# ASSIGNMENT #1

Applied Data Science with ML & AI - Horizons 25

**1. The "Aha!" Moment:**

I had one of my biggest "Aha!" moments while reading '*The Unreasonable Effectiveness of Data*'. I always assumed that the smarter and more complex your ML algorithm is, the more effective and *magical* it becomes. But it turns out that it can be the complete opposite… The paper explained that with enough data, **even simple models** can perform better than fancy, sophisticated ones. That was surprising because it made me realise how powerful data actually is.

Another idea that stood out for me was, **"Every model has assumptions, and they're always wrong—but some are useful." – Pedro Domingos**
It made me realise that models don't have to serve the answer to us on a silver platter to be valuable. They can often help us make better decisions or reveal patterns we couldn't see before.

An imperfect model can still guide us in the right direction.

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

**2. Data is King (or is it?):**

My real-world example: ***Google Translate!***

Earlier, programmers had tried using hand-coded rules to teach the algorithm a language, but it couldn't handle all the nuances and exceptions. However, when it was fed MASSIVE amounts of real-world translated text, it *didn't need to 'understand'* every grammatical rule. It learnt from patterns in the data.

The readings made it clear that it wasn't just the amount of data that mattered, but its variety and even its "messiness".

Real conversations, slang, typos—all of that helped the system become more accurate. Thus, it proved that more data, even if it's imperfect, often leads to better performance than a cleaner but smaller dataset.

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

**3. Humanity in the Loop:**

I was surprised by how dependent machine learning *still* is on humans—especially when it comes to feature engineering. I always assumed models could figure out what matters on their own, but that's **not really the case.** They don't "understand" the data like we do. Humans still have to define the problem clearly, choose what data to include, clean it up, and decide what a "good" result looks like.

This made me realize that people working in data science need to stay very aware. If we're not careful, we can accidentally pass our own **assumptions** or **biases** into the model. And even when a model performs well, that doesn't always mean it's right or unbiased.

In the end, machine learning isn't about *handing off* decisions to machines. It's about working alongside them, guiding them responsibly, and knowing when to step in and question the results.