 **INSTITUTO POLITÉCNICO DE BEJA **

**Escola Superior de Tecnologia e Gestão**

**Licenciatura em Engenharia Informática**

**Trabalho prático nº 1**

José Alves <inserir numero>

Miguel Marmelete, nº 20431

03/11/2023

**INSTITUTO POLITÉCNICO DE BEJA**

**Escola Superior de Tecnologia e Gestão**

**Licenciatura em Engenharia Informática**

**Nome do Projeto/Estágio**

Elaborado por:

José Alves

Miguel Marmelete

Orientado por:

Isabel Brito

Relatório do primeiro trabalho prático da unidade curricular Sistemas de Informação

Escola Superior de Tecnologia e Gestão do Instituto Politécnico de Beja

03/11/2023

**Resumo**

Resumo

Neste trabalho prático, foram combinados dados provenientes de várias tabelas, retirados do dataset “Global Nutrition Adult and Adolescent”, presente no website “Kaggle”. Os dados retirados foram posteriormente submetidos a um processo de limpeza e preparação utilizando o excel. Utilizando o SQL Server, procedemos à criação de uma Data Warehouse, composta por tabelas de dimensão e uma tabela de factos, permitindo assim uma organização coerente e completa dos dados. Posteriormente, os dados foram carregados e analisados através do Power BI, resultando na criação de diversos gráficos e visualizações que proporcionaram uma análise aprofundada dos mesmos. Este relatório detalhará o processo ETL adotado, o modelo de dados implementado e os resultados obtidos durante a análise.

**Palavras-chave:** trabalho prático, dados, tabelas, dataset, Global Nutrition Adult and Adolescent, Kaggle, limpeza, preparação, Excel, SQL Server, Data Warehouse, tabelas de dimensão, tabela de factos, organização, coerente, completa, Power BI, gráficos, visualizações, análise, ETL, modelo de dados, resultados.

**Abstract**

Abstract

In this practical assignment, data from various tables were combined, extracted from the "Global Nutrition Adult and Adolescent" dataset available on the Kaggle website. The extracted data underwent a cleaning and preparation process using Excel. Using SQL Server, we proceeded to create a Data Warehouse, comprised of dimension tables and a fact table, allowing for a coherent and comprehensive organization of the data. Subsequently, the data was loaded and analyzed using Power BI, resulting in the creation of various charts and visualizations that provided an in-depth analysis. This report will detail the adopted ETL process, the implemented data model, and the obtained results during the analysis.

**Keywords:** practical assignment, data, tables, dataset, Global Nutrition Adult and Adolescent, Kaggle, cleaning, preparation, Excel, SQL Server, Data Warehouse, dimension tables, fact table, organization, coherent, comprehensive, Power BI, charts, visualizations, analysis, ETL (Extract, Transform, Load), data model, results.

Índice

[1. Introdução 2](#_Toc60738230)

[2. Caracterização da Entidade de Acolhimento do Estágio 2](#_Toc60738231)

[3. Descrição das Atividades Desenvolvidas 2](#_Toc60738232)

[3.1. Atividade 1 3](#_Toc60738233)

[3.2. Atividade 2 3](#_Toc60738234)

[4. Conclusões e Perspetivas de Trabalho Futuro 3](#_Toc60738235)

**Lista de Figuras**

Figura 2.2 - Organograma da Empresa XX 2

Figura 3.1 – xxx 2

**Lista de Tabelas**

Tabela 3.1 – Tabela de … 2

# Introdução

Descrever aqui sucintamente em que consistiu o seu projeto de trabalho. Local de realização, atividades previstas e respetivo planeamento e faseamento.

O presente projeto representou um esforço colaborativo realizado no conforto do ambiente doméstico, com comunicação facilitada através da plataforma Discord. As atividades desdobraram-se em diversas etapas, com o intuito de explorar e analisar o dataset "Global Nutrition Adult and Adolescent" disponibilizado no renomado repositório Kaggle. Inicialmente, a tarefa primordial consistiu na seleção criteriosa do dataset, seguida pela etapa crucial de limpeza e preparação dos dados, processo esse que foi conduzido de maneira eficaz utilizando a versatilidade do Microsoft Excel.

