Unveiling Cosmetics:

Beauty or Risk?

Presented by Oleksandra Chernyshov

Project Goals

1. Identify Key Chemicals reported.
2. Determine the top 5 brands frequently reporting chemicals.
3. Discover the distribution of toxic chemicals across different product categories.
4. How Did Chemical Reports by Product Category Change Over 10 Years (2009–2019)?
5. Overall Trend of Hazardous Chemical Reports

What’s Inside Your Skincare?

Retinol

Retinol (Vitamin A) above 10,000 IU/day in pregnancy raises birth defect risk by 240%, also harming the heart and brain and central nervous system.

Titanium dioxide

Titanium Dioxide, a UV shield, risks cancer—inhaled particles raised lung cancer in rats by 32%.

Mica

Mica can cause pneumoconiosis via inhaled particles—40% risk in exposed workers; unrefined sources may carry heavy metals.

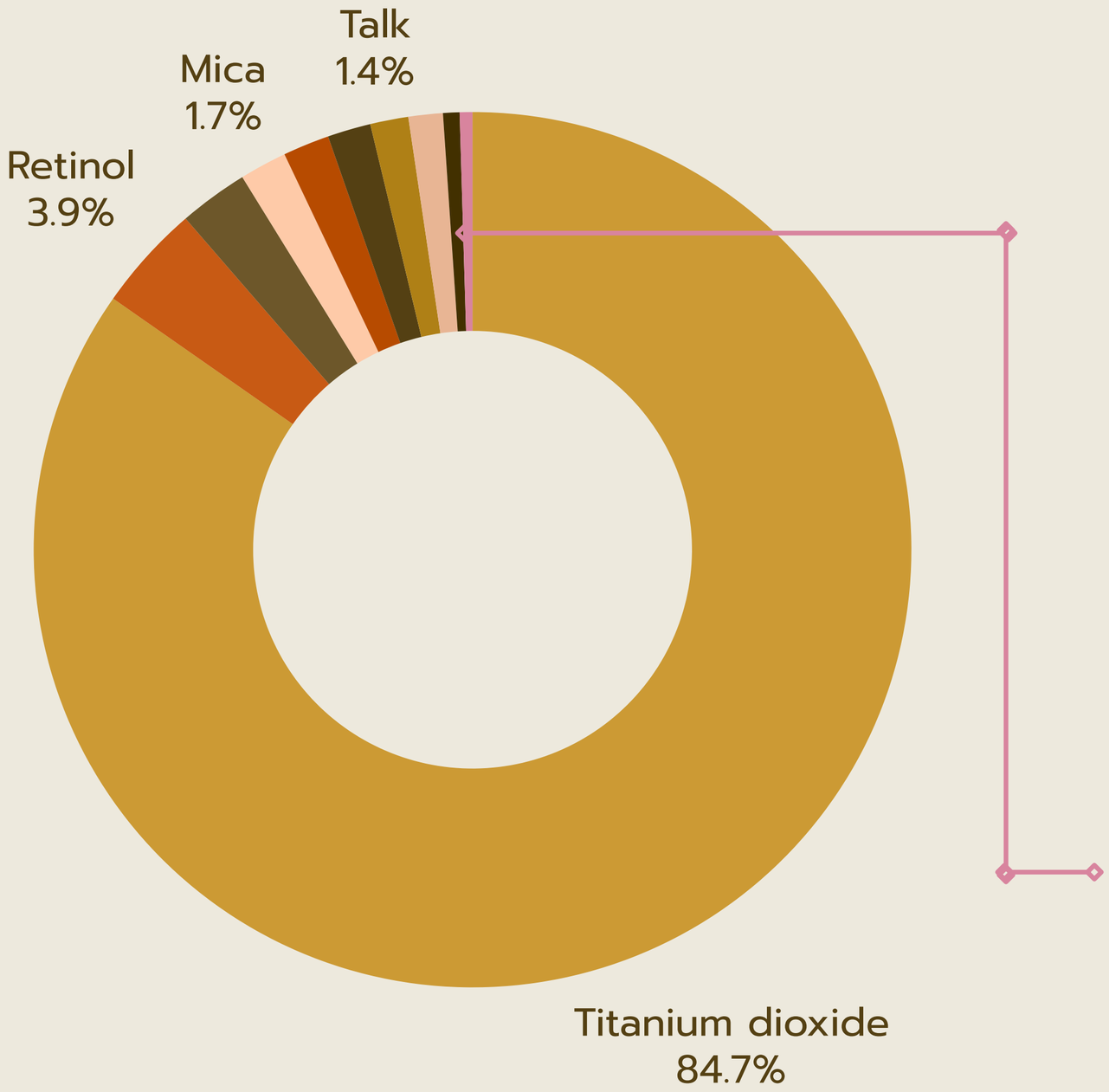
Silica

Silica, often safe in cosmetics, risks silicosis when inhaled—40% risk in workers with long-term exposure.

Reporting for Transparency

Weak regulations let harmful chemicals hide in everyday cosmetics.

Since January 1, 2007, the California Safe Cosmetics Act requires cosmetic companies with annual sales of $1 million or more to report products containing ingredients tied to cancer, birth defects, or developmental/reproductive harm.



