

UNIVERSIDAD DEL VALLE DE GUATEMALA

Minería de Datos

Sección 30

Ing. Leonel Guillen



Excelencia que trasciende

DEL VALLE
GRUPO EDUCATIVO

Proyecto 1

Minería de Datos

| | |
|------------------|--------|
| Diego Leiva | #21752 |
| Maria Ramirez | #21342 |
| Gustavo González | #21438 |
| Gabriel García | #21352 |
| Pablo Orellana | #21970 |

Fecha de entrega: 08 de abril de 2024

Problemática

Las autoridades económicas y los responsables de la política monetaria del país enfrentan desafíos significativos para diseñar estrategias efectivas que controlen la inflación, ajusten las tasas de interés y manejen adecuadamente otras medidas económicas vitales. Estas decisiones requieren un conocimiento profundo y actualizado de las tendencias y variaciones del Índice de Precios al Consumidor (IPC), tanto a nivel regional como por categorías de productos. La comprensión detallada del comportamiento del IPC es esencial para anticipar cambios en la economía y responder con políticas que promuevan la estabilidad y el crecimiento económico, mientras se protege el poder adquisitivo de la población.

Problema Científico

Investigar y comprender las tendencias y patrones del IPC a nivel regional y por categorías de productos en el país durante los últimos 10 años, sin el análisis de factores socioeconómicos externos. Este estudio se enfocará en desglosar el comportamiento del IPC para identificar las variaciones intrínsecas y las tendencias históricas que podrían indicar dinámicas económicas subyacentes importantes para la toma de decisiones.

Objetivo General

Desarrollar un análisis exhaustivo sobre las tendencias y variaciones del Índice de Precios al Consumidor (IPC) a nivel regional y por categorías de productos en el país durante la última década, con el fin de proporcionar a las autoridades económicas y encargados de la política monetaria información relevante que permita la toma de decisiones informadas orientadas a estabilizar la inflación y ajustar las políticas económicas pertinentes.

Objetivos específicos

1. Visualizar y comparar las tendencias del IPC en las diferentes regiones del país.
2. Identificar las categorías de productos que han experimentado mayores aumentos o disminuciones de precios.
3. Identificar las categorías de productos que han experimentado mayores aumentos o disminuciones de precios, analizando los factores que podrían haber contribuido a dichas variaciones

Descripción de los datos

Este conjunto de datos recopila información sobre el Índice de Precios al Consumidor (IPC) en Guatemala entre el 2011 hasta el 2023. Con un total de 45844 registros distribuidos en 13 variables, este dataset proporciona una visión detallada de las tendencias de precios en el país a lo largo de este periodo de tiempo. Las variables incluyen:

1. Año: Representa el año en que se recopilaron los datos.
2. Mes: Indica el mes correspondiente a cada registro.
3. Código: Identificador numérico único para cada producto o categoría de productos.
4. Descripción: Detalle sobre el producto o categoría de productos al que se refiere el índice de precios.
5. Rep: Valor representativo del IPC para toda la república, siendo el promedio de las siguientes variables.
6. Reg I: Valor representativo del IPC para la región Metropolitana.
7. Reg II: Valor representativo del IPC para la región Norte
8. Reg III: Valor representativo del IPC para la región Nororiental
9. Reg IV: Valor representativo del IPC para la región Suroriental
10. Reg V: Valor representativo del IPC para la región Central
11. Reg VI: Valor representativo del IPC para la región Suroccidental
12. Reg VII: Valor representativo del IPC para la región Noroccidental
13. Reg VIII: Valor representativo del IPC para la región de Petén

Análisis Exploratorio

Variables Cuantitativas:

Dentro del conjunto de datos, se identificaron las siguientes variables cuantitativas: República y todas las regiones (Región I, Región II, Región III, Región IV, Región V, Región VI, Región VII, Región VIII).

Variables Categóricas:

Asimismo, se encontraron las siguientes variables categóricas en el dataset: año, mes, código y descripción. Es importante destacar que, aunque tanto el año como el código están compuestos por números, se tratan como variables categóricas debido a su naturaleza en los datos.

Estadística Descriptiva

El análisis de estadística descriptiva reveló las siguientes observaciones sobre nuestras variables cuantitativas:

- **Conteo:** Se constató un número uniforme de observaciones en todas las regiones, con un total de 45,844 registros para cada una, lo que sugiere una recolección uniforme de datos por región.
- **Media:** Se observaron diferencias significativas en las medias de las regiones, que variaron desde 133.16 en la Región I hasta 308.07 en la Región IV. Esta disparidad podría indicar tendencias distintas o condiciones especiales en la Región IV en comparación con otras regiones.
- **Mínimo:** La presencia de valores nulos en algunas regiones sugiere la posible falta de movimiento en los precios de ciertos productos o la omisión de eventos en esos períodos o regiones.
- **Cuartiles:** Los cuartiles exhibieron una distribución de datos heterogénea entre las regiones. Por ejemplo, la mediana (50%) varió menos entre las regiones en comparación con la media, lo que sugiere la influencia de valores atípicos extremos en la media, especialmente notorio en la Región IV.
- **Máximo:** Se registraron valores máximos drásticamente variables, desde 3992.22 en la Región I hasta 68850.73 en la Región IV, lo que indica una variabilidad significativa y la presencia de valores atípicos extremos, particularmente evidentes en la Región IV.
- **Desviación Estándar:** La desviación estándar reveló una variabilidad destacada de los datos, especialmente en la Región IV en comparación con otras regiones, lo que coincide con las observaciones de la media y los valores máximos.
- **Fechas:** El rango de fechas abarcó desde el 1 de diciembre de 2010 hasta el 1 de diciembre de 2023, indicando un período de 13 años. Los cuartiles de la fecha sugirieron una distribución uniforme de registros a lo largo del tiempo.

Gráficos:

Distribución de las regiones:

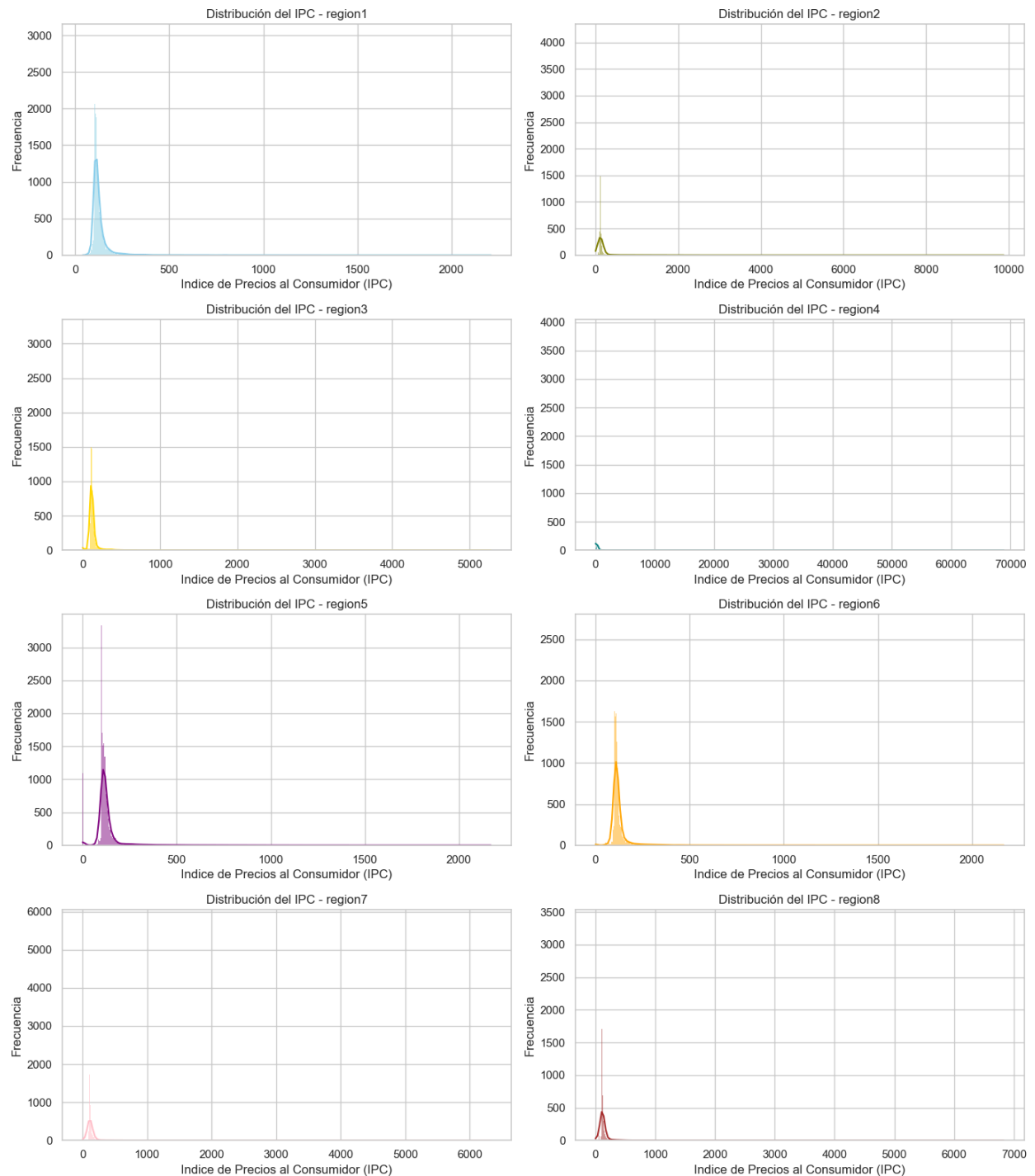


Figura 1. Gráficas de Distribuciones por regiones

En la Figura 1 se puede apreciar claramente la presencia de valores extremadamente altos en varias regiones, particularmente en la región 4, los cuales se apartan notablemente del resto de los datos. Este hallazgo sugiere la posible presencia de valores erróneos o atípicos en el conjunto de datos.

