

---

# ¿Por qué México no crece? Un análisis del estancamiento económico

Por Gabriel González Méndez

“La productividad es la clave del progreso económico a largo plazo.”

- Robert Solow

---

# Contenido

Introducción.....	3
I.    Planteamiento del problema.....	3
II.   Objetivos e Hipótesis .....	5
III.   Marco teórico.....	5
Desarrollo .....	7
I.    Análisis de datos históricos .....	7
II.   Análisis de datos correspondiente al desempeño del sector manufacturero.....	10
Sobre los coeficientes técnicos en el análisis estructural del sector manufacturero .....	14
Análisis sobre los principales componentes .....	15
III.   Ejercicio de Clustering usando el algoritmo K-Means.....	17
El algoritmo K-Means.....	17
Análisis de resultados .....	22
I.    Integración con variables macroeconómicas.....	23
Conclusiones .....	25
Bibliografía .....	26
Anexos .....	26

---

# Introducción

Este trabajo representa sintéticamente el contenido del curso correspondiente a Modelos Macroeconómicos II, contiene una estructura que le es propio al análisis del por qué México no crece como nación, y para esto, se tiene que hacer una diferenciación conceptual de lo que compete el crecimiento económico, al tiempo que se aborda de la misma manera lo que es el desarrollo económico.

De esta forma, el desarrollo del proyecto se elaboró de acuerdo con la temática del curso, específicamente, de la presentación de crecimiento económico en México y Formación Bruta de Capital elaborada de esta parte. En otras palabras, si bien este documento comparte la base analítica, tiene como objetivo ampliar el análisis en un sentido cuantitativo.

## I. Planteamiento del problema

México ha tenido una historia económica de estancamiento económico, siguiendo a Canto (2023), este problema es más que complejo, ya que obedece no solamente a leyes económicas, también corresponde a un periodo histórico con consecuencias sociopolíticas, y desde luego, culturales. En este contexto, el autor problematiza los problemas económicos de México como una herencia cultural y sociopolítica colonial, señala que estructuralmente solo hubo una distribución distinta del poder político y económico, creando una nueva clase predominante.

En términos culturales, la dispersión social podría tener un papel relevante relacionado al crecimiento y desarrollo económico. Siguiendo a Samuelson y Nordhaus (2009) “El crecimiento económico es un aumento sostenido en la producción total de bienes y servicios en lo que respecta a una nación, medido generalmente en PIB real”.

Y refiriéndome ahora al concepto de desarrollo económico, Todaro y Smith (2020) tienen la siguiente definición: “El desarrollo económico debe ser percibido como un proceso multidimensional que implica la reorganización y reorientación de sistemas económicos y

---

sociales completos. Además de mejoras en ingresos y producción, implica cambios institucionales, sociales y culturales, y la expansión de las libertades humanas”

Pero, ¿Cuáles son las diferencias sustanciales entre estos dos conceptos económicos?

El crecimiento económico tiene un corte cuantitativo -y se enfoca a la producción dentro de una sociedad-, y en este documento me enfocaré en ese contexto específicamente, mientras el desarrollo económico es de corte cualitativo y enfocado a la calidad de vida de la población.

El alcance del desarrollo económico tiene un ámbito tanto económico, pero también político y hasta ambiental, y el crecimiento tiene un límite económico, misma definición orienta la elaboración y límite de este trabajo.

Además de la herencia sociopolítica y económica, hay un elemento primordial inscrito en el estancamiento del crecimiento económico de México que Canto (2023) describe como la apropiación de modelos económicos destinados al análisis de la realidad económica mexicana, pero que dichos modelos no fueron diseñados para este tipo de sociedades, ya que su planteamiento y desarrollo teórico se gestaron en una sociedad desarrollada.

Aunado a esto, el sector manufacturero no está siendo trabajado como se debería, para lograr una expansión económica, Aroche (2019). Y para efectos de análisis, en este trabajo se presenta una vía alterna al análisis econométrico convencional, usando un algoritmo de aprendizaje no supervisado y teoría económica, que, en la actualidad, el volumen creciente de los datos demanda la integración analítica convencional, con tendencia del análisis con técnicas de ciencia de datos. Esta perspectiva es complementaria, potente y muy útil en la labor del economista, así como propiamente lo integra Varian (2014) en su análisis económico.

---

## II. Objetivos e Hipótesis

**Hipótesis:** El sector manufacturero puede ser potenciador del crecimiento económico.

**Objetivos:**

- Presentar datos históricos para iniciar una narrativa del contexto del crecimiento económico en México.
- Realizar un análisis de datos en aras de mostrar el desempeño del sector manufacturero mexicano, para que sirva de base demostrativa de la influencia de dicho sector en la economía nacional.
- Emplear un ejercicio de clustering y el algoritmo K-Means, propio de los métodos de análisis de la ciencia de datos, para segmentar la industria en función de las características estructurales, hacer más robusto el análisis de datos usando técnicas interdisciplinarias y así vincularlo satisfactoriamente con el contexto macroeconómico en México.

## III. Marco teórico

Como marco teórico, se usa primordialmente el texto de Chiang y Wainwright (2006: p.112-121) en lo que se refiere al manejo de la teoría sustancial para el análisis de datos correspondiente al análisis de la estructura y desempeño del sector manufacturero en México. Es importante esta justificación porque a partir de aquí se llevan análisis como el que hace Aroche (2019), apuntando a la falta de crecimiento en el sector manufacturero como principal motor económico. Por tal motivo, se toma como base analítica tanto el artículo de Aroche (2019), como el capítulo dedicado al álgebra lineal de Chiang y Wainwright (2006); para llevar a cabo el análisis del desempeño del sector manufacturero mexicano.

---

De este modo, y siguiendo a estos autores, el planteamiento del modelo IP es el siguiente:

$$X = AX + D$$

$$X = (I - A)^{-1}D$$

Dónde:

- X: Es el vector de producción total.
- A: Es la matriz de coeficientes técnicos (insumo-producto).
- D: Es el vector de demanda final
- $(I - A)^{-1}$ : Es la inversa de la matriz de Leontief.

Como conceptos básicos del modelo, y de acuerdo con los autores previamente citados, tenemos que el modelo IP sigue la siguiente lógica: El modelo modela la interdependencia entre sectores en la economía, además de que responde a la pregunta central de la economía ¿cuánto producir?, en términos sectoriales y proyectando la pregunta en cuanto a la satisfacción de la demanda final.

Este modelo también plantea la pregunta: En términos de la demanda intermedia, ¿Qué otros sectores le compran? Y, por último, se muestra el efecto del multiplicador dentro de la economía.