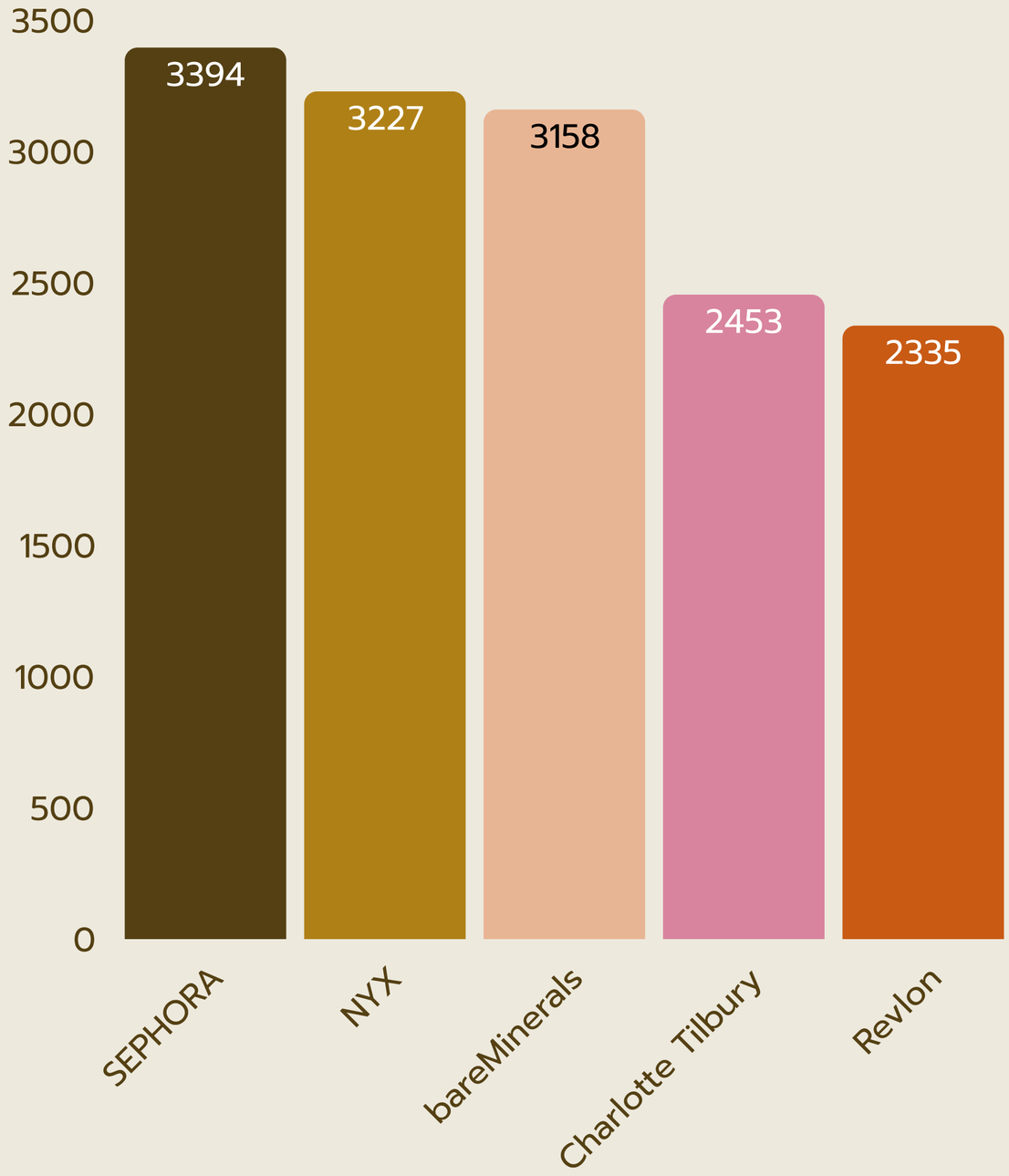
Top 10 most common chemicals reported

Titanium Dioxide is the most reported chemical — found in over 84.7% of products.

Retinol appears in 3.9% of entries, often flagged at high dosages.

