

Aplicación de infracciones en defensa del consumidor

Materia: Modelos analíticos – TP grupal 2do parcial – 2do Cuatrimestre 2023

Profesor: Javier Di Salvo

Estudiantes: Federico Gauna, Jonathan Megra, Christian Toledo y Juan Torossi.

Introducción al análisis

En este proyecto vamos a trabajar y analizar los datos públicos provistos por el gobierno sobre las multas impuestas en defensa de las y los consumidores. Desde una perspectiva general, los principales puntos a analizar serán las empresas con mayor cantidad de sanciones, el monto total de cada infracción y el promedio de sanciones impuestas.

A partir de este análisis, se llevará a cabo un tablero a nivel estratégico, el cual tendrá como objetivo poner en evidencia cuales son las empresas más destacadas con respecto a cantidad de infracciones y su respectivo monto, y cuáles son las infracciones más comunes.

Base de datos y Contexto

El conjunto de datos a utilizar proviene de la página gubernamental datos.gob.ar y es compartido por el Ministerio de Desarrollo Productivo. El dataset se encuentra en el siguiente [link](#).

El mismo contiene la información sobre sanciones aplicadas en defensa de las y los consumidores, siendo estas impuestas por la Dirección Nacional de Defensa del Consumidor y Arbitraje del Consumo, y la Dirección de Protección Jurídica al Consumidor.

El dataset fué publicado el 6 de mayo del 2021 y actualizado por última vez el 26 de diciembre del 2022. Abarca a todas las sanciones aplicadas en el periodo comprendido entre agosto de 2020 y noviembre de 2022.

Este dataset lleva el nombre de “sanciones-aplicadas-defensa-consumidor” y contiene la información de todas las sanciones cuyos procesos judiciales se completaron, detallando el nombre de la empresa, las infracciones cometidas junto con su respectiva descripción, el monto de la infracción y la fecha de imposición de la multa (vista previa a continuación).

N° de Disposición	Empresa	Infracción	Descripción	Monto	Fecha
3	Carsa SA (Musimundo)	Art. 4, 7 y 19	Incumplimiento en la obligación de informar, de la oferta de productos y de la prestación del servicio en las condiciones pactadas	1500000	14/08/2020
5	Gire SA (Rapipago)	Art. 4 y 19	Incumplimiento en la obligación de informar y de la prestación del servicio de cobranza por realizar débitos indebidos	1500000	18/08/2020
6	Cencosud SA	Art. 7	Incumplimiento del Programa "Precios Cuidados" (oferta)	250000	19/08/2020
7	Cencosud SA	Art. 7	Incumplimiento del Programa "Precios Cuidados" (oferta)	250000	20/08/2020
8	Día Argentina SA	Art. 7	Incumplimiento del Programa "Precios Cuidados" (oferta)	250000	20/08/2020
9	Día Argentina SA	Art. 7	Incumplimiento del Programa "Precios Cuidados" (oferta)	450000	25/08/2020
10	Día Argentina SA	Art. 7	Incumplimiento del Programa "Precios Cuidados" (oferta)	450000	25/08/2020
11	Inc SA (Carrefour)	Art. 7	Incumplimiento del Programa "Precios Cuidados" (oferta)	350000	25/08/2020
12	Día Argentina SA	Art. 7	Incumplimiento del Programa "Precios Cuidados" (oferta)	400000	25/08/2020

Estructura de los datos

Con respecto a la estructura de los datos del dataset, su dimensión está caracterizada con 1644 filas y 6 columnas. La descripción de cada una de estas columnas es la siguiente:

Título	Tipo de dato	Descripción
Nº disposición	Entero (integer)	(ID) Nº de disposición de la multa, registro de GDE
Empresa	Texto (string)	Proveedor multado
Infracción	Texto (string)	Referencia al Artículo, ley o práctica por la que se lo multa
Descripción	Texto (string)	Descripción de la infracción
Monto	Texto (string)	Monto de la multa
Fecha	Fecha ISO-8601 (date)	Fecha de firma de la multa

Limpieza de la base de datos

En primera instancia se han normalizado todos los datos de forma tal de mantener cada atributo con el mismo tipo de dato y que sea sencillo agrupar los mismos. Los principales cambios en esta materia fueron sobre las dimensiones de “Monto” y “Fecha”, en donde dejamos los montos expresados en moneda “\$” (pesos) y la Fecha fue normalizada en formato corto (“dd/mm/aaaa”)

Luego comenzamos a realizar una análisis exploratorio sobre los datos para identificar valores faltantes, incorrectos, nulos o duplicados. De esta forma logramos identificar algunos errores de ingreso de datos por los usuarios que fueron corregidos para alinear los registros entre sí y obtener una sólida base de datos.

