



# BASES DE DADOS 2

## Relatório final do projeto

### Membros do Grupo

Gonçalo Pereira - a96550

Joana Pereira - a95539

João Mendes - a97991

Mariana Alves - a97293

Ricardo Pereira - a95925

Grupo 03 - Turno PL6

# INTRODUÇÃO

No âmbito da Unidade Curricular de Bases de Dados 2 foi-nos proposta a realização de um projeto prático, dividido em três partes, em torno de um DataSet escolhido pelo grupo no início do semestre. O nosso DataSet baseia-se nos Jogos Paralímpicos do ano de 2021.

## BASES DE DADOS SQL – MYSQL

Na primeira entrega, trabalhamos com a linguagem SQL (Structured Query Language), uma tecnologia relacional que já nos era conhecida. Apresentamos, então, um Modelo Conceptual e Relacional dos dados, interpretados e analisados, do nosso DataSet, e, ainda, um Dicionário de Dados. No MySQL (SGDB) foram carregados os ficheiros csv, depois de realizarmos todas as alterações necessárias. Por último, criamos queries em SQL e as respetivas soluções.

Nesta fase, as dificuldades foram poucas, uma vez que já estávamos familiarizados com o MySQL.

## BASES DE DADOS NOSQL – CASSANDRA (ORIENTADA A COLUNAS)

No segundo momento, a tecnologia usada foi uma base de dados NoSQL, tecnologia esta que, apesar das semelhanças com a usada anteriormente, é orientada a colunas, e cada família de colunas deve responder a uma consulta específica. Desta forma, pode dizer-se que a flexibilidade em termos de consultas é reduzida. No entanto, possui uma grande escalabilidade horizontal.

Na realização desta fase já tivemos mais algumas dificuldades, uma vez que não conhecíamos esta tecnologia. Uma das questões que inicialmente nos causava alguma dificuldade era o facto de, contrariamente ao MySQL, não existirem chaves estrangeiras, isto é, as tabelas existem autonomamente, o que dificultou a resolução de queries em que era necessária a junção de 2 tabelas.

Apesar destas dificuldades conseguimos adaptar-nos e assim, concluir que esta tecnologia é adequada para grandes bases de dados em que o desempenho das consultas tem que ser grande.

## BASES DE DADOS NOSQL – NEO4J (ORIENTADA A GRAFOS)

Na última entrega voltamos a trabalhar com uma tecnologia NoSQL, mas, desta vez, uma tecnologia orientada a grafos. Esta é ideal para representar dados com grande número de relacionamentos entre si, porém apresenta dificuldades na escalabilidade horizontal.

No que toca a adversidades ocorridas nesta fase, apesar de ser também uma tecnologia desconhecida, não tivemos muitas. Consideramos que esta é uma tecnologia de uso relativamente simples, sendo a única maior dificuldade a análise do modelo de forma a criar os relacionamentos entre as entidades.

## BASES DE DADOS SQL VS BASES DE DADOS NOSQL

SQL	NoSQL
Bases de dados relacionais, com esquema fixo, estático ou pré-definido	Bases de dados não relacionais com esquema dinâmico
Escalabilidade vertical	Escalabilidade horizontal
Baseadas em tabelas	Chave-valor, orientadas a documentos, colunas e grafos
Requisitos ACID (atomicity, consistency, isolation, durability)	Teorema CAP (Consistency, Availability e Partitioning tolerance)
Não adequados para armazenamento hierárquico de dados.	Adequados para armazenamento hierárquico de dados.
Consultas complexas.	Consultas simples