Los supuestos del modelo:

- Los coeficientes técnicos son fijos, es decir, que la tecnología es constante.
- No hay sustitución de insumos.
- Hay rendimientos constantes a escala.
- Como el modelo es de cantidades, y no de precios, estos son fijos.

Una vez planteado el modelo de insumo producto, a manera de sustento en términos de relevancia del análisis del desempeño del sector manufacturero, Aroche (2019) fundamenta en su trabajo la falta de estimulación en este sector, y la pone como principal causa del estancamiento de la economía mexicana. Para estos efectos demostrativos, el autor dentro de su marco teórico utiliza la Ley de Kaldor-Verdoorn en la cual se subraya que:

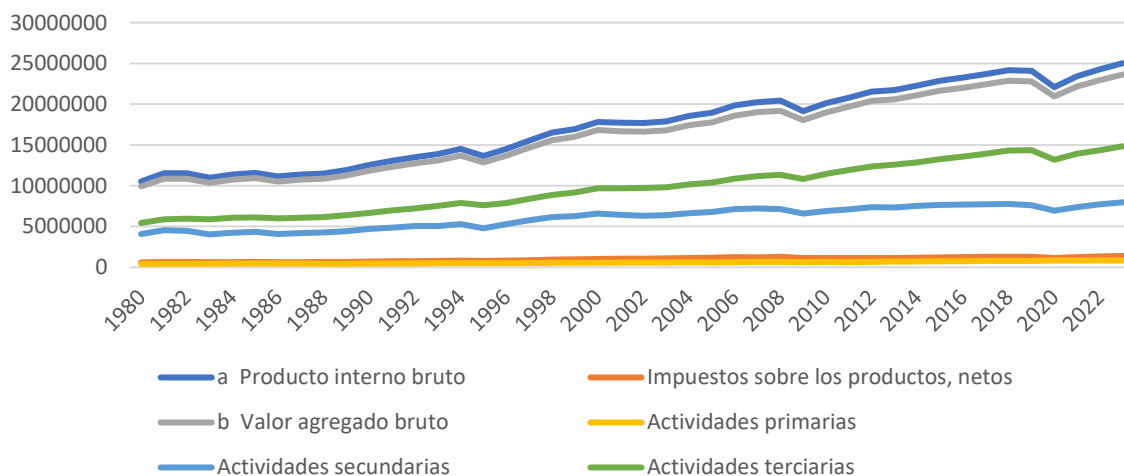
- Se tienen rendimientos constantes a escala.
- Ante un mayor producto manufacturero, hay una mayor productividad dentro de la economía.
- El sector manufacturero, siguiendo el marco teórico del modelo IP, tiene la capacidad de arrastrar el crecimiento de los demás sectores a través de la demanda de insumos.

Aroche (2019) también usa el modelo IP para sustentar su trabajo. En su trabajo se muestra que en las últimas décadas México ha tenido un crecimiento promedio del 2.5% y encuentra que en la economía mexicana hay una tercerización muy fuerte, mismo que se muestra en el análisis de datos de este trabajo, específicamente en el gráfico 2-La composición del PIB.

## Desarrollo

### I. Análisis de datos históricos

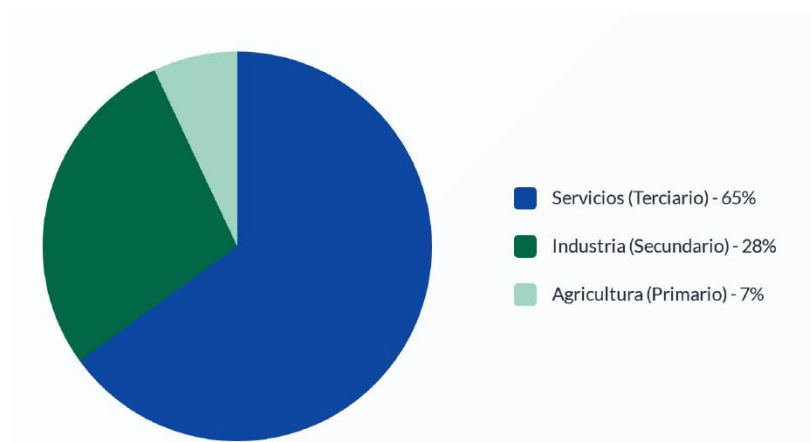
Gráfico 1. Evolución del PIB (1980-2022)



Fuente: Elaboración propia, con datos del INEGI

En el gráfico 1) se muestra que efectivamente, se tiene una tendencia creciente, pero hay periodos de estancamiento y de caída en el PIB, especialmente en periodos tales como en 1995, debido al efecto tequila; en el 2008 y 2009, periodos de crisis financiera en estados unidos en bienes raíces; en 2020, a razón de la pandemia por COVID-19.

Gráfico 2. Composición del PIB



Fuente: Elaboración propia, con datos del INEGI

Ahora, en el gráfico 2) se muestra cómo está compuesto el PIB, y los resultados obtenidos en dicho análisis concuerdan con los resultados que obtuvo Aroche (2019), ya que se comprueba una tercerización generalizada.

Gráfico 3- Actividad económica (IGAE)



Fuente: Elaboración propia, con datos del INEGI



La actividad económica que muestra el gráfico 3) es consistente con lo obtenido en términos del comportamiento del PIB, ya que en el periodo de 1995 hay una caída en la actividad económica; en 2008 se observa también una caída agresiva; y en 2020 se sobrepasa la caída, comparándola con los otros dos periodos.

