t \) and the remainder \( R_t \) components. The seasonality component is estimated using moving averages. Then, linear median regression model is fitted on the seasonally adjusted component, \( y_t - S_t \), using linear trend and the logarithm of the total selling area as explanatory variables. The final forecast value is obtained by adding the values of the seasonality component to the forecast value of the seasonally adjusted component.

**Time series regression models.** A linear model for the log of sales is fitted with a linear trend, seasonality dummy variables and a logarithm of the total selling area as explanatory variables. A median regression model is fitted which is more robust against outliers. It is the same as linear regression, but is estimated by least absolute deviations instead of ordinary least squares.

**Exponential smoothing.** We use additive Holt-Winters’ (aHW) method for the logarithm of sales. Using the taxonomy of exponential smoothing models, it is an \( (A, A) \) model (additive trend and seasonality components). The model consists of a forecast equation for observations \( y_t \) and smoothing equations for level \( \ell_t \), trend \( b_t \) and seasonality \( s_t \) components. An \( h \)-step forecast \( \hat{y}_{t+h|t} \) at time period \( t \) is computed via the following recursive relations:

\[
\begin{align*}
\hat{y}_{t+h|t} &= \ell_t + h b_t + s_{t+1-k} & (2) \\
\ell_t &= \alpha(y_t - s_{t-k}) + (1 - \alpha)(\ell_{t-1} + b_{t-1}) & (3) \\
b_t &= \beta(\ell_t - \ell_{t-1}) + (1 - \beta)b_{t-1} & (4) \\
s_t &= \gamma(s_t - s_{t-(h-1)}) + (1 - \gamma)s_{t-1} & (5)
\end{align*}
\]

where \( m \) denotes the number of seasons, \( k \) is the integer part of \( (h - 1)/m \) and \( \alpha, \beta \) and \( \gamma \) are smoothing parameters.

**ARIMA models.** To be exact, SARIMAX models are used where \( S \) stands for seasonality and \( X \) for external regressors. Taking different specifications, a whole set of models of the same family are fitted, and the best model is selected based on the value of AICc Information criterion (AIC with small-sample correction). Model specification varies by taking different combinations of major parameters: the (seasonal) order of auto-regressive and moving-average parts and the (seasonal) integration order.

To measure model out-of-sample forecast accuracy, we fit a model using data that ends one year prior to the onset of the pandemic. Then, one-year forecast values are subtracted from the actual values. Any aggregate characteristic of these differences (mean absolute error, root-mean-square error) is appropriate to rank the models and choose the best.

It is now a common practice, especially in machine learning, to use ensemble learning techniques (see Polikar (2009) for a short introduction). Very often ensemble models perform better than separate models because the forecast errors of these models cancel each other out. However, at times, these techniques are computationally more expensive than the individual models themselves.

We have tried one of the well-known and simple to implement ensemble learning methods, first, a variant of bagging (bootstrap aggregating) adapted for time series data – moving block bootstrap, introduced in Künsch (1989). It is implemented through the following steps: the time series model is fitted, model residuals are extracted, moving block bootstrap is used to get replicated residuals, replicated time series are obtained, time series models are fitted to each of these replicated data and forecasts are calculated. The average of these forecasts is taken as a final forecast value. We used 100 bootstrap replications for each model with block size 13 covering a time period of one year. While moving block bootstrap alters chronological order of time series (later block might come earlier), it allows to retain time series characteristics (autocorrelation, heteroskedasticity) locally inside a block. Secondly, to make it as simple as it can be, we took the average of all forecast values produced by different models, without any bagging.

### 3.2. Results

We begin this section with the results on forecasting accuracy. For each of the 150 product categories four time series models were fitted using 2 years of data ending in four-week aggregation period 3 of 2019. For each category model, the mean absolute percentage error (MAPE) of one-year forecasts is computed. Then, for every model the median and the inter-quartile range of all 150 MAPEs as well as the percentage of the number of categories with MAPE less than 5% and 10% are computed. The results are presented in Table 1. AVG4B denotes the average of forecast values with bagging, AVG3 denotes ensemble model which takes the average of three (excluding aHW) and AVG4 – all four forecasts, without bagging.

Inspecting the performance of individuals models, it might be tempting to exclude aHW model from averaging. Nevertheless, AVG4 proved to be more accurate. This means there are product categories for which aHW is more accurate than the rest of the models (this is a good example of model diversification). Also, bagging does not bring significant improvement to the accuracy, therefore, we chose the average of all four model forecasts as a final forecast value, which also does not require additional computational resources as bagging does.

Next, forecasting results for several product categories are presented in Fig. 1. We chose categories with different MAPE values: 1.5% for cat food, 3.1% – bananas, 4.9% – baby diapers and 6.6% for cereal. Models are fitted using data that ends in period 2 of 2020 which is denoted by gray vertical line in the figure. Fitted values (in-sample forecasts) are shown to the left of the line, and out-of-sample forecasts are to the right. Forecasts are produced one year ahead up to period 2 of 2021.

We see a very good accordance of model forecasts for cat food and cereal. While for baby diapers the spread of forecasts is big. The aHW shows to the left of the line, and out-of-sample forecasts are to the right. Forecasts are produced one year ahead up to period 2 of 2021.

We see a very good accordance of model forecasts for cat food and cereal. While for baby diapers the spread of forecasts is big. The aHW model forecasts are somewhat at odds with the other models for baby diapers and bananas.

