



ESCOLA SUPERIOR DE TECNOLOGIA E GESTÃO
LICENCIATURA EM ENGENHARIA INFORMÁTICA

SISTEMAS DE INFORMAÇÃO

TRABALHO PRÁTICO N.º 1

Desenvolvimento de um Modelo Multidimensional
e Implementação de Análises OLAP no Contexto
da Habitação e Salários em Portugal

Inês Maria Pereira do Nascimento

Rita Alexandra Lampreia Dias



Beja, novembro de 2025

INSTITUTO POLITÉCNICO DE BEJA
ESCOLA SUPERIOR DE TECNOLOGIA E GESTÃO
LICENCIATURA EM ENGENHARIA INFORMÁTICA
SISTEMAS DE INFORMAÇÃO

TRABALHO PRÁTICO N.º 1

Desenvolvimento de um Modelo Multidimensional e Implementação de Análises OLAP no Contexto da Habitação e Salários em Portugal

Inês Maria Pereira do Nascimento - 24481

Rita Alexandra Lampreia Dias - 23240

ORIENTAÇÃO

Professora Isabel Brito

Beja, novembro de 2025

Índice

1	Resumo	7
2	Abstract	8
3	Introdução.....	9
4	Seleção dos Datasets e Análise das Granularidades.....	10
4.1	Dataset Habitação (INE)	10
4.2	Dataset Salários (PORDATA)	12
4.3	Mapeamento Territorial (NUTS ↔ Municípios)	14
5	Modelo Multidimensional	17
5.1	Perguntas Analíticas que Orientaram o Modelo	17
5.2	Arquitetura Geral do Modelo (Esquema em Constelação)	18
5.3	Componentes do Modelo	20
5.3.1	FACT_Habitação	20
5.3.2	FACT_Salário	20
5.3.3	DIM_Tempo.....	21
5.3.4	DIM_Localização.....	21
5.3.5	DIM_Categoría.....	22
5.4	Relações do Modelo.....	22
6	Processo ETL (Extract, Transform, Load)	23
6.1	Evolução do Processo ETL e Integração no Modelo Multidimensional	23
6.1.1	ETL dos Valores da Habitação (INE)	25
6.1.2	Criação da FACT_Habitação e ligação às dimensões.....	27
6.2	ETL do Salário Médio (PORDATA)	28
6.2.1	Fonte de dados e formato original.....	28
6.2.2	Extract – Importação e limpeza inicial.....	28
6.2.3	Transform – Normalização e mapeamento territorial.....	29
6.2.4	Criação da FACT_Salário	29

6.3	Construção das Dimensões a partir das Queries de ETL	30
6.4	Conclusão do Processo ETL.....	33
7	Construção do Cubo e Operações OLAP.....	33
7.1	Matriz OLAP – Habitação.....	33
7.2	Matriz OLAP – Salários Médios.....	34
7.3	Matriz OLAP Integrada – Habitação + Salários.....	35
7.4	Resultados da Análise OLAP	36
8	Análise Gráfica e Discussão dos Resultados.....	37
8.1	Análise dos Salários Médios	37
8.1.1	Distribuição Regional dos Salários – Tendências Consistentes (2019–2023).....	40
8.1.2	Evolução Temporal dos Salários	41
8.2	Análise dos Valores da Habitação	41
8.2.1	Comparação Regional: Preços por NUTS III e por Categoria	42
8.3	Análise Combinada Habitação vs Salário.....	43
8.3.1	Evolução do Salário e Habitação por Ano	43
8.3.2	Comparação Direta entre o Valor da Habitação ($€/m^2$) e o Salário Médio por NUTS III	47
8.3.3	% Diferença Habitação vs Salário.....	49
9	Demonstração Interativa das Operações OLAP e Gráficos Dinâmicos (Repositório com Vídeos)	53
9.1	Operações OLAP no cubo Habitação.....	53
9.1.1	Roll-up e Drill-down	54
9.1.2	Slice e Dice	54
9.1.3	Pivot	54
9.2	Operações OLAP no cubo Salário.....	55
9.2.1	Roll-up e Drill-down	55
9.2.2	Slice e Dice	55

9.2.3	Pivot	56
9.3	Operações OLAP no cubo Habitação + Salário	56
9.3.1	Roll-up e Drill-down	56
9.3.2	Slice e Dice.....	56
9.3.3	Pivot	57
9.4	Gráficos Dinâmico.....	57
9.4.1	Evolução do Salário e Habitação por Ano	57
9.4.2	Média de Salário por Localização	57
10	Conclusão.....	58
11	Bibliografia.....	60

Índice de Figuras

Figura 1 - Estrutura hierárquica oficial das NUTS I, II e III.	14
Figura 2 - Representação geográfica dos níveis NUTS I, NUTS II e NUTS III em Portugal....	15
Figura 3 - Tabela de correspondência oficial entre NUTS1, NUTS2 e NUTS3.....	16
Figura 4 - Mapeamento de Municípios para NUTS3, NUTS2 e NUTS1	16
Figura 5 - Esquema Modelo Multidimensional.....	19
Figura 6 - Excerto da tabela original do INE contendo os valores medianos das vendas de alojamentos familiares (€/m ²), por NUTS 2024, categoria do alojamento (H1, H11, H12) e trimestre.	23
Figura 7 - Excerto da tabela original da PORDATA com os salários médios mensais por trabalhador, organizados por ano e âmbito geográfico.....	24
Figura 8 – Excerto da Tabela FACT_Habitação.....	28
Figura 9 – Excerto da Tabela FACT_Salário	30
Figura 10 - Excerto da Tabela DIM_Tempo	31

Figura 11 - Excerto da Tabela DIM_Localização	32
Figura 12 - Tabela DIM_Categoria	32
Figura 13 - Matriz OLAP dos valores medianos da habitação (€/m ²) organizada por ano, trimestre e hierarquia territorial NUTS	34
Figura 14 – Matriz OLAP dos salários médios anuais organizados por ano e hierarquia territorial NUTS	35
Figura 15 - Matriz OLAP integrando valores da habitação (€/m ²) e salário médio por região (NUTS1–NUTS3), para os anos 2019-2023	36
Figura 16 - Distribuição da Média de Salário (euros) por região NUTS III no ano de 2019	38
Figura 17 - Distribuição da Média de Salário (euros) por região NUTS III no ano de 2020	38
Figura 18 - Distribuição da Média de Salário (euros) por região NUTS III no ano de 2021	39
Figura 19 - Distribuição da Média de Salário (euros) por região NUTS III no ano de 2022	39
Figura 20 - Distribuição da Média de Salário (euros) por região NUTS III no ano de 2023	40
Figura 22 - Média do Valor da Habitação (€/m ²) por região NUTS III e categoria de alojamento (Existentes, Novos e Total).....	42
Figura 23 - Evolução da Média de Valor da Habitação (€/m ²) e da Média de Salário (€/mês) por Ano no Baixo Alentejo	44
Figura 24 - Evolução da Média de Valor da Habitação (€/m ²) e da Média de Salário (€/mês) por Ano em Viseu Dão Lafões.....	45
Figura 25 - Evolução da Média de Valor da Habitação (€/m ²) e da Média de Salário (€/mês) por Ano na Grande Lisboa	46
Figura 26 - Evolução da Média de Valor da Habitação (€/m ²) e da Média de Salário (€/mês) por Ano na Região Autónoma da Madeira.....	47
Figura 27 - Comparação da Média de Valor da Habitação (€/m ²) e da Média de Salário (€/mês) por Região NUTS III (no período temporal de 2019 a 2023)	48
Figura 28 - Criação de nova medida % Diferença Habitação vs Salário	49
Figura 29 - % Diferença entre o Valor Médio da Habitação (€/m ²) e o Salário Médio (euros) por Região NUTS III	51

1 Resumo

Este trabalho desenvolve um modelo multidimensional para analisar, de forma integrada, a evolução dos salários médios e dos valores de habitação ($\text{€}/\text{m}^2$) nas regiões NUTS III de Portugal entre 2019 e 2023. Para isso, foi construído um processo ETL estruturado, garantindo limpeza, normalização e alinhamento temporal e territorial dos dados. A análise OLAP permitiu responder às questões inicialmente definidas: evolução temporal dos salários e dos preços da habitação, identificação de padrões territoriais e avaliação da sua relação. Os resultados mostram que os salários apresentam uma evolução moderada e relativamente homogénea, enquanto os preços da habitação aumentam de forma mais acentuada e desigual. A comparação entre ambas as variáveis revela assimetrias claras entre litoral e interior, com regiões metropolitanas e turísticas a apresentarem forte pressão imobiliária e menor acessibilidade relativa. Este trabalho demonstra a utilidade do modelo analítico criado para apoiar estudos sobre rendimento, habitação e políticas regionais.

Palavras-chave:

Salários, Habitação, Preços de Imobiliário, NUTS, OLAP, Data Warehouse, ETL,
Acessibilidade Habitacional

2 Abstract

This study develops a multidimensional model to analyse the evolution of average salaries and housing prices ($\text{€}/\text{m}^2$) across Portugal's NUTS III regions between 2019 and 2023. A structured ETL process was implemented to ensure data cleaning, normalisation and consistent temporal and territorial alignment. The OLAP analysis addressed the initial research questions regarding the temporal evolution of salaries and housing values, territorial patterns, and the relationship between these variables. Findings show that salaries follow a moderate and relatively homogeneous growth trend, whereas housing prices increase more sharply and unevenly across regions. The combined analysis highlights strong asymmetries between coastal metropolitan/touristic areas and the interior, with the former showing significant housing pressure and lower affordability. The study demonstrates the value of the proposed analytical model as a support tool for research on income, housing markets and regional policy.

Keywords:

Salaries, Housing, Real Estate Prices, NUTS, OLAP, Data Warehouse, ETL, Housing Affordability

3 Introdução

O presente projeto tem como objetivo a construção de um modelo multidimensional e a realização de análises OLAP aplicadas à evolução do mercado habitacional em Portugal, cruzando esta informação com a dinâmica dos salários médios ao longo do tempo. Para este efeito, foram utilizados dados provenientes de duas fontes oficiais: o Instituto Nacional de Estatística (INE), que disponibiliza indicadores sobre o valor mediano de venda de alojamentos familiares ($\text{€}/\text{m}^2$), e a PORDATA, que reúne estatísticas referentes ao ganho médio mensal dos trabalhadores por conta de outrem.

A motivação inicial do trabalho centrou-se na análise da evolução temporal e espacial dos preços da habitação, procurando identificar tendências, variações regionais e potenciais desequilíbrios territoriais a partir da hierarquia estatística NUTS I, NUTS II e NUTS III. Durante a exploração preliminar dos dados, considerou-se pertinente integrar a informação relativa aos salários médios, permitindo contextualizar a acessibilidade à habitação e avaliar de que forma o rendimento acompanha (ou não) a evolução dos preços. Este cruzamento enriquece a análise, levantando questões relevantes como: “As regiões com habitação mais cara são também as que apresentam salários mais elevados?” ou “A capacidade de compra melhora ou agrava-se ao longo dos anos?”

A necessidade de harmonizar estas fontes tornou evidente a existência de diferentes granularidades temporais (anos vs. trimestres) e geográficas (municípios vs. NUTS), bem como discrepâncias estruturais entre datasets. Isto exigiu um processo de ETL que garantisse a uniformização dos dados, a reconstrução de chaves temporais consistentes, a normalização territorial e a eliminação de duplicações ou valores em falta, preparando assim a informação para um modelo dimensional sólido. O resultado deste processo foi a implementação, em Power BI, de um modelo em estrela composto por duas tabelas de factos — Habitação e Salários — e três dimensões — Tempo, Localização e Categoria — permitindo realizar operações OLAP e construir visualizações analíticas claras e comparáveis.

Este relatório encontra-se organizado da seguinte forma:

1. É apresentada a seleção dos datasets e a análise das suas granularidades, justificando as decisões tomadas no processo de integração.
2. Segue-se a construção do modelo multidimensional, incluindo o diagrama obtido e a função de cada dimensão e tabela de factos.
3. Posteriormente descreve-se o processo ETL implementado, detalhando os tratamentos, transformações e harmonizações realizadas.
4. De seguida é apresentada a exploração OLAP, construção do CUBO.
5. Análise dos resultados e cruzamento entre os vários indicadores.
6. Por fim, são discutidas as principais conclusões.

4 Seleção dos Datasets e Análise das Granularidades

4.1 Dataset Habitação (INE)

A primeira etapa do trabalho consistiu na identificação e seleção dos datasets a integrar no modelo multidimensional. A fonte inicial escolhida foi o Instituto Nacional de Estatística (INE), nomeadamente o indicador **“Valor mediano das vendas de alojamentos familiares nos últimos 12 meses (Metodologia 2022 - €/m²) por Localização geográfica (NUTS 2024) e Categoria do alojamento familiar; Trimestral”**, disponível no portal do INE (Estatística). O site permite configurar a extração dos dados conforme diversas dimensões analíticas, incluindo:

Localização geográfica:

- Município
- NUTS 3
- NUTS 2
- NUTS 1

Tempo

- Trimestre
- Ano (agregado a partir dos trimestres)

Categoria de alojamento

- Total
- Alojamentos novos
- Alojamentos existentes

Durante a exploração inicial dos dados do INE percebemos que a combinação mais detalhada: trimestre x município, tinha muitos valores em falta em vários municípios e em determinados trimestres. Isto acontece porque, em alguns períodos, não existe informação suficiente para calcular o valor mediano, deixando os dados incompletos. Perante este problema, analisámos duas possíveis soluções:

1) Remover os regtos que tinham valores nulos

Esta opção parecia simples, mas tinha um grande inconveniente: ao eliminar esses regtos, várias regiões ficavam com falhas ao longo do tempo. Isso tornaria a análise inconsistente e criaria diferenças artificiais entre regiões, dificultando qualquer comparação.

2) Preencher os nulos usando o valor da região acima (por exemplo, o valor da NUTS3)

Esta abordagem resolveria a falta de valores, mas teria um problema ainda maior: estariam a atribuir aos municípios um valor que não lhes pertence. Isto criaria enviesamento, porque os dados deixavam de refletir a realidade e passavam a ser uma aproximação inventada, podendo levar a interpretações erradas.

Depois de avaliar estas opções, concluímos que nenhuma delas era adequada. Assim, optámos por trabalhar diretamente ao nível das NUTS (1, 2 e 3), que:

- têm todos os valores completos,
- não apresentam nulos,
- permitem análises consistentes entre regiões,
- e garantem que utilizamos sempre dados reais, sem enviesamento.

Além disso, a granularidade NUTS3 continua a oferecer bastante detalhe (26 regiões), o que é mais do que suficiente para suportar as operações OLAP pretendidas no trabalho.

Esta decisão permitiu construir uma base sólida, limpa e íntegra para o cubo multidimensional.

4.2 Dataset Salários (PORDATA)

Depois de concluída a análise inicial dos dados da habitação, percebemos que, embora fosse possível estudar a evolução dos preços ao longo do tempo e comparar regiões, faltava uma perspetiva mais completa sobre o impacto real destes valores na vida das pessoas. O preço da habitação por si só mostra o custo de viver numa região, mas não revela se a população tem ou não capacidade económica para acompanhar esse aumento.

Por essa razão, tornou-se evidente a necessidade de complementar o modelo com um indicador que permitisse analisar também o nível de rendimento das populações. Só assim seria possível avaliar questões como:

- a acessibilidade à habitação,
- a evolução do custo de vida,
- desigualdades regionais em termos de qualidade de vida,
- e a relação entre preços elevados e rendimento médio das pessoas.

Dentro das fontes disponíveis, escolhemos a PORDATA, utilizando o indicador “Ganho médio mensal dos trabalhadores por conta de outrem” (PORDATA), por se tratar de uma base de dados oficial, atualizada e comparável com outros indicadores socioeconómicos. Ao introduzir o salário médio no modelo, passámos a conseguir cruzar diretamente:

- custo da habitação ($\text{€}/\text{m}^2$)
- rendimento médio mensal
- localização
- evolução temporal

permitindo uma leitura mais completa sobre o equilíbrio económico das diferentes regiões.

Tal como no caso da habitação, este dataset comporta:

- Granularidade temporal: apenas anual
- Granularidade geográfica: município

Assim, tornaram-se evidentes duas diferenças estruturais face aos dados da habitação:

1. Tempo:

- Habitação → Trimestres
- Salários → Apenas Ano

2. Localização:

- Habitação → NUTS
- Salários → Municípios

Estas disparidades tornariam impossível um cruzamento direto entre os dois indicadores.

Inicialmente, e durante a fase exploratória, criámos três ETL distintos:

- ETL Habitação
- ETL Salários
- ETL Habitação + Salários (tentativa inicial de integração direta)

Contudo, rapidamente se concluiu que esta abordagem criaria inconsistências e duplicações.

Para resolver este problema estruturámos o modelo final como um esquema de constelação, com duas tabelas de factos independentes (Habitação e Salários), partilhando dimensões comuns.

4.3 Mapeamento Territorial (NUTS ↔ Municípios)

Para garantir uma integração consistente entre os diferentes níveis geográficos presentes nos nossos datasets, tornou-se necessário compreender e representar adequadamente a estrutura oficial de organização territorial utilizada em Portugal. Os dados da habitação seguem a classificação NUTS (Nomenclature of Territorial Units for Statistics), que organiza o país em três níveis hierárquicos (NUTS I, NUTS II e NUTS III) enquanto os dados dos salários se encontram ao nível do município.

As imagens seguintes (A nova configuração de NUTS 2024 - Escola Virtual - Areal Editores) evidenciam a forma como o território português está subdividido segundo a nomenclatura estatística oficial. A Figura 1 mostra a tabela oficial de correspondência entre NUTS I, NUTS II e NUTS III, clarificando quantas unidades existem em cada nível e como estas se agrupam. A Figura 2 é uma representação gráfica dos três níveis NUTS (I, II e III), permitindo visualizar a distribuição territorial de cada nível e a forma como as regiões se organizam hierarquicamente.

NUTS I nível nacional	NUTS II nível regional	NUTS III nível sub-regional
Portugal Continental	Norte	Alto Minho Cávado Ave Área Metropolitana do Porto Alto Tâmega e Barroso Tâmega e Sousa Douro Terras de Trás-os-Montes
	Centro	Região de Aveiro Região de Coimbra Região de Leiria Viseu Dão Lafões Beira Baixa Beiras e Serra da Estrela
	Oeste e Vale do Tejo	Oeste Médio Tejo Lezíria do Tejo
	Grande Lisboa	Grande Lisboa
	Península de Setúbal	Península de Setúbal
	Alentejo	Alentejo Litoral Baixo Alentejo Alto Alentejo Alentejo Central
	Algarve	Algarve
Região Autónoma da Madeira	Região Autónoma da Madeira	Região Autónoma da Madeira
Região Autónoma dos Açores	Região Autónoma dos Açores	Região Autónoma dos Açores
Total: 3 unidades territoriais	Total: 9 unidades territoriais	Total: 26 unidades territoriais

Figura 1 - Estrutura hierárquica oficial das NUTS I, II e III.



Figura 2 - Representação geográfica dos níveis NUTS I, NUTS II e NUTS III em Portugal.

Durante a integração dos dados, tornou-se evidente a necessidade de criar tabelas auxiliares de mapeamento territorial, essenciais para garantir coerência entre fact tables com granularidades diferentes. Foram desenvolvidas duas tabelas fundamentais:

1. Tabela de NUTS (para Habitação)

Contém a hierarquia territorial NUTS1 → NUTS2 → NUTS3, necessária para permitir roll-up e drill-down na dimensão localização.

Nome_NUTS1	Código_NUTS1	Nome_NUTS2	Código_NUTS2	Nome_NUTS3	Código_NUTS3
Continente	1 Norte			11 Alto Minho	111
Continente	1 Norte			11 Cávado	112
Continente	1 Norte			11 Ave	119
Continente	1 Norte			11 Área Metropolitana do Porto	11A
Continente	1 Norte			11 Alto Tâmega e Barroso	11B
Continente	1 Norte			11 Tâmega e Sousa	11C
Continente	1 Norte			11 Douro	11D
Continente	1 Norte			11 Terras de Trás-os-Montes	11E
Continente	1 Centro			19 Região de Aveiro	191
Continente	1 Centro			19 Região de Coimbra	192
Continente	1 Centro			19 Região de Leiria	193
Continente	1 Centro			19 Viseu Dão Lafões	194
Continente	1 Centro			19 Beira Baixa	195
Continente	1 Centro			19 Beiras e Serra da Estrela	196
Continente	1 Oeste e Vale do Tejo	1D		Oeste	1D1
Continente	1 Oeste e Vale do Tejo	1D		Médio Tejo	1D2
Continente	1 Oeste e Vale do Tejo	1D		Lezíria do Tejo	1D3
Continente	1 Grande Lisboa	1A		Grande Lisboa	1A0
Continente	1 Península de Setúbal	1B		Península de Setúbal	1B0
Continente	1 Alentejo	1C		Alentejo Litoral	1C1
Continente	1 Alentejo	1C		Baixo Alentejo	1C2
Continente	1 Alentejo	1C		Alto Alentejo	1C3
Continente	1 Alentejo	1C		Alentejo Central	1C4
Continente	1 Algarve			15 Algarve	150
Região Autónoma dos Açores	2 Região Autónoma dos Açores			20 Região Autónoma dos Açores	200
Região Autónoma da Madeira	3 Região Autónoma da Madeira			30 Região Autónoma da Madeira	300

Figura 3 - Tabela de correspondência oficial entre NUTS1, NUTS2 e NUTS3

2. Tabela de Mapeamento NUTS3 → Municípios (para Salários)

Como os salários estão ao nível do município e a habitação ao nível NUTS3, foi criada uma tabela que associa cada município ao respetivo código NUTS3, permitindo harmonizar o nível geográfico.

Município	Nome_NUTS3	Código_NUTS3	Nome_NUTS2	Código_NUTS2	Nome_NUTS1	Código_NUTS1
Abrantes	Médio Tejo		1D2 Oeste e Vale do Tejo	1D	Continente	1
Águeda	Região de Aveiro		191 Centro	19	Continente	1
Aguiar da Beira	Beiras e Serra da Estrela		196 Centro	19	Continente	1
Alandroal	Alentejo Central		1C4 Alentejo	1C	Continente	1
Albergaria-a-Velha	Região de Aveiro		191 Centro	19	Continente	1
Albufeira	Algarve		150 Algarve	15	Continente	1
Alcácer do Sal	Alentejo Litoral		1C1 Alentejo	1C	Continente	1
Alcanena	Médio Tejo		1D2 Oeste e Vale do Tejo	1D	Continente	1
Alcobaça	Região de Leiria		193 Centro	19	Continente	1
Alcoutim	Algarve		150 Algarve	15	Continente	1
Alenquer	Oeste		1D1 Oeste e Vale do Tejo	1D	Continente	1
Alfândega da Fé	Terras de Trás-os-Montes		11E Norte	11	Continente	1
Alijó	Douro		11D Norte	11	Continente	1
Aljezur	Algarve		150 Algarve	15	Continente	1
Aljustrel	Baixo Alentejo		1C2 Alentejo	1C	Continente	1
Almada	Península de Setúbal		1B0 Península de Setúbal	1B	Continente	1
Almeida	Beiras e Serra da Estrela		196 Centro	19	Continente	1
Almeirim	Lezíria do Tejo		1D3 Oeste e Vale do Tejo	1D	Continente	1
Almodôvar	Baixo Alentejo		1C2 Alentejo	1C	Continente	1
Alpiarça	Lezíria do Tejo		1D3 Oeste e Vale do Tejo	1D	Continente	1
Alter do Chão	Alto Alentejo		1C3 Alentejo	1C	Continente	1
Alvaiázere	Região de Leiria		193 Centro	19	Continente	1
Alvito	Baixo Alentejo		1C2 Alentejo	1C	Continente	1
Amadora	Grande Lisboa		1A0 Grande Lisboa	1A	Continente	1

Figura 4 - Mapeamento de Municípios para NUTS3, NUTS2 e NUTS1

Estas tabelas foram essenciais para possibilitar:

- ligações corretas no modelo dimensional
- análises OLAP cruzadas
- visualizações consistentes entre fact tables
- ausência de duplicações ou ambiguidade geográfica

5 Modelo Multidimensional

A construção do modelo multidimensional foi guiada por um princípio fundamental recomendado em processos de Business Intelligence: antes de modelar, definem-se as perguntas analíticas que o sistema deverá ser capaz de responder. Só depois se identificam as métricas necessárias, as dimensões adequadas e a granularidade correta.

Assim, o modelo criado resulta diretamente dos objetivos analíticos definidos na fase inicial do projeto — compreender a evolução do mercado habitacional português e contextualizar essa evolução com informação sobre salários médios regionais.

5.1 Perguntas Analíticas que Orientaram o Modelo

Antes de qualquer decisão de modelação, identificámos as principais questões que o Data Warehouse deveria responder:

Evolução temporal

- Como variaram os preços de habitação ao longo dos últimos anos?
- A evolução é homogénea entre anos e trimestres?

Análise espacial

- Quais as diferenças entre regiões NUTS I, NUTS II e NUTS III?
- Que regiões apresentam valores mais elevados ou mais baixos?

Análise por categoria

- Habitação nova valoriza mais do que habitação existente?
- Existem padrões consistentes entre as três categorias (Novo, Existente, Total)?

Relação Habitação/Salários

- As regiões mais caras são também as mais bem remuneradas?
- Há zonas onde o custo da habitação evolui mais rápido do que o salário médio?
- Existem regiões em equilíbrio ou em claro desequilíbrio entre rendimento e preço?

Estas perguntas permitiram estruturar claramente quais métricas necessitávamos (valor da habitação, salário médio, percentagens, médias), e quais dimensões seriam indispensáveis (Tempo, Localização, Categoria).

5.2 Arquitetura Geral do Modelo (Esquema em Constelação)

Tendo em conta estes objetivos, optámos por um esquema em constelação, adequado quando existe mais do que uma métrica principal, proveniente de origens diferentes, mas que deve ser analisada num ambiente OLAP comum. O modelo final inclui:

Tabelas de Factos

- FACT_Habitação - núcleo da análise, com valor mediano por m², por trimestre, categoria e região.
- FACT_Salário - tabela complementar, com salário médio anual mapeado para NUTS3.

Dimensões Partilhadas

- DIM_Tempo — comum às duas fact tables.
- DIM_Localização — comum às duas fact tables, garantindo coerência territorial.

Dimensão Exclusiva da Habitação

- DIM_Categoria — Novo, Existente e Total.

No Figura 5 apresenta-se o esquema do modelo multidimensional obtido.

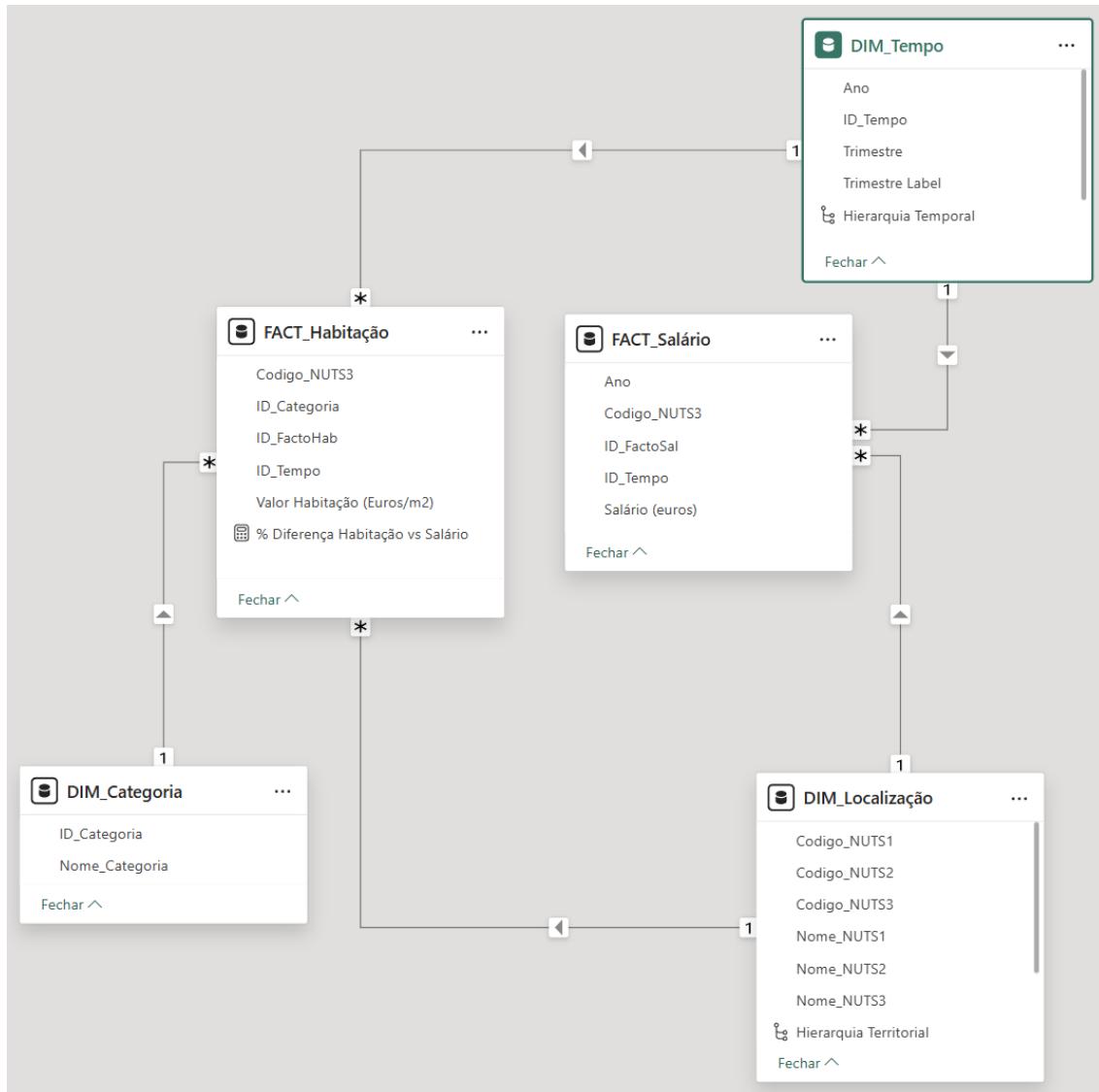


Figura 5 - Esquema Modelo Multidimensional

Este tipo de arquitetura permite:

- integrar datasets com granularidades distintas (trimestres vs ano);
- cruzar métricas de diferentes naturezas sem duplicação;
- garantir que operações OLAP funcionam de forma consistente (drill-down, roll-up, slice, dice);
- manter o modelo expansível para futuras integrações (ex.: indicadores socioeconómicos adicionais).

