



# USA SCHOOL SHOOTINGS FORECASTING

Jose Saúl Vega  
Cristian Picón  
Diego Villamizar  
Oscar Sinuco

## USA SCHOOL SHOOTINGS

En los últimos 10 años han sucedido al menos 180 tiroteos en escuelas estadounidenses. Este fenómeno en escuelas parece repetirse año tras año.

# CONTENIDO

- Objetivos.
- Dataset.
- Modelos a usar.
  - ARIMA.
  - Exponential Smoothing.
- Pre-análisis de datos.
  - Registro de muertes con los años.
  - Muertes por mes.
- Aplicación de ARIMA.
- Aplicación de Exponential Smoothing.
- Conclusiones

# Nuestros objetivos



## Manejo de datos

Discriminar los datos relevantes de los que no lo son, organización de los datos de utilidad para el buen tratamiento de los mismos.



## Predicciones

Se propone predecir la posible cantidad de muertos en próximos meses y las posibles épocas del año en las que sucederán estos eventos para poder prevenirlos.



## Análisis

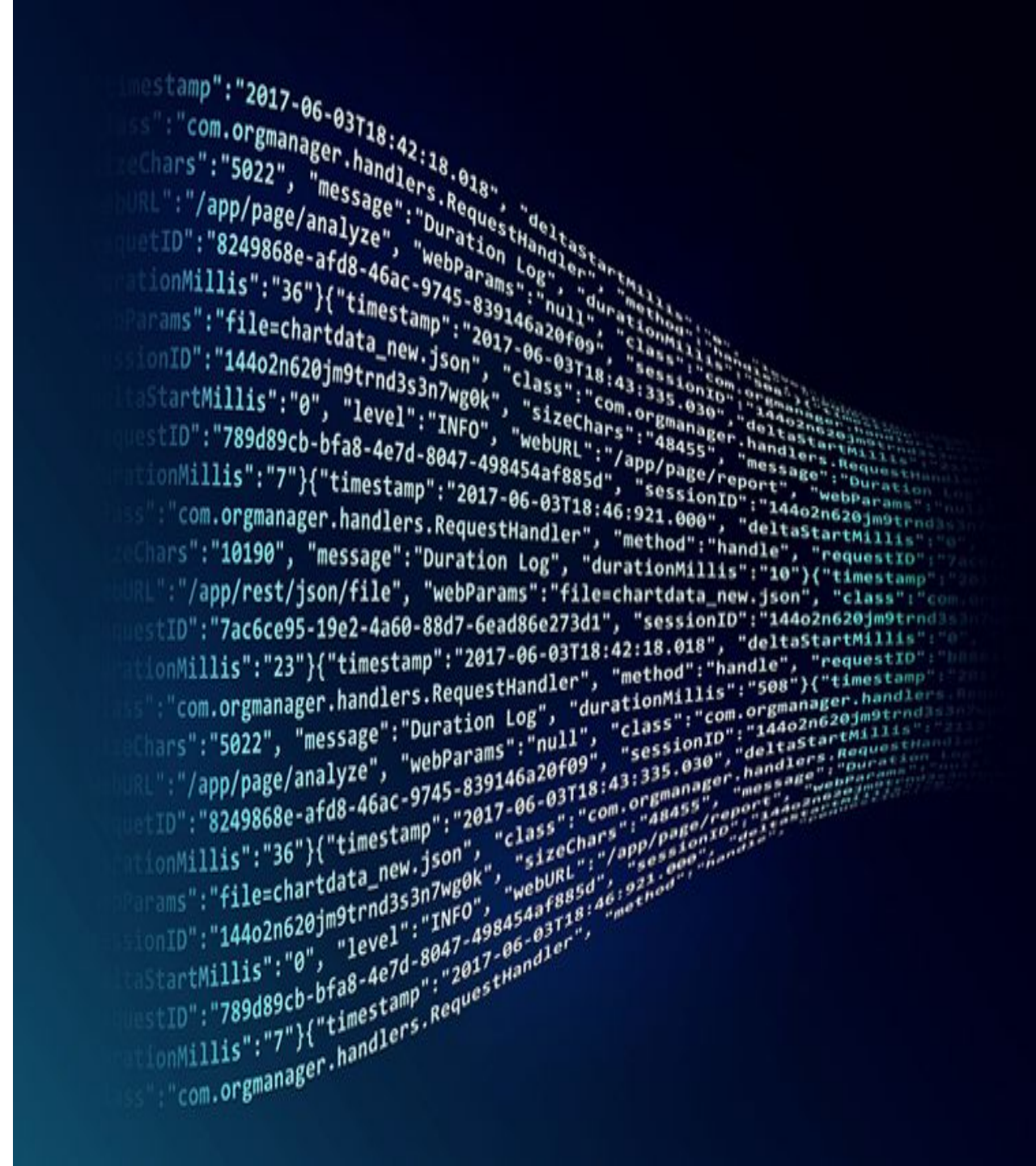
Con respecto a los datos obtenidos con las predicciones se deberá analizar el sentido de los resultados y la coherencia de los mismos.

