

Data Force

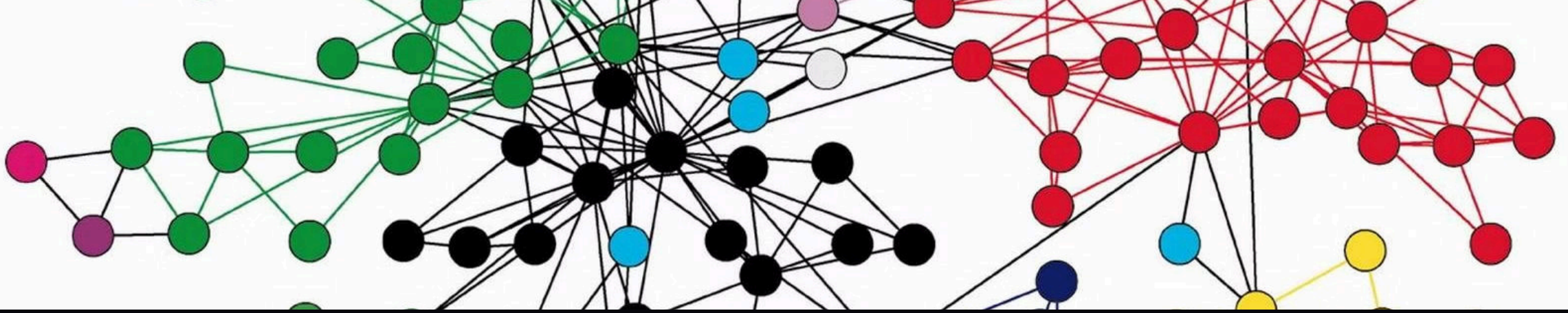
Anna Júlia - 01631263

Kawã Bazante - 01616235

Nicole Galvão - 01602503

Eduardo Oliveira - 01618738

José Paulo - 01622179



Escolha dos Dados

Decidimos trabalhar com um conjunto de dados do Kaggle, selecionando **Amazon Prime Userbase Dataset**, que nos proporcionou uma gama rica de informações para explorar e analisar.

Escolha da Tecnologia NoSQL

Opções Avaliadas

Durante o processo de Escolha da Tecnologia NoSQL, conduzimos uma pesquisa abrangente sobre as diversas opções disponíveis, que incluem Documento, Colunar, Chave e Valor, e Grafo. Avaliamos cuidadosamente os pontos positivos e negativos de cada tecnologia, levando em consideração os requisitos específicos do projeto em questão.

Tecnologia Selecionada: Grafo

Após uma análise ponderada, concluímos que a tecnologia NoSQL mais apropriada para atender às necessidades do projeto é o ****Grafo****. Essa decisão foi baseada em diversos fatores, como sua flexibilidade de esquema e capacidade de lidar eficientemente com dados semiestruturados, entre outros aspectos relevantes que contribuem para a eficácia e sucesso do projeto.

