

IMF BUSINESS SCHOOL  
UNIVERSIDAD CAMILO JOSÉ CELA

MÁSTER EN  
BIG DATA & BUSINESS ANALYTICS

TRABAJO FIN DE MÁSTER

---

**Análisis de productos químicos  
en cosméticos**

---

*Autor:*  
José María  
SÁNCHEZ SALAS

*Tutor:*  
Juan Manuel  
MORENO LAMPARERO

MARZO 2019



*Una de las cosas más fascinantes de los programadores es que no puedes saber si están trabajando o no sólo con mirarlos. A menudo están sentados aparentemente tomando café, chismorreando o mirando a las nubes. Sin embargo, es posible que estén poniendo en orden todas las ideas individuales y sin relación que pululan por su mente.*

Charles M. Strauss



*Como dirían en el cole:  
no hay mejor consejo que el de ser tú mismo,  
a pesar de que al resto no les mole.*

ZPU



## *Agradecimientos*





## *Resumen*



# Índice general

<b>Agradecimientos</b>	<b>VII</b>
<b>Resumen</b>	<b>IX</b>
<b>Índice de figuras</b>	<b>XIII</b>
<b>Índice de tablas</b>	<b>XV</b>
<b>1. Introducción y antecedentes</b>	<b>1</b>
1.1. Introducción . . . . .	1
1.2. Antecedentes . . . . .	1
1.3. Estructura de la memoria . . . . .	1
<b>2. Hipótesis de trabajo y objetivos</b>	<b>3</b>
2.1. Introducción . . . . .	3
2.2. Hipótesis de trabajo . . . . .	3
2.3. Objetivos . . . . .	3
<b>3. Material y métodos</b>	<b>5</b>
3.1. Introducción . . . . .	5
3.2. Material . . . . .	5
3.2.1. Descripción del dataset . . . . .	5
3.3. Métodos . . . . .	6
3.3.1. Subsection 1 . . . . .	6
<b>4. Resultados</b>	<b>7</b>
4.1. Main Section 1 . . . . .	7
4.1.1. Subsection 1 . . . . .	7
4.1.2. Subsection 2 . . . . .	7
4.2. Main Section 2 . . . . .	7
<b>5. Discusión</b>	<b>9</b>
5.1. Main Section 1 . . . . .	9
5.1.1. Subsection 1 . . . . .	9
5.1.2. Subsection 2 . . . . .	9
5.2. Main Section 2 . . . . .	9
<b>6. Conclusiones</b>	<b>11</b>
6.1. Main Section 1 . . . . .	11
6.1.1. Subsection 1 . . . . .	11
6.1.2. Subsection 2 . . . . .	11
6.2. Main Section 2 . . . . .	11
<b>Bibliografía</b>	<b>13</b>



# Índice de figuras



# Índice de tablas





## Capítulo 1

# Introducción y antecedentes

### 1.1. Introducción

La preocupación del ser humano por su aspecto y su cuidado es algo que se ha ido manteniendo a lo largo de los siglos, pues aunque hoy en día la industria cosmética parezca algo tecnológico y sofisticado, las primeras civilizaciones ya se preocupaban por ello. Precisamente, en el Antiguo Egipto se encuentran los primeros vestigios de la elaboración y utilización de diferentes productos cosméticos, utilizando para ello productos naturales, como plantas aromáticas.

La cosmética es uno de los sectores que mayor auge ha vivido durante las últimas décadas. Entre los siglos XVI al XVIII se produjo un gran desarrollo de los cosméticos y se introdujeron numerosos productos nuevos, aún fabricados principalmente a base de plantas. Pero fue a partir de principios del siglo XX cuando los cosméticos se popularizarían, hasta convertirse hoy en un producto casi imprescindible en la mayoría de los hogares.

Hoy en día, los cosméticos han vuelto a incorporar productos químicos dentro de sus fórmulas y se utilizan miles de estos compuestos a los que se le atribuyen multitud de propiedades. Sin embargo, varios científicos y organizaciones han levantado la voz de alerta sobre el impacto de estos compuestos, pues en un alto porcentaje no han sido analizados para saber el daño de estos productos sobre las personas.

### 1.2. Antecedentes

Este trabajo se embarca dentro del contexto de la aplicación de técnicas Big Data e Inteligencia Artificial al sector de la cosmética. Existen varios productos y prototipos de productos que aplican al sector cosmético este tipo de técnicas, como por ejemplo el servicio [Identité](#), que utiliza técnicas de Inteligencia Artificial y Big Data para “Crear paquetes de belleza y productos para el cuidado de la piel muy personalizados.”.

