

ecommerceDB

# Product Description & Methodology

# ecommerceDB provides unrivaled information on stores, marketplaces, companies, and markets

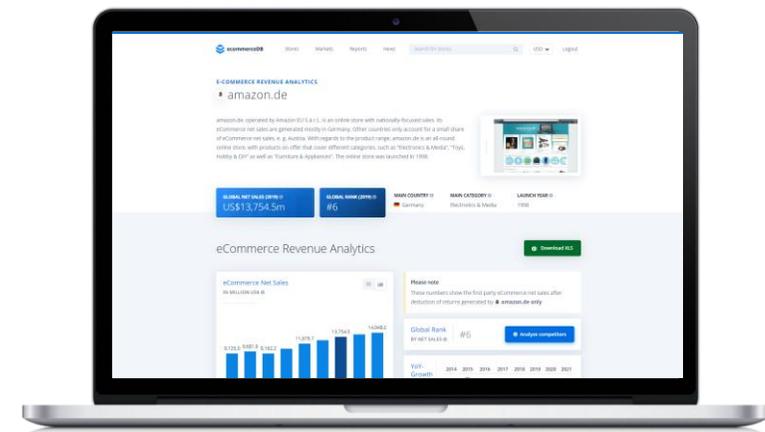
About ecommerceDB (1/2)

## What is ecommerceDB?

The ecommerceDB database provides rankings of online stores, companies and marketplaces in over 50 countries. These rankings cover 42k+ online stores, 350 companies and 300+ online marketplaces. Detailed profiles are available for each player, providing information on 70+ financial and performance KPIs.

Additionally, ecommerceDB provides detailed data on the eCommerce development in 150 national markets. Within each national market, there is a further breakdown of around 250 product markets. The data includes information on market size, growth rates, and analysis of the top players in each market.

In addition to the database, ecommerceDB publishes a wide range of in-depth reports and insights covering analysis of eCommerce markets, customer behavior and market trends.

