Linear Models

Note that in building our kNN model for prediction, we did not compute a closed form for $\hat{F}$. kNN is a **non-parametric** model.

What if we ask the question, "how much more sales do we expect if we double the TV advertising budget?"

We can build a model by first assuming a simple form of f:

$$f(x) = \beta_0 + \beta_1 X$$

where β_1 represents the slope and β_0 represents the intercept. It then follows that our estimate is:

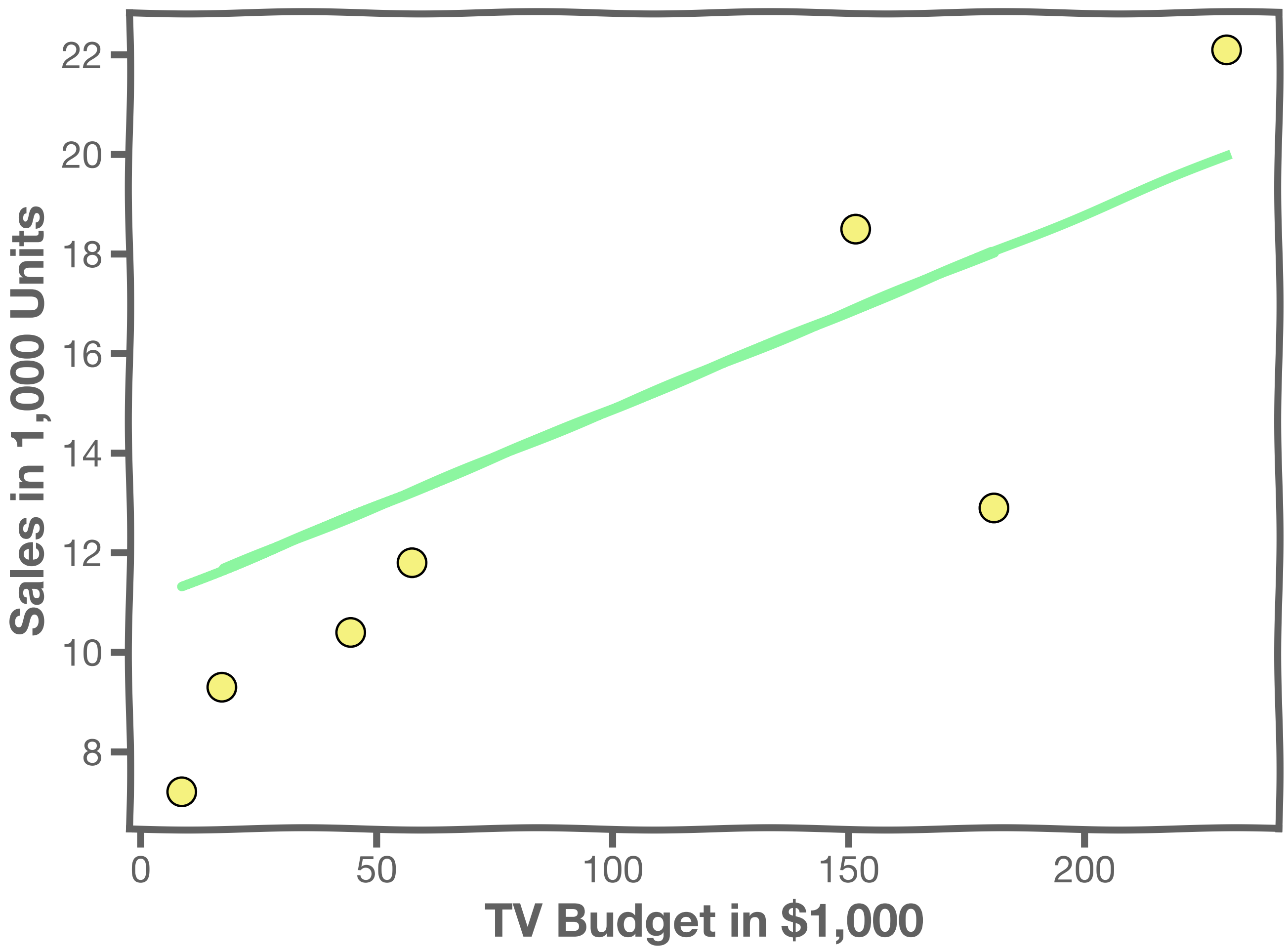
$$\hat{Y} = \hat{f}(x) = \hat{\beta}_0 + \hat{\beta}_1 X$$

where $\hat{\beta}_1$ and $\hat{\beta}_0$ are estimates of β_1 and β_0, respectively, which we compute using observations.

