

## Introduction

Understanding property price dynamics in the housing market is critical for estate agency companies. This study utilizes a dataset developed by University College London (2021), which integrates transaction records from the Land Registry Price Paid Data (PPD) with energy performance and property characteristics from Domestic Energy Performance Certificates (EPCs). By examining these combined datasets, we aim to understand how specific property features influence prices among a set of 10546 houses in **Middlesbrough** using statistical modelling.

## Methods

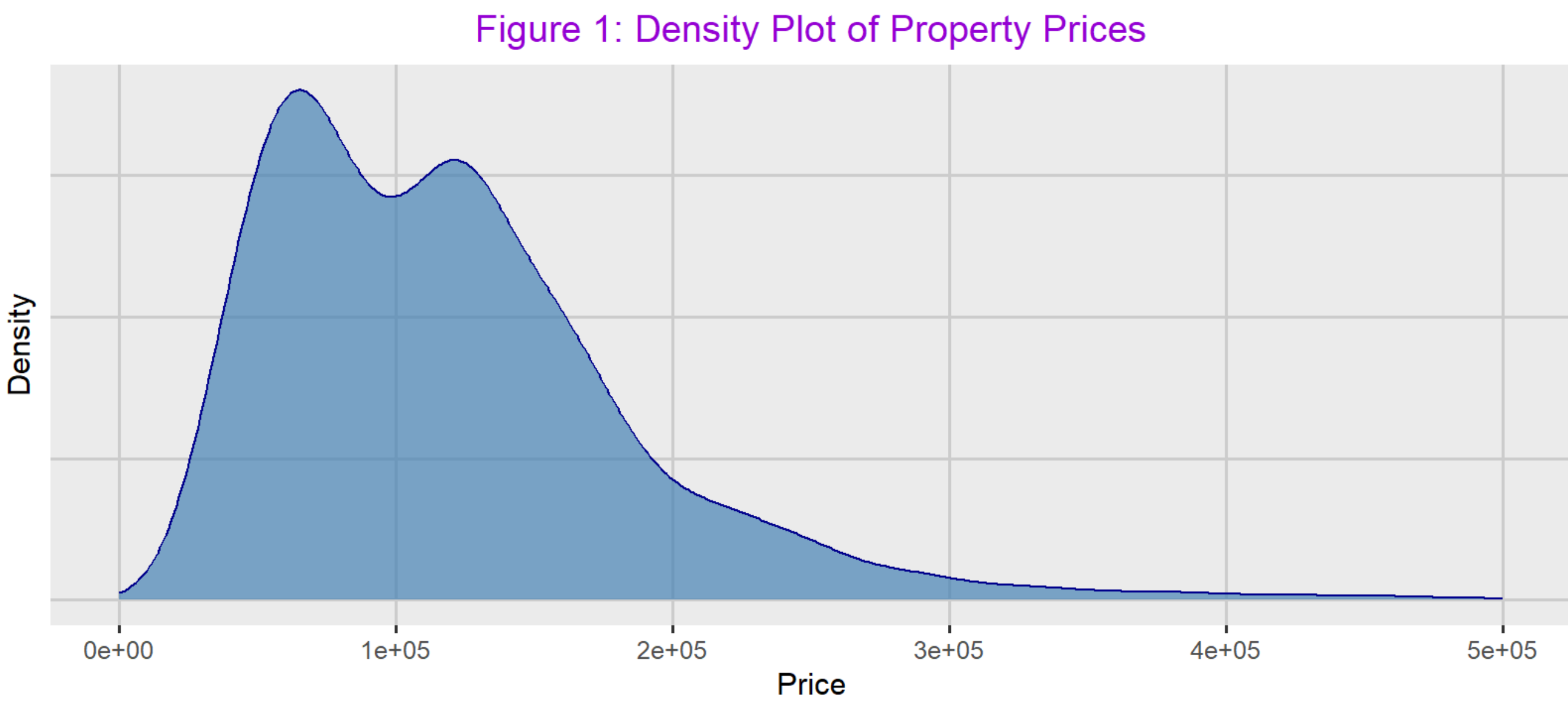
A first selection identified 13 covariates related to the house's size, location, age and energy efficiency, which could be reasonably included in the model. Nine different models were then fitted. The one which could best describe the data (see Figure 1) was identified as a Generalized Additive Gamma Model with log link, which showed **Adjusted  $R^2 = 0.826$** . Below the selected covariates along with their (approximate) p-value.

