

Tecnicatura Superior en Ciencia de Datos e Inteligencia Artificial

PRÁCTICA PROFESIONALIZANTE 1 - 2023

EQUIPO N°19 (Alumnos de ambas comisiones)

Docentes:

Narciso Hector Perez
Carlos Ignacio Charletti
Hector prado

Integrantes:

Lamine Mechedou
Micaela Kler
Liliana Elizabet Lamas
Juan Jure
Sthepany Carol Peves
Rubí Nohemí García Alarcón

Índice

Nombre del proyecto.....	3
Introducción.....	3
Objetivos del proyecto.....	3
Características del objetivo.....	4
Los datos de entrada.....	5
-Evolución del Producto Interno Bruto en Argentina.....	5
Clase de problema al cual nos enfrentamos.....	5
La mejora esperada.....	5
El estado actual de la variable objetivo.....	6
Medir la variable objetivo.....	6
Herramientas y tecnologías utilizadas.....	6
Límites y Alcances de Proyecto.....	6
-Limites.....	6
Alcances:.....	7
Aplicación.....	7
1-Definición del problema:.....	7
2- Recopilación de datos:.....	8
3-Limpieza y preparación de datos:.....	8
4-Exploración de datos:.....	8
5-Análisis de datos:.....	8
6-Interpretación de resultados:.....	16
Conclusiones para el Período 1980-1989.....	16
Conclusiones para el Período 2000-2010.....	17
Conclusiones Generales.....	17
7-Comunicación de resultados.....	17
GitHub.....	17
Tablero Trello.....	17
Video de Presentación.....	18

Nombre del proyecto

Análisis de la Inflación en Argentina y su Impacto en el PIB: Tendencias, Patrones y Perspectivas.

Introducción

En el contexto económico de Argentina, es fundamental comprender cómo la inflación ha influido en su Producto Interno Bruto (PIB) a lo largo de las décadas. Este proyecto tiene como objetivo principal identificar y analizar dos décadas clave en la historia económica del país: la primera, que abarca desde 1990 hasta 2000, se caracteriza por haber registrado la inflación más alta en el período de 1943 a 2023, mientras que la segunda, que se extiende desde 2000 hasta 2010, se distingue por ser la década con la inflación más baja en comparación con el mismo lapso.

Mediante un riguroso análisis de datos históricos, buscamos identificar patrones y factores determinantes que nos permitan arrojar luz sobre la compleja relación entre la inflación y el comportamiento del PIB en Argentina durante estas dos décadas específicas. Estos resultados no solo contribuirán a una comprensión más profunda de la dinámica económica del país en ese período, sino que también ofrecerán valiosas perspectivas sobre la influencia de la inflación en el crecimiento económico y el bienestar de la población.

El proyecto se basará en datos fiables y en un enfoque analítico sólido para proporcionar una visión más completa de cómo la inflación ha impactado la economía argentina en el pasado y, potencialmente, extraer lecciones valiosas para la gestión económica futura.

Objetivos del proyecto

El objetivo principal de este proyecto es identificar dos décadas, una con mayor inflación y otra con menor inflación, utilizando datos que abarcan el período desde 1943 hasta 2023. En particular, nos centramos en evaluar la influencia de la inflación en el comportamiento del Producto Interno Bruto (PIB) de Argentina durante estas dos décadas clave.

La primera de estas décadas se extiende desde 1990 hasta 1999 y se caracteriza por tener el índice de inflación más alto registrado en el período mencionado.

La segunda década abarca desde 2000 hasta 2010 y se destaca por ser la década con la inflación más baja en comparación con el mismo período.

A través del análisis de estos datos históricos, buscamos identificar patrones y factores determinantes que contribuyan a una comprensión más profunda de la dinámica económica de Argentina en estas dos décadas específicas. Los resultados obtenidos se utilizarán para proporcionar entender mejor el comportamiento de la inflación y su impacto en el PIB.

Características del objetivo

1- Enfoque en la inflación y el PIB:

El objetivo principal se centra en el análisis de la relación entre la inflación y el Producto Interno Bruto (PIB) de Argentina a lo largo de dos décadas específicas.

2-Ciclos económicos: Se investigan los ciclos económicos, lo que implica la identificación de patrones recurrentes en el comportamiento de la inflación y el PIB.

2-Identificación de décadas clave(Períodos de alta o baja inflación):

El proyecto busca identificar dos décadas de especial interés, una con alta inflación y otra con baja inflación.

3-Uso de datos históricos extensos:

Para alcanzar el objetivo, se utilizan datos que se remontan desde 1943 hasta 2023, lo que permite un análisis a largo plazo de la evolución de la economía argentina. Usamos también los datos del PIB en Argentina que se remontan desde 1962 hasta 2022.

4- Análisis de correlaciones:

Explorar las relaciones entre la inflación y el PIB a lo largo del tiempo y determinar si existen correlaciones significativas entre la tasa de inflación y el crecimiento del PIB.

5- Identificación de patrones y factores determinantes(El contexto de los datos):

Se busca identificar patrones históricos y factores clave que pueden explicar el comportamiento del PIB en las décadas estudiadas. Estos factores pueden incluir políticas económicas, eventos económicos y sociales, entre otros.

5-Contexto histórico: Tener en cuenta el contexto histórico del país, la región y del mundo ya que los eventos históricos pueden influir en los patrones de inflación y PIB.

6-Interpretación de resultados: Finalmente, el proyecto implicaría la interpretación de los resultados obtenidos a partir del análisis de datos. Esto puede ayudar a comprender mejor la evolución económica y proporcionar información útil para la toma de decisiones económicas y políticas.