We end this section by applying the same procedure to total sales. The result is depicted in Fig. 2. The graph on the left shows very close movement of actual and in-sample forecast values (in-sample MAPE value is 1%). The graph on the right shows in-sample forecast percentage errors to the left of the gray vertical line and percentage pandemic impact values to the right. Percentage pandemic impact is calculated as the difference of actual and forecast values divided by forecast value. We divide by forecast value because it represents “normal” sales, and actual value represent “normal” plus excess sales caused by pandemic.

---

2 See https://otexts.com/fpp2/holt-winters.html for more details.
Table 1
Forecasting accuracy measures.

|               | Decomp. | Regress. | aHW | ARIMA | AVG4B | AVG3 | AVG4 |
|---------------|---------|----------|-----|-------|-------|------|------|
| Median of MAPEs | 6.1     | 6.6      | 8.2 | 6.4   | 5.5   | 6.1  | 5.6  |
| IQR of MAPEs   | 4.9     | 5.6      | 5.3 | 5.1   | 5.6   | 5.0  | 5.2  |
| MAPE < 5%      | 35      | 31       | 21  | 34    | 44    | 37   | 45   |
| MAPE < 10%     | 81      | 75       | 66  | 77    | 80    | 77   | 83   |

Fig. 1. Forecasts of example product categories.

Fig. 2. Pandemic impact on total sales.
Therefore, pandemic impact is measured as a percentage increase of “normal” sales. The first aggregation period of the pandemic covers two weeks before and 2 weeks at the start of the pandemic. Therefore, for this period the value of two-week excess sales is multiplied by two to compare it to four weeks of “normal” sales. Our choice of such aggregation was deliberate as we wanted to separate panic buying effects.

**Remark 1.** We estimate the panic buying effect at 30.6% of excess sales for the two-week panic buying period from March 11 to March 24 of 2020.

The above result is in line with the findings of O’Connell et al. (2021), where yearly increase of more than 25% in quantity bought for staples and household supplies for a period from February 24 to March 22, 2020 was reported. It is also in line with the simplest estimate of a panic buying effect – percentage increase in sales from one week prior to the pandemic onset to the first week of the pandemic at 34.7%.

After the panic buying period, excess sales values vary around 5%, except for a period between the first and the second wave of the pandemic when strict pandemic control measures were loosened. During this period, the average excess sales value is 1.5%.

**Remark 2.** We estimate pandemic impact on total sales at 5.4% of excess sales over the entire pandemic period, excluding the panic buying period.

We have already seen in Fig. 1 pandemic impacts of different magnitudes and various dynamics. It is rational to expect, however, that there are categories with similar impact. It might be that sales of some categories were more alike even before the start of the pandemic and also have similar impact. Conversely, it could be that some categories reacted in a similar fashion to the pandemic while prior to it had quite different sales dynamics. In this situation, a data-driven approach might work best to discover pandemic impact patterns, and we employ clustering techniques to achieve this.

### 4. Clustering product categories by pandemic impact

Clustering helps to summarize pandemic impact results of a large number of product categories. Besides that, it can reveal interesting and valuable insights about consumer behavior during the pandemic. This information might be important for the grocery retailers when deciding on pricing, planning promotions or managing assortment. For example, knowing which categories are most affected by panic buying, retailers can put a limit on allowable purchase quantity or take other actions to balance the demand-supply interplay. Or, they might consider increasing the assortment of product categories that are used in preparing meals. Most likely, people shifted their preference on the features and variety of the products they consume everyday.

#### 4.1. Methods

There are two major clustering approaches: centroid-based and hierarchical clustering (James et al., 2021). In our case, using centroid-based methods has one principal disadvantage: a small number of points (150 product categories) in a high-dimensional space (13 observations over time) leads to the curse of dimensionality problem. This means that with more dimensions, the points in a vector space become sparser, and the centroid-based clustering becomes ineffective. To solve this problem one could either increase the number of points or decrease the number of dimensions. Neither option is suitable in our case, so we opt for hierarchical clustering.

Hierarchical clustering for time series is a well-researched and widely applied technique in data analytics (Aghabozorgi et al., 2015). The main problem for a researcher is to choose a distance measure for two time series, and the number of options is huge (Montero and Villar, 2014). In our case, however, the list of potential distance measures can be reduced. We do not use measures that are based on dynamic time warping algorithms because a four-week aggregation period is long, and we do not expect strong unaligned correlation effects. We also exclude measures that use various transformations (autocorrelations, periodograms) of raw values because the time series are short, and characteristics may not be reliably estimated. The timing in our case is very important, so classical measures (Euclidean, Manhattan) which ignore the chronological order of the values were not considered. With all this in mind, we decided on using dissimilarity index introduced in Chouakria and Nagabhushan (2007) which accounts for both similarity with respect to raw values (magnitude) and the dynamics of the time series (shape). More formally it is defined as follows:

$$D(S_1, S_2) = f[\text{cort}(S_1, S_2)] \times \delta_{\text{min}}(S_1, S_2),$$

where \(\delta_{\text{min}}(S_1, S_2)\) is a conventional distance measure between the raw values of two time series \(S_1\) and \(S_2\) and \(\text{cort}(S_1, S_2)\) is a correlation coefficient of time series changes (temporal correlation):

$$\text{cort}(S_1, S_2) = \frac{\sum(u_{t+1} - u_t)(v_{t+1} - v_t)}{\sqrt{\sum (u_{t+1} - u_t)^2} \sqrt{\sum (v_{t+1} - v_t)^2}}.$$  \(\text{(6)}\)

Function \(f(x) = 2/(1 + \exp(\kappa x))\) modulates (via parameter \(k \geq 0\)) the contribution of temporal correlation to total value of the dissimilarity index. We chose \(k = 1\) which gives more uniform contribution over the whole range of \(\text{cort}(S_1, S_2)\) values. Euclidean distance was chosen as a conventional distance measure.