Covariate	p-value
Total floor area	< 2e-16
Number of rooms	< 2e-16
Energy consumption	1.24e-06
Environmental efficiency rating	8.244e-06
Property Type	< 2e-16
Postcode group	< 2e-16

Splines were applied to the first three covariates to model their non-linear relationships with price.

The following interaction terms were tested for significance and consequently included:

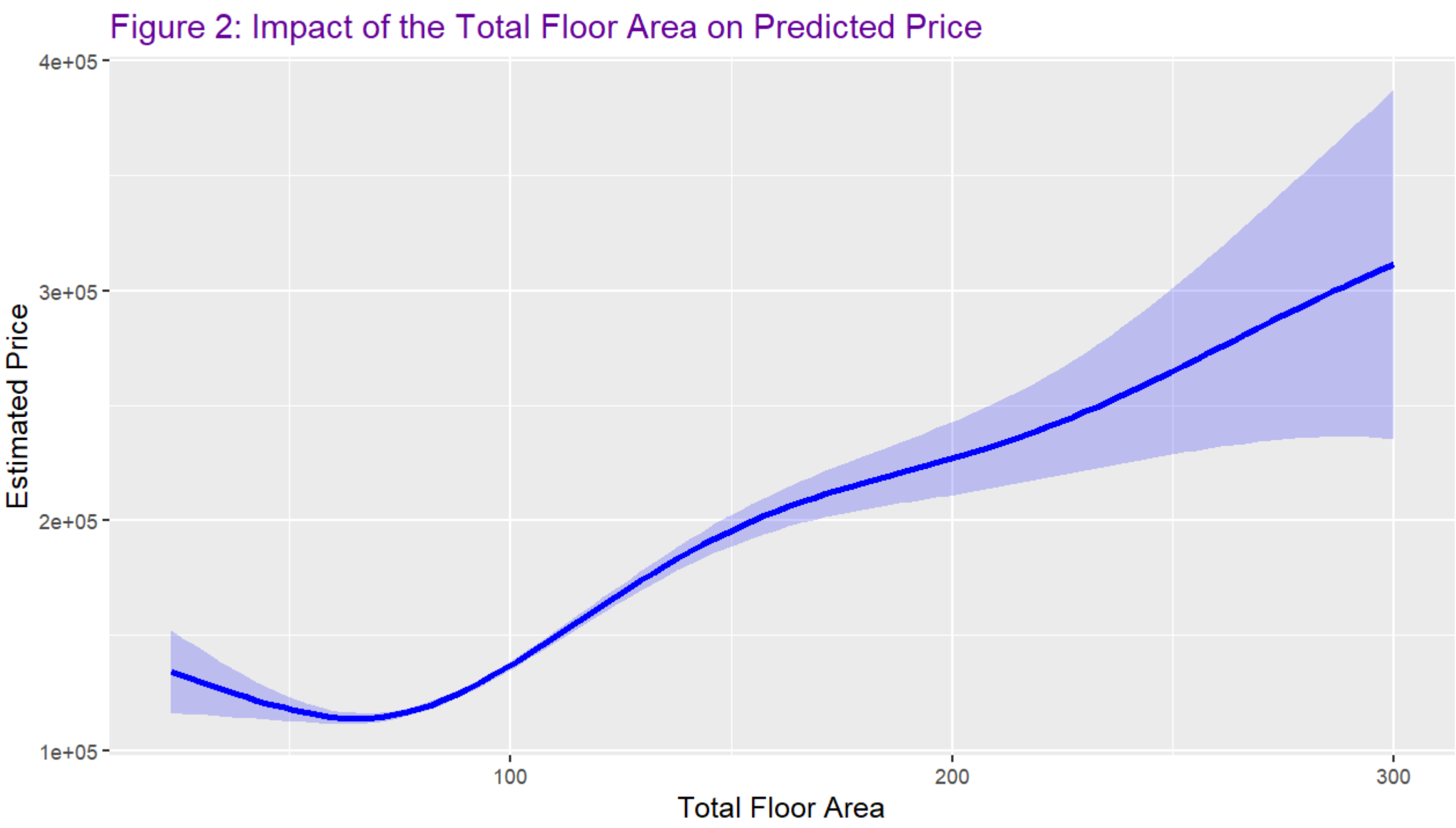
- *Total floor area* \* *Total number of rooms*;
- *Energy consumption* \* *Environmental efficiency*.



## Results

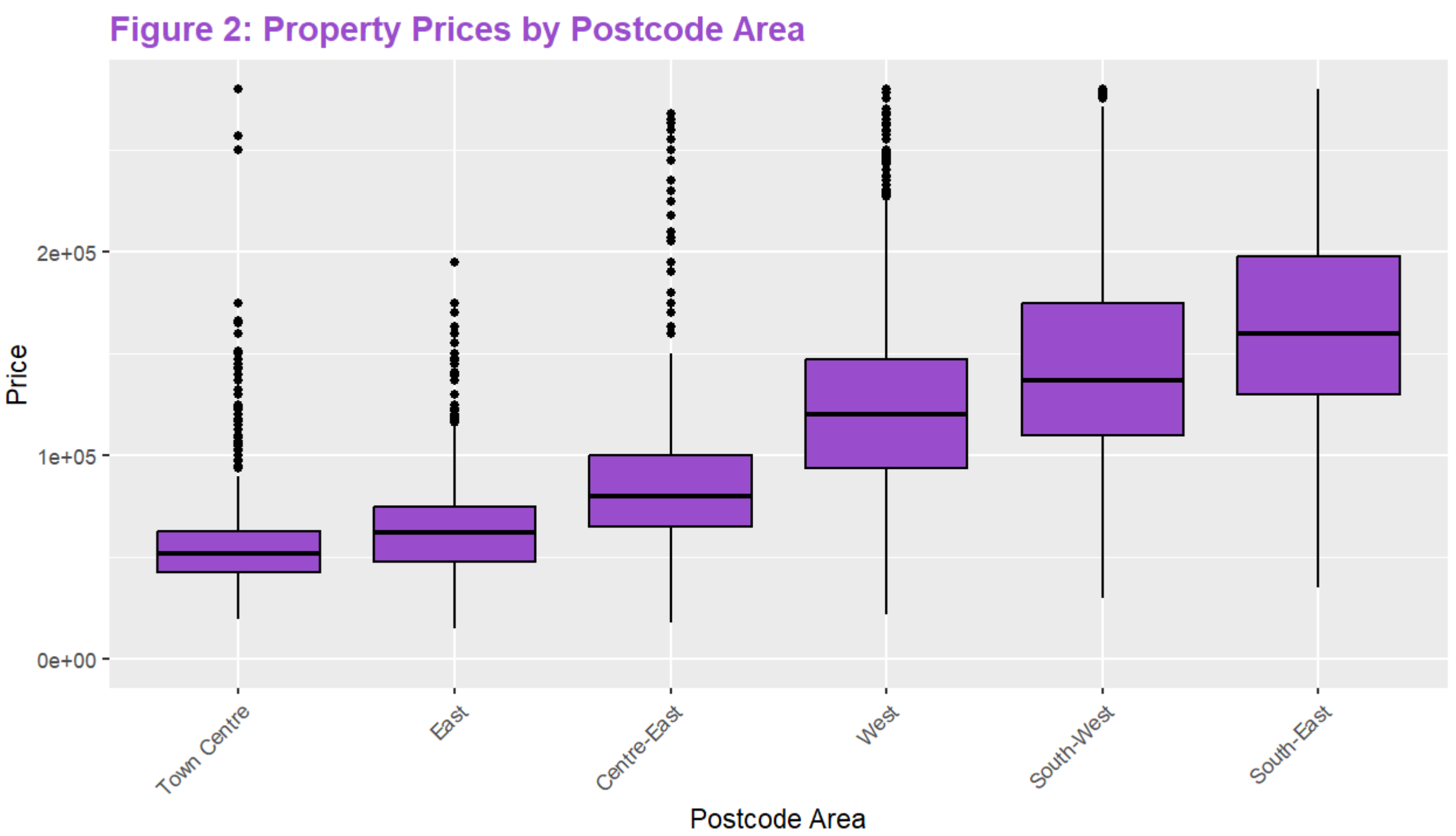
### How do prices vary with the size of the house?

Figure 2 illustrates the predicted prices against the *Total Floor Area*, keeping all the other covariates fixed. The results indicate a clear upward trend. **As the Total Floor Area increases, Property Prices also tend to rise.**



