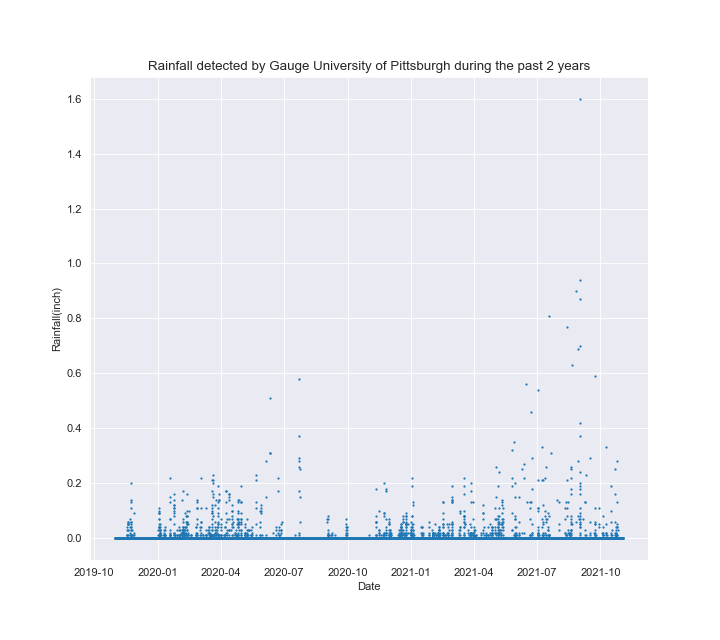
1. Warming up to Time-Series Data Again (15%)

a. 一張含有 文字 的圖片

自動產生的描述



b.

In the rainfall data points, we can see that there seems to be a daily variation in the amount of rainfall. Therefore, to design an equation to model the linear regression, I would propose where is the recorded rainfall, is the parameter, is the time instance with a time sampling interval of 1 day, and is the noise recorded at the time instance.

After setting up the equation, we can now construct the matrix of parameters using and y vector with . Obtaining the estimated linear regression model through the matrix and y vector, we can than calculate the residuals, and the mean and standard derivation of the residuals. To remove the outliers, we may set a range using the mean and standard derivation of residuals. Using the Chauvenet’s criterion, the range is set to be plus and minus three times the standard deviation from the mean, any data point outside this range is considered an outlier. We will repeat the process until all the outliers are identified and removed.

1. Water in the foundation: yikes! (60%)

(a)

一張含有 文字 的圖片

自動產生的描述

= -0.00042984263809532115 micro-strain/min

= -439.50238120742614 micro-strain

= 1.4432743355313284 micro-strains/Celsius

= -53.15723364050697 micro-strains/meter

(b)



q\_3000 = -625.1325035247191 micro-strain

(c)



q\_5y= -1569.1288341219304 micro-strain

(d) 一張含有 文字 的圖片

自動產生的描述

一張含有 桌 的圖片

自動產生的描述



The mean of the residual = 1.6778227974623325e-12

The std of the residual = 3293.952960065999

(e)

一張含有 文字 的圖片

自動產生的描述

The mean of the residual = 1.6778227974623325e-12

The std of the residual = 3293.952960065999

The water level: -9.54688299960992, the Temperature: 4.17978837213146, the row is: 2499

The water level: -8.37139121446426, the Temperature: 2.10399519785036, the row is: 3049

-----------------------------------------------------

The mean of the residual = -5.787398817541758e-14

The std of the residual = 14.009994054339375

The water level: -10.1939760133644, the Temperature: 10.5216587260785, the row is: 598

The water level: -10.2119567730395, the Temperature: 10.8844563878947, the row is: 2199

The water level: -9.8411652681218, the Temperature: 2.07317374340667, the row is: 2959

The water level: -9.08461406356181, the Temperature: 1.39630567837468, the row is: 2969

The water level: -8.68034346724605, the Temperature: -1.51614060916072, the row is: 3199

The water level: -8.75525517694182, the Temperature:1.06486125085496, the row is: 3299

-----------------------------------------------------

The mean of the residual = 3.812267548032659e-14

The std of the residual = 12.456866906148734

The water level: -11.3933741762572, the Temperature: 18.4827137833432, the row is: 1419

The water level: -10.1807982237015, the Temperature: 8.58909008979396, the row is: 2426

-----------------------------------------------------

The mean of the residual = -1.220861124667725e-14

The std of the residual = 12.42769063549792

-----------------------------------------------------

no more outliers

= -3.99452445236156e-05 micro-strain/min

= 12.8355141371718 micro-strain

= 1.5261812447776726 micro-strains/Celsius

= -0.958624064677513 micro-strains/meter

(f) 

q\_3000 = -4.415079381818739 micro-strain

q\_5y= -92.14058847089 micro-strain (g)

(g)



= -8.044421308508439e-05 micro-strain/min

= -37.65036760882104 micro-strain

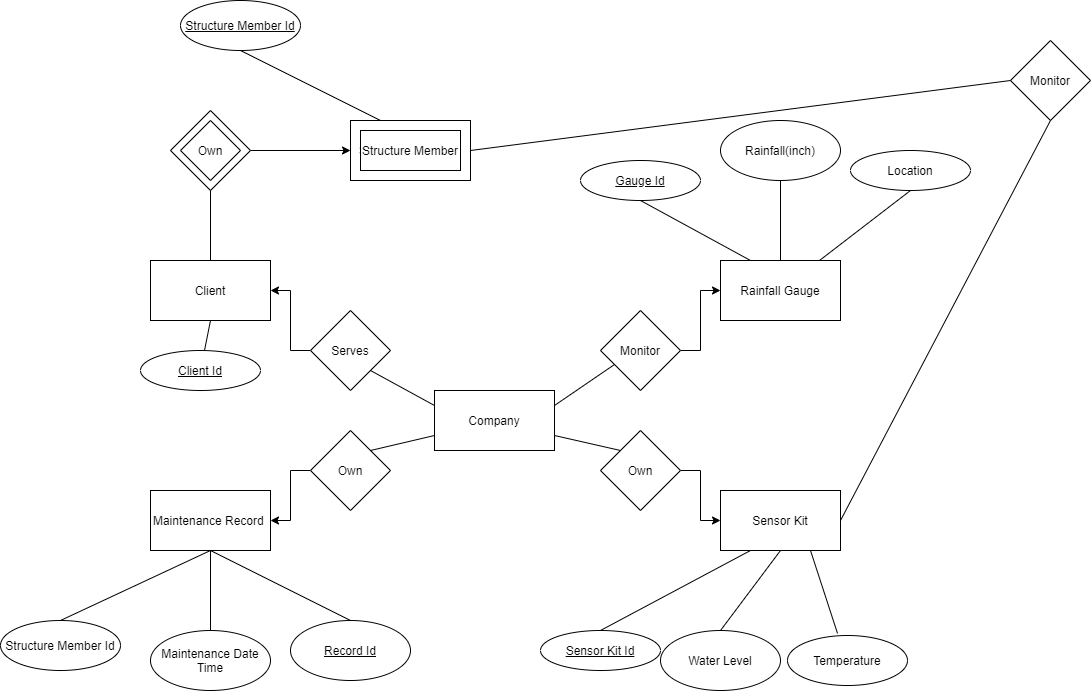
= -9.409107442878295 micro-strains/meter

q\_3000 = -72.39068369489326 micro-strain

q\_5y= -249.05775959642278 micro-strain

Removing the Temperature record, the estimated y of our linear model has a drastic difference, and the predictions are more inaccurate. Especially when we are predicting future values of y. The further into the future we are predicting, the larger the noise will be, therefore the larger the inaccuracy will be.

1. Wet Databases (15%)



1. Set Theory (10%)

Because:

Therefore, the baseball players are a larger group.