

# EDA: Importaciones en Colombia pre, durante y post pandemia de Covid-19

2024-12-7



## REALIZACIÓN A:

- Estadística I
- Universidad Industrial de Santander
- Profesor: Hoover R. Chacón
- Ingeniería de Sistemas
- **Participantes:**
  - Angela Lozano
  - Jose Quintero
  - Nicolas Quinter

## 1. Introducción

- 1.1. Planteamiento

Las importaciones de un país pueden ser el reflejo del estado del mismo, tanto de su economía como de su situación interna, en nuestra exploración de datos observamos que la mayoría son cualitativos y muy pocos son cuantitativos. La base de datos del IDEAM nos proporciona archivos de datos desde 2012 hasta 2024, donde, nos pareció curioso analizar la situación del país en los años de pandemia, 2019 - 2022 podrían darnos una figura de como se manejó el país pre, durante y pos-pandemia; Queremos ver y descubrir ¿cómo este evento global afectó a nuestro país desde el comportamiento de las importaciones?, ¿cómo la afectó a los departamentos y ciudades más relevantes del tema?

Donde nuestras diversas hipótesis son: ¿La pandemia en realidad afectó las importaciones? ¿que tan significativa fue?

- 1.2. Abordamiento

Abordamos el problema desde el análisis y comprensión de los datos que estábamos manejando, entender qué significaban sus siglas, que tipos de datos teníamos para trabajar y posteriormente como podíamos presentarlos de modo que reflejarán la situación del contexto.

La metodología se basó en organizar que tipo de gráficas queríamos llegar a obtener, los cálculos que podríamos lograr con los datos que teníamos y su relevancia para el contexto. Así sucesivamente fuimos revisando la teoría vista en clase porque nuestra misión es aplicar lo visto en el curso, por lo que a medida de que vamos avanzando esperamos obtener más ideas de datos relevantes a analizar.

Por otro lado la creación de estos gráficos y cálculos, se basó en entrar primero al csv original desde una vista de excel e ir identificando que tan limpios estaban los datos y ver que datos debíamos filtrar, cambiar o transformar (de caracteres a numéricos) si era necesario. El proceso se repetía en cada objetivo que teníamos.

- 1.3. Enfoque/técnica analítica

Actualmente nuestra analítica se basa en tener gráficas que nos ayuden a mostrar cómo se comportaron las importaciones durante los años de pandemia. Iniciamos con los datos más básicos que nos brinda la base y seguido de esto esperamos sacar probabilidades y conjeturas de cuales fueron los departamentos donde hubo una mayor reducción de las actividades.

- 1.4. Beneficios al consumidor

Esperamos que nuestro análisis ayude al consumidor a ver como las mercancías se comportan en una época de crisis, en qué ciudades receptoras se incrementaron los aranceles y por ende se incrementaron los precios y cuales fueron esos puntos de quiebre que hicieron que los precios se dispararon para poder tener un mejor manejo sobre importaciones en futuras crisis.

## 2. Preparación de los datos

- 2.1. Obtección original de los datos

La base de datos utilizada durante el estudio fue obtenida a través de los recursos virtuales del DANE, la cual proporciona datos publicos en su servicio de (*ANDA*), donde nos encontramos con un catálogo de nichos de los cuales se proporcionaban archivos de extensión “.CSV” con datos de interés; entre las categorías nos inclinamos por la colección de economía-microdatos, por el nicho de economía internacional, donde entre la lista de opciones de censos nos cautivó las *Estadísticas de Importaciones - IMPO - 2012 A 2024* (<https://microdatos.dane.gov.co/index.php/catalog/473/get-microdata>).

- **2.2.** Explicación de datos originales

Los datos fueron recolectados por el D.A.N.E para garantizar la veracidad y transparencia de la información básica y estratégica para la investigación y los usuarios relacionados con las compras al exterior. Los microdatos registran las importaciones de mercancías legales desde otros países o desde una región portuaria de Colombia al resto del territorio aduanero nacional. Los datos se basan en las en la fecha de presentación.

Los datos se almacenan en una base de datos proporcionada por ANDA, que cuenta con 44 variables, 20 discretas y 24 continuas. Los datos carecen de clasificación *específica* de las mercancías, lo más cercano es la empresa que recibe la mercancía, para la cantidad de variables se proporciona un diccionario de dato donde los nombres de las variables son acrónimos de su significado sencillos de interpretar.

Algo que también nos pareció curioso es que hay dos variables llamadas País de Procedencia (PAISPRO) y País de Compra (PAISCOM) que, en algunos casos, son iguales pero en la mayoría de casos es diferente lo que hace pensar que muchas mercancías son hechas en un país (PAISPRO) y son vendidas desde otros país (PAISCOM), nos pareció curioso.

Variable	Descripción	Variable	Descripción
FECH	Fecha de proceso	CUIDAIMP	Ciudad del importador
ADUA	Código de la aduana	CUIDAEXP	Ciudad del exportador
PAISGEN	País origen	ACTECON	Actividad económica
PAISPRO	País de procedencia	CODADAD	Código administración de aduana
PAISCOM	País de compra	VADUA	Valor aduana
DEPTODES	Departamento destino	VRAJUS	Valor ajuste
VIATRANS	Código vía de transporte	BASEIVA	Base IVA
BANDERA	Bandera	OTROSP	Porcentaje otros
REGIMEN	Código de régimen	OTROSBASE	Base otros
ACUERDO	Código del acuerdo	TOTALIVAYO	Total IVA y otros gastos
PBK	Peso bruto en kilos	SEGUROS	Seguros
PNK	Peso neto en kilos	OTROSG	Otros gastos
CANU	Cantidad de unidades	LUIN	Lugar de ingreso
CODA	Código de unidad	CODLUIN	Código lugar de ingreso
NABAN	Posición arancelaria	DEPIM	Departamento del importador
VAFODO	Valor FOB dólares de la mercancía	COPAEX	Código país del exportador
FLETE	Fletes	TIPOIM	Tipo de importación
VACID	Valor CIF dólares de la mercancía	PORARA	Porcentaje de arancel
VACIP	Valor CIF pesos de la mercancía	NIT	Número de identificación Tributaria
IMP1	Impuesto a las ventas	DIGV	Dígito de Verificación
OTDER	Otros Derechos	RZIMPO	Razón social del importador
CLASE	Clase de importador	DEREL	Derechos Arancelarios

- **2.5.** Información resumida:

Nuestro análisis gráfico abordó todas las variables generales del análisis, sin embargo, un dato que poco estudiamos fue el valor de las Aduana ya que este fluctua por el departamento y su administración política, por lo que de este datos si podemos obtener datos básicos como:

En la obtención de los datos, se puede ver que promedio de los datos de 2019 y 2020 no hubo tanta variación, en promedio, sin embargo, se puede observar que existieron aduanas donde su cobro máximo alcanzó el valor

más alto de los 4 años, reflejando que este costo se llegó a incrementar abismalmente por las medidas de bioseguridad implementadas.

Por este lado vemos que el costo de la mercancía en promedio se mantuvo relativamente estable con una disminución notable en 2020 (inicio de pandemia) y que tuvo un posterior pico en 2022 donde la actividad se incrementó, tras investigaciones nuestra conjetura es que durante pandemia mucha gente empezó negocios online y las compras en línea (principalmente de mercados internacionales) se incrementó al abrir esta ventana digital lo que refleja el impacto de la pandemia que podría llegar a verse como positivo. También podemos ver que los valores más bajos de las mediciones están en 2020 reflejando la recesión económica por el miedo y la incertidumbre de este hecho. Debemos mencionar que la desviación estándar de mayor magnitud es de 2022, lo que indica una gran diferencia de precios con respecto a la mediana, esto podría indicar que estas nuevas importaciones derivadas de las compras online eran más costosas que aquellas que llegaban normalmente.

### 3. Metodología

El primer paso de la limpieza fue ir descargando los archivos “.zip” de año a través del sitio elegido para obtención de datos (DANE), para extraer los “.csv” de cada mes del año correspondiente gracias a que los archivos están muy bien categorizados, nos ahorro algo de tiempo. Seguimos importando y extrayendo los datos necesarios a través de R, donde debimos hacer lectura de “.csv” aclarando que el primer año está separado por “;” y los demás años por “,”, este proceso lo hicimos a través de la definición de una ruta base, lo que agiliza el código, que es la carpeta principal donde se guardaron todas las importaciones de los 4 años, en nuestra carpeta de drive vera una carpeta por cada año, lo ideal es que estas cuatro carpetas sean guardadas en una carpeta específica cuya ruta será la ruta base. **Tener en cuenta que RStudio no le los caracteres de la ruta base como “/” deben cambiarse por “\”.**

Para ellos podemos utilizar el siguiente link correspondiente a la ruta de la carpeta central de datos de nuestro proyecto, esta será la carpeta que usted tendrá que descargar:

Y desde su computador o dispositivo tendrá que captar la ruta de esta carpeta donde haya destinado guardarla y colocarla como guía en la ruta base haciendo el cambio respectivo en la ruta base y, no olvidarlo, **cambiar los caracteres “/” por “\” para una lectura correcta de los archivos.**

Esto permite la extracción hacia R de todos los datos de los “.CSV” al entorno de desarrollo. Una vez hicimos este proceso, continuamos limpiando los data frames originales con las columnas de los datos de interés (PBK, FLETES, VACID, VADUA, “VAFODO”, entre otras posibles), obteniendo los “.CSV” limpios con las columnas necesarias a analizar.

Se aplicó una función de limpieza de columnas donde se convertían los datos en tipo “carácter” o “numérico” tal fuera necesario para el análisis, además de eliminar espacios en blanco o datos vacíos (NA), también se convirtieron los caracteres de tipo “,” a tipo “.” para su procesamiento con las funciones de R.

El paso a continuación fue crear un data frame de cada año (tablas\_año) donde guardamos las columnas de cada mes en una columna mayor de su mes correspondiente, esto nos permitió tener un data frame general para cada año y darnos acceso rápido a cada uno de los meses correspondientes.

Extracción de columnas de interés específico con el fin de aligerar el peso de los data frames anteriores y a su vez filtrar por datos relevantes para este proyecto

## 4. Resultados

- Aduanas

Podemos ver que el valor de las aduanas que es regulado por la DIAN en nuestro país y que se ve afectado por varios factores como acuerdos internacionales y los que más queremos resaltar que fueron los directamente

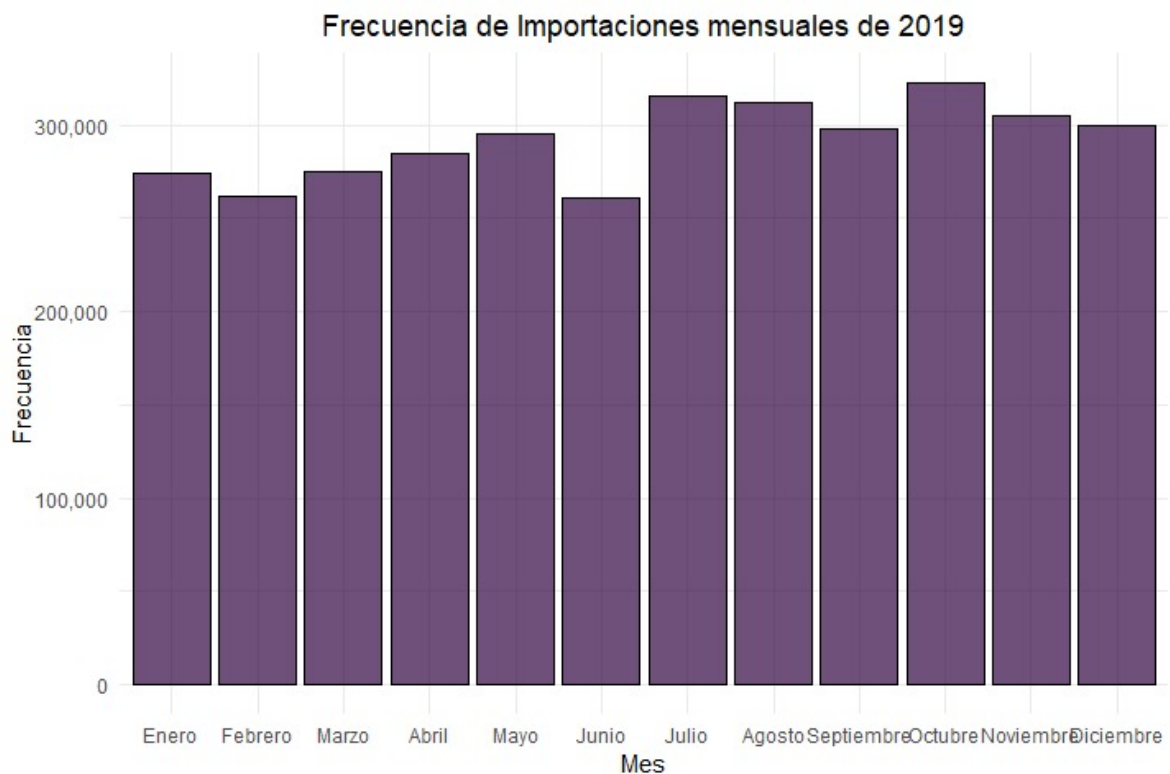
afectados por la pandemia: el precio de compra de las mercancías y gastos relacionados con la carga, descarga y manipulación de las mercancías en el puerto o lugar de importación.

