Context

Studying top products requires more than just product listings. You also need to know what sells well and what does not.

Content

This dataset contains product listings as well as products ratings and sales performance, which you would not find in other datasets.

With this, you can finally start to look for correlations and patterns regarding the success of a product and the various components.

Inspiration

How about trying to validate the established idea of human sensitiveness to price drops ? (discounted price compared to original retail\_price)

You may look for top categories of products so that you know what sells best

Do bad products sell ? How about the relationship between the quality of a product (ratings) and its success ? Does the price factor into this ?

Collection Methodology

The data comes from the Wish platform.

Basically, the products listed in the dataset are those that would appear if you type "summer" in the search field of the platform.

You can browse the Wish website or app to get a feel of the type of information you can get from there and how they are presented. This might give you some ideas and a better understanding.

If you are confused about some columns, you can either look at the column descriptions, browse the Wish website/app,

The data was scraped with french as settings (hence the presence of some non-ascii latin characters such as « é » and « à ») in the title column.

Features and Columns

The data was scraped in the french localisation (hence some non-ascii latin characters such as « é » and « à ») in the title column.

The title\_orig on the other hand contains the original title (the base title) that is displayed by default. When a translation is provided by the seller, it appears in the title column. When the title and title\_orig columns are the same, it generally means that the seller did not specify a translation that would be displayed to users with french settings.

A picture is worth a thousand words. In the following screenshot you see some features and how to interpret them.