Los datos de entrada

Los datos de entrada utilizados para llevar a cabo este proyecto son los siguientes:

-Los registros históricos de la inflación en Argentina, incluyen tasas de inflación anuales a lo largo de los años desde 1943 hasta 2023.

El link: <https://datacraft.com.ar/historia-inflacion-argentina/>

-Evolución del Producto Interno Bruto en Argentina.

El link: [Anexo:Evolución del Producto Interno Bruto en Argentina - Wikipedia, la enciclopedia libre](#)

Clase de problema al cual nos enfrentamos

Nos enfrentamos a un problema de **agrupamiento**. El objetivo es agrupar los datos históricos de inflación en categorías o **clusters** que representan diferentes comportamientos de inflación y PIB en Argentina a lo largo del tiempo, así que en este proyecto se usa el algoritmo de agrupamiento (**clustering**) **K-MEANS**

La mejora esperada

La mejora esperada es una mejor comprensión de los patrones de inflación y su relación con el PIB a lo largo del tiempo en Argentina y caracterizar de manera más efectiva diferentes períodos de inflación.

El estado actual de la variable objetivo

La variable objetivo en este caso es la inflación y el PIB a lo largo del tiempo en Argentina. El estado actual se puede describir mediante los datos históricos disponibles, que representan las tasas de inflación mensuales o anuales desde 1943 hasta 2023 y las tasas de la evolución del PIB desde 1962 hasta 2022.

Medir la variable objetivo

La variable objetivo, que es la inflación, se mide en términos de tasas de inflación. utilizamos medidas estadísticas para describir los patrones, como medias, desviaciones estándar y otros indicadores de tendencia y volatilidad

Herramientas y tecnologías utilizadas

- Python** fue el lenguaje de programación principal seleccionado para este proyecto debido a su amplia adopción en el análisis de datos y la disponibilidad de bibliotecas relevantes.
- Entorno de ejecución elegido:** Google Colab
- Pandas:** Empleado para la manipulación y análisis de datos tabulares.
- NumPy:** Utilizado para realizar operaciones numéricas y matriciales.
- Matplotlib:** Empleado para la creación de visualizaciones gráficas.
- Scikit-Learn:** Utilizado para la implementación de modelos de predicción, en particular, se aplicó el algoritmo de **K-Means** para el análisis de agrupamiento.
- Tableau:** Herramienta de visualización de datos que se utilizó para crear visualizaciones interactivas.
- Para el análisis estadístico usamos **Coefficiente de correlación de Pearson** para determinar si hay correlaciones

Límites y Alcances de Proyecto

-Límites

-**Dependencia de la calidad de los datos:** La calidad de los resultados depende en gran medida de la calidad de los datos recopilados. Errores o inconsistencias en los datos pueden afectar la precisión de las conclusiones.

Simplificación de la complejidad económica: El proyecto se centra en la relación entre la inflación y el PIB, pero la economía es un sistema complejo con múltiples factores que pueden influir en el comportamiento económico.

Análisis correlacional: Aunque se analiza la correlación entre la inflación y el PIB, no se establece una relación causal. Otros factores no considerados en el proyecto pueden influir en los resultados.

Estabilidad de políticas económicas: Las políticas económicas pueden cambiar a lo largo del tiempo, y el proyecto no puede capturar todas las dinámicas políticas que afectan la economía.

Alcances:

Identificación de patrones históricos: El proyecto tiene el potencial de identificar patrones históricos en la relación entre la inflación y el PIB en Argentina, lo que puede ser útil para comprender la dinámica económica.

Caracterización de décadas clave: Se pretende caracterizar las dos décadas específicas con alta y baja inflación para entender cómo afectaron al PIB y cuáles fueron los factores clave en juego.

Información para la toma de decisiones: Los resultados pueden proporcionar información valiosa para la formulación de políticas económicas y la toma de decisiones en Argentina.

Visualización de datos interactiva: La utilización de herramientas como Tableau permite una comunicación efectiva de los resultados a través de visualizaciones interactivas.

Aplicación

Para la aplicación de este proyecto, hemos seguido de manera cuidadosa y rigurosa el ciclo de vida del análisis de datos. Cada etapa de este ciclo se ha abordado con el objetivo de obtener resultados sólidos y significativos que nos permitan comprender la relación entre la inflación y el Producto Interno Bruto (PIB) de Argentina durante las dos décadas clave que hemos identificado. A continuación, describiremos cómo hemos aplicado cada etapa del ciclo de vida del análisis de datos en este proyecto:

1-Definición del problema:

Iniciamos el proyecto identificando claramente nuestro objetivo principal: analizar la influencia de la inflación en el PIB de Argentina durante las dos décadas específicas, una con alta inflación y otra con baja inflación. También hemos establecido las preguntas clave que queremos responder.

Estas son algunas preguntas realizadas durante el brainstorming para entender bien el problema:

-¿Cuál es el impacto de la inflación en el PIB de Argentina?

-¿Existe una correlación significativa entre los niveles de inflación y el crecimiento del PIB en Argentina?

-¿Cómo se comparan los indicadores económicos (inflación y PIB) en la década con mayor inflación con respecto a la década con menor inflación?

-¿Cuáles son los factores económicos y políticos que podrían haber contribuido a las diferencias en los niveles de inflación y su impacto en el PIB en estas dos décadas?

-¿Cuáles son las implicaciones de estos hallazgos para la formulación de políticas económicas en Argentina?

-¿Se pueden identificar patrones a lo largo del tiempo en la relación entre la inflación y el PIB en el país?

2- Recopilación de datos:

Para garantizar la calidad de los datos, hemos recopilado información económica histórica de fuentes confiables, como bancos centrales, organismos gubernamentales y organizaciones internacionales. Hemos utilizado datos de inflación que abarcan desde 1943 hasta 2023 datos del PIB que abarcan desde 1962 hasta 2022

3-Limpieza y preparación de datos:

En esta etapa, hemos realizado una exhaustiva limpieza de los datos, identificando y corrigiendo cualquier posible inexactitud, eliminando valores que se desvían notablemente de la norma y abordando los valores faltantes. Asimismo, hemos estandarizado el formato de los datos de manera coherente para optimizar su idoneidad para el análisis.

La limpieza de datos se ejecutó mediante una herramienta de procesamiento en una hoja de cálculo de Excel, la cual posteriormente fue transferida a un archivo CSV para su importación a la plataforma Colab. Esta secuencia de pasos nos permitió llevar a cabo la exploración de datos de manera eficiente y efectiva.

4-Exploración de datos:

En esta etapa, el objetivo principal es comprender la naturaleza de los datos y obtener una visión general de la relación entre la inflación y el PIB en Argentina a lo largo del tiempo.

5-Análisis de datos:

En primer lugar, nos enfrentamos a un problema de agrupamiento y por eso elegimos el algoritmo de agrupamiento K-MEANS (clustering).

En segundo lugar, tuvimos que elegir el número óptimo de clusters, por eso usamos el método "**codo**" (**Elbow Method**). El método del codo implica ejecutar el algoritmo de agrupamiento para un rango de valores de k (número de clusters) y evaluar la variación intra-cluster en función de k. Luego, se busca el punto en el gráfico donde la variación intra-cluster disminuye significativamente, creando una apariencia de "codo" en la curva.

Esto se llevó a cabo con el siguiente código:

```
import pandas as pd
import numpy as np
from sklearn.cluster import KMeans
import matplotlib.pyplot as plt

# Cargar el archivo CSV en un DataFrame
df = pd.read_csv("inflacion_argentina_43-23.csv")

# Convierte la columna 'Año' a tipo de fecha
df['AÑO'] = pd.to_datetime(df['AÑO'], format='%Y')

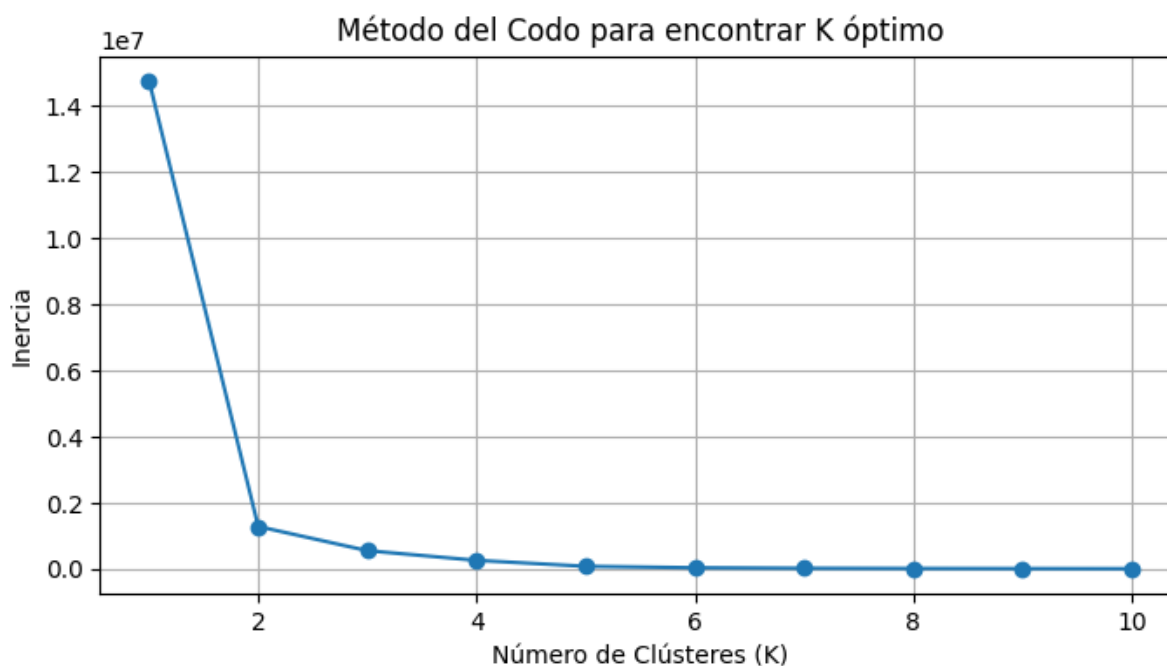
# Selecciona la tasa de inflación como la característica a analizar
X = df[['INFLACION']]

# Entrena el modelo K-Means con diferentes valores de K
k_values = range(1, 11)
inertia = []

for k in k_values:
    kmeans = KMeans(n_clusters=k, random_state=42)
    kmeans.fit(X)
    inertia.append(kmeans.inertia_)

# Grafica del método del codo (Elbow Method) para encontrar el número
# óptimo de clústeres
plt.figure(figsize=(8, 4))
plt.plot(k_values, inertia, marker='o', linestyle='-')
plt.xlabel('Número de Clústeres (K)')
plt.ylabel('Inercia')
plt.title('Método del Codo para encontrar K óptimo')
plt.grid(True)
plt.show()
```

