

Este proyecto nace como un instrumento para poder aplicar los contenidos aprendidos durante el semestre en un problema real. Se tendrá que demostrar los conocimientos adquiridos al realizar un analisis exploratorio sobre datos de interes nacional, poniendo especial énfasis en la correcta visualización de los resultados obtenidos, realizando una correcta interpretación de la misma. En el caso de este grupo en cuestión, se analizarán datos de incautaciones de aduana presentados por la Dirección Nacional de Aduanas, que fueron elegidos Para llevar a cabo estos objetivos, relacionando tipos y el valor las de inacutaciones, cantidades, procedencia, entre otras variables. Se utilizará exclusivamente el lenguaje de programación R, y sus respectivas expansiones, como Markdown, Shiny y RStudio.

```
## here() starts at /home/fede/github/Repositorios_Alumnos/ProyFinal_Sarkisian_Colio_Nunez
## -- Attaching packages ----- tidyverse 1.3.0 --

## v ggplot2 3.3.3      v purrr 0.3.4
## v tibble 3.0.6       v dplyr 1.0.4
## v tidyr 1.1.2        v stringr 1.4.0
## v readr 1.4.0        v forcats 0.5.1

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()

## [1] 78359      16
```

Los datos fueron obtenidos en el catalogo de datos abiertos que puso a disposición el gobierno uruguayo. Más específicamente estos fueron subidos por la Dirección Nacional de Aduanas. Cuentan con 16 variables y 76.792 observaciones. Tres de las variables son de tipo double, mientras que las restantes son de tipo caracter. Cada observación se refiere a una incautación realizada, y ofrece información sobre ella. Las variables que utilizaremos son:

% latex table generated in R 4.1.0 by xtable 1.8-4 package % Wed Jul 21 11:37:57 2021

Variables	Definicion
fecha_incautacion	Fecha en la cual fue incautada la mercaderia
aduana	Ubicacion de la aduana(Local)
infraccion	Tipo de infraccion cometida
país_procedencia	País del cual procede la mercaderia incautada
unidad	Unidades de medida en la cual esta medida la mercaderia
estado_incautacion	Estado en el cual se encuentra dicha incautacion ,(Contensioso(Final o inicial),Falsificador, Valorado)
tipo	Tipo de mercaderia incautada
cantidad	Cantidad de unidades incautadas medidas con respecto a la variable 'unidad'
valor	Valor de dicha de la mercaderia incautada medida en U\$

Table 1: Descripción de las Variables

```
##
## Attaching package: 'data.table'

## The following objects are masked from 'package:dplyr':
##
##   between, first, last

## The following object is masked from 'package:purrr':
##
##   transpose
```

Para realizar la exploración, nos planteamos algunas preguntas para orientar el análisis. En primer lugar, decidimos investigar de qué manera había evolucionado la cantidad total de incautaciones durante el período en cuestión, teniendo en cuenta las diferentes infracciones que se pueden realizar.

```
##
## Attaching package: 'ggpmisc'
## The following object is masked from 'package:ggplot2':
##
##   annotate
##
## Attaching package: 'lubridate'
## The following objects are masked from 'package:data.table':
##
##   hour, isoweek, mday, minute, month, quarter, second, wday, week,
##   yday, year
## The following objects are masked from 'package:base':
##
##   date, intersect, setdiff, union
```