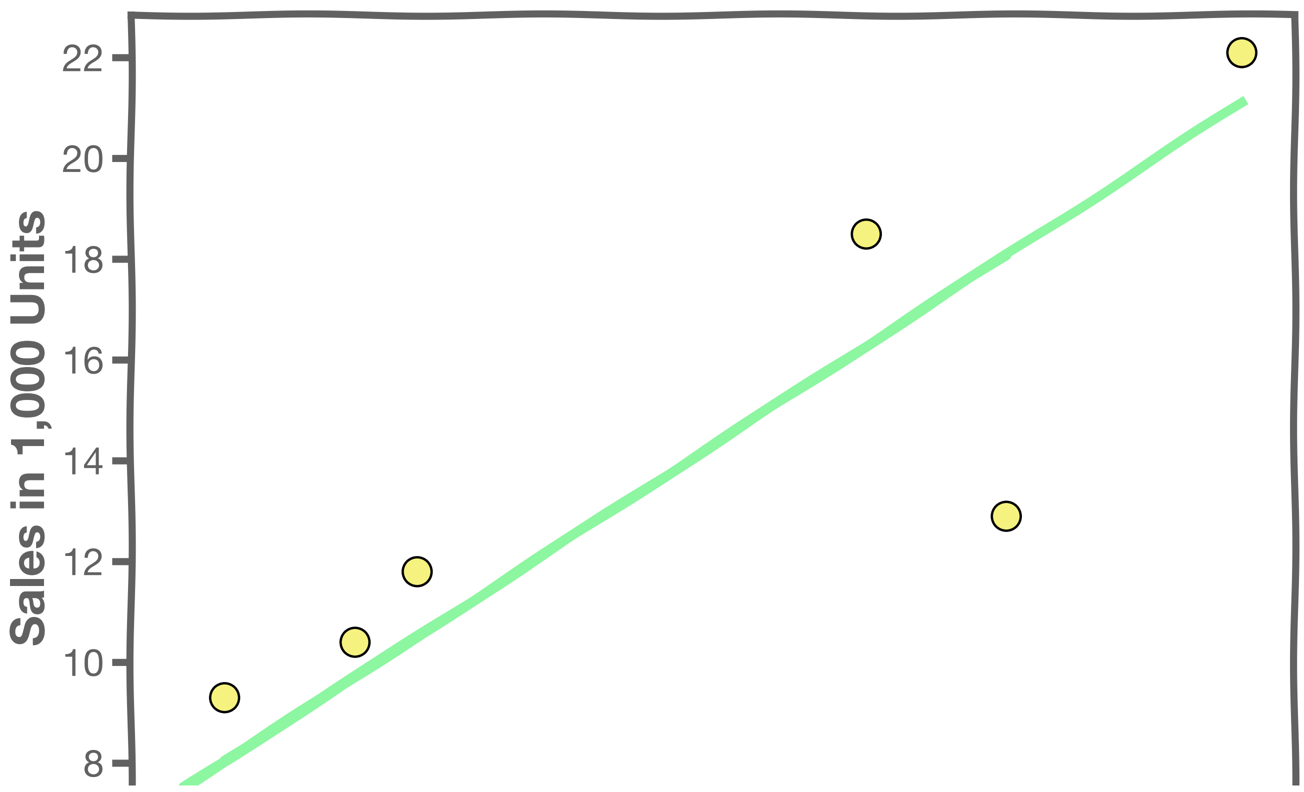
Estimate of the regression coefficients

For a given data set we can draw different lines through the data, we need to determine which line is the best fit.

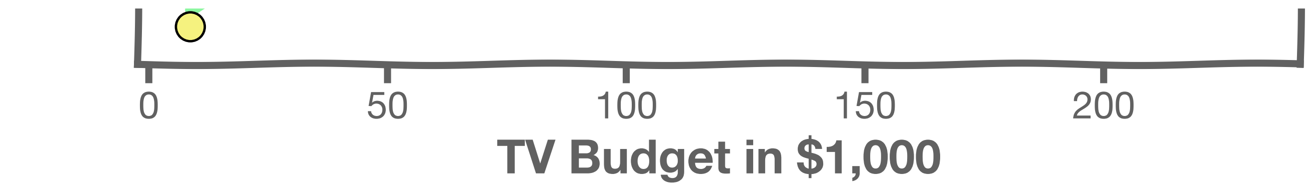
Is this it?



Or this one?

