

Sistema Integrado de Alertas Meteorológicos de Portugal

*ETL com Validação, Análise de Risco
e Geração Automática de Relatórios*

Integração de Sistemas de Informação

Trabalho Prático I

Tiago Ferreira
Número: 27980

Licenciatura em Engenharia de Sistemas Informáticos

Ano Letivo 2025/2026

19 de outubro de 2025

Conteúdo

1	Enquadramento	4
1.1	Identificação do Projeto	4
1.2	Contexto Académico	4
1.3	Repositório Git	4
2	Problema	5
2.1	Descrição Detalhada	5
2.2	Motivação e Relevância	5
3	Estratégia Utilizada	6
3.1	Abordagem Global	6
3.2	Operadores e Processos Envolvidos	6
3.3	Processos ETL Desenvolvidos	6
3.4	Ferramentas e Integrações	7
3.5	Metodologia de Desenvolvimento	7
3.6	Síntese da Estratégia	7
3.7	Componentes Tecnológicos	7
4	Transformações Detalhadas	8
4.1	T1 – Extração e Pré-Processamento (API IPMA)	8
4.1.1	Objetivo	8
4.1.2	Visão Geral	8
4.1.3	Componentes e Funções	8
4.1.4	Descrição das Etapas	9
4.1.5	Resultado	9
4.2	T2 – Validação e Normalização de Dados	10
4.2.1	Objetivo	10
4.2.2	Visão Geral	10
4.2.3	Componentes e Funções	10
4.2.4	Descrição das Etapas Principais	11
4.2.5	Resultado	11
4.3	T3 – Cálculo de Alertas e Relatórios	12
4.3.1	Objetivo	12
4.3.2	Visão Geral	12
4.3.3	Componentes e Funções	13
4.3.4	Descrição das Etapas Principais	13
4.3.5	Resultados	14
5	Job Principal – Orquestração do Processo ETL	17
5.1	Objetivo	17
5.2	Visão Geral	17
5.3	Componentes e Funções	17
5.4	Descrição das Etapas	18
5.4.1	Logs na Base de Dados	19
5.5	Resultado	19

6	Base de Dados	20
6.1	Visão Geral	20
6.2	Schema <code>meteorologia</code>	21
6.3	Schema <code>registos</code>	22
6.4	Integração com o Processo ETL	23
6.5	Resumo	23
7	Diagrama Entidade-Relacionamento (ER)	24
8	Visualização e Monitorização de Dados no Node-RED	25
9	Conclusão	28
9.1	Trabalhos Futuros	28
10	Demonstração do Projeto	29
11	Referências Bibliográficas	30

Lista de Figuras

1	Fluxo geral da transformação T1 – Extração e Pré-Processamento	8
2	Fluxo geral da transformação T2 – Validação e Normalização de Dados . .	10
3	Fluxo geral da transformação T3 – Cálculo de Alertas e Relatórios	12
4	Excerto do ficheiro ALERTAS_URGENTES.xls	15
5	Ficheiro Alerta_Resumo_Cidades.xls	16
6	Fluxo geral do Job Principal – Orquestração das transformações T1, T2 e T3	17
7	Configuração de logs	19
8	Ligação configurada no Pentaho ao PostgreSQL_MeteoDB	20
9	Estrutura geral do schema <code>meteorologia</code>	21
10	Tabela <code>logs</code> no schema <code>registos</code>	22
11	Diagrama Entidade-Relacionamento (ER) do sistema de base de dados. . .	24
12	Interface de alertas meteorológicos no <i>Node-RED</i>	25
13	Visualização de previsões meteorológicas.	26
14	Visualização de previsões meteorológicas.	26
15	Distribuição de alertas por região e nível.	27
16	Gráfico de temperaturas por cidade.	27
17	QR Code para Demonstração.	29

1 Enquadramento

1.1 Identificação do Projeto

Disciplina: Integração de Sistemas de Informação (ISI)

Curso: Licenciatura em Engenharia de Sistemas Informáticos

Ano Letivo: 2025/2026

Docente: Óscar Ribeiro

Aluno: Tiago Ferreira - N^o 27980

Data Entrega: 19 de outubro de 2025

1.2 Contexto Académico

Este trabalho prático insere-se no âmbito da Unidade Curricular de Integração de Sistemas de Informação e tem como objetivo principal a aplicação prática de ferramentas e processos de **ETL (Extract, Transform, Load)** para integração de sistemas ao nível dos dados. O projeto visa consolidar conhecimentos sobre processos de extração, transformação e carregamento de dados, bem como explorar tecnologias de integração de sistemas.

1.3 Repositório Git

Todo o código fonte, transformations Pentaho e documentação estão disponíveis no repositório:

https://github.com/tiagoferreira06/ISI_Fase1_AlertaMeteorologia

2 Problema

2.1 Descrição Detalhada

O projeto consiste no desenvolvimento de um **Sistema Integrado de Alertas Meteorológicos** que automatiza completamente o processo de:

1. **Extração** de dados meteorológicos em tempo real da API oficial do IPMA (Instituto Português do Mar e Atmosfera)
2. **Validação rigorosa** dos dados utilizando expressões regulares, verificação de duplicados e validação de intervalos numéricos
3. **Transformação** dos dados com cálculos de indicadores de risco.
4. **Geração automática** de relatórios por nível de alerta
5. **Persistência** estruturada em base de dados PostgreSQL
6. **Log completo** Logs guardados a cada execução.
7. **Visualização interativa** através de dashboard Node-RED

2.2 Motivação e Relevância

No contexto atual de alterações climáticas e eventos meteorológicos extremos cada vez mais frequentes, sistemas de alerta precoce são fundamentais para:

- **Proteção Civil:** Antecipar situações de risco e ativar planos de contingência
- **Saúde Pública:** Proteger populações vulneráveis (idosos, crianças) de ondas de calor/frio
- **Agricultura:** Planear atividades agrícolas e evitar perdas de colheitas
- **Turismo:** Informar operadores turísticos e visitantes
- **Gestão Municipal:** Apoiar decisões operacionais de câmaras municipais
- **Infraestruturas:** Prevenir danos em estradas, pontes e edifícios

3 Estratégia Utilizada

3.1 Abordagem Global

A estratégia adotada para o desenvolvimento deste projeto baseou-se numa arquitetura modular, organizada em processos de **ETL (Extract, Transform, Load)**. Cada componente do sistema foi construído de forma independente no **Pentaho Data Integration (PDI)**, garantindo clareza e capacidade de manutenção.

O processo completo é composto por três transformações principais (T1, T2 e T3) e um **Job Principal** que as orquestra. O objetivo foi assegurar uma execução sequencial e automatizada de todo o ciclo de tratamento de dados meteorológicos, desde a recolha até à geração de relatórios finais.

3.2 Operadores e Processos Envolvidos

Durante o desenvolvimento foram explorados diferentes operadores de transformação disponibilizados pelo PDI, distribuídos conforme o tipo de operação pretendida:

Tipo de Operação	Operadores Utilizados
Extração de Dados	REST Client, JSON Input, XML Input
Validação e Limpeza	Regex Evaluation, Filter Rows, JavaScript
Transformação e Cálculo	Modified JavaScript, Value Mapper, Calculator
Agregação e Consolidação	Group By, Sort Rows, Unique Rows
Integração e Armazenamento	Table Output (PostgreSQL)
Geração de Relatórios	Excel Output, CSV Output
Orquestração e Controlo	Job Executor, Check Files, Mail
Logs	Job Log Table

Tabela 1: Principais operadores utilizados no processo ETL e respetivas funções

3.3 Processos ETL Desenvolvidos

O sistema foi dividido em três transformações principais, cada uma com objetivos distintos:

- **T1 – Extração e Pré-Processamento:** consulta a API do IPMA, extrai dados meteorológicos em formato JSON, converte-os para XML e calcula indicadores complementares (temperatura média, converte para Kelvin, etc.);
- **T2 – Validação e Normalização:** aplica expressões regulares e verificações de integridade para garantir a qualidade dos registos antes da integração;
- **T3 – Análise e Cálculo de Alertas:** interpreta os dados validados, aplica regras lógicas para definir níveis de alerta, calcula pontuações de risco e gera relatórios em Excel.

O **Job Principal** controla a execução das três transformações, verifica a existência dos ficheiros de saída e envia automaticamente os relatórios por email ao utilizador.

3.4 Ferramentas e Integrações

O projeto recorreu a um conjunto de tecnologias integradas para cobrir todas as fases do processo:

- **Pentaho Data Integration (Kettle):** principal ferramenta de ETL e orquestração de jobs;
- **PostgreSQL + pgAdmin:** base de dados relacional para armazenamento;
- **IPMA REST API:** fonte oficial de dados meteorológicos;
- **Node-RED Dashboard:** visualização interativa dos alertas e previsões;
- **GitHub:** Armazenamento do projeto.

