**Documentação processos de extração de dados SQL**

**Nome extração: Tabela fato – Distribuição CRI**

1. **Objetivo da Extração**

O objetivo desta extração é criar uma tabela fato para alimentar o projeto de BI do setor Jurídico Imobiliário do escritório. A extração dos dados foi realizada a partir das tabelas de eventos históricos (T00069) e tem como intuito identificar os processos ativos que não foram distribuídos dentro do mesmo mês de cadastro. Esse indicador é exigido pelo banco Santander. Os filtros aplicados incluem a seleção de processos ativos que tenham o evento "AF 0.7" e que não possuam o evento "AF 1.1" dentro do mesmo mês. Além disso, foram excluídos os processos que estão em fase de pagamento ou suspensão. Vale ressaltar que existe uma margem de cinco dias para trás, a partir do último dia do mês, que corresponde à meta estipulada de tempo médio (TM) para distribuição. Portanto, processos cadastrados nos últimos cinco dias do mês não são considerados nesta consulta.

**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].T00069 (Alias: a)**

* Colunas selecionadas e chaves:
  + F14474 - Dossiê (dossie)
  + F00385 - Datas de eventos históricos (utilizada para calcular datas dos eventos AF e DP)
  + F04461 - Pasta (pasta)
  + F08501 - Chave usada para realizar JOIN com a tabela T00003 (utilizada em JOIN)
  + F01133 - Chave usada para realizar JOIN com a tabela T00064 (utilizada em JOIN)
  + F02003 - Chave usada para realizar JOIN com as tabelas T00041, T00037, T00030, e T00003 (utilizada em JOIN)

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

* Colunas selecionadas e chaves:
  + ID - Chave primária usada para JOIN com T00069 (utilizada em JOIN)

**Tabela: [ramaprod].[dbo].T00064 (Alias: c)**

* Colunas selecionadas e chaves:
  + F01132 - Descrição do evento histórico (usada para identificar e categorizar os eventos AF e DP)
  + ID - Chave primária usada para JOIN com T00069 (utilizada em JOIN)

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

* Colunas selecionadas e chaves:
  + F14474 - Dossiê (dossie)
  + F35249 - Data de atualização no Benner (data\_atualizacao\_benner)
  + F01187 - Identificador do status do processo (utilizada para filtro na cláusula HAVING)
  + ID - Chave primária usada para JOIN com T00069 (utilizada em JOIN)

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

* Colunas selecionadas e chaves:
  + F01187 - Identificador do status do processo (utilizada para filtro na cláusula HAVING)
  + ID - Chave primária usada para JOIN com T00069 (utilizada em JOIN)

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

* Colunas selecionadas e chaves:
  + F25017 - Status do processo (situacao)
  + F43687 - Chave usada para realizar JOIN com a tabela T02682 (utilizada em JOIN)
  + F00177 - Chave usada para realizar JOIN com a tabela T00037 (utilizada em JOIN)
  + F11578 - Chave usada para realizar JOIN com a tabela T00030 (utilizada em JOIN)
  + ID - Chave primária usada para JOIN com T00069 (utilizada em JOIN)

**Tabela: [ramaprod].[dbo].T02682 (Alias: g)**

* Colunas selecionadas e chaves:
  + F43686 - Subfase do processo (subfase)
  + ID - Chave primária usada para JOIN com T00041 (utilizada em JOIN)

**Tabela: [ramaprod].[dbo].T00037 (Alias: h)**

* Colunas selecionadas e chaves:
  + F00162 - Fase do processo (fase)
  + ID - Chave primária usada para JOIN com T00041 (utilizada em JOIN)

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

* Colunas selecionadas e chaves:
  + F00091 - Advogado interno responsável (advogado\_interno)
  + ID - Chave primária usada para JOIN com T00041 (utilizada em JOIN)

**Tabela: [ramaprod].[dbo].T00064 (Alias: j)**

* Colunas selecionadas e chaves:
  + ID - Chave primária usada para JOIN com T00069 (utilizada em JOIN)

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

* Colunas selecionadas e chaves:
  + ID - Chave primária usada para JOIN com T00069 (utilizada em JOIN)

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

* Colunas selecionadas e chaves:
  + ID - Chave primária usada para JOIN com T00069 (utilizada em JOIN)