Gráfico 4. Tasa de desempleo, series desestacionalizadas



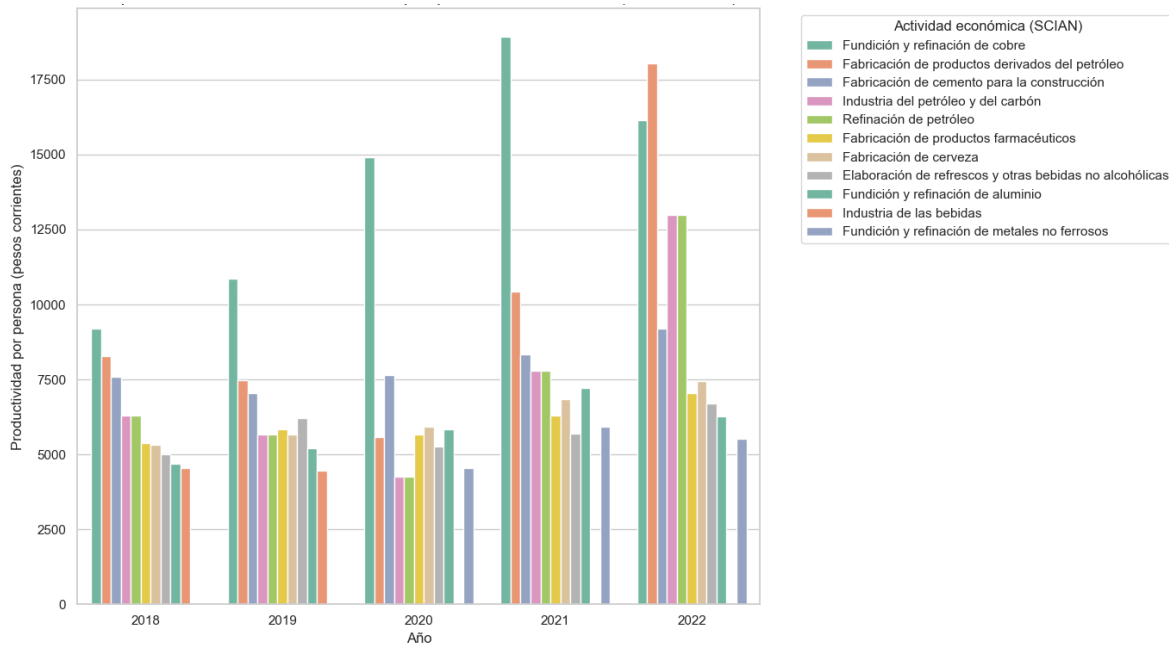
Fuente: Elaboración propia, con datos del INEGI

El análisis en cuanto a tasa de desempleo corresponde a un periodo mucho más corto, pero se puede observar el esbozo consistente con los gráficos anteriores, ya que en 2008 hay una subida en esta tasa, de la misma manera se observa claramente que la tendencia iba hacia la baja y al llegar la pandemia COVID-19 hubo un repunta, que si bien, no fue tan agresivo como en el periodo de 2008, si pasó los 5 puntos porcentuales; y que posteriormente, bajo el intervalo de recuperación, ha ido en descenso.

## II. Análisis de datos correspondiente al desempeño del sector manufacturero

Esta parte del documento se dedica específicamente a la resolución de uno de los objetivos planteados con anterioridad, mismo que corresponde al análisis concreto del desempeño del sector manufacturero.

Gráfico 5. Top 10 de las actividades manufactureras por productividad laboral



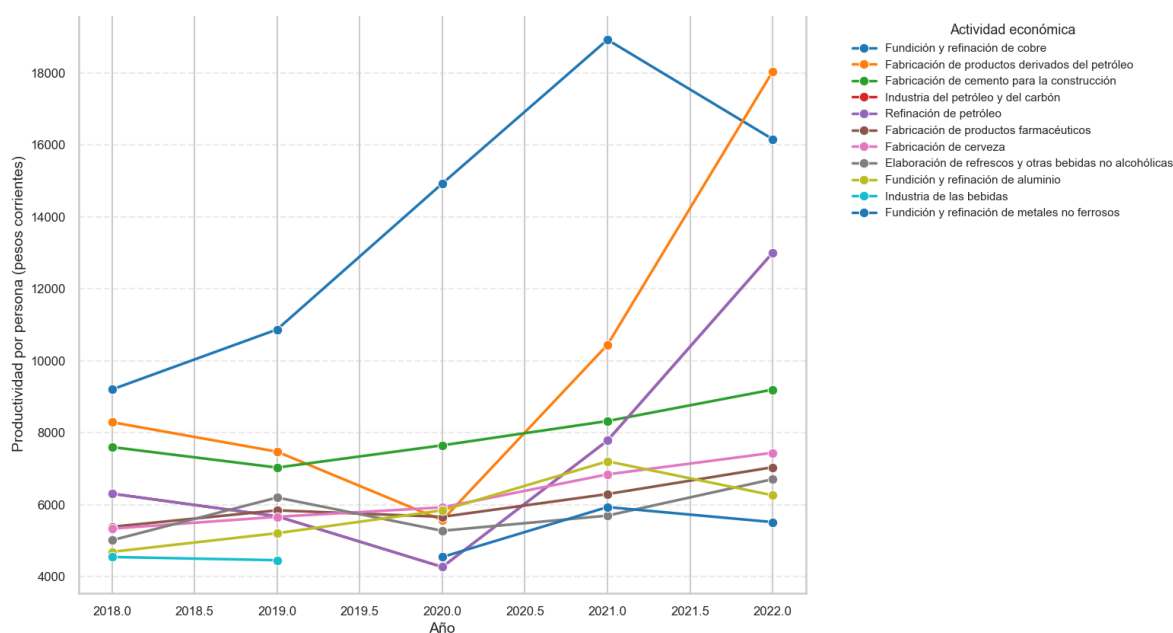
Fuente: Elaboración propia, con datos del INEGI

Para presentar propiamente los hallazgos que le son propios a este análisis, es importante introducir la siguiente identidad, ya que el gráfico 5) expresa un agrupamiento de las actividades más productivas en la economía, en un intervalo de cinco años, esto es, una expresión en términos de la productividad laboral dentro de los sectores:

$$Prod = \frac{Valor\ agregado\ Bruto}{Personal\ Ocupado\ Tot.}$$

Y siguiendo esta idea, se puede observar el comportamiento de la productividad en estos sectores. En este sentido, el comportamiento es muy regular, ya que en términos de productividad la refinación y fundición de metales (cobre) lidera el intervalo. Lo que en este gráfico llama la atención es que, en términos de productividad, todos los sectores van creciendo gradualmente, respetando generalmente la posición que les corresponde, esto se puede deber primordialmente a que el personal ocupado se reduce, o bien a que hay un incremento en la tecnología que está incrementando el valor agregado bruto.

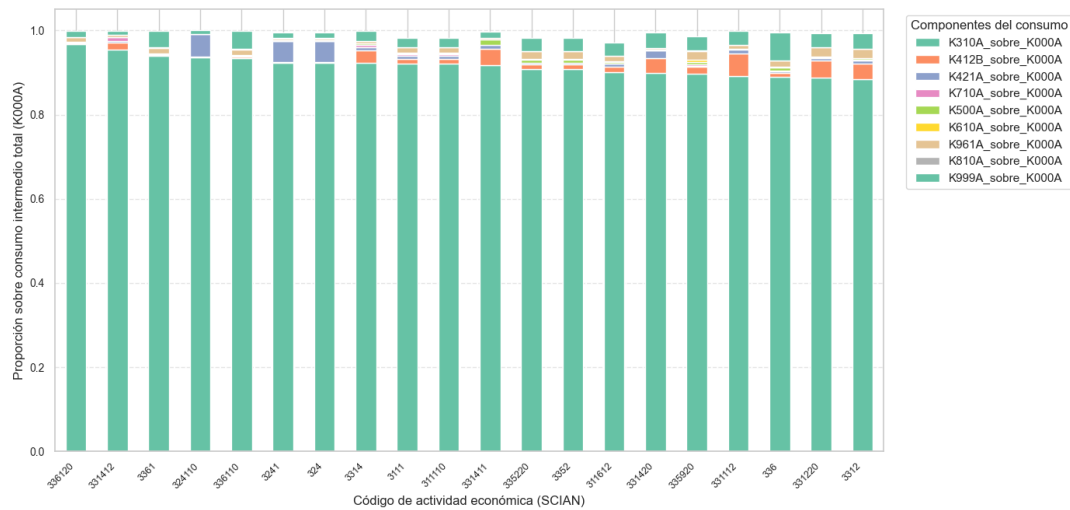
Gráfico 6. Evolución de la productividad laboral por actividad manufacturera