Se encontraron varios registros incorrectos donde figuraba más de una empresa y a los fines de este análisis se decidió agrupar las empresas dentro del grupo corporativo al que pertenecen, por ejemplo, “Jumbo SA” y “VEA SA” se unificaron bajo el grupo “Cencosud SA”. A su vez, en algunos registros figuraba más de una empresa como infractora debido a que la infracción podía ser adjudicada tanto a la empresa como a su proveedor (por ejemplo: por información faltante en la página web o publicidad engañosa) y dado que el dataset no permite realizar una discriminación por cada empresa y su responsabilidad, se optó por dejar como única empresa a la empresa proveedora del servicio que publicó finalmente la información errónea. (por ejemplo: Mercadolibre SRL y Suavestar, se mantuvo a Mercadolibre SRL como infractora)

En cuanto a datos faltantes, tanto en los campos “Monto” y “Fecha” encontramos inconsistencias en los datos. En el caso del campo de Monto, encontramos varios registros con valor 0 y algunos con valores en formato “String” que se decidieron eliminar ya que no aportan información útil en nuestro análisis. Por otro lado bajo el atributo “Fecha” encontramos varios campos nulos y, de acuerdo a la documentación del dataset, este dato refiere a la fecha en la cual es firmada la multa, de tal forma si no se encuentra firmada, nunca fue puesta en vigencia, por lo que también se eliminaron estos registros nulos.

Por último, se eliminaron aquellos registros que se encontraban fuera de la ley de defensa al consumidor 24240 ya que en el análisis sólo tendremos en cuenta aquellos registros que se basen en infracciones de la Ley 24240 de Defensa al Consumidor.

Dentro del análisis exploratorio surgió la necesidad de agregar más columnas al dataset con el fin de poder mejorar la calidad de la información.

En primer lugar se optó por agregar una descripción breve y categórica sobre las descripciones provistas para cada registro de infracciones con el fin de categorizarlas y elaborar una correcta visualización agrupada, dado que existen varios registros que comprende más de un único artículo en infracción. Debajo se detallan [los artículos que se consideran en el Dataset y sus categorías según la página oficial del gobierno “Infoleg”](#).

Articulos	Categoria
Art. 4	Información
Art. 5	Protección al Consumidor.
Art. 6	Cosas y Servicios Riesgosos.
Art. 7	Oferta
Art. 8	Efectos de la publicidad
Art. 8bis	Trato digno. Prácticas abusivas.
Art. 10	Contenido del documento de venta
Art. 10bis	Incumplimiento de la obligación.
Art. 10ter	Modos de Rescisión.
Art. 10quater	Prohibición de cobro.
Art. 11	Garantías.
Art. 12	Servicio Técnico.
Art. 18	Vicios Redhibitorios
Art. 19	Modalidades de Prestación de Servicios
Art. 20	Materiales a Utilizar en la Reparación.
Art. 21	Presupuesto.
Art. 29	Instrumentos y Unidades de Medición.
Art. 32	Venta domiciliaria.
Art. 34	Revocación de aceptación.
Art. 35	Prohibición.
Art. 36	Requisitos.
Art. 37	Interpretación.
Art. 38	Contrato de adhesión.

Las categorías que se decidieron generar para agrupar los registros son:

1. Falta de Información a disposición del consumidor
2. Incumplimiento del servicio
3. Incumplimiento del Programa "Precios Cuidados"
4. Otros

Dentro de la agrupación de “Falta de Información a disposición del consumidor” se incluyen todos aquellos casos donde la empresa no incluyó información necesaria y mandatoria de acuerdo a la ley de defensa al consumidor, como ser el el “Botón de Baja” o “Botón de Arrepentimiento”, falta de información mandatoria del producto o servicio, no incluir los términos y condiciones del contrato a suscribir o una copia del mismo, etc.

