



UNIVERSIDAD TÉCNICA PARTICULAR DE LOJA

CARRERA:

INGENIERÍA EN COMPUTACIÓN

MATERIA:

PROGRAMACIÓN AVANZADA

SEMESTRE:

CUARTO SEMESTRE

DOCENTE:

ING. DANILO JARAMILLO

INTEGRANTES:

CARLOS VALLADARES

DIANA BERMEO

JORGE BELTRÁN

Importaciones Realizadas en el Ecuador

Índice

1.	Introducción	3
2.	Justificación	3
3.	Objetivos	5
4.1.	Incorporar variables contextuales externas	5
4.2.	Facilitar análisis visuales y comparativos	6
4.	Desarrollo.....	6
5.1.	Integración de fuentes externas de datos.....	6
5.2.	Variables seleccionadas	7
5.3.	Utilidad del enriquecimiento del dataset.....	9
5.4.	Implementación en Apache Spark (Zeppelin)	10
5.4.1.	Carga de los archivos.....	13
5.4.2.	Normalización de atributos para integración de datos.....	10
5.4.3.	Agregación de Nuevas Variables al DataFrame Principal.....	¡Error! Marcador no definido.
	Paso 1. Preparación del entorno.	11
	Paso 2. Normalización del campo PAIS_ORIGEN.....	12
	Paso 3. Cálculo de variable CARGA_TRIBUTARIA_TOTAL.	13
	Paso 4. Incorporación de Variables Externas (PIB_PERCAPITA, INFLACION).	¡Error! Marcador no definido.
	Paso 5. Limpieza Final de Valores Nulos. Para asegurar la consistencia de análisis...	¡Error! Marcador no definido.
	Paso 6. Verificación de integración exitosa.....	¡Error! Marcador no definido.
5.6	Análisis exploratorio de datos por medio de consultas.	15
5.6.1.	Consulta 1: ¿Qué países representan, en promedio, una mayor carga tributaria sobre sus importaciones hacia el Ecuador?.....	15
5.	Conclusiones.....	24
6.	Bibliografía	¡Error! Marcador no definido.
7.	Anexos.....	26

1. Introducción

El presente informe tiene como objetivo analizar el comportamiento de las importaciones realizadas en el Ecuador bajo el régimen general durante el primer semestre del año 2024, a partir de los datos abiertos publicados por el Servicio Nacional de Aduana del Ecuador (SENAE). Este conjunto de datos recoge, de forma detallada y estructurada, las declaraciones aduaneras que han sido aprobadas oficialmente para el ingreso de mercancías al país, proporcionando una visión integral de los flujos comerciales en tiempo real.

La información contenida en este archivo es de carácter oficial y público, y se ha puesto a disposición a través del portal nacional de datos abiertos, garantizando así su legitimidad, trazabilidad y utilidad para fines analíticos, institucionales, empresariales y académicos.

El análisis de estas operaciones no solo permite describir el comportamiento del comercio exterior, sino también identificar patrones estratégicos en la dinámica económica del país, evaluar la efectividad de las políticas comerciales vigentes, y anticipar riesgos logísticos, tributarios o de abastecimiento.

2. Justificación

El archivo de declaraciones aduaneras de importación bajo régimen general constituye una fuente crítica de información económica, pues ofrece respuestas concretas a una serie de preguntas clave:

- **¿Qué se está importando?:** A través de las variables DESCRIPCION_ARANCELARIA y SUBPARTIDA, se identifica con precisión qué tipos de productos están ingresando al país, permitiendo clasificar las

importaciones por sectores productivos, bienes de consumo, materias primas o tecnologías.

- **¿Desde dónde provienen estas mercancías?:** La variable PAIS_ORIGEN permite mapear la red internacional de proveedores del Ecuador, identificando socios comerciales prioritarios, posibles dependencias estratégicas y cambios en las rutas comerciales globales.
- **¿Bajo qué condiciones legales se importan?:** Mediante los campos REGIMEN, TRATAMIENTO_PREFERENCIAL y CONVENIO_INTERNACIONAL, es posible analizar el marco jurídico-comercial que regula cada operación, evidenciando el aprovechamiento (o no) de tratados internacionales, preferencias arancelarias y regímenes especiales.
- **¿Qué cargas impositivas enfrentan las importaciones?:** Variables como IVA, FODINFA, ADVALOREM, SALVAGUARDIA, ICE_ADVALOREM, entre otras, ofrecen un panorama completo de los tributos aplicados a cada ítem importado, facilitando el estudio de su impacto fiscal, su peso en el costo final, y la distribución tributaria por sector económico.
- **¿Cuál es el costo real de traer los productos al país?:** Los campos FOB, FLETE, SEGURO y CIF permiten calcular el valor total de importación en aduana, desagregando sus componentes logísticos y comerciales. Este análisis es esencial para entender cómo se forman los precios de los bienes importados y cómo inciden los costos internacionales en la economía local.

- **¿Cuándo y dónde fueron aprobadas estas importaciones?:** Gracias a las variables FEC_INGRESO, ESTADO_DECLARACION y DISTRITO, es posible realizar análisis temporales y geográficos, evaluar la eficiencia operativa de los distritos aduaneros, y establecer tendencias de ingreso de mercancías a lo largo del semestre.

Esta riqueza de variables convierte al dataset en una herramienta insustituible para quienes desean comprender la estructura económica nacional desde una perspectiva de comercio exterior, con un enfoque técnico, actual y basado en evidencia. Por estas razones, se justifica plenamente su elección como insumo para este ejercicio de análisis de datos.

3. Objetivos

El objetivo general de este proyecto es enriquecer y analizar el conjunto de datos de importaciones del Ecuador, bajo el régimen general, correspondiente al primer semestre de 2024(enero - junio). Para esto, se busca no solo comprender los datos internos del comercio registrado, sino también incorporar variables externas y de contexto económico internacional que permitan obtener una visión más completa del fenómeno importador.

4.1. Incorporar variables contextuales externas

Se prevé agregar al dataset principal tres nuevas variables relacionadas con el país de origen de las mercancías importadas:

- Una medida agregada de la carga tributaria total por ítem, basada en la suma de los tributos registrados en la declaración.

- El PIB per cápita del país exportador, como indicador del nivel de desarrollo económico. [1]
- La inflación anual reciente del país exportador, como factor de presión externa sobre los precios. [2]

Estas variables permitirán contextualizar mejor las operaciones de importación, relacionando sus costos y tributos con condiciones económicas externas.

4.2. Facilitar análisis visuales y comparativos

Con el conjunto de datos enriquecido, se espera habilitar consultas y visualizaciones que permitan:

- Comparar carga tributaria entre países y tipos de productos.
- Explorar la relación entre desarrollo económico y valor de importación.
- Representar espacialmente los países exportadores según indicadores relevantes.

4. Desarrollo

5.1. Integración de fuentes externas de datos

Con el fin de enriquecer el análisis de las importaciones registradas en Ecuador bajo régimen general, se identificaron dos fuentes internacionales clave que permitirán contextualizar económicamente los datos existentes:

- **Banco Mundial – PIB per cápita (PPP, US\$ constantes):**

Esta fuente proporciona información sobre el Producto Interno Bruto per cápita,

ajustado por paridad de poder adquisitivo, lo cual permite comparar de forma más precisa el nivel de desarrollo económico de los países exportadores hacia Ecuador.

- **Base Global de Inflación – Base Global de Inflación (2024):**

Este conjunto de datos ofrece tasas anuales de inflación por país. Su inclusión es relevante para estudiar el efecto de contextos económicos inflacionarios en los costos de importación (FOB, CIF), y detectar posibles riesgos derivados de relaciones comerciales con economías inestables.

- **Country-data:**

El archivo en formato tabla contiene información clave para la normalización y vinculación de datos, gracias a la columna alpha3, que almacena el código ISO alfa-3 (de tres letras) de cada país. Esta codificación estandarizada facilita la integración entre diferentes fuentes, permitiendo relacionar correctamente los registros de variables como el PIB per cápita, la inflación y los valores de importación.

El uso combinado de estas fuentes permite incorporar dimensiones macroeconómicas externas al análisis de comercio internacional, fortaleciendo la capacidad explicativa del modelo de datos.

5.2. Variables seleccionadas

Se decidió incorporar al DataFrame general de importaciones, un total de tres nuevas variables, dos de ellas provenientes de fuentes externas y una calculada internamente:

	PIB_PERCAPITA	INFLACION	CARGA_TRIBUTARIA_TOTAL
--	----------------------	------------------	-------------------------------

Origen	Archivo del Banco Mundial sobre PIB per cápita.	Archivo del Banco Mundial sobre Base Global de Inflación.	Derivada de columnas del DataFrame original de importaciones.
Valor agregado	Permite medir el nivel de desarrollo económico del país de origen de la importación, lo cual puede ayudar a analizar si existe relación entre este indicador y los precios, tipos de productos o niveles de tributo aplicados	Proporciona un indicador de presión inflacionaria en el país exportador. Esto es útil para detectar si economías con inflación alta impactan en mayores precios de importación o volatilidad en costos logísticos	Esta métrica resume el impacto fiscal total por ítem importado, facilitando comparaciones directas entre productos, países o regímenes aduaneros, y sirviendo como insumo para decisiones tributarias o evaluaciones de carga impositiva por sector.