Fuente: Elaboración propia, con datos del INEGI

El gráfico 6) muestra que una vez más, la fundición y refinación de cobre tiene una mayor productividad, esto es consistente con el gráfico anterior, ya que nos está indicando que, si hay una actualización tecnológica, ya que este sector es intensivo en el uso de capital. Pero en general, estos sectores están vinculados a una mayor utilización del capital sobre el factor trabajo, es por eso que se obtuvo una mayor productividad en este intervalo temporal.

Gráfico 7. Estructura del consumo promedio por actividad en 2022



Fuente: Elaboración propia, con datos del INEGI

Para poder abordar el contenido del gráfico 7) es bueno precisar al consumo intermedio, y este se refiere a los bienes y los servicios que requiere cierta actividad para producir su producto o servicio, excluyendo al valor agregado, en otras palabras, es lo que se consume en el proceso de producción. Y en este sentido, algunas industrias dependen del transporte, energía y servicios financieros; vinculándolo con el marco teórico planteado, los sectores que tienen alta dependencia con respecto a los insumos externos tienen multiplicadores más altos, ya que su propia expansión genera demanda en los otros sectores.

Tabla 1. Multiplicadores de consumo intermedio<sup>1</sup>

CODIGO_ACTIVIDAD	multiplicador_consumo
1437 324110	0.999961
1468 326211	0.999690
1491 331412	0.999683
1541 336120	0.999578
1527 334310	0.999511
1343 3343	0.999511
1492 331419	0.999293
1425 321112	0.999123
1299 3211	0.998740
1355 3365	0.998739

Fuente: Elaboración propia con datos del INEGI

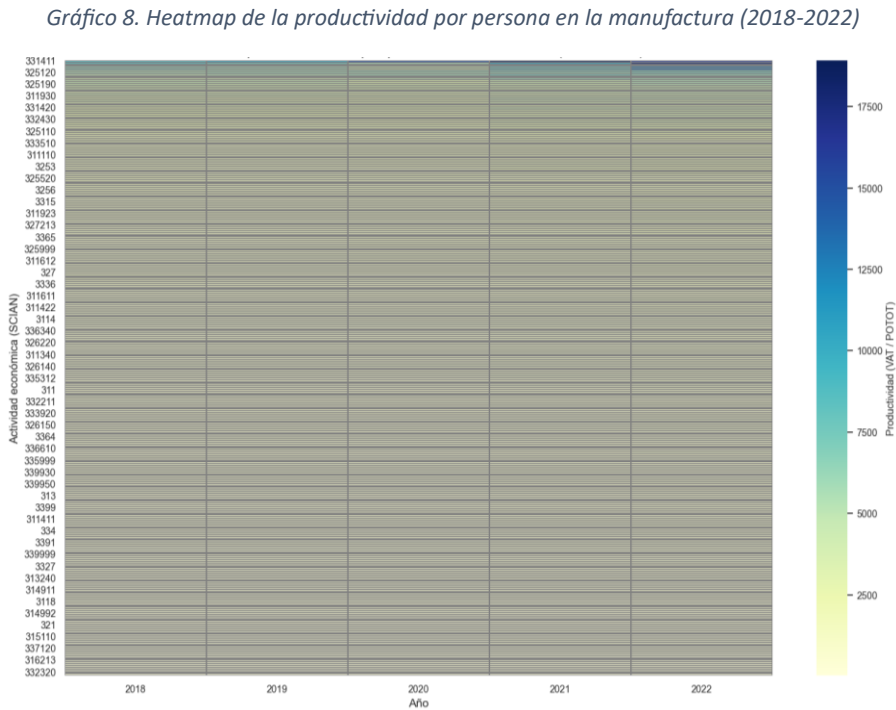
En la tabla 1) se muestran los multiplicadores del consumo intermedio en 2022. El valor de cada multiplicador indica cuánto se expande el consumo intermedio total en el sistema económico cuando una unidad de demanda final se dirige a esa actividad.

Si el multiplicador es cercano a 1, significa que casi toda la producción de esa actividad depende de insumos intermedios. Esta tabla complementa el análisis del gráfico 7) mientras el gráfico muestra la composición de los insumos (energía, servicios, transporte), esta tabla muestra la intensidad total del consumo intermedio.

Actividades como 324110 (fabricación de pinturas y recubrimientos), 326211 (fabricación de empaques plásticos), y 331412 (fundición de metales) tienen multiplicadores de consumo cercanos a 0.999, lo que indica una altísima dependencia de insumos intermedios.

<sup>1</sup> Las actividades con nombre explícito se encuentran en el diccionario proporcionado por el INEGI.

El vínculo de la tabla 1), así como su propia relevancia, es que se muestran los multiplicadores más altos, esto en términos de política industrial es muy significativo, ya que se pueden y se deben buscar activar al máximo las cadenas productivas para impulsar la economía mexicana.



Fuente: Elaboración propia, con datos del INEGI

En el análisis que le corresponde al gráfico 8) se tiene que hay algunas actividades que mantienen alta productividad todos los años, que corresponden a las actividades situadas en la parte superior, los que están en color azul. De la misma manera, se observa una estabilidad estructural, en términos económicos ya que el gráfico esboza que algunos sectores mejoran su eficiencia a través de la automatización y el desarrollo tecnológico.

Y vinculando estos resultados obtenidos con la política sectorial, se puede entender a los sectores en color azul candidatos a inversión estratégica.

