Given that the daily share prices are denoted by

$x_0, x_1, -- x_t, x_{t+1}$,

=> Then we need to develop a linear model which can be used to train a model based on this data value

=> Using share prices on 2 proceeding days, then our predicted y value is given by

$y = f(\vec{x})$

=> Thus the general function for y is given by

$$y = \beta_0 + \beta_1 x_0 + \beta_1 x_1 + --- + \beta_t x_t + \beta_{t-1} x_{t-1} \quad ——$$

where β_0 = y-intercept parameter

$\beta_1 = \beta_2 = --- \beta_t = \beta_{t-1}$ => slope parameter

Given that the daily share prices are denoted by

$x_0, x_1, -- x_t, x_{t+1}$,

=> Then we need to develop a linear model which can be used to train a model based on this data value

=> Using share prices on 2 proceeding days, then our predicted y value is given by

$y = f(\vec{x})$

=> Thus the general function for y is given by

$$y = \beta_0 + \beta_1 x_0 + \beta_1 x_1 + --- + \beta_t x_t + \beta_{t-1} x_{t-1} \quad —— (1)$$

where β_0 = y-intercept parameter

$\beta_1 = \beta_2 = --- \beta_t = \beta_{t-1}$ => slope parameter

ε - Error term

1) From the above general equation (1) then by using simpl

Regression equation

$$E(y) = \beta_0 + \beta_1 x_1 + -- + \beta_t x_t + \beta_{t-1} x_{t-1}$$

is the expected value of y for given

where $E(y)$

Using share prices on 2 proceeding days, then our predicted y value is given by

$y = \bar{f}(x)$

⇒ Thus the general function for y is given by

$$y = \beta_0 + \beta_1 x_0 + \beta_1 x_1 + \cdots + \beta_t x_t + \beta_{t-1} x_{t-1} \quad \text{—— ①}$$

where β_0 = y-intercept parameter

$\beta_1 = \beta_2 = \cdots \beta_t = \beta_{t-1}$ ⇒ slope parameters

ε - Error term

① From the above general equation ① then by using simple Regression equation

$$E(y) = \beta_0 + \beta_1 x_1 + \cdots + \beta_t x_t + \beta_{t-1} x_{t-1}$$

where $E(y)$ is the expected value of y for given values of $x_0, x_1, \cdots x_t, x_{t+1}$

⇒ Since from the given share prices data, the values are in increasing sequence, then our expected value equals to

$$E(y) = \beta_0 + \beta_1 x$$

⇒ So now for $\hat{y}$ which is the mean value of y for given x value will be given as

$$\hat{y} = b_0 + b_1 x$$

Therefore the linear model for prediction will be given as

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

⇒ Thus the general function for y is given by

$$y = \beta_0 + \beta_1 x_0 + \beta_1 x_1 + \cdots + \beta_t x_t + \beta_{t-1} x_{t-1} \quad \text{—— ①}$$

where β_0 = y-intercept parameter

$\beta_1 = \beta_2 = \cdots \beta_t = \beta_{t-1}$ ⇒ slope parameters

ε - Error term

① From the above general equation ① then by using simple Regression equation

$$E(y) = \beta_0 + \beta_1 x_1 + \cdots + \beta_t x_t + \beta_{t-1} x_{t-1}$$

where $E(y)$ is the expected value of y for given values of $x_0, x_1, \cdots x_t, x_{t+1}$

⇒ Since from the given share prices data, the values are in increasing sequence, then our expected value equals to

$$E(y) = \beta_0 + \beta_1 x$$

⇒ So now for $\hat{y}$ which is the mean value of y for given x value will be given as

$$\hat{y} = b_0 + b_1 x$$

⇒ Therefore the linear model for prediction will be given as

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

) Using x_t and x_{t-1} to predict Δ_{t+1}.

