



**MÁSTER EN ESTADÍSTICA APLICADA PARA LA CIENCIA
DE DATOS CON R SOFTWARE**

Trabajo de Fin de Máster

**ANÁLISIS DE LA COMPOSICIÓN EN FILTROS ORGÁNICOS E
INORGÁNICOS DE LAS FORMULACIONES DE
PROTECTORES SOLARES**

By

GUSTAVO PITTALUGA

DIRECTOR: Juan Luis López Garrancho

FECHA: 16/07/2023

ENTIDAD COLABORADORA: Máxima Formación S.L.

Resumen

En los últimos años se ha cuestionado el uso de las sustancias químicas conocidas como filtros solares utilizados en la formulación de los protectores solares. Existen aspectos de regulación, ambientales, políticos y de seguridad que aportan a esta discusión.

Existe literatura que aborda los efectos ambientales de estas sustancias indicando que son bioacumulativos, pseudopersistentes y posiblemente tóxicos para los ecosistemas acuáticos, ambientales e incluso el uso humano. (6)(7)(8)(9).

De estos hechos han surgido nuevas normativas, regulaciones e investigaciones que tratan de definir cómo será el panorama de uso de estos ingredientes en los productos terminados para la protección solar.

En la actualidad no existe una base de información que permita ver a los profesionales químicos o farmacéuticos, a los equivalentes de comercialización o marketing, o al público en general cómo ha evolucionado el uso de los filtros solares como constituyentes de los productos de protección solar. (1)

En este trabajo se expondrán cómo han cambiado su uso mundialmente, por regiones, y que distintos usos individuales y combinados se ha realizado con dichas sustancias para proveer a la sociedad de estos importantes productos para el cuidado de su salud. (1)

Palabras claves: protector solar, fps, filtros solares, oxibenzona, octocrileno, octinoxato, 4-MBC.

Agradecimientos

- A la empresa MINTEL (www.mintel.com) por haberme facilitado la información para la realización de este trabajo.
- A los tutores/profesores que me han acompañado en todo el camino formativo con su cercanía, conocimiento y consejo.
- A mi familia como motor fundamental de apoyo y propósito.

Contents

1 OBJETIVOS	5
2 INTRODUCCIÓN	5
3 MATERIALES Y MÉTODOS	13
4 DESARROLLO	15
5 RESULTADOS	20
6 DISCUSIÓN	23
7 CONCLUSIONES	24
8 REFERENCIAS/BIBLIOGRAFÍA	24
9 ANEXOS	25

1 OBJETIVOS

Describir la evolución del uso de los filtros solares en los protectores solares a nivel general y de los cuestionados en particular en las diferentes regiones.

Analizar si se han variado las características de las combinaciones usadas en productos los productos terminados.

Evaluar si existe alguna relación entre los cambios que se realizan y los precios de las formulaciones.

2 INTRODUCCIÓN

La luz solar visible cubre longitudes de onda entre 400 y 780 nm, lo que representa una pequeña porción del espectro total de radiación emitida por el sol. El llamado espectro infrarrojo corresponde a longitudes de onda por encima de los 780 nm y el ultravioleta se define como el que tiene longitudes de onda por debajo de los 400 nm.

La sobreexposición a los rayos invisibles del sol puede ser dañina para la piel humana.

El daño que produce la radiación solar puede ser:

- Inmediato, con efectos visibles, como eritema y quemaduras solares, lo que conduce a la degradación de células y tejidos.
- A largo plazo, debido al efecto acumulativo de la exposición prolongada, y que da lugar a cambios degenerativos en la piel, como arrugas prematuras y cáncer de piel.

Muchos de los cambios en la piel comúnmente asociados con el envejecimiento son en realidad el resultado del daño por exposición excesiva al sol, conocido como foto envejecimiento.

Radiación ultravioleta (UV)

En términos de protección solar, los rayos UV se dividen en UVA y UVB. Se sabe desde hace mucho tiempo que la radiación UVB (rayos ultravioletas con longitudes de onda entre 315 y 280 nm) es la causa principal de las quemaduras solares agudas y del bronceado. Por lo tanto, los efectos de los filtros solares están dirigidos principalmente contra los rayos UVB. La exposición a la luz UVB también es inmunosupresora, mutagénica y cancerígena. Los rayos UVA (longitud de onda 400-315 nm) provocan reacciones fotobiológicas como el bronceado instantáneo, el bronceado y la

melanogénesis. Aunque los rayos UVA no son la causa principal de las quemaduras solares, se ha demostrado que contribuyen a los cambios degenerativos a largo plazo en la piel, lo que provoca un daño significativo en el tejido conectivo (envejecimiento prematuro de la piel) y el riesgo de desarrollar cáncer.

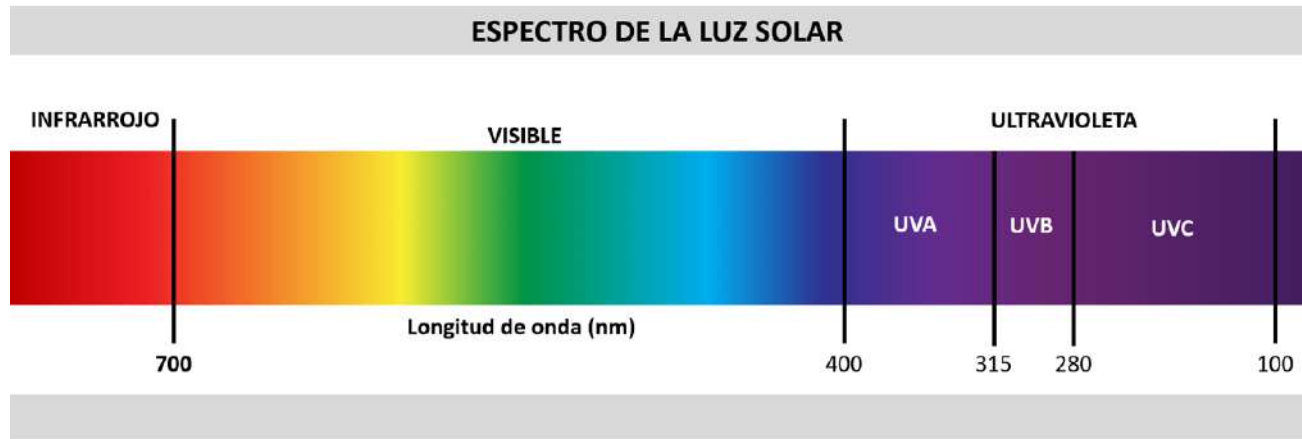


Figura 1: Espectro de la Luz Solar, **Fuente:** Autor

Protector Solar

Los productos fotoprotectores son indispensables para una «práctica responsable» de las exposiciones de la piel a la acción del sol, protegiéndola de los efectos perjudiciales de las diversas radiaciones solares. (2)

La protección contra los efectos de la radiación UV en la piel es proporcionada por moléculas especialmente diseñadas, filtros UV, incorporados en formulaciones de protección solar. Los protectores solares no solo deben proteger contra las quemaduras solares, sino también contra los efectos de los rayos UVA. Los protectores solares no solo contienen químicos que absorben, reflejan y dispersan los rayos UV, sino que además, incluyen ingredientes que previenen los efectos secundarios que producen mediadores dañinos, especies reactivas de oxígeno (llamados radicales libres) dentro de la piel, inflamación, estimulación fotoquímica de varias moléculas y efectos secundarios duraderos que pueden producir las lesiones degenerativas de la piel.

Los efectos según la distinta radiación recibida los podemos ver en la siguiente tabla:

TIPO		LONGITUD ONDA	ACCIÓN	CAPACIDAD DE PENETRACIÓN	EFFECTOS
Ultravioleta	UVC	200-290 nm	Biológica	Epidermis	Carcinogénico
	UVB	290-320 nm		Epidermis	Eritema inmediato Fotoinmunosupresión Bronceado Cáncer no melanoma
	UVA	320-400 nm		Niveles profundos de la dermis	Fotoenvejecimiento Sensibilización Melanoma
Visible		400-800 nm	Luminosa	Hipodermis	Fototoxicidad Fotoalergia
Infrarroja		800-1.800 nm	Calorífica	Hipodermis	Vasodilatación Aumento temperatura Eritema Deshidratación

Tabla 1: Efectos de la Radiación Solar, **Fuente:** (3)

Filtro Solar

Los filtros UV utilizados en las formulaciones cosméticas de protección solar para proteger contra la radiación UV se pueden dividir en dos grupos:

- La materia orgánica tiene la propiedad de absorber la luz ultravioleta.
- Las partículas que reflejan, dispersan o absorben la luz ultravioleta pueden ser inorgánicas (óxidos metálicos) u orgánicas (moléculas de polímero).

Los filtros UV absorben la radiación UV, convirtiéndola en estados electrónicos excitados, desde los cuales la energía se disipa en calor a través de vibraciones moleculares (moléculas orgánicas) y colisiones con las partículas circundantes. Sin embargo, si la energía absorbida no se puede disipar de manera eficiente, el filtro UV dejará de funcionar y puede perder su estructura química. Como resultado, su protector solar puede ser menos efectivo.

Tipos de filtros solares

Según su modo de acción, los filtros solares se clasifican en: físicos, químicos y biológicos.

Físicos

Reflejan la radiación solar. Son de amplio espectro; retienen las radiaciones solares de la zona del ultravioleta y también las del visible y las del infrarrojo. En ocasiones se les denomina «ecran» (pantalla), y se utilizan para evitar tanto el eritema como el bronceado. Una innovación en su desarrollo reside en el empleo de pigmentos micronizados, con un tamaño de partícula entre 10 y 50 nm, formulados en una base adecuada (crema, loción). Su mecanismo de acción no ocasiona perturbación cosmética alguna (no blanquean la piel).

Debido a sus características y mecanismo de acción, los filtros solares físicos se recomiendan en los siguientes casos:

- Niños menores de año y medio.
- Pielles alérgicas y reactivas.
- Dermatitis.
- Intolerancia a los filtros químicos.
- Pielles lesionadas, donde haya riesgo de absorción del filtro.
- Presencia de cicatrices.
- Presencia de eritema.
- Después de tratamientos con láser o realización de peelings químicos.

Químicos

Son moléculas orgánicas con grupos cromóforos en la región UV de estructura electrónica resonante. Actúan por absorción de la radiación solar ultravioleta (captan energía incidente y se emite nuevamente como radiación térmica, inocua para la piel). En función de la longitud de onda absorbida se distinguen entre los filtros UVB, UVA y de amplio espectro. Requieren del orden de 30 minutos para ejercer esta acción, por lo que deben emplearse con la debida antelación antes de la exposición solar. No suelen presentar problemas de formulación, por lo que son los más utilizados. Dado que pueden absorberse por la piel, hay riesgo de intolerancia, por lo que no se recomienda su aplicación en niños, pielles lesionadas donde haya riesgo de absorción (cicatrices recientes, quemaduras, luego de tratamientos con láser o peeling químico), pielles intolerantes, etc. (2)

En la tabla siguiente se exponen las estructuras químicas y la de los respectivos grupos cromóforos de los filtros solares químicos:

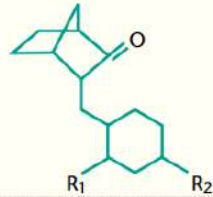
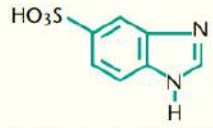
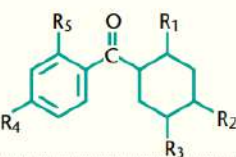
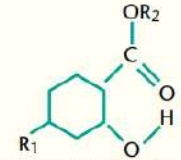
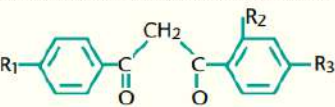
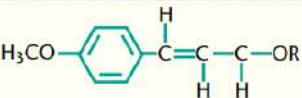
Filtros químicos	Grupo cromóforo
Acrilatos	
Derivados de los benzimidazoles	
Derivados de las benzofenonas	
Derivados de los salicilatos	
Derivados del dibenzoilmetano	
Ésteres cinamatos	

Tabla 2: Grupos químicos cromóforos, **Fuente:** (2)

Entre los filtros químicos se encuentran los orgánicos, que tienen las siguientes características:

- Son pigmentos de origen orgánico e insoluble, y constituyen una nueva generación de filtros solares.
- Mecanismo de acción: combinación de los dos tipos antes indicados, lo que les confiere un gran interés desde la óptica del formulador (protección UVA/UVB, buena cosmética y mínima penetración en la piel [$< 1\%$]).

Biológicos

Son antioxidantes que evitan la formación de radicales libres y, por lo tanto, potencian el subsistema inmunológico cutáneo. Se emplean cada vez con mayor frecuencia, y las vitaminas A (betacarotenos), C y E son las más utilizadas (en forma de acetato o palmitato).

Investigaciones recientes han demostrado que ambas vitaminas, aplicadas por vía tópica, tienen propiedades antirradicales libres y actúan contra el envejecimiento cutáneo y los cánceres de piel fotoinducidos que en numerosas ocasiones son consecuencia, a largo plazo, de una exposición prolongada al sol.

La combinación de filtros químicos UVA/UVB y antioxidantes da lugar a preparados de amplio espectro de protección solar y elevada eficacia. (2).

Como requisitos requeridos para un filtro solar, se encuentra la seguridad (debe ser atóxico, no comedogénico ni alergizante, y estable frente a agentes externos, luz (fotoestabilidad), calor (termolabilidad), pH cutáneo, etc.), la eficacia (debe tener un coeficiente de absorción activo en la zona eritemática y/o del resto del espectro solar UVA/UVB.), Otros factores que pueden influir en su eficacia son el pH, el sistema disolvente empleado, el espesor de la película residual sobre la piel y la estabilidad del producto en el que se vehiculice durante el tiempo de su reposición.

Por otro lado, no olvidemos las relacionadas con la versatilidad en su uso. (proporcionar propiedades cosméticas agradables, no manchar y permitir su formulación en diferentes tipos de excipientes, etc.)

A continuación, se incluye la lista de los filtros solares más utilizados y que abordaremos en este estudio:

NOMBRE INCI	NOMBRE USAN	ABREVIATURA
FILTROS DE AMPLIO ESPECTRO / UVA I		
Bis-Ethylhexyloxyphenol Methoxyphenyl Triazine	Bemotrizinol	BEMT
Butyl Methoxydibenzoylmethane	Avobenzone	BMDBM
Diethylamino Hydroxybenzoyl Hexyl Benzoate	No posee	DHHB
Methylene Bis-Benzotriazolyl Tetramethylbutylphenol	Bisotrizole	MBBT
Terephthalylidene Dicamphor Sulfonic Acid	Ecamsule	TDSA
drometholizole trisiloxane	drometrizole	DMT
Zinc Oxide	Zinc Oxide	ZnO
FILTROS UVB / UVA II		
4-Methylbenzylidene Camphor	Enzacamene	MBC
Benzophenone-3	Oxybenzone	BP3
Benzophenone-4	Sulisobenzone	BP4
Diethylhexyl Butamido Triazone	Iscotrizinol	DBT
Ethylhexyl Methoxycinnamate	Octinoxate	EHMC
Ethylhexyl Salicylate	Octisalate	EHS
Ethylhexyl Triazone	Octyltriazone	EHT
Ethylhexyl dimethyl PABA	Padimate O	EHDP
Homomenthyl Salicylate	Homosalate	HMS
Isoamyl p-Methoxycinnamate	Amiloxate	IMC
Octocrylene	Octocrylene	OCR
PEG-25 PABA	No posee	PABA
Phenylbenzimidazol Sulfonic Acid	Ensulizole	PBSA
Polysilicone-15	No posee	PS15
Titanium Dioxide	Titanium Dioxide aq	TiO2

Tabla 3: Listado de Filtros Solares, **Fuente:** BASF / Autor

Formulación de protector solares

Los filtros UV son los ingredientes funcionales de los protectores solares y la formulación es el proceso de combinarlos en una forma cosmética estable y funcional. El objetivo al formular un protector solar es tener un factor de protección solar (FPS) definido dentro del espectro de longitud de onda, efectivo en términos de protección de los rayos solares y un producto sostenible, es decir, resistente al agua y al frote. La compatibilidad con la piel, la sensación de uso y el precio también son importantes.

Ingredientes de protección solar

En las formulaciones de protección solar, la selección de filtros UV es esencial para lograr el nivel de eficacia UV objetivo, es decir, para identificar el sistema de protección solar más eficaz. Además del filtro solar, el resto de los ingredientes de la fórmula, ayudan a crear un producto cosmético eficaz,

cómodo y estable. Los más importantes se encuentran en los siguientes grupos.

- Los emolientes pueden reducir la sequedad, disolver filtros orgánicos cristalinos y dispersar filtros físicos.
- Los emulsionantes que pueden emulsionar filtros orgánicos y emolientes con la fase acuosa reduciendo la pegajosidad. Hay muchos emulsionantes naturales adecuados para los protectores solares.
- Los modificadores de reología aseguran la textura deseada y la estabilidad dimensional de los cosméticos.
- Los agentes formadores de película que prolongan la resistencia del medio de contraste al agua y al roce.

La combinación de estos ingredientes crea una tecnología que mejora la eficacia y realza la belleza y continúa creando protectores solares innovadores durante generaciones.

Factor de Protección Solar

El FPS (factor de protección solar) es una medida de lo mucho que un protector solar le protegerá de los rayos UVB y algunos rayos UVA.

El número del SPF de un protector solar indica la cantidad de tiempo que el producto le protegerá de las quemaduras solares. Por ejemplo, si comienza a arder la piel 10 minutos después de estar al sol sin protector solar, usar un protector solar con un SPF de 30 aumentaría ese tiempo a 300 minutos (10 minutos x 30 SPF). De todos modos, es necesario que vuelva a aplicarse el protector solar cada 2 horas, incluso si tiene un FPS alto. (3)

Los productos de protección solar tienen la misma fórmula que los cosméticos convencionales (cremas, lociones, barras, etc.), pero contienen hasta un 50% de ingredientes de protección solar. Desde el punto de vista de la formulación, estos ingredientes de protección solar se pueden agrupar en cuatro categorías.

- Aceites polares como homosalato y octocrileno.
- Sólidos cristalinos solubles en aceite como avobenzona y oxibenzona.
- Una sal soluble en agua como el ensulizol (PBSA).
- Polvo/partículas insolubles (como óxido de zinc y dióxido de titanio).

Los de este último grupo suelen usarse en los protectores solares con factores de protección más altos.

Los aceites polares y los filtros solubles en aceite contienen una gran cantidad de fase grasa, que tiende a volver pegajoso el producto.

Los filtros solubles en agua limitan el rendimiento de los ingredientes esenciales en las formulaciones, como los espesantes y emulsionantes poliméricos, que afectan la estabilidad de la formulación.

Finalmente, las partículas insolubles pueden provocar el secado del producto, dando como resultado a menudo un color blanco indeseable en la piel.

Formular un protector solar que sea efectivo contra el sol, el agua y la fricción, compatible, fácil de aplicar, estable y asequible es una tarea compleja.

En el siguiente trabajo se analiza el uso los filtros orgánicos y físicos en las formulaciones de los protectores solares lanzados entre los años 1996-2022. Para ello se utilizó la base de datos de la empresa MINTEL (1) que obtiene la información directamente del producto físico en el mercado. Dicha base de datos es usada por las áreas de comercialización o marketing de las empresas para acceder a las tendencias que los diferentes productos que evalúa siguen en sus ámbitos correspondientes. Por el contrario, esta información no es utilizada por los formuladores de los productos. En este documento, se realizó una modificación en la exposición de la información para que exhiba como ha cambiado el uso de los distintos filtros solares y sus combinaciones a través del tiempo en las distintas regiones. La idea es que sirva como material de referencia ante la estrategia de formulación y lanzamiento de productos, exhiba la reacción de las regulaciones en los cambios realizados y los costos que ellos implica para las empresas y el cuidado del medio ambiente y la sociedad.

3 MATERIALES Y MÉTODOS

Para la realización de este trabajo se ha utilizado una exportación de la base de datos de la empresa MINTEL (1) la cual incluye todos los productos de protección solar lanzados al mercado mundial entre los años 1996 y 2022.

En ella se presentan los siguientes datos:

- **ID:** Número identificador único del producto.

- **PRODUCTO:** Nombre del producto comercializado y empresa (*)
- **INGREDIENTES:** Contiene una lista de los distintos compuestos químicos que forman parte del producto.
- **FPS:** Número que identifica el Factor de Protección Solar del producto.
- **PAIS:** País mercado en que se lanzó dicho producto.
- **REGION:** Región mercado en que se lanzó dicho producto.
- **FECHA:** Fecha en la cual se lanzó dicho producto al mercado.
- **PRECIO_EUROS:** Precio expresado en Euros del producto.
- **PRECIO_DOLLAR:** Precio expresado en dólares del producto.
- **PRECIO_100g_EUROS:** Precio expresado en euros de 100 gramos de la fórmula del producto.
- **PRECIO_100g_DOLARES:** Precio expresado en dólares de 100 gramos de la fórmula del producto.

(*) Esta información no será compartida por razones de confidencialidad.

Para poder utilizar estos datos según nuestro objetivo realizamos los siguientes pasos:

1. Corrección de los nombres de los ingredientes

Se han corregido los siguientes errores:

- Nombres pertenecientes a un filtro solar escritos de diversas formas sin seguir un estándar.
- Nombres de los filtros solares escritos con distinta expresión de las fórmulas químicas
- Nombres de los filtros solares escritos con distintas nomenclaturas. (INCI, USAN, ABR)
- Errores propios de tipeo al ingresar dichos nombres.
- Inclusión de información no relevante y textual en medio de las fórmulas químicas.

Para ello se realizó una lista de conversión para transformar a toda esa información no congruente a nomenclatura ABR (Abreviatura). (**ANEXO IV**)

Esto proceso constituyó realizar 665 cambios de texto que se aplicaron todas vez que este apareciera en los 699.271 componentes presentes en los 17.337 productos evaluados. (**ANEXO IV**)

2. Determinar presencia de un ingrediente

Se agregó a cada registro una columna por cada uno de los filtros solares y designando con un “1” su presencia y con un “0” su ausencia. (**ANEXO IV**)

3. Análisis de la Información

Luego se abordó el análisis que contesta las siguientes preguntas.

1. ¿Cómo ha evolucionado anualmente la cantidad de lanzamiento de productos solares mundialmente?
2. ¿Qué cantidad de lanzamientos de productos solares hubo en las diferentes regiones?
3. ¿Cómo ha evolucionado la cantidad de lanzamientos por región por año?
4. ¿Cuál son los filtros solares cuestionados y cuál es el filtro más crítico en ese sentido?
(ANEXO I)
5. ¿Existe alguna relación entre el Factor de Protección solar y el Precio en los productos que ya no poseen los filtros cuestionados? (ANEXO II)
6. ¿Cómo ha variado el uso de los filtros solares en las formulaciones que no poseen los filtros cuestionados? (ANEXO III)

4 DESARROLLO

De la información procesada podemos obtener en una primera instancia una respuesta a las siguientes preguntas:

1. **¿Cómo ha evolucionado anualmente la cantidad de lanzamiento de productos solares mundialmente?**

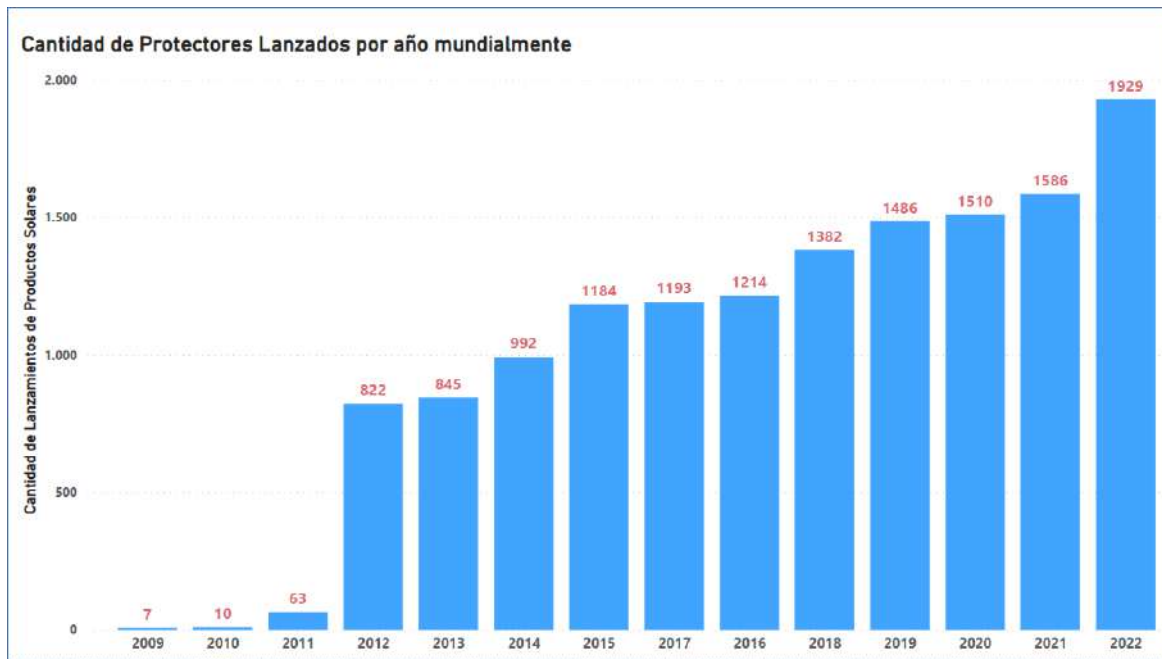


Figura 2: Cantidad de Protectores Lanzados por año en el mundo, **Fuente:** MINTEL y Autor

Es notorio que la conciencia de lo que significa el efecto de la radiación solares en la salud fue en aumento y que era un nicho al que había que atender.

2. ¿Qué cantidad de lanzamientos de productos solares hubo en las diferentes regiones?

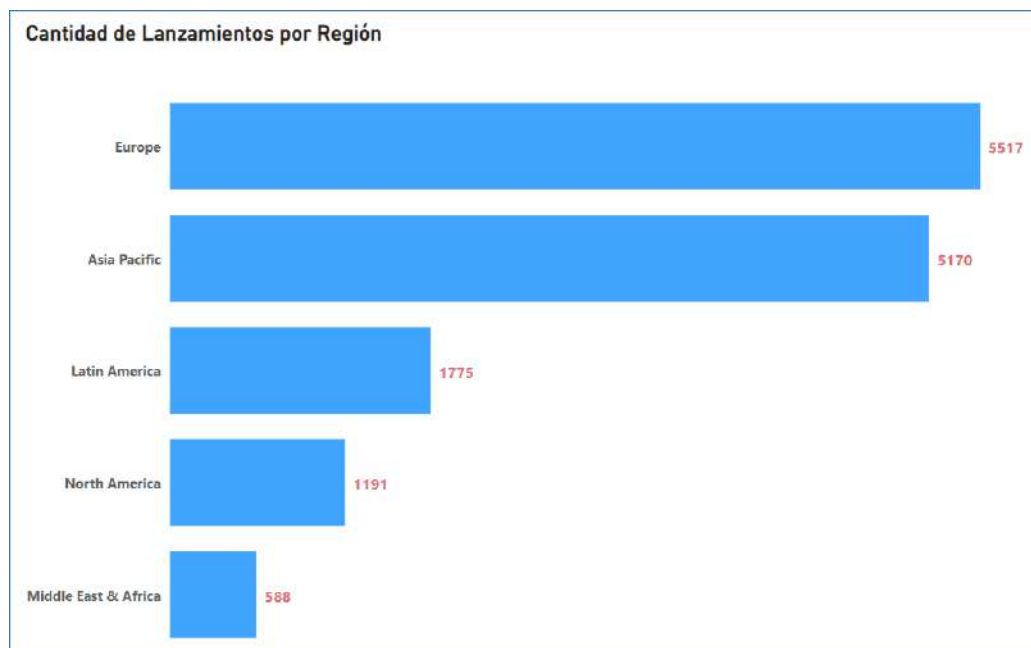


Figura 3: Cantidad de Protectores Lanzados por Región, **Fuente:** MINTEL y Autor

Se observa bastante disparidad en este contexto. En notoria la mayor cantidad de lanzamientos en **Europa y Asia Pacific**.

3. ¿Cómo ha evolucionado la cantidad de lanzamientos por región por año?



Figura 4: Cantidad de Protectores Lanzados por Región Anualmente, **Fuente:** MINTEL y Autor

Vemos, como habíamos observado en punto anterior, que la cantidad de lanzamientos de productos en **Asia Pacific y Europa** es superior a las otras regiones. Aquí también se observa que se fueron incrementando en forma simultánea.

4. ¿Cuál son los filtros solares cuestionados y cuál es el filtro más crítico en ese sentido?

Los filtros solares cuestionados a que nos referimos se exhiben en la siguiente tabla:

Filtro	FPS (UVB)	UVA	Coefficiente UVB/UVA	Eficiencia UVB
Oxibenzona	7,6	2,7	0,36	1,47
Octinoxato	7,9	1,3	0,17	1,58
Octocrileno	6,0	1,8	0,30	1,20
Enzacamene	6,6	1,3	0,19	1,65
DHMB	3,4	16,4	4,81	0,68
Avobenzona	4,5	6,1	1,35	0,90
Bemotrizinol	12,0	8,5	0,71	2,40
Bisotrizole	7,5	8,8	1,17	1,50

Tabla 4: Listado de Filtros Solares, **Fuente:** Autor

Observar las siguientes particularidades:

- Los cuestionados están marcados en color.
- La capacidad que poseen de otorgar Factor de Protección Solar (UVB) es diversa entre ellos.
- La capacidad que poseen de otorgar protección solar UVB y UVA es más diferente aún (desviación estándar de 2.6 en UVB y 5.3 en UVA).
- Los protectores solares en Europa y Mercosur deben tener como mínimo una relación entre la protección UBA y UVA de 33%. (Se deben combinar en el producto final para poder brindar un FPS alto, ejemplo 50 teniendo en cuenta que debe cumplir el porcentaje mencionado).

El tabla siguiente, pueden verse más claramente como se ordenan los filtros según su capacidad de protección en las distintas partes del espectro lumínico.

Filtro	FPS (UVB)	Filtro	UVA
Bemotrizinol	12,0	DHNB	16,4
Octinoxato	7,9	Bisotrizole	8,8
Oxibenzona	7,6	Bemotrizinol	8,5
Bisotrizole	7,5	Avobenzona	6,1
Enzacamene	6,6	Oxibenzona	2,7
Octocrileno	6,0	Octocrileno	1,8
Avobenzona	4,5	Octinoxato	1,3
DHNB	3,4	Enzacamene	1,3

Tabla 5: Listado de Filtros Solares y capacidad UVN/UVA, **Fuente:** Autor

El problema surge en que el filtro más utilizado (y barato) que brinda protección UVA como para que la relación entre UVB y UVA se cumpla es la Avobenzona. Pero si bien, este no es uno de los cuestionados es el disparador de los inconvenientes para realizar el cambio requerido.

Veamos más en detalle. A continuación, se muestran los espectros de absorción de los filtros cuestionados:

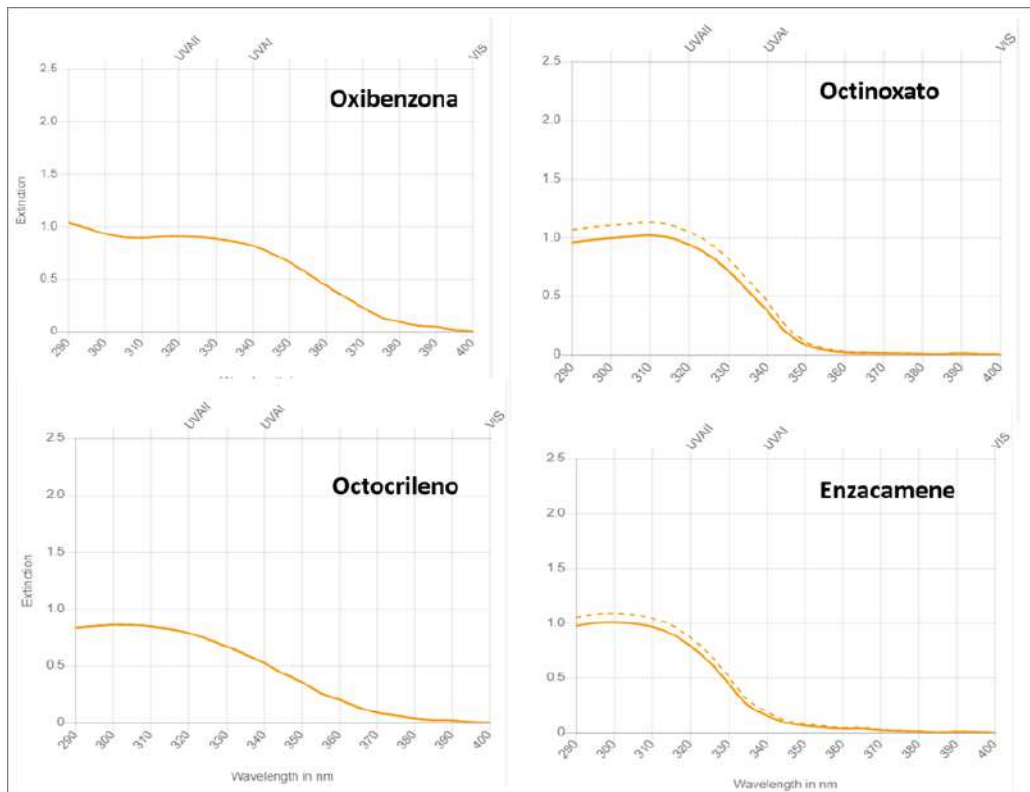


Figura 5: Espectros de Absorción, Fuente: BASF y Autor

Ninguno de ellos presenta una alta protección en ámbito UVA (a la derecha). En la figura de debajo, a la izquierda se puede ver el de la Avobenzona.

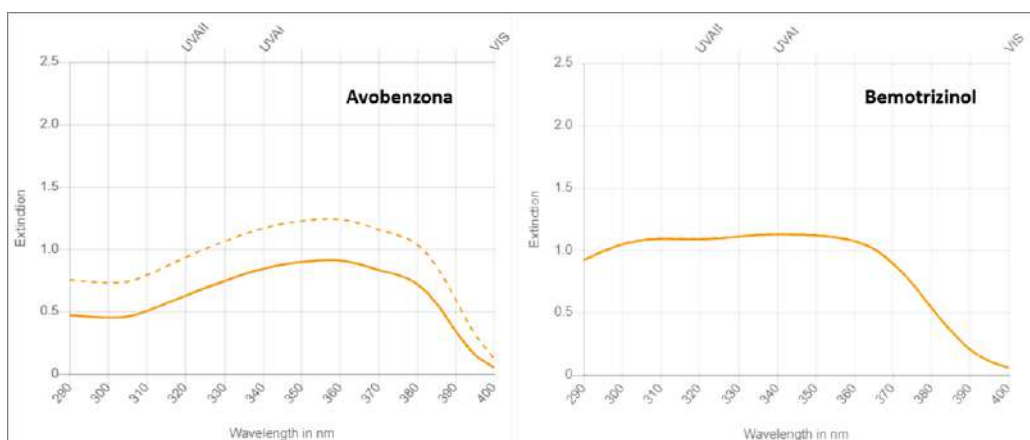


Figura 6: Espectros de Absorción, Fuente: BASF y Autor

Allí, no sólo se nota que posee una diferente protección en esa región UVA que los anteriores mostrados, sino que además, se observa otro detalle. Existe una línea punteada que es la protección

que posee antes de exponerse al sol y la llena, como queda después. Es decir, la Avobenzona, no es fotoestable. Se descompone al exponerse a la luz solar.

Para que esta descomposición no ocurra, se suelen filtros como el Octocrileno. El que haya que reemplazarlo, hace que no se pueda usar Avobenzona y reemplazarla por filtros como el Bemotrizinol (el espectro de la derecha, que puede verse que es fotoestable y con buena protección en el UVA).

Esta situación hace que las fórmulas de los protectores tengan que cambiar drásticamente. Que se dejen de usar los cuatro filtros cuestionados y la Avobenzona (por perder su soporte para no descomponerse) y que sean reemplazada (o parcialmente reemplazada) por otros que son más caros y difíciles de formular para obtener una buena cosmetividad. Ellos son los filtros físicos como el Óxido de Zinc y el Dióxido de Titanio, y los otros orgánicos como el Bemotrizinol y el DHHD. (ver la tabla dónde estaba mostradas sus capacidades UVB y UVA).

La segunda parte de la pregunta que nos realizábamos es si existe alguno de los filtros que sea más crítico, más cuestionado. La respuesta a ese interrogante es: Si. La Oxibenzona es el filtro que se debería reemplazar de las formulaciones al encararse estos cambios. (6)(7)(8)(9)

5 RESULTADOS

Se observa que los filtros cuestionados bajan en su uso en los últimos años y son reemplazados por las alternativas.

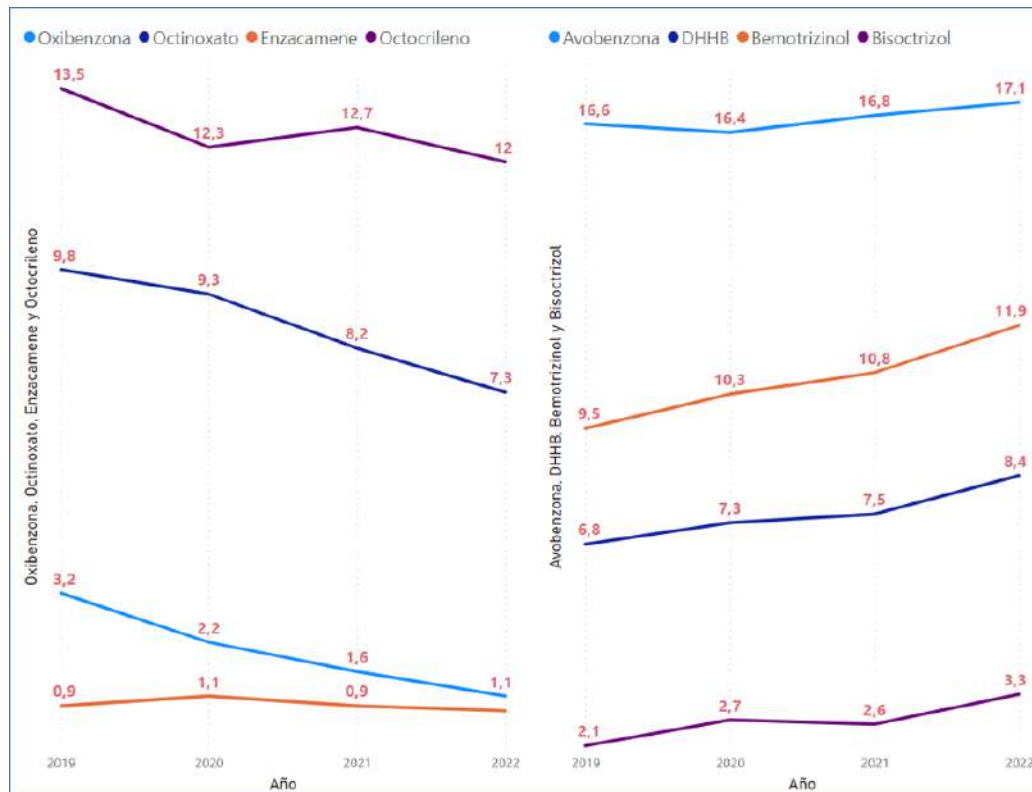


Figura 7: Cambio de uso de los filtros por año, **Fuente:** MINTEL y Autor

Existen diferencias significativas en el uso de la Oxibenzona (BP3) en las distintas regiones analizadas. (ANEXO II)

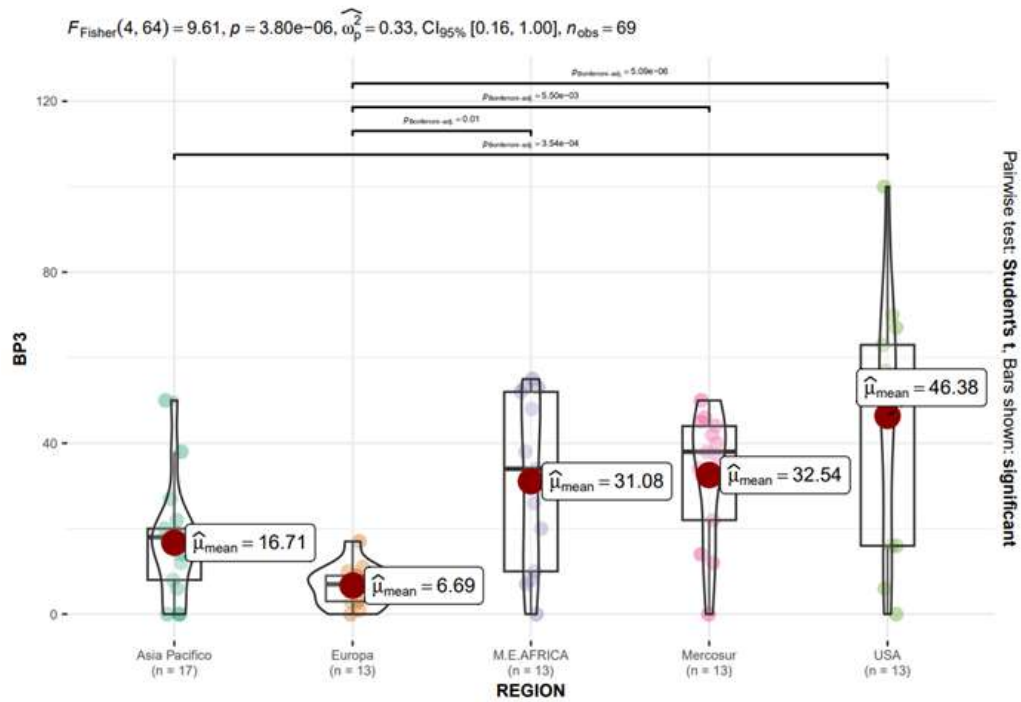


Figura 8: Uso de Oxibenzona según las Regiones, **Fuente:** MINTEL y Autor

La relación entre los precios y los Factores de Protección Solar (FPS) -con el uso de los nuevos filtros que se utilizan para reemplazar los cuestionados- muestra una relación lineal (**ANEXO II**).