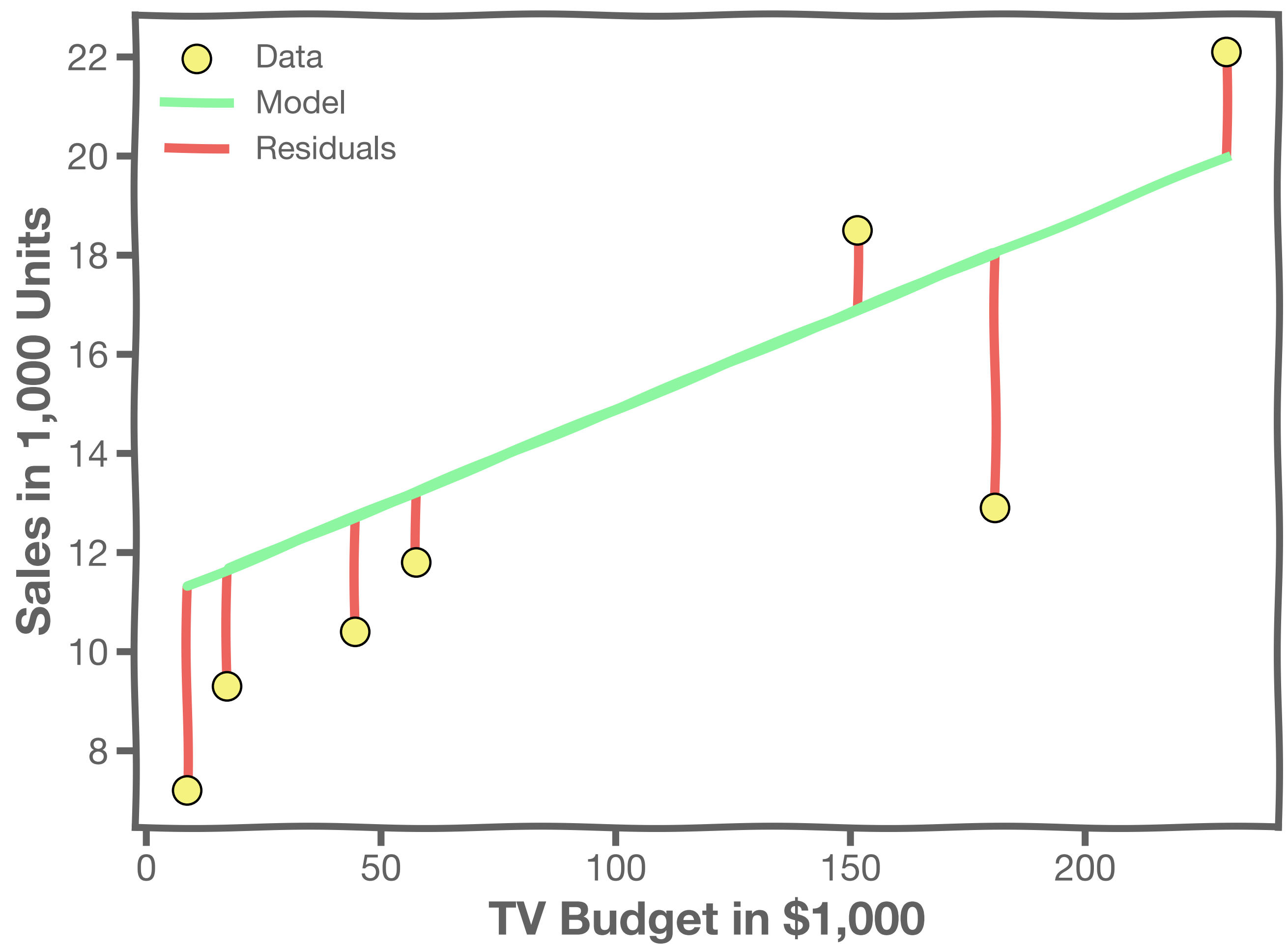


Calculator



Question: Which line is the best?

To decide which is the best line, we can do the same as we did with the kNN model. We estimate the error for every model by looking at the residuals, the difference between the exact value of y and the predicted y. As shown in the plot below, the regression or model line, predicted sales, is the slanted line in green and the residuals to the exact values of sales are the vertical red lines.



As before, for each observation (x_n, y_n), the absolute residuals, $r_i = |y_i - \hat{y}_i|$ quantifies the error at each observation.

Again, we aggregate and use the Mean Squared Error, **MSE**, as our loss function.

$$L(\beta_0, \beta_1) = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

The difference from the kNN model is that there is a specific formula for the $\hat{y}$, which is the linear model.

We choose β_0 and β_1 that minimize the predictive errors made by our model, that is, minimize our loss function.

Then the optimal values, $\hat{\beta}_0$ and $\hat{\beta}_1$, should be:

$$\hat{\beta}_0, \hat{\beta}_1 = \underset{\beta_0, \beta_1}{\arg\min}\, L(\beta_0, \beta_1)$$

We call this "**fitting**" or "**training**" the model.

Discussion Board (External resource)

Haga clic en Aceptar para que su nombre de usuario y dirección de correo electrónico se envíen a una aplicación de terceros.

‹ Previous
Next ›

edX

About
Affiliates
edX for Business
Open edX
Careers
News

Legal

Terms of Service & Honor Code
Privacy Policy
Accessibility Policy
Trademark Policy
Sitemap
Cookie Policy
Your Privacy Choices

Connect

Idea Hub
Contact Us
Help Center
Security
Media Kit

GET IT ON Google play
Download on the App Store

Calculator