⇒ We know that the linear model actual response will be given by

$\beta_1 - \beta_2 - \cdots \beta_t = \beta_{t-1} \rightarrow$ slope parameter

ε - Error term

(1) From the above general equation (i) then by using simple Regression equation

$E(y) = \beta_0 + \beta_1 x_1 t \ \cdots + \beta_t x_t + \beta_{t-1} x_{t-1}$

where $E(y)$ is the expected value of y for given values of $x_0, x_1, \cdots x_t, x_{t+1}$

=> Since from the given share prices data, the values are in increasing sequence, then our expected value equals to

$E(y) = \beta_0 + \beta_1 x$

=> So now for $\hat{y}$ which is the mean value of y for given x value will be given as

$\hat{y} = b_0 + b_1 x$

=> Therefore the linear model for prediction will be given as

$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$

(2) Using x_t and x_{t-1} to predict Δ_{t+1}.

=> We know that the linear model actual response will be given by

$y = f(x) + \varepsilon$

=> Thus to get the bias within the 2 preceeding days we have $e = y - \hat{y}$ which is the residual error

=> Extending this to the shared prices datagen, will be given as

where β_0 = y-intercept parameter

$\beta_1 - \beta_2 - \cdots \beta_t = \beta_{t-1} \rightarrow$ slope parameter

ε - Error term

(1) From the above general equation (i) then by using simple Regression equation

$E(y) = \beta_0 + \beta_1 x_1 t \ \cdots + \beta_t x_t + \beta_{t-1} x_{t-1}$

where $E(y)$ is the expected value of y for given values of $x_0, x_1, \cdots x_t, x_{t+1}$

=> Since from the given share prices data, the values are in increasing sequence, then our expected value equals to

$E(y) = \beta_0 + \beta_1 x$

=> So now for $\hat{y}$ which is the mean value of y for given x value will be given as

$\hat{y} = b_0 + b_1 x$

=> Therefore the linear model for prediction will be given as

$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$

(2) Using x_t and x_{t-1} to predict Δ_{t+1}.

=> We know that the linear model actual response will be given by

$y = f(x) + \varepsilon$

=> Thus to get the bias within the 2 preceeding days we have $e = y - \hat{y}$ which is the residual error

=> Extending this to the shared prices datagen, will be given as

$e = \beta_0 + \beta_1 X + \varepsilon - (\hat{\beta}_0 + \hat{\beta}_1 X)$

$\Rightarrow e = (\beta_0 - \hat{\beta}_0) + (\beta_1 - \hat{\beta}_1)X + \dots + (\beta_t - \hat{\beta}_t)X_t + (\beta_{t-1} - \hat{\beta}_{t-1})X_{t+1} + \varepsilon$

$\Rightarrow$ Thus the arg min function will be written as

$$\arg\min \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$$

$\Rightarrow$ Differentiating this w.r.t β_0, we have

$$\sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0$$

$$\Rightarrow \hat{\beta}_0 = \sum_{i=1}^{n} \frac{y_i}{n} - \hat{\beta}_1 \sum_{i=1}^{n} \frac{x_i}{n}$$

$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$ where $\bar{y}$ and $\bar{x}$ are the share pr

$e = \beta_0 + \beta_1 X + \varepsilon - (\hat{\beta}_0 + \hat{\beta}_1 X)$

$\Rightarrow e = (\beta_0 - \hat{\beta}_0) + (\beta_1 - \hat{\beta}_1)X + \dots + (\beta_t - \hat{\beta}_t)X_t + (\beta_{t-1} - \hat{\beta}_{t-1})X_{t+1} + \varepsilon$

$\Rightarrow$ Thus the arg min function will be written as

$$\arg\min \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$$

$\Rightarrow$ Differentiating this w.r.t β_0, we have

$$\sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0$$

$$\Rightarrow \hat{\beta}_0 = \sum_{i=1}^{n} \frac{y_i}{n} - \hat{\beta}_1 \sum_{i=1}^{n} \frac{x_i}{n}$$

$\Rightarrow \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$ where $\bar{y}$ and $\bar{x}$ are the share prices sample mean scores

$\Rightarrow$ Again, differentiating argmin function w.r.t β_1 we have

$$\sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)(x_i) = 0$$

$$\Rightarrow \sum_{i=1}^{n} x_i y_i - \hat{\beta}_0 \sum_{i=1}^{n} x_i - \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 = 0$$

$\hat{\beta}_1 = \sum^{n} x_i y_i - n \bar{x} \bar{y}$

=> Thus the arg min function will be written as

$$\arg\min \sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$$

=> Differentiating this w.r.t β_0, we have

$$\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0$$

=> $\hat{\beta}_0 = \sum_{i=1}^{n} \frac{y_i}{n} - \hat{\beta}_1 \sum_{i=1}^{n} \frac{x_i}{n}$

=> $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$ where $\bar{y}$ and $\bar{x}$ are the share prices sample mean scores

=> Again, differentiating argmin function w.r.t β_1 we have

$$\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)(x_i) = 0$$

=> $\sum_{i=1}^{n} x_i y_i - \hat{\beta}_0 \sum_{i=1}^{n} x_i - \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 = 0$

=> $\hat{B}_1 = \dfrac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}$

=> Thus, using $\hat{\beta}_0$ and $\hat{\beta}_1$, our linear model function becomes,

$$\hat{y} = \hat{B}_0 + \hat{B}_1 X$$

... Δ_t to predict Δ_{t+1}

+ ε

=> Thus the arg min function will be written as

$$\arg\min \sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$$

=> Differentiating this w.r.t β_0, we have

$$\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0$$

=> $\hat{\beta}_0 = \sum_{i=1}^{n} \frac{y_i}{n} - \hat{\beta}_1 \sum_{i=1}^{n} \frac{x_i}{n}$

=> $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$ where $\bar{y}$ and $\bar{x}$ are the share prices sample mean scores

=> Again, differentiating argmin function w.r.t β_1 we have

$$\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)(x_i) = 0$$

=> $\sum_{i=1}^{n} x_i y_i - \hat{\beta}_0 \sum_{i=1}^{n} x_i - \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 = 0$

=> $\hat{B}_1 = \dfrac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}$

=> Thus, using $\hat{\beta}_0$ and $\hat{\beta}_1$, our linear model function becomes,

$$\hat{y} = \hat{B}_0 + \hat{B}_1 X$$

(3) Using Δ_t to predict Δ_{t+1},

let $y = f(x) + \varepsilon$

where $f(x) = \beta_0 + \beta_1 X_1 + B_2 X_2 + \cdots + B_t X_t$

=> $\hat{y} = \hat{\beta}_0 + \hat{B}_1 X_1 + \cdots + \hat{B}_t X_t$

=> Now, for n number of samples as in our share prices values the number of operations will be equal to argmin $\sum_{i=1}(y_i - \beta_0 - \beta_1 x_i)$

=> Differentiating this w.r.t β_0, we have

$$\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0$$

=> $\hat{\beta}_0 = \sum_{i=1}^{n} \frac{y_i}{n} - \hat{\beta}_1 \sum_{i=1}^{n} \frac{x_i}{n}$

=> $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$ where $\bar{y}$ and $\bar{x}$ are the share prices sample mean scores

=> Again, differentiating argmin function w.r.t β_1 we have

$$\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)(x_i) = 0$$

=> $\sum_{i=1}^{n} x_i y_i - \hat{\beta}_0 \sum_{i=1}^{n} x_i - \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 = 0$

=> $\hat{B}_1 = \dfrac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}$

=> Thus, using $\hat{\beta}_0$ and $\hat{\beta}_1$, our linear model function becomes,

$$\hat{y} = \hat{B}_0 + \hat{\beta}_1 X$$

(3) Using Δ_t to predict Δ_{t+1},

let $y = f(X) + \varepsilon$

where $f(x) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_t X_t$

=> $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \cdots + \hat{\beta}_t X_t$

=> Now, for n number of samples as in our share prices values the number of operations will be equal to

=> $n \times (p-1)^2$

=> From the share prices vals $(x_0, x_1, \cdots x_{t+1})$ we have

$y = X\beta + \varepsilon$ with $\hat{y} = X\hat{\beta}$

=> Finding the estimated $\hat{y}$ for the Δ_{t+1} ~~values~~ we have

$$\hat{y} = X\hat{\beta}_0 + X_1\hat{\beta}_1 + \varepsilon$$

(4) Using Δ_t and x_t to predict Δ_{t+1}

=> From $\hat{y} = \hat{f}(X) = X\hat{\beta}$

=> For variations in both Δ_t and x_t, our y-predicted function model will take the form

