

# Memoria técnica: Exploratory Data Analysis – Mercado Global del Café

Realizado por:

Blanca García Bartomeu  
Diana Hoyos Toscano  
Lucía Luzuriaga Ledesma

## ÍNDICE:

1.	Introducción.....	2
2.	Planteamiento del problema.....	2
3.	Objetivos del análisis.....	2
4.	Hipótesis.....	2
5.	Descripción de los datos.....	3
6.	Análisis univariante.....	3
6.1.	Producción.....	3
6.2.	Distribución global de la producción.....	4
6.3.	Porcentaje acumulado (curva de concentración).....	4
6.4.	Consumo.....	5
6.5.	Exportaciones.....	6
7.	Análisis bivalente.....	6
7.1.	Producción vs. Exportación.....	6
7.2.	Comparación entre Arábica y Robusta.....	7
8.	Análisis multivariante.....	9
8.1.	Producción, consumo y exportaciones.....	9
9.	Limitaciones del análisis.....	10
10.	Conclusiones.....	10
11.	Recomendaciones y propuestas.....	11

# Memoria técnica: Exploratory Data Analysis – Mercado Global del Café

## 1. Introducción:

El café es uno de los productos agrícolas más comercializados a nivel mundial y constituye una fuente de ingresos clave para numerosos países productores. Sin embargo, el mercado global del café presenta fuertes diferencias entre países en términos de volumen de producción y participación en el valor total generado.

El objetivo del EDA es estudiar la distribución de la producción de café entre los países productores y evaluar si dicha distribución es homogénea o si, por el contrario, existe una concentración significativa en un número reducido de países.

Para ello, se analizarán datos históricos de producción de café a nivel país, aplicando técnicas de análisis exploratorio de datos (EDA) que permitan identificar patrones, desigualdades y niveles de concentración en la producción global.

## 2. Planteamiento del problema:

El mercado global del café presenta una elevada relevancia económica para numerosos países productores. Comprender cómo se distribuye la producción y el valor asociado al café resulta clave para identificar posibles desequilibrios entre países.

## 3. Objetivos del análisis:

El objetivo de este análisis exploratorio es estudiar la distribución de la producción de café entre países y evaluar si el valor del café se reparte de forma homogénea.

## 4. Hipótesis:

- H1 :  
El valor total del café no se distribuye de forma homogénea entre los países productores.
- H2:  
Existe una relación positiva entre el volumen de producción de un país y su volumen de exportación de café.
- H3:  
Los países productores de Arábica concentran un mayor valor total que los productores de Robusta.

- H4:  
Los países con mayor producción de café no son necesariamente los mayores consumidores internos.

## 5. Descripción de los datos:

El análisis se basa en un dataset que recoge información anual sobre la producción, consumo y comercio internacional de café por país, incluyendo el tipo de café y el periodo comprendido entre 1990 y 2019.

## 6. Análisis univariante:

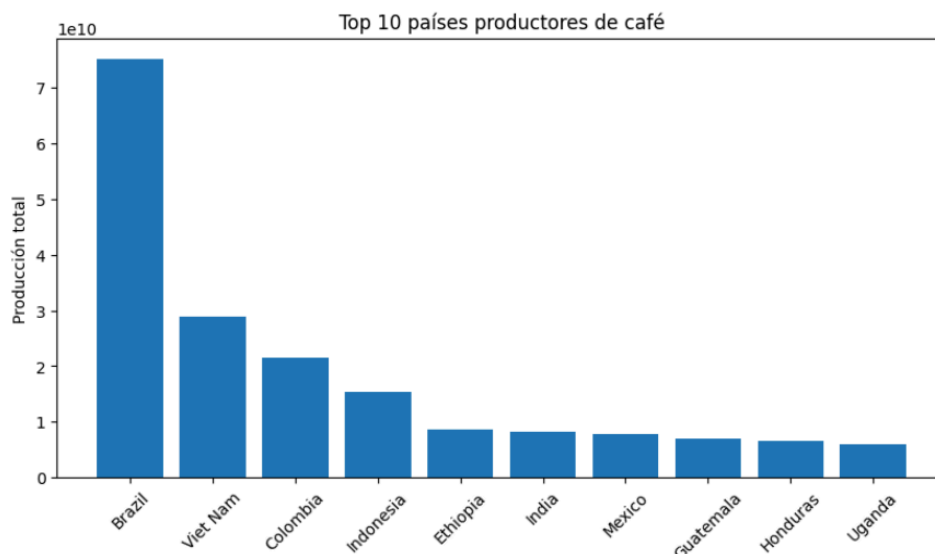
### 6.1 Producción:

La distribución de la producción de café por país muestra una alta asimetría, con un número reducido de países concentrando la mayor parte de la producción total.

