

Universidade do Minho Departamento de Informática

Sistema de Gestão de Recomendações Laboratórios De Informática III Grupo 54

25 de Junho, 2021



Mariana Rodrigues (a93229)



Matilde Bravo (a93246)



Pedro Alves (a93272)

Conteúdo

1		Odução 4 Descrição do Problema 4 Análise da Solução 4
2	Mod	
	2.1	Interfaces
		2.1.1 IUser, IBusiness e IReview
		2.1.2 ICatalog
		2.1.3 IStats
	2.2	Queries
		GestReviews
		2.3.1 Catálogo de Users
		2.3.2 Catálogo de Businesses
		2.3.3 Catálogo de Reviews
		2.3.4 Estatísticas
		2.5.1 Estatisticas
3	Apre	esentação 11
		Interfaces
		3.1.1 IView
		3.1.2 IQueryView
		3.1.3 IQueryViewFX
	3.2	Implementação
	5.2	Implementação
4	Con	rolador 13
		4.0.1 Queries
Α	Feta	tísticas de Execução 15
		Load de ficheiros
		Queries
	Λ.Δ	Queries
В	Diag	rama de Classes 18
	B.1	Modelo
		B.1.1 Exceptions
		B.1.2 Queries
		B.1.3 Negócios, utilizadores e reviews
		B.1.4 Catálogo
		B.1.5 StarsTuple
		B.1.6 Stats
		B.1.7 GestReviews
		B.1.8 Outros
	B.2	Apresentação
	2.2	B.2.1 View principal
		B.2.2 Pop up de abrir ficheiro
		B.2.3 Queries
		Controlador 30
	КY	

С	Scre	hots	34
	B.5	sse principal	33
		.2 Tuples	
		.1 Crono	31
	B.4	ls	31
		.2 Callbacks	31
		.1 Queries	31

Capítulo 1

Introdução

O presente relatório apresenta o projeto realizado no âmbito da Unidade Curricular de Laboratórios de Informática III, ao longo da segunda metade do segundo semestre, do segundo ano do Mestrado Integrado em Engenharia Informática da Universidade do Minho.

Este consiste na elaboração de um programa em Java capaz de ler ficheiros CSV contendo informação de *reviews* de negócios na plataforma Yelp, de armazenar esta informação em ficheiros de objetos, e de realizar *queries* nestes ficheiros.

1.1 Descrição do Problema

Numa fase inicial é necessário carregar 3 ficheiros CSV: *users.csv*, *businesses.csv* e *reviews.csv*, contendo, respetivamente, informação sobre utilizadores, negócios e *reviews*. Um dos principais objetivos deste trabalho foi conseguir armazenar esta informação de uma forma que fizesse sentido na linguagem Java, respeitando o modelo de programação orientada a objetos, para facilitar o mais que possível a sua apresentação e manipulação a pedido do utilizador.

1.2 Análise da Solução

O nosso grupo decidiu criar uma aplicação gráfica utilizando a *framework JavaFX*, organizando a aplicação segundo o modelo MVC, isolando os componentes o máximo possível. Graças à linguagem Java, conseguimos cumprir estes objetivos sem sacrificar uso de memória ou *performance* da aplicação.

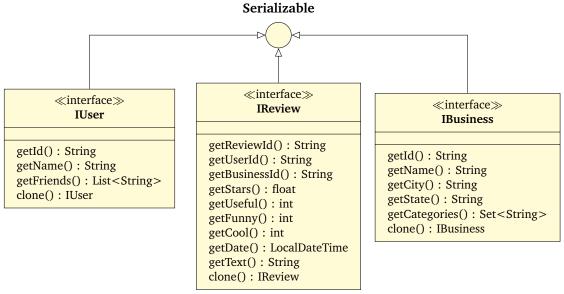
Capítulo 2

Modelo

A camada de módulos contém todas as classes e interfaces utilizadas para modelar os dados dentro da aplicação, os resultados das *queries*, e a lógica de leitura de ficheiros.

2.1 Interfaces

2.1.1 IUser, IBusiness e IReview



Estas três interfaces implementam o necessário para representar os três tipos básicos da aplicação. Qualquer código que lida com dados da aplicação utiliza estas interfaces, em vez de utilizar uma classe diretamente, permitindo a utilização de diferentes implementações. É de notar que todas extendem **Serializable**, necessário para as poder gravar num ficheiro de objetos.

2.1.2 ICatalog

Esta interface define uma estrutura de dados que pode ser utilizada para armazenar informação contida num ficheiro, criando assim um catálogo de um tipo. Particularmente importante é o método **callConstructor**, que chama o construtor do tipo que este catálogo contém com a linha fornecida.

2.1.3 IStats

Esta interface fornece um método para atualizar as estatísticas da aplicação, e é utilizada ao preencher um catálogo a partir de um ficheiro.

2.2 Queries

Os resultados das *queries* implementam todos uma interface vazia, a **IQueryResults**, e especificam dentro de si as estruturas de dados necessárias. Como estas estruturas são todas *wrappers* simples à volta de um tipo, vamos apresentar só os tipos:

- 1. Query1Results List < IBusiness >
- 2. **Query2Results** MyPair<Integer, Integer>
- 3. **Query3Results** List<MyTriple<Integer, Integer, Double»
- 4. **Query4Results** List<MyTriple<Integer, Integer, Double»
- 5. **Query5Results** List<MyPair<String, Integer»¹
- Query6Results Map<Integer, List<MyPair<IBusiness, Integer»>
- Query7Results List<MyTriple<String, String, Integer»
- 8. **Query8Results** List<MyPair<String, Integer»
- Query9Results List<MyPair<String, Double»
- Query10Results Map < String, Map < String, Map < String, StarsTuple» >

≪interface≫ **ICatalog**

getInputFileName() : String

setInputFileName(name: String) : void callConstructor(line: String[]) : T

size(): int

getInvalidCount() : int
getById(id: String) : T
delete(id: String) : void

populateFromFile(stats: Stats, filename: String, catalogoUsers:

CatalogoUsers, catalogoBusinesses: CatalogoBusinesses)



atualiza<T>(newEntity: T, catalogoUsers: CatalogoUsers, catalogoBusinesses: CatalogoBusinesses)

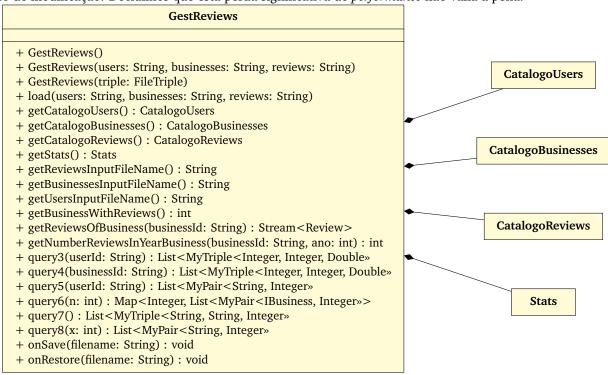
¹Esta classe implementa alguns métodos, ver o diagrama de classes nos anexos para mais detalhes.

2.3 GestReviews

A leitura é feita na classe **GestReviews**, a principal classe do programa. Nesta classe, o método **load** cria três catálogos e utiliza os seus métodos de **populateFromFile** para os preencher.

Esta é também a classe de modelo principal da aplicação. É nesta classe que são implementadas as queries, e é esta que contém todos os dados necessários para a aplicação, através de outras quatro classes que iremos abordar.

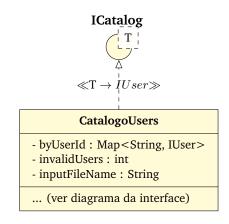
Neste caso em particular, os catálogos não são clonados antes de serem devolvidos - isto porque iria implicar a realocação de vários GB de memória sempre que se quiser aceder, só para evitar o risco de modificação. Decidimos que esta perda significativa de *performance* não valia a pena.



2.3.1 Catálogo de Users

Cada um dos **user**(s) conterá um nome e um id

Como se pode observar o CatalogoUsers conterá um *hash map* em que a chave corresponde ao id de cada utilizador, tendo em atenção que cada utilizador terá um único e exclusivo id. Conterá também um inteiro **invalidUsers**, que servirá para nos informar de quantos utilizadores inválidos foram obtidos ao ler um determinado ficheiro, cujo nome será guardado em **inputFileName**.



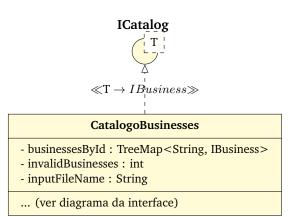
2.3.2 Catálogo de Businesses

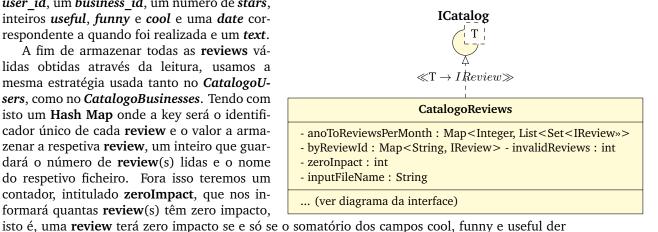
Um **business**, à semelhança de um user, conterá também um identificador único, com isto o catalogoBusiness contém um hash map, onde a sua key será esse identificador e o value o respetivo business, conterá também o número de negócios inválidos lidos e o nome do respetivo ficheiro. Fora isso, face à necessidade de conseguirmos obter a lista ordenada por nome de negócios não avaliados, decidimos armazenar os negócios numa TreeMap. Numa face inicial todos os negócios serão lá colocados, mas à medida que lemos o ficheiro das reviews, todos os businesses avaliados serão removidos dessa mesma àrvore e colocados numa outra estrutura de dados que iremos mostrar posteriormente.

2.3.3 Catálogo de Reviews

Cada review terá um id exclusivo e único, um user id, um business id, um número de stars, inteiros useful, funny e cool e uma date correspondente a quando foi realizada e um text.

A fim de armazenar todas as reviews válidas obtidas através da leitura, usamos a mesma estratégia usada tanto no CatalogoUsers, como no CatalogoBusinesses. Tendo com isto um Hash Map onde a key será o identificador único de cada review e o valor a armazenar a respetiva review, um inteiro que guardará o número de review(s) lidas e o nome do respetivo ficheiro. Fora isso teremos um contador, intitulado zeroImpact, que nos informará quantas review(s) têm zero impacto,





No campo anoToReviewsPerMonth temos um mapa entre um ano e uma lista com um set de

reviews por cada mês do ano.

2.3.4 Estatísticas



Stats

- aaverageReviewByMonth : List<ReviewStarsTuple>
- averageByUserId : Map<String, List<UserStarsTuple»
- averageByBusinessId : Map<String, List<BusinessStarsTuple»
- averageByReviewID : Map<String, List<ReviewStarsTuple»
- averageByStateBusiness: Map<String, Map<String, Map<String, StarsTuple»>
- negociosNuncaAvaliados : TreeMap < String, IBusiness >
- negociosAvaliads : Map < String, IBusiness >
- + Stats()
- + updateStats(review: IReview, business: IBusiness): void
- + updateAveragebyState(review: IReview, business: IBusiness): void
- + updateAverageReviewsByMonth(review: IReview): void
- + updateAverageUser(review: IReview): void
- + updateAverageBusiness(review: IReview): void
- + updateAverageReview(review: IReview): void
- + pairBusinessIdAndTheirReviews(userId: String): List<MyPair<String, Integer»
- + getAllReviews(businessId: String) : Set<String>
- + getNegociosNuncaAvaliadosOrdered(): List<IBusiness>
- + getAverageByStateBusiness(): Map<String, Map<String, Map<String, StarsTuple»>
- + query3(userId: String): List<MyTriple<Integer, Integer, Double»
- + query4(businessId: String): List<MyTriple<Integer, Integer, Double»
- + query7(): List<MyTriple<String, String, Integer»
- + query8(x: int) : List<MyPair<String, Integer»
- + getUsersNaoAvaliados(): int
- + getAverageByMonthReview(): List<MyTriple<Month, Integer, Float»
- + getAverageByMonthReviewAndMore(): List<MyFive<Month, Integer, Float, Integer, Integer»
- + getBusinessWithReviews(): int

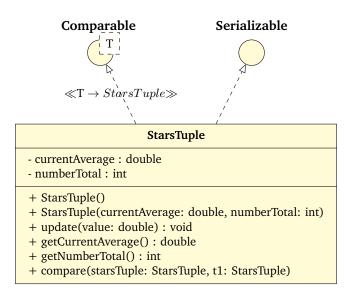
Esta classe encontra-se responsável pelo cálculo e armazenamento das estatísticas dos dados relativos aos três ficheiros lidos. Ao contrário das restantes, não se limita a armazenar os dados obtidos nas estruturas apropriadas, pois à medida que os vai lendo realiza cálculos relativos à média das estrelas. Poder-se-ia ter lido todas as *reviews* e, posteriormente, ter se feito esses mesmos cálculos, mas isso implicaria uma segunda travessia que é evitada mantendo-se um registo da média atual de estrelas com as *reviews* lidas até então, atualizando esse valor sempre que uma linha é lida. Com isto, conseguimos diminuir o tempo de execução sem custos desnecessários.

A fim de armazenar as médias das estrelas lidas em função dos seus identificadores, quer sejam *reviews*, negócios ou utilizadores, e de um determinado mês, decidimos que precisaríamos de 3 maps:

- averageByUserId: Onde a chave será o identificador do *user* e o seu valor será uma lista com objetos da classe UserStarsTuple
- averageByBusinessId: Onde a chave será o identificador do *business* e o seu valor será uma lista com objetos da classe BusinessStarsTuple
- averageByReviewID: Onde a chave será o identificador da *review* e o seu valor será uma lista com objetos da classe ReviewStarsTuple

Visto que existem 12 meses, estabelecemos que essa listas apenas terão 12 elementos, em que cada um dos índices corresponderia a um determinado mês, sendo janeiro o mês 0, fevereiro 1 e por aí adiante, com apenas um objeto da classe respetiva por mês.

Todas as classes encontradas nessas listas são sub-classes de StarsTuple.



É aqui que mantemos a nossa média lida até ao momento. Fora isso, cada uma dessas classes irá armazenar um **set de strings** relativos aos identificadores das reviews.

Um **ReviewStarsTuple** conterá mais dois **sets de strings**, um relativo aos identificadores dos negócios em questão e outro relativo aos utilizadores dessas mesmas *reviews* lidas.

Um **UserStarsTuple** conterá apenas mais um **set de strings** onde armazena todos os identificadores dos negócios relativos as *reviews* desse mesmo utilizador.

Já um **BusinessStarsTuple** conterá apenas mais um **set de strings** onde armazena todos os identificadores dos utilizadores relativos às *reviews* desse mesmo negócio.

Tendo a informação armazenada desta forma, conseguimos com grande eficiência obter pedidos relativos às *queries* 3, 4 e 5.

Face ao desafio proposto na *query* 10, decidiu-se que seria fundamental possuir uma estrutura de dados que fosse de estado para cidades, e de cidade para o seu respetivo **StarTuple**, intitulada **averageByStateBusiness**. Como podemos constatar, temos um Map, onde a sua chave será o nome de um dado estado e o seu valor correspondente será outro Map. Este terá como chave o nome de uma cidade e o seu valor um objeto da classe **StarsTuple**, onde será armazenada a média relativa a essa cidade desse estado correspondente.

Como outrora referido no catálogo de businesses, precisamos de armazenar quais negócios foram ou não analisados. Inicialmente, todos serão colocados num TreeMap, onde a chave corresponderá ao identificador de cada negócio e o seu valor será o negócio em questão. À medida que são lidas *reviews*, verificamos se o identificador do negócio contido nessa *review* encontra-se nessa mesma estrutura; se existir é removido e adiciona-se o mesmo a uma nova estrutura (**Map**), intitulada **negociosAvaliads**. Aqui, a estratégia é a mesma e, portanto, a chave será o identificador do negócio e o seu valor será o negócio em questão.

Constatámos que seria também fundamental armazenarmos a média das estrelas por meses. Face a essa necessidade, conteremos uma lista de **ReviewsStarsTuple**, intitulada **averageReviewByMonth**. Como outrora referido, essa mesma lista terá apenas 12 elementos em que cada um dos índices é atribuído um mês.

Com isto, é-nos possível ter acesso direto a uma grande parte da informação necessária para conseguirmos responder a todos os pedidos do utilizador.

Capítulo 3

Apresentação

A camada de apresentação é constituída pelas interfaces que permitem que o controlador a controle, bem como pelas suas implementações.

3.1 Interfaces

3.1.1 IView

A interface principal da camada de apresentação é a **IView**, definida no *package* **li3.grupo54.Views**. Esta interface possui métodos para apresentar a interface, pedir o nome dos ficheiros, apresentar erros, etc.

É de notar em especial a existência de métodos para definir *callbacks*. Estes *callbacks* são utilizados para facilitar o desenvolvimento de uma aplicação gráfica, pois são mais adequados ao seu modelo de desenvolvimento.

3.1.2 IQueryView

Esta é a interface utilizada para adicionar *queries* a uma classe que implemente **IView**. Uma *view* de uma *query* implementa esta interface e define os métodos **getName** e **getDescription** com o seu nome (por exemplo, "Query 1") e descrição, respetivamente.

O método **showResults** é utilizado para apresentar resultados depois de realizar a *query*. A classe implementadora deve fazer *cast* do objeto **IQueryResults** para o objeto concreto da *query* que lhe diz respeito.

≪interface≫ **IView**

show(): void

getFileTriple(): FileTriple

showError(error: String, description: String): void setOpenCallback(callback: CallbackFileTriple): void

addQuery(query: IQueryView, callback: ExecuteCallback): void

setOnSave(callback: SaveCallback): void

getSaveLocation(): File

disableQueries(disable: boolean): void setOnRestore(SaveCallback callback): void

≪interface≫ **IQueryView**

getName() : String
getDescription() : String

showResults(results: IQueryResults): void

3.1.3 IQueryViewFX

Esta interface é uma extensão da **IQuery-View** que implementa métodos específicos a uma implementação utilizando JavaFX. Se implementássemos uma interface em linha de comandos, haveria uma interface semelhante para esse tipo de interface.

O método mais importante aqui é o **get-ConfigOptionsNode**, pois devolve um conjunto de controlos de JavaFX que são utilizados para apresentar os parâmetros de uma *query*.

Existe também um conjunto de métodos para verificar se o *input* é válido, de modo a ativar o botão de executar apenas quando é possível.

3.2 Implementação

Como foi referido na introdução, a nossa camada de apresentação foi implementada utilizando o **JavaFX**. Esta *framework* permite desenvolver aplicações gráficas de uma forma fácil, e fornece uma forma de definir as nossas interfaces declarativamente através do **FXML**, uma linguagem semelhante a HTML, opcionalmente gerada com a aplicação Scene Builder.

O FXML permite também que, através do decorador **@FXML**, indiquemos que uma variável ou método de instância de uma classe se refere a um elemento com o mesmo id definido no FXML.

A classe principal, que implementa **IView**, é a **DesktopView**, definida no *package* **li3.grupo54.Views**. Ao ser criada, esta classe utiliza o sistema de *resources* do Java para ir buscar o ficheiro **main_ui.fxml**. Quando **show** é chamado, ela itera pelas suas *queries* e adiciona-as ao **queryAccordion**, que foi definido pelo JavaFX quando carregámos o ficheiro.

Um fator importante a notar aqui é que estas *queries* têm o tipo **IQueryViewFX**, não **IQueryView**. Ao serem adicionadas, é verificado se as *IQueryViews* implementam esta interface e, caso não a implementem, são descartadas, pois não há nenhuma maneira de as apresentar.



DesktopView

- stage : Stage
- loader : FXMLLoader
- queries: List<Entry<IQueryViewFX, ExecuteCallback»
- queryAccordion : Accordion
- resultsTabs : TabPane
- nQueries: int
- callback : CallbackFileTriplesaveCallback : SaveCallbackrestoreCallback : SaveCallback
- + DesktopView(s: Stage)
- + addQuery(view: IQueryView, callback: ExecuteCallback): void
- + disableQueries(disable: boolean): void
- + getFileTriple() : FileTriple
- + getSaveLocation(): File
- + setOnRestore(callback: SaveCallback): void
- $+\ setOnSave(callback:\ SaveCallback):\ void$
- + setOpenCallback(callback: CallbackFileTriple): void
- + show(): void
- + showError(error: String, description: String): void
- onRestore() : voidopenFile() : void
- onSave(): void

Capítulo 4

Controlador

A camada de controlador faz a união entre as camadas de modelo e apresentação. Por causa disto, o controlador acaba por ser maioritariamente composto por código de "cola", que não é muito complexo, mas que é fundamental para o funcionamento da aplicação.

Nós decidimos implementar o controlador através de um controlador principal, a classe **Controller**, que implementa uma interface **IController**, que define a API externa de um controlador.

Como o controlador implementa a lógica da aplicação, vemos que esta API é muito simples. Apenas necessita de uma forma de adicionar uma *query* (do qual iremos falar mais à frente) e de um método para executar a aplicação.

Podemos ver que o controlador implementa alguns métodos privados. Estes são utilizados como *callbacks* para ações da camada de apresentação - por exemplo, o método **loadTriple** é executado quando se carrega no botão para abrir um conjunto de ficheiros CSV.

4.0.1 Queries

As *queries* são definidas através da interface **IQueryController**. Como podemos ver pelo diagrama, esta é, também, uma interface bastante simples, servindo para fazer a conexão entre a *view* da *query* e o seu código, implementado no modelo.

Tomando como exemplo a *query* 1, vemos que é muito fácil definir um destes controladores:

≪interface≫ IController + addQuery(controller: IQueryController): void + executa(): void

Controller

- model : GestReviews
- view: IView
- queries : List<IQueryController>
- + Controller(model: GestReviews, view: IView)
- $+\ add Query (controller:\ IQuery Controller):\ void$
- + executa(): void
- saveObjectFile(file: File): void
- loadTriple(triple: FileTriple): void
- readObjectFile(file: File): void

≪interface≫ IQueryController

- + getView(): IQueryView
- + execute(): IQueryResults
- + executeAndShow(): void

```
public class Query1 implements IQueryController {
   private GestReviews model;
   private Query1View view;

public Query1(GestReviews model, Query1View view) {
     this.model = model;
     this.view = view;
}
```

```
@Override
public IQueryView getView() {
    return view;
}

@Override
public IQueryResults execute() {
    return new Query1Results(model.getStats().getNegociosNuncaAvaliadosOrdered());
}
}

Com estas classes conseguimos então criar uma API muito fácil de utilizar, tornando a inicialização da aplicação muito simples:

// Este método start é o equivalente ao método main de uma aplicação de JavaFX
```

```
// Este método start é o equivalente ao método main de uma aplicação de JavaFX
public void start(Stage s) {
    DesktopView view = new DesktopView(s);
    GestReviews model = new GestReviews();
    IController controller = new Controller(model, view);

    controller.addQuery(new Query1(model, new Query1View()));
    // ... mais queries
    controller.addQuery(new Stats(model, new StatsView()));
    controller.executa();
}
```

Apêndice A

Estatísticas de Execução

A.1 Load de ficheiros

Ficheiros: Users (2.7 G), Businesses (17 M), Reviews (700 M)

Comparação de Tempos	Texec (s)	Memória (GB)
Load Friends	9.097	1.80
Load Business	0.902	0.03
Load Reviews	14.433	1.07
Total	24.432	2.90

A.2 Queries

	Argumentos recebidos:		Texec (s)	Núme	ero de
	GestReviews e			dados obtidos	
	Apenas GestReviews		2.936E-6		
Query 1			1.006E-6	132515	
			1.207E-6		
		1	3.172E-6	Reviews: 7578	Business:6944
	Ano 2020 Mês:	5	5.6E-7	Reviews: 3375	Business: 3133
		9	3.16E-7	Reviews: 5194	Business: 4837
	Ano 2012 Mês:	1	8.72E-7	Reviews: 4415	Business: 3823
Query 2		6	4.67E-7	Reviews: 4524	Business: 4837
		11	7.03E-7	Reviews: 3986	Business: 3432
		3	6.05E-7	Reviews: 2357	Business: 1743
	Ano 2009 Mês:	7	4.92E-7	Reviews: 1747	Business: 1441
		12	6.95E-7	Reviews: 1544	Business: 1283
	Ano 1984 Mês:	1	0	Date no	ot found

		"Invalido" "1RCRKuHgP3FskGUVnmFdxg" "J9m48CKaTlXvkqFMs7L6Kg" "I4cfBoRWEhB90HBdjh8z5Q" "ak0TdVmGKo4pwqdJSTLwWw" "6iYb2HFDywm3zjuRg0shjw" "t35jsh9YnMtttm69UCp7gw" "Lye66-VKsO-npfyo3o7qeg"	0 1.458E-6 4.47E-7 7.12E-7 0 2.775E-6 4.6E-6	(Não existe) 4 13 59 (Não existe) 83	
Query 4 Bussiness id Query 5 User id Query 6 N negócios Query 7 Apenas GestRet Query 8 N users "6iYb2HFDywr		"J9m48CKaTlXvkqFMs7L6Kg" "I4cfBoRWEhB90HBdjh8z5Q" "ak0TdVmGKo4pwqdJSTLwWw" "6iYb2HFDywm3zjuRg0shjw" "t35jsh9YnMtttm69UCp7gw" "Lye66-VKsO-npfyo3o7qeg"	4.47E-7 7.12E-7 0 2.775E-6 4.6E-6	4 13 59 (Não existe)	
Query 4 Bussiness id Query 5 User id Query 6 N negócios Query 7 Apenas GestRet Query 8 N users "6iYb2HFDywri		"J9m48CKaTlXvkqFMs7L6Kg" "I4cfBoRWEhB90HBdjh8z5Q" "ak0TdVmGKo4pwqdJSTLwWw" "6iYb2HFDywm3zjuRg0shjw" "t35jsh9YnMtttm69UCp7gw" "Lye66-VKsO-npfyo3o7qeg"	7.12E-7 0 2.775E-6 4.6E-6	59 (Não existe)	
Query 5 User id Query 6 N negócios Query 7 Apenas GestRet Query 8 N users "6iYb2HFDywr		"I4cfBoRWEhB90HBdjh8z5Q" "ak0TdVmGKo4pwqdJSTLwWw" "6iYb2HFDywm3zjuRg0shjw" "t35jsh9YnMtttm69UCp7gw" "Lye66-VKsO-npfyo3o7qeg"	0 2.775E-6 4.6E-6	(Não existe)	
Query 5 User id Query 6 N negócios Query 7 Apenas GestRet Query 8 N users "6iYb2HFDywr		"6iYb2HFDywm3zjuRg0shjw" "t35jsh9YnMtttm69UCp7gw" "Lye66-VKsO-npfyo3o7qeg"	2.775E-6 4.6E-6		
Query 5 User id Query 6 N negócios Query 7 Apenas GestRet Query 8 N users "6iYb2HFDywr		"6iYb2HFDywm3zjuRg0shjw" "t35jsh9YnMtttm69UCp7gw" "Lye66-VKsO-npfyo3o7qeg"	4.6E-6		
Query 5 User id Query 6 N negócios Query 7 Apenas GestRet Query 8 N users "6iYb2HFDywr		"t35jsh9YnMtttm69UCp7gw" "Lye66-VKsO-npfyo3o7qeg"			
Query 6 N negócios Query 7 Apenas GestRet Query 8 N users "6iYb2HFDywr		"Lye66-VKsO-npfyo3o7qeg"	0 (0 (7	
Query 6 N negócios Query 7 Apenas GestRet Query 8 N users "6iYb2HFDywr		1	3.6E-6	26	
Query 6 N negócios Query 7 Apenas GestRet Query 8 N users "6iYb2HFDywr		i invaliuo !	0	(Não existe)	
Query 6 N negócios Query 7 Apenas GestRet Query 8 N users "6iYb2HFDywr		"1RCRKuHgP3FskGUVnmFdxg"	6.57E-7	4	
Query 7 Apenas GestRet Query 8 N users "6iYb2HFDywr		"J9m48CKaTlXvkqFMs7L6Kg"	9.26E-7	11	
Query 7 Apenas GestRet Query 8 N users "6iYb2HFDywr		"I4cfBoRWEhB90HBdjh8z5Q"	9.09E-7	57	
Query 7 Apenas GestRet Query 8 N users "6iYb2HFDywr		1	5.76E-7	18	
Query 7 Apenas GestRet Query 8 N users "6iYb2HFDywr		43	3.76E-7 3.86E-7	742	
Query 7 Apenas GestRet Query 8 N users "6iYb2HFDywr		123	4.28E-7	2102	
Query 7 Apenas GestRet Query 8 N users "6iYb2HFDywr		450	4.26E-7 4.97E-7	7661	
Query 8 N users "6iYb2HFDywr		1000	4.97E-7 5.06E-7	16526	
Query 8 N users "6iYb2HFDywr		1500	5.00E-7 5.35E-7	24237	
Query 8 N users "6iYb2HFDywr		4000	5.627E-7	58409	
Query 8 N users "6iYb2HFDywr		4000		30409	
Query 8 N users "6iYb2HFDywr			2.089E-6	900	
"6iYb2HFDywr	views		1.0396E-6		
"6iYb2HFDywr			1.086-6		
"6iYb2HFDywr		1	2.014E-6	1	
"6iYb2HFDywr		43	9.22E-7	43	
"6iYb2HFDywr		1000	1.254E-7	1000	
"6iYb2HFDywr		1500	8.41E-7	1500	
	N users	4000	8.85E-7	4000	
		10000	1.021E-7	10000	
		100000	9.1E-7	100000	
		600000	8.02E-7	536836	
		5	7.07E-7	5	
		23	8.06E-7	23	
	m3zjuRg0shjw Top:	50	3.74E-7	50	
Query 9 "t35jsh9YnMtt	J J 5 - F	75	3.52E-7	75	
Query 9 "t35jsh9YnMtt		100	2.3E-7	83	
Query 9 "t35jsh9YnMttt		1	0.002383	1	
Query 9 "t35jsh9YnMtt		1	1.58E-7	1	
	"t35jsh9YnMtttm69UCp7gw Top:	3	4.17E-7	3	
1		7	3.08E-7	7	
		10	2.47E-7	10	
		10	0.001157	10	
		5	6.18E-7	5	
"Lye66-VKsO-1		·	3.76E-7	10	
	npfyo3o7qeg Top:	10		1 -	
	npfyo3o7qeg Top:	10 15	5.4E-7	15	
	npfyo3o7qeg Top:		5.4E-7 7.33E-7	25	
Query 10 Apenas GestRe	npfyo3o7qeg Top:	15	7.33E-7		
Tiponus destrict		15			

Todos os resultados aqui apresentados foram obtidos através de um portátil com as seguintes especificações:

OS	Arch Linux x86_64
Host	20TACTO1WW ThinkPad E14 Gen 2
CPU	11th Gen Intel i7-1165G7 (8) @ 4.700GHz
Memory	SoDIMM de 16 GB DDR4 3.200 MHz
First Solid State Drive	512GB SSD NVMe

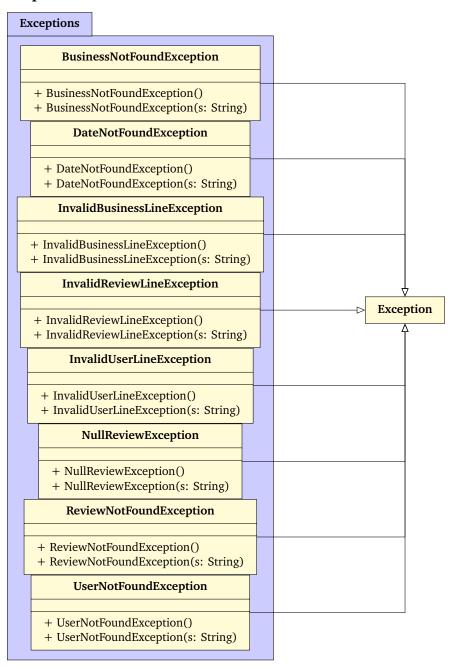
Apêndice B

Diagrama de Classes

De modo a caber nas páginas do formulário, o diagrama será apresentado por secções relacionadas.

B.1 Modelo

B.1.1 Exceptions



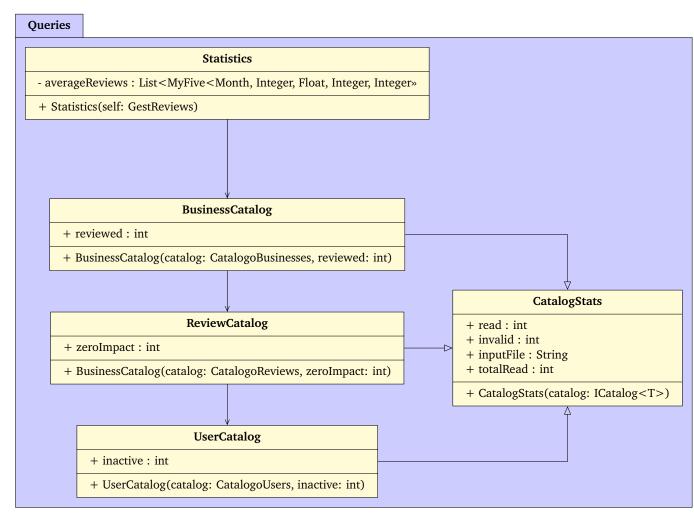
B.1.2 Queries

Queries Query1Results - businesses : List < IBusiness > + Query1Results(businesses: List<IBusiness>) + getBusinesses(): List<IBusiness> + setBusinesses(businesses: List<IBusiness>): void Query2Results - result : MyPair < Integer, Integer > + Query2Results(result: MyPair<Integer,Integer>) + getResults(): MyPair<Integer, Integer> Query3Results - result : List<MyTriple<Integer, Integer, Double» + Query3Results(result: List<MyTriple<Integer, Integer, Double») + getResults(): List<MyTriple<Integer, Integer, Double» \ll interface \gg **IQueryResults** Query4Results - result : List<MyTriple<Integer, Integer, Double» + Query4Results(result: List<MyTriple<Integer, Integer, Double») + getResults(): List<MyTriple<Integer, Integer, Double» Query5Results - result : List<MyPair<String, Integer» + Query5Results(result: List<MyPair<String, Integer») + getResults(): List<MyPair<String, Integer» + allNameReviews(): List<String> + getAllNumber() : List<String> Query6Results - businesses: Map<Integer, List<MyPair<IBusiness, Integer»> + Query6Results(businesses: Map<Integer, List<MyPair<IBusiness, Integer»>) + getBusinesses(): Map<Integer, List<MyPair<IBusiness, Integer»> + setBusinesses(businesses: Map<Integer, List<MyPair<IBusiness, Integer»>): void

(Continua na próxima página)

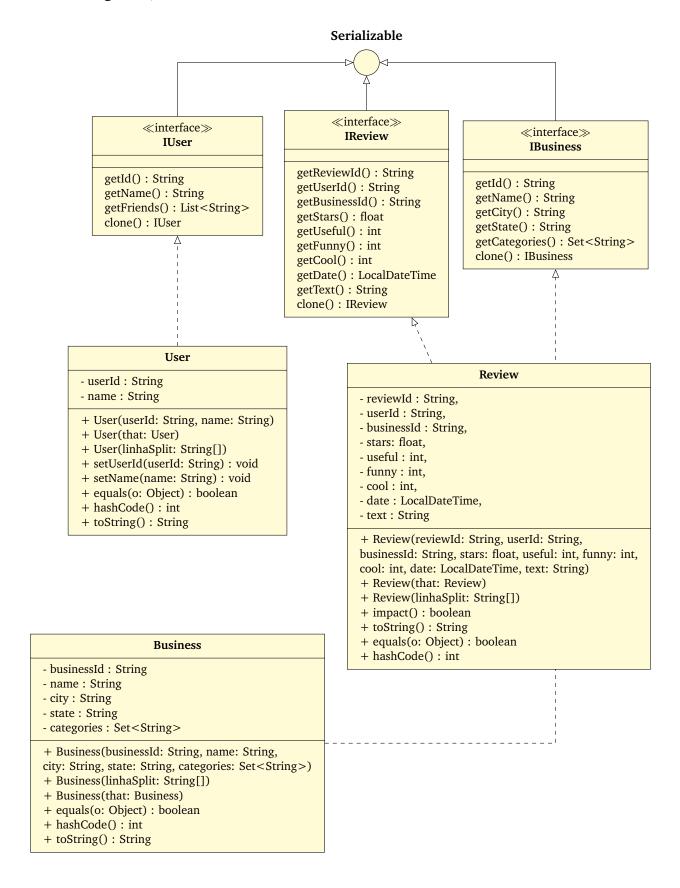
Queries Query7Results - result : List<MyTriple<String, String, Integer» + Query7Results(results: List<MyTriple<String, String, Integer») + getResults(): List<MyTriple<String, String, Integer» + setBusinesses(businesses: List<MyTriple<String, String, Integer»): void Query8Results - result : List<MyPair<String, Integer» + Query8Results(result: List<MyPair<String, Integer») + getResults(): List<MyPair<String, Integer» \ll interface \gg **IQueryResults** Query9Results - result : List < MyPair < String, Double» + Query3Results(result: List<MyPair<String, Double») + getResults(): List<MyPair<String, Double» Query10Results - averageByStateBusiness : Map<String, Map<String, Map<String, StarsTuple»> + Query10Results(averageByStateBusiness: Map<String, Map<String, Map<String, StarsTuple»>) + getResults(): Stream<Result> Result - state : String - city : String - businessId : String - averageRating: double - Result(state: String, city: String, businessId: String, averageRating: double) + getState() : String + getCity() : String + getBusinessId() : String + getAverageRating() : double StatsResults + StatsResults(results: GestReviews) + getResults() : Statistics - results **Statistics**

(Continua na próxima página)

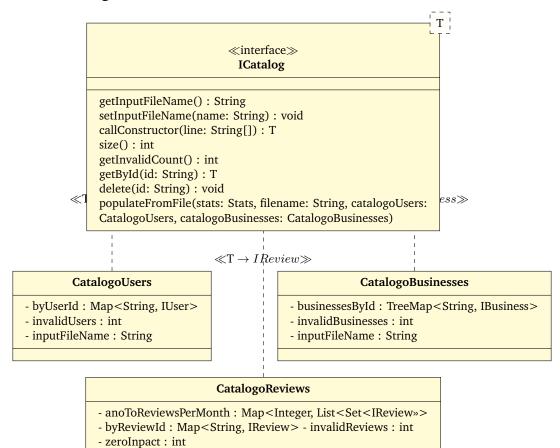


Nota: Os campos são públicos pois são final, e portanto impossíveis de modificar.

B.1.3 Negócios, utilizadores e reviews

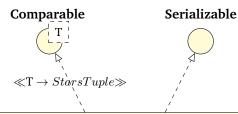


B.1.4 Catálogo



- inputFileName : String

B.1.5 StarsTuple



StarsTuple

- currentAverage : double
- $\hbox{- number Total}: int$
- + StarsTuple()
- + StarsTuple(currentAverage: double, numberTotal: int)
- + update(value: double) : void
- + getCurrentAverage() : double
- + getNumberTotal(): int
- + compare(starsTuple: StarsTuple, t1: StarsTuple)

UserStarsTuple

- reviews : Set<String>business : Set<String>
- + UserStarsTuple()
- + UserStarsTuple(review: Review)
- + updateAverage(review: Review)
- + getReviewsNumber(): int
- + getBusinessNumberDistinct() : int
- + getReviews() : Set < String >
- + getBusiness() : Set<String>

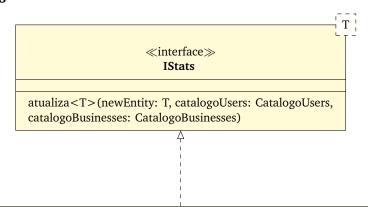
BusinessStarsTuple

- reviews : Set < String>
- users : Set<String>
- + BusinessStarsTuple()
- + BusinessStarsTuple(review: Review)
- + updateAverage(review: Review): void
- $+ \ getReviewsNumber(): int$
- + getUsersNumber() : int
- + getReviews(): Set<String>

ReviewStarsTuple

- reviews : Set<String>users : Set<String>
- business : Set<String>
- + ReviewStarsTuple()
- + ReviewStarsTuple(review: Review)
- + updateAverage(review: Review)
- + getReviewsNumber(): int
- + getBusinessDistincts(): int
- + getUserDistincts(): int
- + getReviews() : Set<String>

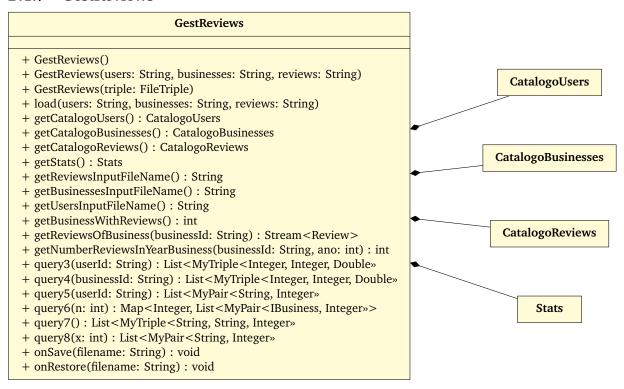
B.1.6 Stats



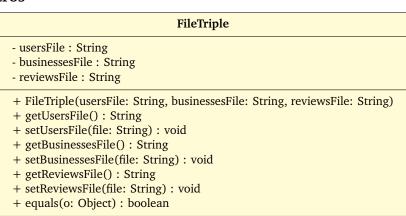
Stats

- aaverageReviewByMonth : List<ReviewStarsTuple>
- averageByUserId: Map<String, List<UserStarsTuple»
- averageByBusinessId: Map<String, List<BusinessStarsTuple»
- averageByReviewID : Map < String, List < ReviewStarsTuple»
- averageByStateBusiness: Map<String, Map<String, Map<String, StarsTuple»>
- negociosNuncaAvaliados : TreeMap < String, IBusiness >
- negociosAvaliads : Map < String, IBusiness >
- + Stats()
- + updateStats(review: IReview, business: IBusiness): void
- + updateAveragebyState(review: IReview, business: IBusiness): void
- + updateAverageReviewsByMonth(review: IReview): void
- + updateAverageUser(review: IReview): void
- + updateAverageBusiness(review: IReview): void
- + updateAverageReview(review: IReview): void
- $+\ pairBusiness Id And Their Reviews (user Id:\ String): List < My Pair < String,\ Integer > 1000 and 1000 an$
- + getAllReviews(businessId: String) : Set<String>
- + getNegociosNuncaAvaliadosOrdered(): List<IBusiness>
- $+ \ getAverageByStateBusiness(): Map < String, Map < String, Map < String, StarsTuple >>$
- + query3(userId: String): List<MyTriple<Integer, Integer, Double»
- + query4(businessId: String) : List<MyTriple<Integer, Integer, Double»
- + query7(): List<MyTriple<String, String, Integer»
- + query8(x: int) : List<MyPair<String, Integer»
- $+\ getUsersNaoAvaliados():int$
- $+ \ getAverageByMonthReview(): List < MyTriple < Month, \ Integer, \ Float >>$
- + getAverageByMonthReviewAndMore(): List<MyFive<Month, Integer, Float, Integer, Integer»
- + getBusinessWithReviews(): int

B.1.7 GestReviews



B.1.8 Outros



Leitura
+ determinaDelimiter(line: String) : char

B.2 Apresentação

B.2.1 View principal

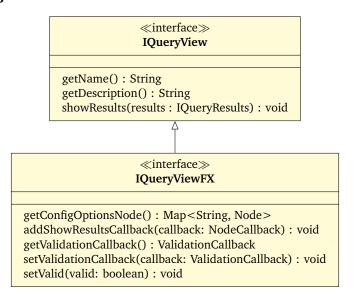
≪interface≫ **IView** show(): void getFileTriple(): FileTriple showError(error: String, description: String): void setOpenCallback(callback: CallbackFileTriple): void addQuery(query: IQueryView, callback: ExecuteCallback): void setOnSave(callback: SaveCallback): void getSaveLocation(): File disableQueries(disable: boolean): void setOnRestore(SaveCallback callback): void DesktopView - stage : Stage - loader : FXMLLoader - queries: List<Entry<IQueryViewFX, ExecuteCallback» - queryAccordion : Accordion - resultsTabs : TabPane - nQueries: int – callback : CallbackFileTriple - saveCallback : SaveCallback - restoreCallback : SaveCallback + DesktopView(s: Stage) + addQuery(view: IQueryView, callback: ExecuteCallback): void + disableQueries(disable: boolean): void + getFileTriple(): FileTriple + getSaveLocation(): File + setOnRestore(callback: SaveCallback) : void + setOnSave(callback: SaveCallback): void + setOpenCallback(callback: CallbackFileTriple): void + show(): void + showError(error: String, description: String): void - onRestore(): void

- openFile() : void
- onSave() : void

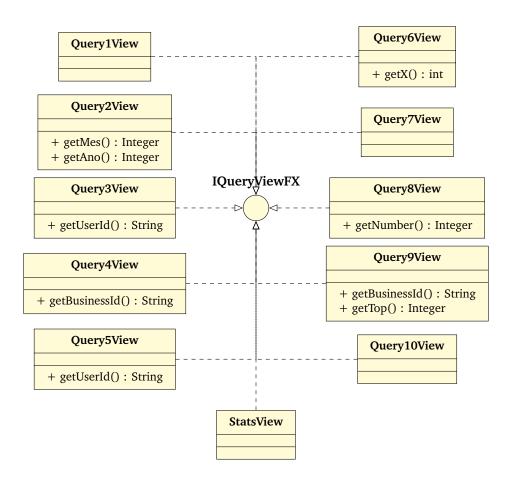
B.2.2 Pop up de abrir ficheiro

OpenFileDialog - loader : FXMLLoader - stage : Stage – users : File – businesses : File - reviews : File - directory: File - ok : boolean - btnUsers: Button btnBusinesses : Button – btnReviews : Button - btnOpen: Button - btnDefault : Button + OpenFileDialog(parent: Stage) + showAndWait() : FileTriple - showDialog(title: String): File - updateUI(): void - pickUsers(): void - pickBusinesses() : void - pickReviews() : void - open(): void - defaultBt(): void

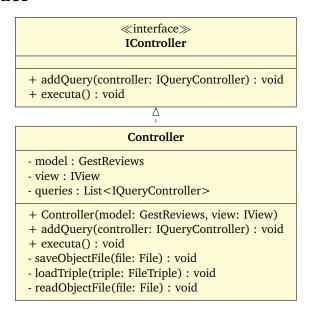
B.2.3 Queries



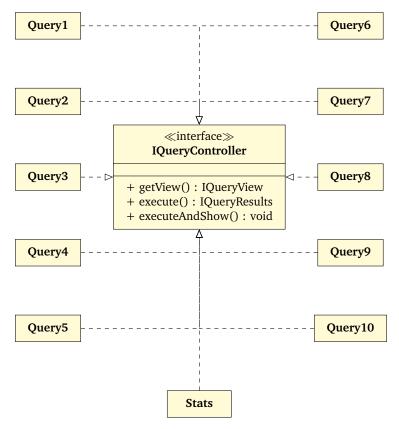
Para simplificar os diagramas, iremos apenas apresentar os getters/setters de campos específicos às views, caso aplicável.



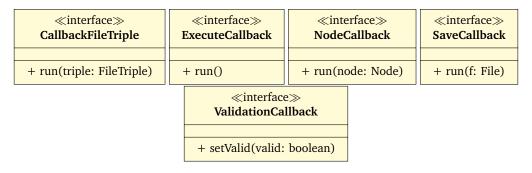
B.3 Controlador



B.3.1 Queries

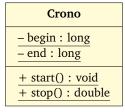


B.3.2 Callbacks

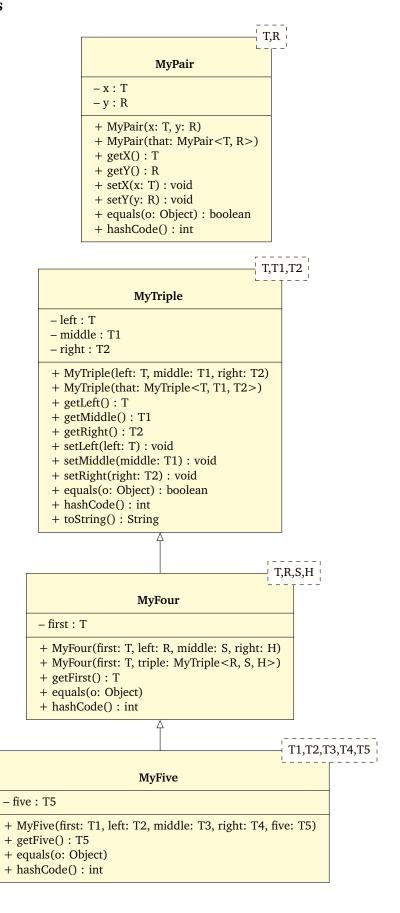


B.4 Utils

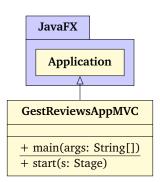
B.4.1 Crono



B.4.2 Tuples



B.5 Classe principal



Apêndice C

Screenshots

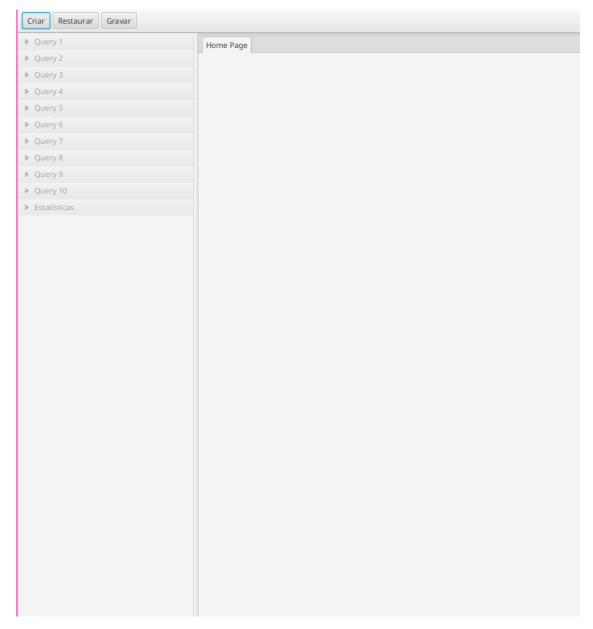


Figura C.1: Estado inicial da aplicação, antes de abrir algo



Figura C.2: Menu para abrir ficheiros CSV

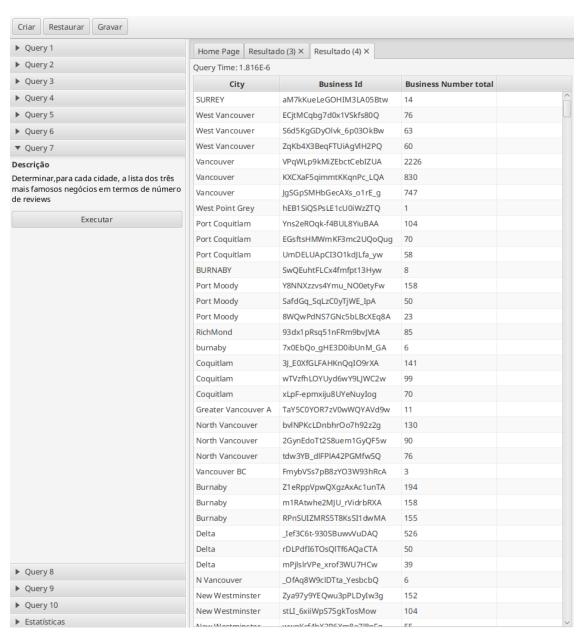


Figura C.3: Resultado de uma query