Housing Realities: Multivariate Insights into Housing, Income, and Perception in Cornwall.

Executive Summary:

This advanced analysis builds on the main findings of the Cornwall Household Survey by applying multivariate techniques to explore deeper patterns in housing, income, employment, and perceptions across the region. Using Binary Logistic Regression (BLR), Ordinal Logistic Regression (OLR), and Cluster Analysis (CA), the report identifies key structural predictors of affordability, work patterns, and satisfaction with housing and the local environment. Tenure and household age structure consistently emerged as the strongest predictors across all models. Owner-occupiers were significantly more likely to report higher incomes, full-time employment, and satisfaction with housing and neighbourhood features. In contrast, social renters and residents of post-industrial areas were more likely to experience affordability constraints, car dependence, and dissatisfaction with environmental quality.

The cluster analysis grouped households into three distinct profiles, reflecting different combinations of affordability, transport reliance, and property characteristics. These clusters provided a broader context for understanding how individual factors intersect to shape housing outcomes. Overall, the advanced analysis confirms that while new housing offers benefits, it does not inherently resolve underlying structural inequalities. These findings reinforce the need for equitable housing strategies that address affordability, transport access, and local identity across all tenures and communities in Cornwall.

1. Overall Introduction

This advanced analysis report supplements the main findings of the 2024–25 Cornwall Household Survey by applying multivariate statistical techniques to uncover deeper relationships between housing conditions, demographic patterns, and economic outcomes. While the primary report identified clear differences in tenure, income, transport behaviour, and environmental perceptions between households in newer and older properties, this report seeks to go further by identifying predictors of income, employment, working-from-home patterns, and subjective perceptions using Binary Logistic Regression (BLR), Ordinal Logistic Regression (OLR), and Cluster Analysis (CA). These techniques not only reinforce earlier descriptive patterns but also reveal structural inequalities and spatial dynamics that demand further policy attention.

2. Aims, Rationale, and Methodology (Summary)

The primary aim of the survey was to compare the characteristics of residents living in housing built in the last ten years with those living in older housing stock. The research sought to explore how factors such as age, income, ethnicity, housing tenure, and migration status interact with affordability, satisfaction, and mobility across Cornwall. Postcodes with significant new development (20+ properties built since 2013) were identified using Land Registry data, allowing for sampling of mixed housing areas. Data collection involved a self-completion survey distributed both in print and online, yielding a sample of 938 households. This comparative approach allowed researchers to test long-held assumptions about who occupies new housing in Cornwall, and how in-migration, affordability, and housing design influence community dynamics. The survey was designed collaboratively with local stakeholders and piloted extensively before deployment.

3. Advanced Statistical Techniques Used

To extend the descriptive insights of the main report, the following statistical methods were applied:

- **Binary Logistic Regression (BLR)** was used to model binary outcomes—such as whether a household falls within a given income band, job status, or working-from-home frequency. BLR allowed us to identify statistically significant predictors of these outcomes across variables such as household age, tenure, ethnicity, property type, and age of dwelling.
- **Ordinal Logistic Regression (OLR)** was used for analysing outcomes measured on an ordinal scale, particularly housing satisfaction, affordability, environmental perception, and mobility attitudes. This approach helped to assess whether structural or demographic variables consistently shape subjective experiences.
- **Cluster Analysis (CA)** was used to segment the population into meaningful groups based on shared characteristics. Through hierarchical and K-Means clustering, the population was grouped into three broad clusters that reflect housing affordability, property features, and transport access. These insights help move beyond individual variables to understand how they combine into lived experiences.

3.1. Key results and insights

3.1.1. Binary Logistic Regression Results: Predictors of Household Income

This section presents the results of binary logistic regression models developed to identify which household and housing characteristics significantly predict different levels of income. By analysing five income brackets, the models offer a deeper understanding of how demographic composition, housing tenure, property type, and identity interact with economic outcomes in Cornwall. These findings supplement and statistically reinforce patterns observed in earlier descriptive analyses.