The final step in hierarchical clustering is to select an algorithm by which time series are joined into clusters, and we chose the agglomerate approach. Initially each element is a cluster of its own. Then, repeatedly, two most similar clusters are merged until all objects end up in one cluster. This process is presented by a dendrogram. There are several methods on how two clusters are merged, and we chose Ward’s method (Murtagh and Legendre, 2014). Using this method, at each agglomeration step, two clusters are selected to merge so that after merging the total within-cluster variance increases by the smallest amount.

Next, the results of clustering 150 product categories using percentage pandemic impact are presented.

#### 4.2. Results

We start by showing the agglomeration process of all 150 product categories in Fig. 3. The earlier two clusters merge (the more similar they are) the closer to the left is the link that joins them together. Clusters are obtained by cutting the dendrogram vertically at a certain level and collecting the elements of each cut branch.

It is rare that one cut gives the final set of clusters unless they are well separated and there are no outliers. Usually, we see some elements forming tight clusters early and others later. Also, some elements stay on their own for a while – these are outliers (can be spotted at the bottom of Fig. 3). It is noteworthy to discuss several of them (forecasting results and pandemic impact are shown in Fig. 4).

**Hand soap.** This product category defines the pandemic and is the biggest outlier. During the two weeks of panic buying, it sold around 9.2 times usual. This excess buying gradually decreased and stayed at about 83.3% above usual level afterward. Even after the pandemic, the sales of this category might stay at a higher level due to a new habit of maintaining hand hygiene.

**Candles.** From the start of the pandemic during the rest of 2020, candles sold on average 40.9% in excess. Business professionals in this market suggest that scented candles (and home fragrance in general) serve as a portal to the dreamland and adds coziness to a place.\(^3\)

**Gambling.** The sales of this product category show very interesting

\(^3\) [https://www.businessofbusiness.com/articles/candles-fragrance-home-scent-trend-diptyque](https://www.businessofbusiness.com/articles/candles-fragrance-home-scent-trend-diptyque).
dynamics. During the first few months of the pandemic, it fell by about 25%. This is the result of less frequent visits to stores – people have a habit of grabbing a lottery ticket at the check-out counter. Another observation is an increase in lottery ticket sales at the end of every year. Comparing lottery ticket sales during the first and second half of December 2019, we observe 118% increase, while this number for 2020 is 282%. This shows just how much people hoped for a better future.

Besides outliers, product categories that are highly seasonal (barbecue), sold mostly for holidays (other breads), form very small clusters (flowers and plants) or overall are hard to model (shampoo), were also not assigned to any cluster. All these are colored in light gray in Fig. 3.

For the rest of the categories, we obtain eight clusters in total. Calinski-Harabasz index (the ratio of the between cluster sum of squares to the within cluster sum of squares) is used to validate the number of clusters. Index values for different number of clusters are depicted in Fig. 5. Although the optimal number of clusters suggested by Calinski-Harabasz index is 2, choosing such a small number of clusters has very limited practical use. The increase in index value at the number of clusters 6–8, validates our choice.

Finally, for each cluster, we give a possible interpretation and show impact graphs of several representatives (see Figs. 6–13). Looking at some of these graphs, it might seem that product categories are not that...
similar. But, let us remember that the chosen distance measure receives contribution not only from the distance with respect to raw values (magnitude) but also from the distance measured by correlation of time series changes (dynamics). The former is recognized more easily and supposedly is perceived as a “true” distance by our brains. The latter requires more effort and can be perceived as a similarity in the shape.

Cluster 1: Survival kit. Product categories in this cluster have the biggest panic buying effect exceeding 100%. The cluster includes toilet paper, preserved food (vegetables, meats, fish), pasta, rice, various meal mixes and soups – basic, dry-packaged foods that do not take up much space (except for toilet paper). After the first two weeks excess sales dropped and stayed at 5–10% throughout the year showing the magnitude of a stay-at-home effect.

Cluster 2: Stay-at-home (stock). This cluster reflects increased activity (baking, tidying) at home. Stay-at-home effect is bigger than that of Cluster 1 and varies around 15–20%. Cluster members are coffee, flour, oils, sugar, frozen fish and vegetables, cleaners and bathroom supplies. These are dry-packaged food ingredients and home/body care products that need more space or special equipment (freezers) to keep.
They also have very big panic buying effect reaching 60–80%.

**Cluster 3: Restaurant at home.** With restaurants closed people seek to replicate the same experience of having a fancy dinner at home. Understanding this need some restaurants and supermarkets even introduced DIY (Do-It-Yourself) meal kits contributing to the boost of insperience (in-home experience) economy. Beef, raw cooking vegetables (asparagus, broccoli), herbs and spices, craft beers and red wine, ice cream form a small but distinct cluster. No panic buying effect is observed, and excess sales vary at around 20% with a drop in between the two pandemic waves.