# Nuestros datos

Contamos con datos que van desde eventos sucedidos en 1990 hasta inicios del año 2020.

Entre los datos tenemos locación, número de víctimas, heridos y fecha exacta del hecho.

[https://www.kaggle.com/ecodan/school-shootings-us-1990present#pah\\_wikp\\_combo.csv](https://www.kaggle.com/ecodan/school-shootings-us-1990present#pah_wikp_combo.csv)





## ARIMA

Permite describir un valor como una función lineal de datos anteriores, además, puede incluir un componente cíclico o estacional.



## Simple Exponential Smoothing

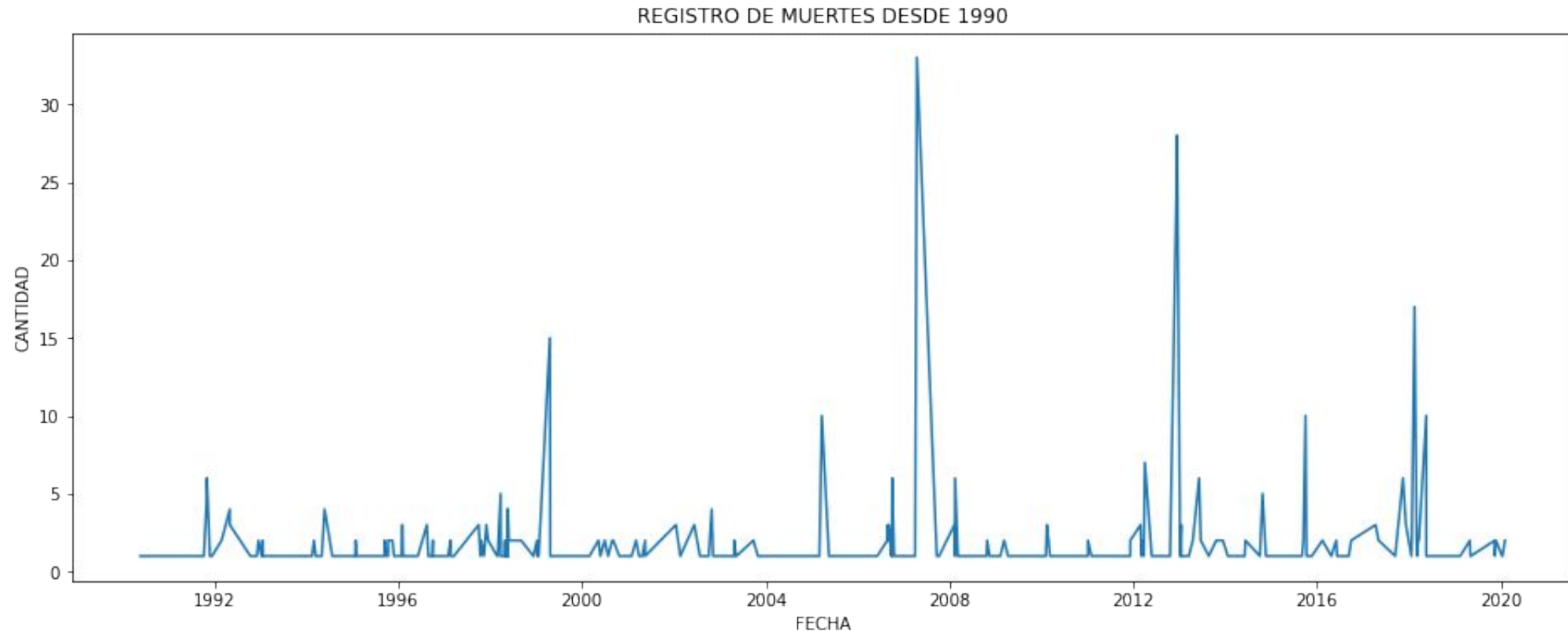
Calcula un promedio ponderado de observaciones pasadas, donde los pesos decaen exponencialmente a medida que las observaciones envejecen.

## Pre-análisis de datos

Antes de aplicar un método de predicción sobre las observaciones reales de nuestro dataset, se procede a realizar algunas anotaciones y evidenciar el comportamiento de los datos de una de manera visual.

## Muertes con los años

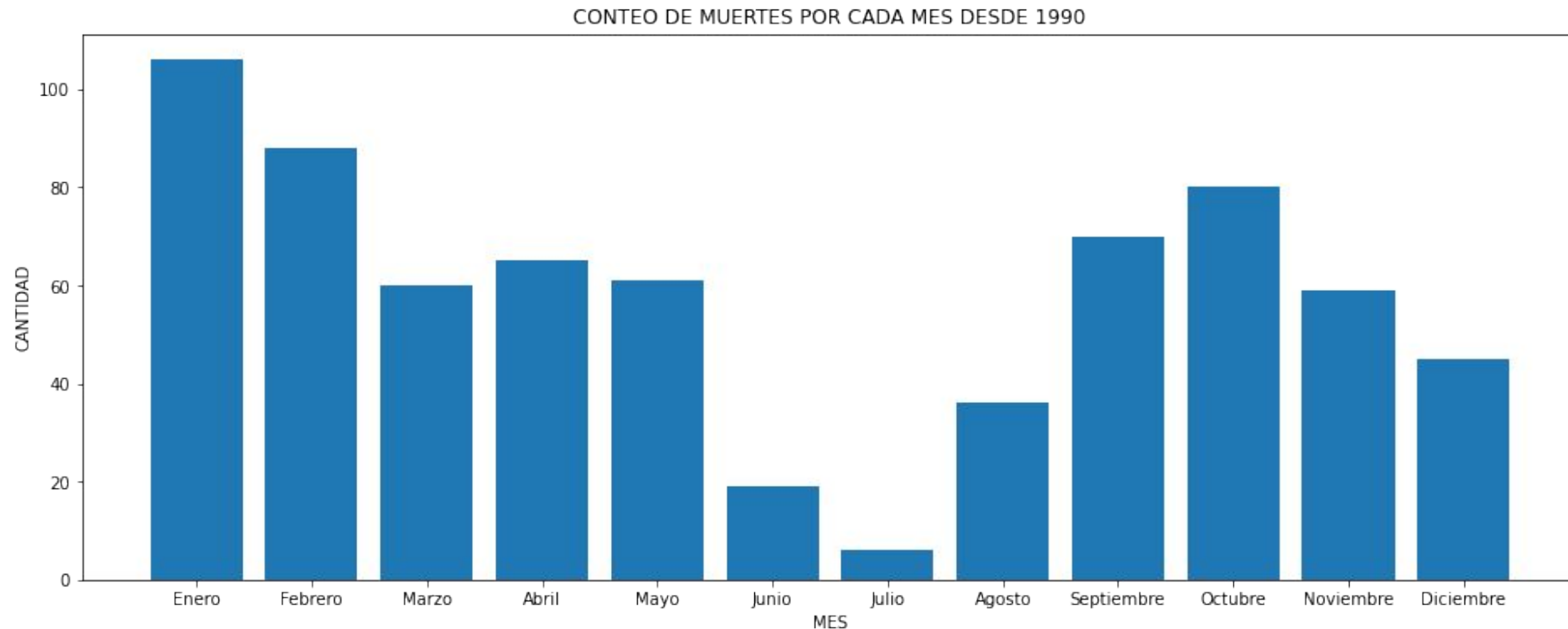
Realizamos una gráfica de fecha vs cantidad de muertes. Como se puede observar existen picos extremos en diversos puntos de la historia.





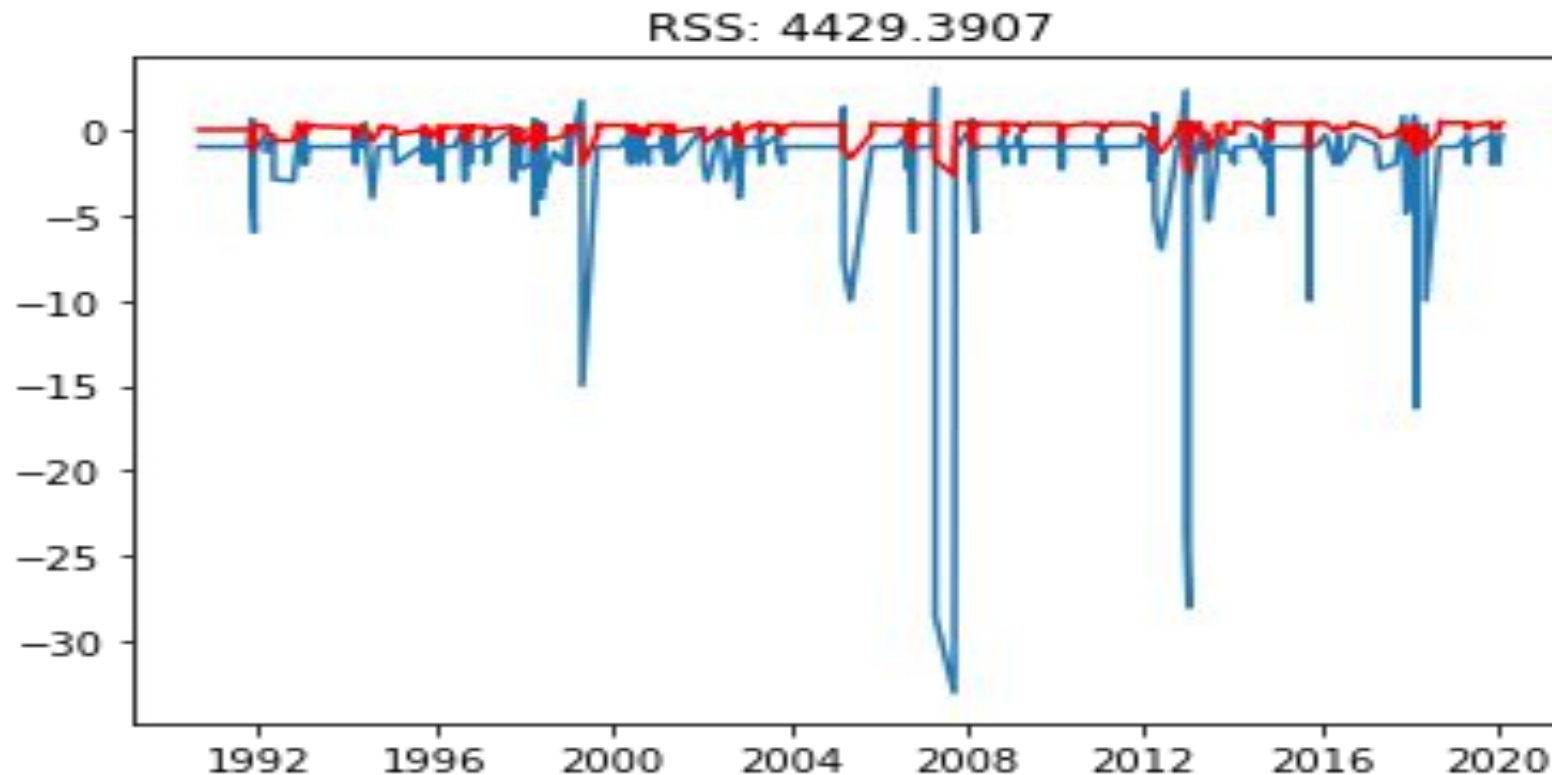
## Conteo de muertes por cada mes

Se suman todas la muertes que se han dado en cada uno de los meses del año a lo largo de la historia. Nótese las pocas que hay a mitad de año, cuando las escuelas están de vacaciones.