3.1.2. Analysis

The binary logistic regression models revealed consistent and important patterns that reinforce and expand upon findings from the main survey report. Across all five income categories analysed—ranging from under £16,000 to over £70,000, certain variables emerged repeatedly as significant predictors of a household's likelihood of falling within specific income bands. These included household age composition, property type, tenure, and ethnicity, while property age (whether the dwelling was built in the last ten years) was not a significant predictor in any model.

Age proved to be the most consistent demographic factor influencing income. Households with members aged 18–35 or 36–50 were significantly less likely to fall into the lowest income band (under £16,000), and more likely to be in the £22,000–£40,000 group. This reflects the broader finding that younger households are more likely to be working full-time and are well represented among newer housing developments. Conversely, those aged 51–65 were associated with the highest income group, over £70,000, pointing to their relative economic security, likely due to accumulated wealth and stable employment or pensions. Property type also emerged as a meaningful differentiator at the lower end of the income scale. Households living in flats were nearly twice as likely to be in the under £16,000 bracket compared to those living in houses. This supports previous findings that smaller property types are more commonly occupied by lower-income households, possibly reflecting both affordability and space constraints.

Property tenure was strongly associated with income in the upper brackets. Owner-occupiers, both those owning outright and those with a mortgage, were significantly more likely to be in the highest income category. This pattern confirms the relationship between wealth accumulation and property ownership, especially among older residents. It also reinforces the cluster analysis finding that homeownership does not guarantee affordability for everyone, but those who do earn more are more likely to be homeowners. Ethnicity had a more targeted effect. British-identified households were more likely to fall within the £41,000–£70,000 income group, compared to Cornish-identified households. This finding aligns with earlier results indicating that Cornish-identified individuals face structural barriers to accessing newer housing and higher incomes. While this does not suggest ethnicity itself causes economic disparity, it reinforces the link between identity and longer-term housing exclusion.

Income Band	Significant Predictor	Odds Ratio (Exp(B))	p-value
Under £16K	Age 18–35	0.65	0.01
Under £16K	Age 36–50	0.68	0.03
Under £16K	Flat (vs. House)	1.92	0.02
£22K–£40K	Age 18–35	1.67	0.04
£41K–£70K	British Identity (vs. Cornish)	1.3	0.03
Over £70K	Owned/Mortgaged (vs. Rent)	2.0	0.001
Over £70K	Age 51–65	1.3	0.02

Table3.1. Summary of Significant Predictors by Income Band

3.1.3. Summary of the model:

Overall, the models highlight that housing outcomes in Cornwall are deeply shaped by demographic and structural factors. While newer housing is often associated with younger and mid-income households, it is not exclusively tied to high-income earners. The results support a dual interpretation: newer homes are occupied both by upwardly mobile younger households and older, more affluent owner-occupiers. At the same time, the underrepresentation of Cornish-identified individuals in higher income groups and newer properties raises important questions around structural inequalities in access to economic security and housing opportunities. The regression analysis confirms that household age, property tenure, and housing type are key factors influencing income distribution, while property age plays a minimal role. The models reveal a clear divide between younger, mid-income, working households and older, higher-income owner-occupiers, reflecting the dual structure of Cornwall's newer housing landscape. Additionally, structural disparities in access to income based on ethnic identity, particularly for Cornish-identified households, remain an important concern for policy and further investigation. The next section will look at the binary logistic regression analysis of job types.

3.1.4. Binary Logistic Regression Results: Job Type Outcomes

This section presents the results of binary logistic regression models used to explore predictors of different job statuses among surveyed households. Five job-related outcomes were examined: full-time employment, multiple jobs, self-employment, full-time caring, and full-time education (for those over 18). Each model included the same set of explanatory variables: household age composition, ethnicity, property tenure, property type, and whether the dwelling was built in the last ten years. These models provide further statistical grounding for the demographic and housing trends discussed earlier in the report.

3.1.5. Analysis

Across the five employment-related binary logistic regression models, several consistent patterns emerged. Age structure was the most powerful and consistent predictor across models. The strongest results came from the model predicting full-time employment, where having someone aged 18–35 in the household increased the odds of being in full-time employment more than 26 times, and age 36–50 increased it by over 27 times. For households with members aged 51–65, the odds were still significantly higher (over 6 times). These figures clearly align with previous findings that younger and mid-aged working-age households are more likely to be engaged in full-time work, particularly in newer housing developments.