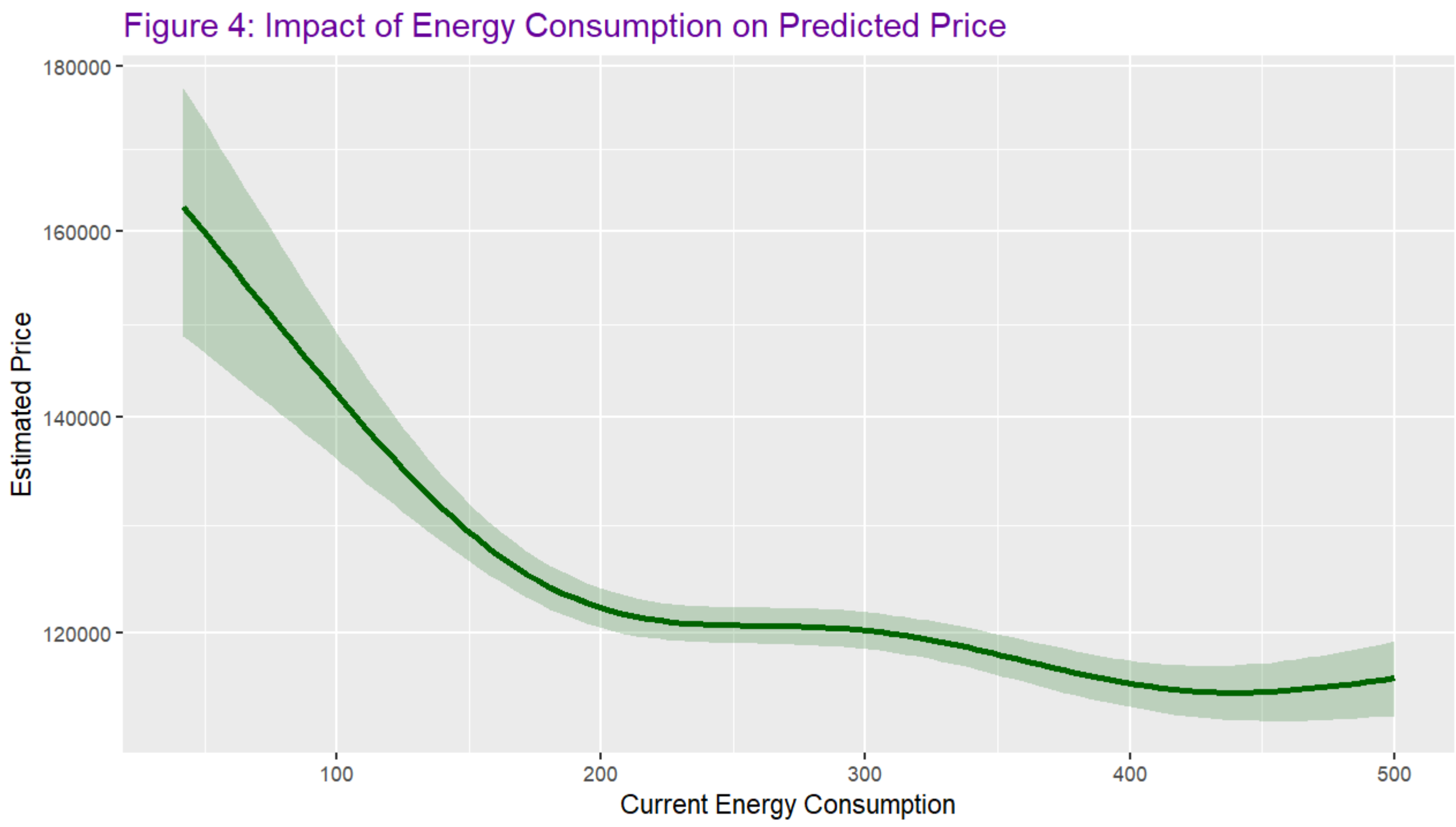
### What are the most affordable and most expensive areas of the county?

To identify Middlesbrough's most expensive areas, we grouped properties by the first three digits of their *Postcodes* and associated each group with specific parts of the city using information from Wikipedia. The analysis showed that **the highest property prices** are concentrated in the **Southeastern Districts** of Middlesbrough, specifically in Marton, Nunthorpe, and Ormesby, while the **Town Centre** has **the lowest average prices** (see Figure 3).



### How do prices vary with the size of the house?

To investigate whether buyers are willing to pay a premium for energy-efficient homes, we analyzed the relationship between property prices and the *Current Consumption of Energy* of each property. Figure 4 shows a downward trend in estimated prices as energy consumption increases. **Energy-efficient homes tend to command higher prices.**



We extended our analysis to explore the effect of *Environmental Efficiency Ratings* on property prices. Predictions across various efficiency levels indicate that **homes with higher energy efficiency ratings generally show higher prices** (see Figure 5). This suggests that buyers are willing to pay a premium for properties with efficient energy consumption and reduced environmental impact.