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{B}_2 X_2 + \cdots + \hat{\beta}_t X_t + \varepsilon$$

where ε is the bias term

=) From the Share prices Vals ($x_0, x_1, \dots x_{t+1}$) we have

$y = X\beta + \varepsilon$ with $\hat{y} = X\hat{\beta}$

=) Finding the estimated $\hat{y}$ for the Δ_{t+1} values we have

$\hat{y} = X\hat{\beta}_0 + X_1\hat{\beta}_1 + \varepsilon$

(4) Using Δ_t and x_t to predict Δ_{t+1}

=) From $\hat{y} = \hat{F}(X) = X\hat{\beta}$

=) For variations in both Δ_t and x_t, our y-predicted function model will take the form

$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \dots + \hat{\beta}_t X_t + \varepsilon$

where ε is the bias term

(5) Models ① and ② are equivalent

=) These 2 models takes in the independent x variables and then integrates them into making predictions with no bias term, unlike in ③ and ④ where the bias term plays a key factor in detecting the

=) From the Share prices Vals ($x_0, x_1, \dots x_{t+1}$) we have

$y = X\beta + \varepsilon$ with $\hat{y} = X\hat{\beta}$

=) Finding the estimated $\hat{y}$ for the Δ_{t+1} values we have

$\hat{y} = X\hat{\beta}_0 + X_1\hat{\beta}_1 + \varepsilon$

(4) Using Δ_t and x_t to predict Δ_{t+1}

=) From $\hat{y} = \hat{F}(X) = X\hat{\beta}$

=) For variations in both Δ_t and x_t, our y-predicted function model will take the form

$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \dots + \hat{\beta}_t X_t + \varepsilon$

where ε is the bias term

(5) Models ① and ② are equivalent

=) These 2 models takes in the independent x variables and then integrates them into making predictions with no bias term, unlike in ③ and ④ where the bias term plays a key factor in detecting the
estimator function for $\hat{y}$

(6) Training the model in part 4

=) First, we generate the samples training datagen, in our case we are using the share prices values,

=) We use the random function to perform shuffling of the data so as to obtain the training and testing data

4) Using Δ_t and x_t to predict Δ_{t+1}

=> From $\hat{y} = \hat{f}(x) = X\hat{\beta}$

=> For variations in both Δ_t and x_t, our y-predicted function model will take the form

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \cdots + \hat{\beta}_t X_t + \varepsilon$$

where ε is the bias term

(5) Models ① and ② are equivalent

=> These 2 models takes in the independent x variables and then integrates them into making predictions with no bias term, unlike in ③ and ④ where the bias term plays a key factor in detecting the estimator function for $\hat{y}$

(6) Training the model in part 4

=> First, we generate the samples, training datagen, in our case we are using the share prices values,

=> We use the random function to perform shuffling of the data so as to obtain the training and testing data

=> Next, we treat the training samples as our independent variables whose real values are known.

=> We then feed forward these training samples ~~on~~ to our linear model function

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \cdots + \hat{\beta}_t X_t + \varepsilon$$

=> For variations in both Δ_t and x_t, our y-predicted function model will take the form

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \cdots + \hat{\beta}_t X_t + \varepsilon$$

where ε is the bias term

(5) Models ① and ② are equivalent

=> These 2 models takes in the independent x variables and then integrates them into making predictions with no bias term, unlike in ③ and ④ where the bias term plays a key factor in detecting the estimator function for $\hat{y}$

(6) Training the model in part 4

=> First, we generate the samples, training datagen, in our case we are using the share prices values,

=> We use the random function to perform shuffling of the data so as to obtain the training and testing data

=> Next, we treat the training samples as our independent variables whose real values are known.

=> We then feed forward these training samples ~~on~~ to our linear model function

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \cdots + \hat{\beta}_t X_t + \varepsilon$$

=) This will yield an estimate of the predicted y value

=) Since our linear model has been trained perfectly on the share price data values, it can make precise predictions ($\hat{y}$) which is the unknown function to be estimated by our linear model algorithm