In the case of full-time education (for over-18s), all three age bands, 18–35, 36–50, and 51–65, were significant predictors, with odds ratios of 12.4, 5.2, and 4.0, respectively. This may appear counterintuitive for older age bands but reflects households where someone (e.g., a parent or mature student) is pursuing further education later in life, or where age bands may overlap with younger adult children. The self-employment model also showed significant results. Having someone aged 36–50 in the household increased the odds of being self-employed by 1.7 times, while age 51–65 raised the odds to 2.6 times. This fits with trends of mid-life or older residents opting for flexible or independent work, particularly in Cornwall's self-employed-heavy economy.

The multiple jobs model yielded fewer significant results. However, households with someone aged 36–50 had significantly higher odds (OR = 2.2) of holding multiple jobs. Property tenure also played a role here, with households renting from a housing association having more than twice the odds (OR = 2.5) of reporting multiple jobs compared to owner-occupiers. This pattern aligns with findings from the income analysis, where affordability pressures among renters were also associated with working multiple roles to meet household needs. The full-time carer model stood out for producing no statistically significant predictors. This suggests that the profile of full-time carers is more widely distributed across housing types, tenure categories, and age groups. It may also indicate that other unmeasured variables, such as disability, household structure, or informal support networks, play a stronger role in determining carer status.

Tenure and ethnicity were less consistently significant across models. However, in the education model, renting from a housing association significantly increased the odds of being in full-time education (OR = 4.6), perhaps pointing to overlaps between low-income households, life transitions, and attempts at social mobility. These findings reinforce the view that economic activity, work pattern, and education are not only strongly shaped by age but also interact with housing conditions in complex ways. In particular, the models underline the dual profile of Cornwall's economically active residents: on one hand, younger and mid-life working professionals in full-time or multiple employment; on the other, older or marginalised households relying on self-employment or accessing education later in life. The absence of significant effects for full-time carers further suggests that this group requires a different lens of analysis, possibly incorporating health or care dependency variables.

Job Type	Significant Predictors	Model Sig.
Full-Time	People 18–35 (OR=26.4), People 36–50 (OR=27.4), People 51–65 (OR=6.4), Older Property (OR=0.64), HA Rent (OR=0.23), Ethnicity: British (OR=0.57)	< .001
Multiple Jobs	People 36–50 (OR=2.19), HA Rent (OR=2.48)	.010
Self-Employed	People 36–50 (OR=1.69), People 51–65 (OR=2.58)	.002
Full-Time Carer	None (Model not significant)	.736
Full-Time Education	People 18–35 (OR=12.4), People 36–50 (OR=5.2), People 51–65 (OR=4.0), HA Rent (OR=4.6)	< .001

Table 3.2: Summary Table: Significant Predictors by Job Type

3.1.6. Brief summary

These regression models reveal that age is the most consistent determinant of employment patterns, but housing tenure, particularly housing association renting, also plays a notable role in specific outcomes such as multiple jobs and full-time education. While full-time carers were not predictable through the variables used, other job statuses aligned clearly with life stage and economic strategy. The insights drawn here support the earlier narrative that Cornwall's working-age population is navigating varied and often complex economic realities, shaped not only by life stage but also by housing conditions and long-term structural inequality. The next section will look binary logistic regression of working from home patterns.

3.1.7. Binary Logistic Regression Results: Working from Home Patterns

This section examines the results of five binary logistic regression models designed to predict the likelihood of households having at least one member working from home at various frequencies: all the time, 4 days a week, 3 days a week, 2 days a week, and never. These models use the same explanatory variables used in the income and job models, including household age composition, property type, tenure, property age, and ethnicity. The aim is to determine whether working-from-home patterns are associated with particular housing or demographic characteristics, contributing to the report's broader understanding of how home-based work is spatially and socially distributed in Cornwall.

3.1.8. Discussion of Findings

Across the five work-from-home models, findings were mixed in strength, with some models producing statistically significant predictors and others performing weakly. Notably, the ability of the models to explain variation in working-from-home patterns was limited overall, as reflected by low R^2 values and non-significant omnibus tests in some cases. This suggests that additional unmeasured factors, such as occupation type, employer policy, or personal preference, likely play important roles in shaping work-from-home behaviour.

3.1.8.1. Working from Home All the Time

The model predicting whether a household had someone working from home full-time produced a few noteworthy findings. Property type was the only significant predictor: households living in bungalows were significantly more likely to report full-time home-based work (OR = 2.015, p = .032). All other variables, including age group, tenure, and ethnicity, were not significant. While the model's goodness-of-fit statistics (Nagelkerke R^2 = .038) indicate weak explanatory power, the finding that housing form matters aligns with broader themes around home layout and remote work feasibility.

3.1.8.2. Working from Home 4 Days a Week

This model produced no statistically significant results. The omnibus test was not significant (p = .374), and all predictors, including household age, property age, type, tenure, and ethnicity, were unrelated to the likelihood of working from home 4 days a week. The model also failed to converge properly in the maximum number of iterations, indicating potential data sparsity or instability in prediction. As such, this model should be interpreted with caution and possibly excluded from visual summary representations of key patterns.

3.1.8.3. Working from Home 3 Days a Week

This model revealed two statistically significant predictors, both related to ethnicity. Households identifying as British or 'Other' (vs. Cornish) were significantly more likely to have someone working from home 3 days a week, with odds ratios of 3.468 (p = .012) and 4.639 (p = .035), respectively. This mirrors broader findings in the report showing that Cornish-identified residents are underrepresented in higher-income and ownership categories and suggests this exclusion may extend into employment flexibility or remote-working capacity.

3.1.8.4. Working from Home 2 Days a Week

This model yielded modest but interpretable findings. Households with a member aged 18–35 were significantly more likely to report this pattern of hybrid working (OR = 1.10, p = .005), and certain property ownership categories were also associated with working from home two days a week (p = .023). The nature of these tenure effects (likely mortgage or owned outright vs. renting) suggests a link between working patterns and housing security or space flexibility.

3.1.8.5. Never Working from Home

The model predicting no remote work produced several significant predictors. Younger households (aged 18–35) had higher odds of never working from home (OR = 2.303, p = .006), and both property type and tenure were associated with non-remote working. Living in flats and renting—particularly from social housing, was correlated with higher odds of having no one working from home. These results support previous findings in the report that residents in flats and rental housing are more likely to face economic insecurity and less flexibility, possibly due to work types that do not allow remote arrangements.

Summary of Key Predictors

Work From Home Frequency	Significant Predictors
All the Time	Property type (bungalow vs. house; OR = 2.015, p = .032)
4 Days a Week	None (model non-significant; poor fit and convergence issues)
3 Days a Week	Ethnicity: British vs Cornish (OR = 3.468, p = .012), Other vs Cornish (OR = 4.639, p = .035)
2 Days a Week	Age 18–35 (OR = 1.10, p = .005), Property ownership (p = .023)
None	Age 18–35 (OR = 2.303, p = .006), Property type (flats), Property ownership, Ethnicity (British and Other)

Table 3.3. Summary of the most important significant predictors identified across the five models.

3.1.9. Brief summary

The work-from-home models underscore that housing type, age, and ethnicity are linked in subtle but significant ways to the capacity for remote work in Cornwall. However, these models also highlight their own limits: several produced weak or non-significant results, and the explanatory power across all models was low. This suggests that while housing and demographic characteristics play a role in shaping remote-working potential, additional occupational and technological variables, such as job sector, employer flexibility, and digital infrastructure, are likely central to understanding who can work from home and how often. These findings still contribute useful nuance to the broader picture of housing, work, and income patterns explored in this report.

