**Análisis de Precios de Medicamentos en 2022: compresión de las variaciones del mercado Farmacéutico**

Los datos analisados pertenecen al conjunto de datos del "Termómetro de Precios de Medicamentos", que facilita la comparación de los precios equivalentes por tableta o cápsula de diversas alternativas en el mercado para un mismo medicamento. Los precios registrados en este conjunto de datos corresponden al año 2022.

Los datos fueron obtenidos de: <https://www.datos.gov.co/Salud-y-Protecci-n-Social/Clicsalud-Term-metro-de-Precios-de-Medicamentos/n4dj-8r7k>

La base de datos incluye la siguiente información:

* *Expediente\_INVIMA:* Número de expediente otorgado por el INVIMA (Instituto Nacional de Vigilancia de Medicamentos y Alimentos) para el medicamento.
* *Principio\_activo:* El principio activo del medicamento, es decir, la sustancia química responsable de su efecto terapéutico.
* *Concentración:* La concentración del principio activo en el medicamento, expresada en unidades específicas.
* *Unidad\_base:* La unidad de medida base en la que se expresa la concentración del principio activo.
* *Unidad\_de\_dispensación:* La unidad en la que el medicamento se dispensa al paciente, como tabletas, cápsulas, ampollas, etc.
* *Nombre\_comercial:* El nombre comercial del medicamento, que es utilizado para su identificación en el mercado.
* *Fabricante:* El nombre del fabricante o empresa responsable de producir el medicamento.
* *Medicamento:* El nombre genérico del medicamento, que se basa en el principio activo y la concentración.
* *Canal:* El canal de distribución o venta del medicamento, que puede incluir institucional o comercial.
* *Precio\_por\_tableta:* El precio unitario por tableta o cápsula del medicamento en el mercado.
* *Factoresprecio:* Factores relacionados con la determinación del precio del medicamento.
* Numerofactor: Número que representa un factor específico en relación con el precio del medicamento.

Estos datos proporcionan una visión detallada de los precios de medicamentos en el mercado durante el año 2022 y permiten realizar análisis y comparaciones significativas para comprender mejor el panorama de precios en la industria farmacéutica. En el siguiente informe, se explorarán los hallazgos claves derivados de este conjunto de datos y su importancia en la toma de decisiones en el campo de la salud y de la química farmacéutica.

Inicialmente el proceso consistió en la limpieza de la base de datos utilizando la librería de Pandas. Se verificó que la base de datos contenía un total de 11,874 registros y no se encontraron datos faltantes en ninguna de las columnas. Se ajustó el tipo de dato en algunas columnas, en particular en "precio\_por\_tableta", para eliminar las comas como separadores de miles y convertir los valores a tipo flotante. No se identificaron filas duplicadas en la base de datos. Además, se agregaron nuevas columnas con valores booleanos según las consideraciones específicas realizadas durante el análisis de los datos. Estos procesos fueron cruciales para permitir el análisis.

El análisis de correlación se realizó utilizando la librería Matplotlib. La columna de interés para el análisis fue la de concentración, la cual presentaba una mezcla de cadenas de caracteres y valores numéricos. Con el objetivo de llevar a cabo las correlaciones utilizando esta columna como variable, se procedió a crear una nueva columna que albergara exclusivamente los valores numéricos, esta nueva columna se llamó 'concentracion\_numero'.

Al analizar la correlación entre las variables 'concentracion\_numero' y 'expediente\_INVIMA' se observa una relación extremadamente débil o prácticamente nula entre ambas en el conjunto de datos. El coeficiente de correlación de 0.0042 indica que no hay una asociación lineal significativa entre estos dos atributos.

Esta falta de correlación lineal sugiere que las variaciones en la concentración de los medicamentos no están acompañadas por patrones predecibles en los valores del expediente INVIMA. Por lo tanto, al realizar análisis o toma de decisiones relacionadas con estas variables, es poco probable que uno pueda usar la información de una variable para hacer inferencias confiables sobre la otra.

En otras palabras, el 0.42% de las variaciones del expediente Invima se debe a las variaciones en la concentración de los medicamentos, el restante 99.58% se debe a otros factores y por lo tanto no hay asociación estadística entre ambas variables.