## Sobre los coeficientes técnicos en el análisis estructural del sector manufacturero

Para analizar esa cuestión, se elabora mediante código en Python 3.11, el cálculo estadístico de los coeficientes:

Tabla 2

Estadísticas descriptivas de los coeficientes:				
	Coef_Insumo	Coef_ValorAgregado	Productividad_Laboral	Suma_Coef
count	314.000000	314.000000	314.000000	314.0
mean	0.575925	0.424075	3839.715315	1.0
std	0.124938	0.124938	5867.590182	0.0
min	0.227774	0.128102	110.105831	1.0
25%	0.517128	0.325648	1009.512757	1.0
50%	0.583218	0.416782	2086.651224	1.0
75%	0.674352	0.482872	3939.142567	1.0
max	0.871898	0.772226	48853.609307	1.0

Fuente: Elaboración propia, con datos del INEGI

Los resultados obtenidos en la tabla 2) se refieren a que, en promedio, 57.6% del valor bruto de producción en las actividades manufactureras se destina a insumos intermedios, hecho que confirma lo que vimos en el gráfico de estructura de consumo intermedio: la mayoría de las actividades son intensivas en insumos, lo que implica multiplicadores altos en el modelo insumo-producto, teniendo en cuenta que las actividades con coeficientes cercanos a 0.87 son las que están muy encadenadas como los sectores de fundición de metales, refinación y empaques.

Por otro lado, el valor agregado representa 42.4% del valor bruto, en promedio, lo que se vincula con los gráficos de productividad laboral: sectores con mayor valor agregado por trabajador tienden a tener coeficientes más altos, como farmacéutica o petróleo, teniendo en cuenta que el mínimo ( $\approx 0.128$ ) indica sectores donde la mayor parte del valor se genera fuera del proceso productivo directo, como en actividades de ensamblaje o maquila.

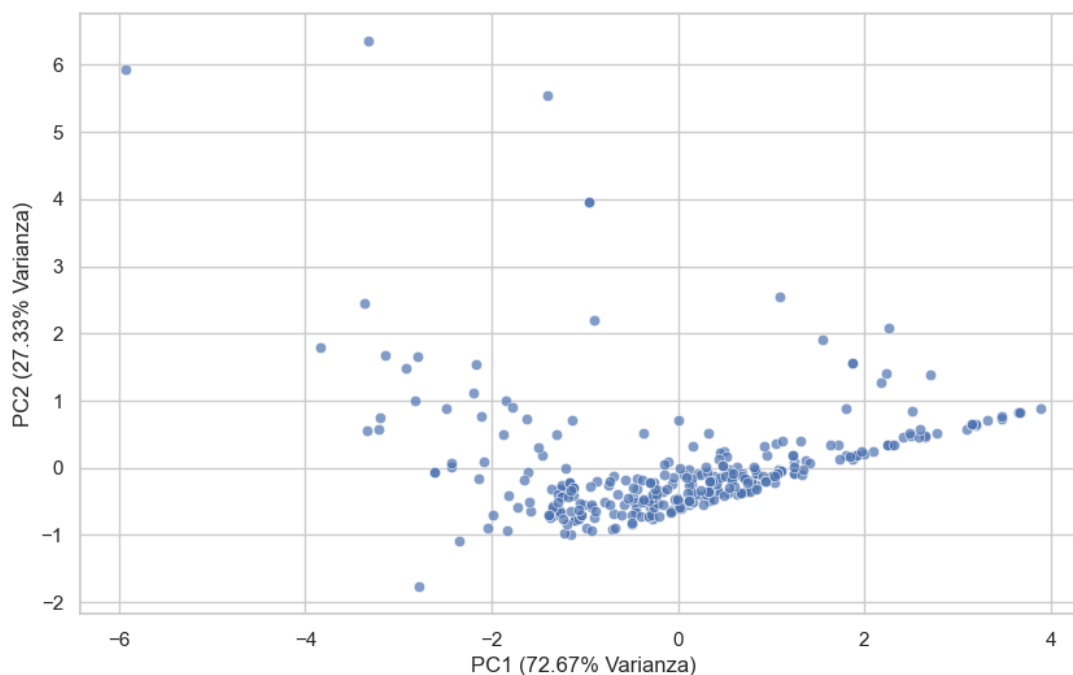
Ahora, en términos de la productividad laboral, existe una altísima dispersión: desde 110 hasta casi 49,000 pesos por persona, esto refleja lo que se mostró en los gráficos de productividad: sectores como refinación de cobre o petróleo tienen valores muy altos, mientras que otros como papel o bebidas están en el rango medio. En cuanto a la desviación estándar ( $\approx 5,867$ ) indica que la heterogeneidad sectorial es enorme, lo que justifica el enfoque multisectorial del modelo insumo-producto.

Estos resultados indican que los sectores con alto coeficiente de insumo y baja productividad pueden ser útiles para generar empleo y activar cadenas, pero no necesariamente van a elevar el PIB. Mientras que los sectores con alto coeficiente de valor agregado y con una alta productividad son ideales para elevar eficiencia y competitividad, aunque su impacto en empleo puede ser limitado.

La relevancia de este análisis es aplicable depende a los objetivos de política económica, ya que permite segmentar de manera estratégica las actividades económicas, lo que se puede lograr con esta matriz es tomar decisiones dependiendo si se quiere estimular el crecimiento vía encadenamientos, o bien, mejorar el ingreso por trabajador (se tendría que mirar a la productividad; o también encontrar una mejora estructural (se tendría que considerar al valor agregado)).

### Análisis sobre los principales componentes

*Gráfico 9. Análisis de componentes principales de actividades industriales*



Fuentes: Elaboración propia, con datos del INEGI

---

Correspondiente a este análisis, se obtuvo el vector varianza explicada por componente:  $[0.72673054, 0.27326946]$ . Esencialmente, esta técnica se usó para reducir la dimensionalidad de los datos a la vez que se mantiene la variabilidad de los mismos. La proyección mostrada corresponde a los sectores industriales, en el plano se muestra (en cuanto a los ejes explícitos):

- $CP_1$ : que corresponde al componente que explica la mayor parte de la variación, con un 72.6730%.
- $CP_2$ : que viene siendo el segundo componente más importante, con un 27.3269%.

En este sentido, cada punto mostrado representa una actividad inscrita dentro del sector manufacturero, cuya posición la determina las características estructurales; para esto es importante denotar que las variables tomadas para este análisis fueron:

- Coeficiente de insumo, mismo que representa la intensidad en términos de consumo intermedio.
- El coeficiente de valor agregado.
- Productividad laboral.

Lo que devela este gráfico 9) la cercanía entre actividades económicas, es decir, las actividades que comparten estructuras similares, términos de insumos, valor agregado y productividad laboral. Esto resulta muy útil para el diseño de políticas sectoriales, ya que el agrupamiento habla de una posible potenciación y encadenamiento en la actividad económica, y en consecuencia, una mejora en el crecimiento del PIB mexicano. Además, permite diferenciar las actividades que no están cerca, y así descartarlas de la política sectorial. De manera directa, este análisis confirma que las variables que se analizaron anteriormente tienen patrones consistentes y que tienen buena capacidad explicativa de la estructura manufacturera.