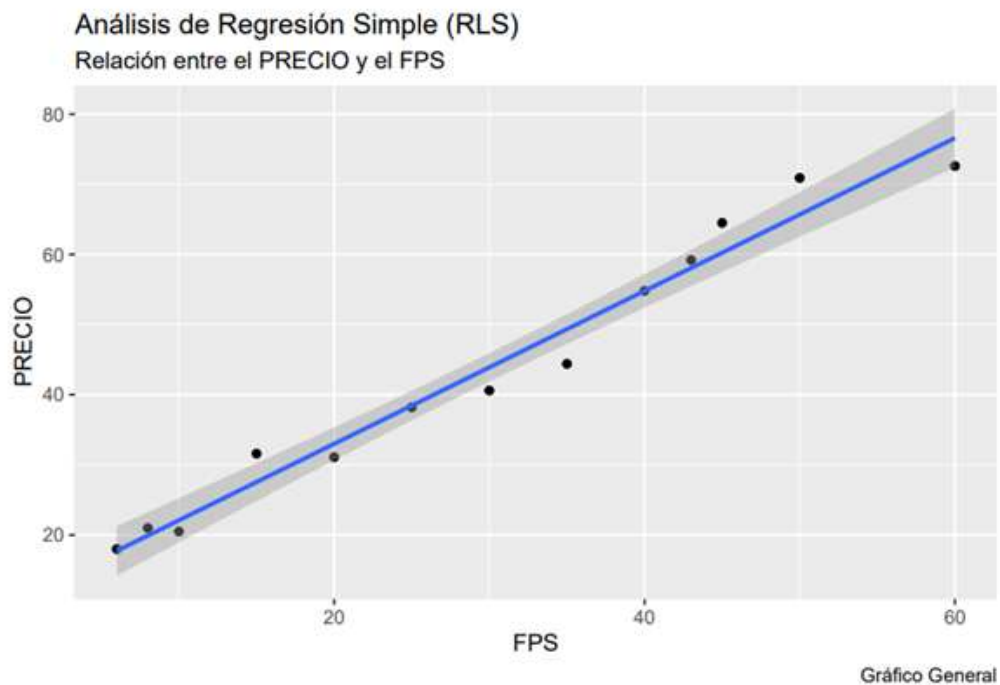


Figura 9: Análisis FPS vs PRECIO, **Fuente:** MINTEL y Autor

La combinación de los filtros antes y después de los reemplazos presenta distintas características de agrupación. (de uso simultáneo). (ANEXO III)

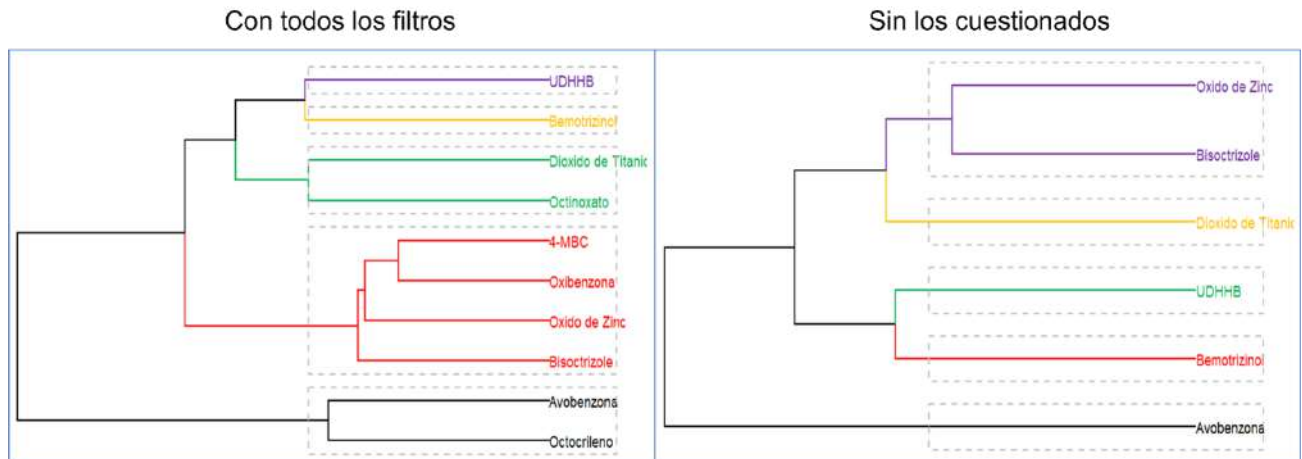


Figura 10: Análisis de Clúster, Dendrograma, **Fuente:** MINTEL y Autor

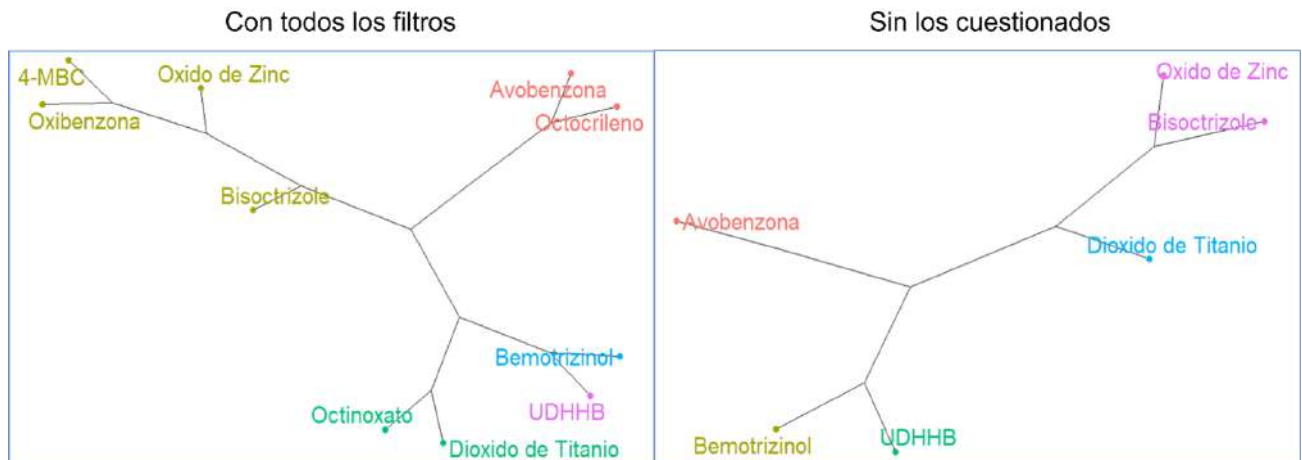


Figura 11: Análisis de Clúster, Filogénico, **Fuente:** MINTEL y Autor

6 DISCUSIÓN

Si bien existen cuestionamientos en algunos de los filtros utilizados aún en las entidades regulatorias se los considera como aptos para su uso. Al no haber una prohibición taxativa de no uso

mundialmente (si los hay en algunos países y regiones) todavía existen formulaciones que los contienen y el declive de su uso como vimos es paulatino.

Se están realizando estudios en varios frentes relacionados con los efectos de estos filtros solares en las personas y en los sistemas marinos. Los resultados futuros nos indicarán que será posible utilizar y como se deberán ajustar dichas formulaciones.

7 CONCLUSIONES

Se ha realizado el análisis y evolución del uso de filtros solares utilizando una base de datos, MINTEL que suele ser usada con propósito de evaluación de tendencia de mercado y no de cambio de uso de ingredientes en los productos. Lo realizado aquí es una importante contribución a los profesionales farmacéuticos y químicos que pueden acceder a importante información de la evolución de las formulaciones y que no es fácil de obtener. (**ANEXO IV**).

Las dificultades de creación de los productos solares en los contextos que aquí expresamos se hacen cada vez más complejas y en algunos casos hasta imposibles. Poder ofrecer una alta protección solar, un equilibrio en el espectro UVA/UBA, con productos que no dejen la piel blanca al aplicárselos, que no manchen la ropa, que sean cosméticamente agradables y que se ofrezcan a un precio razonable, será el gran desafío para enfrentar en los próximos años cuando se analice el costo/beneficio de afectar el ambiente, protegernos del cáncer de piel con respecto a los cambios que surjan de la regulación.

8 REFERENCIAS/BIBLIOGRAFÍA

- (1) Mintel - **A Global Market Intelligence & Research Agency**, <https://es.mintel.com/>
- (2) Deiry Marín, Alfonso del Pozo (2005), **Filtros solares. Características, tipos y requerimientos** (Offarm/Elsevier) <https://www.elsevier.es/es-revista-offarm-4-articulo-filtros-solares-caracteristicas-tipos-requerimientos-13079607>
- (3) Memorial Sloan Kettering Cancer Center, **Información sobre el protector solar**, <https://www.mskcc.org/es/cancer-care/patient-education/understanding-sunscreen>
- (4) Garrote, Antonieta, Bonet, Ramón (2008), **Fotoprotección. Factores de protección y filtros solares** (Offarm/Elsevier) <https://www.elsevier.es/es-revista-offarm-4-articulo-filtros-solares>

fotoproteccion-factores-proteccion-filtros-solares-13120520

- (5) The BASF Personal Care, **Sunscreen Simulator**
<https://www.personal-care.basf.com/products-formulation/sunscreen-simulator>
- (6) Federal Register USA, **Sunscreen Drug Products for Over-the-Counter Human Use** <https://www.federalregister.gov/documents/2019/02/26/2019-03019/sunscreen-drug-products-for-over-the-counter-human-use>
- (7) Journal of the American Medical Association (JAMA), **Effect of Sunscreen Application Under Maximal Use Conditions on Plasma Concentration of Sunscreen Active Ingredients**
<https://jamanetwork.com/journals/jama/fullarticle/2733085>
- (8) Food and Drug Administration (FDA), **Spotlight on CDER Science: New FDA Study Shines Light on Sunscreen Absorption (2019)**
<https://www.fda.gov/drugs/news-events-human-drugs/spotlight-cder-science-new-fda-study-shines-light-sunscreen-absorption> (2019)
- (9) CONICET Argentina / PubMed, **Dermal exposure to the UV filter benzophenone-3 during early pregnancy affects fetal growth and sex ratio of the progeny in mice (2019)** <https://pubmed.ncbi.nlm.nih.gov/32430675/>

9 ANEXOS

- **ANEXO I** – ANOVA: Evaluación de diferencias entre las concentraciones de Oxibenzona en las distintas regiones.
- **ANEXO II** - REGRESION LINEAL SIMPLE: Evaluación de la relación entre las variables Factor de Protección Solar (FPS) y Precio.
- **ANEXO III** - ANALISIS DE CLUSTER JERÁRQUICO: Evaluación de la distinta combinación de filtros solares.
- **ANEXO IV** - OBTENCIÓN, TRANSFORMACION y CARGA DE LOS DATOS (ETL): Obtención de los datos necesarios de la base de datos de MINTEL.

ANEXO I - ANOVA

Contenido

1. ANEXO I - ANOVA	2
1.1. Objetivo	2
1.2. Estructura de los datos	2
1.3. Visualización de los datos	2
1.4. Pregunta de Investigación	3
1.5. Hipótesis Estadísticas	3
1.6. Análisis de Supuestos	3
1.6.1. Estadísticos Descriptivos	3
1.6.2. Identificación de Outliers y Extremos	4
1.6.3. Evaluación de la Normalidad	5
1.6.4. Prueba de Shapiro-Wilk	5
1.6.5. Prueba de Shapiro-Wilk para cada grupo	6
1.6.6. Gráfico QQ para cada grupo	6
1.6.7. Evaluación de la Homocedasticidad	7
1.6.8. Homocedasticidad - Gráfica de Residuos versus Valores Ajustados	7
1.6.9. Homocedasticidad - Prueba de Levene	8
1.7. Realización del ANOVA de una vía de muestras independientes	8
1.8. Pruebas post-hoc (Comparaciones Múltiples)	9
1.9. Graficación de los valores y comparaciones múltiples)	10

1. ANEXO I - ANOVA

Instalación de los Paquetes y cargado de Librerías

Instalamos los paquetes y cargamos las librerías necesarias para realizar el análisis.

```
# Ingresamos los paquetes necesarios
packages <- c("tidyverse", "rstatix", "ggpubr", "ggstatsplot", "WRS2", "asbio", "emmeans",
             "faraway", "ggfortify", "tinytex")

# Instalamos los paquetes no instalados
installed_packages <- packages %in% rownames(installed.packages())
if (any(installed_packages == FALSE)) {
  install.packages(packages[!installed_packages])}

# Cargamos las librerías correspondientes
invisible(lapply(packages, library, character.only = TRUE))

rm(packages); rm(installed_packages) # Los borramos del entorno y quitamos de memoria
```

1.1. Objetivo

Evaluar si existen diferencias significativas entre las concentraciones de **BP3 (oxibenzona)** en las distintas regiones.

Carga de los datos

Cargamos los datos respectivos:

```
# Carga de los datos
rm(list=ls()) # Borrado de todas las variables del Entorno
df <- read.csv("C:/TFM_Gustavo_Pittaluga/DATA/ANEXO_I_ANOVA.csv", sep=";")
```

1.2. Estructura de los datos

Vemos la estructura de los datos.

```
# Estructura de los datos
str(df)
```

```
## 'data.frame': 69 obs. of 2 variables:
## $ BP3 : int 0 17 11 10 7 7 7 9 9 3 ...
## $ REGION: chr "Europa" "Europa" "Europa" "Europa" ...
```

Tenemos dos variables, una categórica y la otra numérica.

1.3. Visualización de los datos

```
# Visualización de los datos
head(df) # Primeros datos
```

```
## BP3 REGION
## 1 0 Europa
## 2 17 Europa
## 3 11 Europa
## 4 10 Europa
## 5 7 Europa
## 6 7 Europa
```

Como solo tenemos **1 factor explicativo** (*BP3*) y **1 respuesta numérica** (*REGION*), vamos a aplicar un **ANOVA de una vía de muestras independientes** (también llamado entre grupos EG).

1.4. Pregunta de Investigación

¿Existen diferencias entre las concentraciones de **BP3** en las distintas regiones?

1.5. Hipótesis Estadísticas

H_o : No existen diferencias.

H_a : Existen diferencias.

1.6. Análisis de Supuestos

En primer lugar, evaluaremos los siguientes supuestos de esta prueba paramétrica:

- Presencia de datos atípicos y extremos.
- Evaluación de la Normalidad
- Evaluación de la Homocedasticidad

1.6.1. Estadísticos Descriptivos

Calculamos los estadísticos descriptivos (cantidad, media y desviación estándar) de la variable **BP3** organizada por la variable **REGION**:

```
# Estadísticos descriptivos
df %>%
  group_by(REGION) %>%
  get_summary_stats(BP3,
    type = "mean_sd")
```

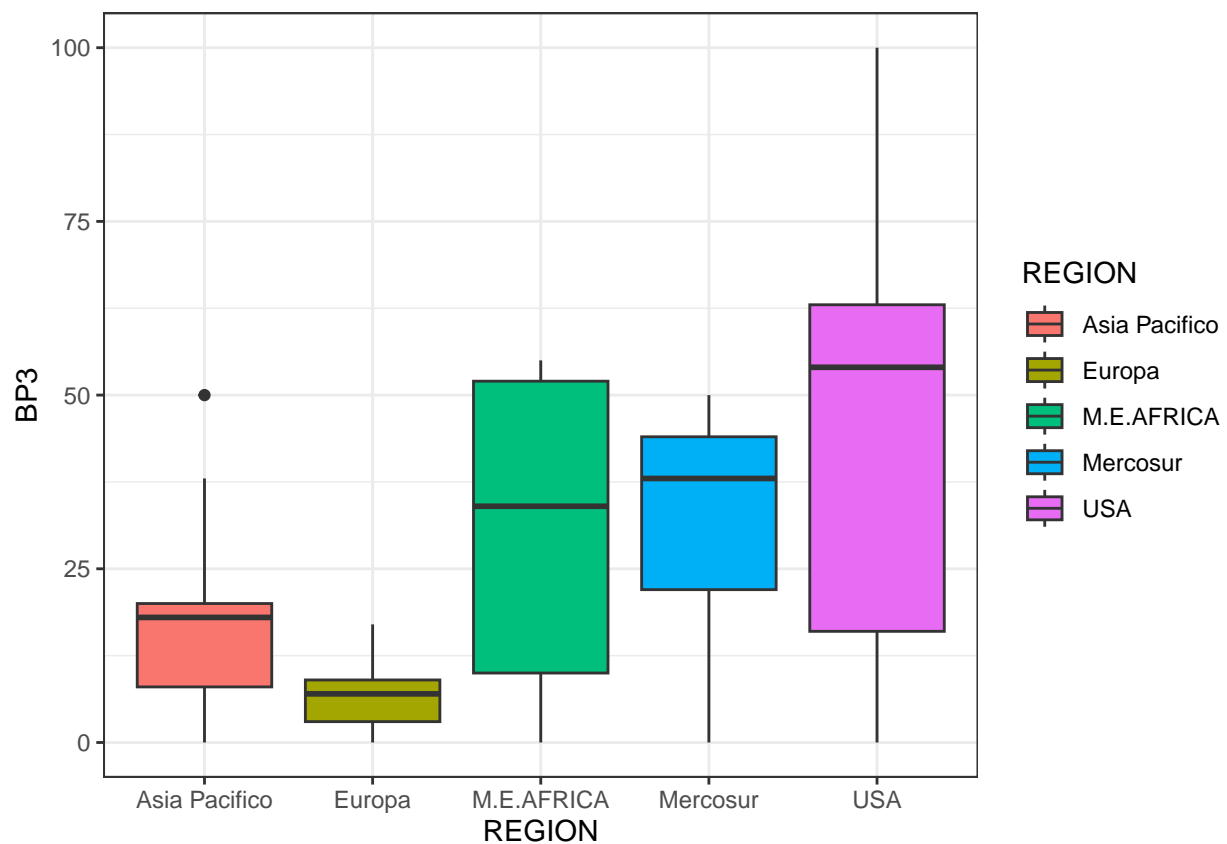
```
## # A tibble: 5 x 5
##   REGION      variable     n mean   sd
##   <chr>      <fct>   <dbl> <dbl> <dbl>
## 1 Asia Pacifico BP3      17 16.7 13.2
## 2 Europa       BP3      13  6.69 4.70
## 3 M.E.AFRICA   BP3      13 31.1 20.4
## 4 Mercosur     BP3      13 32.5 15.5
## 5 USA          BP3      13 46.4 29.0
```

Podemos ver que tanto las **medias**, como la **desviaciones estándar** difieren entre si.

Presencia de datos atípicos y extremos

Analizamos los diagramas de caja (boxplot) para visualizar si se observan datos atípicos y extremos.

```
# Graficación de datos atípicos y extremos
ggplot(data = df, aes(x = REGION, y = BP3)) +
  geom_boxplot(aes(fill = REGION), width = 0.8) + theme_bw()
```



Observamos que existe un dato atípico an **Asia Pacífico**. Evaluamos si es extremo.

1.6.2. Identificación de Outliers y Extremos

Y realizamos el cálculo con las función `identify_outliers` del paquete `rstatix`

```
# Análisis de datos atípicos y extremos
df %>%
  group_by(REGION) %>%
  identify_outliers(BP3)

## # A tibble: 1 x 4
##   REGION      BP3 is.outlier is.extreme
##   <chr>      <int> <lgl>      <lgl>
## 1 Asia Pacifico    50 TRUE      FALSE
```

Vemos que no es extremo. No eliminaremos el dato.

1.6.3. Evaluación de la Normalidad

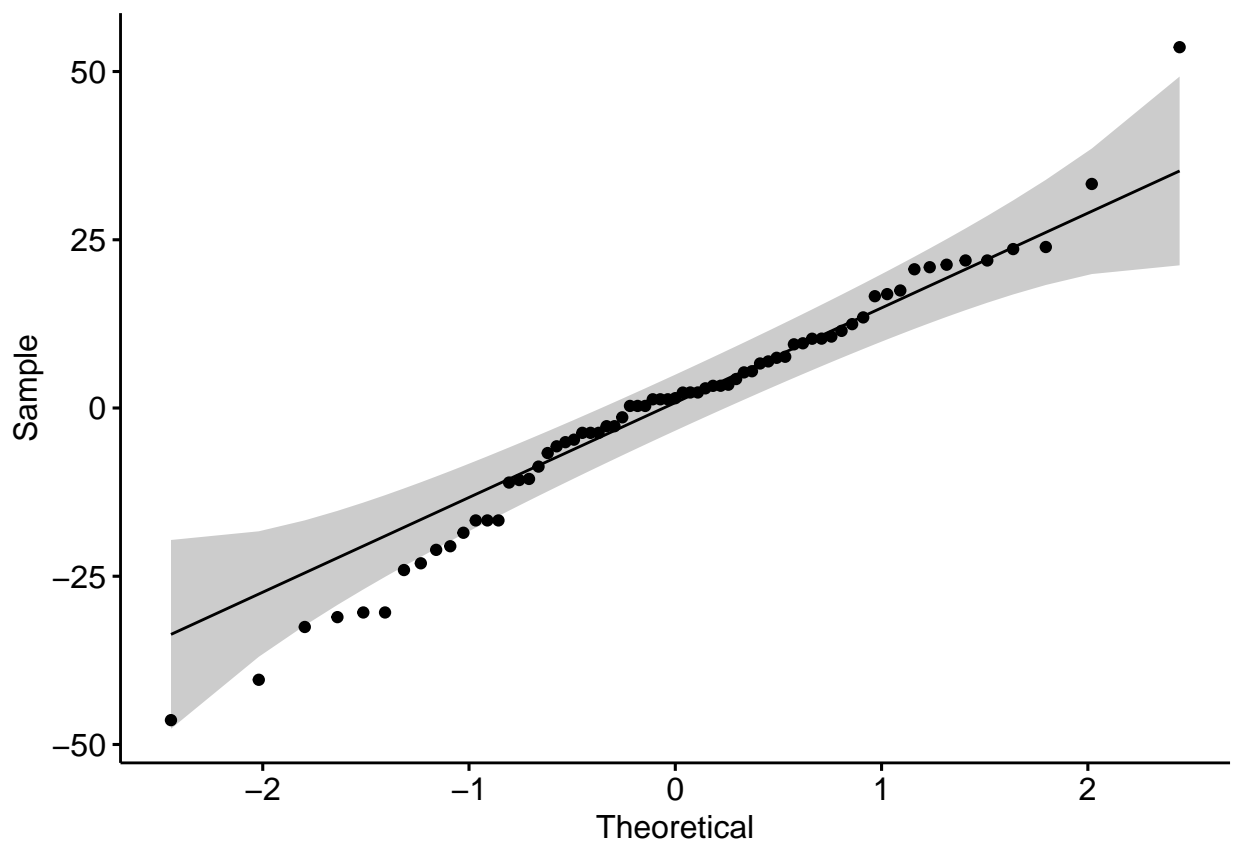
Haremos la comprobación del supuesto de normalidad mediante dos enfoques siguientes:

En el primer caso, analizaremos los residuos del ANOVA para comprobar la normalidad de todos los grupos juntos. Es un método más sencillo y útil cuando se tienen muchos grupos o si hay pocos puntos de datos por grupo. En este caso tenemos 5 grupos y pocos datos por cada grupo.