Se observa también que todas las regiones muestran distribuciones sesgadas, algunas hacia la derecha (regiones 1, 3, 5, 7, 8), mientras que otras presentan un sesgo aún más pronunciado, con valores atípicos considerablemente alejados de la media (regiones 2, 4). Esta asimetría en las distribuciones sugiere que el Índice de Precios al Consumidor (IPC) en todas las regiones no sigue una distribución uniforme y está influenciado por factores que varían regionalmente.

Valores Atípicos (outliers):

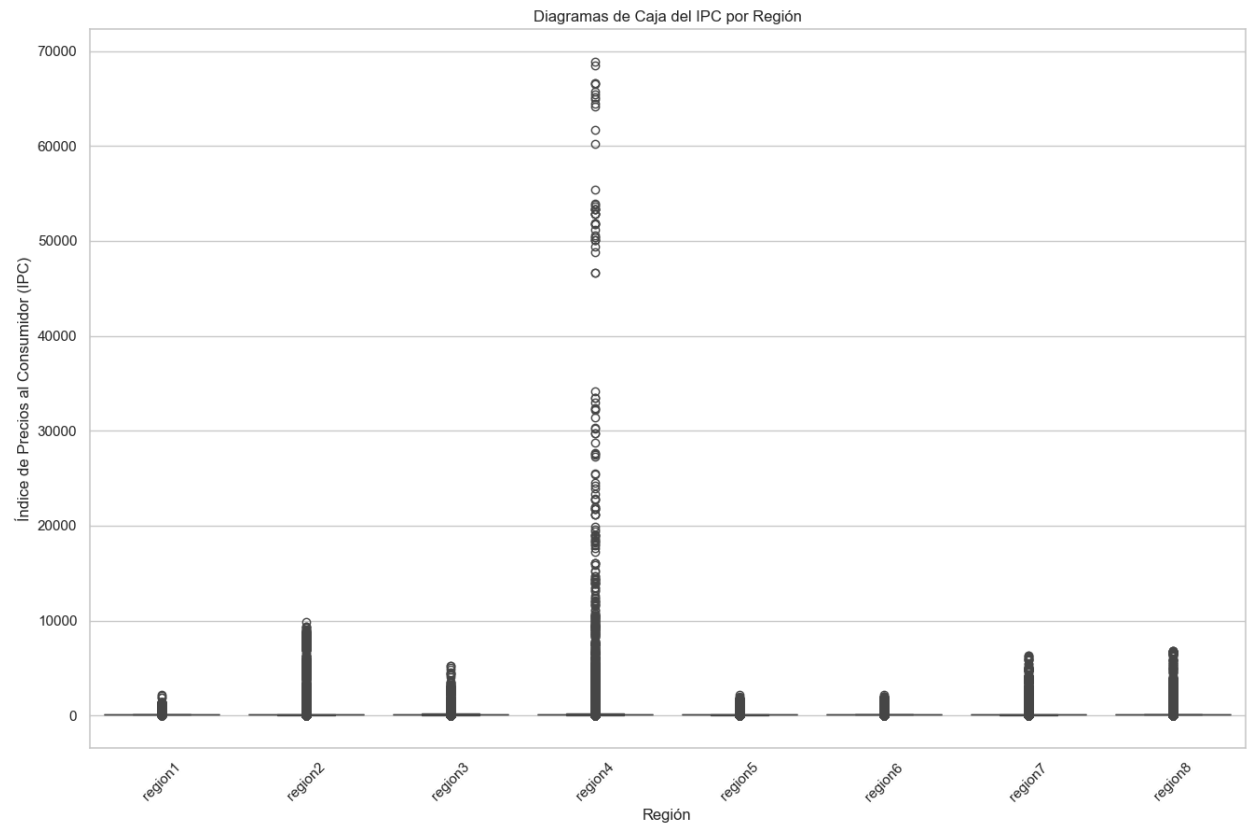


Gráfico 1. Datos Atípicos

El Gráfico 1 revela la presencia significativa de datos atípicos en casi todas las regiones, destacándose la región 4 con la mayor cantidad y los valores más extremos, algunos superando los 60,000.

Además de los valores atípicos, se observa que las medianas de las regiones, con excepción de la región 4, muestran una consistencia notable, lo que sugiere una mediana de IPC similar en diversas partes del país.

En cuanto a la variabilidad, la región 4 exhibe una variabilidad excepcionalmente alta en comparación con las otras regiones, las cuales presentan una variabilidad más contenida.

Por último, se destaca la consistencia en las distribuciones del IPC en todas las regiones, excluyendo la región 4, donde se observan distribuciones más estrechas y consistentes.

Limpieza de Datos Atípicos:

Se busca identificar y eliminar registros con valores atípicos en cualquier región o a nivel de país, asegurando así la consistencia y precisión del análisis al descartar datos extremos que podrían distorsionar los resultados.

Observaciones Posteriores al análisis:

- **Discrepancia Significativa:** Se observa una diferencia notable en la media de la región 4 (20,517.30) en comparación con las medias de otras regiones y la media general del país (220.13). Esta discrepancia sugiere que los valores en la región 4 no solo son inusualmente altos, sino posiblemente erróneos.
- **Comparación con Otras Regiones:** Al contrastar los valores máximos y las medias de la región 4 con otras regiones, se evidencia que los valores en la región 4 son excepcionalmente altos, lo que refuerza la sospecha de errores o datos anómalos.
- **Variabilidad:** La desviación estándar en la región 4 (23,059.95) también sobresale en comparación con otras regiones, indicando una variabilidad extrema poco común que podría ser indicativa de datos erróneos o entradas anómalas.
- **Tipo de Producto:** Además de los hallazgos presentados, es importante destacar que la categoría que registra estos valores anormales corresponde a un producto alimenticio, los "Chiles Picantes". Si bien es cierto que ciertos tipos de chiles pueden tener precios elevados, en un contexto de mercado real, es altamente improbable que este tipo de productos experimenten cambios tan drásticos en sus precios como para justificar un valor de IPC de hasta 68 mil. Esta observación refuerza aún más la sospecha de problemas significativos en los datos de la región 4.

- **Tendencia de Aumento:** Los valores máximos, las medias y las desviaciones estándar en la región 4 muestran una tendencia ascendente año tras año, lo que podría sugerir una acumulación de errores o anomalías a lo largo del tiempo. Está claro que a partir del año 2016, los valores se dispararon.
- **Valores Extremos:** Los valores máximos aumentan dramáticamente con cada año, alcanzando niveles que no parecen ser razonables o típicos para su IPC.
- **Inconsistencia Temporal:** La inconsistencia y la tendencia al alza de los valores en la región 4 no parecen seguir un patrón que se podría esperar de datos genuinos, lo que sugiere que los datos pueden estar comprometidos.

Los análisis indican de manera contundente la presencia de errores significativos o valores atípicos en los datos de la región 4, los cuales no son representativos. Como resultado, se ha tomado la decisión de excluir esta región de los datos con el fin de prevenir posibles distorsiones en los resultados del análisis.

De manera que hemos aplicado un filtro de 4 desviaciones estándar para retener aproximadamente el 99% de los datos, lo que ha permitido una limpieza efectiva del conjunto de datos al eliminar únicamente los valores extremadamente atípicos. Esta medida no ha tenido un impacto significativo en el análisis futuro. Cabe destacar que esta metodología se ha aplicado a un conjunto de datos del cual previamente se excluyó la región 4 debido a sus notables inconsistencias. De esta manera, se garantiza una mayor precisión y fiabilidad en la evaluación de las regiones restantes.

Estadística Descriptiva posterior a la limpieza de valores atípicos:

- **Consistencia en los Conteos:** Se observa una uniformidad en el número de observaciones a través de todas las regiones y el país, totalizando 42,139 registros. Esto indica una base de datos completa y equilibrada para cada región.
- **Promedios Similares:** Las medias de las diferentes regiones y del país son bastante similares, rondando alrededor de 116. Esto sugiere una homogeneidad general en los datos después de la limpieza.
- **Valores Mínimos y Máximos:** Los valores mínimos y máximos muestran una variabilidad razonable entre regiones, sin ningún valor extremadamente alto que sugiera posibles errores o valores atípicos, como era el caso con la región 4 previamente.

- **Dispersión de los Datos:** La desviación estándar en cada región indica la dispersión de los datos alrededor de la media. Estos valores son consistentes y no presentan una variabilidad excesiva.
- **Cuartiles:** Los valores de los cuartiles (25%, 50%, 75%) para cada región y el país indican una distribución de datos coherente y sin distorsiones significativas.

Las estadísticas descriptivas actualizadas muestran ahora un conjunto de datos más coherente y confiable después de la limpieza y la exclusión de la región 4 debido a sus inconsistencias.

Es importante destacar que las diferencias observadas en las categorías reflejan la variabilidad natural de los precios entre distintos productos y servicios. Por ejemplo, es esperable que categorías como vivienda y automóviles presenten precios más elevados en comparación con otros bienes y servicios. Esta disparidad es inherente a la diversidad del mercado, donde los productos y servicios de mayor valor agregado suelen tener precios superiores.

Distribución posterior a la limpieza de valores atípicos:

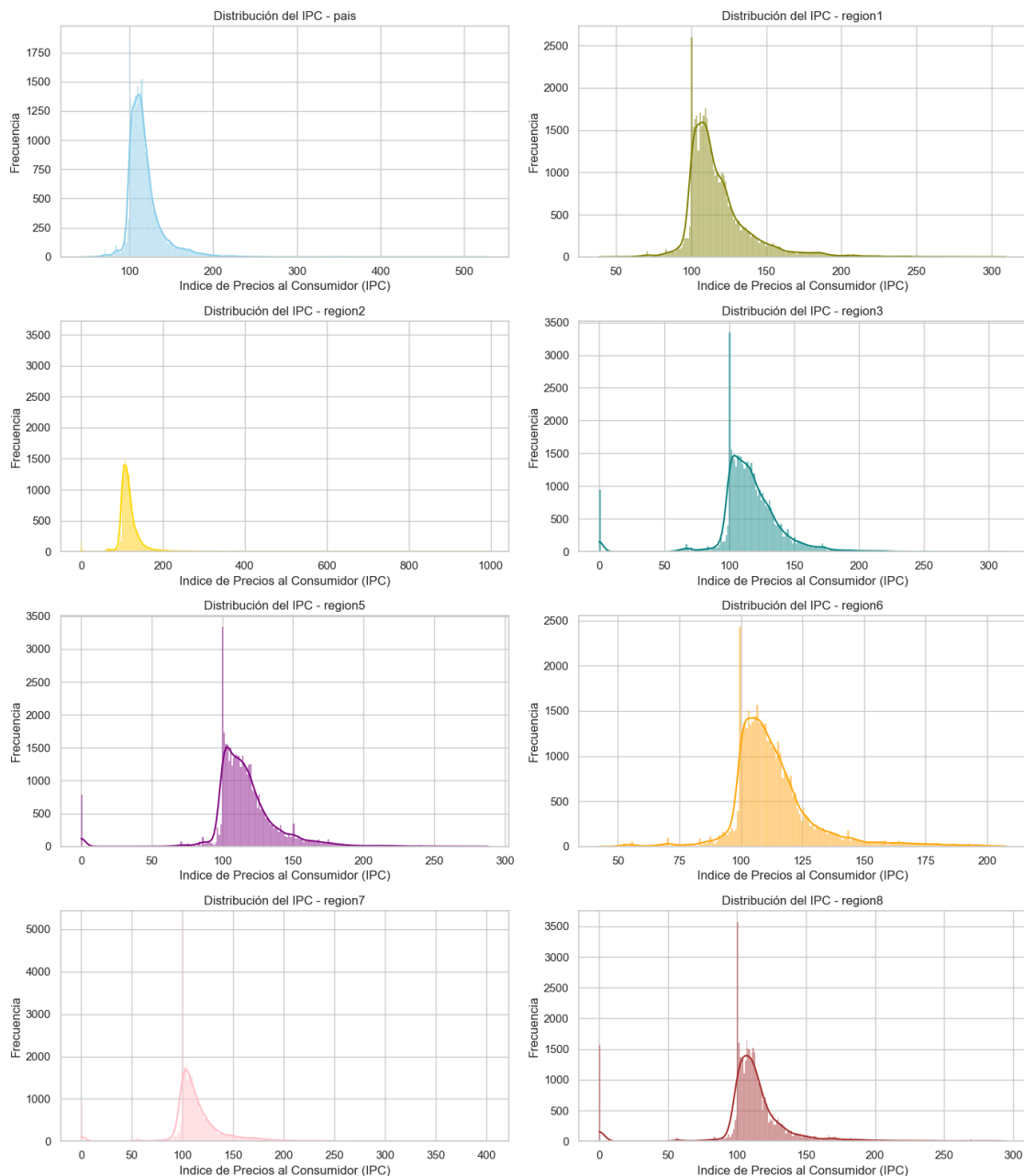
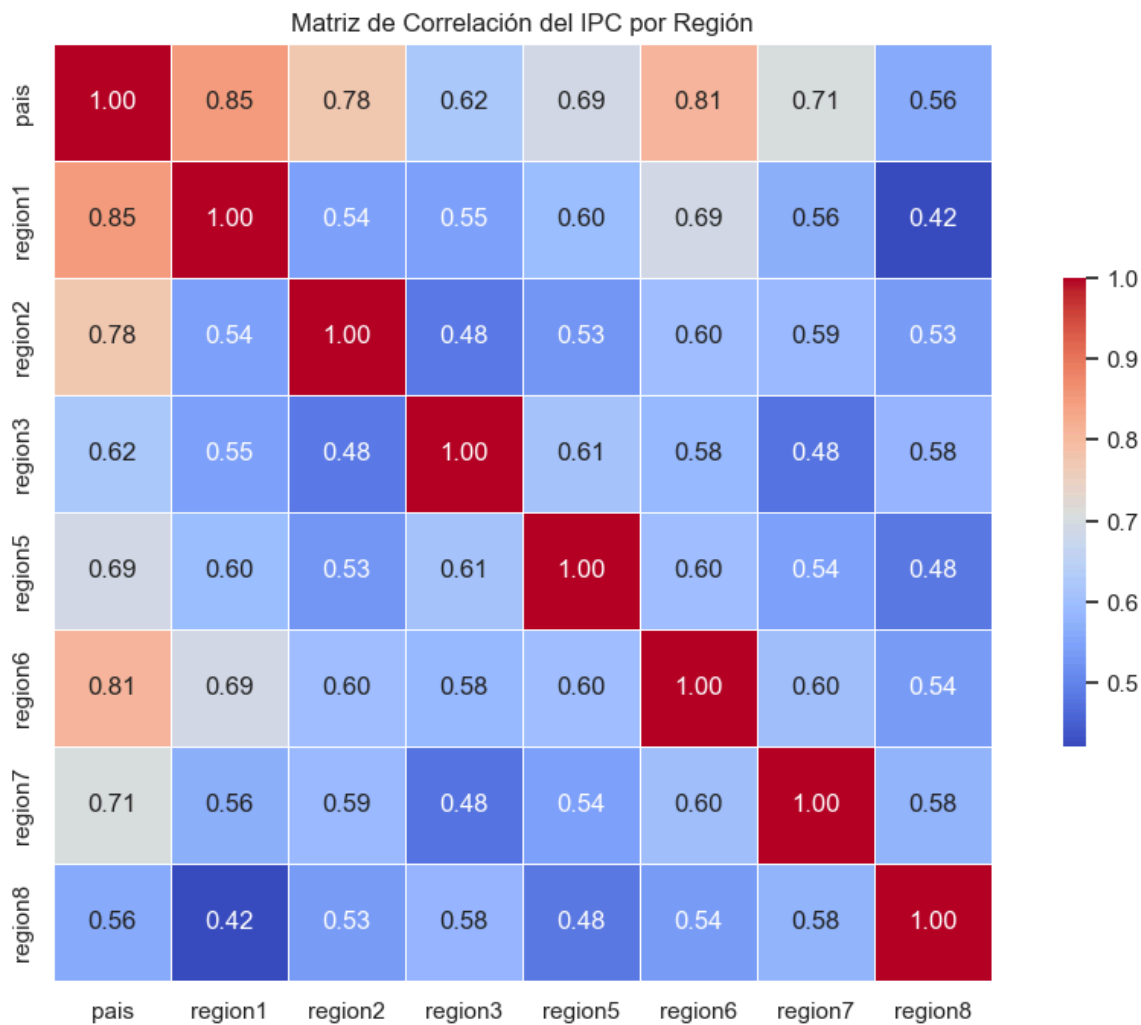


Figura 2. Gráficas de Distribuciones por regiones sin datos atípicos

En la Figura 2, se aprecia una notable diferencia respecto a la Figura 1 tras la limpieza de los datos atípicos. Al comparar entre las regiones, se evidencia una consistencia en las medias y variabilidades. No se observan indicios visuales inmediatos de errores o inconsistencias graves como los que se presentaban anteriormente en la región 4. Cada región exhibe una escala y rango de Índice de Precios al Consumidor (IPC) que se ajusta a un conjunto de datos depurado, sin los extremos que podrían haber estado presentes antes de la limpieza.

