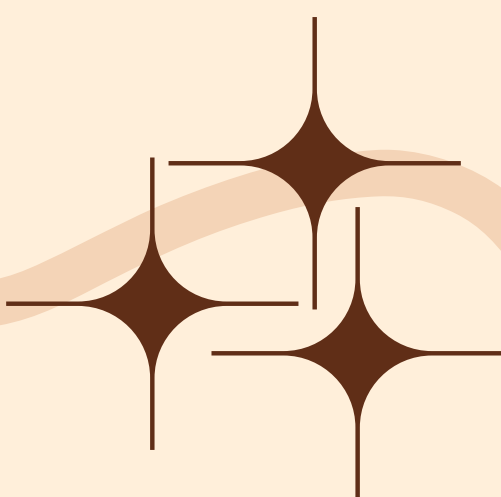


DADOS, PYTHON E ESTATÍSTICA: CAMINHOS PARA STEM



Ana Júlia Amaro Pereira Rocha



SEMANA 1

DIA 5

REVISÃO DA SEMANA

Nesta semana, definimos **dados** como registros de informações sobre o mundo, ou seja, podem representar pessoas, objetos, eventos...

Todos nós geramos dados o tempo todo. A **ciência de dados** usa essas informações para entender padrões e tomar decisões.

Dados podem ser **quantitativos** (representam números e medidas) ou **qualitativos** (descrevem características). Dados também são classificados como **estruturados, semi-estruturados e não estruturados**, como vimos.

REVISÃO DA SEMANA

Sobre os ambientes de armazenamento e gerenciamento de dados, discutimos 3: **banco de dados**, **data warehouse** e **data lake**.

Ambos tem propósitos, estruturas e tipos de dados distintos, sendo o banco de dados focado em transações operacionais (estruturados), o data warehouse em dados históricos e estruturados para análise, e o data lake em grandes volumes de dados brutos (estruturados, semi e não estruturados) para exploração e ciência de dados.

Comentamos também sobre os principais tipos de arquivo que “carregam” os dados, sendo eles **CSV**, **XLSX**, **SQL** e **JSON**.

REVISÃO DA SEMANA

Com relação à computação, entendemos que **hardware** refere-se a tudo que é possível tocar no computador (processador, monitor, teclado...) e **software** é a parte intangível, criada por código (jogos, sites...). E quem gerencia todos esses elementos é o **Sistema Operacional (SO)**.

Sobre programação, aprendemos que existem diferentes linguagens, cada uma é “melhor” a depender da área de uso e programar é uma maneira de se comunicar com máquina, fazendo ideias se transformarem em algo que o computador consegue executar.

(Revisamos termos e operadores de Python em sala, depois fizemos atividade prática de programação.)