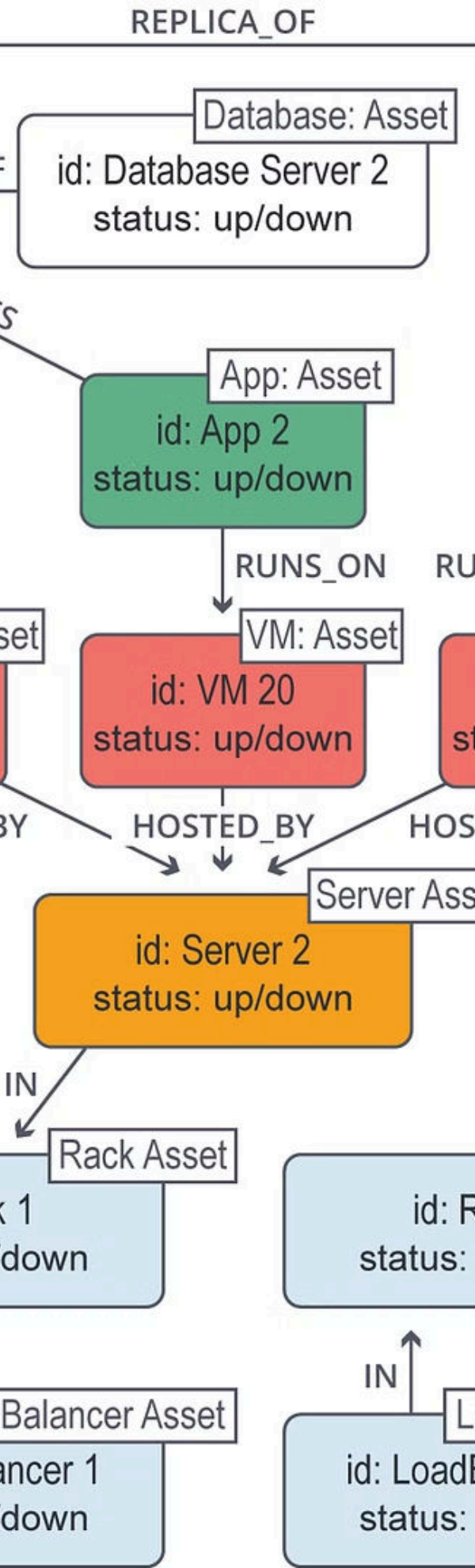
Benefícios do Grafo

A tecnologia de banco de dados de Grafo foi selecionada devido à sua capacidade de representar e analisar facilmente as conexões entre os dados, tornando-se eficiente para descobrir padrões e caminhos de conexão, mesmo em grandes conjuntos de dados. Essa característica se alinha perfeitamente com os requisitos do projeto, permitindo uma compreensão mais profunda das relações entre os elementos do conjunto de dados.

Configuração Inicial

Estabelecemos o repositório DATA FORCE no GitHub como o epicentro para reunir todos os elementos do nosso projeto. Estruturamos cuidadosamente suas seções para abranger não apenas códigos, mas também documentações e outros recursos cruciais. Todos os membros do grupo foram devidamente designados como colaboradores, o que possibilitou o acesso e a participação ativa de cada um no desenvolvimento do projeto.

Além disso, implementamos uma série de ferramentas indispensáveis, como o MySQL Workbench é uma IDE específica para o desenvolvimento de código. Essas medidas visam assegurar um ambiente de trabalho coeso e propício à produtividade, permitindo que todos os envolvidos tenham as condições ideais para contribuir eficazmente para o sucesso do projeto.



Modelagem do Banco de Dados Relacional

1

Análise Minuciosa do Conjunto de Dados

Na etapa de Modelagem do Banco de Dados Relacional, dedicamos tempo e esforço para construir um modelo de dados lógico robusto, utilizando a plataforma MySQL Workbench. Esse processo envolveu uma análise minuciosa do conjunto de dados, identificando entidades, atributos e relacionamentos do conjunto de dados ****Amazon Prime Userbase Dataset****.

2

Aplicação das Regras de Normalização

Adotamos as regras de normalização estudadas em sala de aula para garantir que os dados fossem organizados de maneira eficiente, minimizando redundâncias e evitando anomalias. Esse procedimento visa não apenas garantir a precisão da representação das informações, mas também facilitar a compreensão e a manutenção do banco de dados ao longo do tempo.

3

Documentação Detalhada do Modelo

Além disso, documentamos meticulosamente o modelo de dados em um dicionário detalhado, onde cada entidade e atributo foi descrito minuciosamente, juntamente com suas relações e restrições correspondentes. Esse procedimento visa facilitar a compreensão e a manutenção do banco de dados ao longo do tempo.

Modelagem do Banco de Dados Lógico

Ao elaborar o script do modelo lógico, foi utilizado recursos de aprendizado online, como vídeo aulas, a fim de facilitar a compreensão e agilizar o processo de produção. Durante esse processo, focou-se em identificar as chaves primárias e estabelecer as relações necessárias entre as entidades. Utilizou-se também a plataforma BR Modelo Web para a construção do modelo lógico, aproveitando suas funcionalidades para iterar e ajustar o design até alcançar o resultado desejado.

As chaves primárias de cada entidade foram anotadas cuidadosamente, garantindo a integridade e eficiência do modelo. Além disso, foram estabelecidas as conexões necessárias entre as entidades, construindo relações claras e coerentes para representar adequadamente a estrutura e o fluxo de dados do sistema.

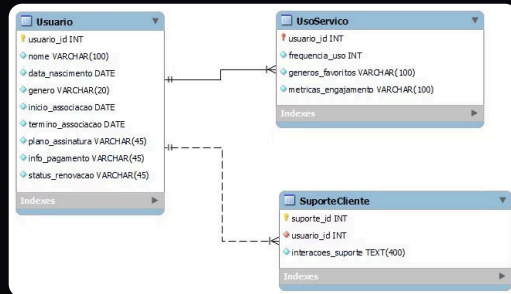
Com o auxílio da plataforma BR Modelo Web, foi possível visualizar e manipular o modelo lógico de forma interativa, realizando ajustes conforme necessário para otimizar a organização e a eficácia do banco de dados. A capacidade de iterar rapidamente sobre o design foi crucial para refinar o modelo até alcançar o resultado final desejado.

