# Friday Quiz 3

## Example 1:

Imagine a horticultural researcher meticulously tracking the growth of various tomato plants in community gardens across the region. They've diligently recorded the hours of direct sunlight each garden receives and the final height of the mature plants. The data reveals a generally positive trend: more sun usually means taller plants. However, one particular garden, despite receiving a seemingly adequate amount of sunlight, yielded a surprisingly short average plant height compared to its peers. This anomaly has the researcher wondering if there might be a specific environmental factor, measurement error, or even a different plant variety at play in that location.

The following plot (Example 1) visualizes the relationship between sunlight exposure and plant height.
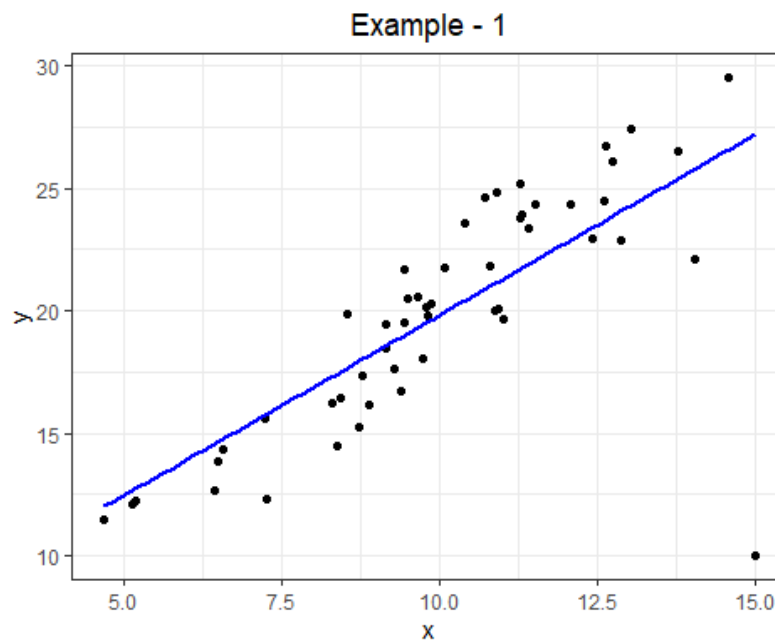
## Example 2:

Consider a marketing consultancy working with numerous small, independent bookstores. They're investigating the effectiveness of different marketing budgets on overall sales revenue. For most bookstores, there's a moderate correlation between spending on advertising and the resulting sales. However, two bookstores stand out. One invested significantly more in a targeted online campaign, while another allocated a larger budget to local community events and partnerships. The consultancy is keen to understand if these higher investments yielded proportionally higher returns compared to the more typical spending patterns.

The following plot (Example 2) illustrates the connection between marketing budget and total sales revenue.
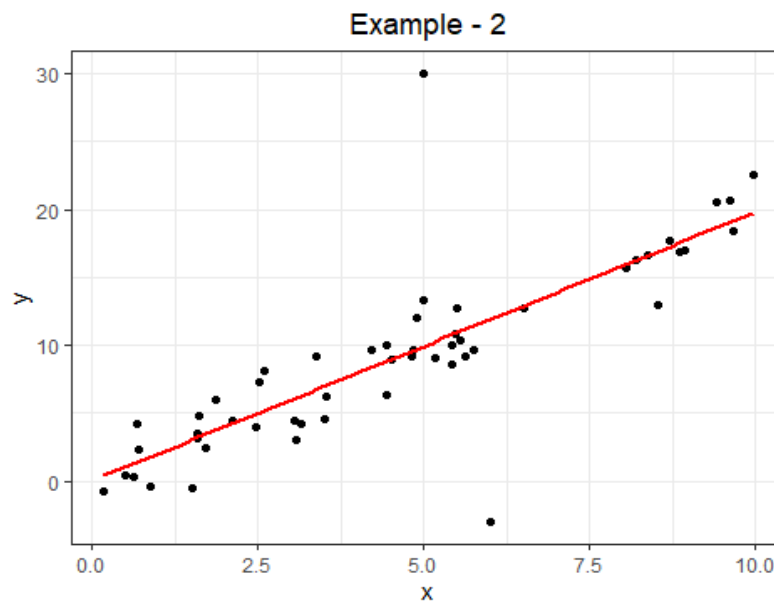
## Example 3:

Picture a customer service analyst at an online electronics retailer examining the relationship between the time it takes for an order to be delivered and the customer satisfaction scores reported after delivery. Generally, faster delivery times correlate with happier customers. However, one particular customer experienced a notably long delivery delay due to unforeseen logistical challenges, yet surprisingly submitted a very high satisfaction rating, praising the exceptional communication and problem-solving efforts of the support team. This unusual instance prompts the analyst to consider if certain service aspects can outweigh the impact of delivery speed on customer perception.

The following plot (Example 3) displays delivery times against customer satisfaction scores.
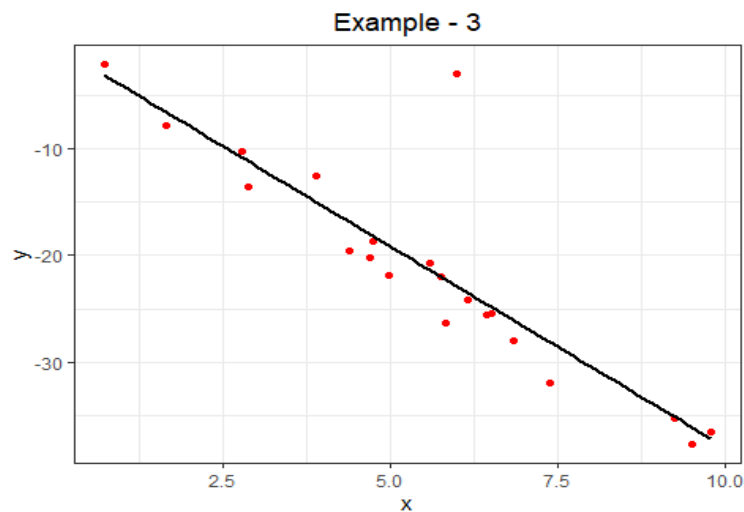
## Example 4:

Envision an agricultural scientist conducting a large-scale experiment across numerous experimental farm plots, testing the impact of varying levels of a new fertilizer on corn yield. After developing a statistical model to predict yield based on fertilizer application, the scientist examines the differences between the actual observed yields and the yields predicted by the model. These differences, or residuals, appear to be randomly scattered above and below zero, with no discernible pattern or clumping, suggesting that the linear model effectively captures the underlying relationship between fertilizer use and crop output.

Example 4 below presents the residuals plotted against the predicted (fitted) yield values and the histogram of these residuals.
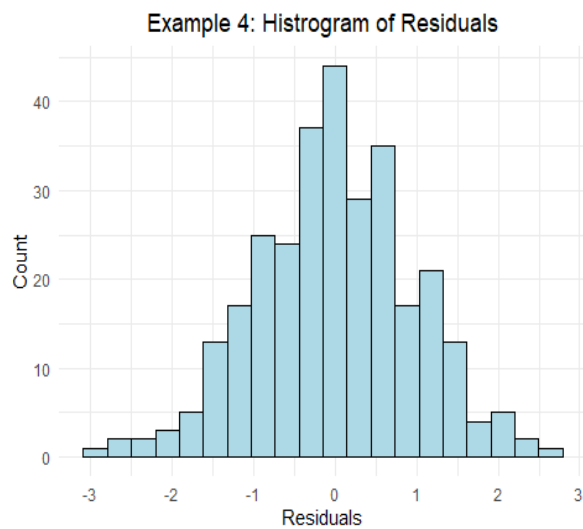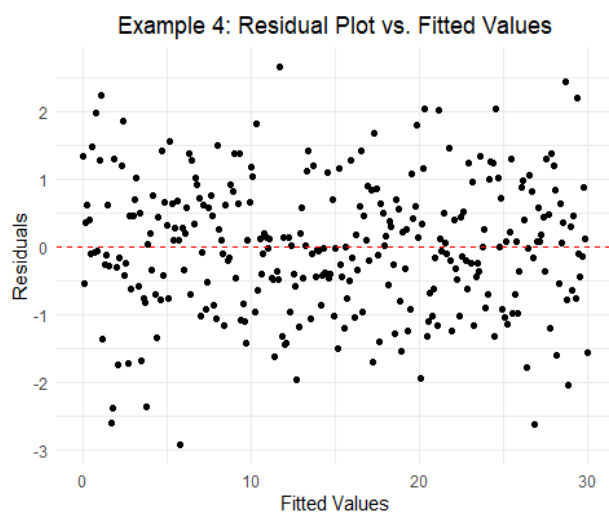
## Example 5:

Imagine researchers in a health and fitness study investigating how the number of hours a person exercises per week relates to their total cholesterol levels. While their statistical model seems to provide a reasonable fit for individuals with lower exercise levels, they observe that as exercise hours increase, the spread or variability in cholesterol levels among participants becomes much wider. Furthermore, they notice that for those who exercise more, the model tends to underestimate their cholesterol levels more frequently than overestimating.

The following plot (Example 5) displays the residuals against the fitted cholesterol values, alongside a histogram of these residuals.