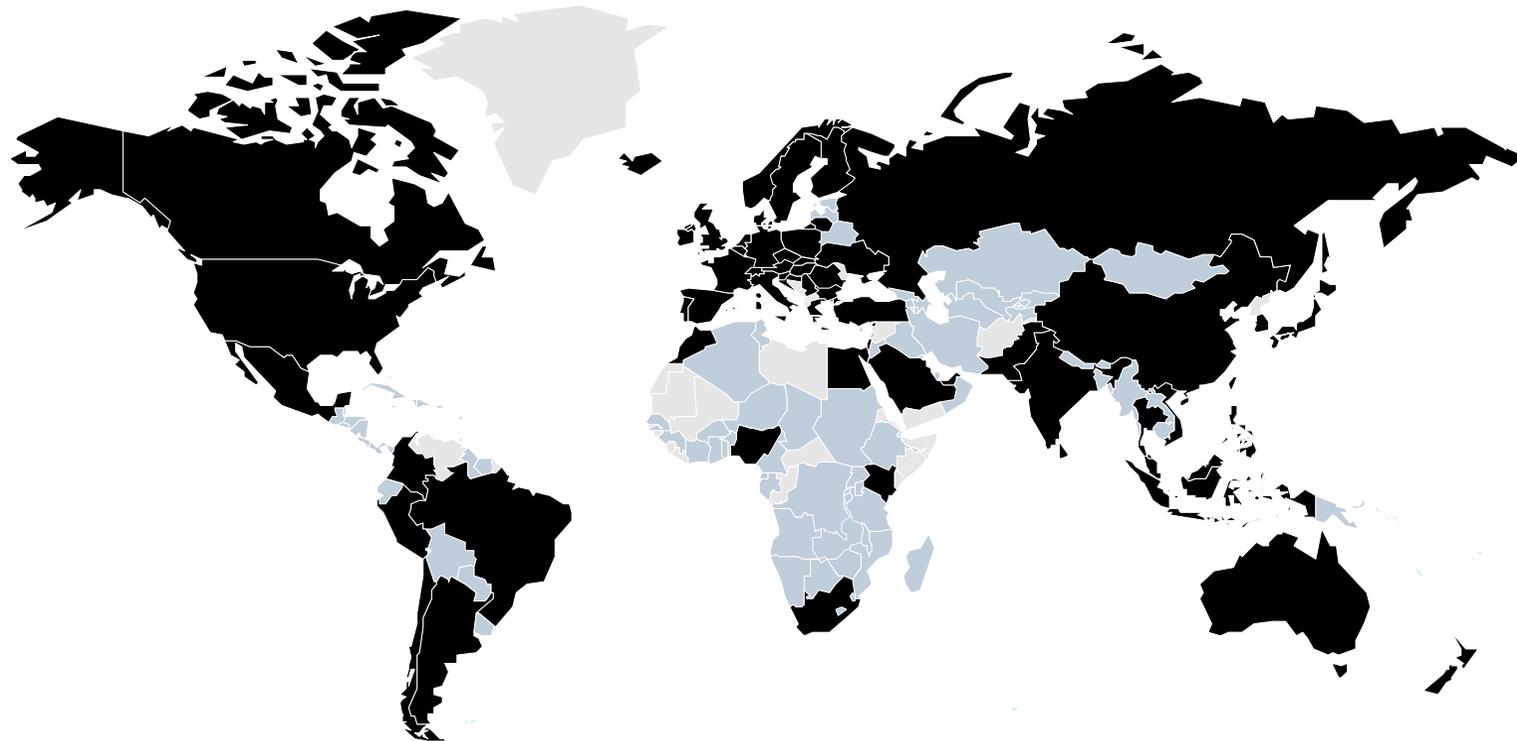


## Key figures

- Store-, company-, and marketplace-rankings in 50+ countries
- Detailed profiles of 42k+ stores, 350 companies and 300+ marketplaces
- 70+ financial and performance KPIs
- eCommerce development in 150 national markets
- Market development for ~250 product categories
- 1,000+ articles, deep dives, and reports

# ecommerceDB has global coverage, offering a comprehensive view of eCommerce worldwide

About ecommerceDB (2/2)



■ Covered: with ranking and market data

■ Covered: with market data

■ Not covered

ecommerceDB

Modeling approach: eCommerce net sales

# Online store revenue modeling is based on diverse data types and expert analysis

Modeling approach overview

## Different approaches depending on data availability

### 1) Historical ecommerce net sales development

When approximating past revenue trends, the modeling approach depends on which types of data are available. Our ecommerceDB experts conduct research about companies based on existing information, i.e., data research on thousands of online stores, annual reports, and secondary sources. This information is supplemented by other store-specific drivers, such as traffic information, offered product range, price level, and shipping information. Depending on data availability, we use two different approaches:

 Top-down: Finding information about the eCommerce turnover of a company or business unit and enhancing it with further store-specific data enables our analysts to determine the revenue for the company's online store.

 Bottom-up: If there is no information on the company's (eCommerce) revenue, we use an algorithm-based approach using store-specific KPIs such as traffic, product range, price level, and average conversion rates to approximate the online store's revenue. The algorithm is based on the correlations between success factors and revenue derived from the top-down approach acquisition of primary data.

### 2) Forecast

 Forecasting the online store's revenue for the running calendar year and the following years takes multiple information factors into account. Our experts evaluate the store's historical revenue development, traffic trends, general market forecasts, and individual store information.

## Underlying data sources

Research on online stores

- eCommerce net sales
- Company information
- Product range and shipping information
- Contact information

Quantitative and financial data from public sources:

- Annual reports and financial statements
- Press releases
- Secondary sources

Store-specific data

- Traffic information, visits in total and by region
- Offered product range
- Price level information
- Shipping information

Studies & third-party data

- Statista Market Insights: Digital
- Statista Market Insights: Consumer
- International organizations, e.g., IMF, World Bank, and OECD

