

UNIVERSIDADE FEDERAL DE PERNAMBUCO CENTRO DE INFORMÁTICA

Juliane Sabrina Magalhães do Nascimento

Quot: Uma ferramenta de suporte à decisão no domínio de imóveis

Trabalho de Graduação

Recife -PE 2017



Universidade Federal de Pernambuco Centro de Informática Graduação em Ciência da Computação

Juliane Sabrina Magalhães do Nascimento

Quot: Uma ferramenta de suporte à decisão no domínio de imóveis

Trabalho apresentado ao Programa de Graduação em Ciência da Computação do Centro de Informática da Universidade Federal de Pernambuco como requisito parcial para obtenção do grau de Bacharel em Ciência da Computação.

Orientador: Luciano de Andrade Barbosa

Recife -PE 2017

Resumo

Com o surgimento da internet, a procura por imóveis foi facilitada por diversos sites que fazem buscas filtradas por região ou características específicas do imóvel como, por exemplo: número de quartos, banheiros, área de lazer, piscina, entre outras. Alguns desses sites, além dos filtros para facilitar a busca, possuem também ferramentas para visualização de imóveis no mapa, porém limitando-se apenas à localização. Este trabalho tem como objetivo propor um sistema web de apoio à decisão, que utiliza técnicas de visualização de dados para auxiliar os usuários a tomarem decisões, seja de compra/venda ou aluguel, baseadas em estatísticas obtidas a partir desses dados.

Palavras-chave: visualização de dados, sistema de apoio à decisão, imóveis, análise de dados.

Abstract

With the internet, real estate search was facilitated by several websites where one can search by filtering region or property specific characteristics such as number of bedrooms, bathrooms, recreation area, swimming pool, among others. Some of these websites, in addition to having these filters to facilitate the search, also have tools for viewing real estate on a map, but limited to location only. This work web-based decision support system that uses data visualization techniques to help users make decisions, for buying/selling or renting, based on statistics obtained from these data.

Keywords: data visualization, decision support system, real estate, data analysis.

Lista de Figuras

Figura 2.1. Área de busca do Zap Imóveis11
Figura 2.2. Área de busca do Expo Imóvel
Figura 2.3. Área de busca do Viva Real11
Figura 2.4. Mecanismo de busca do Zap Imóveis através de um mapa 12
Figura 2.5. Localização de um imóvel no site do Viva Real
Figura 2.6. Localização de um imóvel no site do Expo Imóvel
Figura 2.7. Estimativa de preço por metro quadrado do Zap Imóveis
Figura 2.8. Estimativa de preço por metro quadrado do Expo Imóvel 14
Figura 2.9. Gráfico de barras do preço médio em relação ao número de
quartos
Figura 2.10. Gráfico de barras do preço médio em relação ao tipo de imóvel. 14
Figura 2.11. Gráfico de barras do preço médio em relação à área privativa 15
Figura 2.12. Variação do índice FipeZap de venda de imóveis para a cidade do
Recife
Figura 2.13. Variação do índice FipeZap de locação de imóveis com dois
dormitórios para a cidade do Recife17
Figura 2.14. Índices FipeZap das cidades17
Figura 3.1. Interface do Quot 19
Figura 3.2. Interface do Quot 19
Figura 3.3. Visualização da mediana do preços dos imóveis para a cidade de
São Paulo para o mês de Maio de 2017
Figura 3.4. Visualização da mediana do preços dos imóveis para a cidade de
São Paulo para o mês de Março de 2017
Figura 3.5. Gráficos da mediana dos preços dos imóveis do bairro do Morumbi
da cidade de São Paulo
Figura 3 6. Coordenadas paralelas mapeando os dados dos imóveis com 4
quartos da cidade de Sao Paulo
Figura 4.1. Media de preço dos imoveis do mes de juino de 2017 para a cidade
de Sao Paulo
Figura 4.2. Media de preço dos imoveis do mes de abril de 2017 para a cidade
de Sao Paulo
rigura 4 3. Media de preço dos imoveis do mes de maio de 2017 para a cidade
Ce Sao Paulo
Figura 4.4. Evolução do preço medio dos intoveis do bairro de Moema
Figura 4.6. Evolução do preco médio dos imóveis do bairro de Vila Sônia 30
Figura 4.7 Comparação do preço médio dos imóveis do bairro de Moema po
mês de marco e junho de 2017
Figura 4.8 Evolução do preco médio dos imóveis do bairro de Vila Prudente 30

Lista de Tabelas

2.1. Tabela comparativa dos sites	
-----------------------------------	--

Lista de Acrônimos

- SVG Scalable Vector Graphics
- **JSON** JavaScript Object Notation
- HTML HyperText Markup Language
- DOM Document Object Model

Sumário

1. Int	rodução	. 9
1.1.	Objetivos	. 9
1.2.	Estrutura do trabalho	10
2. Tra	abalhos relacionados	11
3. Qu	ıot	18
3.1.	Tarefas analíticas	18
3.2.	Quot	18
3.2	2.1. Мара	20
3.2	2.2. Gráficos	21
3.2	2.3. Filtros	24
4. Es	tudo de caso do sistema	26
5. Co	nclusão	31
5.1.	Melhorias e trabalhos Futuros	31
Referê	ncias Bibliográficas	33

1. Introdução

Sistemas de suporte à decisão (SSD) são aplicações que ajudam usuários a tomarem decisões mais facilmente em um dado domínio, tópico ou dentro de uma organização. SSDs têm sido usados em diversas áreas como medicina¹, agricultura [1], transportes [2] entre outras. Neste trabalho em particular, estamos interessados em criar um SSD para o domínio de imóveis.