- **Fórmula de cálculo de la variable 3:** $CARGA_TRIBUTARIA_TOTAL = FODINFA + ICE_ADVALOREM + ICE_ESPECIFICO + IVA + ADVALOREM$

Se eligió el año 2023 como referencia, ya que el dataset principal de 2024 aún no cuenta con registros actualizados en varios países. Además, el dataset de 2023 solo incluye datos de enero a junio, lo que limita su precisión para analizar el año 2024

5.3. Utilidad del enriquecimiento del dataset

La incorporación de estas variables externas y calculadas responde a una necesidad analítica concreta: pasar de un enfoque puramente descriptivo a uno explicativo y comparativo.

Con estas nuevas columnas, será posible:

- Evaluar si los países con menor PIB per cápita exportan bienes más baratos o no necesariamente.
- Determinar si la inflación del país exportador se refleja en mayores valores FOB o CIF.
- Comparar la presión fiscal efectiva entre productos o grupos arancelarios.
- Identificar combinaciones de país-producto con características críticas (ej. alta inflación + alta carga tributaria).

Este enriquecimiento sienta las bases para consultas más profundas, visualizaciones interpretativas, y eventualmente, modelos predictivos o clasificaciones por riesgo tributario y logístico.

Cabe destacar que los archivos descargados de fuentes como el Banco Mundial fueron renombrados para facilitar su gestión y manejo ordenado dentro del flujo de trabajo.

Por ejemplo:

- API_FP.CPI.TOTL.ZG_DS2_en_csv_v2_122376.csv → PIB_per_capita.csv
- API_NY.GDP.PCAP.PP.CD_DS2_es_csv_v2_31580.csv →
inflacion_mercancias_pais.csv

Este proceso de renombramiento buscó una mayor claridad en el uso de los datos durante el desarrollo del análisis.

5.4 Implementación en SQL(Workbeanch)

Se replicó el proceso de integración de datos usando SQL estándar dentro de MySQL Workbench, asegurando así la trazabilidad del flujo de transformación en entornos relacionales tradicionales.

5.4.1. Normalización de atributos para integración de datos. Antes de realizar las uniones entre las distintas tablas, fue necesario normalizar ciertos campos clave que permitieran establecer relaciones correctas entre los datos del archivo principal de importaciones y las fuentes externas de PIB per cápita e inflación. Este paso fue fundamental, ya que los nombres de país en los archivos no coinciden exactamente, y trabajar con códigos estándar facilita los cruces.

PAIS_ORIGEN
KR-COREA, REPUBLICA DE
CO-COLOMBIA
US-ESTADOS UNIDOS
CN-CHINA
CN-CHINA
GB-REINO UNIDO DE GRAN BRETAÑA
IT-ITALIA
DE-ALEMANIA

Como se puede observar en la imagen los datos de la tabla principal en la columna PAIS_ORIGEN poseen el nombre del país con diferente estructura e idioma.

Country Name
Aruba
Africa Eastern and Southern
Afghanistan
Africa Western and Central
Angola
Albania
Andorra
Arab World

Esto no sería un problema si tan solo se pudiera solo tomar las primeras dos letras y luego eliminar todo el texto restante de la tabla PAIS_ORIGEN, sin embargo, tanto la tabla encargada del PIB como de la importación, tenían el código del país con tres letras. Volviéndolo algo difícil de comparar por sí solo.

Country Code
ABW
AFE
AFG
AFW
AGO
ALB
AND
ARB
ARE
ARG

Teniendo todo eso en cuenta, por eso se escogió un data set de países [3], porque con la ayuda de su columna alpha3, se podría cambiar el texto del PAIS_ORIGEN para convertirlo a código de 3 letras.

A continuación, se describen los pasos realizados, agrupados en fases lógicas de trabajo:

Paso 1. Preparación del entorno. Se creó una base de datos específica para el proyecto

```
CREATE DATABASE IF NOT EXISTS importacionesZepelin;
USE importacionesZepelin;
```

Posteriormente, se importaron manualmente las cuatro tablas base desde archivos .csv a través del asistente de importación de MySQL Workbench:

- senae_importacion_general_enero_junio_2024
- countries
- PIB_per_capita
- inflacion_mercancias_pais

La tabla principal se renombra a un nombre más manejable

```
ALTER TABLE senae_importacion_general_enero_junio_2024 RENAME TO importacionesData;
```

Paso 2. Normalización del campo PAIS_ORIGEN. Para poder unir los datos con otras tablas externas, fue necesario convertir el campo PAIS_ORIGEN de la tabla importaciones a un formato estandarizado:

- Agregar columna temporal para ISO2:

```
ALTER TABLE importacionesData ADD COLUMN ISO2 VARCHAR(5);
```

- Extraer el código ISO2 desde el texto libre de país:

```
UPDATE importaciones
SET ISO2 = SUBSTRING_INDEX(PAIS_ORIGEN, '-', 1);
```

- Reemplazar PAIS_ORIGEN con su correspondiente código ISO3:

```
UPDATE importacionesData i
JOIN countries c ON i.ISO2 = c.alpha2
SET i.PAIS_ORIGEN = c.alpha3;
```

Paso 3. Cálculo de variable CARGA_TRIBUTARIA_TOTAL. Se creó una nueva columna que resume la carga impositiva de cada ítem:

```
ALTER TABLE importacionesData ADD COLUMN CARGA_TRIBUTARIA_TOTAL DOUBLE;
```

Se usó IFNULL para prevenir errores por valores nulos durante la suma.

```
UPDATE importacionesData
SET CARGA_TRIBUTARIA_TOTAL =
  IFNULL(FODINFA, 0) +
  IFNULL(ICE_ADVALOREM, 0) +
  IFNULL(ICE_ESPECIFICO, 0) +
  IFNULL(IVA, 0) +
  IFNULL(ADVALOREM, 0);
```

5.5. Implementación en Apache Spark (Zeppelin)

La integración de las nuevas variables al DataFrame principal de importaciones se llevó a cabo utilizando Apache Spark dentro del entorno interactivo Zeppelin. A continuación, se describe cada etapa del proceso.

5.5.1. Carga de los archivos. Para iniciar con el proceso, se cargó los datasets revisados a lo largo de este informe, en este punto dado a que los archivos PIB_per_capita y inflación_mercancias_pais contiene metadatos y líneas extra antes del encabezado real, fue

necesario un preprocesamiento, por lo que se limpia eso manualmente y como hay varias columnas innecesarias solo se toma la del año 2023:

```
import org.apache.spark.sql.types._
import org.apache.spark.sql.functions._

// Definir esquema personalizado solo con las columnas que necesitas
val customSchema = StructType(Array(
  StructField("Country Name", StringType, true),
  StructField("Country Code", StringType, true),
  StructField("Indicator Name", StringType, true),
  StructField("Indicator Code", StringType, true),
  StructField("2023", DoubleType, true)
))

// Leer el archivo con ese esquema
val pibDF = spark.read
  .option("header", "true")
  .schema(customSchema)
  .option("delimiter", ",")
  .csv("/media/sf_PA/SQL/PIB_per_capita.csv")

// Reemplazar nulos y multiplicar por 1000
val pibDFlimpio = pibDF.withColumn("2023", coalesce(col("2023"), lit(0.0)) * 1000)

pibDFlimpio.printSchema()
pibDFlimpio.show()
```

La multiplicación por 1000 se realizó debido a que los valores originales están expresados en miles de dólares internacionales (USD constantes), por lo que este ajuste permite obtener una representación en dólares reales, facilitando su interpretación y comparación directa con otras variables económicas en el análisis.

De la misma forma se leen los datos de la tabla encargada de la inflación.

```
import org.apache.spark.sql.types._
import org.apache.spark.sql.functions._

// 1. Definir el esquema solo con las columnas que quieres (puedes incluir Column1 si te da problemas)
val schema2023 = StructType(Array(
  StructField("Country Name", StringType, true),
  StructField("Country Code", StringType, true),
  StructField("Indicator Name", StringType, true),
  StructField("Indicator Code", StringType, true),
  StructField("2023", DoubleType, true),
  StructField("Column1", StringType, true) // por si aparece, luego lo puedes eliminar
))

// 2. Leer solo las columnas definidas
val inflacionDF = spark.read
  .option("header", "true")
  .schema(schema2023)
  .option("delimiter", ",")
  .csv("/media/sf_PA/SQL/inflacion_mercancias_pais.csv")

// 3. Reemplazar nulls en la columna 2023
val inflacionLimpio = inflacionDF
  .withColumn("2023", coalesce(col("2023"), lit(0.0)))
  .drop("Column1") // 4. Eliminar Column1 si existía

// 5. Ver resultado
inflacionLimpio.printSchema()
inflacionLimpio.show()
```

En cuanto a la tabla principal, al haberla cargado ya en el SQL, simplemente se la invoca para utilizarla, guardándola en un dataframe.

```
val url="jdbc:mysql://localhost:3306/importacionesZepelin"
val table="(SELECT * from importacionesData ) AS subconsulta1"
val propiedades = new java.util.Properties()
propiedades.setProperty("user","root")
propiedades.setProperty("password","CarroNuevaVida@1")
propiedades.setProperty("driver","com.mysql.cj.jdbc.Driver")
val data = spark.read.jdbc(url,table,propiedades)
```

5.6 Análisis exploratorio de datos por medio de consultas.

5.6.1. Consulta 1: ¿Quiénes son nuestros principales socios comerciales y qué tan eficientes somos recaudando impuestos de ellos? Como punto de partida para el análisis comercial, se formuló la pregunta: *¿Quiénes son nuestros principales socios comerciales y qué tan eficiente es la recaudación tributaria asociada a ellos?* Para abordarla, se agruparon las

operaciones de importación por país de origen, calculando el valor CIF total, el total de tributos recaudados y el número de operaciones. Con estos datos, se construyó un indicador de eficiencia tributaria, expresado como el porcentaje de tributos respecto al valor CIF.