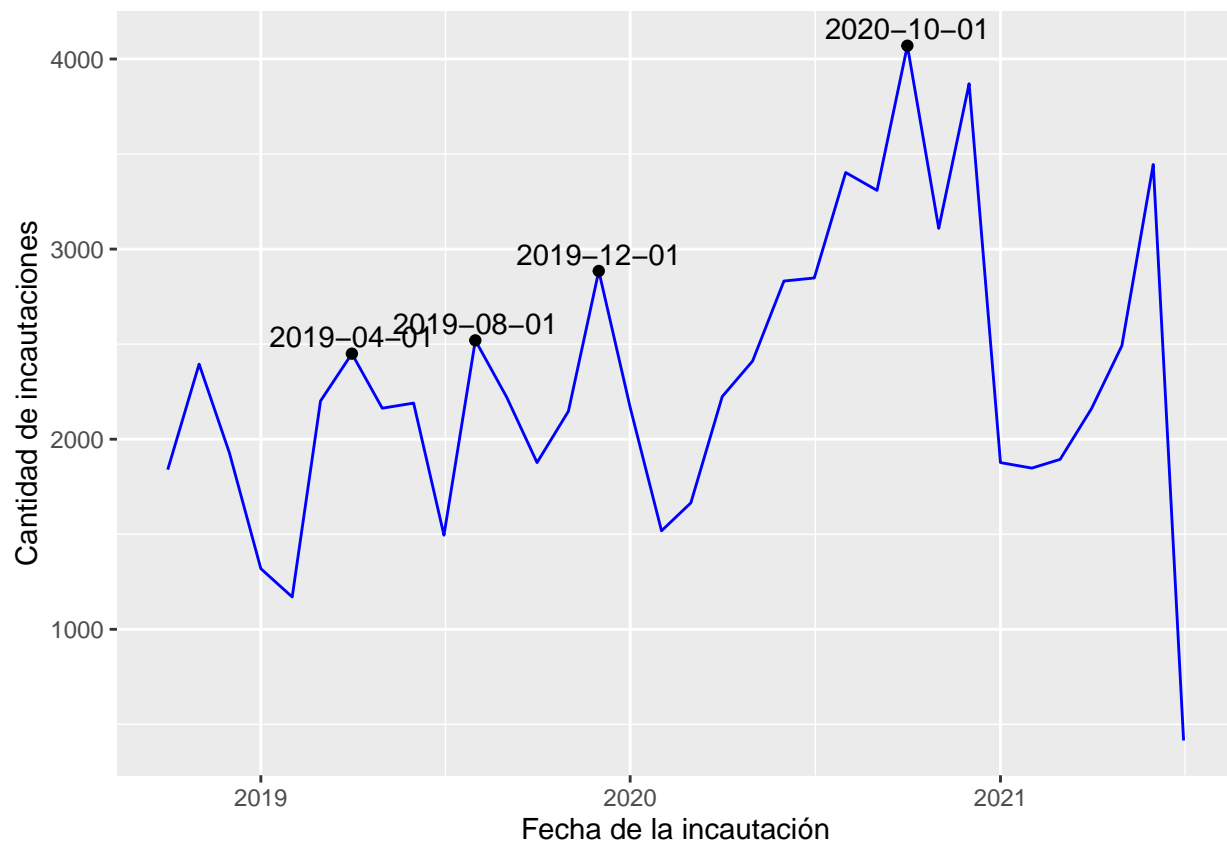


Figure 1: Se puede ver un aumento en la cantidad de incautaciones en el correr del año con un gran descenso en el verano diciembre-enero, puede deberse a un menor control aduanero y un enfoque en el turismo, o por un menor comercio en esta época. Se puede ver también como el covid no disminuyó las incautaciones

```
## 'summarise()' has grouped output by 'infraccion'. You can override using the '.groups' argument.
```

Luego decidimos analizar más en profundidad la variable “Valor”, que indica el valor de la incautación realizada en cada caso. Para hacer esto, iniciamos observando la distribución de la variable, para luego adentrarnos en la relación que tiene el valor (medido en pesos) con la cantidad de incautaciones incautadas.