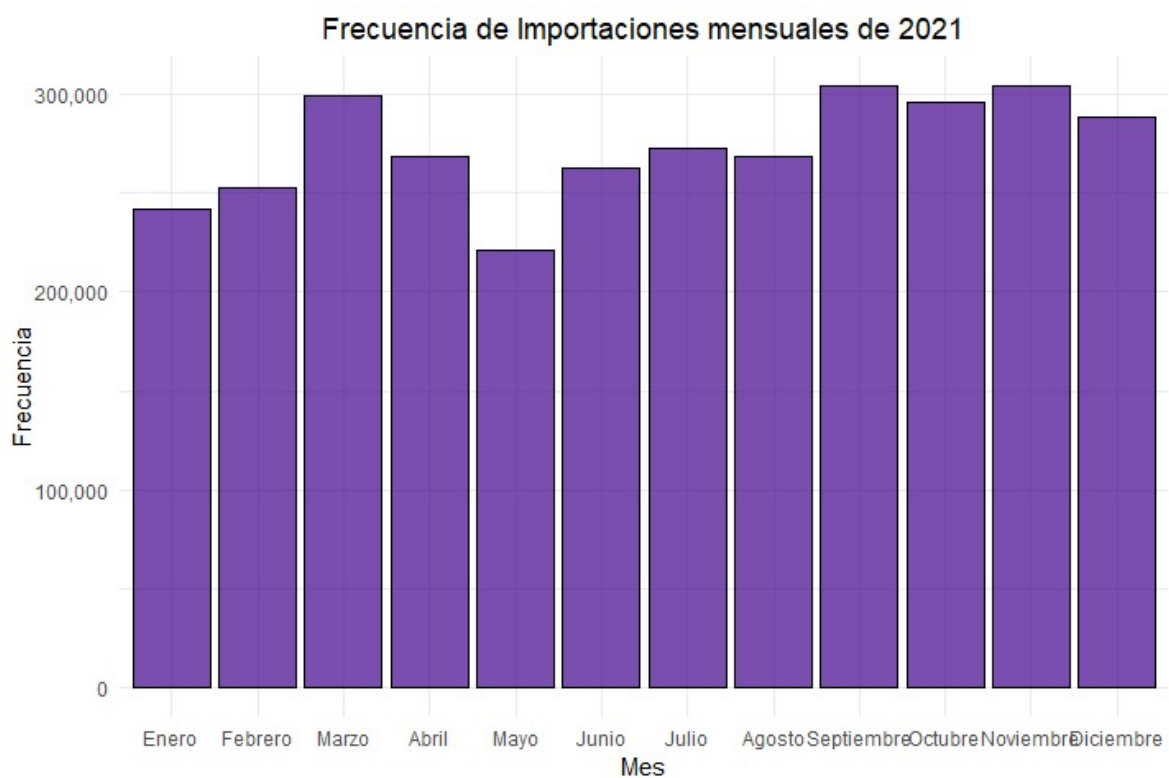
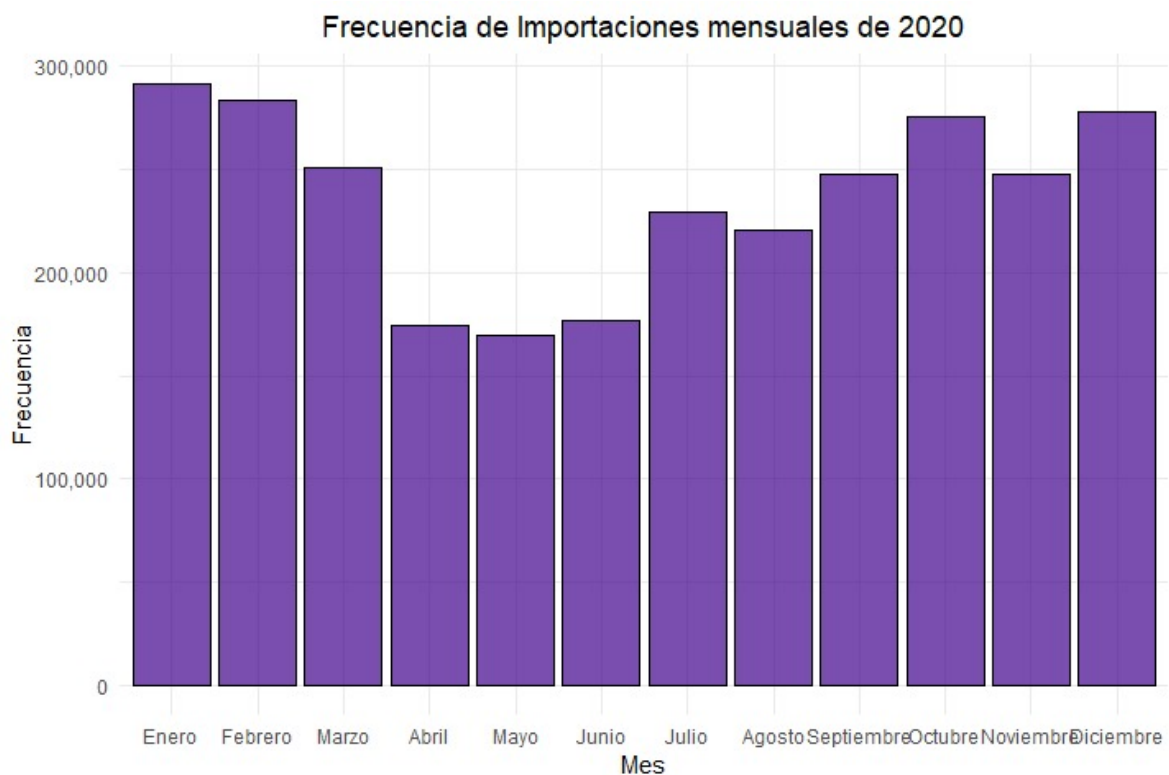
Podemos ver como podemos encontrar mercancía con valor mínimo de 0, lo que puede significar varias cosas, las más comunes exención arancelaria (no paga impuestos) o son donaciones o ayudas, que en el caso de Colombia tuvimos bastantes especialmente entre 2020 y 2021 cuando la pandemia fue más abrasiva, estas donaciones llegan a desviar el promedio o nos muestran como en los años 2019 y 2020 se mantuvieron muy similares, lo que podemos pensar que el promedio del valor de las aduanas se vieron bastante afectadas por estas donaciones, incluso el 2021 hubo una subida leve pero finalmente fue en 2022 donde podemos ver el salto de casi 3000 dolares de diferencia en el promedio que fue el año que dejamos de recibir donaciones magnificadas por la pandemia.