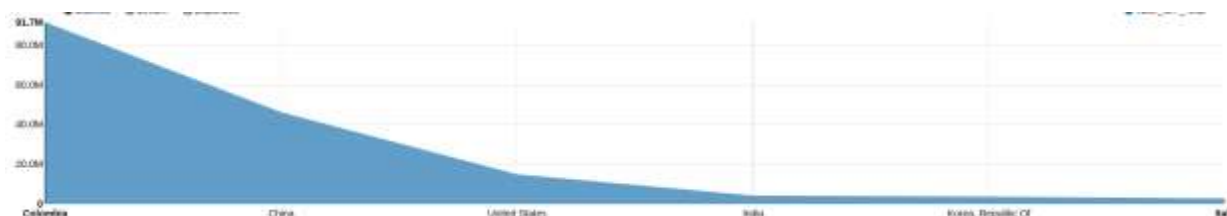
```
// Importa las funciones necesarias
import org.apache.spark.sql.functions.{sum, count, desc, col, round}

// Agrupamos por el nombre del país
val rankingPaíses = dfPuroo.groupBy("name")
  .agg(
    sum("CIF").as("Valor_CIF_Total"),
    sum("CARGA_TRIBUTARIA_TOTAL").as("Tributos_Totales"),
    count("*").as("Numero_Importaciones")
  )
  .withColumn("Porcentaje_Tributario", round(($"Tributos_Totales" / $"Valor_CIF_Total") * 100, 2))
  .orderBy(desc("Valor_CIF_Total"))
  .limit(6)

// Muestra el resultado para visualización
z.show(rankingPaíses)
```

El análisis se centró en los seis países con mayor volumen de importación hacia Ecuador. China se posicionó como el principal socio comercial, con un valor CIF que supera los 90 millones de USD, seguido por Estados Unidos y Colombia. Sin embargo, al observar el indicador de eficiencia tributaria, se evidencian diferencias marcadas: mientras algunos países presentan un alto volumen de comercio con tasas tributarias moderadas, otros tienen una relación más equilibrada entre el valor importado y la carga impositiva generada.

Esta variabilidad en la eficiencia tributaria puede estar influenciada por varios factores, entre ellos la estructura arancelaria, la existencia de acuerdos comerciales preferenciales o el tipo de bienes que predominan en el intercambio. Por ejemplo, países con fuerte presencia de bienes de capital o tecnología suelen beneficiarse de menores cargas impositivas, en comparación con aquellos cuyas exportaciones se componen de productos gravados con impuestos específicos.



Este ranking permite no solo identificar a los principales actores del comercio exterior ecuatoriano, sino también evaluar con mayor precisión qué tan efectiva es la política tributaria en función del país de origen de las mercancías.

5.6.2. Consulta 2: ¿Cómo varía el volumen de importaciones y la recaudación de impuestos mes a mes? Con el objetivo de comprender el comportamiento dinámico del comercio exterior ecuatoriano, se planteó la pregunta: *¿Cómo varía el volumen de importaciones y la recaudación de impuestos mes a mes?* Para ello, se realizó una agregación de los datos por mes calendario, utilizando la fecha de ingreso de las mercancías como referencia temporal.

El análisis comprendió el primer semestre del año 2024. A partir de los resultados obtenidos, se identificaron fluctuaciones notables tanto en el valor CIF total como en el total recaudado por concepto de tributos. En enero, el comercio exterior alcanzó su punto más alto, con importaciones que superaron los 64 millones de USD, lo que también se tradujo en una recaudación tributaria elevada. A partir de febrero, se observa una tendencia decreciente, con caídas progresivas en los valores importados hasta el mes de abril. Este descenso coincide con una contracción proporcional en la recaudación fiscal.

```
import org.apache.spark.sql.functions.{sum, window, date_format}

val evolucionMensual = dfPuroo.groupBy(date_format($"FEC_INGRESO", "yyyy-MM").as("Mes"))
  .agg(
    sum("CIF").as("Valor_CIF_Total"),
    sum("CARGA_TRIBUTARIA_TOTAL").as("Recaudacion_Total")
  )
  .orderBy("Mes")

z.show(evolucionMensual)
```



Durante mayo y junio se evidencian señales de recuperación. El volumen de importaciones muestra un repunte significativo, especialmente en junio, donde el crecimiento es lo suficientemente marcado como para revertir parcialmente la tendencia negativa previa. Sin embargo, estos picos no alcanzan los niveles de enero, lo cual podría estar vinculado a factores estacionales, restricciones logísticas o ajustes de política comercial.