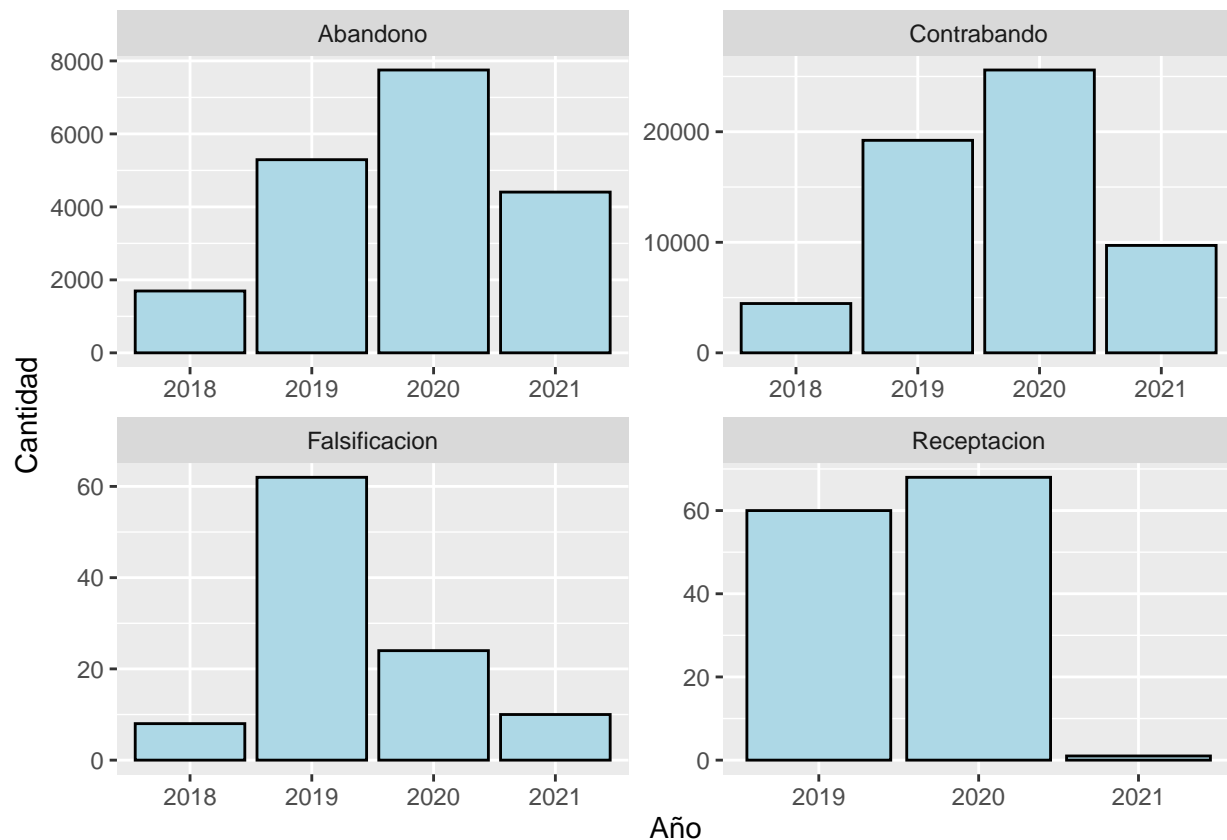


Figure 2: Los bajos datos del año 2021 se debe a que los datos llegan solo hasta mitad de año. Existe un claro aumento desde el 2018 a 2019 en todos los tipos de incautaciones, puede ser por un mayor control en las aduanas o un cambio de politica, hay un gran pico de falsificaciones en 2019, mientras que en el 2020 los otros tipos de incautaciones frecuentaron más

Es importante destacar que “Valor” presenta datos extremos que dificultan la correcta visualización, por lo que se realizaron calculos básicos para determinar y excluir los valores atípicos y poder observar únicamente los datos de interés.