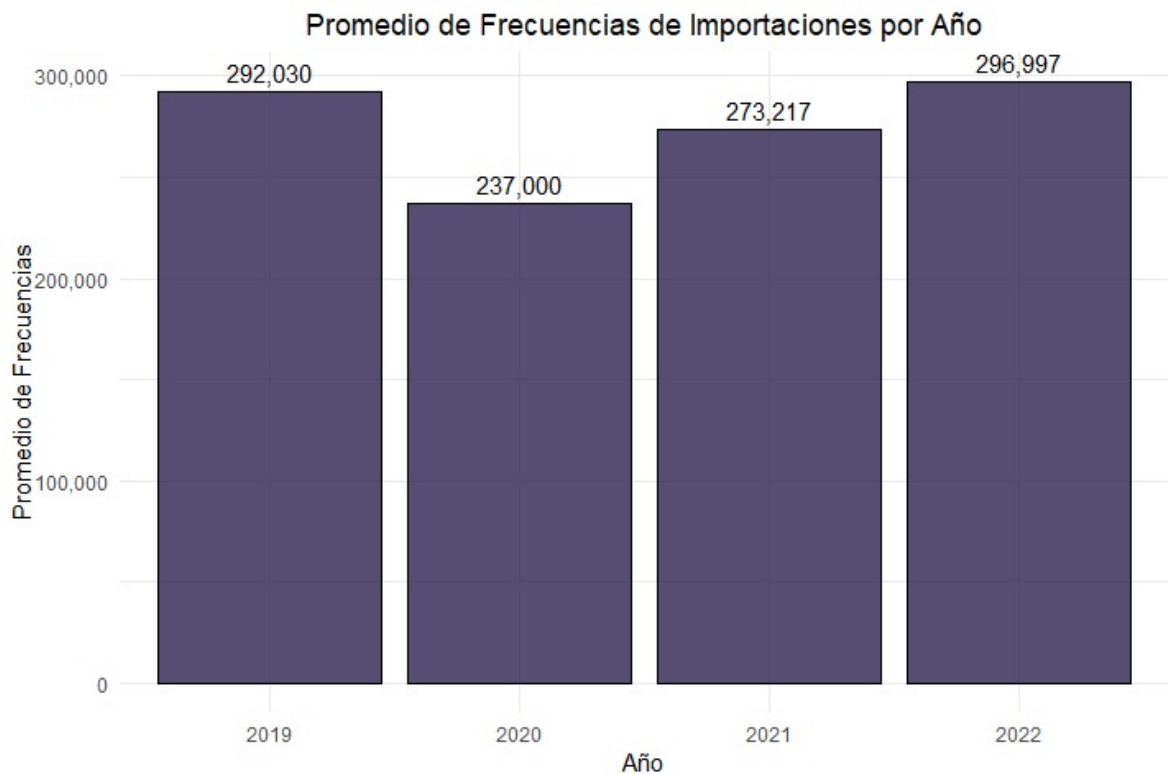
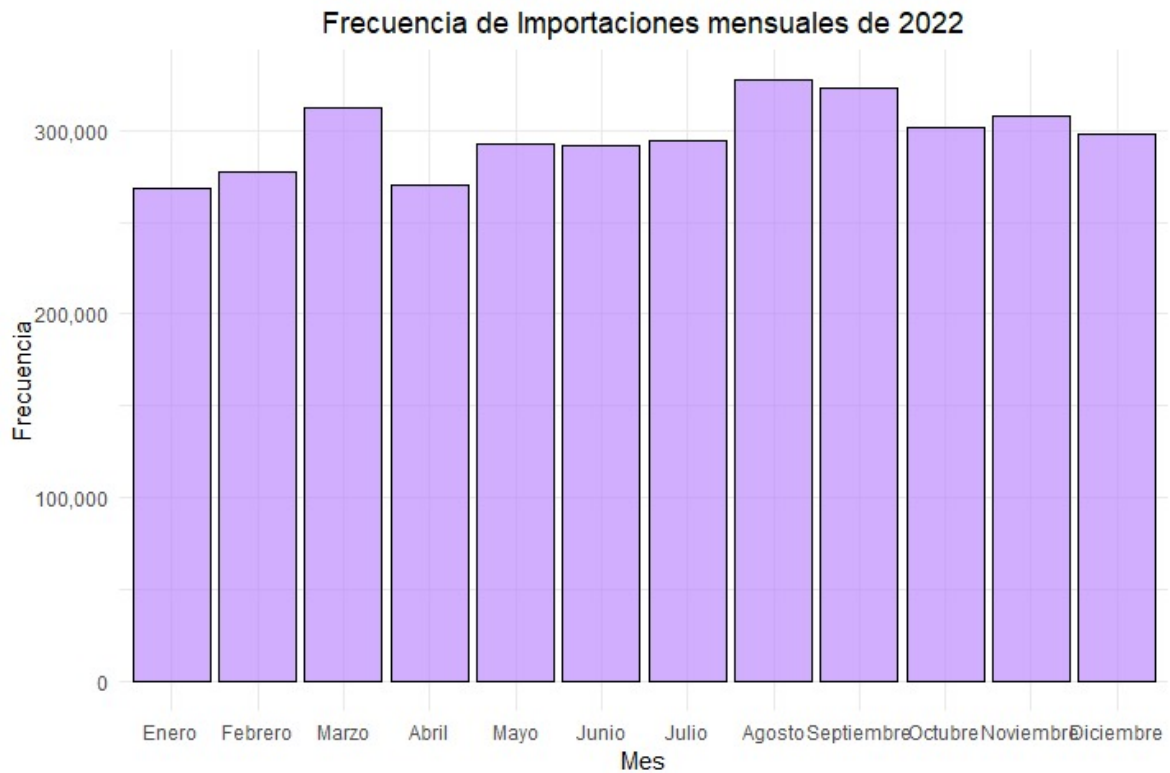
	VADUA	Año	Min	Mean	Max
1		2019	0	14967.81	80546771
2		2020	0	14949.21	100082450
3		2021	0	18529.68	60583780
4		2022	0	21686.36	74717923