Esta arquitetura representada na figura permite não só analisar a habitação e os salários de forma independente, mas também cruzar ambas as métricas, possibilitando indicadores derivados como:

- diferença percentual habitação vs salário;
- comparação entre NUTS III com maior desequilíbrio económico;
- ranking de regiões por custo de vida ajustado ao rendimento médio.

5.3 Componentes do Modelo

5.3.1 FACT_Habitação

A tabela FACT_Habitação é a tabela de factos principal do modelo e armazena todas as métricas relacionadas com o valor mediano da habitação.

Chave primária: ID_FactoHab - Criada no Power Query através de uma coluna de índice, garantindo um identificador único para cada registo.

Chaves estrangeiras:

- Código_NUTS3 → DIM_Localização
- ID_Categoría → DIM_Categoría
- ID_Tempo → DIM_Tempo

Medida: Valor Habitação (Euros/m²) — métrica principal fornecida pelo INE

5.3.2 FACT_Salário

A tabela FACT_Salário complementa a análise com o salário médio anual após mapeamento dos municípios para NUTS3.

Chave primária: ID_FactoSal - Também gerado no Power Query como coluna de índice.

Chaves estrangeiras:

- Código_NUTS3 → DIM_Localização
- ID_Tempo → DIM_Tempo

Medida: Salário (euros) — salário médio anual

5.3.3 DIM_Tempo

Chave primária: ID_Tempo (ex.: “2020_1”)

Atributos:

- Ano (ex.: “2020”)
- Trimestre (ex.: “1”)
- Trimestre Label (ex.: “1º Trimestre”)

Hierarquia Temporal:

- Ano
- Trimestre Label

Suporta drill-down, roll-up e análise temporal multigranular.

5.3.4 DIM_Localização

Chave primária: Código_NUTS3

Atributos:

- Código_NUTS1
- Código_NUTS2
- Código_NUTS3
- Nome_NUTS1
- Nome_NUTS2
- Nome_NUTS3

Hierarquia Territorial:

- Nome_NUTS1
- Nome_NUTS2
- Nome_NUTS3

Fundamental para análises OLAP geográficas.

5.3.5 DIM_Categoria

Chave primária: ID_Categoria

Atributo: Nome_Categoria (Novos, Existentes, Total)

Simples mas necessária para segmentar o mercado habitacional.

5.4 Relações do Modelo

As relações refletem um esquema em constelação, onde duas fact tables partilham dimensões.

FACT_Habitação

- DIM_Tempo → 1:*
- DIM_Localização → 1:*
- DIM_Categoria → 1:*

FACT_Salário

- DIM_Tempo → 1:*
- DIM_Localização → 1:*

6 Processo ETL (Extract, Transform, Load)

6.1 Evolução do Processo ETL e Integração no Modelo Multidimensional

O processo ETL foi, na prática, a parte mais trabalhosa do projeto. Numa fase inicial, depois de exportarmos os dados do INE e da PORDATA, começámos por tratar tudo de forma muito direta no Power Query: limpávamos cada ficheiro e fazíamos vários merges entre tabelas, construindo uma grande tabela combinada com informação de habitação e salários. Esta abordagem produziu rapidamente resultados visuais, mas trazia vários problemas: colunas redundantes, dificuldade em criar hierarquias estáveis, ambiguidade territorial, afastado da ideia de Data Warehouse multidimensional.

As Figuras 6 e 7 (tabelas originais em Excel do INE e da PORDATA) mostram bem o ponto de partida: duas fontes completamente diferentes em estrutura e granularidade. A primeira está em formato “largo”, com uma coluna para cada combinação de ano, trimestre e categoria; a segunda apresenta salários anuais por município. Tentámos inicialmente “forçar” a integração destas duas realidades numa só tabela, o que se revelou pouco robusto quando começámos a pensar em operações OLAP, hierarquias NUTS e reutilização de dimensões.

Localização geográfica (NUTS - 2024) (1)						
	2.º Trimestre de 2025			1.º Trimestre de 2025		
	H1: Total €/ m ²	H11: Novos €/ m ²	H12: Existentes €/ m ²	H1: Total €/ m ²	H11: Novos €/ m ²	H12: Existentes €/ m ²
PT: Portugal	1923	2268	1841	1845	2202	1762
1: Continente	1925	2244	1850	1846	2181	1770
11: Norte	1694	2045	1600	1621	1986	1524
111: Alto Minho	1345	1579	1250	1313	1629	1193
112: Cávado	1657	1807	1600	1586	1730	1533
119: Ave	1420	1538	1359	1322	1432	1254
11A: Área Metropolitana do Porto	2129	2593	1995	2047	2524	1912
11B: Alto Tâmega e Barroso	860	1066	804	833	1020	768
11C: Tâmega e Sousa	1069	1266	984	1010	1241	898
11D: Douro	665	1006	600	640	1041	577
11E: Terras de Trás-os-Montes	787	1176	688	750	1155	635
19: Centro	1121	1515	1035	1082	1488	994
191: Região de Aveiro	1461	2037	1366	1422	1962	1322
192: Região de Coimbra	1291	1570	1229	1245	1530	1184
193: Região de Leiria	1290	1507	1227	1242	1485	1187
194: Viseu Dão Lafões	893	1378	777	847	1361	741
195: Beira Baixa	673	1117	609	619	1155	574

Figura 6 - Excerto da tabela original do INE contendo os valores medianos das vendas de alojamentos familiares (€/ m²), por NUTS 2024, categoria do alojamento (H1, H11, H12) e trimestre

01. Ano	02. Nome Região (Portugal)	03. Âmbito Geográfico	07. Escala	09. Valor
2004	Portugal		euros	877.5
2005	Portugal		euros	907.2
2006	Portugal		euros	934.0
2007	Portugal		euros	963.3
2008	Portugal		euros	1008.0
2009	Portugal		euros	1034.2
2010	Portugal		euros	
2011	Portugal		euros	
2012	Portugal		euros	
2013	Portugal		euros	
2014	Portugal		euros	
2015	Portugal		euros	1094.1
2016	Portugal		euros	1105.6
2017	Portugal		euros	1130.8
2018	Portugal		euros	1166.9
2019	Portugal		euros	1206.3
2020	Portugal		euros	1247.2
2021	Portugal		euros	1289.5
2022	Portugal		euros	1362.4
2023	Portugal		euros	1460.8
2004	Abrantes	Município	euros	800.7
2004	Águeda	Município	euros	740.9
2004	Aguiar da Beira	Município	euros	563.7
2004	Alandroal	Município	euros	642.4
2004	Albergaria-a-Velha	Município	euros	769.3

Figura 7 - Excerto da tabela original da PORDATA com os salários médios mensais por trabalhador, organizados por ano e âmbito geográfico.

Depois desta primeira tentativa, parámos para reavaliar o processo. A coexistência de duas métricas distintas, com granularidades diferentes (habitação: trimestral × NUTS3; salários: anual × município), exigia uma abordagem mais alinhada com a modelação dimensional. Reestruturámos então o ETL com três princípios:

- (i) tratar cada fonte de dados em queries de staging separadas;
- (ii) criar tabelas de mapeamento territorial (NUTS e Municípios→NUTS3) como base das dimensões;
- (iii) derivar, a partir daí, tabelas de factos e dimensões limpas, ligadas por chaves técnicas e com granularidade bem definida.

As restantes subsecções descrevem este processo final.

6.1.1 ETL dos Valores da Habitação (INE)

6.1.1.1 Fonte de dados e formato original

O primeiro dataset utilizado foi o do INE, “Valor mediano das vendas de alojamentos familiares ($\text{€}/\text{m}^2$)”. Cada linha corresponde a uma localização (Portugal, Continente, NUTS I, NUTS II, NUTS III) e cada coluna representa uma combinação de Ano \times Trimestre \times Categoria (H1, H11, H12). É um formato que permite uma leitura rápida e simples de dados estatísticos, mas pouco adequado a um modelo dimensional.

O objetivo do ETL foi transformar esta estrutura larga (na horizontal) numa tabela longa (na vertical), onde cada linha representa uma observação única: Ano, Trimestre, Região NUTS3, Categoria, Valor mediano ($\text{€}/\text{m}^2$).

Esta tabela serviu depois de base para a **FACT_Habitação** e para parte das dimensões de Tempo, Localização e Categoria.

6.1.1.2 Extract – Importação e limpeza inicial

No Power Query foram aplicados os seguintes passos:

1. Importação do ficheiro INE
 - Seleção da folha que contém o quadro com os valores.
2. Remoção de ruído
 - Eliminação de linhas de cabeçalho “decorativas” (títulos, datas de extração, notas, URL do INE) e de linhas de rodapé com observações.
3. Promoção de cabeçalhos
 - Transformação da primeira linha útil em cabeçalhos de coluna.

Nesta fase, a tabela continua em formato largo, mas com apenas a informação relevante.

6.1.1.3 Transform – Normalização temporal e categorial (unpivot)

Os passos seguintes tiveram como objetivo separar claramente tempo e categoria:

1. Seleção da coluna de localização
 - Foi preservada a coluna “Localização geográfica (NUTS – 2024)”.
2. Unpivot das restantes colunas
 - Utilizou-se a operação *Transformar* → *Anular dinamização de outras colunas*, convertendo dezenas de colunas num par:
 - Atributo – texto com ano, trimestre e categoria (ex.: 2019 T4 H11);
 - Valor – valor mediano (€/m²).
3. Divisão da coluna Atributo
 - A coluna foi dividida em:
 - Ano
 - Trimestre
 - ID_Categoría (H1 – Total, H11 – Novos, H12 – Existentes).
4. Conversão de tipos de dados
 - Ano - número inteiro;
 - Trimestre - texto normalizado (1, 2, 3, 4);
 - Valor - número decimal.

6.1.1.4 Transform – Estruturação geográfica (NUTS)

Para integrar a hierarquia territorial NUTS:

1. Separação de código e nome
 - o A coluna do INE contém entradas como "1D3: Lezíria do Tejo".
 - o Foi dividida em duas colunas:
 - Código Territorial (INE)
 - Nome da Localização.
2. Carregamento da tabela NUTS 2024
 - o Tabela auxiliar com:
 - Código/Nome NUTS1
 - Código/Nome NUTS2
 - Código/Nome NUTS3
3. Merge com a tabela NUTS
 - o Left Outer Join entre a tabela de habitação e a tabela NUTS:
 - [Código Territorial INE] = [Código_NUTS3].
4. Expansão das colunas NUTS
 - o Adição dos campos Código_NUTS1, Código_NUTS2, Código_NUTS3, Nome_NUTS1, Nome_NUTS2, Nome_NUTS3.
5. Filtragem para manter apenas NUTS III
 - o Linhas agregadas (Portugal, Continente, NUTS1, NUTS2) surgem com NUTS3 a null e foram removidas.
 - o Desta forma, a granularidade final da fact table ficou em NUTS3.
6. Ordenação final
 - o Ordenação por Código_NUTS3 → Ano → Trimestre → Categoria.

6.1.2 Criação da FACT_Habitação e ligação às dimensões

A query final de habitação, foi usada para gerar a FACT_Habitação (Figura 8):

- Foi criada uma coluna de índice, renomeada para ID_FactoHab, que funciona como chave primária técnica da FACT_Habitação.
- As colunas Ano e Trimestre são usadas para gerar posteriormente a DIM_Tempo;

- A coluna ID_Categoria alimenta a DIM_Categoria e os campos NUTS alimentam a DIM_Localização.

	\sum ID_FactoHab	Avg ID_Categoria	Avg Código_NUTS3	Max ID_Tempo	\sum Valor Habitação (Euros/m2)
1	0 H1	111	2019_4		846
2	1 H11	111	2019_4		1000
3	2 H12	111	2019_4		813
4	3 H1	111	2023_4		1108
5	4 H11	111	2023_4		1280
6	5 H12	111	2023_4		1044
7	6 H1	111	2023_3		1091
8	7 H11	111	2023_3		1326
9	8 H12	111	2023_3		1019
10	9 H1	111	2023_2		1077
11	10 H11	111	2023_2		1348
12	11 H12	111	2023_2		989
13	12 H1	111	2023_1		1073
14	13 H11	111	2023_1		1357
15	14 H12	111	2023_1		986
16	15 H1	112	2020_4		997

Figura 8 – Excerto da Tabela FACT_Habitação

6.2 ETL do Salário Médio (PORDATA)

6.2.1 Fonte de dados e formato original

O segundo dataset utilizado foi o “Ganho médio mensal dos trabalhadores por conta de outrem” (PORDATA). Tal como no caso da habitação, o objetivo foi transformar esta estrutura numa tabela longa do tipo: Município, Ano, Salário médio, para depois a alinhar com a mesma hierarquia NUTS usada no INE.

