Accuracy

An introduction to model accuracy and metrics

What is a model?

A model is a **mathematical representation** of a real-world process or **relationship between variables**.

For example, the **line of best fit is a linear model** used to represent the relationship between two variables.

Real-world situations often involve more complex relationships, requiring more detailed models.

Models often have multiple variables and interactions to better capture the details of the processes they represent.

What is model accuracy?

Model accuracy refers to how well a **model is able to accurately predict values**. We use various **metrics** to **measure model accuracy**.

The importance of model accuracy

- Ensuring that our model is **accurate** helps us **make better predictions and informed decisions**.
- An **inaccurate model** can lead to **incorrect conclusions and poor decision-making**.
- Evaluating model accuracy helps us **identify areas for improvement and refine our model**.

Residuals

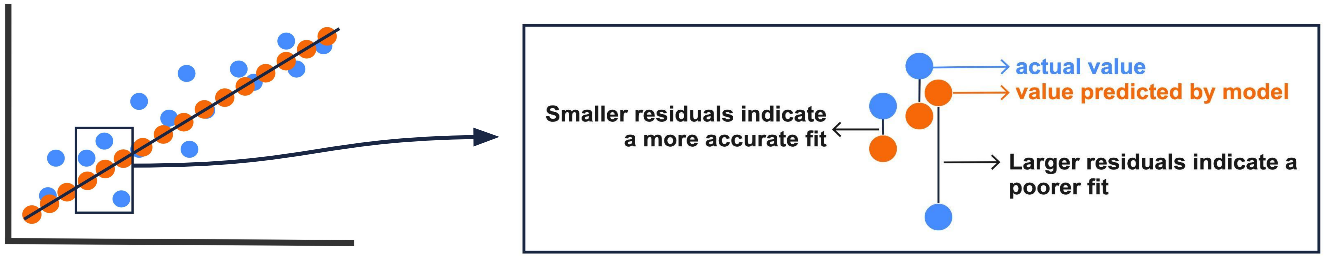
> A residual is the **difference between** the **observed value** and the **predicted value** for a data point. They are also referred to as **errors**.

Residuals help us understand **how well our model is fitting the data**.

By analysing the residuals, we can **identify patterns and trends** that our model may not be capturing.

When to use them?

- Residuals are useful for understanding the **pattern of errors** the model is making and **diagnosing issues** with the model.
- They can also be used to **identify outliers** or **influential data points**.

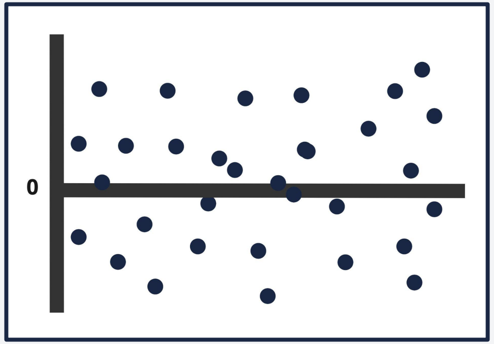


Residual plot

Residual plots are **scatter plots of residual values**. They help determine the **accuracy** of a model.

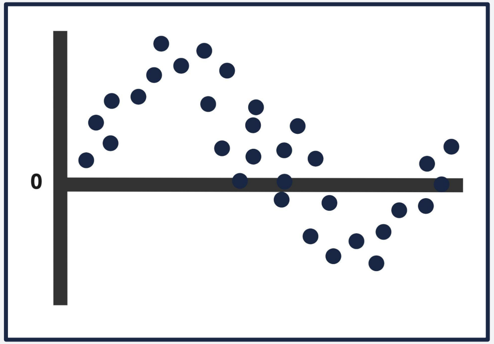
What a residual plot should look like

Residuals are **randomly scattered around zero**, with **no discernible pattern**.



What a residual plot shouldn't look like

Clear patterns or trends in residuals suggest bias or shortcomings in the model.



