# Platform description

Atosa (<https://app.atosa.asia/>) is a company that provides many marketing automation solutions for Shopee, an ecommerce platform in Vietnam. These solutions include Atosa Shopee Ads, Atosa Keyword Tools, and Atosa Agency. When having subscribed to these services, the user will be provided with many tools, one of which is the Market research tool, whose user interface is presented below

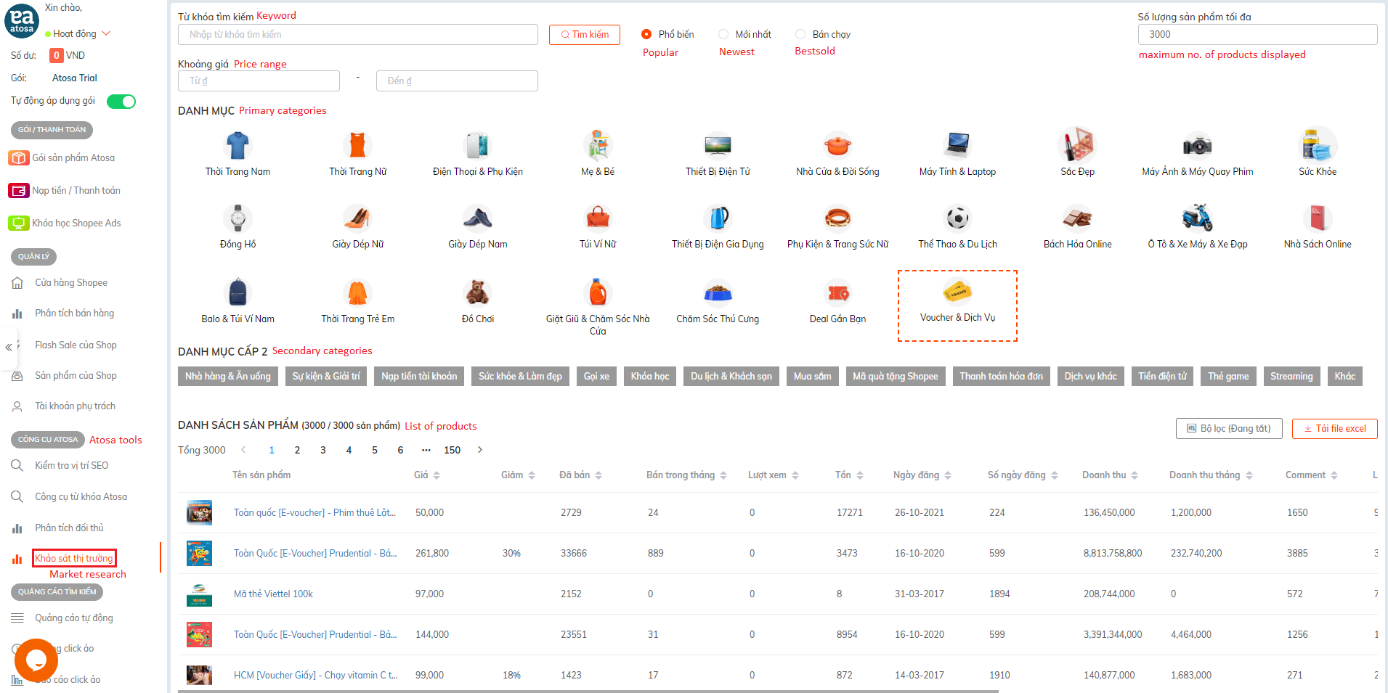


Figure: The user interface of the platform with key elements translated for a better understanding

Since Shopee is an e-commerce platform in Vietnam, the user interface (UI) is in Vietnamese so some texts have been added to the illustration to explain important elements of the UI. Generally speaking, the UI is quite intuitive and there are many options for filtering and ordering data.

The basic function of this tool is to provide a list of products following given preferences of its user. The user can filter the products by choosing a set of categories (primary, secondary, and even tertiary categories) and order the results by popularity, novelty, or bestseller. At maximum, up to 3,000 products can be loaded from the database and displayed in the results list. However, this might not be all the available products following a chosen set of conditions for fetching.

# Data description

The data has many features whose names are displayed in Vietnamese. These features are (after translated if necessary):

* Product name (Tên sản phẩm)
* Price (Giá)
* Discount (Giảm)
* Sold quantity (Đã bán)
* Last month sold quantity (Bán trong tháng)
* Currently in stock (Tồn)
* Posted date (Ngày đăng)
* Posted duration in days (Số ngày đăng)
* Total Revenue (Doanh Thu)
* Last month revenue (Doanh thu tháng)
* Like
* 5 Star
* Rating
* 5 Star
* 4 Star
* 3 Star
* 2 Star
* 1 Star

Fetched data can be saved as an Excel file. However, the downloaded data does not include the last 5 features (5 Star, 4 Star, 3 Star, 2 Star, and 1 Star), which are necessary for understanding the rating distribution. This data is updated constantly so it might change depending on the collected date.

The given data here is collected on June 7, 2022 following a set of certain fetching and ordering conditions. These conditions are as follows

* Ordering by popularity
* Maximum 3,000 products are displayed
* Data is fetched using only primary categories (so there are 27 categories in total)

In total there are maximally 81,000 observations in the final consolidated data before cleaning.