Atualmente, a forma usual usada por um potencial comprador ou inquilino de imóveis que deseja se informar sobre o mercado imobiliário é através dos sites de venda de imóveis. Eles permitem ao usuário buscas filtradas por região do país e características específicas do imóvel como, por exemplo: número de quartos, banheiros, área de lazer, piscina, entre outras. Alguns desses sites, além de dispor desses filtros para facilitar a busca, dispõem também de ferramentas para visualização de imóveis no mapa. Dessa forma o usuário pode visualizar todos os imóveis de um determinado bairro, tendo uma ideia geral de quais bairros possuem imóveis mais baratos, ou quais regiões possuem imóveis que atendem os critérios estabelecidos nos filtros fornecidos pelo usuário.

1.1. Objetivos

O sistema visual humano é capaz de identificar padrões, tendências e relações entre os objetos em uma cena visual. As técnicas de visualização de dados são comumente usadas para comunicar a informação de uma forma clara e efetiva, extraindo informações a partir de um dado conjunto de dados e possibilitando a exploração desses dados de novas formas eficazes.

O objetivo deste trabalho é propor um sistema web de apoio à tomada de decisão, chamado Quot, na área de imóveis (descrita em detalhes no Capítulo 3). A maioria dos sistemas conhecidos atualmente funcionam como portais de classificados, onde o foco principal são anúncios de compra, venda ou aluguel de imóveis. Alguns desses sistemas possuem mecanismos de visualização de dados, como a evolução de preços por metro quadrado de regiões específicas e gráficos que relacionam o preço do metro quadrado com determinadas características do imóvel (como tipo, número de quartos, área privativa, etc.). Assim, esses mecanismos permitem aos usuários analisarem melhor diversos cenários, auxiliando-os na decisão da compra do imóvel. Porém o público alvo desses sites limita-se basicamente aos que estão à procura de um imóvel, seja compra ou aluguel. Diversas imobiliárias e investidores estão constantemente analisando o mercado para avaliarem os seus imóveis e, por isso surgiu a necessidade de desenvolver um sistema que seja voltado inteiramente para a análise e interpretação dos dados de forma mais geral. O Quot possui uma interface interativa com o usuário possibilitando a exploração dos dados de diversas formas. Ele utiliza diversas técnicas de visualização de dados como, coordenadas paralelas e gráficos para auxiliar não só usuários no processo de

¹ http://www.mghlcs.org/projects/dxplain

compra ou aluguel, como também investidores e imobiliárias, através da análise e interpretação dos dados de imóveis.

1.2. Estrutura do trabalho

A descrição da organização deste documento será descrita a seguir. O capítulo 2 apresenta trabalhos relacionados e ferramentas na web disponíveis, como portais de classificados de imóveis e imobiliárias. O capítulo 3 descreve a interface e as tarefas executadas pelo Quot, bem como as técnicas de visualização utilizadas. O capítulo 4 apresenta os casos de uso que descrevem as funcionalidades do Quot. Por fim, no capítulo 5 serão apresentadas as conclusões e considerações finais sobre o trabalho, e quais os possíveis trabalhos futuros, a fim de obter melhorias para a ferramenta.

2. Trabalhos relacionados

Este capítulo é dedicado a análise de diferentes ferramentas de visualização de dados existentes no contexto de imóveis.

Existem diversas ferramentas web que dão suporte aos usuários na busca por um imóvel. Dentre essas ferramentas encontradas na web, podemos citar primeiramente o *Zap Imóveis, Viva Real* e *Expo Imóvel*. Todos esses sites citados tratam-se de ferramentas de buscas, funcionando como portais de classificados de imóveis e imobiliárias, sendo seu principal objetivo anunciar apartamentos, terrenos e imóveis para compra, venda e aluguel.

Tanto o *Zap Imóveis* quanto *o Viva Real* e o *Expo Imóvel,* possuem alguns mecanismos muito semelhantes para a exploração dos dados, como filtros personalizados (tipo de imóvel, localização, etc) como ilustrado nas figuras 2.1, 2.2 e 2.3.

que você prec	isa?		Qual tipo?		Onde?	Ver todas as loc	alidades 🛈
Comprar	Alugar	Lançamentos	Apartamento Padrão	•	Digite o bairro, cidade ou região		
15.201 anún	cios encontra	ados			Busca avança	la - BUS	CAR

Figura 2.1. Área de busca do Zap Imóveis.

