



# FORECASTING CRIME

A Univariate Time Series Analysis  
Nishath Fatima

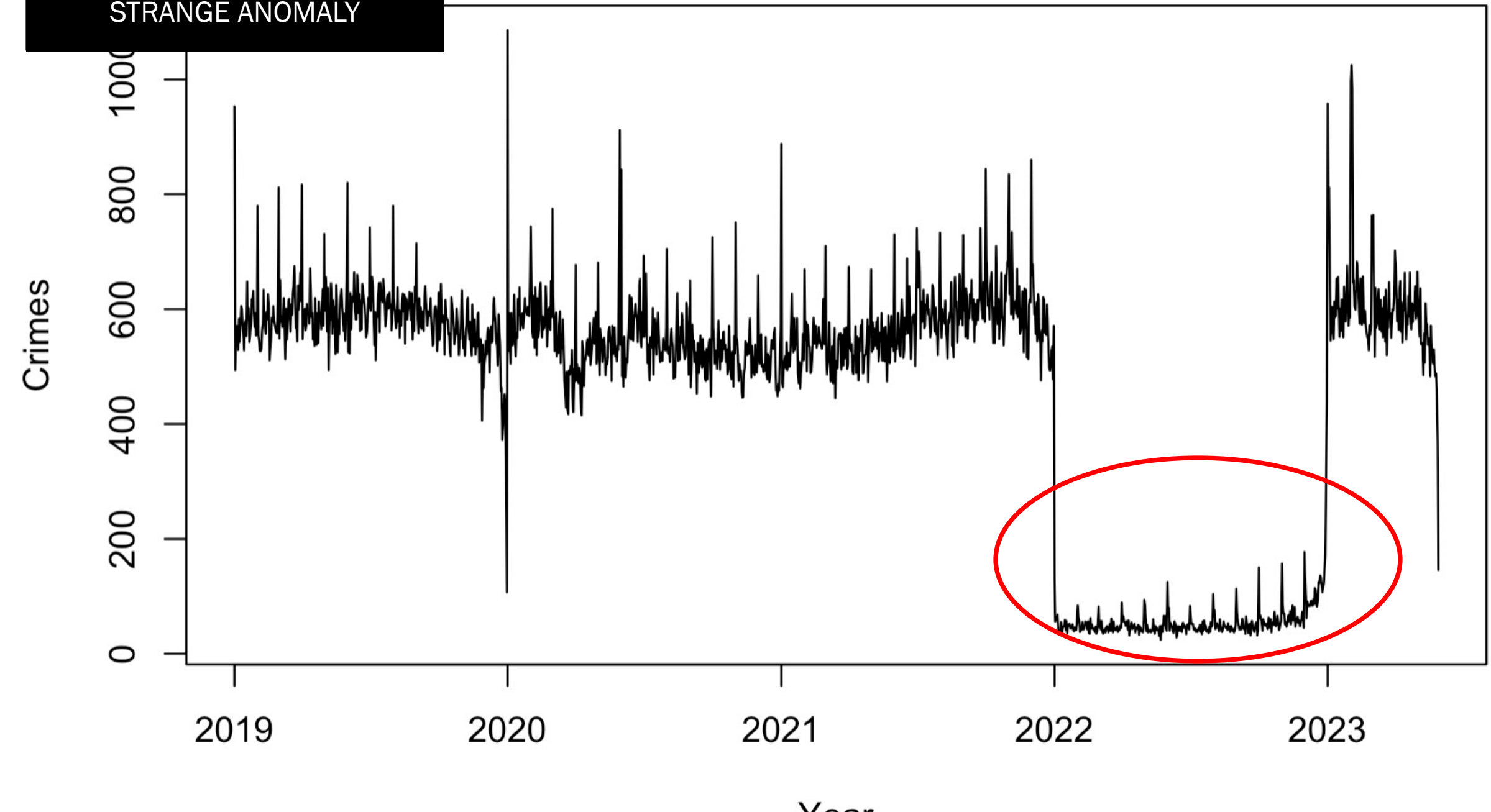




# DATASET: CRIME FROM 2019-2023

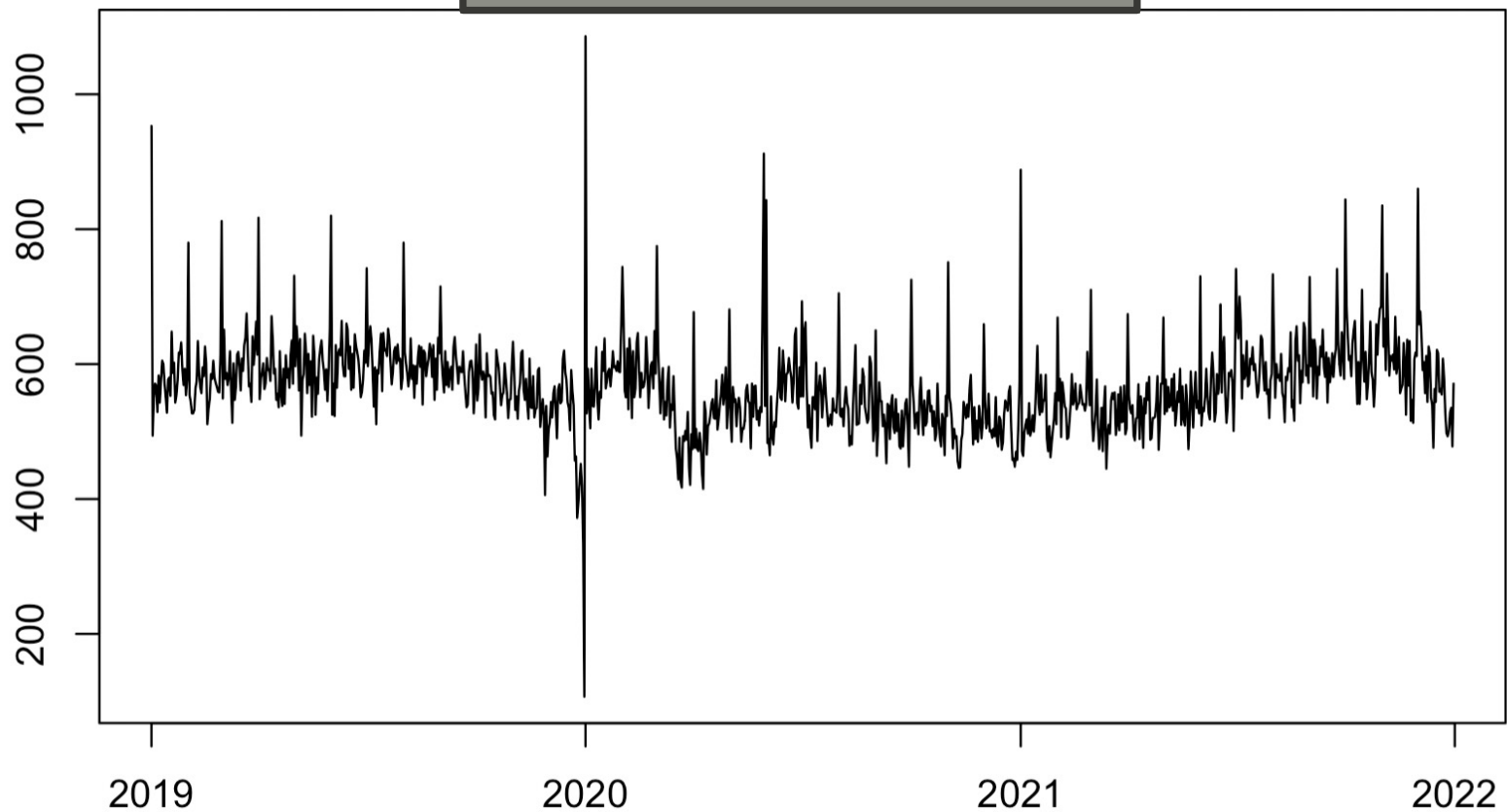
- The dataset reflects daily incidents of crime in the City of Los Angeles dating back to 2019 by neighborhood.