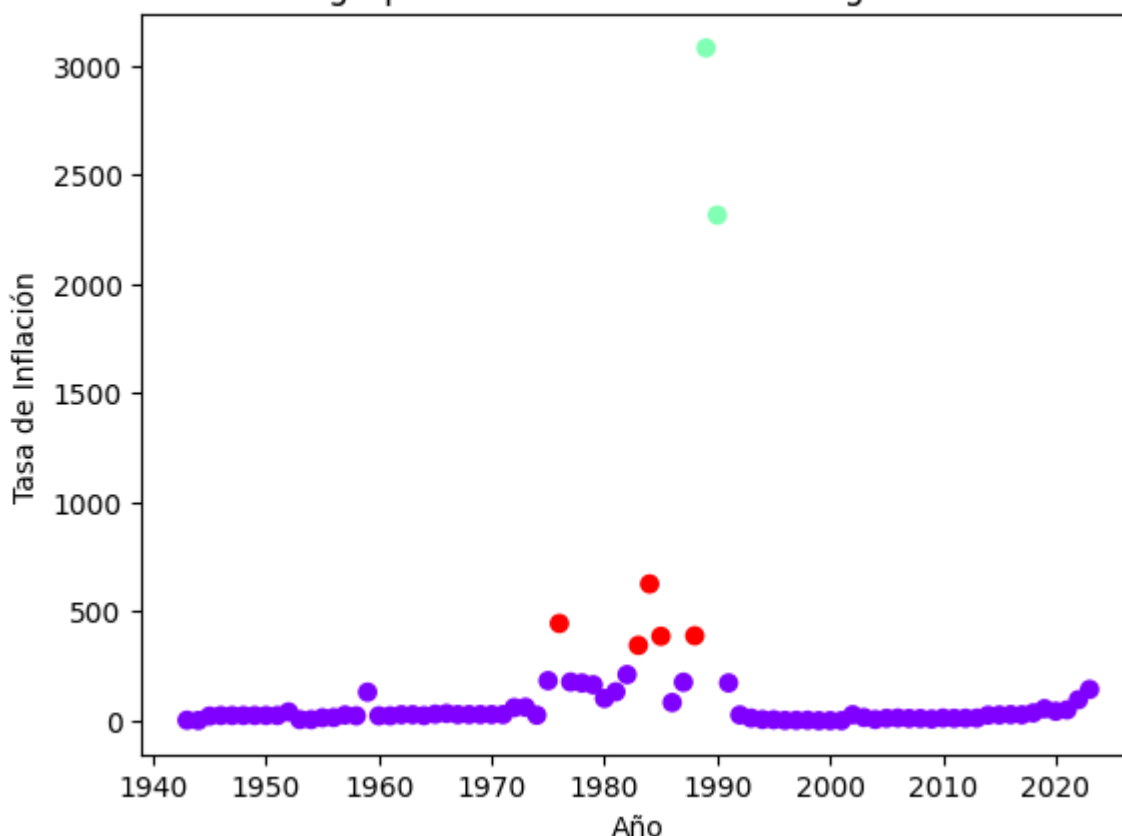
A través del siguiente gráfico, podemos observar que el punto de inflexión de esta curva está en el valor de $k = 2$. Por lo tanto, podemos decir que el número «correcto» de grupos para estos datos es 2



Una vez determinado el número de Clusters óptimos procedemos a usar la biblioteca de python llamada **sklearn** para entrenar nuestro modelo para analizar en el paso siguiente el comportamiento de la inflación.

El siguiente paso es cargar datos de inflación de Argentina, aplicar el algoritmo de K-Means para agrupar los datos en clusters, y luego visualizar los resultados en un gráfico de dispersión donde los puntos de datos se colorean. Esto puede ayudar a identificar patrones en el comportamiento de la inflación a lo largo del tiempo.

Agrupamiento de Inflación en Argentina

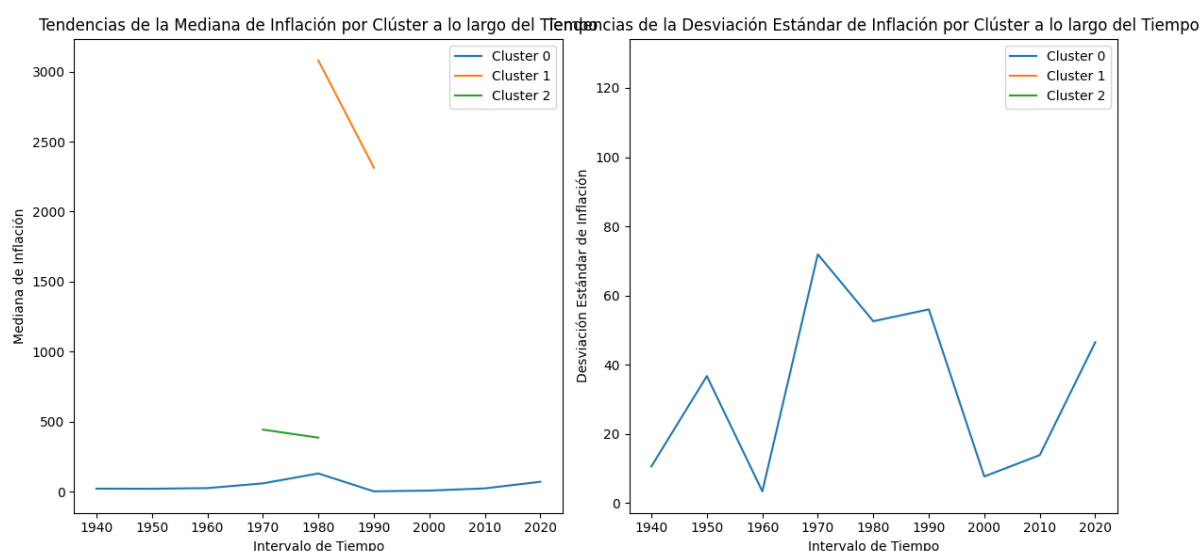


Hemos realizado también **una evaluación del costo computacional** y conseguimos el siguiente resultado: **Tiempo transcurrido para entrenar el modelo K-Means:**

