

Exploração do *Dataset “Lisbon-Houses”*

Visualização de dados

Natalício Culuze – 75204

Ano Letivo

Github link:

2025/2026

https://github.com/NCuluze/information_Visualization

Visualizaçao de dados “Lisbon-Houses”

1. Introdução

A visualização de dados constitui um elemento central no processo de análise e comunicação de informação complexa. Para além da correta representação dos dados, a eficácia de uma visualização depende de decisões de design que influenciam diretamente a percepção, a interpretação e a memorização da informação apresentada. Assim, a visualização deve ser entendida como um processo de design fundamentado, e não apenas como um produto final.

Na fase inicial deste projeto foi realizada uma análise exploratória do *dataset Lisbon-Houses*, permitindo identificar padrões preliminares relacionados com o preço dos imóveis, a sua tipologia e a localização geográfica. No entanto, essa fase revelou limitações associadas à utilização de visualizações isoladas e estáticas, dificultando a análise integrada e comparativa dos dados.

A presente fase do trabalho tem como objetivo ultrapassar essas limitações através do desenvolvimento de um protótipo visual interativo, recorrendo à ferramenta Power BI. O protótipo visa estruturar uma narrativa visual coerente, permitindo ao utilizador explorar relações espaciais e tipológicas de forma dinâmica.

2. Pergunta de investigação e narrativa visual

Com base nos resultados da fase exploratória, foi definida a seguinte pergunta de investigação:

De que forma a localização (freguesia) e a tipologia do imóvel influenciam o preço dos imóveis em Lisboa?

A partir desta questão, estabelecem-se os seguintes objetivos:

- Explorar visualmente a variação espacial dos preços dos imóveis;
- Analisar a relação entre tipologia (número de quartos), área útil e preço;

- Desenvolver um *dashboard* interativo que permita a exploração dinâmica dos dados;

3. Conjunto de dados e preparação

O *dataset* utilizado contém informação sobre imóveis residenciais em Lisboa, incluindo variáveis quantitativas e categóricas relevantes para a análise do mercado imobiliário. As principais variáveis utilizadas neste trabalho são *Price*, *AreaNet*, *Bedrooms* e *Parish*.

No Power BI, os dados foram sujeitos a verificações e transformações básicas, garantindo o correto tipo das variáveis e a eliminação de registo incompletos. Optou-se por não remover *outliers*, uma vez que estes refletem a heterogeneidade real do mercado imobiliário e permitem uma análise mais realista da dispersão dos preços.

Esta etapa corresponde ao nível de abstração de dados, no qual os dados do domínio específico são representados como uma tabela multivariada adequada às tarefas de exploração e comparação (Munzner, 2014).

4. Metodologia de design e enquadramento teórico

4.1 Modelo de quatro níveis de design

O desenvolvimento do protótipo seguiu o modelo de quatro níveis proposto por Munzner (2014):

1. Situação de domínio;
2. Abstração de tarefas e dados;
3. Idiomas de codificação visual e interação;
4. Algoritmo.

Dado o foco do trabalho, a principal contribuição situa-se ao nível da abstração e da codificação visual, sendo estas decisões justificadas teoricamente.

4.2 Minimalismo, ornamentação e eficácia visual

Abordagens clássicas em visualização defendem a redução de elementos não essenciais, frequentemente designados como *chart junk*. No entanto, Bateman et al. (2010) demonstram que visualizações visualmente enriquecidas não apresentam perdas significativas de precisão interpretativa e podem melhorar a memorização da informação a médio prazo.

Com base nesses resultados, o protótipo desenvolvido adota uma estética equilibrada, evitando tanto o excesso decorativo como um minimalismo excessivamente austero, privilegiando a clareza estrutural e a hierarquia visual.

4.3 Uso da cor e consistência visual

A utilização da cor foi orientada por princípios de discriminação perceptiva e consistência. Gramazio et al. (2017) demonstram que paletas cromáticas eficazes devem equilibrar contraste suficiente para distinguir categorias e coerência visual global, de modo a evitar sobrecarga cognitiva.

No *dashboard*, a cor é utilizada de forma controlada para codificar categorias como freguesia, mantendo consistência entre visualizações e evitando múltiplas codificações simultâneas.

4.4 Interatividade e envolvimento do utilizador

A interatividade desempenha um papel central no protótipo desenvolvido. Através de filtros cruzados, o utilizador pode explorar subconjuntos dos dados e observar como diferentes variáveis se relacionam.

Segundo Lan et al. (2025), mecanismos de interação e composição visual contribuem para o envolvimento do utilizador e para a construção de significado, mesmo em visualizações orientadas para análise. Neste trabalho, a interatividade é utilizada como suporte ao raciocínio exploratório e não como mero elemento estético.

5. Descrição do protótipo visual

O protótipo desenvolvido consiste num *dashboard* de página única (ver nexo), integrando um mapa de preço médio por freguesia, gráficos de apoio (barras e dispersão) e cartões de indicadores. Todas as visualizações estão interligadas, permitindo uma exploração dinâmica e contextualizada dos dados.

6. Discussão dos resultados

A análise visual confirma que o preço dos imóveis em Lisboa varia significativamente entre freguesias, evidenciando desigualdades espaciais marcantes. Observa-se também uma relação positiva entre tipologia, área útil e preço, embora com elevada dispersão.

A possibilidade de filtragem interativa permite observar que imóveis com áreas semelhantes podem apresentar preços distintos consoante a freguesia, reforçando o papel da localização como fator determinante.

7. Reflexão crítica e limitações

Apesar das potencialidades do protótipo, existem limitações relevantes. A ausência de variáveis temporais e socioeconómicas restringe a profundidade da análise. Além disso, a utilização de médias pode ser influenciada por valores extremos.

Do ponto de vista do design, reconhece-se que escolhas visuais mais expressivas podem introduzir riscos de enviesamento interpretativo. Lan et al. (2025) alertam para a necessidade de contextualizar cuidadosamente visualizações com maior carga estética, sobretudo em contextos analíticos.

8. Conclusão

Este trabalho demonstrou como a transição de uma análise exploratória inicial para um protótipo visual interativo permite aprofundar a compreensão dos dados e comunicar resultados de forma mais eficaz. A integração de princípios teóricos de visualização revelou-se fundamental para justificar decisões de design e interação.

9. Referências

1. Bateman, S., Mandryk, R. L., Gutwin, C., Genest, A., McDine, D., & Brooks, C. (2010). *Useful junk? The effects of visual embellishment on comprehension and memorability of charts*. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.
2. Gramazio, C. C., Laidlaw, D. H., & Schloss, K. B. (2017). *Colorgorical: Creating discriminable and preferable color palettes for information visualization*. IEEE Transactions on Visualization and Computer Graphics, 23(1), 521–530.
3. Lan, X., Wang, Y., Peng, L., & Ma, X. (2025). *More than beautiful: Exploring design features, practical perspectives, and implications of artistic data visualization*. IEEE Computer Graphics and Applications.
4. Munzner, T. (2014). *Visualization analysis and design*. CRC Press.

ANEXO.