# tjx.com net sales approximation based on top-down approach



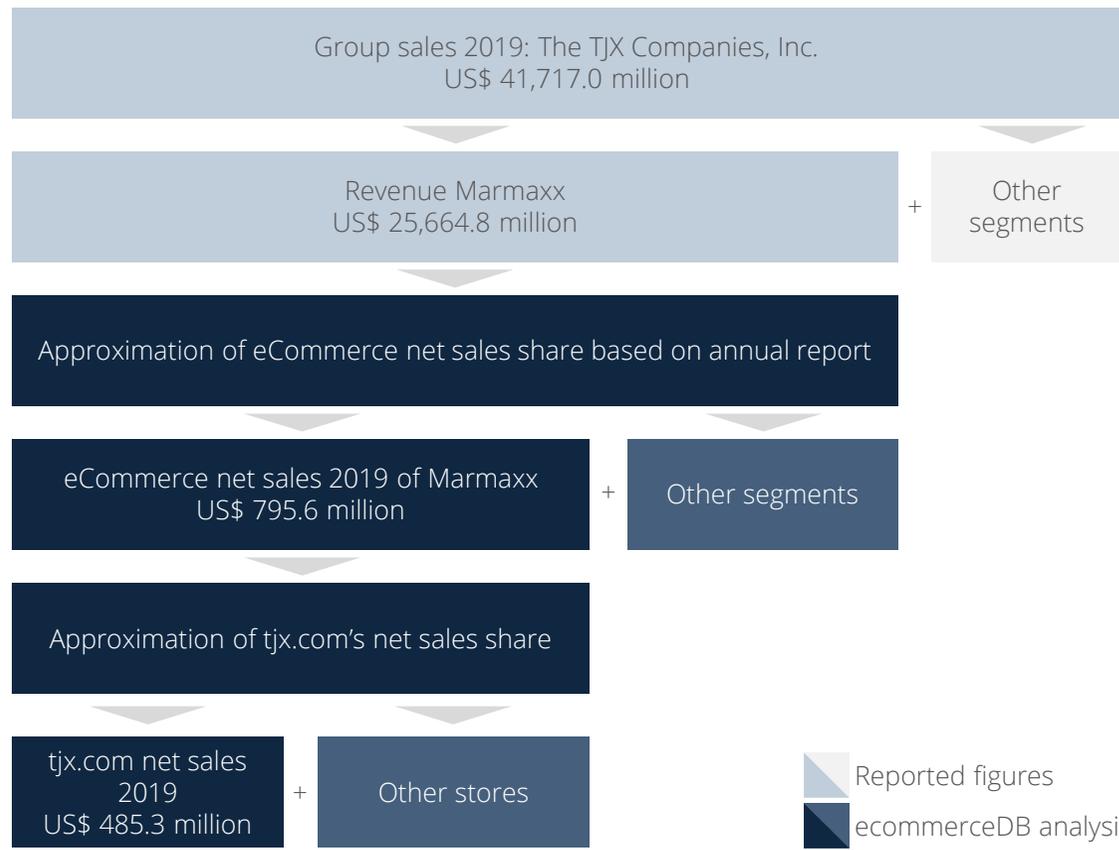
Historical ecommerce net sales development (1/3)

## Top-down net sales approximation

Referring to the annual report of The TJX Companies, Inc., group sales in 2019 amount to US\$ 41,717.0 million<sup>1</sup>. Total group sales include revenue through online and offline retail in various segments. About 61.5% of the Group's revenue is generated by the Marmaxx segment, comprising online and offline retail revenues<sup>1</sup>.

The reported share of online net sales in Marmaxx revenue increased from 2% to 3% between 2015 and 2019<sup>2</sup>. Based on this steady increase and with reference to the positive development of revenue in the eCommerce fashion market<sup>3</sup> and the positively formulated prospects of the management in connection with further investments in the eCommerce segment<sup>1</sup>, we calculate an average eCommerce net sales share of 3% in 2019 to approximate the eCommerce net sales of Marmaxx. As a result, US\$ 795.6 million are generated by Marmaxx's three online stores (tjx.com, sierra.com, and marshall.com).

Considering individual website performance indicators and additional success factors, tjx.com accounts for approximately 61% of Marmaxx eCommerce net sales or US\$ 485.3 million.



# wayfair.com net sales approximation based on top-down approach



Historical ecommerce net sales development (2/3)

## Top-down net sales approximation

The total group sales of Wayfair, Inc. in 2019 amount to US\$ 9,127.1 million<sup>1</sup>. They include revenue through different international online stores, retail partner websites, and media solutions.