COMPRAR				ALUGAR	TEMPORADA		
Apartamento]~	PE	v	Cidade, Bairro, Co	odigo, Nome ou Rua	0	Ø BUSCAR

Figura 2.2. Área de busca do Expo Imóvel.



Figura 2.3. Área de busca do Viva Real.

Além dos filtros presentes, os sites também dispõem de ferramentas de visualizações relativamente simples como mapas e gráficos que mostram a evolução do preço dos imóveis em determinado período ou a relação do preço de acordo com alguma característica do imóvel, como tipo, número de quartos,

área privativa, etc. No caso do Zap, existe a possibilidade de realizar a busca através da manipulação de um mapa, ou seja, os imóveis vão sendo mostrados à medida que o usuário movimenta o mapa com o auxílio do mouse. Esse mecanismo está ilustrado nas figuras 2.4. Já o *Viva Real* e o *Expo Imóvel* dispõem de um mapa apenas para visualizar a localização do imóvel previamente buscado pelo usuário, como é mostrado nas figuras 2.5 e 2.6.



Figura 2.4. Mecanismo de busca do Zap Imóveis através de um mapa.



Figura 2.5. Localização de um imóvel no site do Viva Real.

Em relação à visualização na forma de gráficos, tanto o *Zap Imóveis* quanto o *Expo Imóvel* possuem um mecanismo de série temporal. Essa funcionalidade estima o preço médio do metro quadrado de venda ou aluguel de imóveis para uma determinada região, como mostrado nas figuras 2.7. e 2.8



Figura 2.6. Localização de um imóvel no site do Expo Imóvel.



Figura 2.7. Estimativa de preço por metro quadrado do Zap Imóveis.

O *Expo Imóvel* dispõe ainda de gráficos de barras que permitem analisar o preço médio de acordo com o número de quartos, o metro quadrado, por tipo de imóvel, por área privativa, etc. Alguns desses gráficos são mostrados nas figuras 2.9, 2.10 e 2.11.

A tabela 2.1. mostra uma análise comparativa de alguns aspectos e funcionalidades de visualização de dados encontradas nas ferramentas *Zap Imóveis, Viva Real* e *Expo Imóvel.*



Figura 2.8. Estimativa de preço por metro quadrado do Expo Imóvel.



Figura 2.9. Gráfico de barras do preço médio em relação ao número de quartos.



Figura 2.10. Gráfico de barras do preço médio em relação ao tipo de imóvel.



Este gráfico mostra o preço médio de imoveis para venda de acordo com a área privativa.

Figura 2.11. Gráfico de barras do preço médio em relação à área privativa.

Site	Séries temporais	Gráfico de barras	Мара	Filtros
Zap Imóveis	SIM	NÃO	SIM	SIM
Expo Imóvel	SIM	SIM	SIM	SIM
Viva Real	NÃO	NÃO	SIM	SIM

Tabela 2.1. Tabela comparativa dos sites.

Como podemos perceber, as ferramentas citadas são bastante eficazes no que elas se propõem a fazer, que é servir como um portal de classificados de imóveis com alguns mecanismos de visualização de dados, com o intuito de auxiliar os seus usuários na procura de um imóvel dentro de sua própria base de dados. Porém elas possuem certas limitações em alguns aspectos. Dentre essas limitações podemos citar:

 Diversidade dos dados: a base de dados dos sites de imóveis está usualmente restrita aos anúncios postados no site. Mesmo com uma base grande de imóveis, esses dados podem trazer algum tipo de viés associado, já que os usuários fazem seus anúncios baseados em outros anúncios presentes no site que possuem caraterísticas semelhantes.

 Análise de dados: os sites de imóveis usualmente mostram apenas a localização deles a partir dos filtros previamente selecionados, mas não provêm uma análise mais profunda sobre, por exemplo, disposição desses imóveis na região quanto ao preço ou suas características.

Dessa forma, essas ferramentas não são especificamente direcionadas à tomada de decisão, limitando-se apenas a mostrar os imóveis, seus preços e a sua localização e algumas estatísticas baseadas em seus próprios anúncios. Portanto o objetivo desses sites é de apenas listar os imóveis de acordo com a exigência do usuário, seja quanto à localização, número de quartos, preço, etc. A falta de uma análise mais aprofundada, acaba por confundir mais ainda o usuário, provocando insegurança na compra ou venda de algum imóvel.

Outra ferramenta bastante conhecida e direcionada à tomada de decisão na compra ou venda de um imóvel é a FipeZap. A Fipe (fundação Instituto de Pesquisas Econômicas) é uma organização de direito privado, sem fins lucrativos, criada em 1973 com o objetivo de analisar os fenômenos econômicos e sociais com base no instrumental teórico e metodológico da Economia [4]. A FipeZap é o resultado de uma parceria formada em 2010 entre a Fipe e o portal ZAP [5] e utiliza como fonte de dados os anúncios de venda ou locação de apartamentos prontos cadastrados em diversos sites, sendo a mais relevante em termos de abrangência e tamanho a do Zap Imóveis [6].

A ferramenta é composta por dois gráficos e uma tabela de índices que tem o objetivo de propor uma análise da variação do índice de locação e venda de imóveis. Ambos os gráficos mostrados nas figuras 2.12 e 2.13, apresentam mês a mês a variação do índice no período selecionado. É possível filtrar por cidade, número de quartos, transação (venda ou aluguel) e período (ano atual, últimos 12 meses ou todo o período disponível). Ao passando o mouse sobre os pontos do gráfico, é possível ver os percentuais de variação em relação ao mês anterior e ao início do período selecionado [5]. A figura 2.14 mostra os índices de venda e locação para as cidades disponíveis na base de dados do FipeZap.

Nos últimos anos o mercado imobiliário está sendo muito influenciado pelos preços divulgados pelo FipeZap. Diversas pessoas estão tomando decisões de investimento, compra e venda de imóveis de acordo com a base gerada pela ferramenta. Porém essa ferramenta limita-se apenas a mostrar um índice de histórico de preços, não possuindo uma projeção em um mapa que permitiria visualizar a distribuição de preços de imóveis e fazer uma comparação entre um preço médio em diversas regiões. O filtro utilizado é uma outra limitação encontrada na ferramenta, onde apenas é possível filtrar pela localização e pela quantidade de quartos do imóvel.



Figura 2.12. Variação do índice FipeZap de venda de imóveis para a cidade do Recife.



Figura 2.13. Variação do índice FipeZap de locação de imóveis com dois dormitórios para a cidade do Recife.

Venda					Locação				
CIDADE/LOCAL	NO MÊS	NO ANO	12 MESES	36 MESES	CIDADE/LOCAL	NO MÊS	NO ANO	12 MESES	36 MESES
FIPEZAP VENDA	-0.16	-0.08	0.46	5.51	FIPEZAP LOCACAO	0.03	0.59	-1.68	-6.48
SAO PAULO	0.03	0.68	0.76	7.37	SAO PAULO	0.56	2.08	0.74	-2.87
BELO HORIZONTE	-0.36	3.53	7.32	15.4	BELO HORIZONTE	-0.12	1.69	0.82	
	Mostrar mais cidades 🔹							Mostrar ma	is cidades 🔹

Figura 2.14. Índices FipeZap das cidades.

3. Quot

Neste capítulo será apresentado a interface do Quot, assim como as tarefas analíticas suportadas pelo sistema.

3.1. Tarefas analíticas

O sistema Quot foi desenvolvido para suportar as seguintes tarefas analíticas:

- Tarefa 1 (T1): Mostrar a distribuição de preços de cada região no mapa, permitindo a comparação de preço em uma região com o valor de outra região.
- Tarefa 2 (T2): Permitir visualizar a evolução de preços dos bairros ou de uma cidade.
- Tarefa 3 (T3): Disponibilizar filtros para auxiliar o usuário na exploração mais aprofundada dos dados.

3.2. Quot

Esta seção descreve a interface do Quot. Baseado nas tarefas analíticas listadas anteriormente houve a definição das interações do usuário com a ferramenta e com isso a definição do design da mesma. As figuras 3.1 e 3.2 mostra a interface da ferramenta que possibilita a realização das tarefas analíticas listadas.

O sistema proposto se trata de uma aplicação web, visando facilidade ao acesso à informação pelos usuários e por dispensar qualquer tipo de instalação local da aplicação em um computador. Para implementar o sistema foram usadas duas bibliotecas JavaScript muito utilizadas para apoiar o desenvolvimento de aplicações interativas: D3.js e Leaflet.

A interface é composta por três funcionalidades. A primeira consiste em um mapa que mostra a distribuição da mediana dos preços de acordo com a região pré-selecionada pelo usuário. A razão de se utilizar a mediana de preços e não a média será explicada na próxima seção. A segunda consiste na visualização dos dados dos imóveis na forma de coordenadas paralelas, permitindo que os filtros sejam aplicados. A terceira e última visualização consiste em dois gráficos, sendo um representando uma série temporal dos preços médios das regiões através de um gráfico do tipo *multi-line chart* e o outro gráfico do tipo *scatterplot*, onde é possível analisar o preço médio de cada região em diferentes períodos definidos pelo usuário. Todas essas técnicas são detalhadas nas próximas seções.



Figura 3.1. Interface do Quot.



Figura 3.2. Interface do Quot.

3.2.1. Mapa

A visualização no mapa permite que a tarefa analítica (T1) seja realizada, onde é possível visualizar todas as regiões a partir de um tipo de mapa temático denominado *Choropleth map*. Os mapas *Choropleth* exibem áreas geográficas divididas ou regiões coloridas, sombreadas ou padronizadas em relação a uma variável de dados. Isso fornece uma maneira de visualizar valores em uma área geográfica, que pode mostrar variações ou padrões em toda a localização exibida, representando as diferenças no fenômeno que está sendo mapeado.