3.1.10. Ordinal Logistic Regression Results: Housing, Environment, and Perceptions

This section presents the findings from the ordinal logistic regression (OLR) models used to explore how respondents' perceptions of affordability, housing quality, transport access, and environmental features relate to key structural variables, including tenure, property type, ethnicity, location (postcode), and whether the property was built in the last ten years. The outcomes were measured using ordinal responses ranging from disagreement to agreement and included variables such as whether the new property meets the household's needs, whether it is of better quality than their previous one, whether they face affordability issues, and perceptions of green space availability, clean air, and car dependence. These models were developed to move beyond descriptive statistics and identify consistent patterns in perception shaped by housing and socio-spatial context. However, it is important to note that several models encountered statistical limitations that affect the strength and reliability of their conclusions.

3.1.11. Discussion findings and key insights

Across the models, a consistent and powerful predictor of positive perceptions was property tenure. Households that own their homes, particularly those owning outright or through a mortgage, were significantly more likely to report that their new properties met their needs and were of better quality. In the model exploring whether the property meets needs, ownership emerged as a strong positive predictor, and the model itself demonstrated a reasonable fit, with a significant chi-square value ($p < .001$) and good parallel lines assumption. Similarly, in the model assessing whether the property is of better quality, ownership again increased the likelihood of agreement. These patterns align with earlier findings from descriptive statistics and reinforce the structural advantages experienced by owner-occupiers in accessing satisfactory housing.

Ethnicity and postcode type also played a role in shaping perceptions, though their effects varied across outcomes. For example, households identifying as Cornish were less likely to report satisfaction with their housing quality or affordability, though these effects did not always reach statistical significance. More concretely, respondents living in industrial or post-industrial areas reported poorer experiences, particularly regarding clean air and car dependence. These households were more likely to agree that they depend on cars for travel and less likely to report satisfaction with the local environment. These findings support the wider observation that spatial inequalities persist in Cornwall, with area-level factors shaping both material and perceptual experiences of housing and neighbourhoods. Environmental perception models, however, presented mixed results. For clean air and green spaces, model fit varied considerably. The clean air model identified tenure and area type as significant predictors: renters from housing associations, for example, were significantly less likely to agree that they live in areas with clean air. This points to environmental injustice, where vulnerable groups are more likely to live in locations with poorer environmental quality. In contrast, the green space model failed to produce any significant predictors, despite its thematic importance in the survey. The model's goodness-of-fit indicators were weak, and pseudo R^2 values were low, suggesting limited explanatory power. This outcome may be due to a lack of variation in responses (as green space importance was highly valued across all groups), making it difficult for the model to detect significant differences.

Models related to transport access and cost were similarly uneven. The model assessing whether transport in Cornwall was seen as good value failed to produce any significant predictors and had extremely low explanatory power (Nagelkerke $R^2 < .01$), suggesting that perceptions of transport cost are either evenly distributed across groups or shaped by unmeasured variables such as personal mobility needs, access to public transport, or specific job demands. Meanwhile, the car dependence model produced more interpretable results, with area type again playing a clear role. Respondents in industrial or post-industrial areas had higher odds of agreeing that they depend on a car for most journeys, reinforcing findings in earlier sections of the report regarding the uneven spatial distribution of transport infrastructure and connectivity. Affordability remains one of the clearest areas of perceptual inequality revealed by these models. The model exploring whether respondents had experienced affordability issues showed that those renting from housing associations were significantly more likely to report problems. This echoes earlier descriptive findings and highlights the lived realities of housing insecurity among social renters, even as new developments are often intended to address affordability. The consistency of this result across both affordability and housing quality models further underscores the structural disadvantages that persist for renters in Cornwall.

Despite these valuable insights, it is important to acknowledge the limitations encountered in several of the models. Many of the OLR models had low pseudo R-squared values, generally below 0.10, indicating that much of the variation in the outcome variables remains unexplained. More critically, several models suffered from high levels of sparsity, with large proportions of cells in the design matrix having zero frequencies. For example, the model assessing whether the new property meets needs had nearly half of its cells (46.5%) with zero frequencies and triggered warnings about quasi-complete separation. In such cases, parameter estimates may be unstable or tend toward infinity, reducing confidence in the model's validity. While SPSS continued estimation and produced coefficients, these outputs should be interpreted with caution, and the findings from these models should not be overgeneralised. Moreover, some models, particularly those related to green space importance, transport value, and clean air, produced relatively flat estimates across groups, suggesting that many respondents shared similar views regardless of tenure or housing conditions. In these cases, a regression approach may be less appropriate, as high agreement levels may reflect universal values rather than group-based differences.