	Country	Production
0	Brazil	7.508298e+10
1	Viet Nam	2.880318e+10
2	Colombia	2.159862e+10
3	Indonesia	1.540488e+10
4	Ethiopia	8.669520e+09

### Top 10 Países Productores:

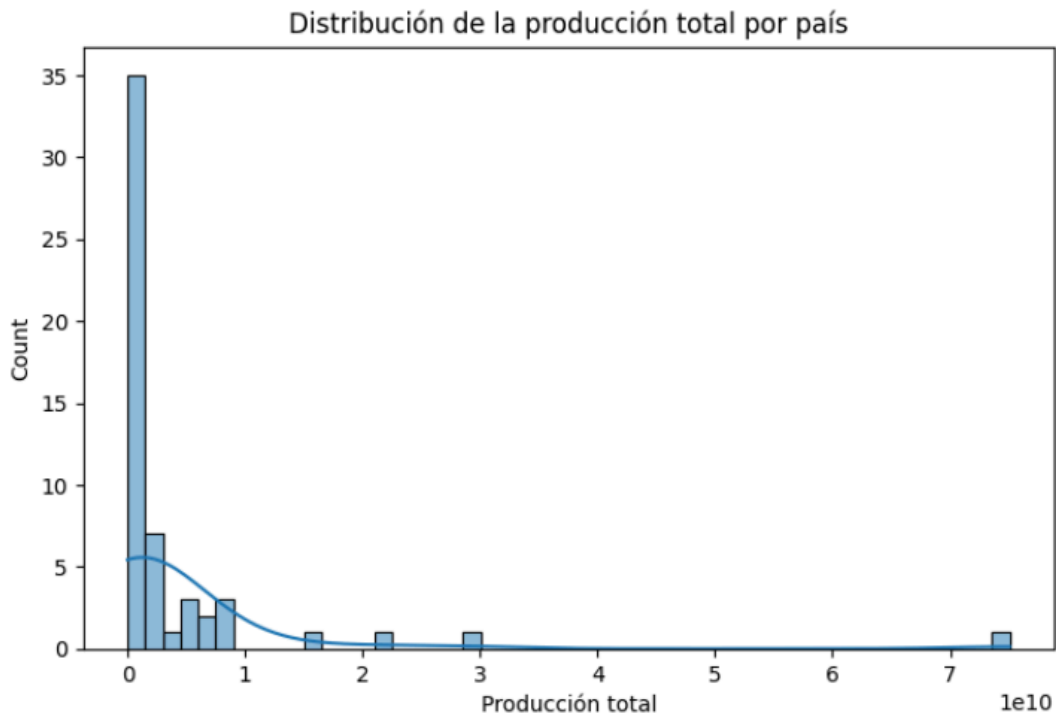
El ranking de producción muestra que un número reducido de países concentra una parte significativa del volumen total de café producido:



Esto confirma la primera hipótesis, la producción no se distribuye de forma homogénea entre los países productores.