Silica appears in 2.6% of product reports.

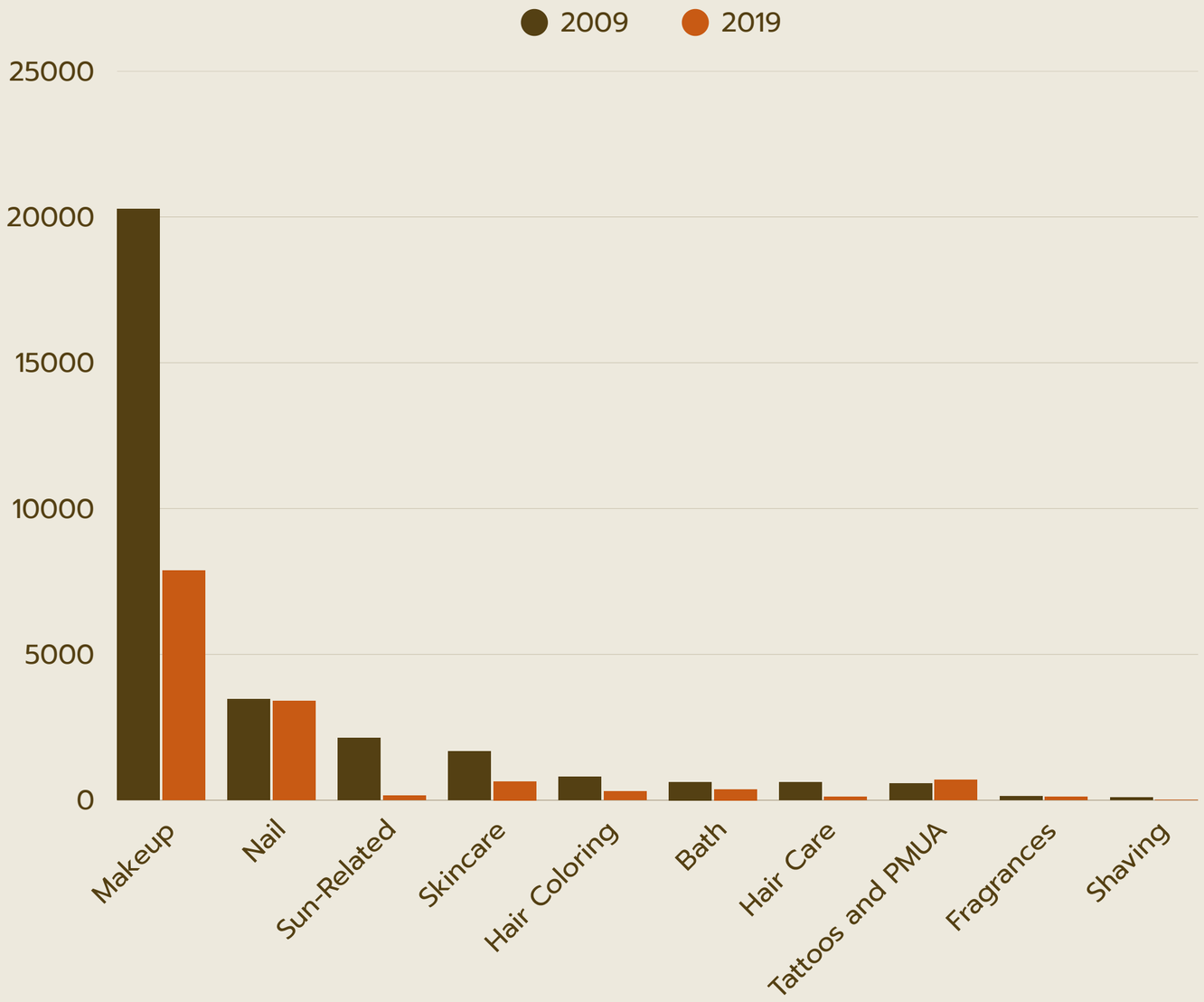
The presence of "Trade secret" (0.6%) indicates hidden ingredients, reducing transparency for consumers (common for fragrances and colorants)



