**Preparação de Dados para Análises Financeiras**

Quando eu iniciei minhas atividades como analista financeiro no Grupo Neoguard, eu fui encarregado de fazer uma análise dos sus dados financeiros, calcular indicadores e preparar painéis interativos através do Power BI. O sistema de informações gerenciais com a sua base de dados estava carregado na nuvem da empresa que o fornecia e eu, inicialmente, só conseguia extrair seus dados por meio de arquivos em Excel, que tinham muitos dados desnecessários e um layout confuso e poluído – eu precisava trabalhar muito os dados no Excel antes de carregá-los no Power BI para realizar as análises.

Contudo, logo no meu segundo mês, após eu trabalhar bastante com os dados brutos e conhecê-los um pouco melhor através de análise exploratória, o sistema e sua base de dados foram baixados e passaram a rodar internamente, no Datacenter da empresa. A base de dados era gerenciada por SQL Server, eu decidi então analisar localizar os dados direto nas tabelas, mas a estrutura das tabelas era extremamente confusa.

Nesse momento, eu decidi criar um Data Warehouse em MySQL – que é mais simples de configurar e manipular que o SQL Server, além de ser gratuito –, extrair os dados da base original do sistema e carregar nele. Eu listei todos os dados que eu iria precisar criando um modelo conceitual dos dados a serem consumidos e solicitei apoio ao suporte do sistema para localizar na base de dados as tabelas que continham esses dados.

Com o DW já criado, identificadas as tabelas e selecionados os atributos dos dados que eu iria extrair, eu criei um Script em Linguagem Python que executava as seguintes tarefas:

- Conectava à Base de Dados no SQL Server

- Acessava as tabelas onde os dados brutos estravam contidos

- Extraía dessas tabelas para Data Frames do Pandas somente os atributos dos dados que eu havia selecionado

- Limpou os dados, converteu eles para os tipos de dados corretos para as análises e renomeou as colunas das tabelas com nomes mais inteligíveis para as análises

- Exportou os dados transformados nos Data Frames para arquivos no formato .csv como backups e para posterior conferência

- Conectou ao Data Warehouse no MySQL e carregou os dados transformados dos Data Frames direto nele como tabelas de dados estruturados

Com os dados formatados carregados no Data Warehouse eu passei a ter os dados prontos para carregá-los ao Power BI ou até mesmo no Excel – via Power Query – para realizar todas as análises financeiras que eu precisasse realizar sem precisar perder tempo manipulando planilhas extraídas direto do sistema. Além disso, para ter os dados atualizados, bastava apenas reexecutar o script.

Competências: Excel, SQL, Python, ETL, Power BI.