Correlación entre variables:



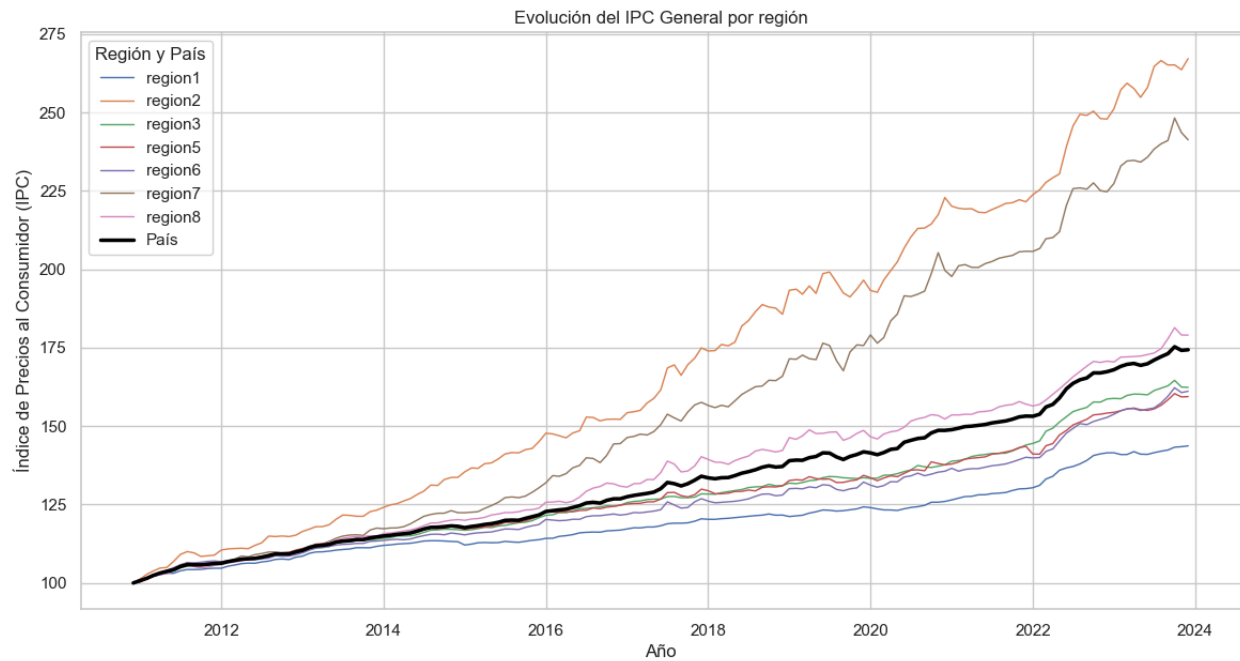
Gráfica 2. Correlación del IPC por regiones

En el análisis de correlación, se destacó que las diversas regiones exhiben variaciones en sus niveles de correlación entre sí, lo que sugiere la posible existencia de tendencias de precios independientes entre regiones. Por ejemplo, la correlación entre la región 1 y la región 2 fue de 0.54, indicando una relación moderada, mientras que la correlación entre la región 1 y la región 6 fue de 0.69, lo que sugiere una relación algo más alta y sugiere que estas regiones tienden a comportarse de manera similar.

Las correlaciones más bajas, como la de la región 1 con la región 8 (0.42), sugieren que el IPC de esas regiones puede estar influenciado por factores más localizados y no necesariamente sigue las mismas tendencias que el país o las demás regiones.

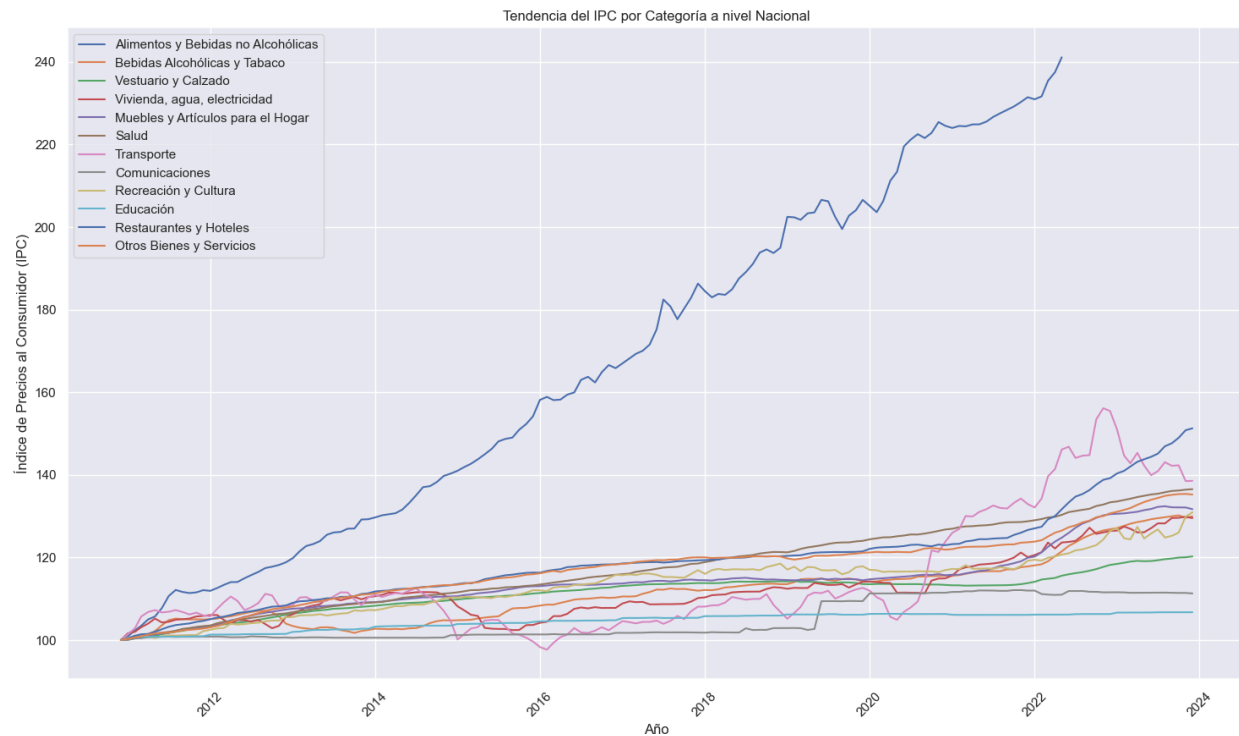
Además, se observó que algunas regiones muestran una alta correlación con el IPC nacional, como la región 1 y la región 6, con valores de 0.85 y 0.81 respectivamente. Esto implica que el IPC de estas regiones tiende a moverse en conjunto con el IPC nacional.

Evaluación de Tendencias:



Gráfica 3. Evolución del IPC General por Región

Tendencia de IPC por categoría a nivel Nacional:

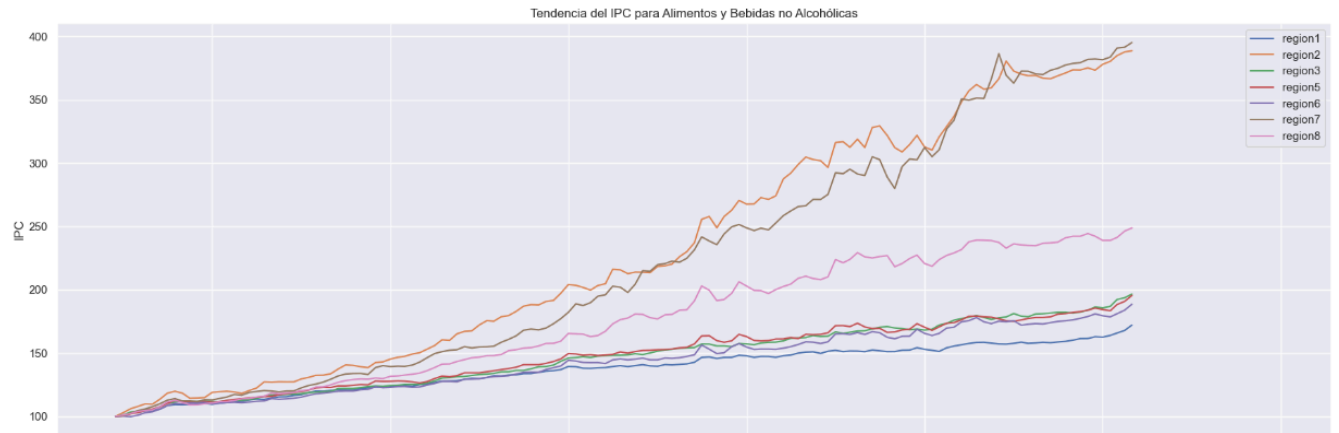


Gráfica 4. Tendencia del ICP por Categoría a nivel nacional

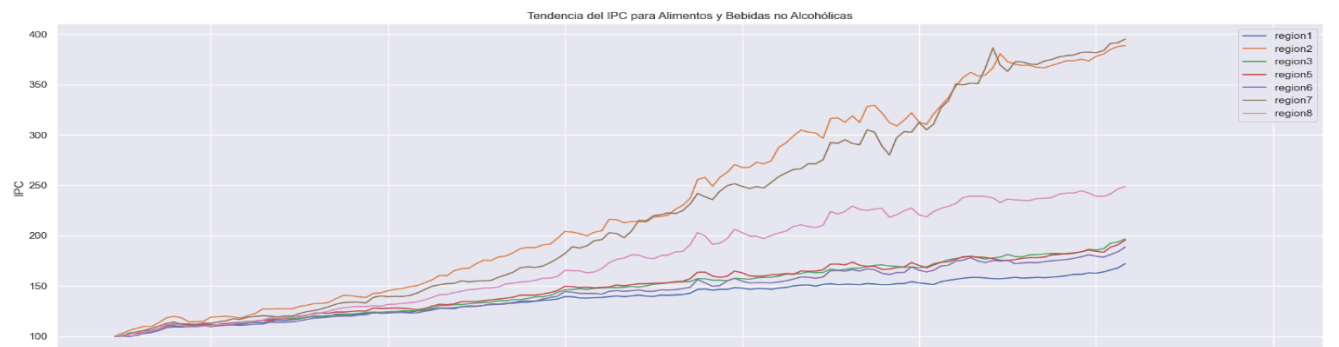
Para analizar la tendencia de la distribución del gasto por región, se han seleccionado cinco de las doce categorías disponibles, dando prioridad a aquellas divisiones de gasto esenciales y fundamentales para el bienestar general de las personas. Estas categorías son las siguientes:

- **Salud:** Esta categoría refleja los costos relacionados con servicios médicos y productos farmacéuticos, que son cruciales para el mantenimiento de la salud física y mental.
- **Educación:** Engloba los gastos en formación y capacitación educativa, que son fundamentales para el desarrollo personal y profesional de las personas.
- **Alimentación:** Representa los gastos en alimentos y bebidas no alcohólicas, que son necesidades primarias para la nutrición y la supervivencia.
- **Vivienda:** Incluye los gastos en alquiler o hipoteca, servicios básicos como agua y electricidad, y otros combustibles, todos vitales para proporcionar un hogar seguro y estable.
- **Ropa:** Comprende el gasto en prendas de vestir y calzado, necesarios para la protección y el confort en la vida cotidiana.

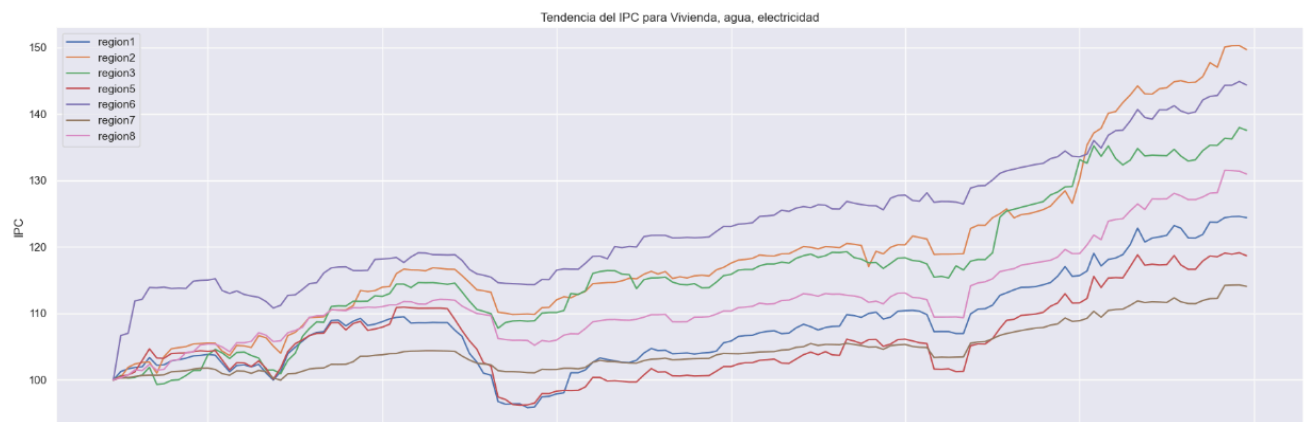
Estas categorías se consideran prioritarias debido a que abarcan las necesidades básicas y son indicativas de la calidad de vida de las personas.



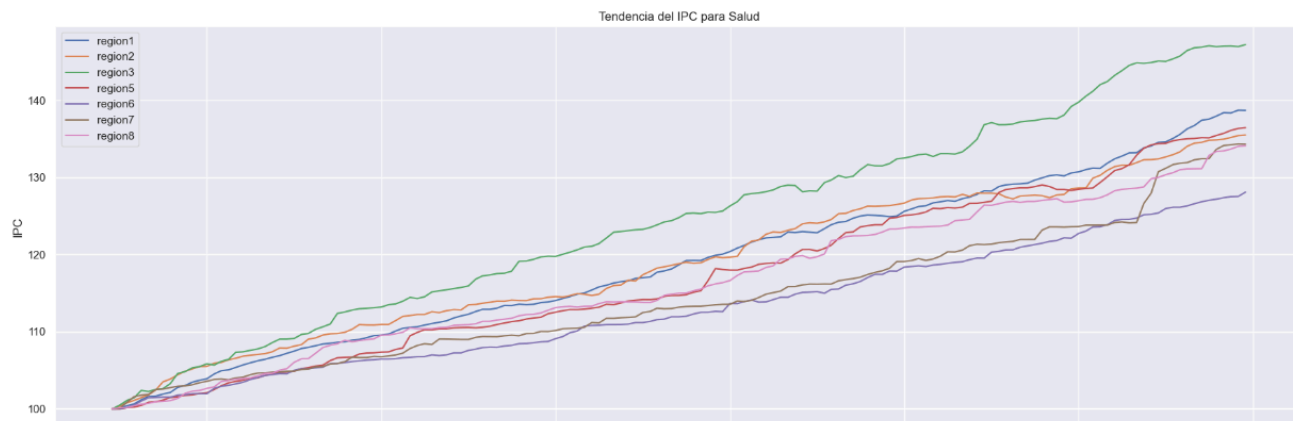
Gráfica 5. Tendencia del IPC para Alimentos y Bebidas no alcohólicas



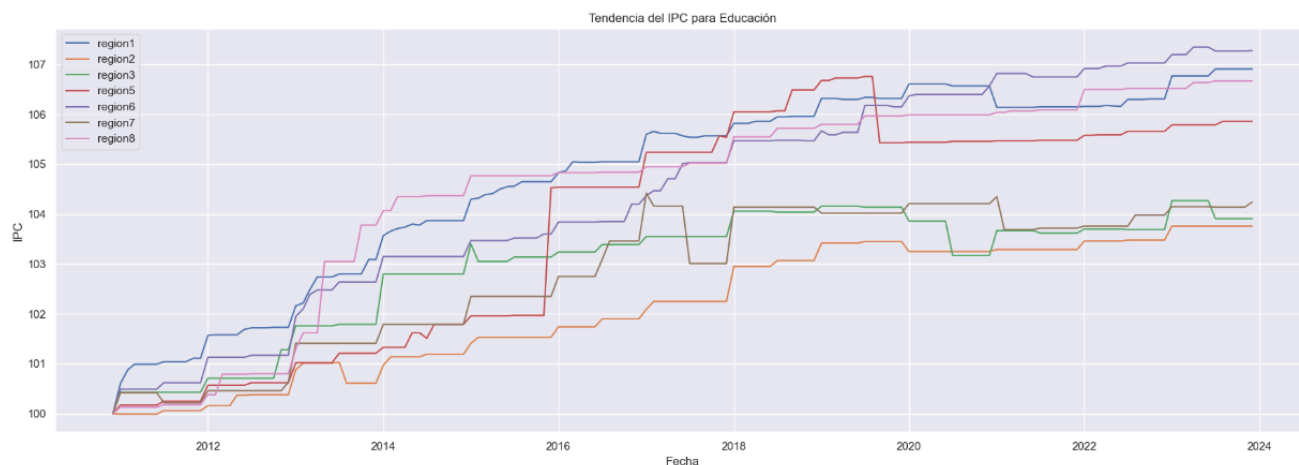
Gráfica 6. Tendencia del IPC para Vestuario y Calzado



Gráfica 7. Tendencia del IPC para vivienda, agua y electricidad



Gráfica 8. Tendencia del IPC para Salud



Gráfica 9. Tendencia del IPC para Educación

Detalle de productos por división de gasto:

1. Categoría de Alimentos y Bebidas no alcohólicas:

Top 3 productos/servicios con mayor IPC

- Aguacate - IPC Promedio: 187.46
- Productos de tortillería - IPC Promedio: 186.64
- Culantro - IPC Promedio: 183.11

Top 3 productos/servicios con menor IPC

- Tomate - IPC Promedio: 101.34
- Productos de repostería - IPC Promedio: 104.77
- Agua envasada - IPC Promedio: 108.24

Top 3 productos/servicios más volátiles:

- Repollo - Variación IPC: 78.99
- Aguacate - Variación IPC: 56.39
- Culantro - Variación IPC: 55.37

Top 3 productos/servicios más estables:

- Productos de repostería - Variación IPC: 2.50
- Agua envasada - Variación IPC: 4.46
- Dulces, chocolates y goma de mascar - Variación IPC: 5.90

2. Categoría de Vestuario y Calzado:

Top 3 productos/servicios con mayor IPC:

- Ropa interior y de dormir para hombre - IPC Promedio: 121.78
- Servicio de reparación de calzado - IPC Promedio: 118.06
- Telas para confección de vestuario - IPC Promedio: 117.99

Top 3 productos/servicios con menor IPC:

- Pijamas para bebé - IPC Promedio: 99.01
- Traje para hombre - IPC Promedio: 99.34
- Servicio de limpieza de prendas de vestir - IPC Promedio: 100.97

Top 3 productos/servicios más volátiles:

- Camiseta para niño - Variación IPC: 10.77
- Servicio de reparación de calzado - Variación IPC: 10.07
- Ropa interior y de dormir para hombre - Variación IPC: 9.30

Top 3 productos/servicios más estables:

- Otras prendas de vestir - Variación IPC: 3.69
- Zapatos para mujer - Variación IPC: 3.76
- Artículos varios de vestuario - Variación IPC: 3.89

3. Categoría de vivienda, agua y electricidad:

Top 3 productos/servicios con mayor IPC:

- Carbón - IPC Promedio: 134.57
- Servicio de mantenimiento de vivienda - IPC Promedio: 124.69
- Leña - IPC Promedio: 121.56

Top 3 productos/servicios con menor IPC:

- Gas propano - IPC Promedio: 97.67
- Alquiler de vivienda - IPC Promedio: 106.53
- Servicio de retiro de basura - IPC Promedio: 107.39

Top 3 productos/servicios más volátiles:

- Carbón - Variación IPC: 21.35
- Gas propano - Variación IPC: 15.08
- Servicio de mantenimiento de vivienda - Variación IPC: 14.41

Top 3 productos/servicios más estables:

- Alquiler de vivienda - Variación IPC: 3.12
- Servicio de retiro de basura - Variación IPC: 4.57
- Artículos para electricidad y grifería - Variación IPC: 4.98

4. Categoría de Salud:

Top 3 productos/servicios con mayor IPC:

- Vendas, parches, apósitos - IPC Promedio: 137.55
- Antigripal - IPC Promedio: 124.26
- Consulta médica - IPC Promedio: 123.10

Top 3 productos/servicios con menor IPC:

- Servicio de hospitalización - IPC Promedio: 108.90
- Vacunas de prevención - IPC Promedio: 112.21
- Examen de radiología y otros - IPC Promedio: 113.17

Top 3 productos/servicios más volátiles:

- Vendas, parches, apósitos - Variación IPC: 23.18
- Antigripal - Variación IPC: 16.41
- Consulta medica - Variación IPC: 15.22

Top 3 productos/servicios más estables:

- Servicio de hospitalización - Variación IPC: 5.17
- Examen de laboratorio clínico - Variación IPC: 6.56
- Lentes de corrección - Variación IPC: 6.68

5. Categoría de Educación

Top 3 productos/servicios con mayor IPC:

- cursos de capacitación - IPC Promedio: 107.77
- servicio de educación preprimaria - IPC Promedio: 103.99
- servicio de educación de primaria - IPC Promedio: 103.93

Top 3 productos/servicios con menor IPC:

- Servicio de educación de la enseñanza media - IPC Promedio: 102.41
- Servicios de enseñanza universitaria - IPC Promedio: 102.54
- Servicio de educación de primaria - IPC Promedio: 103.93

Top 3 productos/servicios más volátiles:

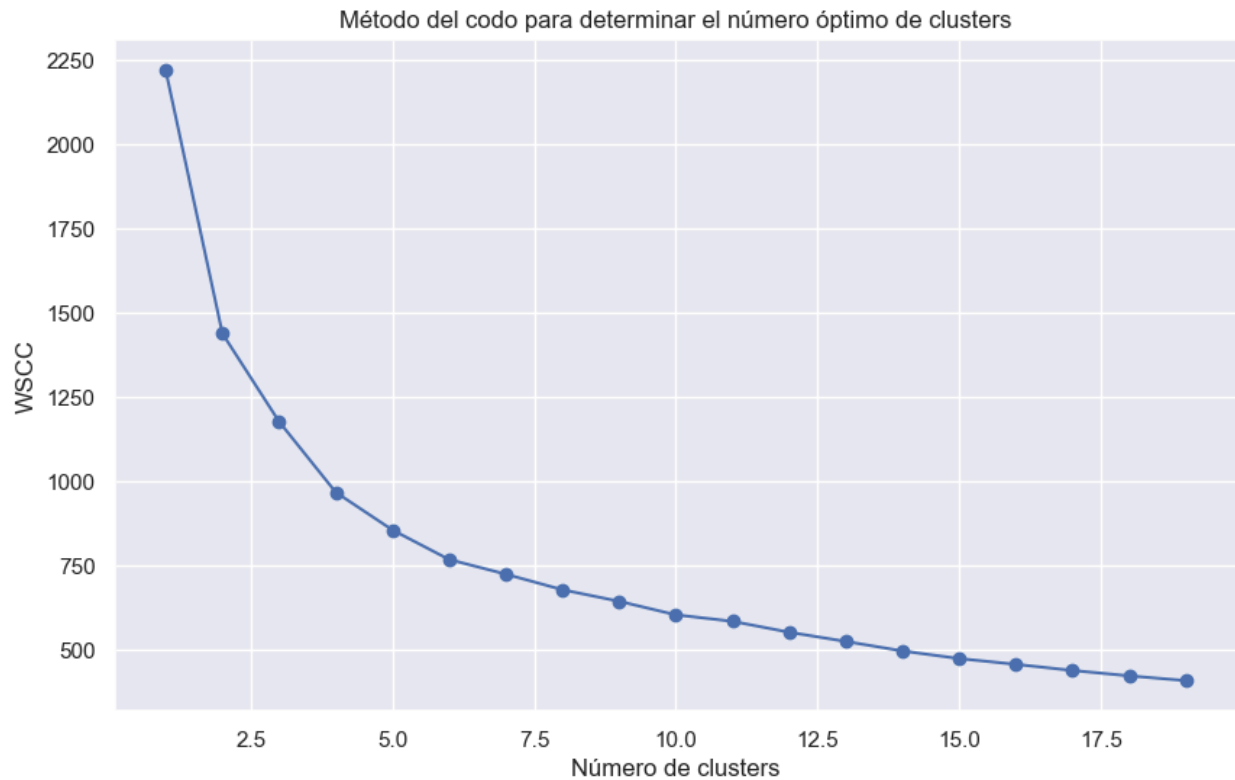
- Cursos de capacitación - Variación IPC: 3.56
- Servicio de educación preprimaria - Variación IPC: 1.99
- Servicio de educación de primaria - Variación IPC: 1.79

Top 3 productos/servicios más estables:

- Servicio de educación de la enseñanza media - Variación IPC: 1.14
- Servicios de enseñanza universitaria - Variación IPC: 1.25
- Servicio de educación de primaria - Variación IPC: 1.79

Clustering de Datos por fechas

Análisis de los grupos:



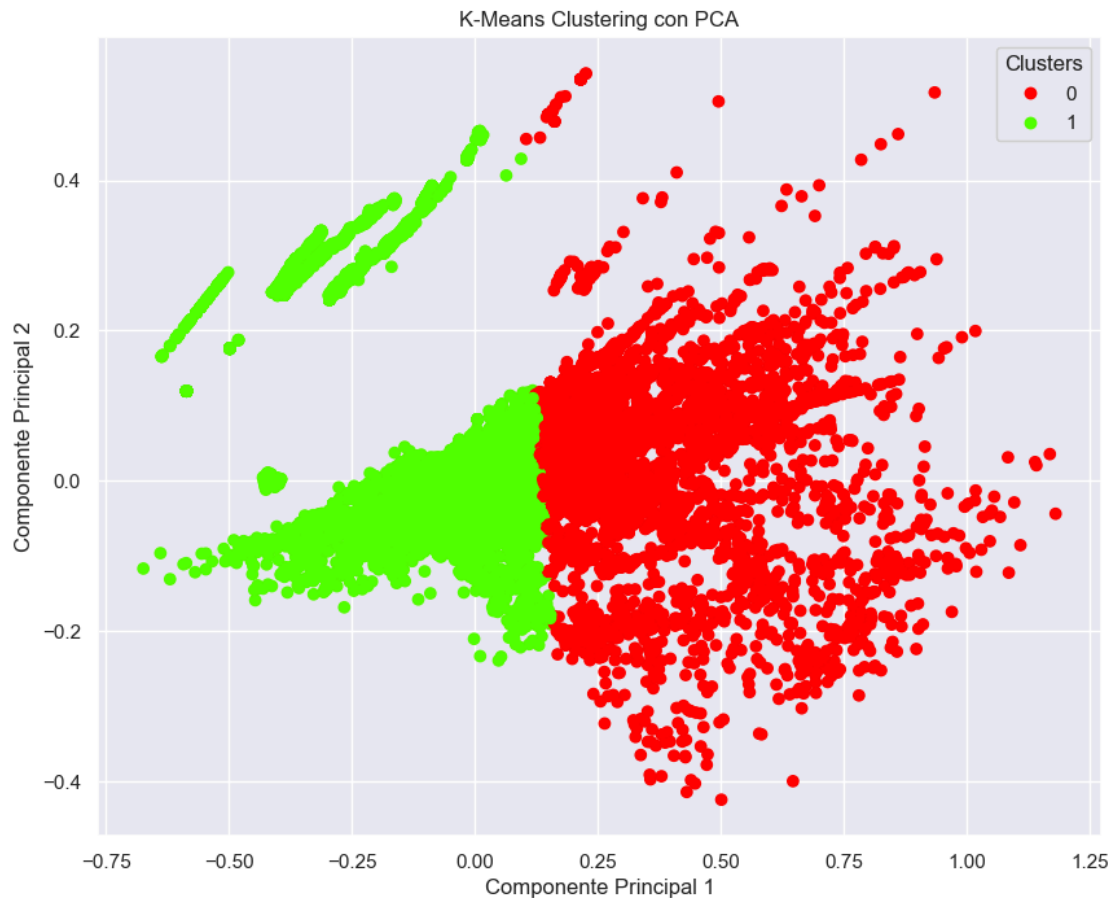
Gráfica 10. Gráfica que muestra los resultados del método del codo

Como se muestra en la Gráfica 10 la curva empieza a aplanarse después del quinto punto en el eje x, indicando que el incremento en la cantidad de clusters más allá de tres no conlleva una mejora sustancial en la inercia. Por ende, se puede inferir que el número óptimo de clusters para este conjunto de datos es cinco.