Posteriormente, voltamos nossa atenção para a construção e compreensão da estrutura da Data Warehouse, concebendo e implementando tabelas no SQL Server para possibilitar o armazenamento e organização coerente dos dados. A visualização e validação da integridade dos dados na Data Warehouse foi um passo crucial para garantir a confiabilidade e consistência das informações.

Em uma etapa subsequente, alavancamos o potencial do Power BI para carregar as tabelas e, a partir disso, gerar uma série diversificada de gráficos e visualizações. Essas representações visuais forneceram uma perspectiva analítica abrangente, permitindo-nos extrair insights valiosos a partir do conjunto de dados em questão.

Ao longo deste processo, cada fase foi meticulosamente planeada e executada com o objetivo de assegurar a qualidade e precisão das análises realizadas. No decorrer deste relatório, serão detalhados de forma mais específica os procedimentos adotados, o modelo de dados implementado e os resultados obtidos durante a análise.

O último parágrafo deste capítulo deve referir a estrutura do relatório, indicando o que se aborda em cada um dos capítulos seguinte.

NOTAS GERAIS

1. Todas as figuras, tabelas, gráficos devem ter numeração sequencial, com o primeiro dígito igual ao número de capítulo, p.e. “Figura 2.1 - Exemplo de Figura”. Neste template isso já está feito
2. A bibliografia deve seguir a norma igual à que tem final deste documento. Todas as referências têm de ser citadas no texto, por exemplo: “… este trabalho segue a metodologia indicada em [1], …”
3. O início de qualquer secção tem de ter sempre algum texto descritivo do que se pretende com essa secção.
4. Todas as imagens, tabelas, etc, têm de estar sempre referidas no texto

# Caracterização da Entidade de Acolhimento do Estágio

Descrição, com historial, áreas de atividade, etc. da empresa onde realizou o estágio.

Deve incluir aqui também a seção/departamento da empresa onde foi inserido para realização do estágio e tarefas/objetivos dessa seção/departamento.

<ex. organograma>

Figura 2.1 - Organograma da Empresa XX

# Descrição das Atividades Desenvolvidas

Descrição detalhada das atividades desenvolvidas no estágio. Aqui devem constar subseções para cada tipo de atividade desenvolvida, ferramentas utilizadas, modelação de bases de dados, estrutura/requisitos de aplicação desenvolvida, etc.

<Exemplo de tabela>

Tabela . – Tabela de …

## Atividade 1

<ex. de figura>

Figura 3.1 – xxx

## Atividade 2

# Conclusões e Perspetivas de Trabalho Futuro

Conclusões do trabalho desenvolvido, com referência comparativa também com os objetivos iniciais (que constam do primeiro capítulo).

A seção de “Perspetivas de Trabalho Futuro” deve indicar o que pode ser desenvolvido futuramente por si, quer no âmbito das atividades realizadas, quer no que este estágio contribuiu para o seu desenvolvimento profissional e o que pode ser o seu percurso futuro.

**Bibliografia**

[1] F. P. J. Brooks, The Mythical Man-Month Essays on Software Engineering Anniversary Edition. Addison-Wesley, 1995.

[2] P. P.-S. Chen, \The entity-relationship model | toward a uni\_ed view of data," ACM Trans. Database Syst., vol. 1, n. 1, pp. 9{36, Mar\_co 1976. [Online]. Disponível: http://doi.acm.org/10.1145/320434.320440

[3] (2013) Android developers: Developer tools. [Online]. Disponível: http://developer.android.com/tools/index.html

[4] J. Hu, H. Wu, A. Jennings, e X. Wang, “Fast and robust equalization: A case study," in Proceedings of the World Multiconference on Systemics, Cybernetics and Informatics, (SCI 2000), Florida, USA, 23-26 July 2000. FL, USA: International Institute of Informatics and Systemics, 2000, pp. 398.

<Outras referências>

**Anexo 1**

Imagens, Código, etc

**Anexo 2**

Imagens, Código, etc