The direct retail net revenue is also reported in the annual report, but there is no further information on the revenue distribution by country. But based on last year's annual report, we assume that non-ecommerce revenue is subsumed under "U.S. revenue" and that "International" revenue is eCommerce only. By subtracting that from the total direct retail net revenue, the U.S. eCommerce revenue amounts to US\$ 7,726 million.

eCommerce net sales in the U.S. are generated mainly via wayfair.com but also via jossandmain.com, allmodern.com, birchland.com, and perigold.com.

Considering individual website performance indicators and additional success factors, wayfair.com accounts for approximately 61% of the U.S. eCommerce net sales or US\$ 6,211.4 million.



Reported figures  
ecommerceDB analysis

# Algorithm-based bottom-up revenue modeling is driven by a broad range of success factors

Historical ecommerce net sales development (3/3)

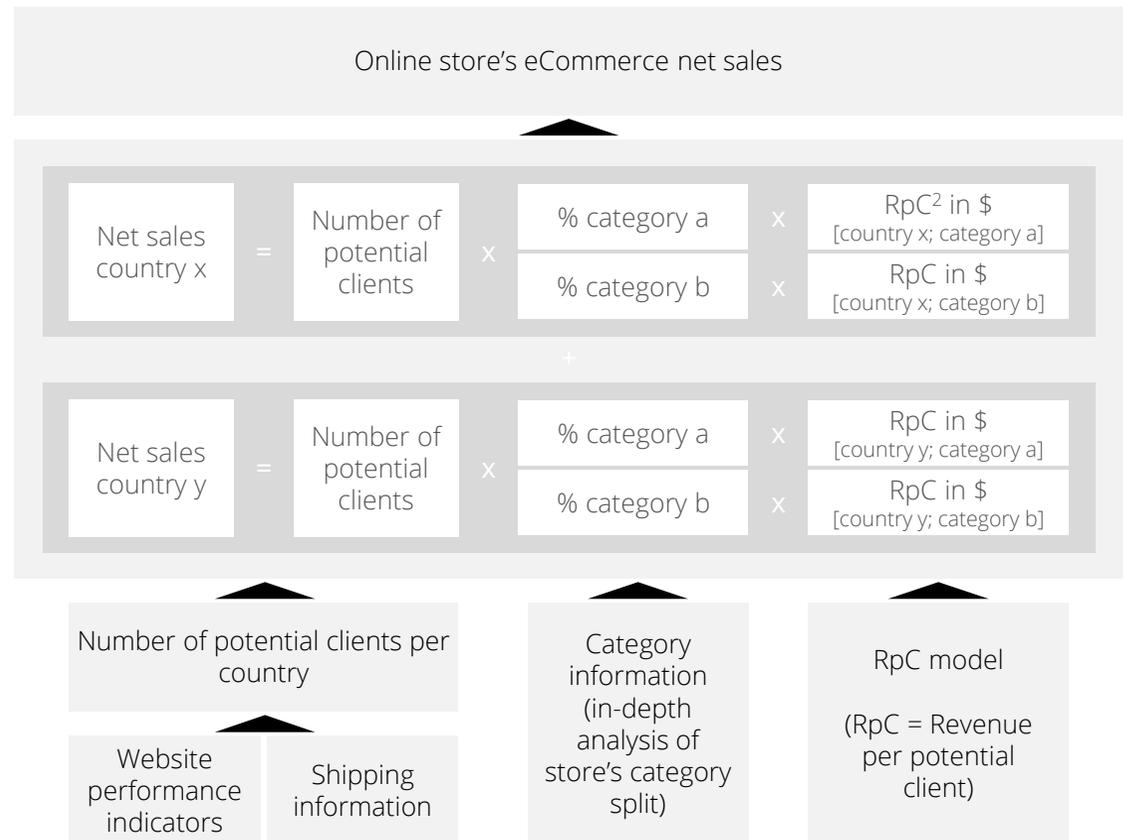
## Bottom-up net sales approximation

If there is no information about the company's (eCommerce) revenue, we apply our algorithm-based approach using a broad range of KPIs to approximate the online store's eCommerce net sales.