Se utiliza entonces, el **gráfico QQ** y la **prueba de normalidad de Shapiro-Wilk**.

El **gráfico QQ** dibuja la correlación entre un dato dado y la distribución normal.

```
# Evaluación de la Normalidad
fit <- lm(BP3 ~ REGION, data = df)
ggqqplot(residuals(fit))
```



En la **gráfica QQ**, los puntos caen aproximadamente a lo largo de la línea de referencia, por lo cual **podemos asumir la normalidad**. Como hay algunos fuera en los extremos realizamos como comprobación la Prueba de Shapiro-Wilk.

1.6.4. Prueba de Shapiro-Wilk

```
# Prueba de Shapiro-Wilk
shapiro_test(residuals(fit))
```

```
## # A tibble: 1 x 3
##   variable      statistic p.value
##   <chr>         <dbl>   <dbl>
## 1 residuals(fit) 0.974   0.167
```

Como puede verse el **p-valor** es **no significativo** ($p > 0.05$, p-valor: 0.1670543).

En consecuencia, como habíamos asumido, corroboramos **que se cumple el supuesto de normalidad** de los residuos.

Como puede verse el **p-valor** es no significativo ($p > 0.05$, p-valor: 0.1670543). En consecuencia, como habíamos asumido, corroboramos **que se cumple el supuesto de normalidad** de los residuos.

En el segundo caso, podemos verificar la normalidad para cada grupo por separado.

Este enfoque se puede utilizar cuando solo tiene unos pocos grupos y muchos puntos de datos por grupo. Si bien, no es nuestro caso, ya que no tenemos muchos datos por grupo, adicionaremos esta manera de cálculo.

Aquí se calcula la **prueba de Shapiro-Wilk** para cada nivel de grupo y el **gráfico QQ** dibuja la correlación entre un dato dado y la distribución normal.

1.6.5. Prueba de Shapiro-Wilk para cada grupo

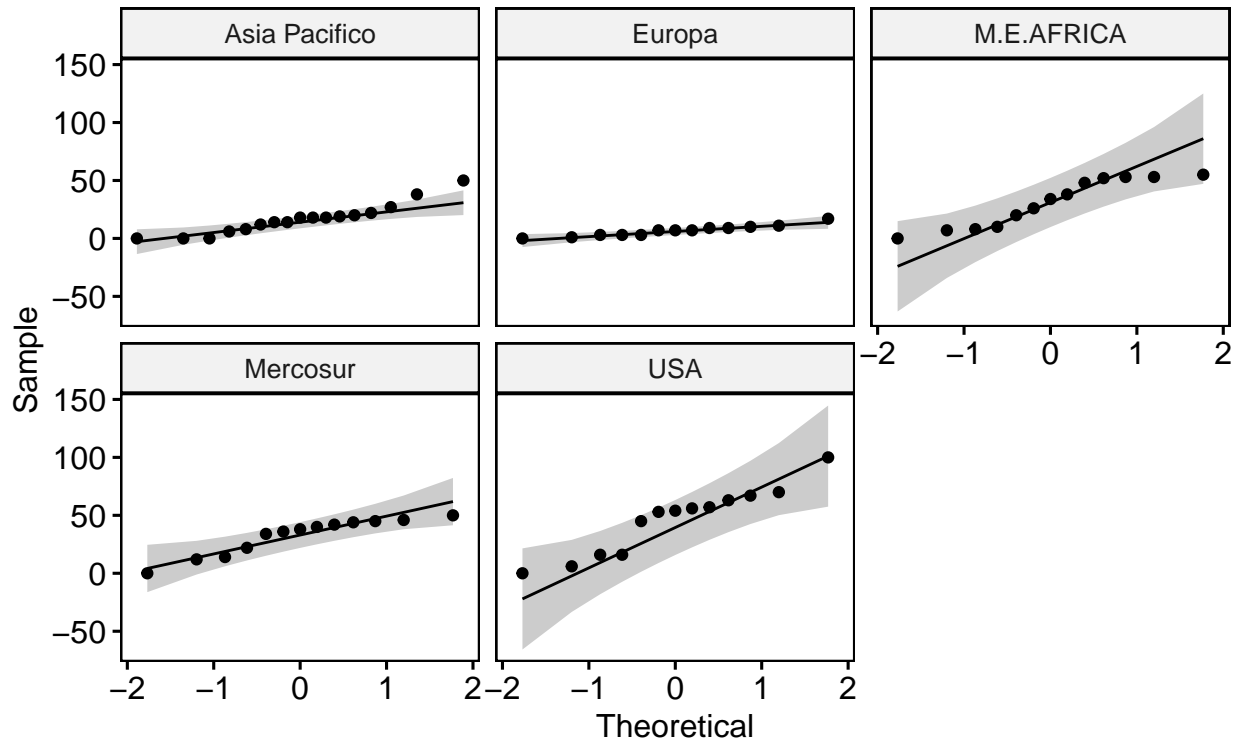
```
# Normalidad
df %>%
  group_by(REGION) %>%
  shapiro_test(BP3)
```

```
## # A tibble: 5 x 4
##   REGION      variable statistic      p
##   <chr>      <chr>         <dbl>   <dbl>
## 1 Asia Pacifico BP3           0.912 0.106
## 2 Europa       BP3           0.944 0.507
## 3 M.E.AFRICA   BP3           0.888 0.0905
## 4 Mercosur     BP3           0.878 0.0661
## 5 USA          BP3           0.928 0.320
```

En todos los casos se observa el cumplimiento del supuesto.

1.6.6. Gráfico QQ para cada grupo

```
# Gráfico QQ para cada grupo
ggqqplot(df, "BP3", facet.by = "REGION")
```



Tanto en los **gráficos QQ** como en el cálculo según la **prueba de Shapiro-Wilk**, los datos para cada grupo se **distribuyen normalmente** ($p > 0.05$).

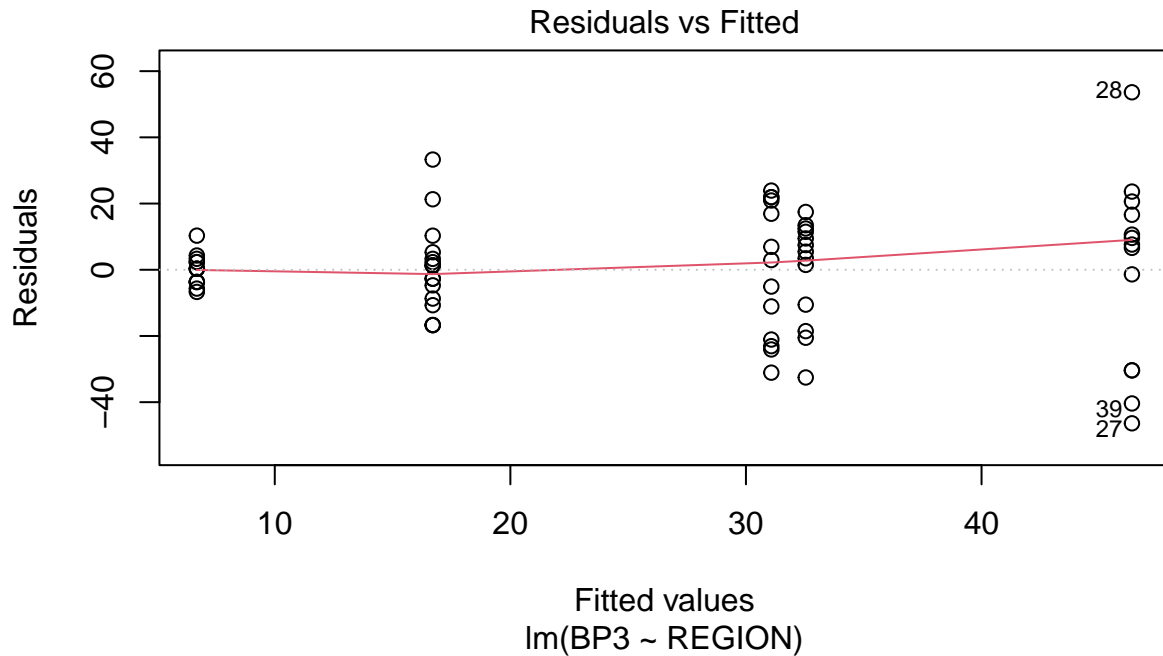
1.6.7. Evaluación de la Homocedasticidad

Utilizaremos la visualización de la **gráfica de residuos versus valores ajustados** y la **Prueba de Levene** para evaluar la **homocedasticidad**.

1.6.8. Homocedasticidad - Gráfica de Residuos versus Valores Ajustados

A continuación mostramos la gráfica

```
# Homocedasticidad - Gráfica de Residuos versus Valores Ajustados
plot(fit, 1)
```



1.6.9. Homocedasticidad - Prueba de Levene

A continuación realizamos el cálculo:

```
# Homocedasticidad - Prueba de Levene
df %>%
  levene_test(BP3 ~ REGION)
```

```
## # A tibble: 1 x 4
##   df1 df2 statistic      p
##   <int> <int>   <dbl>   <dbl>
## 1     4    64     4.55 0.00269
```

Según lo observado en el **gráfico entre los residuos y los valores ajustados** (la media de cada grupo) y el resultado de la **Prueba de Levene** concluimos que **se cumple el supuesto de homogeneidad de varianza entre los grupos**. ($p < 0.05$)

1.7. Realización del ANOVA de una vía de muestras independientes

Generamos el modelo:

```
# ANOVA de una vía de muestras independientes
df %>%
  anova_test(BP3 ~ REGION)
```

```
## ANOVA Table (type II tests)
```

```
##
## Effect DFn DFd F p p<.05 ges
## 1 REGION 4 64 9.605 3.8e-06 * 0.375
```

Se observa un efecto significativo de la **REGION**- sobre el **BP3** (ges= 0.375).

Este valor significa que se explica que un **37.5%** del cambio en la variable *BP3* puede explicarse por un cambio en la variable *REGION*.

Vemos además que tenemos un **p < 0.05** (3.8×10^{-6}) lo que indica que **existen diferencias significativas entre los grupos** (diferentes regiones).

1.8. Pruebas post-hoc (Comparaciones Múltiples)

Teniendo en cuenta que el **ANOVA de una vía de muestras independientes** ha sido significativo, nos interesa evaluar entre que grupos (REGION) se ha observado diferencias en el efecto de la variable BP3.

Para ello podemos realizar a **Pruebas post-hoc de Tukey** que realiza múltiples comparaciones por pares entre los grupos.

```
# Pruebas post-hoc (Comparaciones Múltiples)
tukey_hsd(fit)

## # A tibble: 10 x 9
##   term group1 group2 null.value estimate conf.low conf.high p.adj
## * <chr> <chr> <chr> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 REGION Asia Pacifico Europa 0 -10.0 -28.7 8.72 5.66e-1
## 2 REGION Asia Pacifico M.E.AFRI~ 0 14.4 -4.36 33.1 2.11e-1
## 3 REGION Asia Pacifico Mercosur 0 15.8 -2.90 34.6 1.36e-1
## 4 REGION Asia Pacifico USA 0 29.7 10.9 48.4 3.33e-4
## 5 REGION Europa M.E.AFRI~ 0 24.4 4.44 44.3 9.01e-3
## 6 REGION Europa Mercosur 0 25.8 5.90 45.8 4.83e-3
## 7 REGION Europa USA 0 39.7 19.7 59.6 4.97e-6
## 8 REGION M.E.AFRICA Mercosur 0 1.46 -18.5 21.4 1 e+0
## 9 REGION M.E.AFRICA USA 0 15.3 -4.64 35.3 2.1 e-1
## 10 REGION Mercosur USA 0 13.8 -6.10 33.8 3.03e-1
## # i 1 more variable: p.adj.signif <chr>
```

También es posible realizar las pruebas post-hot mediante:

```
df %>%
  pairwise_t_test(BP3 ~ REGION, p.adjust.method = "bonferroni", pool.sd=TRUE)

## # A tibble: 10 x 9
##   .y. group1 group2 n1 n2 p p.signif p.adj p.adj.signif
## * <chr> <chr> <chr> <int> <int> <dbl> <chr> <dbl> <chr>
## 1 BP3 Asia Pacifico Europa 17 13 1.38e-1 ns 1 e+0 ns
## 2 BP3 Asia Pacifico M.E.AF~ 17 13 3.51e-2 * 3.51e-1 ns
## 3 BP3 Europa M.E.AF~ 13 13 1.05e-3 ** 1.05e-2 *
## 4 BP3 Asia Pacifico Mercos~ 17 13 2.07e-2 * 2.07e-1 ns
## 5 BP3 Europa Mercos~ 13 13 5.5 e-4 *** 5.5 e-3 **
## 6 BP3 M.E.AFRICA Mercos~ 13 13 8.38e-1 ns 1 e+0 ns
## 7 BP3 Asia Pacifico USA 17 13 3.54e-5 **** 3.54e-4 ***
```

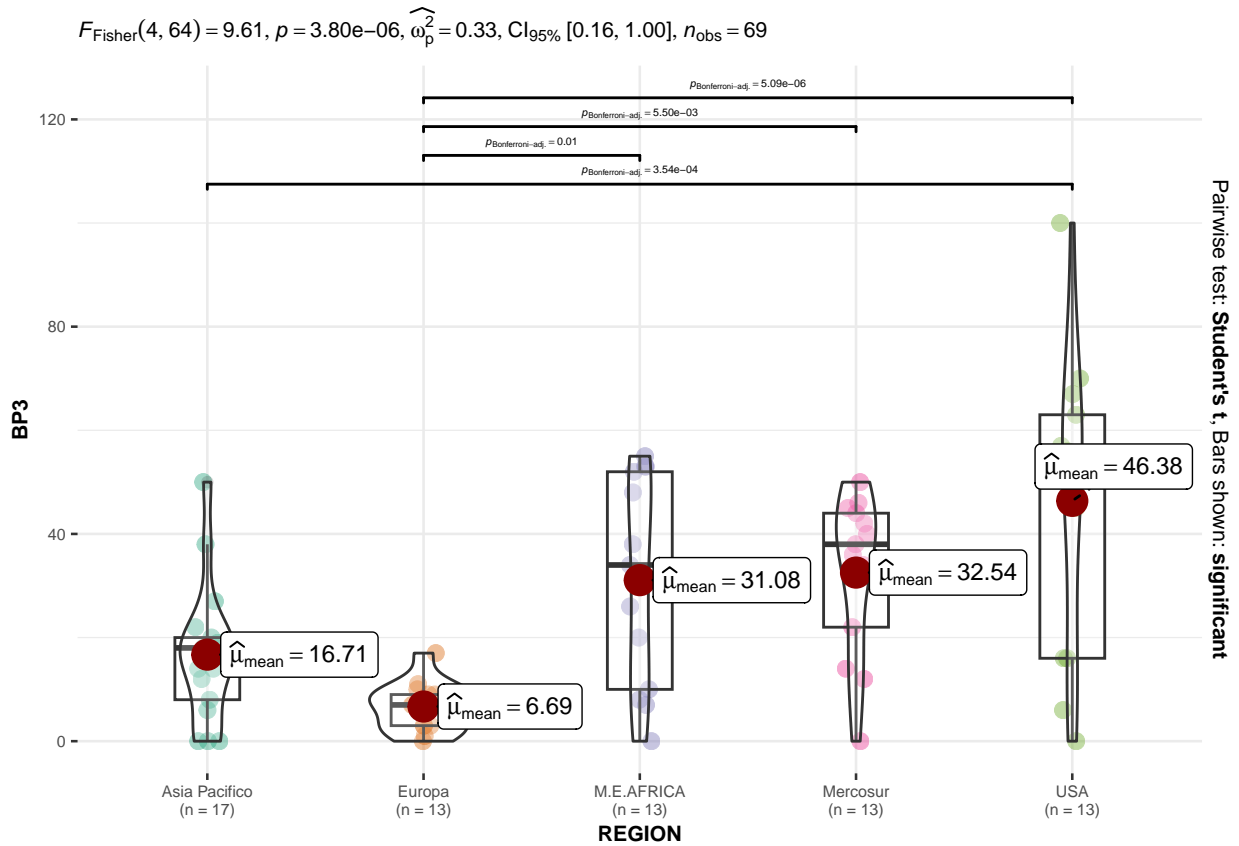
##	8	BP3	Europa	USA	13	13	5.09e-7	****	5.09e-6	****
##	9	BP3	M.E.AFRICA	USA	13	13	3.5 e-2	*	3.5 e-1	ns
##	10	BP3	Mercosur	USA	13	13	5.57e-2	ns	5.57e-1	ns

En ambos casos, podemos apreciar que no hay diferencias significativas entre las dietas las regiones de *Asia Pacífico con Europa, con M.E.Africa o con Mercosur*, ni tampoco lo hay entre *USA y M.E.Africa y el Mercosur*

Las demás poseen diferencias significativas de diferentes magnitudes (observar los valores p.adj y los “**”).

1.9. Graficación de los valores y comparaciones múltiples)

```
ggbetweenstats(x = REGION, y = BP3, data = df, p.adjust.method = "bonferroni",
  bf.message = FALSE, var.equal = TRUE,
  ggsignif.args = list(textsize = 1.5, tip_length = 0.01)) +
  theme(text = element_text(size = 8), plot.subtitle = element_text(size=8) )
```



En el gráfico anterior podemos ver resumido lo que habíamos ya calculado.

Un ANOVA significativo ($p < 0.05$), un tamaño de efecto que explica el 33% de la variación de la variación de la **BP3** en función de las diferentes **REGIONES** y la significatividad entre las regiones explicada anteriormente.

ANEXO II - REGRESION LINEAL SIMPLE

Contenido

2. ANEXO II - REGRESION LINEAL SIMPLE	2
2.1. Objetivo	2
2.2. Instalación de los Paquetes y cargado de Librerías	2
2.3. Carga de los datos	2
2.4. Análisis exploratorio y visualización de la relación entre las variables.	2
2.4.1. Dimensión de los datos	2
2.4.2. Vista del encabezado de los datos	3
2.4.3. Vista de la estructura de los datos	3
2.4.4. Vista de Bloxplot	3
2.4.5. Gráfico exploratorio de los datos	4
2.5. Realización del modelo de regresión simple (RLS)	5
2.6. Interpretación la bondad de ajuste del modelo.	6
2.7. Calcular la tasa de error del modelo.	6
2.7.1. Calculo del error estándar residual (RSE)	6
2.7.2. Calculamos la tasa de error	6
2.7.3. Interpretamos el RSE y la Tasa de Error	6
2.8. Interpretar los coeficientes del modelo.	6
2.9. Calculo de los intervalos de confianza para los coeficientes.	7
2.10. Evaluación de los supuestos del modelo.	7
2.10.1. Linealidad	8
2.10.2. Normalidad	8
2.10.3. Homosedasticidad	8
2.11. Graficación de la Regresión	8

2. ANEXO II - REGRESION LINEAL SIMPLE

Utilizando los datos que incluyen los promedios de los precios de los productos que poseen los filtros que reemplazan a los cuestionados, evaluamos la linealidad entre estos y el Factor de Protección Solar (FPS)

2.1. Objetivo

Confirmar que existe una relación lineal por Regresión Lineal Simple (RLS) entre ambas variables **PRECIO** y **FPS**

2.2. Instalación de los Paquetes y cargado de Librerías

Instalamos los paquetes y cargamos las librerías necesarias para realizar el análisis.

```
# Ingresamos los paquetes necesarios
packages <- c("ggplot2", "GGally", "corrplot", "ppcor", "relaimpo", "MASS", "car", "lmtest")

# Instalamos los paquetes no instalados
installed_packages <- packages %in% rownames(installed.packages())
if (any(installed_packages == FALSE)) {
  install.packages(packages[!installed_packages])}

# Cargamos las librerías correspondientes
invisible(lapply(packages, library, character.only = TRUE))

rm(packages); rm(installed_packages) # Los borramos del entorno y quitamos de memoria
```

2.3. Carga de los datos

Accedemos a los datos de estudio.

```
rm(list=ls()) # Borrado de todas las variables del Entorno
df <- read.csv("C:/TFM_Gustavo_Pittaluga/DATA/ANEXO_II_REGRESION.csv")
```

2.4. Análisis exploratorio y visualización de la relación entre las variables.

```
# Vemos la dimensión de los datos
dim(df)
```

2.4.1. Dimensión de los datos

```
## [1] 13 2
```

Son dos dimensiones correspondientes a *PRECIO* y *FPS* con 13 registros.

```
# Vemos el encabezado de los datos  
head(df)
```

2.4.2. Vista del encabezado de los datos

```
##   FPS PRECIO  
## 1    6  18.0  
## 2    8  21.0  
## 3   10  20.5  
## 4   15  31.6  
## 5   20  31.1  
## 6   25  38.2
```

```
# Vemos la estructura de los datos  
str(df)
```

2.4.3. Vista de la estructura de los datos

```
## 'data.frame':   13 obs. of  2 variables:  
## $ FPS      : int  6 8 10 15 20 25 30 35 40 43 ...  
## $ PRECIO: num  18 21 20.5 31.6 31.1 38.2 40.6 44.4 54.8 59.2 ...
```

Observamos que las variables son numéricas.