Análisis de componente principal (PCA):

Dado que se cuenta con múltiples características, la tarea de representar cada una de ellas en un solo gráfico se torna complicada. Por consiguiente, es necesario recurrir al Análisis de Componentes Principales (PCA) para reducir la cantidad de variables a analizar y visualizar. Este método implica la creación de un conjunto de nuevas variables que representan de manera óptima las variables originales.

Clusters formados:



Gráfica 11. Resultados del Clustering

Observando los resultados del clustering (Gráfica 11), podemos interpretar los grupos de la siguiente manera:

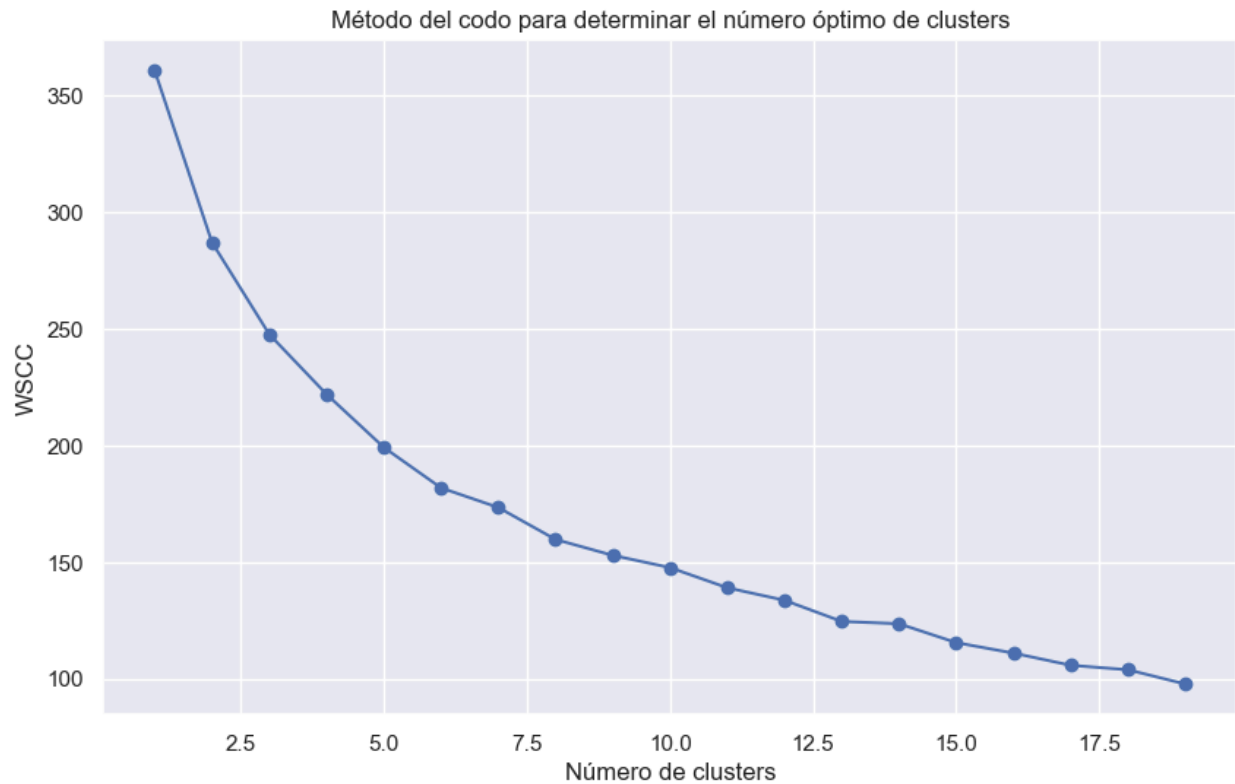
1. Cluster 0 (Fechas con IPC Alto):

Este grupo representa las fechas con un índice de precios al consumidor generalmente más alto en comparación con otras fechas. Indica que durante estos periodos, se experimenta un mayor costo de vida, lo que sugiere una mayor inflación y precios más elevados en una amplia gama de productos y servicios.

2. Cluster 1 (Fechas con IPC moderado):

Este grupo incluye fechas con un IPC más moderado en comparación con el grupo anterior. Aunque el IPC durante estos periodos es más bajo que en el primer cluster, aún se encuentra por encima del promedio nacional. Esto sugiere que el costo de vida durante estos periodos puede ser más accesible en comparación con las fechas del cluster 0.

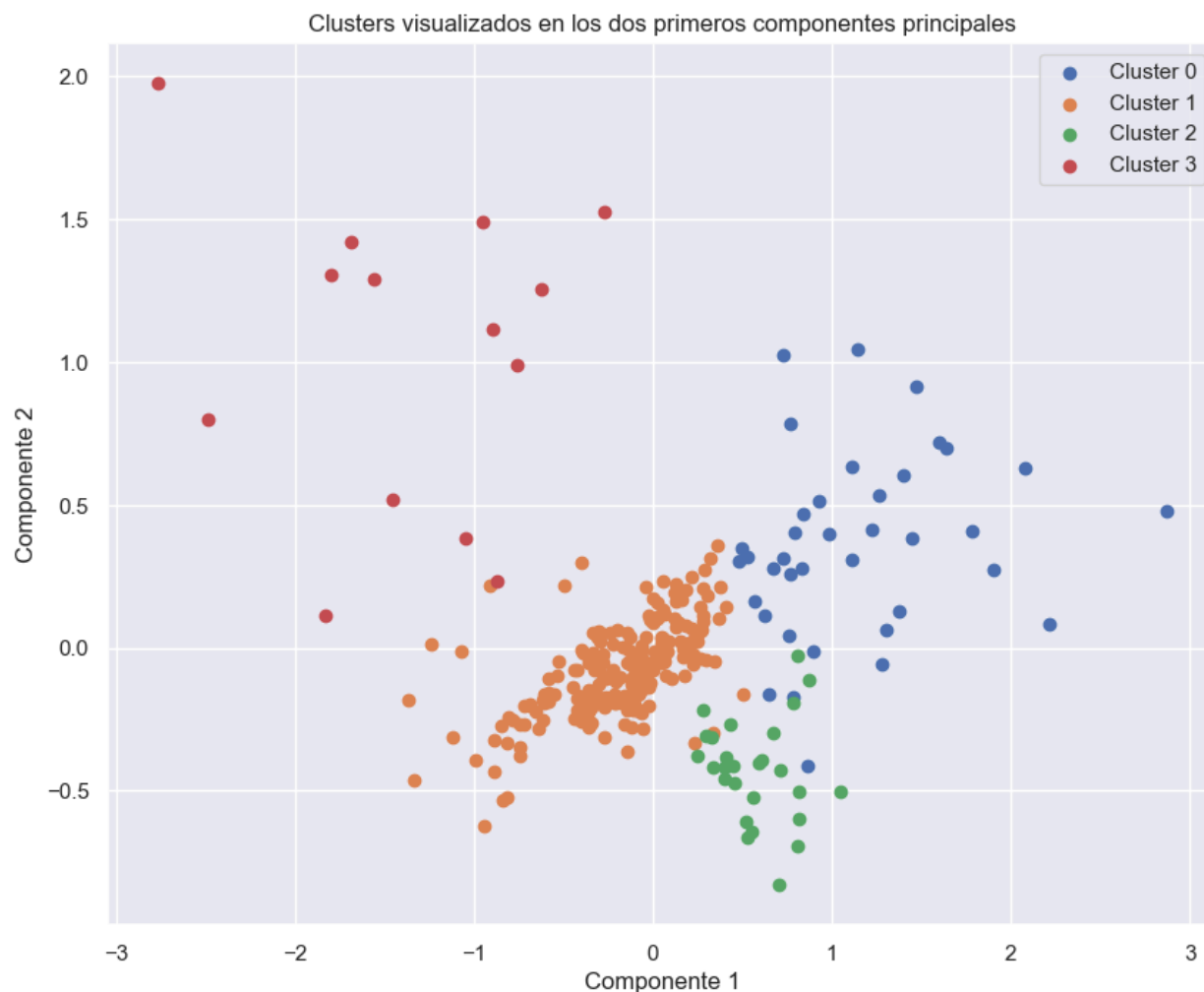
Clustering de Datos por Codigo:



Gráfica 12. Codo para clustering por codigo

Al realizar nuevamente nuestro análisis de datos con el objetivo de llevar a cabo una clusterización enfocada en el código de cada región, podemos observar en la Gráfica 12 que el método del codo nos indica que el número adecuado de clusters debería situarse entre 4 y 7. Esto garantizaría una representación significativa de la estructura subyacente de los datos.