0.011434793472290039 segundos

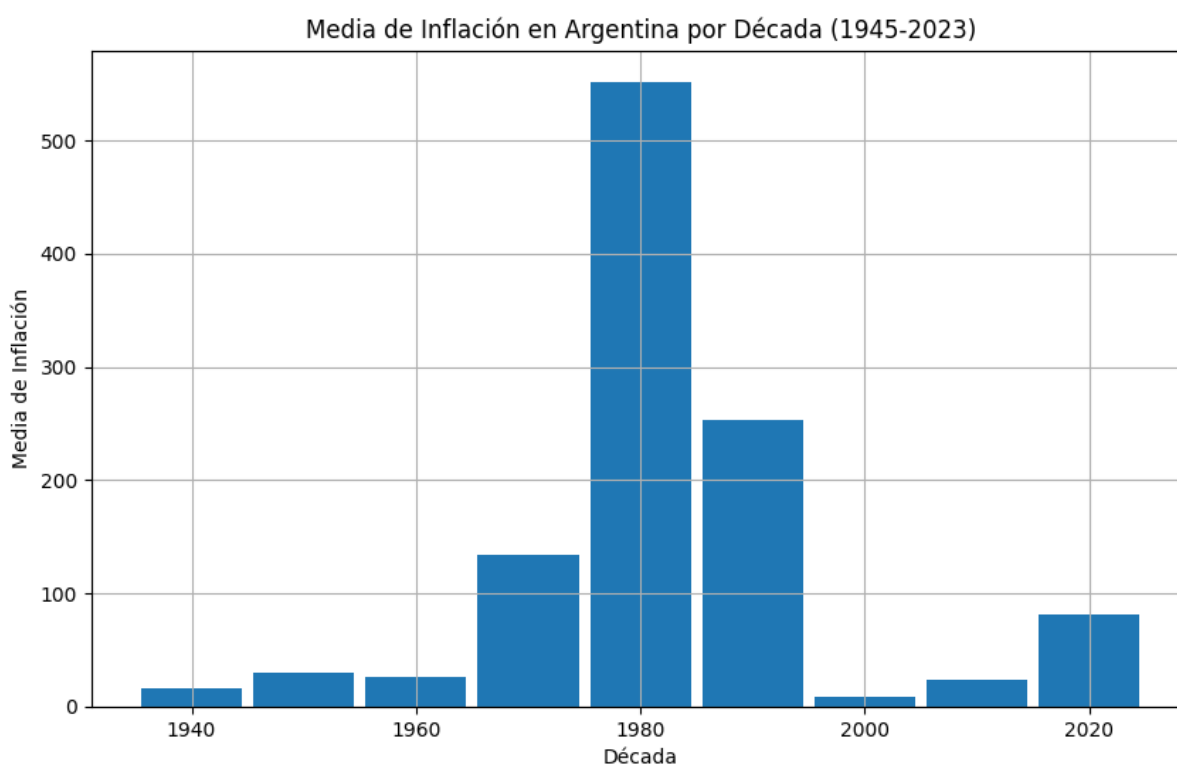
-Para poder hacer un análisis eficiente hemos realizado los siguientes cálculos:

-calcular la **media** de la inflación y **desviación estándar** a lo largo de todas las décadas que van desde 1943 hasta 2023 como se ve en las siguientes gráficas:



-Encontrar la década con el promedio **mayor** de inflación

-Encontrar la década con el promedio menor de inflación

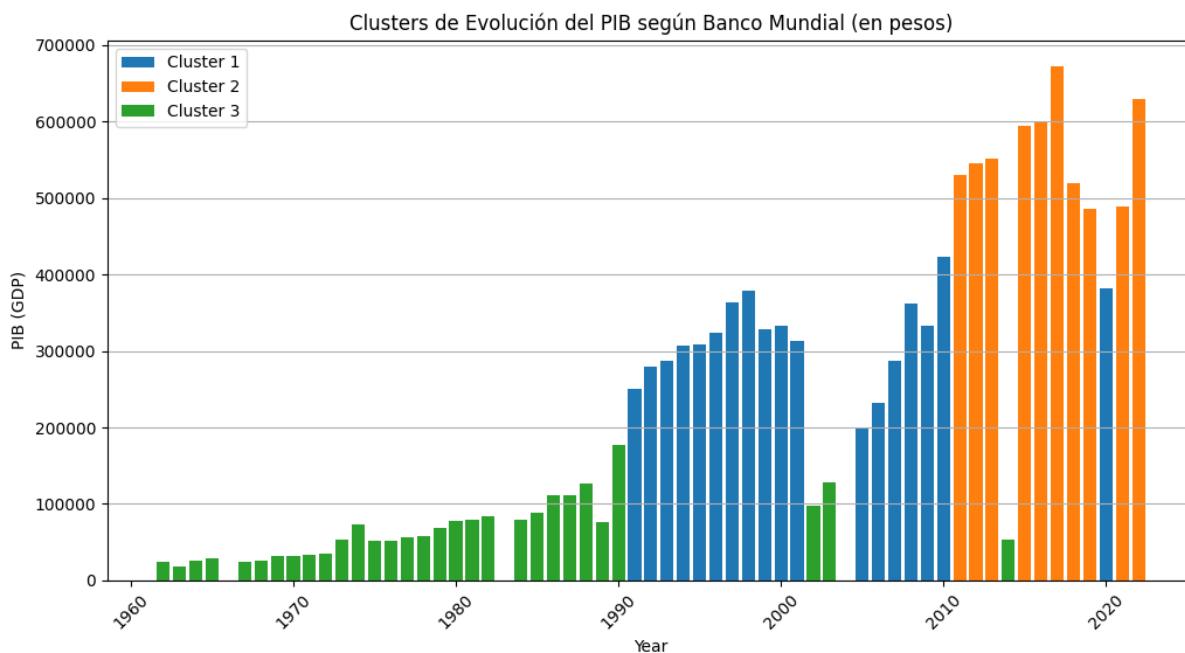


-El año con mayor inflación

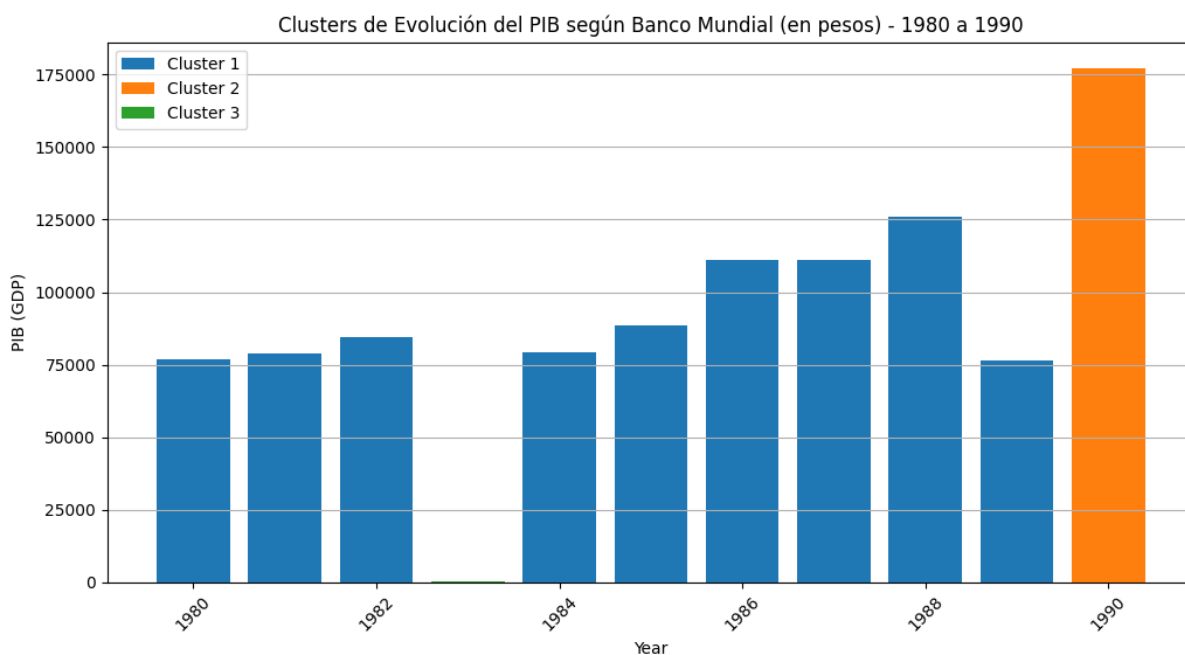
-El año con menor inflación

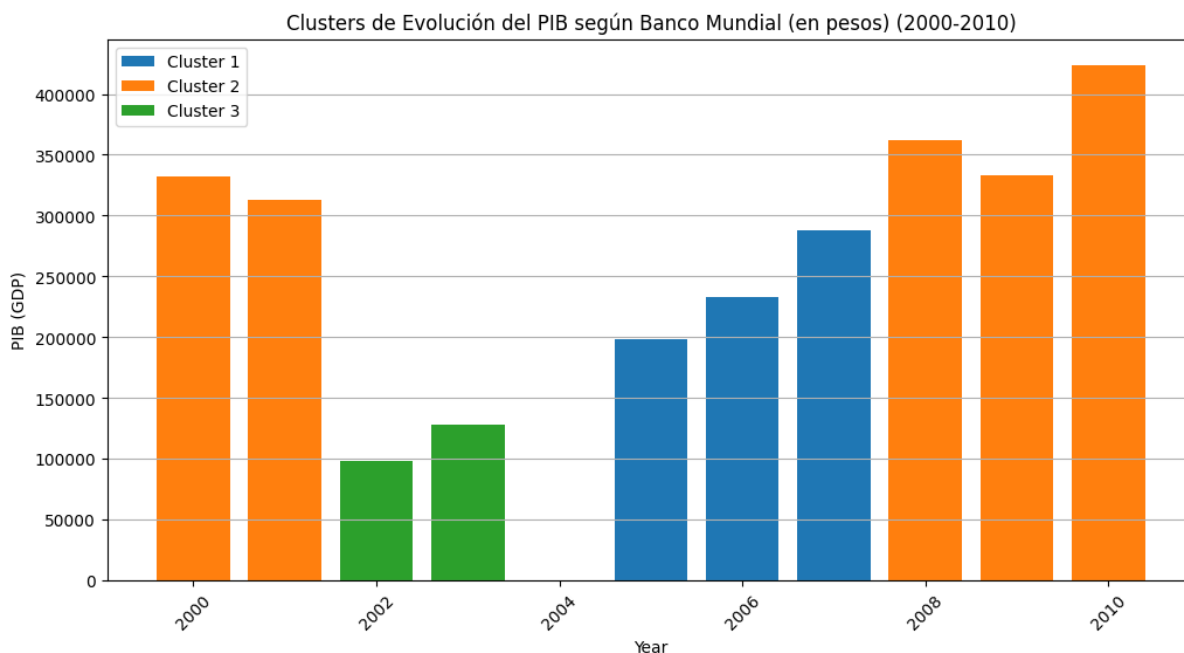
Como podemos notar la década con mayor inflación es la de 1980-1990 y la que tiene una media de inflación menor es la de 2000-2010, así que son las dos décadas que van a ser analizadas comparándolas con el comportamiento del PIB.