- According to [data.lacity.org](https://data.lacity.org), the data is transcribed from original crime reports that are typed on paper and therefore there may be *some inaccuracies in the data*.

DATA VISUALIZATION: A  
STRANGE ANOMALY

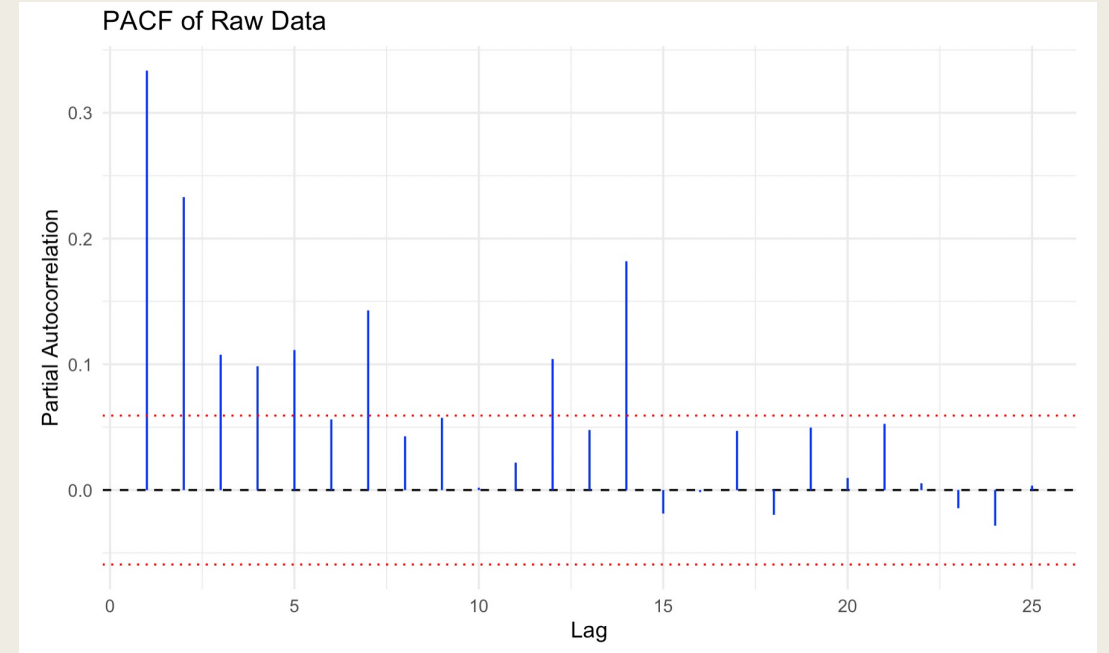
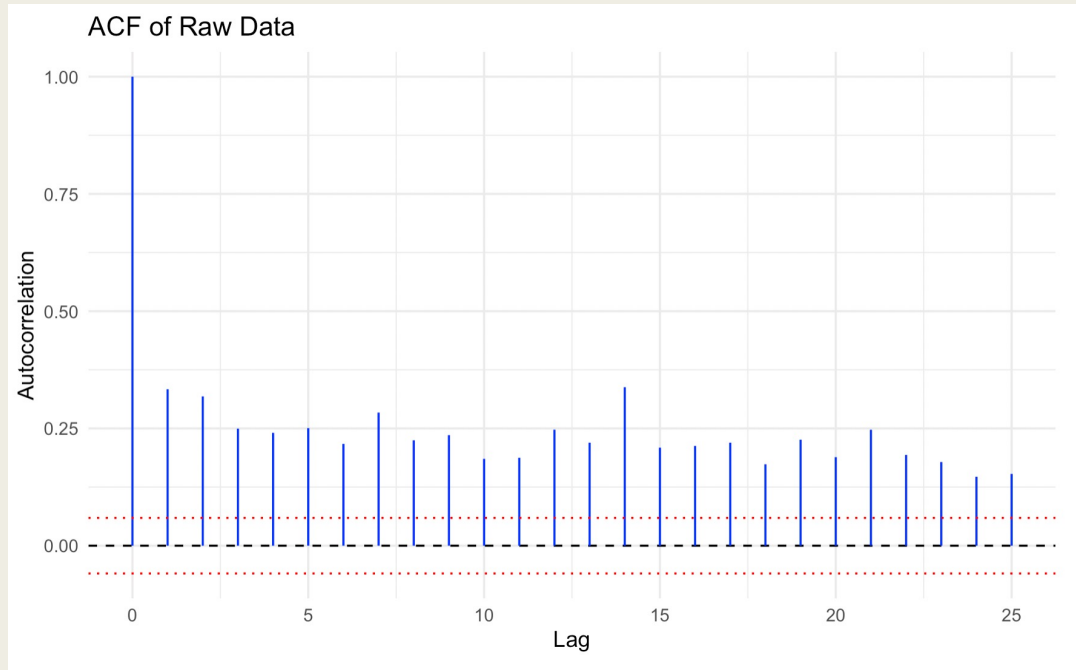


# PREREQUISITES: DATA STRUCTURE AND CONTEXT

Removed 2022-2023

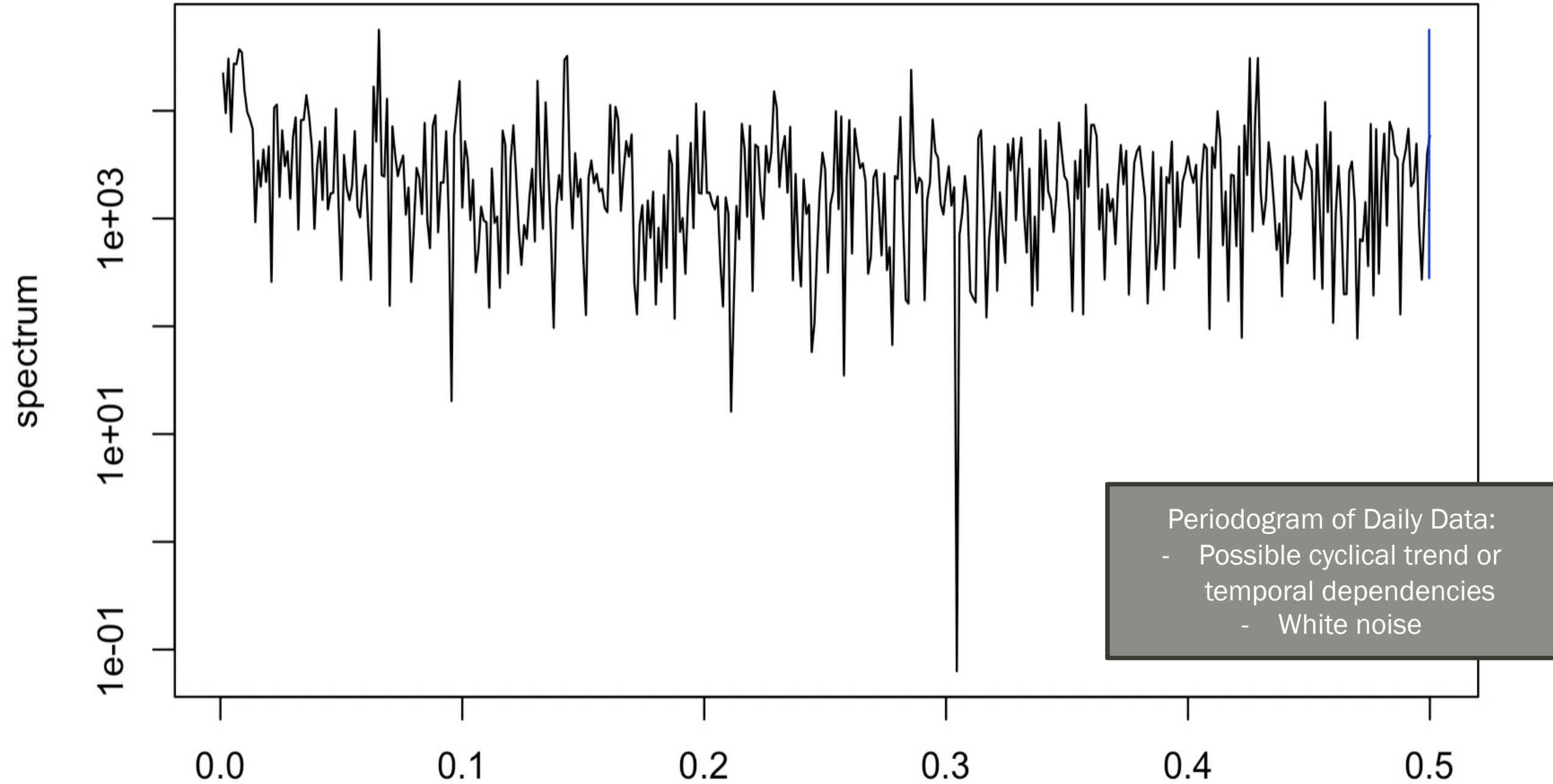


	▲ date ▼	total_crimes ▼
1	2019-01-01	953
2	2019-01-02	494
3	2019-01-03	550
4	2019-01-04	571
5	2019-01-05	568
6	2019-01-06	529
7	2019-01-07	583
8	2019-01-08	543
9	2019-01-09	575
10	2019-01-10	605
11	2019-01-11	600
12	2019-01-12	572
13	2019-01-13	551

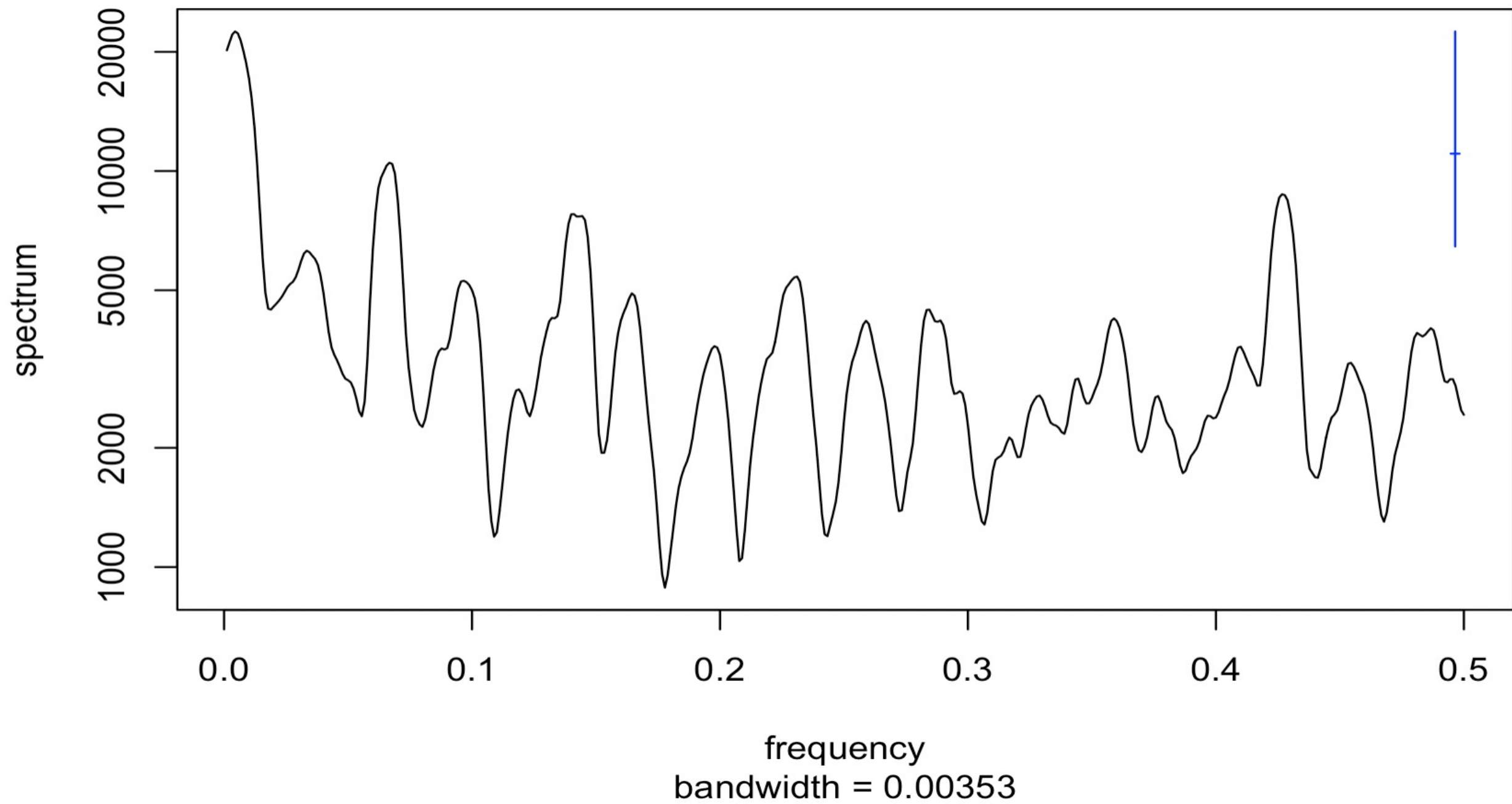


SAMPLE ACF AND PACF

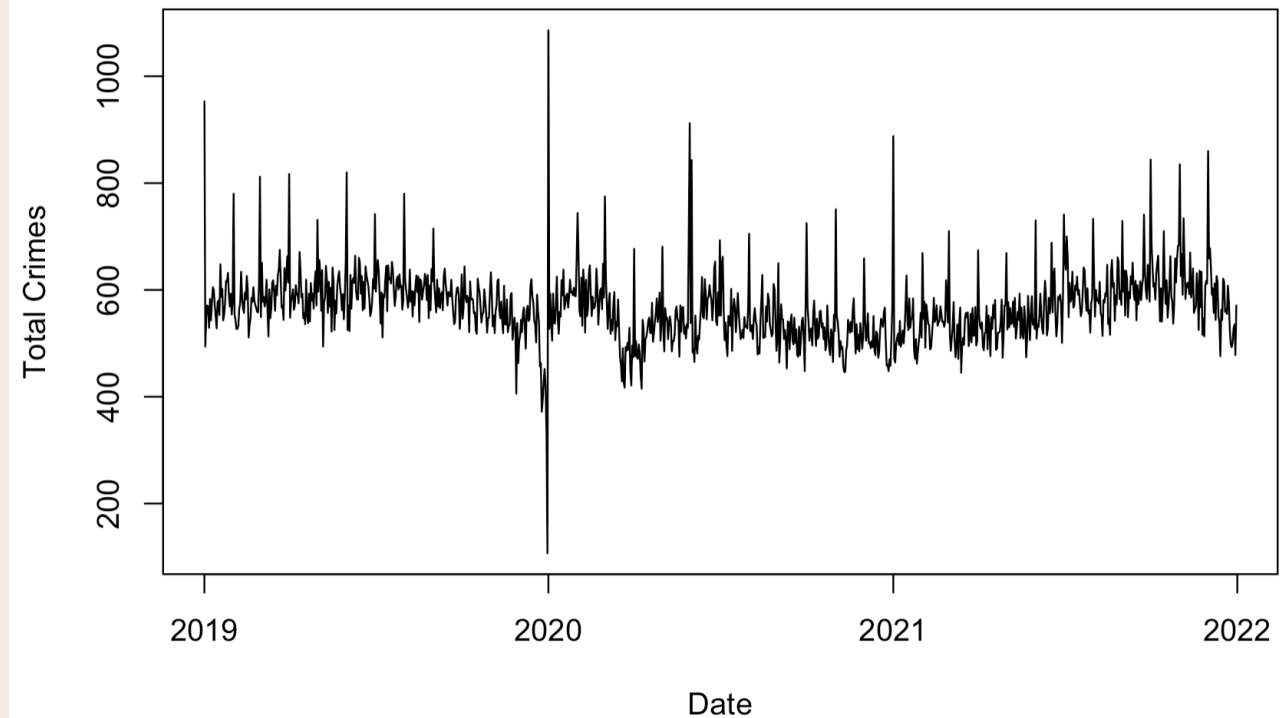
# Periodogram of Raw Data



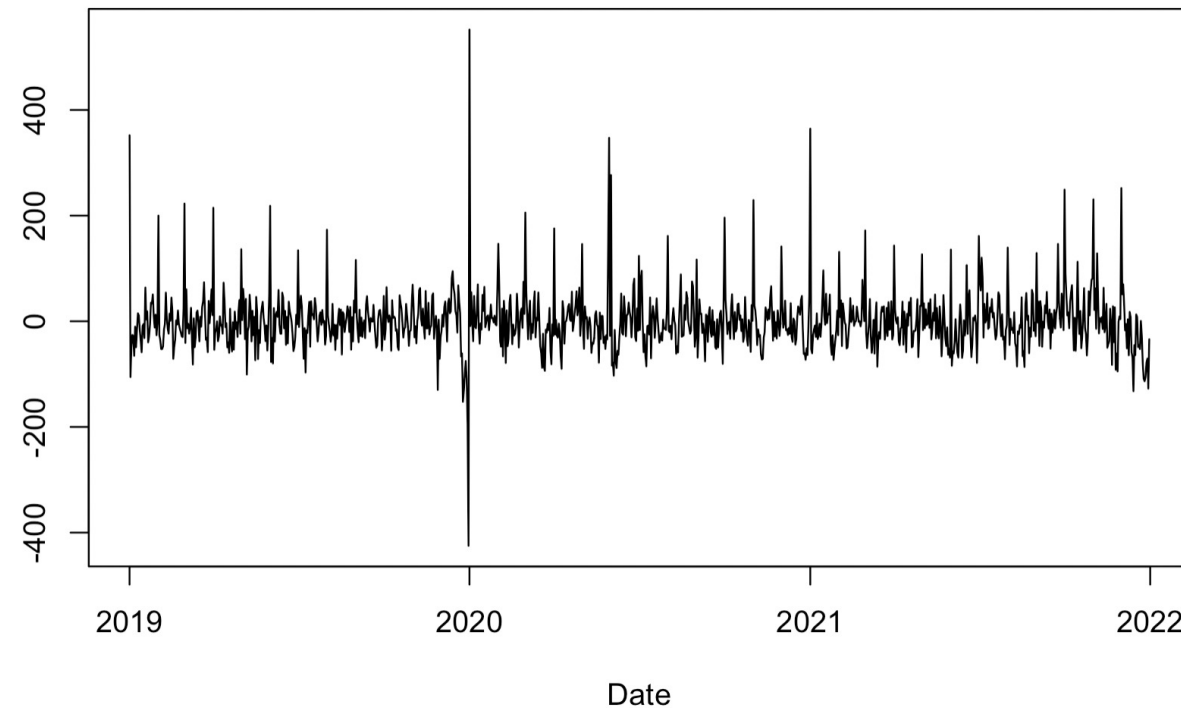