- Frecuencias

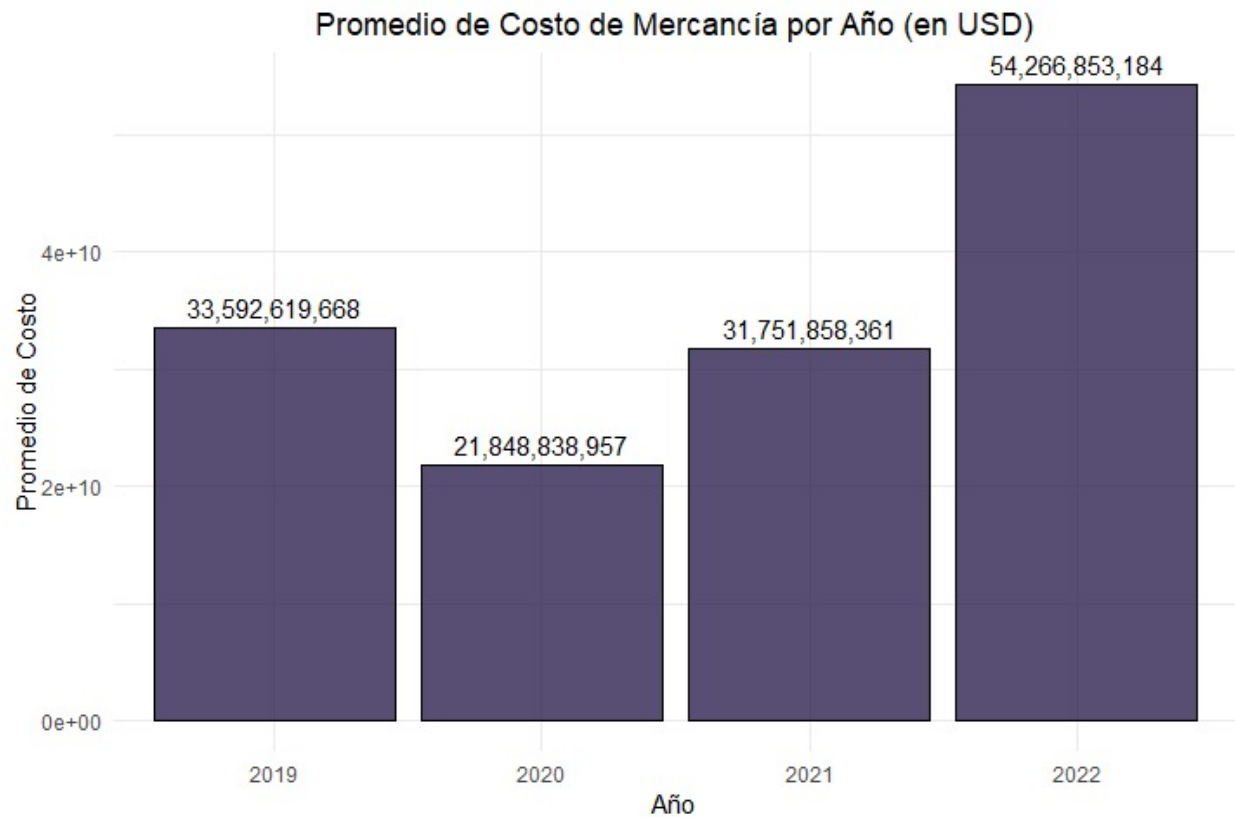
En las siguientes graficas se puede observar el numero de importaciones que se realizaron todos los meses de los años 2019 al 2022. En el grafico del año 2020 se puede observar como la cantidad de importaciones disminuyó drásticamente los meses de Abril, Mayo y Junio, que es cercano a cuando se declaró la pandemia en el país (12 de marzo, 2020), de igual modo se puede observar como eventualmente el número de importaciones empieza a aumentar hasta volver a la normalidad como se puede observar en la última grafica con el promedio de importaciones mensuales de cada año, en la que en el año 2020 cayó casi 60000 puntos y al año siguiente aumento hasta que en el 2022 se volvió (y se superó) la cifra del año 2019.







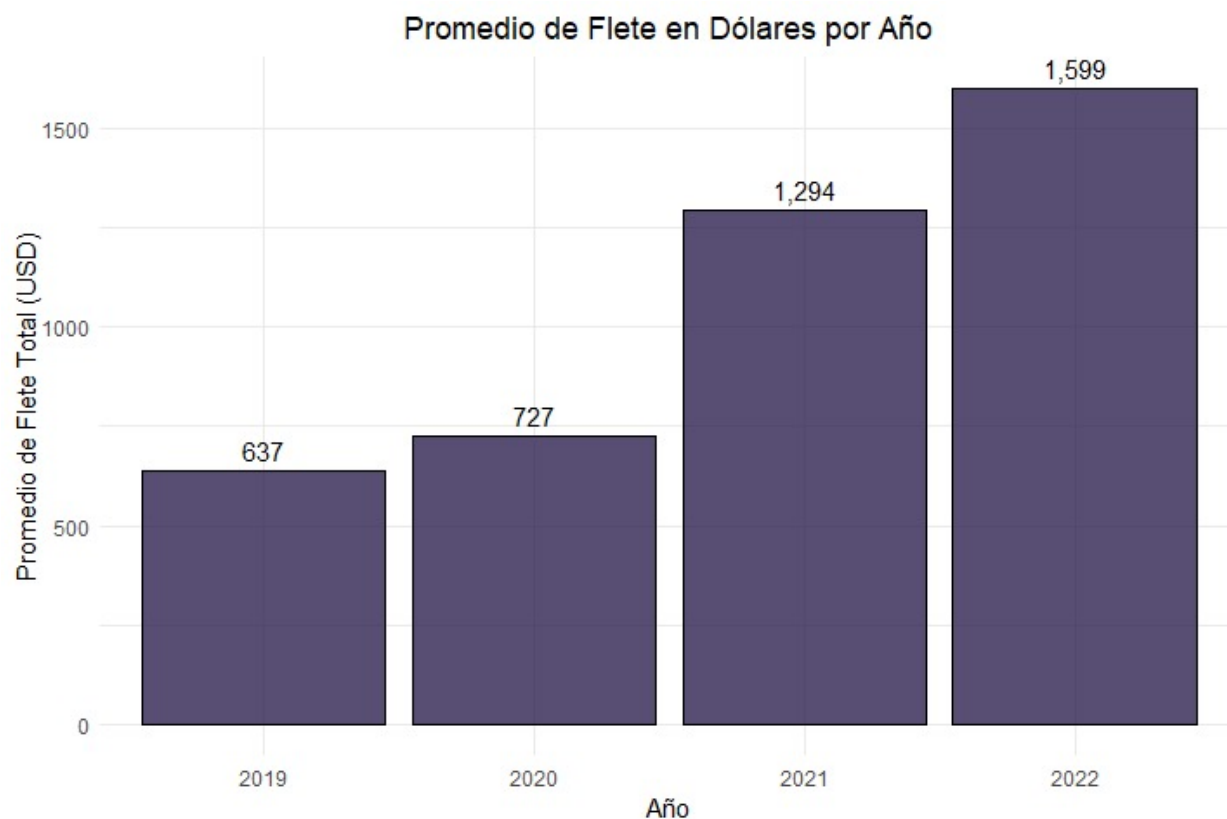
- Valor importado En la grafica se puede observar lo que se mencionaba anteriormente, como la actividad de importación desde el exterior vivió un momento de recesión por el momento histórico que se vivió (la pandemia por COVID-19) y como esta actividad se recupera poco a poco. Esto destaca la resiliencia del comercio internacional después de haber pasado por un evento disruptivo.



