European Urban Affordability Index (EUAI)

Dublin City

 

Module: Data Analysis and Visualization

Author: Marco Ladeira

Part 0: Methodology

**Methodology**

The European Urban Affordability Index (EUAI) evaluates cost-of-living pressures for Dublin's young professionals (22-35) through six key steps:

1. Data Collection
   * Source: Numbeo's March 2025 Dublin dataset
   * Variables: 7 metrics across Housing (40%), Essentials (30%), Lifestyle (20%), and Income (10%)
2. Normalization
   * Converted all values to 0–1 scale higher = better affordability
   * Costs: Inverted min-max scaling (e.g., rent €2,038.71 → 0.7028)
   * Income: Scaled €3,000–4,000 range (€3,439.98 → 0.4400)
3. Weighting
   * Weights reflect Numbeo's expenditure patterns:
     + Housing: 40% (matches 38.4% actual expense share)
     + Essentials: 30% (groceries + utilities)
     + Lifestyle: 20% (dining + transport)
     + Income: 10%
4. Index Calculation

EUAI = sum(Normalized × Weight) # Final score: 0.6642

1. Validation
   * Internal: PCA (79.4% variance explained) and cluster analysis
   * External: Compared with Numbeo COL (0.2755) and CPI (0.4800)

Part 1: Theoretical Framework

# Introduction

The **European Urban Affordability Index(EUAI)** is a composite indicator designed to evaluate cost-of-living pressures for young professionals (aged 22-35), in this Index I am going over the data of city of Dublin, Ireland specifically. Unlike broad economic metrics, the EUAI focuses on **Essential Expenditures** (Housing, groceries, transportation etc) and weight that against local earning power, providing actionable insights for:

* **Recent Graduates** assessing Job Markets
* **Employers** determining relocation packages
* **Policymakers** identifying affordability crises

# Index Rational

Dublin ranks among Europe’s top 5 most expensive cities for rent (Numbeo, 2025), with housing consuming **38.4%** of average monthly expenses. Traditional indices like the Consumer Price Index (CPI) fail to:

* Weight housing costs proportionally
* Adjust for disposable income
* Target youth demographics

# Sub-Indices and Variables

The EUAI combines **4 sub-indices** derived exclusively from Numbeo’s Dublin dataset (March 2025):