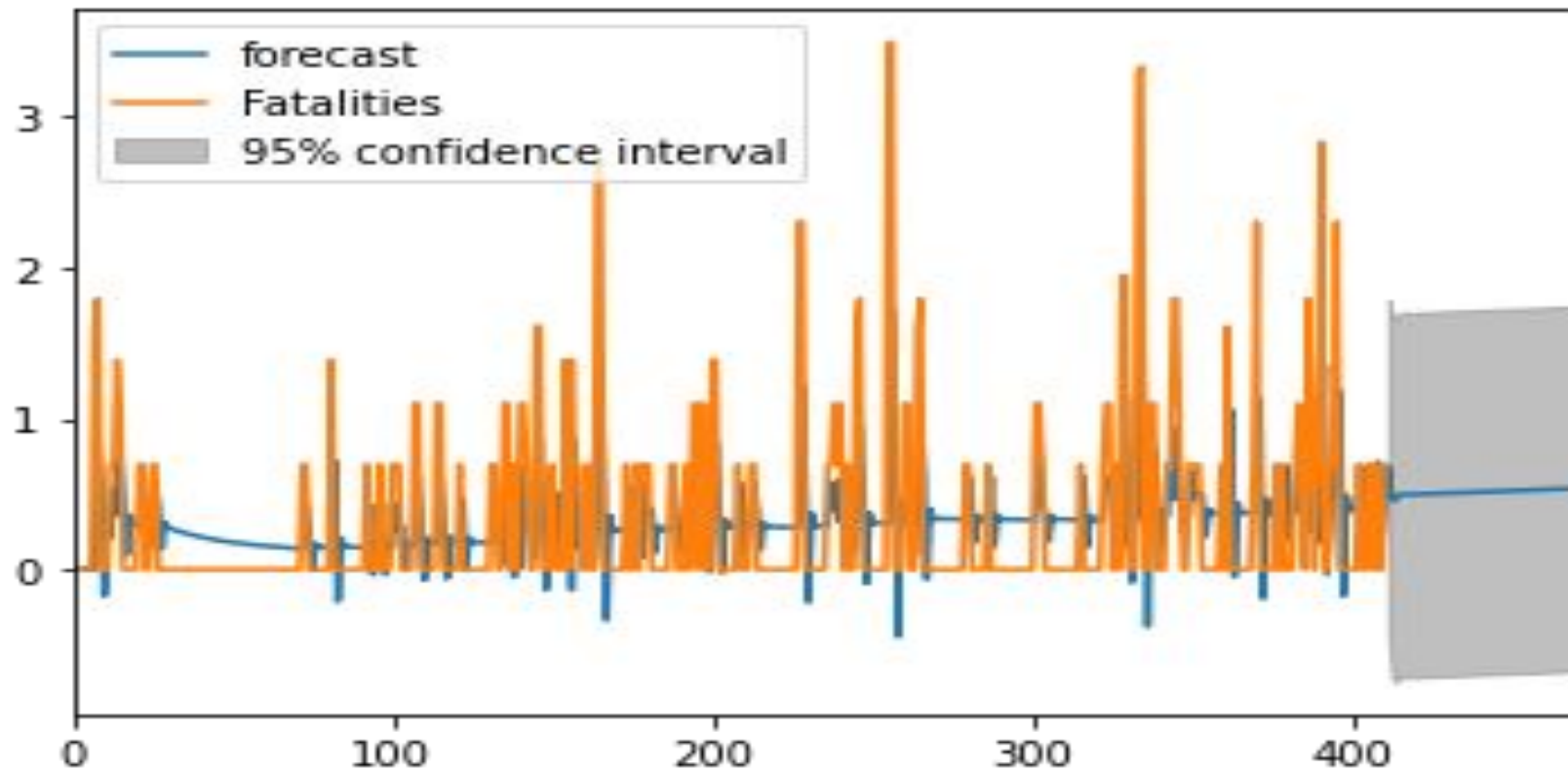
## Aplicación de ARIMA

Los picos en algunos puntos del gráfico se pueden explicar por los picos en los registros de muertes, pero se puede notar que los valores entrenados tienen un mismo comportamiento que los valores originales.