### 1.3. Estructura de la memoria



## Capítulo 2

# Hipótesis de trabajo y objetivos

### 2.1. Introducción

En este capítulo se va a exponer la hipótesis de este trabajo así como los objetivos propuestos para el mismo.

### 2.2. Hipótesis de trabajo

Los cosméticos son unos de los productos que más se consumen en los hogares, hasta tal punto que en algunos casos se han vuelto imprescindibles. Muchos productos químicos diferentes se utilizan en la fabricación de cosméticos. Los grupos de defensa de consumidores y trabajadores están preocupados porque algunos productos cosméticos contienen sustancias químicas que se sabe o se sospecha que causan cáncer, defectos de nacimiento o daños al sistema reproductivo. Aquellos que trabajan con cosméticos, incluidos los peluqueros, los estilistas y el cuidado de la piel, el cuidado del cuerpo y los trabajadores de salones de uñas, pueden ser más vulnerables a los efectos adversos para la salud que presentan estos productos porque manejan con mayor frecuencia grandes cantidades de cosméticos (State of California, 2019).

Con esta idea en mente, la hipótesis de este trabajo es aplicar técnicas de Inteligencia Artificial y Big Data al sector de la cosmética para poder encontrar patrones y relaciones entre los cosméticos y los productos químicos que contienen, así como encontrar una clasificación de dichos productos químicos para poder tener un mayor control sobre de qué están compuestos los cosméticos que consumimos.

### 2.3. Objetivos

Los objetivos principales de este trabajo son:

- Encontrar una clasificación de los productos químicos presentes en los cosméticos.
- Encontrar qué productos químicos son los más frecuentes en los cosméticos, así como los cosméticos que mayor productos químicos presentan.
- Obtener una predicción de la cantidad de productos químicos dañinos que contendrán los futuros cosméticos.



## Capítulo 3

# Material y métodos

### 3.1. Introducción

### 3.2. Material

Para la realización de este trabajo se ha hecho uso del dataset público **Chemicals in Cosmetics** ([State of California, 2019](#)) proporcionado por **HealthData** ([Health Data, 2019](#)).

Para todos los productos cosméticos vendidos en California, la Ley de Cosméticos Seguros de California requiere que el fabricante, empacador y/o distribuidor nombrado en la etiqueta del producto proporcione una lista al Programa de Cosméticos Seguros de California (*California Safe Cosmetics Program (CSCP)*) de todos los productos cosméticos que contengan cualquier ingrediente conocido o sospechoso de causar cáncer, defectos de nacimiento u otros daños al desarrollo o reproductivos. El CSCP mantiene una lista de ingredientes “reportables”. Las compañías con ingredientes reportables en sus productos deben enviar información al Programa de Cosméticos Seguros de California si la compañía:

- Tiene ventas anuales agregadas de productos cosméticos de un millón de dólares o más, y
- Ha vendido productos cosméticos en California a partir del 1 de enero de 2007.

El CSCP mantiene un sistema de informes *online* para que las empresas informen sobre productos e ingredientes reportables, generando así el dataset utilizado para este trabajo. Los datos reflejan información que ha sido reportada al CSCP. No se incluyen todos los productos que contengan carcinógenos o tóxicos para el desarrollo o la reproducción, debido a que las compañías no los informan.

#### 3.2.1. Descripción del dataset

El dataset proporcionado por el CSCP es un dataset en formato CSV con un histórico desde el año 2009 y que se va actualizando cada 3 o 4 días. En las fechas de realización de este trabajo (última fecha reportada: 21/02/2019), el dataset consta de:

- 97.760 registros.
- 22 columnas.

A continuación, se describen cada una de las columnas:

COLUMN NAME	FORMAT	DEFINITION
CDPHId	Plain text	The California Department of Public Health (CDPH) internal identifier.
ProductName	Plain text	Product name as entered by manufacturer, packer and/or distributor.
CSFId	Plain text	CDPH internal identification number for a color/scent/flavor.
CSF	Plain text	Color, scent and/or flavor as entered by manufacturer, packer and/or distributor.
CompanyId	Plain text	CDPH internal identification number for company.
CompanyName	Plain text	Company name as entered by manufacturer, packer and/or distributor.
BrandName	Plain text	Brand name as entered by manufacturer, packer and/or distributor.
PrimaryCategoryId	Plain text	CDPH internal identification number for category.
PrimaryCategory	Plain text	Type of product (13 primary categories: Baby Products, Bath Products, etc.).
SubCategoryId	Plain text	CDPH internal identification number for subcategory.
SubCategory	Plain text	Type of product within one of the 13 primary categories. See attachment for details.
CASId	Plain text	CDPH internal identification number for chemical.
CasNumber	Plain text	Chemical identification number (Chemical Abstracts Service registration number).
ChemicalId	Plain text	CDPH internal identification number for this chemical's record spreadsheet.
ChemicalName	Plain text	Name of chemical substance in cosmetic product. Note that chemical name may be different from product name.
InitialDateReported	Date	Date the product profile was created by manufacturer, packer and/or distributor.
MostRecentDateReported	Date	Date the product profile was last modified by manufacturer, packer and/or distributor.
DiscontinuedDate	Date	If applicable, date the product was discontinued.
ChemicalCreatedAt	Date	The date that this chemical was first reported to CDPH for this product.
ChemicalUpdatedAt	Date	Date this chemical report was last modified by manufacturer, packer and/or distributor.
ChemicalDateRemoved	Date	If applicable, date the chemical was removed from product, i.e., no longer used.
ChemicalCount	Number	Total number of current chemicals reported for this product. This count includes chemicals that have been removed from the product.

### 3.3. Métodos

#### 3.3.1. Subsection 1

## Capítulo 4

# Resultados

### 4.1. Main Section 1

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aliquam ultricies lacinia euismod. Nam tempus risus in dolor rhoncus in interdum enim tincidunt. Donec vel nunc neque. In condimentum ullamcorper quam non consequat. Fusce sagittis tempor feugiat. Fusce magna erat, molestie eu convallis ut, tempus sed arcu. Quisque molestie, ante a tincidunt ullamcorper, sapien enim dignissim lacus, in semper nibh erat lobortis purus. Integer dapibus ligula ac risus convallis pellentesque.

#### 4.1.1. Subsection 1

Nunc posuere quam at lectus tristique eu ultrices augue venenatis. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Aliquam erat volutpat. Vivamus sodales tortor eget quam adipiscing in vulputate ante ullamcorper. Sed eros ante, lacinia et sollicitudin et, aliquam sit amet augue. In hac habitasse platea dictumst.

#### 4.1.2. Subsection 2

Morbi rutrum odio eget arcu adipiscing sodales. Aenean et purus a est pulvinar pellentesque. Cras in elit neque, quis varius elit. Phasellus fringilla, nibh eu tempus venenatis, dolor elit posuere quam, quis adipiscing urna leo nec orci. Sed nec nulla auctor odio aliquet consequat. Ut nec nulla in ante ullamcorper aliquam at sed dolor. Phasellus fermentum magna in augue gravida cursus. Cras sed pretium lorem. Pellentesque eget ornare odio. Proin accumsan, massa viverra cursus pharetra, ipsum nisi lobortis velit, a malesuada dolor lorem eu neque.

### 4.2. Main Section 2

Sed ullamcorper quam eu nisl interdum at interdum enim egestas. Aliquam placerat justo sed lectus lobortis ut porta nisl porttitor. Vestibulum mi dolor, lacinia molestie gravida at, tempus vitae ligula. Donec eget quam sapien, in viverra eros. Donec pellentesque justo a massa fringilla non vestibulum metus vestibulum. Vestibulum in orci quis felis tempor lacinia. Vivamus ornare ultrices facilisis. Ut hendrerit volutpat vulputate. Morbi condimentum venenatis augue, id porta ipsum vulputate in. Curabitur luctus tempus justo. Vestibulum risus lectus, adipiscing nec condimentum quis, condimentum nec nisl. Aliquam dictum sagittis velit sed iaculis. Morbi tristique augue sit amet nulla pulvinar id facilisis ligula mollis. Nam elit libero, tincidunt ut aliquam at, molestie in quam. Aenean rhoncus vehicula hendrerit.





## Capítulo 5

# Discusión

### 5.1. Main Section 1

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aliquam ultricies lacinia euismod. Nam tempus risus in dolor rhoncus in interdum enim tincidunt. Donec vel nunc neque. In condimentum ullamcorper quam non consequat. Fusce sagittis tempor feugiat. Fusce magna erat, molestie eu convallis ut, tempus sed arcu. Quisque molestie, ante a tincidunt ullamcorper, sapien enim dignissim lacus, in semper nibh erat lobortis purus. Integer dapibus ligula ac risus convallis pellentesque.

#### 5.1.1. Subsection 1

Nunc posuere quam at lectus tristique eu ultrices augue venenatis. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Aliquam erat volutpat. Vivamus sodales tortor eget quam adipiscing in vulputate ante ullamcorper. Sed eros ante, lacinia et sollicitudin et, aliquam sit amet augue. In hac habitasse platea dictumst.