Hacemos un resumen:

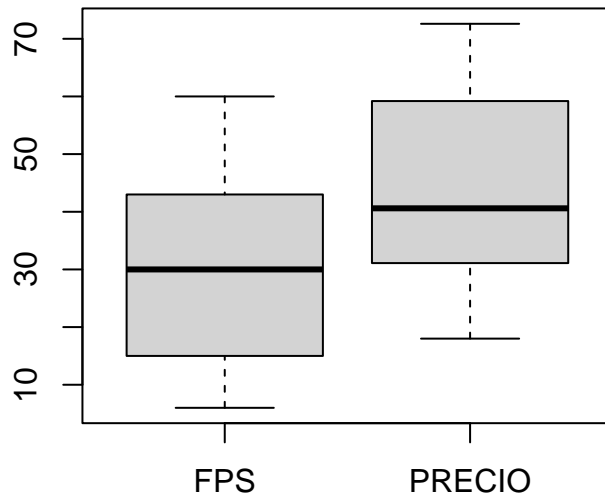
```
summary(df)
```

```
##           FPS           PRECIO  
## Min.      : 6.00   Min.      :18.00  
## 1st Qu.   :15.00   1st Qu.   :31.10  
## Median   :30.00   Median   :40.60  
## Mean     :29.77   Mean     :43.65  
## 3rd Qu.  :43.00   3rd Qu.  :59.20  
## Max.     :60.00   Max.     :72.60
```

Observamos que no hay datos faltantes (NAs).

2.4.4 Vista de Bloxplot

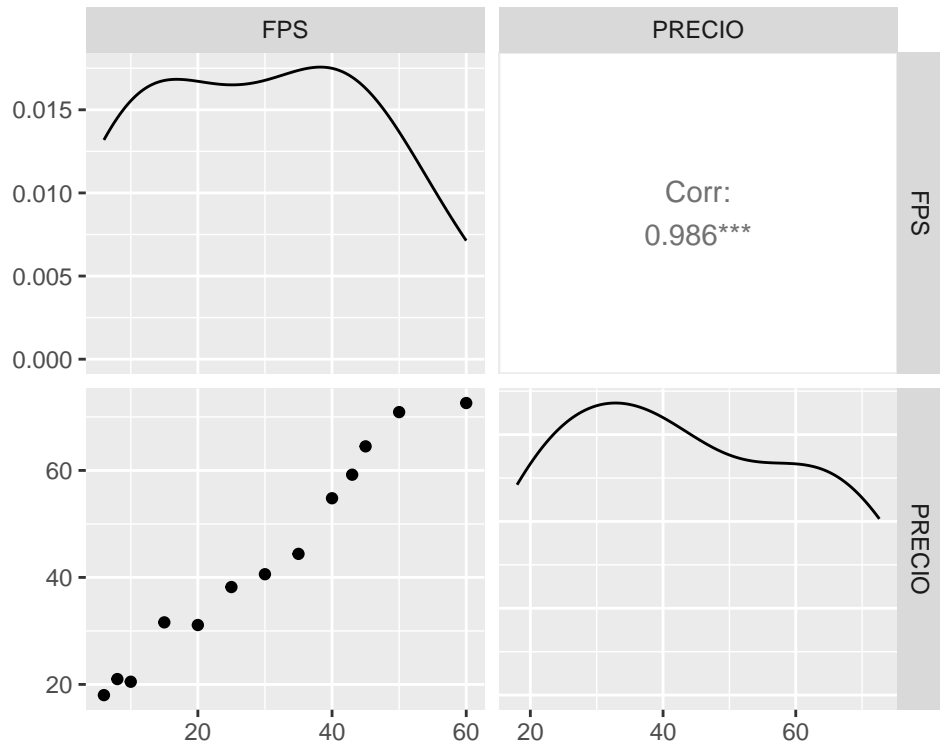
```
# Vista de Bloxplot  
attach(df)  
boxplot(df)
```



No se observan valores atípicos ni extremos.

2.4.5. Gráfico exploratorio de los datos

```
# Gráfico Exploratorio de los Datos  
# library(GGally) Fue instalada al principio.  
ggpairs(df)
```



Vemos que la distribución de puntos se asemeja a una recta y que también se presenta una alta correlación entre las variables.

2.5. Realización del modelo de regresión simple (RLS)

Realizamos un modelo de regresión simple por medio de la función `lm()` para las variables `PRECIO` que vamos a analizar en función del `FPS`.

```
# Modelo de Regresión simple
model <- lm(PRECIO ~ FPS, data=df)
summary(model)

##
## Call:
## lm(formula = PRECIO ~ FPS, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4.9465 -1.8999  0.0046  1.1353  5.2069
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  11.20439    1.88457   5.945 9.66e-05 ***
## FPS           1.08977    0.05521  19.740 6.15e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.325 on 11 degrees of freedom
## Multiple R-squared:  0.9725, Adjusted R-squared:  0.97
## F-statistic: 389.7 on 1 and 11 DF, p-value: 6.15e-10
```

2.6. Interpretación la bondad de ajuste del modelo.

Para interpretar la bondad de ajuste observamos los resultados finales del modelo:

- **Prueba de F global:** El modelo de regresión predice el *PRECIO* significativamente bien (**F(1, 11) = 389.7, p < .001**). (p-valor < 0.05)

2.7. Calcular la tasa de error del modelo.

```
# Calculo del error estándar residual (RSE)
sigma(model)
```

2.7.1. Calculo del error estándar residual (RSE)

```
## [1] 3.325393
```

```
# Calculo de la tasa de error
attach(df)
sigma(model)/mean(PRECIO)
```

2.7.2. Calculamos la tasa de error

```
## [1] 0.07618984
```

2.7.3. Interpretamos el RSE y la Tasa de Error El error estándar residual (RSE) es de **3.33** y la tasa del error es de **7.6 %**. Lo que es un error adecuado para el análisis que estamos haciendo.

2.8. Interpretar los coeficientes del modelo.

Obtenemos la tabla de coeficientes:

```
# Tabla de Coeficientes del Modelo
summary(model)$coefficients
```

```
##           Estimate Std. Error  t value    Pr(>|t|)
## (Intercept) 11.204395  1.8845656  5.945346 9.656950e-05
## FPS         1.089775  0.0552066 19.739939 6.149699e-10
```

- El **intercepto** vale **11.204395**, en este caso no tiene sentido interpretar su valor ya que es representaría el precio de un protector solar de SPF 0. (o sea, no sería un protector solar).
- El **coeficiente de regresión** para el predictor **FPS** vale **1.089775** y representa el cambio en el **PRECIO** cuando la variable **FPS** permanece constante.

En ambos casos el $p < 0.05$ lo que indica que son significativamente distintos de cero.

2.9. Cálculo de los intervalos de confianza para los coeficientes.

Obtenemos los intervalos de confianza correspondientes a un 95%

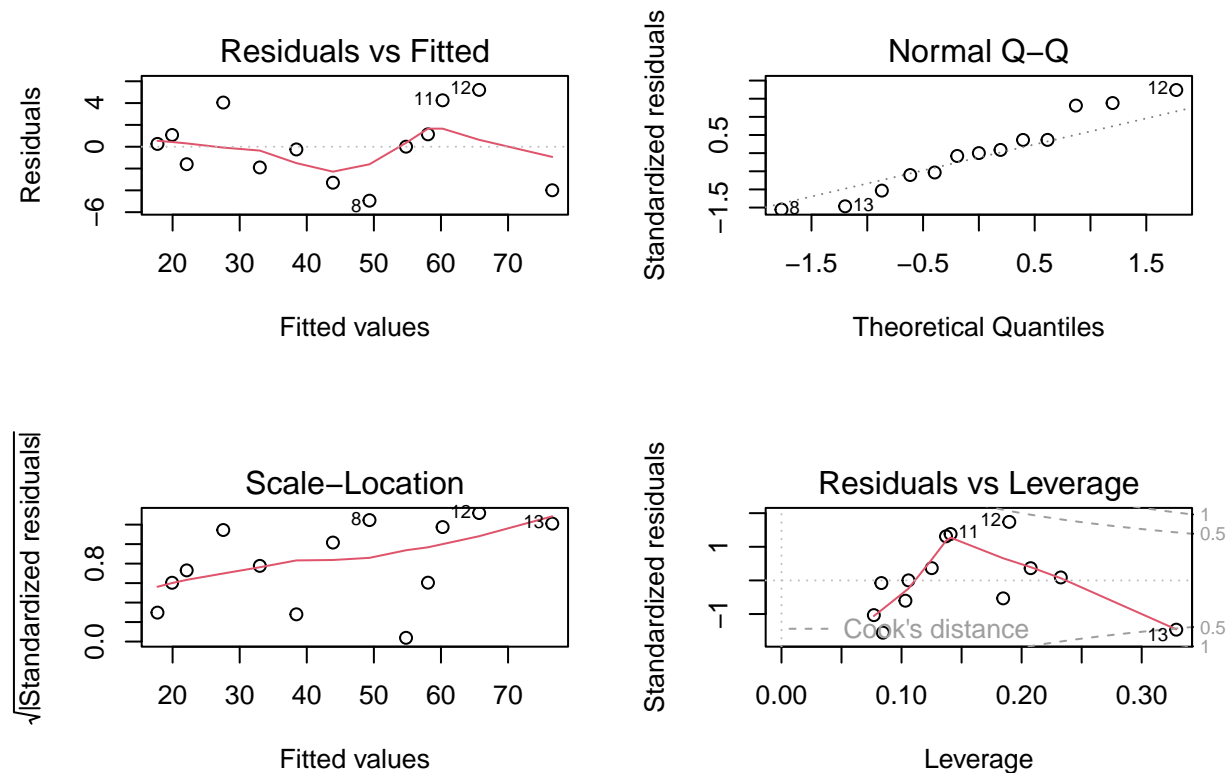
```
# "Cálculo de los intervalos de confianza para los coeficientes"  
confint(model)
```

```
##           2.5 %    97.5 %  
## (Intercept) 7.056494 15.352296  
## FPS         0.968266  1.211284
```

2.10. Evaluación de los supuestos del modelo.

Realizamos los gráficos diagnósticos del modelo.

```
# Graficos del Modelo  
par(mfrow = c(2,2))  
plot(model)
```



Los gráficos nos permiten tener una visión aproximada de lo que sucede con el modelo.

En ellos podemos observar.

1. Los residuos aumentan y disminuyen en el rango de valores ajustados (ver la línea roja) con lo cual no se cumpliría el supuesto de homogeneidad de la varianza (Homocedasticidad). (Residuals vs Fitted)

2. Los residuos muestran distribución bastante normal, se apartan las colas de la diagonal. (Normal Q-Q)
3. Similar a lo mencionado en el punto 1.
4. No se observan valor influyentes fuera de la distancia de Cook.

Estimamos oportuno realizar las comprobaciones pertinentes en forma analítica para validar nuestras visión anterior.

2.10.1. Linealidad

```
# Linealidad
mean(model$residuals)
```

```
## [1] -6.245005e-17
```

Se observa que se cumple que el promedio de los residuos debe ser cercano a cero.

2.10.2. Normalidad

Hacemos la conmbroación utilizando el **Non-constant Variance Score Test**

```
# Normalidad Shapiro-Wilk
# library(MASS) Cargada al inicio.
sresid <- studres(model)
shapiro.test(sresid)
```

```
##
## Shapiro-Wilk normality test
##
## data:  sresid
## W = 0.95932, p-value = 0.7431
```

El p-valor > 0.05 indica que la distribución de los residuos es normal.

2.10.3. Homosedasticidad

Hacemos la comprobación utilizando el **Non-constant Variance Score Test**

```
# Hacemos el Non-constant Variance Score Test
ncvTest(model)
```

```
## Non-constant Variance Score Test
## Variance formula: ~ fitted.values
## Chisquare = 1.897927, Df = 1, p = 0.16831
```

El p-valor > 0.05 indica que la distribución de los residuos es homocedástica (varianza constante).

2.11. Graficación de la Regresión

```

library(ggplot2)
ggplot(df, aes(y = PRECIO, x = FPS)) +
  geom_point() +
  stat_smooth(method="lm") +
  labs(
    x = "FPS",
    y = "PRECIO",
    title = "Análisis de Regresión Simple (RLS)",
    subtitle = "Relación entre el PRECIO y el FPS",
    caption = "Gráfico General"
  ) +
  theme_gray()

```

Análisis de Regresión Simple (RLS)

Relación entre el PRECIO y el FPS

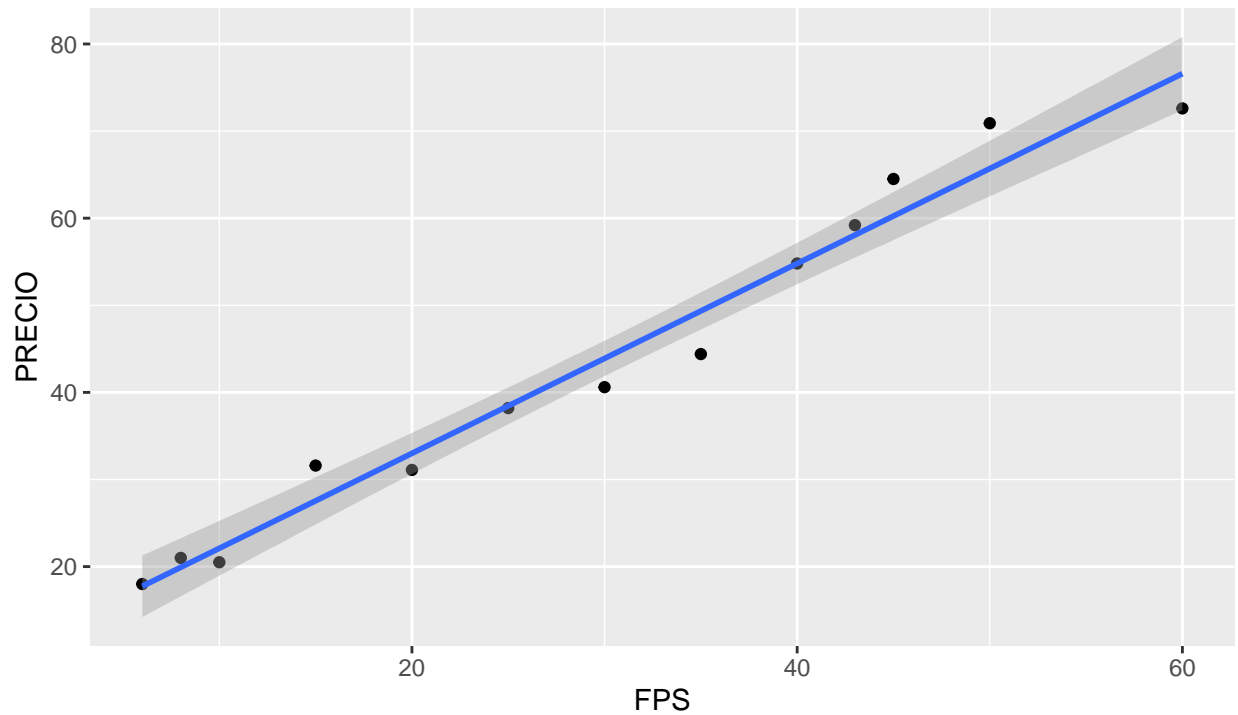


Gráfico General

ANEXO III - ANALISIS DE CLUSTER JERÁRQUICO

Contenido

3. ANALISIS DE CLUSTER JERÁRQUICO	2
3.1. Instalación de los Paquetes y cargado de Librerías	2
3.2. Objetivo	2
3.3. Carga y estandarización de los datos	2
3.4. Estructura de los datos	2
3.5. Exploración de los datos	3
3.6. Cálculo del Cluster Jerárquico	3

3. ANALISIS DE CLUSTER JERÁRQUICO

3.1. Instalación de los Paquetes y cargado de Librerías

Instalamos los paquetes y cargamos las librerías necesarias para realizar el análisis.

```
# Ingresamos los paquetes necesarios
packages <- c("tidyverse", "clustertend", "factoextra", "NbClust", "knitr", "rstatix")

# Instalamos los paquetes no instalados
installed_packages <- packages %in% rownames(installed.packages())
if (any(installed_packages == FALSE)) {
  install.packages(packages[!installed_packages])}

# Cargamos las librerías correspondientes
invisible(lapply(packages, library, character.only = TRUE))

rm(packages); rm(installed_packages) # Los borramos del entorno y quitamos de memoria
```

3.2. Objetivo

Evaluar como se combinan los distintos filtros en los Protectores Solares

3.3. Carga y estandarización de los datos

```
rm(list=ls()) # Borrado de todas las variables del Entorno

df <- read.csv("C:/TFM_Gustavo_Pittaluga/DATA/ANEXO_III_CLUSTER.csv")
origen <- df
df <- df[, c(9, 10, 11, 12, 14, 15, 18, 24, 27, 28)]
colnames(df) <- c("Oxibenzona", "Octinoxato", "Octocrileno", "4-MBC",
                 "Bemotrizinol", "Avobenzona", "UDHBB", "Bisoctrizole",
                 "Dioxido de Titanio", "Oxido de Zinc")

df_2019_2022 <- df[7732:14241,]
df_sin_filtros <- df_2019_2022[,c(-1,-2,-3,-4)]
```

3.4. Estructura de los datos

Vemos cual es la estructura de los datos:

```
# Estructura de los datos
str(df)

## 'data.frame': 14241 obs. of 10 variables:
## $ Oxibenzona : int 0 0 0 0 0 0 0 0 0 0 ...
## $ Octinoxato : int 0 0 1 1 0 1 1 1 0 0 ...
## $ Octocrileno : int 1 1 1 1 1 0 1 1 1 1 ...
## $ 4-MBC : int 1 1 0 0 1 0 0 0 0 0 ...
```

```
## $ Bemotrizinol      : int  0 0 1 0 0 0 0 0 1 0 ...
## $ Avobenzona        : int  1 1 0 1 1 0 1 0 1 1 ...
## $ UDHHB             : int  0 0 0 0 0 0 0 0 0 0 ...
## $ Bisotrizole       : int  0 0 0 0 0 0 0 0 0 0 ...
## $ Dioxido de Titanio: int  1 1 1 0 1 1 0 1 1 1 ...
## $ Oxido de Zinc     : int  0 0 0 0 0 0 0 0 0 0 ...
```

Los datos se encuentran en la manera correcta para operar sobre ellos.

3.5. Exploración de los datos

Vemos un resumen de los datos respectivos:

```
# Exploración de los datos
summary(df)
```

```
##      Oxibenzona      Octinoxato      Octocrileno      4-MBC
## Min.   :0.0000   Min.   :0.0000   Min.   :0.0000   Min.   :0.00000
## 1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:1.0000   1st Qu.:0.00000
## Median :0.0000   Median :0.0000   Median :1.0000   Median :0.00000
## Mean   :0.1616   Mean   :0.4594   Mean   :0.7701   Mean   :0.04494
## 3rd Qu.:0.0000   3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:0.00000
## Max.   :1.0000   Max.   :1.0000   Max.   :1.0000   Max.   :1.00000
##      Bemotrizinol      Avobenzona      UDHHB      Bisotrizole
## Min.   :0.0000   Min.   :0.0000   Min.   :0.0000   Min.   :0.0000
## 1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.0000
## Median :0.0000   Median :1.0000   Median :0.0000   Median :0.0000
## Mean   :0.4539   Mean   :0.6304   Mean   :0.3055   Mean   :0.1259
## 3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:0.0000
## Max.   :1.0000   Max.   :1.0000   Max.   :1.0000   Max.   :1.0000
## Dioxido de Titanio Oxido de Zinc
## Min.   :0.0000   Min.   :0.0000
## 1st Qu.:0.0000   1st Qu.:0.0000
## Median :0.0000   Median :0.0000
## Mean   :0.3842   Mean   :0.1207
## 3rd Qu.:1.0000   3rd Qu.:0.0000
## Max.   :1.0000   Max.   :1.0000
```

Vemos que no hay valores faltantes (NAs).

3.6. Cálculo del Cluster Jerárquico

En este documento aplicaremos una **metodología aglomerativa** (AGNES) de análisis jerárquico, en la cual se parte de tantos grupos como individuos y se van agrupando hasta tener a todos los individuos en un sólo grupo. Es una agrupación Botton-Up o ascendente que se utiliza para grupos pequeños.

