MVP SPRINT II – ENGENHARIA DE DADOS

PEDRO ZAPPA

Descrição

Breve análise das falências bancárias nos EUA, comparativamente ao caso mais recente de falência do SVB.

Objetivo

O texto abaixo foi construído a partir de perguntas ao BARD, ferramenta de chat generativo do Google:

O Silicon Valley Bank (SVB) era um banco comercial de destaque, sediado em Santa Clara, Califórnia, conhecido por focar em tecnologia, ciências da vida e clientes de capital de risco. Ele oferecia uma variedade de serviços financeiros, incluindo empréstimos, tesouraria, gestão de caixa, banco de investimento, entre outros.

O banco ganhou destaque pela sua estreita relação com startups, inovadores e a indústria de tecnologia no Vale do Silício. Era reconhecido por compreender as necessidades financeiras únicas das empresas de tecnologia e empresas emergentes, muitas vezes fornecendo suporte financeiro e orientações adaptadas às suas necessidades específicas.

O SVB faliu em 09/03/2023. Foi o maior banco a falir nos Estados Unidos desde a crise financeira de 2008.

A falência do SVB foi causada por uma combinação de fatores, mas o principal fator de destaque foi a aposta do banco uma estratégia de investimento agressiva, que incluía a compra de títulos de longo prazo com taxas de juros fixas. Quando o FED começou a aumentar as taxas de juros, o valor desses títulos caiu, o que levou a um aumento das perdas financeiras do banco.

A falência do SVB causou um impacto significativo no ecossistema de startups dos Estados Unidos. O banco era um importante fornecedor de financiamento e serviços para startups, e sua falência deixou muitas startups sem acesso a capital.

Impacto financeiro: O SVB tinha mais de US\$ 200 bilhões em ativos, o que o torna um dos maiores bancos a falir na história dos Estados Unidos. A falência do banco causou perdas significativas para seus clientes, incluindo investidores, startups e empresas de capital de risco.

A falência do SVB levantou preocupações sobre a saúde do sistema bancário dos Estados Unidos.

O objetivo do trabalho é confirmar as seguintes informações, extraídas do texto acima:

- Qual foi a representatividade da falência do SVB em relação a outros casos nos EUA no período de 30 anos de análise?
- Qual o tamanho do SVB em termos de ativos e depósitos?
- Qual a concentração de ativos e depósitos do SVB em relação ao total de ativos e depósitos? Ela é representativa?
- Analisando exclusivamente este caso isolado, devemos ter preocupações sobre a saúde bancária dos EUA?

Detalhamento

1. Busca pelos dados

A base de dados utilizada neste trabalho foi obtida em https://banks.data.fdic.gov/bankfind-suite/ e Financial Institutions - Catalog (data.gov)

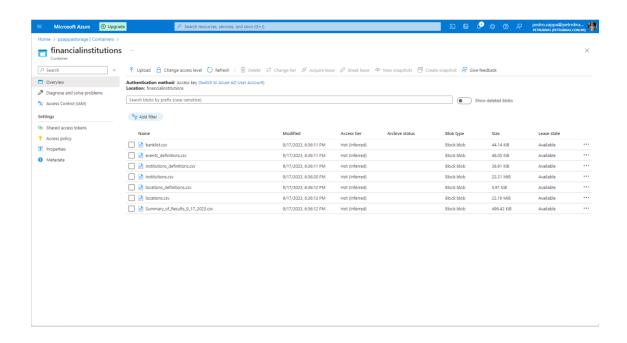
Foram obtidos dois datasets:

- a. Bank Failures & Assistance Data: dataset no qual encontramos informações sobre falências e eventos de assistência a bancos americados e – origem da nossa tabela de fatos;
- b. Financial Institutions: dataset em que encontramos detalhes sobre as instituições financeiras origem de nossa tabela de dimensões.

Além disso, foi obtido um dataset de auxílio com detalhes das colunas do dataset de Financial Institutions, os chamamos de institutions_definitions. Não houve necessidade de dataset auxiliar para o conjunto de dados obtido para o dataset Bank Failures & Assistance Data.

