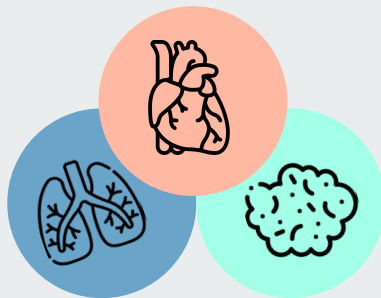


Dinámica de Defunciones en Argentina (2005 - 2019)



G28: Barrios, A.; de Otazua, A.; Gutiérrez, J.; Lumsden, D.
Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires
Laboratorio de Datos - 1C 2021



Esquema

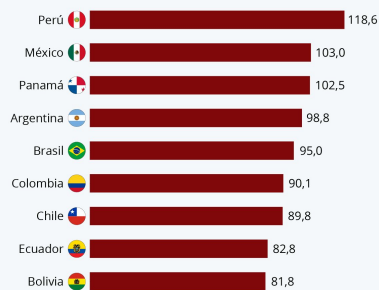
- Motivación
- Definición de objetivos
- Obtención de datos
- Dataset
- Análisis exploratorio
- Predicciones 2019
- Conclusiones
- Propuestas a futuro

Motivación



La tasa de mortalidad del coronavirus en Latinoamérica

Países con el mayor número de muertes por COVID-19 cada 100.000 habitantes*

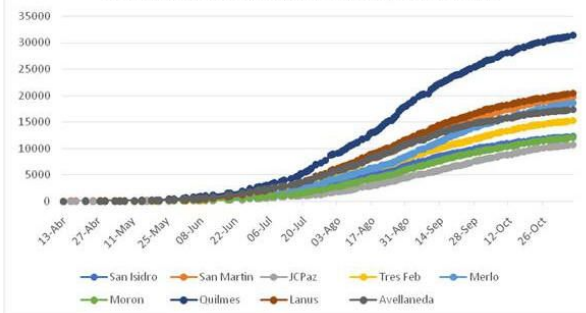


* Datos del 8 de enero de 2021 a las 11h CET.
Fuente: Universidad Johns Hopkins

Cantidad de casos nuevos de COVID en Argentina
(media móvil 3 días)



Casos acumulados en algunos municipios del AMBA



Motivación

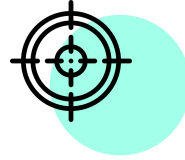


Definición de Objetivos





Definición de Objetivos



- Detectar si existen diferencias entre categorías (provincia, sexo, etc)
- Predecir cantidad de defunciones totales y por causa para el año 2019

Obtención de datos



- Dirección de Estadísticas e Información de la Salud (DEIS)

www.argentina.gob.ar/salud/deis/datos/defunciones

Dataset



Año



2005 - 2019



Causa de muerte



>1450 causas



Ubicación



26 lugares



Grupo de edad



Rangos de 5 años (ej.: 50 a 54 años)



Sexo



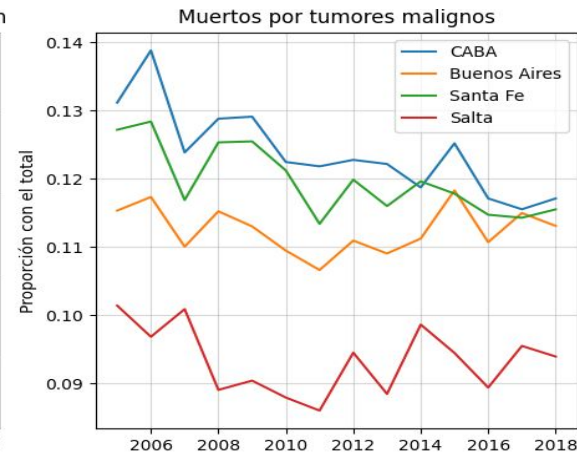
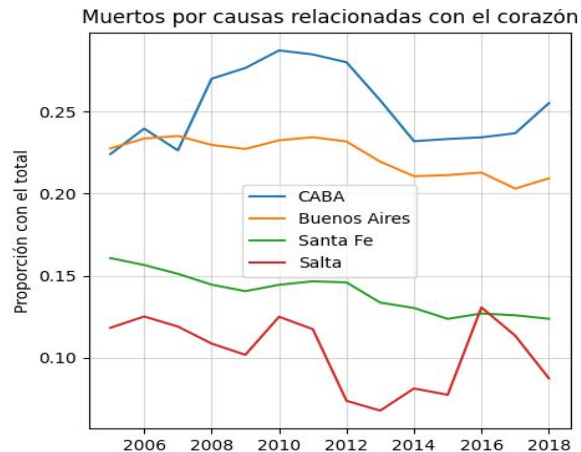
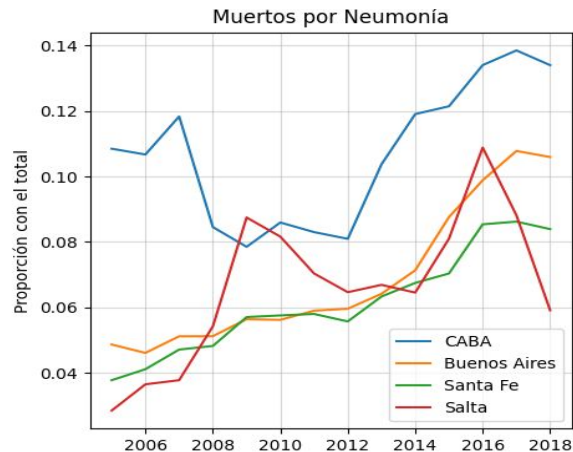
'Varón', 'Mujer', 'Sin especificar'

www.argentina.gob.ar/sites/default/files/2021/03/descdef1.xlsx

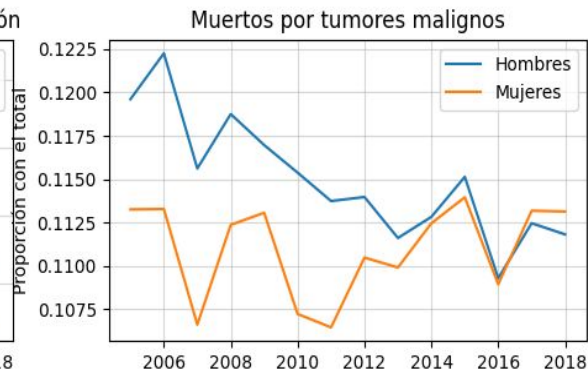
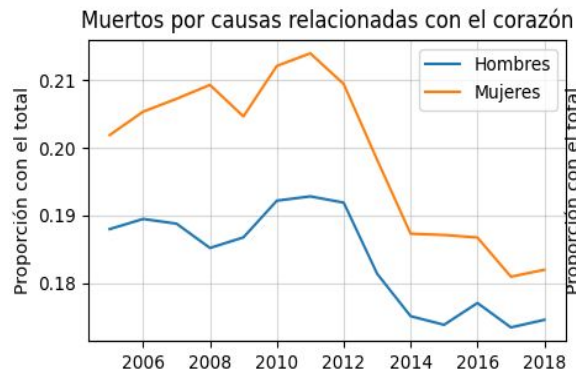
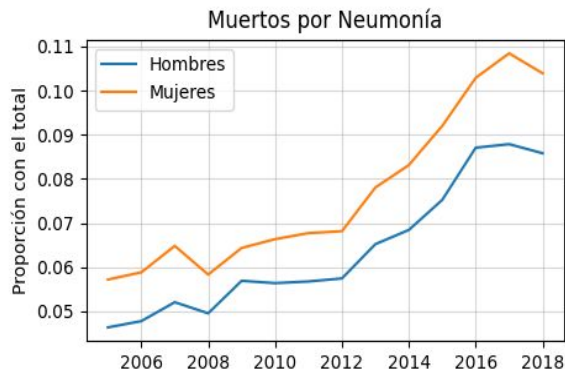
Breve análisis exploratorio entre categorías



Comparación entre provincias



Comparación entre sexos (todo Argentina)



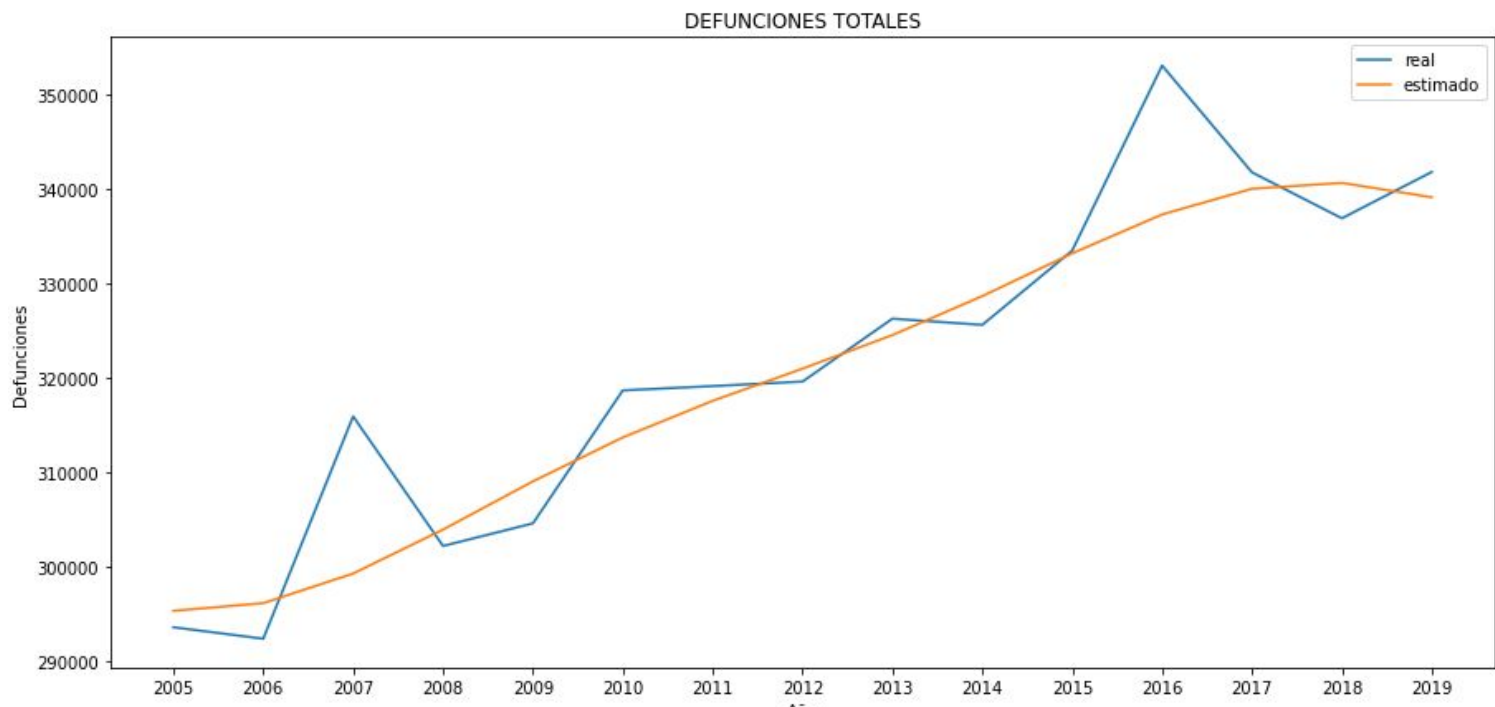
Predicción 2019





Predicción de defunciones totales

R2_score:
0,8619





Datos, muchos datos para mi SVR...

- 1455 Causas , 18 Edades , 3 Sexos ...
 - 2005 - 2018 = 675264 Filas
 - 2019 = 49820 Filas
- SVR es $O(n^2)$
- 10000 samples o menos
- Dividir y conquistar



Dividir

- Si tenemos menos de 5 filas
 - El mejor estimador es repetir el último año.
- Si al estandarizar da todo 0 no hacemos estimador
 - El mejor estimador es repetir el último año.
- Construimos un estimador por Provincia y Causa (SVR)
 - Optimizamos lo mejor posibles el estimador haciendo un GridSearch
- Persistimos cada estimador
 - Nos permite apagar la máquina
 - Podemos corregir un solo estimador



Conquistar

- 37830 Estimadores
 - Repetir 11393
 - SVR
 - Lineal 7530
 - RBF 3864



Resultados

- Causas Cardiacas
 - Insuficiencia cardíaca
 - Infarto agudo del miocardio
 - Cardiomiopatía
 - Enfermedad cardíaca hipertensiva
 - Enfermedad isquémica crónica del corazón
 - Otras arritmias cardiacas
 - Complicaciones y descripciones mal definidas de enfermedad cardíaca

R2_Score = 0.8920



Resultados

- Causas Tumorales
 - Tumor maligno del estómago
 - Tumor maligno del colon
 - Tumor maligno del páncreas
 - Tumor maligno de los bronquios y del pulmón
 - Tumor maligno de la mama
 - Tumor maligno del cuello del útero
 - Tumor maligno de la próstata
 - Tumor maligno de sitios no especificados

R2_Score = 0.9137



Resultados

- Causas Neumonía
 - Neumonía viral no clasificada en otra parte
 - Neumonía debida a *Streptococcus pneumoniae*
 - Neumonía debida a *Haemophilus influenzae*
 - Neumonía bacteriana no clasificada en otra parte
 - Neumonía debida a otros microorganismos infecciosos no clasificados en otra parte
 - Neumonía organismo no especificado

R2_Score = 0.9841



Conclusiones

- La predicción realizada predice muy bien las defunciones durante el 2019, en general como causas tumorales, cardíacas y neumonías.
- Se puede ver que no hay diferencias significativas entre la proporción de muertos por sexo masculino y femenino entre las distintas causas.
- Entre la proporción de muertos de cada provincia es llamativa la diferencia entre entre las mismas categorías (caso CABA vs Salta).



Propuestas a futuro

- Utilizar el algoritmo propuesto para predecir las defunciones para 2020 y comparar con la publicación del DEIS para el 2020, teniendo en cuenta el COVID-19

(Exceso de mortalidad por la pandemia de COVID-19 durante 2020 en la provincia de Buenos Aires, Pesci et. al., Revista argentina de salud pública, suplemento COVID, ju 2021.)

- Hacer un análisis por categorías de muerte en provincias para profundizar el trabajo exploratorio realizado.
- Disponer de otras librerías para utilizar mejor la capacidad de cómputo.

GRACIAS



El fumar es perjudicial para la salud, si tomas no manejes, beber con moderación, si cruzas la calle mira para los dos lados