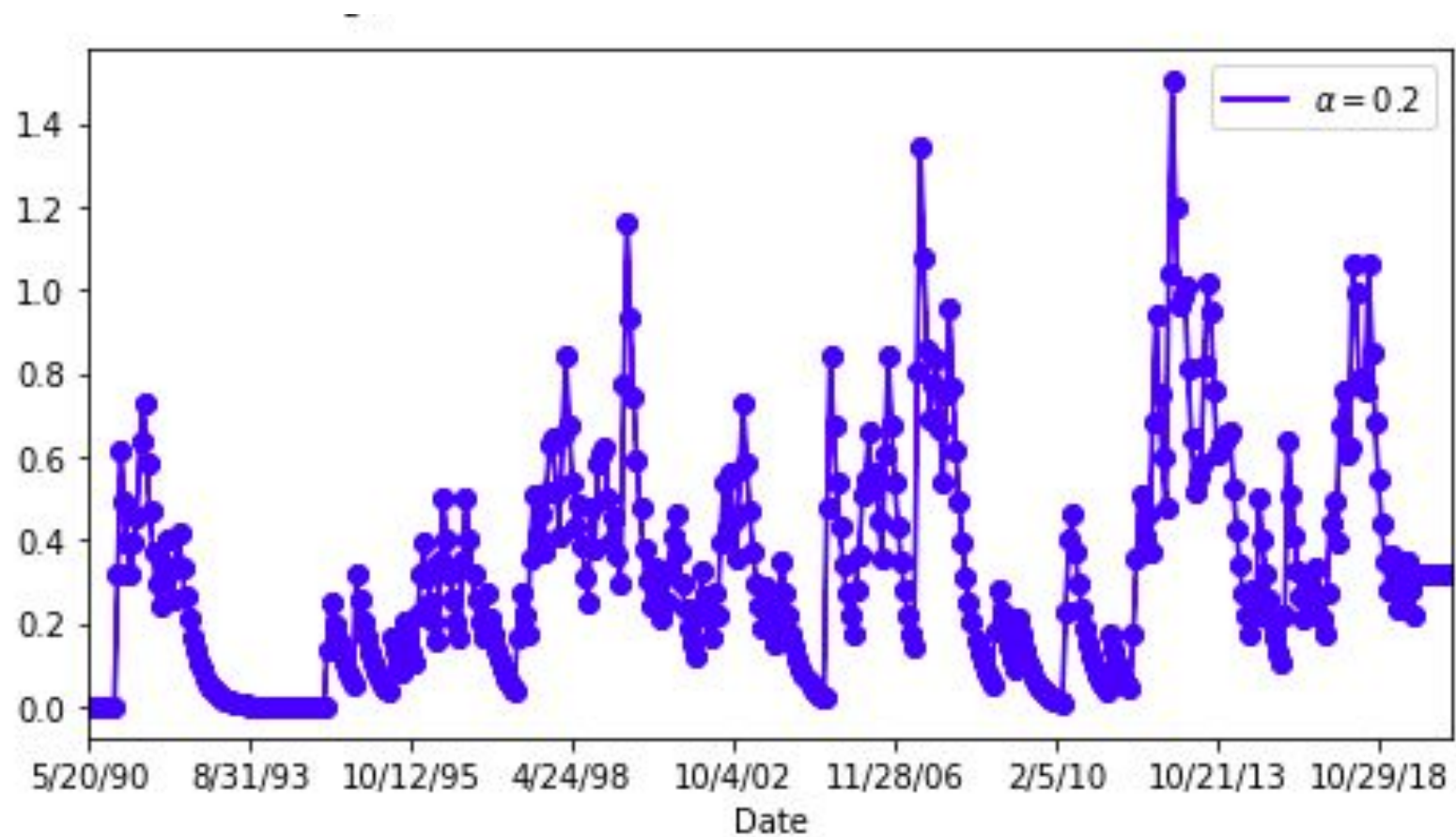
## Aplicación de ARIMA

Aunque las predicciones se ven afectadas por la presencia de las situaciones extremas, podemos lograr una predicción de algunos meses en el año actual, que nos indica la subsiguiente subida de muertes en instituciones educativas.



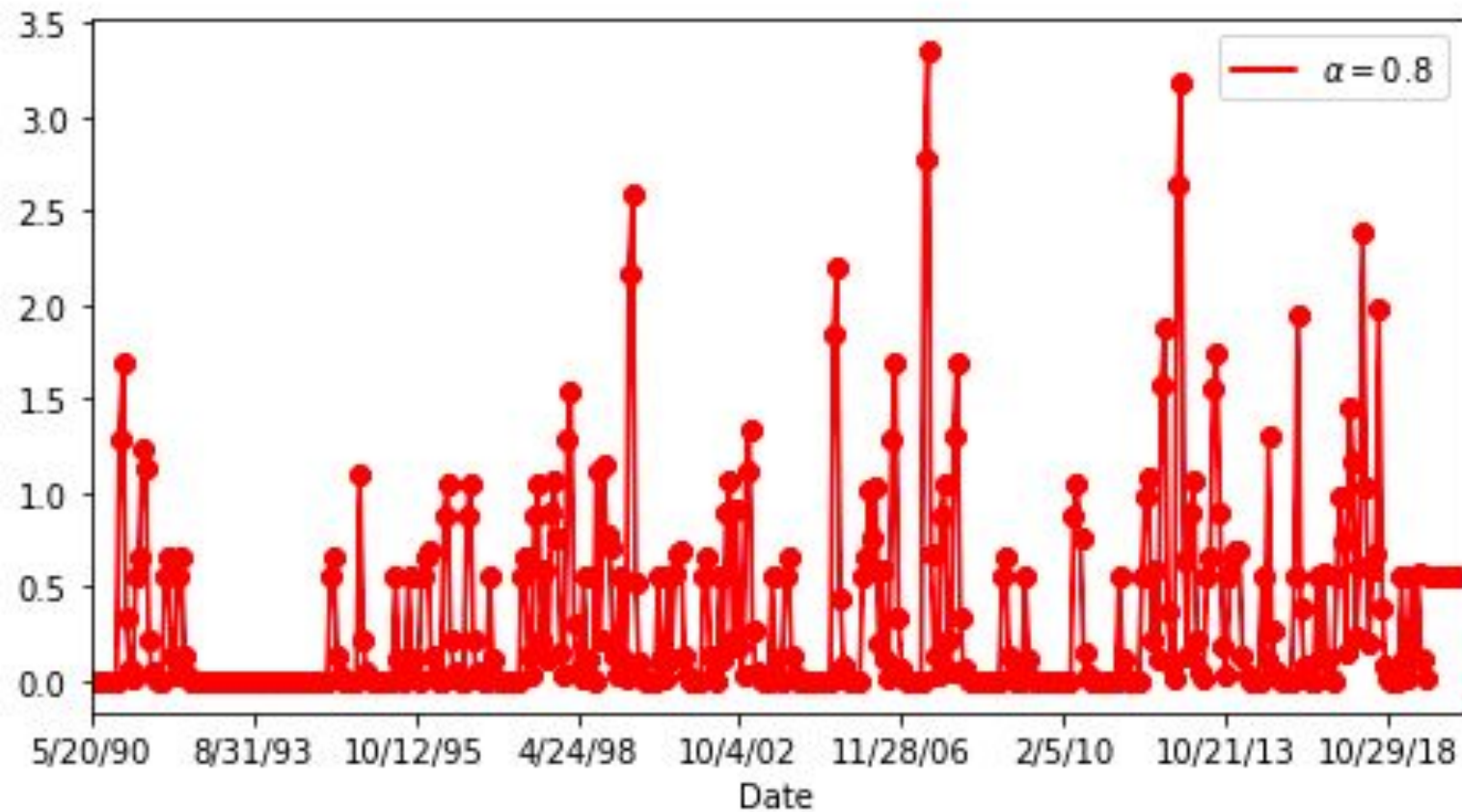
# Aplicación de Exponential Smoothing

Observaciones menos recientes con mayor peso.



# Aplicación de Exponential Smoothing

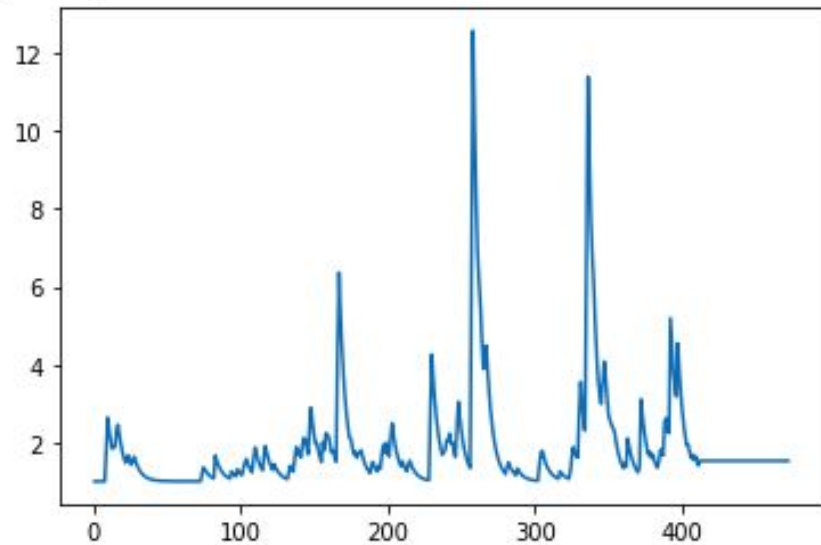
Observaciones más recientes con mayor peso.



## Aplicación de Exponential Smoothing

Se observa que al darle menor peso a las observaciones más recientes se disminuye un poco los cambios abruptos en las observaciones y se infiere que en nuestro caso los resultados mejoran cuando los datos más recientes tienen una mayor importancia o peso.

# Predicciones - Exponential Smoothing



```
413 1.526158
414 1.526158
415 1.526158
416 1.526158
417 1.526158
418 1.526158
419 1.526158
420 1.526158
421 1.526158
422 1.526158
423 1.526158
424 1.526158
425 1.526158
426 1.526158
427 1.526158
428 1.526158
429 1.526158
430 1.526158
431 1.526158
432 1.526158
433 1.526158
434 1.526158
435 1.526158
436 1.526158
437 1.526158
438 1.526158
439 1.526158
440 1.526158
441 1.526158
```

```
[105] 415 1.806658
      416 1.806658
      417 1.806658
      418 1.806658
      419 1.806658
      420 1.806658
      421 1.806658
      422 1.806658
      423 1.806658
      424 1.806658
      425 1.806658
      426 1.806658
      427 1.806658
      428 1.806658
      429 1.806658
      430 1.806658
      431 1.806658
      432 1.806658
      433 1.806658
      434 1.806658
      435 1.806658
      436 1.806658
      437 1.806658
      438 1.806658
      439 1.806658
      440 1.806658
      441 1.806658
      442 1.806658
      443 1.806658
      444 1.806658
      445 1.806658
```

## Conclusiones - resultados

Con la aplicación de ambos métodos de predicción se obtienen resultados similares, la predicción se vuelve constante para muchas de las fechas futuras, lo que no nos permite dar una predicción acertada a fechas más lejanas a 1 mes.

Los resultados obtenidos no permiten dar una predicción confiable ya que el comportamiento de los datos predichos por medio de los métodos no asemejan las observaciones reales.





Gracias