#### 5.1.2. Subsection 2

Morbi rutrum odio eget arcu adipiscing sodales. Aenean et purus a est pulvinar pellentesque. Cras in elit neque, quis varius elit. Phasellus fringilla, nibh eu tempus venenatis, dolor elit posuere quam, quis adipiscing urna leo nec orci. Sed nec nulla auctor odio aliquet consequat. Ut nec nulla in ante ullamcorper aliquam at sed dolor. Phasellus fermentum magna in augue gravida cursus. Cras sed pretium lorem. Pellentesque eget ornare odio. Proin accumsan, massa viverra cursus pharetra, ipsum nisi lobortis velit, a malesuada dolor lorem eu neque.

### 5.2. Main Section 2

Sed ullamcorper quam eu nisl interdum at interdum enim egestas. Aliquam placerat justo sed lectus lobortis ut porta nisl porttitor. Vestibulum mi dolor, lacinia molestie gravida at, tempus vitae ligula. Donec eget quam sapien, in viverra eros. Donec pellentesque justo a massa fringilla non vestibulum metus vestibulum. Vestibulum in orci quis felis tempor lacinia. Vivamus ornare ultrices facilisis. Ut hendrerit volutpat vulputate. Morbi condimentum venenatis augue, id porta ipsum vulputate in. Curabitur luctus tempus justo. Vestibulum risus lectus, adipiscing nec condimentum quis, condimentum nec nisl. Aliquam dictum sagittis velit sed iaculis. Morbi tristique augue sit amet nulla pulvinar id facilisis ligula mollis. Nam elit libero, tincidunt ut aliquam at, molestie in quam. Aenean rhoncus vehicula hendrerit.



## Capítulo 6

# Conclusiones

### 6.1. Main Section 1

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aliquam ultricies lacinia euismod. Nam tempus risus in dolor rhoncus in interdum enim tincidunt. Donec vel nunc neque. In condimentum ullamcorper quam non consequat. Fusce sagittis tempor feugiat. Fusce magna erat, molestie eu convallis ut, tempus sed arcu. Quisque molestie, ante a tincidunt ullamcorper, sapien enim dignissim lacus, in semper nibh erat lobortis purus. Integer dapibus ligula ac risus convallis pellentesque.

#### 6.1.1. Subsection 1

Nunc posuere quam at lectus tristique eu ultrices augue venenatis. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Aliquam erat volutpat. Vivamus sodales tortor eget quam adipiscing in vulputate ante ullamcorper. Sed eros ante, lacinia et sollicitudin et, aliquam sit amet augue. In hac habitasse platea dictumst.

#### 6.1.2. Subsection 2

Morbi rutrum odio eget arcu adipiscing sodales. Aenean et purus a est pulvinar pellentesque. Cras in elit neque, quis varius elit. Phasellus fringilla, nibh eu tempus venenatis, dolor elit posuere quam, quis adipiscing urna leo nec orci. Sed nec nulla auctor odio aliquet consequat. Ut nec nulla in ante ullamcorper aliquam at sed dolor. Phasellus fermentum magna in augue gravida cursus. Cras sed pretium lorem. Pellentesque eget ornare odio. Proin accumsan, massa viverra cursus pharetra, ipsum nisi lobortis velit, a malesuada dolor lorem eu neque.

### 6.2. Main Section 2

Sed ullamcorper quam eu nisl interdum at interdum enim egestas. Aliquam placerat justo sed lectus lobortis ut porta nisl porttitor. Vestibulum mi dolor, lacinia molestie gravida at, tempus vitae ligula. Donec eget quam sapien, in viverra eros. Donec pellentesque justo a massa fringilla non vestibulum metus vestibulum. Vestibulum in orci quis felis tempor lacinia. Vivamus ornare ultrices facilisis. Ut hendrerit volutpat vulputate. Morbi condimentum venenatis augue, id porta ipsum vulputate in. Curabitur luctus tempus justo. Vestibulum risus lectus, adipiscing nec condimentum quis, condimentum nec nisl. Aliquam dictum sagittis velit sed iaculis. Morbi tristique augue sit amet nulla pulvinar id facilisis ligula mollis. Nam elit libero, tincidunt ut aliquam at, molestie in quam. Aenean rhoncus vehicula hendrerit.



# Bibliografía

Health Data. Healthdata.gov, 2019. URL <https://healthdata.gov/>. (última visita 09/03/2019).

State of California. Chemicals in cosmetics, 2019. URL <https://healthdata.gov/dataset/chemicals-cosmetics>. (última visita 09/03/2019).