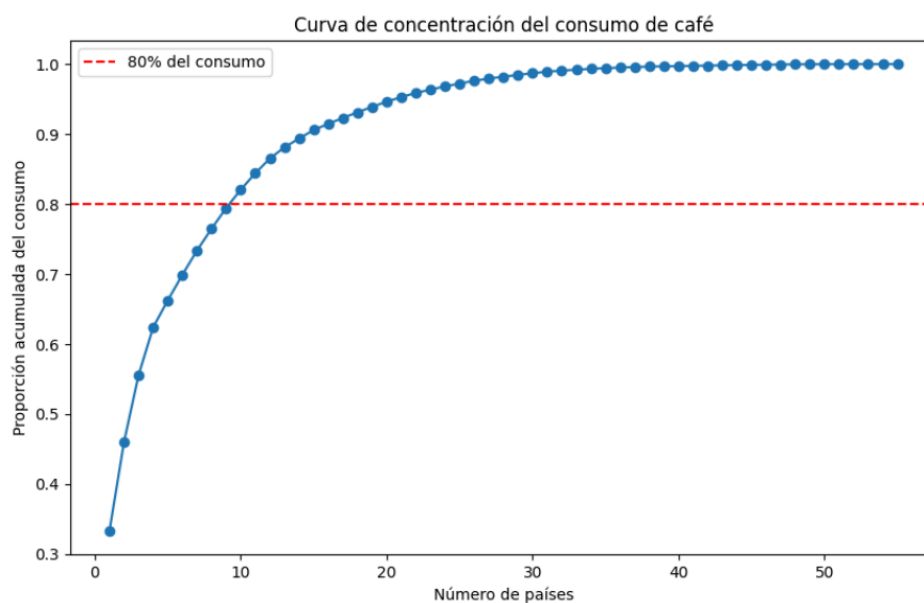
## 6.2 Distribución global de la producción:

La distribución presenta una fuerte asimetría positiva, indicando que la mayoría de países produce volúmenes reducidos frente a unos pocos grandes productores.



## 6.3 Porcentaje acumulado (curva de concentración):

Se observa que un pequeño número de países concentra un porcentaje elevado de la producción total, lo que evidencia una distribución no homogénea.



Aquí vemos que el Top 5 países concentra el 66.3% de la producción mundial y el Top 10 países el 82.1% de la producción mundial

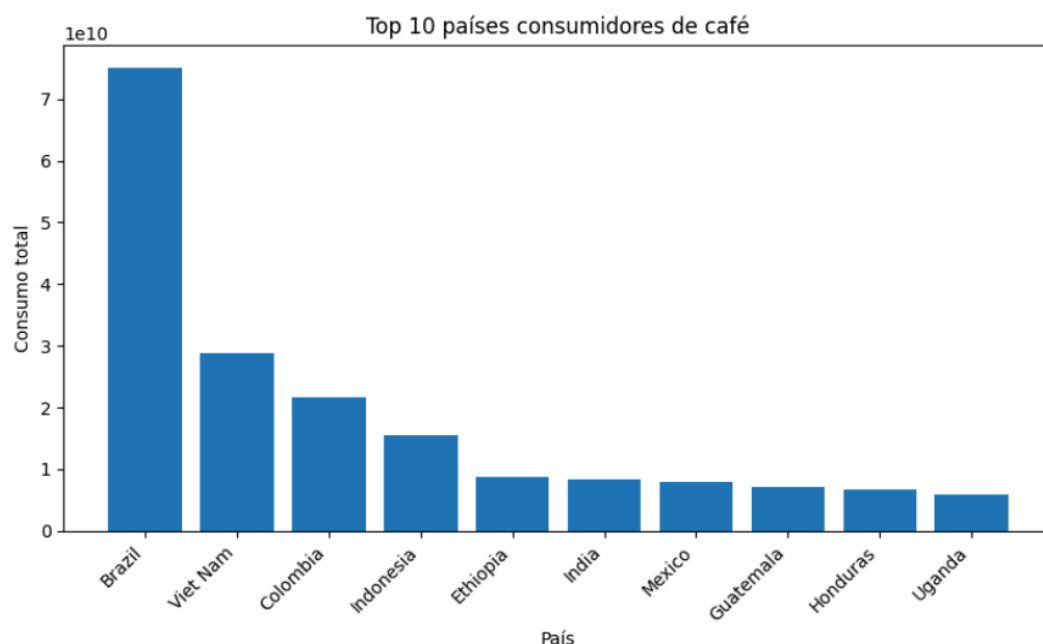
Para evaluar el grado de desigualdad en la distribución de la producción de café entre países, se ha utilizado el coeficiente de variación (CV), que relaciona la desviación estándar con la media. Este indicador permite comparar la dispersión relativa de los datos independientemente de su escala.