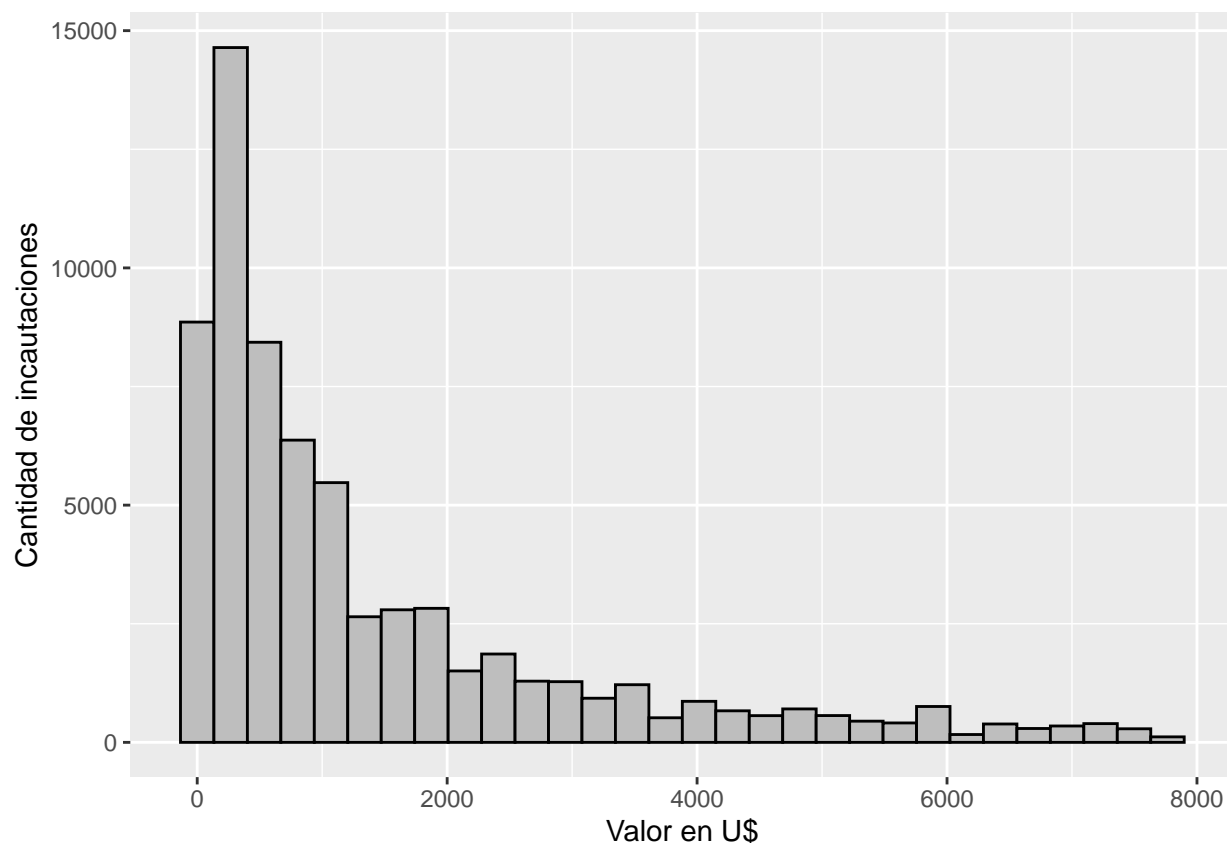


Figure 3: Se puede observar que la mayoría de las incautaciones son de un bajo valor

```
##
## Attaching package: 'plotly'
## The following object is masked from 'package:ggplot2':
##
##   last_plot
## The following object is masked from 'package:stats':
##
##   filter
## The following object is masked from 'package:graphics':
##
##   layout
## [1] 0.1082492
```

A continuación nos enfocamos en la variable “aduana”, que hace referencia al lugar dentro del Uruguay que se realizó el acto de incautar. Como primer paso se investigó de que manera se distribuyen las incautaciones totales dentro del territorio nacional. Sucesivamente, nos interesó saber el destino de cada uno de estos actos, sabiendo también el origen.

Como último objetivo, decidimos investigar el estado en el que se encuentra cada incautación, intentando