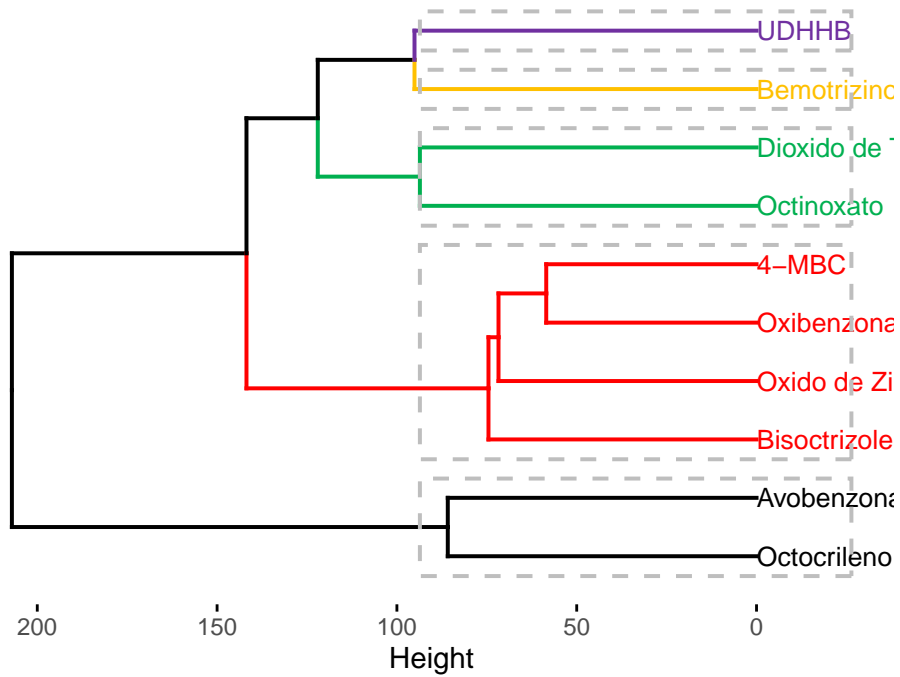
```
# Cálculo del Cluster Jerárquico - Dendograma
cluster_2019_2022 <- hcut(t(df_2019_2022),
  k = 5,
  stand = TRUE,
  hc_func = c("hclust", "agnes", "diana"),
```

```

    hc_method = "ward.D2",
    hc_metric = "euclidean",)
fviz_dend(cluster_2019_2022, rect = TRUE, cex = 0.7,
    k_colors = c("#000000", "#FF0000", "#00B050", "#FFC000", "#7030A0"), horiz = TRUE)

```

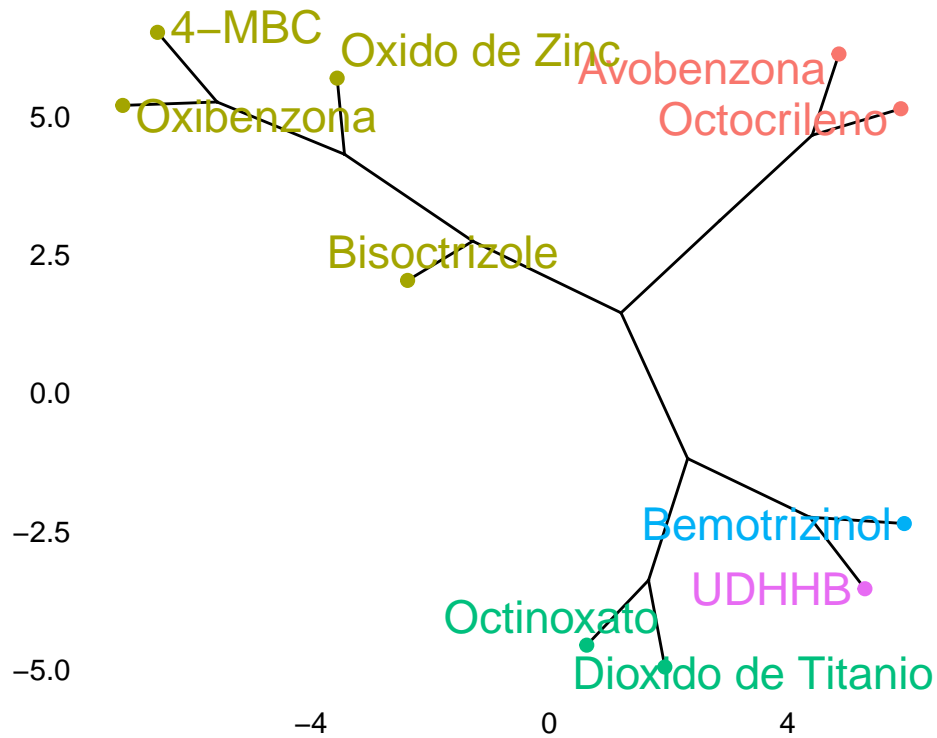
Cluster Dendrogram



```

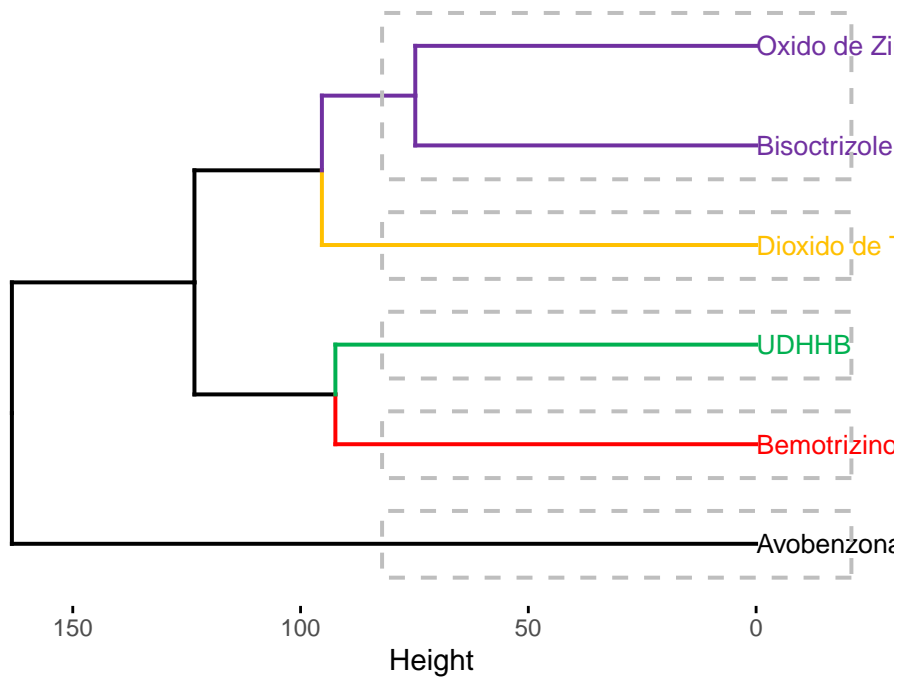
fviz_dend(cluster_2019_2022, k=3, rect = TRUE, cex=1.5, type = "phylogenetic",
    color_labels_by_k=T, repel=T)

```

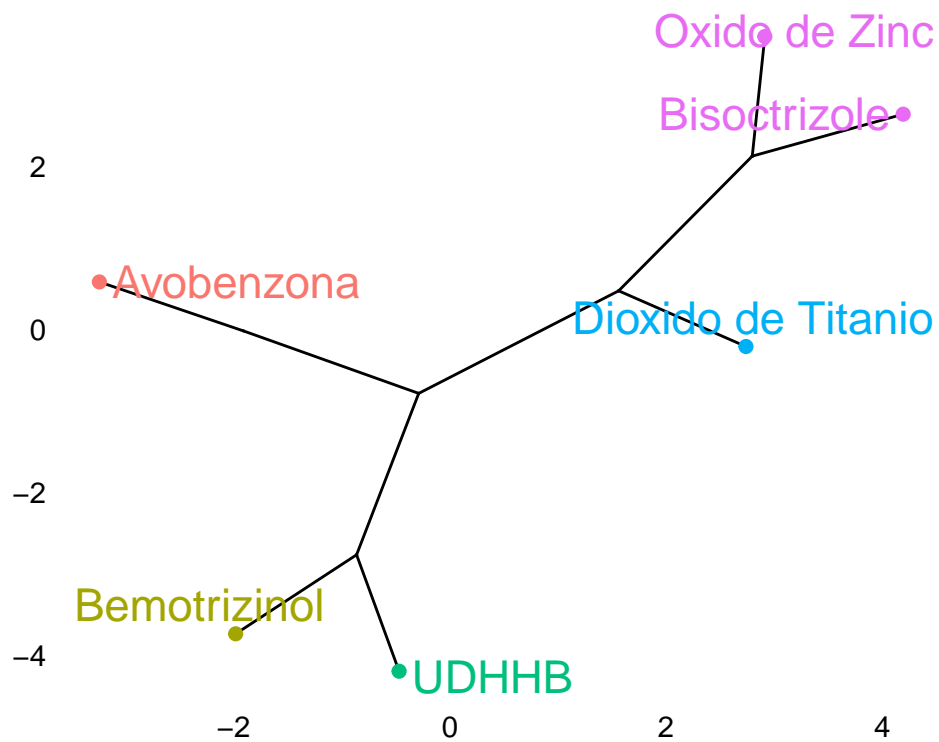


```
# Cálculo del Cluster Jerárquico - Filogenético
cluster_sin_filtros <- hcut(t(df_sin_filtros),
  k = 5,
  stand = TRUE,
  hc_func = c("hclust", "agnes", "diana"),
  hc_method = "ward.D2",
  hc_metric = "euclidean",)
fviz_dend(cluster_sin_filtros, rect = TRUE, cex = 0.7,
  k_colors = c("#000000", "#FF0000", "#00B050", "#FFC000", "#7030A0"), horiz = TRUE)
```

Cluster Dendrogram



```
fviz_dend(cluster_sin_filtros, k=5, rect = TRUE, cex=1.5, type = "phylogenic",  
          color_labels_by_k=T, repel=T)
```



ANEXO IV - OBTENCIÓN, TRANSFORMACION y CARGA DE LOS DATOS (ETL)

Contenido

4. OBTENCIÓN, TRANSFORMACION y CARGA DE LOS DATOS (ETL)	2
4.1. Instalación de los Paquetes y cargado de Librerías	2
4.2. Carga de los datos	2
4.3. Resumen de los datos	2
4.3. Corrección de los nombres de los ingredientes	3
4.4. Determinar presencia de un ingredientes	16
4.5. Guardado el archivo transformado	18

4. OBTENCIÓN, TRANSFORMACION y CARGA DE LOS DATOS (ETL)

4.1. Instalación de los Paquetes y cargado de Librerías

Instalamos los paquetes y cargamos las librerías necesarias para realizar el análisis.

```
# Ingresamos los paquetes necesarios
packages <- c("tidyverse","stringr")

# Instalamos los paquetes no instalados
installed_packages <- packages %in% rownames(installed.packages())
if (any(installed_packages == FALSE)) {
  install.packages(packages[!installed_packages])}

# Cargamos las librerías correspondientes
invisible(lapply(packages, library, character.only = TRUE))

rm(packages); rm(installed_packages) # Los borramos del entorno y quitamos de memoria
```

4.2. Carga de los datos

```
rm(list=ls()) # Borrado de todas las variables del Entorno

origen <- read.csv("C:/TFM_Gustavo_Pittaluga/DATA/ANEXO_IV_ETL.CSV", encoding="UTF-8", sep=";", header = TRUE)

df <- origen[,c(3,4,5,6,10,11,12)] # Eliminación de las columnas innecesarias
colnames(df)[5] <- "PRECIO" # Renombrado de la columna a PRECIOS

df$PRECIO <- gsub(",", ".", df$PRECIO)
df$PRECIO <- as.numeric(df$PRECIO)

df <- df %>% filter(!df$INGREDIENTES == "") # Eliminar celdas vacías
df <- df %>% filter(df$FPS > 5) # Eliminar valores con errores
df <- df %>% filter(!df$PRECIO == "") # Eliminar valores con errores
df <- df %>% filter(df$PRECIO > 0) # Eliminar valores con errores

df <- df[order(df$ANIO,df$MES),]

rm(origen) # Eliminar datos originales
```

4.3. Resumen de los datos

Como comprobación que se han cargado correctamente hacemos un resumen.

```
# Resumen de los datos
summary(df)
```

```
## INGREDIENTES          FPS          PAIS          REGION
```

```
## Length:17337      Min.   : 6.00   Length:17337      Length:17337
## Class :character  1st Qu.: 30.00  Class :character  Class :character
## Mode  :character  Median  : 50.00  Mode  :character  Mode  :character
##                               Mean   : 42.12
##                               3rd Qu.: 50.00
##                               Max.   :130.00
##      PRECIO          ANIO          MES
## Min.   : 0.16   Min.   :2001   Min.   : 1.000
## 1st Qu.: 6.86   1st Qu.:2015   1st Qu.: 4.000
## Median : 14.67  Median :2018   Median : 6.000
## Mean   : 33.30  Mean   :2018   Mean   : 6.069
## 3rd Qu.: 34.70  3rd Qu.:2021   3rd Qu.: 8.000
## Max.   :1539.29 Max.   :2022   Max.   :12.000
```

4.3. Corrección de los nombres de los ingredientes

En los datos obtenidos de las bases de datos se han encontrado diversos errores:

- Nombres pertenecientes a un filtro solar escritos de diversas formas sin seguir un estándar.
- Nombres de los filtros solares escritos con distinta expresión de las fórmulas químicas
- Nombres de los filtros solares escritos con distintas nomenclaturas. (INCI, USAN, ABR)
- Errores propios de tipeo al ingresar dichos nombres.
- Inclusión de información no relevante textual en medio de las fórmulas químicas.

Para ello se realizó una lista de conversión que transformó a toda esa información no congruente a nomenclatura ABR.

Se realizaron 665 cambios de textos de este tipo que se aplicaron todas las veces que apareciera en los 699.271 componentes presentes en los 17.337 productos.

```
# Corrección de los nombres de los ingredientes

# install.packages("stringr", dependencies = T)
library(stringr)

df$INGREDIENTES <- tolower(df$INGREDIENTES)
colnames(df)[1] <- "ING"

df$ING <- str_replace(df$ING, "z-cote", "ZnO")
df$ING <- str_replace(df$ING, "zinc oxide", "ZnO")
df$ING <- str_replace(df$ING, "z-cote", "ZnO")
df$ING <- str_replace(df$ING, "zinclear", "ZnO")
df$ING <- str_replace(df$ING, "zinci oxidum", "ZnO")
df$ING <- str_replace(df$ING, "zinc dioxide", "ZnO")
df$ING <- str_replace(df$ING, "zinc ocide", "ZnO")
df$ING <- str_replace(df$ING, "benzophenone-4", "BP4")
df$ING <- str_replace(df$ING, "4-benzophenone", "BP4")
df$ING <- str_replace(df$ING, "oxibenzona", "BP3")
df$ING <- str_replace(df$ING, "oxibenzone", "BP3")
df$ING <- str_replace(df$ING, "benzophenome-3", "BP3")
df$ING <- str_replace(df$ING, "benzophenon-3", "BP3")
df$ING <- str_replace(df$ING, "benzophenone-3", "BP3")
df$ING <- str_replace(df$ING, "3-benzophenone", "BP3")
```

```

df$ING <- str_replace(df$ING,"benzophenon -3", "BP3")
df$ING <- str_replace(df$ING,"benzophenone", "BP3")
df$ING <- str_replace(df$ING,"benzophenone-", "BP3")
df$ING <- str_replace(df$ING,"benzophenone trisiloxane", "BP3")
df$ING <- str_replace(df$ING,"benzophenone3 ", "BP3")
df$ING <- str_replace(df$ING,"benzophenone-3 ", "BP3")
df$ING <- str_replace(df$ING,"benzophenone-8", "BP3")
df$ING <- str_replace(df$ING,"benzophenones", "BP3")
df$ING <- str_replace(df$ING,"benzopheone-3", "BP3")
df$ING <- str_replace(df$ING,"benzophonone 3.", "BP3")
df$ING <- str_replace(df$ING,"oxybenzoin", "BP3")
df$ING <- str_replace(df$ING,"oxybenone ", "BP3")
df$ING <- str_replace(df$ING,"oxybenxone", "BP3")
df$ING <- str_replace(df$ING,"oxybenzene", "BP3")
df$ING <- str_replace(df$ING,"oxybenzoate", "BP3")
df$ING <- str_replace(df$ING,"oxybenzon", "BP3")
df$ING <- str_replace(df$ING,"oxybenzone", "BP3")
df$ING <- str_replace(df$ING,"benxophenone-3", "BP3")
df$ING <- str_replace(df$ING,"oxybenzone-3", "BP3")
df$ING <- str_replace(df$ING,"oxyenzone", "BP3")
df$ING <- str_replace(df$ING,"ocybenzone ", "BP3")
df$ING <- str_replace(df$ING,"phenylbenzimidazole-5-sulfonic acid ", "PBSA")
df$ING <- str_replace(df$ING,"benzimidazole", "PBSA")
df$ING <- str_replace(df$ING,"phenylbenzimidazole", "PBSA")
df$ING <- str_replace(df$ING,"ensulizone ", "PBSA")
df$ING <- str_replace(df$ING,"phenylbenzimidazole sulfonic acid ensulizole", "PBSA")
df$ING <- str_replace(df$ING,"uvsorb-heb", "DBT")
df$ING <- str_replace(df$ING,"diethyl hexyl butamido tryazone", "DBT")
df$ING <- str_replace(df$ING,"diethyl hexyl butamido triazone", "DBT")
df$ING <- str_replace(df$ING,"diethyl hexyl butamido tryazone", "DBT")
df$ING <- str_replace(df$ING,"diethylhexye butamido trizone", "DBT")
df$ING <- str_replace(df$ING,"diethylexyl butamido triaxone", "DBT")
df$ING <- str_replace(df$ING,"diethylhexyl bedmate triazone", "DBT")
df$ING <- str_replace(df$ING,"diethylhexyl betamido triazone", "DBT")
df$ING <- str_replace(df$ING,"diethylhexyl but amide triazone", "DBT")
df$ING <- str_replace(df$ING,"diethylhexyl butamido tiazone ", "DBT")
df$ING <- str_replace(df$ING,"diethylhexyl butamido triazone", "DBT")
df$ING <- str_replace(df$ING,"diethylhexyl butamido trizane", "DBT")
df$ING <- str_replace(df$ING,"diethylhexyl butamido trizone", "DBT")
df$ING <- str_replace(df$ING,"diethylhexyl butamidoi triazone", "DBT")
df$ING <- str_replace(df$ING,"diethylhexyl butamindo triazone", "DBT")
df$ING <- str_replace(df$ING,"diethylhexyl butamine triazone", "DBT")
df$ING <- str_replace(df$ING,"diethylhexyl butamine triazone", "DBT")
df$ING <- str_replace(df$ING,"diethylhexyl butamine triazone", "DBT")
df$ING <- str_replace(df$ING,"diethylhexyl butamine triazone", "DBT")
df$ING <- str_replace(df$ING,"diethylhexyl butamine triazone", "DBT")
df$ING <- str_replace(df$ING,"diethylhexyl butamine triazone", "DBT")
df$ING <- str_replace(df$ING,"dihexyl butamido triazone", "DBT")
df$ING <- str_replace(df$ING,"diethllhexyl butamido triazone", "DBT")
df$ING <- str_replace(df$ING,"diethylhexyl butamido triazone", "DBT")
df$ING <- str_replace(df$ING,"triazine methoxy phenyl bis-ethylhexyloxyphenol", "BEMT")
df$ING <- str_replace(df$ING,"triazone ethylhexyl", "BEMT")
df$ING <- str_replace(df$ING,"triazone octyl", "BEMT")

```

```

df$ING <- str_replace(df$ING,"ethylhexyloxyphenol methoxyphenyl", "BEMT")
df$ING <- str_replace(df$ING,"ethylhexyloxyphenol methoxyphenyl triazine", "BEMT")
df$ING <- str_replace(df$ING,"indicated bis ethyloxyphenol methoxyphenyl triazine", "BEMT")
df$ING <- str_replace(df$ING,"indicated bis-ethylhexyloxyphenol methoxyphenyl triazine", "BEMT")
df$ING <- str_replace(df$ING,"key bis-ethylhexyloxyphenol methoxyphenyl triazine", "BEMT")
df$ING <- str_replace(df$ING,"methoxyphenyl triazine", "BEMT")
df$ING <- str_replace(df$ING,"methoxyphenyl triazine bis-ethylhexyl oxyphenol", "BEMT")
df$ING <- str_replace(df$ING,"methoxyphenyl triazine bis-ethylhexyloxyphenol", "BEMT")
df$ING <- str_replace(df$ING,"methoxyphenyl triazine bis-ethylhexyloxyphenol", "BEMT")
df$ING <- str_replace(df$ING,"methoxyphenyl triazine", "BEMT")
df$ING <- str_replace(df$ING,"methoxyphenyl triazine bis ethyl hexyl oxyphenol", "BEMT")
df$ING <- str_replace(df$ING,"2,4-bis4-2-ethylhexyloxy -2-hydroxyphenyl-6-4-methoxyphenyl-1,3,5-triazin", "BEMT")
df$ING <- str_replace(df$ING,"polymethyl methacrylate bis-ethylhexyloxyphenol methoxyphenyl triazine", "BEMT")
df$ING <- str_replace(df$ING,"bis-ethylhexyloxyphenol methoxyphenyl triazine", "BEMT")
df$ING <- str_replace(df$ING,"sodium dicocoyl ethylendiamine methoxyphenyl triazine", "BEMT")
df$ING <- str_replace(df$ING,"methoxy phenyl triazine bis ethyl hexyl oxyphenol", "BEMT")
df$ING <- str_replace(df$ING,"methoxyphenyl triazine", "BEMT")
df$ING <- str_replace(df$ING,"methoxydibenzoylmethane", "BMDBM")
df$ING <- str_replace(df$ING,"4-butyl-4 metoxydebenzoilmethanme", "BMDBM")
df$ING <- str_replace(df$ING,"4-butyl-4-methoxydibenzoilmethane", "BMDBM")
df$ING <- str_replace(df$ING,"4-butyl-4-methoxydibenzoyl methane", "BMDBM")
df$ING <- str_replace(df$ING,"4-butylmethoxybenzoylmethane", "BMDBM")
df$ING <- str_replace(df$ING,"t-butyl methoxy dibenzoyl methane", "BMDBM")
df$ING <- str_replace(df$ING,"t-butylmethoxy dibenzoylmethane", "BMDBM")
df$ING <- str_replace(df$ING,"t-butylmethoxydibenzoylmethane", "BMDBM")
df$ING <- str_replace(df$ING,"methoxydibenzoyl-methane", "BMDBM")
df$ING <- str_replace(df$ING,"methoxybenzoyl methane butyl", "BMDBM")
df$ING <- str_replace(df$ING,"methoxybenzoylmethane", "BMDBM")
df$ING <- str_replace(df$ING,"methoxydibezoylmathane butyl", "BMDBM")
df$ING <- str_replace(df$ING,"octocrulene", "OCR")
df$ING <- str_replace(df$ING,"triazone", "OCR")
df$ING <- str_replace(df$ING,"triazone ethylhexyl", "OCR")
df$ING <- str_replace(df$ING,"triazone octyl", "OCR")
df$ING <- str_replace(df$ING,"octocrylene", "OCR")
df$ING <- str_replace(df$ING,"octocylene", "OCR")
df$ING <- str_replace(df$ING,"2-ethylhexyl-2-cyano-3,3-diphenylacrylate", "OCR")
df$ING <- str_replace(df$ING,"octicryline", "OCR")
df$ING <- str_replace(df$ING,"ocotocrilene", "OCR")
df$ING <- str_replace(df$ING,"octocrytene", "OCR")
df$ING <- str_replace(df$ING,"octocyclene", "OCR")
df$ING <- str_replace(df$ING,"octocylene", "OCR")
df$ING <- str_replace(df$ING,"octocylene ", "OCR")
df$ING <- str_replace(df$ING,"octodrylene", "OCR")
df$ING <- str_replace(df$ING,"octorcylene", "OCR")
df$ING <- str_replace(df$ING,"octorylene", "OCR")
df$ING <- str_replace(df$ING,"parsol 30", "OCR")
df$ING <- str_replace(df$ING,"octrocrylene", "OCR")
df$ING <- str_replace(df$ING,"2-ethylhexyl-p-dimethylamino benzoate", "OCR")
df$ING <- str_replace(df$ING,"oxtocylene", "OCR")
df$ING <- str_replace(df$ING,"odocrylene", "OCR")
df$ING <- str_replace(df$ING,"octyocylene", "OCR")
df$ING <- str_replace(df$ING,"octyocrylene", "OCR")
df$ING <- str_replace(df$ING,"octcrylene", "OCR")

```

```

df$ING <- str_replace(df$ING,"octocrileno", "OCR")
df$ING <- str_replace(df$ING,"octocrylane", "OCR")
df$ING <- str_replace(df$ING,"octacrylene", "OCR")
df$ING <- str_replace(df$ING,"2-ethylhexyl-p-methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethylhexyl homosalatemethoxysinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethylhexyl methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octyl methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octyl methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octyl methozycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethylhexyl methxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethyhexyl methxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"2-ethyhexyl 4-methoxycinnamate ", "EHMC")
df$ING <- str_replace(df$ING,"2-ethyl hexyl p-methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"2-ethylhexyl 4-methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"2-ethylhexy-1-4-methoxycinnamate ", "EHMC")
df$ING <- str_replace(df$ING,"2-ethylhexyl paramethoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"2-ethylhexyl4-methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"2-ethylhexyl-4-methoxycinnamate ", "EHMC")
df$ING <- str_replace(df$ING,"4-methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"4-methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"8. ethylhexyl-p-methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"coetyl methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"denat. methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ectyl methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethlhexyl methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethyl exyl methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethyl hexyl dimethyl paraisoamyl p-methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethyl methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethylehxl methyl methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethylene methoxycinnamate 7.50%", "EHMC")
df$ING <- str_replace(df$ING,"ethylexic methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethylezyl methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethylhexy methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethylhexy methoxycinnamate octinoxate", "EHMC")
df$ING <- str_replace(df$ING,"ethyl-hexyl methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethylhexyl para methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethylhexyl-methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethylhexyl-p methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethylhexyl-p-methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethylhexyl-p-methoxycinnamate 1.2 w/w", "EHMC")
df$ING <- str_replace(df$ING,"ethylhexyl-p-methoxycinnamate 7.", "EHMC")
df$ING <- str_replace(df$ING,"ethylhexy-p-methoxycinnamate 7.", "EHMC")
df$ING <- str_replace(df$ING,"ethylhezyl methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethyhexyl p-methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"glyceryl ethylhexanoate dimethoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"hexyl ethyl methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"in % 7.5 ethylhexyl-p-methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"butyl methoxydibenxoylmethane", "EHMC")
df$ING <- str_replace(df$ING,"methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octyl p-methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octylmethoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octyl-methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octymethoxycinnamate", "EHMC")

```

```

df$ING <- str_replace(df$ING,"oktyl methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"oxtyl methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octinoate 7", "EHMC")
df$ING <- str_replace(df$ING,"octylmethoxycinnamte", "EHMC")
df$ING <- str_replace(df$ING,"2-ethylhexyl p-metoxycynnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethyl hexyl cinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethyl hexyl methoxycinnamte", "EHMC")
df$ING <- str_replace(df$ING,"ethyl hexyl methoxycinnmate", "EHMC")
df$ING <- str_replace(df$ING,"ethylhexyl methoxcinnarmate", "EHMC")
df$ING <- str_replace(df$ING,"ethylhexyl methoxy cinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethylhexyl methoxycinamate", "EHMC")
df$ING <- str_replace(df$ING,"ethylhexyl methoxycinammate", "EHMC")
df$ING <- str_replace(df$ING,"ethylhexyl methoxycinnamae", "EHMC")
df$ING <- str_replace(df$ING,"ethylhexyl methoxycinnmate", "EHMC")
df$ING <- str_replace(df$ING,"ethylhexyl methoxyinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethylhexyl methoycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethylhexyl metoxicionamate", "EHMC")
df$ING <- str_replace(df$ING,"ethylhexyl p-methoxycinnagate", "EHMC")
df$ING <- str_replace(df$ING,"octinoxato", "EHMC")
df$ING <- str_replace(df$ING,"octinoxate", "EHMC")
df$ING <- str_replace(df$ING,"octinoxalate", "EHMC")
df$ING <- str_replace(df$ING,"octionoxate", "EHMC")
df$ING <- str_replace(df$ING,"octionoxate", "EHMC")
df$ING <- str_replace(df$ING,"octionxate", "EHMC")
df$ING <- str_replace(df$ING,"ethylhexyl-p-methoxy cinnamate", "EHMC")
df$ING <- str_replace(df$ING,"methoxy octyl cinnamate", "EHMC")
df$ING <- str_replace(df$ING,"methoxyethyl cinnamate ester", "EHMC")
df$ING <- str_replace(df$ING,"metoxy octyl cinnamate", "EHMC")
df$ING <- str_replace(df$ING,"metoxyoctyl cinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octadecyl di-t-butyl-4-hydroxyhydrocinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octyl cinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octyl mehoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octyl methocycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octyl methoxy cinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octyl methoxy cinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octyl methyl cinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octyl methylcinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octyl methyloxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octyl metoxcinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octylmethocinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octylmethoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octylmetoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"oxymethyl cinnamate", "EHMC")
df$ING <- str_replace(df$ING,"2-ethylhexyl cinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethyl hexyl p-methoxy cinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethyl hexylcinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethylhexyl cinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethylhexylcinnamate", "EHMC")
df$ING <- str_replace(df$ING,"methoxycinnamic acid octil", "EHMC")
df$ING <- str_replace(df$ING,"octyl methoxycinnamamte", "EHMC")
df$ING <- str_replace(df$ING,"octyl methoxycinnammate", "EHMC")
df$ING <- str_replace(df$ING,"octyl methoxycinnamte", "EHMC")
df$ING <- str_replace(df$ING,"octyl methoxycinnate", "EHMC")

```

```

df$ING <- str_replace(df$ING,"2-ethylhexyl-p-methoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octyl dimethoxybenzylidene dioxoimidazolidine propionate", "EHMC")
df$ING <- str_replace(df$ING,"octyl dimethoxybenzylidenedioxoimidazolizine propionate", "EHMC")
df$ING <- str_replace(df$ING,"octyl dimethoxybenzylideneoxoimidazolidine propionate", "EHMC")
df$ING <- str_replace(df$ING,"octyl methoxycinnmate", "EHMC")
df$ING <- str_replace(df$ING,"octyl methoxyconnamate", "EHMC")
df$ING <- str_replace(df$ING,"octyl methoxyenamatho", "EHMC")
df$ING <- str_replace(df$ING,"ethylhexyl-p-methoxy cinnamate", "EHMC")
df$ING <- str_replace(df$ING,"methoxy octyl cinnamate", "EHMC")
df$ING <- str_replace(df$ING,"methoxyethyl cinnamate ester", "EHMC")
df$ING <- str_replace(df$ING,"metoxy octyl cinnamate", "EHMC")
df$ING <- str_replace(df$ING,"metoxyoctyl cinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octadecyl di-t-butyl-4-hydroxyhydrocinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octyl cinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octyl mehocycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octyl methocycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octyl methoxy cinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octyl methoxy cinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octyl methyl cinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octyl methylcinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octyl methyloxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octyl metoxcinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octylmethocinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octylmethoxicinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octylmetoxycinnamate", "EHMC")
df$ING <- str_replace(df$ING,"oxymethyl cinnamate", "EHMC")
df$ING <- str_replace(df$ING,"2-ethylhexyl cinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethyl hexyl p-methoxy cinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethyl hexylcinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethylhexyl cinnamate", "EHMC")
df$ING <- str_replace(df$ING,"ethylhexylcinnamate", "EHMC")
df$ING <- str_replace(df$ING,"octinoate", "EHMC")
df$ING <- str_replace(df$ING,"octinoate", "EHMC")
df$ING <- str_replace(df$ING,"octinovate", "EHMC")
df$ING <- str_replace(df$ING,"octilmetoxicinamato", "EHMC")
df$ING <- str_replace(df$ING,"octyl methocycinamate", "EHMC")
df$ING <- str_replace(df$ING,"octyl methoxicnnamate ", "EHMC")
df$ING <- str_replace(df$ING,"octyl p-methoxycinnamate ", "EHMC")
df$ING <- str_replace(df$ING,"ethylhexyl methoxyicnmmate", "EHMC")
df$ING <- str_replace(df$ING,"iso amyl p-methoxycinnamate", "IMC")
df$ING <- str_replace(df$ING,"isoamlyl p-methoxycinnamate", "IMC")
df$ING <- str_replace(df$ING,"isoamyl p-methoxycinnamate", "IMC")
df$ING <- str_replace(df$ING,"isomil p-methoxycinnamate", "IMC")
df$ING <- str_replace(df$ING,"isoamil p-methoxycinnamate", "IMC")
df$ING <- str_replace(df$ING,"3. 4-methyl benzylidene camphor", "MBC")
df$ING <- str_replace(df$ING,"3-4-methylbenzylidene camphor", "MBC")
df$ING <- str_replace(df$ING,"3-4-methylbenzylidene camphor camphor", "MBC")
df$ING <- str_replace(df$ING,"3-benzylidene camphor", "MBC")
df$ING <- str_replace(df$ING,"4 methyl benzylidene camphor", "MBC")
df$ING <- str_replace(df$ING,"4-methylbenzylidene camphor", "MBC")
df$ING <- str_replace(df$ING,"4-methylbenzylidene camphor 40mg/g", "MBC")
df$ING <- str_replace(df$ING,"4-methyl benzylidene camphor", "MBC")
df$ING <- str_replace(df$ING,"4-methyl benzylidene camphor 30mg/g", "MBC")

```



```

df$ING <- str_replace(df$ING,"terephthalylidene dicamphor sulphonic acid", "TDSA")
df$ING <- str_replace(df$ING,"terephthalylidene dicamphor sulphoric acid", "TDSA")
df$ING <- str_replace(df$ING,"terephthalylidenedicamphor sulfonic acid", "TDSA")
df$ING <- str_replace(df$ING,"terephthylylidene dicamphor sulfonic acid", "TDSA")
df$ING <- str_replace(df$ING,"terephthylylidene dicamphor sulfonic acid", "TDSA")
df$ING <- str_replace(df$ING,"terephthalydenedicamphor sulphuric acid", "TDSA")
df$ING <- str_replace(df$ING,"terephthalylidene dicamphor sulfonic acid", "TDSA")
df$ING <- str_replace(df$ING,"tereththalydene dicamphor sulfonic acid", "TDSA")
df$ING <- str_replace(df$ING,"threphthalylidene dicamphor sulfonic acid", "TDSA")
df$ING <- str_replace(df$ING,"terephthalylidene dl-camphor sulfonic acid", "TDSA")
df$ING <- str_replace(df$ING,"terephthalylidene dl-camphor sulphonic acid", "TDSA")
df$ING <- str_replace(df$ING,"terephthalylidine", "TDSA")
df$ING <- str_replace(df$ING,"eusolex t aqua", "TDSA")
df$ING <- str_replace(df$ING,"eusolex t-aqua 1 w/w", "TDSA")
df$ING <- str_replace(df$ING,"ecamsule", "TDSA")
df$ING <- str_replace(df$ING,"Diethylamino hydroxybenzoyl hexyl benzoate", "DHHB")
df$ING <- str_replace(df$ING,"diethylamino hydroxybenzoyl hexyl benzoate", "DHHB")
df$ING <- str_replace(df$ING,"diethyl aminohydroxybenzoyl hexyl benzoate", "DHHB")
df$ING <- str_replace(df$ING,"diethylamidno hydroxybenzoyl hexyl benzoate", "DHHB")
df$ING <- str_replace(df$ING,"diethylamine hydroxybenzoyl hexyl benzoate", "DHHB")
df$ING <- str_replace(df$ING,"diethylamino hydroxybenzoyl hexyl benzoate", "DHHB")
df$ING <- str_replace(df$ING,"diethylamino hydroxybenzoyl hexyl benzoate ", "DHHB")
df$ING <- str_replace(df$ING,"diethylamino hydroxybenzoyl hexyl benzoate alcohol", "DHHB")
df$ING <- str_replace(df$ING,"diethylamino hydroxybenzoyl hexyl benzoate all aqua", "DHHB")
df$ING <- str_replace(df$ING,"diethylamino hydroxybenzoyl hexyl benzoate aqua", "DHHB")
df$ING <- str_replace(df$ING,"diethylamino hydroxybenzoyl hexyl benzoate aqua deionized", "DHHB")
df$ING <- str_replace(df$ING,"diethylamino hydroxybenzoyl hexyl benzoate aqua purified", "DHHB")
df$ING <- str_replace(df$ING,"diethylamino hydroxybenzoyl hexyl benzoate cyclopentasiloxane", "DHHB")
df$ING <- str_replace(df$ING,"diethylamino hydroxybenzoyl hexyl benzoate dimethicone", "DHHB")
df$ING <- str_replace(df$ING,"diethylamino hydroxybenzoyl hexyl benzoate ethylhexyl stearate", "DHHB")
df$ING <- str_replace(df$ING,"diethylamino hydroxybenzoyl hexyl benzoate filtered", "DHHB")
df$ING <- str_replace(df$ING,"diethylamino hydroxybenzoyl hexyl benzoate granules", "DHHB")
df$ING <- str_replace(df$ING,"diethylamino hydroxybenzoyl hexyl benzoate inaqua", "DHHB")
df$ING <- str_replace(df$ING,"diethylamino hydroxybenzoyl hexyl benzoate inparaffinum liquidum liquid", "DHHB")
df$ING <- str_replace(df$ING,"diethylamino hydroxybenzoyl hexyl benzoate liquefied petroleum gas liquid", "DHHB")
df$ING <- str_replace(df$ING,"diethylamino hydroxybenzoyl hexyl benzoate other ingredients ethylhexyl s", "DHHB")
df$ING <- str_replace(df$ING,"diethylamino hydroxybenzoyl hexyl benzoate preservatives phenoxyethanol", "DHHB")
df$ING <- str_replace(df$ING,"diethylamino hydroxybenzoyl hexyl benzoate tan: zinc oxide", "DHHB")
df$ING <- str_replace(df$ING,"diethylamino hydroxybenzoyl hexyl benzoate titanium dioxide", "DHHB")
df$ING <- str_replace(df$ING,"diethylamino hydroxybenzoyl hexyl benzoate uvinul a+ ", "DHHB")
df$ING <- str_replace(df$ING,"diethylamino hydroxybenzoyl hexyl benzoate uvnal a plus", "DHHB")
df$ING <- str_replace(df$ING,"diethylamono hydroxybenzoyl hexyl benzoate", "DHHB")
df$ING <- str_replace(df$ING,"dimethyl amino hydroxybenzoyl hexyl benzoate 1.00%", "DHHB")
df$ING <- str_replace(df$ING,"dimethylamine hydroxybenzoyl hexyl benzoate", "DHHB")
df$ING <- str_replace(df$ING,"dimethylamino hydroxybenzoyl hexyl benzoate", "DHHB")
df$ING <- str_replace(df$ING,"dyethylamine hydroxybenzoyl hexyl benzoate", "DHHB")
df$ING <- str_replace(df$ING,"ethylamino hydroxybenzoyl hexyl benzoate", "DHHB")
df$ING <- str_replace(df$ING,"indicated diethylamino hydroxybenzoyl hexyl benzoate", "DHHB")
df$ING <- str_replace(df$ING,"key ingredients diethylamino hydroxybenzoyl hexyl benzoate", "DHHB")
df$ING <- str_replace(df$ING,"main diethylamino hydroxybenzoyl hexyl benzoate", "DHHB")
df$ING <- str_replace(df$ING,"2-4-diethylamino-2-hydroxybenzoyl benzoic acid", "DHHB")
df$ING <- str_replace(df$ING,"2-4-diethylamino-2-hydroxybenzoyl benzoic acid hexyl ester", "DHHB")
df$ING <- str_replace(df$ING,"diethylamino hydroxybenzol hexyl benzoate", "DHHB")

```

```

df$ING <- str_replace(df$ING,"diethylamino hydroxybenzol hexyl benzoate 2.00%", "DHHB")
df$ING <- str_replace(df$ING,"diethylamino hydroxybenzoyl hexyl benzoate", "DHHB")
df$ING <- str_replace(df$ING,"diethylamino hydroxybenzoyl", "DHHB")
df$ING <- str_replace(df$ING,"diethylamino hydroxybenzoyl benzoate", "DHHB")
df$ING <- str_replace(df$ING,"diethylamino hydroxybenzoylhexyl benzoate", "DHHB")
df$ING <- str_replace(df$ING,"diethylamino hydroxybezoyl hexyl benzoate", "DHHB")
df$ING <- str_replace(df$ING,"diethylamino hydroxybenzol hexyl benzoate", "DHHB")
df$ING <- str_replace(df$ING,"diethylaminohydroxybenzoylhexyl benzoate", "DHHB")
df$ING <- str_replace(df$ING,"main diethylamino hydroxybnzoyl hexyl benzoate dhhb ", "DHHB")
df$ING <- str_replace(df$ING,"p-methoxycinnamic acid 2-ethyl hexyl ester 2-4diethylamino 2-hydroxybenzo", "DHHB")
df$ING <- str_replace(df$ING,"DHHB", "DHHB")
df$ING <- str_replace(df$ING,"diethyl amino hydroxy benzoyl hexybenzoate", "DHHB")
df$ING <- str_replace(df$ING,"diethyl amino hydroxybenzoyl", "DHHB")
df$ING <- str_replace(df$ING,"diethyl amino hydroxybenzyl benzoate", "DHHB")
df$ING <- str_replace(df$ING,"methylene bis-benzotriazolyl tetramethylbutylphenol", "MBBT")
df$ING <- str_replace(df$ING,"dicaprylyl carbonate methylene bis-benzotriazolyl tetramethylbutylphenol", "MBBT")
df$ING <- str_replace(df$ING,"diethylhexyl bis-benzotriazolyl tetramethylbutylphenol", "MBBT")
df$ING <- str_replace(df$ING,"ethylene bis-benzotriazolyl tetramethylbutylphenol", "MBBT")
df$ING <- str_replace(df$ING,"ethylene bis-benzotriazolyl tetramethylbutylphenol nano", "MBBT")
df$ING <- str_replace(df$ING,"glycol decyl glucoside xanthan gum methylene bis-benzotriazolyl tetramethylbutylphenol", "MBBT")
df$ING <- str_replace(df$ING,"main methylene bis-benzotriazolyl tetramethylbutylphenol nano", "MBBT")
df$ING <- str_replace(df$ING,"mehtylene bis-benzotriazolyl tetramethylbutylphenol", "MBBT")
df$ING <- str_replace(df$ING,"methyl bis-benzotriazolyl tetramethylbutylphenol", "MBBT")
df$ING <- str_replace(df$ING,"methylene bis benzoatriazolyl tetramethylbutylphenol/decyl glucoside/xanthan", "MBBT")
df$ING <- str_replace(df$ING,"methylene bis- benzotriazolyl tetramethylbutylphenol", "MBBT")
df$ING <- str_replace(df$ING,"methylene bis- benzotriazolyl tetramethylbutylphenol", "MBBT")
df$ING <- str_replace(df$ING,"methylene bis-benzotriazolyl tetramethylbutylphenol", "MBBT")
df$ING <- str_replace(df$ING,"methylene bis-benzotriazolyl tetramethylbutylphenol", "MBBT")
df$ING <- str_replace(df$ING,"methylene bis-benzotriazolyl tetramethylbutylphenol", "MBBT")
df$ING <- str_replace(df$ING,"methylene bisbenzotriazolyl tetramethylbutylphenol", "MBBT")
df$ING <- str_replace(df$ING,"methylene bis-benzotriazolyl tetramethylbutylphenol", "MBBT")
df$ING <- str_replace(df$ING,"methylene bis-benzotriazolyltetramethylbutylphenol", "MBBT")
df$ING <- str_replace(df$ING,"methylene bis-benzotriazolyltetramethylbutylphenol,parfum.", "MBBT")
df$ING <- str_replace(df$ING,"methylene bisbenzotriazolyl tetramethylbutylphenol", "MBBT")
df$ING <- str_replace(df$ING,"methylene bis-benzotriazolyl tetramethylbutylphenol", "MBBT")
df$ING <- str_replace(df$ING,"methylene-bis benzotriazolyl tetramethylbutylphenol", "MBBT")
df$ING <- str_replace(df$ING,"methylenebis-benzotriazolyl tetramethylbutylphenol", "MBBT")
df$ING <- str_replace(df$ING,"metylene bis-benzotriazolyl tetramethylbutylphenol", "MBBT")
df$ING <- str_replace(df$ING,"tetramethylbutylphenol", "MBBT")
df$ING <- str_replace(df$ING,"tetramethylbutylphenol potassium cetyl phosphate", "MBBT")
df$ING <- str_replace(df$ING,"tris-biphenyl triazine", "MBBT")
df$ING <- str_replace(df$ING,"methyl bis-benzotriazolyl tetramethylbutyl phenol", "MBBT")
df$ING <- str_replace(df$ING,"methylen bis benzotriazolyl", "MBBT")
df$ING <- str_replace(df$ING,"methylene bis-benzotriazolyl", "MBBT")
df$ING <- str_replace(df$ING,"methylene bis-benzotriazolyl tetra methybutylphenyl", "MBBT")
df$ING <- str_replace(df$ING,"methylene bis-benzotriazolyl tetra methyl butylphenol", "MBBT")
df$ING <- str_replace(df$ING,"methylene bis-benzotriazolyl tetra methylbutyl-phenol", "MBBT")
df$ING <- str_replace(df$ING,"methylene bis-benzotriazolyl tetramethyl", "MBBT")
df$ING <- str_replace(df$ING,"methylene bis-benzotriazolyl tetramethyl butylphenyl", "MBBT")
df$ING <- str_replace(df$ING,"methylene bis-benzotriazolyl tetramethylbutylphenon", "MBBT")
df$ING <- str_replace(df$ING,"methylene bis-benzotriazolyl tetramethylbutylphenyl", "MBBT")
df$ING <- str_replace(df$ING,"methylene bis-benzotriazolyl tetramethylbutylphenol glyceryl stearate", "MBBT")
df$ING <- str_replace(df$ING,"methylene bis-benzotriazolyl tetramethylphenol", "MBBT")