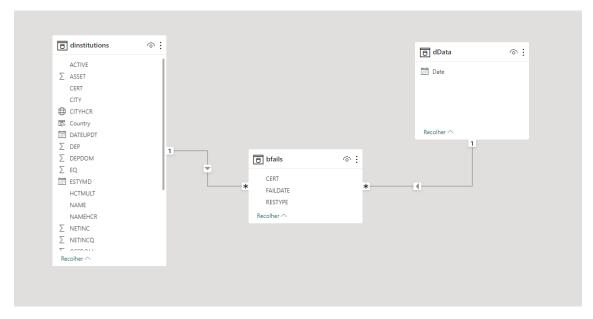
2. Coleta

Os dados foram baixados para o notebook pessoal e subidos para um container no Microsoft Azure Blob Storage, conforme pode ser visto no print a seguir:



3. Modelagem

Foi criado um modelo de dados simples em esquema estrela, conforme a seguir:



Catálogo de Dados

Tabela	Nome da Coluna	Descrição	Tipo	Formato	Dados de Referência	Propriedades	Relacionamentos	Mantenedor	Disponibilidade	Linhagem/Proveniência
dInstitutions	ACTIVE	Situação da instituição	bit		null ou 1	Um número que indica a situação de uma instituição. 1='Instituições que estão atualmente abertas e seguradas pelo FDIC'		FDIC	Público	Financial Institutions - Catalog (data.gov)
dInstitutions	ASSET	Ativos totais	int		USD mil	A soma de todos os ativos de propriedade da instituição, incluindo caixa, empréstimos, títulos, instalações bancárias e outros ativos.		FDIC	Público	Financial Institutions - Catalog (data.gov)
dInstitutions	CERT	Número de certificado FDIC	smallint			Um número exclusivo atribuído pelo FDIC usado para identificar instituições e para a emissão de certificados de seguro.	PK - Relacionado com: bFails	FDIC	Público	Financial Institutions - Catalog (data.gov)
dInstitutions	CITY	Cidade	nvarchar			A cidade na qual uma instituição ou agência está fisicamente localizada.		FDIC	Público	Financial Institutions - Catalog (data.gov)
dInstitutions	CITYHCR	Cidade do detentor principal	nvarchar			Cidade na qual a sede do detentor principal regulatório da instituição está fisicamente localizada.		FDIC	Público	Financial Institutions - Catalog (data.gov)
dInstitutions	DATEUPDT	Última atualização	date	MM/DD/AAAA		A data da última atualização de dados.		FDIC	Público	Financial Institutions - Catalog (data.gov)
dinstitutions	DEP	Depósitos totais	int		USD mil	A soma de todos os depósitos		FDIC	Público	Financial Institutions - Catalog (data.gov)
dinstitutions	DEPDOM	Depósitos mantidos em escritórios domésticos	int		USD mil	A soma de todos os depósitos em escritórios domésticos		FDIC	Público	Financial Institutions - Catalog (data.gov)
dInstitutions	EQ	Capital próprio	int		USD mil	Capital próprio total (inclui ações preferenciais e ordinárias, superávit e lucros não distribuídos).		FDIC	Público	Financial Institutions - Catalog (data.gov)
dinstitutions	ESTYMD	Data de estabelecimento	date	MM/DD/AAAA		A data na qual a instituição ou agência iniciou suas operações.		FDIC	Público	Financial Institutions - Catalog (data.gov)
dinstitutions	HCTMULT	Tipo de holding bancária	bit		0 ou 1	Um sinalizador usado para indicar se uma instituição é membro de uma holding bancária múltipla 1=sim, 0=não		FDIC	Público	Financial Institutions - Catalog (data.gov)
dinstitutions	NAME	Nome da instituição	nvarchar			O título legal ou nome da instituição.		FDIC	Público	Financial Institutions - Catalog (data.gov)
dInstitutions	NAMEHCR	Holding bancária (detentor principal regulatório)	nvarchar			O detentor principal regulatório é atribuído pelo Federal Reserve Board com base nas porcentagens de propriedade e controle.		FDIC	Público	Financial Institutions - Catalog (data.gov)
dInstitutions	NETINC	Renda líquida	int			Renda líquida de juros mais receita total não relacionada a juros mais ganhos (perdas) realizados sobre títulos e itens extraordinários, menos despesa total não relacionada a juros, provisões para perdas com empréstimos e impostos de renda.		FDIC	Público	<u>Financial Institutions - Catalog</u> (data.gov)
dinstitutions	NETINCQ	Renda líquida - trimestral	int			Renda líquida de juros trimestral		FDIC	Público	Financial Institutions - Catalog (data.gov)