3.5 Metodologia de Desenvolvimento

A construção do sistema seguiu uma abordagem **iterativa e incremental**, testando e validando cada transformação isoladamente antes da sua integração. Foram utilizados dados reais da API IPMA e testes simulados com dados de exceção (valores inválidos ou nulos) para garantir robustez.

3.6 Síntese da Estratégia

A estratégia global garantiu:

- Verificação automática de erros e notificações por email;
- Persistência e rastreabilidade total no PostgreSQL;
- Geração de relatórios automatizada;
- Integração com dashboard interativo para análise visual.

Assim, a abordagem adotada cumpre integralmente os requisitos definidos no enunciado: integração de APIs externas, utilização de expressões regulares, operações de normalização, geração de logs, manipulação de múltiplos formatos e automatização de jobs.

3.7 Componentes Tecnológicos

Componente	Tecnologia
Ferramenta ETL	Pentaho Data Integration (Kettle) 9.4 CE
Base de Dados	PostgreSQL com pgAdmin
Linguagens	SQL, JavaScript, Expressões Regulares
API Externa	IPMA REST API (Instituto Português do Mar e Atmosfera)
Formatos Dados	JSON, XML, CSV, Excel (XLS)
Dashboard	Node-RED + node-red-dashboard
Controlo Versões	Git + GitHub
Sistema Operativo	Windows 10/11

Tabela 2: Componente Tecnológica do Projeto

4 Transformações Detalhadas

4.1 T1 – Extração e Pré-Processamento (API IPMA)

4.1.1 Objetivo

A transformação **T1** tem como objetivo extrair previsões meteorológicas da API do IPMA, validar e enriquecer os dados com cálculos e classificações automáticas, e armazená-los em ficheiro XML. Esta etapa constitui a base de todo o processo ETL.

4.1.2 Visão Geral

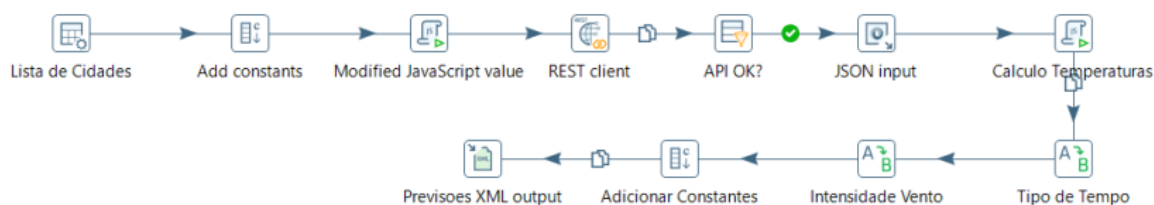


Figura 1: Fluxo geral da transformação T1 – Extração e Pré-Processamento

4.1.3 Componentes e Funções

Componente	Função
Data Grid	Lista de cidades IPMA
Modified JavaScript	Gera URLs API
REST Client	Faz chamadas HTTP
Filter Rows	Verifica resposta 200
JSON Input	Extrai campos JSON
Modified JavaScript	Calcula e classifica
Value Mapper	Traduz códigos tempo
Add Constants	Adiciona variaveis
XML Output	Cria ficheiro XML

Tabela 3: Componentes da T1 e respectivas funções resumidas

4.1.4 Descrição das Etapas

1. Extração e Validação da API IPMA A transformação inicia-se com a criação da lista de cidades (**Data Grid**), cada uma com o seu código IPMA. O componente **Modified JavaScript** gera dinamicamente as URLs de consulta, e o **REST Client** executa as chamadas HTTP. O **Filter Rows** garante que apenas respostas com estado 200 (OK) prosseguem.

2. Interpretação e Normalização dos Dados O **JSON Input** extrai os campos principais: data, temperaturas, precipitação, vento etc.

São utilizados caminhos (*JSONPath expressions*) para aceder aos campos do objeto devolvido pela API do IPMA.

Por exemplo: `.data[*].forecastDate`

3. Enriquecimento e Classificações No segundo **Modified JavaScript**, são calculadas novas métricas como temperatura média, converter para Kelvin e classificações (Frio, Quente, etc.). O **Value Mapper** traduz códigos meteorológicos em descrições legíveis, e o **Add Constants** adiciona a fonte dos dados e tipo de previsão.

4. Armazenamento Por fim, exporta resultados em ficheiro XML para uso em etapas seguintes.

4.1.5 Resultado

A T1 produz previsões meteorológicas limpas, normalizadas e enriquecidas com classificações automáticas, prontas para as fases de validação e análise subsequentes.

4.2 T2 – Validação e Normalização de Dados

4.2.1 Objetivo

A transformação **T2** tem como objetivo validar e normalizar os dados meteorológicos extraídos na T1, assegurando a consistência e qualidade dos registos antes da sua integração final. São aplicadas expressões regulares e correções automáticas sobre os valores importados.

4.2.2 Visão Geral

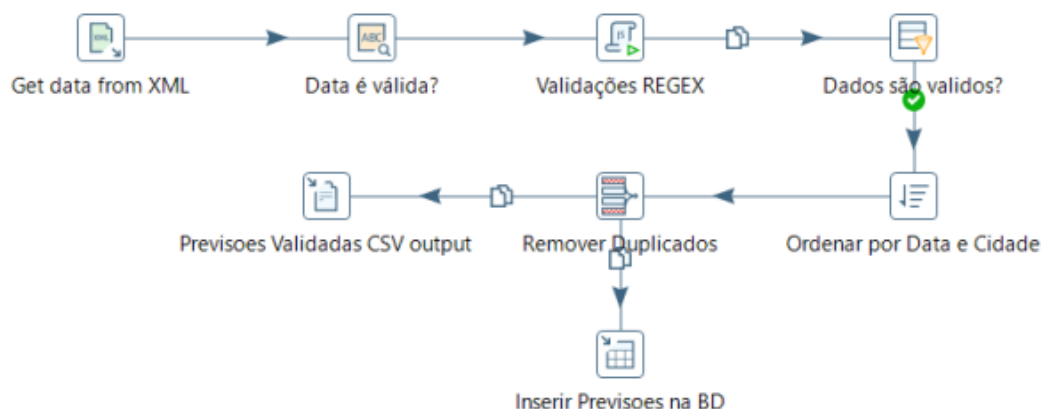


Figura 2: Fluxo geral da transformação T2 – Validação e Normalização de Dados

4.2.3 Componentes e Funções

Componente	Função
Get data from XML	Lê ficheiro XML de previsões
Regex Evaluation	Valida formato da data
Modified JavaScript	Aplica validações e normalizações
Filter Rows	Separa registos válidos e inválidos
Sort Rows	Ordena por data e cidade
Unique Rows	Remove duplicados
Table Output	Inserir dados validados na BD
CSV Output	Exporta previsões validadas

Tabela 4: Componentes da T2 e respetivas funções resumidas

4.2.4 Descrição das Etapas Principais

1. Leitura de Dados XML A transformação inicia-se com o componente **Get data from XML**, que importa os dados previamente exportados pela T1. O ficheiro XML contém as previsões meteorológicas estruturadas por cidade e data:

C:\Users\bntia\Desktop\Projeto_ISI\data\dados_ipma_previsoes.xml

2. Validação de Formato de Data (Regex Evaluation) Nesta etapa, o componente **Regex Evaluation** valida o campo `data_previsao` com a expressão regular:

```
^\d{4}-\d{2}-\d{2}$
```

Esta expressão assegura que o formato da data está conforme o padrão **YYYY-MM-DD**. O resultado é armazenado no campo `data_valida`, com valores booleanos (verdadeiro/-falso).

3. Validações e Normalizações (Modified JavaScript) O componente **Modified JavaScript** aplica várias verificações e transformações:

- **Nome da cidade:** Remoção de acentos e conversão para maiúsculas;
- **Temperaturas:** Verificação de intervalos plausíveis (-50°C a 60°C);
- **Precipitação:** Limitação a 0–100%;
- **Direção do vento:** Validação de siglas (N, S, E, W, NE, SW, etc.);
- **Estado geral:** Geração de campo `validacao_status` com valores “VALIDO” ou “INVALIDO”.