Outcome Variable	Significant Predictors
New Property Meets Needs	Homeownership (↑), $p < .001$
New Property Better Quality	Homeownership (↑), $p = .037$
No Affordability Issues	Housing association renters (↓), $p < .001$
Clean Air Satisfaction	Housing association renters (↓), $p = .009$
Car Dependence for Travel	Industrial/post-industrial area (↑), $p < .05$
Transport Too Expensive	Private renters (↑), $p < .05$
Ease of Walking in Local Area	Coastal/Inland towns (↑), $p < .05$
Green Spaces Importance	None

Table 3.4. Summary of Significant Predictors from OLR Models

3.1.12. Brief summary

The ordinal logistic regression analysis reinforces key themes in the housing report: owner-occupiers generally report better outcomes in terms of housing satisfaction and perceived environmental quality, while renters, particularly those in housing association properties, and residents of post-industrial areas tend to face structural disadvantages. However, the explanatory power of these models was modest, and several encountered technical limitations that affect confidence in their estimates. These issues underscore the complexity of modelling perceptions and the need for caution when interpreting regression outputs based on categorical predictors and ordinal outcomes. Nonetheless, the analysis provides valuable context to the lived experiences of different housing groups in Cornwall and highlights areas, especially affordability, satisfaction, and car dependence, where structural inequalities remain pronounced.

3.1.13. Cluster Analysis: Interpretation and Limitations

This section presents the interpretation of the final cluster solution developed through hierarchical and K-Means cluster analysis. The aim of the analysis was to identify meaningful groupings of households based on housing affordability, transport choices, property characteristics, and demographic indicators. Despite extensive efforts in data imputation and cleaning, a small number of cases remained unclassified. Nevertheless, the resulting three-cluster solution offers robust insights into key structural and lifestyle differences among residents in Cornwall.

3.1.14. Discussion and key insights

The clustering process was informed by multiple methodological steps to ensure accuracy and minimise data bias. Initially, missing data were addressed using multiple imputation in SPSS, generating five complete datasets. However, SPSS concatenated rather than replaced the original data, inflating the dataset to over 10,000 cases. After applying a filter to retain only the first imputed dataset, the final working sample was reduced to 938 complete cases. Hierarchical Cluster Analysis (HCA) was used as a starting point to explore the structure of the data. Using squared Euclidean distance and average linkage, both three- and four-cluster solutions were explored, with the three-cluster option selected for further refinement using K-Means analysis. The final K-Means solution successfully classified 95.8% of cases, with only 4.2% (39 respondents) remaining unclassified. This low exclusion rate is acceptable in cluster analysis and does not substantially undermine the reliability of the solution. However, the existence of unclassified cases does warrant consideration and is addressed in detail below.

The three final clusters showed clear differentiation, particularly along dimensions of housing affordability, main transport mode, and property type. These three variables were highly significant in the clustering process ($p < .001$), indicating strong influence on the formation of cluster groups. For instance, affordability concerns were the most statistically significant factor, underscoring the persistent stratification in Cornwall's housing market. Similarly, main transport mode effectively separated residents based on car dependence versus public or active transport use, which in turn intersected with geographic and tenure-based patterns. Property year built also contributed significantly ($p < .001$) to cluster formation, suggesting that the age of housing stock influences how households group together in terms of lifestyle and need. In contrast, property ownership status ($p = .697$) and ethnicity ($p = .001$) played weaker roles in shaping clusters. Ownership did not appear to be a central factor distinguishing groups, which may reflect Cornwall's mixed ownership-rental landscape. Ethnicity showed only marginal significance, likely due to the relatively homogeneous ethnic makeup of the region.