En segundo lugar, la dimensión de “Incumplimiento del servicio” agrupa aquellas infracciones donde la empresa haya incumplido con el acuerdo pactado, como por ejemplo: no tramitar las bajas solicitadas por los usuarios, entrega incompleta de un producto o servicio, que la empresa no garantice proveedores de reparaciones de su producto, cancelación indebida de una compra, etc.

La única descripción que decidimos mantener sin modificaciones es la de incumplimiento del programa de “Precios Cuidados” ya que esta descripción corresponde a casi la mitad de las infracciones incluidas dentro del dataset.

Varias descripciones agrupadas fueron incluidas dentro de “Otros” dado que sus descripciones eran muy específicas y no se podían agrupar en otras categorías. Esta categoría incluye infracciones varias como ser inconvenientes con los pagos y remuneraciones, cláusulas abusivas, publicidades engañosas, reclamos por trato indigno hacia los clientes y productos que puedan atentar contra la salud del usuario.

Para poder evaluar los artículos que determinaban la infracción, se realizó una extracción del campo "Infracción" y se incorporaron la cantidad de columnas como artículos obtenidos y se imputaron valores binarios según correspondiera (similar sklearn.preprocessing.LabelEncoder). A continuación captura y link al código en Python utilizado.

```
sanciones_def_cons.py > ...
1  # -*- coding: utf-8 -*-
2  import re
3
4  import pandas as pd
5  import numpy as np
6  import matplotlib.pyplot as plt
7
8
9  archivo = ''
10 df = pd.read_excel(archivo)
11 df.drop(df.columns[10:], axis=1, inplace=True)
12
13 # Corrección de formato de fecha
14 df['Fecha'] = df['Fecha'].dt.strftime('%d/%m/%Y') # Drive
15
16
17 def capturar_nros(texto):
18     """capturar_nros
19     Args:
20         texto (str): cadena a explorar.
21     Returns:
22         str: números enteros de la cadena recibida.
23     """
24     num = re.findall(r'[0-9]+', texto)
25     resultado = []
26
27     for i in range(len(num)):
28         num[i] = int(num[i])
29
30         if num[i] not in resultado and num[i] < 41:
31             resultado.append(num[i])
32
33     resultado.sort()
34     resultado = [str(x) for x in resultado]
35     return ' '.join(resultado)
36
37 # Creación de DataFrame temporal
38 temp = df['Infraccion'].apply(lambda x: capturar_nros(x))
39 art_infrac = pd.DataFrame(temp)
40
```

```

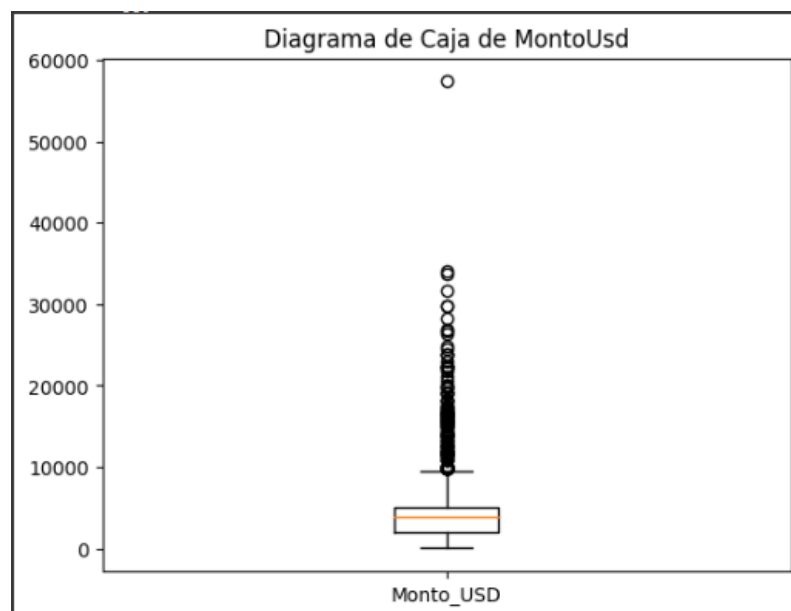
40
41 # Captura de los valores únicos de toda la serie
42 unicos_str = ''
43
44 for values in art_infrac['Infraccion'].str.split():
45
46     for value in values:
47
48         if f'{value}' not in unicos_str:
49             unicos_str = unicos_str + ' ' + value
50
51 unicos = [int(x) for x in unicos_str.split()]
52 unicos.sort()
53
54 # Creación de DataFrame con los artículos
55 df_infracciones = pd.DataFrame(columns=['Art. ' + str(x) for x in unicos])
56
57 # Asignación de valores por fila
58 for i in range(len(art_infrac.index)):
59     fila = [1 if str(x) in art_infrac['Infraccion'][i] else 0 for x in unicos]
60     df_infracciones.loc[len(df_infracciones.index)] = fila
61
62 # Asignación para valores no contemplados
63 df_infracciones['Otros arts.'] = [1 if sum(df_infracciones.loc[i]) == 0 \
64     else 0 for i in range(len(df_infracciones.index))]
65
66 sanciones_prepro = pd.concat([df, df_infracciones], axis=1)
67
68 # Exportación
69 ruta_archivo = ''
70 sanciones_prepro.to_excel(ruta_archivo, index=False, sheet_name='Preprocesado')

```

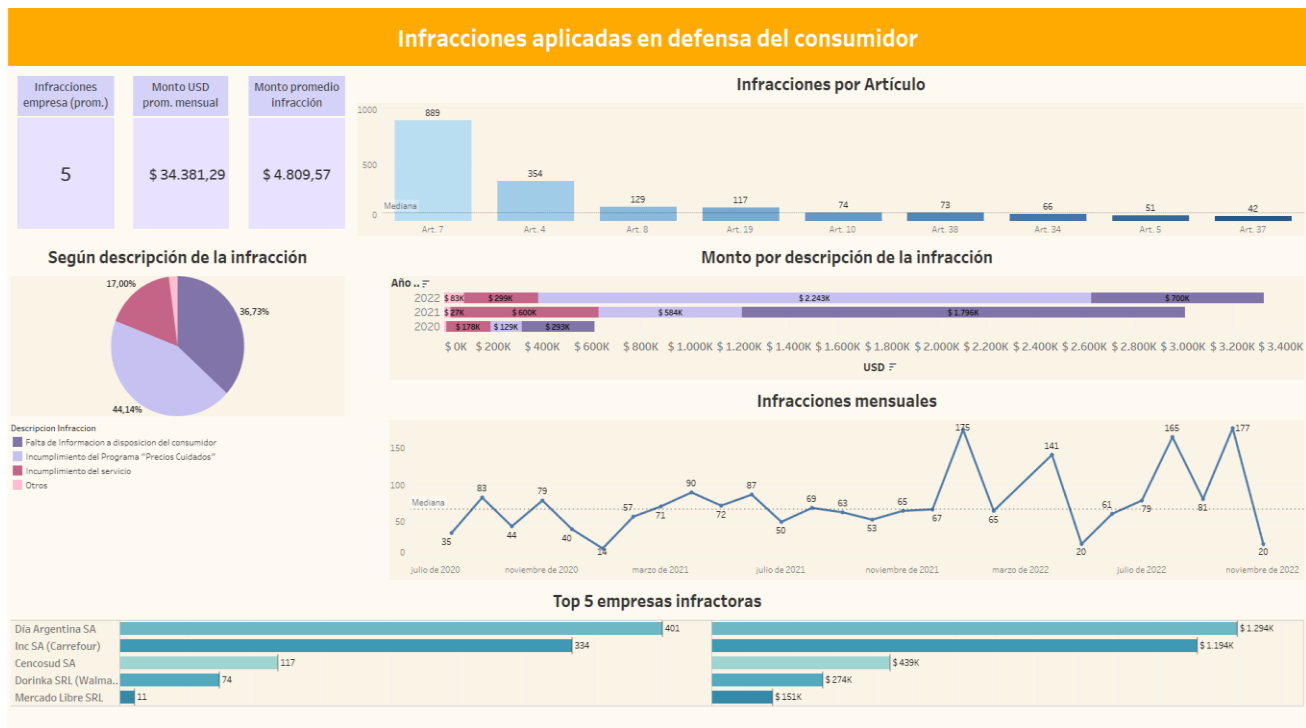
[Link al código en Google Collaboratory](#)

Una de las columnas que se agregaron al dataset fue “TipoCambio_USD” que contiene la cotización del dólar por día, con el fin de elaborar una nueva dimensión “Monto USD”, donde se encuentren los montos abonados calculados en dólares. De esta forma podemos determinar la evolución de los montos de las multas dependiendo del año en el cual fueron firmadas y poder así comparar entre los distintos años en base a una moneda más estable.