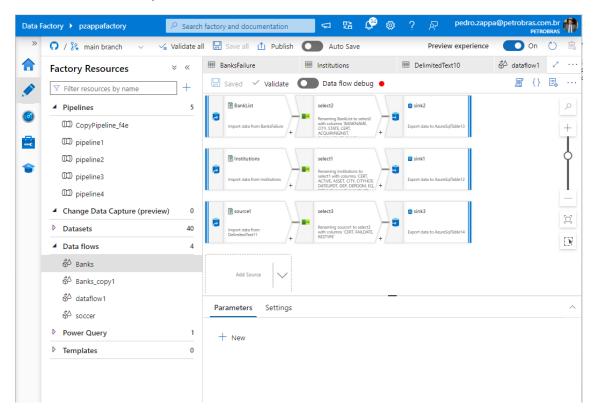
dInstitutions	OFFDOM	Número de escritórios domésticos	smallint			O número de escritórios domésticos (incluindo a sede) operados por instituições ativas nos 50 estados dos EUA.		FDIC	Público	Financial Institutions - Catalog (data.gov)
dInstitutions	OFFFOR	Número de escritórios estrangeiros	smallint			O número de escritórios estrangeiros (fora dos EUA) operados pela instituição.		FDIC	Público	Financial Institutions - Catalog (data.gov)
dInstitutions	OFFICES	Escritório	smallint			Número total de escritórios		FDIC	Público	Financial Institutions - Catalog (data.gov)
dinstitutions	ROE	Retorno sobre o patrimônio líquido (ROE)	float		% ou NA caso negativo	Renda líquida anualizada como um percentual do patrimônio líquido médio em uma base consolidada.		FDIC	Público	Financial Institutions - Catalog (data.gov)
dinstitutions	STALP	Código alfa do estado	nvarchar			A abreviatura do estado da localização da sede da instituição.		FDIC	Público	Financial Institutions - Catalog (data.gov)
dinstitutions	STNAME	Nome do estado	nvarchar			Estado em que a instituição ou agência está fisicamente localizada.		FDIC	Público	Financial Institutions - Catalog (data.gov)
dInstitutions	STNUM	Número do estado	smallint			O código do Padrão Federal de Processamento de Informações usado para identificar estados.		FDIC	Público	Financial Institutions - Catalog (data.gov)
dinstitutions	ZIP	Código postal	int			Código postal completo representando a localização física da instituição ou de uma de suas agências.		FDIC	Público	Financial Institutions - Catalog (data.gov)
bFails	CERT	Número de certificado FDIC	smallint			Um número exclusivo atribuído pelo FDIC usado para identificar instituições e para a emissão de certificados de seguro.	FK - Relacionado com: dInstitutions	FDIC	Público	Financial Institutions - Catalog (data.gov)
bFails	FAILDATE	Data do evento de falha	date	MM/DD/AAAA		Data em que ocorreu o evento de falha do banco		FDIC	Público	https://banks.data.fdic.gov/bankfind- suite/
bFails	RESTYPE	Tipo do evento	nvarchar		ASSISTANCE ou FAILURE	Pode assumir os seguintes valores "ASSISTANCE" ou "FAILURE"		FDIC	Público	https://banks.data.fdic.gov/bankfindsuite/

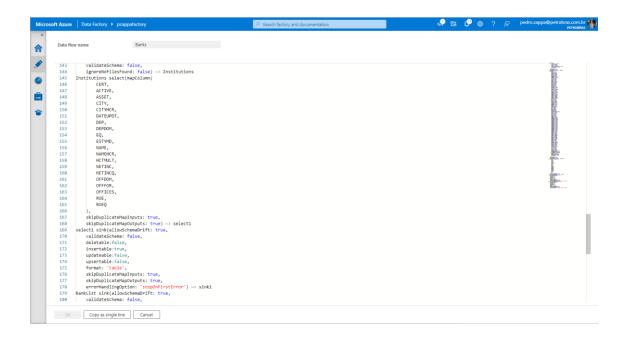
4. Carga

Foi escolhido o Azure SQL Database para realizar a carga de dados.