MAE, MSE, and RMSE

> Three commonly used measures of accuracy are MAE, MSE, and RMSE. These measures **quantify the errors or residuals** of a model's predictions and provide a way to **compare the accuracy of different models.**

Mean Absolute Error (MAE)

The **average of the absolute values of the residuals**. It represents the average distance between the predicted values and the actual observed values.

Mean Squared Error (MSE)

The **average of the squared residuals**. It emphasises larger errors by squaring them, making it more sensitive to outliers.

Root Mean Squared Error (RMSE)

The **square root of the MSE**. It represents the average distance between the predicted values and the actual observed values, similar to MAE, but with more emphasis on larger errors.

Understanding MAE for model accuracy

The **average of the absolute values of the residuals**.

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n}$$

where y_i is the predicted value, x_i is the observed value, and n is the number of data points.

Analysis

- It gives an idea of how **far the predictions are**, on average, **from the true values**.
- A **lower** MAE value indicates a **more accurate** model. For example, if the MAE is 2, it means that, on average, the model's predictions are off by 2 units.
- MAE is **easy to interpret** and **less sensitive to outliers**.

Understanding MSE for model accuracy

The **average of the squared residuals**.

$$MSE = \frac{\sum^{n}_{i=1}(y_i - x_i)^2}{n}$$

where y_i is the predicted value, x_i is the observed value, and n is the number of data points.

Analysis

- It is the **average of the squared differences** between the predicted values and the actual values of the target variable.
- A **lower** MSE value indicates a more **accurate** model. For example, if the MSE is 4, it means that, on average, the model's predictions are off by 4 **squared** units.
- MSE **penalises larger errors** more heavily than smaller errors as errors are squared and, therefore, larger errors have a greater impact on the overall value.

Understanding RMSE for model accuracy

The **square root of the MSE**.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - x_i)^2}{n}}$$

where y_i is the predicted value, x_i is the observed value, and n is the number of data points.

Analysis

- It represents the **average distance between the predicted values and the actual observed values**, similar to MAE, but with more **emphasis on larger errors**.
- A **lower** value of RMSE indicates a more accurate model.
- RMSE is more interpretable than MSE, as it is in **the same unit as the original data**. For example, if the RMSE is 2, it means that, on average, the model's predictions are off by 2 units, in the same units as the target variable.

Interpreting accuracy in context

> To **determine the accuracy** of our model, it's crucial to **interpret metrics within the context** of the scenario we are investigating.

Question: We have calculated an MAE of 4 units. **Is our model accurate?** Let's consider two scenarios:

Predicting a student's final grade

If we aim to **predict a student's final grade**, an **MAE of 4 units** indicates our model is relatively **accurate**. This is because we consider a **predicted mark of 80% quite close to an actual mark of 84%**.

Predicting shoe size

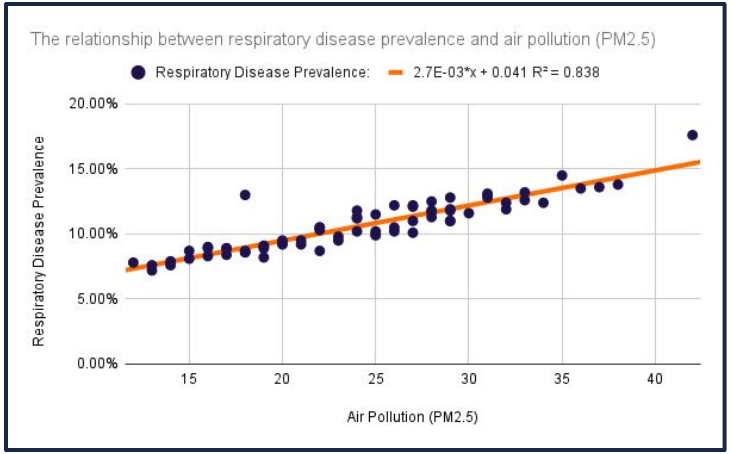
If we aim to **predict the shoe size** someone will buy, an **MAE of 4 units** indicates high **inaccuracy**. This is because a shoe size difference of 4 units, such as a **size 5 to size 9, is significantly different**.