Top 5 Brands reporting Chemicals

SEPHORA leads with 3,394 products containing hazardous chemicals, followed by NYX (3,227) and bareMinerals (3,158), showing these popular brands frequently report risky ingredients. This highlights the need for consumer awareness when choosing cosmetics from major players.

Chemical Reports per Primary Category: 2009 vs. 2019



Over 10 years, chemical reports in cosmetics dropped by over 60%.

Makeup alone decreasing from 20,266 to 7,877 — signaling either safer formulations or increased underreporting, especially in high-usage categories..

Trend of Hazardous Chemical Reports

The trend in the graph shows a sharp decline in reported chemical incidents after 2009, followed by fluctuating but overall lower levels through 2020.

Analysis revealed 12,920 products had chemicals discontinued over 11 years, led by Titanium Dioxide (10,291 reports) and Retinol (721 reports), signaling safer cosmetics.

Few More reasons behind the decline in reporting

Product Reformulations: Many companies removed toxic ingredients after regulatory or public pressure.

Regulatory Changes: Shifts in reporting thresholds or database updates may impact data trends.

Underreporting: Chemicals may still be used but not reported due to trade secret labeling or reporting loopholes.

Safer Manufacturing: Improved protocols and closed systems reduced exposures and incidents. (OSHA/NIOSH)

Overview of the Dataset

Source:

The dataset used in this analysis is from the California Safe Cosmetics Program (CSCP) in the California Department of Public Health

Name:

Chemicals in Cosmetics

Content:

This dataset tracks cosmetic products and their chemicals, reported to California's Safe Cosmetics Program from 2009 to 2020.
It has 22 columns and 114635 rows.

Key columns:

Product Info:
- Company_Name
- Brand_Name
- Product_Name,
- Primary_Category
- Sub_Category

Chemical Info:
- Chemical_Name
- Cas_Number
- Chemical_Count

Reporting Dates:
- Initial_Date_Reported
- Most_Recent_Date_Reported
- Discontinued_Date
- Chemical_Date_Removed

Methodologies:

Utilized Pandas, Numpy, Matplotlib and Seaborn.
Converted date related columns to datetime format using to_datetime method.
Renamed columns for improved readability.
Inspected the structure, identified missing values in CSF_ID and CSF columns with specific registration numbers which were considered as not relevant and left unchanged.
Centered the analysis around chemical names, brands, product types and relevant dates to derive insights from dataset.

Thank you very much!