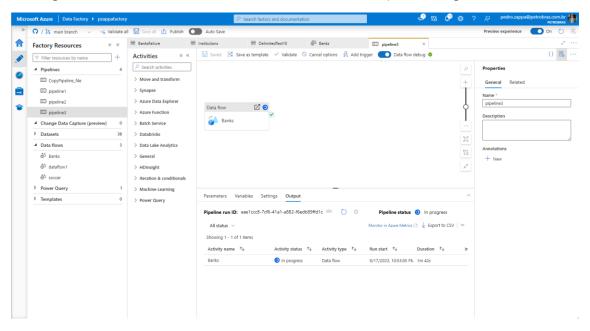
A ferramenta de ETL (Extração, Transformação e Carga) Data Factory foi realizada para realizar a extração, transformação e carga, conforme prints a seguir.

As principais transformações feitas foram a formatação das colunas (integer, varchar, date etc) e o drop de colunas indesejadas. De um total de 127 colunas no dataset Institutions, escolhemos 23 e do dataset Banks Failures, do total de 13 colunas, escolhemos 3.

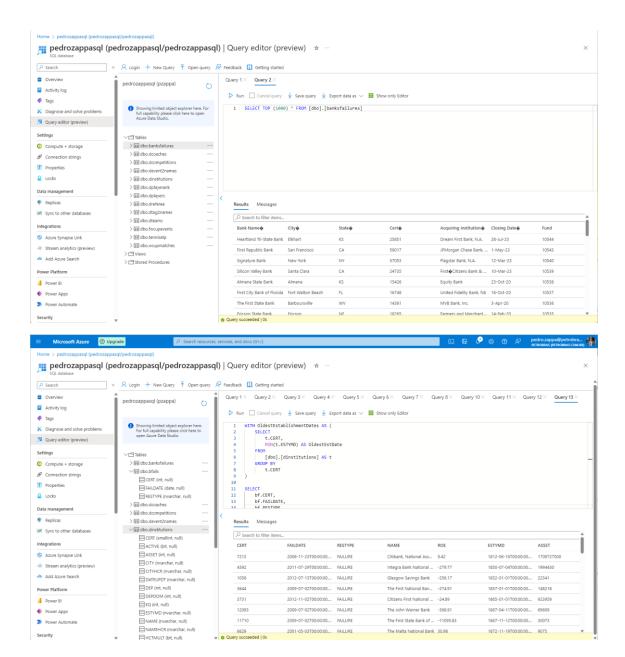




A carga então foi feita conforme fluxo de dados do print a seguir:



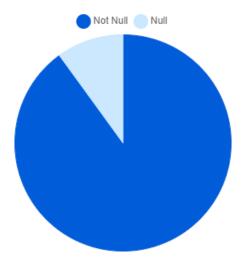
Exemplos de consultas SQL de teste no banco de dados Azure SQL:



5. Análise

a. Qualidade de dados

Identificamos a ausência de valores na coluna CERT da tabela que denominamos **bfails**, conforme gráfico a seguir:



Esses códigos são foreign key para a tabela dInstitutions e necessários para identificar nomes das instituições financeiras, assim como outras dimensões importantes para a análise.

Uma possível solução seria identificar outra chave no conjunto de dados que pudesse ser utilizada para criar o relacionamento entre ambas as tabelas, mas não encontramos tal chave.