**4. Código SQL Completo**

*WITH subquery AS (*

*SELECT*

*d.F14474 AS dossie,*

*MAX(f.F35249) AS data\_atualizacao\_benner,*

*MAX(a.F04461) AS pasta,*

*MAX (*

*CASE*

*WHEN f.F25017 = 1 THEN 'Ativo'*

*WHEN f.F25017 = 2 THEN 'Encerrado'*

*WHEN f.F25017 = 3 THEN 'Acordo'*

*WHEN f.F25017 = 4 THEN 'Em encerramento'*

*ELSE 'Em precatório (Ativo)'*

*END ) AS situacao,*

*MAX (h.F00162) AS fase,*

*MAX (g.F43686) AS subfase,*

*MAX (i.F00091) AS advogado\_interno,*

*CAST (MAX(CASE WHEN c.F01132 = 'AF 0.7 - Pendência regularizada: Prosseguir Distribuição CRI' OR c.F01132 = 'Na esteira de ajuizamento com kit OK' THEN a.F00385 ELSE NULL END) AS DATE) AS 'AF\_07',*

*CAST (MAX(CASE WHEN c.F01132 = 'AF 1.1 - Distribuição via CRI' OR c.F01132 = 'Protocolo do Requerimento de Intimação' THEN a.F00385 ELSE NULL END) AS DATE) AS 'AF\_11',*

*CAST (MAX(CASE WHEN c.F01132 = 'AF 2.8 - Todas as Intimações Positivas' THEN a.F00385 ELSE NULL END) AS DATE) AS 'AF 2.8',*

*CAST (MAX(CASE WHEN c.F01132 = 'AF 5.1 - Consolidação da Propriedade' OR c.F01132 = 'DP 1.3.1 - Averbação da dação' THEN a.F00385 ELSE NULL END) AS DATE) AS 'AF 5.1',*

*CAST (MAX(CASE WHEN c.F01132 = 'AF 6.2 - Documentação aceita' OR c.F01132 = 'DP 1.5.2 - Documentação aceita' THEN a.F00385 ELSE NULL END) AS DATE) AS 'AF 6.2',*

*CAST (MAX(CASE WHEN c.F01132 = 'AF 6.1 - Envio do Imóvel ao Real Estate' THEN a.F00385 ELSE NULL END) AS DATE) AS 'AF 6.1',*

*CAST (MAX(CASE WHEN c.F01132 = 'AF 9.3 - Leilão Negativo' THEN a.F00385 ELSE NULL END) AS DATE) AS 'AF 9.3',*

*CAST (MAX(CASE WHEN c.F01132 = 'AF 12.1 - Pagamento' THEN a.F00385 ELSE NULL END) AS DATE) AS 'AF\_121'*

*FROM [ramaprod].[dbo].T00069 AS a*

*LEFT JOIN [ramaprod].[dbo].T00003 AS b ON a.F08501 = b.ID*

*LEFT JOIN [ramaprod].[dbo].T00064 AS c ON a.F01133 = c.ID*

*LEFT JOIN [ramaprod].[dbo].T00041 AS d ON a.F02003 = d.ID*

*LEFT JOIN [ramaprod].[dbo].T00041 AS e ON a.F02003 = e.ID*

*LEFT JOIN [ramaprod].[dbo].T00041 AS f ON a.F02003 = f.ID*

*LEFT JOIN [ramaprod].[dbo].T02682 AS g ON f.F43687 = g.ID*

*LEFT JOIN [ramaprod].[dbo].T00037 AS h ON f.F00177 = h.ID*

*LEFT JOIN [ramaprod].[dbo].T00030 AS i ON f.F11578 = i.ID*

*LEFT JOIN [ramaprod].[dbo].T00064 AS j ON a.F01133 = j.ID*

*LEFT JOIN [ramaprod].[dbo].T00041 AS k ON a.F02003 = k.ID*

*LEFT JOIN [ramaprod].[dbo].T00041 AS l ON a.F02003 = l.ID*

*GROUP BY a.F04461, d.F14474*

*HAVING*

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

*MAX (h.F00162) LIKE '%AF%' OR MAX (h.F00162) LIKE '%DP%' OR MAX (h.F00162) LIKE '%Alienação%'*

*)*

*SELECT*

*dossie,*

*AF\_07,*

*AF\_11*

*FROM subquery*

*WHERE*

*dossie IS NOT NULL AND*

*AF\_07 IS NOT NULL AND*

*(AF\_11 IS NULL OR MONTH(AF\_11) <> MONTH(AF\_07) OR YEAR(AF\_11) <> YEAR(AF\_07)) AND*

*AF\_121 IS NULL AND*

*AF\_07 <= DATEADD(DAY, -5, EOMONTH(AF\_07))*

*ORDER BY AF\_07 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.