---

### III. Ejercicio de Clustering usando el algoritmo K-Means

#### El algoritmo K-Means

En esta parte del trabajo introduzco un análisis interdisciplinario entre la economía aplicada y la ciencia de datos, esto porque, es una manera de mostrar la intersección de los métodos de análisis en ciencia de datos en el análisis económico. De esta manera, ampliando el análisis, no solo se demuestra la hipótesis planteada, sino que se evidencia la importancia de la aplicación de estos métodos en el trabajo como analista económico/de datos.

Con respecto al planteamiento teórico de este algoritmo, y siguiendo a Boschetti y Massaron, (2018, p. 269), este algoritmo pertenece al aprendizaje no supervisado (no tiene una variable objetivo o datos etiquetados para llevar a cabo el trabajo analítico. De esta manera, es pertinente enunciar que el análisis económico tradicional (con métodos econométricos) es complementario al análisis que se hace en ciencia de datos, ya que la regresión lineal -una herramienta fundamental en el análisis económico-, pertenece al aprendizaje supervisado, y también se puede entender como una red neuronal de una sola neurona.

El planteamiento teórico matemático de este algoritmo responde a un problema de optimización, ya que se busca minimizar la varianza intra-cluster, y que, en términos de este trabajo, minimizar la varianza intra-actividad. Este proceso de minimización es una suma de errores cuadráticos.

Formalmente:

Sea  $X = \{x_1, x_2, x_3, \dots, x_n\}$  un conjunto de datos, expresado como vector en un espacio  $d$ -dimensional  $x_i \in \mathbb{R}^d$

Y se tienen que definir ciertos criterios, tales como:

- $k$ : que representa el número de clusters.
- $C = \{\mu_1, \mu_2, \mu_3, \dots, \mu_k\}$ : son los centroides y  $\mu_j$  representa el vector de la media, o centro del cluster  $j$ .
- Cada punto  $x_i \in S_j$ , en donde  $S_j$  es un cluster.

Se tiene una función objetivo:

$$J = \sum_{j=1}^k \sum_{x \in S_j} |x_i - \mu_j|^2$$

Dónde:

$|x_i - \mu_j|^2$  es la distancia euclidiana entre la observación  $x_i$  y  $\mu_j$ , en otras palabras, entre un punto y el centroide definido perteneciente a un cluster.

Este algoritmo es iterativo, ya que se tiene que encontrar una solución óptima. Siguiendo esta lógica analítica, el algoritmo primero selecciona centroides iniciales aleatorios, para posteriormente se realiza el proceso de asignación para cada  $x_i$  en relación con su centroide más cercano.

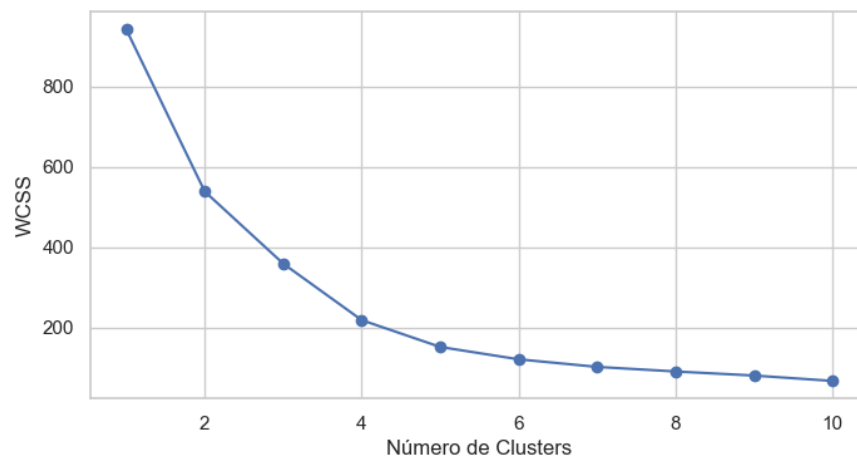
$$S_j = \{x_i: |x_i - \mu_j|^2 \leq |x_i - \mu_l|^2, \forall l, 1 \leq l \leq k\}$$

En la parte de la maximización se recalculan los centroides, a partir del promedio de los puntos asignados a cada cluster:

$$\mu_j = \frac{1}{|S_j|} \sum_{x_i \in S_j} x_i$$

Como es un proceso iterativo, se repiten los pasos hasta que los centroides en cuestión no cambien significativamente. Y en el caso de este documento de investigación, se utilizó el método de codo, mismo que grafica la inercia (J) en función de k; esto es, a medida que k (número de clusters) disminuye la inercia, entendiendo como punto óptimo el punto en donde la gráfica hace codo. Como se puede observar, el número óptimo de clusters se esboza propiamente en k=4.

Gráfico 10. Cálculo de clusters óptimo



Fuente: Elaboración propia en Python

De la misma manera, se presenta el algoritmo<sup>2</sup>, en código Python, para obtener el número óptimo de clusters (k=4).

```
# Determinamos el número óptimo de clusters (Método del Codo)
wcss = []
for i in range(1, 11):
    kmeans = KMeans(n_clusters=i, init='k-means++', random_state=42)
    kmeans.fit(X_scaled)
    wcss.append(kmeans.inertia_)

plt.figure(figsize=(8, 4))
plt.plot(range(1, 11), wcss, marker='o')
plt.title('Método del Codo')
plt.xlabel('Número de Clusters')
plt.ylabel('WCSS')
plt.show()

# Aplicamos K-Means con k=4 (hipótesis inicial)
k = 4
kmeans = KMeans(n_clusters=k, init='k-means++', random_state=42)
clusters = kmeans.fit_predict(X_scaled)

# Agregamos los clusters al dataframe original
df_2022['Cluster'] = clusters
pca_df['Cluster'] = clusters

# Visualización clusters en el espacio PCA
plt.figure(figsize=(10, 6))
sns.scatterplot(x='PC1', y='PC2', hue='Cluster', data=pca_df,
palette='viridis', s=100, alpha=0.8)
plt.title(f'Clusters de Sectores Industriales (k={k})')
plt.show()
```

