**Documentação processos de extração de dados SQL**

**Nome extração: Tabela fato – Providências**

**1.Objetivo da Extração**

O objetivo desta extração é criar uma tabela fato para ser utilizada no projeto de BI do setor Jurídico Imobiliário do escritório. A extração foi realizada a partir da tabela de providências históricas (T00076) e tem como principal finalidade monitorar a performance dos advogados do setor, com base em suas providências ou tarefas sistêmicas. Os dados extraídos incluem a quantidade de providências realizadas e os tempos relativos a esses procedimentos, permitindo uma análise detalhada da eficiência e produtividade dos advogados.

**2. Descrição do Ambiente**

* **Banco de Dados:** SQL Server - ramaprod.
* **Gateway:** Power BI Gateway conectado ao Power BI Service.
* **Ferramenta de Extração:** Fluxo de Dados do Power BI.
* **Permissões:** A consulta SQL tem acesso apenas às colunas e tabelas liberadas dentro do gateway.

**3. Fontes de Dados**

* **Sistema de Extração:** Sistema Espaider Rama
* **Periodicidade da Extração/Atualização:** Extração diária agendada para às 18:30

A consulta faz uso das seguintes tabelas e colunas do banco de dados:

**Tabela: [ramaprod].[dbo].T00076 (Alias: a)**

* Colunas Selecionadas e Chaves:
  + F00445: Utilizada na cláusula CASE para determinar a coluna situacao
  + F05342: Selecionada como a coluna data
  + F00453: Utilizada para calcular data\_inicio e na cláusula HAVING
  + F17598: Utilizada para calcular data\_inicio\_contagem e dias\_efetivos
  + F00454: Selecionada como a coluna prazo
  + F00455: Utilizada para calcular data\_final, dias\_mesclado, e dias\_atrasados
  + F06337: Selecionada como as colunas cumprido\_em, providencia\_futura, e cumprido\_em\_hora
  + F00451: Utilizada para calcular dias\_mesclado, dias\_efetivos, e dias\_atrasados
  + F05633: Selecionada como a coluna resp e utilizada em JOIN com a tabela [ramaprod].[dbo].T00557
  + F06982: Chave utilizada em JOIN com a tabela [ramaprod].[dbo].T00041
  + F00447: Chave utilizada em JOIN com a tabela [ramaprod].[dbo].T00077
  + F05341: Chave utilizada em JOIN com a tabela [ramaprod].[dbo].T00003

**Tabela: [ramaprod].[dbo].T00077 (Alias: b)**

* Colunas Selecionadas e Chaves:
  + F00446: Selecionada como a coluna providencia
  + ID: Chave primária da tabela [ramaprod].[dbo].T00077, utilizada em JOIN com a coluna F00447 da tabela [ramaprod].[dbo].T00076

**Tabela: [ramaprod].[dbo].T00041 (Alias: c)**

* Colunas Selecionadas e Chaves:
  + F14474: Selecionada como a coluna dossie
  + ID: Chave primária da tabela [ramaprod].[dbo].T00041, utilizada em JOIN com a coluna F06982 da tabela [ramaprod].[dbo].T00076
  + F01187: Utilizada na cláusula HAVING para filtrar os resultados

**Tabela: [ramaprod].[dbo].T00003 (Alias: d)**

* Colunas Selecionadas e Chaves:
  + F00689: Selecionada como a coluna criado\_por
  + ID: Chave primária da tabela [ramaprod].[dbo].T00003, utilizada em JOIN com a coluna F05341 da tabela [ramaprod].[dbo].T00076

**Tabela: [ramaprod].[dbo].T00557 (Alias: e)**

* Colunas Selecionadas e Chaves:
  + F05200: Chave utilizada em JOIN com a tabela [ramaprod].[dbo].T00030
  + ID: Chave primária da tabela [ramaprod].[dbo].T00557, utilizada em JOIN com a coluna F05633 da tabela [ramaprod].[dbo].T00076

**Tabela: [ramaprod].[dbo].T00030 (Alias: f)**

* Colunas Selecionadas e Chaves:
  + F00091: Selecionada como a coluna resp2 e utilizada para calcular providencia\_futura e tempo\_entre\_prov
  + ID: Chave primária da tabela [ramaprod].[dbo].T00030, utilizada em JOIN com a coluna F05200 da tabela [ramaprod].[dbo].T00557

**4. Código SQL Completo**

*SELECT*

*c.F14474 AS dossie,*

*a.F05633 AS resp,*

*f.F00091 AS resp2,*

*b.F00446 AS providencia,*

*CASE*

*WHEN MAX(a.F00445) = 1 THEN 'Inserido'*

*WHEN MAX(a.F00445) = 2 THEN 'Sugestão'*

*WHEN MAX(a.F00445) = 3 THEN 'A cumprir'*

*WHEN MAX(a.F00445) = 4 THEN 'Cumprido'*

*WHEN MAX(a.F00445) = 5 THEN 'Cancelado'*

*END AS situacao,*

*a.F05342 AS data,*

*MAX(d.F00689) AS criado\_por,*

*CAST(MAX(a.F00453) AS DATE) AS data\_inicio,*

*CAST(MAX(a.F17598) AS DATE) AS data\_inicio\_contagem,*

*MAX(a.F00454) AS prazo,*

*CAST(MAX(a.F00455) AS DATE) AS data\_final,*

*a.F06337 AS cumprido\_em,*

*LAG(a.F06337) OVER (PARTITION BY f.F00091 ORDER BY a.F06337) AS providencia\_futura,*