Exemplo de código aplicado:

```
1 // Validar temperaturas (entre -50 e 60)
2 var temp_valida = (temp_min >= -50 && temp_min <= 60 &&
3   temp_max >= -50 && temp_max <= 60 && temp_min <=
   temp_media && temp_media <= temp_max);
```

Listing 1: Validação da Temperatura

4. Filtragem e Ordenação O passo **Filter Rows** direciona apenas registos com `validacao_status = "VALIDO"` para a sequência principal. Estes são depois ordenados por data e cidade (**Sort Rows**) e verificados para duplicados (**Unique Rows**).

5. Saída e Registo de Resultados Os registos válidos são:

- Inseridos na base de dados através de **Table Output**;
- Exportados para CSV (**Previsoes Validadas CSV output**);

4.2.5 Resultado

A T2 produz um conjunto de previsões meteorológicas **validadas, corrigidas e normalizadas**, prontas para integração na etapa T3. Os registos inválidos são descartados, garantindo integridade e consistência nos dados utilizados para cálculo de alertas.

4.3 T3 – Cálculo de Alertas e Relatórios

4.3.1 Objetivo

A transformação **T3** tem como objetivo integrar os dados meteorológicos validados (provenientes da T2), calcular níveis de alerta com base em regras estabelecidas e gerar relatórios detalhados e agregados por cidade e região. Esta etapa é responsável por transformar dados limpos em informação útil.

4.3.2 Visão Geral

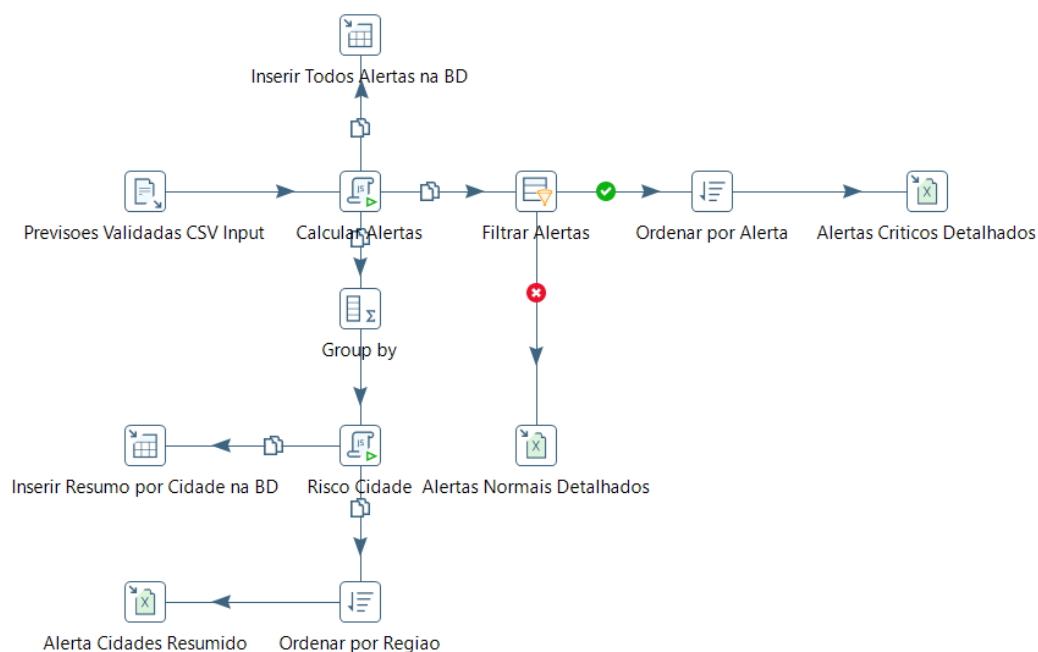


Figura 3: Fluxo geral da transformação T3 – Cálculo de Alertas e Relatórios

4.3.3 Componentes e Funções

Componente	Função
CSV Input	Lê previsões validadas da T2
Modified JavaScript (Calcular Alertas)	Define níveis de alerta e mensagens
Filter Rows	Separa alertas críticos dos normais
Sort Rows	Ordena por nível de alerta
Excel Output	Gera relatórios detalhados (críticos/normais)
Group By	Calcula médias e totais por cidade
Modified JavaScript (Risco Cidade)	Classifica risco final da cidade
Sort Rows (Região)	Ordena por região geográfica
Excel Output	Exporta resumo de alertas por cidade
Table Output	Insere resultados na BD

Tabela 5: Componentes da T3 e respectivas funções resumidas

4.3.4 Descrição das Etapas Principais

1. Leitura e Preparação dos Dados A transformação inicia-se com o **CSV Input**, que lê o ficheiro **Previsoes Validadas.csv** gerado na T2. Os campos incluem cidade, data, temperaturas, precipitação e direção do vento. Estes dados servem de base para o cálculo de risco e geração de alertas.

2. Cálculo de Alertas (Modified JavaScript) O componente **Modified JavaScript Value** é responsável por aplicar as regras e definir os níveis de alerta, tipo de evento e mensagem descritiva. A lógica avalia as condições meteorológicas e atribui classificações com base em thresholds predefinidos.

Exemplo de código:

```

1 var nivel_alerta = "";
2 var tipo_alerta = "";
3 var mensagem = "";
4 var acao_recomendada = "";
5
6 // Alerta 1: Temperatura
7 if (temp_media > 35) {
8     nivel_alerta = "CRITICO";
9     tipo_alerta = "Calor Extremo";
10    mensagem = "Temperatura muito elevada: " + temp_media + " C ";
11    acao_recomendada = "Risco de ondas de calor";
12 } else if (temp_media < 5) {
13     nivel_alerta = "CRITICO";
14     tipo_alerta = "Frio Extremo";
15     mensagem = "Temperatura muito baixa: " + temp_media + " C ";
16     acao_recomendada = "Ficar em ambientes quentes";
17 }

```

Listing 2: Cálculo dos níveis de alerta da temperatura no JavaScript

O resultado inclui campos como **nivel_alerta**, **tipo_alerta**, **mensagem**, **acao_recomendada** e **score_risco**, que serão utilizados nas etapas seguintes.

3. Filtragem e Relatórios Detalhados O `Filter Rows` separa os registos em dois fluxos:

- **Alertas Críticos** – Nível = “CRITICO” ou ”ALTO”(ALERTAS_URGENTES.xls);
- **Alertas Normais** – Restantes níveis (ALERTAS_NORMAIS.xls).

Ambos os conjuntos são ordenados por nível de alerta e cidade, garantindo uma leitura clara e hierarquizada.

4. Agregação e Classificação por Cidade (Group By e JavaScript) O `Group By` agrega os resultados por cidade, calculando estatísticas:

- Média, mínimo e máximo das temperaturas;
- Média e máximo do score de risco;
- Totais de alertas críticos, altos e médios;
- Número de dias analisados;

O passo seguinte (Modified JavaScript – Risco Cidade) define o risco final da cidade com base nos totais e score médio:

```
1 if (total_alertas_criticos > 0) risco_cidade = "CRITICO";  
2 else if (total_alertas_altos > 2) risco_cidade = "ALTO";  
3 else if (score_medio >= 40) risco_cidade = "MEDIO";  
4 else risco_cidade = "BAIXO";
```

Listing 3: Cálculo de risco por cidade

5. Exportação e Armazenamento São gerados múltiplos outputs:

- ALERTAS_URGENTES.xls – lista detalhada de eventos de alto risco;
- ALERTAS_NORMAIS.xls – previsões sem risco significativo;
- Alerta_Resumo_Cidades.xls – resumo por cidade e região;
- Inserção dos resultados na base de dados (via `Table Output`).

4.3.5 Resultados

A T3 gera uma visão dos níveis de risco meteorológico, com classificações automáticas por cidade e região. Os relatórios produzidos permitem identificar rapidamente zonas críticas e tendências.

Para ilustrar os resultados obtidos pela transformação T3, apresentam-se de seguida excertos dos ficheiros gerados automaticamente pelo processo de cálculo de alertas.

Ficheiro ALERTAS_URGENTES.xls A Figura 4 mostra um excerto do ficheiro ALERTAS_URGENTES.xls

Guardar Automaticamente

ALERTAS_URGENTES.xls - Modo de Compatibilidade

Procurar

Ficheiro

Base

Inserir

Desenhar

Esquema da Página

Fórmulas

Dados

Rever

Ver

Automatizar

Ajuda

Acrobat

Colar

Arial

10

A⁺

A⁻

N

I

S

Área de Transfer...

Tipo de Letra

Alinhamento

Número

Estilos

Células

Formatação Condicional

Inserir

Eliminar

Formatar

Formatar como Tabela

Estilos de Célula

A1

nome_cidade_normalizado

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	nome_cidade	data_previsao	temp_min	temp_max	temp_med	temp_med	temp_med	precipitacao	vento_dire	classificacao	descricao	descricao	fonte_dados	tipo_previsao	nivel_alerta
2	AVEIRO	19/10/2025 00:00	17,70	21,90	19,80	67,60	292,90	100,00	S	Agradável	Chuva/agu	Moderado	IPMA	Diária	ALTO
3	AVEIRO	21/10/2025 00:00	17,60	21,80	19,70	67,50	292,80	100,00	SW	Agradável	Chuva/agu	Moderado	IPMA	Diária	ALTO
4	AVEIRO	22/10/2025 00:00	18,50	22,30	20,40	68,70	293,50	100,00	SW	Agradável	Chuva frac	Moderado	IPMA	Diária	ALTO
5	AVEIRO	23/10/2025 00:00	18,60	22,90	20,80	69,30	293,90	94,00	SW	Agradável	Chuva frac	Moderado	IPMA	Diária	ALTO
6	BRAGA	19/10/2025 00:00	16,50	19,40	17,90	64,30	291,10	100,00	S	Fresco	Chuva/agu	Moderado	IPMA	Diária	ALTO
7	BRAGA	20/10/2025 00:00	15,00	20,60	17,80	64,00	290,90	100,00	SW	Fresco	Aguaceiro Fraco	IPMA	Diária	ALTO	
8	BRAGA	21/10/2025 00:00	15,20	19,80	17,50	63,50	290,60	100,00	SW	Fresco	Chuva/agu	Fraco	IPMA	Diária	ALTO
9	BRAGA	22/10/2025 00:00	17,10	19,80	18,50	65,20	291,60	100,00	SW	Agradável	Chuva frac	Moderado	IPMA	Diária	ALTO
10	BRAGA	23/10/2025 00:00	16,90	19,70	18,30	64,90	291,40	96,00	SW	Agradável	Chuva/agu	Fraco	IPMA	Diária	ALTO
11	COIMBRA	19/10/2025 00:00	16,80	23,40	20,10	68,20	293,20	100,00	SW	Agradável	Chuva/agu	Moderado	IPMA	Diária	ALTO
12	COIMBRA	21/10/2025 00:00	16,50	19,70	18,10	64,60	291,20	100,00	SW	Agradável	Chuva/agu	Moderado	IPMA	Diária	ALTO
13	COIMBRA	22/10/2025 00:00	17,90	22,50	20,20	68,40	293,30	100,00	SW	Agradável	Aguaceiro Fraco	IPMA	Diária	ALTO	
14	LEIRIA	19/10/2025 00:00	16,90	24,70	20,80	69,40	293,90	100,00	SW	Agradável	Chuva/agu	Fraco	IPMA	Diária	ALTO
15	LEIRIA	21/10/2025 00:00	17,10	22,30	19,70	67,50	292,80	100,00	SW	Agradável	Chuva/agu	Moderado	IPMA	Diária	ALTO
16	LEIRIA	22/10/2025 00:00	18,60	22,90	20,80	69,30	293,90	100,00	W	Agradável	Chuva frac	Fraco	IPMA	Diária	ALTO
17	LISBOA	19/10/2025 00:00	18,20	23,50	20,90	69,50	294,00	96,00	SW	Agradável	Aguaceiro Moderado	IPMA	Diária	ALTO	
18	LISBOA	21/10/2025 00:00	16,80	23,60	20,20	68,40	293,30	96,00	SW	Agradável	Chuva frac	Moderado	IPMA	Diária	ALTO
19	PORTO	19/10/2025 00:00	16,70	19,90	18,30	64,90	291,40	100,00	S	Agradável	Chuva/agu	Moderado	IPMA	Diária	ALTO
20	PORTO	20/10/2025 00:00	15,30	19,30	17,30	63,10	290,40	98,00	SW	Fresco	Aguaceiro Moderado	IPMA	Diária	ALTO	
21	PORTO	21/10/2025 00:00	16,40	19,50	17,90	64,30	291,10	100,00	SW	Fresco	Chuva/agu	Moderado	IPMA	Diária	ALTO
22	PORTO	22/10/2025 00:00	17,80	19,70	18,80	65,80	291,90	100,00	SW	Agradável	Chuva frac	Moderado	IPMA	Diária	ALTO
23	PORTO	23/10/2025 00:00	17,40	19,80	18,60	65,50	291,80	98,00	SW	Agradável	Chuva/agu	Moderado	IPMA	Diária	ALTO
24	SETUBAL	19/10/2025 00:00	15,20	23,50	19,40	66,80	292,50	94,00	SW	Agradável	Aguaceiro Fraco	IPMA	Diária	ALTO	
25	SETUBAL	21/10/2025 00:00	15,00	24,30	19,60	67,40	292,80	92,00	SW	Agradável	Chuva frac	Fraco	IPMA	Diária	ALTO
26															
27															
28															

Figura 4: Excerto do ficheiro ALERTAS_URGENTES.xls

Ficheiro Alerta_Resumo_Cidades.xls A Figura 5 apresenta o relatório organizado por cidade, Alerta_Resumo_Cidades.xls, onde são agregadas as médias de temperatura, os totais de alertas e o risco global de cada cidade.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	nome_cidade	regiao	dias_analise	temp_med	temp_min	temp_max	score_med	score_max	total_alerta	total_alerta	total_alerta	risco_cidade	
2	COIMBRA	Centro	5	19,2	15,1	23,4	24,0	30,0	0	3	2	ALTO	
3	LEIRIA	Centro	5	20,1	16,0	24,7	24,0	30,0	0	3	2	ALTO	
4	LISBOA	Centro	5	20,5	16,2	23,8	15,0	30,0	0	2	1	BAIXO	
5	SETUBAL	Centro	5	19,9	14,7	25,2	15,0	30,0	0	2	1	BAIXO	
6	AVEIRO	Norte	5	19,9	16,4	22,9	27,0	30,0	0	4	1	ALTO	
7	BRAGA	Norte	5	18,0	15,0	20,6	30,0	30,0	0	5	0	ALTO	
8	PORTO	Norte	5	18,2	15,3	19,9	30,0	30,0	0	5	0	ALTO	
9	BEJA	Sul	5	21,5	15,2	29,2	6,0	15,0	0	0	2	BAIXO	
10	EVORA	Sul	5	21,0	13,9	28,1	6,0	15,0	0	0	2	BAIXO	
11	FARO	Sul	5	21,4	17,0	26,5	0,0	0,0	0	0	0	BAIXO	
12													
13													
14													
15													
16													

Figura 5: Ficheiro Alerta_Resumo_Cidades.xls

5 Job Principal – Orquestração do Processo ETL

5.1 Objetivo

O **Job Principal** tem como finalidade coordenar a execução das três transformações principais (T1, T2 e T3), garantindo a correta sequência de execução, tratamento de erros e notificação automática por email ao utilizador. Desta forma, todo o processo de extração, validação e análise de previsões meteorológicas é automatizado e monitorizado.

5.2 Visão Geral

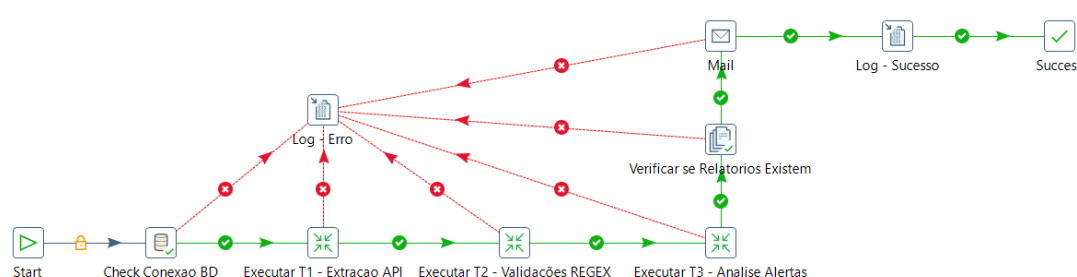


Figura 6: Fluxo geral do Job Principal – Orquestração das transformações T1, T2 e T3

5.3 Componentes e Funções

Componente	Função
Start	Inicia a execução do job
Check Conexão BD	Verifica se a base de dados está acessível
Executar T1 – Extração API	Chama a transformação T1 (extração XML da API)
Executar T2 – Validações REGEX	Executa a T2 (validação e limpeza de dados)
Executar T3 – Análise Alertas	Executa a T3 (análise e geração de relatórios)
Verificar se Relatórios Existem	Confirma a criação dos ficheiros de saída
Mail	Envia email automático com os relatórios em anexo
Log – Sucesso	Regista a execução bem-sucedida do processo
Log – Erro	Envia alerta de erro
Success	Finaliza o processo com estado “completo”

Tabela 6: Componentes do Job Principal e respetivas funções resumidas

5.4 Descrição das Etapas

1. Check Conexão BD Antes de iniciar as transformações, o job utiliza o componente **Check DB Connections** para confirmar a disponibilidade da ligação à base de dados PostgreSQL:

- Nome da conexão: `PostgreSQL_MeteoDB`

Se a ligação falhar, o processo é interrompido e um log de erro é registado.

2. Execução Sequencial das Transformações Após a validação da base de dados, o job executa as três transformações principais:

1. **Executar T1 – Extração API:** recolhe os dados da API meteorológica.
2. **Executar T2 – Validações REGEX:** limpa e valida os dados (datas, temperaturas, precipitação, vento).
3. **Executar T3 – Análise Alertas:** calcula níveis de alerta, risco e gera relatórios (.xls).

Cada etapa depende do sucesso da anterior, garantindo a integridade da cadeia de execução.

3. Verificação de Relatórios Após a T3, o job utiliza o componente **Check if Files Exist** para confirmar a geração dos ficheiros:

- `Alerta_Resumo_Cidades.xls`
- `Alertas_NORMAIS.xls`
- `Alertas_URGENTES.xls`

Caso algum dos ficheiros não exista, a execução é desviada para o fluxo de erro.

4. Envio Automático de Email O componente **Mail** envia automaticamente os relatórios ao utilizador. As configurações incluem:

- **Remetente:** Sistema de Alertas Meteorológicos
(`alertas@meteorologicos.pt`)
- **Destinatário:** `tiagonferr14@gmail.com`
- **Servidor SMTP:** `smtp.gmail.com`, Porta 587 (TLS)
- **Assunto:** [ALERTA METEOROLÓGICO] Ficheiros
- **Mensagem:** resumo automático indicando os ficheiros anexados:

Em anexo encontram-se os ficheiros relativos à previsão meteorológica de hoje e dos próximos 4 dias:

- `ALERTAS_NORMAIS`
- `ALERTAS_URGENTES`
- `ALERTAS_RESUMO_CIDADES`

5. Logs e Monitorização O job possui dois fluxos de monitorização:

- **Log – Sucesso:** gera um registo indicando o término correto de todas as transformações e envio de email;
- **Log – Erro:** gera um registo indicando erro de alguma etapa na execução do Job Principal.

5.4.1 Logs na Base de Dados

A execução do job é totalmente monitorizada através do registo de logs numa tabela específica da base de dados. As configurações encontram-se na aba Log, conforme ilustrado na Figura 7.

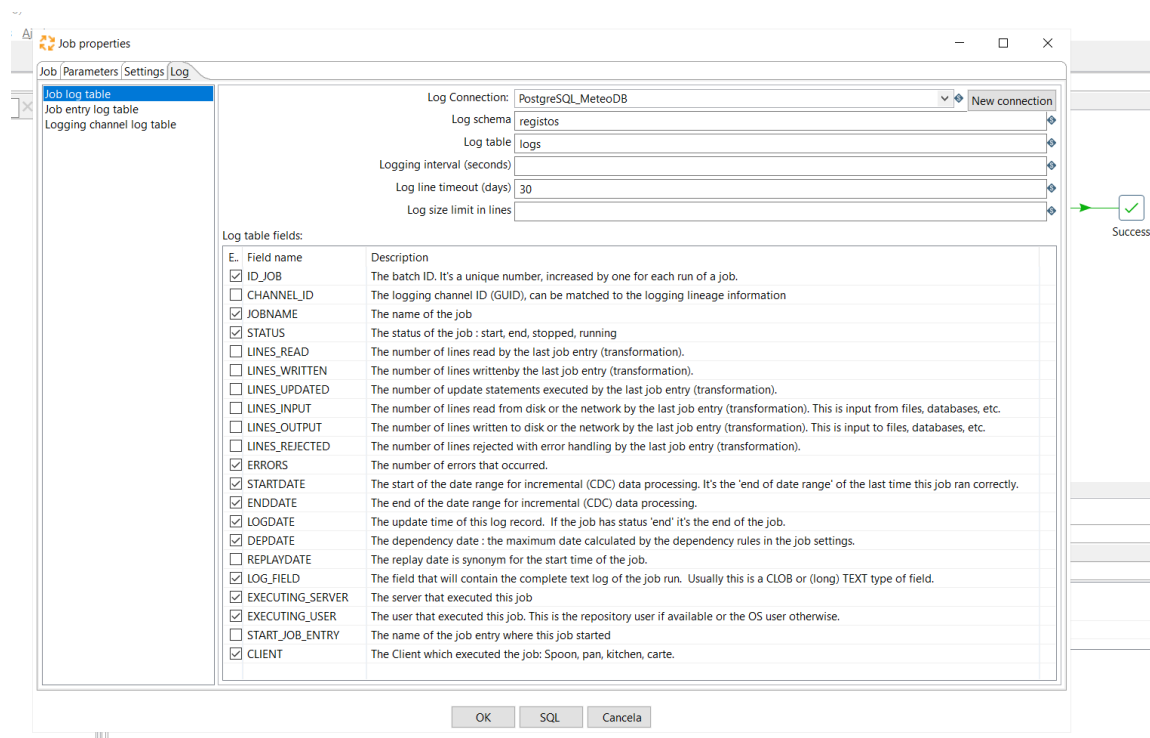


Figura 7: Configuração de logs

5.5 Resultado

O Job Principal garante a execução coordenada, segura e automatizada de todo o ETL do sistema de alertas meteorológicos. Através deste fluxo, assegura-se a:

- Disponibilidade dos dados na base de dados PostgreSQL;
- Geração e distribuição automática de relatórios;
- Monitorização de erros e notificações imediatas.

O resultado é um sistema robusto e totalmente automatizado, com alertas e relatórios diários enviados por email.

6 Base de Dados

6.1 Visão Geral

A base de dados utilizada no projeto é o **PostgreSQL**, configurado localmente no servidor `localhost:5432` com o nome `meteo_db`. A ligação foi definida no Pentaho sob o nome `PostgreSQL_MeteoDB`, utilizando PostgreSQL como tipo de conexão, conforme ilustrado na Figura 8.

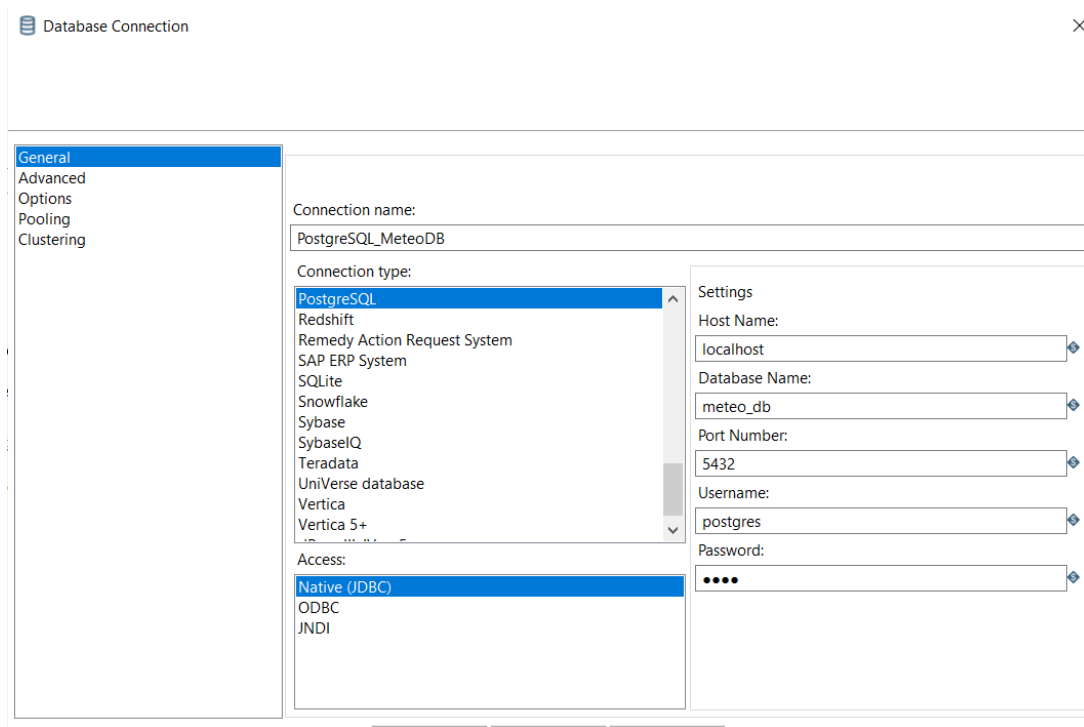


Figura 8: Ligação configurada no Pentaho ao PostgreSQL_MeteoDB

O modelo de dados foi projetado de forma simples e eficiente, dividido em dois **schemas**:

- **meteorologia** – contém os dados processados (previsões, alertas e resumos);
- **registos** – armazena os logs de execução.

6.2 Schema meteorologia

O schema `meteorologia` é o núcleo do sistema e contém três tabelas principais:

- **`previsoes_meteorologicas`** – resultados da T2 com previsões limpas e validadas;
- **`alertas_meteorologicos`** – dados finais da T3 com cálculos de risco e classificações;
- **`resumo_cidades`** – estatísticas agregadas de temperatura e risco por cidade e região.

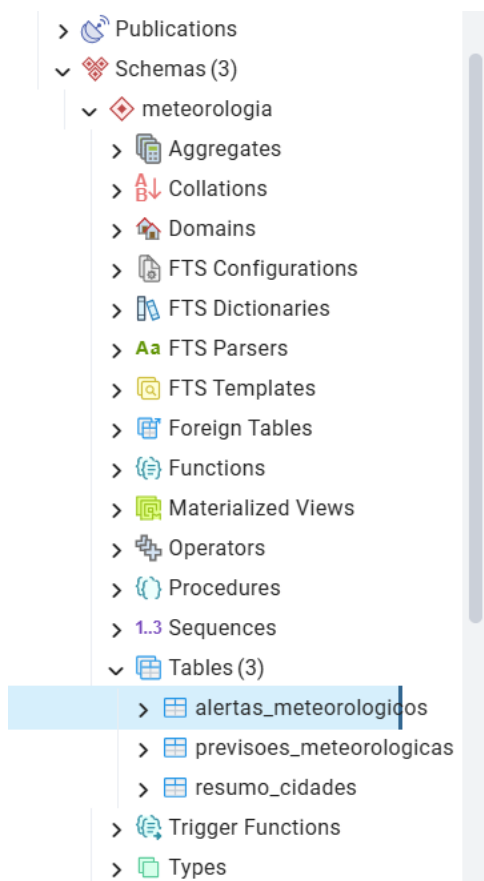


Figura 9: Estrutura geral do schema `meteorologia`

Cada tabela segue um formato coerente, com uma chave primária `id`, campos de identificação da cidade, valores meteorológicos, etc.

6.3 Schema registros

O schema `registros` contém a tabela `logs`, utilizada pelo Pentaho para monitorizar todas as execuções dos jobs. Esta tabela inclui informações como:

- **jobname** – nome do job executado;
- **status** – estado da execução (SUCCESS, ERROR, etc.);
- **errors** – número de erros registados;
- **startdate/enddate** – datas de início e fim;
- **executing_user/server** – utilizador e servidor que efetuaram a execução.

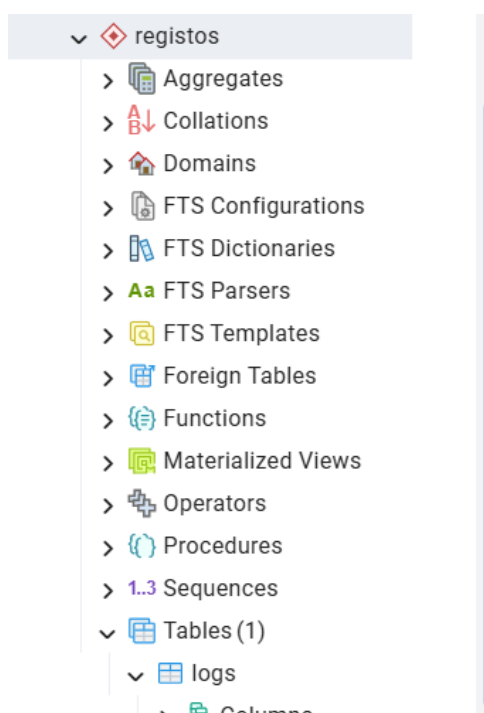


Figura 10: Tabela logs no schema registros

6.4 Integração com o Processo ETL

As transformações interagem com a base de dados da seguinte forma:

- **T2** insere registos em `meteorologia.previsoes_meteorologicas`;
- **T3** grava resultados em `meteorologia.alertas_meteorologicos` e `meteorologia.resumo_cidades`;
- O **Job Principal** escreve logs de execução em `registos.logs`.

Esta estrutura assegura a integridade dos dados.

6.5 Resumo

A arquitetura da base de dados garante:

- Persistência e histórico completo de previsões e alertas;
- Separação entre dados operacionais e de logs;
- Facilidade de consulta e exportação para relatórios;

7 Diagrama Entidade-Relacionamento (ER)

A Figura 11 apresenta o diagrama entidade-relacionamento (ER) do sistema de base de dados desenvolvido no âmbito deste projeto. As tabelas estão organizadas por esquema (*meteorologia* e *registos*).

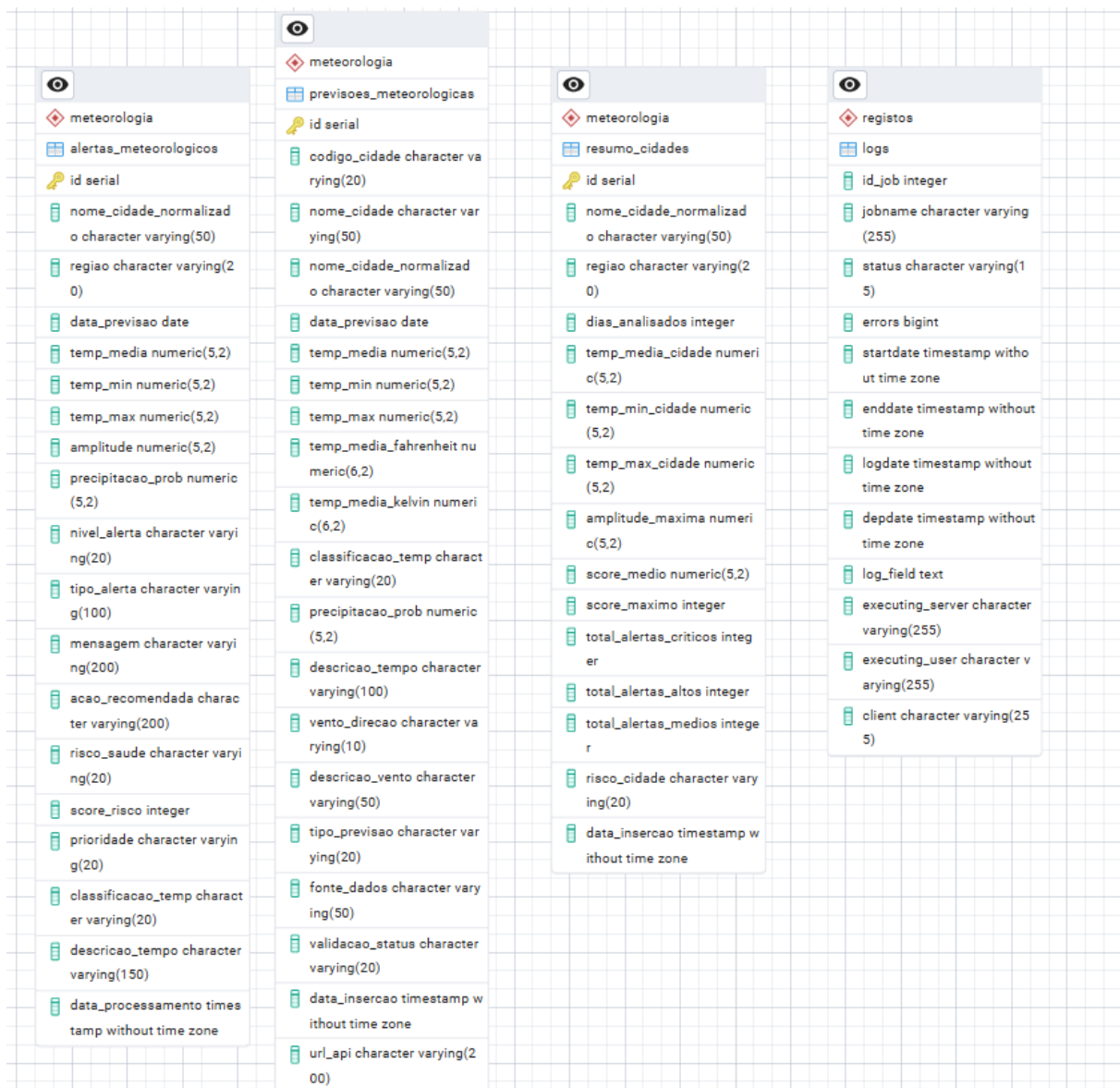
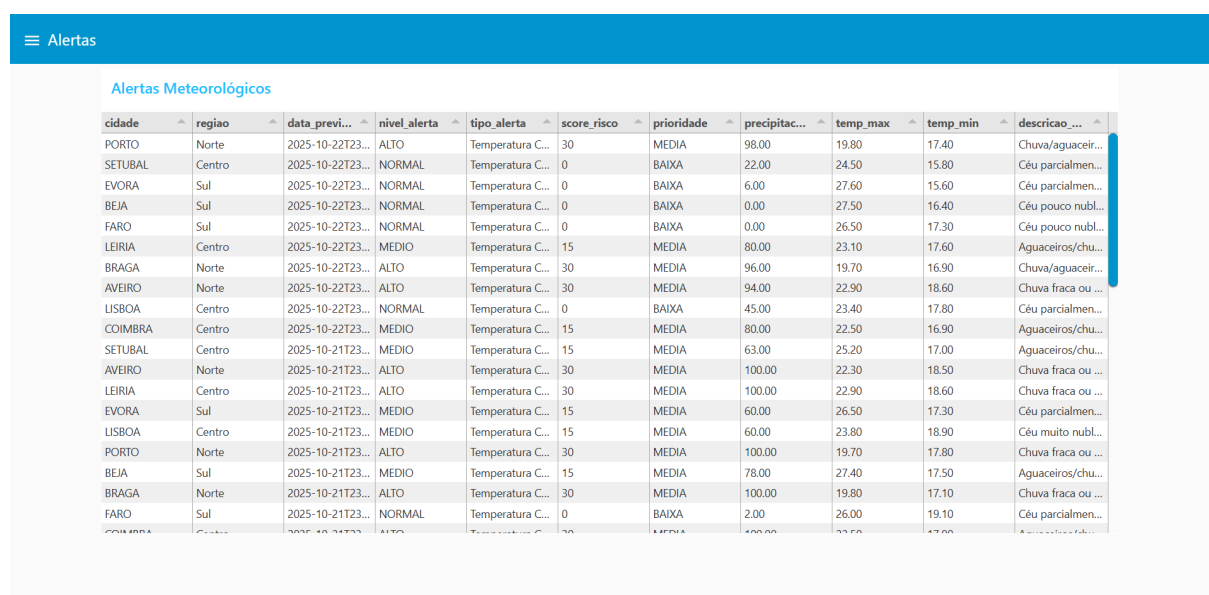