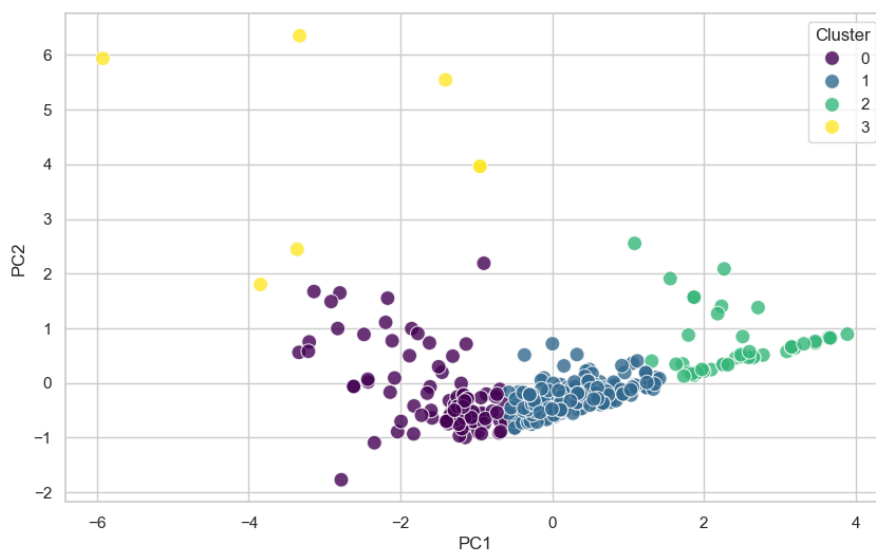
Código para la construcción de número óptimo de clusters, realizado en Python

Siguiendo esta optimización de clusters, en el gráfico 11) se muestra la clasificación que corresponde al análisis de segmentación de actividades económicas, misma que se basa en el análisis que corresponde al vector de varianza que se presenta en la parte anterior.

---

<sup>2</sup> Este código para la obtención de clusters óptimos se basa en el código presentado por Boschetti y Massaron, (2018, p. 269)

Gráfico 11. Clusters de actividades industriales con k=4



Fuente: Elaboración propia, en Python

A manera de precisión, las variables que se tomaron para la construcción de los clusters fueron las características estructurales:

- Coeficiente de insumo.
- Valor agregado.
- Productividad laboral.

En cuanto a los valores de los clusters:

Tabla 3. Valores de clusters

	Coeficiente de insumo	Valor agregado	Productividad laboral
<b>Cluster 0</b>	0.71	0.29	5,859
<b>Cluster 1</b>	0.56	0.44	1,828
<b>Cluster 2</b>	0.35	0.65	2,486
<b>Cluster 3</b>	0.67	0.33	33,240

Fuente: Elaboración Propia

---

Pero, ¿Qué es lo que indican los valores que le corresponden a los clusters y cuál es su relevancia en la política en aras del crecimiento económico? En este sentido, el análisis de datos se torna a una interpretación económica de los resultados, que no otra forma de mostrar el complemento analítico.

- En el cluster<sub>0</sub>: se tiene un valor que revela la dependencia de los insumos intermedios, además de que se tiene un bajo valor agregado (en comparación de los demás clusters). En términos de productividad, este segmento tuvo elementos medios. Estas actividades pueden tender a un alto encadenamiento para activar cadenas productivas, y así impulsar el empleo.
- En el cluster<sub>1</sub>: En este cluster se encuentra una dependencia moderada de insumos intermedios, el valor agregado es mayor al del cluster<sub>0</sub>, pero tiene la productividad más baja de las actividades económicas del sector manufacturero. Este conjunto de actividades económicas se presta para la apertura de innovación en el proceso de producción, ya que, si bien absorben algo de la oferta de trabajo, su fuerte es el valor agregado, no la generación de empleo.
- En el cluster<sub>2</sub>: Este conjunto tiene una baja dependencia a los insumos intermedios con un alto valor agregado y tiene una productividad baja.
- En el cluster<sub>3</sub>: Este conjunto del sector manufacturero es estratégico, ya que genera buen valor agregado, y tiene una productividad muy alta, es la clave para la competitividad y política de exportaciones, este cluster se identifica como el conjunto de actividades económicas clave para la potenciación del crecimiento económico, que es el objeto de este trabajo de investigación.

Y poniéndolo en términos de lo que ya se ha analizado, haciendo un análisis completo de datos que integra toda la narrativa construida, se tiene que las actividades que se encuentran en color azul dentro del gráfico 8) son las actividades clave dentro del sector manufacturero que sirven como potenciación del crecimiento económico de México.

A continuación, se ofrece un esbozo de la tabla 3) a manera de heatmap, para poder comparar los clusters en términos de productividad por colores.

Gráfica 12. Presentación de tabla 3, en formato Heatmap



Fuente: Elaboración propia, en Python

## Análisis de resultados

El análisis de clusters permitió segmentar la industria en grupos con características homogéneas estructurales, mismas que ya se tipificaron con anterioridad, y se obtuvo de manera satisfactoria y significativa que la manera óptima de segmentar el sector manufacturero es en 4 clusters ( $k=4$ ), y a partir de ahí, asignarles una relevancia estructural en aras de potenciar el crecimiento económico de México. Dichas asignaciones, propias de la política económica son las siguientes:

- Cluster<sub>0</sub>: Encadenamientos de actividad/sectorial.
- Cluster<sub>1</sub>: Dedicado a la política laboral (ideal para disminuir la tasa de desempleo).
- Cluster<sub>2</sub>: Política de innovación.
- Cluster<sub>3</sub>: Política de exportaciones y competitividad económica (cluster clave para el crecimiento del PIB).

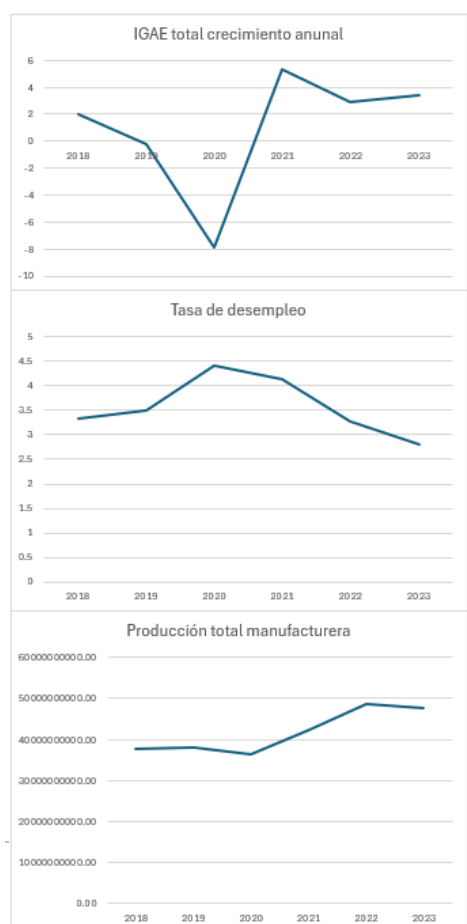
# I. Integración con variables macroeconómicas