- Flete

En la grafica se puede observar como el valor promedio del flete de por año cambia y va en aumenta, pero como se puede observar no es un cambia brusco, por que este cambio podemos supones, que no estan influenciados unica y principalmente por la pandemia, sino que más bien, se debe a una mezcla de factores como lo fueron la pandemia la política nacional e internacional, así como también la producción del pais, exportaciones, etc... (PIB).





- Tipo de transporte

En las graficas se puede observar el tipo de transporte que más se usa para realizar importaciones en las que se puede observar que el tipo de transporte preferido es por via maritima, terrestre y aerea. Así también, se puede observar que el tipo de transporte terrestre disminuyo desde el mes de abril del año 2020 debido a un control más estricto por parte de las autoridades en las fronteras; además, se puede observar como el uso de esta forma de transporte se fue recuperando poco a poco.

1 MARITIMO

2 FERREO

3 TERRESTRE

4 AEREO

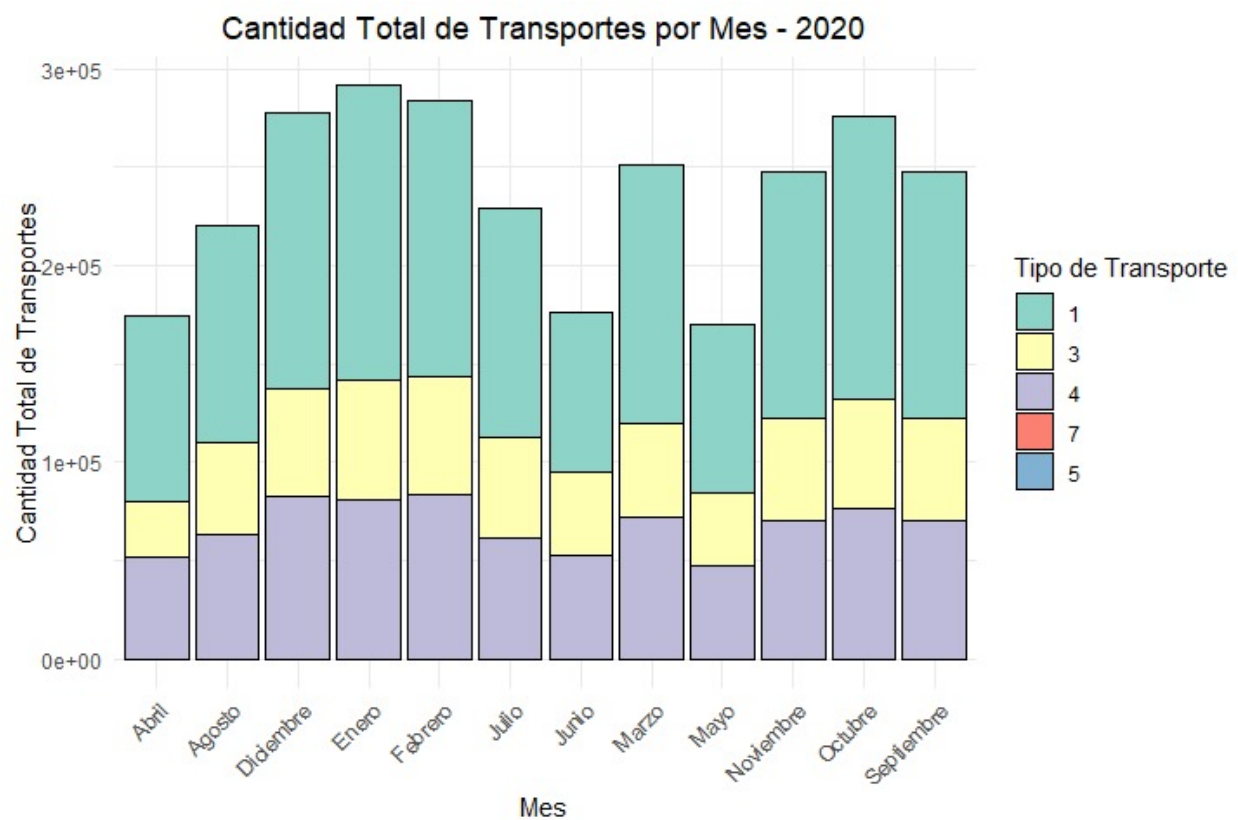
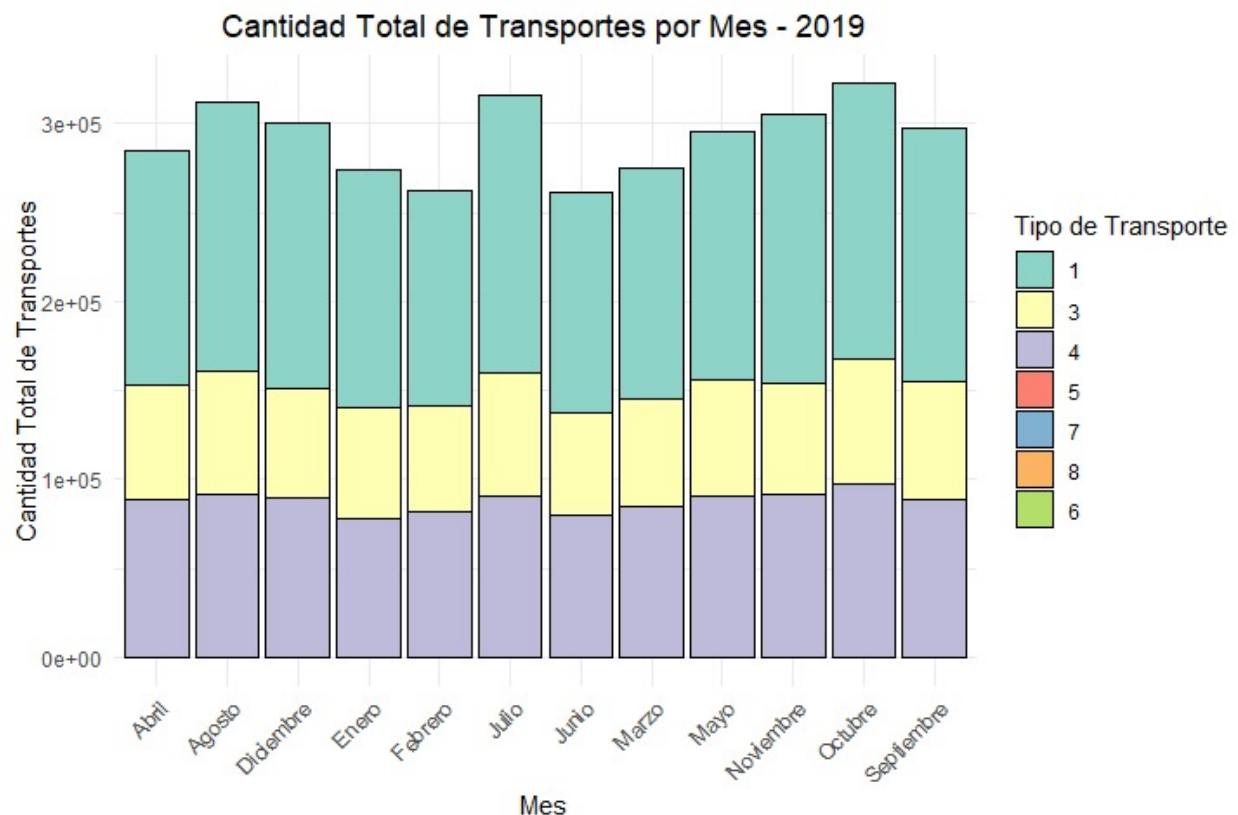
5 CORREO

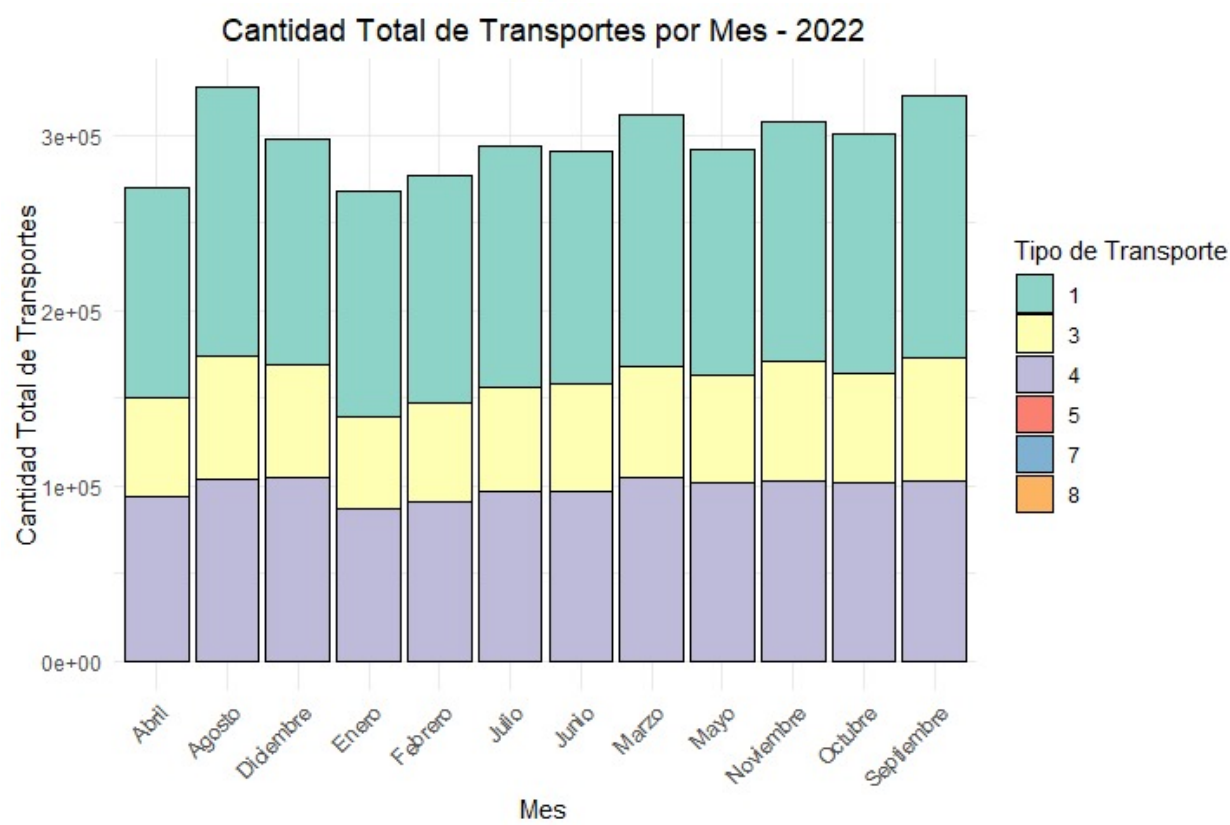
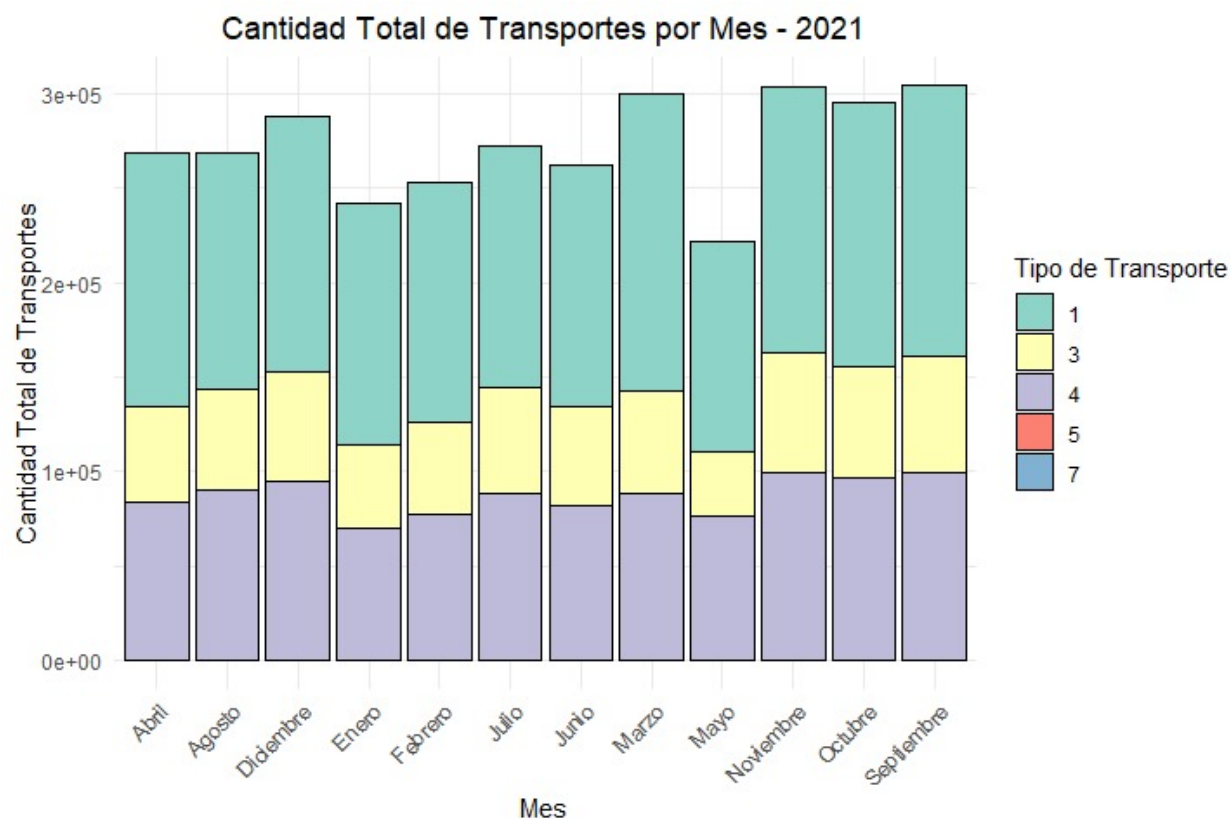
6 MULTIMODAL