Hemos cargado en el siguiente paso los datos del PIB para analizar su comportamiento y tendencias que van desde 1962 hasta 2022 y realizamos los mismos pasos para determinar el número óptimo de clusters



En el siguiente paso analizamos las tasas del PIB en las décadas **1980-1990 y 2000-2010** para compararlas con las tasas de inflación en las mismas décadas.

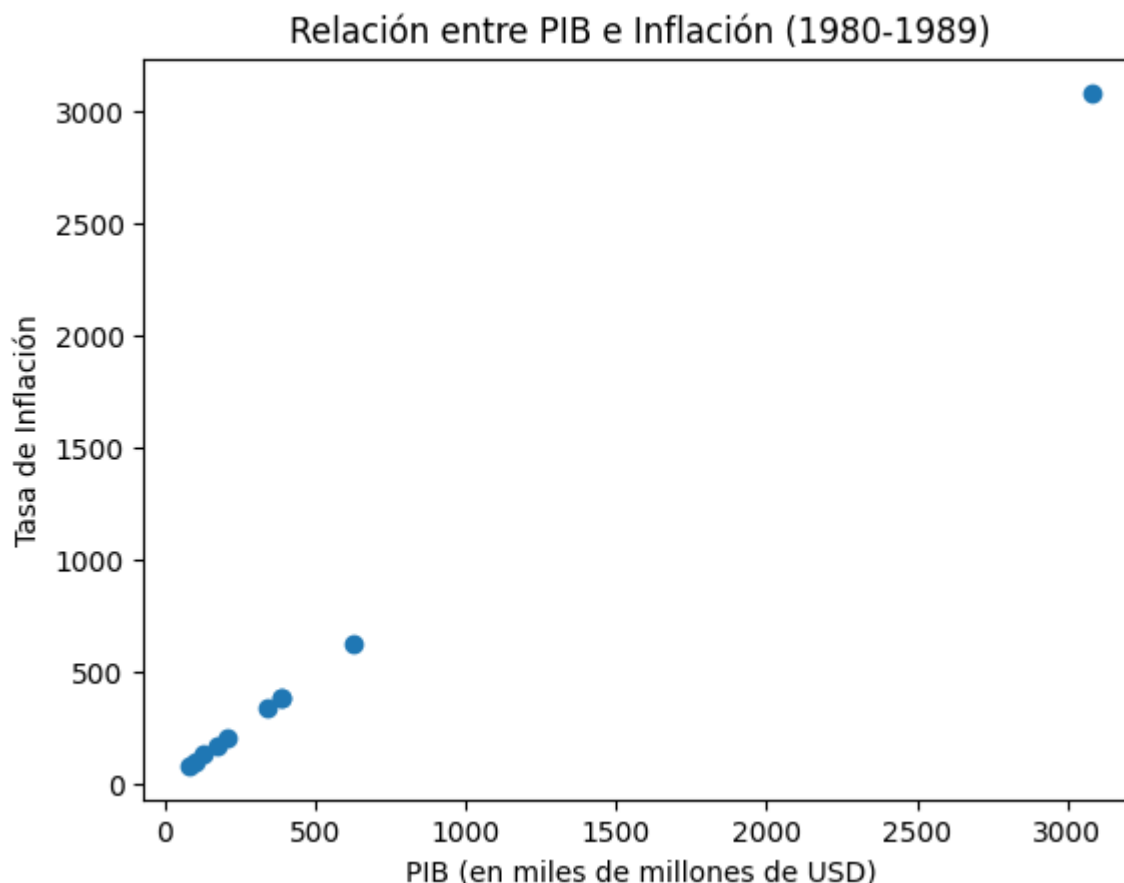




Para determinar si existe una **correlación** entre el Producto Interno Bruto (PIB) y la tasa de inflación durante el período de **1980 a 1989** en Argentina, llevamos a cabo un análisis exploratorio inicial y **calcular la correlación** entre estas dos variables.

Para eso usamos **El coeficiente de correlación de Pearson** te dará una medida de la relación lineal entre las dos variables.