Foi utilizada uma progressão de cores para se representar em cada região do mapa os valores relativos à mediana do preço, como mostrado na figura 3.1. Dessa forma é possível observar quais regiões que possuem imóveis mais valorizados, utilizando-se apenas o mapa. Um estudo empírico mostrou que, na maioria dos casos, os leitores do mapa assumem que "o escuro significa mais e a claro significa menos" [7]. Por isso, foi utilizado o padrão de cores mais escuras para representar as regiões que possuem uma maior mediana de preço dos imóveis e um padrão mais claro para representar as regiões que possuem uma mediana de preço menor.

Outra forma de interação no mapa é através de um filtro que faz buscas por períodos, e reflete essa busca alterando as cores de cada região. Quando o período é alterado, o sistema recalcula a mediana dos preços dos imóveis e utiliza métodos estatísticos para fazer uma discretização desses preços, gerando uma nova escala de cores para o mapa e refletindo essa mudança na coloração das regiões. Utilizamos a mediana por ela ser uma estatística robusta de centralidade, ou seja, ela é menos sensível a *outliers* (em estatística, *outliers* são valores que fogem muito da tendência central, ou seja, valores muito altos ou muito baixos que não são representativos do todo) do que a média, por exemplo.

Essa funcionalidade faz com que o mapa se comporte de forma interativa e auxilia o usuário a entender quais as regiões estão crescendo mais no aspecto imobiliário nos últimos meses ou anos, auxiliando-o na tomada de decisão de compra, venda ou investimentos futuros. Esse comportamento é ilustrado na figura 3.3 e 3.4.

O mapa foi implementado utilizando a biblioteca *javascript* Leaflet.js, amplamente utilizada usada para criar aplicativos web com mapas interativos. Essa biblioteca possui inúmeras funcionalidades como, por exemplo, aplicar elementos em SVG (Scalable Vector Graphics), carregar dados de recursos de arquivos GeoJSON, modelar e criar camadas interativas, como marcadores com *popups* e legendas para mapas no formato HTML (HyperText Markup Language. Além dessas funcionalidades, a biblioteca também disponibiliza o uso de eventos, como *mouseover, mouseout* e *click,* que promovem ainda mais a interação com o usuário e com outras bibliotecas, como a biblioteca *javascript* D3.js, que foi utilizada na construção dos gráficos da aplicação. Os eventos de *mouseover* e *mouseout* sobre o mapa, permite o usuário identificar mais facilmente a região que está sendo analisada, bem a mediana de preço dos imóveis para aquela região, como mostra a figura 3.3. O evento de *click*, foi utilizado para dar um zoom automático no mapa e para funcionar como uma interface de busca, dispensando o uso do seletores de localização (Figura 3.1),



tornando a pesquisa mais rápida para os usuários que possuem um conhecimento maior na localização de regiões em mapas.

Figura 3.3. Visualização da mediana do preços dos imóveis para a cidade de São Paulo para o mês de Maio de 2017.

3.2.2. Gráficos

A visualização através de gráficos permite que a tarefa analítica (T2) seja realizada, onde é possível visualizar a evolução da mediana dos preços de cada região quando o usuário interage com o gráfico através do mouse. As próximas seções descrevem os dois tipos de gráficos utilizados na aplicação: *multi-line chart* e *scatterplot*.

3.2.2.1. Multi-line chart e Scatterplot

Um gráfico de linhas ou *line chart* é muitas vezes usado para visualizar dados que mudam ao longo do tempo e é composto por dois eixos, um vertical e outro horizontal, e por uma linha que mostra a evolução de um fenômeno ou processo [9]. A visualização através do gráfico de linha, representa uma série temporal que permite visualizar a evolução da mediana dos preços dos imóveis de uma região específica e comparar com a mediana geral de todas as regiões, como mostrado na figura 3.4. O objetivo da análise de séries temporais é identificar padrões não aleatórios na série temporal de uma variável de interesse, e a observação deste comportamento passado pode permitir fazer previsões sobre o futuro, orientando a tomada de decisões [8]. Então um possível investidor, inquilino ou comprador pode, através dessa análise, decidir qual região está mais valorizada em aspectos imobiliários ou se está no momento de comprar ou vender um apartamento, ou optar por comprar ou alugar um imóvel, caso os padrões encontrados nos gráficos indique um aumento acelerado do preço em relação ao tempo.



Figura 3.4. Visualização da mediana do preços dos imóveis para a cidade de São Paulo para o mês de Março de 2017.