**Cluster 4: A drink or two.** This is another cluster with no panic buying effect and another example of experiences transformed into insperiences. It consists of alcoholic drinks (beer, white wine, spirits),
Cluster 5: Breakfast and lunch. Sandwich spreads, pastes and toppings, cereal, butter and margarine, bread substitutes, minced meats, poultry, frozen potatoes, chocolate are in this cluster. Except for a minced meat and poultry, these categories can be easily stored. Thus, panic buying effect is large and is around 50%. The cluster is quite like Cluster 2, but categories in this cluster are consumed daily and are purchased more frequently.

Cluster 6: In excess. This is the smallest cluster with five product categories in it: laundry detergent, baby food and diapers, oral and intimate hygiene. It has a very big panic buying effect of 60–70% and is distinct by a plunge to negative impact values right after the end of a panic buying period. This might have the following explanation: categories in this cluster have long consumption cycle due to big packaging (baby diapers) or small one-time dose (toothpaste) and/or limited time period of consumption (baby formula). While we see a slow subsequent recovery, the impact stays below zero throughout the year. Similar dynamics was also observed for the cat and dog food which were assigned to Cluster 8, however.

Cluster 7: School’s out. This is another cluster with negative pandemic impact except for two panic buying weeks. Looking at the list of cluster members – iced tea, snack bars, candy, convenience (pizza, pancakes), frozen meals, diet (lactose-free, slimming, sport food), cosmetics (shaving, hair), juices – suggests that consumers of these products could be schoolchildren and students. With the closure of schools, schoolchildren had less opportunities to consume soft drinks and snacks. While the closure of universities (and campuses) left some students no
choice but to return to their parents’ house where they had healthier food. In April 2020, for example, the sales of the snack bars was 30%, and those of the convenience category – 12% below the expected.

**Cluster 8: Solitary.** Product categories in this cluster are used for meals that do not require much cooking creativity and effort. Pre-packaged cheese and meats for a sandwich, prepackaged sausages with cold sauces (ketchup), unprocessed vegetables (tomato, cucumber), dairy food and drink, nuts. These are perishable products and have small panic buying effect. Stay-at-home effect varies around 5%. Pandemic impact is closest to that of Cluster 5 (“Breakfast and lunch”) as seen in the dendrogram (Fig. 3). While categories in Cluster 5 could represent family with children, this cluster might be more about singles, couples without kids or empty-nesters. Another interpretation could be that categories in these clusters are used for meals that perform different functions – more elaborate meal when one has more time to prepare (Cluster 5) versus basic dish with no time to spend in the kitchen (Cluster 8).

Pandemic impact for all 150 product categories is presented in Table 2 in Appendix.

### 4.3. Clustering summary

To summarize clustering results, Fig. 14 depicts all product categories as points on a xy-scatter plot, where on x-axis we have the panic buying effect and on y-axis the median of pandemic impact (excluding panic buying period). For each cluster, a bivariate t-distribution is fitted, and an ellipse region is added which covers 80% of a probability mass of that distribution. We can use this plot as a visual tool for cluster validation, and it is yet another evidence that our clusters are well separated.

Cluster 1 immediately draws attention with the biggest panic buying effect. Yet, afterward, sales stayed moderately elevated at around 5–10%, similarly to Cluster 5 and 8. Product categories in these three clusters are mostly used in everyday meal preparation and make up around 50% of total sales. Cluster 4 with total sales proportion at around 14% also has similar median pandemic impact, however panic buying effect is closer to zero.

Clusters 2 and 3 have very similar median pandemic impact at around 15–20%. However, lower panic buying effect of Cluster 3 lets it distinguish itself as a separate group of product categories. It is interesting to note that, although small, there is a persistent and significant increase of sales proportion for Cluster 3 from 5.1% in one year prior pandemic to 5.7% of total sales in one pandemic year. This shows people acquiring gusto, starting to cook for themselves and creating the restaurant atmosphere at home.

Cluster 6 and 7 differentiate themselves with negative median pandemic impact, and differ one from the other a lot in the magnitude of panic buying effect.

To see pandemic impact time dynamics, for each cluster, at each time point (excluding panic buying period), Fig. 15 depicts the median value of all pandemic impact values in that cluster at that time point. For clusters 2 and 3 we see a large decline in values towards the summer season, when the tight pandemic control measures were relaxed, and a sharp increase towards the cold season, when the pandemic situation worsened, and people stayed more inside. Similar dynamics is observed for Cluster 8 but in a smaller amplitude. Neighbouring Cluster 5 has similar shape too. Pandemic impact for Cluster 1 stays almost constant throughout a year except for the rise in October 2020. This indicates a small panic buying effect before a partial lockdown was introduced due to the onset of the second wave of the pandemic.

Product categories in all these five clusters are mostly used for cooking, baking and preparing all kinds of meals. With the lifestyles changing (working more from home) and the pandemic strengthening the trend of home cooking and baking, we expect elevated sales of these product categories from now on. Our projection is supported by the fact that the pandemic impact stays well above zero even during the pause between the two pandemic waves for these clusters.

For the remaining clusters (Clusters 4, 6 and 7) we can not find a support for sales increase after the end of pandemic. For Cluster 4 (“A
drink or two”) no panic buying effect is observed, and increased pandemic impact is observed only during the summer season. This can be explained by people choosing to spend their vacation at home country. As for Cluster 6 (“In excess”), we see pandemic impact approaching zero from the negative side as people slowly clear off their stocks. With schools and universities fully open, and students returning to campuses, we expect sales for product categories in Cluster 7 (“School’s out”) also return to normal levels.

