Criar apresentação para o repositório - README

Este projeto foi usado docker com WSL2 para contruir a infraestrutura de clusters.

Instalação Docker - Windows

Link para instalação do Docker Desktop no **Windows**

* [https://hub.docker.com/editions/community/docker-ce-desktop-windows/ (Links para um site externo.)](https://hub.docker.com/editions/community/docker-ce-desktop-windows/)

Link de Atualização Windows 10 (20.04)

* [https://www.microsoft.com/pt-br/software-download/windows10 (Links para um site externo.)](https://www.microsoft.com/pt-br/software-download/windows10)

Download WSL 2

* <https://wslstorestorage.blob.core.windows.net/wslblob/wsl_update_x64.msi>

Instalação Docker - Mac

Link para instalação do Docker Desktop no **Mac**

* <https://hub.docker.com/editions/community/docker-ce-desktop-mac/>

Instalação Docker - Linux

Link para instalação do Docker Engine

* [https://docs.docker.com/engine/install/ (Links para um site externo.)](https://docs.docker.com/engine/install/)

Link para instalação do Docker Compose

* <https://docs.docker.com/compose/install/>

\* Baixar conteúdo do Cluster

git clone https://github.com/rodrigo-reboucas/dockerbigdata.git spark

\* Baixar as imagens

docker-compose -f docker-compose-completo.yml pull

\* Listar as imagens

docker image ls

\* Iniciar todos os serviÃ§os

docker-compose -f docker-compose-completo.yml up -d

\* Baixar conteúdo do Cluster

$ git clone https://github.com/rodrigo-reboucas/dockerbigdata.git

\* Baixar Cluster de Big Data - Parcial

Arquivo: docker-compose-parcial.yml

â¢ Remover ~8GB

o Baixar as imagens

docker-compose -f docker-compose-parcial.yml pull

o Listar as imagens

docker image ls

o Iniciar todos os serviÃ§os

docker-compose -f docker-compose-parcial.yml up -d

Ponto 2 do exercício

Otimizar todos os dados do hdfs para uma tabela Hive particionada por município

Aqui pensando em fazer uma contagem para saber quantos municípios tem e assim definir o numero de particionamentos

Quantos municípios e quais, para assim particionar.

Pronto 3 do exercício

Criar as 3 visualizações pelo Spark com os dados enviados para o HDFS: