

UNIVERSIDADE DO MINHO

MESTRADO EM ENGENHARIA INFORMÁTICA



Implementação de Sistema de Business Intelligence para o suporte à decisão

Análise de Dados

Maria Silva, A83830
Mariana Marques, A85171
Miguel Solino, A86435
Paulo Lima, A89893
Grupo 6

Janeiro, 2021

Conteúdo

1	Introdução	4
2	Contextualização das Fontes de Informação	5
3	Base de dados	6
3.1	Dimensões	7
3.1.1	Dimensão <i>User</i>	7
3.1.2	Dimensão <i>Place</i>	14
3.1.3	Dimensão <i>Rating</i>	18
3.2	Factos	19
3.3	Esquematização do Esquema Dimensional	20
4	Tratamento dos dados	20
5	Povoamento do Data Warehousing	21
6	Criação dos Indicadores	26
6.1	Indicadores por contagem de id_users	26
6.2	Indicadores por contagem de id_place	34
6.3	Indicadores por contagem de id_fact_rating	40
7	Conclusões	42

Lista de Figuras

1	Dimensão User	7
2	Dimensão Dress	8
3	Dimensão Kids	8
4	Dimensão Interest	8
5	Dimensão Drink	9
6	Dimensão Year	9
7	Dimensão Marriage	9
8	Dimensão Personality	10
9	Dimensão Color	10
10	Dimensão Religion	10
11	Dimensão Activity	11
12	Dimensão Transport	11
13	Dimensão Weigth	11
14	Dimensão Height	12
15	Dimensão Smoker	12
16	Dimensão Ambience	12
17	Dimensão Budget	13
18	Dimensão Coordinates	13
19	Dimensão Place	14
20	Dimensão City	14
21	Dimensão State	15
22	Dimensão Name Restaurant	15
23	Dimensão Address	15
24	Dimensão Alcohol	16
25	Dimensão Parking Lot	16
26	Dimensão Dress Code	16
27	Dimensão Services	17
28	Dimensão Smoking Area	17
29	Dimensão Accessibility	17
30	Dimensão Area	18
31	Dimensão Rating	18
32	Dimensão Rating	19
33	Esquema do Data Warehouse	20
34	Esquema auxiliar usado para facilitar o povoamento	21
35	Funcionalidade de import do MySQL Workbench	22
36	Método auxiliar usado para facilitar o povoamento	23
37	Método auxiliar usado para facilitar o povoamento	24
38	Método auxiliar usado para facilitar o povoamento	25
39	Quantidade de Users por preferência de cor	26
40	Quantidade de Users por nível de bebida	27
41	Quantidade de Users por preferência de vestuário	27
42	Quantidade de Users por personalidade	28

43	Quantidade de Users por religião	28
44	Quantidade de Users por fumador	29
45	Quantidade de Users por situação familiar	29
46	Quantidade de Users por tipo de transporte utilizado	30
47	Quantidade de Users por peso	30
48	Quantidade de Users por ano de nascimento	31
49	Quantidade de Users por ambiente que prefere na ida a restaurantes	31
50	Quantidade de Users por alturas do indivíduo	32
51	Quantidade de Users por estado civil do individuo	32
52	Quantidade de Users por atividade	33
53	Quantidade de Users por budget	33
54	Quantidade de restaurantes por nome	34
55	Quantidade de restaurantes por tipo de álcool servido	34
56	Quantidade de restaurantes por acessibilidade	35
57	Quantidade de restaurantes por área de fumadores	36
58	Quantidade de restaurantes por código de vestuário	37
59	Quantidade de restaurantes por Código de vestuário por prestação de outros serviços	37
60	Quantidade de restaurantes por tipo de parque de estacionamento	38
61	Quantidade de restaurantes por tipo de espaço	38
62	Quantidade de restaurantes por budget	39
63	Quantidade de cada rating do serviço dos restaurantes	40
64	Quantidade de cada rating da comida dos restaurantes	40
65	Quantidade de cada rating geral dos restaurantes	41

1 Introdução

Este trabalho foi realizado no âmbito do Trabalho Prático da Unidade Curricular de Análise de Dados. Tendo este como objetivo a criação de um sistema de Data Warehousing, bem como um sistema de Business Intelligence para o suporte à decisão.

Para desenvolvimento do projeto proposto é necessário realizar um trabalho de análise, planeamento, e implementação tendo como base o dataset escolhido (Restaurant Data with Consumer Ratings). Deve ser desenvolvido primeiramente o modelo que suporte o dataset escolhido, bem como o respetivo povoamento da base de dados e ainda estruturas necessárias à sua atualização de forma incremental e/ou diferencial. O sistema de Business Intelligence poderá ser desenvolvido na plataforma Microsoft Power BI Desktop ou Tableau Desktop, no entanto o nosso grupo optou por escolher a plataforma Power BI Desktop.

2 Contextualização das Fontes de Informação

Para este projeto vamos apenas usar uma fonte de dados, o "Restaurant Data with Consumer Ratings".

Este conjunto de dados foi usado para um estudo em que a tarefa era gerar uma lista dos melhores restaurantes de acordo com as preferências do consumidor e encontrar as características significativas.

Este pode ser usado para criar um sistema de recomendação ou determinar que restaurantes uma pessoa terá mais probabilidade de visitar, com base nos seus gostos e situações de vida.

3 Base de dados

Foi utilizado o MySQL para a criação da base de dados do sistema. Foi criada uma estrutura dimensional que armazena as informações recolhidas a partir dos ficheiros disponibilizados para o trabalho.

De forma a implementar o nosso *Data Warehouse* corretamente, tivemos de o planejar e criar o nosso modelo Dimensional.

É ainda de notar que para a realização deste esquema dimensional, optámos por retirar alguns atributos do dataset que, na nossa opinião, não teriam grande importância para o estudo em causa e que mais tarde na parte da criação dos indicadores não teriam grande importância informativa.

Estes atributos são: dias e horas úteis, métodos de pagamento e tipos de cuisine

3.1 Dimensões

O nosso modelo dimensional apresenta tabelas de dimensão, que apresentaremos de seguida.

Como o grupo optou por desenvolver um diagrama sob a forma de floco de neve cada tabela de dimensão apresenta ainda mais tabelas dimensão.

3.1.1 Dimensão *User*

Nesta dimensão é apresentado através de chaves estrangeiras todas as características do *user* da base de dados.

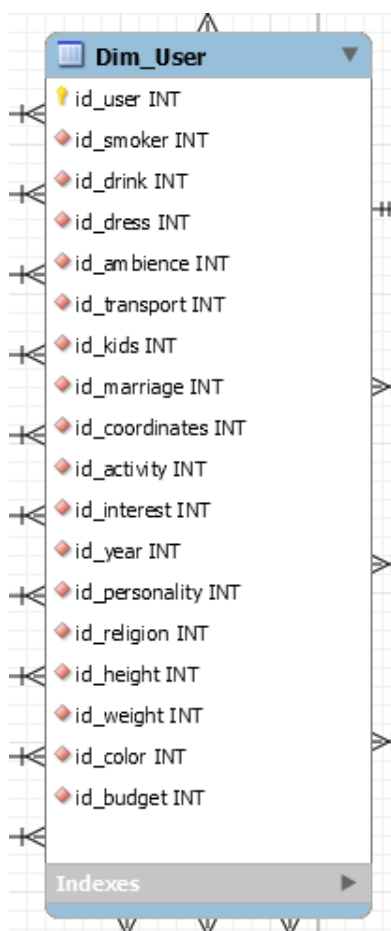


Figura 1: Dimensão User

- **Dimensão *Dress***

Esta dimensão indica a preferência do *user* em termos de vestuário, indicando se este prefere vestir-se formalmente, informalmente ou se simplesmente nem tem preferência.

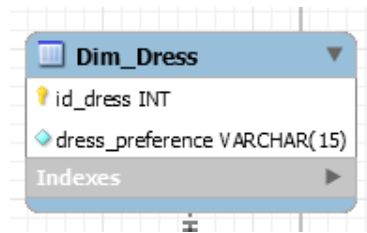


Figura 2: Dimensão Dress

- **Dimensão *Kids***

Esta dimensão pretende indicar se um *user* tem filhos, se é dependente ou independente.

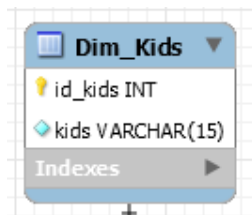


Figura 3: Dimensão Kids

- **Dimensão *Interest***

Esta dimensão apresenta possíveis interesses que os *users* têm.

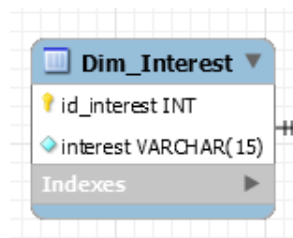


Figura 4: Dimensão Interest

- **Dimensão *Drink***

Esta dimensão apresenta a regularidade com que o *user* bebe álcool.

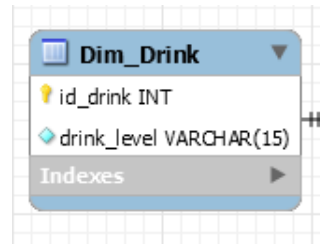


Figura 5: Dimensão Drink

- **Dimensão *Year***

Esta dimensão apresenta o ano de nascimento do *user*.

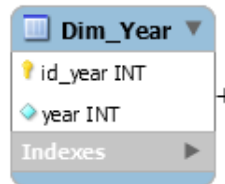


Figura 6: Dimensão Year

- **Dimensão *Marriage***

Esta dimensão indica o estado matrimonial do *user*, ou seja, se este é casado, se já foi ou se é solteiro.

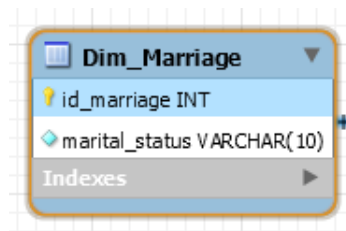


Figura 7: Dimensão Marriage

- **Dimensão *Personality***

Esta dimensão apresenta dentro de algumas possibilidades, como sendo estas "hard-worker", "thrifty-protector", "hunter-ostentations", "conformist" os traços de personalidade de cada *user*.

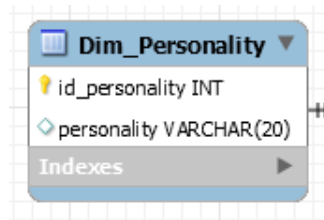


Figura 8: Dimensão Personality

- **Dimensão *Color***

Esta dimensão apresenta a cor favorita de cada *user*.



Figura 9: Dimensão Color

- **Dimensão *Religion***

Esta dimensão indica a religião de cada *user*.

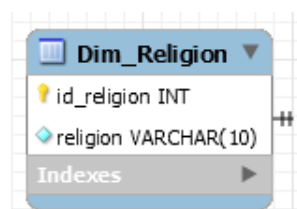


Figura 10: Dimensão Religion

- **Dimensão *Activity***

Esta dimensão apresenta se o *user* tem atividade profissional, se é estudante ou desempregado.

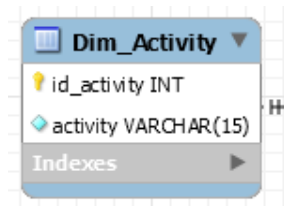


Figura 11: Dimensão Activity

- **Dimensão *Transport***

Esta dimensão apresenta se o *user* se desloca de carro, transportes públicos ou a pé.

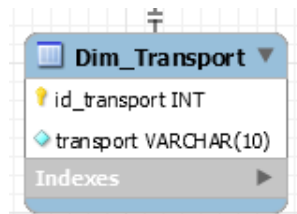


Figura 12: Dimensão Transport

- **Dimensão *Weight***

Esta dimensão apresenta o peso de cada indivíduo.

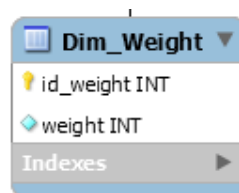


Figura 13: Dimensão Weight

- **Dimensão *Height***

Esta dimensão apresenta a altura de cada *user*.

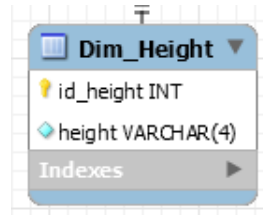


Figura 14: Dimensão Height

- **Dimensão *Smoker***

Esta dimensão indica se um *user* é ou não um fumador.

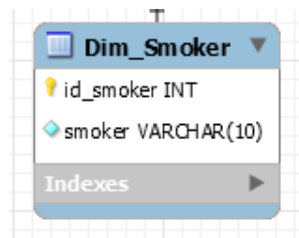


Figura 15: Dimensão Smoker

- **Dimensão *Ambience***

Esta dimensão é partilhada por duas tabelas de dimensão , tanto a do "Dim.User" como a do "Dim.Place"(que falaremos mais tarde). No primeiro caso esta dimensão pretende caracterizar o *user* na forma como se apresentou no restaurante, com família, amigos ou sozinho. No segundo caso pretende representar o restaurante em si, como sendo um restaurante familiar ou um restaurante mais calmo.

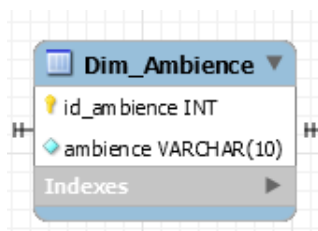


Figura 16: Dimensão Ambience

- **Dimensão *Budget***

Mais uma vez, esta dimensão é partilhada por duas tabelas de dimensão , tanto a do "Dim_User" como a do "Dim_Place". No primeiro caso, indica o budget de cada **user**. No segundo caso, indica a descrição de preços do restaurante, isto é, se o restaurante em causa é considerado ter preços baixos, médios ou altos.

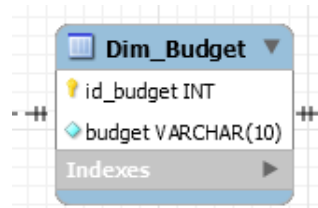


Figura 17: Dimensão Budget

- **Dimensão *Coordinates***

Mais uma vez, esta dimensão é partilhada por duas tabelas de dimensão , tanto a do "Dim_User" como a do "Dim_Place". No primeiro caso, esta dimensão apenas apresenta a latitude e longitude do local de residência do *user*. No segundo caso, indica a longitude e latitude de um certo restaurante.

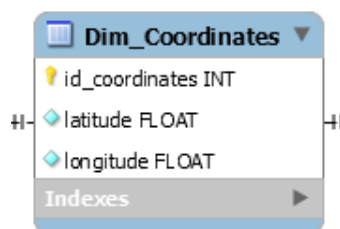


Figura 18: Dimensão Coordinates

3.1.2 Dimensão *Place*

Nesta dimensão é apresentado através de chaves estrangeiras todas as características de um restaurante da base de dados.

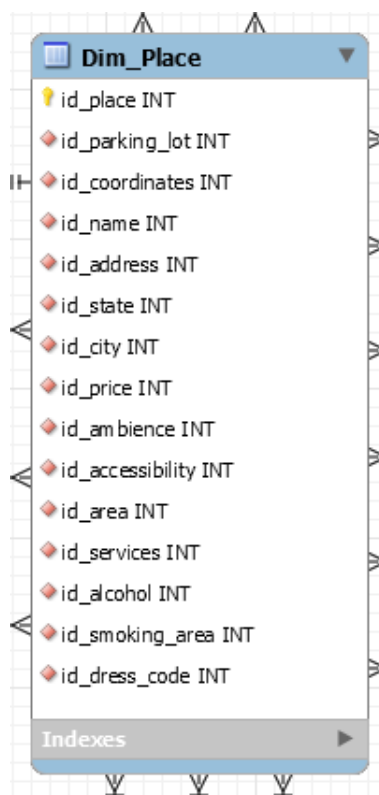


Figura 19: Dimensão Place

- Dimensão *City*

Esta dimensão apresenta a cidade onde cada restaurante se encontra.

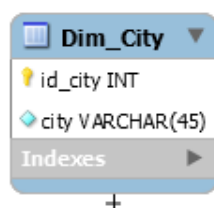


Figura 20: Dimensão City

- **Dimensão *State***

Esta dimensão indica o estado a que pertence um determinado restaurante.

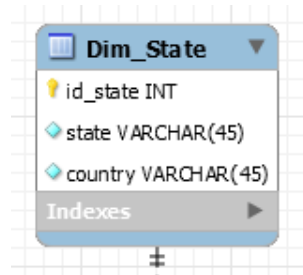


Figura 21: Dimensão State

- **Dimensão *Name***

Esta dimensão apresenta o nome de cada restaurante.

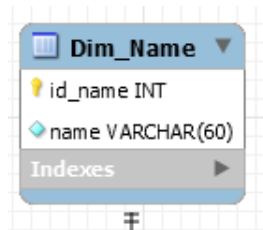


Figura 22: Dimensão Name Restaurant

- **Dimensão *Address***

Esta dimensão apresenta o endereço dos restaurantes.

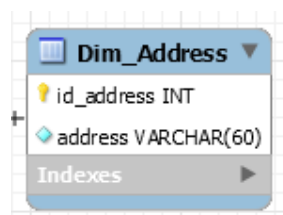


Figura 23: Dimensão Address

- **Dimensão *Alcohol***

Esta dimensão indica o tipo de serviço de venda de álcool que um restaurante oferece.

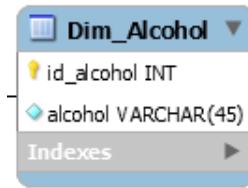


Figura 24: Dimensão Alcohol

- **Dimensão *Parking Lot***

Esta dimensão indica se um restaurante apresenta ou não parque de estacionamento.

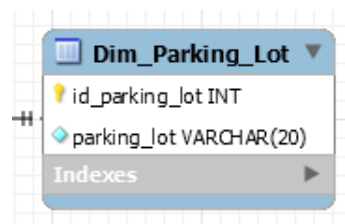


Figura 25: Dimensão Parking Lot

- **Dimensão *Dress Code***

Esta dimensão indica o tipo de vestuário dos empregados de um determinado restaurante. Isto é, se estes de vestem formalmente, informalmente ou casualmente.

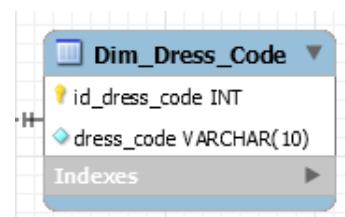


Figura 26: Dimensão Dress Code

- **Dimensão *Services***

Esta dimensão indica se o restaurante em causa apresenta outros serviços para além dos esperados de um restaurante.

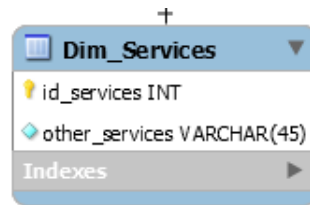


Figura 27: Dimensão Services

- **Dimensão *Smoking Area***

Esta dimensão indica se o restaurante tem ou não uma zona de fumadores.

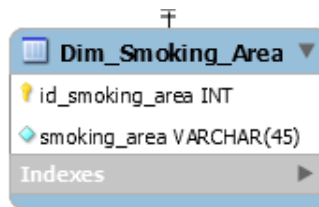


Figura 28: Dimensão Smoking Area

- **Dimensão *Accessibility***

Esta dimensão pretende indicar a acessibilidade de um dado restaurante.

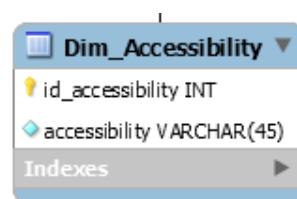


Figura 29: Dimensão Accessibility

- **Dimensão *Area***

Esta dimensão apenas indica se o restaurante tem um espaço aberto ou fechado.



Figura 30: Dimensão Area

3.1.3 Dimensão *Rating*

Esta dimensão apenas apresenta os ratings dados. Esta tabela apresenta três ligações à tabela de facto para representar os atributos relativos ao rating da comida do restaurante, ao rating do serviço do restaurante e ao rating geral do restaurante.

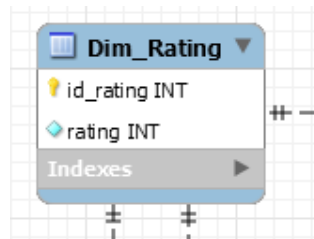


Figura 31: Dimensão Rating

3.2 Factos

Na realização deste modelo, o grupo optou por construir apenas uma tabela de factos. Esta tabela está ligada diretamente às tabelas de dimensão principais (como seria de esperar num modelo em floco de neve), sendo estas a "Dim_Place" e "Dim_User".

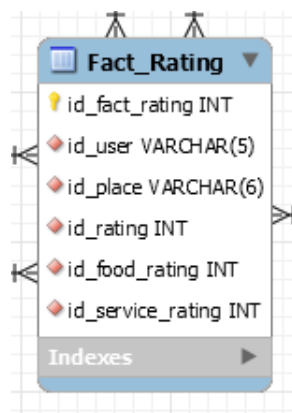


Figura 32: Dimensão Rating

3.3 Esquematização do Esquema Dimensional

Depois de definidas as tabelas de dimensão e de factos, partimos para a definição do nosso esquema dimensional, usamos um esquema Floco de neve, como já tínhamos referido anteriormente, isto é, um refinamento do esquema em estrela onde parte da hierarquia dimensional é normalizada num conjunto de tabelas dimensão mais pequenas, numa forma similar a um floco de neve.

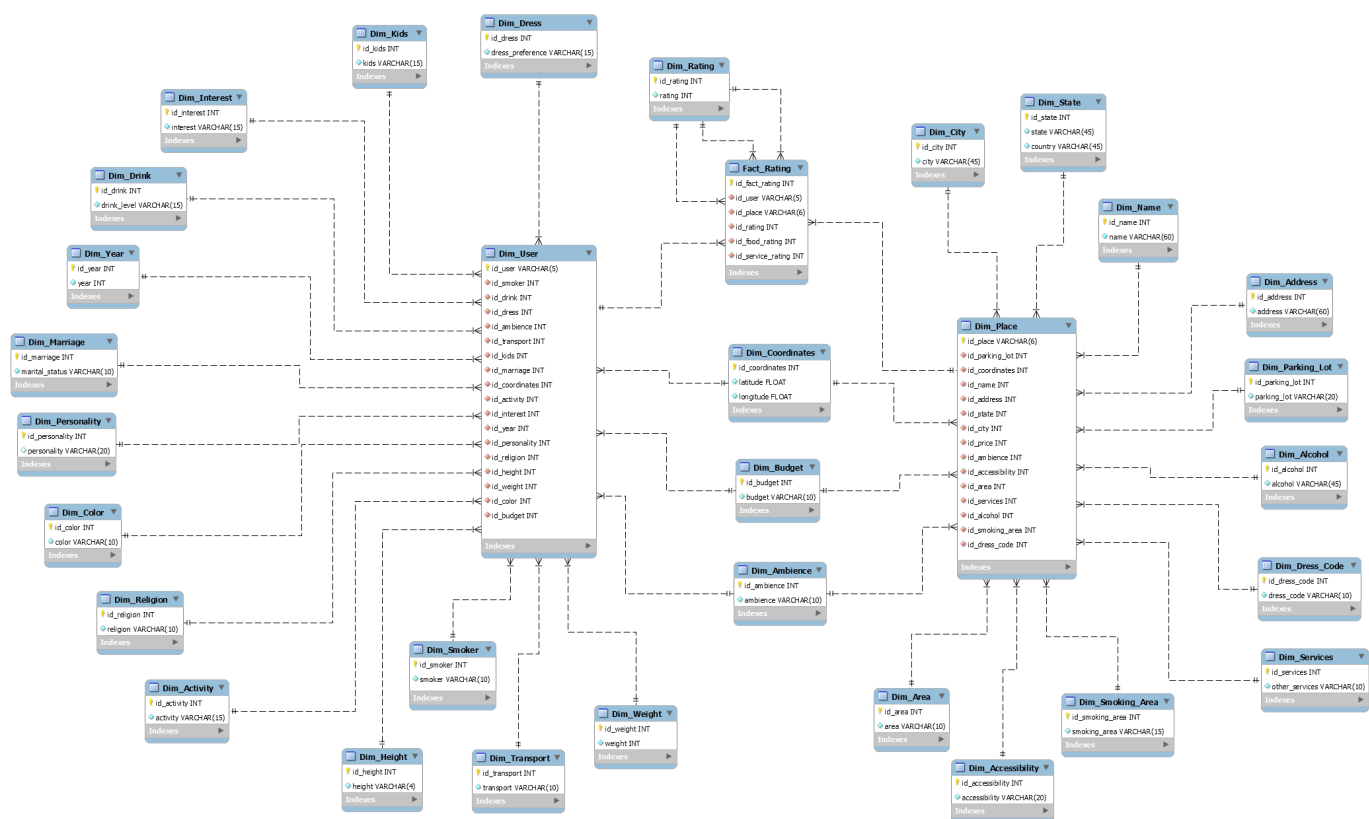


Figura 33: Esquema do Data Warehouse

4 Tratamento dos dados

Nesta parte, os ficheiros selecionados foram:

- userprofile.csv
- rating_final.csv
- geoplaces2.csv
- chefmozparking.csv

Os dados obtidos neste dataset, em geral, são "bons". Contém alguns "?", mas através de um 'replace' com regex, conseguimos substituir automaticamente para "X.NONE.X", e desta forma não apresentou qualquer tipo de problemas para o nosso Data Warehouse.

Eliminámos apenas 4 linhas, por ter uma quantidade excessiva de campos em falta:

```
-- apaga users com demasiados nulls
delete from Aux_User
where id_user = 'U1024'
   or id_user = 'U1083'
   or id_user = 'U1122'
   or id_user = 'U1130';
```

5 Povoamento do Data Warehousing

Para proceder ao povoamento da nossa base de dados, optámos por fazê-lo através de vários passos, incluindo a funcionalidade de import do MySQL Workbench, tabelas auxiliares, e queries SQL.

Parte 1 (Imports) - Criámos tabelas auxiliares onde pudemos colocar todos os dados considerados úteis do dataset fornecido.

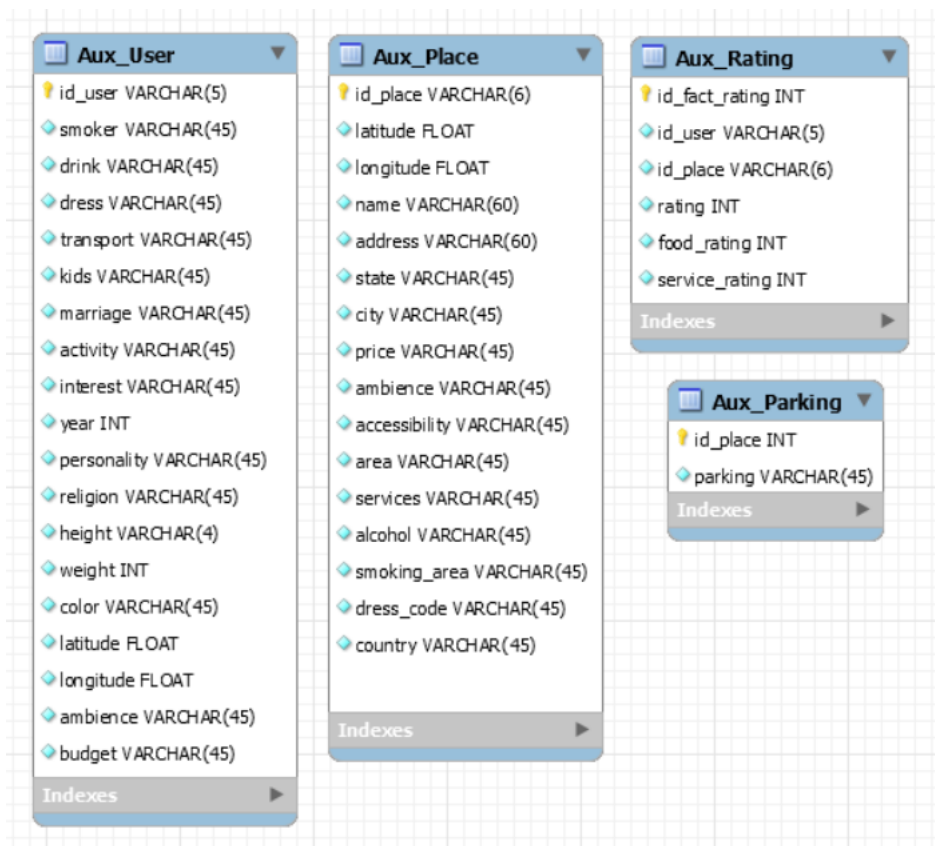


Figura 34: Esquema auxiliar usado para facilitar o povoamento

Depois, foi feito o seguinte processo para os ficheiros selecionados:

- 1) Aproveitámos a funcionalidade do MySQL Workbench que facilita os imports.
- 2) Escolhemos o ficheiro a fazer o import.
- 3) Escolhemos a tabela onde adicionar os dados em questão.
- 4) Selecionámos os atributos a inserir na respetiva tabela, e corrigimos os campos errados, que sendo preenchidos automaticamente podem estar errados quando o atributo a inserir não tem o mesmo nome que no ficheiro original.

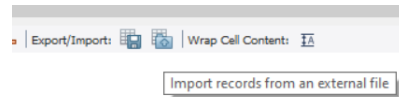


Figura 35: Funcionalidade de import do MySQL Workbench

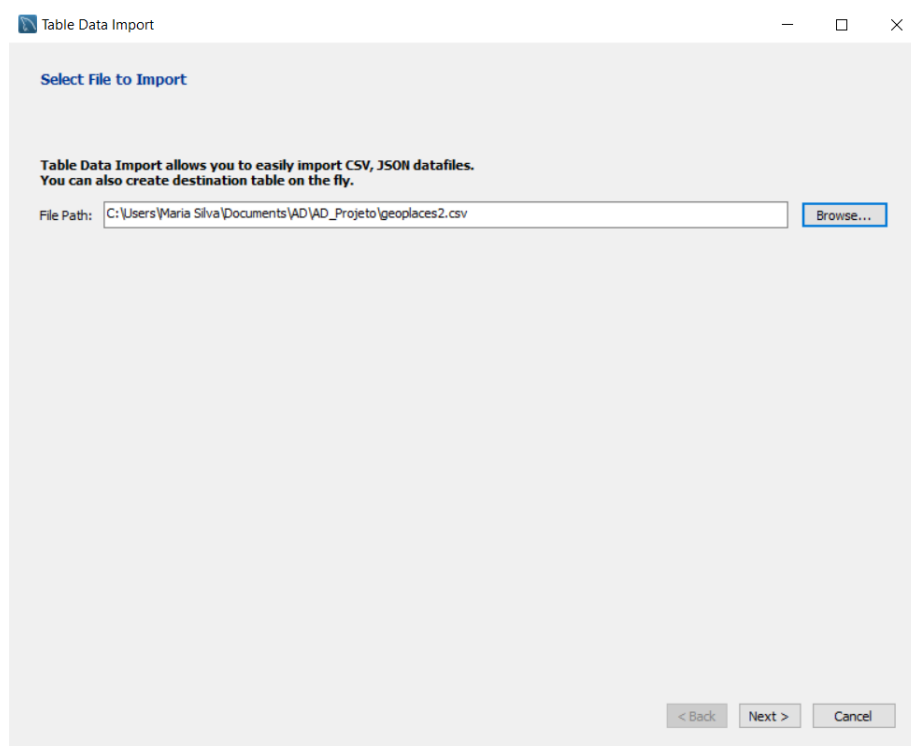


Figura 36: Método auxiliar usado para facilitar o povoamento

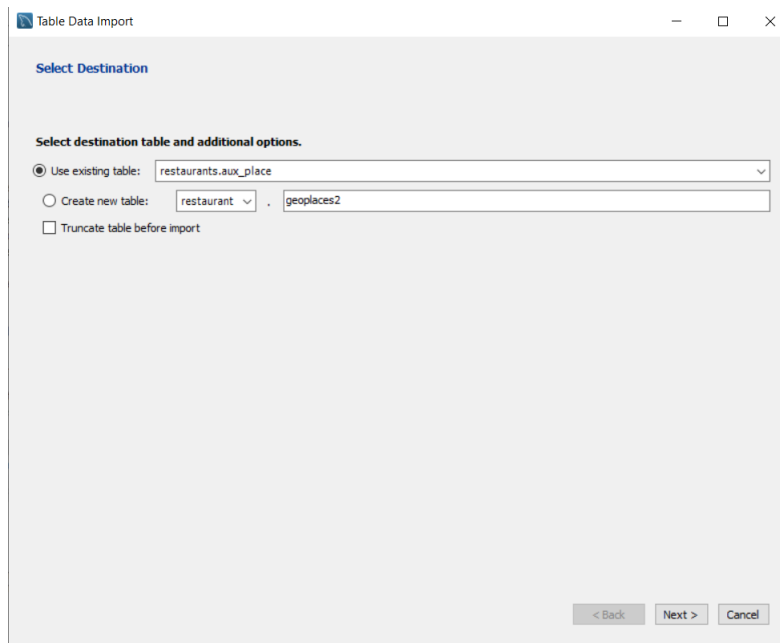


Figura 37: Método auxiliar usado para facilitar o povoamento

Parte 2 (Inserts) - Esta fase foi dividida em duas partes. Nos ficheiros fornecidos existia um README, com alguns tipos especificados, por exemplo:

```
4 chefmozparking.csv
Instances: 702
Attributes: 2
placeID: Nominal
parking_lot:Nominal, 7 [public,none,yes,valet_parking,
free,street,validated_parking]
```

Deste ficheiro, conseguimos perceber que existiam dois tipos de informações sobre os atributos: atributos com tipos fixos entre parêntesis (como no exemplo dado de parking_lot), e atributos sem tipos fixos (como no exemplo dado de placeID).

Para o primeiro tipo de atributos, os inserts foram feitos manualmente:

```
insert into Dim_Area (area) values ('open'),
                                   ('closed');
```

Para o segundo tipo de atributos, e não sendo o mesmo método possível, fizemos um novo tipo de queries baseado na seleção de todos os valores distintos existentes na dada coluna da tabela auxiliar:

```
insert into Dim_Address (address)
select distinct address from Aux_Place;
```

Parte 3 (Inserts Complexos) - Para as tabelas de dimensão principais (Dim_User e Dim_Place), foram feitos inserts num novo formato, de forma a ter os ID's de cada dimensão correspondente.

Table Data Import

Configure Import Settings

Detected file format: csv

Encoding: utf-8

Columns:

<input checked="" type="checkbox"/>	accessibility	accessibility
<input checked="" type="checkbox"/>	price	price
<input type="checkbox"/>	url	id_place
<input checked="" type="checkbox"/>	Rambience	ambience
<input type="checkbox"/>	franchise	id_place
<input checked="" type="checkbox"/>	area	area
<input checked="" type="checkbox"/>	other_services	services

placeID	latitude	longitude	the_geom...	name	address	city	state	country	fax
134999	18.915421	-99.184871	010100002...	Kiku Cuern...	Revolucion	Cuernavaca	Morelos	Mexico	X_NON
132825	22.1473922	-100.983092	010100002...	puesto de t...	esquinasan...	s.l.p.	mexico	X_NON	X_NON
135106	22.1497088	-100.97609...	010100002...	El Rincón...	Universidad...	San Luis Po...	San Luis Po...	Mexico	X_NON
132667	23.7526973	-99.1633594	010100002...	little pizza...	calle emilio...	victoria	tamaulipas	X NONE X	X NON

< Back Next > Cancel

Figura 38: Método auxiliar usado para facilitar o povoamento

Parte 4 (Tabela de Facto) - Foi feita uma query semelhante às da parte 3, embora menos complexa, devido ao número inferior de atributos a seleccionar.

6 Criação dos Indicadores

Por último, e já com a implementação terminada, temos de passar à criação de diversos indicadores referentes ao nosso modelo dimensional. Para isso, foram criadas diferentes dashboards com o intuito de fazer uma análise mais cuidada de cada indicador, podendo servir propósitos diferentes, como para tentar satisfazer as preferências dos clientes, ou simplesmente para fins estatísticos.

6.1 Indicadores por contagem de id_users

Com este indicador podemos analisar a preferência a nível da cor dos users que frequentaram os restaurantes da nossa base de dados. Desta forma, através da análise do histograma conseguimos concluir que a cor que no geral mais users gostam é azul. Podemos ainda acrescentar que a análise das preferências dos clientes é importante, pois desta forma poderá servir de base para melhoramentos nos vários restaurantes.

Quantidade de Users por preferência de cor

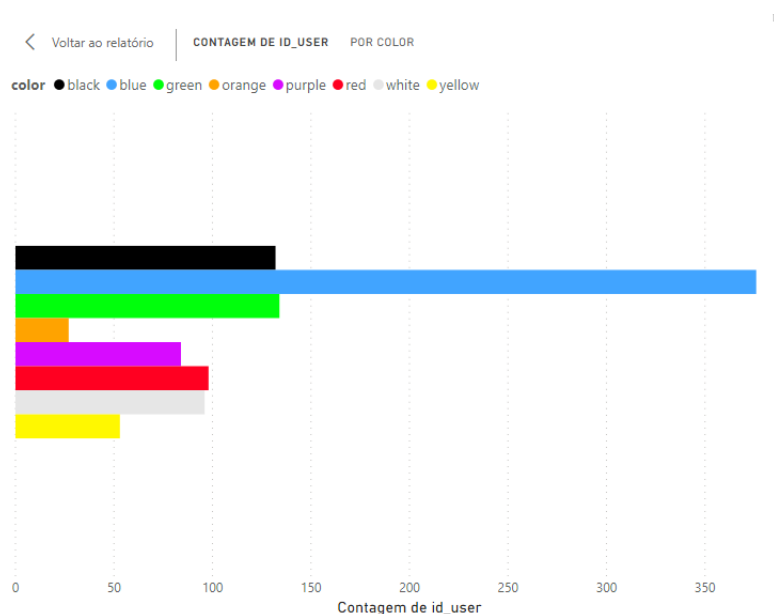


Figura 39: Quantidade de Users por preferência de cor

- **Quantidade de Users por nível de bebida**

A partir deste indicador, podemos concluir que grande parte dos users se consideram "casual drinker", o que poderá ser útil por parte dos restaurantes para saber que tipos de serviços fornecer e bebidas oferecer.

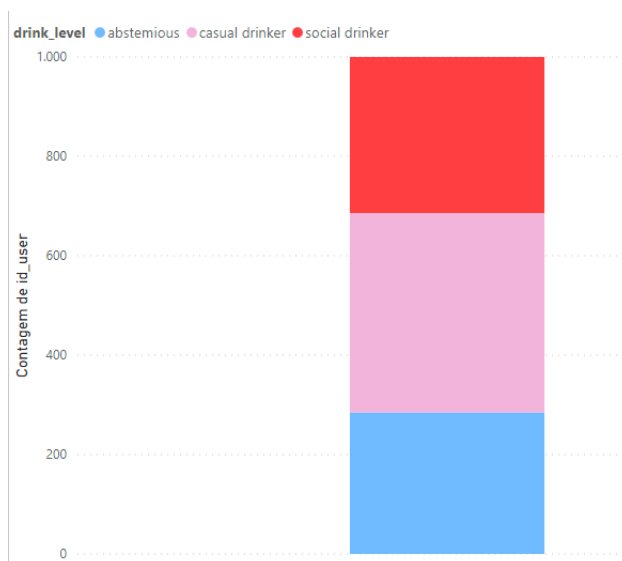


Figura 40: Quantidade de Users por nível de bebida

- **Quantidade de Users por preferência de vestuário**

Através da análise deste indicador podemos concluir que uma boa parte dos Users da nossa base de dados não têm preferências no seu vestuário.

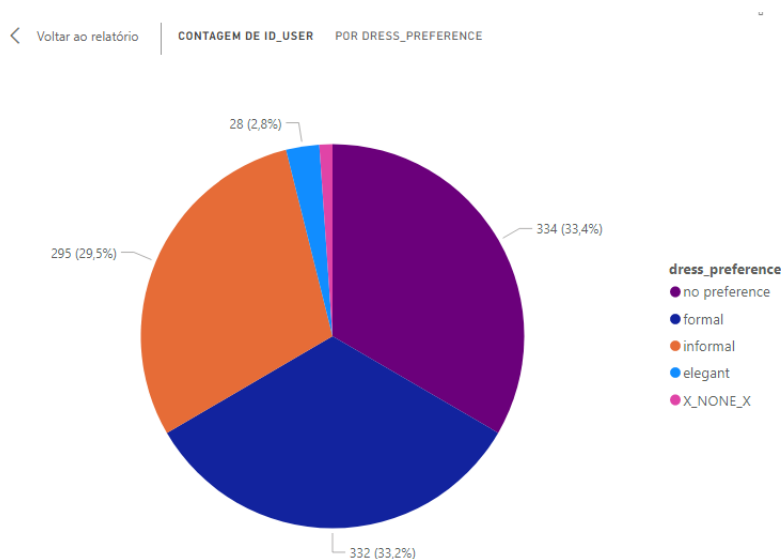


Figura 41: Quantidade de Users por preferência de vestuário

- **Quantidade de Users por personalidade**

Ao analisar este indicador concluímos que grande parte dos users se considera trabalhador.

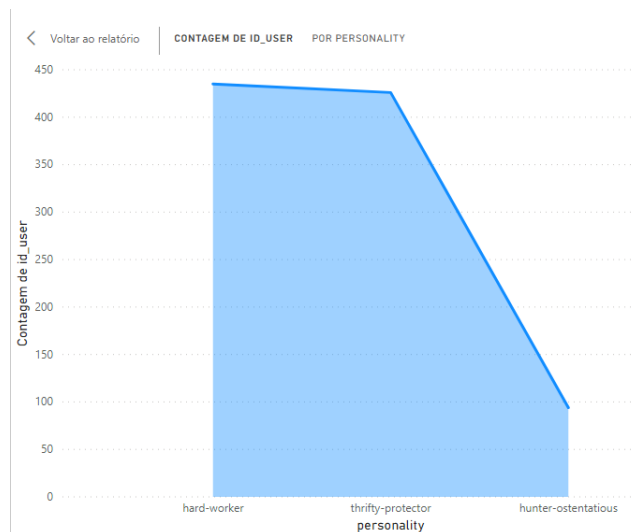


Figura 42: Quantidade de Users por personalidade

- **Quantidade de Users por religião**

Este indicador, permite-nos concluir que a maior parte dos Users são de religião católica.

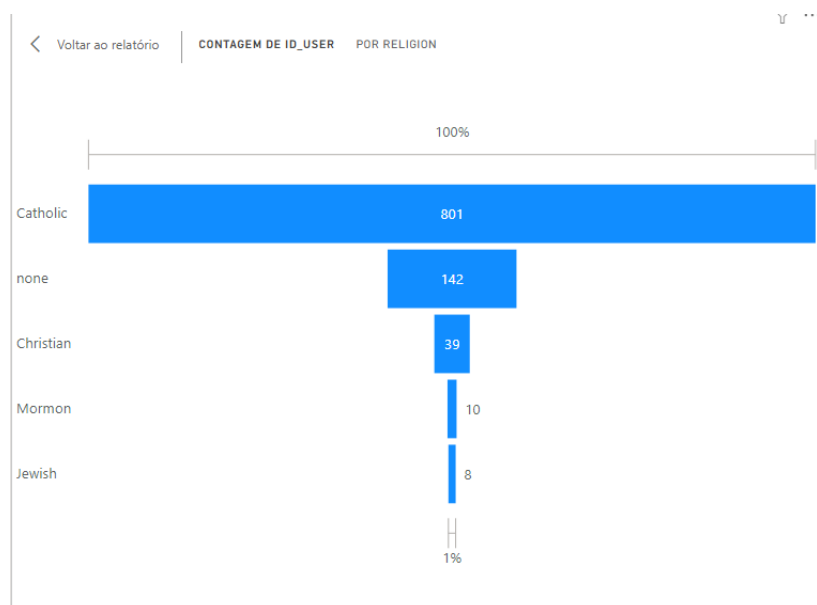


Figura 43: Quantidade de Users por religião

- **Quantidade de Users por fumador**

Com este indicador podemos concluir que a maior parte dos Users não são fumadores. Com estas informações os restaurantes poderão por exemplo, tomar decisões sobre a construção de possíveis áreas para fumadores, ou ainda sobre a instalação de máquinas de venda de tabaco.

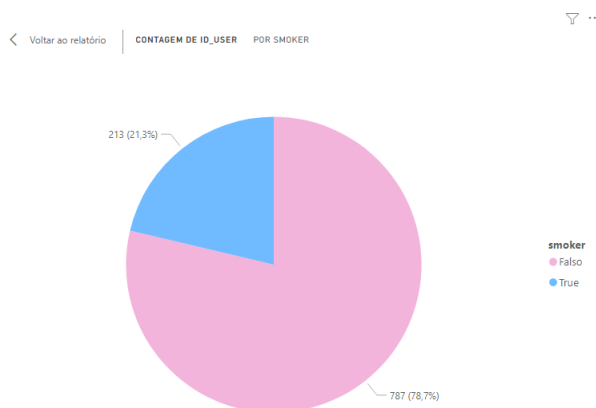


Figura 44: Quantidade de Users por fumador

- **Quantidade de Users por situação familiar**

Com este indicador conseguimos concluir que há uma maior percentagem de Users que são independentes do que a percentagem de Users com filhos. Este indicador é de valor informativo, pois desta forma cada restaurante pode, por exemplo, tomar decisões sobre campanhas promocionais para adultos em vez de menus infantis.

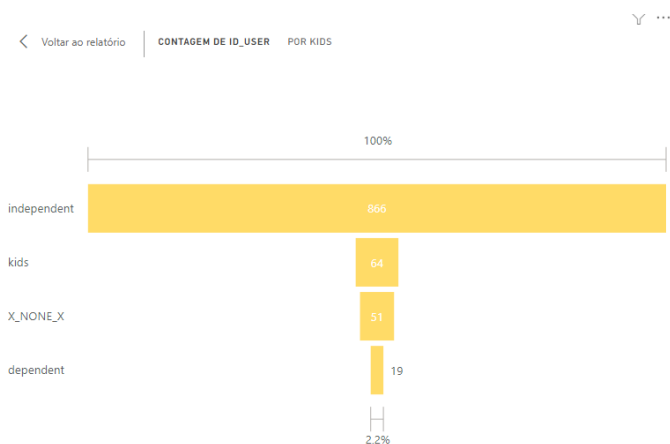


Figura 45: Quantidade de Users por situação familiar

- **Quantidade de Users por tipo de transporte utilizado**

Com este indicador conseguimos perceber que a maior parte dos Users na nossa base de dados se transporta normalmente de transportes públicos.

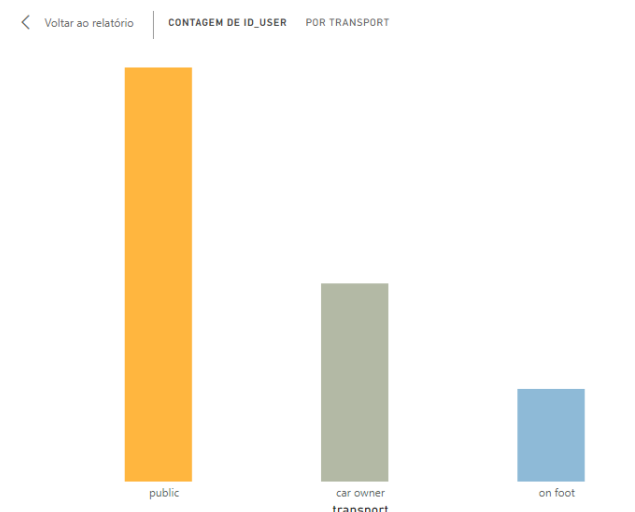


Figura 46: Quantidade de Users por tipo de transporte utilizado

- **Quantidade de Users por peso**

Através da análise deste indicador, podemos concluir que a maior parte dos Users que frequentaram os restaurantes desta base de dados não se encontram numa fasquia de peso demasiado acima da média. A maior parte dos indivíduos pesa cerca de 65kg, o que em princípio quer dizer que não estão em obesidade.

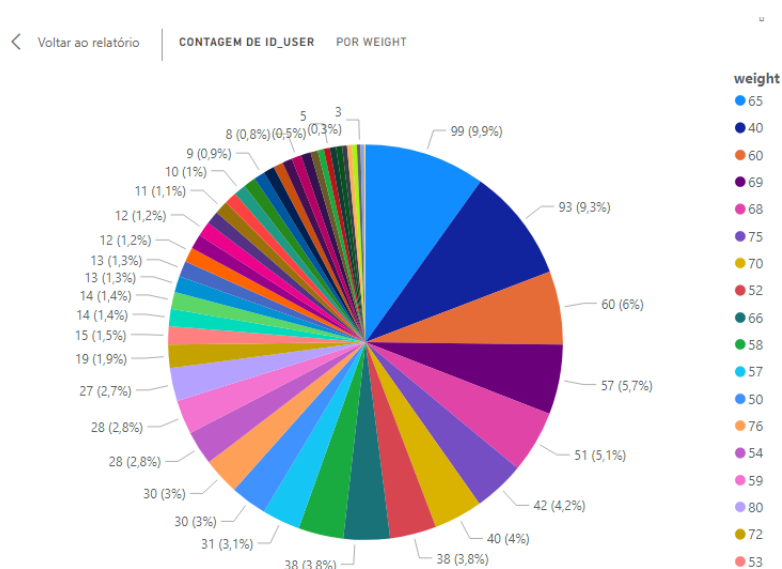


Figura 47: Quantidade de Users por peso

- **Quantidade de Users por ano de nascimento**

Com este indicador cada restaurante consegue avaliar a faixa etária de clientes que passa pelos seus restaurantes. Ao analisar este indicador percebemos que os Users que frequentaram os restaurantes da nossa base de dados são, na sua maioria, jovens adultos, pois a maioria dos Users nasceram sensivelmente entre 1990 e 1995.

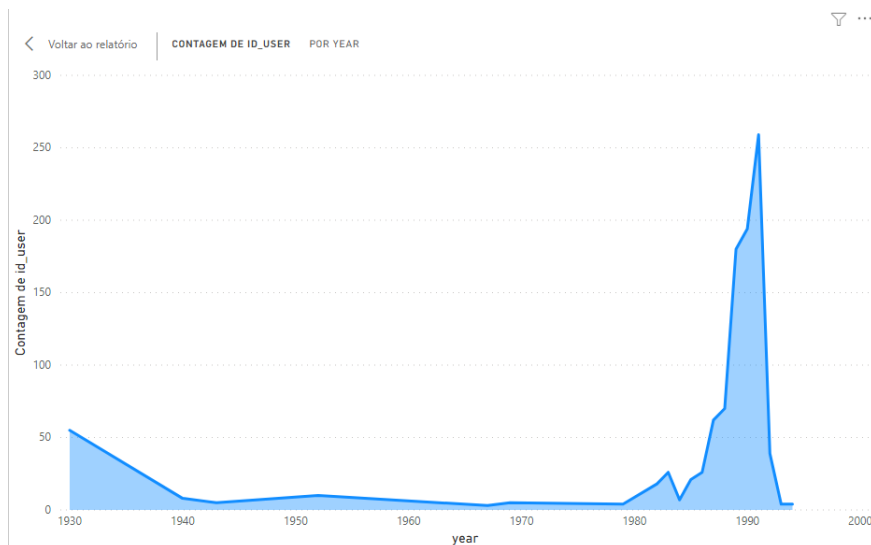


Figura 48: Quantidade de Users por ano de nascimento

- **Quantidade de Users por ambiente que prefere na ida a restaurantes**

Com a análise deste indicador, concluímos que a maior parte dos Users quando procuram um restaurante preferem fazê-lo com as suas famílias.

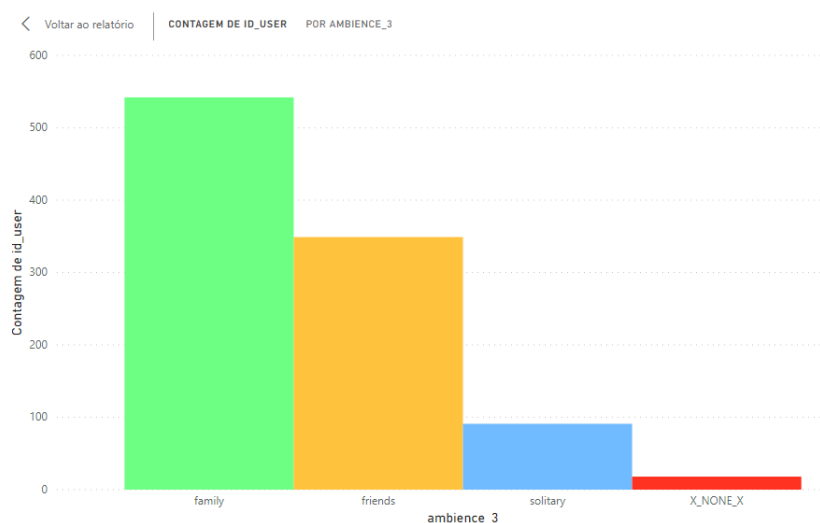


Figura 49: Quantidade de Users por ambiente que prefere na ida a restaurantes

- **Quantidade de Users por alturas do individuo**

Este indicador serve para fins estatísticos e ao observá-lo podemos concluir que a maioria dos clientes que frequentaram os restaurantes da nossa base de dados tem aproximadamente 1,6m de altura.

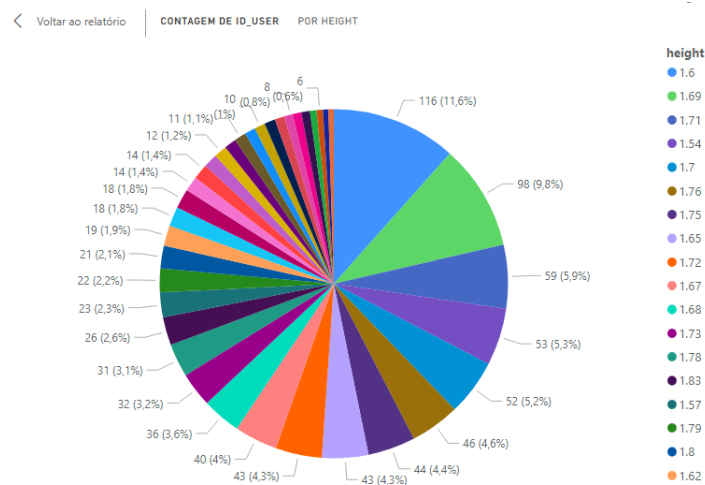


Figura 50: Quantidade de Users por alturas do indivíduo

- **Quantidade de Users por estado civil do indivíduo**

Através da análise deste indicador conclui-se que a maioria dos users são solteiros.

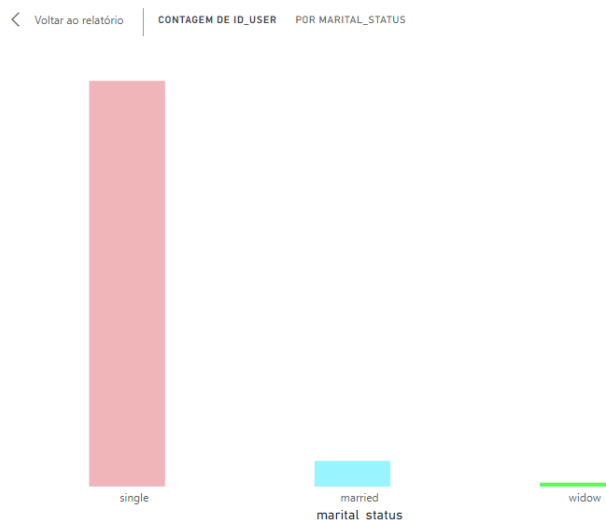


Figura 51: Quantidade de Users por estado civil do individuo

- **Quantidade de Users por estado atividade**

Ao analisar este indicador percebe-se que a grande maioria de users que frequentam os restaurantes desta base de dados são estudantes e que por outro lado, apenas uma pequena percentagem é desempregado.

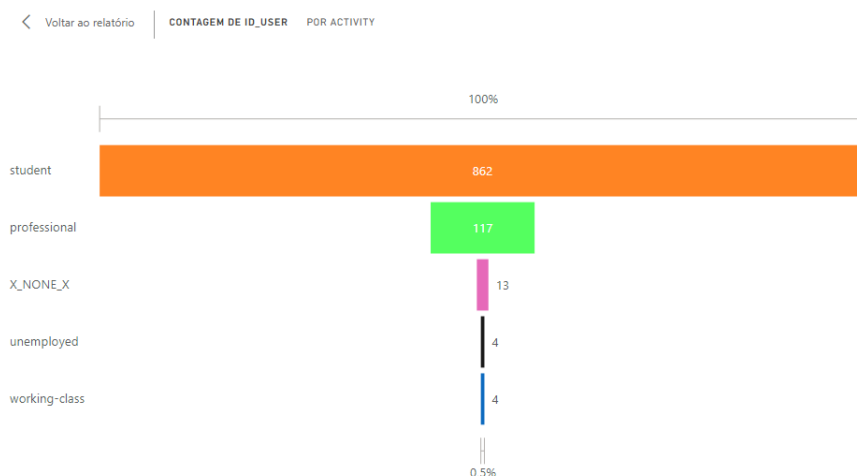


Figura 52: Quantidade de Users por atividade

- **Quantidade de Users por budget**

Com este indicador podemos concluir que a maior parte dos users desta base de dados tem um saldo médio para gastar nos restaurantes em questão.

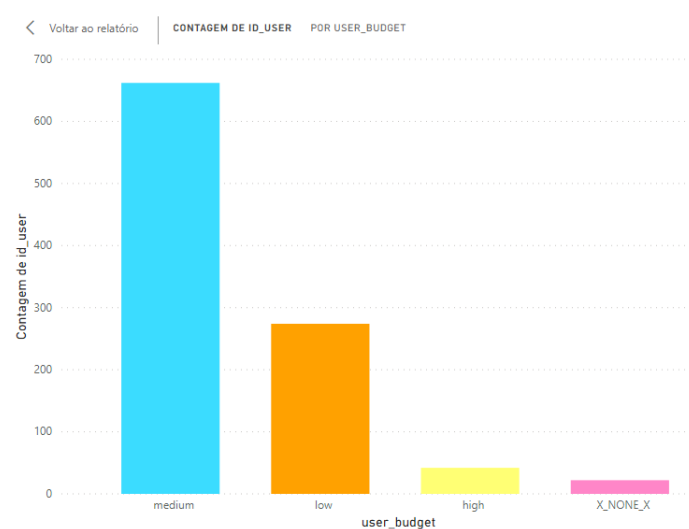


Figura 53: Quantidade de Users por budget

6.2 Indicadores por contagem de id_place

- **Quantidade de restaurantes por nome**

Com este indicador podemos concluir que o restaurante com mais estabelecimentos (da nossa base de dados) no México é o "Tortas Locas Hipocampo".

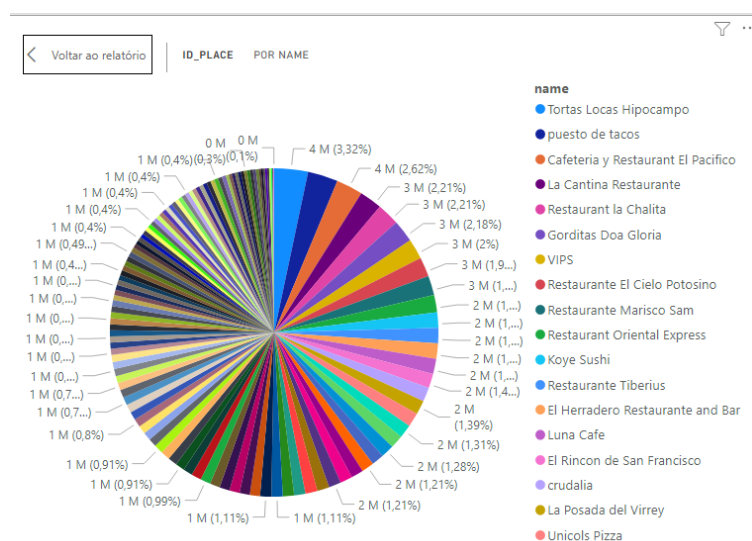


Figura 54: Quantidade de restaurantes por nome

- **Quantidade de restaurantes por tipo de álcool servido**

Através da análise deste indicador conseguimos concluir que, embora as três percentagens estejam muito próximas, há uma maior percentagem de restaurantes com bar aberto. Esta análise é importante para podermos comparar com a regularidade com que os clientes bebem, de modo a perceber quais restaurantes escolher.

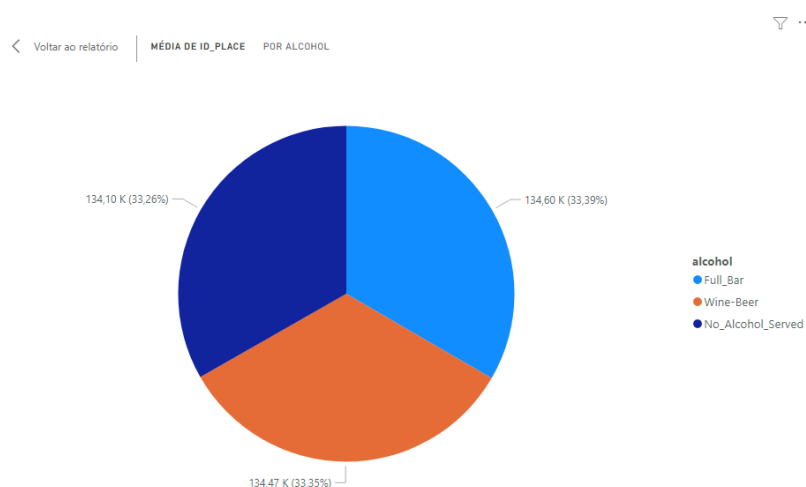


Figura 55: Quantidade de restaurantes por tipo de álcool servido

- **Quantidade de restaurantes por acessibilidade**

Com este indicador conclui-se que a maior parte dos restaurantes não tem facilidade de acesso para pessoas com deficiências.

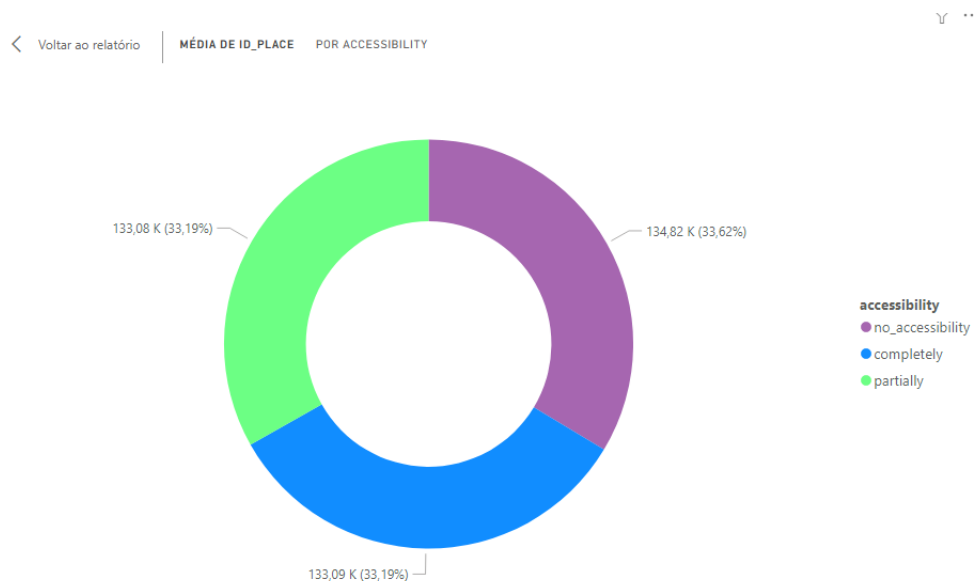


Figura 56: Quantidade de restaurantes por acessibilidade

- **Quantidade de restaurantes por área de fumadores**

Através da análise deste indicador concluímos que a maior parte dos restaurantes apenas permite os clientes fumarem na zona de bar. Mais uma vez esta análise é importante para a escolha, podendo assim comparando os users que são ou não fumadores, com os restaurantes que assim o permitam.

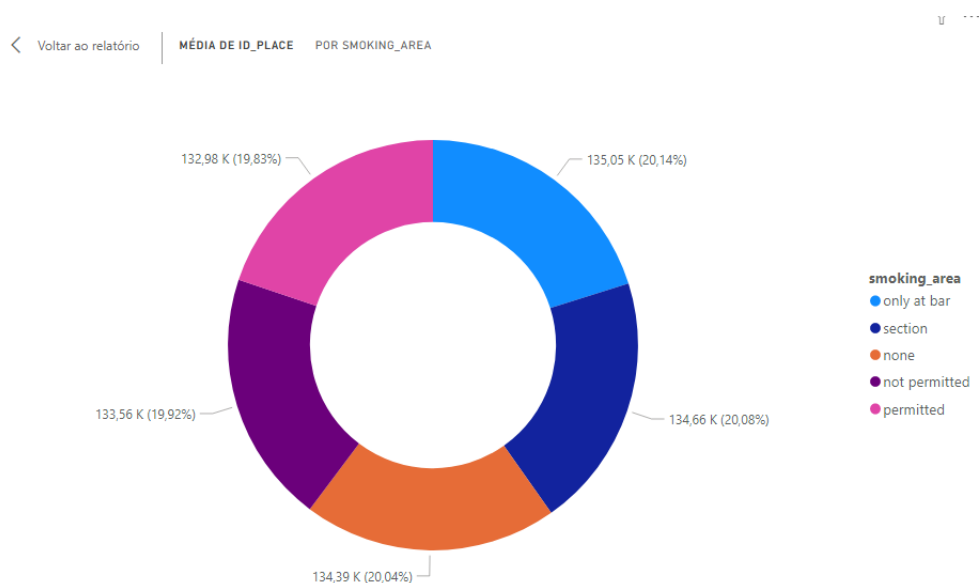


Figura 57: Quantidade de restaurantes por área de fumadores

- **Quantidade de restaurantes por código de vestuário**

Com este indicador percebemos que a maior parte dos restaurantes têm como expectativa os clientes usarem roupas informais, sendo pouca a diferença para os que esperam roupas formais ou casuais.

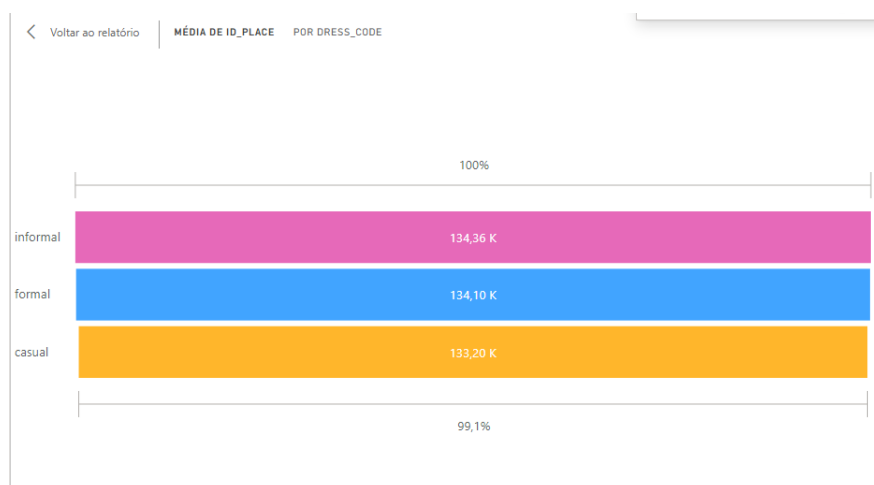


Figura 58: Quantidade de restaurantes por código de vestuário

- **Quantidade de restaurantes por prestação de outros serviços**

A partir deste indicador podemos concluir quais os restaurantes que prestam serviços além dos lógicos, o que será útil para a tomada de decisão por parte dos clientes. Assim conclui-se que na nossa base de dados a maior parte dos restaurantes disponibiliza internet aos seus clientes, apesar de estarem quase igualmente equilibrados.

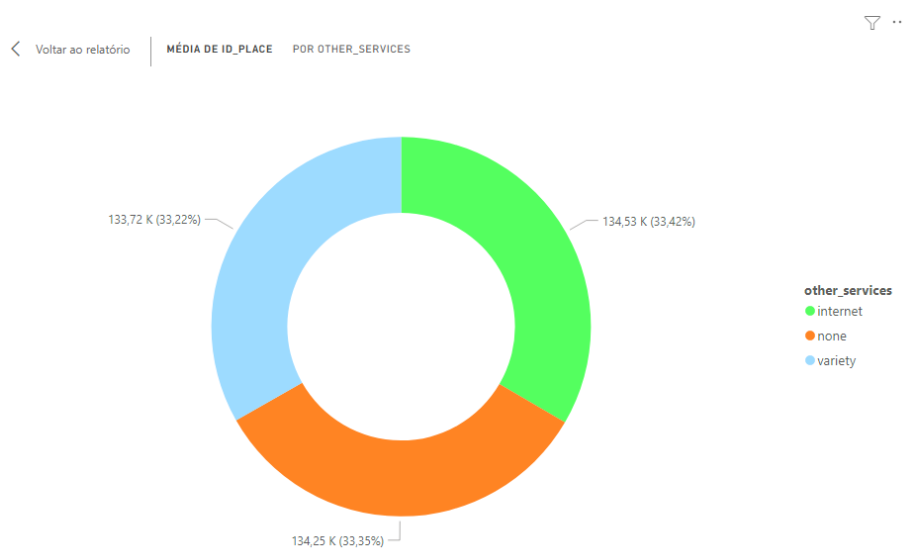


Figura 59: Quantidade de restaurantes por Código de vestuário por prestação de outros serviços

- **Quantidade de restaurantes por tipo de parque de estacionamento**

Embora a diferença entre percentagens seja muito pequena, ao analisarmos este indicador percebemos que a maior parte dos restaurantes da nossa base de dados tem parque de estacionamento público.

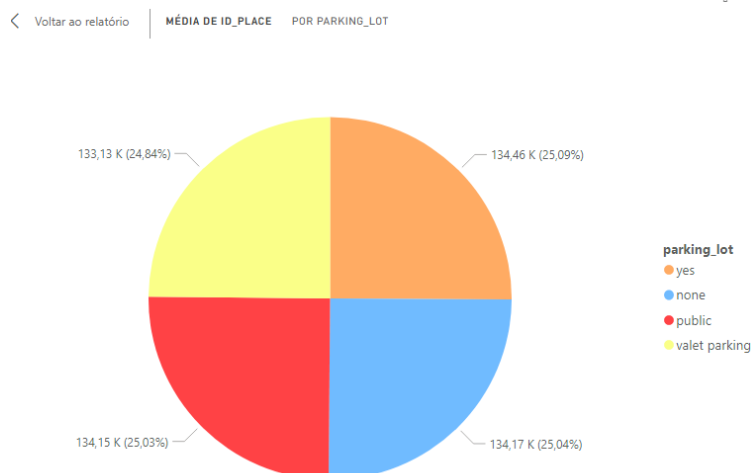


Figura 60: Quantidade de restaurantes por tipo de parque de estacionamento

- **Quantidade de restaurantes por tipo de espaço**

Com este indicador podemos concluir que, embora as duas percentagens estejam muito próximas, a maior parte dos restaurantes tem estabelecimentos de espaço fechado.

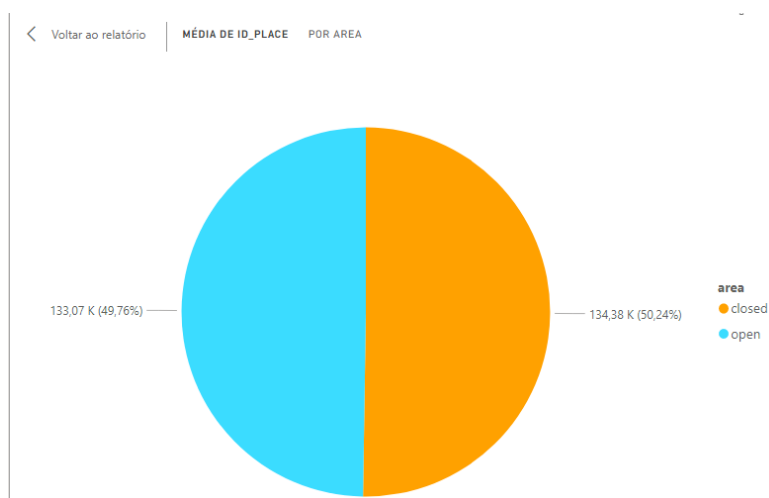


Figura 61: Quantidade de restaurantes por tipo de espaço

- **Quantidade de restaurantes por budget**

Através deste indicador conseguimos concluir que a maior parte dos restaurantes da nossa base de dados apresenta valores considerados altos.

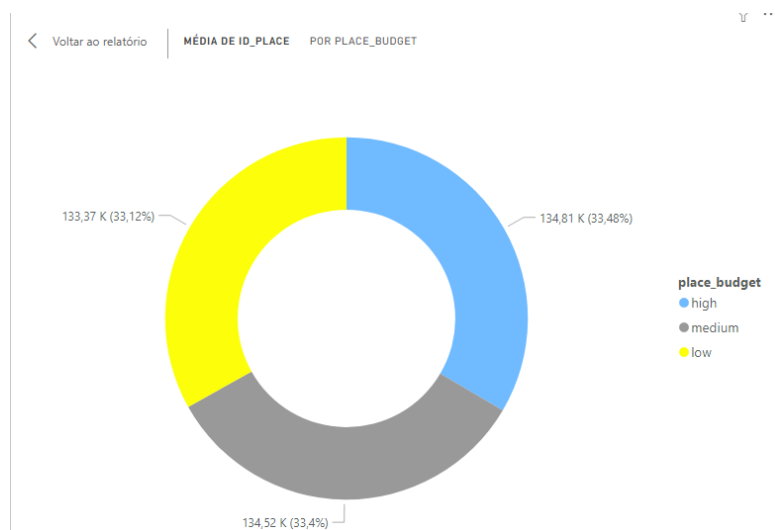


Figura 62: Quantidade de restaurantes por budget

6.3 Indicadores por contagem de id_fact_rating

Ao analisar este indicador conclui-se que a maior parte dos users da nossa base de dados deram ratings de 2 (máximo) pelos serviços prestados nos vários restaurantes.

- **Quantidade de cada rating do serviço dos restaurantes**

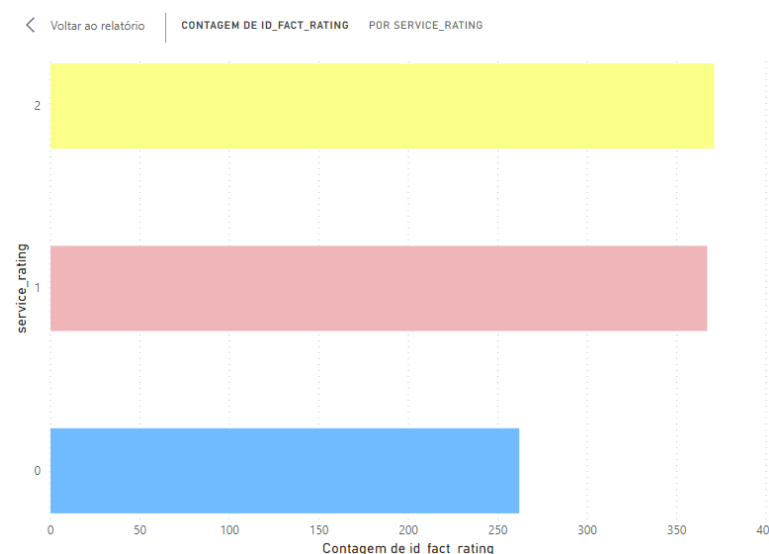


Figura 63: Quantidade de cada rating do serviço dos restaurantes

- **Quantidade de cada rating da comida dos restaurantes**

Ao analisar este indicador conclui-se que a maior parte dos users da nossa base de dados deram ratings de 2 (máximo) pela comida servida nos vários restaurantes.

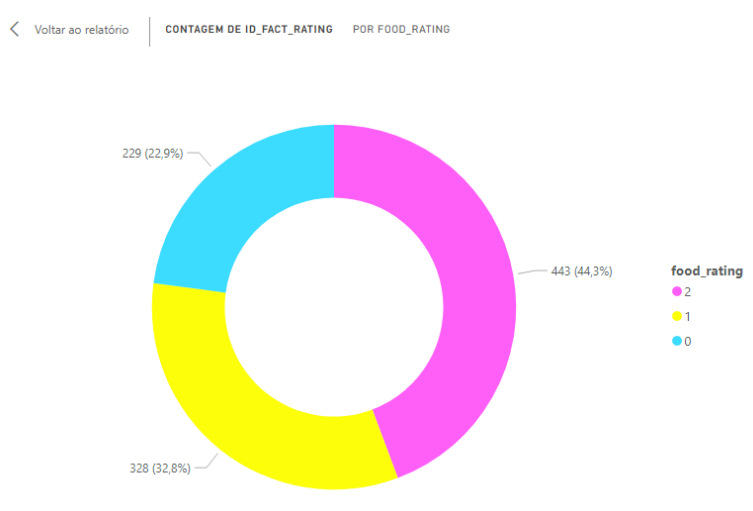


Figura 64: Quantidade de cada rating da comida dos restaurantes

- **Quantidade de cada rating geral dos restaurantes**

Ao analisar este indicador conclui-se que a maior parte dos users da nossa base de dados deram ratings de 2 (máximo) ao restaurante no geral.

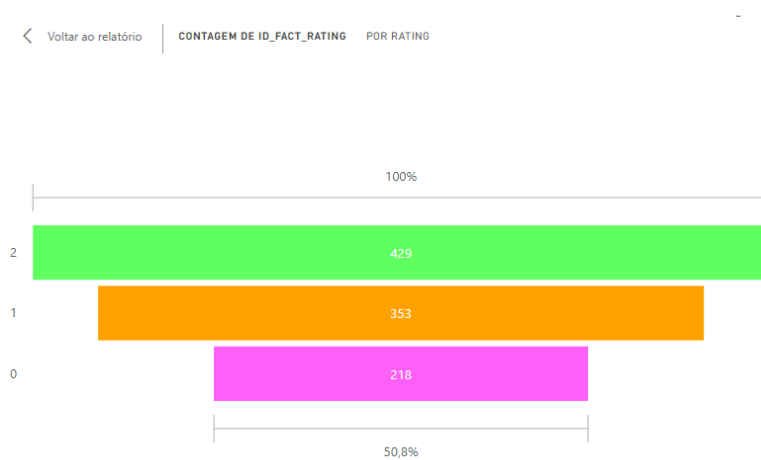


Figura 65: Quantidade de cada rating geral dos restaurantes

7 Conclusões

Inicialmente com a realização deste trabalho enfrentámos alguns problemas, nomeadamente na escolha do dataset, pois todos os que escolhíamos apresentavam bastante inconsistência nos dados ou então apresentavam pouca manobra para mais tarde criarmos indicadores sobre este dataset. Desta forma, consideramos esta parte a mais difícil do trabalho e a que levou mais tempo, pois muitas vezes só nos apercebíamos deste problema depois de avançados no trabalho.

Depois de escolhido o dataset que usamos no resto do trabalho prático, procedemos ao resto da realização deste sem grandes contratemplos ou dificuldades.

Este trabalho foi importante, pois ajudou-nos a perceber a dificuldade da implementação de sistemas de *Business Intelligence* para o suporte à decisão. Foi útil ainda para nos habituarmos a utilizar plataformas como o *Microsoft Power BI Desktop*. Permitiu-nos também aplicar os conhecimentos sobre um sistema de *Data Warehousing* com a criação de uma base de dados usando um modelo sob a forma de floco de neve.