The algorithm is based on the correlations between success factors and revenue, which are derived using primary data from the top-down approach.

The main success factors are:

- the number of potential clients per country, derived from general website performance indicators (visits, time on site, etc.<sup>1</sup>) in combination with information on conversion rates by category and research on the store's shipping destinations
- the online store's offered product range and the average price level
- the information on average revenue per potential client for each country and category, which has been determined by a regression analysis, based on the primary data from the top-down approach



# The revenue forecast of an online store is driven by four indicators

Future ecommerce net sales development (1/2)

## Forecast approach

Our forecast approach to anticipate future eCommerce net sales development of an online store considers a variety of underlying data.

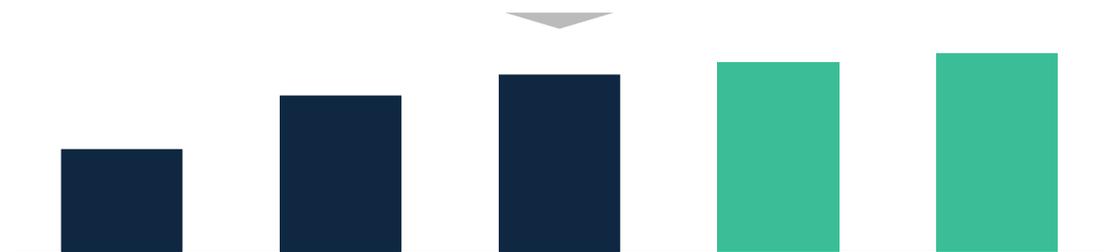
On the one hand, we examine internal and external historical information, e.g., previous net sales and traffic development of the online store<sup>1</sup>.

On the other hand, we also take indications of future developments into account. Here, we consider internal and external information such as statements by the company on its assessment of future developments and market forecasts, as conducted by Statista Market Insights team<sup>2</sup>.

These four indicators are individually weighed by our eCommerce analysts to assure the best possible forecast based on the available information.

	Internal information	External information
Historical information	<b>1</b> Store's net sales development	<b>2</b> Traffic & conversion development
Future development indicators	<b>3</b> Store's individual statements regarding its continued net sales development	<b>4</b> Statista Market Insights: Digital