6.2.2 Extract – Importação e limpeza inicial

No Power Query:

- Foram removidas colunas desnecessárias (metadados, códigos internos da PORDATA, etc.);
- O salário foi convertido para formato numérico, tratando a vírgula como separador decimal.

6.2.3 Transform – Normalização e mapeamento territorial

1. Normalização dos nomes de municípios
 - o Alguns nomes não coincidiam exatamente com a tabela de mapeamento (acentos, hífenes, etc.).
 - o Foram ajustados manualmente para garantir correspondência no *merge*.
2. Mapa Município → NUTS3
 - o Foi construída uma tabela auxiliar com as colunas: Município, Nome_NUTS3, Código_NUTS3, Nome_NUTS2, Código_NUTS2, Nome_NUTS1, Código_NUTS1.
3. Merge com a tabela de mapeamento
 - o Left Outer Join entre o dataset de salários e a tabela Município → NUTS3, usando o nome do município.
 - o Expansão das colunas NUTS para associar cada salário ao respetivo NUTS3 (e níveis superiores).
4. Conversão de tipos de dados
 - o Ano → inteiro;
 - o Salário → decimal;
 - o Códigos NUTS → texto.

6.2.4 Criação da FACT_Salário

A query final de salários, foi usada para gerar a FACT_Salário (Figura 9):

- Foi adicionada uma coluna de índice, renomeada para ID_FactoSal, que serve de chave primária técnica;
- Os campos Ano foram mapeados para a DIM_Tempo, através de linhas anuais específicas (ID_Tempo do tipo 2019_0, 2020_0, etc);
- Os campos Código_NUTS3 ligam-se diretamente à DIM_Localização.

¹ ₂ ₃ ID_FactoSal	A ^B _C	Código_NUTS3	¹ ₂ ₃ Ano	1.2 Salário (euros)	A ^B _C	ID_Tempo
0	111		2019	948,97	2019_0	
1	111		2019	948,97	2019_0	
2	111		2019	948,97	2019_0	
3	111		2023	1157,51	2023_0	
4	111		2023	1157,51	2023_0	
5	111		2023	1157,51	2023_0	
6	111		2023	1157,51	2023_0	
7	111		2023	1157,51	2023_0	
8	111		2023	1157,51	2023_0	
9	111		2023	1157,51	2023_0	
10	111		2023	1157,51	2023_0	
11	111		2023	1157,51	2023_0	
12	111		2023	1157,51	2023_0	
13	111		2023	1157,51	2023_0	
14	111		2023	1157,51	2023_0	
15	112		2020	1006,62	2020_0	
16	112		2020	1006,62	2020_0	
17	112		2020	1006,62	2020_0	
18	112		2020	1006,62	2020_0	
19	112		2020	1006,62	2020_0	
20	112		2020	1006,62	2020_0	
21	112		2020	1006,62	2020_0	
22	112		2020	1006,62	2020_0	
23	111		2020	987,86	2020_0	

Figura 9 – Exerto da Tabela FACT_Salário

6.3 Construção das Dimensões a partir das Queries de ETL

As dimensões usadas no modelo em constelação resultam diretamente das queries de ETL descritas acima:

- DIM_Tempo (Figura 10)
 - Gerada a partir dos anos e trimestres presentes nos dados de habitação e salários;
 - Inclui linhas trimestrais (2019_1 ... 2023_4) e linhas anuais agregadas (2019_0 ... 2023_0 com “Todos” no Trimestre Label);
 - Suporta a hierarquia “Ano, Trimestre Label”.

A ^B _C ID_Tempo	1 ² ₃ Ano	A ^B _C Trimestre	A ^B _C Trimestre Label
2019_4	2019	4	4º Trimestre
2020_1	2020	1	1º Trimestre
2020_2	2020	2	2º Trimestre
2020_3	2020	3	3º Trimestre
2020_4	2020	4	4º Trimestre
2021_1	2021	1	1º Trimestre
2021_2	2021	2	2º Trimestre
2021_3	2021	3	3º Trimestre
2021_4	2021	4	4º Trimestre
2022_1	2022	1	1º Trimestre
2022_2	2022	2	2º Trimestre
2022_3	2022	3	3º Trimestre
2022_4	2022	4	4º Trimestre
2023_1	2023	1	1º Trimestre
2023_2	2023	2	2º Trimestre
2023_3	2023	3	3º Trimestre
2023_4	2023	4	4º Trimestre
2019_0	2019	Todos	Todos
2020_0	2020	Todos	Todos
2021_0	2021	Todos	Todos
2022_0	2022	Todos	Todos
2023_0	2023	Todos	Todos

Figura 10 - Excerto da Tabela DIM_Tempo

- DIM_Localização (Figura 11)
 - Obtida a partir da tabela NUTS 2024, para garantir um registo por Código_NUTS3;
 - Contém Código/Nome NUTS1, NUTS2 e NUTS3;
 - Utilizada como dimensão comum para FACT_Habitação e FACT_Salário.

A ^B _C Código_NUTS3	A ^B _C Nome_NUTS3	A ^B _C Código_NUTS2	A ^B _C Nome_NUTS2	A ^B _C Código_NUTS1	A ^B _C Nome_NUTS1
111	Alto Minho	11	Norte	1	Continente
11D	Douro	11	Norte	1	Continente
112	Cávado	11	Norte	1	Continente
11A	Área Metropolitana do Porto	11	Norte	1	Continente
119	Ave	11	Norte	1	Continente
11E	Terras de Trás-os-Montes	11	Norte	1	Continente
11C	Tâmega e Sousa	11	Norte	1	Continente
11B	Alto Tâmega e Barroso	11	Norte	1	Continente
192	Região de Coimbra	19	Centro	1	Continente
195	Beira Baixa	19	Centro	1	Continente
194	Viseu Dão Lafões	19	Centro	1	Continente
193	Região de Leiria	19	Centro	1	Continente
196	Beiras e Serra da Estrela	19	Centro	1	Continente
1D1	Oeste	1D	Oeste e Vale do Tejo	1	Continente
1D2	Médio Tejo	1D	Oeste e Vale do Tejo	1	Continente
1D3	Lezíria do Tejo	1D	Oeste e Vale do Tejo	1	Continente
1A0	Grande Lisboa	1A	Grande Lisboa	1	Continente
1C3	Alto Alentejo	1C	Alentejo	1	Continente
1B0	Península de Setúbal	1B	Península de Setúbal	1	Continente

Figura 11 - Excerto da Tabela DIM_Localização

- DIM_Categoria (Figura 12)
 - Extraída dos valores distintos da coluna ID_Categoria (H1, H11, H12);
 - Mapeia estes códigos para a descrição “Total”, “Novos” e “Existentes”.

A ^B _C ID_Categoria	A ^B _C Nome_Categoria
1	H1
2	H11
3	H12

Figura 12 - Tabela DIM_Categoria

6.4 Conclusão do Processo ETL

O processo ETL evoluiu de uma abordagem inicial baseada em *merges* pouco estruturados e tabelas não normalizadas para um método organizado, assente em tabelas auxiliares de mapeamento territorial e derivação correta de tabelas de facto e dimensões.

Após a limpeza e normalização das duas fontes, foi possível alinhar as granularidades — Habitação em *Ano* × *Trimestre* × *NUTS3* × *Categoria* e Salários em *Ano* × *NUTS3* — garantindo consistência na integração. A hierarquia NUTS foi aplicada de forma uniforme e a dimensão de tempo foi construída para suportar simultaneamente períodos trimestrais e anuais.

Com isto, FACT_Habitação, FACT_Salário, DIM_Tempo, DIM_Localização e DIM_Categoria ficaram prontas para alimentar o modelo multidimensional em constelação e permitir análises OLAP fráveis na secção seguinte.

7 Construção do Cubo e Operações OLAP

Após a conclusão do processo ETL e a criação do modelo multidimensional, foi possível avançar para a exploração analítica através da implementação de cubos OLAP no Power BI. O objetivo desta fase consistiu em transformar as tabelas de factos e dimensões — previamente normalizadas e integradas — em estruturas analíticas navegáveis, permitindo aplicar operações OLAP como **drill-down, roll-up, slice, dice e pivot**.

Foram construídas três matrizes OLAP distintas: uma para os valores de habitação, outra para os salários e uma matriz integrada que combina ambas as métricas para análise comparativa.

7.1 Matriz OLAP – Habitação

A primeira matriz OLAP foi construída a partir da FACT_Habitação, utilizando as dimensões Tempo, Localização (NUTS) e Categoria.

Na área das linhas foi colocada a combinação das hierarquias temporal e territorial (Ano - Trimestre / NUTS I - NUTS II - NUTS III), enquanto a dimensão Categoria foi usada nas colunas. A medida apresentada corresponde ao valor mediano €/m².

Esta matriz permite explorar o dataset através de drill-down temporal e territorial, além de comparar o valor da habitação entre categorias (Total, Novos, Existentes).

Ano		H1: Total	H11: Novos	H12: Existentes	Total
■ 2019		847.92	1056.46	809.35	904.58
■ 4.º Trimestre		847.92	1056.46	809.35	904.58
■ Continente		835.54	1046.50	797.54	893.19
■ Alentejo		679.00	879.25	655.00	737.75
Alentejo Central		732.00	953.00	700.00	795.00
Alentejo Litoral		1010.00	1135.00	998.00	1047.67
Alto Alentejo		438.00	680.00	415.00	511.00
Baixo Alentejo		536.00	749.00	507.00	597.33
■ Algarve		1696.00	1855.00	1654.00	1735.00
■ Centro		719.00	944.00	676.00	779.67
■ Grande Lisboa		1820.00	2417.00	1744.00	1993.67
■ Norte		766.88	940.13	729.25	812.08
■ Oeste e Vale do Tejo		743.33	898.33	711.00	784.22
■ Península de Setúbal		1142.00	1447.00	1100.00	1229.67
■ Região Autónoma da Madeira		1214.00	1411.00	1167.00	1264.00
■ Região Autónoma dos Açores		779.00	941.00	735.00	818.33
■ 2020		894.31	1120.88	848.63	954.61
■ 2021		944.78	1205.29	895.09	1015.05
■ 2022		1056.82	1333.82	1002.37	1131.00
■ 2023		1161.19	1455.98	1097.25	1238.14
■ 2024		1255.93	1588.78	1181.75	1342.15
■ 2025		1390.17	1702.52	1314.96	1469.22
Total		1081.76	1360.02	1023.46	1155.08

Figura 13 - Matriz OLAP dos valores medianos da habitação (€/m²) organizada por ano, trimestre e hierarquia territorial NUTS

7.2 Matriz OLAP – Salários Médios

A segunda matriz OLAP utiliza a FACT_Salário, estruturada com base na mesma hierarquia territorial, mas com granularidade temporal apenas anual.

As linhas incluem NUTS I, NUTS II e NUTS III e as colunas representam os anos. A medida utilizada é o Salário Médio (€).

Esta matriz suporta operações de drill-down/roll-up nos diferentes níveis territoriais e permite observar a evolução anual dos salários.

Nome_NUTS1	2019	2020	2021	2022	2023	Total
Continente	1.001,44	1.036,87	1.076,26	1.128,01	1.215,67	1.107,57
Alentejo	1.024,08	1.059,30	1.102,23	1.159,51	1.261,76	1.138,55
Alentejo Central	978,74	1.016,61	1.051,62	1.108,41	1.184,94	1.083,83
Alentejo Litoral	1.160,24	1.185,08	1.231,04	1.296,78	1.430,08	1.278,36
Alto Alentejo	935,42	966,67	1.007,94	1.058,16	1.152,97	1.039,90
Baixo Alentejo	1.021,92	1.068,85	1.118,32	1.174,68	1.279,04	1.152,09
Algarve	964,44	1.004,44	1.041,96	1.091,09	1.182,08	1.073,10
Algarve	964,44	1.004,44	1.041,96	1.091,09	1.182,08	1.073,10
Centro	964,52	1.001,00	1.043,99	1.097,72	1.180,13	1.073,88
Grande Lisboa	1.305,86	1.348,86	1.396,27	1.465,53	1.563,54	1.435,45
Norte	943,32	976,53	1.016,44	1.062,27	1.143,29	1.043,38
Oeste e Vale do Tejo	1.016,74	1.052,48	1.084,34	1.134,54	1.221,47	1.116,95
Península de Setúbal	1.247,07	1.282,76	1.302,39	1.359,12	1.456,31	1.344,08
Região Autónoma da Madeira	1.005,29	1.030,53	1.072,40	1.107,58	1.183,18	1.092,94
Região Autónoma dos Açores	1.034,36	1.077,26	1.084,00	1.155,86	1.257,18	1.137,15
Total	1.002,91	1.038,23	1.076,42	1.128,30	1.216,03	1.108,17

Figura 14 – Matriz OLAP dos salários médios anuais organizados por ano e hierarquia territorial NUTS

7.3 Matriz OLAP Integrada – Habitação + Salários

A terceira matriz integra as duas fact tables num único cubo analítico. Para garantir comparabilidade, ambos os datasets foram previamente uniformizados para a granularidade Ano x NUTS III.

Foram colocadas na área dos valores duas medidas:

- Média do Valor de Habitação (€/m²)
- Média de Salário (€)

Nas linhas permanece a hierarquia territorial; nas colunas, os anos. Esta matriz permite analisar simultaneamente os dois indicadores, possibilitando operações OLAP completas: comparação entre métricas, análise temporal, slice por categoria de habitação e drill-down por região.

Ano Nome_NUTS1	2019		2020	
	Média de Valor Habitação (Euros/m ²)	Média de Salário (euros)	Média de Valor Habitação (Euros/m ²)	Média de Salário (euros)
Continente	985,54	1.001,44	985,54	1.036,87
Alentejo	789,96	1.024,08	789,96	1.059,30
Alentejo Central	836,71	978,74	836,71	1.016,61
Alentejo Litoral	1259,18	1.160,24	1259,18	1.185,08
Alto Alentejo	474,18	935,42	474,18	966,67
Baixo Alentejo	589,76	1.021,92	589,76	1.068,85
Algarve	2079,94	964,44	2079,94	1.004,44
Centro	754,29	964,52	754,29	1.001,00
Grande Lisboa	2225,76	1.305,86	2225,76	1.348,86
Norte	902,78	943,32	902,78	976,53
Oeste e Vale do Tejo	903,00	1.016,74	903,00	1.052,48
Península de Setúbal	1499,29	1.247,07	1499,29	1.282,76
Região Autónoma da Madeira	1465,76	1.005,29	1465,76	1.030,53
Região Autónoma dos Açores	937,53	1.034,36	937,53	1.077,26
Total	1002,83	1.002,91	1002,83	1.038,23

Figura 15 - Matriz OLAP integrando valores da habitação (€/m²) e salário médio por região (NUTS1–NUTS3), para os anos 2019–2023

7.4 Resultados da Análise OLAP

A utilização destas três matrizes OLAP permitiu:

- observar a evolução temporal dos preços de habitação e dos salários;
- comparar regiões ao longo das hierarquias NUTS;
- distinguir comportamentos entre categorias de habitação;
- relacionar crescimento salarial com evolução dos preços da habitação;
- identificar desequilíbrios territoriais e tendências estruturais.