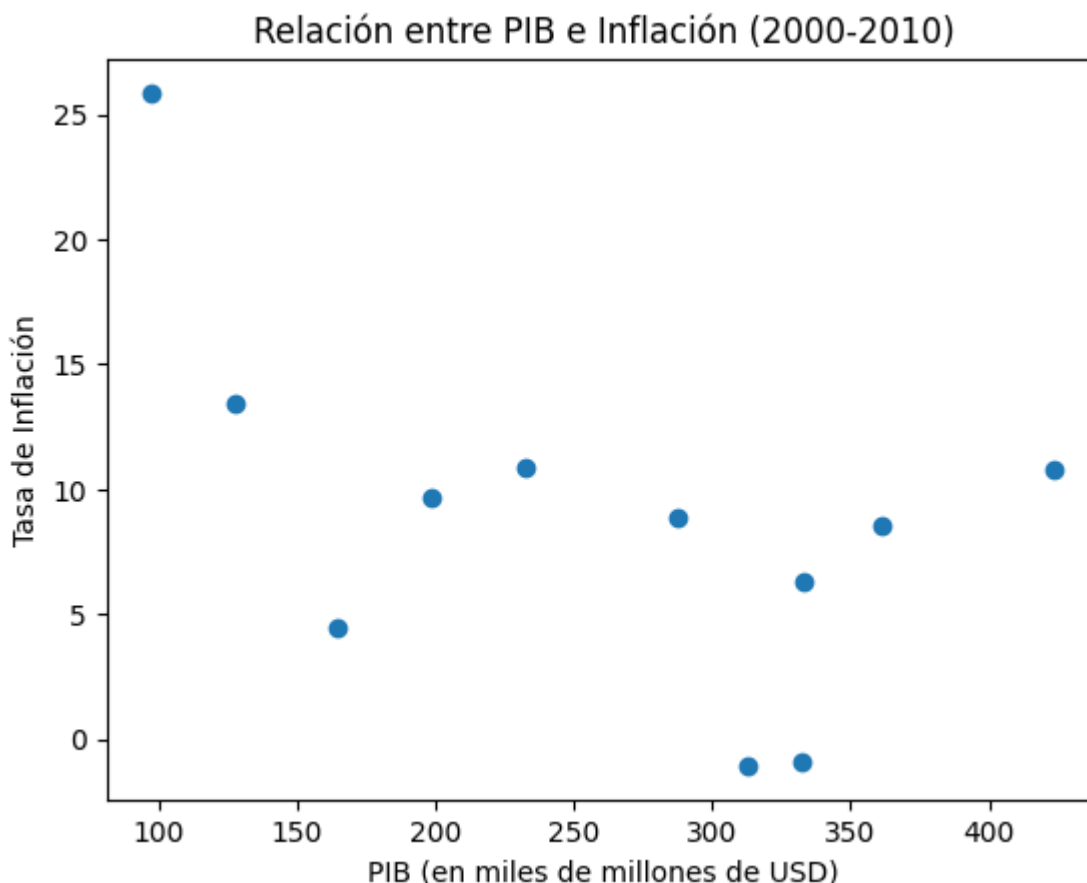
La gráfica obtenida:



El resultado del **coeficiente de correlación de Pearson de 0.9999999999999999** para el período de **1980-1989** indica una correlación positiva **extremadamente fuerte** entre el Producto Interno Bruto (PIB) y la tasa de inflación en Argentina durante ese período. Esto significa que a medida que el PIB **aumenta**, la tasa de inflación también tiende a **aumentar**, y viceversa. Es importante señalar que un coeficiente de correlación tan cercano a 1 sugiere una relación casi perfecta entre estas dos variables en ese período.

-Ahora analizamos si hay una correlación entre el PIB y la inflación en la década que representa la menor media de tasa de inflación que va desde 2000-2010.

el gráfico obtenido es:



El coeficiente de correlación de Pearson de -0.5460 indica una correlación **moderada** y negativa entre el PIB y la inflación durante el período de 2000 a 2010 en Argentina. Esto sugiere que, en general, existe una relación inversa entre estas dos variables en ese período.

6-Interpretación de resultados:

Conclusiones para el Período 1980-1989

Coeficiente de correlación de Pearson: **0.9999999999999999**

En el período de 1980 a 1989, el coeficiente de correlación de Pearson es de **0.9999999999999999** indica una correlación positiva **extremadamente fuerte** entre el Producto Interno Bruto (PIB) y la tasa de inflación en Argentina durante ese período. Esto significa que a medida que el PIB **aumenta**, la tasa de inflación también tiende a **aumentar**, y viceversa. Es importante señalar que un coeficiente de correlación tan cercano a 1 sugiere una relación casi perfecta entre estas dos variables en ese período.

Los datos sugieren que hubo **una relación inversa** entre el crecimiento económico (PIB) y la tasa de inflación en Argentina en ese período.

El conocimiento de esta correlación puede ser útil para la formulación de políticas económicas y la toma de decisiones.

Conclusiones para el Período 2000-2010

Coefficiente de correlación de Pearson: -0.5460386189661915

En el período de 2000 a 2010, el coeficiente de correlación de Pearson

-0.5460386189661915 indica una correlación negativa **moderada** entre el PIB y la tasa de inflación en Argentina. Esto sugiere que, en general, existe una relación inversa entre estas dos variables en ese período, aunque es menos intensa que en el período anterior.

Durante este período, también hubo fluctuaciones en la inflación y el PIB, aunque la relación **no es tan fuerte** como en la década anterior.

La economía argentina mostró una **correlación inversa**, lo que significa que cuando el PIB **crecía**, la inflación tendía a **disminuir** y viceversa.

Los cambios en la política económica y otros factores pueden haber influido en esta relación menos intensa en comparación con el período anterior.

Conclusiones Generales

Ambos períodos muestran una **correlación** entre el PIB y la inflación, lo que indica una relación inversa en la que el crecimiento económico se asocia con una menor inflación.

A pesar de la correlación negativa, es importante recordar que la correlación no implica causalidad. Otros factores económicos y políticos pueden influir en estas variables.

Estos resultados pueden proporcionar información útil para la formulación de políticas y la toma de decisiones en el ámbito económico en Argentina pero sabemos que muchos otros factores influyen en la inflación y el objetivo de este trabajo es dar información valiosa sobre el impacto del PIB en la inflación.

7-Comunicación de resultados

Hemos preparado informes detallados y visualizaciones efectivas para comunicar nuestros hallazgos de manera clara y accesible a los interesados.

Usamos Tableau como herramienta principal para la visualización de datos.

GitHub

Link: [garciarubi/Grupo19-PP1 \(github.com\)](https://github.com/garciarubi/Grupo19-PP1)

Tablero Trello

Link: <https://trello.com/b/Wij5Ek3W/grupo-19>

Video de Presentación

Link:

https://drive.google.com/file/d/1euYybov-rGcqU2zAycj2vBOZ9k2HCvSc/view?usp=drive_link