Finally, let us discuss how the methods and results in this research could be applied and how it could be extended.

5. Discussion

Our research has several potential implications for other researchers and business practitioners. The strategy that we chose to measure pandemic impact could be applied in any situation where it is important to estimate the discrepancy between normal (expected) and unusual (actual) observations in time. Such strategy may contribute to decision support systems of various supermarket operations from stocking, inventory planning, promotion effectiveness evaluation to marketing, etc. However, application is limited to short time horizon because forecasting farther into the future loses accuracy. To extend the analysis even further, a researcher will have to choose an alternative strategy – regression modeling incorporating pandemic information. Forecasting sales also requires modeling and forecasting pandemic variables

(Simchi-Levi et al., 2020).

We have shown how different product categories can be grouped using pandemic impact data. Using product level sales data and the same methodology, researchers could identify small groups of complementary products. Similar goals are set when performing market basket analysis, and this approach could be used to support it. Clustering has an advantage in that it requires significantly smaller amount of computational resources. Also, when comparing two products, it can take sales dynamics into account, contrary to the static nature of the classical algorithms. Using a similar approach, Li et al. (2021) proposed temporal correlation of sales time series to measure product similarity.

Using thorough methodology, we have shown the magnitude and the shape of the pandemic impact for different product categories. Employing this information, business practitioners could take timely actions in anticipation of a demand change in similar future events. To benefit more people they could impose a limit on the allowable purchase quantity or adjust prices in certain categories. Finally, knowing product categories which might have elevated sales in the future can help retailers in other merchandising and marketing functions such as deciding the changes to product assortment and planning for promotion, for example.

Declaration of competing interest

None.

Appendix

Table 2

Pandemic impact of all product categories.
The columns “Panic”, “Post-panic”, and “Average” represent the effect of the pandemic as percent change in sales for the two-week panic buying period, the four-week immediately following the two weeks of panic buying, and the one-year of pandemic excluding the first six weeks, respectively.

| Cluster          | Product category         | Panic  | Post-panic | Average |
|------------------|--------------------------|--------|------------|---------|
| 1: Survival kit  | Meal mixes               | 105.9  | 4.9        | 2.8     |
|                  | Over-the-counter medicine| 153.5  | –11.0      | 7.5     |
|                  | Pasta                    | 197.1  | 7.9        | 14.5    |
|                  | Pasta sauces             | 154.3  | 8.8        | 11.2    |
|                  | Preserved fish           | 135.2  | 14.7       | 7.5     |
|                  | Preserved meats          | 130.5  | 10.5       | 3.4     |
|                  | Preserved vegetables     | 159.7  | 10.4       | 7.8     |
|                  | Rice, noodles, oriental  | 159.0  | 10.2       | 13.0    |
|                  | Soups                    | 103.8  | 1.0        | 1.6     |
|                  | Stationery               | 176.2  | –3.3       | 5.5     |
| 2: Stay-at-home  | Bags and foil            | 82.8   | 18.9       | 16.1    |
|                  | Bath and shower          | 79.2   | 23.8       | 21.6    |
|                  | Bathroom supplies        | 81.2   | 37.0       | 23.6    |
|                  | Cleaners                 | 135.3  | 31.0       | 18.3    |
|                  | Coffee                   | 65.0   | 11.2       | 12.5    |
|                  | Flour and baking products| 123.5  | 56.1       | 23.8    |
|                  | Frozen fish              | 71.6   | 21.6       | 19.6    |
|                  | Frozen vegetables        | 104.2  | 19.0       | 19.2    |
|                  | International assortment | 71.3   | 42.2       | 33.9    |
|                  | Nuts and southern fruits | 50.8   | 17.2       | 30.2    |
|                  | Oils                     | 69.0   | 16.2       | 14.9    |
|                  | Pork                     | 74.0   | 23.9       | 21.0    |
|                  | Preserved fruits         | 73.7   | 18.0       | 12.0    |
|                  | Regional products (meat) | 86.6   | 10.0       | 13.9    |
|                  | Shelf-stable dairy       | 61.5   | 7.4        | 11.3    |
|                  | Sugar                    | 85.3   | 23.1       | 19.3    |
|                  | Sweet treat (bread pastries)| 41.9  | 0.2        | 21.9    |
|                  | Tea                      | 69.9   | 6.5        | 17.3    |
|                  | Unbaked bread (prepackaged)| 116.1 | 27.6       | 26.7    |
|                  | White fats               | 318.4  | 31.1       | 16.3    |
| 3: Restaurant at home | Air fresheners        | 23.7   | 7.7        | 11.0    |
|                  | Beef                     | 31.5   | 12.8       | 14.1    |
|                  | Craft beers              | 19.3   | 24.1       | 22.4    |
|                  | Herbs and spices         | 43.1   | 29.8       | 18.1    |
|                  | Ice cream (buckets)      | 13.3   | 23.4       | 21.3    |
|                  | Mediterranean            | 28.4   | 7.2        | 18.7    |
|                  | Raw cooking vegetables   | 29.8   | 23.8       | 18.0    |