A matriz integrada demonstra de forma clara que, na maioria das regiões, o preço da habitação tem crescido significativamente mais rápido do que os salários, sugerindo uma perda gradual de poder de compra habitacional.

8 Análise Gráfica e Discussão dos Resultados

O culminar de todo o processo de construção do modelo multidimensional — desde o ETL à definição das dimensões, hierarquias e tabelas de factos — traduz-se na capacidade de gerar visualizações que respondem de forma direta às questões colocadas no início do trabalho. O objetivo desta fase é transformar o modelo num instrumento analítico, permitindo observar padrões, comparar regiões e perceber a evolução dos fenómenos estudados.

Para isso foram produzidos três conjuntos principais de gráficos:

- **Gráficos dedicados exclusivamente aos salários médios**, permitindo analisar a distribuição regional, desigualdades territoriais, diferenças entre anos e tendências ao longo do tempo.
- **Gráficos focados apenas nos valores da habitação ($\text{€}/\text{m}^2$)**, organizados por ano, trimestre, territorialidade e categoria (Novos, Existentes, Total), possibilitando compreender a dinâmica do mercado imobiliário português com detalhe.
- **Gráficos combinados que integram salários e habitação**, construídos a partir da granularidade comum Ano \times NUTS III, permitindo avaliar a relação entre rendimento médio e custo de habitação, uma das questões mais relevantes do estudo.

Estes gráficos materializam o propósito do cubo OLAP: **permitir exploração flexível**, aplicar operações como drill-down, roll-up, slice, dice e pivot, e transformar dados dispersos em conhecimento estruturado. A partir daqui é possível interpretar os fenómenos estudados de forma fundamentada, identificando tendências, padrões e desequilíbrios territoriais que motivaram este trabalho.

8.1 Análise dos Salários Médios

A análise dos salários médios por região NUTS III baseia-se numa série de gráficos de barras (Figura 16 a 20) filtrados por ano (2019–2023). A utilização desta visualização permite comparar rapidamente o posicionamento relativo das regiões, identificar padrões consistentes e perceber a evolução temporal das disparidades regionais. Em todos os anos analisados, observa-se uma grande estabilidade na ordenação das regiões, o que indica que as desigualdades territoriais no rendimento médio não são acidentais, mas sim estruturais.

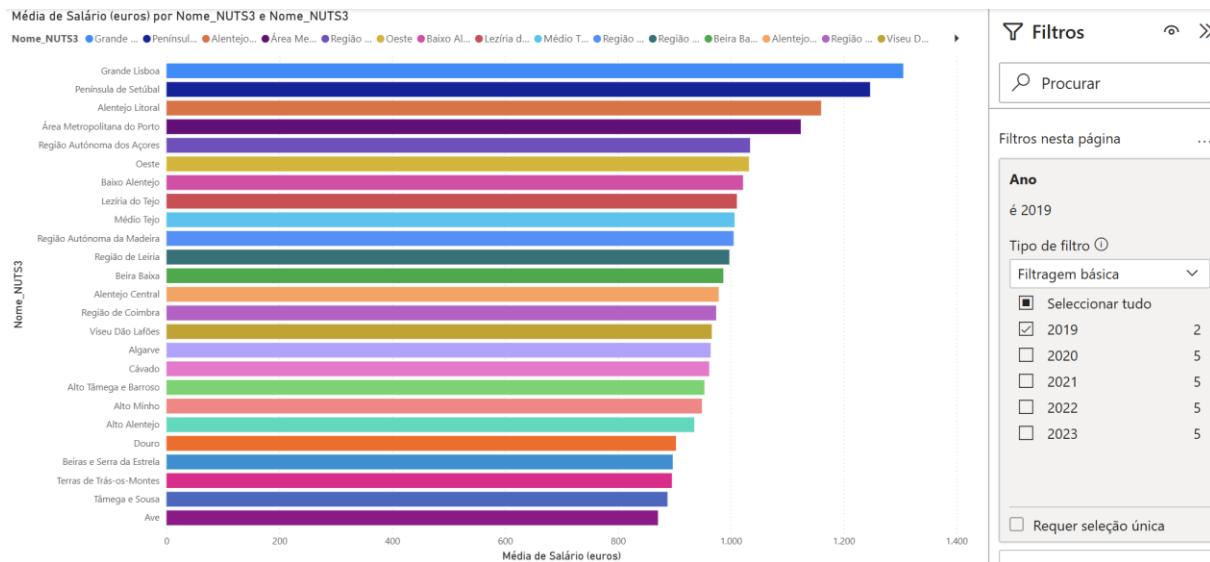


Figura 16 - Distribuição da Média de Salário (euros) por região NUTS III no ano de 2019

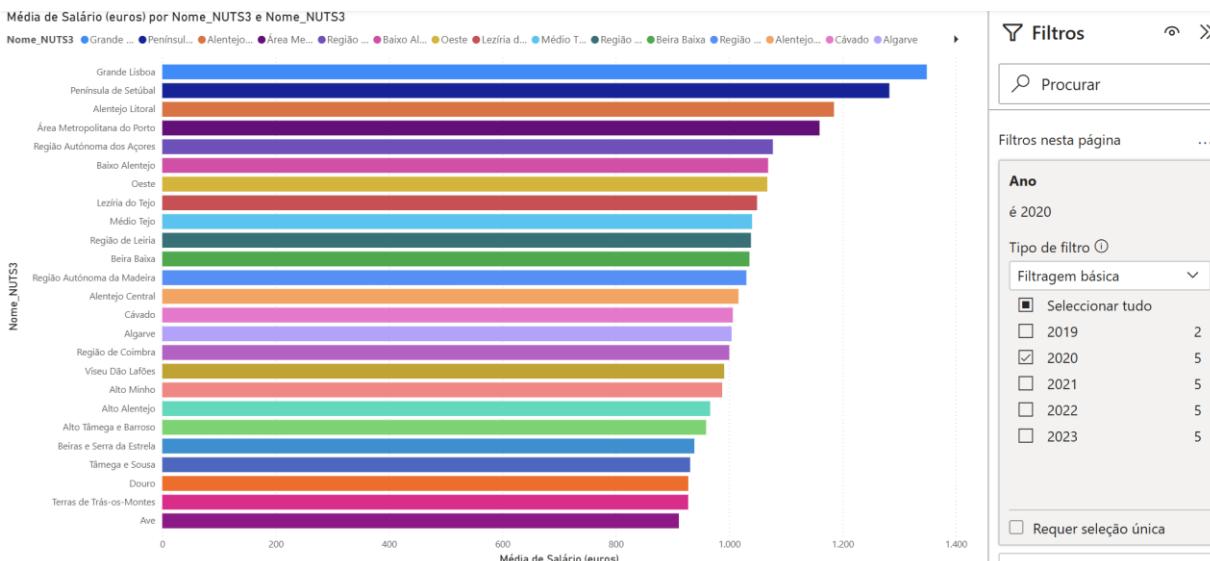


Figura 17 - Distribuição da Média de Salário (euros) por região NUTS III no ano de 2020

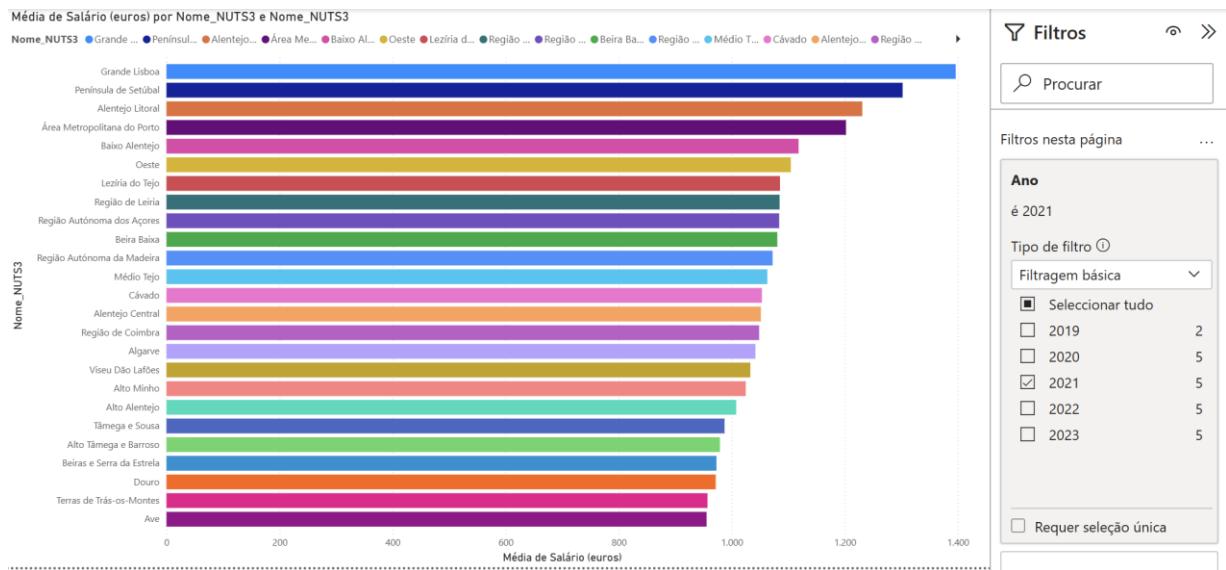


Figura 18 - Distribuição da Média de Salário (euros) por região NUTS III no ano de 2021

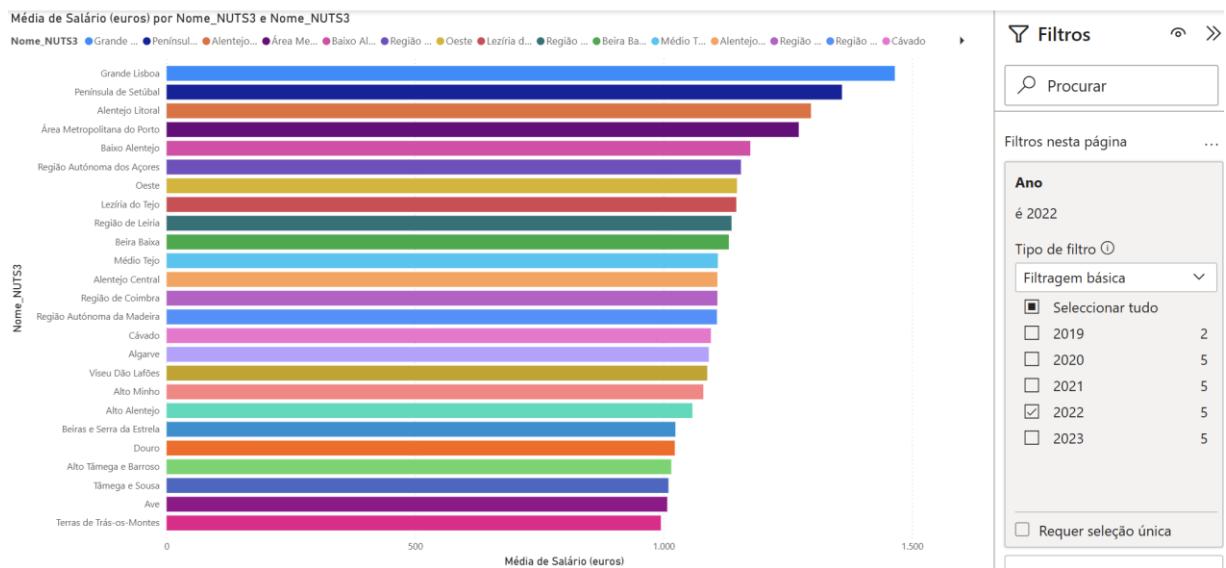


Figura 19 - Distribuição da Média de Salário (euros) por região NUTS III no ano de 2022

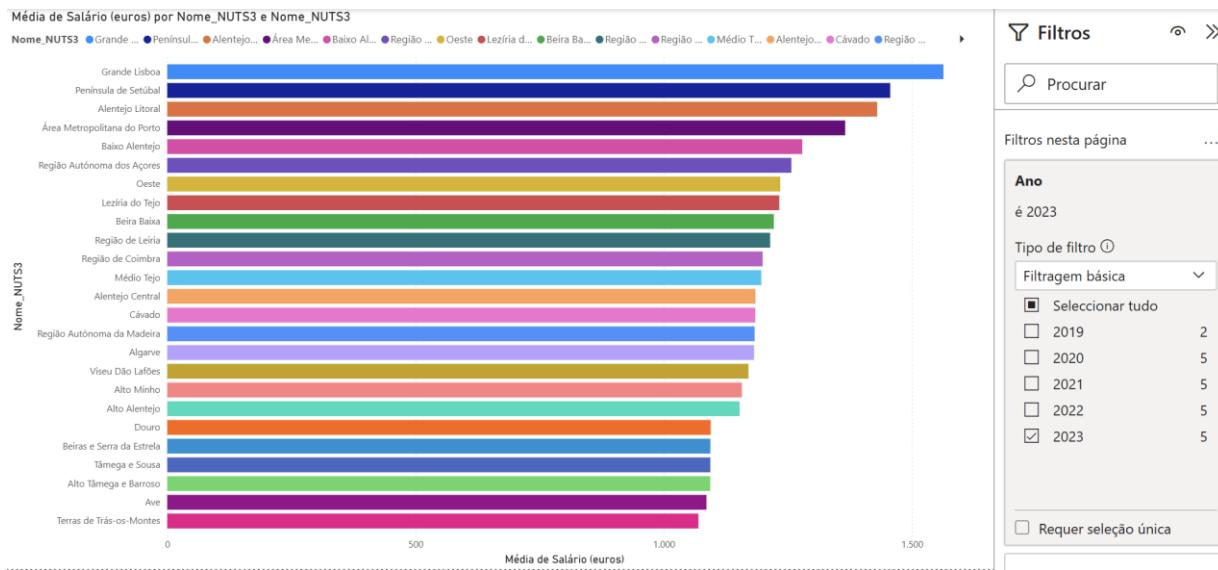


Figura 20 - Distribuição da Média de Salário (euros) por região NUTS III no ano de 2023

8.1.1 Distribuição Regional dos Salários – Tendências Consistentes (2019–2023)

Os gráficos mostram que, ao longo de todo o período analisado, existe um conjunto de regiões que se destaca sistematicamente com salários médios superiores. Em particular:

- Grande Lisboa apresenta, de forma consistente, os salários médios mais elevados do país.
- Península de Setúbal, Alentejo Litoral e Área Metropolitana do Porto surgem logo a seguir, sendo regiões com forte concentração de atividade económica e setores com maior valor acrescentado.