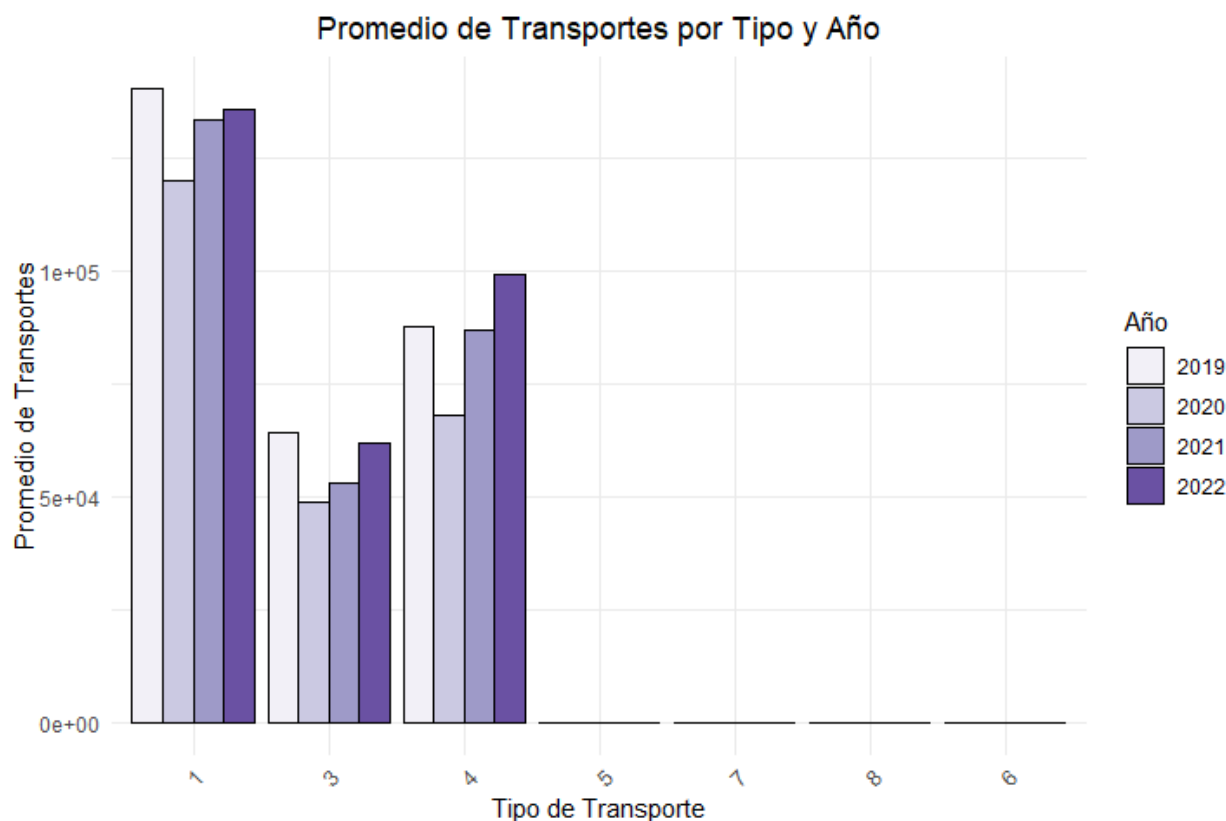
7 INSTALACIONES DE TRANSPORTE FIJAS (TUBERIAS, CABLE, ETC.)

8 VIAS NAVEGABLES INTERIORES

## 9 OTRO MODO DE TRANSPORTE







## 5. Conclusiones

Através de nuestro estudio logramos observar que efectivamente la pandemia comprometió las actividades de importación, siendo reflejado en una **disminución** de frecuencia, del precio total de mercancía y del costo del flete durante el 2020 que fue el año central del suceso, donde podemos llamar al acontecimiento como una recesión económica y después se refleja una alza exponencial dentro de 2021 y 2022 después de pandemia donde los items analizados se dispararon , principalmente en 2022 donde, como comentábamos, tenemos la conjetura que fue por el crecimiento de las compras online lo que llevo a la llegada de muchas más importaciones, probablemente de mayor valor a diferencia de las acostumbradas a llegar.

Por otro lado, debemos mencionar que el costo de los fletes que es el costo total del valor del transporte de cada mercancía era bastante bajo en 2019 y tuvo un crecimiento leve en 2020 pero se fue al alza con rapidez en 2021-2022, la conjetura del motivo son las implementaciones de medidas de bioseguridad y revisión de las cargas por parte de entidades de salubridad para el manejo y disipación del virus.

También hacemos mención a que las importaciones nos ayudaron a reflejar por completo el suceso de pandemia puesto que mostro la escasez a finales de 2019 y todo el 2020, el crecimiento inestable de 2020 y el crecimiento exponencial de 2021 y 2022. Entendemos que las importaciones también se pudieron ver afectadas por otros factores como el ambiental el político etc...