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | **Sub-Index** | **Variables** | **Weight** | **Data Source** | **Operationalization** | | --- | --- | --- | --- | --- | | **Housing (40%)** | 1-bedroom rent (city center) | 20% | Numbeo "Rent Per Month" | Monthly cost in € | |  | Price per m² to buy apartment | 20% | Numbeo "Buy Apartment Price" | €/m² in city center | | **Essentials (30%)** | Groceries (single person) | 15% | Numbeo "Summary"\* | 60% of €1,087.50 (excl. rent) = €652 | |  | Basic utilities (85m² apartment) | 15% | Numbeo "Utilities" | Electricity/water/garbage in € | | **Lifestyle (20%)** | Meal at inexpensive restaurant | 10% | Numbeo "Restaurants" | Cost of 1 meal in € | |  | Monthly public transport pass | 10% | Numbeo "Transportation" | Regular price in € | | **Income (10%)** | Average monthly net salary | 10% | Numbeo "Salaries" | Post-tax income in € |    Weighting Justification **EUAI Theoretical Framework**  **Concept**: European Urban Affordability Index (EUAI)  **Primary Purpose**: Measure cost-of-living pressures for young professionals (22-35) in Dublin  **Dimensions**:   | **Component** | **Weight** | **Measurement Focus** | | --- | --- | --- | | Housing Stress | 40% | Rent and property purchase costs | | Essential Costs | 30% | Groceries, utilities | | Lifestyle Flexibility | 20% | Dining, transportation | | Income Relief | 10% | Net salary after tax |   **Scientific Rationale**:   * Dublin ranks top 5 most expensive European cities (Numbeo 2025) * Housing consumes 38.4% of average monthly expenses * Traditional indices (CPI) don't weight housing appropriately   Weights reflect **expenditure patterns** from Numbeo’s Dublin data:   * **Housing (40%)**: Largest expense (38.4% of total costs) * **Essentials (30%)**: Non-negotiable survival costs * **Lifestyle (20%)**: Discretionary spending capacity * **Income (10%)**: Contextualizes costs against earnings |
|  |
|  |
| Part 2 : Data Selection Data Source Overview All the data used is sourced from (Numbeo, 2025)(March 2025). The data is Crowdsourced(User-reported) and covers all of the required variables for living in Dublin city. The data is up to date with March of 2025, this is my following data displayed in Jupyter:   Data Quality Assessment **Strengths**   * **Completeness**: No missing values for selected variables * **Transparency**: Numbeo provides ranges (e.g., rent: €1,700-2,500) * **Timeliness**: March 2025 data reflects current conditions   **Limitations**   * **Self-reported bias**: User-submitted data may skew high/low * **No suburb-specific salary data**: Used city-wide average * **Static snapshot**: No historical trends  Selected Variables The following 7 variables were chosen based on the Theoretical Framework (Part 1):   | **Category** | **Variable** | **Value (Dublin)** | **Numbeo Section** | **Notes** | | --- | --- | --- | --- | --- | | **Housing** | 1-bedroom rent (city center) | €2,038.71 | "Rent Per Month" | Range: €1,700-2,500 | |  | Price per m² to buy (city center) | €6,813.00 | "Buy Apartment Price" | Range: €5,000-10,000 | | **Essentials** | Groceries (single person) | €652.50\* | "Summary" | \*60% of €1,087.50 | |  | Basic utilities (85m² apartment) | €247.08 | "Utilities" | Electricity/water/garbage | | **Lifestyle** | Meal, inexpensive restaurant | €20.00 | "Restaurants" | Range: €15-30 | |  | Monthly transport pass | €115.00 | "Transportation" | Range: €70-173.91 | | **Income** | Average net monthly salary | €3,439.98 | "Salaries" | After-tax |   Single person monthly costs (excl. rent) = €1,087.50 × 60% (assumed groceries share) = €652.50  Part 3 : Imputation of Missing Data **Introduction to Imputation** Imputation is the process of replacing missing data with estimated values to ensure a complete dataset. Since our Dublin dataset from Numbeo (March 2025) is fully populated, we will:  **Confirm no missing values exist** in the selected variables.  **Outline a robust imputation strategy** for hypothetical missing data to demonstrate methodological rigor.  **Data Completeness Verification**  All **7 key variables** from Part 2 are complete. Below is the verification:   | **Variable** | **Value (Dublin)** | **Source** | **Missing?** | **Data Range (if applicable)** | | --- | --- | --- | --- | --- | | 1-bedroom rent (city center) | €2,038.71 | Rent Per Month | No | €1,700–2,500 | | Price per m² (city center) | €6,813.00 | Buy Apartment Price | No | €5,000–10,000 | | Groceries (single person) | €652.50\* | Summary | No\* | N/A (proxy) | | Basic utilities (85m²) | €247.08 | Utilities | No | €150–450 | | Meal, inexpensive restaurant | €20.00 | Restaurants | No | €15–30 | | Monthly transport pass | €115.00 | Transportation | No | €70–173.91 | | Avg. net monthly salary | €3,439.98 | Salaries | No | N/A (single value) | |
| The output on my jupyter notebook is as follows |
|  |
|  |
|  |

# Hypothetical Imputation Strategy

If future updates introduce missing data, we would apply the following methods:

Method 1: Median Imputation

Method 2: Mean Imputation(if no range)

Method 3: Geographic Adjustment

Method 4:Future Time Adjustment

Part 4 : Multivariate Analysis

The multivariate analysis section explores the relationships between variables in the European Urban Affordability Index (EUAI) through three key techniques: correlation analysis, principal component analysis (PCA), and cluster analysis. These methods validate the theoretical framework and reveal hidden patterns in Dublin's affordability landscape for young professionals.

# **Correlation Analysis**

The correlation matrix reveals several important relationships:

* **Strong positive correlation (0.82)** between normalized values and weighted scores, confirming our weighting scheme effectively amplifies important variables
* **Moderate negative correlation (-0.65)** between weights and z-values, indicating higher-weighted variables tend to have more extreme original values
* **Weak correlation (0.12)** between weights and normalized values, suggesting proper normalization before weighting

A screenshot of a graph

AI-generated content may be incorrect.

The strong correlation between normalized values and final scores validates our transformation pipeline, while the weight-zvalue relationship confirms we've appropriately weighted the most impactful cost drivers.

# **Principal Component Analysis**

A graph with blue squares

AI-generated content may be incorrect.

The PCA yielded two significant components:

1. **PC1 (79.4% variance)**: Dominated by normalized values and weights
2. **PC2 (20.6% variance)**: Primarily reflects z-scores of original values

Scree Plot Interpretation

The scree plot shows PC1 captures most variance, suggesting our index could potentially be simplified without major information loss. However, maintaining all components preserves nuanced differences between housing and lifestyle factors.

