Telas para MVP - ajuste

Escopo do MVP

Interface gráfica do usuário, Texto, Aplicativo, Email

Descrição gerada automaticamente

Interface gráfica do usuário, Texto, Aplicativo, Email

Descrição gerada automaticamente

Interface gráfica do usuário, Texto, Aplicativo, Carta, Email

Descrição gerada automaticamente

Interface gráfica do usuário, Texto, Aplicativo, Email

Descrição gerada automaticamente

Interface gráfica do usuário, Texto, Aplicativo

Descrição gerada automaticamente

Interface gráfica do usuário, Texto

Descrição gerada automaticamente

Realizei uma tentativa de utilizar imagens capturadas de telas e incluir no notebook do Databricks mas a imagem não ficou nítida. Daí resolvi utilizar o Word para inserir as imagens

Interface gráfica do usuário, Texto, Aplicativo

Descrição gerada automaticamente

Inclusão da imagem no notebook

Interface gráfica do usuário, Texto, Aplicativo

Descrição gerada automaticamente

Objetivo e Detalhamento: feitos no Notebook

**Azure Databricks**

Azure Databricks é uma plataforma de análise Apache Spark baseada em nuvem, integrada com o Azure para fornecer uma solução única para big data e análise avançada. Ele foi projetado em colaboração entre Microsoft e Databricks, otimizado para o Microsoft Azure. O serviço é destinado a simplificar o processo de criação, configuração e gerenciamento de ambientes Apache Spark para análise de dados e processamento de fluxo de trabalho.

**Iniciando o Workspace no Azure Databricks**

Um workspace no Azure Databricks é uma área de trabalho colaborativa onde os usuários podem criar notebooks, clusters, Jobs e bibliotecas. É um espaço dedicado para uma organização, e dentro dele, diferentes usuários e grupos podem ter permissões variadas para acessar e modificar recursos.

Ao clicar em "Iniciar o Workspace", acesso o workspace que criou (**WS\_Teste**).

Uma vez dentro do workspace, é possível:

1. **Criar e gerenciar clusters Spark**: Provisionar e gerenciar clusters Apache Spark que podem ser usados para executar notebooks, jobs e outros fluxos de trabalho.
2. **Desenvolver notebooks**: Os notebooks são documentos interativos que permitem escrever, executar e visualizar código, bem como visualizar resultados, gráficos e tabelas.
3. **Agendar e gerenciar jobs**: Jobs são tarefas ou séries de tarefas a fim de criar agenda para serem executadas em um cluster.
4. **Instalar bibliotecas**: É possível adicionar bibliotecas de terceiros ou criar próprias para expandir as funcionalidades dos notebooks e jobs.
5. **Gerenciar permissões e colaborar**: É possível conceder permissões para diferentes usuários e grupos dentro do workspace, permitindo colaboração e controle de acesso.

Interface gráfica do usuário, Texto, Aplicativo, Word

Descrição gerada automaticamente

Acessar o Databricks

Interface gráfica do usuário, Texto, Aplicativo, Email

Descrição gerada automaticamente

Coleta:

Ingestão de tabelas para o Azure Databricks

Interface gráfica do usuário, Aplicativo

Descrição gerada automaticamente

A ingestão de dados é uma etapa fundamental ao trabalhar com qualquer plataforma de análise de dados, incluindo o Azure Databricks. Vamos mergulhar um pouco mais na área de ingestão de dados no Databricks, especificamente a ingestão de arquivos CSV.

### Ingestão de Dados no Azure Databricks

No contexto do Azure Databricks, a ingestão refere-se ao processo de importar dados para serem utilizados nos clusters Spark. Esta etapa é essencial porque, antes de poder analisar ou processar dados com o Databricks, você primeiro precisa ter os dados disponíveis na plataforma.

#### **Adicionando Arquivos CSV**

A interface do Azure Databricks permite uma ingestão fácil de arquivos, incluindo arquivos CSV:

1. **Arrastar e Soltar**: O Databricks oferece uma interface intuitiva que permite aos usuários simplesmente arrastar e soltar arquivos CSV (ou outros formatos compatíveis) de seus computadores ou diretamente de fontes de dados conectadas, facilitando a importação.
2. **Criação de Tabelas**: Uma vez que os arquivos CSV são carregados no ambiente do Databricks, você pode transformá-los em tabelas. Ao criar uma tabela a partir de um CSV, você tem a flexibilidade de escolher o esquema, especificar opções de leitura, e muito mais.
3. **Opções de Formatação**: Databricks permite ajustar várias opções durante a ingestão, como delimitadores, codificação de caracteres, tratamento de linhas de cabeçalho, entre outros. Isso garante que os dados sejam lidos corretamente, mantendo a integridade e a formatação desejada.
4. **Integração com Azure Blob Storage e ADLS**: Além de permitir que você carregue arquivos diretamente, o Azure Databricks integra-se perfeitamente com outras soluções de armazenamento de dados da Microsoft, como Azure Blob Storage e Azure Data Lake Storage (ADLS). Isso facilita a ingestão de grandes volumes de dados ou a integração com pipelines de dados existentes.

### Relatório:

**Título**: Ingestão de Dados no Azure Databricks

**Resumo**: A plataforma Azure Databricks fornece uma interface user-friendly para a ingestão de dados, facilitando a importação de arquivos CSV e sua transformação em tabelas utilizáveis. Os usuários podem simplesmente arrastar e soltar arquivos, definir opções de formatação e integrar-se com soluções de armazenamento em nuvem da Microsoft para uma experiência de ingestão de dados flexível e eficiente.

**Detalhes da tela**: [Descreva os elementos visuais específicos da tela, como botões, ícones, áreas de texto, opções disponíveis e quaisquer mensagens ou status exibidos.]

Espero que isso ajude na criação do seu relatório! Se você precisar de informações adicionais ou detalhes sobre a área de ingestão ou qualquer outro aspecto do Azure Databricks, sinta-se à vontade para perguntar.

As tabelas seguintes foram criadas conforme abaixo:

Tela de computador com texto preto sobre fundo branco

Descrição gerada automaticamente

Catálogo:

Dados das colunas com seu tipo e comentário

Tela de computador com texto preto sobre fundo branco

Descrição gerada automaticamente

Tela de computador com texto preto sobre fundo branco

Descrição gerada automaticamente

Dados de exemplo:  
Tela de computador com texto preto sobre fundo branco

Descrição gerada automaticamente

Tela de computador com texto preto sobre fundo branco

Descrição gerada automaticamente

Cluster que estamos usando

Interface gráfica do usuário, Texto, Aplicativo

Descrição gerada automaticamente

O processo de criação do notebook foi feito por aqui e já apresentou a tabela

Tela de computador com texto preto sobre fundo branco

Descrição gerada automaticamente

Interface gráfica do usuário, Aplicativo

Descrição gerada automaticamente