

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE INFORMÁTICA



Análise Geográfica de uma Rede de Agências Bancárias

Mestrado em Gestão de Informação
Gestão e Análise de Dados

Inês Maria Gomes Ferreira

Trabalho de Projeto orientado por:
Professor Doutor António Manuel Silva Ferreira

2015

Agradecimentos

Ao Professor Doutor António Ferreira pela sua orientação, disponibilidade e contributo ao longo de todo o projeto.

À Caixa Geral de Depósitos pela disponibilização de informação interna, representada pelo Paulo Semblano que sempre se mostrou disponível.

Aos professores que contribuíram de alguma forma para a realização deste projeto e aos que estiveram presentes no meu percurso académico pela partilha de conhecimentos.

Aos meus pais, irmãos e padrinhos pela presença e apoio que sempre deram.

Aos meus amigos que me acompanharam na vida académica e me ajudaram quando necessário.

Resumo

A análise de uma rede comercial é um fator de grande importância para um negócio, em particular no mercado bancário onde os investimentos são avultados. Neste contexto, a escolha da localização de uma nova agência, ou o encerramento de uma existente, passam por um processo decisório complexo que considera fatores internos e externos, sendo fundamental acautelar o sucesso futuro. Portanto é cada vez mais necessário fazer avaliações regulares e estruturadas ao desenho da rede, sendo que neste trabalho é considerada a rede de agências da Caixa Geral de Depósitos (CGD).

Neste relatório descreve-se uma metodologia suportada num sistema de informação geográfica (SIG) e num modelo gravitacional que possibilita a análise da rede de agências da CGD, tendo por objetivo classificar as agências em função da sua localização e outras características.

Começou-se pelo conhecimento dos dados relativos às agências e à sua localização, que consistiu na realização de transformações, para poderem ser representados em mapas. Nomeadamente foi feita conversão de moradas em coordenadas geográficas com recurso a um serviço da Google e foram consultadas várias fontes com dados sobre os municípios e distritos.

De seguida, desenvolveu-se um programa em Javascript para obter matrizes de distâncias e tempos de percursos realistas entre clientes e agências, que têm em conta dois modos de deslocação, a pé e de carro. Para este efeito foi usado o serviço Distance Matrix da biblioteca Google Maps.

Com base na matriz de distâncias foram desenvolvidos três programas em R para calcular percentagens do modelo gravitacional de Huff: um primeiro que só considera as agências da CGD, um segundo que inclui as agências da concorrência e um terceiro que procura a percentagem de Huff máxima no caso de existirem vários pontos de origem de clientes.

Por último, a informação obtida do modelo de Huff foi visualizada em mapas usando um SIG, permitindo a avaliação da estrutura e características da rede de agências da CGD.

Através do trabalho realizado, é também contemplada a capacidade de analisar cenários resultantes da conjugação e/ou manipulação de fatores internos e externos simulados ou com valores previstos para o futuro, que podem provocar possíveis alterações na rede de agências, indo ao encontro de um leque mais alargado de necessidades de um decisor.

Palavras-chave:

Análise de Rede de Agências, Sistemas de Informação Geográfica, Modelo de Huff, *Geomarketing*.

Abstract

The analysis of a commercial network is an extremely important factor to a business, in particular when it comes to the bank market where the investments are usually big. In this context, choosing the location of a new agency or closing an already existing one is something that goes through a complex deciding process that takes into account both internal and external factors, thus being of particular importance to take care of the future success. It is more and more necessary to undergo regular and structured evaluations regarding the scheme of the network; in this paper we're looking into the Caixa Geral de Depósitos (CGD) agency network.

In this report is described a methodology supported by a geographical information system (GIS) and a gravitational model that makes the CGD agency network analysis possible, so which ultimate goal is to classify the agencies according to their locations and other characteristics.

The project started with a data agencies knowledge and their location, consisting on transformations so it could appear on maps. A conversion of addresses into geographical coordinates was made by using a Google service. Also most of the information about the different counties and districts was consulted from different sources.

Based on the distance matrix three 'R' programs were developed to calculate percentages off of the Huff gravitational model: the first one, which only takes into account the CGD agencies; the second one, including all the other competing agencies; and the third one, that looks for the maximum Huff percentage, in case there are several points of clients source.

At last, the information obtained from Huff model was displayed on maps using a GIS, allowing the evaluation of structure and characteristics of CGD's agencies network.

Through the work, it is also considered the ability to analyze scenarios resulting from the combination and/or manipulation of a pretending internal and external factors or predicted values for the future, that could cause possible changes in the agencies network, going to meet a wider range of needs of a decision maker.

Keywords:

Agencies Network Analysis, Geographic Information Systems, Huff model, Geomarketing.

Lista de Siglas

ACTIVOBANK - Banco Activobank, SA

BANIF - Banco Internacional do Funchal, SA

BANIFM - Banco Banif Mais, SA

BARCLAYS - Barclays Bank, PLC

BBVA - Banco Bilbao Vizcaya Argentaria (Portugal), SA

BCP - Banco Comercial Português, SA

BEST - Banco Electronico De Servico Total, SA

BIC - Banco Bic Portugues, SA

BIG - Banco De Investimento Global, SA

BP – Banco de Portugal

BPI - Banco BPI, SA

BPP - Banco Popular Portugal, SA

CAOP - Carta Administrativa Oficial de Portugal

CCA - Caixa De Credito Agricola Mutuo, CRL

CCCAM - Caixa Central - Caixa Central De Credito Agricola Mutuo, CRL

CEMAH - Caixa Economica Da Misericordia De Angra Do Heroismo

CEMG - Caixa Economica Montepio Geral

CGD – Caixa Geral de Depósitos

CTT – Correios e Telecomunicações de Portugal

DBA - Deutsche Bank Aktiengesellschaft - Sucursal Em Portugal

DG Território – Direção Geral do Território

INVEST – Banco Invest, SA

NB – Novo Banco, SA

SIG - Sistema de Informação Geográfica (*GIS – Geographic information system*)

TOTTA – Banco Santander Totta, SA

Glossário

Área de influência ou *trade area*: Também denominada amplitude, diz respeito à distância máxima que os clientes estão dispostos a percorrer na busca por um produto ou serviço ^[1]. É uma área geográfica que contém consumidores de bens ou serviços específicos de uma empresa particular ou grupo de empresas ^[2].

Áreas administrativas: Áreas delimitadas pelos limites administrativos oficiais ao nível da freguesia, concelho, e distrito/ilha ^[3].

Atratividade (modelo de Huff): Medida da capacidade de uma agência para atrair clientes, tendo em conta um ou mais fatores, tais como o tamanho da agência, o número de lugares de estacionamento, entre outros.

Coordenadas geográficas: Dizem respeito à latitude e à longitude, definidas na superfície terrestre ou numa superfície de referência ^[3].

Densidade populacional: Intensidade do povoamento expressa pela relação entre o número de habitantes de uma área territorial determinada e a superfície desse território, habitualmente expressa em número de habitantes por quilómetro quadrado ^[4].

Ganho médio mensal: É o montante que o empregado recebe de facto todos os meses. Para além da remuneração de base, inclui outras pagas pelo empregador, como horas extra, subsídio de férias, e prémios ^[4].

Geomarketing: Também conhecido como *marketing* geográfico, é uma ferramenta que auxilia a tomada de decisão através da utilização de abordagens geográficas direcionadas à vertente do *marketing*, por exemplo a observação de mapas geográficos anotados com variáveis importantes ao *marketing* ^[1].

Informação externa: Informação relativa à população residente e à existência ou não de concorrência ao nível do município.

Informação interna: Informação de negócio fornecida pela CGD, como o número de clientes.

Mercado: Local onde agentes económicos procedem à troca de bens por uma unidade monetária ou por outros bens.

Percentagem de Huff: Probabilidade de um cliente numa determinada localização ir a uma determinada agência.

Índice

| | |
|--|------|
| Agradecimentos | iii |
| Resumo..... | v |
| Abstract | vii |
| Lista de Siglas | ix |
| Glossário..... | xi |
| Índice..... | xiii |
| Lista de Figuras..... | xv |
| Lista de Tabelas | xix |
| Capítulo 1 - Introdução | 1 |
| 1.1. Motivação..... | 1 |
| 1.2. Objetivos | 2 |
| 1.3. Principais resultados obtidos | 3 |
| 1.4. Notação adotada..... | 3 |
| 1.5. Organização do documento | 3 |
| Capítulo 2 – Conceitos..... | 5 |
| 2.1. SIG – Sistemas de Informação Geográfica | 5 |
| 2.2. Modelos gravitacionais | 6 |
| 2.3. Áreas de influência | 9 |
| 2.4. Discussão de conceitos..... | 10 |
| 2.5. Sumário do capítulo | 11 |
| Capítulo 3 - Trabalho realizado | 13 |
| 3.1. Ambiente de trabalho | 13 |
| 3.2. Concretização dos objetivos..... | 13 |
| 3.2.1. Conhecimento e transformação de dados..... | 14 |
| 3.2.2. Obtenção da matriz de distâncias | 35 |
| 3.2.3. Aplicação do modelo de Huff | 39 |
| 3.2.4. Visualização dos dados..... | 45 |
| 3.3. Cenários de análise | 49 |
| 3.4. Sumário do capítulo | 58 |
| Capítulo 4 – Conclusões | 59 |

| | |
|---|----|
| 4.1. Principais contribuições | 59 |
| 4.2. Competências adquiridas | 59 |
| 4.3. Dificuldades encontradas | 60 |
| 4.4. Trabalho Futuro..... | 60 |
| Referências..... | 63 |
| Apêndices | 67 |
| Apêndice A - Obtenção da matriz de distâncias | 69 |
| Apêndice B - Cálculo das percentagens de Huff para agências da CGD..... | 75 |
| Apêndice C - Cálculo das percentagens de Huff para agências da CGD e da concorrência | 79 |
| Apêndice D - Pesquisa das percentagens de Huff máximas | 83 |

Lista de Figuras

| | |
|---|----|
| Figura 1 - Fórmula de cálculo do modelo de Converse ^[21] | 8 |
| Figura 2 - Fórmula de cálculo da percentagem de Huff..... | 8 |
| Figura 3 – Etapas da metodologia proposta. | 14 |
| Figura 4 - Processo de conversão das moradas da concorrência em coordenadas geográficas.17 | |
| Figura 5 - Diagrama da ligação das fontes de dados utilizadas..... | 18 |
| Figura 6 - Adição de uma camada vetorial no QGIS..... | 20 |
| Figura 7 - Adição de uma camada de texto delimitado no QGIS. | 21 |
| Figura 8 - Passos executados para abertura da janela de agregação. | 21 |
| Figura 9 - Janela onde é escolhida a camada, o campo de agregação e o ficheiro de saída. | 22 |
| Figura 10 - Escolha do local e do nome do ficheiro resultante da agregação é guardado. | 22 |
| Figura 11 - Resultado da agregação por distrito, para o distrito de Lisboa. a) Representação original por freguesia, b) Agregação por distrito e c) Resultado após limpeza da agregação.... | 23 |
| Figura 12 – Representação da localização de agências no mapa de Portugal. a) Diz respeito a Portugal só com as agências da CGD representadas, b) São representadas apenas as agências da concorrência, e c) Sobreposição de a) b). | 25 |
| Figura 13 - Mapa do município de Lisboa com a diferenciação da concorrência. a) Representação de todas as agências e b) Sobreposição de a) com Google Maps..... | 26 |
| Figura 14 - Percentagem de agências da CGD por distrito e município com sobreposição da localização das agências. a) Percentagem de agência por distrito, b) Percentagem de agências por município, distrito de Lisboa e c) Percentagem de agências por município com sobreposição das agências, município de Lisboa..... | 27 |
| Figura 15 - Distribuição do negócio das agências da CGD em 2013/12 pelos a) distritos e b) municípios, <i>zoom</i> no distrito de Lisboa..... | 28 |
| Figura 16- Média do negócio em 2013 das agências da CGD, distrito de Lisboa..... | 29 |
| Figura 17 - Negócio por cliente em 2013/12, a) nos distritos e b) nos municípios, <i>zoom</i> no distrito de Lisboa..... | 30 |
| Figura 18 - Localização das agências face à densidade populacional. a) Portugal, b) distrito de Lisboa e c) município de Lisboa..... | 31 |
| Figura 20 - Município de Lisboa com o negócio 2013/12 de cada uma das agências. | 31 |
| Figura 20 - Ganho Mensal médio sob as Localização das agências da CGD. a) Portugal e b) distrito de Lisboa..... | 32 |

| | |
|--|----|
| Figura 22 - Número de clientes face ao número de agregados familiares, representação de vários níveis de detalhe. a) a nível dos distritos de Portugal, b) A nível dos municípios de Lisboa e c) Município de Lisboa com sobreposição das agências e com a representação do negócio. 33 | |
| Figura 22 - % Poder de compra vs Localização das agências da CGD. | 34 |
| Figura 23 - Interface de cálculo da matriz de distâncias, com <i>input</i> em coordenadas geográficas e os vários outputs possíveis, mapa com pontos assinalados e linhas com informação da distância, tempo e modo de deslocação entre dois pontos. | 36 |
| Figura 24 - Processo de preenchimento dos blocos dos destinos, análogo para origens. | 37 |
| Figura 25 - Processo do cálculo da matriz de distâncias. | 38 |
| Figura 26 - Função de cálculo da atratividade. Neste caso a atratividade é o NEGOCIO_201312 do <i>data.frame</i> das agências CGD. | 40 |
| Figura 27 - Processo do cálculo da % de Huff para agências da CGD. | 41 |
| Figura 28 - Processo de cálculo da % de Huff para agências da CGD e concorrência. Os passos com o fundo cinzento são novos face aos que constam na Figura 27. | 42 |
| Figura 29 - Processo do cálculo da % máxima de Huff. | 43 |
| Figura 30 - Desempenho do programa de cálculo das percentagens de Huff relativamente a um ponto de origem, e 5 a 400 destinos, em que a soma representa o tempo total, ou seja, tempo de cálculo do modelo de Huff e o carregamento das tabelas com a informação correspondente das agências. | 44 |
| Figura 31- Desempenho do programa de cálculo das percentagens de Huff relativamente a um ponto de origem, 5 a 500 destinos, em que a soma representa o tempo total, ou seja, tempo de cálculo do modelo de Huff e o carregamento das tabelas com a informação correspondente das agências. | 45 |
| Figura 32 - Processo para visualização dos dados obtidos da aplicação do modelo de Huff. | 45 |
| Figura 33 - Processo de filtragem por coluna no QGIS de uma tabela de atributos da camada 'MatrizDistanciasHuffCGD.csv'..... | 46 |
| Figura 34 – Passos para a criação de uma nova camada <i>Shape</i> com os dados resultantes da filtragem: a) Botão onde carregar para guardar os elementos selecionados, dando origem a b) janela que aparece para concluir a gravação dos elementos selecionados. | 47 |
| Figura 35 - Definição na representação da percentagem de Huff no mapa, escolha do atributo a ser representado pelo diagrama de texto..... | 47 |
| Figura 36 - Definição na representação da percentagem de Huff no mapa, escolha da posição do diagrama de texto. | 48 |
| Figura 37 - Definição na representação da percentagem de Huff no mapa, escolha do tamanho do diagrama. | 48 |

| | |
|---|----|
| Figura 38 – a) Município de Lisboa com a localização das agências da CGD representadas por bolas azuis e a localização de um cliente representado por uma estrela vermelha (mais ou menos a meio à direita), b) Equivalente a a) mas com um retângulo por cima da área de análise, que se podem observar em, c) Agências representadas por diagramas circulares e graduados com a % de Huff relativamente ao ponto selecionado..... | 49 |
| Figura 39 - Representação das agências da CGD por pontos a azul e do ponto de origem por uma estrela vermelha (aproximadamente no centro da figura). | 51 |
| Figura 40 – a) Seleção do raio de análise e b) Agências consideradas para análise a amarelo.. | 51 |
| Figura 41 - Cálculo da matriz de distâncias e resultados, para uma origem e três destinos..... | 52 |
| Figura 42 - Processo de análise da rede de agências..... | 53 |
| Figura 43 - Rede original de agências com assinalação da agência a eliminar da rede. a) Município de Lisboa com a localização das agências da CGD representadas por bolas azuis e a localização de um cliente representado por uma estrela vermelha (mais ou menos a meio à direita), b) Equivalente a a) mas com um retângulo por cima da área de análise, que se podem observar em c) Agências representadas por diagramas circulares e graduadas com a % Huff relativamente ao ponto selecionado, onde o X representa a agência que vai ser eliminada da área de análise e d) resultado da eliminação da agência representada por um ponto azul da rede (a mesma que tem a cruz em c)..... | 56 |
| Figura 44- Áreas de influência circular, raio de 1 km em linha reta. | 57 |

Lista de Tabelas

| | |
|---|----|
| Tabela 1 - Diferentes dimensões de áreas de influência. | 10 |
| Tabela 2 - Comparação entre os quatro modelos gravitacionais. | 11 |
| Tabela 3 - Exemplo de dados recolhidos da lista de agências bancárias do BP ^[8] | 14 |
| Tabela 4 - Exemplo de códigos dos municípios recolhidos dos CTT ^[9] | 15 |
| Tabela 5 - Exemplo de códigos postais recolhidos dos CTT ^[9] | 15 |
| Tabela 6 - Exemplo de dados fornecidos pela CGD com a localização e atributos de cada agência. | 15 |
| Tabela 7 - Exemplo de dados da CAOP 2014 recolhidos da DG Território ^[3] | 16 |
| Tabela 8 - Exemplo de dados de densidade populacional recolhidos da Pordata ^[4] | 16 |
| Tabela 9 – Informação usada para o mapa de localização das agências bancárias. | 25 |
| Tabela 10 - Percentagem de agências da CGD em Portugal. | 27 |
| Tabela 11 - Negócio das agências da CGD em 2013/12 | 28 |
| Tabela 12 – Negócio das agências da CGD por cliente. | 30 |
| Tabela 13 - Localização das agências sobre a densidade populacional. | 31 |
| Tabela 14 - Localização das agências da CGD sobre o ganho mensal médio..... | 32 |
| Tabela 15 - Quantidade de clientes das agências da CGD face ao número de agregados familiares..... | 33 |
| Tabela 16 - Localização das agências da CGG sobre a percentagem de poder de compra. | 34 |
| Tabela 17 - Ficheiro com o tipo de resultados obtidos através do cálculo da matriz de distâncias..... | 37 |
| Tabela 18 - Campos do ficheiro que resulta da aplicação do modelo de Huff. | 41 |
| Tabela 19 - Cenários de utilização considerados neste trabalho..... | 50 |

Capítulo 1 - Introdução

Este relatório descreve o trabalho realizado de desenvolvimento de uma metodologia de análise geográfica da rede de agências bancárias da Caixa Geral de Depósitos (CGD), no âmbito da disciplina de Projecto em Gestão de Informação do Mestrado em Gestão de Informação, ramo de Gestão e Análise de Dados.

1.1. Motivação

Num contexto de mudança constante dos fatores externos relevantes para a localização de uma agência bancária, como a concorrência, negócios, população, infraestruturas de comunicação e estado, e dos fatores internos das agências como o número de clientes das agências entre outros, é cada vez mais necessário fazer avaliações regulares e estruturadas ao desenho da rede de agências.

É, aliás, conhecido que “os três fatores básicos para o sucesso são: localização, localização, localização” ^[5,6]. Posto isto, é importante saber escolher o melhor local para instalar ou retirar uma agência, sendo desejável a utilização de ferramentas de apoio à tomada de decisão que possibilitem fazer análises de cenários alternativos e escolher os que maximizam os benefícios.

Sendo a localização um fator crítico para o sucesso de um negócio, deve ter-se em conta a distância de uma agência a pontos de interesse ou de grande concentração de potenciais clientes, tais como universidades, centros de negócios, centros comerciais, ou transportes. Quanto mais perto estiver a agência dos pontos de interesse, tendencialmente mais clientes irá conseguir captar, seguindo a lógica de que a atração de um cliente passa por três fatores principais: o princípio do menor esforço, a distância, e o tipo de bens ou serviços ^[7].

Relativamente à distância, esta pode não ser interpretada da mesma forma por todos e pode variar por diversos fatores. Por exemplo, o que pode parecer um caminho curto olhando para dois pontos num mapa em linha reta, sem se ter em consideração a rede viária, pode ser na realidade um caminho longo, se existir por exemplo um rio entre os dois pontos, que obrigue a travessar uma ponte, que pode não estar perto. Portanto, é desejável o conhecimento da rede viária para entender melhor a realidade dos percursos. Assim, mais do que assumir um caminho a direito a unir dois pontos no mapa, é pertinente acomodar estas variações do que pode ser considerado o percurso mais curto na análise da localização das agências.

Para ir ao encontro destas necessidades existem os modelos gravitacionais ^[7], os quais possibilitam a análise de uma rede de agências segundo vários critérios de distância, mas também a atratividade do local em que uma agência se encontra, que reflete outros fatores de decisão, nomeadamente externos, de modo a observar quais as agências melhores posicionadas.

É portanto importante conseguir visualizar dados relativos às agências num mapa geográfico, facilitar a sua análise e a tomada de decisão. Ao ser visualizada a rede de agências juntamente com os seus atributos, por exemplo o seu negócio ou a densidade populacional do local onde

estão inseridas, consegue ter-se uma visão mais contextualizada e aproximada à da realidade, podendo assim ser feitas interpretações que noutras circunstâncias não seriam possíveis.

1.2. Objetivos

Neste trabalho, pretende-se utilizar um Sistema de Informação Geográfica (SIG) para visualizar a rede de agências da CGD e da concorrência, e conseguir analisar a rede existente ou o impacto de fazer alterações à rede, isto é, suportar simulações de cenários de mudanças.

O trabalho de desenvolvimento da metodologia de análise geográfica da rede de agências da CGD foi organizado em torno de quatro objetivos principais:

1. **Conhecimento e transformação de dados** – Consistiu na análise dos dados disponibilizados pela CGD e na identificação de informação adicional proveniente de fontes externas na *Web*, por exemplo dados demográficos. Os dados das várias fontes tiveram de ser transformados para poderem ser utilizados e visualizados num SIG, e também para possibilitar o seu uso num modelo gravitacional. A avaliação deste objetivo foi feita segundo o enquadramento e utilidade dos dados na metodologia.
2. **Obtenção de matriz de distâncias realistas** – Reflete o interesse em obter dados realistas de deslocação, tendo em conta os modos a pé e de automóvel segundo caminhos existentes, bem como duas medidas de distância: os metros e o tempo para os percorrer. A avaliação deste objetivo passou pela capacidade de desenvolver uma aplicação informática para obter automaticamente, de um serviço na *Web*, uma matriz de distâncias para várias agências bancárias e posições de clientes, e formatar esses dados para que pudessem ser usados num modelo gravitacional.
3. **Aplicação de um modelo gravitacional** – Este objetivo compreendeu a escolha e aplicação de um modelo gravitacional capaz de combinar a informação interna e externa sobre a atratividade de cada agência (do Objetivo 1) e a matriz de distâncias realistas (do Objetivo 2) para estimar a possibilidade de um cliente num certo local se dirigir a cada agência, e tendo também em conta a concorrência. A avaliação foi feita considerando a produção dos resultados pretendidos e a facilidade de uso do modelo em vários cenários.
4. **Visualização de dados num SIG** – Corresponde ao interesse de possibilitar a visualização dos dados resultantes dos Objetivos 1 e 3 em mapas geográficos usando um SIG, de modo a facilitar a tomada de decisão. A avaliação deste objetivo incide sobre se a visualização permite compreender e comparar a importância das agências, e se permite tirar conclusões interessantes. É também realizada a apresentação de vários cenários de utilização da metodologia desenvolvida. Que têm por objetivo demonstrar algumas possibilidades de análise à rede de agências da CGD.

1.3. Principais resultados obtidos

Para o Objetivo 1 foram utilizadas duas ferramentas distintas: o LibreOffice Calc para fazer a transformação dos dados, tanto de dados externos recolhidos de diversas fontes (Banco de Portugal (BP) ^[8], Direção Geral do Território ^[3], Pordata ^[4], CTT ^[9]), como de dados internos fornecidos pela CGD; e o GPSVisualizer ^[10] para conversão das moradas oriundas da lista de agências extraída do BP em coordenadas geográficas, sendo esta necessária para a aplicação de um modelo gravitacional. Estes dados puderam ser usados e visualizados no sistema de informação geográfica QGIS, tendo sido alcançadas todas as propostas deste objetivo.

Relativamente ao Objetivo 2, foi desenvolvida uma aplicação em JavaScript que tira partido do serviço Distance Matrix da biblioteca Google Maps, para obter automaticamente uma matriz de distâncias, considerando percursos reais, duas medidas de distância, metros e minutos, e dois modos de deslocação, a pé e de carro. Para a obtenção desta matriz, é necessária informação sobre a localização dos clientes ou potenciais clientes e a localização das agências.

Para o Objetivo 3 foi utilizada a linguagem R para programar o cálculo do modelo gravitacional de Huff, o qual foi escolhido por aceitar uma matriz de distâncias realistas (em vez de assumir linhas a direito entre pontos) e por considerar de forma sistemática a concorrência. O programa desenvolvido permite a um decisor ajustar facilmente uma função de atratividade das agências e outros parâmetros do modelo, para obter como resultado a probabilidade de um cliente se deslocar a cada uma das agências em seu redor. Ainda relativamente a este objetivo, foram realizados testes de desempenho, que mostram que o cálculo da percentagem de Huff para as agências da CGD tem um tempo de execução linear até às 400 agências.

Face ao Objetivo 4, voltou a recorrer-se ao QGIS para a representação das agências num mapa geográfico, em que cada agência é rodeada por um círculo de tamanho proporcional à percentagem de Huff (sendo esta também mostrada em forma numérica). Pretendeu-se desta forma aumentar a rapidez de identificação da agência mais importante a partir da localização de um cliente. Consegue-se também a partir de cenários fazer simulações, por exemplo o encerramento de uma agência, que possibilitam a visualização do impacto de alterações na rede de agências.

1.4. Notação adotada

No texto principal deste documento todas as palavras que não sejam portuguesas são escritas em itálico. As sequências de passos necessárias para executar uma determinada tarefa usando uma ferramenta são apresentadas sob a forma de 1º passo -> 2º passo -> 3º passo, e assim sucessivamente.

1.5. Organização do documento

Este documento tem quatro capítulos principais. No Capítulo 1 é abordada a motivação, os objetivos do projeto e os principais resultados obtidos. No Capítulo 2 são explanados conceitos que foram importantes para a realização deste trabalho. No Capítulo 3 apresenta-se o trabalho

realizado relativamente aos quatro objetivos principais: realiza-se o conhecimento e depuração dos dados, é demonstrado o método de cálculo da matriz de distâncias, e do modelo gravitacional, e demonstro como podem ser visualizados os resultados num mapa geográfico. No Capítulo 4 descrevo quais as contribuições principais conseguidas com este trabalho, dificuldades encontradas, aptidões adquiridas, e propostas de trabalho futuro.

Capítulo 2 – Conceitos

Neste capítulo são expostos conceitos relevantes para uma melhor compreensão deste trabalho, incluindo o que é um SIG, uma análise comparativa de 4 modelos gravitacionais (para justificar a escolha do mais adequado), e ainda alguns tipos de áreas de influência.

2.1. SIG – Sistemas de Informação Geográfica

Um sistema de informação geográfica, SIG, é uma ferramenta que permite exibir, armazenar e manipular dados que estão distribuídos geograficamente, possibilitando a consulta, análise e o auxílio na tomada de decisão. A sua utilização tem crescido exponencialmente, pode ser utilizado em várias áreas, pois é composto por diversas ferramentas possibilitam realizar diversos tipos de análise, como análise geográfica, topológica e cartográfica, e até mesmo decisões relacionadas com o *marketing (geomarketing)* ^[11].

Os SIG são generalizações do mundo real ^[1,12] e podem ser utilizados em áreas, como por exemplo análises de áreas comerciais ^[13], planeamento de percursos ^[14,15] e identificação de áreas de influência ^[16]. Estas ferramentas permitem a análise de grandes áreas e produzir resultados quantitativos, a partir de relações entre *layers* ou camadas. Uma camada tem informação relativa a determinada entidade. Por exemplo as agências da CGD estão numa camada enquanto a concorrência, o CAOP2014 (limites administrativos de Portugal) estão noutras camadas diferentes ^[17].

A informação geográfica pode ser guardada sob duas formas distintas: *raster* ou vetorial. Quanto à informação na forma de *raster*, é relativa às propriedades do espaço, sendo constituído por células. O espaço é representado por uma grelha, que quanto mais fina for maior é o detalhe. Por exemplo uma grelha 10*10 apresenta menos detalhe que uma grelha 30*30 ^[12].

Quanto à informação vetorial, é focada na precisão da localização, e dispõe de três formas: linhas podem representar estradas e rios; pontos podem ser, por exemplo, as agências bancárias, e polígonos podem representar cidades, municípios ^[12].

É através das camadas com dados que é possível gerar mapas geográficos. Estas camadas são compostas por atributos, que podem ser a localização em coordenadas geográficas ou outra informação relativa à camada. Estes atributos são essenciais para a aplicação de filtros. Por exemplo, supondo uma camada relativa às agências, um dos atributos são as suas coordenadas, para visualizar a sua localização no mapa, e outro atributo pode ser o tamanho da agência; um filtro pode servir para visualizar aquelas que têm um tamanho superior a 20 m² ^[12].

Os SIG's possibilitam estabelecer relações geográficas entre os elementos num contexto geográfico, facilitando a resposta a algumas questões interessantes que, no caso do planeamento de expansão, redução ou remapeamento de uma rede de agências bancárias, podem ser ^[19]:

- Onde estão localizadas as agências bancárias da própria rede?
- Onde estão localizadas as agências bancárias da concorrência?

- Onde estão localizados os polos geradores de tráfego?
- Onde estão localizadas as áreas sem cobertura ou com sobreposição de atendimento?
- Onde se localizam as áreas de maior poder de compra da população?
- Onde se localizam as áreas de maior densidade demográfica?
- Onde se localizam as agências com maiores vendas?
- Quais as zonas que têm mais vendas em função do rendimento da população?

A apresentação de dados referentes às localizações das agências bancárias e postos de atendimento, tanto da própria rede como da concorrência, bem como, dados socioeconômicos, como ganho mensal, densidade populacional, entre outros, que quando disponíveis nas instituições bancárias se encontram apenas em forma de relatórios, mapas em papel, dificultam as análises geográficas. Pois tais que tais análises requerem a combinação de diversos dados. O uso de um SIG facilita as análises geográficas, uma vez que possibilita a combinação de dados de diferentes fontes, e ainda representá-los graficamente, de modo a permitir a sua visualização ^[11]. Sendo que algumas das maneiras de utilizar um SIG no caso do retalho do mercado bancário são: perfil de mercado, seleção de uma nova agência ou realocação de uma existente, entre outras ^[18].

2.2. Modelos gravitacionais

Os modelos gravitacionais são, por analogia, baseados na teoria da atração gravitacional de Newton: “dois corpos são atraídos entre si na proporção direta de suas massas e na proporção inversa ao quadrado da distância entre eles” ^[19].

De seguida, são apresentados quatro modelos, no entanto o primeiro modelo não é um modelo gravitacional mas sim um modelo que está na base dos modelos gravitacionais. É por isso relevante a sua presença. Quanto aos modelos restantes são efetivamente gravitacionais, mas têm um contexto de aplicação diferente, o que é explicado de seguida.

- **Christaller – Teoria do Lugar Central**

Christaller desenvolveu a teoria do lugar central, que é mais conhecida na ótica do crescimento urbano ^[19]. É utilizada para fazer previsões referentes ao número, tamanho e âmbito das cidades de uma região, tendo por base a extensão da análise de áreas de influência. Uma vez que os padrões de localização variam de setor para setor as áreas de influência também, dependendo da procura e de economias de escala.

A teoria do lugar central, a partir da conjugação de padrões de localização de diferentes sectores, apresenta a organização de um sistema regional de cidades ^[5], sendo uma cidade o centro de atividades de vendas a retalho. Os principais conceitos da teoria são os seguintes:

- Área de influência ou amplitude: distância máxima que os consumidores estão dispostos a percorrer para adquirir um produto ou serviço, correspondente ao limite máximo do ponto de vista da procura;

· Limiar do produto: distância mínima em torno de um ponto de venda que garante a viabilização da oferta, corresponde ao limite mínimo sob o ponto de vista da oferta;

Assim, quando a área de influência é maior que o limiar o produto é comercializado ^[5].

O objetivo principal desta teoria, considerando as cidades como centros de serviços, é explicar como se desenrola a organização geográfica das cidades e as suas áreas de influência, tendo depois dado origem a novos estudos sobre os fatores que influenciam a decisão do consumidor de escolher determinada organização, a chamada atratividade ^[20]. Surgiram então os modelos gravitacionais baseados nesta teoria e no menor esforço de locomoção ^[7].

- **Reilly**

William Reilly, com a sua lei da gravitação do comércio a retalho de 1929, foi o primeiro a mostrar a aplicação dos modelos gravitacionais no *geomarketing* ou *marketing* geográfico ^[1], adotando os modelos gravitacionais para analisar a atração de círculos comerciais para clientes ^[21]. A lei da gravitação de Reilly é mais focada no movimento populacional do que na localização de empresas ^[22].

Fundamentando-se na teoria gravitacional de Newton, Reilly desenvolveu a teoria da interação geral, uma teoria de movimento, e aplicou-a a diferentes tipos de análise, como movimento populacional, fluxo de tráfego, dimensionamento de área de influência e potencialidade de mercados. Segundo esta teoria para a determinação da atratividade de um centro relativamente a outro, existem dois eixos fundamentais: a população, que funciona como elemento de atração, e a distância, que funciona como repulsão. Logo, cidades maiores e de melhor acesso teriam uma maior força de atração sobre os clientes: “duas cidades atraem negócios de uma vila intermediária situada próxima dos limites de suas áreas de influência, em proporção direta das populações nas cidades e inversa dos quadrados das distâncias das cidades à vila intermediária” ^[22,23].

- **Converse**

Reformulações da lei de Reilly tornaram-na mais indicada para ser utilizada em estudos de centros comerciais. Surgindo assim, o modelo do ponto de equilíbrio de Paul Converse. O ponto de equilíbrio é representado pela distância em unidades métricas entre centros comerciais, onde o consumidor tanto optará por como por outro ^[1].

Neste modelo, a grande diferença relativamente ao modelo anterior foi a criação de um ponto de indiferença também conhecido como ponto de equilíbrio, ou seja, um local em que o cliente não tem preferência de consumo entre a loja A ou B ^[20,24]. Portanto, se o cliente estiver num local em que a distância é inferior ao ponto de equilíbrio relativamente a um dos centros comerciais (por exemplo, centro comercial A) significa que irá realizar compras no centro comercial A. O modelo permite assim delimitar áreas de influência comercial para os pontos em estudo ^[21].

$$\text{Distância à cidade B} = \frac{\text{Distância entre cidades}}{1 + \sqrt{\frac{\text{População da cidade A}}{\text{População da cidade B}}}}$$

Figura 1 - Fórmula de cálculo do modelo de Converse ^[21]

- **Huff**

David Huff, através do seu modelo, foi o primeiro a sugerir que as áreas de influência são complexas, contínuas e probabilísticas, ao contrário das áreas resultantes da teoria do lugar central e da Lei de Reilly, que eram geométricas sem sobreposição ^[2,1,25].

Este modelo permite calcular a probabilidade de um consumidor residente em determinada área de influência comprar ou ir a uma loja (ou agência bancária) específica.

Huff (1964) afirmou que o modelo por ele proposto deve ser ajustado mediante o tipo de negócio ^[5], sendo que no modelo original eram necessários os seguintes dados:

- Tamanho da loja em metros quadrados, considerada a variável de atração, com base no raciocínio de que quanto maior a loja, maior a variedade de produtos que ela oferece, e portanto mais atrativa ela se torna para o consumidor;
- Distância da loja até aos locais de origem dos consumidores, considerada a variável de repulsão ^[1], logo funcionando como um fator inibidor ^[25].

O modelo de Huff pode ser adaptado considerando diferentes medidas de atratividade e de repulsão, por exemplo, em vez da distância da deslocação ser utilizado o tempo de deslocação. Para o cálculo da atratividade, as variáveis a utilizar podem ser obtidas estatisticamente e recorrendo a diversos fatores, e daí fazer uma ponderação a cerca da sua importância ^[5]. Este modelo possibilita também a consideração da concorrência ^[26].

Relativamente ao contexto deste trabalho, este modelo pode ser utilizado para estudar o impacto da abertura de uma nova agência da CGD ou da concorrência. Pode também ser usado para a criação da área de influência de uma agência e para se perceber qual a agência que tem maior probabilidade de ser escolhida por um cliente ou potencial cliente. Portanto, para aplicar este modelo são necessários alguns dados sobre as agências e a localização de clientes.

Pode observar-se na Figura 2 a fórmula de cálculo da percentagem de Huff, que corresponde à probabilidade de um cliente num certo local ir a uma agência de referência. Sendo a agência de referência aquela para a qual o cliente pensa ir, num universo de várias.

$$\text{Percentagem de Huff} = \frac{(\text{Atratividade da agência de referência})^\alpha}{(\text{Distância do cliente à agência de referência})^\beta} \frac{1}{\sum_{i=1}^n \frac{(\text{Atratividade da agência } i)^\alpha}{(\text{Distância do cliente à agência } i)^\beta}}$$

Figura 2 - Fórmula de cálculo da percentagem de Huff.

Os parâmetros constantes na Figura 2 são os seguintes:

- A atratividade das agências, que pode ser a área da loja, o espaço de estacionamento, o número de clientes, ou o que se entender pertinente para a análise. Portanto, a atratividade pode ser uma multiplicidade de atributos relacionados com a agência^[27].
- A distância entre o cliente e as agências, que pode ser em metros ou em minutos.
- A, ou coeficiente de atratividade, estimado a partir de observações empíricas, tomando o valor de 1, por ser o mais usual^[28,21].
- B, ou Coeficiente de decaimento, que normalmente é 2^[28,21], podendo tomar valores entre 0 e 5. Quando $\beta > 2$ existe grande aversão dos consumidores em percorrer maiores distâncias para ir comprar num dado local. Assim, quanto maior for o β maiores são as probabilidades de Huff para as agências mais próximas e menores para as agências mais afastadas^[6].
- n, corresponde ao número de agências consideradas na análise.

Portanto, através dos modelos descritos, percebe-se a sua aplicabilidade, quais os mais simples ou complexos, podendo assim escolher-se aquele que se revelar mais adequado à análise que se pretenda realizar.

2.3. Áreas de influência

A área de influência representa aquilo que um cliente está disposto a percorrer para chegar a uma certa loja^[29], sendo um conceito fundamental nos estudos de localização comercial, pois reflete a dimensão geográfica da procura de mercado do retalho. Costuma-se defini-la como sendo a “área geográfica que contém a maior parte dos clientes de uma loja”^[25].

E existem diferentes metodologias para o cálculo de áreas de influência, umas mais simples de aplicar do que outras, gerando resultados um pouco diferentes principalmente a nível visual. A dimensão da área de influência é influenciada por diversos fatores como, por exemplo, a densidade populacional^[2]. Algumas destas metodologias de cálculo de áreas de influência são: circular, percentagem de clientes, tempo/distância de deslocação, e o modelo de Huff.

Relativamente à área de influência circular, é simples de aplicar, consistindo na criação de um círculo em torno da agência com um raio pré definido. No entanto, este método apresenta algumas limitações como o facto de não ter em consideração a concorrência ou limites geográficos^[30,31,32]. Tradicionalmente, neste tipo de áreas de influência, o raio é calibrado utilizando resultados de inquéritos a clientes^[33].

Quanto à área de influência por percentagem de clientes, a área é delimitada tendo em conta a percentagem de clientes que determinada loja tem. Por exemplo, a área em redor da loja pode ser obtida por forma a conter 60% dos clientes. Este método é mais preciso que o circular mas é um modelo descritivo e não preditivo^[30,31]. Neste modelo de área de influência são consideradas três subdivisões: a área primária, a área secundária, e a área terciária cada uma com a sua percentagem de clientes, como se pode observar na Tabela 1.

Tabela 1 - Diferentes dimensões de áreas de influência.

| | Área Primária | Área Secundária | Área terciária |
|---|---------------|-----------------|----------------|
| Applebaum (1966) ^[5,2,25,29] | 60% - 75% | 15% - 25% | ≈ 10% |
| Applebaum (1976) ^[7,1] | 50% - 70% | 20% - 30% | 10% - 20% |

A área de influência calculada através do tempo ou distância de deslocação tem o intuito de perceber o impacto da acessibilidade da loja tanto seja através de estradas, de carro ou caminhos pedestres. Por norma este tipo de método gera áreas de influência irregulares pois tem em consideração a rede viária ^[30,31,32].

Por último, a área de influência delimitada pelo modelo de Huff, que é um modelo gravitacional como já foi descrito a cima. Essencialmente tem por objetivo mapear qual a probabilidade de clientes em determinados locais irem para uma certa loja, tendo em conta diversos fatores, nomeadamente a presença de concorrência. Este modelo poderá ser útil na visualização do impacto da alteração de uma loja tanto a nível de atributos, por exemplo o seu tamanho ou lugares de estacionamento, como a nível de localização ou de produtos e serviços ^[30,31].

2.4. Discussão de conceitos

O processo de decisão sobre qual a melhor localização de um ponto comercial, é muito importante em qualquer negócio, logo deve ser estudado cuidadosamente ^[11]. Para uma análise mais fiel deve utilizar-se um modelo capaz de incluir diversas variáveis como: a proximidade a meios de transporte, características topográficas, concorrência, demografia, influências culturais. Pois são variáveis identificadas na literatura que causam impacto na decisão de compra do consumidor ^[20].

Sendo o modelo de Huff adequado/preparado para incluir vários fatores na sua análise como a concorrência ^[26], características populacionais, entre outros, e descrevendo também uma área de influência de acordo com a probabilidade de atrair clientes para um determinado negócio ^[21].

Ainda relativamente ao modelo de Huff foi efetuado um estudo por Lima e Teixeira (2008), sobre a sua utilização na tomada de decisões envolvendo dados geográficos, cujo objetivo era testar a eficiência do modelo na escolha da localização da loja de preferência e na identificação da área de influência. Ao comparar as respostas obtidas diretamente dos clientes quanto à loja de preferência, com a loja que teria maior probabilidade obtida através do modelo de Huff, observou-se que os resultados obtidos eram semelhantes. A área de influência permite identificar a região em torno da loja onde está localizada a maioria dos clientes, esta foi identificada através das preferências dos entrevistados e calculada com o modelo de Huff, sendo que também neste caso o modelo revelou-se bastante adequado ^[6].

Foi realizada uma comparação entre os quatro modelos descritos no ponto 2.3, tendo como objetivo perceber as suas diferenças relativamente ao contexto de utilização de cada um como se pode observar pela Tabela 2.

Tabela 2 - Comparação entre os quatro modelos gravitacionais.

| Autor | Concorrência? | Múltiplos fatores? | Área de influência | Foco |
|----------------------------|----------------------|---------------------------|-------------------------------------|--|
| Christaller [20] | Não | Não | Cidades como centros de serviços | Organização geográfica das cidades |
| Reilly [22] | Não | Não | Geométrica sem sobreposição | Movimento da população |
| Converse [1] | Não | Não | Geométrica sem sobreposição | Estudos de centros comerciais |
| Huff [5] | Sim | Sim | Complexa, contínua e probabilística | Probabilidade de um cliente num local ir a uma loja de referência. |

Pode observar-se na Tabela 2 que o modelo de Huff tem maior versatilidade de utilização, podendo incluir mais informação, como considerar a concorrência e possibilitar a utilização de diversas medidas tanto de atratividade como de repulsão (medidas de distância em metros ou segundos considerando caminhos reais em vez de distância medida em linha reta). É também o que apresenta as características mais úteis para o tipo de análise que se pretende realizar no projeto documentado neste trabalho.

2.5. Sumário do capítulo

Deste capítulo pode-se perceber alguns conceitos relativamente aos modelos gravitacionais, de onde se pode observar que o modelo de Huff aparenta ser o mais indicado para o caso em estudo. É também exposto o conceito de área de influência, e apresentam-se algumas aplicabilidades dos SIG nomeadamente na área do mercado bancário, das quais algumas são apresentadas no capítulo seguinte.

Capítulo 3 - Trabalho realizado

Neste capítulo são apresentadas as ferramentas utilizadas no ambiente de trabalho, e explicada a forma como foram concretizados os objetivos do projeto, sendo expostos todos os passos necessários para os alcançar. No final, são demonstrados vários cenários de aplicação da metodologia desenvolvida.

3.1. Ambiente de trabalho

No decorrer deste projeto, no sistema operativo Windows 7, recorreu-se à utilização de diversas ferramentas. Uma das ferramentas utilizadas foi o LibreOffice Calc^[34], tendo sido necessário para trabalhar a informação relativa a bancos da concorrência. Para o processamento destes dados também foi usado o GPSVisualizer^[35] para converter moradas em coordenadas geográficas.

Foram utilizadas ferramentas *open source*, incluindo o sistema de informação geográfica (SIG) QGIS^[10], o R^[38] e o LibreOffice Calc^[34], tendo este último, na versão 3.6, servido como auxiliar para processar tabelas de dados para serem utilizadas pelo QGIS. Relativamente ao R, na versão 3.0.3, foi usado para a concretização do Objetivo 3, ou seja, a aplicação do modelo de Huff.

Quanto ao QGIS, na versão 2.6.1, este é composto por diversos módulos que podem ser instalados de acordo com as necessidades do utilizador. Para este projeto foi considerado pertinente instalar o módulo PinPoint^[10], para que fosse possível gravar um ou mais pontos num mapa. Pois, para a aplicação do modelo de Huff é necessário ter informação sobre a distância entre clientes e agências, portanto este módulo vai permitir que se tenha acesso às coordenadas de um ou mais pontos escolhidos no mapa, que pode ser referente a clientes ou agências.

Para a realização do Objetivo 2, foi utilizado o serviço Distance Matrix da biblioteca Google Maps^[39], sob a forma de JavaScript num *browser* Web, pois permite a obtenção de percursos reais entre dois pontos, considerando dois modos de deslocação e duas medidas de distância, que é informação necessária ao Objetivo 3.

Tanto para a realização do Objetivo 1, para a conversão de moradas em coordenadas geográficas, como do Objetivo 2, para o cálculo da distância entre dois pontos foi também necessário utilizar também o serviço JavaScript da biblioteca Google Maps.

3.2. Concretização dos objetivos

Neste capítulo descreve-se a metodologia aplicada para alcançar os objetivos propostos (Figura 3). Na Subsecção 3.2.1 é feita a descrição dos dados, das transformações efetuadas aos dados, e ainda da sua representação usando um SIG. Nas subsecções seguintes é explicado o processo da obtenção da matriz de distâncias, do cálculo da percentagem de Huff, e da sua representação num mapa.

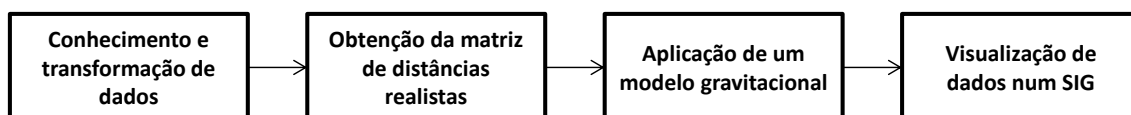


Figura 3 – Etapas da metodologia proposta.

Na Figura 3 podem observar-se as etapas definidas na metodologia proposta, as quais estão de acordo com os objetivos propostos no Capítulo 1.

3.2.1. Conhecimento e transformação de dados

Na realização deste projeto, começou-se por analisar os dados enviados pela CGD, que outros dados estavam disponíveis para recolha e quais os necessários para a execução dos objetivos seguintes. A CGD forneceu os dados internos, sobre as suas agências. Foram extraídos dados de várias fontes externas, aos quais foram aplicadas transformações de modo a poderem ser utilizados por um SIG. As fontes de dados externas foram: o Banco de Portugal (BP) ^[8], a Pordata ^[4], a Direção Geral do Território ^[3], e os CTT ^[9].

De seguida são descritos sucintamente quais os dados extraídos de cada fonte e o processo de transformação por que passaram:

- **BP:** Foi obtida a lista de agências bancárias ^[8] de vários bancos, de onde foram extraídos os endereços postais das várias agências dos bancos, tendo estes passado por um filtro do qual só ficaram para consideração os que têm mais de dez agências. Após esta seleção de dados realizou-se a passagem das moradas das agências para coordenadas geográficas através do GPSVisualizer ^[35], que utiliza dados fornecidos pela Google.

Tabela 3 - Exemplo de dados recolhidos da lista de agências bancárias do BP ^[8].

| Denominação da Instituição | Tipo de Estabelecimento | Cód. AF | Denominação do Estabelecimento | País de Estabelecimento | Cód. Agência | Denominação da Agência | Sucursal/Balcão sede | Morada | Localidade | Código Postal | Concelho | Distrito | País |
|---|--------------------------|---------|---|-------------------------|--------------|--|----------------------|--|------------|------------------|----------|----------|----------|
| AFTAB CURRENCY EXCHANGE LIMITED | Rede | 8832 | AFTAB CURRENCY EXCHANGE LIMITED - SUCURSAL EM PORTUGAL | Portugal | 8832 | AFTAB CURRENCY EXCHANGE LIMITED - SUCURSAL EM PORTUGAL | Sim | RUA ARCO MARQUES DO ALEGRETE, LOJA 4F, PALACIO DOS ABOIM | LISBOA | 1100-034 LISBOA | Lisboa | Lisboa | Portugal |
| AGÊNCIA DE CÂMBIOS - J.R. PEIXE REI & COMPANHIA LIMITADA (SUCESSORES) | Rede | 742 | AGÊNCIA DE CÂMBIOS - J.R. PEIXE REI & COMPANHIA LIMITADA (SUCESSORES) | Portugal | 1 | SEDE - PORTO | Sim | RUA RAMALHO ORTIGÃO, 10 | PORTO | 4000-407 PORTO | Porto | Porto | Portugal |
| AGÊNCIA DE CÂMBIOS CENTRAL, LDA | Rede | 832 | AGÊNCIA DE CÂMBIOS CENTRAL, LDA | Portugal | 1 | SEDE - SETUBAL | Sim | AV LUISA TODI, 226 | SETUBAL | 2900-452 SETUBAL | Setúbal | Setúbal | Portugal |
| AX MARKETS LIMITED | Escritório Representação | 176 | AX MARKETS LIMITED PORTUGAL | Portugal | 176 | AX MARKETS LIMITED PORTUGAL | Sim | AVILA BUSINESS CENTER, AVENIDA JOÃO CRISÓSTOMO, N.º 30-5.º | LISBOA | 1050-127 LISBOA | Lisboa | Lisboa | Portugal |

A Tabela 3 mostra algumas linhas da lista de agências obtida do BP. Para cada agência, uma por linha, há informação referente à sua localização (as últimas seis colunas), ao tipo de estabelecimento (coluna dois), que pode ser 'Rede' ou 'Escritório de Representação', tendo sido apenas consideradas as agências do tipo 'Rede'. A coluna oito indica se a agência é Sucursal ou Balcão sede e existem ainda colunas referentes à denominação da instituição, do estabelecimento e da agência.

- **CTT:** Foram recolhidos os códigos postais com os respetivos municípios e distritos de modo a poderem-se associar as agências da CGD aos municípios (ver Tabela 4).

Tabela 4 - Exemplo de códigos dos municípios recolhidos dos CTT ^[9].

31;10;São Vicente
31;01;Calheta (Madeira)

A informação retirada dos CTT vem em três ficheiros de texto, em que um se refere aos códigos dos distritos, outro aos dos municípios (na Tabela 4) e o terceiro tem códigos postais (ver Tabela 5).

Quanto à Tabela 4 pode observar-se que é composta por três colunas, em que a primeira diz respeito ao código do distrito, a segunda ao código do município e a terceira ao nome do município. A tabela referente aos códigos é semelhante a esta mas tem apenas duas colunas, uma com o código e outra com o nome do distrito.

Tabela 5 - Exemplo de códigos postais recolhidos dos CTT ^[9].

01;01;249;Alcafaz;;;;;;;;;;3750;011;AGADÃO
01;01;250;Caselho;;;;;;;;;;3750;012;AGADÃO

Relativamente à Tabela 5, esta é composta por diversas colunas, em que as consideradas mais importantes foram as duas antes da última, que correspondem ao código postal de 4 mais 3 dígitos, bem como as duas primeiras colunas, que são referentes aos códigos do distrito e do município, respetivamente.

- **CGD:** Foram considerados dados sobre as suas agências, que associados aos dados dos CTT, podem ser agrupados em dois níveis de detalhe distintos: distrito e município. Destes dados foram ainda escolhidos os indicadores a considerar na metodologia proposta: o negócio (NEGOCIO_201312) e a quantidade de clientes (Q_CLIENTES_201312) em 2013 (ver Tabela 6).

Tabela 6 - Exemplo de dados fornecidos pela CGD com a localização e atributos de cada agência.

| C_AG | NM_AG | CODIGO_PO STAL | LATITUDE | LONGITUDE | NEGOCIO_20 1112 | NEGOCIO_20 1212 | NEGOCIO_20 1312 | NEGOCIO_P A_201112 | NEGOCIO_P A_201212 | NEGOCIO_P A_201312 | Q_CLIENTES_ 201112 | Q_CLIENTES_ 201212 | Q_CLIENTES_ 201312 |
|------|-----------------------------|-------------------|-----------|-----------|--------------------|--------------------|--------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 1 | CENTRAL SEDE - LISBOA | 1000-300 | 38.741528 | -9.141671 | 1097666 | 1377392 | 1804864 | 612033 | 951558 | 1259273 | 3997 | 4130 | 4181 |
| 3 | ABRANTES | 2200-358 | 39.462352 | -8.197930 | 332277 | 351875 | 347997 | 20762 | 33245 | 32256 | 4115 | 4011 | 3841 |

A primeira coluna (C_AG) corresponde ao código único de cada agência, as quatro colunas seguintes referem-se à localização de cada agência, e as restantes colunas são atributos de negócio de cada agência (na Tabela 6 aparecem apenas alguns dos atributos de negócio disponibilizados).

- **DG Território:** Foram extraídos dados relativos às áreas administrativas de Portugal continental (CAOP 2014) em formato shapefile, que fornecem informação sobre a organização de Portugal a nível da freguesia (ver Tabela 7). Estes são os dados base para fazer agregações de informação em dois níveis de detalhe: distrito e município.

Tabela 7 - Exemplo de dados da CAOP 2014 recolhidos da DG Território ^[3].

| DICOFRE | Freguesia | Município | Distrito | TAA | Area_EA_Ha | Area_T_Ha | Des_Simpli | Loc_Postal |
|---------|-----------|---------------|----------|-----------------|------------|-----------|------------|------------|
| 081504 | Sagres | VILA DO BISPO | FARO | ÁREA SECUNDÁRIA | 0.04 | 3436.91 | Sagres | Sagres |
| 081504 | Sagres | VILA DO BISPO | FARO | ÁREA SECUNDÁRIA | 0.17 | 3436.91 | Sagres | Sagres |

Portanto pela Tabela 7 pode verificar-se o formato dos dados extraídos da DG Território, em que cada linha corresponde a uma freguesia distinta. Para cada freguesia tem-se o seu código (DICOFRE), o distrito e o município onde se insere, a sua área (Area_EA_Ha) e o tipo de área, que pode ser secundária ou principal. Pelo que é possível representar esta informação num mapa recorrendo a um SIG.

- **Pordata:** Foram extraídos indicadores ao nível do município, como a densidade populacional (na Tabela 8), poder de compra, e ganho mensal. Estes dados foram cruzados com os recolhidos da DG Território, o que possibilitou a agregação da informação nos níveis de detalhe distrito e município, e permitiu a sua visualização num SIG.

Tabela 8 - Exemplo de dados de densidade populacional recolhidos da Pordata ^[4].

| Territórios | | N.º médio de indivíduos por Km ² | |
|-------------------|------------------|---|------|
| Âmbito Geográfico | Anos | 2001 | 2013 |
| Município | Melgaço | 41.9 | 37.2 |
| Município | Monção | 94.3 | 89.6 |
| Município | Paredes de Coura | 69.1 | 65.4 |

Os dados recolhidos desta fonte de dados apresentam-se na Tabela 8 em que para cada município (linha) possui o valor do indicador escolhido em diferentes anos (colunas).

Após a apresentação geral dos dados e das transformações efetuadas aos mesmos, segue-se uma explicação mais detalhada das mesmas.

3.2.1.1. Conversão das moradas das agências em coordenadas geográficas

É importante a representação das agências num mapa pois permite, por exemplo, visualizar zonas com maior ou menor densidade de agências, tanto da CGD como da concorrência, sendo também fundamental para a aplicação do modelo de Huff saber onde estão localizadas as agências, pois só assim é possível calcular as distâncias entre agências e os clientes.

Portanto, numa primeira fase, com o objetivo de se poder representar num mapa as agências da CGD e da concorrência, foi necessário conhecer as respetivas coordenadas geográficas (latitude, longitude). Relativamente às agências da CGD, as coordenadas já se encontravam no ficheiro disponibilizado, contrariamente às agências da concorrência que constavam na lista de agências bancárias ^[35] do BP.

Pode observar-se na Figura 4 uma síntese do processo de conversão das moradas das agências da concorrência em coordenadas geográficas.

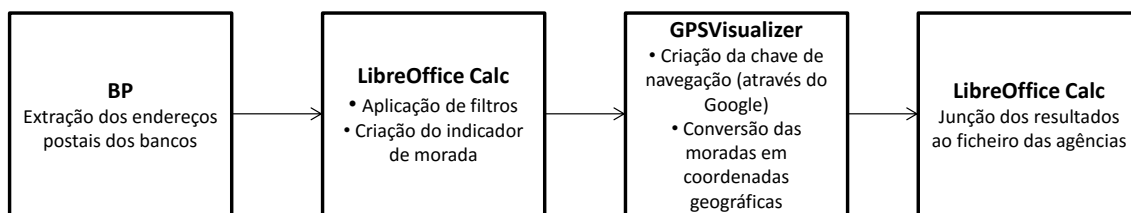


Figura 4 - Processo de conversão das moradas da concorrência em coordenadas geográficas.

Tendo em conta que os dados extraídos do BP têm a informação da morada em colunas diferentes, ou seja, código-postal numa, rua noutra e assim sucessivamente (Tabela 3), para que a conversão fosse a mais correta possível, ou seja, evitar que a coordenada obtida seja referente a um local distinto do procurado, pois existem por exemplo ruas com nomes iguais em diferentes distritos, é por isso importante utilizar o máximo de informação possível, foi criado um identificador da morada correspondente. Este identificador resulta da junção de toda a informação referente à morada da agência é um exemplo de chave de morada: “RUA ARCO MARQUES DO ALEGRETE, LOJA 4F, PALÁCIO DOS ABOIM|LISBOA|1100-034 LISBOA|Lisboa|Lisboa|Portugal”.

Este identificador é depois usado no GPSVisualizer^[35], que converte moradas em coordenadas geográficas. Esta ferramenta utiliza dados da Google, sendo que necessita de uma chave para poder ser usada a Google Maps JavaScript API^[36].

Assim, para a utilização do GPSVisualizer deve-se, na página inicial, escolher ‘Geocode Addresses’ e de seguida ‘2. Geocode multiple addresses’. Para obter a chave, deve primeiro fazer-se *login* na Google Developers Console, criar um novo projeto, ativar a Google Maps JavaScript API, e, por último, criar uma nova chave de navegador (*Browser Key*). Este processo é explicado na página que aparece após carregar em ‘2. Geocode multiple addresses’, à direita onde aparece ‘You can choose from three different sources of coordinates: [Bing Maps](#), [MapQuest Open](#), or [Google](#)’, clicando em Google.

Após a inserção das moradas (identificador de morada) e da chave do projeto Google nos campos de texto respetivos escolhe-se a opção ‘Start geocoding’, e quando o processamento terminar, faz-se clique em ‘create a GPX file’, um ficheiro GPX apresenta-se em formato texto em XML, que guarda conjuntos de pontos^[37]. De seguida, seleciona-se ‘Return to the “convert” form’ para que os resultados possam ser guardados num ficheiro de texto para que se fizesse a correspondência entre os identificadores de morada inseridos com os que constam no ficheiro das agências dos bancos concorrentes, de modo a atribuir a cada uma das agências as suas coordenadas geográficas.

3.2.1.2. Agregação da informação em dois níveis de detalhe distintos – distrito e município

Com o objetivo de representar dados sobre agências e população em mapas geográficos, revelou-se pertinente considerar dois níveis de detalhe, município e distrito, pois assim seria possível visualizar a informação de uma maneira mais global (distritos) e mais ao pormenor (municípios). Para que as agregações fossem executadas com sucesso, ou seja, sem haver perda de informação, foi necessário ter em atenção alguns fatores devido à existência de várias fontes, tais como os nomes de ligação e o preenchimento correto dos atributos. Os dados precisam de

ser unidos através de campos com significado comum, os quais, em concreto, são os mostrados entre parênteses na Figura 5. Para este processamento dos dados foi utilizado o LibreOffice Calc, pois esta ferramenta reconhece o formato DBF usado na CAOP da DG Território.

Das transformações e uniões efetuadas resultaram dois ficheiros que correspondem aos dados da CAOP unidos aos dados populacionais e aos dados das agências da CGD, cada um relativo a um nível de detalhe, ou seja, um dos ficheiros diz respeito à informação dos distritos e o outro à dos municípios. Estamos agora em condições de utilizar o SIG escolhido, que neste caso foi o QGIS 2.6.1, para fazer a agregação de acordo com o nível de detalhe pretendido.

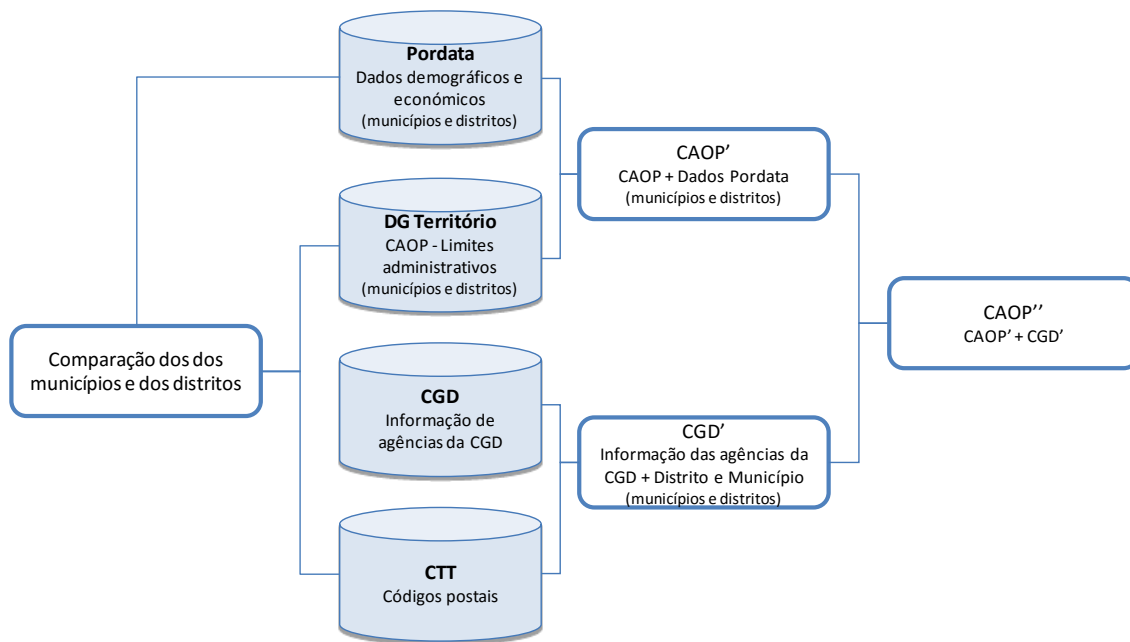


Figura 5 - Diagrama da ligação das fontes de dados utilizadas.

De modo a explicar como foram realizadas as transformações e uniões entre os diversos dados, passo a explicar como foi executada cada uma das etapas para cada caso. Sendo que para esta manipulação dos dados foi utilizado o LibreOffice Calc ^[34].

- **União dos dados da Pordata com os dados da DG Território**

O primeiro detalhe a ter em conta neste caso foi se os nomes dos municípios eram iguais nas duas fontes, neste caso coincidiam mas foi necessário passar os nomes dos municípios da Pordata para maiúsculas. O passo seguinte foi unir as tabelas utilizando a função PROCV(), a qual permite procurar um valor comum em duas tabelas e obter um dado valor de uma coluna pretendida. Este é o processo utilizado para unir o CAOP aos dados da Pordata a nível do município, pois pretende-se que todas as linhas de um dado município tenham a mesma informação, para quando são agregados no QGIS não haver perda de informação.

No caso dos distritos, existe uma transformação antes da união aos dados de ambas as fontes, ou seja, para cada distrito é somada a informação dos municípios correspondentes. Assim quando agregados no QGIS a informação resultante ser correta, ou seja, cada distrito fica com a sua informação correspondente. Como resultado desta união tem-se CAOP', que possui os

dados da CAOP aumentados com a informação da Pordata sobre a população. Desta união resultam dois ficheiros distintos cada um correspondente a um nível de detalhe, distrito e município. Portanto, foi um processo diferente para cada nível de detalhe, em que ambos tinham a mesmos dados base.

- **União dos dados da CGD com os dados dos CTT**

A razão desta união é o facto de os dados da CGD não possuírem diretamente informação sobre qual o município ou distrito onde estão as agências. Ou seja, sabe-se onde estão as agências, pois tem-se acesso às suas coordenadas geográficas e código-postal, mas não se têm colunas com os municípios e distritos as quais são necessárias para posteriormente unir a informação das agências ao CAOP'. A primeira tarefa realizada foi a verificação dos nomes dos dados (municípios e distritos) que se pretendem unir, compararam-se os nomes relativos aos municípios e distritos dos CTT com os do CAOP.

Através do código postal, e utilizando novamente a função PROCV(), pôde-se procurar os códigos de distrito e município correspondentes às agências da CGD no ficheiro dos CTT. Tendo guardados estes códigos em duas colunas extra no ficheiro da CGD, para se acrescentar a coluna dos nomes dos distritos volta a ser utilizada a função PROCV(). No entanto para acrescentar a coluna dos municípios foi necessário concatenar os códigos de distrito obtidos no passo anterior aos códigos de município, pois um município não é identificado apenas por um valor mas sim pela junção do seu código ao código de distrito. Depois dos valores estarem concatenados procurou-se esse valor na tabela referente aos municípios para se ter acesso aos seus nomes, utilizando a mesma lógica que nos passos anteriores. Como resultado tem-se CGD', que corresponde à informação da CGD acrescida de duas colunas importantes uma referente aos nomes dos municípios e outra aos nomes dos distritos.

- **União dos dados da CGD' com os dados da CAOP'**


Esta união teve como intuito poder-se representar mais globalmente a informação relativa à CGD, permitindo representar, por exemplo, não só o número de clientes de uma agência, como também o número de clientes nesse município, e relacionar com a população. Nesta fase só se juntou a informação relativamente às agências, deixando-se de parte a informação sobre sua localização (morada, latitude, longitude).

Foi necessária a transformação prévia dos dados para esta união, ao nível de município a informação foi obtida através da soma da informação das agências de um dado município. Por exemplo, a informação sobre o número de clientes do município de Lisboa é a soma dos clientes de todas agências no município de Lisboa. A mesma lógica foi aplicada aos distritos, cada distrito corresponde à soma de informação dos seus municípios. O resultado designa-se CAOP'', que são duas tabelas como toda a informação – CAOP, Pordata, CGD – uma com informação a nível do município e outra a nível do distrito, que podem ser agregadas consoante o nível de detalhe pretendido respetivamente e representadas pelo QGIS.

- **Agregação da informação do CAOP”**

Originalmente a informação da CAOP é fornecida até ao nível da freguesia, o que permite visualização de informação neste nível de detalhe. Portanto, que para fazer a representação de informação a nível do município não seja o seu nome é necessário que a informação esteja associada ao município, ou seja, que cada município deve corresponder a uma linha da tabela, e essa linha conter informação sobre vários indicadores a nível do município. Analogamente, aplica-se o mesmo processo ao nível do distrito.

Tendo os dados preparados para se realizar a agregação no QGIS, segue-se uma demonstração do processo realizado para a agregação num nível, por exemplo distrito, sendo análogo para o outro nível:

1. Carregaram-se os novos dados no QGIS:
 - a. Quando o formato do ficheiro é ‘shape (.shp)’, como é o caso dos dados da CAOP, clicando em ‘Adicionar Camada Vetorial’ , daí aparece uma janela onde se escolhe o ficheiro pretendido, como mostrado na Figura 6;

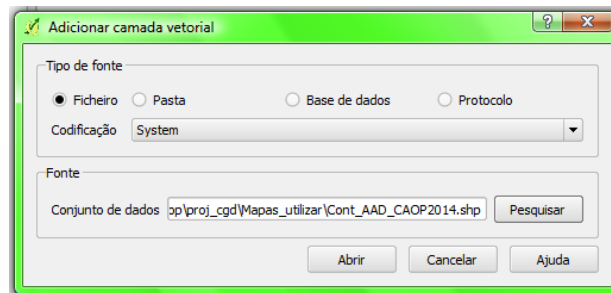



Figura 6 - Adição de uma camada vetorial no QGIS.

Neste caso será utilizada esta opção visto os dados serem trabalhados sobre uma camada *shape*, esta é também a alternativa utilizada para os ficheiros resultado das uniões descritas acima.

- b. Quando o formato do ficheiro é ‘valores separados por virgulas (.csv)’, como no caso da informação das agências bancárias, clicando em ‘Adicionar Camada de Texto Delimitado’ , de onde surge uma janela como a da Figura 7.

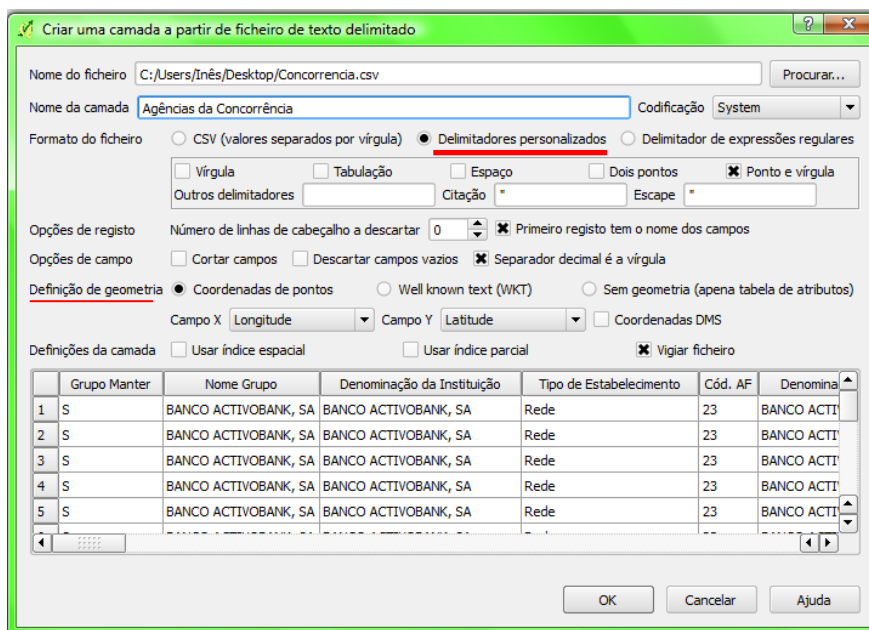


Figura 7 - Adição de uma camada de texto delimitado no QGIS.

Nesta janela, escolhe-se o tipo de separação de valores existente no ficheiro, sendo melhor escolher a opção ‘delimitadores personalizados’, assinalado com uma linha carregada na Figura 7. Também deve ser definido o tipo de geometria em ‘Definição de geometria’, assinalado com uma linha fina na mesma figura, de acordo com o ficheiro, se diz respeito a um ficheiro com coordenadas de pontos com o intuito de representar a sua localização ou se não tem geometria sendo por exemplo uma tabela de informação auxiliar, e de seguida clicou-se em ‘OK’.

- Na barra do menu clicou-se em ‘Vector’ -> ‘Ferramentas de geoprocessamento’ -> ‘Agregar...’, que abriu a janela na Figura 9;

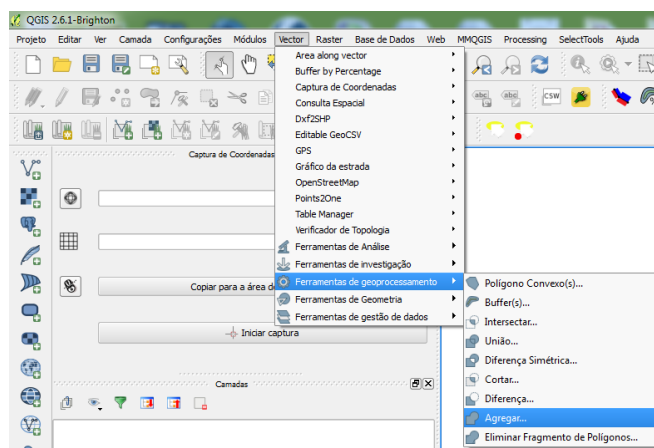


Figura 8 - Passos executados para abertura da janela de agregação.

A Figura 8 mostra como é a interface na execução dos pontos anteriores, sendo que na figura aparecem os pontos 2, 3 e 4.

3. A janela obtida a partir do ponto anterior pode ser observada na Figura 9. Aí escolheu-se a camada a agregar neste caso foi a camada resultante das uniões executadas acima a nível do distrito, o campo de agregação, que vai ser neste caso o neste caso a coluna 'Distrito' pois o objetivo foi agregação por distrito e o ficheiro de saída, que corresponde ao ficheiro resultante da agregação por distrito.

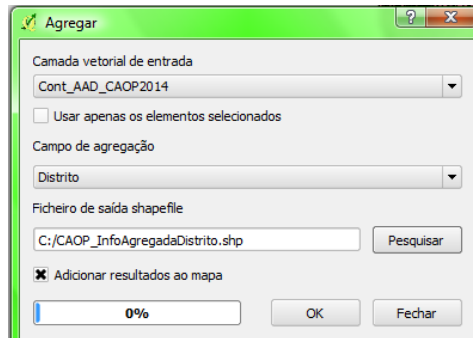


Figura 9 - Janela onde é escolhida a camada, o campo de agregação e o ficheiro de saída.

Para a inserção do campo 'Ficheiro de saída shapefile' clicou-se em 'Pesquisar' na Figura 9, de seguida escolheu-se o local onde o ficheiro foi guardado como se pode visualizar na Figura 10.

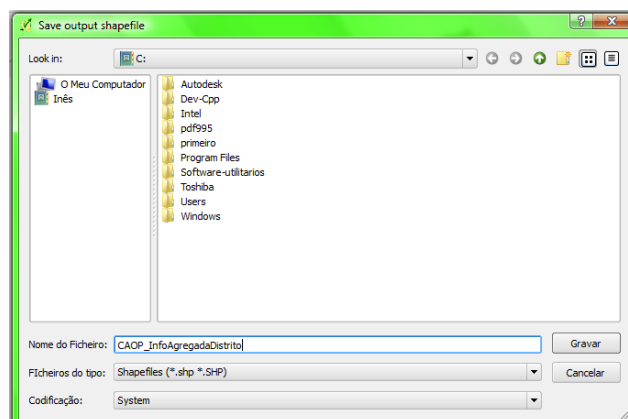


Figura 10 - Escolha do local e do nome do ficheiro resultante da agregação é guardado.



Após o nome e o local onde o ficheiro resultante da agregação escolhidos, clicou-se em gravar onde é visualizado na Figura 10.

4. Pressionou-se 'ok' na Figura 9.

Portanto desta agregação resultam os dois ficheiros já referidos anteriormente, em que um corresponde a uma tabela com toda a informação por distrito, ou seja, em cada linha da tabela existe um distrito distinto e cada coluna representa um indicador distinto. O outro ficheiro contém uma tabela semelhante, mas neste caso cada linha é referente a um município diferente.



Figura 11 - Resultado da agregação por distrito, para o distrito de Lisboa. a) Representação original por freguesia, b) Agregação por distrito e c) Resultado após limpeza da agregação.

Como se pode observar através da Figura 11, a agregação nem sempre origina resultados perfeitos, portanto é necessário fazer uma limpeza dos elementos sobrantes visíveis na Fig.10b). Para realizar esta tarefa de limpeza no QGIS o processo é o seguinte: ‘Vector’ -> ‘Ferramentas de geometria’ -> ‘Multipartes para partes simples’, depois entrar em modo edição clicando no lápis , adicionar um elemento clicando em , e seleccionar no mapa os elementos sobrantes a eliminar. Após os elementos seleccionados basta juntá-los ao elemento que queremos que se mantenha clicando em ‘editar’ -> ‘juntar elementos seleccionados’ -> marcar o campo que se pretende juntar -> ‘ok’. A Figura 11c) mostra o resultado final deste processo de limpeza.

Toda a informação que se encontra disponível diz respeito ao distrito. Portanto perde-se informação relativa aos outros níveis, sendo por isso importante, as transformações anteriores.

3.2.1.3. Cenários de recarregamento de dados para a agregação da informação em dois níveis de detalhe distintos – distrito e município

Após este processo, com as alterações de informação ou acesso a nova informação, por exemplo da população, deve ser possível atualizar a informação de modo a ser representada. Tendo em conta tais assunções, vou expor três cenários de alterações a realizar de modo a manter atualizada a informação:

Cenário 1: Existe informação nova sobre as agências da CGD. Não é necessário executar todo o processo, sendo apenas preciso retomar ao ponto ‘**União dos dados da CGD com os dados CTT**’ e seguir os passos sequencialmente como descrito em cima.

Cenário 2: Novos dados interessantes na Pordata. Neste caso é necessário refazer o ponto ‘**União dos dados da Pordata com dados do DG Território**’. Podendo-se seguir logo para o ponto ‘**Agregação da informação do CAOP**’, uma vez que não foram feitas alterações aos dados da CDG. O que acontece é a substituição da informação dos indicadores da Pordata ou o seu incremento.

Cenário 3: Atualização da informação da CAOP. Para este cenário é preciso verificar as mudanças existentes e o impacto nos cálculos a realizar, sendo provavelmente necessário repetir todas as transformações a partir do primeiro ponto, pois podem ter existido, por exemplo, junção de municípios.

Para que se possa fazer uma análise no QGIS com informação sobre as agências têm de se considerar mais dois ficheiros: um diz respeito à localização das agências da concorrência e o outro à localização das agências da CGD com os respetivos campos de cada agência. Estes campos são indicadores que podem ser representados graficamente e utilizados nas fases seguintes do trabalho.

3.2.1.4. Representação da informação disponível para o projeto

De seguida, são apresentados vários mapas relativos a Portugal continental, através dos quais é possível conhecer visualmente os dados, e relacionar as agências da CGD com outros fatores através de indicadores simples ou complexos.

Os indicadores simples dizem respeito a dados relativos apenas a agências da CGD, enquanto os indicadores complexos aos dados internos das agências ponderados com dados externos, populacionais ou da concorrência. Através destes mapas representativos de indicadores pode-se, por exemplo, auxiliar a identificação de zonas com uma cobertura mais elevada ou menos elevada de agências face à concorrência.

Para estes indicadores podem ser definidos objetivos genéricos por “intervalo de valores” ou específicos para cada agência ou ainda objetivos a nível dos distritos, município ou freguesia. Por exemplo, pretende-se saber quais os distritos ou municípios em que a CGD tem mais de 10% do total das agências.

O tipo de informação obtido através destes mapas (da Figura 12 até à 22) pode também ser útil para a escolha de uma zona a analisar com mais detalhe. Por exemplo, através do modelo de Huff pode ser relevante saber em que município uma agência se insere e as suas características, o ganho mensal (ver Tabela 14 e Figura 20). Posto isto, de seguida são apresentados os mapas com a representação deste tipo de informação, onde para cada um dos mapas existe uma tabela (ver Tabelas 9 à 16) que explica a informação apresentada.

Tabela 9 – Informação usada para o mapa de localização das agências bancárias.

| | |
|--------------------------|--|
| Nome | Localização das agências |
| Tipo de indicador | Simple |
| Objetivo | Visualizar a localização das agências da CGD e da concorrência |
| Metodologia | Mapeamento das agências da CGD a azul e das agências da concorrência a laranja. E <i>zoom</i> com as agências da concorrência identificadas por instituição bancária |
| Fontes | DG Território, BP e CGD |

Na Tabela 9 encontra-se um breve resumo da descrição das figuras seguintes, a partir desta tabela percebe-se o tipo de dados que são utilizados e como são relacionados.

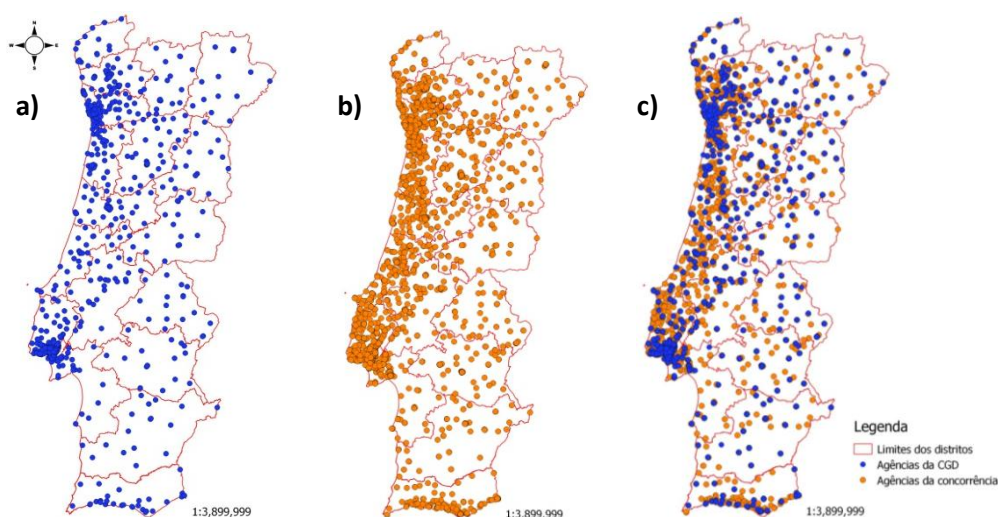


Figura 12 – Representação da localização de agências no mapa de Portugal. a) Diz respeito a Portugal só com as agências da CGD representadas, b) São representadas apenas as agências da concorrência, e c) Sobreposição de a) b).

Através da Figura 12 pode visualizar-se a distribuição tanto das agências da CGD como da concorrência por Portugal, onde se observa que no litoral é onde existem mais agências. A Figura 13 diz respeito ao município de Lisboa, na qual também se pode visualizar a localização das agências, com a particularidade de que as agências da concorrência em vez de se encontrarem todas representadas a laranja, recebem uma cor diferente por instituição.

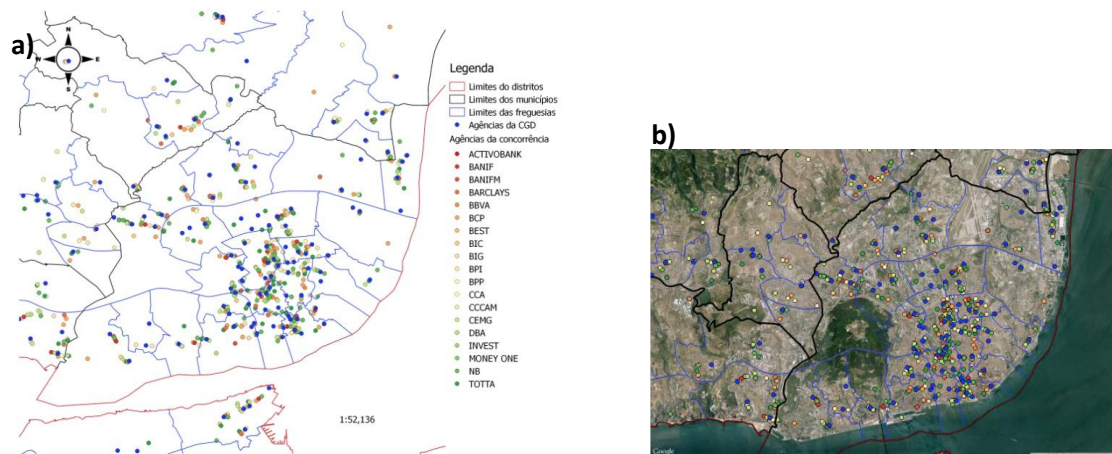


Figura 13 - Mapa do município de Lisboa com a diferenciação da concorrência. a) Representação de todas as agências e b) Sobreposição de a) com Google Maps.

Tabela 10 - Percentagem de agências da CGD em Portugal.

| | |
|--------------------------|---|
| Nome | Percentagem de agências da CGD face à concorrência |
| Tipo de indicador | Complexo |
| Objetivo | Estudar a cobertura das agências face à concorrência |
| Metodologia | N° de agências da CGD / (N° de agências da CGD + N° agências da concorrência) |
| Fontes | DG Território, BP e CGD |

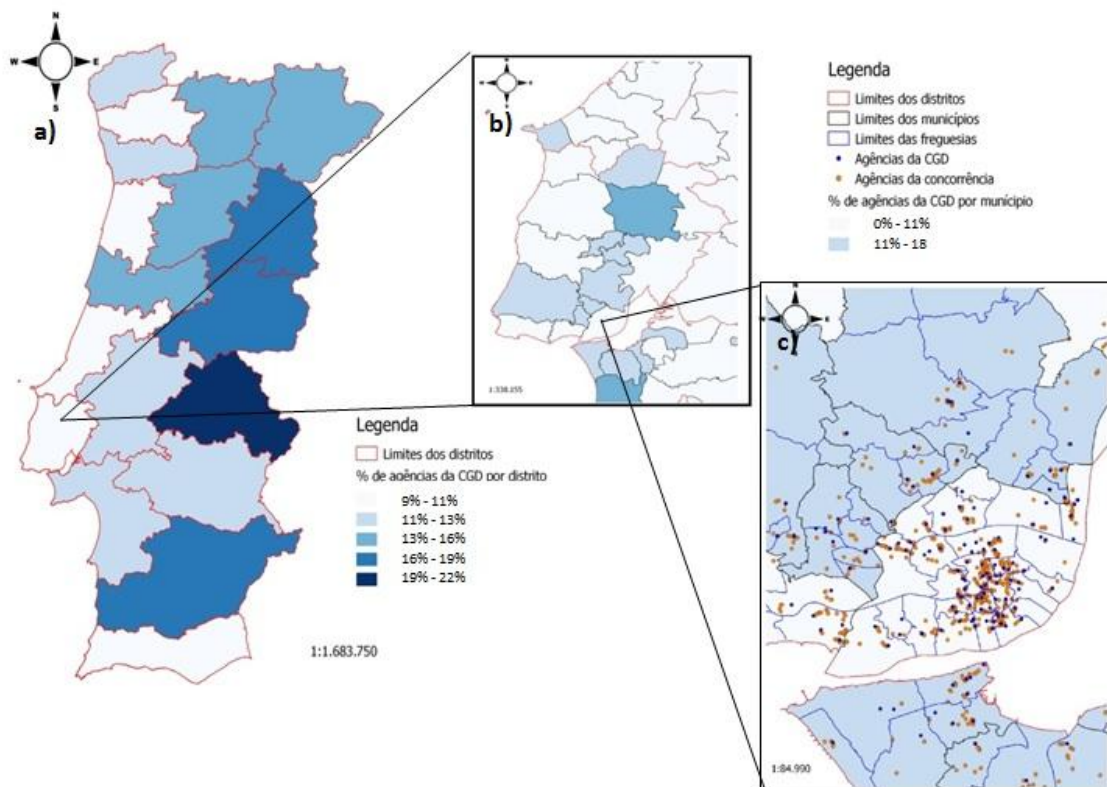


Figura 14 - Percentagem de agências da CGD por distrito e município com sobreposição da localização das agências. a) Percentagem de agência por distrito, b) Percentagem de agências por município, distrito de Lisboa e c) Percentagem de agências por município com sobreposição das agências, município de Lisboa.

Na Figura 14 é possível ver quais os municípios com uma maior percentagem de agências da CGD, verificando-se que é no interior que a percentagem é mais elevada face ao litoral.

Tabela 11 - Negócio das agências da CGD em 2013/12

| | |
|--------------------------|---|
| Nome | Negócio em 2013/12 por distrito e município |
| Tipo de indicador | Simple |
| Objetivo | Estudar a distribuição do negócio agregado das agências ao nível do distrito e do município. Classificar quais estão acima ou abaixo da média |
| Metodologia | Agregação do negócio em 2013/12 das agências de acordo com os dois níveis de detalhe – distrito e município. |
| Fontes | DG Território e CGD |

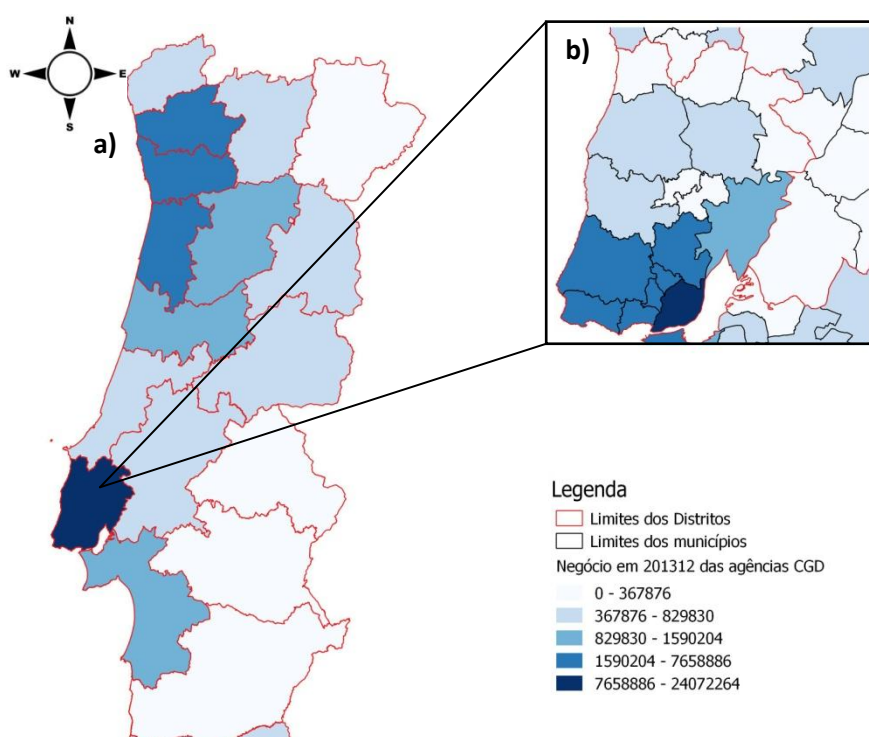


Figura 15 - Distribuição do negócio das agências da CGD em 2013/12 pelos a) distritos e b) municípios, zoom no distrito de Lisboa.

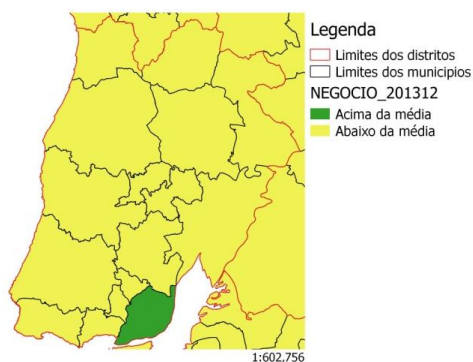


Figura 16- Média do negócio em 2013 das agências da CGD, distrito de Lisboa.

Através da Figura 15 pode visualizar-se a distribuição do negócio da CGD por Portugal, onde em Lisboa se verifica um valor mais elevado. Quanto à Figura 16 observa-se que apenas o município de Lisboa se encontra acima da média, e que todos os restantes municípios de Lisboa abaixo.

Tabela 12 – Negócio das agências da CGD por cliente.

| | |
|--------------------------|---|
| Nome | Negócio em 201312 por cliente por município |
| Tipo de indicador | Simple |
| Objetivo | Estudar a distribuição do negócio por cliente, agregados a nível do distrito e município, ver qual a área em que se gera mais negócio por cliente |
| Metodologia | Cálculo do $\text{Negócio}_{201312} / \text{Q_Clientes}_{201312}$ e agregação em dois níveis de detalhe - distrito e município |
| Fontes | DG Território e CGD |

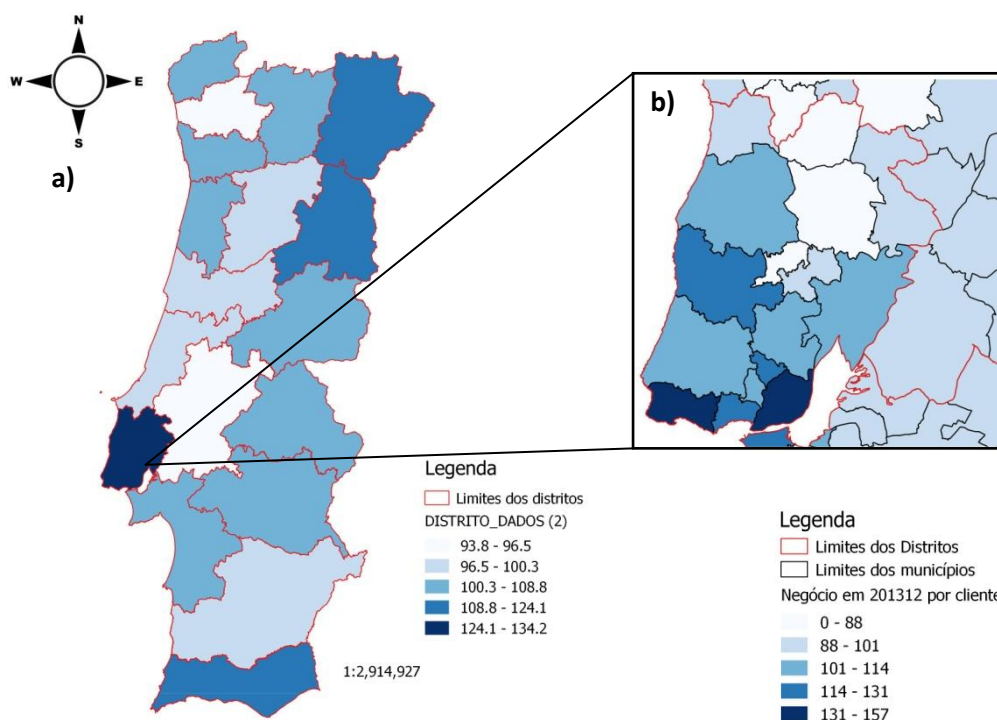


Figura 17 - Negócio por cliente em 2013/12, a) nos distritos e b) nos municípios, zoom no distrito de Lisboa.

Na Tabela 12 tem-se a explicação do objetivo e obtenção da Figura 17a), na podem visualizar-se os distritos em que se gera mais negócio por cliente, de onde sobressai o distrito de Lisboa. Quanto aos municípios do distrito de Lisboa (ver Figura 17b)) é gerado mais negócio por cliente em Lisboa e Cascais.

Tabela 13 - Localização das agências sobre a densidade populacional.

| | |
|--------------------------|---|
| Nome | Localização das agências face à densidade populacional |
| Tipo de indicador | Complexo |
| Objetivo | Verificar a relação entre o número de agências e a densidade populacional nos dois níveis de detalhe – distrito e município |
| Metodologia | Visualização da densidade populacional nos diferentes níveis de detalhe e sobreposição da localização das agências |
| Fontes | DG Território, Pordata, CGD e BP |

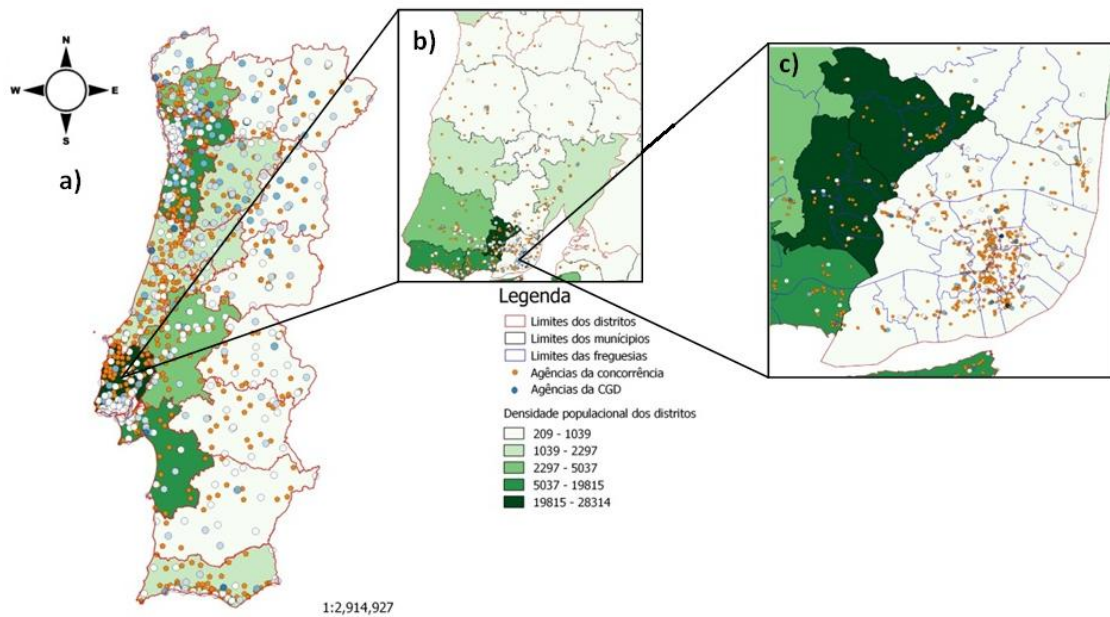


Figura 18 - Localização das agências face à densidade populacional. a) Portugal, b) distrito de Lisboa e c) município de Lisboa.

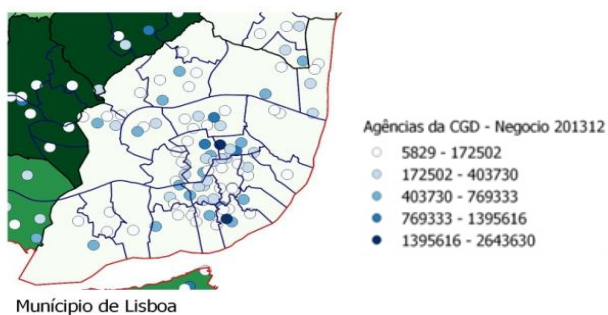


Figura 19 - Município de Lisboa com o negócio 2013/12 de cada uma das agências.

Na Figura 18c) pode observar-se que apesar do município de Lisboa não ser um dos municípios com densidade populacional mais elevada, é aquele que apresenta maior número de agências relativamente aos circundantes. Quanto à Figura 19 visualiza-se o negócio em 2013 de cada uma das agências no município de Lisboa.

Tabela 14 - Localização das agências da CGD sobre o ganho mensal médio.

| | |
|--------------------------|--|
| Nome | Ganho Mensal sob a Localização das agências da CGD |
| Tipo de indicador | Complexo |
| Objetivo | Observar a distribuição das agências face a ganho medio mensal nos municípios |
| Metodologia | Visualização do ganho mensal médio por municípios sobrepostos pela localização das agências da CGD |
| Fontes | DG Território, Pordata e CGD |

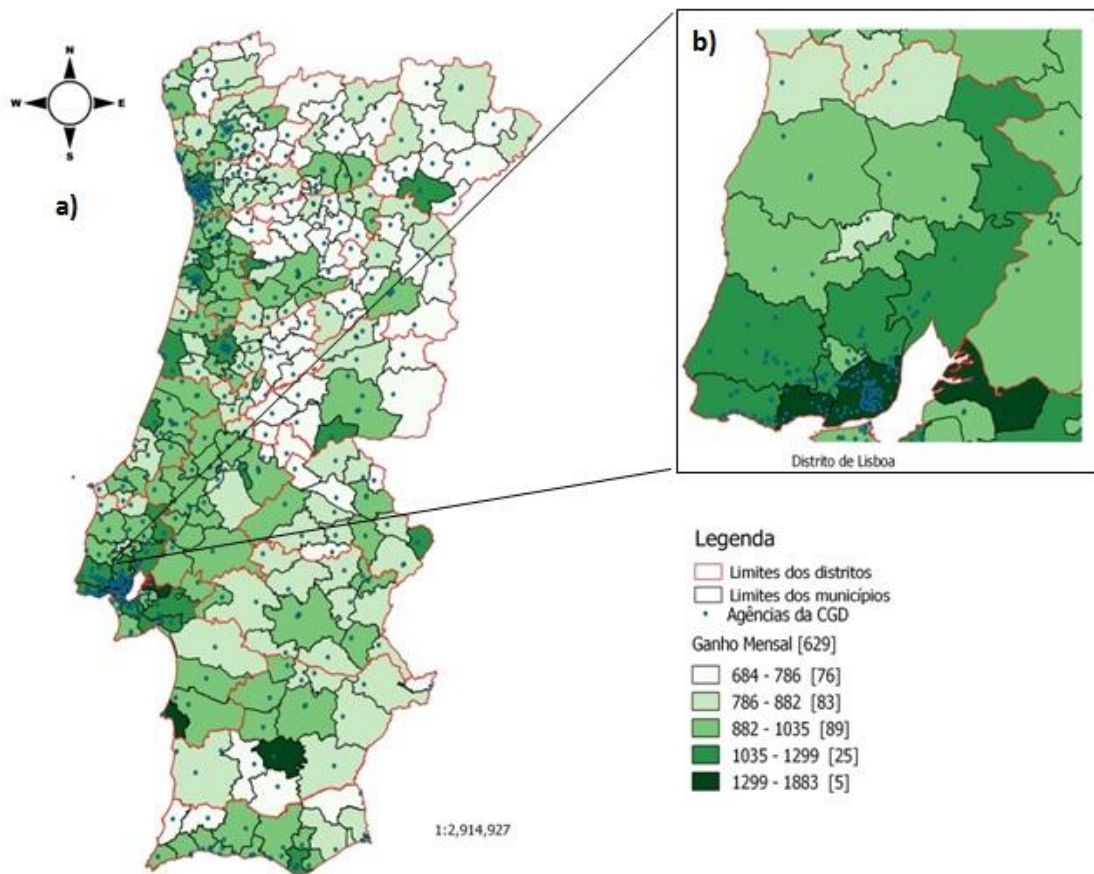


Figura 20 - Ganho Mensal médio sob as Localização das agências da CGD. a) Portugal e b) distrito de Lisboa.

Na Figura 20 pode ver-se que existe uma maior concentração de agências da CGD nas áreas onde o ganho médio mensal é mais elevado, principalmente no Porto e em Lisboa.

Na página seguinte, a Tabela 15 refere-se à explicação do mapa visualizado na Figura 21, através da qual se pode observar o número de clientes das agências da CGD a dividir pelo número de agregados familiares nos dois níveis de detalhe, distrito e município, sendo mais elevado no interior norte e menor no sul. Quanto ao distrito de Lisboa, o município de Lisboa foi aquele que apresentou valores mais elevados.

Tabela 15 - Quantidade de clientes das agências da CGD face ao número de agregados familiares.

| | |
|--------------------------|---|
| Nome | Quantidade de clientes face ao número de agregados familiares |
| Tipo de indicador | Complexo |
| Objetivo | Observar a distribuição das agências e seu negócio face à relação entre o número de clientes e o número de agregados familiares |
| Metodologia | Divisão do número de clientes das agências pelo número de agregados familiares e a sua visualização sobreposta pela localização das agências da CGD com possibilidade de visualização do negócio de cada agência. |
| Fontes | DG Território, Pordata e CGD |

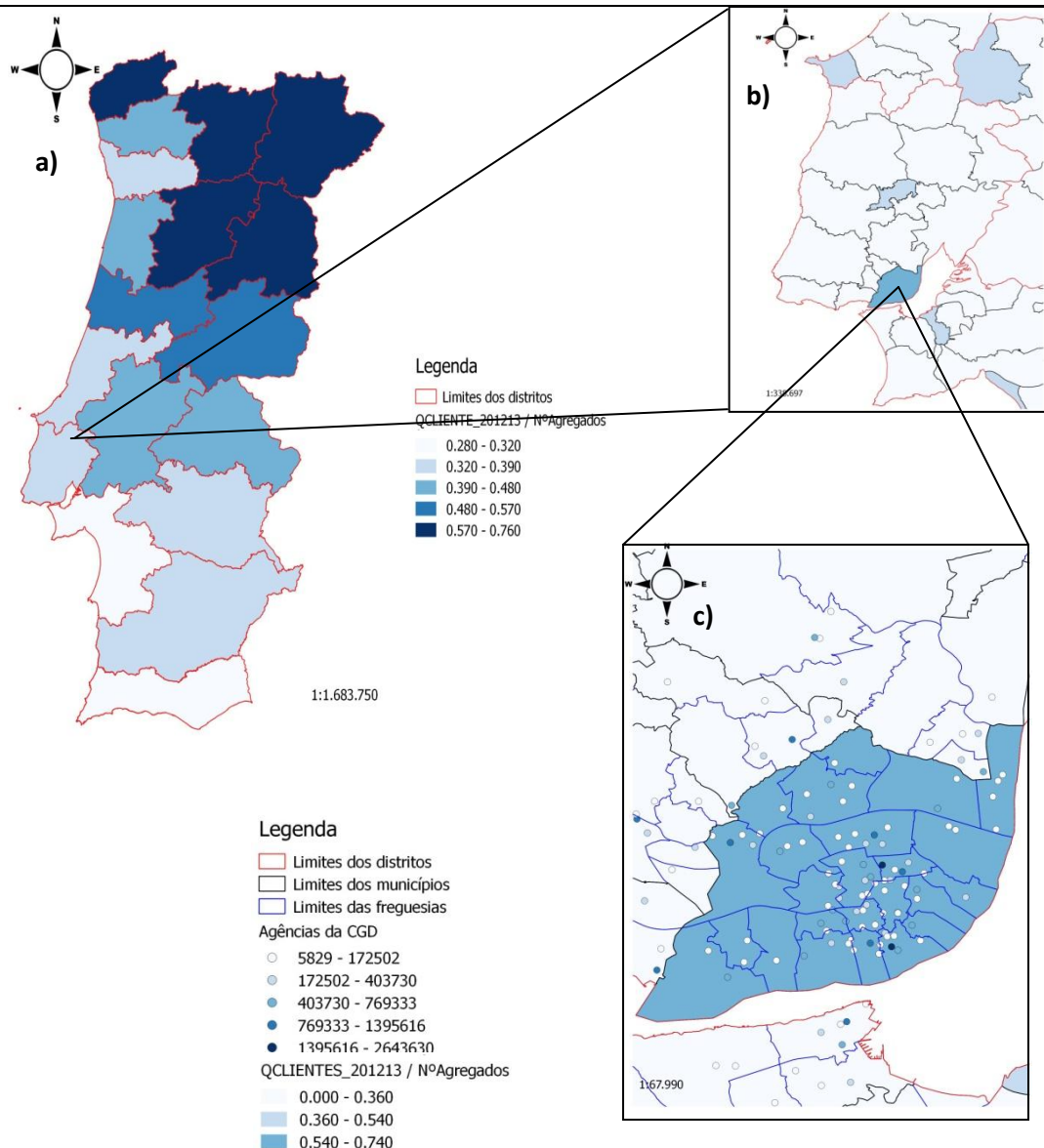


Figura 21 - Número de clientes face ao número de agregados familiares, representação de vários níveis de detalhe. a) a nível dos distritos de Portugal, b) A nível dos municípios de Lisboa e c) Município de Lisboa com sobreposição das agências e com a representação do negócio.

Tabela 16 - Localização das agências da CGG sobre a percentagem de poder de compra.

| | |
|--------------------------|---|
| Nome | Localização das agências da CGD sobre % Poder de compra |
| Tipo de indicador | Complexo |
| Objetivo | Observar a distribuição das agências face ao poder de compra nos municípios |
| Metodologia | Visualização do poder de compra por municípios sobrepostos pela localização das agências da CGD |
| Fontes | DG Território, Pordata e CGD |

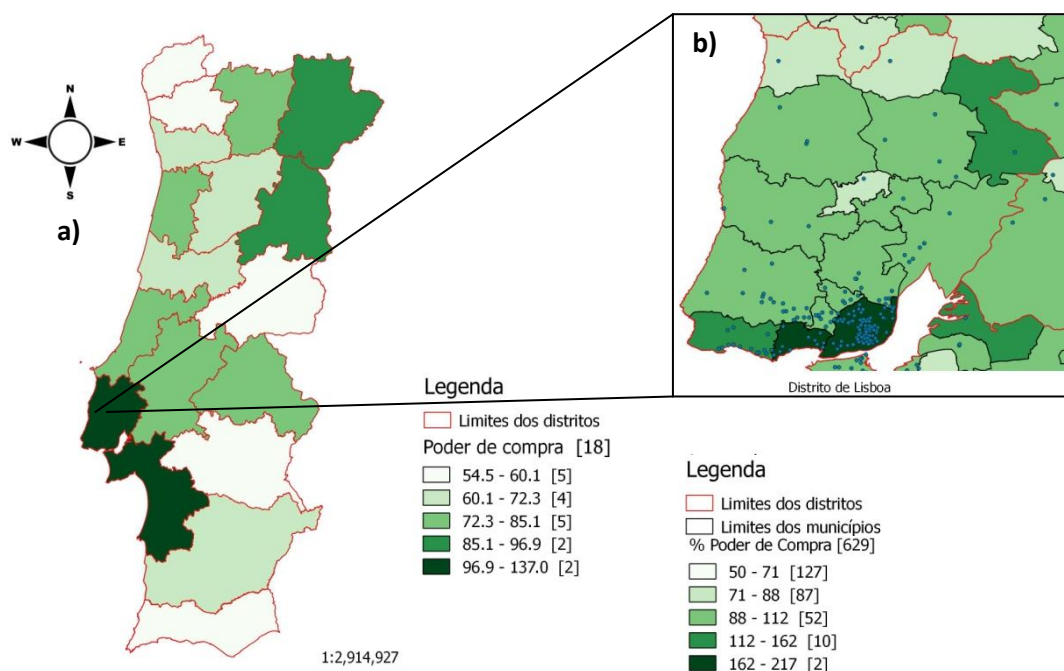


Figura 22 - % Poder de compra vs Localização das agências da CGD.

Pela Figura 22 observa-se que município de Lisboa é aquele que apresenta uma maior percentagem de poder de compra, e também uma maior concentração de agências da CGD.

É então possível visualizar a informação através deste tipo de mapas temáticos/geográficos, podendo tirar-se informações relevantes, como por exemplo, quais as zonas com maior ou menor densidade de agências ou qual a zona que está pior classificada segundo algum parâmetro ou que agência está a ter mais lucro que as restantes. Este tipo de informação pode ser condutor para outros tipos de análise, por exemplo, focar a atenção numa zona problemática.

3.2.2. Obtenção da matriz de distâncias

O objetivo do cálculo desta matriz é a recolha de informação necessária – distâncias entre dois pontos - à aplicação do modelo de Huff, e representa a distância entre clientes e agências. Portanto, são consideradas as origens como a localização de clientes ou potenciais clientes, e os destinos referem-se às agências. Esta matriz de distâncias dispõe de duas medidas de distância, metros e minutos de percursos realistas entre dois pontos, são ainda contemplados dois modos de deslocação, a pé e de carro. Para a criação da matriz de distâncias foi utilizado o serviço Distance Matrix da biblioteca Google Maps ^[39].

- **Funcionamento do serviço Distance Matrix da Google Maps JavaScript API:** Este serviço fornece distâncias e tempos entre dois pontos num mapa, funciona da seguinte forma: inserem-se os pontos de origem e os pontos de destino manualmente nos seus vetores no código do tutorial em JavaScript, da google Distance Matrix, e de seguida abre-se o ficheiro.html no *browser web*; aí carrega-se no botão *'calculate distances'*. São retornados num mapa os pontos correspondentes às origens a amarelo e os pontos correspondentes aos destinos a vermelho. Este serviço calcula as distâncias das várias origens para os vários destinos.

Quanto aos dados de entrada—origens e destinos—estes podem ser introduzidos na forma de coordenadas geográficas ou de morada, sendo a sua nomenclatura distinta: no caso das moradas é só inseri-las no vetor, mas com coordenadas é necessário complementar com código semelhante a `new google.maps.LatLng(38.741528, -9.141671)`.

Nota: Na versão gratuita, o serviço Distance Matrix da biblioteca Google Maps possui limites de utilização, só permitindo que sejam utilizados 100 elementos por iteração/*query* e por cada 10 segundos, os elementos correspondem a combinações de origens e destinos, outra restrição é só permitir 25 destinos ou origens de cada vez, para as combinações não ultrapassarem o limite de 100 por iteração pode-se fazer uma pesquisa de por exemplo, 4 origens para 25 destinos.

- **Automação do cálculo da matriz de distâncias:** Para que fosse possível automatizar o cálculo das distâncias de modo a gerar a matriz de distâncias mais rápida, versátil e facilmente, procedeu-se a alterações no código do tutorial do serviço Distance Matrix, recorrendo a um manual de JavaScript ^[40] as principais alterações foram:
 - Criação de um vetor com todas origens e respetivo bloco
 - Criação de um vetor com todos os destinos e respetivo bloco
 - Criação de um vetor com o código de todos os destinos, que é o indentificador dos destinos
 - Criação de um vetor com os modos de deslocação
 - Entrada dos dados através de formulário

As vantagens destas alterações foram essencialmente o aumento da facilidade e versatilidade e rapidez do cálculo da matriz de distâncias. A criação dos vetores com todas as origens e destinos permite fazer-se a inserção de qualquer número de origens e destinos, ou seja, pode-se

inserir por exemplo, o dobro do limite permitido pela API. Este processo é facilitado pela criação do formulário que evita fazer alterações no código claro que ao permitir-se a existência de um maior número de origens ou destinos que o permitido foi imperativa a criação de uma sequência de validações de preenchimentos dos dados de cada pedido ao servidor Distance Matrix para que não se ultrapassem os limites impostos e tudo funcione devidamente. Sendo que através dos blocos criados (ver Figura24) é pesquisado apenas o número de elementos permitidos por *query*.

A criação do vetor com os códigos dos destinos é importante pois torna possível a identificação do destino após a sua utilização pela API para posterior tratamento dos dados, uma vez que a sua identificação através da morada ou ponto geográfico pode ficar condicionada devido ao facto de os nomes no *output* aparecerem diferentes dos colocados inicialmente. Sendo depois esta informação necessária nomeadamente para utilização do modelo de Huff.

A criação do vetor dos modos de deslocação tem como objetivo calcular as distâncias entre dois pontos usando diferentes modos de deslocação sem que seja necessário estar a alterar no programa. Este vetor permite também a versatilidade de escolha do modo de deslocação. Em termos de utilização é necessário recarregar a página após a inserção dos valores referentes às origens, destinos e códigos de destino, para serem lidos pelo programa de cálculo da matriz de distâncias. Este programa apresenta-se sob a forma de uma página html, como se pode observar na Figura 23.

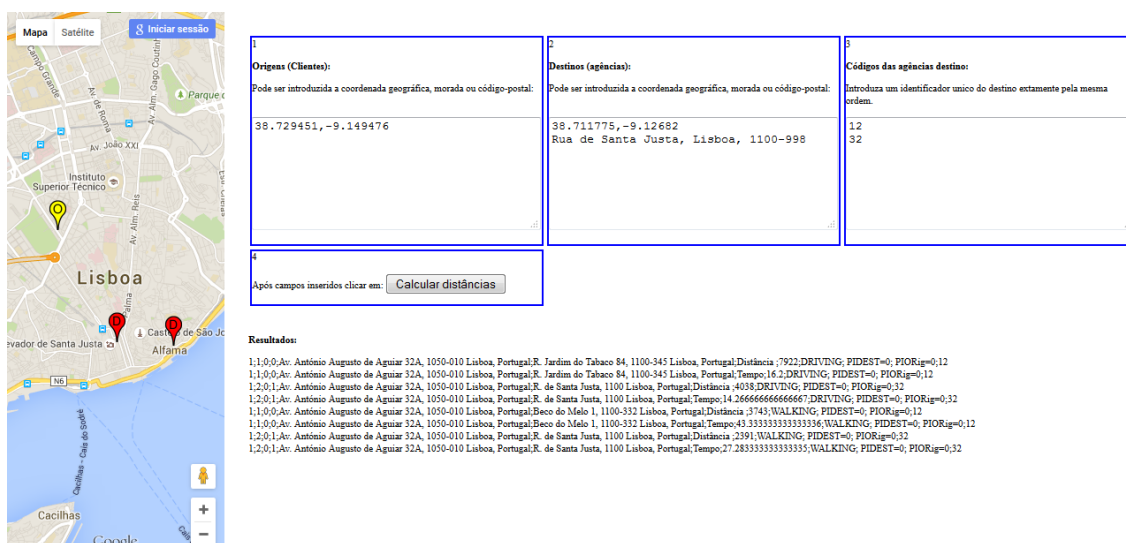


Figura 23 - Interface de cálculo da matriz de distâncias, com *input* em coordenadas geográficas e os vários outputs possíveis, mapa com pontos assinalados e linhas com informação da distância, tempo e modo de deslocação entre dois pontos.

Através da Figura 23 pode visualizar-se a página do cálculo da matriz de distâncias com um caso de exemplo, no primeiro campo (1), no campo superior esquerdo diz respeito à informação relativa à localização das origens, sendo neste exemplo o campo preenchido com as coordenadas geográficas da origem. No segundo campo (2), à direita insere-se a localização dos destinos, sendo aqui preenchida com informação referente a dois destinos, um sob a forma de

coordenadas geográficas e outro sob a forma de morada. O terceiro campo (3) é onde se insere o código dos destinos, pela ordem utilizada no campo (2). Estando todos os campos preenchidos, recarrega-se a página para serem lidos pelo programa, depois carrega-se em 'Calcular distâncias'. O resultado é obtido diretamente na página como se pode observar na mesma figura, nas diversas linhas de texto, em que cada uma representa uma célula da matriz de distâncias.

Resumidamente o cálculo da matriz de distância procede-se da seguinte forma: são recebidos dados correspondentes à localização das origens e destinos, que vão ser inspecionados de modo a perceber se correspondem a moradas ou coordenadas geográficas, pois o modo como é inserido no vetor de pesquisa difere. Quanto aos modos de deslocação, o vetor já se encontra preenchido internamente após os vetores de origens, destinos e códigos de destinos preenchidos, procede-se ao cálculo das distâncias, onde em cada iteração é feita a verificação se os vetores com todos os destinos, origens e modos de deslocação já foram todos percorridos para que o programa termine. Pode observar-se o processo de cálculo descrito mais detalhadamente na Figura 25.

No final, é devolvido um *script* no ecrã com a informação resultante, ou seja, são devolvidas distâncias em metros e minutos tendo em conta o modo de deslocação. Esta informação deve ser guardada como 'MatrizDistancias.csv', pois esta matriz vai ser necessária à aplicação do modelo de Huff. Pode visualizar-se na Tabela 17 um exemplo de informação contida na matriz de distâncias.

Tabela 17 - Ficheiro com o tipo de resultados obtidos através do cálculo da matriz de distâncias.

| Índice do vetor de pesquisa das origens | Índice do vetor de pesquisa dos destinos | Índice do vetor global das origens | Índice do vetor global dos destinos | Origens | Destinos | Medida de distância | Distância | Modo de deslocação | Posição inicial do vetor global dos destinos | Posição inicial do vetor global das origens | Código dos destinos |
|---|--|------------------------------------|-------------------------------------|----------|-----------|---------------------|-----------|--------------------|--|---|---------------------|
| 1 | 1 | 0 | 0 | Av. Antó | R. Jardim | Distância | 7922 | DRIVING | PIDEST=(| PIORig=(| 12 |
| 1 | 1 | 0 | 0 | Av. Antó | R. Jardim | Tempo | 16.2 | DRIVING | PIDEST=(| PIORig=(| 12 |

Os modos de deslocação considerados são: de carro (*DRIVING*) e a pé (*WALKING*). As medidas de distância são: a distância do percurso entre dois pontos em metros e o tempo do percurso entre os dois pontos em minutos. Como já referido anteriormente, as origens dizem respeito a pontos de clientes ou potenciais clientes, e os destinos referem-se às agências, sendo o código do destino o código das agências.

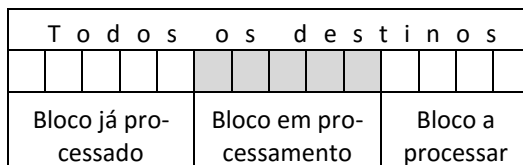


Figura 24 - Processo de preenchimento dos blocos dos destinos, análogo para origens.

Portanto na Figura 24 é apresentada a maneira como se utilizam os blocos de destino e origem criados e a sua interação com os vetores que recebem todos os dados.

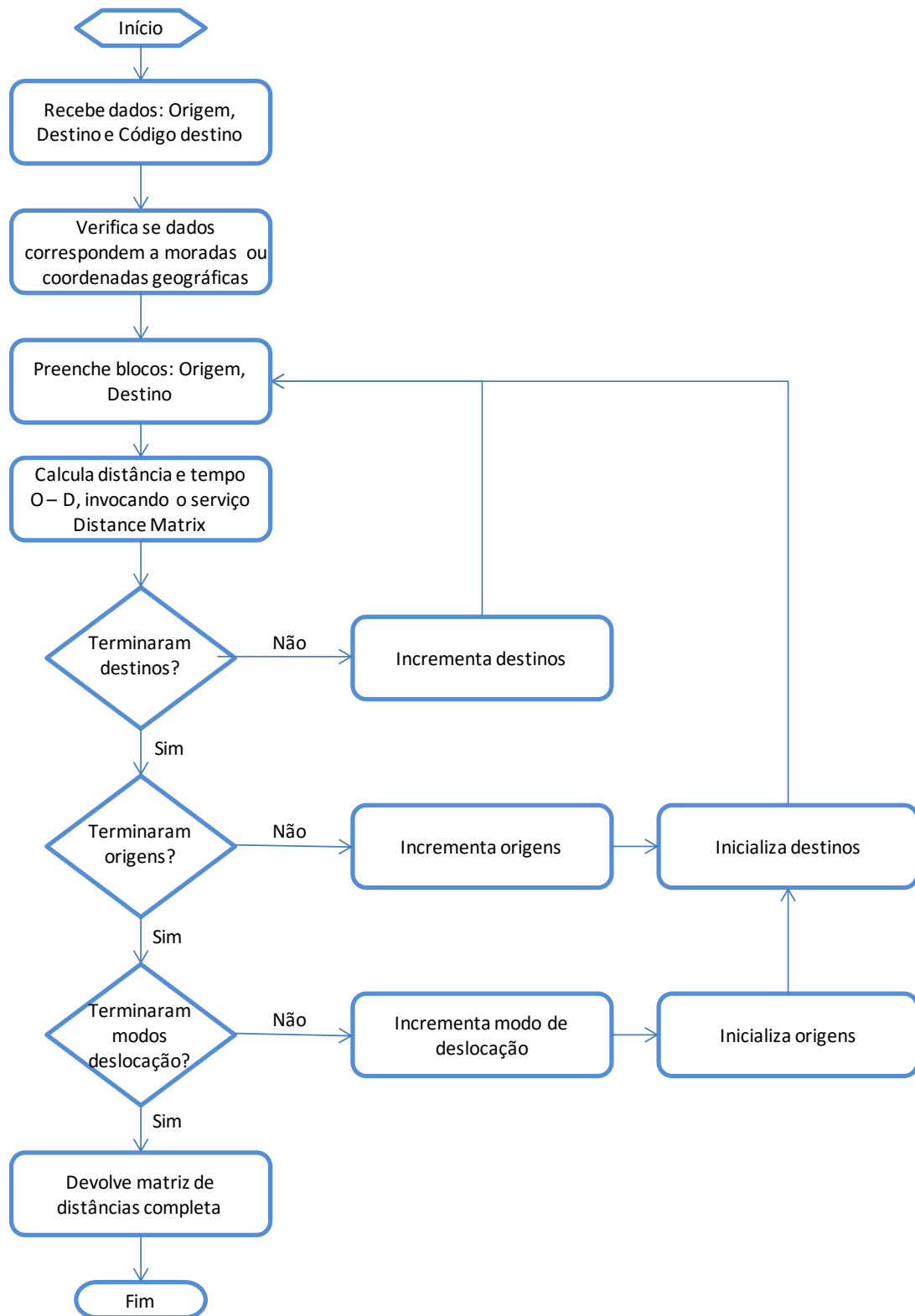


Figura 25 - Processo do cálculo da matriz de distâncias.

Através da Figura 25 é possível observar todo o processo por detrás do cálculo da matriz de distâncias, assim como a informação necessária. E esta matriz é essencial para a aplicação do modelo de Huff no passo seguinte.

3.2.3. Aplicação do modelo de Huff

Para a aplicação deste modelo é necessário ter em atenção dois pontos importantes referentes à calibragem do modelo: possuir informação sobre a variável de repulsão e informação sobre a variável de atratividade. A variável de repulsão diz respeito à distância, que pode ser dada em metros ou minutos pela matriz de distâncias calculada na fase anterior. A escolha da medida de distância é efetuada de acordo com os objetivos do decisor, assim como o modo de deslocação, também presente na matriz de distâncias.

Quanto à variável de atratividade, sendo uma medida que varia com o negócio, deve poder ser adaptada de acordo com os objetivos de análise, como já foi descrito no Capítulo 2. Portanto, para a construção desta variável tive em consideração os diversos fatores, que podem ser internos ou externos ou até uma ponderação destes dois tipos de fatores.

Relativamente ao modelo de Huff, este atribui uma percentagem de potencial interesse de um cliente, numa dada localização, a cada uma das agências da zona considerada. Sendo que a esta percentagem passada a referir-se como percentagem de Huff.

Antes de se efetuar o cálculo da percentagem de Huff é preciso calcular a atratividade de cada agência, sendo necessário ter já num ficheiro a informação que irá ser considerada na fórmula de atratividade. Logo, se para as agências da CGD se pretender que a atratividade seja expressa, por exemplo, pelo negócio, é necessária a existência de uma coluna com essa informação no ficheiro com a informação das agências da CGD. No caso de se pretender considerar múltiplos fatores nem todos internos, é também necessário que essa informação esteja em colunas para depois, na função de cálculo da atratividade (ver Figura 26), se aplique a fórmula pretendida aos fatores considerados, para cada agência.

Posto isto, para a aplicação deste modelo foram criados três programas distintos utilizando o *software* R: o primeiro calcula a percentagem de Huff das agências da CGD relativamente a um cliente; o segundo calcula a percentagem de Huff das agências da CGD e concorrência, também relativamente a um cliente; o terceiro procura a percentagem máxima de Huff para o caso de existir mais de um cliente, sendo que mais de um cliente significa por exemplo, que são dois focos distintos geograficamente de clientes ou potenciais clientes. Cada um destes programas possui pressupostos diferentes, de forma geral inicialmente é calculada a percentagem de Huff para todas as agências destino. De seguida, se existir mais de um ponto de origem é procurada percentagem máxima de Huff para cada uma das agências e é guardado ponto que gerou esse máximo.

Neste trabalho, quando a análise realizada é feita tendo apenas em consideração agências da CGD, é utilizada como variável/medida de atratividade o negócio em 2013/12, enquanto, quando a análise engloba também as agências da concorrência e por não se ter acesso a dados internos das agências da concorrência, a atratividade passa a ser um indicador do município em que as agências estão inseridas, por exemplo o poder de compra, passando assim a distância a ter um maior impacto nos resultados, pois algumas das agências vão ter a medida de atratividade igual por estarem no mesmo município.

Dada a necessidade da variável da atratividade poder ser modificada ou incluir vários indicadores/fatores, esta pode ser alterada nos programas que calculam a percentagem de Huff sem que seja necessário fazer alterações para além da função `calcula.atratividade`. Pode observar-se através da Figura 26 o local no programa onde é possível alterar a atratividade, ou seja, a função `calcula.atratividade`.

```
# Fórmula de atratividade pode ser alterada pelo utilizador.  
# Podem ser usadas na fórmula todas as colunas do data.frame das agências.  
calcula.atratividade <- function(dados.agencia) {  
  return(dados.agencia$NEGOCIO_201312);  
}
```

Figura 26 - Função de cálculo da atratividade. Neste caso a atratividade é o NEGOCIO_201312 do *data.frame* das agências CGD.

➤ **Processo de aplicação do modelo de Huff considerando apenas agências da CGD**

Na sequência da decisão sobre a fórmula de cálculo da atratividade, é executado o cálculo da percentagem de Huff para todas as agências consideradas na análise tendo em conta os vários modos de deslocação e as medidas de distância. O resultado deste cálculo possibilita diferentes análises, pois cada agência vai possuir quatro percentagens de Huff diferentes correspondentes às diferentes combinações de medida de distância e modo de deslocação. Pode observar-se na Figura 27 o processo de cálculo da percentagem de Huff tendo em conta apenas as agências da CGD.

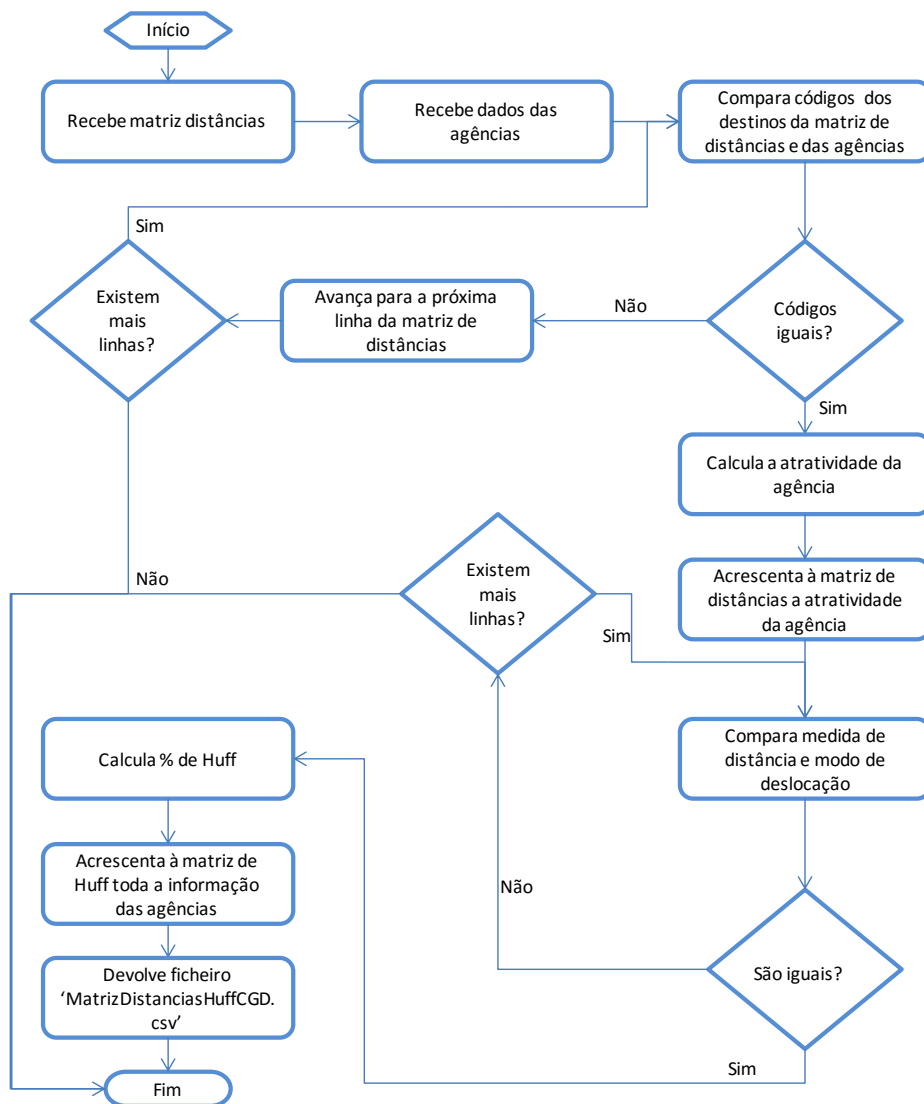


Figura 27 - Processo do cálculo da % de Huff para agências da CGD.

Deste processo resulta um ficheiro, com o nome 'MatrizDistanciasHuffCGD.csv' (ver tabela 18), contendo a percentagem de Huff, a medida de distância, o modo de deslocação, a origem, o destino entre outros dados importantes, nomeadamente a localização. Esta informação é manipulada na fase seguinte de modo a possibilitar a visualização dos resultados obtidos.

Tabela 18 - Campos do ficheiro que resulta da aplicação do modelo de Huff.

| IVO | IVD | Origem | Destino | medida.d istanca | TD | modo.de slocacao | C_AG | atratividade de | Huff | NM_AG | NEGOCIO _201312 | Q_CLIENT ES_20131 2 | ANTG_A G | CODIGO _POSTAL | LATITUDE | LONGITUDE |
|-----|-----|--|---------------------------------|---------------------|-------|---------------------|------|--------------------|-------|--|--------------------|---------------------------|-------------|-------------------|----------|-----------|
| 112 | 47 | Rua Lisboa, 2765 Estoril, Portugal | 1700-178 Lisboa, Portugal | Tempo | 21728 | WALKING | 82 | 218700 | 0.233 | ESTADOS UNIDOS DA AMERICA - LISBOA | 218700 | 1159 | 4 | 1700-178 | 38.74968 | -9.141658 |
| 112 | 48 | Rua Lisboa, 2765 Estoril, Portugal | 1700-337 Lisboa, Portugal | Distância | 29577 | WALKING | 125 | 98371 | 0.116 | RIO DE JANEIRO - LISBOA | 98371 | 737 | 3 | 1700-337 | 38.75615 | -9.139501 |

Através da tabela 18 é possível observar os principais campos de saída da aplicação do modelo Huff. Os campos mais relevantes para a representação num mapa da percentagem de Huff

obtida para cada uma das agências consideradas são a latitude e longitude, para que as agências possam ser representadas geograficamente num mapa, a percentagem de Huff, para que se visualizem os valores obtidos de cada agência, e o modo de deslocação e a medida de distância.

➤ **Processo de aplicação do modelo de Huff considerando agências da CGD e concorrência**

Na Figura 28 é apresentado o processo do cálculo da percentagem de Huff tendo em conta as agências da CGD e as agências da concorrência, de onde, resulta o ficheiro 'MatrizDistancia-sHuffCGD_Concorrencia.csv'. O objetivo deste cálculo é observar se a percentagem de Huff é muito afetada pela existência da concorrência.

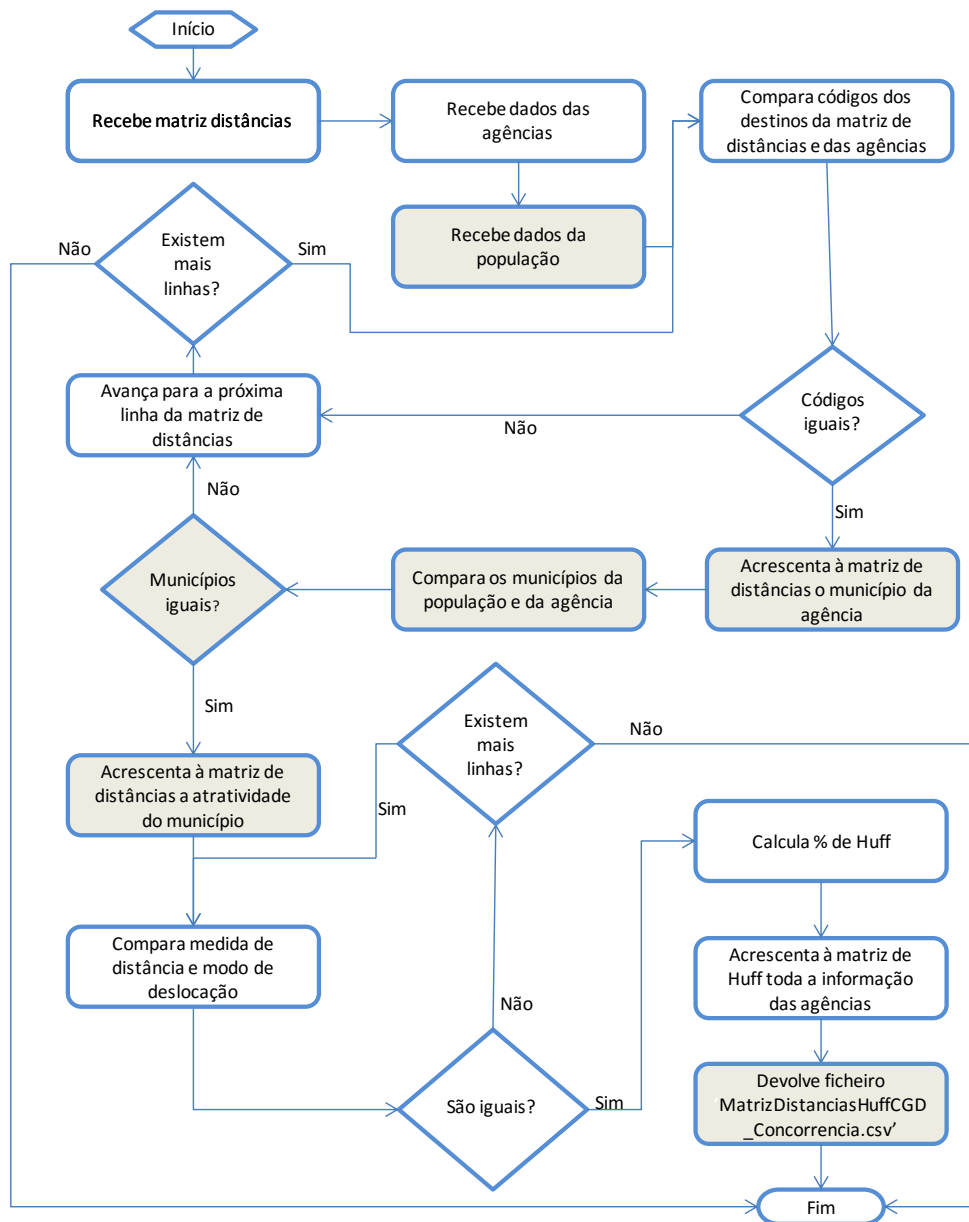


Figura 28 - Processo de cálculo da % de Huff para agências da CGD e concorrência. Os passos com o fundo cinzento são novos face aos que constam na Figura 27.

➤ **Processo de aplicação do modelo de Huff considerando vários pontos de origem**

Dos processos anteriores quando se considera um ponto de origem para chegar às agências, ou seja, um ponto geográfico onde existem ou podem vir a existir clientes, resulta um conjunto de quatro valores diferentes de percentagem de Huff para cada agência, que correspondem às diferentes combinações de modos de deslocação e medidas de distância. Por outro lado, quando são considerados vários pontos de origem, resultam vários conjuntos de quatro valores para cada agência. Por exemplo, se forem considerados dois pontos de origem o resultado são dois conjuntos de quatro valores para cada agência. Portanto, no caso de existirem vários pontos de origem para além de utilizar um dos programas anteriores, cálculo da % de Huff para agências da CGD ou cálculo da % de Huff para agências da CGD e concorrência, é realizada a procura da percentagem máxima de Huff de cada uma das agências resultantes desse cálculo (ver Figura 29).

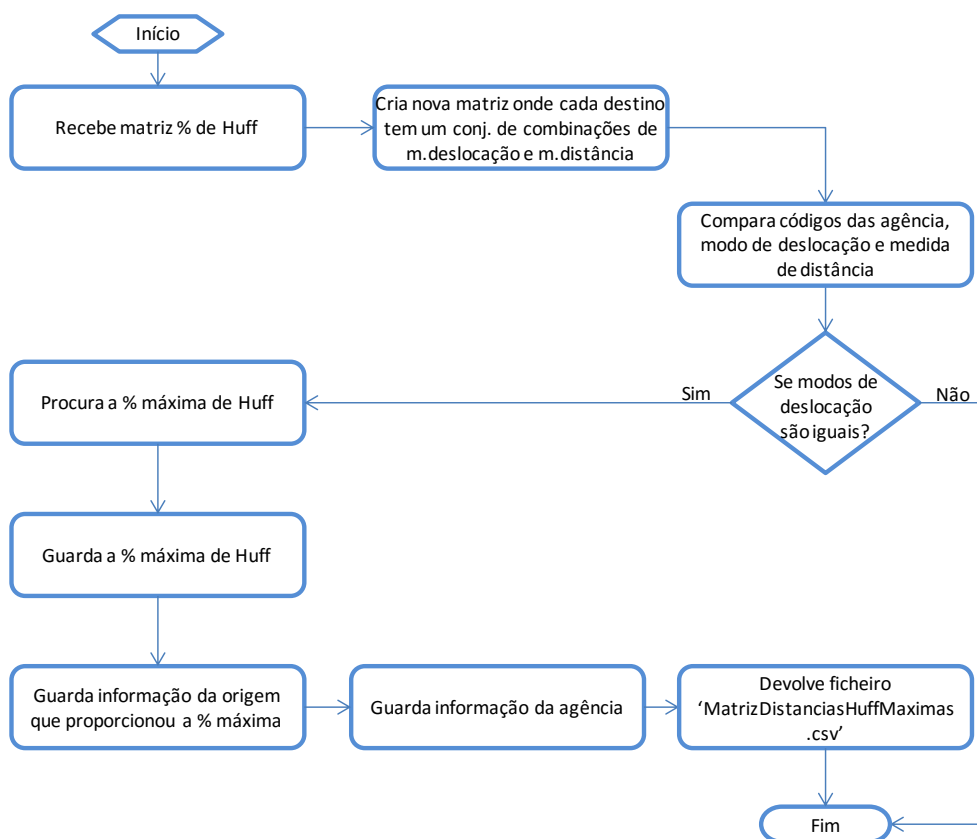


Figura 29 - Processo do cálculo da % máxima de Huff.

Deste processo resulta a percentagem máxima de cada agência assim como a origem associada e toda a informação das agências, para que possam ser localizadas geograficamente, sendo estes dados guardados num ficheiro 'MatrizDistanciasHuffMaximas.csv'.

Em qualquer um dos dois primeiros programas referidos - o cálculo da percentagem de Huff apenas para a CGD e agências da CGD e da concorrência - pode alterar-se facilmente o valor do β , pois, como foi apresentado no Capítulo 2, este parâmetro deve ser o mais adequado para cada caso de negócio. Tendo sido aqui utilizado o valor mais usual $\beta=2$.

- **Análise do desempenho do modelo de Huff**

Com o intuito de apurar o desempenho do programa R que calcula a percentagem de Huff para as agências da CGD, foi feita a análise do tempo de execução para diferentes números de destinos relativamente a um ponto de origem. Para esta análise utilizou-se o Toshiba Portégé com o processador Intel i5 de 1.70GHz, com 4GBytes de memória no sistema operativo Windows 10 de 64 Bits, e um disco de 100 GBytes. O processo usado foi o seguinte:

1. Criou-se um ficheiro com a amostra referente a 500 destinos, que deu origem a sete subconjuntos, com 5, 10, 50, 100, 200, 300 e 400 destinos.
2. Para cada uma das amostras, correu-se o programa uma vez. Mediu-se o tempo de cálculo da percentagem de Huff, executando-se cinco vezes este processo para cada amostra. O primeiro valor obtido foi desprezado em todas amostras por ser um valor consideravelmente diferente dos restantes obtidos. Processo analogo para o carregamento da informação das agências.
3. Fez-se a média dos resultados obtidos para cada uma das amostras e representaram-se graficamente (ver Figura 30).

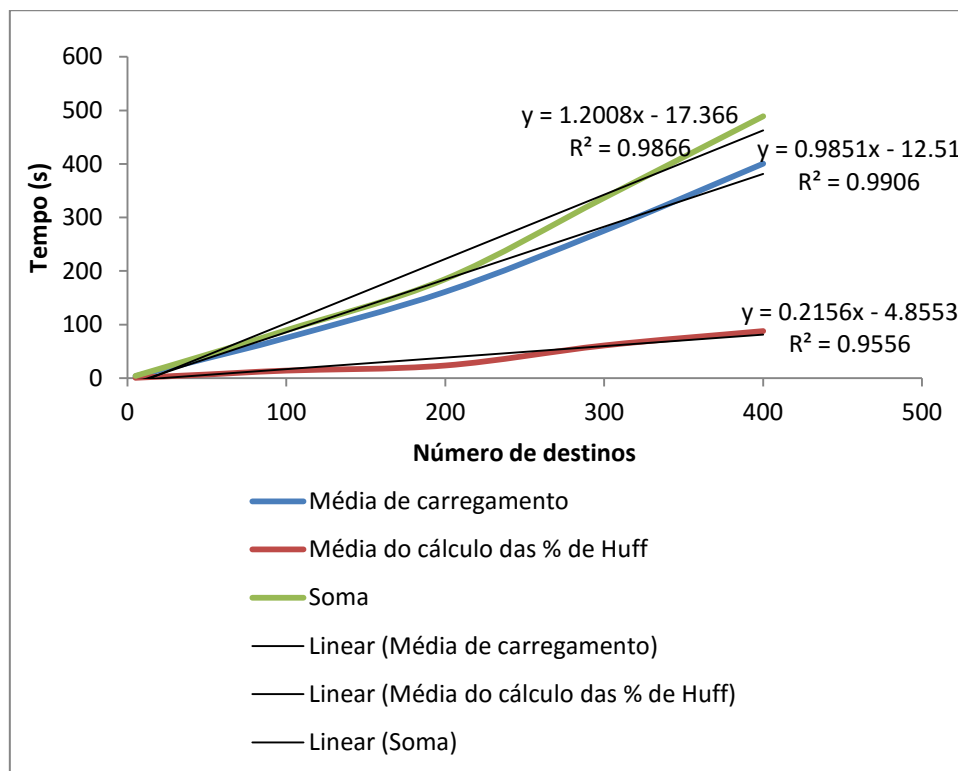


Figura 30 - Desempenho do programa de cálculo das percentagens de Huff relativamente a um ponto de origem, e 5 a 400 destinos, em que a soma representa o tempo total, ou seja, tempo de cálculo do modelo de Huff e o carregamento das tabelas com a informação correspondente das agências.

Observando a Figura 30, nota-se que de 5 a 400 destinos o tempo de carregamento é praticamente linear ($R^2=0.99$), segundo a fórmula $0.9851 \cdot \text{destinos} - 12.51$. O tempo passa a ser maior do que o esperado, de 400 para 500 destinos como se pode visualizar na Figura 31, tendo sido duas vezes superior conforme a fórmula $1.6695 \cdot \text{destinos} - 61.619$.

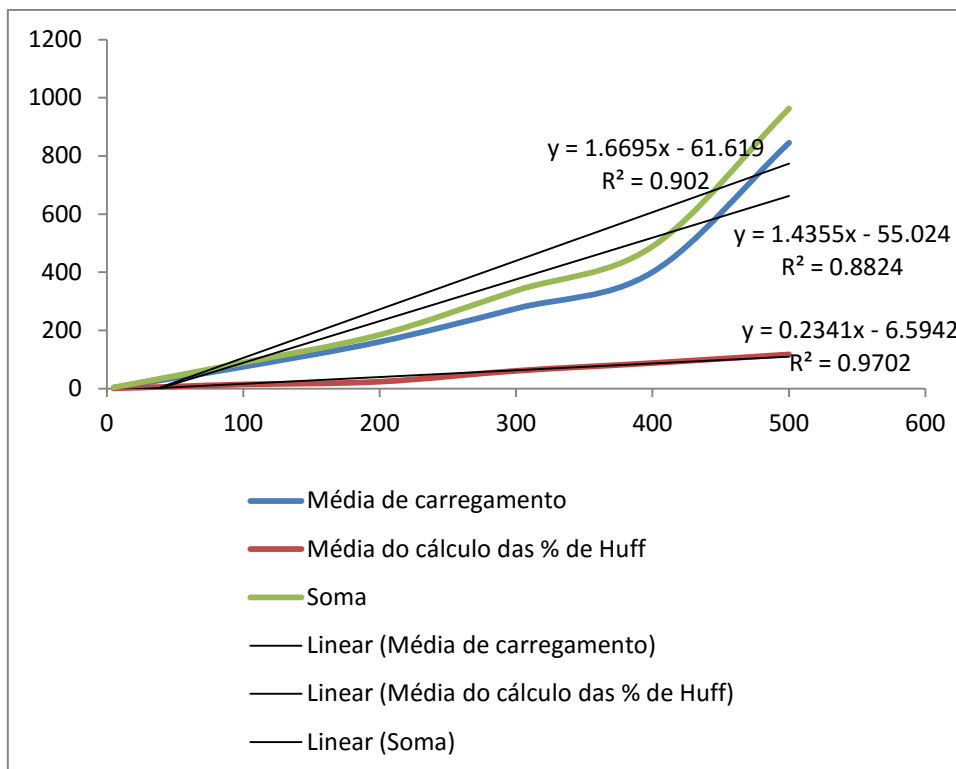


Figura 31- Desempenho do programa de cálculo das percentagens de Huff relativamente a um ponto de origem, 5 a 500 destinos, em que a soma representa o tempo total, ou seja, tempo de cálculo do modelo de Huff e o carregamento das tabelas com a informação correspondente das agências.

Verifica-se também que a maior parte do tempo é gasto no carregamento da informação das agências no ficheiro gerado.

3.2.4. Visualização dos dados

Após o cálculo das percentagens de Huff no passo anterior, passa a ser possível recorrer a um SIG para se representar os resultados obtidos, nomeadamente no QGIS. Através desta ferramenta consegue-se uma visualização da informação de modo mais interativo e natural que permite o auxílio à tomada de decisão. Na Figura 32 pode observar-se o processo realizado para a visualização dos dados obtidos através da aplicação do modelo de Huff.

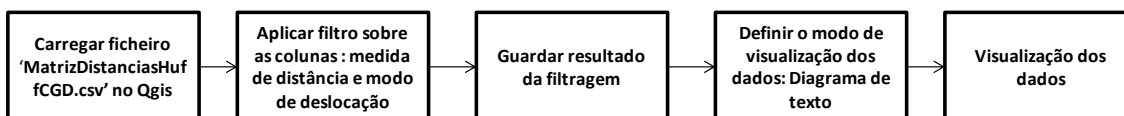



Figura 32 - Processo para visualização dos dados obtidos da aplicação do modelo de Huff.

Para que se faça uma representação correta dos dados, é necessário carregar o ficheiro obtido do programa de cálculo das percentagens e aplicar o filtro do que se pretende observar, visto os resultados obtidos terem com todas as opções possíveis, ou seja, para cada agência tem-se todas as combinações modo de deslocação e medida de distância, o que leva a que cada agência tenha mais de uma percentagem de Huff associada. Portanto, para fazer uma análise correta é necessário aplicar um filtro sobre as colunas medida de distância e modo de deslocação,

por exemplo, escolher como modo de deslocação “DRIVING” e medida de distância “Distancia” que é representada em metros, analisando-se assim a percentagem de Huff considerando a distância que o cliente percorre deslocando-se à agência de carro.

Para executar o filtro, abre-se a tabela de atributos de, por exemplo, da camada ‘MatrizDistanciasHuffCGD.csv’ clicando em , de seguida carrega-se no canto inferior esquerdo onde aparece ‘mostrar todos os elementos’ e depois clica-se em ‘filtro por coluna’ e seleciona-se o campo, como é demonstrado na Figura 33.

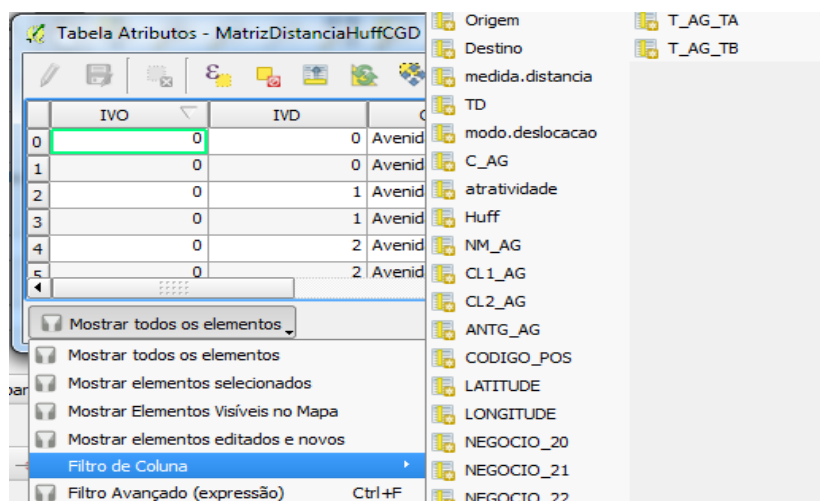


Figura 33 - Processo de filtragem por coluna no QGIS de uma tabela de atributos da camada ‘MatrizDistanciasHuffCGD.csv’.

Após o filtro aplicado é necessário guardar a informação num novo ficheiro do tipo shapefile, que gera uma nova camada, para tal seleciona-se os elementos resultantes da filtragem, de seguida, em cima da camada correspondente pressiona-se o botão direito do rato, e logo de seguida escolher ‘Guardar Como’. Abre-se uma janela, sendo apenas necessário escolher o nome que se pretende dar ao novo ficheiro e selecionar ‘Guardar apenas elementos seleccionados’. Podem-se observar estes passos na Figura 34.

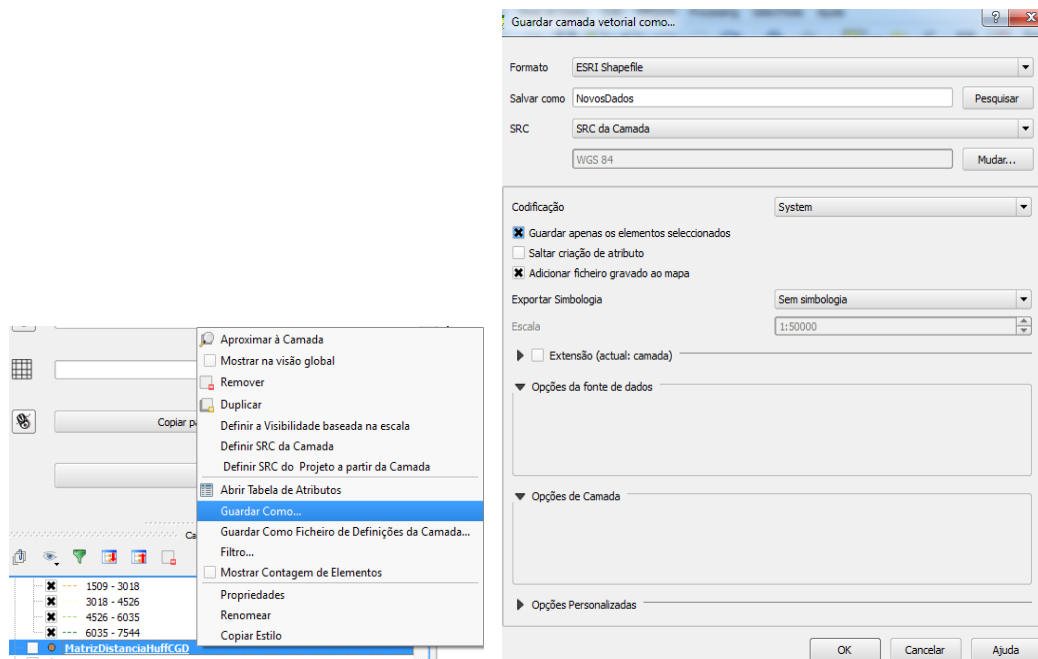


Figura 34 – Passos para a criação de uma nova camada Shape com os dados resultantes da filtragem: a) Botão onde carregar para guardar os elementos seleccionados, dando origem a b) janela que aparece para concluir a gravação dos elementos seleccionados.

Terminado o conjunto de passos anterior o novo ficheiro é adicionado ao mapa, pois a opção ‘Adicionar ficheiro gravado ao mapa’ está por defeito seleccionada (ver Figura 34), encontrando-se os dados em condições de serem representados.

Para a visualização da percentagem de Huff num mapa é necessário seguir alguns passos: primeiro seleccionam-se as propriedades da camada ‘MatrizDistanciasHuffCGD.csv’ clicando com no botão direito do rato na camada em questão; depois clicar em ‘Propriedades’ -> ‘Diagramas’ e seleccionar ‘Mostrar diagramas’, depois escolhida a representação pretendida, neste caso ‘Diagrama de texto’ sendo ainda necessário seleccionar o atributo a ser mostrado que é o ‘Huff’. Pode-se observar na Figura 35 o local onde se escolhe este atributo.

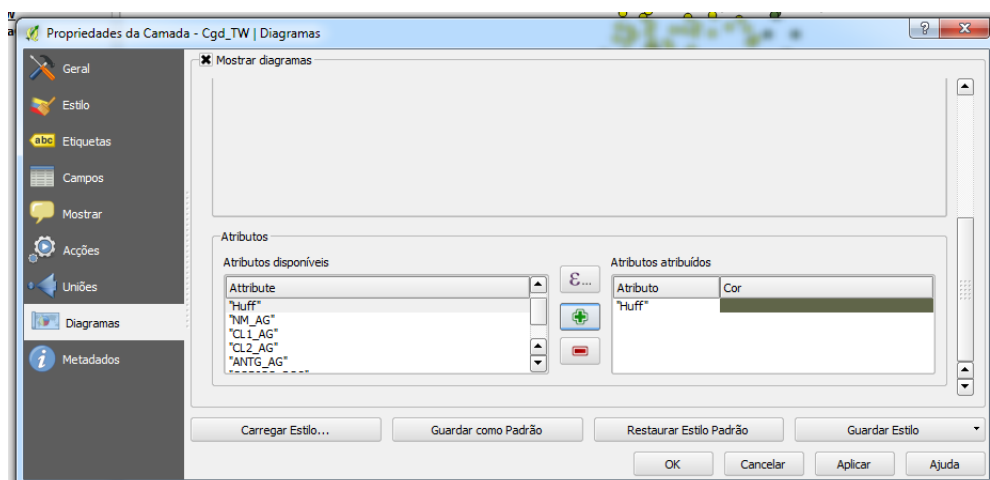


Figura 35 - Definição na representação da percentagem de Huff no mapa, escolha do atributo a ser representado pelo diagrama de texto.

Depois da seleção do atributo, é importante definir a posição onde se pretende que o diagrama de texto se localize, sendo escolhido 'Sobre o ponto', como se pode visualizar na Figura 36.

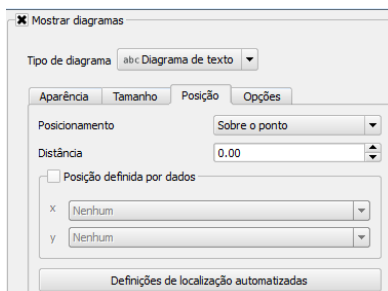


Figura 36 - Definição na representação da percentagem de Huff no mapa, escolha da posição do diagrama de texto.

Por fim, escolhe-se o tamanho desejado para o diagrama de texto, que é variável consoante a percentagem de Huff. Para isso, ainda na mesma janela mas no separador 'Tamanho' preenchem-se os campos de acordo com a Figura 37.

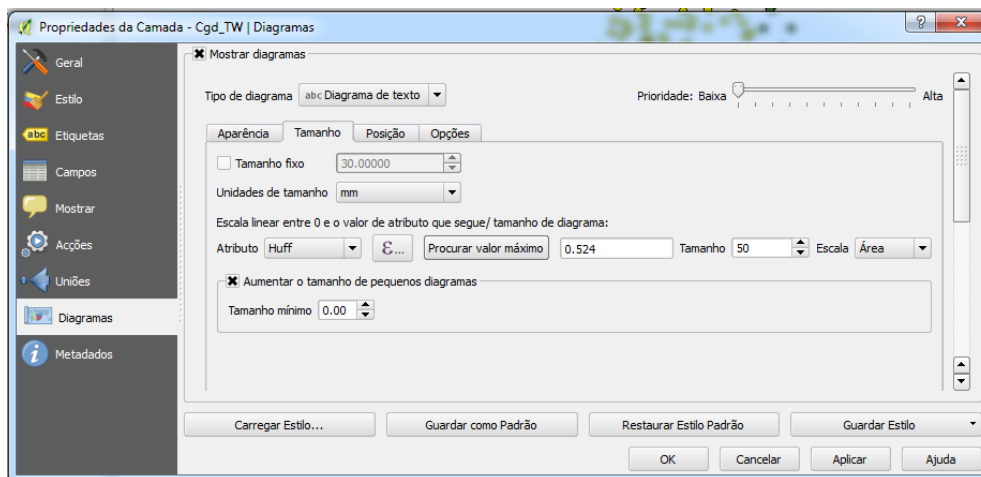


Figura 37 - Definição na representação da percentagem de Huff no mapa, escolha do tamanho do diagrama.

Para finalizar a definição do diagrama de representação das percentagens de Huff, clica-se em 'OK', através da Figura 38c) pode-se visualizar os resultados obtidos, ou seja, a percentagem de Huff de cada uma das agências consideradas.

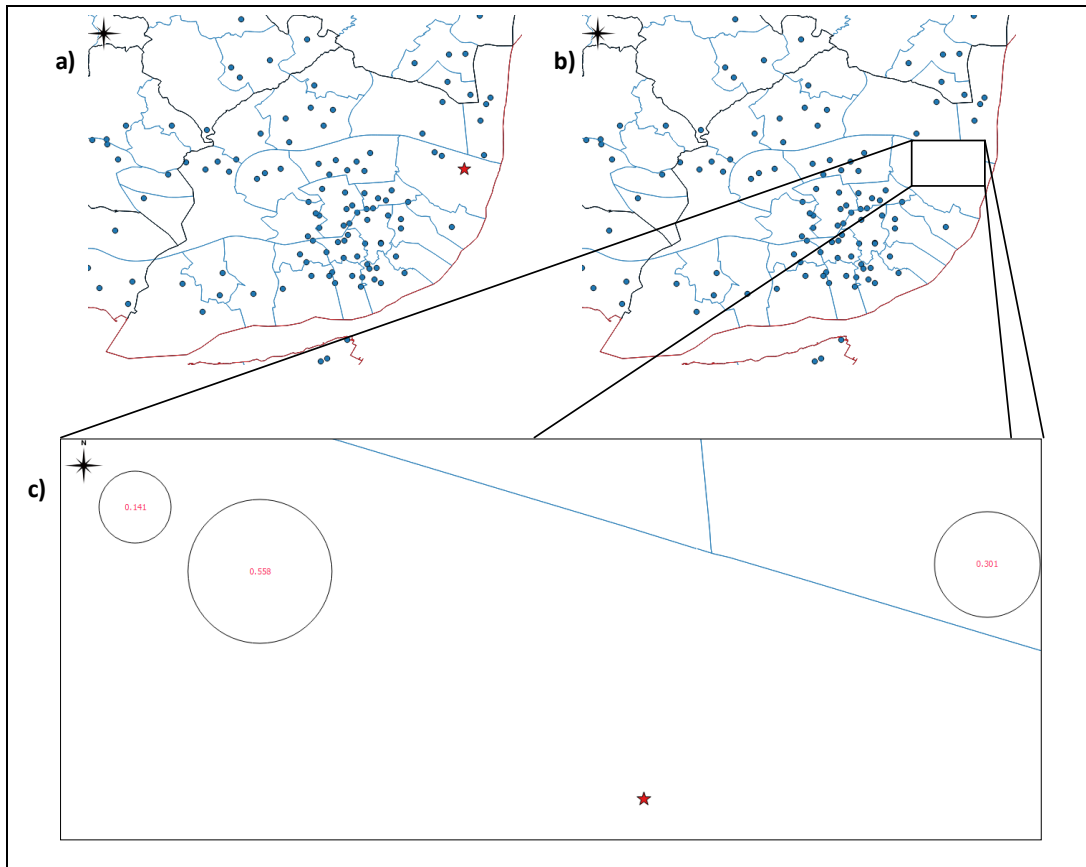


Figura 38 – a) Município de Lisboa com a localização das agências da CGD representadas por bolas azuis e a localização de um cliente representado por uma estrela vermelha (mais ou menos a meio à direita), b) Equivalente a a) mas com um retângulo por cima da área de análise, que se podem observar em, c) Agências representadas por diagramas circulares e graduados com a % de Huff relativamente ao ponto selecionado.

Pode verificar-se através da Figura 38, que a representação da percentagem de Huff das agências selecionadas é feita através de diagramas circulares anotados com a percentagem de Huff no centro. Está nesta figura o município de Lisboa com a localização das agências da CGD representadas por círculos azuis no canto superior esquerdo, ao seu lado direito encontra-se uma imagem semelhante, mas sobreposta por um quadrado branco. Este quadrado diz respeito à área de análise aqui usada, sendo necessário escolher uma área, local ou dimensão de análise, pois se forem consideradas todas as agências de norte a sul por exemplo, os valores da percentagem de Huff serão zero em praticamente todas as agências.

Pode observar-se através do *zoom* feito Figura 38c), que as agências ficam representadas por diagramas maiores ou menores consoante a percentagem de Huff da agência, sendo neste caso a percentagem mais elevada de 56% e a menor de 14%. A estrela vermelha da figura diz respeito ao ponto de origem, ou seja, à posição de um cliente ou potencial cliente.

3.3. Cenários de análise

Neste ponto são expostos alguns cenários possíveis (Tabela 19) de análise tendo em conta o modelo de Huff e as ferramentas disponíveis.

Tabela 19 - Cenários de utilização considerados neste trabalho.

| Cenário | Objetivo |
|---------|---|
| A | Analisar a rede de agências relativamente à localização de um cliente, desprezando a concorrência. |
| B | Analisar a rede de agências relativamente à localização de um cliente, considerando a concorrência. |
| C | Analisar a rede de agências relativamente a várias localizações de clientes. |
| D | Analisar a rede de agência relativamente à instalação de uma nova agência. |
| E | Analisar a rede de agência relativamente ao encerramento permanente de uma agência. |
| F | Analisar a rede de agência relativamente ao encerramento pontual de uma agência. |
| G | Analisar a rede de agência partindo da sobreposição de áreas de influência. |

Os cenários são de seguida apresentados, sendo que o primeiro é descrito com maior detalhe, pois os seus passos vão ser também executados nos cenários seguintes, estando nesses mencionadas as diferenças existentes.

A. Cenário de análise da rede de agências da CGD relativamente a um ponto de origem

Este cenário tem por objetivo analisar apenas a rede de agências da CGD de uma certa zona, desprezando a existência de concorrência. Toma como pressuposto que o ponto de origem diz respeito à localização de um cliente ou de potenciais clientes. Por exemplo, sabendo-se que vai ser construído um centro comercial, pretende-se saber qual a agência que está melhor ou pior posicionada relativamente a esse ponto, ou seja, pretende-se observar qual a que tem menor ou maior percentagem de Huff relativamente ao centro comercial.

Este cenário consiste numa sequência de passos, que de forma geral passa pela escolha do ponto do cliente, definir o raio da análise, calcular a matriz de distâncias, aplicar o modelo de Huff e por fim representar os resultados obtidos para facilitar a tomada de decisão.

De seguida é explicado cada passo no cenário de análise à rede de agências, expondo um exemplo de aplicação desta metodologia utilizando o QGIS. Assumindo que toda a informação necessária já foi carregada no QGIS (localização das agências e mapa de Portugal), têm-se 5 principais passos: escolher o ponto de origem, selecionar raio de análise, calcular matriz de distâncias, aplicar o modelo de Huff, representação dos resultados obtidos.

1. Escolher o ponto de origem

Neste passo vai-se realizar a escolha da posição do cliente a analisar, posição esta marcada no mapa clicando em 'Palce a Pin' (ver Figura 39) e de seguida apontando no mapa o ponto de interesse. No caso de já existir um ponto pré-escolhido, também é possível encontrá-lo no mapa através das suas coordenadas e marcá-lo.

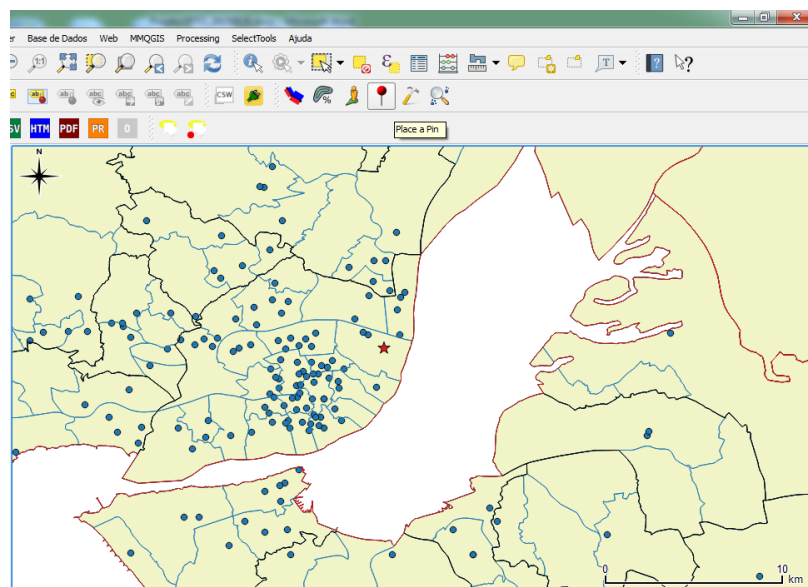



Figura 39 - Representação das agências da CGD por pontos a azul e do ponto de origem por uma estrela vermelha (aproximadamente no centro da figura).

Na Figura 39 pode observar-se o local onde clicar para se poder marcar um ponto no mapa, ficando este representado por uma estrela vermelha, neste caso no município de Lisboa, enquanto as agências da CGD são representadas por pontos azuis.

2. Selecionar o raio de análise

Neste passo realiza-se a escolha das agências que vão ser consideradas na análise, visto ser pouco provável que um cliente com agências próximas vá percorrer uma quantidade maior de quilómetros para chegar a uma agência mais afastada. Sendo assim, considerou-se relevante a seleção de um raio de análise para tornar a aplicação do modelo mais fiável. Para selecionar as agências de acordo com um raio a partir do ponto do cliente seleciona-se 'Selecionar elementos por raio' clicando em , de seguida com rato em cima do ponto do cliente carrega-se no botão esquerdo do rato e arrasta-se até se obter o raio pretendido.

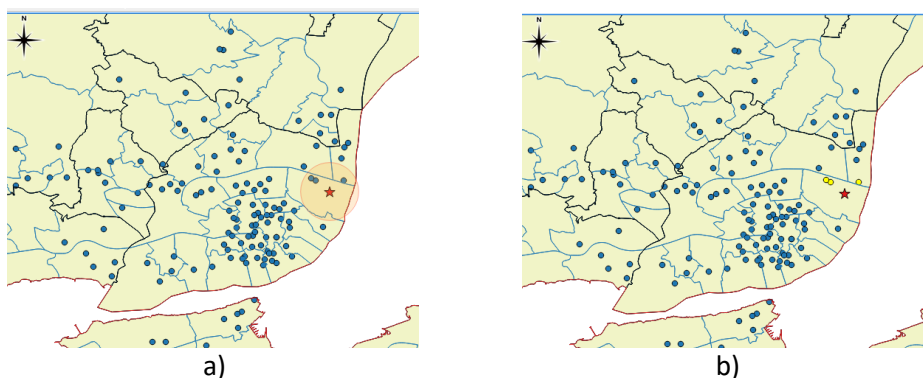



Figura 40 – a) Seleção do raio de análise e b) Agências consideradas para análise a amarelo.

Pela Figura 40a), observa-se o raio de análise selecionado e na Figura 40b) mostra todas as agências que estiverem contidas nesse raio estão agora representadas a amarelo; estas serão as agências a considerar para a análise através do modelo de Huff.

A sua informação das agências selecionadas deve ser guardada num ficheiro devido ao facto de ser necessária para programas externos ao QGIS, no passo seguinte. Para guardar a informação abre-se a tabela de atributos clicando em , de seguida carrega-se no canto inferior esquerdo onde aparece 'mostrar todos os elementos' -> 'Mostrar elementos selecionados' e por fim copia-se a informação por exemplo para uma folha de LibreOffice Calc como demonstrado na Secção 3.2.4..

3. Cálculo da matriz de distâncias

Para que o cálculo da matriz de distâncias é necessária informação sobre a localização do ponto do cliente, localização das agências, e ainda o código das agências, tendo esta sido guardada na etapa anterior. Para carregar os dados no programa de cálculo da matriz de distâncias, utiliza-se a informação obtida no passo anterior inserindo-a no respetivo campo. Após a matriz de distâncias obtida, deve ser guardado como 'MatrizDistancias.csv', pois é necessária no passo seguinte.

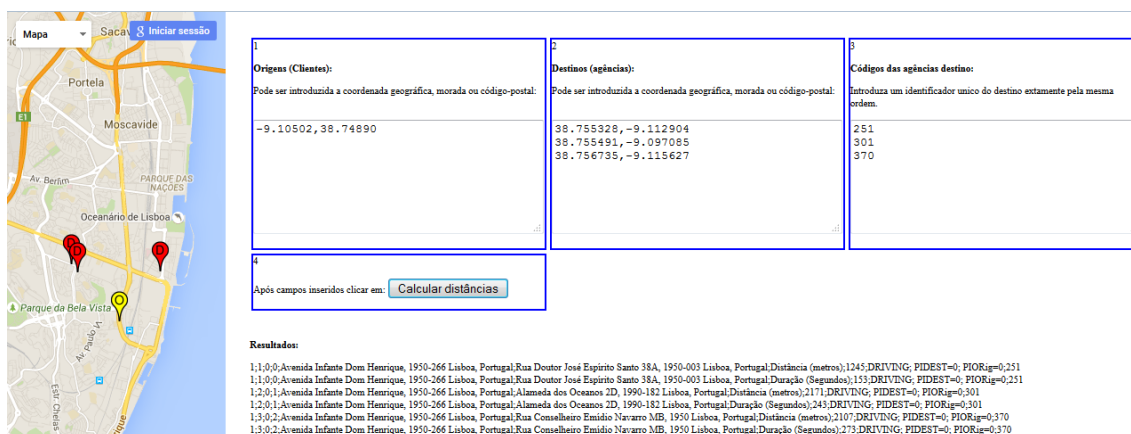


Figura 41 - Cálculo da matriz de distâncias e resultados, para uma origem e três destinos.

A Figura 41 representa o cálculo da matriz de distâncias com os campos preenchidos com o exemplo, e também os resultados obtidos, onde as várias linhas obtidas resultam das combinações dos modos de deslocação e medidas de distância.

4. Aplicação do modelo de Huff

Para a aplicação deste modelo recorre-se ao programa em R que calcula a percentagem de Huff para as agências da CGD. Nesta fase já se possui toda a informação necessária à aplicação da fórmula de cálculo da percentagem de Huff, ou seja, tem-se informação sobre as agências, o que possibilita o cálculo da atratividade e as distâncias em duas medidas, metros e minutos, dois modos de deslocação, a pé e de carro.

Para a correta execução do programa é necessário que os ficheiros 'MatrizDistancias.csv' e 'InfoCGD.csv' fiquem guardados na mesma pasta (o separador dos campos deve ser ';'). No final do programa é gerado um ficheiro CSV com o nome 'MatrizDistanciasHuffCGD' que fica guardado na mesma pasta.

5. Representação dos resultados obtidos

Para representar os resultados obtidos no passo anterior recorre-se à ferramenta QGIS. Em primeiro lugar carrega-se a 'MatrizDistanciasHuff.csv', como descrito na Secção 3.2.2, e aplica-se um filtro sobre os resultados, pois estes incluem todas as combinações de modo deslocação e medida de distância. O processo de representação das percentagens de Huff está descrito na Secção 3.2.4.

B. Cenário de análise da rede de agências da CGD e da concorrência relativamente a um ponto de origem

Este cenário é realizado na ótica de visualizar e tentar compreender o impacto da concorrência na rede de agências da CGD. O processo adotado passa por três fases distintas, ver Figura 42: primeiro executa-se uma análise apenas à rede de agências da CGD, portanto sem ter em consideração as agências da concorrência. Na fase seguinte incluem-se as agências da concorrência na análise e por último comparam-se os resultados obtidos de modo a observar o impacto da concorrência na rede de agências da CGD.

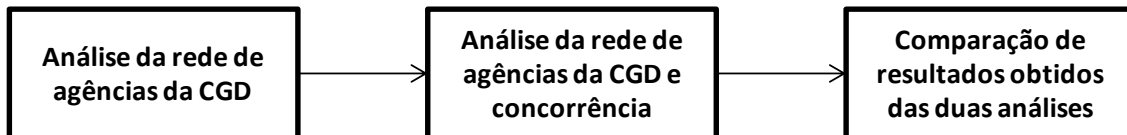


Figura 42 - Processo de análise da rede de agências.

- Primeiro analisam-se apenas as agências da CGD

Nesta primeira fase realiza-se uma análise similar à executada no cenário **A**. O objetivo de analisar apenas as agências da CGD é possuir uma base de referência, de modo a se possibilitar observar o impacto das agências da concorrência na percentagem de Huff que as agências da CGD possuíam inicialmente.

- Segundo, analisam-se as agências da CGD juntamente com as da concorrência

O processo de análise é semelhante ao realizado no cenário **A**, com a diferença de a atratividade considerada ter em conta o local onde está inserida a agência, em vez de uma característica específica das agências em si, para que seja possível utilizar a mesma medida de atratividade tanto para as agências da CGD como da concorrência.

Nesta análise utiliza-se o programa de cálculo das percentagens de Huff para agências da CGD e concorrência, sendo necessários, para além dos ficheiros usados no cenário **A**, o ficheiro com os dados relativos à população, 'DadosPopulacao.csv' e informação relativa às agências da concorrência, 'InfoConcorrenca.csv'. Este programa dá origem ao ficheiro com o nome 'MatrizDistanciasHuffCGD_Concorrenca.csv', que deverá ser carregado no QGIS de modo a possibilitar a visualização dos dados.

- Terceiro, compara-se o resultado da análise considerando apenas as agências da CGD com a análise que considera também as agências da concorrência

O objetivo deste passo análise é observar se a existência de concorrência influencia ou causa alterações nas percentagens de Huff calculadas para as agências da CGD e concorrência nos passos anteriores.

C. Cenário de análise das agências relativamente a vários pontos de origem

Neste cenário, são considerados vários pontos de origem, pretendendo-se analisar dois pontos de potenciais clientes e ver qual o que proporciona uma maior percentagem de Huff a uma agência, ou seja, ver qual dos pontos poderá ser o mais relevante e para que agência.

Esta vertente também possibilita a criação de uma grelha regular com as percentagens de Huff em cada ponto da grelha, sendo esses pontos representativos de clientes, e as agências representadas com a percentagem máxima obtida a partir de um dos clientes. Esta podia ser uma análise interessante para se perceber quais seriam as percentagens obtidas de clientes de vários pontos.

Para este cenário seria usado o programa de procura das percentagens de Huff máximas, que vai procurar no resultado obtido pelo programa de cálculo da percentagem Huff os clientes que originaram maior percentagem de Huff nas diferentes combinações de modo de deslocação e medida de distância.

D. Cenário de análise das agências relativamente à instalação de uma nova agência

Para que esta análise seja feita de um modo fiável, é necessário comparar as agências através dos mesmos atributos, ou seja, a fórmula da atratividade deve ser igual para todas as agências tanto para a nova como para as restantes já existentes. Por exemplo se para a análise inicial da rede de agências se tomou como atratividade o negócio, então para a nova agência, e visto ser impossível saber-se o valor certo relativamente ao negócio que iria gerar, é necessário fazer uma estimativa desse valor de modo a ficar ao mesmo nível comparativo que as restantes.

Só após serem estimados todos os valores dos indicadores utilizados na fórmula da atratividade das agências já existentes para a nova agência é que se podem seguir para os passos seguintes:

1. Análise da rede inicial de agências: Onde se executa o processo análogo ao do cenário **A**.
2. Estimativa da atratividade da nova agência: tendo em conta os fatores considerados realizar uma estimativa para cada um deles e aplicar a fórmula que está a ser utilizada nas restantes agências (integrada no programa do cálculo da percentagem de Huff).
3. Escolha da localização para a nova agência: Selecionar pontos possíveis para a nova agência, e de seguida correr o programa que calcula da percentagem de Huff. Comparar qual das novas agências tem maior percentagem de Huff e escolhe-la.
4. Análise da rede considerando a nova agência: Processo análogo ao do cenário **A**.
5. Comparação com a rede inicial, com visualização do impacto nas percentagens das agências existentes e da percentagem que a nova agência obteve. E daí avaliar se

as condições apresentadas são ou não favoráveis à instalação de uma nova agência relativamente a um certo ponto.

E. Cenário de análise das agências relativamente ao encerramento permanente de uma agência

Neste caso o processo é mais simples que o anterior, bastando eliminar essa agência da rede e comparar com a rede inicial. A informação que se pode observar vai passar pelo deslocamento dos clientes da agência eliminada para as agências que se mantêm na rede. Pode observar-se o estado inicial da rede com a rede a ser eliminada com uma cruz por cima através da Figura 43c) e o resultado após a eliminação de uma agência na Figura 43d).

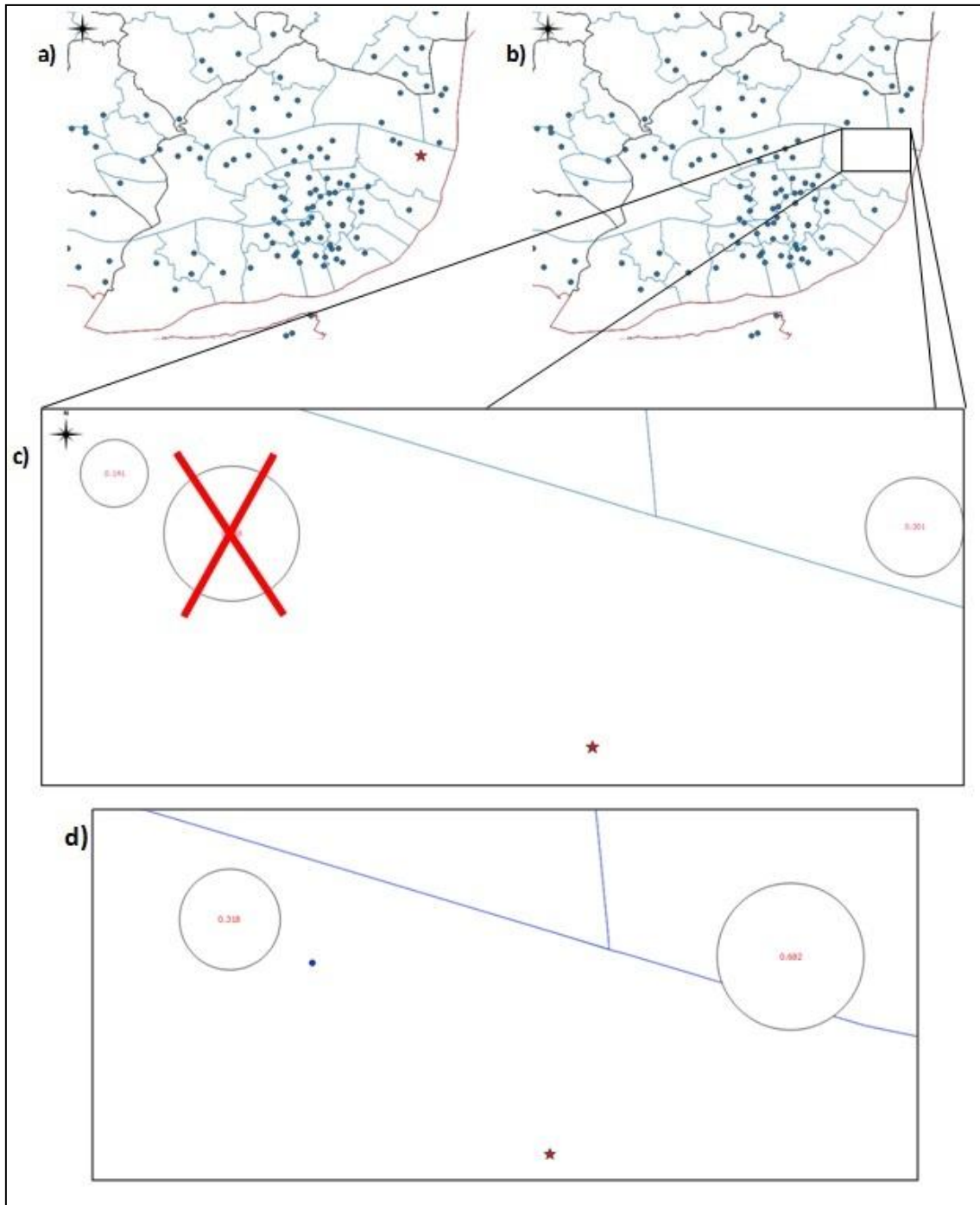


Figura 43 - Rede original de agências com assinalação da agência a eliminar da rede. a) Município de Lisboa com a localização das agências da CGD representadas por bolas azuis e a localização de um cliente representado por uma estrela vermelha (mais ou menos a meio à direita), b) Equivalente a a) mas com um retângulo por cima da área de análise, que se podem observar em c) Agências representadas por diagramas circulares e graduadas com a % Huff relativamente ao ponto selecionado, onde o X representa a agência que vai ser eliminada da área de análise e d) resultado da eliminação da agência representada por um ponto azul da rede (a mesma que tem a cruz em c).

Assim, após a remoção da agência, verificou-se um deslocamento dos clientes, para as duas agências restantes, passando estas a ter aproximadamente o dobro da percentagem de Huff que tinham inicialmente.

F. Cenário de análise das agências relativamente ao encerramento temporário de uma agência

Neste cenário assume-se que é necessário encerrar temporariamente uma agência, por exemplo para manutenção, e que se pretende encaminhar os clientes dessa agência para outra nas redondezas que seja considerada atrativa para os clientes, ou porque é perto ou porque o serviço é rápido ou por uma junção de características que iriam satisfazer o cliente encaminhado.

Assim, para calcular qual das agências mais próximas seria a mais indicada para encaminhar os clientes, procede-se à aplicação do modelo de Huff da mesma forma como foi exemplificado no cenário **A**, mas neste caso o ponto de origem seria a agência que iria estar encerrada, calculando assim as distâncias às restantes agências consideradas como potenciais alvos.

G. Cenário de análise de sobreposição de áreas de influência circulares

Este é um cenário um pouco diferente, em que se realiza inicialmente uma análise da rede através de áreas de influencia, desta análise pode observar-se zonas em que há mais ou menos sobreposição de agências. Daí definir qual a zona deve ser analisada através do modelo de Huff, por exemplo pode-se verificar pela Figura 44 o município de Lisboa sobreposto pelas agências da CGD, representadas a azul, sendo as três em análise preenchidas com cores diferentes, cada uma com a sua de área de influência circular com 1 km, observa-se então que duas das agências estão praticamente na mesma área havendo uma grande sobreposição. Pode considerar-se esta situação relevante e significativa à utilização do modelo de Huff. Pois assim pode calcular-se qual das agências tem a maior e a menor percentagem de ser visitada por um cliente, e utilizar esta percentagem como indicador, por exemplo para o encerramento se existir alguma que se verifique ter uma percentagem muito pequena.

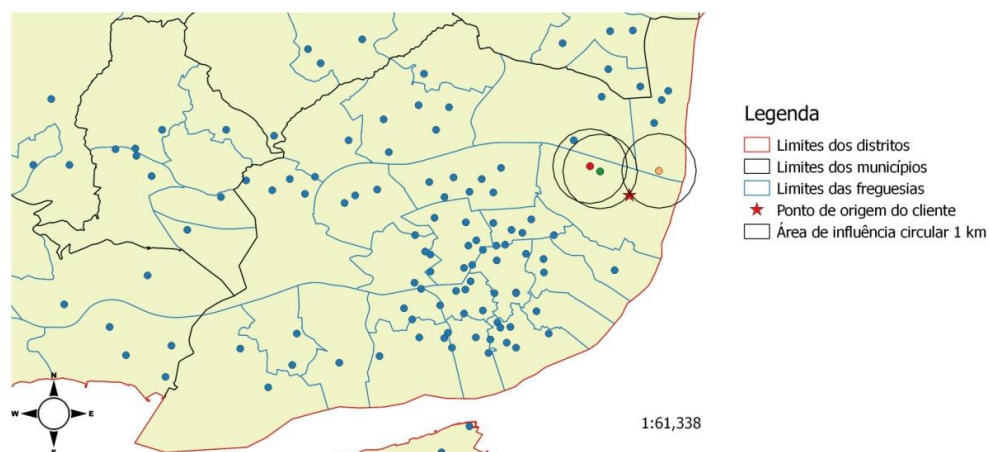


Figura 44- Áreas de influência circular, raio de 1 km em linha reta.

Em síntese, os cenários expostos, procuram refletir algumas das utilizações práticas em que a abordagem metodológica proposta neste trabalho, pode ser utilizada na gestão da rede de agências bancárias da CGD.

3.4. Sumário do capítulo

Neste capítulo foram descritos os passos para a concretização de cada dos objetivos propostos inicialmente: conhecimento dos dados, que consistiu na transformação de dados e a sua visualização por meio de mapas, obtenção da matriz de distâncias que fornece a distância e tempo de percursos realistas a pé ou de carro entre dois pontos, aplicação do modelo gravitacional de Huff, que faculta a percentagem de potencial de interesse por uma agência e visualização dos dados, onde são visualizadas as percentagens de Huff por diagramas circulares. Para além dessa explicação foram expostos alguns cenários de utilização da metodologia proposta, com o intuito de mostrar a sua possível utilidade na gestão da rede de agências de uma agência bancária.

Capítulo 4 – Conclusões

No presente capítulo são descritas as principais contribuições resultantes deste projeto e competências adquiridas. Ao longo na sua execução salientam-se também algumas dificuldades encontradas e ideias de trabalho futuro.

4.1. Principais contribuições

Através do Objetivo 1, foi necessário transformar dados de várias fontes, possibilitando-se a elaboração de mapas temáticos geográficos, que tornam possível a visualização da localização das agências da CGD relativamente às agências concorrentes e também ao mercado onde estão inseridas, conseguindo-se assim um conhecimento dos mesmos visualmente.

Foram também desenvolvidos quatro programas, um deles para o Objetivo 2, que diz respeito ao cálculo da matriz de distâncias recorrendo ao serviço Distance Matrix da Google, o qual fornece automaticamente tempo e distância entre origens e destinos, referentes aos clientes e às agências respetivamente, usando diferentes modos de deslocação, a pé e de carro, de percursos realistas, que de outro modo, teria de ser obtida manualmente e possivelmente com recurso a simplificações, como o uso de distâncias lineares.

Os restantes três programas foram para o Objetivo 3, dois dos programas, dizem respeito ao cálculo da percentagem de Huff, considerando ou não as agências da concorrência, e um programa que encontra a percentagem máxima de Huff para o caso de serem considerados vários pontos de origem. Ficou disponível a alteração da medida ou fórmula de cálculo de atratividade, uma vez, que esta variável pode divergir de acordo com o negócio ou interesse do decisor.

Relativamente ao Objetivo 4, conseguiu-se recorrendo à utilização de um SIG - QGIS - a visualização de mapas com os resultados encontrados através do modelo de Huff por meio de diagramas circulares. Foram ainda apresentados alguns cenários que mostram como pode ser utilizada a metodologia proposta na análise da rede de agências da CGD.

Estas contribuições foram validadas numa reunião com o representante da CGD no projeto. Atualmente a CGD está a desenvolver abordagens semelhantes com as mesmas tecnologias, sendo admissível que esta implementação possa ser utilizada num contexto profissional.

4.2. Competências adquiridas

Este projeto foi bastante enriquecedor, pois possibilitou o aumento de contato e experiência com diversas ferramentas e linguagens de programação, nomeadamente R e JavaScript. Uma competência na programação foi a construção de código modular, tornando relativamente fácil alterações/adaptações por decisores, por exemplo, a função de cálculo da atratividade. Também aprendi a aceder a bibliotecas *online*, a Google Distance Matrix e Javascript API, ambas utilizadas para o cálculo da matriz de distâncias e a última usada para a conversão de

moradas em coordenadas geográficas. Para além destas competências, também planeei e executei testes de desempenho.

Outra das principais competências adquiridas foi o conhecimento e a manipulação de um SIG, especialmente o QGIS, para a visualização de informação geograficamente distribuída e escolha de pontos, que depois serviram de *input* para os programas de cálculo de matriz de distâncias e cálculo de percentagens de Huff. Neste âmbito adquiri também conhecimentos relacionados com o modelo gravitacional aplicado, incluindo a noção de fatores de atratividade.

4.3. Dificuldades encontradas

Pelo facto de necessitar de recorrer a bibliotecas da Google, tanto para a conversão de moradas em coordenadas geográficas como para o cálculo das distâncias realistas entre clientes e agências, uma das maiores limitações encontradas foram os limites de utilização impostos por pela biblioteca Distance Matrix da Google a qual só permite 100 elementos por interações de 10 em 10 segundos, e um máximo diário de 2500 elementos. Como consequência destas limitações foi necessário fazer alterações aos programas, para serem feitos pedidos em blocos de dados de menor dimensão e com pausas entre pedidos. Inevitavelmente o tempo total de execução dos programas ficou maior que o esperado, mas ainda assim comportável.

Outra dificuldade encontrada foi a conversão correta de códigos postais de 7 dígitos em coordenadas geográficas, tendo sido necessária a utilização de moradas completas quanto possível ou escrever também o nome do país. Ainda assim houve pontos que deram coordenadas geográficas na América do Sul devido ao facto de existir uma rua chamada Portugal.

4.4. Trabalho Futuro

Como proposta de trabalho futuro, um ponto a abordar seria otimizar o cálculo da matriz de distâncias de maneira a que as mensagens dos limites excedidos deixassem de aparecer. Também seria vantajoso, em vez de se inserir a localização dos destinos e os códigos dos destinos, inserir apenas o código dos destinos e toda a informação necessária ao cálculo das distâncias ser obtida internamente, passando a preencher-se apenas dois campos. Da Figura 23 ficariam apenas os campos (1) e (3).

Outro ponto importante a abordar, seria a calibragem do modelo de Huff, uma vez que poderia ser enriquecedor ter mais variáveis para a medida de atratividade, tanto com fatores internos como externos.

Alguns fatores internos poderiam ser a atratividade dos produtos e serviços, o número de lugares de estacionamento, a quantidade de clientes, o tamanho do negócio, a tipologia de serviços, entre outros. No cálculo destes fatores, poderiam ser utilizadas ponderações, principalmente para adequar o modelo de cálculo com parâmetros específicos (por exemplo, escalões) para desenvolver as análises pretendidas em função dos objetivos estratégicos.

Quanto a fatores externos, poderiam incluir-se variáveis representativas da vizinhança, nomeadamente os pontos comerciais, clientes, comércio (proximidade e tipologia), concorrência, população, infraestruturas de comunicação e instituições públicas, escolas, centros comerciais, parques industriais ou comerciais. Quanto mais informação sobre a envolvente externa à agência for incluída no modelo, melhor a aderência à realidade, produzindo análises mais fiáveis e melhor apoio à decisão.

Ainda relativamente ao modelo de Huff, incluir as agências da concorrência na análise, poderia ser realizada do modo complementar aos parágrafos anteriores, criando um escalão que represente o impacto das agências concorrentes no cálculo da atratividade das agências da CGD.

Com a maturidade da metodologia, poderiam começar a ser definidos critérios de utilização ou de decisão. Por exemplo, relativamente à atratividade, do ponto de vista da inserção de uma nova agência, podia-se considerar a percentagem de Huff mínima para a inserção da nova agência, em que X é uma percentagem previamente estipulada, por exemplo maior que 30%.

Quanto à visualização das percentagens de Huff, essas poderiam ser representadas no mapa sob a forma de polígonos em redor das agências, de modo a visualizar áreas de influência.

Referente aos três programas utilizadas nos processos de cálculo de matriz de distâncias, aplicação do modelo de Huff e visualização dos dados resultantes, para a sua utilização profissional deveriam ser integrados num sistema para possibilitar um fluxo de trabalho fluido.

Referências

- [1] Aranha, F.; Figoli, S. (2001). *Geomarketing: Memórias de viagem*. Disponível em: http://mundogeo.com/webinar/geofusion/memoriasdeviagem_FranciscoAranha_SusanaFigoli.pdf. Consultado em: 05-09-2015
- [2] Parente, J.; Kato, H. (2001). Área de influência: Um estudo no varejo de supermercados. *Revista de Administração de Empresas*. 41:2, 46-53.
- [3] Direção Geral do Território. *Carta administrativa oficial de Portugal*. Disponível em: http://dgterritorio.pt/cartografia_e_geodesia/cartografia/carta_administrativa_oficial_de_portugal_caop/. Consultado em: 11-02-2015.
- [4] Base de dados da Pordata. Disponível em: <http://www.pordata.pt/Municipios>. Consultado em: 12-02-2015.
- [5] Martins, E. (2011). *Estratégia de localização varejista: Aplicação do modelo de Huff a um survey na rua de comércio especializado de material eletro-eletrônico de Belo Horizonte*. Relatório de Mestrado em Administração, Faculdade Novos Horizontes.
- [6] Lima, A.; Teixeira, D. (2008). Localização varejista: Um estudo sobre a utilização do modelo de Huff para a tomada de decisões sobre localização. *Revista Gestão & Tecnologia*, 8:1, 1-19.
- [7] Ramos, J. (2012). *Mini-curso de Geomarketing e aplicações*. Conferência e Feira de Geomática e Soluções Geoespaciais.
- [8] Banco de Portugal. *Lista de Agências bancárias*. Disponível em: <https://www.bportugal.pt/pt-PT/Supervisao/Paginas/Instituicoesautorizadas.aspx>. Consultado em: 21-01-2015.
- [9] CTT. *Códigos postais*. Disponível em: http://www.ctt.pt/feapl_2/app/open/postalCodeSearch.jspx?lang=def. Consultado em: 16-06-2015.
- [10] Documentação do QGIS. Disponível em: http://www.qgis.org/pt_PT/site/. Consultado em: 16-02-2015.
- [11] Carnasciali, A.; Delazari, L. (2011). A localização geográfica como recurso organizacional: Utilização de sistemas especialistas para subsidiar a tomada de decisão locacional do setor bancário. *RAC*, Curitiba. 15:1, 103-125.
- [12] Almeida, C. (2015). *Gis solution for energy & utilities*. Relatório de Mestrado em Engenharia Informática, Faculdade de Ciências, Universidade de Lisboa.
- [13] Murad, A. (2003). Creating a GIS application for retail centers in Jeddah city. *International Journal of Applied Earth Observation and Geoinformation*. 4, 329-338.
- [14] Snyder, S.; et al (2008). Ecological criteria, participant preferences and location models: A GIS approach toward ATV trail planning. *Applied Geography*. 28, 248-258.

- [15] Xiang, W. (1996). A GIS based method for trial alignment planning. *Landscape and Urban Planning*. 35, 11-23.
- [16] Narumalani, S.; Zhou, Y.; Jensen, J. (1997). Application of remote sensing and geographic information systems to the delineation and analysis of riparian buffer zones. *Aquatic Botany*. 58, 393-409.
- [17] Atkinson, S.; Canter, L. (2011). Assessing the cumulative effects of projects using geographic information systems. *Environmental Impact Assessment Review*. 31, 457-464.
- [18] MacDonald, E. (2011). Top 10 ways to use GIS in retail banking. *Market Forté*. White paper. Disponível em: <http://www.instantsiteintelligence.com/MFL3/Whitepapers/WhitePaper-MarketForte-GISinBanking2.pdf>. Consultado em:17-05-2015
- [19] Carnasciali, A. (2007). *Integração de sistemas de informações geográficas e sistemas especialistas visando auxiliar a tomada de decisão locacional do setor bancário*. Relatório de Pós-graduação em Ciências Geodésicas, Universidade Federal do Pará.
- [20] Oliveira, S. (2012). *Fatores de atratividade: Por que estudantes escolhem o curso de administração da PUC Minas em Belo horizonte?*. Seminários em Administração.
- [21] Pan, H.; Li, Y.; Dang, A. (2013) Application of network Huff model for commercial network planning at suburban – Taking Wujin district, Changzhou as a case, *Annals of GIS*, 19:3, 131-141, DOI: 10.1080/19475683.2013.806356.
- [22] Martins, R. (2012). *Construção de um modelo de análise espacial em sig, que determine a localização ótima de equipamentos sociais para idosos, no concelho de Lisboa*. Relatório de Mestrado em Gestão do Território na Área de Especialização em Detecção Remota e Sistemas de Informação Geográfica, Faculdade de Ciências Sociais e Humanas, Universidade Nova de Lisboa.
- [23] Santiago, M. (2008). *Gestão de Marketing*. IESDE Brasil, ISBN: 978-85-387-2166-6.
- [24] Mattar, F. (2011). *Administração de Varejo*. Elsevier, ISBN: 978-85-352-3219-6.
- [25] Amholdt, H. (2014). *O conceito da área de influência no varejo*. Disponível em: <http://www.arnholdt.com.br/o-conceito-da-area-de-influencia-no-varejo/>. Consultado em: 12-03-2015
- [26] Fernández, j.; Hendrix, M. (2013). Recent insights in Huff-like competitive facility location and design. *European Journal of Operational Research*. 227, 581-584.
- [27] Kuçukaydin, H.; Aras, N.; Altinel, I. (2011). Competitive facility location problem with attractiveness adjustment of the follower: A bilevel programming model and its solution. *European Journal of Operational Research*. 208, 206-220.
- [28] Directions Magazine. *Retail trade area analysis using the Huff model*. Disponível em: <http://www.directionsmag.com/entry/retail-trade-area-analysis-using-the-huff-model/23411>. Consultado em: 01-03-2015.

- [29] Elias, D. (2008). Localização de lojas no varejo: Uma análise de perfis de consumidores e proposição de modelo de cobertura. Relatório de Pós-Graduação em Administração de Organizações. Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto. Universidade de São Paulo.
- [30] Única 360. *Áreas de influências, tipos e aplicações em geomarketing*. Disponível em: <http://www.unica360.com/areas-de-influencia-tipos-y-aplicaciones-en-geomarketing>. Consultado em: 11-06-2015.
- [31] Directions Magazine. Not your father's trade area ... beyond block group and census tract. Disponível em: <http://www.directionsmag.com/entry/not-your-fathers-trade-area-...-beyond-block-group-and-census-tract/250615>. Consultado em: 11-06-2015.
- [32] Directions Magazine. *Retail trade area analysis: concepts and new approaches*. Disponível em: <http://www.directionsmag.com/entry/retail-trade-area-analysis-concepts-and-new-approaches/129766>. Consultado em: 11-06-2015.
- [33] Gomes, N. et al (2010). Áreas de influência e população servida por centros comerciais na área metropolitana de Lisboa: uma análise espaço-temporal. Encontro de utilizadores de informação geográfica.
- [34] LibreOffice Calc 3.6. Disponível em: <https://pt-br.libreoffice.org/baixe-ja/notas-da-versao/nova-pagina-4/>. Consultado em: 16-06-2015.
- [35] GPSVisualizer. Disponível em: <http://www.gpsvisualizer.com>. Consultado em: 26-04-2015.
- [36] Exemplos de utilização da Google Maps JavaScript API. Disponível em: <https://developers.google.com/maps/documentation/javascript/examples/map-simple>. Consultado em: 22-06-2015.
- [37] Topografix. *GPX the GPS Exchange format*. Disponível em: <http://www.topografix.com/gpx.asp>. Consultado em: 27-09-2015.
- [38] R e sua documentação. Disponível em: <https://www.r-project.org/>. Consultado em: 16-06-2015.
- [39] Documentação da Google Maps JavaScript API: Distance Matrix Service. Disponível em: <https://developers.google.com/maps/documentation/javascript/distancematrix>. Consultado em: 22-06-2015.
- [40] Flanagan, D. (2011). *JavaScript: the definitive guide*. O'Reilly, ISBN: 978-0-596-80552-4.

Apêndices

Apêndice A - Obtenção da matriz de distâncias

```
<!-- -----
# Projeto CGD - Inês Ferreira - julho 2015

# Calcula distâncias com diversos modos de deslocação e de distância.
# ----- -->
<!DOCTYPE html>
<html>
  <head>
    <meta http-equiv="Content-Type" content="text/html; charset=utf-8"/>
    <title>Distance Matrix service</title>
    <script
src="https://maps.googleapis.com/maps/api/js?v=3.exp&signed_in=true"></script>

    <style>
      html, body {
        height: 100%;
        margin: 0;
        padding: 0;
      }

      #map-canvas {
        height: 100%;
        width: 18%;
      }
      #content-pane {
        float:right;
        width:80%;
        padding-left: 2%;
        padding-top: 2%;
      }
      #outputDiv {
        font-size: 11px;
        padding-top: 30%;
      }
      #inputs {
        font-size: 11px;
      }
    </style>
  </head>

  <body>
    <div id="content-pane">
      <div id="inputs">

        <div style="border: 2px solid BLUE; margin: 2px; height: 20%; width: 30%; float:
left;">
          1
          <p><b>Origens (Clientes):</b></p>
          <p>Pode ser introduzida a coordenada geográfica, morada ou código-postal:
</p><br />
          <textarea id="origens" cols="38" rows="7" type="text" >38.729451,-
9.149476</textarea> <br /><br />
          </div>

          <div style="border: 2px solid BLUE; margin: 2px; height: 20%; width: 30%; float:
left;">
            2
            <p><b>Destinos (agências):</b></p>
            <p>Pode ser introduzida a coordenada geográfica, morada ou código-postal:
</p><br />
            <textarea id="destinos" cols="38" rows="7" type="text" >38.711775,-9.12682
Rua de Santa Justa, Lisboa, 1100-998</textarea><br /><br />
            </div>

          <div style="border: 2px solid BLUE; margin: 2px; height: 20%; width: 30%; float:
left;">
            3
```

```

        <p><b>Códigos das agências destino:</b></p>
        <p>Introduza um identificador unico do destino extamente pela mesma ordem. </p>
        <textarea id="codigoAgencia" cols="38" rows="7">12
32</textarea><br /><br />
    </div>

    <div style="border: 2px solid BLUE; margin: 2px; height: 30%; width: 30%; float:
left;">
        4
        <p>Após campos inseridos clicar em:
        <button type="button" onclick="calculateDistances();">Calcular
distâncias</button></p>
    </div>
    </div>

    <div id="outputDiv"><p><b>Resultados:</b></p></div>
</div>

<div id="map-canvas"></div>
</body>

<!-- ----- -->

<script>
var map;
var geocoder;
var bounds = new google.maps.LatLngBounds ();
var markersArray = [];

var destinationIcon =
'https://chart.googleapis.com/chart?chst=d_map_pin_letter&chld=D|FF0000|000000';
var originIcon =
'https://chart.googleapis.com/chart?chst=d_map_pin_letter&chld=O|FFFF00|000000';

// -----
// BLOCO DE VARIÁVEIS E CONSTANTES USADAS PARA O CÁLCULO DA MATRIZ DISTÂNCIAS.

var todasOrigensOriginal = [
    new google.maps.LatLng(38.7188450122431, -9.1383259876333)
]

var todosDestinosOriginal = [
    new google.maps.LatLng(38.711775,-9.12682),
    new google.maps.LatLng(38.729451,-9.149476),
    new google.maps.LatLng(38.722776,-9.135472),
    new google.maps.LatLng(38.71199,-9.153875),
    new google.maps.LatLng(38.723571,-9.14923),
    new google.maps.LatLng(38.722698,-9.141294),
    new google.maps.LatLng(38.723304,-9.151692),
    new google.maps.LatLng(38.710322,-9.142478),
    new google.maps.LatLng(38.725903,-9.14778),
    new google.maps.LatLng(38.731368,-9.1407),
    new google.maps.LatLng(38.734337,-9.144447),
    new google.maps.LatLng(38.728265,-9.128041),
    new google.maps.LatLng(38.717827,-9.130076),
    new google.maps.LatLng(38.735705,-9.138428),
    new google.maps.LatLng(38.710954,-9.146306),
    new google.maps.LatLng(38.73186,-9.128244),
    new google.maps.LatLng(38.717305,-9.149562),
    new google.maps.LatLng(38.708089,-9.135497),
    new google.maps.LatLng(38.719361,-9.155927),
    new google.maps.LatLng(38.709474,-9.13805),
    new google.maps.LatLng(38.714809,-9.140587),
    new google.maps.LatLng(38.708156,-9.152719),
    new google.maps.LatLng(38.722684,-9.13547),
    new google.maps.LatLng(38.735386,-9.140942),
    new google.maps.LatLng(38.718118,-9.144453),
    new google.maps.LatLng(38.706576,-9.142922),
    new google.maps.LatLng(38.733218,-9.132949),
    new google.maps.LatLng(38.713715,-9.137052),
    new google.maps.LatLng(38.713455,-9.139831),
    new google.maps.LatLng(38.730184,-9.147298)]

```

```

var separadorRegistro = "\n";
var separadorCampo = ",";

var todasOrigensTextoPorOmissao = "38.7188450122431,-9.1383259876333";

var todosDestinosTextoPorOmissao =
"38.711775,-9.12682" + separadorRegistro +
"38.729451,-9.149476" + separadorRegistro +
"38.722776,-9.135472";

var todosC_AGTextoPorOmissao = "41" + separadorRegistro + "97" + separadorRegistro +
"100";

// Indicadores da posição inicial dos vetores origins e destinations.
var posicaoInicialOrigem = 0; // 133
var posicaoInicialDestino = 0; // 50

// API Google permite 4 origens e 25 destinos em cada cálculo na matriz distâncias.
const TAMANHO_BLOCO_ORIGENS = 1; // 4
const TAMANHO_BLOCO_DESTINOS = 25; // 25

// Modos de deslocação a usar no cálculo da matriz distâncias.
const MODOS_DESLOCACAO = [ google.maps.TravelMode.DRIVING, google.maps.TravelMode.WALKING ];
var posicaoModoDeslocacao = 0;

// Tempo entre invocações sucessivas do cálculo das distâncias.
const TEMPO_PAUSA = 10000; // 10 segundos.

var expressaoRegularCoordenadas = /^s*(\d+(\.\d+)?)s*,s*(\d+(\.\d+)?)s*$/;

// "1,2".search(expressaoRegularCoordenadas);

// ----- Testes de formulário

var todasOrigensTexto = document.getElementById("origens").value;

if (todasOrigensTexto == "") {
    todasOrigensTexto = todasOrigensTextoPorOmissao;
}

var todosDestinosTexto = document.getElementById("destinos").value;

if (todosDestinosTexto == "") {
    todosDestinosTexto = todosDestinosTextoPorOmissao;
}

var todosC_AGTexto = document.getElementById("codigoAgencia").value;

if (todosC_AGTexto == "") {
    todosC_AGTexto = todosC_AGTextoPorOmissao;
}

var C_AG = todosC_AGTexto.split(separadorRegistro);

//var todasOrigensTexto = auxiliarVetores[0];
//var todosDestinosTexto = auxiliarVetores[1];
//var C_AGTexto = auxiliarVetores[2];

// ----- Fim de teste formulário.

var todosDestinos = converteTextoParaCoordenadas(todosDestinosTexto);
var todasOrigens = converteTextoParaCoordenadas(todasOrigensTexto);
//var C_AG = document.getElementById("codigoAgencia").value.split("\n");

// -----

function initialize() {
    var opts = {

```

```

        center: new google.maps.LatLng(38.741528, -9.141671), // Lisboa
        zoom: 8
    };
    map = new google.maps.Map(document.getElementById('map-canvas'), opts);
    geocoder = new google.maps.Geocoder();
}

function calculateDistances() {
    var service = new google.maps.DistanceMatrixService();

    // -----

    // Origens e destinos a considerar para a invocação atual do cálculo da matriz distâncias.
    var blocoOrigens = [];
    var blocoDestinos = [];
    var numeroOrigensBloco = TAMANHO_BLOCO_ORIGENS;
    var numeroDestinosBloco = TAMANHO_BLOCO_DESTINOS;

    if ( (todasOrigens.length - posicaoInicialOrigem) < TAMANHO_BLOCO_ORIGENS ) {
        numeroOrigensBloco = todasOrigens.length - posicaoInicialOrigem;
    }

    for (var indiceOrigem = 0; indiceOrigem < numeroOrigensBloco; indiceOrigem++) {
        blocoOrigens[indiceOrigem] = todasOrigens[posicaoInicialOrigem + indiceOrigem];
    }

    console.log(blocoOrigens)

    if ( (todosDestinos.length - posicaoInicialDestino) < TAMANHO_BLOCO_DESTINOS ) {
        numeroDestinosBloco = todosDestinos.length - posicaoInicialDestino;
    }

    for (var indiceDestino = 0; indiceDestino < numeroDestinosBloco; indiceDestino++) {
        blocoDestinos[indiceDestino] = todosDestinos[posicaoInicialDestino + indiceDestino];
    }

    console.log(blocoDestinos)

    // for (modoDeslocacao in MODOS_DESLOCACAO) { // Mudar: passa a ser controlado pela
    calculaProximaMatrizDistancias.
        service.getDistanceMatrix(
            {
                origins: blocoOrigens,
                destinations: blocoDestinos,
                travelMode: MODOS_DESLOCACAO[posicaoModoDeslocacao],
                unitSystem: google.maps.UnitSystem.METRIC,
                avoidHighways: false,
                avoidTolls: false
            }, recebeMatrizDistancias);

    // }

    // -----
}

function recebeMatrizDistancias(response, status) {
    if (status != google.maps.DistanceMatrixStatus.OK) {
        alert('Erro foi: ' + status);
    } else {
        var origins = response.originAddresses;
        var destinations = response.destinationAddresses;
        var outputDiv = document.getElementById('outputDiv');
        // outputDiv.innerHTML = '';
        deleteOverlays();

        for (var indiceOrigem = 0; indiceOrigem < origins.length; indiceOrigem++) {
            var results = response.rows[indiceOrigem].elements;

            addMarker(origins[indiceOrigem], false);

            for (var indiceDestino = 0; indiceDestino < results.length; indiceDestino++) {

```

```

        addMarker(destinations[indiceDestino], true);

        // Mostra a linha atual da matriz distâncias.
        outputDiv.innerHTML += [indiceOrigem+1] + ';' + [indiceDestino+1] + ';' + [indiceOrigem + posicaoInicialOrigem] + ';' + [indiceDestino+posicaoInicialDestino] + ';' + origins[indiceOrigem] + ';' + destinations[indiceDestino] + ';' + 'Distância ;' + results[indiceDestino].distance.value + ';' + MODOS_DESLOCACAO[posicaoModoDeslocacao] + ';' + PIDEST=' + posicaoInicialDestino + ';' + PIORig=' + posicaoInicialOrigem + ';' + C_AG[indiceDestino+posicaoInicialDestino] + '<br>' + [indiceOrigem+1] + ';' + [indiceDestino+1] + ';' + [indiceOrigem + posicaoInicialOrigem] + ';' + [indiceDestino+posicaoInicialDestino] + ';' + origins[indiceOrigem] + ';' + destinations[indiceDestino] + ';' + 'Tempo;' + results[indiceDestino].duration.value/60 + ';' + MODOS_DESLOCACAO[posicaoModoDeslocacao] + ';' + PIDEST=' + posicaoInicialDestino + ';' + PIORig=' + posicaoInicialOrigem + ';' + C_AG[indiceDestino+posicaoInicialDestino] + '<br>';
    }
}

// Deixa passar algum tempo até fazer o próximo pedido de cálculo de matriz
// Distâncias por restrição da Google.
setTimeout(calculaProximaMatrizDistancias, TEMPO_PAUSA);
}

function addMarker(location, isDestination) {
    var icon;
    if (isDestination) {
        icon = destinationIcon;
    } else {
        icon = originIcon;
    }
    geocoder.geocode({'address': location}, function(results, status) {
        if (status == google.maps.GeocoderStatus.OK) {
            bounds.extend(results[0].geometry.location);
            map.fitBounds(bounds);
            var marker = new google.maps.Marker({
                map: map,
                position: results[0].geometry.location,
                icon: icon
            });
            markersArray.push(marker);
        } else {
            alert('Geocode was not successful for the following reason: ' + status);
        }
    });
}

function deleteOverlays() {
    for (var i = 0; i < markersArray.length; i++) {
        markersArray[i].setMap(null);
    }
    markersArray = [];
}

google.maps.event.addDomListener(window, 'load', initialize);

// Ciclo de destinos 10 em 10
function contatodos() {
    while (posicaoInicialDestino < 90)
    {
        calculateDistances();
        posicaoInicialDestino = posicaoInicialDestino + 25 ;
    }
}

```

```

    }
}

// Auxiliar para testes. A retirar em breve.
function proximoBlocoOrigens() {
    posicaoInicialOrigem += TAMANHO_BLOCO_ORIGENS;
}

// Auxiliar para testes. A retirar em breve.
function proximoBlocoDestinos() {
    posicaoInicialDestino += TAMANHO_BLOCO_DESTINOS;
}

function calculaProximaMatrizDistancias() {

    posicaoInicialDestino += TAMANHO_BLOCO_DESTINOS;
    if (posicaoInicialDestino >= todosDestinos.length) {
        posicaoInicialDestino = 0;
        posicaoInicialOrigem += TAMANHO_BLOCO_ORIGENS; // ver proximo if().
    }

    if (posicaoInicialOrigem < todasOrigens.length) {
        calculateDistances();
    }
    else {
        // Avança para o proximo modo deslocação se possivel.
        posicaoInicialDestino = 0;
        posicaoInicialOrigem = 0;
        posicaoModoDeslocacao += 1;
        calculateDistances();
    }
}

// Preenche o vetor de origens os destinos de acordo com a informação do ponto inserida.
function converteTextoParaCoordenadas(coordenadasEmTexto) {
    // Assume-se que cada linha guarda um local.
    var textoEmArray = coordenadasEmTexto.split(separadorRegisto);
    var arrayCoordenadas = [];

    // Só aplicável caso sejam usadas coordenadas.
    var parXY = [];
    var x = 0; // Longitude, que para Portugal é negativa.
    var y = 0; // Latitude, que para Portugal é positiva.

    for (var indiceLinha = 0; indiceLinha < textoEmArray.length; indiceLinha++) {
        // Caso a linha guarde um par x = longitude ,y = latitude do local.
        if (textoEmArray[indiceLinha].search(expressaoRegularCoordenadas) == 0) {
            // Registo com coordenadas.
            parXY = textoEmArray[indiceLinha].split(separadorCampo);
            if (parXY[0] < 0) x = parXY[0]; else y = parXY[0];
            if (parXY[1] < 0) x = parXY[1]; else y = parXY[1];

            arrayCoordenadas[indiceLinha] = new google.maps.LatLng(y, x);
        }
        else { // Caso a linha guarde uma morada local.
            arrayCoordenadas[indiceLinha] = textoEmArray[indiceLinha];
        }
    }

    return arrayCoordenadas;
}

</script>
</html>

```

Apêndice B - Cálculo das percentagens de Huff para agências da CGD

```
# -----
# Projeto CGD - Inês Ferreira - julho 2015
# Calcular % Huff para agências da CGD.

# Funciona para diversos modos de deslocação e de distância e junção das tabelas.
# beta também pode ser alterado.
# -----
getwd()
# Caminho para a o sitio onde a informação necessária está guardada.
setwd("C:/Users/Maria/Desktop/Apres_CGD")

matriz.distancias <- read.table("MatrizDistancias.csv", header = F, sep = ";");
head(matriz.distancias)
matriz.distancias[c(7,9)]

CGD<-read.table("InfoCGD.csv",header=T,sep=";")
CGD$Município<- as.character(CGD$Município);
CGD$Distrito<- as.character(CGD$Distrito);
CGD$NM_AG <- as.character(CGD$NM_AG);
CGD$CODIGO_POSTAL <- as.character(CGD$CODIGO_POSTAL);
# temporario

names(matriz.distancias)

# -----
names(matriz.distancias) <- c("OP", "DP", "IVO", "IVD", "Origem", "Destino",
  "medida.distancia", "TD", "modo.deslocacao",
  "pid", "pio", "C_AG");

# Medidas de distância podem ser obtidas automaticamente ou escolhidas pelo utilizador.
# MEDIDAS.DISTANCIA <- levels(matriz.distancias$medida.distancia);
MEDIDAS.DISTANCIA <- c("Distância", "Tempo");

# Modos de deslocação podem ser obtidos automaticamente ou escolhidos pelo utilizador.
# MODOS.DESLOCACAO <- levels(matriz.distancias$modo.deslocacao);
MODOS.DESLOCACAO <- c("DRIVING", "WALKING");

CASAS.DECIMAIIS <- 3;

# coeficiente distância
beta <- 2;

matriz.distancias$atratividade <- 0; # Criação da coluna onde estão inseridas as atrati-
vidade das agências.
matriz.distancias$Huff <- 0; # Criação da coluna onde vão ser inseridas as % de Huff.

#for (colunas.cgd in 1:ncol(CGD)){
# matriz.distancias$CGD[,colunas.cgd] <- "";
#}

# -----

# Fórmula de atratividade pode ser alterada pelo utilizador.
# Podem ser usadas na fórmula todas as colunas do data.frame das agências.
calcula.atratividade <- function(dados.agencia) {
  return(dados.agencia$NEGOCIO_201312);
}

# -----

##preencher posições do Huff com atratividade o NEGOCIO 201312.
calcula.huff <- function(medida.distancia, modo.deslocacao) {

  # Numerador do modelo de Huff para a medida atratividade e distância e modo de deslo-
  cação.
  # Atratividade / ( valor medida de distancia )^2

  matriz.distancias$Numerador <- 0; # coluna auxiliar para o calculo da % de Huff.
```

```

# Procurar na tabela CGD o valor correspondente ao da matriz.distancias.
for (linha.matriz in 1:nrow(matriz.distancias)) {
  for (linha.CGD in 1:nrow(CGD)) {
    if ((matriz.distancias$C_AG[linha.matriz]) == (CGD$C_AG[linha.CGD])) {
      # matriz.distancias$Numerador[linha.matriz] <-
CGD[[medida.atratividade]][linha.CGD] / (matriz.distancias$TD[linha.matriz]^beta);
      matriz.distancias$Numerador[linha.matriz] <- calcula.atratividade(CGD[linha.CGD,
]) / (matriz.distancias$TD[linha.matriz]^beta);
      matriz.distancias$atratividade[linha.matriz] <<- calcula.atratividade(CGD[linha.CGD, ]);
    }
  }
}

# Denominador do modelo de Huff para a medida atratividade e distância e modo de des-
locação.
denominador <- 0;

for (numero.linha in 1:nrow(matriz.distancias)) {
  linha.atual <- matriz.distancias[numero.linha, ];
  if (linha.atual$medida.distancia == medida.distancia && linha.atual$modo.deslocacao
== modo.deslocacao)
    denominador <- denominador + linha.atual$Numerador;
}

# Calculo da % de Huff. Numerador(i)/denominador.
for (numero.linha in 1:nrow(matriz.distancias)) {
  # for (numero.linha in 1:5) {
  linha.atual <- matriz.distancias[numero.linha, ];
  if (linha.atual$medida.distancia == medida.distancia && linha.atual$modo.deslocacao
== modo.deslocacao)
    matriz.distancias$Huff[numero.linha] <<- round( linha.atual$Numerador / denomi-
nador, CASAS.DECIMAIAS);
  # print(linha.atual);
  # print(matriz.distancias$Huff[numero.linha]);
}
# MatrizDistanciaHuff <- matriz.distancias;
# return(MatrizDistanciaHuff)
}

# -----

# Chamada da função calculaHuff para os vários modos de deslocção e MEDIDAS de distân-
cia.

for (indice.deslocacao in 1:length(MODOS.DESLOCACAO)){ # Para cada modo de deslocção.
  for (indice.distancia in 1:length(MEDIDAS.DISTANCIA)){ # Vê todas as MEDIDAS de dis-
tância.
    calcula.huff(medida.distancia = MEDIDAS.DISTANCIA[indice.distancia],
modo.deslocacao = MODOS.DESLOCACAO[indice.deslocacao]);
    # print(MEDIDAS.DISTANCIA[indice.distancia]);
    # print(MODOS.DESLOCACAO[indice.deslocacao]);
  }
}

#write.table(matriz.distancias, "MatrizDistanciaHuffhuff0.csv", row.names = F, sep =
";", quote = F);

# Limpeza da tabela.
matriz.distancias= matriz.distancias[,c(-1, -2, -10, -11)]
# Cria um ficheiro com a matriz distancias mais a %huff.
#write.table(matriz.distancias, "MatrizDistanciaHuff.csv", row.names = F, sep = ";",
quote = F);

# Acrescentar todas a colunas correspondentes ao ficheiro da CGD.
# Começa na coluna 2 para não ser substituído o C AG.
for (colunas.cgd in 2:ncol(CGD)){
  nome.coluna<-names(CGD)[colunas.cgd] ;

```



```

print(nome.coluna);
matriz.distancias[[nome.coluna]] <-"";
}

# Preencher as colunas com o valores correspondentes.

for (colunas.cgd in 2:ncol(CGD)){
  for (coluna.matriz.distancias in 10:ncol(matriz.distancias)){
    nome.coluna.MD <- names(matriz.distancias)[coluna.matriz.distancias];
    nome.coluna.CGD <- names(CGD)[colunas.cgd];
    if (nome.coluna.MD == nome.coluna.CGD) {
      for (linha.matriz in 1:nrow(matriz.distancias)) {
        for (linha.CGD in 1:nrow(CGD)) {
          if ((matriz.distancias$C_AG[linha.matriz]) == (CGD$C_AG[linha.CGD]) ) {
            matriz.distancias[[nome.coluna.MD]][linha.matriz] <-
CGD[[nome.coluna.CGD]][linha.CGD];
          }
        }
      }
    }
  }
}

write.table(matriz.distancias, "MatrizDistanciasHuffCGD.csv", row.names = F, sep = ";",
quote = F);

```


Apêndice C - Cálculo das percentagens de Huff para agências da CGD e da concorrência

```
# ----
# Projeto CGD - Inês Ferreira - julho 2015
# Calcular % Huff para agências da CGD e da concorrência.

# Funciona para diversos modos de deslocação e de distancia e junção das tabelas.
# beta também pode ser alterado.
# ----
getwd()
# Caminho para a o sitio onde a informação necessária está guardada.

setwd("C:/Users/Maria/Desktop/Apres_CGD")

matriz.distancias <- read.table("MatrizDistancias.csv", header = F, sep = ";");
head(matriz.distancias)
names(matriz.distancias) <- c("OP", "DP", "IVO", "IVD", "Origem", "Destino",
                             "medida.distancia", "TD", "modo.deslocacao",
                             "pid", "pio", "C_AG");

CGD <- read.table("InfoCGD.csv", header=T, sep=";")
concorrencia <- read.table("InfoConcorrencia.csv", header=T, sep=";");
#concorrencia = concorrencia[,c(-7, -8, -11, -12, -13, -14, -15, -17, -18,
#                               -24, -25, -26, -27, -28, -29, -30, -31, -33,
#                               -35, -40)];
# Mudança do nome da coluna Concelho para Município.
names(concorrencia) <-
  c("i..Grupo.Manter", "Nome.Grupo",
    "Denominação.da.Instituição", "Tipo.de.Estabelecimento",
    "Cód.AF", "Denominação.do.Estabelecimento",
    "País.de.Estabelecimento", "Cód.Agência",
    "Sucursal.Balcão.sede", "Localidade",
    "Código.Postal", "Município",
    "Distrito", "País",
    "Banco.Sigla", "Sigla.Agencia",
    "ChaveMorada", "Latitude",
    "Longitude", "Idnt.Agencia");

dados.populacao <- read.table("DadosPopulacao.csv", header=T, sep="");
names(dados.populacao) <-
  c("Chave.Geográfico",
    "Âmbito.Geográfico",
    "Município",
    "Densidade.populacional",
    "Habitantes.por.caixa.automática.multibanco",
    "Total.Receitas",
    "Receitas.efectivas",
    "Activos.financeiros",
    "Passivos.financeiros",
    "Despesas",
    "Receitas",
    "Saldo",
    "Ganho.mensal.total",
    "Ganho.mensal.masculino",
    "Ganho.mensal.feminino",
    "Poder.de.compra",
    "População.residente",
    "Índice.de.envelhecimento.racio",
    "Nº.agregados",
    "Nº.agregados.de.1",
    "Nº.agregados.de.2",
    "Nº.agregados.de.3.5",
    "Nº.agregados.de.6.",
    "Taxa.de.desemprego.total")
# temporario

names(matriz.distancias)

# Para que possam ser comparados para o cálculo da atratividade.
```

```

CGD$Município<- as.character(CGD$Município);
concorrençia$Município<- as.character(concorrençia$Município);
dados.populacao$Município<- as.character(dados.populacao$Município);
CGD$NM_AG <- as.character(CGD$NM_AG);
CGD$CODIGO_POSTAL <- as.character(CGD$CODIGO_POSTAL);

CGD$Distrito<- as.character(CGD$Distrito);
concorrençia$Distrito<- as.character(concorrençia$Distrito);

concorrençia$Idnt.Agencia<-as.character(concorrençia$Idnt.Agencia)
CGD$C_AG<-as.character(CGD$C_AG)
matriz.distancias$C_AG<-as.character(matriz.distancias$C_AG)
# -----

# Medidas de distância podem ser obtidas automaticamente ou escolhidas pelo utilizador.
# MEDIDAS.DISTANCIA <- levels(matriz.distancias$medida.distancia);
MEDIDAS.DISTANCIA <- c("Distância", "Tempo");

# Modos de deslocação podem ser obtidos automaticamente ou escolhidos pelo utilizador.
# MODOS.DESLOCACAO <- levels(matriz.distancias$modo.deslocacao);
MODOS.DESLOCACAO <- c("DRIVING", "WALKING");

CASAS.DECIMAIIS <- 3;

# coeficiente distância.
beta <- 2;

matriz.distancias$atratividade <- 0; # Criação da coluna onde estão inseridas as atrati-
vidade das agências.
matriz.distancias$Huff <- 0; # Criação da coluna onde vão ser inseridas as % de Huff.
concorrençia$atratividade_ag <- 0;
CGD$atratividade_ag <- 0;

# -----

# Formula de atratividade pode ser alterada pelo utilizador consoante o calculo a ser
feito.
# Nomes de colunas que podem ser utilizados na atratividade são os nomes das colunas do
data.frame das agencias.
# verificar em que municipio esta a agencia e atribuir a atratividade correspondente.
# Vai ser considerado como medida de atratividade o poder de compra do municipio.
calcula.atratividade <- function(dados.agencia) {
  for (linha.agencia in 1:nrow(dados.agencia)) {
    for (linha.dados.populacao in 1:nrow(dados.populacao)) {
      if(dados.populacao$Município[linha.dados.populacao] ==
dados.agencia$Município[linha.agencia]){
        dados.agencia$atratividade_ag[linha.agencia] <-
dados.populacao$Poder.de.compra[linha.dados.populacao];
      }
    }
  }
  return(dados.agencia$atratividade_ag)
}

# -----

##preencher posições do Huff com atratividade o NEGOCIO 201312.
calcula.huff <- function(medida.distancia, modo.deslocacao) {

  # Numerador do modelo de Huff para a medida atratividade e distância e modo de deslo-
cação.
  # Atratividade / ( valor medida de distancia )^2

  matriz.distancias$Numerador <- 0; # coluna auxiliar para o cálculo da % de Huff.

  # Procurar na tabela CGD o valor correspondente ao da matriz.distancias.
  for (linha.matriz in 1:nrow(matriz.distancias)) {
    for (linha.CGD in 1:nrow(CGD)) {
      if ((matriz.distancias$C_AG[linha.matriz]) == (CGD$C_AG[linha.CGD])) {
        # matriz.distancias$Numerador[linha.matriz] <-
CGD[[medida.atratividade]][linha.CGD] / (matriz.distancias$TD[linha.matriz]^beta);
      }
    }
  }
}

```

```

        matriz.distancias$Numerador[linha.matriz] <- calcula.atratividade(CGD[linha.CGD,
]) / (matriz.distancias$TD[linha.matriz]^beta);
        matriz.distancias$atratividade[linha.matriz] <<- calcula.atratividade(CGD[linha.CGD,
]);
    }
}

# Procurar na tabela concorrência o valor correspondente ao da matriz.distancias.
for (linha.matriz in 1:nrow(matriz.distancias)) {
    for (linha.concorrencia in 1:nrow(concorrencia)) {
        if ((matriz.distancias$C_AG[linha.matriz]) == (concorrencia$Idnt.Agencia[linha.concorrencia])) {
            # matriz.distancias$Numerador[linha.matriz] <- concorrencia[[medida.atratividade]][linha.CGD] / (matriz.distancias$TD[linha.matriz]^2);
            matriz.distancias$Numerador[linha.matriz] <- calcula.atratividade(concorrencia[linha.concorrencia,
]) / (matriz.distancias$TD[linha.matriz]^2);
            matriz.distancias$atratividade[linha.matriz] <<- calcula.atratividade(concorrencia[linha.concorrencia,
]);
        }
    }
}

# Denominador do modelo de Huff para a medida atratividade e distância e modo de deslocamento.
denominador <- 0;

for (numero.linha in 1:nrow(matriz.distancias)) {
    linha.atual <- matriz.distancias[numero.linha,
];
    if (linha.atual$medida.distancia == medida.distancia && linha.atual$modo.deslocacao == modo.deslocacao)
        denominador <- denominador + linha.atual$Numerador;
}

# Calculo da % de Huff. Numerador(i)/denominador.
for (numero.linha in 1:nrow(matriz.distancias)) {
    # for (numero.linha in 1:5) {
    linha.atual <- matriz.distancias[numero.linha,
];
    if (linha.atual$medida.distancia == medida.distancia && linha.atual$modo.deslocacao == modo.deslocacao)
        matriz.distancias$Huff[numero.linha] <<- round( linha.atual$Numerador / denominador, CASAS.DECIMAIS);
    # print(linha.atual);
    # print(matriz.distancias$Huff[numero.linha]);
}
# MatrizDistanciaHuff <- matriz.distancias;
# return(MatrizDistanciaHuff)
}

# -----

# Chamada da função calcula Huff para os vários modos de deslocamento e MEDIDAS de distância.

for (indice.deslocacao in 1:length(MODOS.DESLOCACAO)){ # Para cada modo de deslocamento.
    for (indice.distancia in 1:length(MEDIDAS.DISTANCIA)){ # Vê todas as MEDIDAS de distância.
        calcula.huff(medida.distancia = MEDIDAS.DISTANCIA[indice.distancia],
modo.deslocacao = MODOS.DESLOCACAO[indice.deslocacao]);
        # print(MEDIDAS.DISTANCIA[indice.distancia]);
        # print(MODOS.DESLOCACAO[indice.deslocacao]);
    }
}

#write.table(matriz.distancias, "MatrizDistanciaHuffCGD CONCO.csv", row.names = F, sep = ";", quote = F);

# Limpeza da tabela.
matriz.distancias= matriz.distancias[,c(-1, -2, -10, -11)]

```

```

# Cria um ficheiro com a matriz distancias mais a %huff.
#write.table(matriz.distancias, "MatrizDistanciaHuffCGD_CONC1csv", row.names = F, sep =
";", quote = F);

matriz.distancias$Longitude <- "";
matriz.distancias$Latitude <- "";
matriz.distancias$Município <- "";
matriz.distancias$Distrito <- "";

for (linha.matriz in 1:nrow(matriz.distancias)) {
  for (linha.CGD in 1:nrow(CGD)) {
    if ((matriz.distancias$C_AG[linha.matriz]) == (CGD$C_AG[linha.CGD]) ) {
      matriz.distancias$Longitude[linha.matriz] <- CGD$LONGITUDE[linha.CGD];
      matriz.distancias$Latitude[linha.matriz] <- CGD$LATITUDE[linha.CGD];
      matriz.distancias$Município[linha.matriz] <- CGD$Município[linha.CGD];
      matriz.distancias$Distrito[linha.matriz] <- CGD$Distrito[linha.CGD];
    }
  }
}

for (linha.matriz in 1:nrow(matriz.distancias)) {
  for (linha.concorrencia in 1:nrow(concorrencia)) {
    if ((matriz.distancias$C_AG[linha.matriz]) == (concorrencia$Idnt.Agencia[linha.concorrencia]) ) {
      matriz.distancias$Longitude[linha.matriz] <- concorrencia$Longitude[linha.concorrencia];
      matriz.distancias$Latitude[linha.matriz] <- concorrencia$Latitude[linha.concorrencia];
      matriz.distancias$Município[linha.matriz] <- concorrencia$Município[linha.concorrencia];
      matriz.distancias$Distrito[linha.matriz] <- concorrencia$Distrito[linha.concorrencia];
    }
  }
}

write.table(matriz.distancias, "MatrizDistanciasHuffCGD_Concorrencia.csv", row.names =
F, sep = ";", quote = F);

```

Apêndice D - Pesquisa das percentagens de Huff máximas

```
# -----
# Projeto CGD - Inês Ferreira - julho 2015
# Calcular % Huff máxima de cada agência para os clientes existentes.
# -----

# Ler ficheiro resultante do calculo das percentagens de Huff.
percentagens.huff <- read.table("MatrizDistanciaHuffCGD.csv", header = T, sep = ";");
percentagens.huff$C_AG <- factor(percentagens.huff$C_AG);
percentagens.huff$Município <- as.character(percentagens.huff$Município);
percentagens.huff$Distrito <- as.character(percentagens.huff$Distrito);

# data.frame onde vão ser guardadas as % máximas de huff.
percentagens.huff.maximas <- percentagens.huff[ , c("C_AG", "medida.distancia",
"modo.deslocacao")];

# Eliminar duplicados.
percentagens.huff.maximas <- unique(percentagens.huff.maximas);

# -----

# Codigos das agencias.
AGENCIAS <- levels(percentagens.huff$C_AG);

# Medidas de distância podem ser obtidas automaticamente ou escolhidas pelo utilizador.
MEDIDAS.DISTANCIA <- levels(percentagens.huff$medida.distancia);

# Modos de deslocação podem ser obtidos automaticamente ou escolhidos pelo utilizador.
MODOS.DESLOCACAO <- levels(percentagens.huff$modo.deslocacao);

# -----

# Colunas referentes à percentagem máxima de Huff.
percentagens.huff.maximas$huff.max <- NA;
percentagens.huff.maximas$cliente <- NA;
percentagens.huff.maximas$valor.distancia <- NA;
percentagens.huff.maximas$atratividade <- NA;

# -----

calcula.huff.max <- function(agencia, medida.distancia, modo.deslocacao) {
  percentagens.huff.pertinentes <-
    percentagens.huff[percentagens.huff$C_AG == agencia &
      percentagens.huff$medida.distancia == medida.distancia &
      percentagens.huff$modo.deslocacao == modo.deslocacao, ];
  return(max(percentagens.huff.pertinentes$huff));
}

# -----

# Para cada agência e cada modo de deslocação e para cada medida de distância:
for (numero.linha in 1:nrow(percentagens.huff.maximas)) {
  linha.atual <- percentagens.huff.maximas[numero.linha, ];
  # Para cada agência, obter as percentagens de Huff resultantes de todas as origens,
  e escolher a máxima.
  percentagens.huff.maximas[numero.linha, "huff.max"] <-
    calcula.huff.max(agencia = linha.atual$C_AG,
      medida.distancia = linha.atual$medida.distancia,
      modo.deslocacao = linha.atual$modo.deslocacao);
  # ver a que cliente corresponde a percentagem máxima.
  cliente.com.percentagem.max <-
    percentagens.huff[percentagens.huff$C_AG == linha.atual$C_AG &
      percentagens.huff$medida.distancia == linha.atual$medida.distancia
&
      percentagens.huff$modo.deslocacao == linha.atual$modo.deslocacao &
      percentagens.huff$huff == percentagens.huff.maximas[numero.linha,
"huff.max"], ][1, ];
  percentagens.huff.maximas[numero.linha, "cliente"] <- cliente.com.percentagem.max$IVO;
  percentagens.huff.maximas[numero.linha, "valor.distancia"] <- clien-
te.com.percentagem.max$TD;
```

```

percentagens.huff.maximas[numero.linha, "atratividade"] <- clien-
te.com.percentagem.max$atratividade;

}

#write.table(percentagens.huff.maximas, "MatrizDistanciasHuffMaximas.csv", row.names =
F, sep = ";", quote = F);
# Para cada agência, obter as percentagens de Huff resultantes de todas as origens, e
escolher a máxima.

# varias hipoteses.
for (indice.deslocacao in 1:length(MODOS.DESLOCACAO)){ # Para cada modo de deslocação.
  for(indice.distancia in 1:length(MEDIDAS.DISTANCIA)){ # Vê todas as MEDIDAS de dis-
tância.
    considerar <- percentagens.huff.maximas[percentagens.huff.maximas$medida.distancia
== MEDIDAS.DISTANCIA[indice.distancia] &
percentagens.huff.maximas$modo.deslocacao ==
MODOS.DESLOCACAO[indice.deslocacao], ]
    print(considerar);
  }
}

# Acrescentar todas a colunas correspondentes às agencias na tabela percentagens.huff..
# Começa na coluna 11 pois é onde começam os dados extra.
for (colunas.huff in 11:ncol(percentagens.huff)){
  nome.coluna<-names(percentagens.huff)[colunas.huff] ;
  print(nome.coluna);
  percentagens.huff.maximas[[nome.coluna]] <-"";
}

# Preencher colunas com os valores respetivos.
for (colunas.huff in 11:ncol(percentagens.huff)){
  for (coluna.huff.max in 1:ncol(percentagens.huff.maximas)){
    nome.coluna.MDM <- names(percentagens.huff.maximas)[coluna.huff.max];
    nome.coluna.PH <- names(percentagens.huff)[colunas.huff];
    if (nome.coluna.MDM == nome.coluna.PH) {
      for (linha.matriz in 1:nrow(percentagens.huff.maximas)) {
        for (linha.PH in 1:nrow(percentagens.huff)) {
          if ((percentagens.huff.maximas$C_AG[linha.matriz]) == (percenta-
gens.huff$C_AG[linha.PH]) ) {
            percentagens.huff.maximas[[nome.coluna.MDM]][linha.matriz] <- percenta-
gens.huff[[nome.coluna.PH]][linha.PH];
          }
        }
      }
    }
  }
}

write.table(percentagens.huff.maximas, "MatrizDistanciasHuffMaxima.csv", row.names = F,
sep = ";", quote = F);

```