(continued on next page)
| Cluster | Product category | Panic | Post-panic | Average |
|---------|------------------|-------|------------|---------|
| Red wine | 12.7             | 10.8  | 16.0       |
| Rose wine | 29.6             | 28.0  | 26.5       |
| 4: A drink or two | Alcohol | 5.3 | 7.1 | 10.5 |
| | BO Sweet treat | 8.2 | -14.4 | 4.1 |
| | Beer | -5.0 | -5.1 | 4.9 |
| | Cheese (foreign, serviced) | 0.0 | -10.8 | 7.6 |
| | Coasters | 17.1 | -5.1 | 2.6 |
| | Cola | 12.4 | -3.8 | -1.3 |
| | Digestive drinks | 0.7 | 4.2 | 15.5 |
| | Dry sausage service | 8.6 | -5.5 | 9.8 |
| | Fish | -13.7 | -6.9 | -3.5 |
| | Frozen cakes and pastry | 17.8 | -1.0 | 6.0 |
| | Functional drinks | 3.8 | -5.2 | 1.9 |
| | Pastries | -24.5 | -28.8 | -4.6 |
| | Premixed drinks | -23.3 | 3.1 | 16.1 |
| | Salads (prepackaged) | -1.6 | -0.7 | -1.6 |
| | Soda | 12.2 | -4.2 | 0.0 |
| | Soft fruit (exotic) | 2.3 | 15.2 | 2.9 |
| | Taps | -19.3 | -11.2 | 2.6 |
| | White wine | 6.3 | 8.4 | 9.1 |
| 5: Breakfast and lunch | Bread substitutes | 59.6 | -0.8 | 8.6 |
| | Butter and margarine | 57.1 | 7.8 | 3.8 |
| | Cereal | 46.5 | 7.3 | 2.0 |
| | Chocolate | 58.8 | 11.3 | 6.7 |
| | Citrus fruits | 69.9 | 23.8 | 6.0 |
| | Deodorant | 50.7 | 9.0 | 0.1 |
| | Frozen potatoes | 62.5 | 14.7 | 2.0 |
| | Frozen snacks | 52.0 | 14.8 | 7.8 |
| | Meal packets | 49.3 | 10.6 | 4.9 |
| | Meat substitutes | 43.9 | -6.1 | -1.0 |
| | Minced meats | 48.4 | 11.7 | 2.8 |
| | Organic | 34.7 | -2.7 | 0.6 |
| | Oriental foods | 62.5 | 15.8 | 8.0 |
| | Potatoes and onions | 50.9 | 9.5 | -3.6 |
| | Poultry | 42.7 | 7.5 | 5.6 |
| | Sandwich spreads, pastes and toppings | 44.4 | -1.1 | 0.9 |
| | Surinamese and Antillean products | 44.0 | 11.2 | 10.7 |
| | Warm sauces | 51.1 | 15.5 | 8.6 |
| 6: In excess | Baby diapers | 84.6 | -15.0 | -5.9 |
| | Baby food | 52.0 | -22.6 | -5.5 |
| | Intimate hygiene | 75.8 | -15.8 | -9.0 |
| | Laundry detergent | 62.9 | -11.2 | -2.2 |
| | Mouth hygiene | 64.2 | -5.7 | -5.1 |
| 7: School’s out | Baked at home bread | 0.3 | 28.5 | -15.6 |
| | Banana | 21.2 | 4.8 | 1.4 |
| | Bread (prepackaged) | 19.0 | -3.1 | 2.6 |
| | Bread (serviced) | 21.9 | -1.5 | -1.6 |
| | Candy (sugar) | 19.6 | -7.1 | -1.8 |
| | Cheese (domestic, serviced) | 10.8 | -12.2 | -7.8 |
| | Chilled sodas | -9.1 | -22.6 | -20.8 |
| | Chips | 20.7 | 2.1 | 2.8 |
| | Convenience | 8.7 | -11.9 | -9.0 |
| | Cookie | 19.2 | 0.7 | 0.8 |
| | Cookies (directly from supplier) | 18.6 | -3.6 | -5.2 |
| | Diet | 21.9 | -16.8 | -2.3 |
| | Eggs | 19.7 | 10.0 | -2.7 |
| | Fresh juices | 16.6 | 2.6 | -1.9 |
| | Frozen meals | 31.8 | -17.9 | -18.5 |
| | Frozen pizzas | 27.0 | -9.1 | -8.1 |
| | Ice tea | 8.1 | -17.8 | -14.8 |
| | Juices | 15.5 | -8.5 | -1.7 |
| | Men’s cosmetics | 8.8 | -2.7 | -6.6 |
| | Pickles | 24.9 | 3.1 | -1.1 |
| | Ready-made potatoes and vegetables | 14.0 | -2.8 | -0.2 |
| | Savory snack | -13.6 | -37.8 | -21.6 |
| | Small bread (prepackaged) | 1.7 | -16.8 | -9.5 |
| | Snack bars (not chocolate) | -2.0 | -36.1 | -12.9 |
| | Styling cosmetics | 4.2 | -19.5 | -12.7 |
| | Water | 24.1 | -14.9 | -5.7 |
| 8: Solitary | Cat food | 38.0 | -8.8 | 1.6 |
| | Cheese (prepackaged) | 26.9 | 6.6 | 3.6 |
| | Chilled bread | 22.6 | 9.0 | 5.7 |
| | Cleaning products (HG brand) | 20.3 | 5.6 | 7.0 |
| | Coffee milk | 39.0 | 2.4 | 6.3 |
| | Cold sauces | 31.8 | 11.2 | 7.4 |
| | Dairy drink | 20.6 | 11.2 | 5.4 |