These four indicators are weighed by our eCommerce analysts

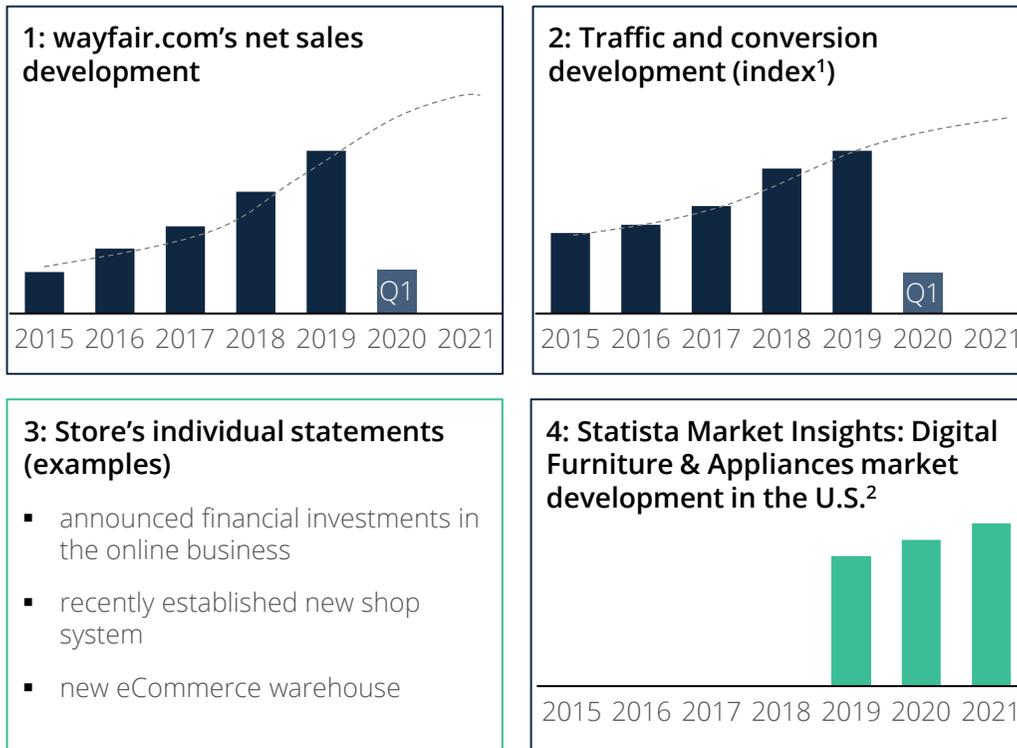


<sup>1</sup> SimilarWeb <sup>2</sup> Statista Market Insights: Digital (2021)

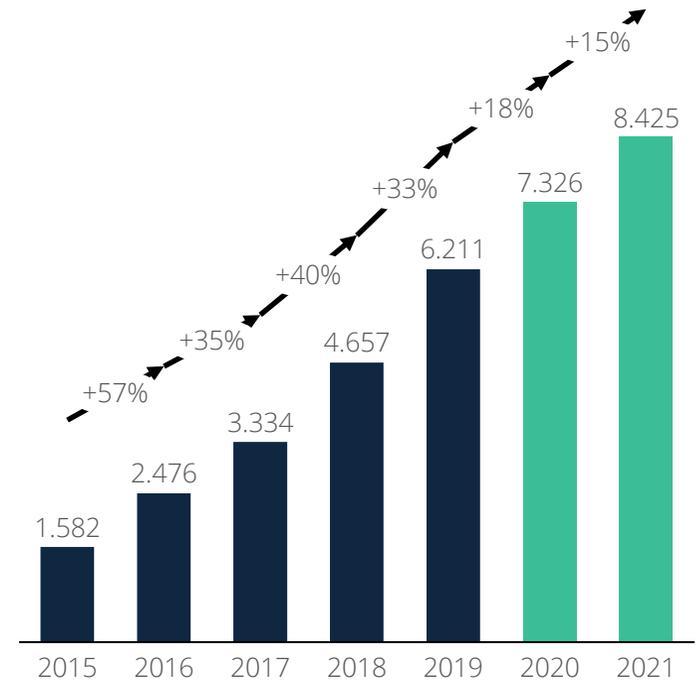
# Forecasting wayfair.com's eCommerce net sales by considering historical and future indicators

Future ecommerce net sales development (2/2)

## Input data



## wayfair.com net sales development in mUS\$



1: SimilarWeb 2: Statista Market Insights: Digital (2021)  
Source: [ecommercedb.com](https://ecommercedb.com) 2021

ecommerceDB

# Modeling approach: eCommerce GMV

# GMV modeling uses a similar approach to store net sales modeling, but with brand-level data points

Modeling approach overview

## A similar approach to store net sales modeling

To model the GMV data for online marketplaces, we follow an approach similar to that described in the previous chapter (see [“Modeling approach: eCommerce net sales”](#)). The top-down approach is used to approximate a marketplace’s historical GMV trend when the marketplace operator publishes marketplace-related data. The bottom-up approach is used when a marketplace operator does not provide information about its marketplace business or success factors. The GMV forecast follows the same approach as the net sales forecast.

## Differences compared with store net sales modeling

One difference between marketplace GMV modeling and store net sale modeling is the underlying data sources: GMV modeling uses marketplace-specific data as input (see list on right).

In addition, we always provide the brand-level data for a marketplace (e.g., total GMV of Walmart), along with the key marketplace URLs (e.g., GMVs for walmart.com, walmart.ca, and walmart.com.mx, respectively). In the top-down approach, we break down the published brand-level data to individual URLs. In the bottom-up approach, we first estimate GMV for individual URLs and then aggregate them to obtain brand-level data. The brand-level data point is not available for stores’ net sales.

## Underlying data sources

Marketplace-specific data:

- Traffic information
- Product range and main product category
- Price level information
- Number of third-party vendors
- Type of commission
- Commission rates
- Geographical operating area
- Scope of the marketplace
- Type of marketplace

Quantitative and financial data from public sources:

- Annual reports and financial statements
- Press releases
- Secondary sources

Studies & third-party data

- Statista Market Insights: Digital
- Statista Market Insights: Consumer
- International organizations, e.g., IMF, World Bank, and OECD

ecommerceDB

Modeling approach: consumer behavior  
insights

# The average order value analysis combines both actual and estimated data

Average order value (AOV)

## Definition

Average order value (AOV) is the average amount customers spend on each order, inclusive of VAT and exclusive of returns and cancellations. To model AOV values, our approach involves two steps.

## Actual AOV data as a benchmark

In the first step, we research actual AOV data from companies' published documents, such as annual reports and financial statements. If a company has reported an AOV, we assign that value to all of its online stores. For example, the AOV reported by the company Zalando SE is assigned to zalando.de, zalando.pl, zalando.ch, etc.

In addition to store-specific AOV, we also research AOV by country and product category. All collected actual data is then used as a benchmark to build an AOV estimation model for stores where AOV data is not directly available.

## Estimation based on influencing factors

Using the actual data collected in the first step, we develop an AOV estimation model that considers several potential influencing factors that can affect a store's AOV. These factors include:

- country and category shares combined with price indices (e.g., a Swiss shopper buying consumer electronics spends more than an Indian shopper buying books);
- shopping method, i.e., via mobile device or desktop computer;
- free shipping policies of stores;
- specialist groups under which the stores are classified (e.g., music stores are part of our hobby segment, but instruments tend to be more expensive than other hobby-related items).

The factors are weighted based on their correlations with AOV. A store's unknown AOV is then approximated based on the estimation model, using the store's actual data on these influencing factors as input.

# The conversion rate analysis combines both actual and estimated data

Conversion rate (CVR)

## Definition

Conversion rate (CVR) is the percentage of website visits that result in orders. Similar to AOV, we model the CVR of online stores in two steps.

## Actual CVR data as a benchmark

In the first step, we research actual CVR data from companies' published documents, such as annual reports and financial statements. If a company has reported a CVR, we assign that value to all of its online stores. For example, the CVR reported by the company Zalando SE is assigned to zalando.de, zalando.pl, zalando.ch, etc.

In addition to store-specific CVR, we also research CVR by country and product category. All collected actual data is then used as a benchmark to build a CVR estimation model for stores where CVR data is not directly available.

## Estimation based on influencing factors

Using the actual data collected in the first step, we develop a CVR estimation model that considers several potential influencing factors that can affect a store's CVR. These factors include:

- country and category shares (e.g., conversion rates for food and beverage are higher than those for furniture or homeware);
- shopping method, i.e., via mobile device or desktop computer;
- channels through which shoppers access the online stores: direct access, search, email, referrals, display ads, or social media channels;
- presence of brick-and-mortar (an online store can also be used as a catalog for consumers before they shop in physical stores).

The factors are weighted based on their correlations with CVR. A store's unknown CVR is then approximated based on the estimation model, using the store's actual data on these influencing factors as input.

ecommerceDB

# Modeling approach: eCommerce markets

# A bottom-up approach and a driver model are used to model eCommerce market revenues

The eCommerce market development research and modeling is conducted in partnership with [Statista Market Insights: Digital](https://www.statista.com).

A bottom-up approach is used to approximate the eCommerce market size in countries where data availability is high, while a driver model is developed to estimate the eCommerce market size in countries where data availability is low.

You can find the detailed methodology [here](#).

**We use a combined top-down and bottom-up approach for our sizing of the market's status quo**  
Market sizing (1/1)

**Detailed status quo analysis of selected countries**  
We obtain the data from the Statista Digital Market Outlook for 56 countries, such as the United States, China, and Germany, through an in-depth analysis of the markets. To evaluate the market, we use the latest data from various country-specific data and analytics resources, survey results from our primary research in the [Statista Global Consumer Survey](#). Primary studies and reports, as well as our analytics knowledge. Review of the annual performance, interpretation, and analysis that flows from the Statista Digital Market Outlook, a detailed representation of the data sources for each data point is not possible. Market size is determined using a bottom-up approach based on an individual top-down approach.

