



BlueShift

Business Blueprint

Caieiras, 06 de julho de 2022

CONTROLE DE VERSÃO			
Autor	Versão	Data	Descrição
Felipe Gustavo de Abreu	1.0	05/07/2022	Criação do Documento.

1. Introdução

Este documento visa detalhar as necessidades do cliente PoccoBank, do ponto de vista técnico referente ao procedimento de extração da cotação do câmbio diretamente do Banco Central, e convertê-los e armazená-los diariamente em uma tabela final.

2. Solicitação

O Cliente PoccoBank deseja uma solução para extração e processamento de dados para gerar alguns relatórios semanais para seus principais investidores, e para isso eles necessitam da cotação diária do dólar. A entrega do projeto deverá ser dividida em três partes. A primeira consistirá na extração via Azure Data Factory para o Azure SQL database. A segunda parte consistirá na modelagem dos dados, contendo a conversão dos mesmos, via Stored Procedure dentro do Azure Data Factory. A terceira e última parte consistirá na cópia dos dados do Azure SQL databases para o Blob como Parquet.

3. Premissas da Solução

A seção abaixo apresenta as premissas da solução

Origem e especificação dos dados

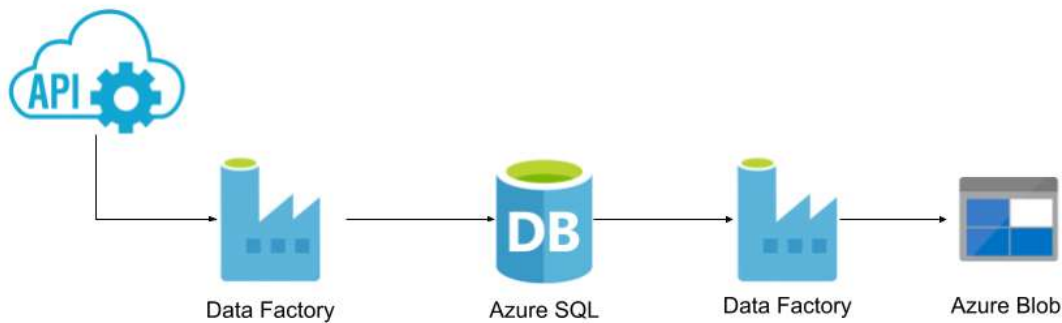
- Os dados estão sendo disponibilizados por uma API pelo Banco Central.

Ambiente de desenvolvimento

- O cliente deverá disponibilizar acessos aos ambientes de desenvolvimento em todas as ferramentas específicas na arquitetura proposta neste documento para a Blueshift Brasil.

4. Modelo da arquitetura sugerida

A imagem abaixo apresenta a arquitetura da solução proposta com base no levantamento de requisitos e entendimento do negócio.



5. Extração dos dados da API no Azure Data Factory

Será desenvolvido uma Pipeline no Azure Data Factory, e nele será inserido primeiramente um Copy Data do link da API informado abaixo, na qual irá receber os dados da cotação diretamente do Banco Central. Após o consumo dos dados da API, essas informações serão inseridas no banco de dados criado no Azure SQL.

- API do Banco Central para coleta dos dados:

[https://olinda.bcb.gov.br/olinda/servico/PTAX/versao/v1/odata/CotacaoDolarPeriodo\(dataInicial=@dataInicial,dataFinalCotacao=@dataFinalCotacao\)?@dataInicial='01-01-2019'&@dataFinalCotacao='12-31-2025'&\\$top=9000&\\$format=text/csv&\\$select=cotacaoCompra,cotacaoVenda,dataHoraCotacao](https://olinda.bcb.gov.br/olinda/servico/PTAX/versao/v1/odata/CotacaoDolarPeriodo(dataInicial=@dataInicial,dataFinalCotacao=@dataFinalCotacao)?@dataInicial='01-01-2019'&@dataFinalCotacao='12-31-2025'&$top=9000&$format=text/csv&$select=cotacaoCompra,cotacaoVenda,dataHoraCotacao)

6. Modelagem dos Dados

Para este projeto estamos considerando uma tabela de stage para armazenar os dados extraídos da API, e uma tabela final que irão receber os dados convertidos através da Storage Procedure. A estrutura das tabelas estão listadas abaixo.

- Tabela Stage:

Campo	Tipo
cotacaoCompra	varchar(40)
cotacaoVenda	varchar(40)
dataHoraCotacao	varchar(40)

- Tabela Final:

Campo	Tipo
cotacaoCompra	float
cotacaoVenda	float
dataHoraCotacao	datetime

7. Storage Procedure

Os dados extraídos da API foram inseridos no banco de dados no formato “VARCHAR”, na qual será necessário fazer a conversão dos dados. Para realizar a conversão, será criada uma “STORED PROCEDURE” no Azure SQL Database, e introduzindo-a no Data Factory após o primeiro Copy Data. A Storage Procedure irá converter os dados da tabela Stage que estão todos em varchar para os respectivos formatos mostrados abaixo, e que serão inseridos na tabela final.

Campo	Tabela Stage	Tabela Final
cotacaoCompra	varchar(40)	float
cotacaoVenda	varchar(40)	float
dataHoraCotacao	varchar(40)	datetime

8. Azure Storage

A última etapa do projeto, será criado o segundo Copy Data e será adicionado após o Stored Procedure na pipeline, sendo este Copy Data com a função de copiar os dados da Tabela Final já convertida para o Azure Storage no formato Parquet.