El gráfico correspondiente refleja visualmente esta evolución. Cada área representa un mes distinto, permitiendo identificar de forma clara los ciclos de alza y baja. Esta visualización también permite anticipar posibles correlaciones entre el flujo de comercio internacional y la política fiscal aplicada a las importaciones.

Este tipo de análisis resulta crucial para las autoridades tributarias y económicas, ya que permite una planificación más informada de las metas de recaudación y la formulación de políticas comerciales ajustadas al comportamiento real del mercado.

5.6.3. Consulta 3: ¿Qué distritos aduaneros concentran el mayor volumen de importaciones y cómo varían en eficiencia logística?: Para analizar la distribución geográfica de las importaciones en Ecuador y comprender los patrones logísticos que sustentan el comercio exterior, se planteó la pregunta: ¿Qué distritos aduaneros concentran el mayor volumen de importaciones y cómo varía su eficiencia operativa? Esta consulta permite identificar no solo los puntos críticos de entrada de mercancías, sino también evaluar el desempeño operativo según el volumen y la cantidad de operaciones procesadas.

```
import org.apache.spark.sql.functions.{sum, count, avg, desc}

val analisisDistritos = dfPuroo.groupBy("DISTRITO")
  .agg(
    sum("CIF").as("Valor_Total_Importado"),
    count("*").as("Numero_Operaciones"),
    avg("PESO_NETO").as("Peso_Neto_Promedio_Kg")
  )
  .orderBy(desc("Valor_Total_Importado"))
z.show(analisisDistritos)
```

El análisis se realizó agrupando las operaciones de importación por distrito aduanero, considerando tres variables clave: valor CIF total importado, número de operaciones y peso neto promedio por operación (en kilogramos). Estos indicadores permiten establecer una relación entre el volumen económico, la frecuencia de actividad y la capacidad logística de cada distrito.



Principales hallazgos

Los resultados muestran una alta concentración del comercio exterior en unos pocos distritos:

Guayaquil - Marítimo (028) se posiciona como el principal centro logístico del país, con un valor CIF que supera los 79 millones de USD, gestionando más de 2.800 operaciones. Su peso neto promedio, cercano a los 8.793 kg por operación, sugiere un perfil de carga consolidada o a granel, típica de grandes volúmenes.

Le sigue Tulcán (073), con un valor de más de 3,7 millones de USD y un peso promedio de 4.897 kg, lo que refleja una relevancia significativa del comercio terrestre en frontera norte, probablemente de productos perecibles o de consumo masivo.

Guayaquil - Aéreo (019) destaca por su alto valor importado con bajo peso promedio (50 kg por operación), lo que indica la presencia de mercancías de alto valor unitario como tecnología, farmacéuticos o repuestos especializados.

Otros distritos como Manta (037) y Cuenca (091) también muestran un volumen relevante, aunque con características logísticas distintas. Manta, por ejemplo, evidencia un peso promedio extremadamente alto (81.747 kg), lo cual sugiere operaciones concentradas en carga masiva, posiblemente vinculadas a sectores industriales o agroexportadores.

5.6.4. Consulta 4: ¿Qué volumen de comercio se realiza bajo convenios internacionales y cómo afecta esto a la carga tributaria promedio? En el marco del análisis del comercio exterior, se planteó la pregunta: ¿Qué volumen de comercio se realiza bajo convenios internacionales y cómo afecta esto a la carga tributaria promedio? Esta consulta busca evaluar el impacto real de los acuerdos internacionales en los niveles de recaudación tributaria asociados a las importaciones.

Para ello, se clasificaron las operaciones de importación en dos grupos: aquellas “con convenio”, que aplican beneficios derivados de tratados o acuerdos internacionales, y aquellas “sin convenio”, que no se acogen a tales beneficios. Posteriormente, se calcularon tres indicadores clave para cada grupo:

- Valor CIF Total Importado
- Tributo Promedio por Operación
- Número Total de Operaciones

```
import org.apache.spark.sql.functions.{when, lit, sum, avg, count, desc}

// Creamos una columna para diferenciar si se usó un convenio o no
val dfConConvenio = dfPuroo.withColumn("Aplica_Convenio",
  when(col("CONVENIO_INTERNACIONAL") != "NO APLICA", "Con Convenio")
  .otherwise("Sin Convenio")
)

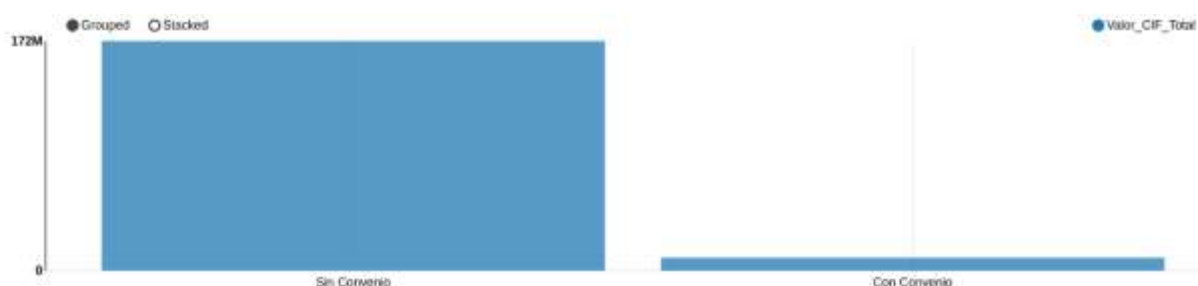
// Agrupamos para comparar las importaciones con y sin convenio
val impactoConvenios = dfConConvenio.groupBy("Aplica_Convenio")
  .agg(
    sum("CIF").as("Valor_CIF_Total"),
    avg("CARGA_TRIBUTARIA_TOTAL").as("Tributo_Promedio_por_Operacion"),
    count("*").as("Numero_de_Operaciones")
  )
  .orderBy(desc("Valor_CIF_Total"))

z.show(impactoConvenios)
```

Las operaciones sin convenio concentran la gran mayoría del comercio exterior en términos de valor, con un valor CIF total superior a 171 millones de USD, representando más del 94% del total importado. Estas operaciones generan, en promedio, una carga tributaria de 318 USD por operación.

En contraste, las operaciones con convenio internacional representan apenas 9,7 millones de USD en valor CIF, pero sorprendentemente, tienen una carga tributaria promedio más alta, de 450 USD por operación.

Esta diferencia sugiere que, aunque los convenios internacionales suelen estar diseñados para reducir o eliminar barreras arancelarias, su efecto no siempre se traduce en una menor carga tributaria promedio, posiblemente debido al tipo de mercancías importadas bajo este régimen, su valor unitario o la aplicación limitada de exenciones específicas.



El gráfico de barras incluido permite visualizar con claridad esta desigualdad: el volumen importado sin convenio es abrumadoramente mayor que el registrado con convenio. Esta desproporción plantea interrogantes sobre el grado de aprovechamiento de los acuerdos internacionales vigentes y sobre la necesidad de estrategias para fomentar su uso más efectivo entre los operadores de comercio exterior.

Este análisis aporta elementos clave para la evaluación de la política comercial ecuatoriana. Comprender el impacto real de los convenios vigentes permite a las autoridades ajustar su aplicación, promover capacitaciones específicas para importadores y optimizar los mecanismos de facilitación comercial, todo con el fin de aumentar la eficiencia tributaria sin comprometer la competitividad.

5.6.5. Consulta 5: ¿En qué productos gastamos más dinero y cuál es la tasa impositiva real que se les aplica? Con el fin de comprender en qué tipos de productos se concentra el mayor gasto en importaciones y qué nivel de carga fiscal efectiva representan para el país, se formuló la pregunta: ¿en qué productos gastamos más dinero y cuál es la tasa

impositiva real que se les aplica? Este análisis permite identificar no solo los bienes más relevantes en términos de valor económico, sino también evaluar qué tan significativa es su contribución a la recaudación tributaria, lo que resulta clave para la toma de decisiones en materia de política comercial y fiscal.

```
import org.apache.spark.sql.functions.{sum, desc, round}

val productosRelevantes = dfPuroo.groupBy("SUBPARTIDA", "DESCRIPCION_ARANCELARIA")
  .agg(
    sum("CIF").as("Valor_CIF_Total"),
    sum("CARGA_TRIBUTARIA_TOTAL").as("Tributos_Totales")
  )
  .withColumn("Tasa_Impositiva_Promedio", round(($"Tributos_Totales" / $"Valor_CIF_Total") * 100, 2))
  .orderBy(desc("Valor_CIF_Total"))
  .limit(20)

z.show(productosRelevantes)
```

La metodología consistió en agrupar los datos por subpartida arancelaria y descripción del producto, calculando el valor CIF total importado, los tributos totales recaudados y una tasa impositiva promedio resultante del cociente entre ambos valores, expresado en porcentaje. Con esta estructura, se seleccionaron los veinte productos más relevantes por valor económico para su análisis.

Los resultados muestran que la energía eléctrica, clasificada bajo la subpartida 2716000000, encabeza el listado con un valor CIF superior a 87 millones de dólares, lo cual refleja su importancia estratégica dentro del sistema energético nacional, sobre todo en contextos de déficit o estacionalidad en la oferta interna. Le sigue la importación de motocicletas en modalidad CKD (desensambladas) para ensamblaje local, con un valor cercano a los 28 millones de dólares, lo que evidencia la importancia del régimen industrial automotriz que opera con incentivos y tarifas arancelarias especiales. Otros productos destacados incluyen algodón de fibra larga, maquinaria pesada y perfiles metálicos, todos vinculados directa o indirectamente a sectores productivos o industriales clave para la economía ecuatoriana.

En cuanto a la carga tributaria efectiva, se identifican diferencias significativas entre productos. Algunos, como los perfiles metálicos bajo la subpartida 7604292000, alcanzan una tasa impositiva promedio superior al 23%, lo cual representa una carga considerable para el importador y una fuente relevante de ingresos fiscales. Otros productos, especialmente maquinaria y materias primas, presentan tasas impositivas moderadas, en rangos que van del 14% al 19%, lo cual podría estar asociado a beneficios arancelarios parciales, políticas de fomento productivo o consideraciones técnicas de clasificación arancelaria.

SUBPARTIDA	DESCRIPCION ARANCELARIA	Valor CIF Total	Tributos Totales	Tasa Impositiva Promedio
5201000000	variable para modelos y versiones autorizadas para el ensamble por el Ministerio rector de la Política Industrial, MCE >= 18%	9290329.778999997	1395226.5699999996	14.59
8429590000	-- Las demás	7463692.760999999	1095506.06	14.68
3923290000	--- Las demás	3405570.0010000025	505203.4800000014	14.83

La visualización correspondiente permite observar claramente la relación entre el valor económico importado por producto y su respectiva carga tributaria, destacando aquellos bienes que, por su volumen o tasa, generan un mayor impacto en la recaudación. Esta información resulta fundamental para que las autoridades puedan orientar políticas que favorezcan la competitividad sin comprometer los ingresos fiscales, ajustar aranceles en función de objetivos sectoriales, o fortalecer los mecanismos de control aduanero en productos de alto valor y alta carga impositiva.

5. Conclusiones

El análisis realizado a las importaciones del Ecuador bajo el régimen general durante el primer semestre de 2024 permitió obtener una visión integral y fundamentada del comportamiento del comercio exterior del país. Gracias a la incorporación de variables externas,

como el PIB per cápita y la inflación del país exportador, y al enriquecimiento del dataset con métricas calculadas como la carga tributaria total, fue posible ir más allá de una simple descripción de datos y alcanzar interpretaciones más profundas y contextualizadas.

Los resultados evidencian que el comercio exterior ecuatoriano está altamente concentrado en pocos socios estratégicos, siendo China, Estados Unidos y Colombia los principales proveedores. Asimismo, se identificaron diferencias notables en la eficiencia tributaria según el país de origen, lo cual está condicionado por factores como acuerdos internacionales, tipos de productos y estructuras arancelarias.

A nivel logístico, el distrito Guayaquil - Marítimo destaca como el principal nodo de ingreso de mercancías al país, tanto en valor económico como en número de operaciones, mientras que otros distritos como Tulcán o Guayaquil - Aéreo tienen roles más específicos y complementarios dentro de la cadena logística nacional. Además, se pudo constatar que la mayoría de importaciones no se acogen a convenios internacionales, lo que abre una oportunidad de mejora en la difusión y aprovechamiento de estos instrumentos comerciales.

Por otra parte, el análisis de productos más relevantes en términos económicos reveló que rubros como energía eléctrica, motocicletas CKD y maquinaria pesada no solo tienen un alto peso financiero, sino que también enfrentan tasas impositivas variadas, reflejando políticas diferenciadas según el sector y el tipo de bien. Esto permite concluir que el sistema tributario aplicado a las importaciones en Ecuador no es uniforme, sino que responde a objetivos específicos de política industrial, recaudación o protección económica.

Finalmente, este estudio demuestra que el uso de herramientas como Apache Spark, SQL y la integración de fuentes externas mejora significativamente la calidad del análisis de datos,

aportando una mirada más estratégica al entendimiento del comercio internacional. La metodología aplicada puede servir como base para futuras investigaciones que busquen fortalecer la toma de decisiones en áreas como política fiscal, logística aduanera o planificación económica nacional.

6. Bibliografía

[1] World Bank. (2017). *GDP per capita, PPP (current international \$)*.

<https://datos.bancomundial.org/indicador/NY.GDP.PCAP.PP.CD>

[2] World Bank. (2017). *Inflation, consumer prices (annual %)*.

<https://data.worldbank.org/indicator/FP.CPI.TOTL.ZG>

[3] OpenBookPrices. (2015). *Country data (countries.csv)* [Dataset]. GitHub.

<https://github.com/OpenBookPrices/country-data/blob/master/data/countries.csv>