

Desafio de Projeto: Coleta e Processamento de dados com Power BI

Formação Power BI Analyst

Juliana Mascarenhas

Tech Education Specialist DIO / Owner @Simplificandoredes
e @SimplificandoProgramação

Mestre em modelagem computacional | Cientista de dados

@in/juliana-mascarenhas-ds/

Objetivo Geral

- Configurar o setup de banco de dados na Azure
- Popular o servidor com script fornecido
- Integre o MySQL com Power BI
- Realize as transformações indicadas
- Vá além....

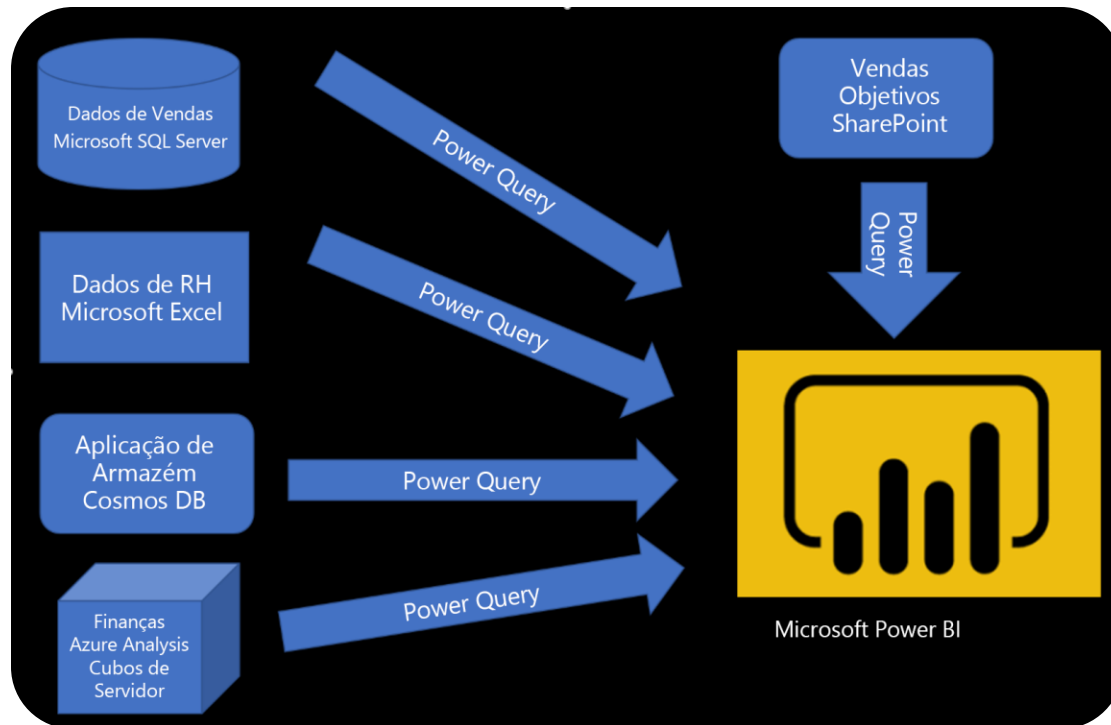


Etapa 1

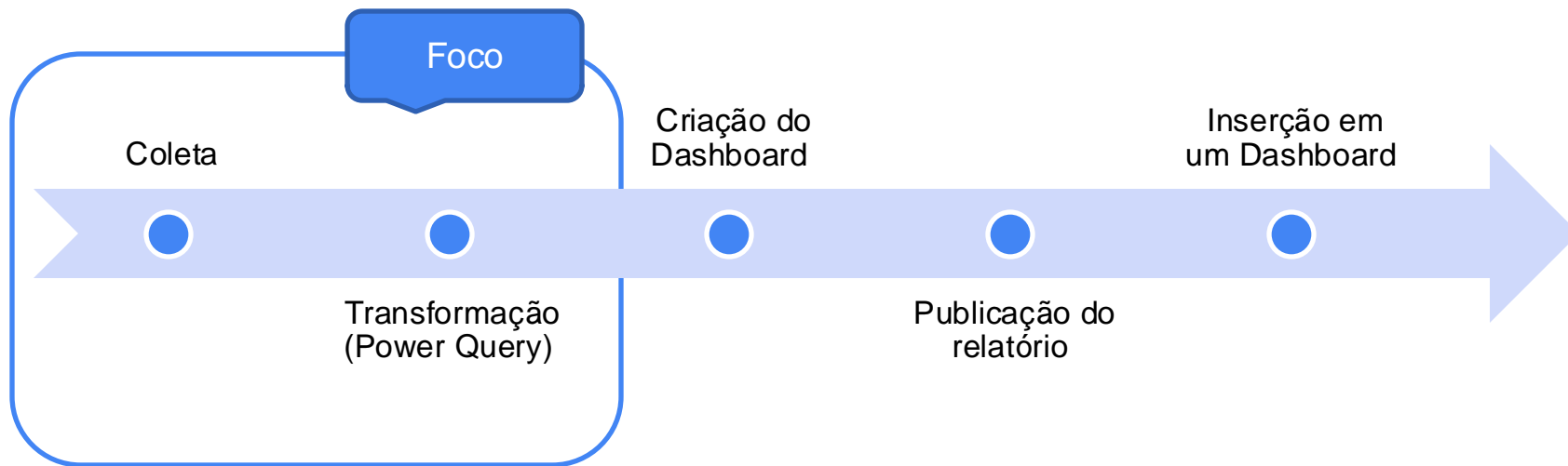
Transformando dados com Power BI

// Power BI Analyst

Realidade – Fontes Heterogêneas



Fluxo de Dados no Power BI



O que fazer?

Base de dados (teste) Company

Crie um relatório para verificar as informações e possíveis anomalias

- Criação de uma instância na Azure para MySQL
- Criar o Banco de dados com base disponível no github
- Integração do Power BI com MySQL no Azure
- Verificar problemas na base a fim de realizar a transformação dos dados

Transformando Dados

Passos a serem verificados na base do Power BI

- Verifique os cabeçalhos e tipos de dados
- Modifique os valores monetários para o tipo double preciso
- Verifique a existência dos nulos e analise a remoção
- Os employees com nulos em Super_ssn podem ser os gerentes.
Verifique se há algum colaborador sem gerente

Transformando Dados

Passos a serem verificados na base do Power BI

- Verifique se há algum departamento sem gerente
- Se houver departamento sem gerente, suponha que você possui os dados e preencha as lacunas
- Verifique o número de horas dos projetos
- Separar colunas complexas

Transformando Dados

Passos a serem verificados na base do Power BI

- Mesclar consultas employee e departament para criar uma tabela employee com o nome dos departamentos associados aos colaboradores. A mescla terá como base a tabela employee.
- **OBS:** Fique atento, essa informação influencia no tipo de junção
- Neste processo elimine as colunas desnecessárias.

Transformando Dados

Passos a serem verificados na base do Power BI

- Realize a junção dos colaboradores e respectivos nomes dos gerentes. Isso pode ser feito com consulta SQL ou pela mescla de tabelas com Power BI. Caso utilize SQL, especifique no README a query utilizada no processo.
- Mescle as colunas de Nome e Sobrenome para ter apenas uma coluna definindo os nomes dos colaboradores

Transformando Dados

Passos a serem verificados na base do Power BI

- Mescle os nomes de departamentos e localização. Isso fará que cada combinação departamento-local seja único. Isso irá auxiliar na criação do modelo estrela em um módulo futuro.
- Explique por que, neste caso supracitado, podemos apenas utilizar o mesclar e não o atribuir.

Transformando Dados

A^B_C Gerente	A^B_C Colaborador
3 distinto(s), 0 exclusivo(s)	9 distinto(s), 9 exclusivo(s)
	John Smith
John Smith	Amanda Teixeira
	Franklin Wong
Franklin Wong	Joyce English
Franklin Wong	Ramesh Narayan
	James Borg
John Smith	Jennifer Wallace
John Smith	Ahmad Jabbar
Franklin Wong	Alicia Zelaya

Passos a serem verificados na base do Power BI

A^B_C Store	t^2_3 Dnumber	A^B_C Mgr_ssn
5 distinto(s), 5 exclusivo(s)	3 distinto(s), 2 exclusivo(s)	3 distinto(s), 2 exclusivo(s)
Houston Headquarters	1	888665555
Stafford Administration	4	987654321
Bellaire Research	5	333445555
Houston Research	5	333445555
Sugarland Research	5	333445555

Transformando Dados

Passos a serem verificados na base do Power BI

- Agrupe os dados a fim de saber quantos colaboradores existem por gerente
- Elimine as colunas desnecessárias, que não serão usadas no relatório, de cada tabela