# Smoothed Periodogram



Original Data



Detrended Data

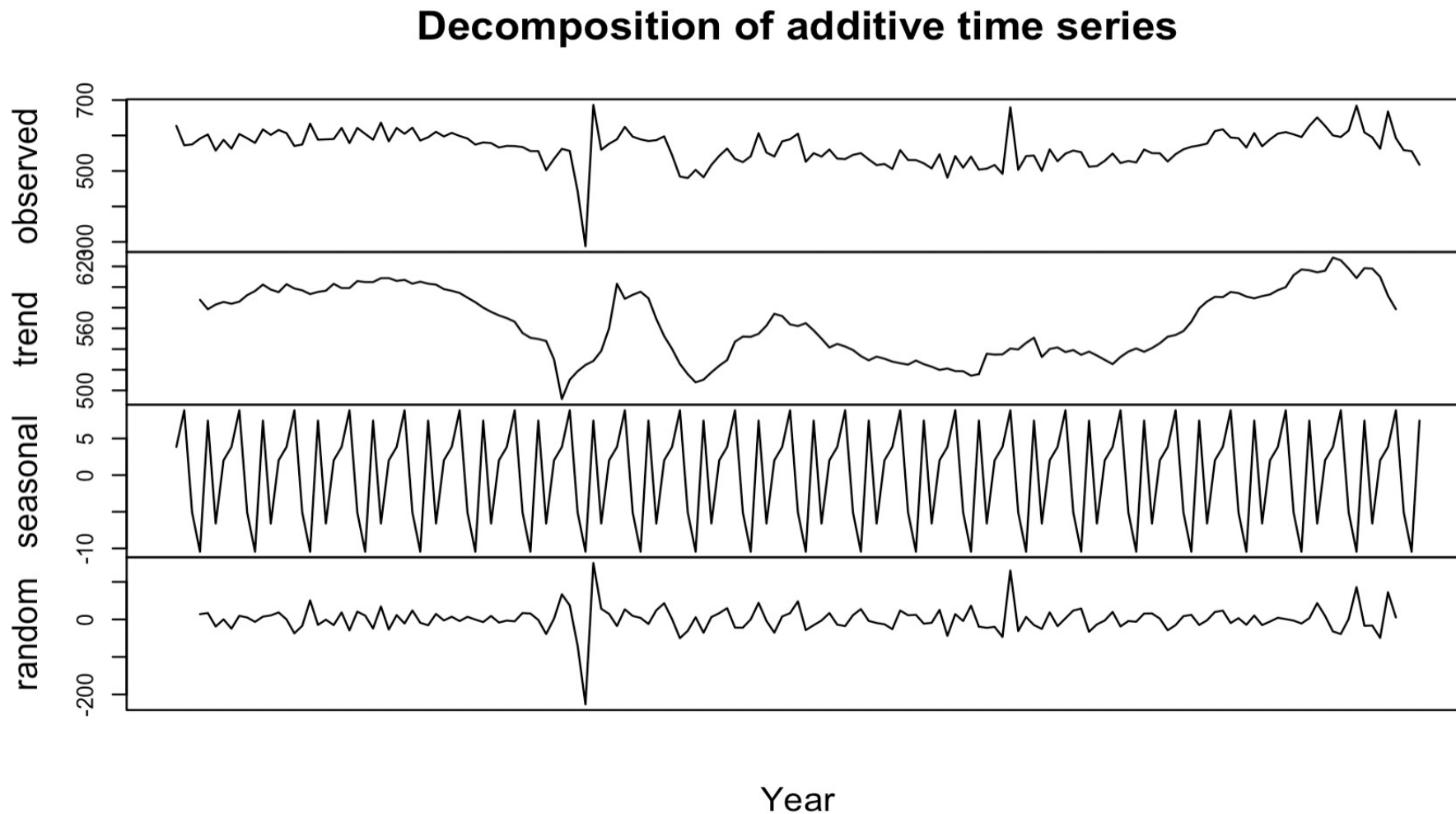


# ORIGINAL VS. DETRENDED DATA

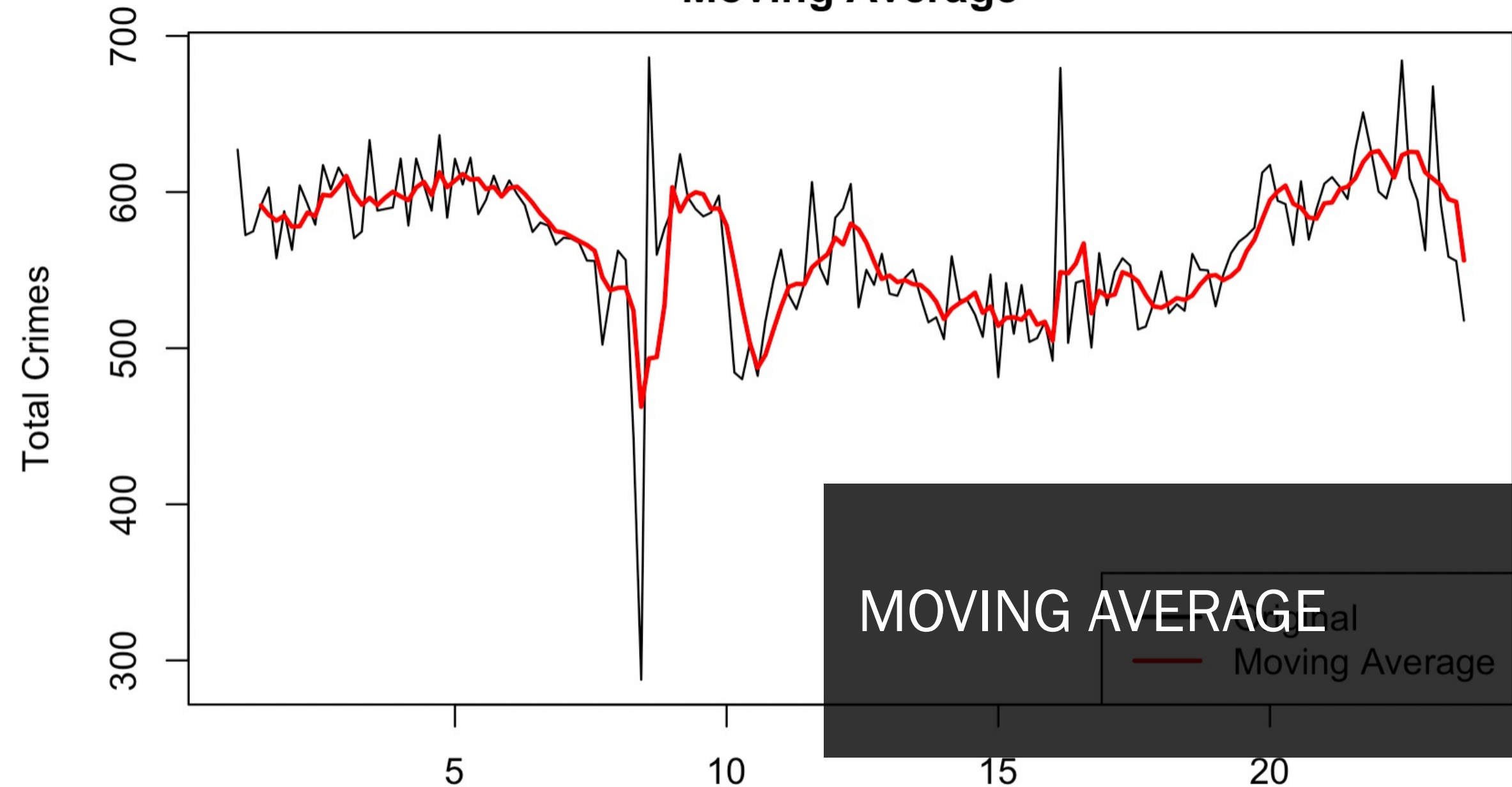
*Note that there was some white noise present, but the most apparent fact is that the data can be aggregated and modeled much better if we could look at weekly cycles and the average*