**Market research**  
Digital consumer profiling - the [Statista Global Consumer Survey](#)  
- Exclude representative ad hoc surveys in selected countries on specific current topics

**Key player analysis & monitoring**  
- Company profiles and key performance indicators  
- Product and price monitoring  
- News and trends

**Macroeconomic indicators**  
- Country-specific statistical offices and census data  
- International organizations and associations

**Stocks & third-party data**  
- Market analysis and analyst opinions  
- Acquired reports, industry analyses, and public relations

**We use a combined top-down and bottom-up approach for our sizing of the market's status quo**  
Market sizing (2/1)

**Driver-based forecast of market data to maximize with low data availability**  
The market data for countries with low data availability is generated using an algorithmic model. To maximize the use of available data to estimate the country's potential, we use the performance rates of all countries with a similar situation and/or similar conditions as benchmark values. There are only a few algorithm-based indicators to create market KPI estimations, using the country's key market indicators as drivers. Over 100 data sets for 150 countries and territories have been collected from a variety of sources, including the International Monetary Fund (IMF), the International Telecommunication Union (ITU), and the World Bank. The data set includes historical data from 2008 to 2020. The data are categorized and assigned to the market or the driver based on the correlation coefficient. The driver model is used to estimate market KPIs, using data from the country of base and the country key market indicators as drivers

**Country market using exemplified by the Hungarian Fashion eCommerce market in 2021**  
- Driver 1: Online share rate country? Hungary  
- Driver 2: Current key market indicators? Hungary  
Population: 10,366,765  
GDP: 150,545,000,000 USD  
Consumer spending per capita: 14,545 USD  
Consumer spending on online and mobile: 1,243 USD

**Results**  
- Hungary: 100%  
- Hungary: 100%  
- Hungary: 100%  
- Hungary: 100%

**The application of the Bass innovation diffusion model exemplified by Online Food Delivery market growth**  
Forecasting (2/1)

**Innovation diffusion curve (2021)**

The diffusion of innovations graph shows successive groups of consumers adopting a new technology. The graph shows the penetration rate of selected countries. Diffusion is the rate and volume at which innovations spread among their users (or adopters) over time (100% is the steady state). Given the steady high diffusion rates, the Online Food Delivery market is likely to grow most strongly in the next few years.

**The eCommerce market as an example of the bottom-up approach**  
Market sizing (2/1)

**Bottom-up approach**  
Retail apparel revenue + Retail footwear revenue + Retail accessories revenue = eCommerce revenue share  
Share of online apparel shoppers + Share of online footwear shoppers + Share of online accessories shoppers = Population

**Source and benchmark**  
- Fashion (through third-party studies, national surveys, the [Statista Global Consumer Market Outlook](#), expert interviews)  
- Retail revenue data from the [Statista Global Consumer Market Outlook](#)  
- Third-party studies, national statistics, primary research, Google Trends search data, key market indicators

**Our market growth forecasting is based on parameterized forecasts**  
Forecasting (1/1)

To create forecasts for up to 5 years, we combine trend modeling with statistical and mathematical forecasting methods. Digital penetration and revenue are not only used as a substitute for the same but also as a driver. This market analysis is based on a driver model. The Bass diffusion model is used to estimate market potential. Values can be identified through regression analysis based on long time series and they define the market potential. The degree of knowledge penetration can be represented as a graph.

Once the status quo has been established, we assess the current market growth and the macroeconomic environment of the country or territory in question and its impact. We then move on to trend modeling, making use of historical and current data to estimate the market's future growth.

A market growth forecast is generated with the help of our tool, which combines historical market capacity parameters regarding the development of the market drivers by creating an S-curve function. The S-curve is a special case of the logistic function, which is used to model the growth of a market over time.

The Bass model is suitable for making predictions for all products despite possible differences in product characteristics and complexity. The curve can shift in time and its compression ratio, but it stays in shape under.

**Parametric forecasts: the S-curve function**  
Forecasting (2/1)

**Parameters that describe the projected market development**

A: Base level - starting point of market development (market threshold)  
B: Jumping height - remaining growth until maximum market penetration  
C: S-curve function - progression of market diffusion  
D: Inflection point - development speed (velocity of growth)  
E: Inflection point - point of maximum in starting market saturation