No extremo oposto:

- Tâmega e Sousa, Terras de Trás-os-Montes, Ave, Alto Tâmega e Barroso e Beiras e Serra da Estrela apresentam, em todos os anos, os salários médios mais baixos. Estas regiões caracterizam-se por menor densidade económica, setores industriais com menor remuneração média ou maior dependência de atividades primárias.

A consistência destas posições ao longo dos anos reforça a existência de um enviesamento territorial persistente no rendimento, com um fosso claro entre as áreas metropolitanas e o interior.

8.1.2 Evolução Temporal dos Salários

Apesar da estabilidade relativa entre regiões, os gráficos permitem observar um crescimento progressivo dos salários médios em praticamente todas as NUTS III, com ritmos diferentes:

- Entre 2019 e 2023, o aumento simultâneo em todas as regiões confirma uma tendência nacional de ajustamento salarial.
- As regiões que já eram mais remuneradas mantêm as mesmas posições, mas ampliam ligeiramente a distância relativa para algumas regiões do interior.
- As regiões do interior registam aumentos, mas em geral com valores absolutos menores, o que contribui para a manutenção do desfasamento.

Destaques Específicos por Ano - Pontos-chave observados nos gráficos:

- **2019:** diferenças mais pronunciadas entre Grande Lisboa e o restante país; muitas regiões abaixo dos 1.000€.
- **2020–2021:** ligeiro abrandamento no crescimento, possivelmente relacionado com o impacto económico da pandemia.
- **2022–2023:** crescimento mais forte, especialmente nas regiões metropolitanas; persistência das desigualdades territoriais.

Esta análise temporal evidencia que a dinâmica de evolução salarial não é homogénea e tende a favorecer regiões com maior atividade económica, reforçando a importância de cruzar esta informação com o custo habitacional nas secções seguintes.

8.2 Análise dos Valores da Habitação

A análise dos preços da habitação baseia-se nas três categorias disponibilizadas pelo INE: Total (H1), Novos (H11) e Existentes (H12).

8.2.1 Comparação Regional: Preços por NUTS III e por Categoria

O gráfico territorial, baseado na média dos valores ao longo dos anos disponíveis (Figura 22), evidencia diferenças claras entre regiões:

Grande Lisboa, Algarve, Península de Setúbal e Área Metropolitana do Porto destacam-se como as zonas com valores médios mais elevados, tanto para imóveis novos como existentes.

As regiões do interior - como Terras de Trás-os-Montes, Beiras e Serra da Estrela, Tâmega e Sousa e Alto Alentejo apresentam os valores mais baixos, mantendo um fosso significativo relativamente às áreas metropolitanas.

É visível também que, em praticamente todas as regiões, os preços dos Novos são substancialmente superiores aos dos Existentes, o que reforça a segmentação do mercado entre construção recente e parque habitacional mais antigo.

Além disso, a amplitude entre regiões metropolitanas e regiões do interior evidencia uma clara assimetria territorial na acessibilidade à habitação em Portugal.

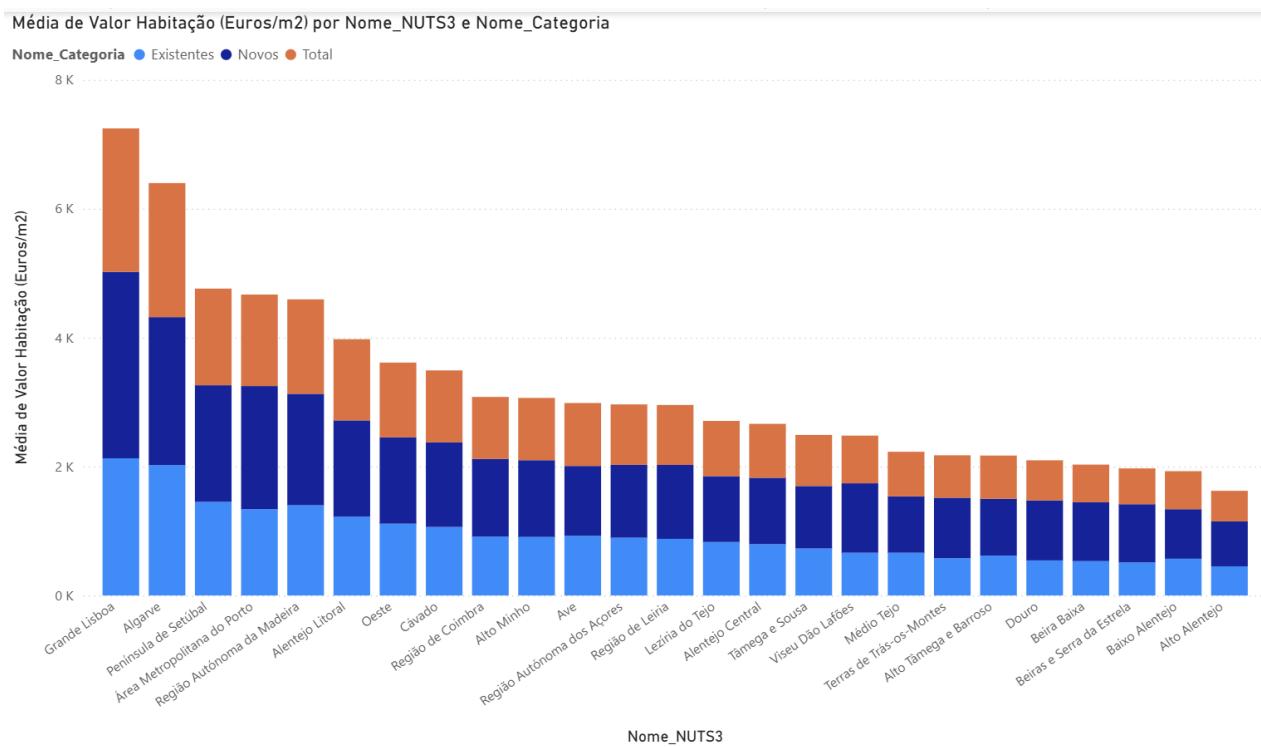


Figura 21 - Média do Valor da Habitação (€/m²) por região NUTS III e categoria de alojamento (Existentes, Novos e Total)

8.3 Análise Combinada Habitação vs Salário

A análise combinada entre os valores da habitação e os salários teve como objetivo principal compreender o desfasamento real entre o rendimento médio disponível e o custo médio de aquisição de habitação em Portugal. Até aqui, os dados tinham sido analisados separadamente (primeiro isolando os preços da habitação, depois os salários) mas só com a junção das duas métricas é possível avaliar de forma objetiva a acessibilidade habitacional nas diferentes regiões NUTS III.

8.3.1 Evolução do Salário e Habitação por Ano

Os gráficos seguintes representam, para cada região NUTS III, a evolução temporal dos valores médios da habitação e dos salários entre 2019 e 2023. A leitura comparada das curvas verde (salários) e vermelha (valor da habitação por m²) nos gráficos regionais permite identificar situações onde existe uma inversão clara das posições relativas entre rendimento e custo da habitação. Essa inversão é relevante pois indica se, numa dada região, o nível salarial supera largamente o custo da habitação ou se ocorre o oposto — um desfasamento acentuado em desfavor do poder de compra.

Regiões onde os salários superam claramente os preços da habitação

1. Baixo Alentejo

- A linha verde (salário) mantém-se consistentemente muito acima da linha vermelha (habitação).
- Isto evidencia uma região onde o poder de compra relativo é favorável e a habitação permanece proporcionalmente acessível.
- A diferença reduzida entre os valores reforça a ideia de um mercado menos pressionado.

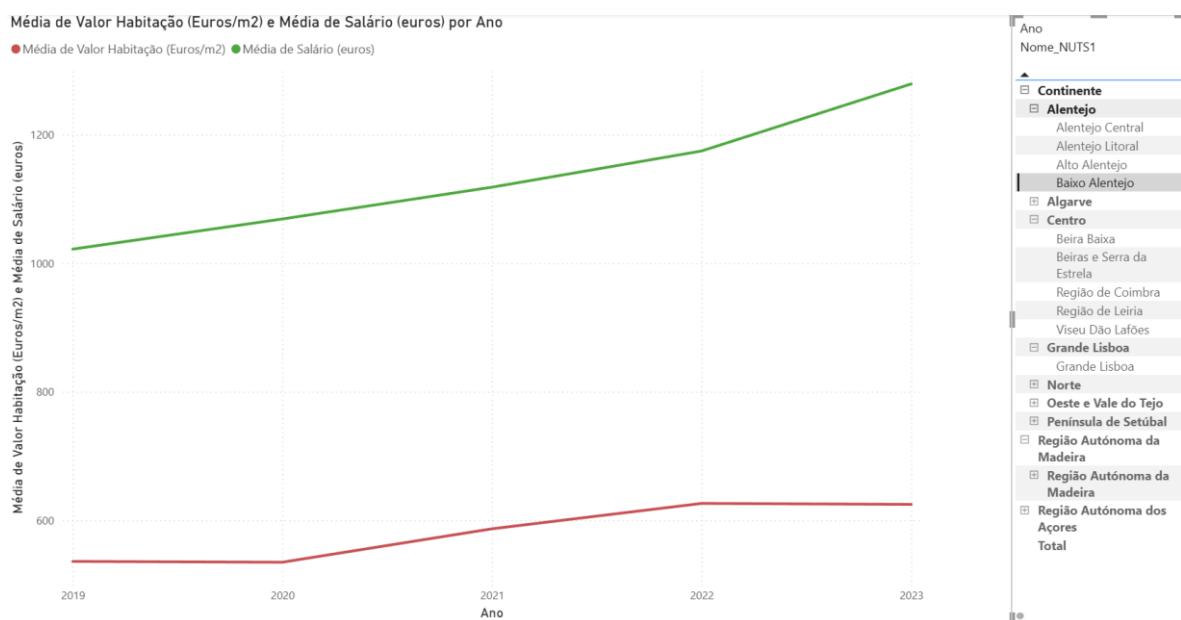


Figura 22 - Evolução da Média de Valor da Habitação (€/m²) e da Média de Salário (€/mês) por Ano no Baixo Alentejo

2. Viseu Dão Lafões

- A mesma tendência verifica-se: o salário médio está substancialmente acima do valor médio da habitação.
- A evolução temporal também é equilibrada, sem disparidades abruptas.
- Trata-se de uma região onde o mercado imobiliário se mantém alinhado com os rendimentos, garantindo maior acessibilidade.

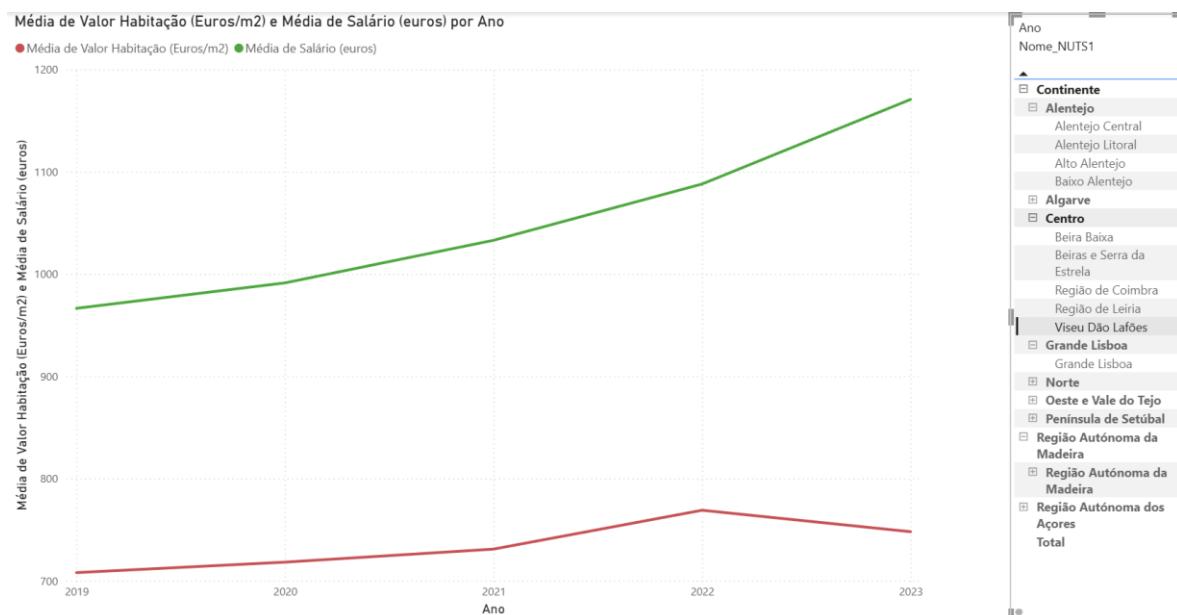


Figura 23 - Evolução da Média de Valor da Habitação (€/m²) e da Média de Salário (€/mês) por Ano em Viseu Dão Lafões

Regiões onde a habitação supera claramente os salários

3. Grande Lisboa

- Aqui a leitura é inversa: a linha vermelha (habitação) encontra-se muito acima da linha verde (salário).
- Esta diferença aumenta de forma marcada ao longo dos anos, especialmente após 2021.
- É um caso paradigmático de forte pressão imobiliária, com clara erosão da acessibilidade.