La correlación entre 'concentracion\_numero' y 'precio\_por\_tableta' es de -0.0018, este resultado sugiere una relación extremadamente débil. Un valor negativo indica que, en general, cuando una variable aumenta, la otra tiende a disminuir ligeramente, y viceversa. Sin embargo, el valor es tan cercano a cero que esta relación es prácticamente insignificante. Por lo tanto, no existe una relación lineal apreciable entre estas dos variables en el conjunto de datos analizado.

Así mismo, la correlación calculada de 0.0190 entre 'precio\_por\_tableta' y 'expediente\_invima' sugiere una relación extremadamente débil entre estas dos variables.

Con la ayuda de la librería Seaborn, se logró realizar análisis más detallados:

Se puede observar que las tabletas o cápsulas son las que más se dispensan. Esta observación sugiere que los medicamentos en forma de tabletas o cápsulas son preferidos o prescritos con mayor frecuencia en comparación con otras formas de presentación, como soluciones líquidas, inyecciones u otras. La popularidad de las tabletas o cápsulas puede deberse a su comodidad en términos de dosificación, almacenamiento y administración. Este análisis también podría indicar que los medicamentos en forma de tabletas o cápsulas podrían ser más eficientes en términos de producción y distribución en comparación con otras formas de presentación.

Se realizó también un mapa de calor para analizar la correlación entre “precio\_por\_tableta” y “concentracion\_numero” El mapa de calor es una matriz cuadrada con dos filas y dos columnas. Las filas y columnas están etiquetadas como “precio\_por\_tableta” y “concentracion\_numero”. Los elementos diagonales de la matriz son rojos y tienen un valor de 1. Los elementos fuera de la diagonal de la matriz son azules y tienen un valor de -0.0018. Esto indica una correlación muy cercana a cero entre las dos variables. Esta correlación casi nula sugiere que no hay una relación lineal fuerte entre el precio por tableta y la concentración numérica. Este análisis ya se había realizado con las gráficas de Matplotlib.

Los resultados muestran que el canal institucional dispensa una cantidad significativamente mayor de medicamentos en comparación con el canal comercial. Esta discrepancia sugiere que los medicamentos se distribuyen principalmente en entornos médicos y clínicos, donde los tratamientos especializados son más frecuentes. Por otro lado, el canal comercial parece estar orientado al público general y a condiciones menos complejas. Esta diferencia en la distribución refleja una posible división entre tratamientos especializados y más comunes, influenciada por regulaciones, demanda y preferencias del sector de la salud.

El gráfico de cajas muestra la distribución de concentraciones por principio activo para los 2 principios más comunes, los cuales son ibuprofeno y acetaminofén. Estos gráficos permiten visualizar la mediana, los cuartiles y posibles valores atípicos en cada principio activo. El 50% de las dispensaciones (mediana) de ibuprofeno es aproximadamente 450, mientras que para el acetaminofén es cercana a 400. Esto sugiere que las concentraciones típicas de ambos medicamentos están ligeramente diferenciadas, con el ibuprofeno teniendo una mediana ligeramente más alta. Al considerar el tercer cuartil, que representa el 75% de las dispensaciones, notamos una diferencia más marcada. El tercer cuartil de las concentraciones de ibuprofeno alcanza alrededor de 800, lo que indica que la mayoría de las dispensaciones para este medicamento presentan concentraciones relativamente más altas. En contraste, para el acetaminofén, el tercer cuartil se sitúa en torno a 500, lo que sugiere una concentración más uniforme en comparación con el ibuprofeno. Además, la presencia de valores atípicos en ambos boxplots resalta la existencia de algunas dispensaciones con concentraciones inusuales.

El análisis de datos con la librería numpy permitió hacer las siguientes observaciones:

La media del precio por tableta, que es de 169090.34 unidades, proporciona un valor central

alrededor del cual se agrupan los precios individuales. Sin embargo, la desviación estándar significativamente alta, de 2801158.67 unidades, indica una dispersión amplia en torno a la media. Esto sugiere que los precios por tableta varían considerablemente entre los diferentes medicamentos. La existencia de una desviación estándar tan amplia podría estar influenciada por una serie de factores, como la presencia de medicamentos especializados con precios extremadamente altos, así como opciones más asequibles.