# **Biplot Insights**

The biplot visualization reveals:

* Housing variables cluster in high-weight, high-normalized space
* Income appears as an outlier, confirming its unique role in affordability
* Lifestyle and essentials form an intermediate group

A graph with points and arrows

AI-generated content may be incorrect.

# **Cluster Analysis**

The K-means clustering analysis revealed three distinct clusters with the following characteristics:

Cluster Profiles

| **Cluster** | **Avg. Normalized Value** | **Std Dev** | **Avg. Weight** | **Std Dev** | **Dominant Category** | **Avg. Weighted Score** |
| --- | --- | --- | --- | --- | --- | --- |
| 2 | 0.703 | - | 0.200 | - | Housing | 0.141 |
| 0 | 0.965 | 0.041 | 0.125 | 0.029 | Essentials | 0.120 |
| 1 | 0.220 | 0.311 | 0.150 | 0.071 | Housing |  |

A graph of different sizes and colors

AI-generated content may be incorrect.

Cluster Analysis Findings

Three distinct clusters emerged from the analysis, each revealing important patterns in Dublin's affordability landscape:

**1. Primary Cost Drivers (Cluster 2)**  
This cluster contains the most influential variables, predominantly housing-related metrics. Key characteristics:

* Normalized values average 0.70, indicating significant cost pressure
* Variables carry substantial 20% weights in the index calculation
* Generates the highest weighted score contribution (0.141)

These results empirically validate housing costs as the paramount affordability challenge for young professionals in Dublin's urban core. The data confirms what policy analysts have observed anecdotally - that accommodation expenses disproportionately impact this demographic.

**2. Core Living Expenses (Cluster 0)**  
Representing essential, non-discretionary costs, this cluster shows:

* Exceptionally high normalized values (0.96), reflecting universal expense burdens
* Lower relative weights (12.5%) due to less variability between individuals
* Moderate score impact (0.120), consistent with baseline living costs

The tight clustering of these variables (evidenced by low standard deviations) suggests remarkable consistency in essential expenditures across Dublin's young professional population.

**3. Secondary Factors (Cluster 1)**  
This cluster presents an interesting counterpoint:

* Surprisingly low normalized values (0.22) for housing-related variables
* Moderate weighting (15%) fails to compensate for low base values
* Minimal score contribution (0.022) indicates limited practical impact

The presence of housing variables in both primary and secondary clusters suggests important nuances in accommodation costs that merit further investigation.

# **Key Findings from Multivariate Analysis**

- 1. Housing variables show the highest weighted impact on the EUAI score

- 2. PCA reveals Normalized values and Weights explain most of the variance

- 3. Cluster analysis identifies distinct affordability profiles:

- - High impact housing variables

- - Balanced essentials/lifestyle

- - Income-related variables

- 4. The analysis confirms the weighting scheme effectively captures

- different dimensions of urban affordability

Part 5 : Data Normalization Methodology

**Introduction**

To ensure comparability across diverse cost variables, normalization was applied to transform raw monetary values into a standardized 0–1 scale. This process adjusts for differences in magnitude between variables (e.g., rent vs. grocery costs) while preserving their relative impact on affordability.

How I Normalized the Data

1. For Cost Variables (rent, groceries, utilities, meals, transport):
   * We identified the lowest cost (€20 for a meal) and highest cost (€6,813 for property per m²)
   * Each cost was converted using:  
     [1 minus (value minus lowest cost) divided by (highest cost minus lowest cost)]
   * Example for rent:  
     (1 - (2038.71 - 20)/(6813 - 20)) = 0.7028
   * This means higher costs get lower scores (worse affordability)
2. For Income (salary):
   * We used a range of €3,000-€4,000 based on typical salaries
   * The calculation was:  
     (salary minus 3000) divided by (4000 - 3000)
   * Example for €3,440 salary:  
     (3439.98 - 3000)/1000 = 0.4400
   * Higher salaries get higher scores (better affordability)

Coding this on Jupyter Resulted in this final Normalization Results Table

|  | **Variable** | **Value** | **Normalized** |
| --- | --- | --- | --- |
| 0 | 1-bedroom rent (city center) | 2038.71 | 0.7028 |
| 1 | Price per m² to buy (city center) | 6813.00 | 0.0000 |
| 2 | Groceries (single person) | 652.50 | 0.9069 |
| 3 | Basic utilities (85m²) | 247.08 | 0.9666 |
| 4 | Meal, inexpensive restaurant | 20.00 | 1.0000 |
| 5 | Monthly transport pass | 115.00 | 0.9860 |
| 6 | Average net monthly salary | 3439.98 | 0.4400 |