Extração e Carga dos Dados no Banco de Dados NoSQL

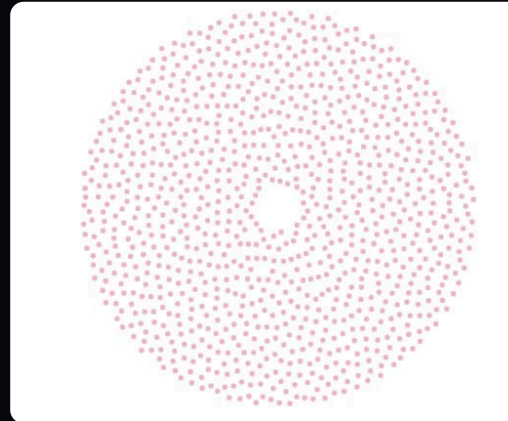
Na etapa de Extração e Carga dos Dados no Banco de Dados NoSQL, empregamos os scripts que desenvolvemos anteriormente para extrair os dados do banco de dados relacional MySQL e importá-los para o banco de dados NoSQL, utilizando Python como a ferramenta principal. Durante esse processo, foi necessário transformar os dados para que se adequassem ao modelo de dados NoSQL e executar operações de carga em larga escala.

Após a conclusão da carga dos dados, realizamos uma validação minuciosa para garantir sua integridade. Isso incluiu a comparação dos resultados obtidos com os dados originais e a realização de testes abrangentes de integridade. Esse procedimento rigoroso é essencial para assegurar a qualidade e a consistência dos dados no novo ambiente NoSQL, proporcionando uma base sólida para as etapas subsequentes do projeto.

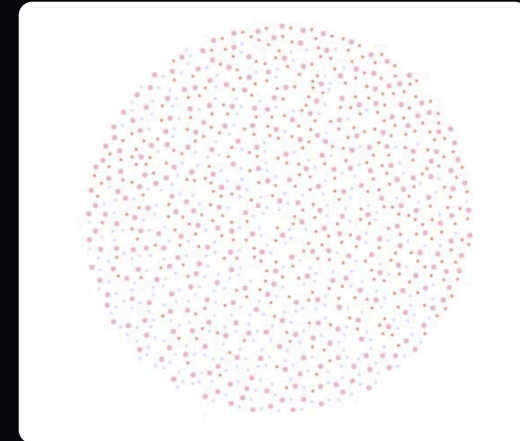
Repositório de Imagens



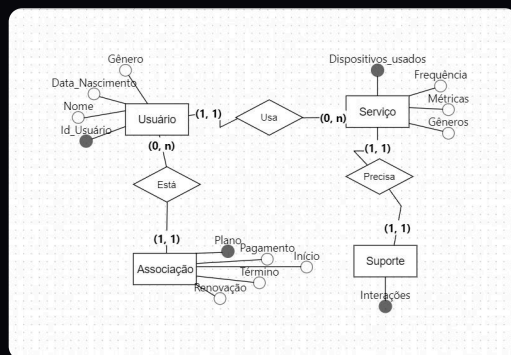
Modelo Relacional



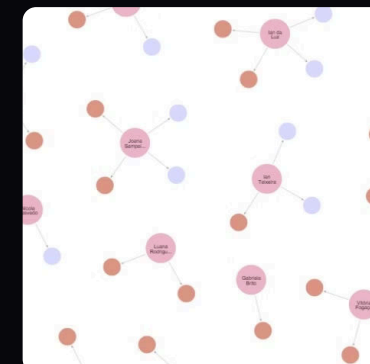
GRAFO, USANDO O: **MATCH**
(n) RETURN n



GRAFO, usando: **MATCH**
(u:Usuario)-[r]->(n) RETURN
u, r, n



Modelo Lógico



Conclusão

Com grande satisfação, a equipe DATA FORCE concluiu de forma bem-sucedida o projeto de aplicação de banco de dados, garantindo o cumprimento de todos os critérios estabelecidos pela atividade de avaliação. Agradecemos a oportunidade de aprendizado.

Fim