Scatterplots são úteis para interpretar tendências em dados estatísticos. Cada ponto no gráfico possui duas coordenadas. A primeira corresponde à primeira porção de dados no par (que representa a coordenada *x*) e a segunda coordenada corresponde à segunda porção de dados no par (coordenada *y*). O ponto que representa essa observação é colocado na interseção das duas coordenadas. A visualização através do gráfico de *Scatterplot* além de permitir visualizar a evolução da mediana dos preços dos imóveis de uma região específica em relação a dois períodos especificados no eixo *x* e no eixo *y* do gráfico, também mostra a distribuição desse valor comparado a outras regiões nesse mesmo período. O sistema interage com o usuário através da manipulação do mouse sobre os pontos do gráfico que representam as regiões escolhidas pelo usuário na busca. A figura 3.5 mostra a evolução da mediana dos preços dos imóveis do simóveis do mouse sobre os pontos do gráfico que representam as regiões escolhidas pelo usuário na busca. A figura 3.5 mostra a evolução da mediana dos preços dos imóveis do simóveis do bairro do Morumbi da cidade de São Paulo.



Figura 3.5. Gráficos da mediana dos preços dos imóveis do bairro do Morumbi da cidade de São Paulo.

Para a construção dos gráficos foi utilizada a biblioteca javascript D3.js, uma ferramenta muito utilizada produzir visualizações de dados dinâmicas e interativas em aplicações da web e manipular dados do DOM (*Document Object Model*) usando HTML, SVG e CSS. O SVG (*Scalable Vectorial Graphics*) é um formato de imagem criado pelo *W3C* e consiste em um arquivo XML que contém tags específicas para gerar uma imagem vetorizada na sua aplicação. Dessa forma, é possível criar imagens de alta qualidade vetorizadas [10].

3.2.3. Filtros

Os filtros que a aplicação dispõe permitem que a tarefa analítica (T3) seja realizada, proporcionando ao usuário uma melhor exploração dos dados e uma maior interação com o sistema. Os filtros foram implementados utilizando uma técnica bastante conhecida e que está entre os assuntos mais comuns de trabalhos acadêmicos na área de visualização de dados, as coordenadas paralelas. Essa técnica é uma ferramenta muito poderosa para a compreensão de conjuntos de dados numéricos multidimensionais. Para implementar as coordenadas paralelas foi utilizada a biblioteca javascript D3.js e elementos SVG HTML. A técnica agrupa os dados dos imóveis de uma determinada região em colunas ou dimensões que são mapeadas em eixos verticais, como ilustrado na figura 3.6. Para uma maior exploração dos dados foi utilizada a técnica de *brushing*, que permite filtrar as informações baseando-se nos eixos verticais, como ilustrado na figura 3.6.

Para que a técnica seja implementada de maneira eficiente o número de itens de dados precisa ser limitado (quando o número é muito alto há uma grande quantidade de *overplotting*) e o número de dimensões precisa ser moderado, caso contrário tornar-se quase impossível visualizar e fazer alguma análise dos dados. Essas limitações também estão diretamente relacionadas ao desempenho da aplicação, já que um número elevado de objetos DOM torna a interface do usuário bastante lenta.

Como o conjunto de dados é considerado grande, foi utilizado o método requestAnimationFrame do javascript para impedir que a renderização das coordenadas paralelas provoque lentidão ou bloqueie a interface com o usuário. Esse método executa código javaScript relacionado à alguma animação que faz mudanças na tela do usuário de forma eficiente e otimizada, já que o código de animação é chamado quando o computador do usuário está realmente pronto para fazer alterações na tela. Atualmente diversos navegadores modernos como o *Internet Explorer 10* +, *FireFox* 11 +, Chrome e Safari já suportam a utilização desse método.



Figura 3 6. Coordenadas paralelas mapeando os dados dos imóveis com 4 quartos da cidade de São Paulo.

4. Estudo de caso do sistema

O objetivo desta seção é apresentar uma análise comparativa entre diferentes regiões por meio de técnicas de visualização de dados utilizadas pelo sistema.

O primeiro cenário analisado foram dos imóveis da cidade de São Paulo a princípio sem a utilização dos filtros. Analisando a figura 4.1, percebe-se que os bairros com os imóveis mais caros para o mês de junho de 2017 são Anhanguera, com um preço médio de 2 milhões e 500 mil reais, e Capão Redondo, com 3 milhões de reais.



Figura 4.1. Média de preço dos imóveis do mês de julho de 2017 para a cidade de São Paulo.

Fazendo a mesma análise para a cidade de São Paulo para o mês de maio de 2017, percebe-se que houve uma grande redução da média de preços dos bairros Anhanguera e Capão Redondo, com valores de 260 mil e 292 mil respectivamente, como mostrado na figura 4.2.



Figura 4.2. Média de preço dos imóveis do mês de abril de 2017 para a cidade de São Paulo.