Al agregar la columna con los montos en dólares se logró realizar también un análisis de valores atípicos en base a un Diagrama de Caja sobre el dataset que hemos analizado. Debido a que el objetivo del proyecto es identificar a los mayores infractores tanto en monto como en frecuencia junto con aquellas infracciones que son más comunes, se mantuvieron los registros atípicos del dataset.



Visualización de datos



[Link al tablero en la web.](#)

Para la visualización de los datos, optamos por desarrollar un Dashboard a través de la herramienta de Tableau que nos permite realizar un análisis más detallado sobre el dataset.

Para esto realizamos algunas **métricas** para el análisis del dataset.

❖ **Promedio de cantidad de infracciones por Empresa**

Fórmula: Cantidad de Infracciones / Cantidad de Empresas

❖ **Monto Promedio de infracción por empresa**

Fórmula: SUMA (Infracciones) / Cantidad Empresas

❖ **Promedio por Infracción**

Fórmula: SUMA (Infracciones) / Cantidad Infracciones

En este mismo Dashboard se incluyeron distintos tipos de gráfico de forma tal que se reduzca la carga cognitiva del lector y permita un mejor análisis de los datos.

- **Gráfico de Barras Verticales:** Top Cantidad Infracciones por Artículo

En este gráfico se visualizan los artículos con mayor cantidad de infracciones.

- **Gráfico de Torta:** Representación por Descripción de la Infracción

Visualiza la distribución de cada Descripción Agrupada en base al porcentaje del total.

- **Gráfico de Barras Horizontales:** MontoUSD por Descripción de la Infracción

Gráfica de la distribución anual en cuanto al MontoUSD de cada descripción agrupada y evidencia la evolución anual de cada uno.

- **Gráfico de Líneas:** Cantidad de Infracciones por Mes.

Representa la cantidad total de infracciones por mes dentro del periodo de tiempo que abarca el análisis (ago 2020 - nov 2022).

- **Gráfico de Barras Horizontales:** Top 5 Empresas Infractoras (Por Cantidad y MontoUSD)

El gráfico se encuentra dividido por la dimensión calculada, en primera instancia por la cantidad de infracciones cometidas por empresa y en segundo lugar por el MontoUSD por empresa.

Conclusiones:

Luego de analizar el dataset de “Sanciones aplicadas en defensa de las y los consumidores” pudimos identificar claramente aquellos Artículos que más frecuentemente fueron infringidos por las empresas de bienes y servicios, siendo ampliamente superiores los artículos 7 y 4, y en menor medida los Artículos 8, 19 y 10.

Por otro lado, pudimos identificar aquellas empresas que más infracciones tienen, tanto por la cantidad como por el monto, y son: Día Argentina SA, INC SA, Cencosud SA y Dorinka SRL, todas pertenecientes al segmento de Supermercados y Retail y se vieron fuertemente afectadas por sus infracciones al Incumplimiento del Programa de Precios Cuidados. Ante esta información, se recomienda realizar acciones correctivas y punitivas sobre las empresas de este segmento para que mejoren su nivel de servicio y acoplamiento al régimen de “Precios Cuidados”.

El resto de las empresas tienen una distribución equitativa entre el incumplimiento del servicio y falta de información a disposición del consumidor, por lo que se recomienda realizar un seguimiento más extensivo sobre estas compañías para que no promocionen incorrectamente sus servicios y cumplan con lo pactado.

La frecuencia y gravedad de las multas sugieren una necesidad urgente de mejorar el cumplimiento normativo en varias empresas, por lo que enfatizamos en la importancia de fortalecer la supervisión y cumplimiento normativo, especialmente en áreas críticas como el Programa "Precios Cuidados" y servicios de información clara a los clientes, para garantizar la protección de los derechos de los consumidores.