

# Final Project – Subject 1

---

## Sales of summer clothes in E-commerce Wish

### Source

[Sales of summer clothes in E-commerce Wish | Kaggle](#)

### Content

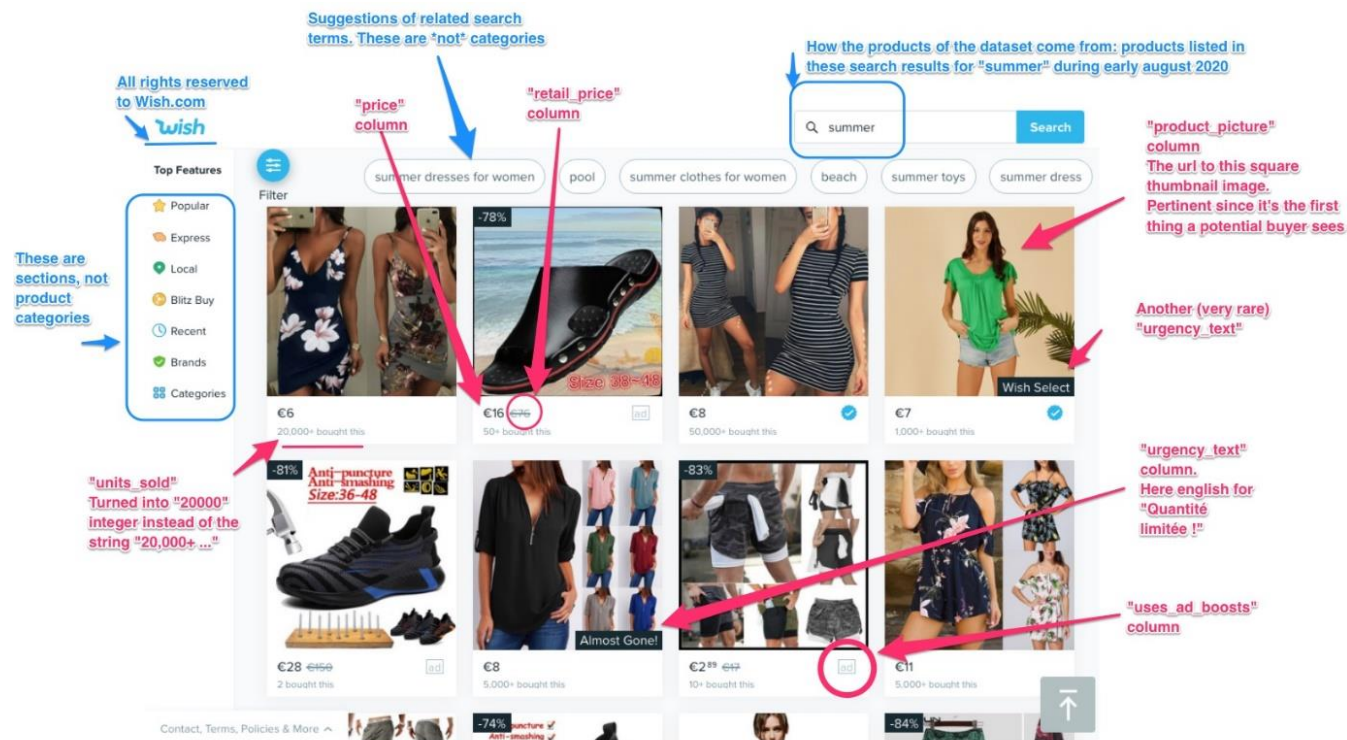
The data comes from the Wish platform and contains product listings as well as product ratings and sales performance

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, or try to find answers in online forums.

# Final Project – Subject 1



All in all, we see that the website actually does not have the concept of categories. Rather, it uses tags (keywords/compound keywords) sellers can set. When the keyword is searched for (like "summer"), products related to that tag appear.

## Column's description

1. title - title for localized for European countries. It May be the same as title\_orig if the seller did not offer a translation
2. title\_orig - original English title of the product
3. price- the price you would pay to get the product
4. retail\_price - reference price for similar articles on the market, or in other stores/places. Used by the seller to indicate a regular value
5. currency\_buyer - the currency of the prices
6. units\_sold - the number of products sold during the whole "lifespan" of the product listing.
7. uses\_ad\_boosts - whether the seller paid to boost his product within the platform (highlighting, better placement, or whatever).  
0=No boost, 1=boosts.
8. rating - mean product rating

# Final Project – Subject 1

---

9. rating\_count - total number of ratings of the product
10. rating\_five\_count - number of 5-star ratings
11. rating\_four\_count - number of 4-star ratings
12. rating\_three\_count - number of 3-star ratings
13. rating\_two\_count - number of 2-star ratings
14. rating\_one\_count - number of 1-star ratings
15. badges\_count - number of badges the product or the seller has
16. badge\_local\_product - a badge that denotes the product as a local product. Conditions may vary (being produced locally, or something else). Some people may prefer buying local products rather than.  
1 means Yes, has the badge
17. badge\_product\_quality - badge awarded when many buyers consistently gave good evaluations 1 means Yes, has the badge
18. badge\_fast\_shipping - badge awarded when this product's order is consistently shipped rapidly
19. tags - tags set by the seller
20. product\_color - product's main color
21. product\_variation\_size\_id - one of the available size variations for this product
22. product\_variation\_inventory - inventory the seller has. Max allowed quantity is 50
23. shipping\_option\_name
24. shipping\_option\_price - shipping price
25. shipping\_is\_express - whether the shipping is express or not.  
1 for True
26. countries\_shipped\_to - number of countries this product is shipped to. Sellers may choose to limit where they ship a product to
27. inventory\_total - total inventory for all the product's variations (size/color variations for instance)
28. has\_urgency\_banner - whether there was an urgency banner with an urgency

## Final Project – Subject 1

---

- 29.urgency\_text - a text banner that appears over some products in the search results
- 30.origin\_country
- 31.merchant\_title - merchant's displayed name (show in the UI as the seller's shop name)
- 32.merchant\_name - merchant's canonical name. A name not shown publicly. Used by the website under the hood as a canonical name. Easier to process since all lowercase without white space
- 33.merchant\_info\_subtitle - the subtitle text as shown on a seller's info section to the user. (raw, not preprocessed). The website shows this to the user to give an overview of the seller's stats to the user. Mostly consists of `%<positive\_feedbacks> (<rating\_count> reviews)` written in french
- 34.merchant\_rating\_count - number of ratings of this seller
- 35.merchant\_rating - merchant's rating
- 36.merchant\_id - merchant unique id
- 37.merchant\_has\_profile\_picture - Convenience boolean that says whether there is a `merchant\_profile\_picture` URL
- 38.merchant\_profile\_picture - Custom profile picture of the seller (if the seller has one). Empty otherwise
- 39.product\_url - URL to the product page. You may need to log in to access it
- 40.product\_picture
- 41.product\_id - product identifier. You can use this key to remove duplicate entries if you're not interested in studying them.
- 42.Theme - the search term used in the website's search bar to get these search results.
- 43.crawl\_month - meta: for info only