**Semana 1 – Fundamentos de Dados**

* **Conceitos: O que é Engenharia de Dados?**
* **Diferença entre Engenheiro, Cientista e Analista de Dados.**
* **Tipos de dados: estruturados x não estruturados.**
* **Sistema de arquivos x Bancos de dados.**

**🛠️ Prática:**

* **Criar conta no GitHub**
* **Curso introdutório gratuito de SQL e Python (ex: SQL do Mode, Python do Kaggle)**

**📅 Semana 2 – SQL e Banco de Dados Relacional**

**📚 Conteúdos:**

* **SELECT, WHERE, JOIN, GROUP BY, HAVING, Subqueries.**
* **Modelagem de dados (1NF, 2NF, 3NF).**
* **Ferramentas: MySQL, PostgreSQL, DBeaver.**

**🧠 Prática Guiada:**

1. **Acesse e resolva exercícios nos sites:**
   * [**SQLZoo**](https://sqlzoo.net/)
   * [**DataLemur**](https://datalemur.com/)
   * **LeetCode - SQL**
2. **Use o** [**DBeaver**](https://dbeaver.io/) **para praticar com bancos locais.**
3. **Faça download do banco Sakila ou Northwind e rode no MySQL/PostgreSQL.**

**🛠️ Projeto prático sugerido:**

**Projeto: “Mini Sistema de Biblioteca”**

* **Modele um banco de dados com tabelas: Livros, Autores, Empréstimos, Usuários.**
* **Aplique diferentes tipos de JOINs e subqueries.**
* **Crie ao menos 8 consultas SQL cobrindo: filtros, agrupamentos e ordenações.**

**📅 Semana 3 – Python para Dados**

**📚 Conteúdos:**

* **Bibliotecas: Pandas, NumPy, openpyxl, json, os.**
* **Manipulação de DataFrames.**
* **Funções lambda, map, filter, apply.**

**🧠 Prática Guiada:**

1. **Faça desafios do** [**DataWars**](https://www.codewars.com/) **ou Kaggle Notebooks.**
2. **Baixe um dataset realista (como de vendas, clima ou finanças).**

**🛠️ Projeto prático sugerido:**

**Projeto: “Painel de Análise de Vendas”**

* **Faça o ETL de um dataset .csv com Pandas.**
* **Use groupby, merge e fillna para limpar e agrupar dados.**
* **Gere gráficos com Matplotlib e Seaborn.**
* **Extra: automatize o envio diário por e-mail com smtplib.**

**💡 Dica extra:**

**Inclua nesse módulo uma automação simples, como:**

* **RPA que abre arquivos Excel e envia e-mail com resumo.**
* **Extração de dados da web com requests e BeautifulSoup.**

📅 **Semana 4 – Data Warehousing e ETL** **Tópicos:**

* Conceito de Data Warehouse (OLAP vs OLTP)
* Ferramentas: Airbyte, DBT, Apache Airflow
* Tabelas Fato e Dimensão

**🛠️ Prática:**

* Assistir vídeos sobre Kimball vs Inmon
* **Projeto prático:** Criar um mini Data Warehouse com tabelas Fato de Vendas e Dimensão Cliente/Produto, e simular ETL com Python + agendamento

📅 **Semana 5 – Big Data e Arquitetura em Nuvem** **Tópicos:**

* HDFS, Hive, Spark
* Google Cloud / AWS / Azure
* Serviços: BigQuery, Redshift, Athena

**🛠️ Prática:**

* Explorar datasets públicos no BigQuery
* **Projeto prático:** Criar pipeline de ingestão de dados para GCP usando Python + upload em bucket

📅 **Semana 6 – Engenharia de Dados na Prática** **Tópicos:**

* Monitoramento de pipelines
* Logs, erros e alertas
* Testes e versionamento de dados

**🛠️ Prática:**

* **Projeto prático:** Simular um pipeline completo: coletar dados de uma API (ex: IBGE), transformar com Pandas e armazenar em banco SQL
* Usar Git para versionar tudo e programar alertas por e-mail/log