

**Universidade do Minho**

Mestrado Integrado em Engenharia Informática

**Unidade Curricular de**

**Bases de Dados**

Ano Lectivo de 2018/2019

**Arquivo de Revistas da Ritinha**

**Bárbara Andreia Cardoso Ferreira A80453**

**José Augusto Ferreira Alves A82885**

**Miriam Miranda Pinto A42040**

**Rui Filipe Brito Azevedo A80789**

Janeiro, 2019

**BD**

|  |  |
| --- | --- |
| Data de Recepção |  |
| Responsável |  |
| Avaliação |  |
| Observações |  |

**Arquivo de Revistas da Ritinha**

**Bárbara Andreia Cardoso Ferreira A80453**

**José Augusto Ferreira Alves A82885**

**Miriam Miranda Pinto A42040**

**Rui Filipe Brito Azevedo A80789**

Janeiro, 2019

# Resumo

O presente relatório tem por objetivo descrever as várias fases que foram abordadas para o desenvolvimento de uma base de dados relacional e posteriormente não relacional. O objetivo nas duas fases do trabalho foi o armazenamento de informação digital, ou seja, o armazenamento digital de informação relativa a revistas, facilitando a procura e atualização do registo das mesmas.

Na primeira fase é feita a descrição do contexto do problema, fundamentada a necessidade da base de dados e analisada a viabilidade do processo. De seguida foi feito o levantamento e análise dos requisitos junto do cliente, a amável Dona Ritinha, que posteriormente foram divididos em três categorias: descrição, exploração e controlo. Após a sua análise foram redigidos numa linguagem mais técnica que por sua vez levou à identificação e caracterização das entidades, atributos e relacionamentos necessários para a passagem à fase seguinte, ou seja, a construção do modelo conceptual.

Com o modelo conceptual estabelecido foi gerado o modelo lógico onde se averiguou a normalização do mesmo e a sua capacidade de resposta às interrogações do cliente. Nesta fase do processo é preciso ter em atenção ao conjunto de operações que só podem ser executadas como um todo, no caso de uma falhar não se executa nenhuma. Este conjunto de operações dão origem a transações.

Após a validação do modelo lógico é gerado o modelo físico recorrendo à ferramenta *Forward Engineer* presente no *MySQL* *Workbench*. Nesta fase as interrogações do cliente e as transações passam a estar descritas em código *SQL*, são escolhidos os índices, é feita uma estimativa do espaço em disco da base de dados e da sua taxa de crescimento anual e são definidas e caracterizadas as vistas de utilização do *SQL*.

Na segunda fase do trabalho foi feita a migração da base de dados relacional para um sistema não relacional e foram descritas em código *MongoDB* algumas interrogações do cliente.

Após a concretização de todas estas fases pensamos estar perante uma base de dados relacional e uma não recional operacionais, consistentes, uteis e apropriadas para o projeto da nossa cliente.

**Área de Aplicação:** Criação de Sistemas de Bases de Dados

**Palavras-Chave:** Modelo Conceptual, Modelo Lógico, Base de Dados Relacional, *MySQL*, NoSQL, Entidade, Relacionamento, Atributo, Chave Primária, Chave Estrangeira, Coleção, Documento

Índice

Resumo i

Índice de Figuras vi

Índice de Tabelas vii

1. Definição do sistema 1

1.1. Contexto de aplicação do sistema 1

1.2. Fundamentação da implementação da base de dados 1

1.3. Análise da viabilidade do processo 2

1.4. Estrutura do Relatório 2

2. Levantamento e Análise de Requisitos 3

2.1. Método de levantamento e de análise de requisitos adotado 3

2.2. Requisitos levantados 3

2.2.1. Requisitos de Descrição 4

2.2.2. Requisitos de Exploração 4

2.2.3. Requisitos de Controlo 5

2.3. Análise geral dos requisitos 5

3. Modelação conceptual 6

3.1. Apresentação da abordagem de modelação realizada 6

3.2. Identificação e caracterização das entidades 6

3.3. Identificação e caracterização dos relacionamentos 7

3.4. Identificação e caracterização das associações dos atributos com as entidades e relacionamentos 10

3.5. Detalhe ou generalização de entidades 12

3.6. Apresentação e explicação do diagrama ER 12

3.7. Validação do modelo de dados com o utilizador 13

4. Modelação Lógica 14

4.1. Construção e validação do modelo de dados lógico 14

4.2. Desenho do modelo lógico 15

4.3. Validação do Modelo através da Normalização 16

4.3.1. Primeira forma normal (1NF) 16

4.3.2. Segunda forma normal (2NF) 16

4.3.3. Terceira forma normal (3NF) 16

4.4. Validação do modelo com interrogações do utilizador 17

4.4.1. Permitir a consulta das informações de uma empresa 17

4.4.2. Listar os contactos de uma dada empresa 17

4.4.3. Permitir a consulta das informações de um anúncio 17

4.4.4. Permitir a consulta das informações de um artigo 17

4.4.5. Listar todos as revistas de uma dada empresa 18

4.4.6. Listar todos os artigos contidos numa revista 18

4.4.7. Listar todos os anúncios contidos numa revista 18

4.4.8. Listar todos os artigos escritos por um dado escritor 18

4.4.9. Listar todas as revistas editadas por um dado editor 18

4.4.10. Identificar a quantidade de revistas visualizadas de uma empresa 19

4.4.11. Listar todos os escritores que têm pelo menos um artigo escrito numa determinada revista 19

4.4.12. Listar todas as revistas de uma dada empresa numa determinada data ou intervalo de datas 19

4.4.13. Listar todos os escritores que já contribuíram para a escrita de revistas de uma dada empresa 20

4.4.14. Listar todas as revistas em que aparecerem os anúncios de uma determinada empresa 20

4.4.15. Listar as N revistas com maior número de consultas 20

4.4.16. Listar as categorias por frequências relativas de consultas 21

4.5. Validação do modelo com as transações estabelecidas 21

4.6. Revisão do modelo lógico com o utilizador 22

5. Implementação Física 23

5.1. Seleção do sistema de gestão de bases de dados 23

5.2. Tradução do esquema lógico para o sistema de gestão de bases de dados escolhido em SQL 23

5.3. Tradução das interrogações do utilizador para SQL (alguns exemplos) 23

5.3.1. Listar todos os escritores que já contribuíram para a escrita de revistas de uma dada empresa 24

5.3.2. Calcular o número total de consultas das revistas de uma empresa 24

5.3.3. Listar as categorias por frequências relativas de consultas 25

5.4. Tradução das transações estabelecidas para SQL (alguns exemplos) 25

5.4.1. Inserir uma empresa 26

5.4.2. Inserir um artigo 27

5.5. Escolha, definição e caracterização de índices em SQL (alguns exemplos) 28

5.6. Estimativa do espaço em disco da base de dados e taxa de crescimento anual 28

5.6.1. Estimativa do espaço em disco da base de dados 28

5.6.2. Taxa de crescimento anual 31

5.7. Definição e caracterização das vistas de utilização do SQL (alguns exemplos) 31

5.7.1. Vista da Tabela Escritor 31

5.7.2. Vista das empresas e respetivas visualizações das suas revistas 32

5.7.3. Vista do *Top 5* de escritores 32

5.8. Definição e caracterização dos mecanismos de segurança em SQL (alguns exemplos) 33

5.9. Revisão do sistema implementado com o utilizador 34

6. Migração da Base de Dados 35

6.1. Fundamentação da Migração da Base de Dados – (Identificação e descrição dos objetivos da base de dados em termos de aplicações e de utilizadores) 35

6.2. Identificação e explicação do tipo de questões/necessidades que serão realizadas sobre o sistema de dados NoSQL 36

6.3. Definição da estrutura base para o sistema de dados NoSQL 36

6.4. Identificação dos objetos de dados no sistema SQL que serão utilizados para alimentar o novo sistema 41

6.5. Mapeamento do processo de migração de dados, descrevendo o processo de conversão dos vários objetos de dados 41

6.6. Explicação do processo de migração de dados, explicando de modo detalhado as suas principais etapas – extração, transformação e carregamento 41

6.7. Apresentação e descrição da implementação do processo de migração de dados 41

6.8. Apresentação da forma como as questões identificadas anteriormente podem ser satisfeitas com o novo sistema, utilizando a linguagem de interrogação do sistema NoSQL 41

7. Conclusões (REVER ---🡪 NoSQL) 46

Referências 47

Lista de Siglas e Acrónimos 53

# Índice de Figuras

Figura 1 - Relacionamento Revista – Anúncio. 7

Figura 2 - Relacionamento Revista – Artigo. 8

Figura 3 - Relacionamento Revista – Empresa. 8

Figura 4 - Relacionamento Revista – Editor. 8

Figura 5 - Relacionamento Empresa – Anúncio. 9

Figura 6 - Relacionamento Artigo – Escritor. 9

Figura 7 - Modelo Conceptual 12

Figura 8 - Modelo Lógico 15

Figura 9 - Tradução do RE5 para *SQL* 24

Figura 10 - Tradução do RE10 para *SQL* 24

Figura 11 - Tradução do RE15 para *SQL* 25

Figura 12 - Tradução da transação de inserção de uma empresa em *SQL* 26

Figura 13 - Tradução de transação da inserção de um artigo em *SQL* 27

Figura 14 - Vista sobre a tabela Escritor 31

Figura 15 - Vista das empresas e respetivas visualizações totais das suas revistas 32

Figura 16 – Vista do *Top 5* de escritores 32

Figura 17 - Código *SQL* para a criação da vista de utilização e respetivas permissões do utilizador admin 33

Figura 18 - Código *SQL* para a criação da vista de utilização e respetivas permissões do utilizador app 33

Figura 19 – Informação da Revista *“Fofoquices com a Vizinha”* 36

Figura 20 - Informação da Revista *“Fofoquices com a Vizinha”* (continuação) 37

Figura 21 - Informação da Revista *“Fofoquices com a Vizinha”* (continuação) 38

Figura 22 - Informação da Revista *“Fofoquices com a Vizinha”* (continuação) 39

Figura 23 - Informação da Revista *“Fofoquices com a Vizinha”* (continuação) 40

Figura 24 – Query que apresenta os artigos contidos na Revista com ID 10 42

Figura 25 – Resultados obtidos à Query anterior (Figura 24) 42

Figura 26 - Resultados obtidos à Query anterior (Figura 24) (continuação) 43

Figura 27 - Resultados obtidos à Query anterior (Figura 24) (continuação) 44

Figura 28 - Resultados obtidos à Query anterior (Figura 24) (continuação) 44

Figura 29 – Query que lista as n revistas com maior número de consultas 45

Figura 30 - Resultados obtidos à Query anterior (Figura 29) 45

Figura 31 - Povoamento da tabela Empresa 48

Figura 32 - Povoamento da tabela Contactos 48

Figura 33 - Povoamento da tabela Anúncio 49

Figura 34 - Povoamento da tabela Editor 49

Figura 35 - Povoamento da tabela Revista 50

Figura 36 - Povoamento da tabela Artigo 50

Figura 37 - Povoamento da tabela Escritor 51

Figura 38 - Povoamento da tabela AnuncioRevista 51

Figura 39 - Povoamento da tabela EscritorArtigo 52

# Índice de Tabelas

[Tabela 1 - Dicionário de dados das entidades 7](#_Toc535329698)

[Tabela 2 - Dicionário de dados dos relacionamentos. 9](#_Toc535329699)

[Tabela 3 - Identificação dos atributos 12](#_Toc535329700)

[Tabela 4 - Espaço por entrada da tabela Revista 28](#_Toc535329701)

[Tabela 5 - Espaço por entrada da tabela Empresa 29](#_Toc535329702)

[Tabela 6 - Espaço por entrada da tabela Contacto 29](#_Toc535329703)

[Tabela 7 - Espaço por entrada da tabela Artigo 29](#_Toc535329704)

[Tabela 8 - Espaço por entrada da tabela Escritor 29](#_Toc535329705)

[Tabela 9 - Espaço por entrada da tabela Anuncio 30](#_Toc535329706)

[Tabela 10 - Espaço por entrada da tabela Editor 30](#_Toc535329707)

[Tabela 11 - Espaço por entrada da tabela AnuncioRevista 30](#_Toc535329708)

[Tabela 12 - Espaço por entrada da tabela ArtigoEscritor 30](#_Toc535329709)

1. Definição do sistema

Nesta secção será feita uma contextualização e fundamentação da implementação da base de dados, tendo em mente a viabilidade do processo. No final é feito um pequeno resumo da estrutura geral do relatório.

* 1. Contexto de aplicação do sistema

O crescimento da informação acompanha a constante mudança de um mundo turbulento e inquieto e o formato em que esta é apresentada pode surgir em diversos tipos, tais como digital ou em papel.

Com o avanço das tecnologias surgiu um enorme crescimento no ‘consumo’ de informação em formato digital, que apresenta um vasto leque de artigos variados, redigidos por diversos escritores. Toda essa informação encontra-se dispersa por diferentes revistas.

Neste contexto é apropriado organizar os dados relevantes das revistas, tais como dados acerca dos artigos e anúncios que estas contêm, dos editores responsáveis pela edição das várias revistas, dos escritores responsáveis pelos artigos e das empresas, que podem ser produtoras de anúncios e/ou produtoras de revistas, de modo a fornecer as informações relativas às preferências dos leitores.

* 1. Fundamentação da implementação da base de dados

A realização deste estudo é feita por solicitação da Dona Ritinha, uma senhora simpática e singela que requisitou os serviços de quatro heroicos futuros engenheiros da Universidade do Minho para que organizassem os dados relativamente às revistas vendidas da sua livraria digital.

A Dona Ritinha tinha por objetivo obter informações acerca das preferências dos leitores e com base nas mesmas ter um arquivo que facilitasse o acesso a informações gerais e específicas relativamente às revistas. Perante esta necessidade decidiu organizar-se os dados de forma a que seja possível consultar detalhadamente revistas armazenados digitalmente, os seus editores, os seus artigos, anúncios, empresas e escritores.

As revistas são caracterizadas por nome, categoria, edição e n.º de consultas. O editor de uma ou várias revistas possui um nome. Os artigos possuem data, título, corpo, n.º de consultas e categoria. Um artigo pode ser redigido por um ou mais escritores os quais são caracterizados por nome. Os anúncios têm categoria, título, conteúdo, n.º de consultas e contato. Uma empresa está caraterizada por nome, morada, contato e quantidade de revistas. Cada empresa pode ter uma ou mais revistas associadas e pode publicar anúncios na revista.

* 1. Análise da viabilidade do processo

Este sistema foi pensado com o intuito de satisfazer dois tipos de utilizadores: as pessoas que estão constantemente em busca de informação digital e a Dona Ritinha que pretende melhorar o seu serviço.

Como sabemos, atualmente, cada vez mais o ser humano sente a necessidade de estar informado acerca do que se passa à sua volta. As pessoas procuram respostas, fontes de inspiração ou uma mera distração na Web e nem sempre é fácil obter o que realmente se pretende.

Perante este cenário pensamos ser adequado que a informação das revistas da Dona Ritinha esteja organizada de um modo simples e completo permitindo uma procura de informação cómoda e pouco dispendiosa. Apesar de uma revista impressa ser uma grande ‘relíquia’, nos dias correntes, poucas são as pessoas que se deslocam a um espaço físico para a adquirir e folhear em procura de um dado assunto. O tempo tornou-se algo escasso e tudo o que pode ser feito num clique de rato faz com que se descarte qualquer outra hipótese mais dispendiosa em tempo e dinheiro.

Este sistema de base de dados visa auxiliar qualquer tipo de aplicação que seja utilizada pelos dois utilizadores referidos no início desta secção e satisfazer as suas necessidades neste contexto.

* 1. Estrutura do Relatório

Após a definição do sistema feita ao longo desta secção, será apresentada a metodologia, o levantamento e a análise dos diversos tipos de requisitos ao longo do capítulo 2. No capítulo 3 será exposto o modelo conceptual elaborado com base nos dados recolhidos e expostos até então. O modelo lógico surgirá no capítulo 4, onde será explicada a sua elaboração abordada a sua validade. No penúltimo capítulo surge o modelo físico com recurso ao *MySQL* *Workbench* e as interrogações do cliente assim como as transações passam a estar descritas em código *SQL*. No último capítulo será feita uma reflexão relativamente ao trabalho elaborado, ao que se pode melhorar e o que se gostaria de fazer no futuro.

1. Levantamento e Análise de Requisitos

Nesta secção será abordado o método de levantamento e de análise dos vários requisitos levantados e no final será feita uma análise geral dos mesmos.

* 1. Método de levantamento e de análise de requisitos adotado

Quando um profissional se encontra perante um projeto/desafio, o primeiro passo a ser executado é compreender o que o cliente que solicitou o mesmo pretende. Com esta ideia em mente o nosso primeiro intuito foi abordar a Dona Ritinha através de uma entrevista para o levantamento de requisitos. Com essa entrevista conseguimos determinar que a Dona Ritinha pretende um site onde os seus clientes possam aceder a informações relativas a diversas revistas, tais como artigos e respetivos escritores, editor, anúncios publicados e respetivas empresas anunciantes ou a própria empresa da revista em causa.

A Dona Ritinha quer armazenar, relativamente a cada revista, o nome, a edição, a categoria e o número de consultas. Para cada artigo, que pode ser redigido por um ou mais escritores, ela pretende que seja armazenada a data, a categoria, o número de consultas, o corpo e o título. Relativamente a cada escritor dos artigos e editores de revistas apenas pretende guardar o seu nome como informação.

No decorrer da entrevista a Dona Ritinha frisou que quer manter um registo dos anúncios que são publicados nas várias revistas e quer manter registado a sua categoria, o contato das empresas anunciantes, o conteúdo, o título e o número de consultas. A estes anúncios quer também que estejam associadas as empresas que os publicam e que seja possível aceder ao seu nome, à sua morada e contatos. A Dona Ritinha manifestou também que quer manter o registo das empresas responsáveis pelas revistas tendo disponível no site as informações sobre o seu nome, a morada, os seus contatos e a quantidade de revistas produzidas por elas.

* 1. Requisitos levantados

De seguida serão apresentados os requisitos que foram levantados através da entrevista à nossa cliente e que deverão ser suportados pelo sistema de base de dados a desenvolver para cada um dos diferentes tipos (exploração e administração).

* + 1. Requisitos de Descrição

RD1: Uma empresa é caracterizada pelo nome, a morada e os contactos.

RD2: Uma revista é caracterizada pelo nome, a edição, a categoria e número de consultas (visualizações).

RD3: Os anúncios são caracterizados pelo seu título, o conteúdo, a categoria, o contacto do anunciante e o número de consultas.

RD4: Um artigo é caracterizado pelo título, a data em que foi redigido, o corpo(conteúdo), a categoria e o número de consultas.

RD5: Um escritor é caracterizado pelo seu nome.

RD6: Um editor é caracterizado pelo seu nome.

* + 1. Requisitos de Exploração

RE1: Permitir a consulta das informações de uma empresa.

RE2: Listar os contatos de uma dada empresa

RE3: Permitir a consulta das informações de um anúncio.

RE4: Permitir a consulta das informações de um artigo.

RE5: Listar todas as revistas de uma dada empresa.

RE6: Listar todos os artigos contidos numa revista.

RE7: Listar todos os anúncios contidos numa revista.

RE8: Listar todos os artigos escritos por um dado escritor.

RE9: Listar todas as revistas editadas por um dado editor.

RE10: Identificar a quantidade de revistas visualizadas de uma empresa.

RE11: Listar todos os escritores que têm pelo menos um artigo escrito numa determinada revista.

RE12: Listar todas as revistas de uma dada empresa numa determinada data ou intervalo de datas.

RE13: Listar todos os escritores que já contribuíram para a escrita de revistas de uma dada empresa.

RE14: Listar todas as revistas em que aparecerem os anúncios de uma determinada empresa.

RE15: Listar as N revistas com maior número de consultas.

RE16: Listar as categorias por frequência relativas de consultas.

* + 1. Requisitos de Controlo

RC1: O gerente da aplicação deverá ter permissão para adicionar revistas e todas as informações das mesmas.

RC2: Atualizar visualizações associadas a uma revista, anúncio ou artigo sempre que um destes seja consultado.

RC3: Impedir a criação de uma revista com uma data de edição inferior à última edição criada.

RC4: O utilizador da aplicação apenas poderá consultar informações.

* 1. Análise geral dos requisitos

Após o levantamento dos requisitos, os quatro candidatos a bons e dedicados engenheiros decidiram reunir-se para elaborar uma estratégia de ataque/ação face ao problema que a Dona Ritinha havia colocado.

Foram necessárias algumas reuniões até que o conjunto final de requisitos ficasse finalmente estabelecido e com a concordância de todos os membros.

O sistema de base de dados em questão deve ser utilizado com o auxílio a uma aplicação *web* ou *mobile* desenvolvida neste contexto.

A administradora, Dona Ritinha, deve conseguir inserir novas revistas e os respetivos dados associados à mesma. Por outro lado, os utilizadores devem apenas conseguir consultar todos os artigos e anúncios contidos no sistema.

1. Modelação conceptual

Nesta secção começaremos por apresentar a abordagem de modelação realizada e de seguida identificar e caracterizar as entidades, os relacionamentos e as associações dos atributos com as entidades e relacionamentos. Será também apresentado o modelo conceptual e feita a sua validação com o utilizador.

* 1. Apresentação da abordagem de modelação realizada

A abordagem de modelação realizada envolve duas vistas de análise de desenvolvimento: o utilizador da aplicação que só poderá consultar os dados e a Dona Ritinha que poderá aceder, inserir e manipular as informações armazenadas na base de dados.

* 1. Identificação e caracterização das entidades

Com base na análise dos requisitos que foram apresentados anteriormente, identificamos a necessidade da existência das seguintes seis entidades:

* **Empresa:** representa todas as empresas responsáveis pela produção das revistas e/ou patrocínio de anúncios;
* **Revista:** representa todas as revistas existentes no arquivo;
* **Anúncio:** representa todos os anúncios publicados;
* **Artigo:** representa todos os artigos publicados;
* **Editor:** representa todos os editores de revistas;
* **Escritor:** representa todos os escritores dos artigos publicados.

Apresentamos agora o dicionário de dados que representa as entidades com algumas informações das mesmas:

|  |  |  |
| --- | --- | --- |
| **Nome da Entidade** | **Descrição** | **Ocorrências** |
| Empresa | Termo geral que descreve todas as empresas existentes, produtoras de revistas e/ou patrocinadores de anúncios. | Cada empresa é responsável pela produção das revistas e/ou pelo patrocínio de anúncios. |
| Revista | Termo geral que descreve as revistas existentes no arquivo. | Cada revista é produzida por uma empresa e contém artigos e anúncios. |
| Anúncio | Termo geral que descreve todos os anúncios publicados. | Cada revista tem anúncios publicados. |
| Artigo | Termo geral que descreve todos os artigos publicados. | Cada revista tem artigos publicados. |
| Editor | Termo geral que descreve todos os editores de revistas. | Cada editor é o responsável por fazer a edição das revistas. |
| Escritor | Termo geral que descreve todos os escritores dos artigos publicados. | Cada escritor é o responsável pela escrita dos artigos publicados. |

Tabela 1 - Dicionário de dados das entidades

* 1. Identificação e caracterização dos relacionamentos

Tendo em consideração todos os requisitos apresentados anteriormente e as entidades existentes, iremos agora a apresentar os seis relacionamentos que identificamos entre as mesmas.

Visto que a entidade Revista é o elemento central do nosso arquivo, esta tem vários relacionamentos. Nomeadamente, uma revista contém artigos e anúncios, pertence a uma empresa e é editada por um editor. Quanto aos restantes, um artigo é redigido por um escritor e uma empresa financia anúncios.

* **Relacionamento Revista - Anúncio**



Figura 1 - Relacionamento Revista – Anúncio.

Este relacionamento identifica os anúncios publicados nas revistas.

A cardinalidade apresentada foi baseada no facto de uma revista ter vários anúncios e de um anúncio poder estar publicado em várias revistas.

* **Relacionamento Revista – Artigo**



Figura 2 - Relacionamento Revista – Artigo.

Este relacionamento identifica os artigos publicados em cada revista.

A multiplicidade deste relacionamento é 1:N, visto que uma revista tem vários artigos publicados e um artigo é publicado em uma revista.

* **Relacionamento Revista – Empresa**

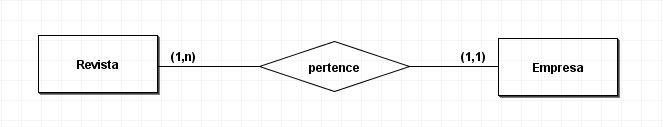


Figura 3 - Relacionamento Revista – Empresa.

O relacionamento agora apresentado identifica as revistas que pertencem a cada empresa.

Este relacionamento apresenta cardinalidade N:1, ou seja, significa que uma revista pertence a uma e uma só empresa, e uma empresa pode ser produtora de muitas revistas.

* **Relacionamento Revista – Editor**



Figura 4 - Relacionamento Revista – Editor.

Este relacionamento indica-nos as revistas que foram editadas por cada um dos editores.

Este relacionamento apresenta, mais uma vez, uma cardinalidade N:1, tendo em conta que uma revista é editada por um e um só editor, e um editor edita várias revistas.

* **Relacionamento Empresa – Anúncio**

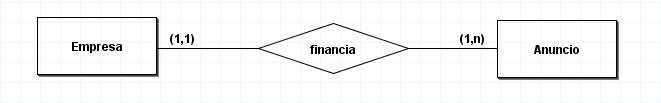


Figura 5 - Relacionamento Empresa – Anúncio.

O relacionamento entre a empresa e o anúncio identifica quais são os anúncios que são financiados por uma empresa.

A multiplicidade apresentada deve-se ao facto de uma empresa poder financiar muitos anúncios e um anúncio ser financiado por uma só empresa.

* **Relacionamento Artigo - Escritor**



Figura 6 - Relacionamento Artigo – Escritor.

Este relacionamento associa os artigos aos seus respetivos escritores.

A cardinalidade deste relacionamento apresentada é de N:N, visto que um artigo pode ser redigido por vários escritores e um escritor é autor de vários artigos.

Apresentamos agora o dicionário de dados que representa os relacionamentos com as informações dos mesmos:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Entidade** | **Multiplicidade** | **Relação** | **Multiplicidade** | **Entidade** |
| Revista | N | tem publicado | N | Anúncio |
| 1 | tem publicado | N | Artigo |
| N | pertence | 1 | Empresa |
| N | é editada | 1 | Editor |
| Empresa | 1 | Financia | N | Anúncio |
| Artigo | N | é redigido | N | Escritor |

Tabela 2 - Dicionário de dados dos relacionamentos.

* 1. Identificação e caracterização das associações dos atributos com as entidades e relacionamentos

Tendo em consideração os requisitos, as entidades e os relacionamentos estabelecidos anteriormente, determinamos agora os atributos referentes a cada uma das entidades.

Na caraterização da entidade Empresa foi utilizado o atributo composto Morada, que é composto por Cidade e Rua. Foi também utilizado o atributo multi-valor Contacto, pois a empresa tem vários contactos associados. Para além deste, ainda para a entidade Empresa, decidimos também criar o atributo derivado QuantidadeRevistas, que indica o número de revistas que uma empresa tem.

Cada uma das entidades tem um atributo definido como a chave primária. Como esta tem que ser única, decidimos criar um Id para cada uma delas como sendo o identificador único da entidade.

Na entidade Revista poderíamos ter usado como chave primária a Edição, que é a data da mesma, mas não seria única, pois as revistas são de várias empresas e por este motivo, seriam várias as revistas com a mesma data. Logo, poderiam existir chaves primárias iguais. Definindo o Id como chave primária, garantimos que não haverá este problema.

A identificação dos atributos encontra-se na seguinte tabela:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Entidade** | **Atributo** | **Descrição** | **Domínio** | **Nulo** | **Composto** | **Multivalor** | **Derivado** |
| Revista | Id | Identificador da revista. | INT | Não | Não | Não | Não |
| Nome | Nome da revista. | VARCHAR(45) | Não | Não | Não | Não |
| Nr consultas | Número de consultas feitas à revista. | INT | Não | Não | Não | Não |
| Categoria | Categoria da revista. | VARCHAR(45) | Não | Não | Não | Não |
| Edicao | Data da revista. | DATE | Não | Não | Não | Não |
| Anúncio | Id | Identificador do anúncio. | INT | Não | Não | Não | Não |
| Categoria | Categoria do anúncio. | VARCHAR(45) | Não | Não | Não | Não |
| Conteúdo | Conteúdo do anúncio. | VARCHAR(2500) | Não | Não | Não | Não |
| Nr consultas | Número de consultas feitas ao anúncio. | INT | Não | Não | Não | Não |
| Contacto | Contacto da empresa que publicou o anúncio. | VARCHAR(45) | Sim | Não | Não | Não |
| Titulo | Título do anúncio. | VARCHAR(45) | Não | Não | Não | Não |
| Artigo | Id | Identificador do artigo. | INT | Não | Não | Não | Não |
| Título | Título do artigo. | VARCHAR(45) | Não | Não | Não | Não |
| Categoria | Categoria do artigo. | VARCHAR(45) | Não | Não | Não | Não |
| Corpo | Conteúdo do artigo. | VARCHAR(2500) | Não | Não | Não | Não |
| Data | Data de quando o artigo foi escrito. | DATE | Não | Não | Não | Não |
| Nr consultas | Número de consultas feitas ao artigo. | INT | Não | Não | Não | Não |
| Empresa | Id | Identificador da empresa. | INT | Não | Não | Não | Não |
| Nome | Nome da empresa. | VARCHAR(45) | Não | Não | Não | Não |
| Contacto | Contactos da empresa. | VARCHAR(45) | Não | Não | Sim | Não |
| Morada | Morada da empresa. | VARCHAR(45) | Não | Sim | Não | Não |
| Composto por:  Cidade  Rua |  | | | | | |
| Cidade da empresa. | VARCHAR(45) | Não | Não | Não | Não |
| Rua da empresa. | VARCHAR(45) | Não | Não | Não | Não |
| QuantidadeRevistas | Quantidade de revistas que pertencem a uma empresa. | INT | Não | Não | Não | Sim |
| Editor | Id | Identificador do editor. | INT | Não | Não | Não | Não |
| Nome | Nome do editor. | VARCHAR(45) | Não | Não | Não | Não |
| Escritor | Id | Identificador do escritor. | INT | Não | Não | Não | Não |
| Nome | Nome do escritor. | VARCHAR(45) | Não | Não | Não | Não |

Tabela 3 - Identificação dos atributos

* 1. Detalhe ou generalização de entidades

Com base em toda a informação recolhida dos requisitos e identificadas as nossas entidades, decidimos que não seria necessário detalhar nem fazer uma generalização das mesmas.

* 1. Apresentação e explicação do diagrama ER

Tendo em conta as entidades, relacionamentos e atributos identificados e como forma de apresentação do nosso modelo ao utilizador, desenvolveu-se o diagrama conceptual apresentado de seguida.

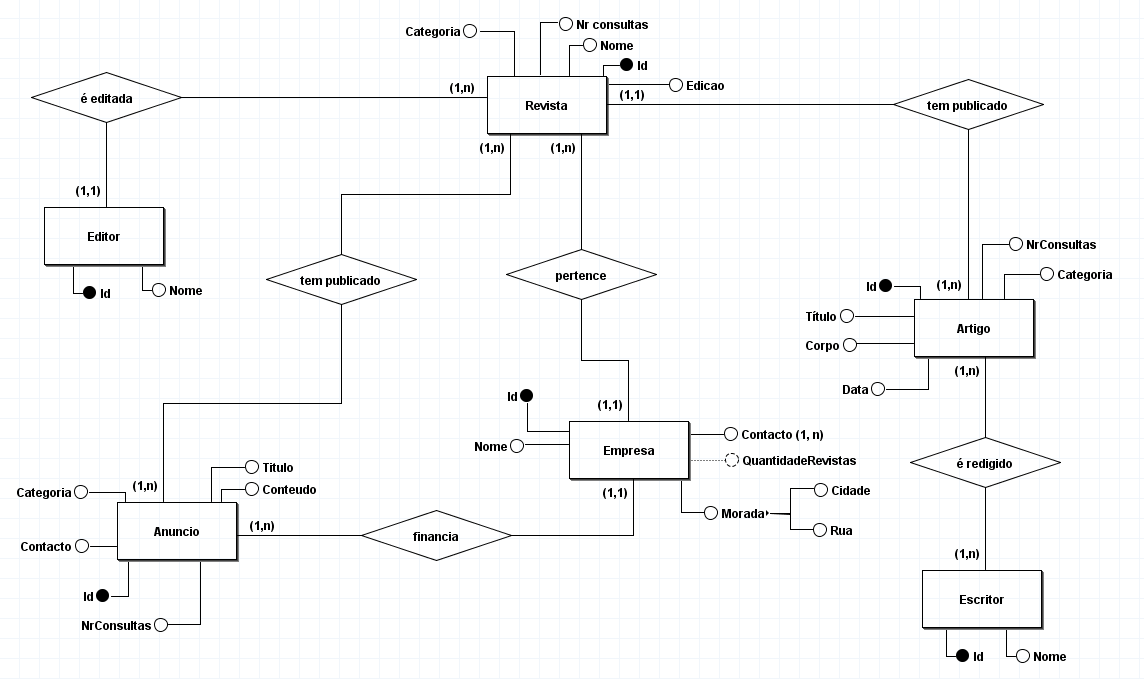


Figura 7 - Modelo Conceptual

* 1. Validação do modelo de dados com o utilizador

Concluído o modelo conceptual, achamos por bem pedir a aprovação da nossa cliente e perceber o que concorda e o que não concorda da organização deste projeto.

É importante que o nosso modelo represente todos os requisitos que foram solicitados pela Dona Ritinha.

Assim, expôs-se o nosso modelo à Dona Ritinha esclarecendo cada uma das decisões tomadas ao longo de toda a elaboração do modelo.

A cliente avaliou a opção que lhe foi proposta e verificou então que todos os requisitos solicitados estão todos representados de forma correta.

Dado que todos os requisitos são satisfeitos, a cliente autorizou a continuação do projeto, agora a elaboração do modelo lógico.

1. Modelação Lógica

Nesta secção iremos descrever a construção e validação do nosso modelo lógico, apresentando o desenho do mesmo com recurso ao programa *MySQL Workbench*.

No processo de validação do nosso modelo verificamos que este se encontrava normalizado e que era capaz de responder às interrogações e transações por parte de cada utilizador.

* 1. Construção e validação do modelo de dados lógico

A obtenção do modelo lógico foi feita a partir do modelo conceptual através da aplicação de variados passos específicos. De acordo com esses passos, criou-se uma tabela para cada uma das entidades, para cada relacionamento de N para N e para cada atributo multivalorado representados no modelo conceptual já apresentado anteriormente.

De seguida iremos enumerar as nossas relações, representando as chaves primárias usando um sublinhado contínuo e as chaves estrangeiras usando um sublinhado tracejado. Para as chaves que são simultaneamente primárias e estrangeiras utilizamos o duplo tracejado.

* Revista: ID, Nome, Edição, Categoria, NrConsultas, Editor\_ID, Empresa\_ID
* Artigo: ID, Título, Conteúdo, Data, Categoria, NrConsultas, Revista\_ID
* Escritor: ID, Nome
* Editor: ID, Nome
* Empresa: ID, Nome, Cidade, Rua, QuantidadeRevistas
* Anuncio: ID, Contacto, Categoria, Título, Conteúdo, NrConsultas, Empresa\_ID
* AnuncioRevista: Anuncio\_ID, Revista\_ID
* EscritorArtigo: Artigo\_ID, Escritor\_ID
* Contactos: Empresa\_ID, Contacto

Na tabela Empresa, o atributo composto Morada presente no modelo conceptual é traduzido pelos seus atributos simples (Cidade e Rua) no modelo lógico.

Na tabela Revista, os relacionamentos de N para 1 entre esta entidade e o Editor e entre esta entidade e a Empresa levaram à criação de duas chaves estrangeiras, Editor\_ID e Empresa\_ID.

Na tabela Artigo, o relacionamento de N para 1 entre esta entidade e a Revista levou à criação da chave estrangeira Revista\_ID.

Na tabela Anúncio, o relacionamento de N para 1 entre esta a entidade Empresa levou à criação da chave estrangeira Empresa\_ID.

As relações AnuncioRevista e EscritorArtigo representam, respetivamente, os relacionamentos entre as entidades Anuncio e Revista e entre Escritor e Artigo que no modelo conceptual eram relacionamentos de N para N. Nas tabelas AnuncioRevista e EscritorArtigo podemos constatar que cada uma tem uma cópia das chaves primárias de ambas as entidades envolvidas, que servem de chaves estrangeiras. Assim, estas chaves vão fazer parte da chave primária da tabela do relacionamento.

Houve necessidade da criação da tabela Contactos, pois este (Contacto) era um atributo multivalorado da empresa, o que levou então à criação de uma tabela com duas chaves estrangeiras que também são primárias: Empresa\_id e o próprio atributo contacto.

* 1. Desenho do modelo lógico

Com recurso ao *MySQL Workbench* obtivemos o seguinte modelo lógico:

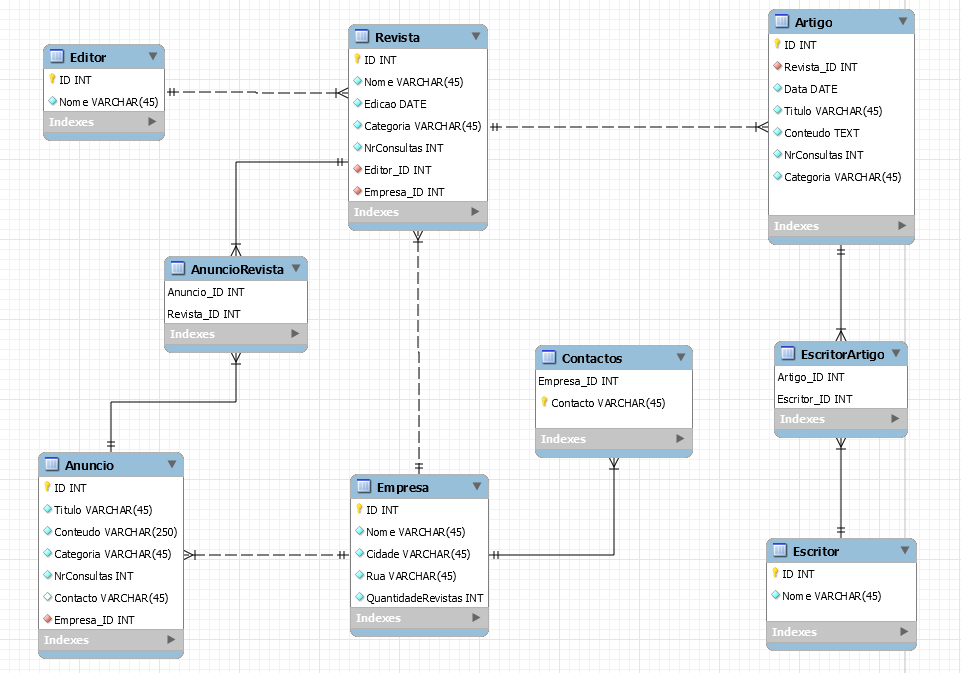


Figura 8 - Modelo Lógico

* 1. Validação do Modelo através da Normalização
     1. Primeira forma normal (1NF)

Segundo a definição, presente em Connoly (Connoly,2015), podemos afirmar que se uma relação na qual a interseção de cada linha e coluna contém um e um só elemento, está na primeira forma normal.

O nosso modelo lógico obedece a esta regra logo nada temos a alterar.

* + 1. Segunda forma normal (2NF)

Todos os atributos de uma tabela que não são chaves primárias têm que ser totalmente e funcionalmente dependentes da chave primária da mesma (Connoly, 2015).

O nosso modelo lógico cumpre esta regra, pois não existem dependências funcionais parciais.

* + 1. Terceira forma normal (3NF)

Uma relação que está na primeira e segunda forma normal e em que todos os atributos, que não são chaves primárias, são transitivamente dependentes da chave primária está na terceira forma normal (Connoly, 2015).

O nosso modelo lógico cumpre a terceira forma normal e para se fazer esta verificação analisamos as dependências funcionais de cada relação do modelo lógico com o objetivo de se encontrar dependências transitivas. As dependências funcionais dessas relações são:

Revista: ID → Nome, Edição, Categoria, NrConsultas, Editor\_ID, Empresa\_ID

Editor: ID → Nome

Artigo: ID → Título, Corpo, Data, Categoria, NrConsultas, Revista\_ID

Escritor: ID → Nome

Empresa: ID → Nome, Cidade, Rua, QuantidadeRevistas

Anuncio: ID → Contacto, Categoria, Título, Conteúdo, NrConsultas, Empresa\_ID

Concluímos assim que todos os atributos, à exceção daqueles que são chave primária, dependem apenas da respetiva chave primária da relação e não de outros atributos.

* 1. Validação do modelo com interrogações do utilizador

Ao entrevistarmos a Dona Ritinha facilmente concluímos que ela pretendia que a aplicação fosse capaz de fazer vários tipos de pesquisas considerando diversos parâmetros. Tendo em conta os desejos da nossa cliente elaboramos os requisitos de exploração, e apresentamos agora como é suposto a base de dados “responder” a estas.

* + 1. Permitir a consulta das informações de uma empresa

As informações relativas a uma empresa podem ser obtidas através da consulta da tabela Empresa.

* + 1. Listar os contactos de uma dada empresa

A partir da relação Empresa e da relação Contacto, somos capazes de obter os contactos da empresa em questão.

* + 1. Permitir a consulta das informações de um anúncio

As informações relativas a um anúncio podem ser obtidas através da consulta da tabela Anuncio.

* + 1. Permitir a consulta das informações de um artigo

As informações relativas a um artigo podem ser obtidas através da consulta da tabela Artigo.

* + 1. Listar todos as revistas de uma dada empresa

De forma a obter as revistas de uma dada empresa, basta consultar as revistas cujo ao atributo Empresa\_ID da mesma seja igual ao atributo ID da empresa para a qual pretendemos obter a informação.

* + 1. Listar todos os artigos contidos numa revista

Para obter os artigos contidos numa revista, basta consultar os artigos cujo ao atributo Revista\_ID seja igual ao atributo ID da revista para a qual pretendemos obter a informação.

* + 1. Listar todos os anúncios contidos numa revista

Para obter os anúncios contidos numa revista, basta consultar as relações da tabela AnuncioRevista para as quais o atributo Revista\_ID seja igual ao ID da revista em questão.

* + 1. Listar todos os artigos escritos por um dado escritor

De forma a obter os artigos escritos por um escritor, basta consultar as relações da tabela ArtigoEscritor para as quais o atributo Escritor\_ID seja igual ao ID do escritor em questão.

* + 1. Listar todas as revistas editadas por um dado editor

De forma a obter as revistas editadas por um editor, basta percorrer as revistas e verificar se o editor da mesma coincide com o pretendido.

* + 1. Identificar a quantidade de revistas visualizadas de uma empresa

Para obter a quantidade de revistas visualizadas de uma empresa, basta somar o número de consultas (NrConsultas) de cada revista que a empresa em questão publicou.

* + 1. Listar todos os escritores que têm pelo menos um artigo escrito numa determinada revista

Para listar todos os escritores que têm pelo menos um artigo escrito numa determinada revista, temos de percorrer os artigos publicados na revista pretendida e de seguida verificar os escritores que contribuíram na redação desse artigo.

* + 1. Listar todas as revistas de uma dada empresa numa determinada data ou intervalo de datas

De forma a obter a listagem de todas as revistas de uma dada empresa numa determinada data ou intervalo de datas, basta percorrer todas as revistas de uma dada empresa e verificar se o atributo Data se encontra dentro dos parâmetros pretendidos.

* + 1. Listar todos os escritores que já contribuíram para a escrita de revistas de uma dada empresa

De forma a obter a listagem de todos os escritores que já contribuíram para a escrita de revistas de uma dada empresa, temos de percorrer as revistas publicadas por uma empresa, verificar os artigos publicados na mesma para depois verificarmos os escritores contribuintes na sua redação.

* + 1. Listar todas as revistas em que aparecerem os anúncios de uma determinada empresa

Para obter todas as revistas em que aparecerem os anúncios de uma determinada empresa, basta verificar a relação entre Anuncio e Empresa, através do campo Empresa\_ID, para posteriormente verificarmos as revistas nas quais esses anúncios estão presentes.

* + 1. Listar as N revistas com maior número de consultas

Para a resolução desta interrogação basta ordenar decrescentemente as revistas por número de consultas (NrConsultas) e aplicando um limite (N) ao número de resultados devolvidos.

* + 1. Listar as categorias por frequências relativas de consultas

Para calcular as frequências relativas das consultas das categorias, primeiro calculamos o número de consultas totais, ou seja, o número de consultas de todas as revistas. De seguida, agrupamos a tabela das revistas por categoria e somamos o número de consultas das revistas da mesma categoria. Por fim, dividimos esse valor pelo número total de consultas multiplicando por cem de modo a obter as frequências relativas.

* 1. Validação do modelo com as transações estabelecidas

Para que as informações que constam na base de dados sejam facilmente acedidas, optamos por desenvolver operações que permitem executar tarefas básicas na aplicação onde irá operar a base de dados. A este conjunto de operações dá-se o nome de transações. Quando as operações de uma transação são todas executadas, a transação é executada. No caso de ocorrer um erro numa dada operação, todas as outras serão anuladas e a transação não ocorre.

Consideramos essencial o desenvolvimento de transações que permitam a inserção de uma nova revista, uma vez que este era o principal objetivo da Dona Ritinha. As transações que permitem remover e atualizar dados não deverão ser executadas com muita frequência tendo em conta o objetivo do projeto.

Para inserir uma revista é necessário que a transação tenha acesso a todas as novas informações da mesma, tais como nome, edição e categoria (o número de consultas começará sempre igual a zero). Primeiro são preenchidos todos os atributos do novo registo na tabela Revista com a nova informação e, de seguida, completa-se a tabela Artigo com os artigos que constam na Revista recém adicionada. Na tabela Escritor serão inseridos todos os que ainda não constarem da base de dados e efetua-se o relacionamento dos mesmos com os Artigos, adicionando na tabela EscritorArtigo. Na tabela Anuncio são inseridos os novos registos e após isso é preenchida a tabela que traduz o relacionamento entre Anuncio e Revista, inserindo os respetivos identificadores. Na tabela Empresa preenche-se o que ainda não constar da base de dados, assim como na tabela Editor.

* 1. Revisão do modelo lógico com o utilizador

Após a elaboração do modelo lógico, entrevistamos novamente a Dona Ritinha para podermos expor o novo modelo e obter o seu parecer relativamente ao mesmo. Tendo em conta que a interpretação do modelo lógico já requer mais conhecimentos de base de dados, tentamos passar a informação de um modo mais funcional e simplificado para que a Dona Ritinha pudesse ter uma visão geral do funcionamento.

Após ter compreendido o modelo a Dona Ritinha verificou que todos os requisitos que havia estabelecido no início estavam presentes e que os atributos de cada entidade eram capazes de armazenar toda a informação que se pretendia. De seguida a Dona Ritinha questionou como seria possível responder às várias questões que surgem ao pesquisar os vários dados. No seguimento desta questão apresentamos as questões e respetivas respostas presentes na secção 4.4 as quais deram uma visão do método de pesquisa e tiveram a sua concordância.

Perante esta situação tivemos autorização por parte da Dona Ritinha para avançar com o projeto.

1. Implementação Física

Nesta secção iremos abordar os aspetos relativos à metodologia para a construção do modelo físico bem como alguns exemplos. Será detalhado o processo de transição do modelo lógico para o modelo físico, utilizando para esse efeito uma ferramenta de gestão de Base de Dados, nomeadamente o MySQL.

* 1. Seleção do sistema de gestão de bases de dados

O Sistema de Gestão de Base de Dados utilizado para desenvolver a base de dados relacional para este projeto foi o MySQL. Esta ferramenta apresenta várias vantagens, as quais passamos de seguida a enumerar algumas delas:

* É um software grátis;
* É fácil de utilizar;
* Tem um elevado desempenho;
* Possui compatibilidade cruzada, isto é, pode ser instalado em todos os sistemas operativos principais sem perda de desempenho.

Como nenhum sistema é perfeito o MySQL também tem algumas desvantagens, mas tendo em conta todos os fatores considerou-se que o MySQL seria uma boa opção para o projeto em questão.

* 1. Tradução do esquema lógico para o sistema de gestão de bases de dados escolhido em SQL

O programa utilizado para o desenvolvimento desta base de dados foi o “MySQL Workbench”. Esta ferramenta possibilita a utilização de uma opção, Forward Engineer, após o modelo lógico estar definido. Este comando transforma o modelo lógico desenvolvido em físico, facilitando a implementação física da base de dados ao gerar automaticamente o código SQL.

* 1. Tradução das interrogações do utilizador para SQL (alguns exemplos)

Todos os requisitos de exploração definidos na secção 2.2.2 foram criados através de procedimentos com exceção do requisito RE9 que foi criado como uma função. Essa decisão baseou-se no facto de se pretender procurar informações na base de dados através de parâmetros, pelo que apenas é necessário projetar informação das tabelas. No requisito RE9 optamos por criar uma função pois o objetivo é devolver um número (número total de consultas das revistas de uma determinada empresa).

Para além dos conceitos abordados acima, foram também usadas as funções elementares do *SQL* como *SELECT*, *WHERE*, *ORDER BY*, *GROUP BY*, *INNER JOIN*.

De seguida encontram-se alguns exemplos do código *SQL* criado de forma a responder aos requisitos de exploração.

* + 1. Listar todos os escritores que já contribuíram para a escrita de revistas de uma dada empresa

Este procedimento faz a pesquisa em cinco tabelas, daí serem usados quatro join’s para obter a tabela resultante. É ainda feita uma restrição ao nome da empresa.

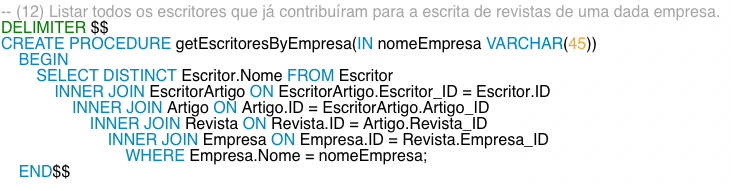


Figura 9 - Tradução do RE5 para *SQL*

* + 1. Calcular o número total de consultas das revistas de uma empresa

Esta função foi a única que necessitamos de criar pois é o único requisito que devolve apenas um valor. Primeiro, tabelámos a coluna que contém o número de consultas das revistas, selecionando as revistas que referenciam a empresa pretendida. De seguida, aplicámos a função SUM do SQL para somar o número de consultas obtendo o resultado pretendido.

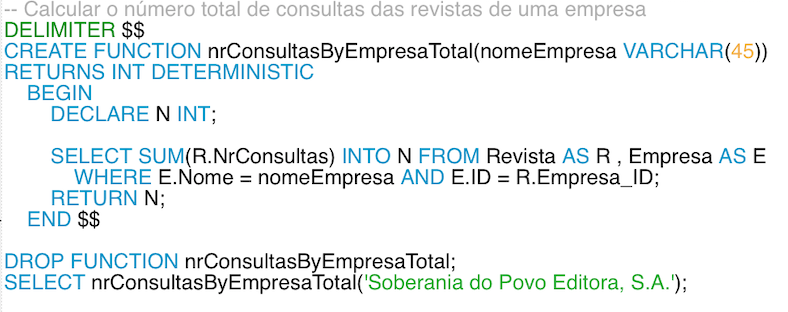


Figura 10 - Tradução do RE10 para *SQL*

* + 1. Listar as categorias por frequências relativas de consultas

Este procedimento traduz o requisito que pede para calcular as frequências relativas de consultas por categoria. Numa primeira fase, calculámos a tabela que contém as categorias e os respetivos números de consultas totais usando a função GROUP BY, que permite agrupar as categorias pelo respetivo somatório do número de consultas, eliminando as repetições. De seguida, a cada elemento da coluna que contém a soma total do número de consultas de uma certa categoria, vamos dividi-lo pelo número total de visualizações de todas as revistas e multiplicámos por cem, obtendo assim as frequências relativas.

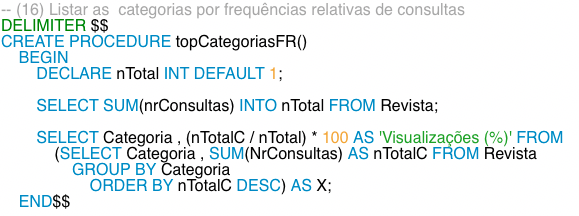


Figura 11 - Tradução do RE15 para *SQL*

* 1. Tradução das transações estabelecidas para SQL (alguns exemplos)

Nesta seção apresentamos as transações criadas para a base de dados. Foram desenvolvidas transações para inserção de dados em todas as tabelas do arquivo de revistas. De seguida, apresentamos alguns exemplos dessas transações.

* + 1. Inserir uma empresa

O procedimento recebe como parâmetros o nome, a cidade e a rua da empresa. Primeiro é calculado o índice que esta nova entrada na tabela vai ter como chave primária. De seguida, testamos se existem empresas com o mesmo nome. Caso não haja, a nova empresa é adicionada à tabela usando o comando *COMMIT* e é lançada uma mensagem de sucesso. Caso haja uma empresa com o mesmo nome, através do comando *ROLLBACK*, desfazemos as alterações feitas e é lançada uma mensagem de erro.

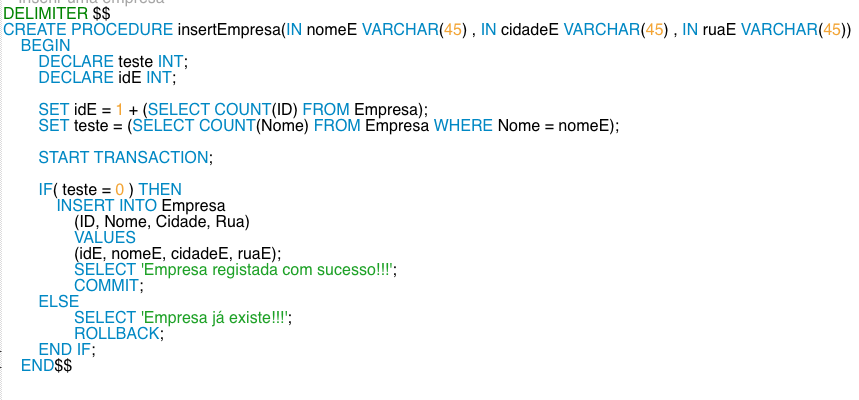


Figura 12 - Tradução da transação de inserção de uma empresa em *SQL*

* + 1. Inserir um artigo

O procedimento recebe como parâmetros o nome da revista à qual pertence o artigo a inserir, a data em que foi escrito, o seu título, o corpo (conteúdo), a categoria em que está inserido e, por fim, o seu escritor. Numa primeira fase é calculado os *ID’s* da revista e do escritor do artigo bem como o *ID* que a nova entrada irá ter. De seguida, caso os *ID’s* da revista e do escritor existam , é adicionado o artigo à tabela com referência à respetiva revista, e é atualizada a tabela EscritorArtigo, adicionando o novo relacionamento entre o artigo adicionado e o seu escritor através do comando  *COMMIT*. Caso os *ID’s* da revista ou do escritor do artigo não existam, a inserção fica sem efeito aplicando o comando *ROLLBACK.*

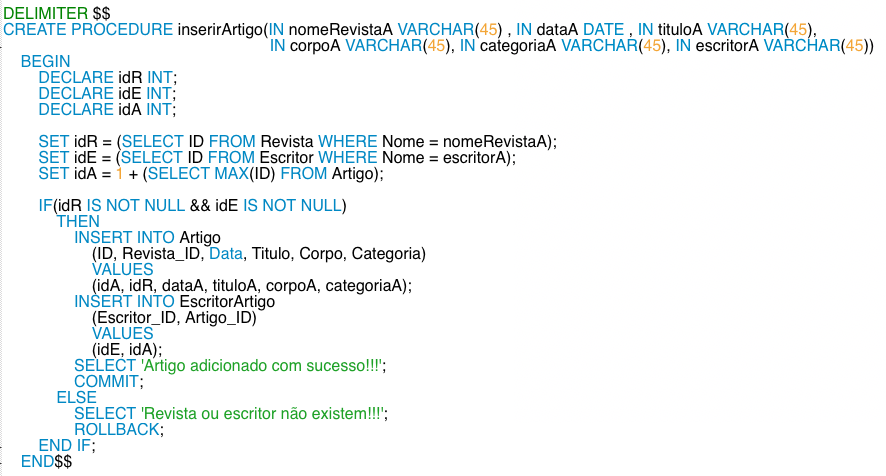
****

Figura 13 - Tradução de transação da inserção de um artigo em *SQL*

* 1. Escolha, definição e caracterização de índices em SQL (alguns exemplos)

O *MySQL* faz a atribuição automática de índices aos atributos que sejam chaves primárias, pelo que todas as chaves primárias envolvidas nas tabelas da base de dados são índices. Para além disso, todos os atributos que sejam restringidos como *UNIQUE*, também são indexados.

Face às traduções das interrogações dos utilizadores, podemos verificar que a maior parte das consultas às tabelas são feitas através do atributo ‘Nome’ das diferentes entidades. Com isto, considerámos restringir esses mesmos atributos como *UNIQUE* para que, consequentemente, a pesquisa sobre as tabelas fosse mais eficiente. No entanto, tal decisão não fazia sentido dado o contexto do problema pois é permitido haver, por exemplo, duas empresas com o mesmo nome ou duas revistas com o mesmo nome.

* 1. Estimativa do espaço em disco da base de dados e taxa de crescimento anual
     1. Estimativa do espaço em disco da base de dados

Cada em entrada em cada relação vai ocupar um certo número de bytes. Comecemos por calcular esse espaço:

**Relação Revista**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Nome da Tabela | Atributos | Tipo | Espaço ocupado (bytes) | Total |
| Revista | ID  Nome  Edicao  Categoria  NrConsultas  Editor\_ID  Empresa\_ID | INT  VARCHAR(45)  DATE  VARCHAR(45)  INT  INT  INT | 4  45  3  45  4  4  4 | 109 |

Tabela 4 - Espaço por entrada da tabela Revista

**Relação Empresa**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Nome da Tabela | Atributos | Tipo | Espaço ocupado (bytes) | Total |
| Empresa | ID  Nome  Cidade  Rua  QuantidadeRevistas | INT  VARCHAR(45)  VARCHAR(45)  VARCHAR(45)  INT | 4  45  45  45  4 | 143 |

Tabela 5 - Espaço por entrada da tabela Empresa

**Relação Contacto**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Nome da Tabela | Atributos | Tipo | Espaço ocupado (bytes) | Total |
| Contacto | Empresa\_ID  Contacto | INT  VARCHAR(45) | 4  45 | 49 |

Tabela 6 - Espaço por entrada da tabela Contacto

**Relação Artigo**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Nome da Tabela | Atributos | Tipo | Espaço ocupado (bytes) | Total |
| Artigo | ID  Revista\_ID  Data  Titulo  Corpo  NrConsultas  Categoria | INT  INT  DATE  VARCHAR(45)  VARCHAR(2500)  INT  VARCHAR(45) | 4  45  250  45  4  45  4 | 2605 |

Tabela 7 - Espaço por entrada da tabela Artigo

**Relação Escritor**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Nome da Tabela | Atributos | Tipo | Espaço ocupado (bytes) | Total |
| Escritor | ID  Nome | INT  VARCHAR(45) | 4  45 | 49 |

Tabela 8 - Espaço por entrada da tabela Escritor

**Relação Anuncio**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Nome da Tabela | Atributos | Tipo | Espaço ocupado (bytes) | Total |
| Anuncio | ID  Titulo  Conteudo  Categoria  NrConsultas  Contacto  Empresa\_ID | INT  VARCHAR(45)  VARCHAR(250)  VARCHAR(45)  INT  VARCHAR(45)  INT | 4  45  250  45  4  45  4 | 397 |

Tabela 9 - Espaço por entrada da tabela Anuncio

**Relação Editor**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Nome da Tabela | Atributos | Tipo | Espaço ocupado (bytes) | Total |
| Editor | ID  Nome | INT  VARCHAR(45) | 4  45 | 49 |

Tabela 10 - Espaço por entrada da tabela Editor

**Relação AnuncioRevista**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Nome da Tabela | Atributos | Tipo | Espaço ocupado (bytes) | Total |
| AnuncioRevista | Anuncio\_ID  Revista\_ID | INT  INT | 4  4 | 8 |

Tabela 11 - Espaço por entrada da tabela AnuncioRevista

**Relação ArtigoEscritor**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Nome da Tabela | Atributos | Tipo | Espaço ocupado (bytes) | Total |
| ArtigoEscritor | Artigo\_ID  Escritor\_ID | INT  INT | 4  4 | 8 |

Tabela 12 - Espaço por entrada da tabela ArtigoEscritor

Necessitamos agora de uma estimativa do espaço total ocupado. Para tal iremos ter como base revistas apenas do presente ano. Iremos considerar o seguinte caso: 250 revistas, 830 artigos, 30 empresas, 35 anúncios, 20 editores e 17 escritores. Iremos também considerar que cada empresa terá 2 contactos, existirão 10 anúncios por revista, as revistas apenas terão 1 editor e os artigos irão ter 3 escritores.

Assim, para este tipo de povoamento, o espaço ocupado seria cerca de:

250 \* 109 + 830 \* 2605 + 30 \* 143 + 35 \* 397 + 20 \* 49 + 17 \* 49 + 30 \* 49 \* 2 + 250 \* 10 \* 8 + 250 \* 8 + 830 \* 8 \* 3 = 2239513 *bytes* = 2.135 *MB*.

* + 1. Taxa de crescimento anual

Para determinar a taxa de crescimento anual, vamos supor que diariamente são inseridas 150 revistas, com 25 anúncios e 70 artigos cada uma. Iremos também assumir que cada revista terá 1 editor, os artigos irão ter 3 escritores cada um e que serão inseridas 0.1 empresas novas, 0.2 editores e 0.3 escritores. Assim sendo e tendo em conta que um ano tem 365 dias, o espaço ocupado aumenta anualmente da seguinte forma:

365 \* (150 \* 109 + 25 \* 397 + 70 \* 2605 + 70 \* 8 + 3 \* 70 \* 8 + 0.2 \* 8 + 0.3 \* 8 + 0.1 \* 143) = 76972404.5 *bytes =* 73.4 *MB*

* 1. Definição e caracterização das vistas de utilização do SQL (alguns exemplos)

As interrogações traduzidas para *SQL* baseiam-se todas em pesquisas personalizadas pelo que necessitam quase sempre de parâmetros de entrada para poderem ser executadas. Consequentemente, decididos criar vistas sobre todas as tabelas com informação relevante tanto para o visitante do site como para o utilizador da aplicação. Foram então criadas vistas sobre todas as entidades da base de dados com exceção das tabelas que surgiram do relacionamento entre entidades. Para além das vistas básicas sobre as entidades, foram também criadas mais duas vistas, uma para a consulta das empresas e o número total de consultas das suas revistas e outra para listar o *Top 5* de escritores, ou seja, os escritores cujos artigos têm mais visualizações. De seguida, encontram-se alguns exemplos do código *SQL* de algumas vistas.

* + 1. Vista da Tabela Escritor

Vista muito simples que permite aceder ao nome de todos os escritores que estão na base de dados do arquivo de revistas.



Figura 14 - Vista sobre a tabela Escritor

* + 1. Vista das empresas e respetivas visualizações das suas revistas

Vista que permite visualizar o nome de todas as empresas e respetivas visualizações das suas revistas. Para tal, juntámos as duas tabelas Revista e Empresa, associando uma empresa às respetivas revistas. De seguida, aplicámos a função *GROUP BY* de modo a eliminar as repetições e somar o número de consultas. O resultado é apresentado ordenado por ordem decrescente, para tal, usámos a função *ORDER BY*.

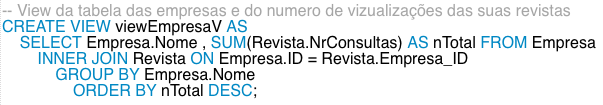


Figura 15 - Vista das empresas e respetivas visualizações totais das suas revistas

* + 1. Vista do *Top 5* de escritores

Vista que permite visualizar o *Top 5* de escritores. Para a vista ser criada, simplesmente juntámos as tabelas Escritor e EscritorArtigo e agrupámos a tabela pelo nome de escritor contando o número de ocorrências que um dado escritor tem na última tabela mencionada. Para o resultado final pretendido, usámos a função de ordenação *ORDER BY* para ordenar as entradas pelo número total de artigos e limitando a tabela em cinco entradas através da função *LIMIT* com o parâmetro cinco.



Figura 16 – Vista do *Top 5* de escritores

* 1. Definição e caracterização dos mecanismos de segurança em SQL (alguns exemplos)

O site desenvolvido para a Dona Ritinha é utilizado por visitantes (ações de consulta) e pela própria Dona Ritinha (ações de inserção, atualização, remoção e consulta). Dado isto, decidimos criar apenas dois tipos de utilizadores, o administrador (admin) da base de dados e o da aplicação (app), visto que é esta que vai gerir os movimentos da base de dados.

O administrador da base de dados tem todas as permissões possíveis sobre a base de dados visto que é este que vai gerir a mesma.

O utilizador da aplicação tem permissões de seleção, inserção, execução, atualização e de remoção sobre todas as tabelas com exceção das tabelas que tiveram origem das ligações (N:N).

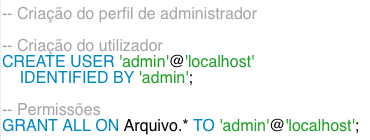


Figura 17 - Código *SQL* para a criação da vista de utilização e respetivas permissões do utilizador admin

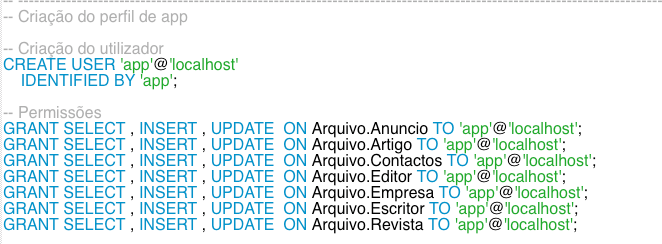


Figura 18 - Código *SQL* para a criação da vista de utilização e respetivas permissões do utilizador app

* 1. Revisão do sistema implementado com o utilizador

Terminada a implementação do modelo físico, a equipa reuniu-se com a Dona Ritinha uma última vez para garantir que todos os requisitos foram cumpridos e que o projeto desenvolvido se encontra apto para ser lançado.

Tendo em conta que a Dona Ritinha não possui conhecimentos de informática, a equipa decidiu realizar uns requisitos extra que virão a ser necessários para a aplicação que irá gerir a base de dados, como por exemplo, as definições das vistas sobre as tabelas mais importantes da base de dados, assim como um trigger que atribui o número total de revistas de uma empresa ao atributo que guarda tal informação na entidade Empresa.

No geral, a Dona Ritinha está satisfeita com o trabalho desenvolvido para o seu arquivo de jornais e considera que todas as suas exigências foram cumpridas.

1. Migração da Base de Dados

Nesta secção será abordada a migração de base de dados de SQL para NoSQL e algumas questões associadas a esta.

* 1. Fundamentação da Migração da Base de Dados – (Identificação e descrição dos objetivos da base de dados em termos de aplicações e de utilizadores)

Após a implementação da base de dados relacional, a D.Ritinha dedicou-se bastante e sentiu o crescimento substancial do seu arquivo de jornais e a procura por parte dos utilizadores. Analisando este crescimento é evidente o aumento do número de relacionamentos em comparação aos restantes dados, o que origina uma extraordinária redução do desempenho na pesquisa de informações, devido à enorme quantidade de junções envolvidas, que são operações pesadas a nível de desempenho para a base de dados.

Nas bases de dados relacionais os esquemas são fixos, ou seja, por exemplo, quando não se quer inserir um determinado atributo para um registo, tem de se o colocar a Null, o que gasta memória desnecessária. Por outro lado, as bases de dados não relacionais (NoSQL) são sem esquema, ou seja, admitem esquemas variáveis. Assim, quando não se pretende inserir um determinado atributo, simplesmente não é necessário inserir.

Assim, optou-se por migrar a nossa base de dados relacional original para uma não relacional orientada a documentos (MongoDB).

As bases de dados não relacionais orientadas a documentos armazenam documentos flexíveis, o que significa que os campos podem variar de documento para documento e a estrutura de dados pode ser alterada ao longo do tempo, o que permite uma melhoria do desempenho na execução de pesquisa de informações. As suas caraterísticas permitem que as aplicações modelem informações de modo muito mais natural, pois os dados podem ser [aninhados](https://pt.wikipedia.org/w/index.php?title=Aninhar&action=edit&redlink=1) em hierarquias complexas e continuar a ser [indexáveis](https://pt.wikipedia.org/wiki/Indexar) e fáceis de aceder.

As bases de dados não relacionais orientadas a documentos disponibilizam uma maior escalabilidade e flexibilidade que tornam mais eficiente o desempenho da consulta/pesquisa de informações na base de dados. O modelo de documento mapeia os objetos no código do seu aplicativo, facilitando o trabalho com os dados.

* 1. Identificação e explicação do tipo de questões/necessidades que serão realizadas sobre o sistema de dados NoSQL
  2. Definição da estrutura base para o sistema de dados NoSQL

O grupo de trabalho juntou-se durante vários dias numa reunião e chegaram à decisão de utilizar apenas uma coleção que contém um conjunto de documentos. Cada documento é composto pela informação de uma revista (artigos publicados, anúncios publicados, empresa, editor e escritor dos artigos).

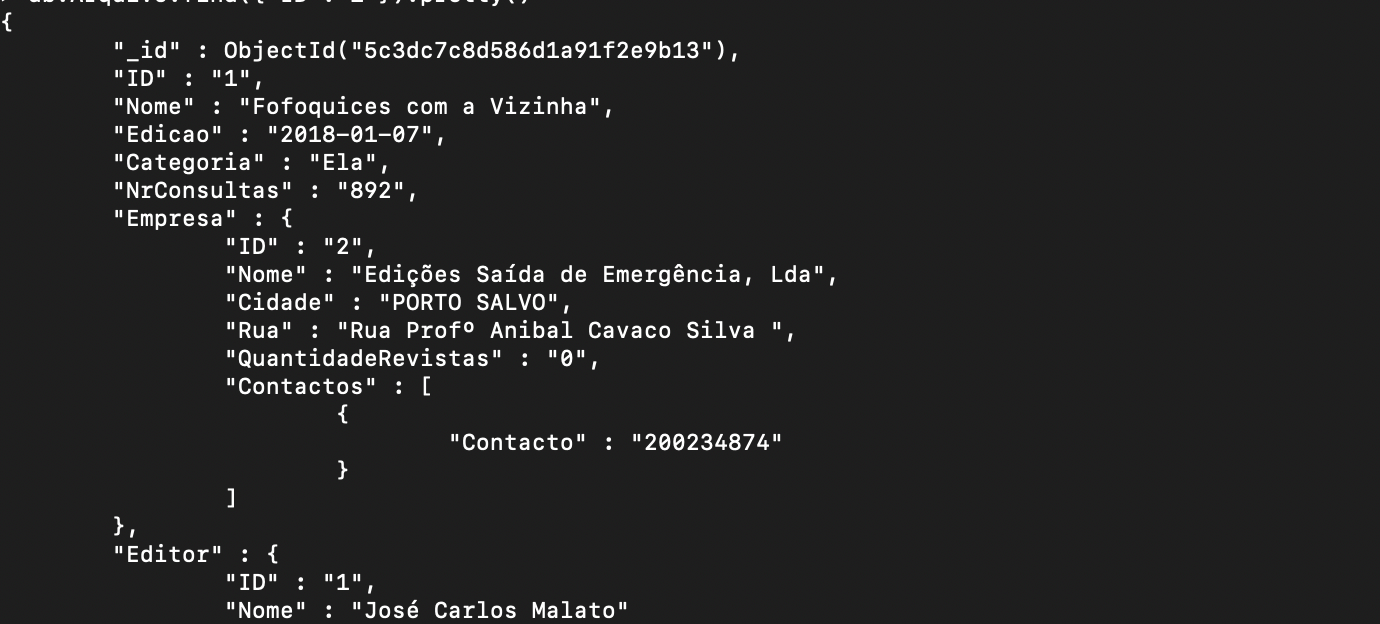


Figura 19 – Informação da Revista *“Fofoquices com a Vizinha”*

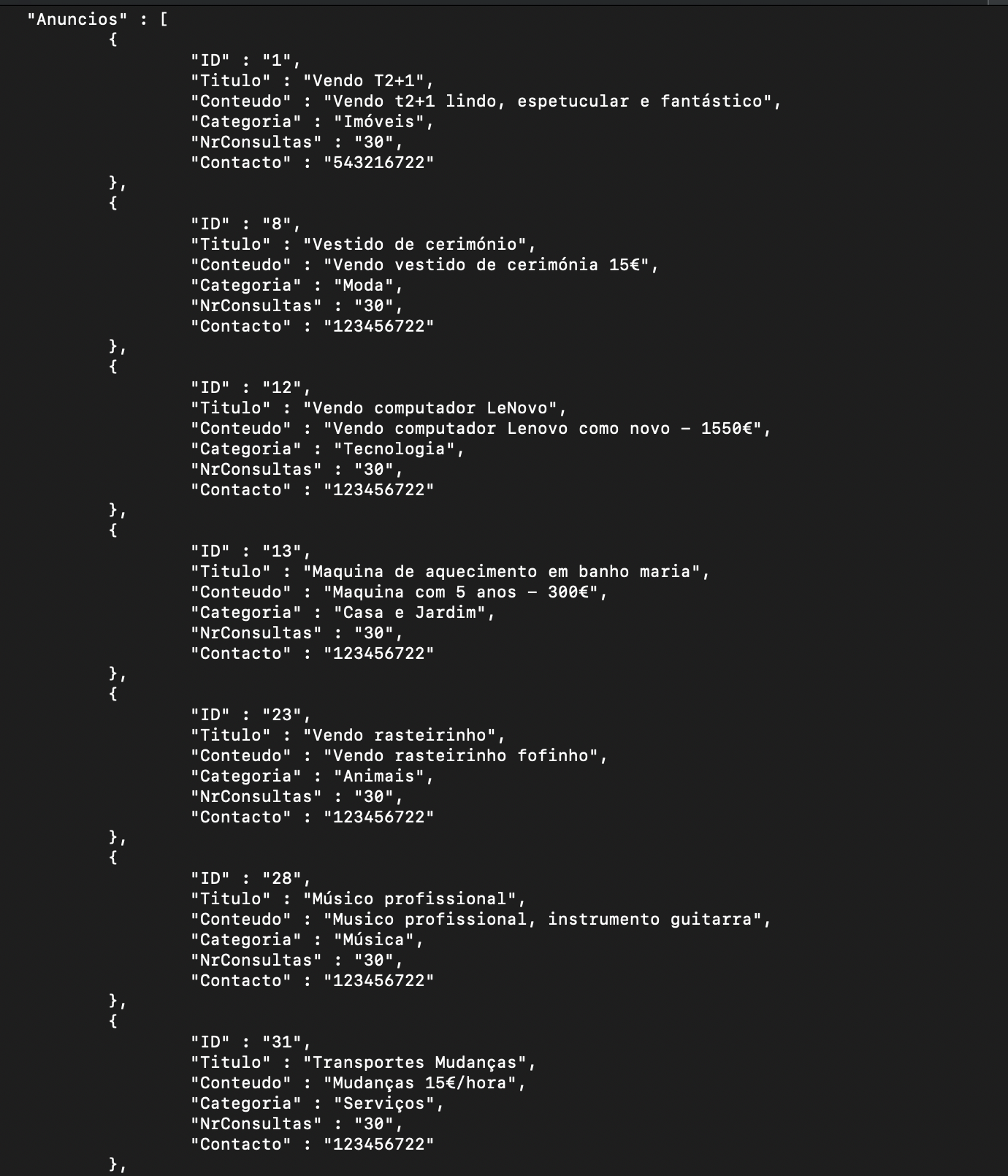


Figura 20 - Informação da Revista *“Fofoquices com a Vizinha”* (continuação)



Figura 21 - Informação da Revista *“Fofoquices com a Vizinha”* (continuação)

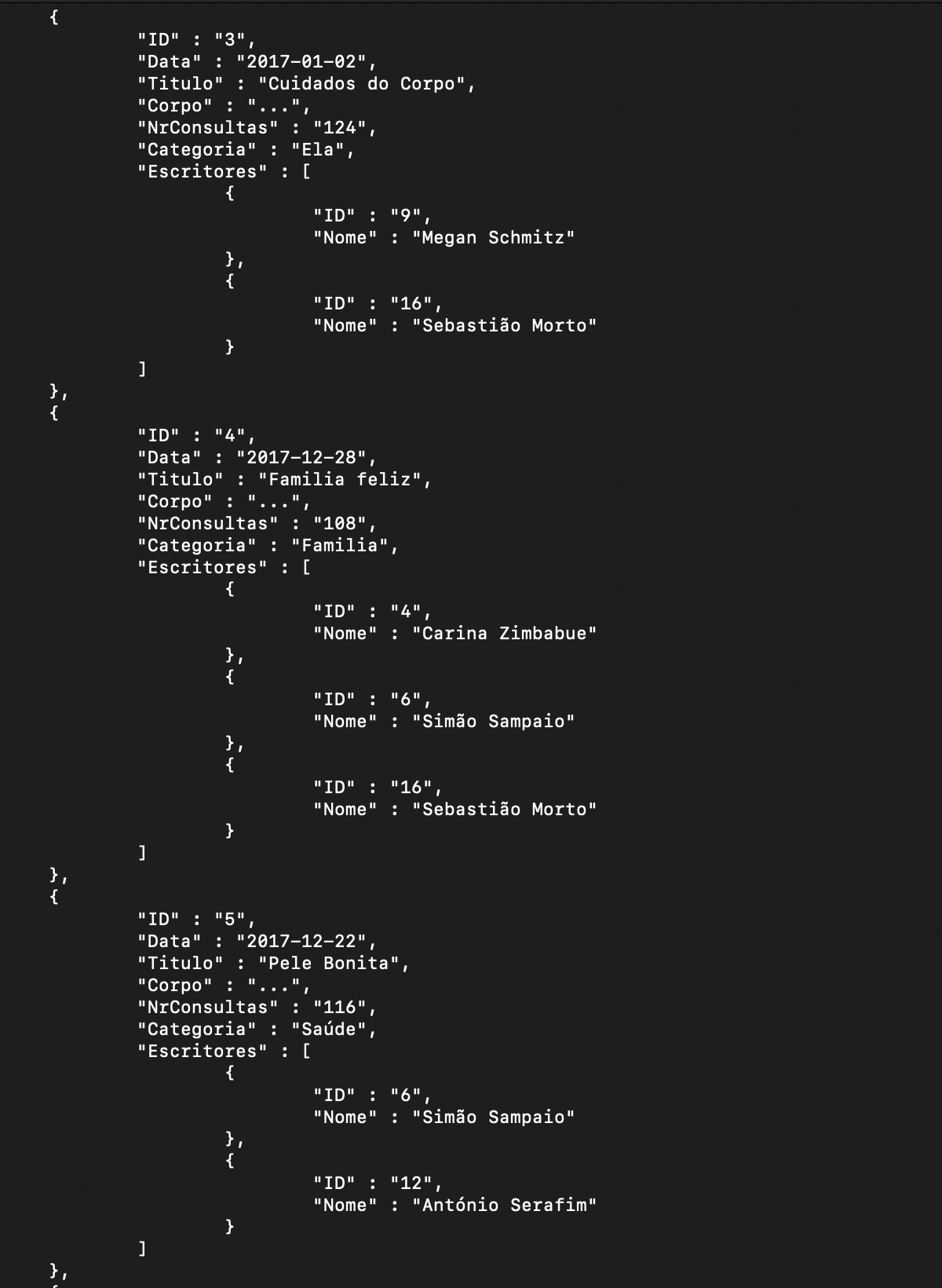


Figura 22 - Informação da Revista *“Fofoquices com a Vizinha”* (continuação)

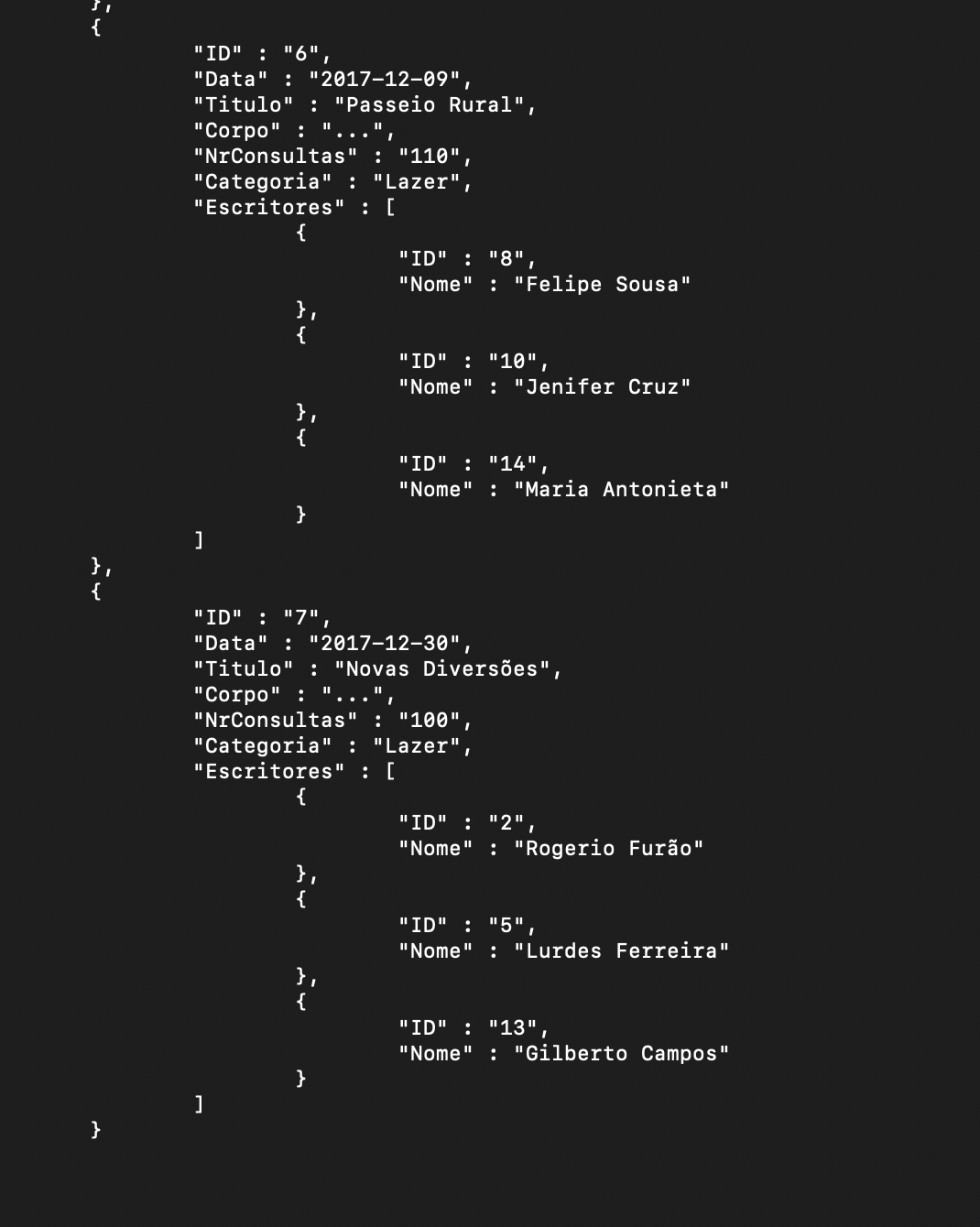


Figura 23 - Informação da Revista *“Fofoquices com a Vizinha”* (continuação)

* 1. Identificação dos objetos de dados no sistema SQL que serão utilizados para alimentar o novo sistema

Perante as necessidades da Dona Ritinha e para as pesquisas dos utilizadores, o grupo decidiu que era necessário ter disponível toda a informação relativa a cada revista, deste modo foi necessário fazer a migração de todos os dados do sistema SQL para o sistema NoSQL.

* 1. Mapeamento do processo de migração de dados, descrevendo o processo de conversão dos vários objetos de dados
  2. Explicação do processo de migração de dados, explicando de modo detalhado as suas principais etapas – extração, transformação e carregamento
  3. Apresentação e descrição da implementação do processo de migração de dados

Na implementação do processo de migração de dados foi utilizada a linguagem *JAVA*.

Nesse processo foi considerada a Revista como foco principal, onde foram agregadas todas as informações.

Cada documento Revista tem contido a informação da mesma, os Artigos nela publicados assim como os Escritores que os redigiram, os seus Anúncios, Empresa a que pertence e o seu Editor.

O conjunto de documentos Revista formam a coleção Arquivo.

* 1. Apresentação da forma como as questões identificadas anteriormente podem ser satisfeitas com o novo sistema, utilizando a linguagem de interrogação do sistema NoSQL

RE6:



Figura 24 – Query que apresenta os artigos contidos na Revista com ID 10



Figura 25 – Resultados obtidos à Query anterior (Figura 24)



Figura 26 - Resultados obtidos à Query anterior (Figura 24) (continuação)



Figura 27 - Resultados obtidos à Query anterior (Figura 24) (continuação)

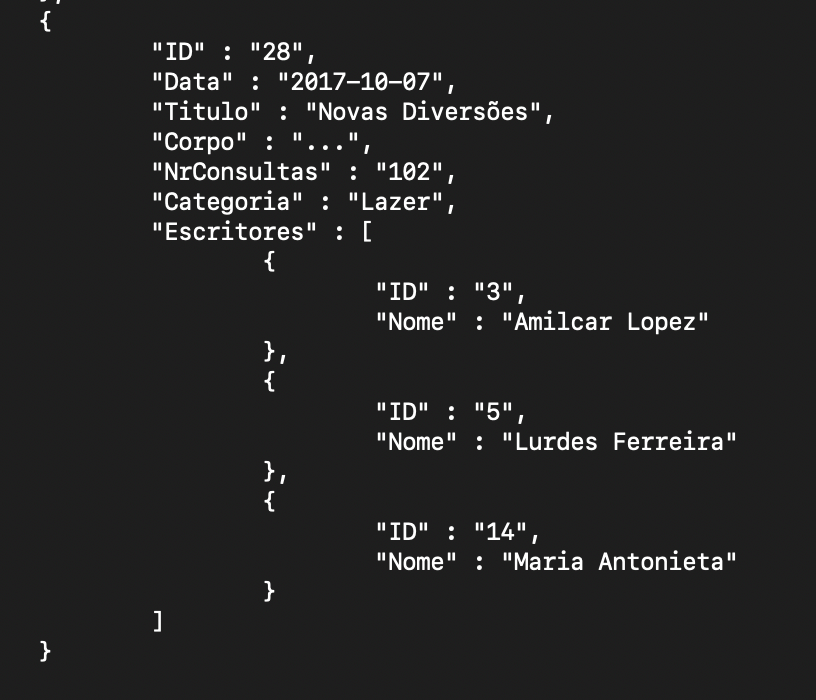


Figura 28 - Resultados obtidos à Query anterior (Figura 24) (continuação)

RE15:

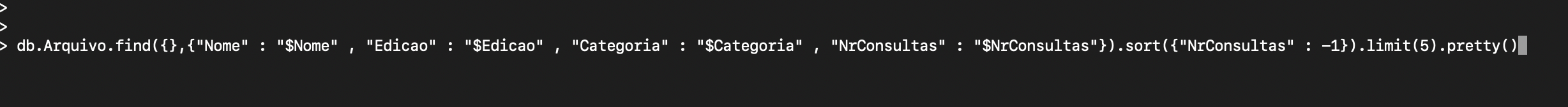


Figura 29 – Query que lista as n revistas com maior número de consultas

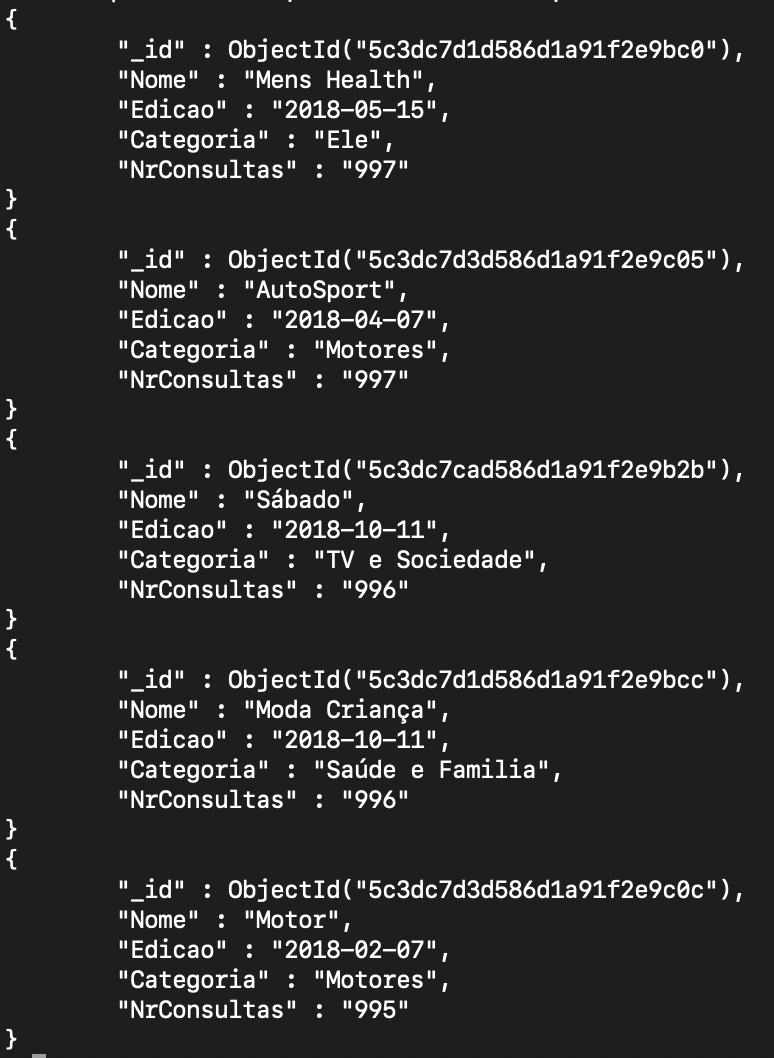


Figura 30 - Resultados obtidos à Query anterior (Figura 29)

RE12:

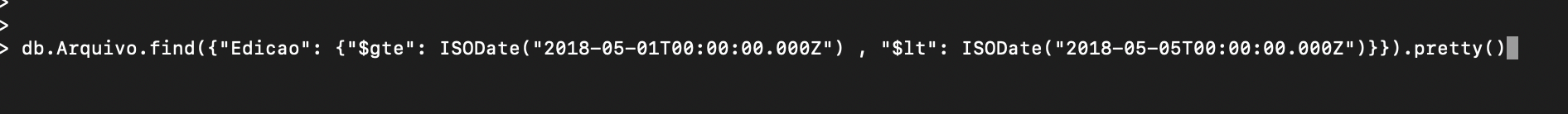


Figura 31 – Query que lista todas as revistas de uma dada empresa num interval de datas

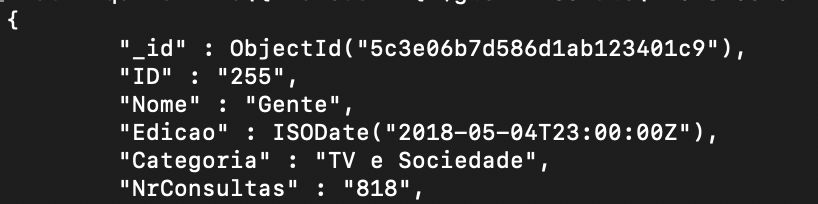


Figura 32 – Resultados obtidos à Query anterior (Figura 31)

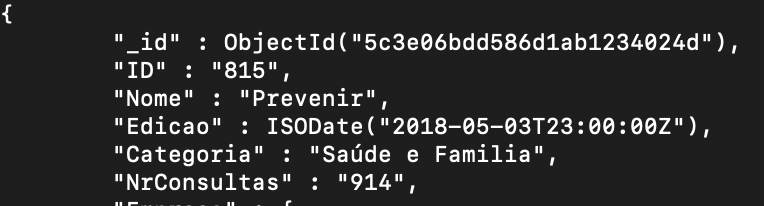


Figura 33 - Resultados obtidos à Query anterior (Figura 31) (continuação)

1. Conclusões (REVER ---🡪 NoSQL)

De uma forma geral, os quatro corajosos aspirantes a engenheiros informáticos avaliam de uma forma positiva o trabalho desenvolvido para a criação de uma base de dados para armazenamento digital de revistas. Na nossa opinião, um dos aspetos mais relevantes para o desenvolvimento desta base de dados foi desde início a enorme preocupação em sermos consistentes e evitar redundâncias, o que levaria a erros que mais tarde seriam difíceis ou impossíveis de resolver.

Tendo em mente os requisitos a que nos propusemos no início do projeto perante a nossa querida cliente, tivemos um excelente *feedback* da sua parte e um bolinho de chocolate como agradecimento, deduzimos assim que atingimos os nossos objetivos.

Num futuro trabalho pretendemos desenvolver uma base de dados maior e elaborar interrogações e transações mais complexas e diversificadas, no caso do negócio da Dona Ritinha se expandir e ganhar proporções que o exijam. Estamos certos que o tempo, a experiência e as aulas de Base de Dados nos tornarão aptos para isso.

Referências

* Connolly, T.M. and Begg C. E., 2005. Database Systems A Practical Approach to Design, Implementation, and Management. 6th Edition, England: Pearson Education Limited.

**Anexos**

****

Figura 34 - Povoamento da tabela Empresa

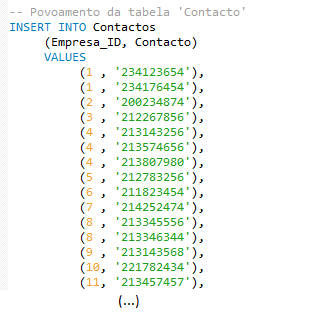
****

Figura 35 - Povoamento da tabela Contactos

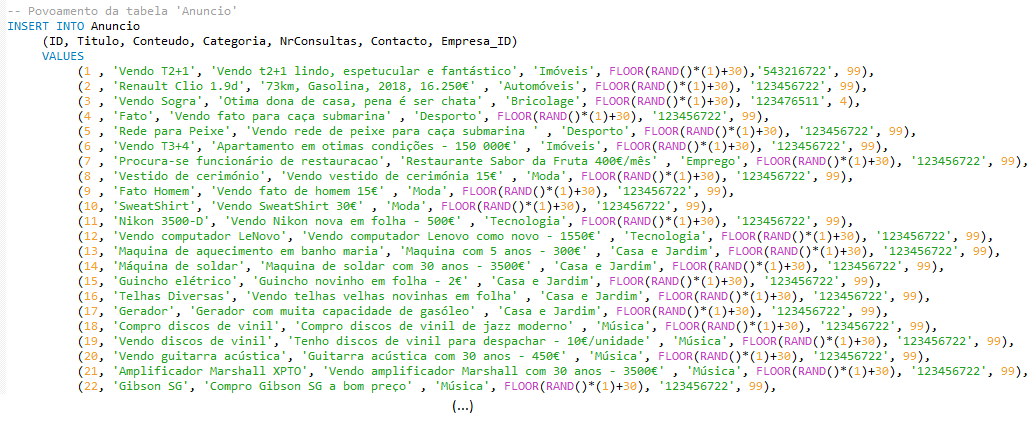
****

Figura 36 - Povoamento da tabela Anúncio

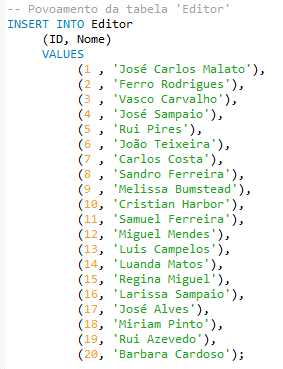
****

Figura 37 - Povoamento da tabela Editor

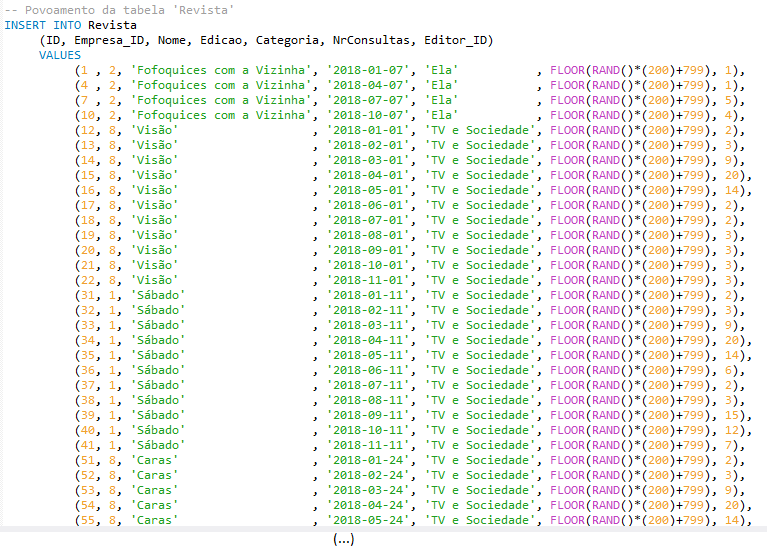
****

Figura 38 - Povoamento da tabela Revista

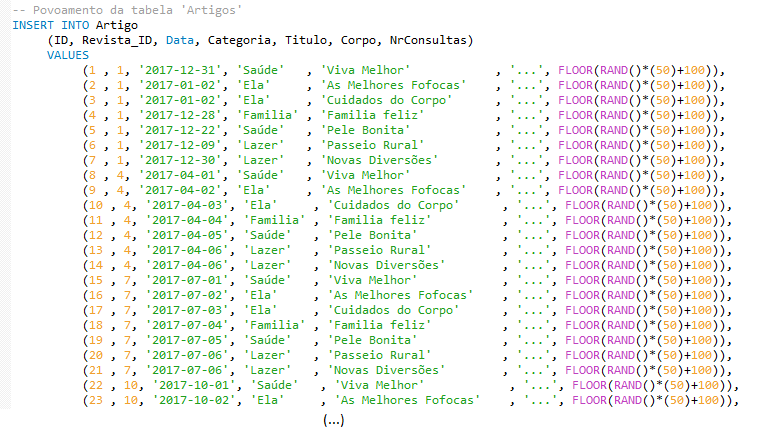
****

Figura 39 - Povoamento da tabela Artigo

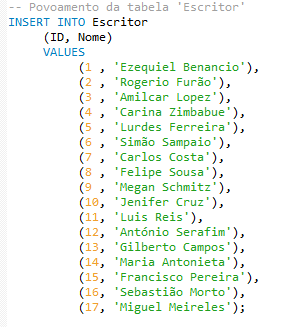
****

Figura 40 - Povoamento da tabela Escritor

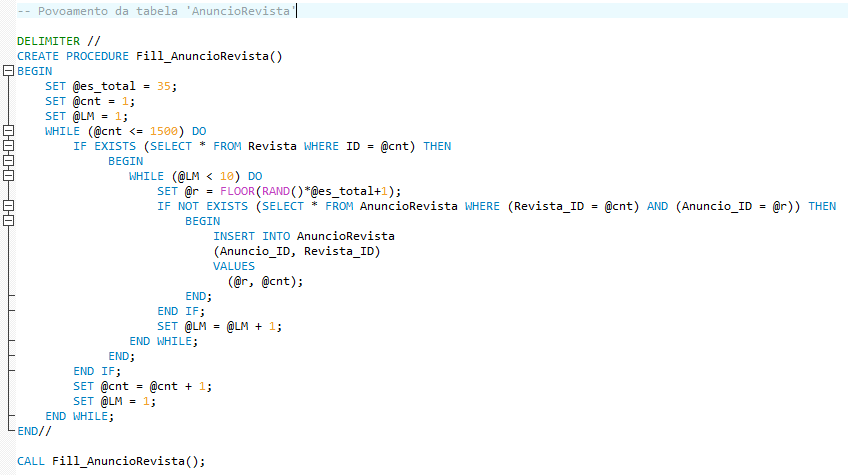
****

Figura 41 - Povoamento da tabela AnuncioRevista

****

Figura 42 - Povoamento da tabela EscritorArtigo

Lista de Siglas e Acrónimos

SQL Structured Query Language

NoSQL Not Only Structured Query Language