(continued on next page)
Table 2 (continued)

| Cluster                        | Product category        | Panic   | Post-panic | Average |
|--------------------------------|-------------------------|---------|------------|---------|
| Dairy food                     | 21.8                    | 5.6     | 4.6        |
| Dog food                       | 29.9                    | –2.2    | 3.6        |
| Fresh nuts                     | 14.8                    | 3.9     | 8.4        |
| Frozen petfood                 | 15.3                    | –6.8    | 3.7        |
| Meats (prepackaged)            | 35.5                    | 4.5     | 3.8        |
| Meats (serviced)               | 10.6                    | –3.8    | 4.5        |
| Mexican                        | 35.8                    | 7.1     | 7.3        |
| Nuts (long shelf life)         | 31.2                    | 5.3     | 6.0        |
| Salad essentials               | 21.2                    | 3.2     | 6.3        |
| Sausages (prepackaged)         | 26.5                    | 3.0     | 4.6        |
| Tobacco                        | 14.8                    | 0.7     | 9.1        |
| Unprocessed vegetables         | 19.4                    | 8.0     | 6.4        |
| Not assigned                   |                         |         |            |
| Barbecue (nonfood)             | –98.2                   | 23.8    | 12.6       |
| Candles                        | 51.5                    | 33.0    | 71.2       |
| Chilled beer                   | –19.2                   | –19.0   | –30.6      |
| Chilled fruit                  | 6.0                     | –25.0   | 17.6       |
| Cleaning products (directly from supplier) | 92.2 | 41.3 | 20.1 |
| Counter products (snacks)      | –11.1                   | –44.5   | –30.7      |
| Disposables                    | 25.0                    | –8.8    | 37.7       |
| Electronics                    | 50.5                    | 32.0    | 53.5       |
| Flowers and plants (bouquets)  | –33.1                   | –15.7   | 2.6        |
| Flowers and plants (shelf)     | –15.0                   | –8.0    | 18.5       |
| Fresh pizza                    | –28.3                   | –27.6   | –24.1      |
| Gambling                       | –19.6                   | –44.0   | 21.6       |
| Grill                          | –15.3                   | –21.4   | –22.9      |
| Hand soap                      | 922.7                   | 195.4   | 102.1      |
| Hard fruits                    | 21.0                    | 4.9     | 9.2        |
| Magazines                      | 23.3                    | 14.8    | 13.2       |
| Multipack ice cream            | 0.1                     | 5.3     | 2.3        |
| Other breads                   | 26.6                    | 20.9    | 11.0       |
| Plants                         | –43.6                   | –22.6   | 12.2       |
| Salads (serviced)              | –30.4                   | –8.2    | –34.9      |
| Seasonal meats                 | –11.9                   | 49.1    | 21.5       |
| Shampoo                        | 42.7                    | 15.9    | 3.0        |
| Sparkling wines                | –1.7                    | 16.1    | 46.3       |
| Syrups                         | 34.7                    | 16.3    | 0.2        |
| Vegetables meal pack           | –12.0                   | –18.0   | –32.1      |

References

Aghabozorgi, S., Seyed Shirkhorshidi, A., Ying Wah, T., 2015. Time-series clustering – a decade review. Inf. Syst. 53, 16–38.

Andersen, A.L., Hansen, E.T., Johansen, N., Sheridan, A., 2020. Consumer responses to the COVID-19 crisis: evidence from bank account transaction data (Working Paper No. 18/20). Center for Economic Behavior & Inequality. Retrieved from http://www.econ.ku.dk/cebi/publikationer/working-papers/CEBI_WP_18-20_rev.pdf.

Baarsma, B., Groenewegen, J., 2021. COVID-19 and the demand for online grocery shopping: empirical evidence from the Netherlands. De Economist 169, 407–421. https://doi.org/10.1007/s10645-021-09389-y.

Bennett, G., Young, E., Butler, I., Coo, S., 2021. The impact of lockdown during the COVID-19 pandemic on dietary habits in various population groups: a scoping review. Front. Nutr. 8, 62642.

Billore, S., Anisimova, T., 2021. Panic buying research: a systematic literature review and future research agenda. Int. J. Consum. Stud. 45 (4), 777–804.

Carvalho, Â.M., Garcia, J.R., Hansen, S., Ortiz, Álvaro, Rodrigo, T., Morã, J.V.R., Ruiz, J., 2020. Tracking the COVID-19 crisis with high-resolution transaction data (Working Paper No. 2030). University of Cambridge, Faculty of Economics. Retrieved from https://www.econ.cam.ac.uk/research/files/repe/cam/pdf/cwp2030.pdf.

Chen, H., Qian, W., Wen, Q., 2021. The impact of the COVID-19 pandemic on consumption: learning from high-frequency transaction data. AEA Papers and Procedings 111, 307–311. https://doi.org/10.1257/pandp.202111003.

Chouakria, A.D., Nagababhan, P.N., 2007. Adaptive dissimilarity index for measuring series proximity. Adv. Data Anal. Classificat. 1 (1), 5–21.

Eger, L., Komárom, I., Egervá, D., Mišk, M., 2021. The effect of COVID-19 on consumer shopping behaviour: generational cohort perspective. J. Retailing Consum. Serv. 61, 102542.

Filides, R., Ma, S., Kolassa, S., 2019. Retail forecasting: research and practice. Int. J. Forecast.

Hwang, T., Rabheru, K., Peisah, C., Reichman, W., Ikeda, M., 2020. Loneliness and social isolation during the COVID-19 pandemic. Int. Psychogeriatr. 32 (10), 1217-1220.

Hyndman, R.J., Athanasopoulos, G., 2021. Forecasting: Principles and Practice, third ed. OTexts, Melbourne Australia https://otexts.com/fpp3. (Accessed 16 August 2021).

Islam, T., Pitha, A.H., Arya, V., Wang, Y., Akhtar, N., Mubark, S., Xiao, L., 2021. Panic buying in the COVID-19 pandemic: a multi-country examination. J. Retailing Consum. Serv. 59, 102357.

James, G., Witten, D., Hastie, T., Tibshirani, R., 2021. An Introduction to Statistical Learning: with Applications in R, second ed. Springer, New York, https://www.gremlin.com/essays/dependencies.html. (Accessed 16 August 2021).

Jung, E., Sung, H., 2017. The influence of the Middle East Respiratory Syndrome Syndrome outbreak on online and offline markets for retail sales. Sustainability 9 (3), 411, Kaushik, M., Guleria, N., 2020. The panic buying behavior of consumers during the COVID-19 pandemic: the stimulus-organism-response approach. J. Retailing Consum. Serv. 62, 102600.

Letzel, V., Pozas, M., Schneider, C., 2020. Energetic students, stressed parents, and nervous teachers: a comprehensive exploration of inclusive homeschooling during the COVID-19 crisis. Open Education Studies 2 (1), 159–170.

Li, H., Wu, Y.J., Zhang, S., Zhou, J., 2021. Temporary rules of retail product sales time series based on the matrix profile. J. Retailing Consum. Serv. 60, 102431.

Montero, P., Villar, J.A., 2014. TScst: an R package for time series clustering. J. Stat. Software 62 (1), 1–43.

Murtugh, F., Legendre, P., 2014. Ward’s hierarchical agglomerative clustering method: which algorithms implement Ward’s criterion? J. Classif. 31 (3), 274–295.

O’Connell, M., de Paula, Á., Smith, K., 2021. Preparing for a pandemic: spending dynamics and panic buying during the COVID-19 first wave. Fisc. Stud. 42 (2), 249–264.

Omar, N.A., Nazri, M.A., Ali, M.H., Alam, S.S., 2021. Panic buying behavior of consumers during the COVID-19 pandemic: examining the influences of uncertainty, perceptions of severity, perceptions of scarcity, and anxiety. J. Retailing Consum. Serv. 62, 102224.

Panzone, L., Larcom, S., She, P., 2021. Estimating the impact of the first COVID-19 lockdown on UK food retailers and the restaurant sector. Global Food Secur. 28, 100495.

Polikar, R., 2009. Ensemble learning. Scholarpedia 4 (1), 2776 (revision #186097). Simchi-Levi, D., Sun, R., Wu, M.K., Zhu, R., August 9, 2020. Calibrating sales forecast in a pandemic using competitive online non-parametric regression. Available at SSRN: https://ssrn.com/abstract=3670264 or. http://dx.doi.org/10.2139/ssrn.3670264.
Tandon, A., Kaur, P., Bhatt, Y., Mintynaki, M., Dhir, A., 2021. Why do people purchase from food delivery apps? A consumer value perspective. J. Retailing Consum. Serv. 63, 102667.

Vall Castelló, J., Lopez Casasnovas, G., 2021. The effect of lockdowns and infection rates on supermarket sales. Econ. Hum. Biol. 40, 100947.

Wang, Y., Xu, R., Schwartz, M., Ghosh, D., Chen, X., 2020. COVID-19 and retail grocery management: insights from a broad-based consumer survey. IEEE Eng. Manag. Rev. 48 (3), 202–211.

Danas Zuokas, a Senior Data Scientist at NielsenIQ received his Ph.D. in Mathematics from the Vilnius University in Lithuania. He also has a bachelor’s degree in Mathematics with minor degree in Business Management and a master’s degree in Mathematical methods in Economics from the Vytautas Magnus University in Lithuania. Dr. Zuokas has 15 years of experience in applying various statistical and machine learning methods to solve problems in retail (customer segmentation, market basket analysis), telecommunication (churn prediction, customer lifetime value), and investment management (portfolio optimization) industries.

Evren Gul, a Data Science Advisor at NielsenIQ received his Master of Statistics and Ph.D. in Industrial Engineering with concentration in Statistics from the Georgia Institute of Technology. He also obtained bachelor’s and master’s degrees from the Middle East Technical University in Turkey. Dr. Gul has 10 years of experience in research and development of novel statistical methods for solving complex problems and is currently focused on solving retail merchandising problems. He is a recipient of the Statistical Partnerships Among Academe, Industry & Government (SPAIG) Award from the American Statistical Association in 2020 and the Lloyd S. Nelson Award from the American Society for Quality in 2021.

Alvin Lim, the Chief Science Officer of NielsenIQ received his Ph.D. in Mathematical Sciences from the Johns Hopkins University. He also has a Bachelor of Science in Mathematics and a Master Science in Applied Mathematics from the University of the Philippines, Diliman. Dr. Lim has over 25 years of research and development experience and specializes in the solution of pricing, promotion, product assortment, and marketing mix optimization problems in retail and B2B applications. He is an industry leader in the research, development and application of pricing, revenue management and marketing analytics.