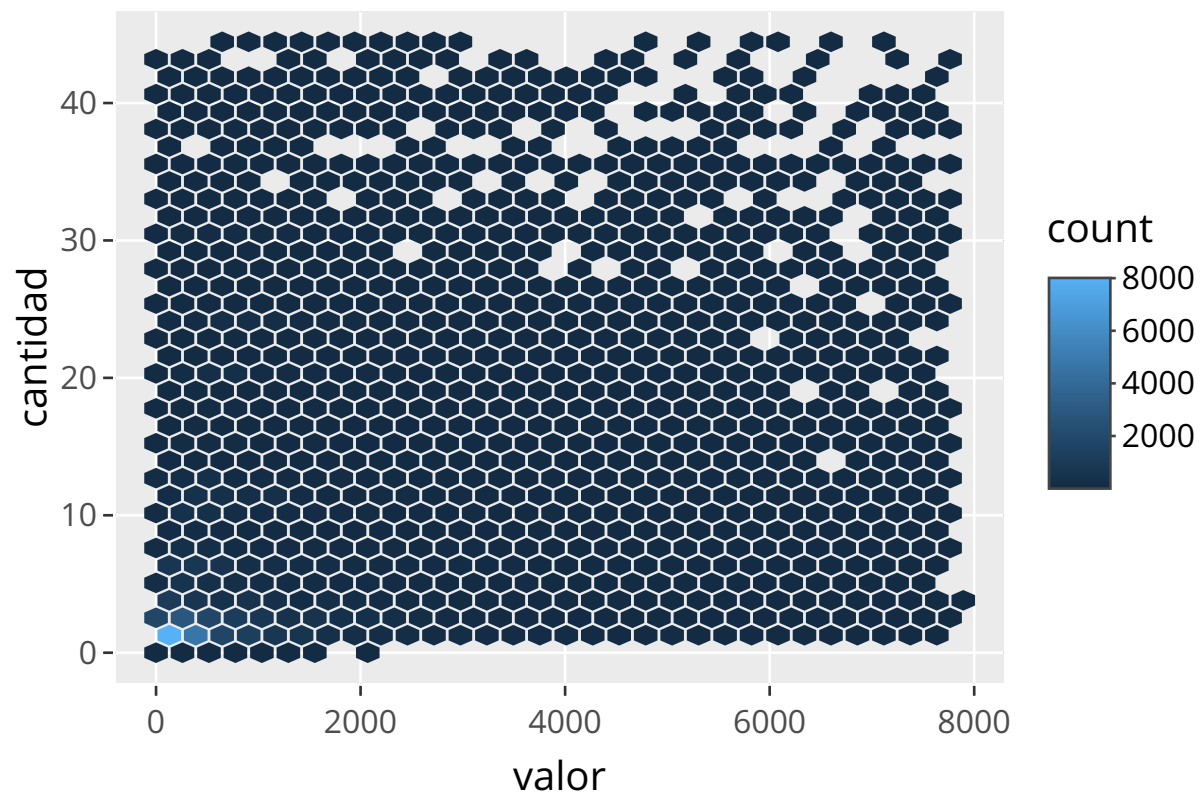


Figure 4: Se observa una correlacion casi nula entre las variables, pero si una clara tendencia a que las incautaciones sean de bajo valor y en poca cantidad