Clusters formados:



Gráfica 13. Resultados de clusters enfocado al código de las regiones

Como lo muestra la Gráfica 13 obtenemos 4 clusters los cuales podemos interpretar como:

1. Cluster 0: Este cluster parece representar productos con un IPC relativamente alto y una estabilidad moderada. Estos productos están ubicados hacia la derecha en el espacio de los componentes principales, lo que indica un IPC mayor, y hacia la parte inferior, lo que sugiere una estabilidad moderada en los precios.
2. Cluster 1: Este cluster abarca una cantidad significativa de productos, con un IPC moderado y una estabilidad variable. Los productos en este cluster están distribuidos de manera dispersa en el espacio de los componentes principales, lo que sugiere una variabilidad en el IPC y la estabilidad de los precios.
3. Cluster 2: Este cluster representa productos con un IPC bajo y una estabilidad moderada a alta. Estos productos están ubicados hacia la izquierda en el espacio de los componentes

principales, lo que indica un IPC más bajo, y hacia la parte inferior, lo que sugiere una estabilidad moderada a alta en los precios.

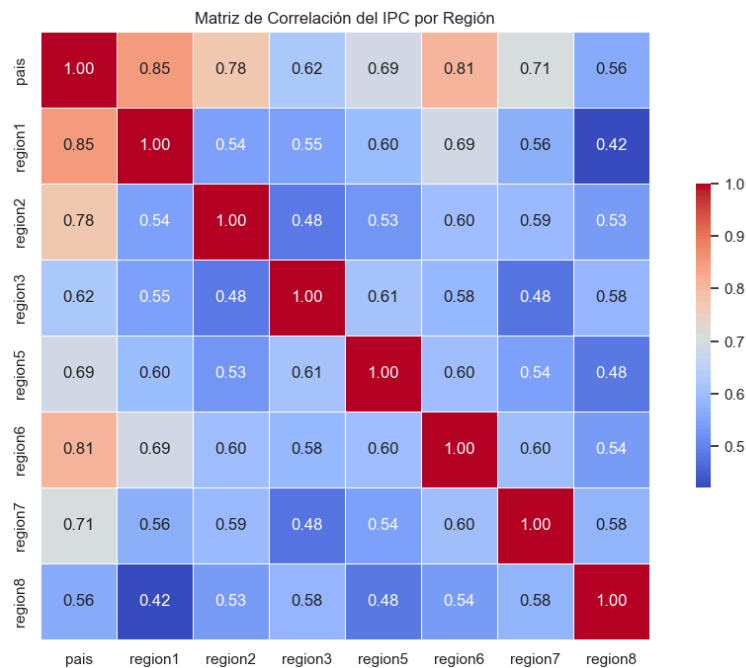
4. Cluster 3: Este cluster incluye productos con un IPC relativamente bajo y una estabilidad alta. Estos productos están ubicados hacia la izquierda en el espacio de los componentes principales, lo que indica un IPC más bajo, y hacia la parte inferior, lo que sugiere una estabilidad alta en los precios.

Hallazgos y conclusiones:

Aunado a lo anterior descrito en el apartado de análisis exploratorio, tenemos que la presencia de ceros en algunas regiones y valores extremadamente altos en otras (como el valor máximo de 68,850.73 en la región 4) sugiere anomalías en los datos, posiblemente debido a errores de medición o eventos extraordinarios no capturados adecuadamente. Estos valores atípicos afectan significativamente la media y la desviación estándar, destacando la importancia de revisar y, si es necesario, corregir los datos antes de realizar análisis más profundos. Luego de realizar dicha corrección en data, tenemos lo siguiente:

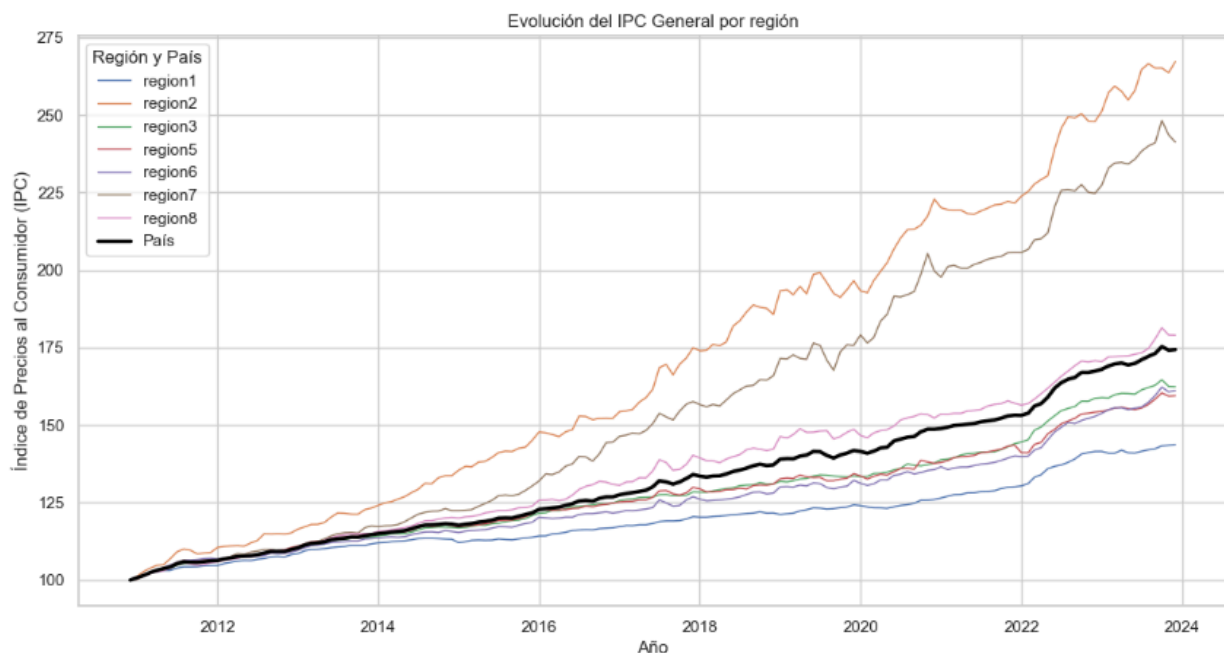
Iniciando con los hallazgos luego de anular la región 4, podemos decir que las gráficas de distribución del IPC nos muestran una notable consistencia en su forma, dicha uniformidad sugiere un patrón estable de precios en las regiones, con variaciones en las medianas y medias que reflejan diferencias regionales en los niveles de los precios. La falta de picos o caídas abruptas en dichas distribuciones nos indican la ausencia de fluctuaciones extremas o errores, que podrían llegar a distorsionar el análisis de los niveles de precios

Por otro lado, la matriz de correlación del IPC por región desvela la existencia de diversas relaciones entre las distintas zonas. Correlaciones moderadas, como la observada entre las regiones 1 y 2, sugieren que aunque compartan algunas tendencias económicas, existen factores distintivos que afectan a cada región de manera diferente. En contraste, la relación más fuerte entre las regiones 1 y 6 puede ser un indicativo de similitudes en las actividades económicas o en las políticas que influyen en los precios.



También se puede decir que las regiones con alta correlación con el IPC nacional, como la región 1 y la región 6, pueden ser vistas como representativas de la tendencia económica general del país. Estas regiones podrían ser utilizadas como termómetros de la inflación nacional, mientras que las regiones con correlaciones más bajas, como la región 1 con la región 8, sugieren influencias económicas más localizadas que podrían requerir políticas regionales específicas para su gestión.

Ahora bien, tomando en cuenta el IPC general tenemos una tendencia creciente a lo largo del tiempo, con notables diferencias en la trayectoria de crecimiento entre las categorías. Esto refleja cómo ciertos sectores pueden ser más sensibles a las variaciones económicas, políticas de precios, o incluso a eventos globales como la inflación importada o choques de oferta y demanda.



Pasando a temas de productos, tenemos que productos como el repollo y el aguacate presentan una alta volatilidad en su IPC, lo que podría atribuirse a la variabilidad en la producción agrícola, cambios estacionales, y sensibilidad a las dinámicas de mercado como la oferta y la demanda. Estos productos requieren una vigilancia continua y estrategias de gestión de riesgos para los consumidores y los productores.

Y también se tienen los productos que muestran un IPC promedio elevado, como la ropa interior y el servicio de reparación de calzado, pueden indicar un aumento en los costos de producción, la existencia de pocos sustitutos o una menor sensibilidad al precio por parte de los consumidores. Este conocimiento es valioso para entender las prioridades de gasto y podría informar las decisiones de política económica.

En resumen, nuestro estudio no solo descubre las tendencias y cambios en el IPC, sino que también nos da pistas sobre qué las está provocando y qué significan estas variaciones para quienes compran y venden estos productos, así como para quienes están al mando de dirigir la economía del país. Estos hallazgos proporcionan una base para el desarrollo de estrategias y políticas específicas que busquen mitigar los efectos de la inflación y mejorar la estabilidad económica.

Referencias

INE. (2023). Índice de precios al consumidor. Guatemala: INE. Recuperado de <https://www.ine.gob.gt/indice-de-precios-al-consumidor/>