*DATEDIFF(*

*SECOND,*

*LAG(a.F06337) OVER (PARTITION BY f.F00091 ORDER BY a.F06337),*

*a.F06337*

*) AS tempo\_entre\_prov,*

*CAST(MAX(a.F06337) AS TIME) AS cumprido\_em\_hora,*

*CASE*

*WHEN MAX(a.F00445) = 4 THEN DATEDIFF(DAY,MAX(a.F00455),MAX(a.F00451))*

*WHEN MAX(a.F00445) = 3 THEN DATEDIFF(DAY,MAX(a.F00455),GETDATE())*

*END AS dias\_mesclado,*

*DATEDIFF(DAY,MAX(a.F17598),MAX(a.F00451)) AS dias\_efetivos,*

*CASE*

*WHEN DATEDIFF(DAY, MAX(a.F00455), MAX(a.F00451)) < 0 THEN 0*

*ELSE DATEDIFF(DAY, MAX(a.F00455), MAX(a.F00451))*

*END AS dias\_atrasados*

*FROM [ramaprod].[dbo].T00076 AS a*

*LEFT JOIN [ramaprod].[dbo].T00077 AS b ON a.F00447 = b.ID*

*LEFT JOIN [ramaprod].[dbo].T00041 AS c ON a.F06982 = c.ID*

*LEFT JOIN [ramaprod].[dbo].T00003 AS d ON a.F05341 = d.ID*

*LEFT JOIN [ramaprod].[dbo].T00557 AS e ON a.F05633 = e.ID*

*LEFT JOIN [ramaprod].[dbo].T00030 AS f ON e.F05200 = f.ID*

*GROUP BY c.F14474, a.F05633,f.F00091, b.F00446, a.F05342, a.F06337*

*HAVING*

*YEAR(CAST(MAX(a.F00453) AS DATE)) >= 2022 AND*

*MAX(c.F01187) IN (32,34,14,36,35)*

*ORDER BY cumprido\_em DESC;*

**5. Destinos dos Dados Extraídos**

Os dados extraídos por meio desta consulta SQL são carregados diretamente no Power BI Service através do fluxo de dados configurado no Power BI. Esses dados são armazenados em datasets no Power BI, onde são utilizados em dashboards e relatórios para análise e visualização.

* **Destino:** Projeto de BI Jurídico Imobiliário.
* **Formato dos Dados:** Os dados são carregados em formato tabular no Power BI Service, prontos para visualização em relatórios e dashboards.

**6. Qualidade e Segurança**

Antes que os dados sejam considerados prontos para uso nos dashboards, são aplicados os seguintes critérios de qualidade:

* **Integridade dos Dados**
* A integridade dos dados é verificada e atestada através do banco de dados relacional, com uso das chaves primárias de cada tabela que participa de junções.
* Todo dataset oriundo da extração passa por uma análise exploratória para validação das colunas. Valores nulos e/ou inconsistentes são tratados da maneira que a regra de negócio do escopo permite.
* **Precisão dos Dados:**
* A precisão dos dados é checada através de validações diretamente no sistema Espaider Rama e através de planilhas de cada setor. Cada coluna é validada usando esses comparativos citados.
* **Versionamento:**
* Todos os códigos de extrações e documentações são armazenados localmente na pasta: C:\Users\pedro.cecere\Documents\Projetos\A - Jurídico Imobiliário\A - 06 - Queries BI. Esta pasta é adicionada remotamente ao repositório criado no GitHub: A-04-Queries\_BI.
* Qualquer alteração no código é postada no diretório do GitHub, onde temos acesso ao histórico de versões deste código, mantendo assim uma linha cronológica de todas versões do código desenvolvido.

Sobre segurança e permissões de acessos aos dados, são tomadas as seguintes medidas:

* **Acessos e Permissões:**
  + A extração de dados é realizada de acordo com as permissões estabelecidas pela Smart Suport, empresa terceirizada responsável por hospedar o banco de dados. O acesso é gerenciado através do gateway do Power BI, que controla quais colunas e tabelas podem ser acessadas. Somente usuários autorizados pelo administrador do gateway têm permissão para modificar ou executar essa consulta.
* **Conformidade com Políticas de Privacidade:**
  + Os dados extraídos são restritos a informações profissionais e não incluem dados sensíveis. Todo o processo de extração é realizado em conformidade com as normas e diretrizes da LGPD, bem como as exigências estabelecidas pelo Banco Santander.
* **Logs e Auditoria:**
  + O Power BI Service registra todas as modificações no fluxo de dados, detalhando quem fez as alterações e em que momento. Esses logs são mantidos para fins de auditoria e conformidade. Da mesma forma, o versionamento realizado no GitHub permite rastrear todas as alterações e identificar os responsáveis por cada modificação.
* **Encriptação de Dados:**
  + Os dados são encriptados, por padrão, em trânsito e em repouso no Power BI Service, garantindo que as informações estão protegidas contra acessos não autorizados.