En cuanto a la concentración de los medicamentos, el valor mínimo de 0.0 sugiere la presencia de medicamentos con concentraciones extremadamente bajas o incluso nulas. Esto podría ser el resultado de medicamentos en presentaciones líquidas altamente diluidas o de productos donde la concentración no es relevante para su efectividad. El valor máximo de 30000000.0 indica la presencia de medicamentos con concentraciones muy altas. Es importante considerar que tales concentraciones extremadamente altas podrían estar relacionadas con productos altamente concentrados, como soluciones concentradas o extractos, que requieren administraciones cuidadosas y especializadas.

Se hallaron los 3 medicamentos más costosos. Los resultados sugieren que estos medicamentos están asociados con tratamientos especializados o condiciones médicas que requieren opciones terapéuticas más avanzadas y, por ende, más costosas. Spinraza es un medicamento inyectable que se usa para tratar la atrofia muscular espinal, una enfermedad genética que afecta a los nervios que controlan los músculos. El costo de Spinraza es de $750,000 por año, lo que lo convierte en el medicamento más caro del mundo. Remodulin es un medicamento intravenoso que se usa para tratar la insuficiencia cardíaca congestiva, una condición en la que el corazón no puede bombear suficiente sangre. El costo de Remodulin es de $100,000 por año. Juxtapid es un medicamento oral que se usa para tratar la enfermedad renal policística, una condición en la que los riñones se llenan de quistes. El costo de Juxtapid es de $60,000 por año.

También se encontró que los medicamentos en forma de tabletas o cápsulas tienden a ser más asequibles debido a su proceso de fabricación más simple, facilidad de almacenamiento y transporte, y comodidad para los pacientes al tomarlos. En contraste, los medicamentos en forma de soluciones o suspensiones inyectables suelen ser más costosos debido a la complejidad en su fabricación, necesidad de cuidados especiales en almacenamiento y transporte, y la exigencia de equipos especializados para su administración.

Se desarrolló un tablero interactivo en Power BI que permite visualizar los hallazgos de manera más dinámica y detallada. Se puede acceder al tablero Power BI en el siguiente enlace: <https://app.powerbi.com/view?r=eyJrIjoiMTJmMzI1ODEtYTI0Yi00ZmE2LWIwZDQtZTUxMGY2MzNkYmZlIiwidCI6Ijk5ZTFlNzIxLTcxODQtNDk4ZS04YWZmLWIyYWQ0ZTUzYzFjMiIsImMiOjR9>

La base de datos revela una diversidad notable en la industria farmacéutica en el año 2022. Se registraron un total de 1,821 principios activos distintos, evidenciando la amplia variedad de componentes farmacéuticos utilizados en los medicamentos disponibles en el mercado. Asimismo, se identificaron 445 fabricantes, con Tecnoquímicas, Procaps y MegaLabs sobresaliendo como los principales, lo que indica una competencia sustancial en la industria.

La abundancia de opciones para pacientes y profesionales de la salud se refleja en los 7,178 medicamentos distintos registrados en la base de datos. Esta variedad puede ser beneficiosa al permitir una mayor selección y personalización de tratamientos.

En términos de distribución, el 52% de los medicamentos se obtienen a través del sector comercial, mientras que el 48% restante se suministra a través del canal institucional. Esta división es crucial para comprender cómo se accede y se adquieren la mayoría de los medicamentos, lo que podría influir en la disponibilidad y asequibilidad de los tratamientos.

Entre los medicamentos de mayor costo se encuentran Spinraza, Ramodulin y Juxtapid, lo que plantea consideraciones importantes sobre el acceso a la atención médica y los costos terapéuticos. Además, se observa que Biogen lidera en términos de precios por tableta de medicamento, seguido de cerca por Ferrer y Roche.

El canal institucional presenta precios más elevados por tableta en comparación con el canal comercial. Esta diferencia podría atribuirse a las dinámicas de precios y los acuerdos de compra que operan en entornos institucionales, como hospitales y clínicas.

Es importante notar que la tableta o cápsula es la unidad de dispensación más comúnmente utilizada. Esto indica que la mayoría de los medicamentos se presentan en formas de dosificación oral, lo que facilita la administración y el cumplimiento por parte de los pacientes.