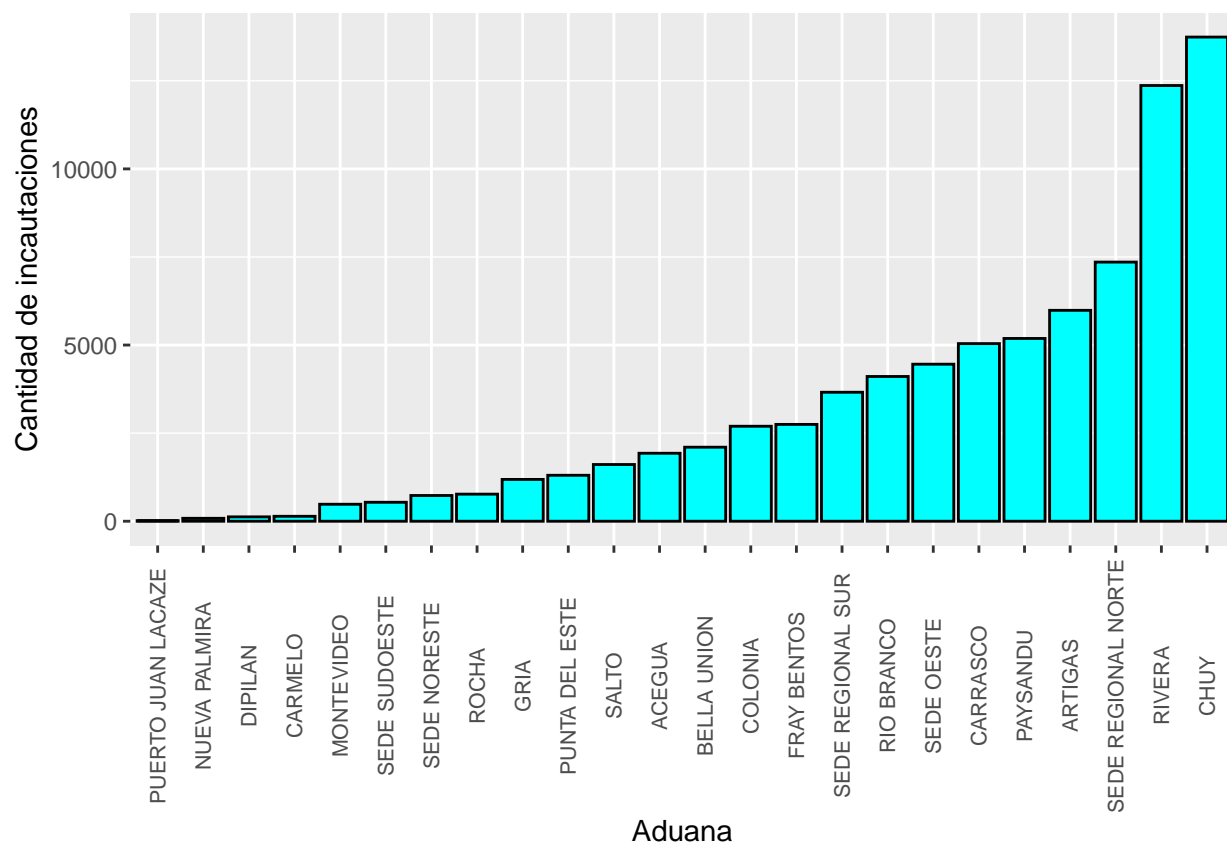


Figure 5: Hay una clara mayor cantidad de incautaciones en las aduanas del este y norte, se debe al mercado con el pais fronterizo Brasil, el cual se diferencia claramente de las aduanas del oeste

