

Imputation Project

I began the project with tidying the given data frame based on my needs, using Pandas functions to generate a dataset with daily dates, indices for each row, and an NA type value for missing data, that would go one week into the future from the last date given. My approach for imputing data in a column of the dataset, was to iterate through said column, until a missing value is found, then I would use an appropriate equation, to estimate its value, and replace it with the NA. It is important to note, that I *only used original values to approximate and not the estimated values*, to avoid a growing error margin.

Linear Approximation

For **Linear Approximation**, I decided to approximate values based on their previous, and next closest points. In the case of the week into the future, I used the last two previous points instead. This is what the estimations looked like:



