 <small>ESCOLA SUPERIOR DE TECNOLOGIA E GESTÃO</small>	Tipo de Prova Trabalho prático Época Normal – Enunciado 2	Ano letivo 2025/2026	Data
	Curso <b>Mestrado em Engenharia Informática</b>	Hora	
	Unidade Curricular Tópicos Avançados de Armazenamento de Dados	Duração	

## Enunciado do Trabalho Prático

### 1 Destinatários e Objetivos

Este trabalho destina-se a todos os estudantes inscritos na unidade curricular de Tópicos Avançados de Armazenamento de Dados que pretendam obter aprovação à Unidade Curricular. Com a realização do trabalho prático, pretende-se que os alunos ponham em prática todos os conhecimentos adquiridos nos diversos modelos de armazenamento de dados existentes, não só na implementação de sistemas operacionais, mas também de sistemas de suporte à decisão.

### 2 Caso de estudo


Os incêndios florestais constituem um problema crítico em Portugal e em diversos países, com impactos sociais, económicos e ambientais significativos. A análise de dados relacionados com incêndios como a sua ocorrência, extensão, causas, meios de combate, danos humanos e materiais é fundamental para apoiar decisões estratégicas de prevenção, resposta e recuperação. Neste contexto, pretende-se que desenvolva um **sistema analítico** que suporte a análise de dados de incêndios. Os **datasets a utilizar deverão ser escolhidos pelos estudantes** e comunicados ao docente para validação. **Não serão aceites trabalhos diferentes com o mesmo dataset.**

O sistema deverá ser capaz de:

- Recolher e integrar dados provenientes de diferentes fontes (ex.: relatórios oficiais, bases de dados governamentais, dados meteorológicos, sensores, dados geográficos, entre outros);
- Estruturar a informação num **Data Warehouse** com pelo menos **duas tabelas de factos** (sujeito a validação do docente);
- Facultar uma análise multidimensional, permitindo comparações por localização, tempo, intensidade, tipologia do incêndio, causas, recursos utilizados, entre outros;
- Explorar, se necessário, **outras tecnologias de armazenamento** (Data Lake, bases de dados de grafos, vetoriais, etc.) para complementar o sistema de suporte à decisão.

#### 2.1 Componentes do trabalho

O sistema a desenvolver deverá estar caracterizado em cinco componentes:


 <small>ESCOLA SUPERIOR DE TECNOLOGIA E GESTÃO</small>	Tipo de Prova Trabalho prático Época Normal – Enunciado 2	Ano letivo 2025/2026	Data
	Curso <b>Mestrado em Engenharia Informática</b>	Hora	
	Unidade Curricular Tópicos Avançados de Armazenamento de Dados	Duração	

- **Caracterização das Fontes de Dados:** identificação e descrição dos datasets escolhidos, bem como da estratégia de extração e integração. Sempre que possível, a extração deverá ser automatizada (via APIs, scraping, downloads programáticos, etc.).
- **Componente de Integração de Dados:** preparação dos dados (auditoria, limpeza, transformação, enriquecimento, etc.), alinhada com os requisitos identificados para os data marts.
- **Implementação do(s) Data Mart(s):** criação de uma arquitetura de Data Warehouse que suporte a análise multidimensional. Devem ser apresentadas pelo menos duas tabelas de factos devidamente justificadas, bem como as respetivas dimensões.
- **Outras tecnologias de armazenamento de dados:** possibilidade de complementar o sistema com Data Lakes, bases de dados vetoriais para LLM ou bases de dados de grafos, explorando diferentes paradigmas de suporte à decisão.
- **Componente de Visualização de Dados:** desenvolvimento de dashboards e relatórios que apresentem os resultados da análise.

Deverá ainda ser elaborado um relatório em formato de artigo, abordando todos os passos do desenvolvimento do sistema de acordo com a metodologias e métodos apresentados ao longo das aulas. O artigo deverá ser elaborado utilizando o *template* LNCS<sup>1</sup> e deverá, no mínimo, conter os seguintes conteúdos (podem ser reorganizados em termos de estrutura):

- **Introdução:** enquadramento do tema dos incêndios e a relevância da sua análise com sistemas de Business Intelligence e tecnologias complementares.
- **O Sistema de Business Intelligence:**
  - Caso de aplicação: descrição do caso de estudo e fundamentação da relevância do sistema.
  - Arquitetura do sistema: requisitos, análise das fontes, estrutura do Data Warehouse (modelação concetual, lógica e física), bem como tecnologias utilizadas.
  - Processo de povoamento: caracterização da área de retenção e descrição do processo de ETL. A modelação concetual dos processos de povoamento deverá ser apresentada em BPMN (processo inicial e incremental).
  - Análise e Visualização de Dados: exemplos de dashboards e relatórios construídos.

<sup>1</sup> <https://www.springer.com/gp/computer-science/lncs/conference-proceedings-guidelines>

 <small>ESCOLA SUPERIOR DE TECNOLOGIA E GESTÃO</small>	Tipo de Prova Trabalho prático Época Normal – Enunciado 2	Ano letivo 2025/2026	Data
	Curso <b>Mestrado em Engenharia Informática</b>	Hora	
	Unidade Curricular Tópicos Avançados de Armazenamento de Dados	Duração	

- Conclusões e Trabalho Futuro: análise crítica do trabalho.

Como referência, o artigo deverá ter até 25 páginas.

### 3 Datas e considerações

Os alunos devem formar grupos de trabalho de no máximo **3** elementos. O trabalho deve ser entregue até ao dia 18 de Dezembro, devendo a entrega ser feita através da página da unidade curricular em <http://moodle.estg.ipp.pt>.

### 4 Entrega e apresentação

Todos os elementos do grupo deverão submeter o trabalho no *link* respetivo, devendo entregar em formato zip e separado por pastas:

- As fontes utilizadas no trabalho prático;
- Processo de povoamento e estruturas de suporte;
- Backup e script de desenvolvimento do sistema;
- Ficheiros necessários para reproduzir as *dashboards*/relatórios produzidos;
- Relatório abordando todos os tópicos apresentados no modelo de relatório disponível na plataforma moodle.
- Um documento com os passos para a reprodução de uma demonstração do trabalho tendo em consideração a inexistência de dados no *Data Warehouse*, o seu povoamento inicial, exemplo de povoamento incremental e visualização de dados.

Será ainda realizada uma apresentação de no máximo 15 minutos, apresentando o caso de estudo, pressupostos associados, arquitetura da solução desenvolvida e ainda uma breve demonstração do trabalho desenvolvido. A data para a apresentação será posteriormente comunicada.