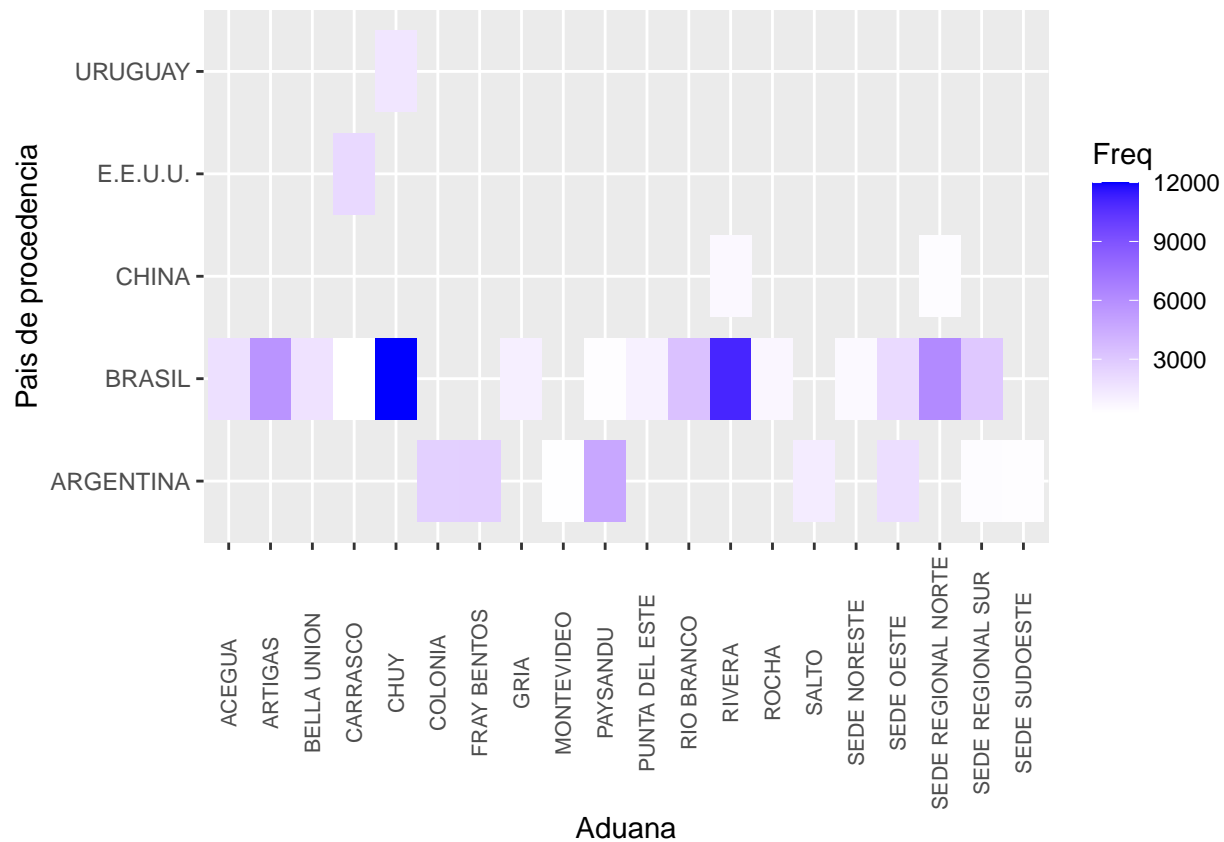
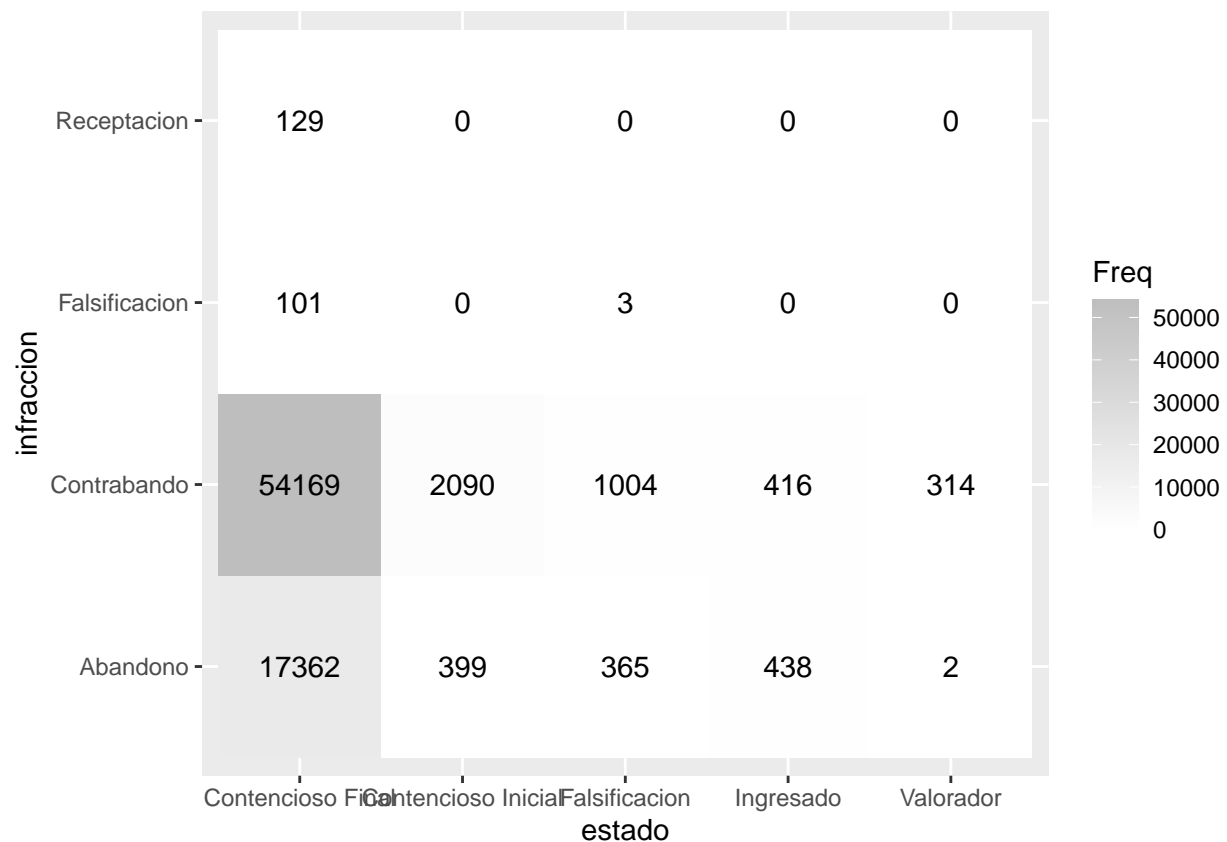