The same logic applies to **residuals**, **MSE**, **and RMSE**. Therefore, it is important to always **consider the context** of the problem when interpreting our results.

Accuracy

Calculating MAE, MSE, and RMSE – high accuracy

Let's calculate MAE, MSE, and RMSE to assess the **accuracy** of a **linear model** (the line of best fit) representing the relationship between **air pollution and respiratory disease prevalence** in African cities.



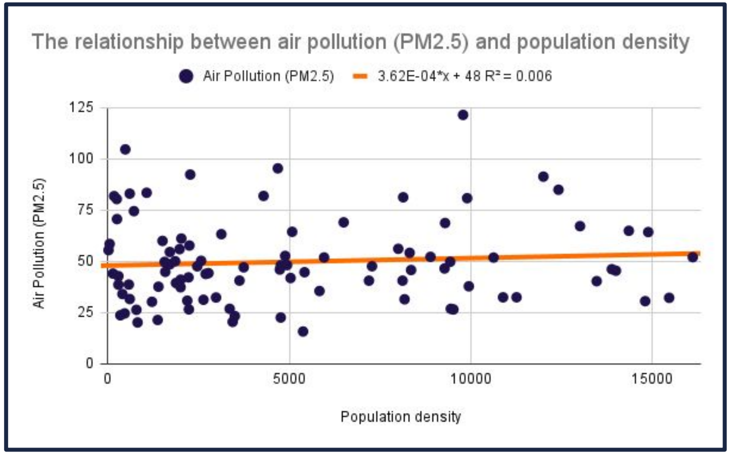

```
MAE = 1.3308
MSE = 2.2809
RMSE = 1.5104
```

- An **MAE of 1.3308** suggests, on average, the **predicted** respiratory disease prevalence values **deviate from the actual values by 1.33 percentage points**. The relatively **low MAE** suggests that the linear model provides reasonably **accurate predictions**.
- The **MSE of 2.2809** suggests that there is **some variability** in the predicted values compared to the actual values, although it is **small**.
- The **RMSE of 1.5104** indicates that, on average, the predicted respiratory disease prevalence values deviate from the **actual values by approximately 1.51 percentage points**. The low value indicates our model is making **accurate predictions**.

Accuracy

Calculating MAE, MSE, and RMSE – low accuracy

Let's calculate MAE, MSE, and RMSE to assess the **accuracy** of a **linear model** (the line of best fit) representing the relationship between **population density and air pollution levels** in African cities.



MAE = 15.7641
MSE = 419.9802
RMSE = 20.4934

- An **MAE of 15.76** suggests that, on average, the **predicted** air pollution values **deviate from the actual values by 15.76 units**. The relatively **high MAE** suggests that the model's predictions have a moderate level of error and may **not be accurate**.
- An **MSE of 419.98** is relatively **high**. This suggests there are **considerable differences** between the values, indicating the model's predictions may **not be accurate**.
- An **RMSE of 20.49** is **high.** This suggests that the model's predictions have a **moderate to high level of deviation** from the actual values.

Comparison of accuracy metrics

Residuals

By examining the residuals, we can assess if there are any **patterns or trends** in the model's **predictions**.

We can examine the **residual plot** to assess the **accuracy** of the model.

MAE

Easy to interpret and **less sensitive to outliers**, which means it provides a good measure of average prediction error.

However, it **doesn't punish large errors** more than small ones. This can be a drawback when large errors are considered critical.

MSE

Punishes large errors more than small errors because the errors are squared. This is useful when we want to emphasise large errors.

However, MSE is **more sensitive to outliers** and it is **not as simple to interpret**.

RMSE

Combines the **advantages of both MAE and MSE**.

It **punishes large errors** more than small ones, similar to MSE. It is also more **interpretable as it is in the same units as the target variable**.

However, it remains **sensitive to outliers**.