Com relação à tabela dinstitutions, de um total de 27.822 registros:

```
Query 1 × Query 2 × Query 3 ×

D Run □ Cancel query ₺ Save query ₺ Export data as ∨ ₩ Show only Editor

1 SELECT COUNT(*) FROM dinstitutions;
2
3

Results Messages

D Search to filter items...

27822
```

Encontramos valores nulos nas colunas ASSET e DEP.

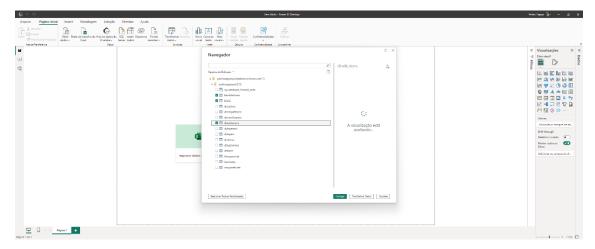


Os valores contidos em ASSET e DEP serão utilizados para medir a representatividade dos bancos onde ocorreram falências e são importantes para o modelo.

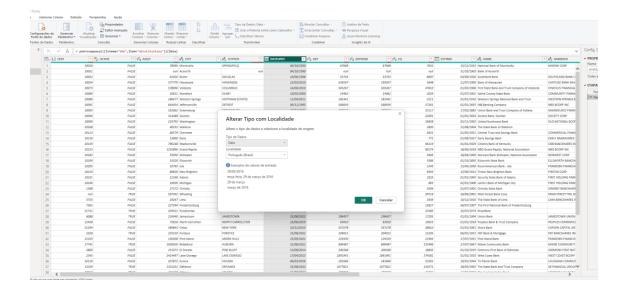
Com exceção da questão dos dados nulos, não foram verificados em ambas as tabelas problemas de qualidade de dados. Datas estão em formato correto, assim como nomes e chaves de identificação.

b. Solução do problema

Para a solução do problema, iremos utilizar um relatório de PowerBi. As tabelas foram importadas diretamente do Azure SQL, conforme telas a seguir:



E foi feita novamente a etapa de ETL dentro do PowerBi para ajustes pontuais:



O relatório será criado com as seguintes visualizações:

- Tendência de Falência de Bancos ao Longo do Tempo: um gráfico de linhas mostrando a tendência das falências de bancos ao longo do tempo, usando o eixo x como data e a contagem de falências de bancos no eixo y.
- Falências de Bancos por Estado/Cidade: uma visualização de mapa para mostrar a distribuição das falências de bancos por estado e cidade.
- Ativos do Banco versus Depósitos: um gráfico de dispersão para comparar "ativos" e "depósitos" de cada banco, com "ativos" em um eixo, "depósitos" em outro e o "nome do banco" como rótulo
- Resumo de Falências de Bancos: relatório resumido com uma tabela exibindo o "nome do banco," "data de constituição," "data de fechamento" e "ativos para cada falência de banco.

Foram também adicionadas interações e filtros apropriados para permitir que os usuários explorem e analisem os dados de forma dinâmica.



Autoavaliação

- Como pudemos verificar pelo painel criado, a falência do SVB foi representativa em relação a outros casos no período de 30 anos de análise. O SVB é o segundo maior banco neste período a falir, ficando logo atrás do Washington Mutual Bank no gráfico de dispersão Ativos x Depósitos.
- O SVB possuía ativos na ordem de USD 210 bilhões e depósitos de pouco mais de 175 bilhões. No entanto, nota-se que o mercado americano de bancos é bem desconcentrado, com vários bancos locais, o que torna a preocupação com a saúde bancária americana menor face a um evento isolado.
- Qual a concentração de ativos e depósitos do SVB em relação ao total de ativos e depósitos? Ela é representativa? Não conseguimos responder esta pergunta, pois não obtivemos dados anuais de ativos e depósitos por banco, portanto não teríamos a informação a ser analisada ano a ano.

Futuros trabalhos recomendáveis seriam incorporar outras bases de dados que podem ser obtidas em https://banks.data.fdic.gov/bankfind-suite/ para responder à última pergunta, assim como dar maiores insights sobre risco de contágio de crises entre bancos americanos, como por exemplo comparação de peers.