# Analyzing impact of Covid-19 Pandemic and the effect of quarantine in Argentina

**Applied Data Science Capstone** 

Martín Alejandro Heredia

# Introduction

### The Problem

From the beginning of Covid-19 Pandemic, people's life has changed drastically due to the fast and exponential growth of cases, which leads governments to take decisions on how to prevent the expansion rate of the disease, health system saturation and economic impact.

The common denominator is quarantine. This helps to reduce infection rate, giving more time to governments to prepare health system. In opposition to this positive consequence of quarantine is the economic, social and industrial impact that quarantine has on a country. This positive and negative effect of quarantine makes it a "trade-off" solution.

The analysis of growth of cases, distribution of disease on the territory, impact of quarantine, among others; is fundamental for governments to take decisions. Also, viewing the evolution of disease and its influence on indicators such like economics or industrial ones, gives governments tools to evaluate and reorganize its own decisions.

This study is focused on Argentinian case. First, statistical data of Covid-19 on Argentina will be analyzed: how numbers of cases are distributed on its geography and where are the most affected cities. Then, the effect of quarantine in other areas (such like transport, industry, economic) will be studied and compared with the same period of the previous year.

### The Data

To analyze the effect of Covid-19 on the territory, the data from the following website of Argentinian Government will be used:

https://sisa.msal.gov.ar/datos/descargas/covid-19/files/Covid19Casos.csv

In [3]:	]: url='https://sisa.msal.gov.ar/datos/descargas/covid-19/files/Covid19Casos.csv' cases=pd.read_csv(url,encoding='utf-16') cases.head()											
Out[3]:	i	d_evento_caso	sexo	edad	edad_años_meses	residencia_pais_nombre	residencia_provincia_nombre	residencia_depar				
	0	672064	М	52.0	Años	Argentina	Buenos Aires					
	1	717629	F	46.0	Años	Argentina	Buenos Aires					
	2	717926	F	41.0	Años	Argentina	CABA					
	3	718029	F	52.0	Años	Argentina	Buenos Aires					
	4	718055	F	34.0	Años	Argentina	CABA	S				
	5 rov	vs × 25 columr	ıs									

All the categories (columns) of this dataset are depicted below:

```
In [13]: for category in cases.columns.sort values():
             print(category)
         Clasificacion
         asistencia respiratoria mecanica
         carga provincia id
         carga provincia nombre
         clasificacion resumen
         cuidado intensivo
         edad
         edad años meses
         fallecido
         fecha apertura
         fecha cui intensivo
         fecha diagnostico
         fecha_fallecimiento
         fecha_inicio_sintomas
         fecha internacion
         id evento caso
         origen financiamiento
         residencia_departamento_id
         residencia departamento nombre
         residencia_pais_nombre
         residencia provincia id
         residencia provincia nombre
         sepi_apertura
         sexo
         ultima_actualizacion
```

This columns provides information such like:

- Age and gender of the patients
- Dates (diagnosis date, hospitalization date, date of death, among others)
- Geographical information (province name, department name)
- Classification (active, suspect, cured, among others)
- ID of each case

With this dataset, we can visualize the impact of this disease with indicators such like:

- Total of cases
- Total of deaths
- Total of recovered patients
- Total of active cases
- Total number of hospitalized patients
- Total of cases on intensive therapy

In addition to this data, the use of Foursquare will help us to visualize the distribution of cases on the country.

After this preliminary study, the work will be focused on analyze and visualize the impact of quarantine on the following areas:

- Economic activity:

  <a href="https://www.indec.gob.ar/ftp/cuadros/economia/sh\_emae\_mensual\_base2004.xls">https://www.indec.gob.ar/ftp/cuadros/economia/sh\_emae\_mensual\_base2004.xls</a>

  <a href="https://www.indec.gob.ar/ftp/cuadros/economia/sh\_emae\_actividad\_base2004.xls">https://www.indec.gob.ar/ftp/cuadros/economia/sh\_emae\_actividad\_base2004.xls</a>
- Industrial activity: <u>https://www.indec.gob.ar/ftp/cuadros/economia/sh\_ipi\_manufacturero\_2020.xls</u>
- Transport (railway):
   https://servicios.transporte.gob.ar/gobierno\_abierto/descargar.php?
   t=trenes&d=pasajeros

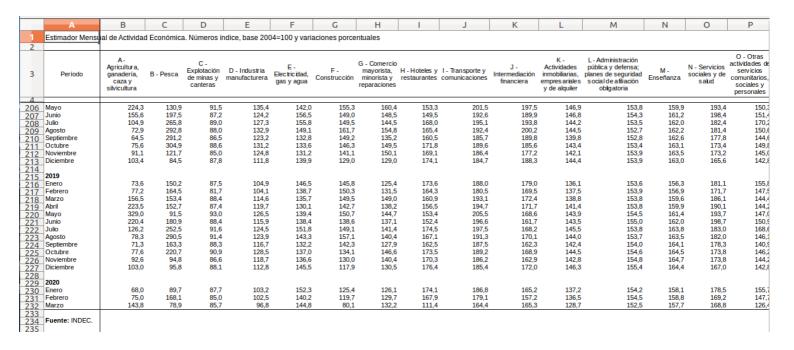
Next, each of the previous datasets will be introduced.