Tabla 4. Datos macroeconómicos integrados

Año	Producción total manufacturera	IGAE total crecimiento anual	Tasa de desempleo
2018	37577980016.00	2.015	3.324632
2019	37968416192.00	-0.204167	3.490511
2020	36353539348.00	-7.851667	4.409108
2021	42407950408.00	5.314167	4.136034
2022	48601107808.00	2.8925	3.27043
2023	47830989364.00	3.408	2.795143

Fuente: Elaboración propia, con datos del INEGI

Gráfico 13

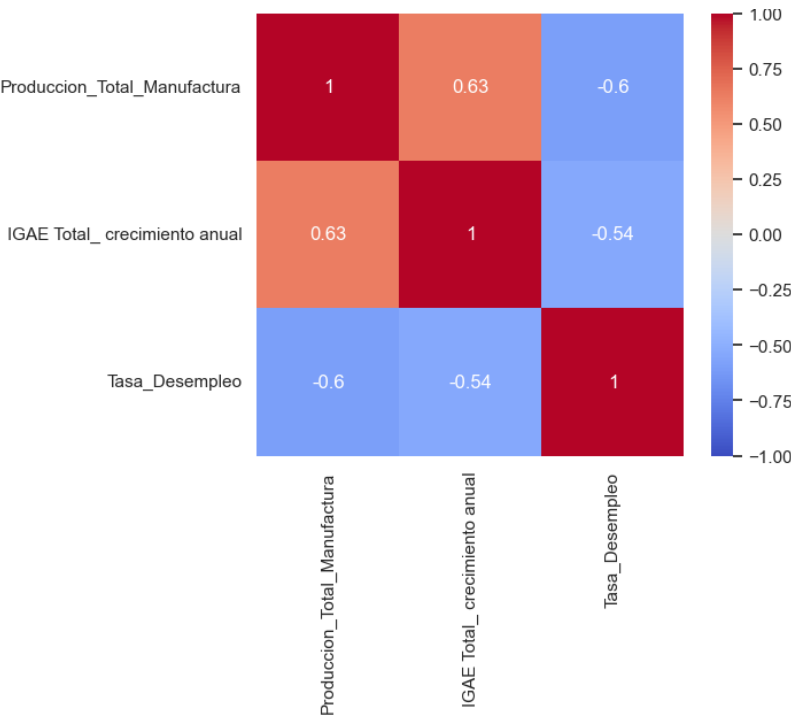


Fuente: Elaboración propia, con datos del INEGI

En la tabla 4) se vuelven a presentar los datos obtenidos, pero con un intervalo más corto, para verlo con más detalle. En 2020, la economía mexicana sufrió una contracción del 7%, seguida de una recuperación del 5.3% el vínculo explicativo que le precede es que la correlación que guarda con el sector manufacturero es positiva (0.63).

En lo que se refiere al coeficiente de  $-0.54$  (en el gráfico) indica que el crecimiento económico reduce el desempleo, pero no tan directamente como lo hace la manufactura, pero se tiene que tomar con reserva, ya que no todo crecimiento genera empleo si está concentrado en sectores capital-intensivos, como se demostró en este trabajo, ya que es fácil concluir que el crecimiento económico crea pleno empleo, esa idea no es del todo cierta, de acuerdo con los resultados.

Gráfico 14. Correlación de la producción manufacturera con el conjunto macroeconómico



Fuente: Elaboración Propia en Python



---

## Conclusiones

En este documento están presentes los cimientos analíticos desde la óptica interdisciplinaria de la economía como ciencia social y el proceder técnico de la ciencia de datos, ambas herramientas orientadas a la labor de la economía aplicada. Asimismo, se logra demostrar la complementariedad analítica de dichas disciplinas, y lo más importante, mostrar una vía para potenciar el crecimiento económico.

Identificar los clusters mediante un proceso iterativo de optimización representa una contribución significativa al estudio del crecimiento económico, abordando tanto el ámbito académico como la labor analítica requerida en el mercado laboral. Esta metodología tiene la capacidad de ser aplicada en la formulación de políticas económicas, así como en el desarrollo de modelos alternativos de teoría económica.

En este análisis, la utilización del algoritmo K-means permitió descomponer las actividades económicas pertenecientes al sector manufacturero. De esta manera, se evitó incurrir en el error teórico de suponer que estimular de manera uniforme al sector manufacturero conduce automáticamente al crecimiento económico. Los resultados obtenidos muestran la existencia de clusters óptimos, a los cuales se les asignó un rol específico conforme a sus características estructurales, permitiendo así la toma de decisiones más eficientes y basadas en datos concretos.

Adicionalmente, la construcción de una narrativa estructurada contribuyó a que la información disponible dejara de ser un conjunto de datos aislados y adquiriera un contexto relevante, transformando los datos en información significativa. Este enfoque deja abierta la puerta para que, en etapas posteriores, la segmentación lograda mediante la clusterización se utilice en el desarrollo del modelo de insumo-producto, tomando como base teórica la segmentación previamente obtenida y los aportes específicos de esta investigación.

---

## Bibliografía

- Aroche Reyes, F. (2019). Estructura productiva y crecimiento económico en México: una perspectiva multisectorial. *Investigación Económica*, 78(309), 3–26. <https://doi.org/10.22201/fe.01851667p.2019.309.70117>
- Boschetti, A., & Massaron, L. (2018). *Python data science essentials* (3rd ed.). Packt Publishing Ltd.
- Canto Sáenz, R. (2023). *Las políticas públicas en México y la insuficiencia de la teoría*. Revista Estado y Políticas Públicas, 20(mayo-septiembre), 75–89. ISSN 2310-550X.
- Chiang, A. C., & Wainwright, K. (2006). *Fundamentos de economía matemática* (4ª ed.). McGraw-Hill.
- Galindo, L. M., & Ros, J. (2006). *Banco de México: política monetaria de metas de inflación*. EconomíaUNAM, 3(9), 81–88.
- Varian, H. R. (2014). Big data: New tricks for econometrics. *Journal of Economic Perspectives*, 28(2), 3–28.
- Samuelson, P. A., & Nordhaus, W. D. (2009). *Macroeconomía* (19ª ed.). McGraw-Hill.
- Todaro, M. P., & Smith, S. C. (2020). *Economic Development* (13th ed.). Pearson Education.

## Anexos

Repositorio en GitHub en donde está el cuaderno de trabajo donde se desarrolló el proyecto, además de que se ofrece un dashboard interactivo con datos simulados:

[https://github.com/GabrielGM153/Modelo\\_IP\\_Clustering](https://github.com/GabrielGM153/Modelo_IP_Clustering)