Part 6: Weighting and Index Aggregation

The final European Urban Affordability Index (EUAI) score was calculated through a three-step process:

1. **Variable Weighting**:
   * Each normalized variable was multiplied by its predetermined weight
   * Formula: Weighted Score = Normalized Value × Weight
   * Example: Rent (Normalized: 0.7028 × Weight: 0.20) = 0.1406
2. **Sub-Index Calculation**:
   * Weighted scores were summed within each category:
     + Housing Stress (40%)
     + Essential Costs (30%)
     + Lifestyle Flexibility (20%)
     + Income Relief (10%)
3. **Final Index Computation**:
   * All sub-indices were aggregated to produce the comprehensive EUAI score

Results

**Final EUAI Score**

A screenshot of a computer

AI-generated content may be incorrect.

Part 6: Comparative to other Indices

**Methodology**

I validated the EUAI by comparing it with two widely used economic indicators:

1. **Numbeo Cost of Living Index** (Dublin score: 72.45) (Numbeo, 2025)
   * Lower Numbeo scores indicate better affordability
   * Converted to our 0-1 scale by subtracting from 100 and dividing by 100
   * Calculation: (100 - 72.45)/100 = 0.2755
2. **Ireland Consumer Price Index** (Annual inflation: 5.2%) (CPI, 2025)
   * Higher inflation indicates worsening affordability
   * Converted to 0-1 scale using 10% inflation as worst-case scenario
   * Calculation: (10 - 5.2)/10 = 0.4800

All scores were normalized so higher values always indicate better affordability.

Key Findings

| **Index** | **Normalized Score** | **Original Value** | **Interpretation** |
| --- | --- | --- | --- |
| EUAI | 0.6642 | 0.6642 | Our base measurement |
| Numbeo COL | 0.2755 | 72.45 | Confirms cost pressures |
| CPI Inflation | 0.4800 | 5.2% annual increase | Shows rising living costs |

For visual reference from my Jupyter notebook

A graph of a bar chart

AI-generated content may be incorrect.

**Advantages of EUAI**

* More targeted than Numbeo for young professionals
* Incorporates income data missing from CPI
* Provides actionable breakdowns by expense category

**Limitations**

* Currently limited to Dublin (unlike global indices)
* Snapshot view without historical trends
* Relies on crowdsourced data for some inputs

Part 7: Visualisation of Results

A close-up of a graph

AI-generated content may be incorrect.

# **Composite Index Breakdown**

**Pie Chart of EUAI Composition by CategoryA pie chart with numbers and text

AI-generated content may be incorrect.**

This pie chart illustrates how each weighted category contributes to the total score, as you can see Housing dominates with a 42.1%, followed by Essentials at 31.3%.

# **Variable-Level Analysis**

**Horizontal Bar Chart / Weighted Contributions of ALL variables**

A graph of a bar chart

AI-generated content may be incorrect.

Ranks variables by their normalized score × weight

# **EUAI vs Standard Indices**

A graph of different sizes and colors

AI-generated content may be incorrect.

The **EUAI (0.68)** scores significantly higher than Numbeo's Cost of Living index (0.28), reflecting:

* + My index's adjustment for local salaries
  + Exclusion of luxury goods from calculations
  + Demographic targeting

**CPI Inflation (0.48)** and **Eurostat Housing (0.42)** (EuroStat, 2025) fall midway, showing:

* + General inflation underrepresents housing-specific pressures
  + Standard indices miss critical disposable income effects

# **Weighing Sensitivity Analysis**

A graph of a bar chart

AI-generated content may be incorrect.

**Scenario Comparison**:

| **Scenario** | **EUAI Score** | **Change vs. Original** |
| --- | --- | --- |
| Original | 0.6642 | Baseline |
| Equal Weights | 0.5921 | ▼ 10.9% |
| Housing Focus | 0.7243 | ▲ 9.0% |
| Income Focus | 0.6105 | ▼ 8.1% |

**Critical Takeaways**:

1. **Housing Weight Matters** : Increasing housing weight boosts scores, confirming its outsized impact
2. **Income's Limited Leverage**: Even focused weighting can't compensate for low base salaries
3. **Validation** : Original weights perform optimally between extremes

Part 8: Extras(Uncertainty and Sensitivity Analysis)

# **Uncertainty and Sensitivity Analysis**

A blue and green rectangular shapes

AI-generated content may be incorrect.

A graph of blue bars

AI-generated content may be incorrect.

**Weight Sensitivity**

Tested four weighting scenarios to validate the robustness of the EUAI:

| **Scenario** | **EUAI Score** | **Change vs. Original** | **Key Adjustment** |
| --- | --- | --- | --- |
| Original | 0.6642 | Baseline | Housing (40%), Income (10%) |
| Equal Weights | 0.5921 | ▼ 10.9% | All variables weighted equally |
| Housing Focus | 0.7243 | ▲ 9.0% | Housing weight increased to 70% |
| Income Focus | 0.6105 | ▼ 8.1% | Income weight doubled to 20% |

The original weighting scheme provides optimal balance between housing dominance and income effects.

**Normalization Robustness**

A green and blue rectangular bars

AI-generated content may be incorrect.

Compared normalization methods:

| **Method** | **EUAI Score** | **Note** |
| --- | --- | --- |
| Min-Max | 0.6642 | Original approach (0–1 scale) |
| Z-score | 0.7151 | Alternative statistical normalization |

**Finding**: Z-score normalization moderately inflates scores but preserves rank order.

**Outlier Impact**

* Original EUAI: **0.6642**
* Without outliers (|z-score| > 2): **0.6642**  
  **Conclusion**: Outliers have negligible impact on results.

A key takeaway in this analysis is that Min and Max and Z-score normalisation yield directionally consistent results

**A screenshot of a computer

AI-generated content may be incorrect.**

Part 9: Results

==================================================

FINAL EUAI SCORE

0.6642

==================================================

SUBCATEGORY CONTRIBUTIONS:

Category

Essentials 0.2810

Housing 0.1406

Income 0.0440

Lifestyle 0.1986

DETAILED VARIABLE SCORES:

Variable Category Normalized Weight Weighted\_Score

Basic utilities (85m²) Essentials 0.9666 0.1500 0.1450

1-bedroom rent (city center) Housing 0.7028 0.2000 0.1406

Groceries (single person) Essentials 0.9069 0.1500 0.1360

Meal, inexpensive restaurant Lifestyle 1.0000 0.1000 0.1000

Monthly transport pass Lifestyle 0.9860 0.1000 0.0986

Average net monthly salary Income 0.4400 0.1000 0.0440

Price per m² to buy (city center) Housing 0.0000 0.2000 0.0000

**Final EUAI Score (Dublin): 0.6642**  
This places Dublin in a mid-high-cost affordability for young professionals, the wage is quite high however soaring housing costs depress overall affordability. When housing weight increases to 70%, the score jumps to 0.7243. It’s clear that housing is the main affordability barrier.

I would also like to mention that I have done other visualization work on my Jupiter Page that I have not implemented into this documentation as it does not suit however you can view this code on GitHub.

**Code Summary**

This analysis was carried out in Python using a Jupyter Notebook. All data was manually gathered from Numbeo (as of March 2025) and entered into the notebook as Python dictionaries or arrays.

**Data Input**  
Key variables—including rent, groceries, utilities, and average salary—were manually input using values directly taken from Numbeo.

**Normalization Methods**  
To make the monetary values comparable, both min-max scaling and z-score normalization were applied. Cost-related metrics were inverted so that lower costs translated into higher affordability scores.

**Weighting & Aggregation**  
Each variable was assigned a predefined weight (e.g., housing at 40%) to reflect its relative impact on the budgets of young professionals. These weights were used to calculate weighted scores through element-wise multiplication.

**Composite Score Calculation**  
The final affordability score, referred to as the EUAI, was calculated by summing the weighted values. This resulted in a single index value for Dublin.

**Analysis Techniques**  
To better understand the relationships between variables and to assess the robustness of the weighting system, the notebook included correlation analysis, Principal Component Analysis (PCA), and K-Means clustering.

**Visualizations**  
Visual outputs were created using matplotlib and seaborn to show how the index was constructed, how each variable contributed, and to illustrate the resulting clusters.

References/Citations (Harvard)

CPI, 2025. *CPI IRELAND.* [Online]   
Available at: https://www.cso.ie/en/statistics/prices/consumerpriceindex/  
[Accessed 6 5 2025].

EuroStat, 2025. *EuroStat.* [Online]   
Available at: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Housing\_price\_statistics\_-\_house\_price\_index  
[Accessed 6 5 2025].

Numbeo, 2025. *Numbeo.* [Online]   
Available at: https://www.numbeo.com/cost-of-living/in/Dublin  
[Accessed 3 4 2025].

Version Control

All code is on my Jupyter notebook that I have pushed to git, to access the git repository follow this link: https://github.com/fristtysteam/EUAI-Index-DAV