En términos interpretativos, valores bajos del coeficiente de variación ( $CV < 0.5$ ) indican una distribución relativamente homogénea, mientras que valores intermedios ( $0.5 \leq CV < 1$ ) reflejan una dispersión moderada. Por el contrario, valores elevados ( $CV \geq 1$ ) señalan una alta desigualdad o concentración de la producción en un número reducido de países.

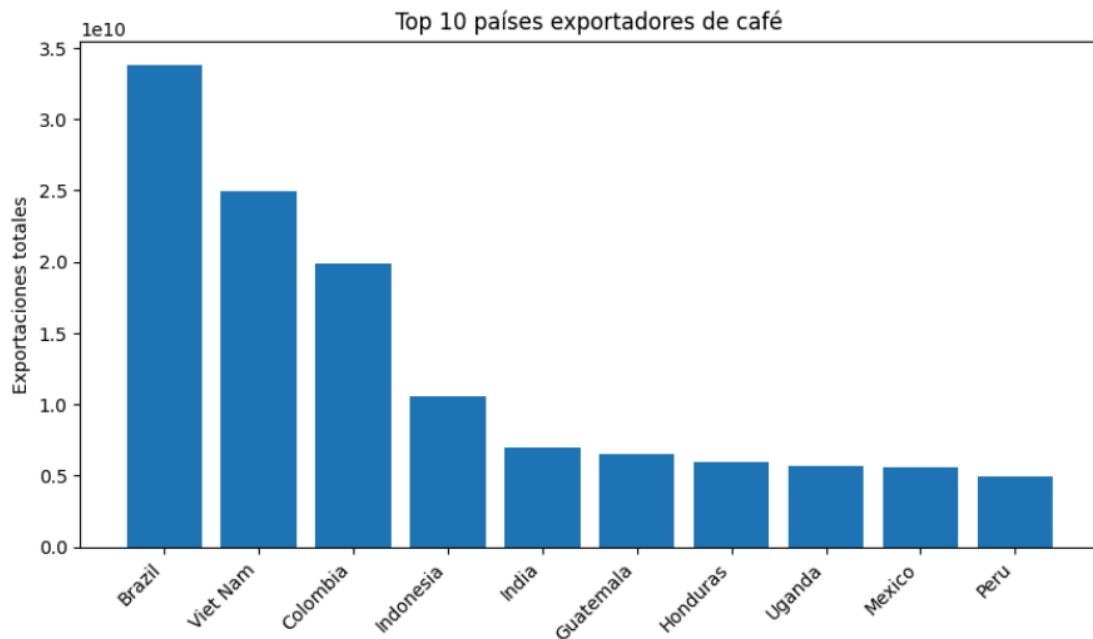
En este caso el coeficiente de variación de la producción total por país es 2.71 presentando un valor elevado, lo que indica una alta dispersión relativa. Este resultado evidencia una alta concentración de la producción de café, confirmando que el mercado está dominado por un conjunto limitado de países productores.

#### 6.4 Consumo:

El consumo también está concentrado en pocos países, hay desigualdad en la demanda.



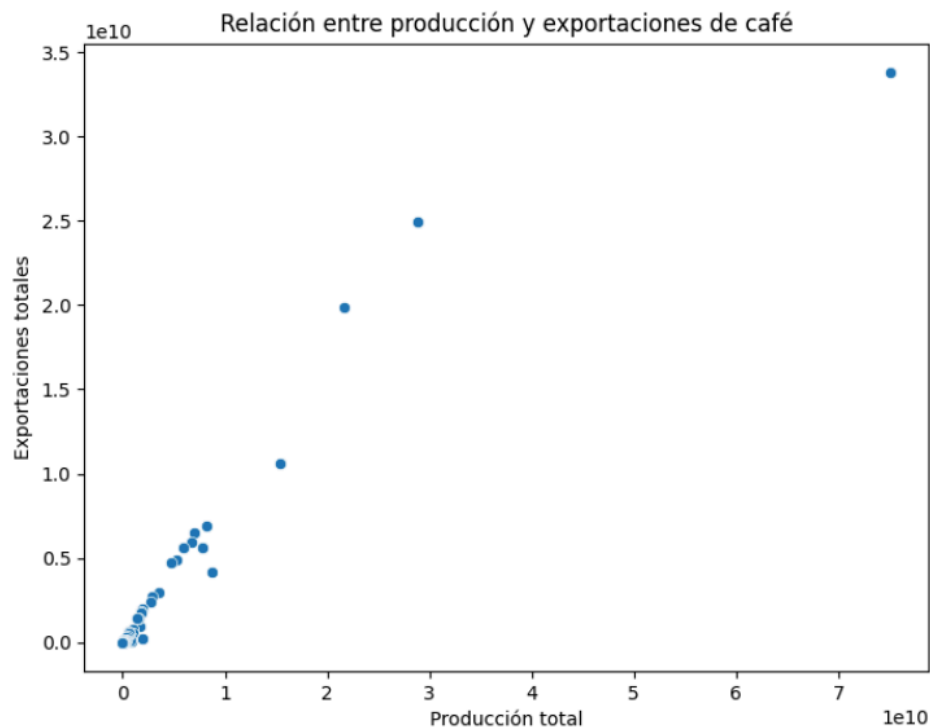
## 6. 5 Exportaciones:



El análisis del consumo y de las exportaciones muestra patrones de concentración similares a los observados en la producción. En ambos casos, un número reducido de países concentra una proporción significativa del total, reforzando la existencia de una distribución no homogénea del valor del café.

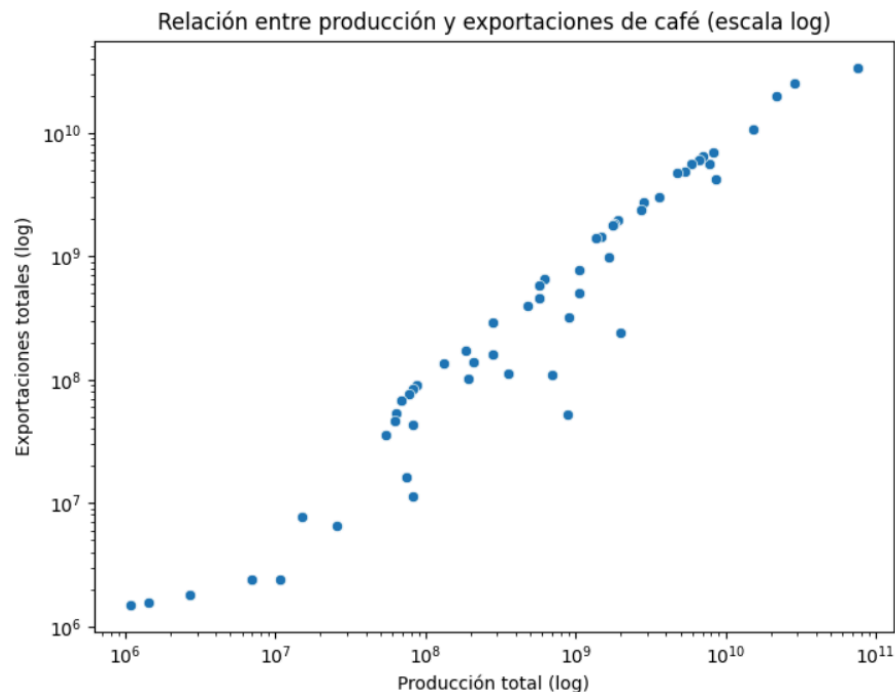
## 7. Análisis bivalente:

### 7.1 Producción vs. Exportación:



El coeficiente de correlación es del 94%, lo que indica una correlación positiva muy fuerte, es decir, los países que producen más café tienden a exportar mucho más café.

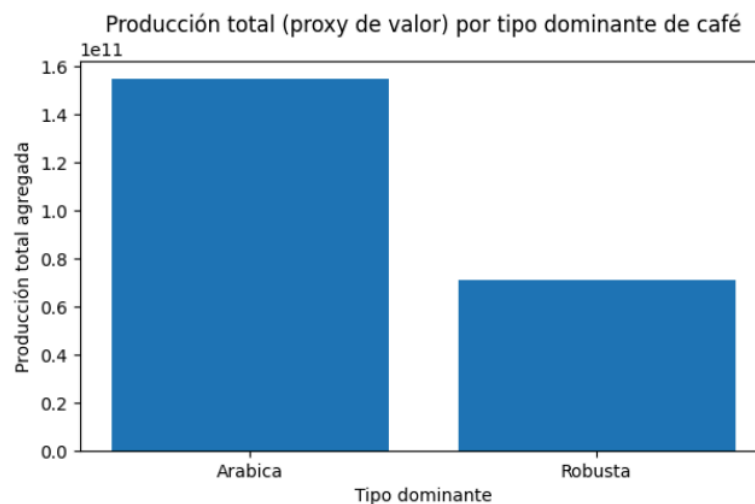
Escala logarítmica de relación entre producción y exportación:



Al utilizar una escala logarítmica se reduce la influencia de los valores extremos, permitiendo observar con mayor claridad la relación entre producción y exportaciones para la mayoría de los países.

## 7.2 Comparación entre arábica y robusta:

Con el objetivo de analizar si existen diferencias relevantes entre los principales tipos de café, se comparó el valor total asociado a la producción de café Arábica y Robusta. Dado que el conjunto de datos no incluye información de precios, se utilizó la producción total agregada como proxy del valor económico.

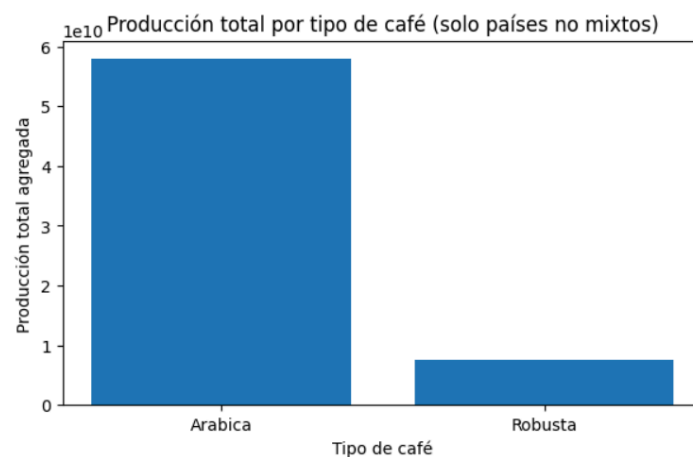




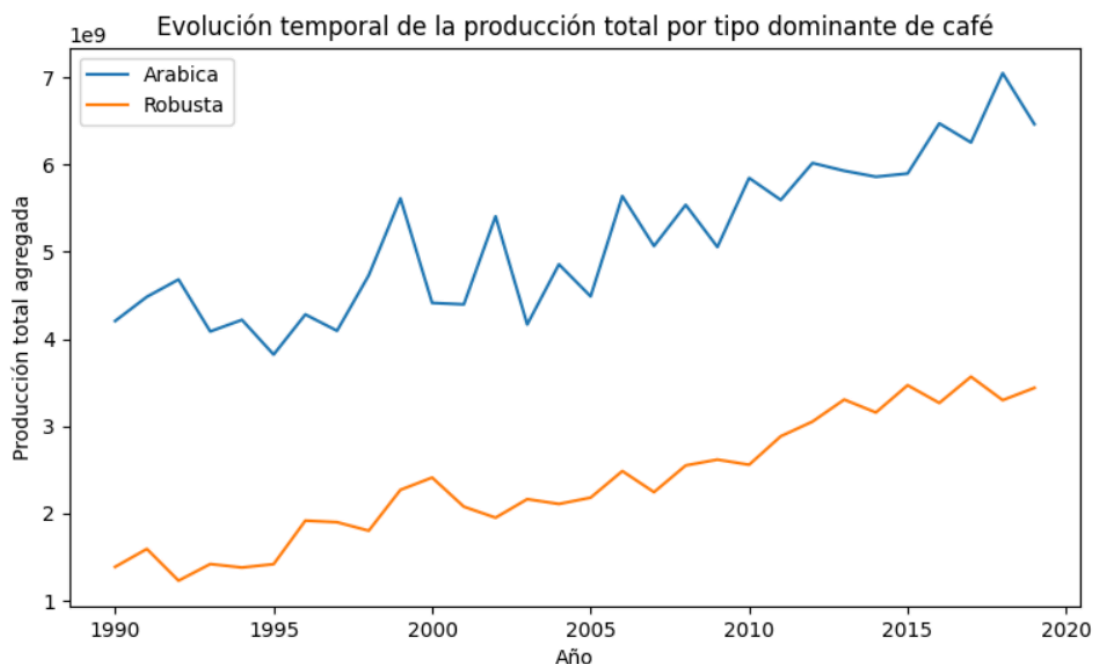
Tras agregar la producción total por tipo dominante de café, se observa que el valor total asociado al café Arábica es mayor que el correspondiente al café Robusta.

Para evitar el doble conteo de países con producción mixta, se asignó un tipo dominante de café a cada país.

Para reforzar los resultados, se repitió el análisis excluyendo los países con producción mixta de Arábica y Robusta. Los resultados obtenidos mantienen la misma conclusión que el análisis principal, lo que refuerza la validez de la hipótesis planteada:



Al analizar la evolución temporal de la producción total por tipo dominante de café, se observa una tendencia creciente tanto en el café Arábica como en el Robusta a lo largo del período analizado:



No obstante, el café Arábica mantiene de forma consistente un volumen de producción agregado superior al del café Robusta en todos los años considerados.

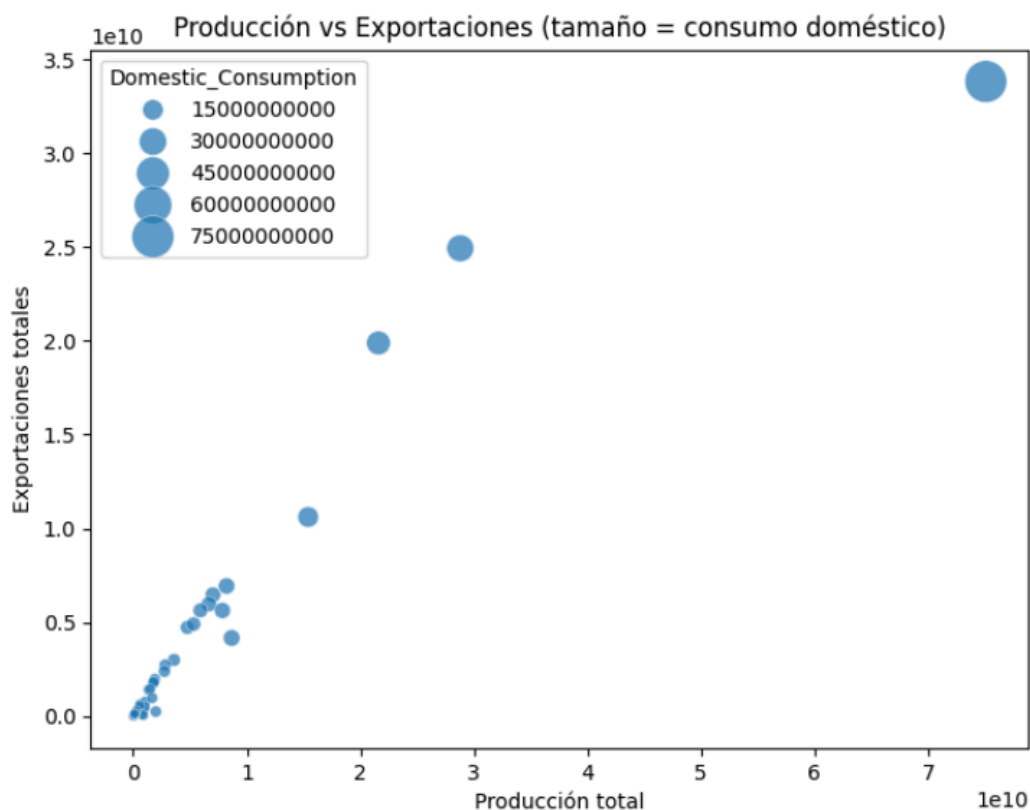
La distancia entre ambas curvas se mantiene relativamente estable en el tiempo, lo que indica que la mayor concentración de valor asociada al café Arábica no responde a un fenómeno puntual, sino que constituye una característica estructural del mercado del café. Aunque el café Robusta presenta un crecimiento progresivo, este no resulta suficiente para cerrar la brecha existente con el Arábica.

Este comportamiento refuerza los resultados obtenidos en el análisis agregado, confirmando que el predominio del café Arábica en términos de producción total se sostiene de manera consistente a lo largo del tiempo.

## 8. Análisis multivariante:

### 8.1 Producción consumo y exportaciones:

Para analizar la orientación del valor del café hacia el mercado internacional, se realizó un análisis multivariante a nivel país considerando producción, consumo doméstico y exportaciones



Los resultados muestran que los países con mayores niveles de producción tienden a concentrar también mayores volúmenes de exportación, mientras que su consumo interno es

relativamente menor en comparación con su producción total. Este patrón sugiere una clara orientación exportadora en los principales países productores.

### 9. Limitaciones del análisis:

El análisis se basa en datos agregados a nivel país, lo que impide capturar diferencias internas o regionales. No se consideran variables como precios internacionales o costes de producción.

### 10. Conclusiones:

- H1: Los resultados obtenidos confirman que la producción de café no se distribuye de forma homogénea entre los países productores, ya que un reducido grupo de países concentra la mayor parte del volumen total producido.

Esto significa que:

- Solo 5 países producen casi dos tercios del café del mundo, y solo 10 países producen más del 80%.
  - Los cinco principales países productores concentran aproximadamente el 63% de la producción total, confirmando una elevada concentración del mercado.
- H2: El análisis bivalente confirma la existencia de una relación positiva entre producción y exportaciones de café, A medida que aumenta el nivel de producción de un país, también tienden a incrementarse sus volúmenes de exportación, lo que sugiere que una parte significativa del café producido se destina al mercado internacional.
  - H3: El análisis comparativo entre los principales tipos de café muestra que los países productores de café Arábica concentran un mayor valor total que aquellos especializados en café Robusta. Utilizando la producción total agregada como proxy del valor económico, los resultados indican que el Arábica mantiene una ventaja clara en términos de volumen agregado.

Esta diferencia se mantiene de forma consistente a lo largo del tiempo, lo que sugiere que no se trata de un fenómeno puntual, sino de una característica estructural del mercado del café. Aunque el café Robusta presenta una tendencia creciente, su nivel de producción total no alcanza al del Arábica, reforzando el papel predominante de este último en el mercado global.

- H4: El análisis conjunto de la producción y el consumo interno de café muestra que los países con mayores niveles de producción no son necesariamente aquellos con un mayor consumo doméstico. En muchos casos, los principales países productores destinan una parte significativa de su producción al mercado internacional, lo que reduce el volumen de café consumido internamente.

Este patrón pone de manifiesto la existencia de roles diferenciados dentro del mercado global del café, donde algunos países actúan principalmente como proveedores para la exportación, mientras que otros, con menor capacidad productiva, concentran un elevado consumo interno. Esta divergencia subraya la importancia de considerar el consumo doméstico junto a la producción para comprender el funcionamiento del mercado del café.

#### 11. Recomendaciones y propuestas:

A partir de los resultados obtenidos en el análisis exploratorio, se proponen varias líneas de trabajo que permitirían profundizar en la comprensión del mercado global del café y mejorar la interpretación económica de los resultados.

En primer lugar, sería recomendable incorporar información de precios para complementar los datos de producción, exportaciones y consumo. Esto permitiría estimar el valor económico real del café, diferenciando no solo por volumen, sino también por el impacto de los distintos tipos de café y de las fluctuaciones del mercado internacional.

En segundo lugar, un análisis más detallado de la evolución temporal permitiría identificar cambios estructurales en el mercado, como variaciones en los patrones de consumo o en la orientación exportadora de determinados países. La incorporación de datos más recientes ayudaría a evaluar si las tendencias observadas se mantienen o están evolucionando.

Asimismo, se podría ampliar el análisis incorporando factores externos, como variables climáticas o políticas comerciales, que influyen de forma directa en la producción y el comercio del café. Este enfoque permitiría contextualizar mejor los resultados y explicar posibles desviaciones observadas en algunos países.

Por último, sería interesante segmentar el análisis por tipologías de países, diferenciando entre productores-exportadores, productores con alto consumo interno e importadores netos. Esta clasificación facilitaría la comparación entre países con perfiles similares y aportaría una visión más estratégica del funcionamiento del mercado del café.