# DECOMPOSITION WEEKLY AVERAGE

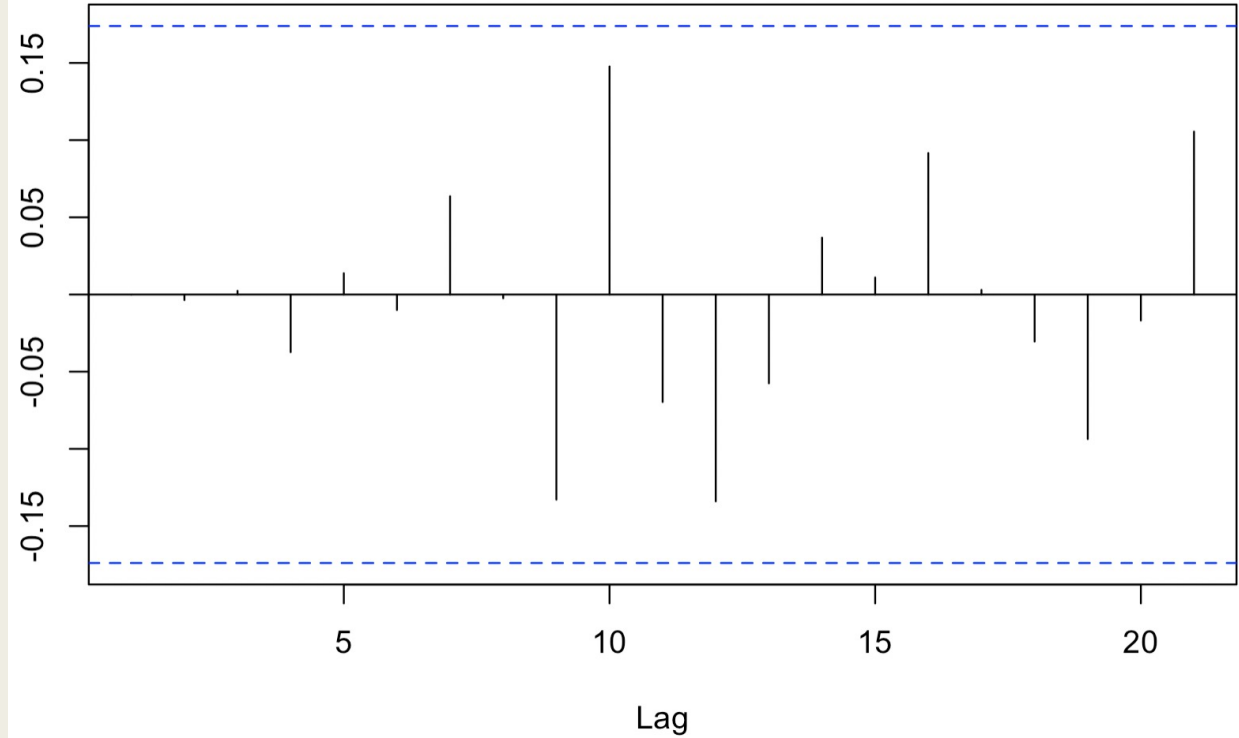
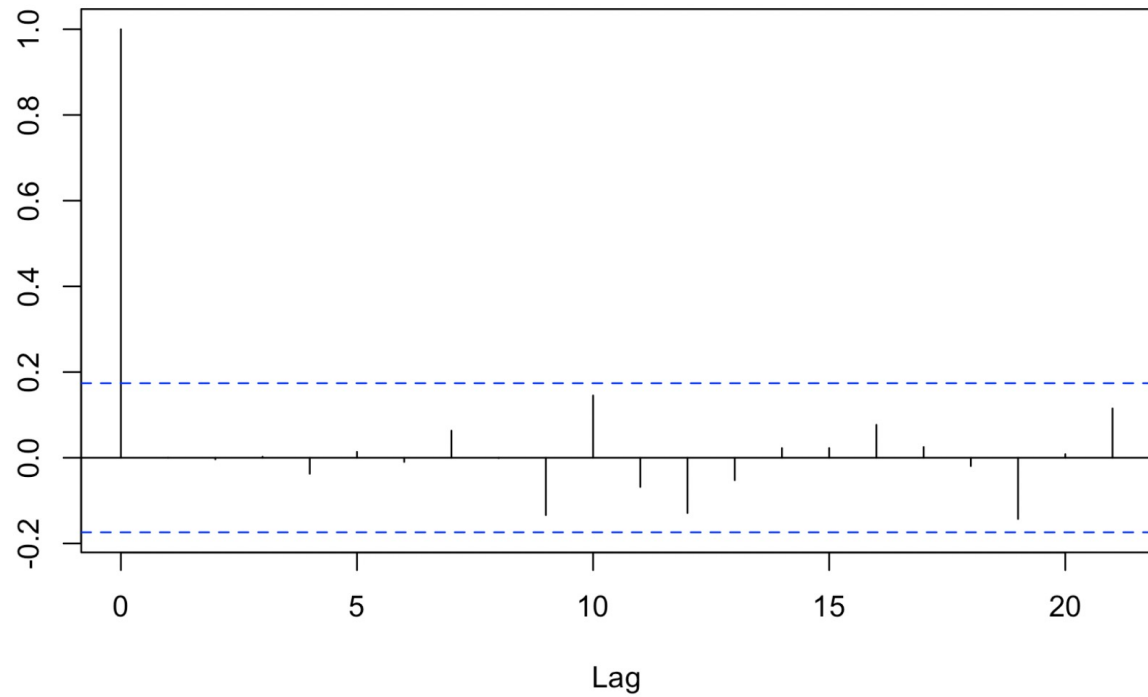


## Moving Average



MOVING AVERAGE

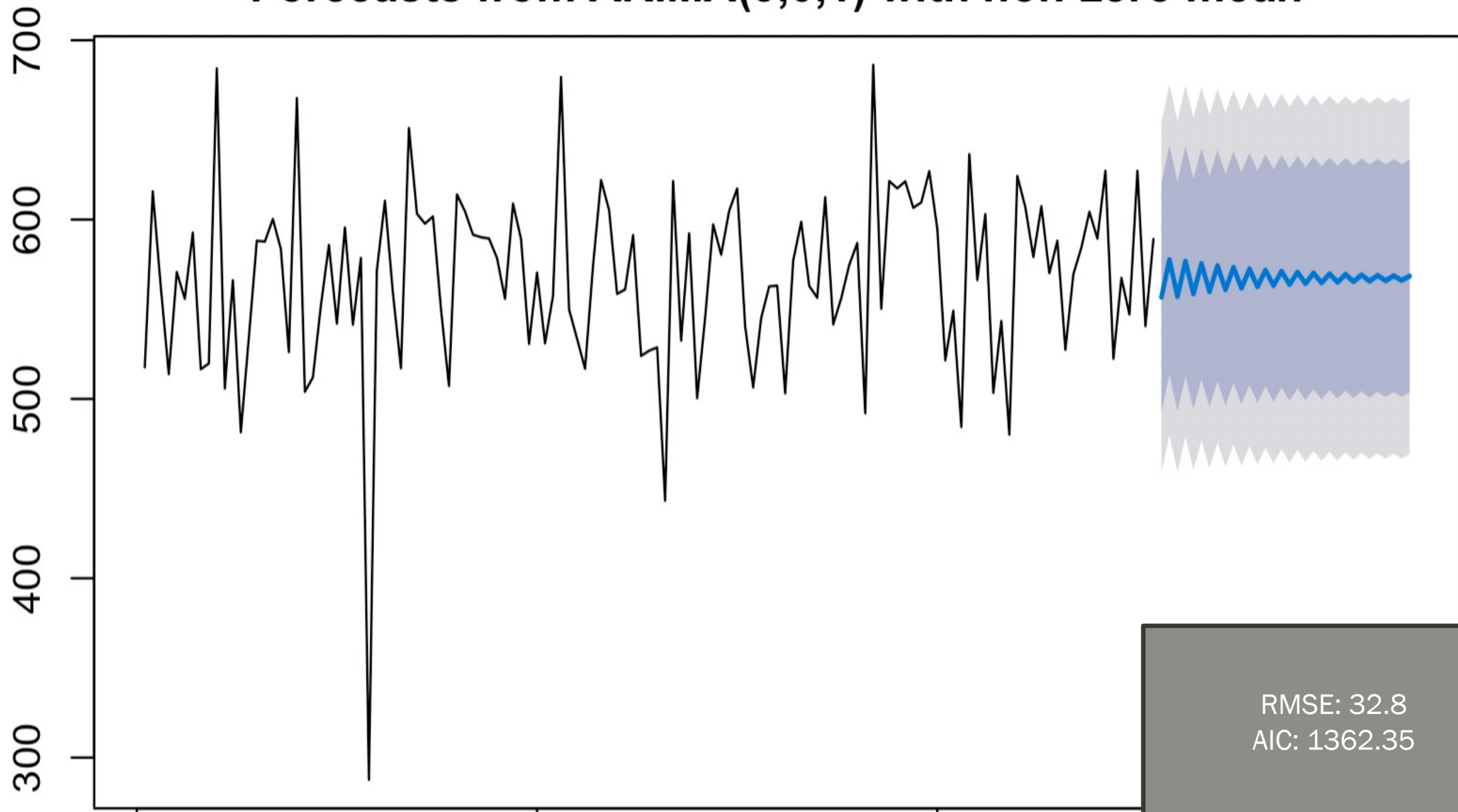
— Signal  
— Moving Average



# ARIMA RESIDUALS: ACF AND PACF

*The residuals don't look like they will be super helpful in determining ARIMA( $p, q, d$ ) for model preparation. Though some points on the lag look better than others, we already know there isn't a huge trend to follow from previous data understanding*

**Forecasts from ARIMA(3,0,1) with non-zero mean**



# THANK YOU!