Figura 11: Diagrama Entidade-Relacionamento (ER) do sistema de base de dados.

8 Visualização e Monitorização de Dados no Node-RED

Por fim, foi desenvolvida uma interface de visualização no *Node-RED*, que permite consultar de forma interativa os dados meteorológicos processados e armazenados na base de dados.

O painel inclui tabelas com previsões, alertas e resumos por cidade (Figuras 12 , 13 e 14), bem como gráficos de distribuição de alertas por região (Figura 15) e temperaturas médias por cidade (Figura 16). Estes elementos permitem ao utilizador acompanhar em tempo real a evolução das condições meteorológicas e o nível de risco associado.



The screenshot shows a web interface titled "Alertas" with a sub-header "Alertas Meteorológicos". It displays a table with 11 columns: cidade, regio, data_previ..., nivel_alerta, tipo_alerta, score_risco, prioridade, precipitac..., temp_max, temp_min, and descricao.... The table lists alerts for 18 cities, including Porto, Setúbal, Évora, Beja, Faro, Leiria, Braga, Aveiro, Lisboa, Coimbra, Setúbal, Aveiro, Leiria, Évora, Lisboa, Porto, Beja, Braga, and Faro. Each row provides specific data for these categories, such as alert levels (ALTO, NORMAL, MEDIO), risk scores, and weather descriptions.

cidade	regio	data_previ...	nivel_alerta	tipo_alerta	score_risco	prioridade	precipitac...	temp_max	temp_min	descricao...
PORTO	Norte	2025-10-22T23...	ALTO	Temperatura C...	30	MEDIA	98.00	19.80	17.40	Chuva/aguaceir...
SETUBAL	Centro	2025-10-22T23...	NORMAL	Temperatura C...	0	BAIXA	22.00	24.50	15.80	Céu parcialmen...
EVORA	Sul	2025-10-22T23...	NORMAL	Temperatura C...	0	BAIXA	6.00	27.60	15.60	Céu parcialmen...
BEJA	Sul	2025-10-22T23...	NORMAL	Temperatura C...	0	BAIXA	0.00	27.50	16.40	Céu pouco nubl...
FARO	Sul	2025-10-22T23...	NORMAL	Temperatura C...	0	BAIXA	0.00	26.50	17.30	Céu pouco nubl...
LEIRIA	Centro	2025-10-22T23...	MEDIO	Temperatura C...	15	MEDIA	80.00	23.10	17.60	Aguaceiros/chu...
BRAGA	Norte	2025-10-22T23...	ALTO	Temperatura C...	30	MEDIA	96.00	19.70	16.90	Chuva/aguaceir...
AVEIRO	Norte	2025-10-22T23...	ALTO	Temperatura C...	30	MEDIA	94.00	22.90	18.60	Chuva fraca ou ...
LISBOA	Centro	2025-10-22T23...	NORMAL	Temperatura C...	0	BAIXA	45.00	23.40	17.80	Céu parcialmen...
COIMBRA	Centro	2025-10-22T23...	MEDIO	Temperatura C...	15	MEDIA	80.00	22.50	16.90	Aguaceiros/chu...
SETUBAL	Centro	2025-10-21T23...	MEDIO	Temperatura C...	15	MEDIA	63.00	25.20	17.00	Aguaceiros/chu...
AVEIRO	Norte	2025-10-21T23...	ALTO	Temperatura C...	30	MEDIA	100.00	22.30	18.50	Chuva fraca ou ...
LEIRIA	Centro	2025-10-21T23...	ALTO	Temperatura C...	30	MEDIA	100.00	22.90	18.60	Chuva fraca ou ...
EVORA	Sul	2025-10-21T23...	MEDIO	Temperatura C...	15	MEDIA	60.00	26.50	17.30	Céu parcialmen...
LISBOA	Centro	2025-10-21T23...	MEDIO	Temperatura C...	15	MEDIA	60.00	23.80	18.90	Céu muito nubl...
PORTO	Norte	2025-10-21T23...	ALTO	Temperatura C...	30	MEDIA	100.00	19.70	17.80	Chuva fraca ou ...
BEJA	Sul	2025-10-21T23...	MEDIO	Temperatura C...	15	MEDIA	78.00	27.40	17.50	Aguaceiros/chu...
BRAGA	Norte	2025-10-21T23...	ALTO	Temperatura C...	30	MEDIA	100.00	19.80	17.10	Chuva fraca ou ...
FARO	Sul	2025-10-21T23...	NORMAL	Temperatura C...	0	BAIXA	2.00	26.00	19.10	Céu parcialmen...

Figura 12: Interface de alertas meteorológicos no *Node-RED*.

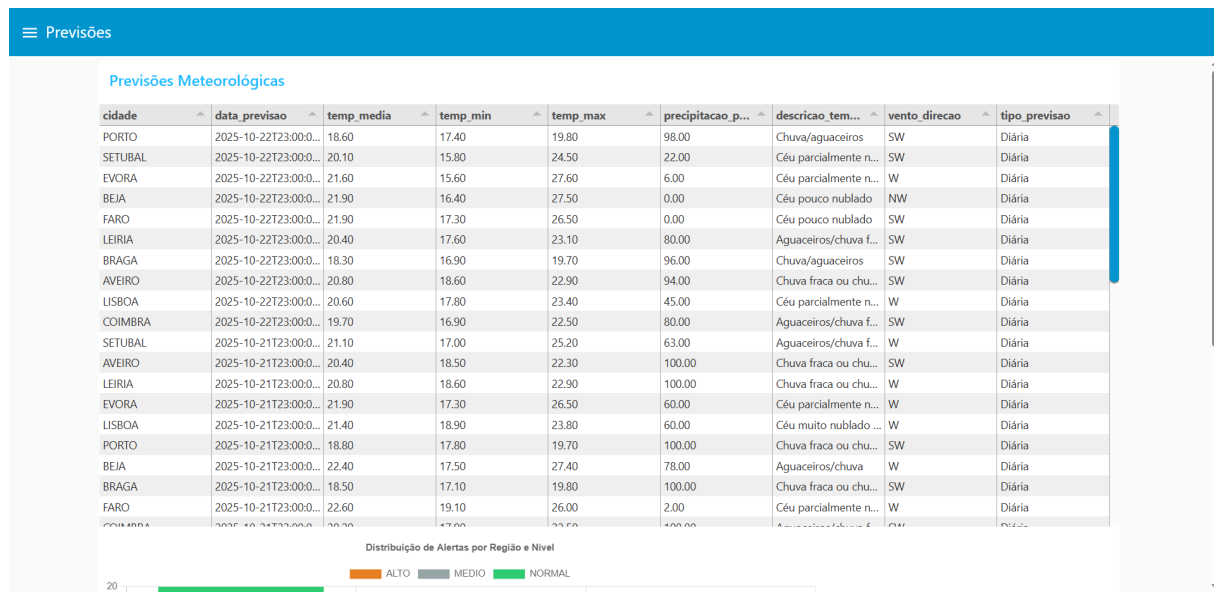


Figura 13: Visualização de previsões meteorológicas.