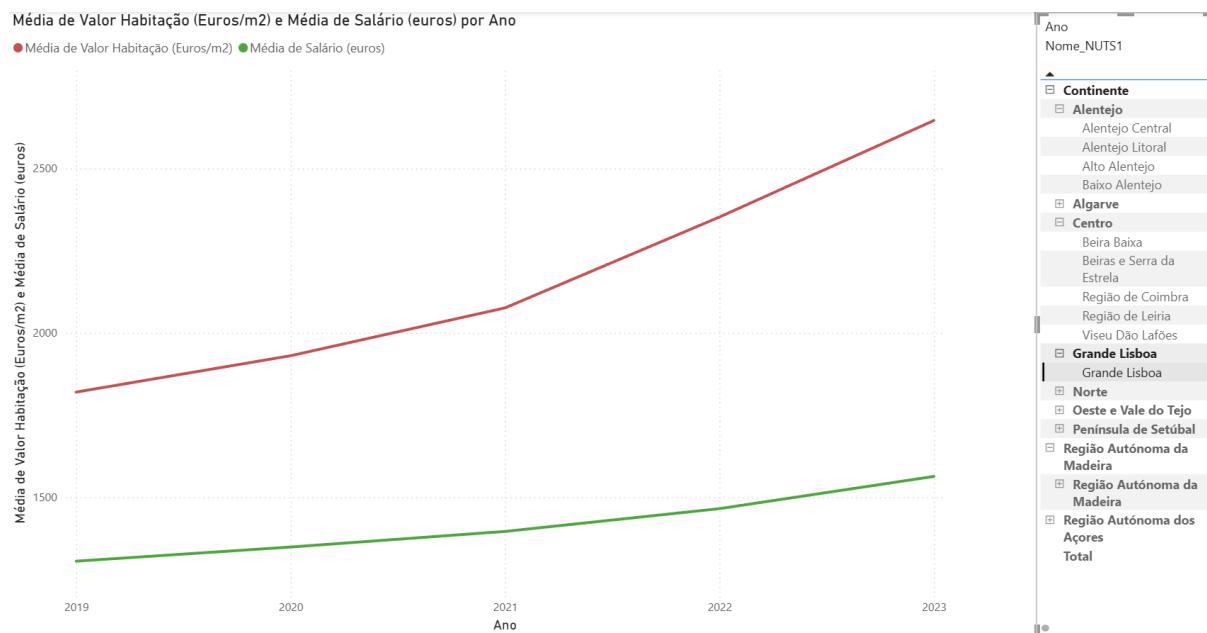


Figura 24 - Evolução da Média de Valor da Habitação (€/m²) e da Média de Salário (€/mês) por Ano na Grande Lisboa

4. Região Autónoma da Madeira

- O comportamento é semelhante ao de Lisboa: a linha vermelha cresce de forma acentuada e mantém uma distância evidente em relação aos salários.
- A discrepância entre rendimento e custo da habitação acentua-se ao longo do período.
- Confirma um mercado onde o custo por m² ultrapassa significativamente a capacidade económica média dos residentes.

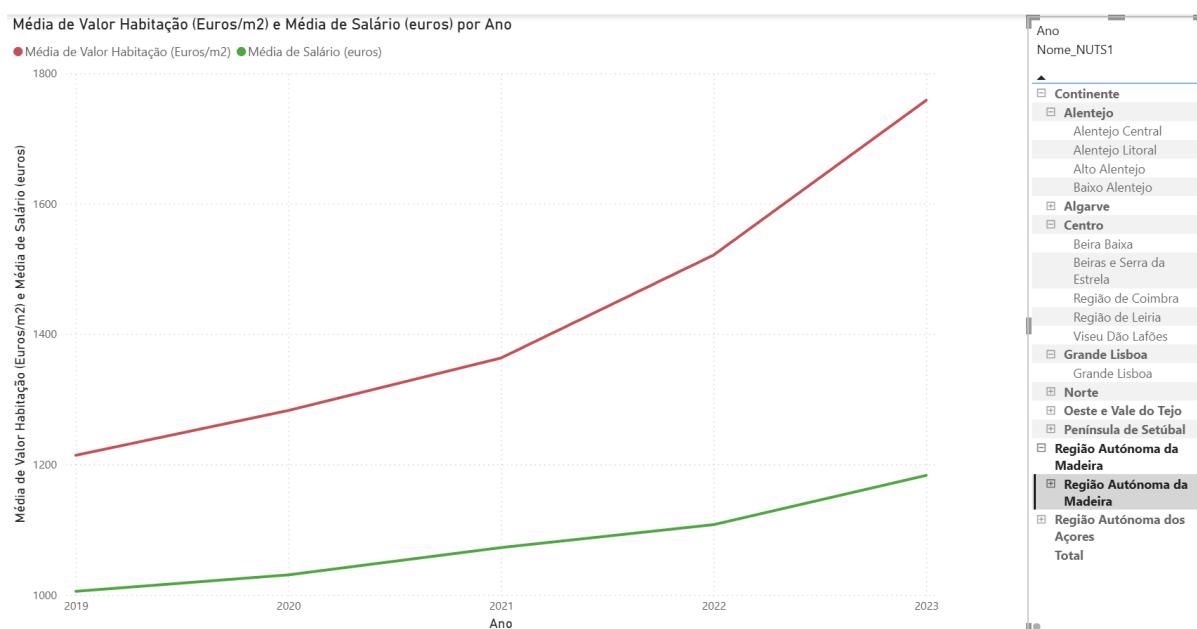


Figura 25 - Evolução da Média de Valor da Habitação (€/m²) e da Média de Salário (€/mês) por Ano na Região Autónoma da Madeira

8.3.2 Comparação Direta entre o Valor da Habitação (€/m²) e o Salário Médio por NUTS III

Para complementar a análise temporal realizada anteriormente, é essencial observar diretamente como o valor médio da habitação por metro quadrado se compara ao salário médio em cada região NUTS III. Esta visão lado a lado permite identificar rapidamente situações de desequilíbrio estrutural entre rendimento e custo de acesso à habitação, evidenciando territórios onde o mercado imobiliário está acima da capacidade económica dos residentes e regiões onde existe maior acessibilidade relativa.

Esta comparação absoluta é relevante porque torna visível o desequilíbrio entre custos e rendimentos, ou seja, mostra diretamente quem está mais próximo de poder comprar casa e quem está estruturalmente afastado desse objetivo.

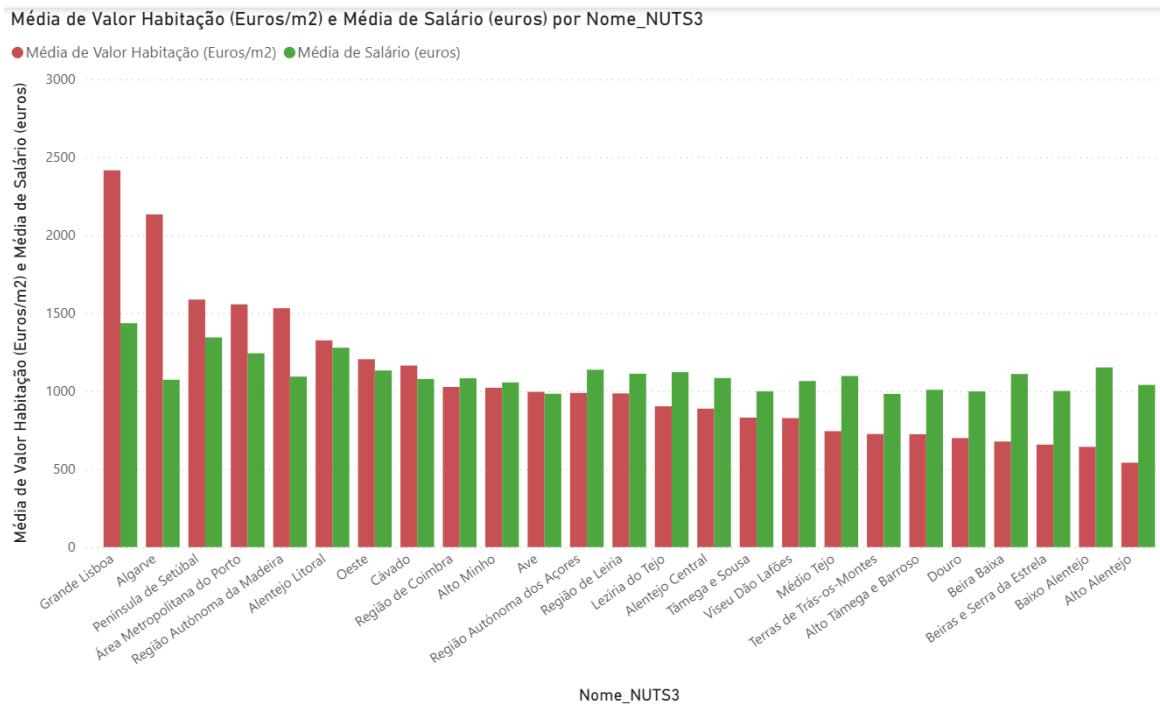


Figura 26 - Comparação da Média de Valor da Habitação (€/m²) e da Média de Salário (€/mês) por Região NUTS III (no período temporal de 2019 a 2023)

A leitura do gráfico evidencia contrastes muito marcados entre o litoral urbano/turístico e o interior. Regiões como Grande Lisboa, Algarve, Península de Setúbal, Área Metropolitana do Porto e Região Autónoma da Madeira apresentam valores de habitação muito superiores aos salários médios locais — em alguns casos, com diferenças que ultrapassam os mil euros por metro quadrado. São territórios onde o preço da habitação se afastou significativamente da realidade salarial, resultando em baixa acessibilidade habitacional.

Em contraste, regiões do interior como Beira Baixa, Alto Alentejo, Baixo Alentejo, Douro e Médio Tejo mostram uma relação inversa: os salários superaram os valores por m², refletindo mercados imobiliários menos pressionados e maior capacidade de aquisição relativa. Este comportamento confirma a divisão estrutural já observada noutras análises:

- Litoral e regiões insulares mais dinâmicas - preços da habitação muito elevados face ao rendimento.
- Interior continental - maior equilíbrio ou vantagem salarial na comparação direta.

Entre estes dois extremos surgem regiões intermédias - como Oeste, Cávado, Região de Coimbra ou Alto Minho - onde ambos os indicadores se aproximam, revelando mercados ainda relativamente sustentáveis, embora já com sinais de pressão moderada.

No conjunto, esta análise reforça que a acessibilidade habitacional em Portugal é profundamente segmentada, o problema não é homogéneo, mas concentrado em regiões metropolitanas, turísticas e com forte atração económica, enquanto o interior mantém níveis de acessibilidade muito superiores.

8.3.3 % Diferença Habitação vs Salário

Para uma análise mais profunda foi criada uma métrica % Diferença Habitação vs Salário (utilizando o menu Modelação - > Nova medida – Figura 28) ((Microsoft, s.d.-a) (Microsoft, s.d.-b) (Microsoft, s.d.-c)) que quantifica diretamente a diferença percentual entre o preço de compra de habitação por m² e o salário médio da mesma região. Esta abordagem permite medir a pressão económica que cada território enfrenta, identificar desigualdades territoriais e perceber até que ponto o crescimento do mercado imobiliário se tem afastado da evolução do rendimento médio. A análise que se segue traduz precisamente esses contrastes, destacando regiões onde a habitação se tornou proporcionalmente mais cara e outras onde permanece relativamente acessível.

Esta medida compara o preço de compra de habitação com o rendimento, permitindo identificar regiões onde a habitação é proporcionalmente mais inacessível (valores positivos elevados) ou onde os salários são, relativamente, mais equilibrados com o custo da habitação (valores próximos de zero ou negativos).

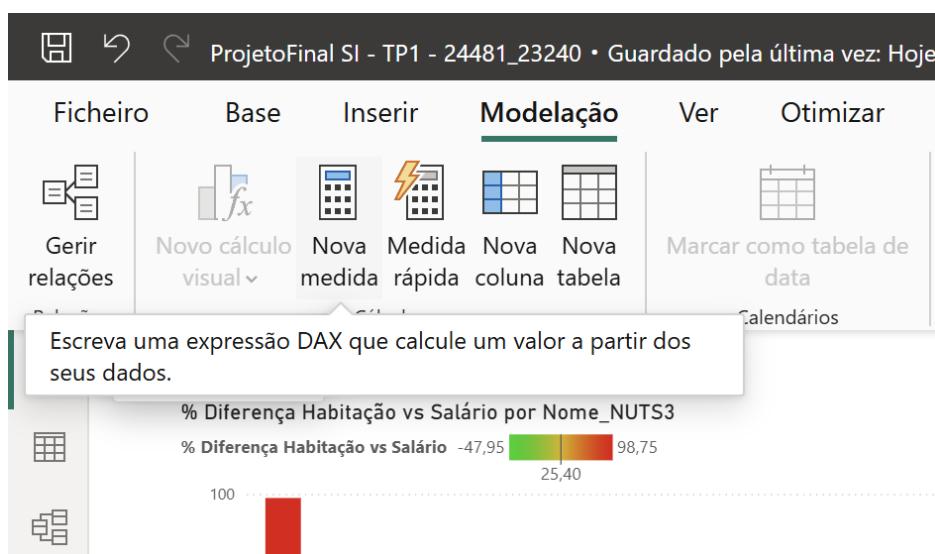


Figura 27 - Criação de nova medida % Diferença Habitação vs Salário

Como interpretar a fórmula:

$$\frac{\text{Média do Valor da Habitação por m}^2 - \text{Média do Salário}}{\text{Média do Salário}} \times 100$$

```
1 % Diferença Habitação vs Salário =
2 DIVIDE(
3     AVERAGE('FACT_Habitação'[Valor Habitação (Euros/m2)]) -
4     AVERAGE('FACT_Salário'[Salário (euros)]),
5     AVERAGE('FACT_Salário'[Salário (euros)])
6 ) * 100
```

- **Valor positivo** - Habitação mais cara do que o salário justificaria; maior pressão económica.
- **Valor negativo** - Salário médio superior ao custo da habitação; maior acessibilidade relativa.
- **Valor próximo de zero** - Tendência de equilíbrio.

Antes de analisar individualmente cada grupo de regiões, é importante compreender o que este gráfico representa. A métrica foi calculada para todas as regiões NUTS III, ordenada da maior para a menor diferença percentual, permitindo identificar imediatamente quais os territórios onde o custo habitacional está mais ou menos alinhado com o rendimento disponível.

O gráfico da Figura 23 evidencia, portanto, a distribuição territorial do desfasamento entre preços da habitação e salários, facilitando a interpretação das disparidades regionais. A partir desta visualização torna-se possível distinguir três grandes padrões, que são analisados de seguida.