Essa diferença exagerada da média de preço do mês de abril e junho de 2017 reflete um dos problemas enfrentados pela ferramenta que é a limitação dos dados. Os dados utilizados pelo Quot são baseados em anúncios coletados em diversos sites. Dessa forma, a quantidade de dados pode tanto aumentar ou diminuir a robustez do sistema. No caso do estudo apresentado acima, concluímos que os dados coletados sobre os bairros de Anhanguera e Capão Redondo foram insuficientes para realizar uma análise confiável. Esse problema também impacta o desempenho da ferramenta FipeZAp, citada no capítulo 2, pois os índices do FipeZap assim como os gráficos e visualizações que Quot dispõe, são baseados em anúncios e dependem completamente da qualidade e da quantidade desses dados coletados para que o sistema proporcione uma análise confiável para os usuários. A figura 4.3, ilustra o cenário de imóveis da cidade de São Paulo para o mês de maio de 2017 e observa-se que não existe dados para o bairro de Anhanguera, comprometendo assim a análise dos gráficos que acompanham a evolução dos preços dos imóveis em relação ao tempo. Como o bairro de Anhanguera não possui dados para o mês de março, seu preço médio é representado com o valor 0, assim o ponto que o representa no gráfico de dispersão se encontra no eixo y e não no espaço cartesiano.

Devido aos problemas encontrados na ferramenta, as análises apresentadas neste capítulo irão desconsiderar alguns bairros que se encaixam



nos cenários onde a insuficiência ou ausência de alguns dados comprometem a interpretação dos mesmos.

Figura 4 3. Média de preço dos imóveis do mês de maio de 2017 para a cidade de São Paulo.

Analisando os imóveis da cidade de São Paulo, que possuem de dois a três guartos e com preço variando de 119 mil e 486 reais, como mostrado na figura 4.5, é possível observar através da escala de cores refletidas no mapa Choropleth que os bairros de Pinheiros, Moema e Jardim Paulista possuem um maior preço médio de imóveis comparado aos demais bairros da cidade de São Paulo. Analisando a evolução do preço médio do bairro de Moema desde o mês de março à junho de 2017, percebe-se que a média de preço está seguindo aproximadamente a mesma variação do preço médio geral para todos os bairros, como mostrado na figura 4.6. Porém o bairro de Moema apresenta uma média de preço mais elevada comparada a média de preço dos bairros em geral, e a partir da interpretação e análise desses dados podemos concluir que Moema é um dos bairros que possuem os imóveis mais valorizados da cidade de São Paulo. A figura 4.6 mostra o gráfico de dispersão para os bairros de São Paulo e nota-se que houve um aumento significativo na média de preço dos imóveis de Moema, variando de 754 mil e 500 reais no março de 2017 para 790 mil no mês de junho de 2017.

Analisando o mesmo cenário descrito anteriormente para os bairros de Vila Sônia e Vila Prudente, através da figura 4.7 e 4.8 pode-se observar que a evolução dos preços dos imóveis de Vila Sônia não foram significativas em relação à média de preço geral dos bairros, já os preços dos imóveis de Vila Prudente sofreram uma queda ao longo do período de março à abril, ficando um pouco abaixo do preço médio geral dos demais bairros. Dessa forma, podemos concluir que os preços dos imóveis do bairro de Vila Sônia encontram-se estagnados nesse período, e portanto não seria um bom momento para se investir em um imóvel nesta região. Quanto ao bairro de Vila Prudente, devido a queda do preço dos imóveis, não seria um bom momento para vender, porém seria uma boa época para comprar ou alugar um imóvel nesta região.



Figura 4.4. Imóveis com filtros aplicados.



Figura 4.5. Evolução do preço médio dos imóveis do bairro de Moema.



Figura 4.6. Comparação do preço médio dos imóveis do bairro de Moema no mês de março e junho de 2017.



Figura 4.7. Evolução do preço médio dos imóveis do bairro de Vila Sônia.



Figura 4.4. Evolução do preço médio dos imóveis do bairro de Vila Prudente.

5. Conclusão

Este trabalho propôs uma ferramenta de visualização de dados de imóveis do estado de São Paulo com o intuito de auxiliar processos de tomada de decisão. A ferramenta se trata de uma aplicação web que dispensa a necessidade de instalações locais e provê facilidade ao acesso por parte dos usuários do sistema. Através das técnicas de visualização implementadas no Quot (o uso de gráficos e a distribuição de preços no mapa), foi dado um suporte mais elaborado aos usuários como investidores, imobiliárias e compradores no cenário de análise de padrões e tendências futuras do mercado imobiliário. O Quot possui um design simples, porém cumpre bem o que lhe foi proposto. Com a intenção de melhorar ainda mais as funcionalidades do sistema, a próxima seção descreve possíveis melhorias e trabalhos futuros, que podem aumentar ainda mais a compreensão dos dados e o apoio à tomada de decisão. Apesar do trabalho ter sido voltado para a análise dos imóveis do estado de São Paulo, a ferramenta é plenamente capaz de fazer a mesma análise para qualquer região, sendo preciso apenas coletar os dados de imóveis para as demais regiões do país.

5.1. Melhorias e trabalhos Futuros

Para que a ferramenta se torne mais precisa e forneça uma boa interpretação, deve-se investir cada vez mais na melhoria nos dados. Como citado no capítulo 4, o desempenho da ferramenta é totalmente dependente da qualidade e da quantidade de informações que lhe é fornecida, e quanto maior a quantidade e a precisão desses dados, maior a quantidade de informação entregue ao usuário e melhor será a ferramenta de apoio à decisão.