Figure 6: En el grafico se puede observar claramente como la mayor cantidad de incautaciones de dan en las aduanas del este y norte, las de procedencia brasilera. Y como tambien por parte de argentina pasa lo mismo, pero con aduanas del oeste, en cuanto a las incautaciones provenientes de EEUU todas son incautadas en Carrasco algo que puede tener que ver con el tipo de bien y el nivel adquisitivo

observar cuál infracción demora más tiempo en resolverse.

'summarise()' has grouped output by 'infraccion'. You can override using the '.groups' argument.

```
## # A tibble: 13 x 3
## # Groups:   infraccion [4]
##   infraccion estado_incautacion conteo
##   <chr>      <chr>          <int>
## 1 Abandono   Contencioso Final    18447
## 2 Abandono   Contencioso Inicial    250
## 3 Abandono   Falsificacion         410
## 4 Abandono   Ingresado             20
## 5 Abandono   Valorador             17
## 6 Contrabando Contencioso Final   55917
## 7 Contrabando Contencioso Inicial  1630
## 8 Contrabando Falsificacion       1189
## 9 Contrabando Ingresado            96
## 10 Contrabando Valorador           150
## 11 Falsificacion Contencioso Final    101
## 12 Falsificacion Falsificacion         3
## 13 Receptacion Contencioso Final    129
```



Al terminar con la exploración, el siguiente paso fue crear una Shiny web App con el objetivo de transmitir de una manera amigable con el usuario las principales exploraciones conseguidas.

###link

Dentro de la app, lo primero que se quiso comunicar fue información exclusiva de cada tipo de incautación.

La forma de lograr esto fue incorporar un menú desplegable, de donde el usuario puede elegir el tipo de incautación que más le interese. Se muestran dos gráficos con respecto a ellos. El primero es similar a la evolución total de incautaciones realizada anteriormente, pero específico para cada tipo de incautación. Y el segundo permite visualizar de qué países provinieron esas incautaciones.

En el siguiente panel, se observa la cantidad total de incautaciones según la infracción realizada, permitiendo observar individualmente la evolución de cada una de ellas.

Más adelante, se permite delimitar el rango de valores que se quieren observar dentro del gráfico que muestra las incautaciones que ha realizado cada aduana durante los últimos años. Es una gran forma de observar en qué aduanas se concentran las incautaciones de mayor o menor valor.

Para finalizar, se agregó el gráfico que relaciona valor y cantidad de artículos, pudiendo seleccionar según cuál variable categórica se pretende pintar el interior. Solo se seleccionaron las variables que el gráfico resultante permitiera sacar algún tipo de conclusión.

Luego de esta exploración pudimos observar como la aparición del covid-19 no disminuyó la cantidad de incautaciones, sino que hasta aumentaron en gran medida en comparación con el 2019, también que la mayoría son de un muy bajo valor, una clara mayor cantidad de incautaciones de procedencia brasileña en comparación con el otro país fronterizo (Argentina), lo cual indica un mayor mercado ilegal o un mayor control por parte de las autoridades en las aduanas del este.