### **Economics**

There are two datasets in this category. Both of them shows the evolution of EMAE estimator ("Mensual Estimator of Economic Activity" in English). The first dataset, shows the evolution of this estimator, related to economic activity of year 2004:

206 207 208 209 210 211	Abril Mayo Junio Julio	153,3 159,6	0,1	147,2	-3,2		
207 208 209 210 211	Junio		4.7		-3,2	147,8	-0,6
208 209 210 211		140.0	-4,7	143,7	-2,4	146,8	-0,6
209 210 211	Julio	149,0	-6,5	142,2	-1,0	145,9	-0,6
210 211	Julio	145,6	-2,9	142,9	0,5	145,0	-0,6
211	Agosto	146,2	-1,8	145,9	2,1	144,2	-0,5
	Septiembre	137,8	-6,3	143,3	-1,8	143,6	-0,4
212	Octubre	143,0	-4,0	144,2	0,7	143,1	-0,3
	Noviembre	140,9	-7,3	141,9	-1,6	142,7	-0,2
213	Diciembre	136,8	-7,0	142,1	0,1	142,5	-0,2
214							
215	2019						
216	Enero	134,9	-5,7	143,4	0,9	142,4	-0,1
217	Febrero	132,4	-4,7	143,4	0,0	142,4	0,0
218	Marzo	144,8	-7,0	142,0	-1,0	142,4	0,0
219	Abril	150,7	-1,7	142,3	0,2	142,5	0,0
220	Mayo	162,8	2,0	142,3	0,0	142,5	0,0
221	Junio	148,4	-0,4	141,7	-0,4	142,5	0,0
222	Julio	146,2	0,4	144,7	2,1	142,4	-0,1
223	Agosto	140,9	-3,6	144,0	-0,4	142,2	-0,1
224	Septiembre	134,9	-2,1	141,6	-1,7	142,0	-0,2
225	Octubre	141,6	-0,9	143,9	1,6	141,7	-0,2
226	Noviembre	137,9	-2,2	141,3	-1,8	141,3	-0,3
227	Diciembre	136,4	-0,3	141,1	-0,1	140,8	-0,3
228							
229	2020						
230	Enero	132,3	-1,9	141,1	0,0	140,3	-0,3
231	Febrero	129,2	-2,4	139,2	-1,3	139,8	-0,4
232	Marzo	128,2	-11,5	125,6	-9,8	139,3	-0,3
233							
234	Fuente: INDEC.						

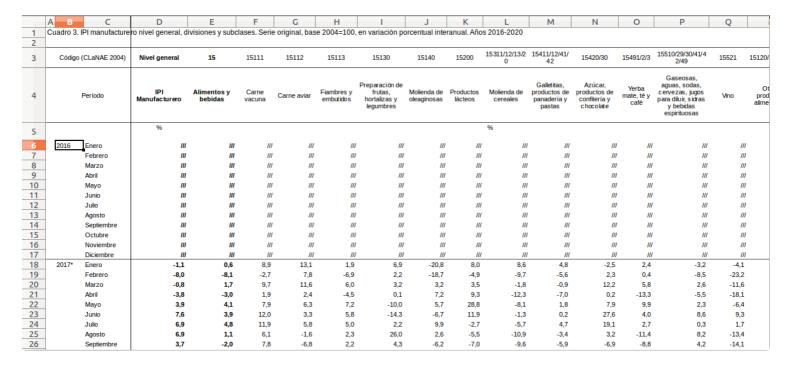
The second one, shows the same estimator for different activities:



This two datasets will help us to compare the estimator at the first month of quarantine vs the estimator at the same month of previous years, viewing which activities were more affected by quarantine.

### **Industrial**

This datasets shows the IPI (Industrial Production Index) of Argentina referred to the last 4 years, and it's also divided by activities:



It will help us to evaluate how much industries were affected on the first month of quarantine.

## **Transport**

This dataset shows the total amount of tickets by month and by train station for every Railway Line on the region of Buenos Aires. This information will be used to compare the first month of quarantine with same month of the previous year and see how much were railway mobility reduced:

