ABOUT THE DATASET

This dataset contains information on the chemicals used in cosmetics, including the name of the chemical, the company that manufactures it, the primary category it is used in, and the date it was first reported.

The dataset comes with 3 different but related files which include Chemicals-in-cosmetics (main file), chemical-in-cosmetics Data Dictionary and Chemicals-in-cosmetics (subcategory)

The chemicals-in-cosmetics csv file contains one table. This table contains 114,299 rows including the Headers and 22 columns dropping the index column.

The Chemical-in-cosmetics Data Dictionary explains what each column in the chemical-in-cosmetics represents and how they are related to each other. While the chemical-in-cosmetics (subcategory) explains how the primary category and the subcategory are related i.e. which subcategories make up a particular primary category

The file we are going to make use of throughout Data analysis is the Main file, the remaining files are for reference sake.

This dataset was Downloaded from [Chemicals in Cosmetics: What's Really in Your? | Kaggle](https://www.kaggle.com/datasets/thedevastator/chemicals-in-cosmetics-what-s-really-in-your)

DATA EXPLORATION

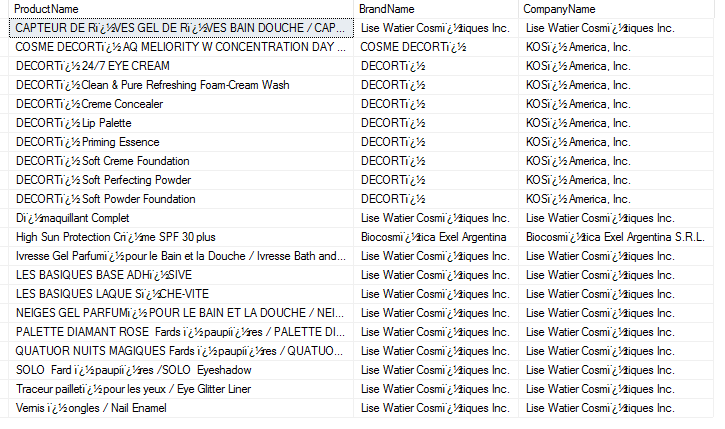
During the exploration phase of the Analysis, the following issues were uncovered:

* All the date columns were of datetime data type.

Graphical user interface, table

Description automatically generated

* Product Name, Company Name and Brand Name columns contained a strange character 'ï¿½'.

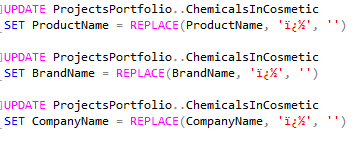


* Inconsistent names in the Company Name column



DATA CLEANING

* Updated the Product Name, Company Name and Brand Name columns by removing the strange character 'ï¿½'.

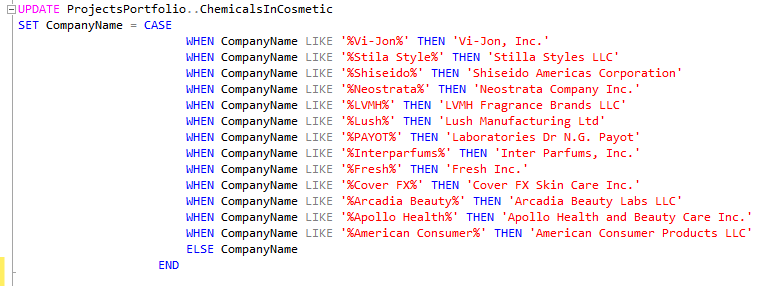


* Converted all date columns from datetime date type to date.

Text

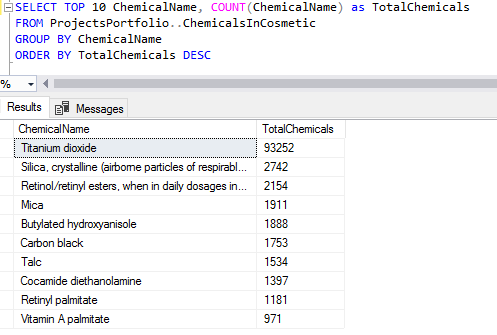
Description automatically generated

* Made consistent all inconsistent names in the Company Name column.

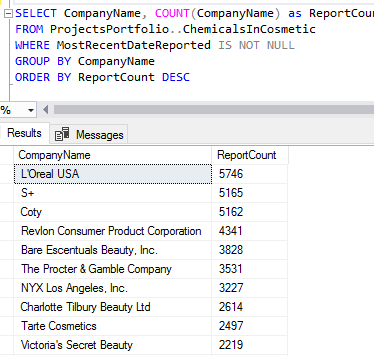


BUSINESS QUESTIONS

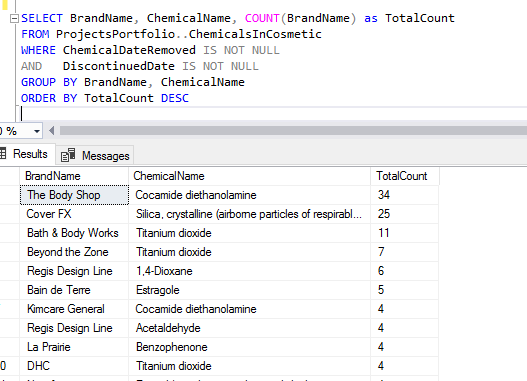
1. Find out which chemicals were used the most in cosmetics and personal care products.



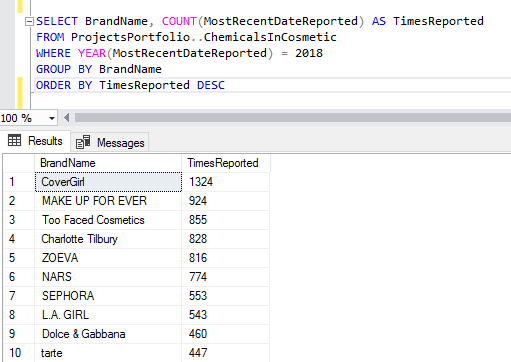
1. Find out which companies used the most reported chemicals in their cosmetics and personal care products.



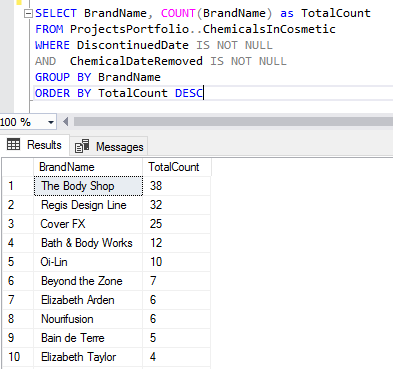
1. Which brands had chemicals that were removed and discontinued? Identify the chemicals.



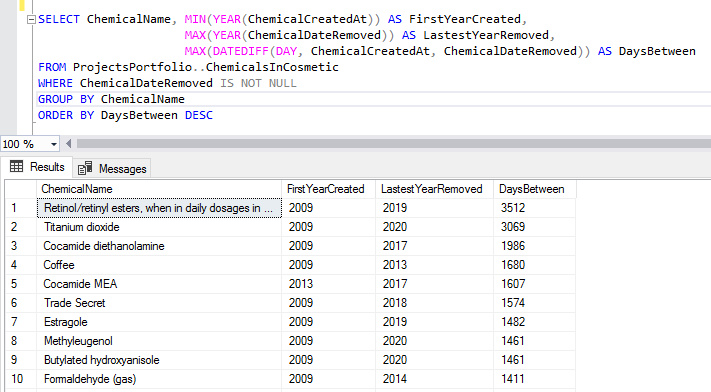
1. Identify the brands that had chemicals which were mostly reported in 2018.



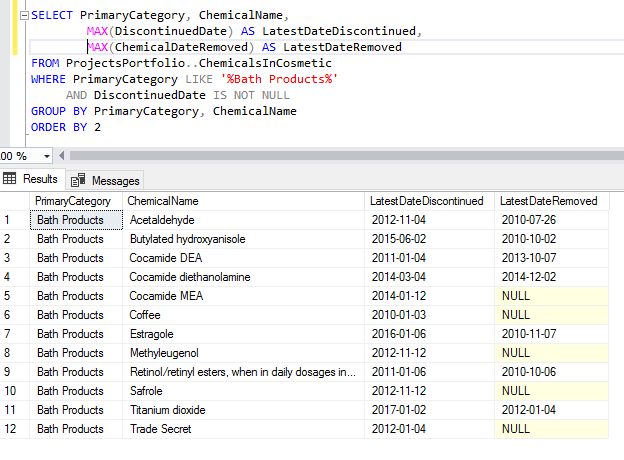
1. Which brands had chemicals discontinued and removed?



1. Identify the period between the creation of the removed chemicals and when they were removed.



1. Can you tell if discontinued chemicals in bath products were removed.



1. How long were removed chemicals in baby products used? (Tip: Use creation date to tell)

Text

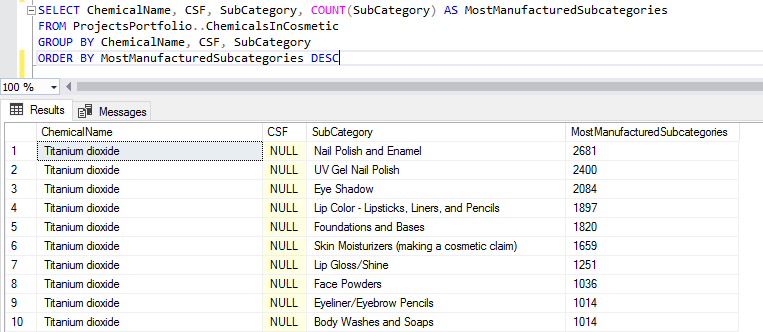
Description automatically generated with low confidence

1. Identify the relationship between chemicals that were mostly recently reported and discontinued. (Does most recently reported chemicals equal discontinuation of such chemicals?)

Table

Description automatically generated

1. Identify the relationship between CSF and chemicals used in the most manufactured sub categories. (Tip: Which chemicals gave a certain type of CSF in sub categories?)



DATA VISUALIZATION WITH POWER BI