Resumo de Cidades

cidade	regiao	temp_media	temp_min	temp_max	score_medio	score_max...	total_alert...	total_alert...	total_alert...	risco_cidade
AVEIRO	Norte	19.92	16.40	22.90	27.00	30	0	4	1	ALTO
BEJA	Sul	21.54	15.20	29.20	6.00	15	0	0	2	BAIXO
BRAGA	Norte	18.00	15.00	20.60	30.00	30	0	5	0	ALTO
COIMBRA	Centro	19.22	15.10	23.40	24.00	30	0	3	2	ALTO
EVORA	Sul	21.00	13.90	28.10	6.00	15	0	0	2	BAIXO
FARO	Sul	21.36	17.00	26.50	0.00	0	0	0	0	BAIXO
LEIRIA	Centro	20.06	16.00	24.70	24.00	30	0	3	2	ALTO
LISBOA	Centro	20.54	16.20	23.80	15.00	30	0	2	1	BAIXO
PORTO	Norte	18.18	15.30	19.90	30.00	30	0	5	0	ALTO
SETUBAL	Centro	19.92	14.70	25.20	15.00	30	0	2	1	BAIXO

Figura 14: Visualização de previsões meteorológicas.

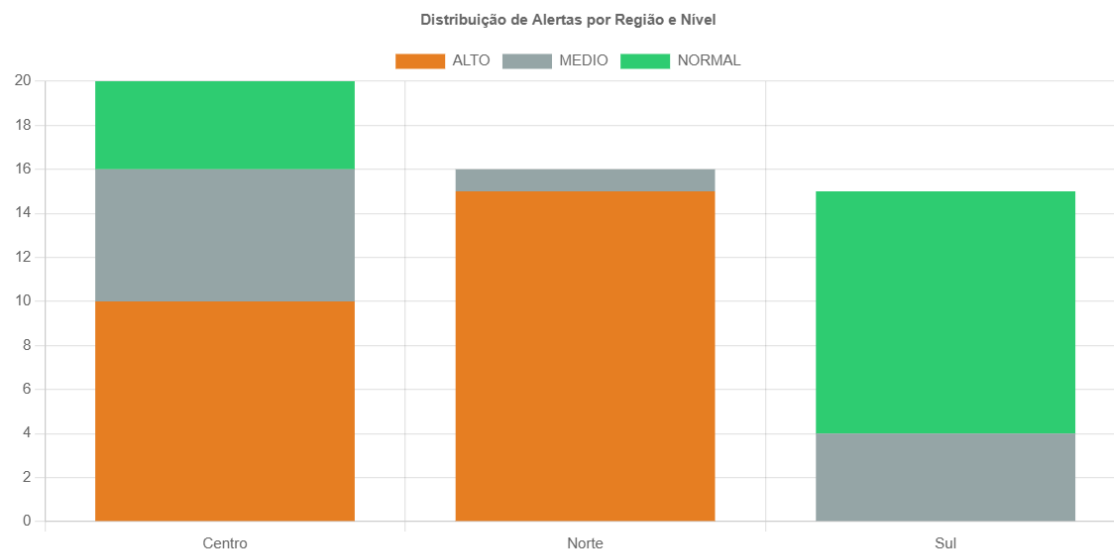


Figura 15: Distribuição de alertas por região e nível.

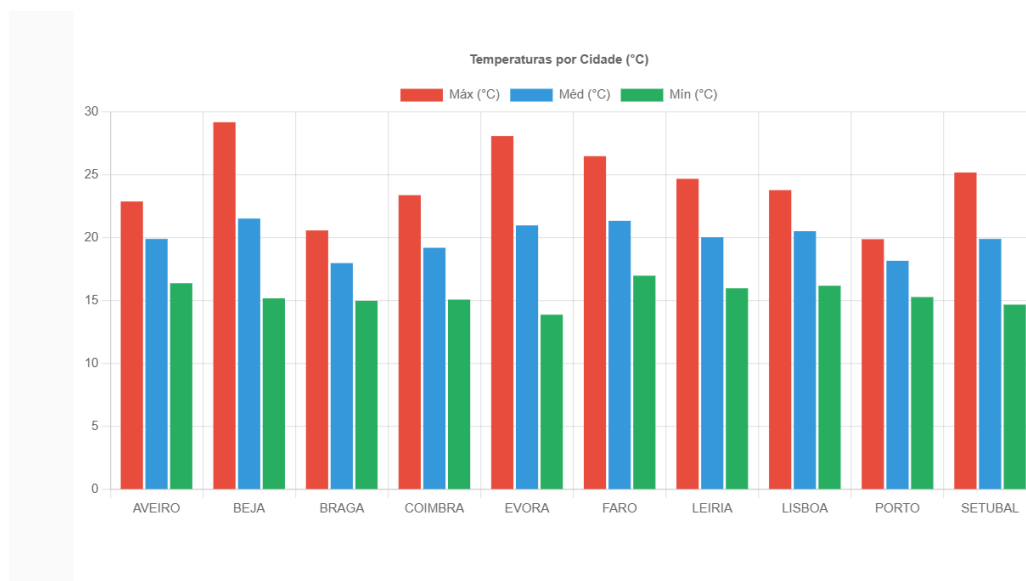


Figura 16: Gráfico de temperaturas por cidade.

A integração com o *Node-RED* reforça a utilidade prática do sistema, tornando os resultados acessíveis e facilmente interpretáveis para utilizadores.

9 Conclusão

O projeto desenvolvido permitiu compreender de forma prática todo o ciclo de vida de um sistema ETL, desde a extração de dados externos, passando pela validação e transformação, até à geração automática de relatórios e integração numa base de dados relacional. O uso do Pentaho Data Integration revelou-se uma ferramenta eficaz para a orquestração de processos e automação de fluxos de dados complexos. O sistema final implementado cumpre os objetivos definidos, garantindo:

- Integridade e rastreabilidade total dos dados meteorológicos;
- Automatização completa das operações de extração e análise;
- Emissão diária de alertas e relatórios via email.

9.1 Trabalhos Futuros

1. **Alertas via SMS:** Sistema de notificações push via SMS
2. **API RESTful:** Disponibilizar dados processados via API própria
3. **Análise Geoespacial:** Integração com mapas interativos
4. **Histórico Alargado:** Análise de tendências climáticas anuais
5. **App:** Aplicação móvel para consulta de alertas

10 Demonstração do Projeto

Por fim, foi elaborado um vídeo de demonstração. O mesmo encontra-se disponível através do *QR Code* apresentado abaixo, permitindo visualizar o processo completo de extração, transformação e visualização dos dados desenvolvidos neste projeto.



Figura 17: QR Code para Demonstração.

11 Referências Bibliográficas

1. IPMA - Instituto Português do Mar e Atmosfera. *API de Dados Abertos*. Disponível em: <https://api.ipma.pt>
2. Pentaho Documentation. *Pentaho Data Integration*. Disponível em: <https://help.hitachivantara.com/Documentation/Pentaho>
3. PostgreSQL Global Development Group. *PostgreSQL 16 Documentation*. Disponível em: <https://www.postgresql.org/docs/16/>
4. Node-RED. *Node-RED Documentation*. Disponível em: <https://nodered.org/docs/>
5. Regular Expressions Info. *Regular Expression Tutorial*. Disponível em: <https://www.regular-expressions.info/>
6. node-red-dashboard. *node-red-dashboard 3.6.6*. Disponível em: <https://flows.nodered.org/node/node-red-dashboard>
7. What is ETL? *What is ETL (extract, transform, load)?*. Disponível em: <https://www.ibm.com/think/topics/etl>