Os mapas de choropleth passam uma impressão ao usuário de que o fenômeno de interesse é contínuo, ou seja, presente em toda a região, devido à mudança abrupta das colorações entre as fronteiras da área da região. Outra questão é que as regiões maiores aparecem mais enfatizadas que as menores, dominando a aparência visual do mapa, podendo fazer com que o usuário tenha uma percepção exagerada da importância de algumas regiões específicas em relação a outras no mapa. Devido a isso, a implementação de outro tipo de visualização no mapa de forma que possa mostrar algo que não seja tão geral como, por exemplo, um tipo de mapa de calor que demonstra a intensidade dos dados em cada ponto geográfico de uma região específica. Esse tipo de visualização seria bastante útil para a melhor compreensão e análise dos dados, tornando o processo de tomada de decisão mais elaborado e eficiente.

Seguindo pela mesma idéia do mapa de calor, seria interessante ter uma abordagem de previsão de preços em pontos desconhecidos do mapa, de forma que o usuário pudesse ter uma idéia geral dos preços nas regiões que possuem esses pontos. Para implementar essa funcionalidade, existe um método usado em geoestatística chamado de Krigagem que usa regressão para aproximar ou interpolar dados. A técnica de Krigagem parte do princípio que pontos próximos no espaço tendem a ter valores mais parecidos do que pontos mais afastados, ou seja, que os dados recolhidos de uma determinada população se encontram correlacionados no espaço [11].

Uma das perguntas mais relevantes quando alguém quer comprar um imóvel é saber se esse imóvel é bem localizado. Para muitos, para um imóvel ser bem localizado significa que ele é próximo de estabelecimentos como: padarias, shoppings, mercados, escolas, estação de metrô, hospitais, delegacias, etc, onde as pessoas não teriam que se deslocar por longas distâncias para fazer tarefas simples, como ir ao mercado, ou enfrentar engarrafamentos para levar seus filhos à escola [3]. Seria interessante para a aplicação adicionar os pontos de interesse citados anteriormente de forma que o sistema possa filtrar os imóveis próximos a esses pontos, tornando o mecanismo de busca ainda mais robusto.

Referências Bibliográficas

[1] V. Rony e P. Taco. Sistema de apoio à decisão de informações espaciais via web para transporte público urbano. Pós-Graduação em Transportes - Universidade de Brasília – UnB. Brasília. Disponível em : http://www.anpet.org.br/xxviiianpet/anais/documents/RT316.pdf. Acessado em

http://www.anpet.org.br/xxviiianpet/anais/documents/RT316.pdf. Acessado em 10/04/2017.

[2] N. Marcelo; S. Amarindo; P. Danilo e N. Irenilza. Banco de Dados e Sistema de Apoio à Decisão para as Culturas de Milho e Soja. 25f. Campinas: Embrapa Informática Agropecuária, 2003. Disponível em:

https://ainfo.cnptia.embrapa.br/digital/bitstream/CNPTIA/10203/1/doc36.pdf. Acessado em 10/04/2017.

[3] S. DÉBORA. Comportamento do consumidor: Critérios de escolha adotados para compra de imóveis residenciais no plano piloto em Brasília DF. 51f. Monografia (Graduação em Administração) - Faculdade de Tecnologia e Ciências Sociais Aplicadas. Brasília. Disponível em:

http://repositorio.uniceub.br/bitstream/235/9984/1/20707164.pdf. Acessado em 08/03/2017.

[4] Fundação Instituto de pesquisas econômicas - Institucional. Disponível em: http://www.fipe.org.br/pt-br/institucional/sobre-a-fipe/. Acessado em 1 de julho de 2017.

[5] FipeZap – O índice do FipeZap é. Disponível em: http://fipezap.zapimoveis.com.br/. Acessado em 1 de julho de 2017.

[6] Índice FipeZap de preços de Imóveis anunciados – Notas metodológicas. Disponível em:

http://downloads.fipe.org.br/content/downloads/indices/fipezap/metodologia/Fip eZAP_RevMetodologia_v20140218.pdf. Acessado em 1 de julho de 2017.

[7] McGranaphan, M. 1993. A cartographic view of spatial data quality. Cartographica, 30(2): 8-19.

[8] Reis, Marcelo. Análise de Séries temporais. Departamento de informática e Estatística. Universidade Federal de Santa Catarina. Disponível em: http://www.inf.ufsc.br/~marcelo.menezes.reis/Cap4.pdf. Acessado em 2 de julho de 2017.

[9] Neil J. Salkind. 2006. Statistics for People who (Think They) Hate Statistics: The Excel Edition. Página 106.

[10] Araújo, H. Entendendo e usando o SVG. Disponível em: http://www.devmedia.com.br/entendendo-e-usando-o-svg/19773. Acessado em 4 de julho de 2017.

[11] Bohling, G. 2005. Kriging. Disponível em: http://people.ku.edu/~gbohling/cpe940/Kriging.pdf. Acessado em 5 de julho de 2017.