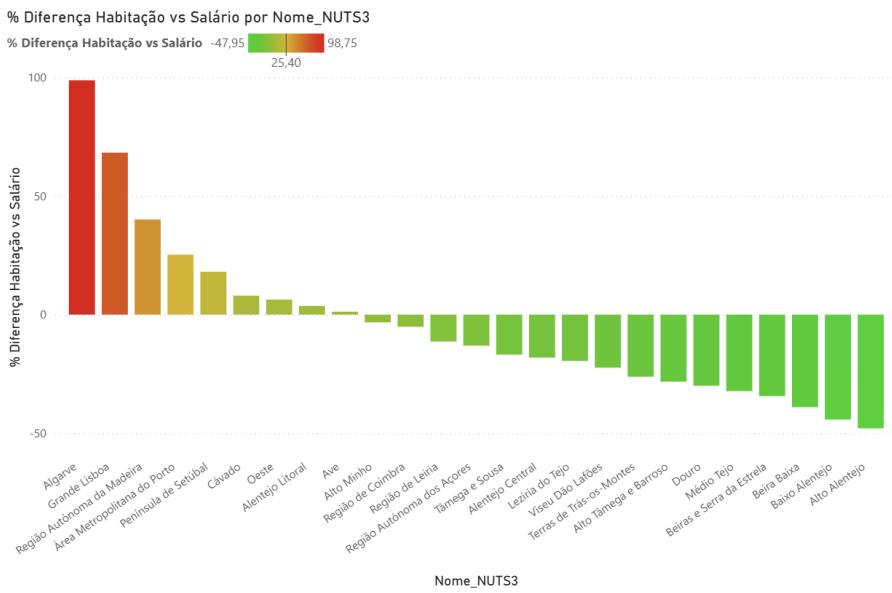


Figura 28 - % Diferença entre o Valor Médio da Habitação ($\text{€}/\text{m}^2$) e o Salário Médio (euros) por Região NUTS III

Principais conclusões observadas no gráfico:

1. Regiões onde a habitação pesa muito mais que o salário

No extremo superior, surgem regiões como:

- Algarve
- Grande Lisboa
- Região Autónoma da Madeira
- Área Metropolitana do Porto
- Península de Setúbal

Estas regiões apresentam diferenças percentuais positivas muito intensas - chegando a valores próximos dos 100%, significando que o custo por m^2 é praticamente o dobro do salário médio disponível. Isto confirma uma realidade económica já esperada: são zonas urbanas ou turísticas onde existe maior especulação imobiliária, maior pressão demográfica e maior procura externa.

Interpretação direta: nestas regiões, a aquisição de casa está muito desfasada do poder de compra médio dos residentes.

2. Regiões com valores moderados ou próximos do equilíbrio

Regiões como:

- Oeste
- Cávado

- Alentejo Litoral
- Região de Coimbra
- Alto Minho

apresentam percentagens entre pouco acima de zero até cerca de **20–30%**. Estas áreas mantêm algum desfasamento entre preço e salário, mas sem os extremos observados nas regiões metropolitanas ou turísticas.

Interpretação direta: nestas zonas o mercado imobiliário cresce, mas ainda se mantém relativamente ajustado ao rendimento da população residente.

3. Regiões onde a habitação pesa muito mais que o salário

No extremo inferior surgem:

- Beiras e Serra da Estrela
- Baixo Alentejo
- Alto Alentejo
- Beira Baixa
- Douro
- Médio Tejo

Aqui a % Diferença é negativa, chegando a valores próximos dos - 40% a - 50%. Significa que nestas regiões o salário médio é, relativamente, superior ao preço de compra de habitação.

Interpretação direta: São zonas do interior onde os preços da habitação continuam mais baixos e o mercado é menos pressionado, tornando a habitação proporcionalmente mais acessível.

No global o gráfico revela uma clara divisão litoral–interior:

- Litoral urbano e turístico - elevada pressão imobiliária, baixa acessibilidade.
- Interior e regiões menos densas - habitação mais acessível em termos relativos.

Além disso, mesmo com o crescimento simultâneo dos salários e dos valores da habitação, a velocidade de crescimento da habitação é consistentemente superior, refletindo-se nas percentagens altamente positivas nos grandes centros populacionais.

9 Demonstração Interativa das Operações OLAP e Gráficos Dinâmicos (Repositório com Vídeos)

Para complementar a análise desenvolvida ao longo do trabalho, foi preparada uma demonstração interativa (disponibilizada em formato de vídeo) onde são apresentadas as principais operações OLAP aplicadas ao modelo multidimensional construído, bem como exemplos de gráficos dinâmicos que permitem explorar os dados de forma visual e flexível.

Repositório com os vídeos demonstrativos:

https://drive.google.com/drive/folders/1Q6L8WRvn_dZDnFF2IB_0ZESUjpD_uD3y?usp=drive_link

9.1 Operações OLAP no cubo Habitação

O cubo da Habitação é o mais completo dos três, pois integra três dimensões distintas:

- 1. Hierarquia Temporal**
 - Ano – Trimestre
- 2. Hierarquia Territorial (NUTS)**
 - NUTS I - NUTS II - NUTS III
- 3. Categoria da Habitação**
 - Existente
 - Novo
 - Total

A presença simultânea destas três dimensões confere ao cubo todas as características de um cubo OLAP 3-dimensional. Contudo, como uma matriz só permite visualizar duas dimensões em simultâneo, foi necessário aninhá-las: a hierarquia territorial foi colocada dentro da hierarquia temporal, permitindo que a matriz representasse tempo e localização enquanto a dimensão Categoria permanecia nas colunas.

9.1.1 Roll-up e Drill-down

- **Drill-down:** permite expandir um nível da hierarquia para um nível mais detalhado.
Exemplo: Ano -> Trimestre -> NUTS I -> NUTS II -> NUTS III
- **Roll-up:** o movimento inverso, agregando níveis inferiores em níveis mais gerais.

9.1.2 Slice e Dice

- **Slice:** seleção de um corte específico de uma dimensão.
Ex.: escolher apenas o ano 2023 ou apenas a categoria “Novos”.
- **Dice:** aplicação simultânea de múltiplas condições.
Ex.: analisar apenas o 1.º trimestre de 2020 para as regiões do Norte e Centro.

No Power BI, estas operações são realizadas com filtros de página e selecções na matriz.

9.1.3 Pivot

No Power BI é possível trocar linhas por colunas, alterar a ordem das dimensões dentro da mesma área, por exemplo reorganizando a hierarquia temporal antes da territorial, e reestruturar a matriz para comparar categorias por território ou territórios por categoria. Esta rotação lógica do cubo permite observar relações sob perspectivas completamente distintas, reforçando a flexibilidade analítica do modelo multidimensional.

9.2 Operações OLAP no cubo Salário

O cubo do Salário é mais simples do que o da Habitação, pois assenta apenas em duas dimensões principais:

1. **Tempo - Ano**
2. **Hierarquia Territorial (NUTS) – NUTS 1/NUTS2/NUTS3**

Como apenas existem duas dimensões relevantes, a matriz permite representá-las diretamente: o tempo é colocada nas colunas e a hierarquia territorial nas linhas, facilitando a comparação da evolução dos salários no tempo e entre regiões.

9.2.1 Roll-up e Drill-down

- **Drill-down:** expande a hierarquia territorial para níveis mais detalhados.

Ex.: NUTS I → NUTS II → NUTS III

- **Roll-up:** o movimento inverso, agregando regiões detalhadas em níveis superiores.

9.2.2 Slice e Dice

- **Slice:** seleção de um corte específico de uma dimensão.

Ex.: escolher apenas o ano 2023.

- **Dice:** combinação de vários filtros.

Ex.: analisar apenas NUTS III selecionados num conjunto restrito de anos.

No Power BI, estas operações são realizadas sobretudo com filtros de página e selecções na matriz.

9.2.3 Pivot

No cubo Salário, o pivot consiste simplesmente em trocar as linhas pelas colunas, permitindo inverter a perspetiva da análise: por exemplo, visualizar regiões nas colunas e anos nas linhas. Esta rotação altera o ponto de vista da análise e facilita comparações alternativas entre tempo e território.

9.3 Operações OLAP no cubo Habitação + Salário

O cubo conjunto Habitação + Salário integra duas medidas distintas: Média do Valor Habitação ($\text{€}/\text{m}^2$) e Média de Salário (euros) analisadas em simultâneo sobre as mesmas dimensões. Este cubo combina:

1. **Tempo** - Ano
2. **Hierarquia Territorial (NUTS)** - NUTS I/ NUTS II /NUTS III

Nesta matriz, a dimensão temporal é colocada nas colunas, enquanto a hierarquia territorial organiza as linhas. As duas medidas são apresentadas lado a lado para cada combinação de ano e região, permitindo uma leitura direta dos contrastes.

9.3.1 Roll-up e Drill-down

- **Drill-down:** permite expandir um nível da hierarquia para um nível mais detalhado.
Exemplo: NUTS I -> NUTS II -> NUTS III
- **Roll-up:** o movimento inverso, agregando níveis inferiores em níveis mais gerais.

9.3.2 Slice e Dice

- **Slice:** isolamento de um subconjunto simples, como um ano específico ou apenas uma medida.
Ex.: escolher apenas o ano 2023.
- **Dice:** combinação de múltiplos filtros para uma análise mais segmentada.
Ex.: analisar apenas o 1.º trimestre de 2020 para as regiões do Norte e Centro.

9.3.3 Pivot

Neste cubo, o pivot corresponde a trocar linhas por colunas, invertendo a forma como tempo e território são apresentados.

9.4 Gráficos Dinâmico

Além dos vídeos das matrizes OLAP, foram criados dois vídeos para ilustrar a visualização dos gráficos dinâmicos Evolução do Salário e Habitação por Ano e Média de Salário por Localização.

9.4.1 Evolução do Salário e Habitação por Ano

O utilizador pode selecionar qualquer região NUTS III, como Baixo Alentejo, Viseu Dão Lafões, Grande Lisboa ou Região Autónoma da Madeira e observar automaticamente a evolução temporal do salário (linha verde) e a evolução temporal do valor da habitação (linha vermelha).

9.4.2 Média de Salário por Localização

O utilizador pode selecionar o ano (2023 a 2019).

10 Conclusão

O desenvolvimento deste trabalho permitiu construir um modelo multidimensional sólido, capaz de integrar duas fontes de informação distintas - preços da habitação (INE) e salários médios (PORDATA) - e de as analisar de forma conjunta através de operações OLAP. O processo completo, desde a extração e harmonização dos dados até à criação das tabelas de factos e dimensões, demonstrou a importância de normalizar granularidades e garantir coerência territorial para possibilitar análises comparativas fiáveis.

O objetivo inicial deste trabalho foi construir um modelo multidimensional que permitisse analisar, de forma integrada, a evolução dos valores da habitação e dos salários em Portugal, respondendo às questões levantadas no início do projeto:

Como evoluíram os salários médios nas diferentes regiões NUTS III ao longo do tempo?

Como evoluíram os valores de venda da habitação (€/m²) nas mesmas regiões?

Existem padrões territoriais comuns entre salários e valores de habitação?

Qual a relação entre o salário médio e o preço médio da habitação?

Quais as regiões onde a habitação é mais (ou menos) acessível em termos relativos?

Com a implementação do modelo multidimensional e as análises OLAP realizadas estas questões foram respondidas, com evidência visual e quantitativa obtida através dos gráficos desenvolvidos.

A análise temporal dos salários mostrou um crescimento moderado e relativamente homogéneo entre regiões, mantendo uma estrutura estável ao longo dos anos. Por outro lado, a evolução dos valores de habitação revelou um comportamento muito mais dinâmico e desigual, com aumentos bastante acentuados nas regiões metropolitanas e turísticas.

A análise combinada permitiu identificar padrões territoriais semelhantes aos observados nos mercados imobiliário e laboral portugueses:

Litoral urbano e turístico - salários relativamente elevados, mas preços da habitação muito superiores, gerando forte pressão e baixa acessibilidade.

Interior - salários mais baixos, mas com habitação proporcionalmente muito mais acessível.

A métrica criada (% Diferença Habitação vs Salário) permitiu quantificar esse desfasamento, confirmado que as maiores dificuldades de acessibilidade se concentram em Grande Lisboa, Algarve, Madeira, enquanto regiões como Baixo Alentejo, Alto Alentejo ou Beiras e Serra da Estrela apresentam valores negativos, indicando maior equilíbrio entre rendimento e custo da habitação.

Assim, o trabalho atingiu o seu objetivo: criou um modelo robusto, permitiu explorar os dados de forma multidimensional e respondeu de forma clara e fundamentada às perguntas inicialmente colocadas, revelando padrões estruturais relevantes e consistentes para análises futuras em políticas habitacionais ou desenvolvimento regional.

11 Bibliografia

Areal Editores. (s.d.). *A nova configuração de NUTS 2024 – Escola Virtual.*

https://app.escolavirtual.pt/fs/c/6/8/d/configuracao_nuts_2024.pdf

Instituto Nacional de Estatística. (s.d.). *Valor mediano das vendas de alojamentos familiares nos últimos 12 meses (Metodologia 2022 - €/m²) por Localização Geográfica (NUTS – 2024) e Categoria do alojamento familiar* [Base de dados].
https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&contecto=pi&indOcorrCod=0012234&selTab=tab0&xxlang=pt

PORDATA. (s.d.). *Ganho médio mensal: Qual é o salário bruto médio dos empregados em Portugal, com horas extra ou subsídios?*

<https://www.pordata.pt/pt/estatisticas/salarios-e-pensoes/salarios/ganho-medio-mensal>

Microsoft. (s.d.). *DIVIDE function (DAX).*

<https://learn.microsoft.com/pt-pt/dax/divide-function-dax>

Microsoft. (s.d.). *Create measures in Power BI Desktop.*

<https://learn.microsoft.com/en-us/power-bi/transform-model/desktop-measures>

Microsoft. (s.d.). *Calculated columns in Power BI Desktop.*

<https://learn.microsoft.com/en-us/power-bi/transform-model/desktop-calculated-columns>

Brito, I. S. (2025). *Sistemas de Informação e Negócios* [Slides de aula, documento PDF]. ESTIG/IPBeja.
(Documento fornecido no âmbito da UC.)

Brito, I. S. (2025). *Business Intelligence* [Slides de aula, documento PDF]. ESTIG/IPBeja.
(Documento fornecido no âmbito da UC.)

Brito, I. S. (2023). *ETL 2023* [Material de apoio, documento PDF]. ESTIG/IPBeja.
(Documento fornecido no âmbito da UC.)

Brito, I. S. (2024). *OLAP 2024* [Material de apoio, documento PDF]. ESTIG/IPBeja.
(Documento fornecido no âmbito da UC.)

Brito, I. S. (2024). *Data Warehouse 12024* [Material de apoio, documento PDF]. ESTIG/IPBeja.
(Documento fornecido no âmbito da UC.)