**DIÁRIO SOBRE O PRIMEIRO PROJETO ENVOLVENDO COMPETÊNCIAS**

**Introdução:** Neste diário, vou relatar minha jornada enquanto trabalhava em um projeto de programação orientada a objetos (POO) e ciência de dados com o uso das ferramentas Power BI e Python. Ao longo desse processo, enfrentei diversos desafios, desde questões lógicas até aspectos relacionados à infraestrutura do sistema operacional. Esta documentação busca registrar minhas experiências, compartilhar aprendizados e destacar soluções encontradas para os problemas encontrados. Vamos começar!

**Dia 1: Configuração do Ambiente e Exploração Inicial**

* Configuração do ambiente de desenvolvimento com as ferramentas necessárias: Python, VSCode, Anaconda, SQL Developer, SQL Data Modeler, Data Stage, Redshift e Power BI.
* Instalação do Oracle Database Express Edition (Oracle XE) como um banco de dados gratuito para fins de desenvolvimento e aprendizado.
* Detalhes sobre como baixar e instalar o Oracle XE, incluindo os parâmetros de conexão necessários para acessar o banco de dados.
* Início do planejamento do design da aplicação em Python, considerando a conexão com o Oracle e a lógica de inserção nas tabelas de produtos e compras.

**Dia 2: Desenvolvimento do Código SQL e Conexão ao Oracle**

* Problemas encontrados ao importar códigos e solução encontrada ao manter todos os arquivos no mesmo diretório.
* Necessidade de conceder as permissões adequadas no Oracle XE, incluindo as permissões necessárias para executar as stored procedures e CRUD (Create, Read, Update, Delete).
* Criação de um usuário no Oracle XE e inserção das stored procedures dentro desse usuário para evitar erros de "callproc".
* Execução de consultas no banco de dados Oracle usando a biblioteca cx\_Oracle em Python para obter dados da tabela "t\_produto\_vendido".

**Dia 3: Problemas de Processamento e Análise Exploratória de Dados**

* Exploração dos datasets disponíveis e ideias para criar um sistema de recomendação personalizado usando técnicas como filtragem colaborativa ou filtragem baseada em conteúdo.
* Identificação de problemas de processamento devido ao grande número de dados dummy durante a análise de cluster após as etapas de PCA (Principal Component Analysis) e StandardScaler.
* Testes realizados para otimizar o processamento, removendo dados irrelevantes para a análise de cluster.
* Conclusão de etapas relacionadas ao classificador e sistema de recomendação em POO, assim como a análise visual no Power BI.

**Dia 4: Análise de Dados e Estatísticas Descritivas**

* Utilização de gráficos (Displot, Boxplot, Barplot) para realizar análises de dados exploratórias dentro do sistema de compras em Python.
* Conceitos de variância de padrão amostral e desvio padrão, bem como sua aplicação em comparações e análises de dados.
* Uso do groupby(), std() e loc[] para calcular o desvio padrão de acordo com diferentes categorias.
* Discussão sobre probabilidade, amostragem de dados e teste de hipóteses com base na distribuição dos dados.

**Dia 5: Conexão do Power BI com o Oracle e Finalização do Projeto**

* Configuração da conexão do Power BI com o Oracle usando o Oracle ODBC Driver Configuration e a definição de parâmetros como usuário, senha e TNS Service Name.
* Alternativas de conexão entre o Power BI e o Oracle, incluindo o uso da linguagem M e a execução de comandos SQL diretamente.
* Finalização do projeto com o desenvolvimento do dash finalizado e criação de apresentações visualmente atraentes com o Power BI.

**Conclusão:** Durante o desenvolvimento deste projeto, enfrentei diversos desafios relacionados à configuração do ambiente, conexão com o Oracle, processamento de dados e análise exploratória. No entanto, com persistência e busca por soluções, consegui superar esses obstáculos e concluir o projeto com sucesso. Aprendi muito sobre POO, ciência de dados e o uso das ferramentas Power BI e Python, e agora tenho um sistema funcional de vendas e um dashboard interativo para análise de dados. Estou animado para continuar aprendendo e explorando novas possibilidades nessa área fascinante da tecnologia.

Base de Dados Kaggle: <https://www.kaggle.com/datasets/nosbielcs/brazilian-delivery-center>

Linkedin: <https://www.linkedin.com/in/perfil-pedro-batista-t%C3%A9cnico/>