```

```

df$ING <- str_replace(df$ING,"methylene bis-benzotriazolyltetramethylbutylphenol", "MBBT")
df$ING <- str_replace(df$ING,"methylene bis- benzothiazolyl", "MBBT")
df$ING <- str_replace(df$ING,"methylene bis-benzotriazolyl tetramethylbutylphenol", "MBBT")
df$ING <- str_replace(df$ING,"methylene bis-benzotriazolyl tetramethylbutylphenol polydecane", "MBBT")
df$ING <- str_replace(df$ING,"methylene bis-benzotriazolyl tetramethyl butyl phenol ", "MBBT")
df$ING <- str_replace(df$ING,"methylene bis-benzotriazolyltetramethylbutyl phenol", "MBBT")
df$ING <- str_replace(df$ING,"methylene bisbenzotriazolyl tetraethylbutylphenol", "MBBT")
df$ING <- str_replace(df$ING,"methylene bis-benzotriazolyl tetramethylbutylphenol", "MBBT")
df$ING <- str_replace(df$ING,"methylene bis-benzotriazolyl tetramethylbutylphenol", "MBBT")
df$ING <- str_replace(df$ING,"methylene bis-benzotriazolyl tetra methylbutylphenol", "MBBT")
df$ING <- str_replace(df$ING,"tinosorb m", "MBBT")
df$ING <- str_replace(df$ING,"uvinul t-150", "EHT")
df$ING <- str_replace(df$ING,"uvinul t150 broad spectrum uvb-filter basf", "EHT")
df$ING <- str_replace(df$ING,"tris-ethylhexyl oxy carbonyl anilynotriazine", "EHT")
df$ING <- str_replace(df$ING,"tris-ethylhexyl oxycarbonyl anilynotriazine", "EHT")
df$ING <- str_replace(df$ING,"tris-ethylhexyl oxycarbonyl anilynotriazine", "EHT")
df$ING <- str_replace(df$ING,"ethylhexyl triazone", "EHT")
df$ING <- str_replace(df$ING,"ethyl hexyl triazone", "EHT")
df$ING <- str_replace(df$ING,"octyl triazone", "EHT")
df$ING <- str_replace(df$ING,"octyutriazone", "EHT")
df$ING <- str_replace(df$ING,"ethylesiltriazone", "EHT")
df$ING <- str_replace(df$ING,"ethylexil triazone", "EHT")
df$ING <- str_replace(df$ING,"ethylexyl triazone", "EHT")
df$ING <- str_replace(df$ING,"ethylhecyll triazone", "EHT")
df$ING <- str_replace(df$ING,"ethylhexil triazone", "EHT")
df$ING <- str_replace(df$ING,"ethylhexyl triazone", "EHT")
df$ING <- str_replace(df$ING,"ethyl-hexyl triazone", "EHT")
df$ING <- str_replace(df$ING,"ethylhexyl triazone", "EHT")
df$ING <- str_replace(df$ING,"ethylhexyl triaxone", "EHT")
df$ING <- str_replace(df$ING,"ethylhexyl triazine", "EHT")
df$ING <- str_replace(df$ING,"ethylhexyl triazione", "EHT")
df$ING <- str_replace(df$ING,"ethylhexyl triazon", "EHT")
df$ING <- str_replace(df$ING,"ethylhexyl triazo-ne", "EHT")
df$ING <- str_replace(df$ING,"ethylhexyl triazonine", "EHT")
df$ING <- str_replace(df$ING,"octyltriaazona", "EHT")
df$ING <- str_replace(df$ING,"octyl taiazone", "EHT")
df$ING <- str_replace(df$ING,"octyl tirazone", "EHT")
df$ING <- str_replace(df$ING,"octyl triaxzone", "EHT")
df$ING <- str_replace(df$ING,"octiltoriazon", "EHT")
df$ING <- str_replace(df$ING,"uvinul a", "DHHB")
df$ING <- str_replace(df$ING,"uvinul easy", "DHHB")
df$ING <- str_replace(df$ING,"uvinol ", "DHHB")
df$ING <- str_replace(df$ING,"parsol tx", "TiO2")
df$ING <- str_replace(df$ING,"titanium dioxide", "TiO2")
df$ING <- str_replace(df$ING,"titanium dioxide", "TiO2")
df$ING <- str_replace(df$ING,"titanium", "TiO2")
df$ING <- str_replace(df$ING,"titanium 2.0%", "TiO2")
df$ING <- str_replace(df$ING,"titanium bi oxide 2.5", "TiO2")
df$ING <- str_replace(df$ING,"titanium dixide", "TiO2")
df$ING <- str_replace(df$ING,"titanium dioxid", "TiO2")
df$ING <- str_replace(df$ING,"titanium dioxidem", "TiO2")
df$ING <- str_replace(df$ING,"tin oxide", "TiO2")

```

```

df$ING <- str_replace(df$ING,"tinanium oxide", "TiO2")
df$ING <- str_replace(df$ING,"tinii oxide", "TiO2")
df$ING <- str_replace(df$ING,"drometholizole trisiloxane", "DMT")
df$ING <- str_replace(df$ING,"dromethorizol trisiloxane", "DMT")
df$ING <- str_replace(df$ING,"dromethorizoltrisiloxane", "DMT")
df$ING <- str_replace(df$ING,"dromethricole trisiloxane", "DMT")
df$ING <- str_replace(df$ING,"dromethrizole trisiloxane", "DMT")
df$ING <- str_replace(df$ING,"dromethyzol parkii", "DMT")
df$ING <- str_replace(df$ING,"drometizole trisiloxane", "DMT")
df$ING <- str_replace(df$ING,"drometizole trisoloxane", "DMT")
df$ING <- str_replace(df$ING,"drometorizole trisiloxane", "DMT")
df$ING <- str_replace(df$ING,"drometrixole trisiloxane", "DMT")
df$ING <- str_replace(df$ING,"drometrizol trisiloxane", "DMT")
df$ING <- str_replace(df$ING,"drometrizole", "DMT")
df$ING <- str_replace(df$ING,"drometrizole trisil oxane", "DMT")
df$ING <- str_replace(df$ING,"drometrizole trisilloxane", "DMT")
df$ING <- str_replace(df$ING,"drometrizole trisilocane", "DMT")
df$ING <- str_replace(df$ING,"drometrizole trisiloxame", "DMT")
df$ING <- str_replace(df$ING,"drometrizole trisiloxane", "DMT")
df$ING <- str_replace(df$ING,"drometrizole tri-siloxane", "DMT")
df$ING <- str_replace(df$ING,"dro-metrizole trisiloxane", "DMT")
df$ING <- str_replace(df$ING,"drometrizole trisiloxane ", "DMT")
df$ING <- str_replace(df$ING,"drometrizole trisilozane", "DMT")
df$ING <- str_replace(df$ING,"drometrizole trisloxane", "DMT")
df$ING <- str_replace(df$ING,"drometrizole trisolixane", "DMT")
df$ING <- str_replace(df$ING,"drometrizole trisoloxane", "DMT")
df$ING <- str_replace(df$ING,"drometrizole trixiloxane", "DMT")
df$ING <- str_replace(df$ING,"drometrizole trixyloxane", "DMT")
df$ING <- str_replace(df$ING,"drometrizolettrisiloxane", "DMT")
df$ING <- str_replace(df$ING,"drometrizoltrisiloxane", "DMT")
df$ING <- str_replace(df$ING,"dromettrizole trisiloxane", "DMT")
df$ING <- str_replace(df$ING,"dromiceius oil", "DMT")
df$ING <- str_replace(df$ING,"drommetrizole trisiloxane", "DMT")
df$ING <- str_replace(df$ING,"dromtrizole trisolcane", "DMT")
df$ING <- str_replace(df$ING,"ethyl hexylsalicylate", "EHS")
df$ING <- str_replace(df$ING,"ethyl salicylate", "EHS")
df$ING <- str_replace(df$ING,"ethylhexyl salicylate", "EHS")
df$ING <- str_replace(df$ING,"octylsalicylate", "EHS")
df$ING <- str_replace(df$ING,"octyl salicylate", "EHS")
df$ING <- str_replace(df$ING,"ocytl salicylate", "EHS")
df$ING <- str_replace(df$ING,"oxtyl salicylate", "EHS")
df$ING <- str_replace(df$ING,"2 ethylhexyl salicilate", "EHS")
df$ING <- str_replace(df$ING,"2 ethylhexyl saliclate", "EHS")
df$ING <- str_replace(df$ING,"actisalate", "EHS")
df$ING <- str_replace(df$ING,"ethyhexyl salicylate", "EHS")
df$ING <- str_replace(df$ING,"octly salicylate", "EHS")
df$ING <- str_replace(df$ING,"ethyheyl silicylate", "EHS")
df$ING <- str_replace(df$ING,"octilsalate ", "EHS")
df$ING <- str_replace(df$ING,"ethylhexyl silacylate", "EHS")
df$ING <- str_replace(df$ING,"ethylhexyl silylate ", "EHS")
df$ING <- str_replace(df$ING,"ethylhexyl silylated", "EHS")
df$ING <- str_replace(df$ING,"ethylhexyl silylated", "EHS")
df$ING <- str_replace(df$ING,"ethylhexyl silylated ", "EHS")

```

```

df$ING <- str_replace(df$ING,"ethyl hexyl salicate", "EHS")
df$ING <- str_replace(df$ING,"ethylhexyl salicyclate", "EHS")
df$ING <- str_replace(df$ING,"2-ethylhexylsalixylate ", "EHS")
df$ING <- str_replace(df$ING,"octysalate", "EHS")
df$ING <- str_replace(df$ING,"octilisalate", "EHS")
df$ING <- str_replace(df$ING,"octyl salicilate", "EHS")
df$ING <- str_replace(df$ING,"octyl salicyclate", "EHS")
df$ING <- str_replace(df$ING,"octyl silicate", "EHS")
df$ING <- str_replace(df$ING,"octyl silyate ", "EHS")
df$ING <- str_replace(df$ING,"octyl silylate ", "EHS")
df$ING <- str_replace(df$ING,"octyl silylated ", "EHS")
df$ING <- str_replace(df$ING,"octisalate", "EHS")
df$ING <- str_replace(df$ING,"octisalate", "EHS")
df$ING <- str_replace(df$ING,"octisilate ", "EHS")
df$ING <- str_replace(df$ING,"octislalate", "EHS")
df$ING <- str_replace(df$ING,"octislate ", "EHS")
df$ING <- str_replace(df$ING,"octisolate", "EHS")
df$ING <- str_replace(df$ING,"octlsalate ", "EHS")
df$ING <- str_replace(df$ING,"octuisalate", "EHS")
df$ING <- str_replace(df$ING,"octosalate ", "EHS")
df$ING <- str_replace(df$ING,"ocisalate ", "EHS")
df$ING <- str_replace(df$ING,"octilsalato", "EHS")
df$ING <- str_replace(df$ING,"octyl dimethyl p-aminobenzoic acid", "EHDP")
df$ING <- str_replace(df$ING,"ethylhexyl p-dimethylamino benzoate", "EHDP")
df$ING <- str_replace(df$ING,"octyldimethyl p-aminobenzoic acid", "EHDP")
df$ING <- str_replace(df$ING,"dimethyl octyl paba", "EHDP")
df$ING <- str_replace(df$ING,"ethyl hexyl dimethyl paba", "EHDP")
df$ING <- str_replace(df$ING,"ethylhexyl dimethyl paba", "EHDP")
df$ING <- str_replace(df$ING,"ethyhexyl dimethyl paba", "EHDP")
df$ING <- str_replace(df$ING,"ethylhexyl dimethyl paba", "EHDP")
df$ING <- str_replace(df$ING,"octyldimethyl paba", "EHDP")
df$ING <- str_replace(df$ING,"pentyl dimethyl paba", "EHDP")
df$ING <- str_replace(df$ING,"pentyl dimethyl paba", "EHDP")
df$ING <- str_replace(df$ING,"padimate", "EHDP")
df$ING <- str_replace(df$ING,"homomenthyl salicylate", "HMS")
df$ING <- str_replace(df$ING,"homomethyl salicylate", "HMS")
df$ING <- str_replace(df$ING,"homo salate", "HMS")
df$ING <- str_replace(df$ING,"homo salted", "HMS")
df$ING <- str_replace(df$ING,"homosalate", "HMS")
df$ING <- str_replace(df$ING,"homosalato", "HMS")
df$ING <- str_replace(df$ING,"homosalute", "HMS")
df$ING <- str_replace(df$ING,"homoslaate ", "HMS")
df$ING <- str_replace(df$ING,"homoslate", "HMS")
df$ING <- str_replace(df$ING,"homoslate oxybenzoate", "HMS")
df$ING <- str_replace(df$ING,"homosolate", "HMS")
df$ING <- str_replace(df$ING,"homsalate", "HMS")
df$ING <- str_replace(df$ING,"honosalate", "HMS")
df$ING <- str_replace(df$ING,"phenyl banzimidazole", "PBSA")
df$ING <- str_replace(df$ING,"phenyl benximidazole", "PBSA")
df$ING <- str_replace(df$ING,"phenyl benziimidazole sulfonic acid", "PBSA")
df$ING <- str_replace(df$ING,"phenyl benzimidazol sulfonate", "PBSA")
df$ING <- str_replace(df$ING,"phenyl benzimidazol sulfonic acid", "PBSA")
df$ING <- str_replace(df$ING,"phenyl benzimi-dazole-5-sulfonic acid", "PBSA")

```



```
df <- df %>% filter(!df$total == "0") # Eliminar celdas vacías
df$total <- NULL
```

```
# Resumen de los datos obtenidos
summary(df)
```

```
## INGREDIENTES          FPS          PAIS          REGION
## Length:14241         Min.   : 6.00   Length:14241   Length:14241
## Class :character     1st Qu.: 30.00   Class :character Class :character
## Mode  :character     Median : 50.00   Mode  :character Mode  :character
##                               Mean  : 42.56
##                               3rd Qu.: 50.00
##                               Max.   :130.00
##      PRECIO          ANIO          MES          BP3
## Min.   : 0.16       Min.   :2002     Min.   : 1.000   Min.   :0.0000
## 1st Qu.: 6.33       1st Qu.:2015     1st Qu.: 4.000   1st Qu.:0.0000
## Median : 12.67      Median :2018     Median : 6.000   Median :0.0000
## Mean   : 28.34      Mean   :2018     Mean   : 6.062   Mean   :0.1616
## 3rd Qu.: 29.93     3rd Qu.:2020     3rd Qu.: 8.000   3rd Qu.:0.0000
## Max.   :1083.33     Max.   :2022     Max.   :12.000   Max.   :1.0000
##      EHMC          OCR          MBC          HMS
## Min.   :0.00000     Min.   :0.00000   Min.   :0.00000   Min.   :0.0000
## 1st Qu.:0.00000     1st Qu.:1.00000   1st Qu.:0.00000   1st Qu.:0.0000
## Median :0.00000     Median :1.00000   Median :0.00000   Median :0.0000
## Mean   :0.4594      Mean   :0.7701    Mean   :0.04494    Mean   :0.3234
## 3rd Qu.:1.00000     3rd Qu.:1.00000   3rd Qu.:0.00000   3rd Qu.:1.0000
## Max.   :1.00000     Max.   :1.00000   Max.   :1.00000   Max.   :1.0000
##      BEMT          BMDBM          BP4          DBT
## Min.   :0.00000     Min.   :0.00000   Min.   :0.000000   Min.   :0.00000
## 1st Qu.:0.00000     1st Qu.:0.00000   1st Qu.:0.000000   1st Qu.:0.00000
## Median :0.00000     Median :1.00000   Median :0.000000   Median :0.00000
## Mean   :0.4539      Mean   :0.6304    Mean   :0.002036    Mean   :0.06889
## 3rd Qu.:1.00000     3rd Qu.:1.00000   3rd Qu.:0.000000   3rd Qu.:0.00000
## Max.   :1.00000     Max.   :1.00000   Max.   :1.000000   Max.   :1.00000
##      DHHB          DMT          EHDP          EHS
## Min.   :0.00000     Min.   :0.000000   Min.   :0.000000   Min.   :0.0000
## 1st Qu.:0.00000     1st Qu.:0.000000   1st Qu.:0.000000   1st Qu.:0.0000
## Median :0.00000     Median :0.000000   Median :0.000000   Median :1.0000
## Mean   :0.3055      Mean   :0.07675    Mean   :0.002879    Mean   :0.5629
## 3rd Qu.:1.00000     3rd Qu.:0.000000   3rd Qu.:0.000000   3rd Qu.:1.0000
## Max.   :1.00000     Max.   :1.000000   Max.   :1.000000   Max.   :1.0000
##      EHT          IMC          MBBT          PBSA
## Min.   :0.000000     Min.   :0.0000000   Min.   :0.0000     Min.   :0.0000
## 1st Qu.:0.000000     1st Qu.:0.0000000   1st Qu.:0.0000     1st Qu.:0.0000
## Median :0.000000     Median :0.0000000   Median :0.0000     Median :0.0000
## Mean   :0.001966     Mean   :0.0002809    Mean   :0.1259     Mean   :0.1713
## 3rd Qu.:0.000000     3rd Qu.:0.0000000   3rd Qu.:0.0000     3rd Qu.:0.0000
## Max.   :1.000000     Max.   :1.0000000   Max.   :1.0000     Max.   :1.0000
##      TDSA          TiO2          ZnO
## Min.   :0.00000     Min.   :0.0000     Min.   :0.0000
## 1st Qu.:0.00000     1st Qu.:0.0000     1st Qu.:0.0000
## Median :0.00000     Median :0.0000     Median :0.0000
## Mean   :0.05814     Mean   :0.3842     Mean   :0.1207
```

```
## 3rd Qu.:0.00000 3rd Qu.:1.0000 3rd Qu.:0.0000
## Max. :1.00000 Max. :1.0000 Max. :1.0000
```

4.5. Guardado el archivo transformado

Con lo realizado ya se posee disponible el archivo preparado para la realización del *Trabajo de Fin de Máster*

```
# Guardar el archivo transformado para la realización del TFM
```

```
write.csv(df, "C:/Users/gusta/OneDrive/__CURSOS/Nebrija/99_Tesis/TFM_Gustavo_Pittaluga/DATA/df_COMPLETA
```

```
# Primero registros
```

```
head(df[,-1])
```

```
##   FPS      PAIS      REGION PRECIO ANIO MES BP3 EHMC OCR MBC HMS
## 1  25      Greece      Europe   6.41 2002  10  0  0  1  1  0
## 2  20      Greece      Europe   7.60 2002  10  0  0  1  1  0
## 3  30 South Africa Middle East & Africa 10.96 2002  11  0  1  1  0  0
## 4  22      Japan      Asia Pacific 23.38 2003   4  0  1  1  0  0
## 5  50  Argentina      Latin America  4.34 2005  12  0  0  1  1  0
## 6  56  Thailand      Asia Pacific 22.93 2006   2  0  1  0  0  0
##   BEMT BMDBM BP4 DBT DHHB DMT EHDP EHS EHT IMC MBBT PBSA TDSA TiO2 ZnO
## 1   0     1  0  0  0  0  0  0  0  0  0  0  0  1  0
## 2   0     1  0  1  0  0  0  0  0  0  0  1  0  1  0
## 3   1     0  0  0  0  0  0  0  0  0  0  1  0  1  0
## 4   0     1  0  0  0  0  0  0  0  0  0  0  0  0  0
## 5   0     1  0  0  0  0  0  0  0  0  0  1  0  1  0
## 6   0     0  0  0  0  0  0  0  0  0  0  0  0  1  0
```