Despite the strength of the three-cluster solution, several limitations were identified and should be acknowledged transparently. Most importantly, the presence of 39 unclassified cases suggests that a portion of the sample either did not fit neatly into any centroid or had inconsistent or outlier values. This may have stemmed from three interrelated issues: First, SPSS's handling of categorical variables within K-Means analysis, designed primarily for continuous data, can be problematic, even when categorical variables are numerically recoded. Some respondents' characteristics may not have aligned cleanly with any one cluster, making it difficult for the algorithm to assign them. Second, it is possible that some households represented outlier cases, such as those with unusual transport arrangements or atypical affordability responses. These cases may fall between clusters, resulting in their exclusion during centroid refinement. Finally, while multiple imputation addressed most missing data, the classification process may have been affected by underlying inconsistencies introduced during imputation. If imputed values created combinations of characteristics rarely observed in the raw data, this could have confused the clustering algorithm and reduced model fit for certain records.

3.1.15. Brief summary

The final three-cluster solution provides a statistically valid and conceptually meaningful segmentation of residents in Cornwall. It highlights clear divisions in housing affordability, transport reliance, and property conditions, with implications for planning, service provision, and housing policy. However, the methodological constraints, particularly the presence of unclassified cases and the use of categorical variables in a technique optimised for continuous inputs, should be acknowledged as limitations. The results remain robust for most of the population, but further analysis using alternative clustering approaches or complementary qualitative insight could help clarify the position of the 4.2% of respondents who were not assigned to any cluster.

Variable	**Statistical Significance (p-value)**	**Interpretation**
Property Type	< 0.001	Strongly differentiates clusters
Main Transport Mode	< 0.001	Clear divider between car-dependent and others
Affordability Concern	< 0.001	Most significant factor separating groups
Property Year Built	< 0.001	Reflects lifestyle, affordability, and location
Property Ownership	0.697	Not a significant driver of cluster membership
Ethnicity	0.001	Weak influence, likely due to regional homogeneity

Table 3.5. Summary Table: Key Drivers of Cluster Formation

4. Cross-Method Discussion: Integrating Key Findings

The advanced analyses confirmed many of the patterns identified in the primary survey report but also brought new complexity and depth. Most notably, housing tenure and age composition of households emerged as the strongest and most consistent predictors across all regression models. BLR showed these factors were central in shaping both income and employment patterns. Owner-occupiers, especially those with mortgages, were more likely to have higher incomes and be engaged in full-time work or remote working arrangements. The OLR models reinforced the importance of tenure in shaping perceptions of affordability, quality, and environmental conditions. Owner-occupiers generally expressed more positive views, while social renters, particularly in post-industrial areas, were more likely to report dissatisfaction with affordability and air quality. These perceptual outcomes aligned well with the BLR's more objective economic outcomes, suggesting both structural and subjective inequalities.

Cluster Analysis offered a holistic view, combining multiple variables to define three distinct resident groups: financially secure homeowners, affordability-constrained renters, and mixed-status residents with moderate pressures. This classification not only echoed the regression outputs but helped reveal hidden dynamics, such as the fact that homeownership alone does not guarantee affordability, particularly for those in transitional stages. Taken together, the advanced methods suggest that Cornwall's housing inequality is layered: demographic composition, tenure, ethnicity, and geography intersect to produce varied experiences of economic security, housing satisfaction, and mobility. The findings suggest that the original research goals, understanding who lives in new housing, and whether newer housing meets local needs, have been addressed in full. However, the data also demonstrate that a deeper and more enduring divide exists between long-term residents and recent in-migrants, and between those in financial stability and those under pressure, regardless of housing age.

5. Limitations of the Advanced Analysis

While the advanced statistical techniques applied in this report provide valuable insight into housing and demographic patterns in Cornwall, several limitations should be acknowledged for transparency. Some binary and ordinal logistic regression models exhibited low explanatory power, as indicated by low pseudo R-squared values, suggesting that important influencing factors may lie outside the variables included in the models. A number of ordinal models also showed signs of sparse data and quasi-complete separation, particularly where categorical variables interacted with ordinal outcomes. These issues can lead to unstable estimates and reduced generalisability. In cluster analysis, a small proportion of cases (4.2%) could not be classified into any cluster, likely due to the complexity of categorical data and the presence of outliers. While these limitations do not invalidate the findings, they highlight the importance of interpreting results with caution and considering complementary qualitative or contextual data where appropriate.

6. Overall Conclusion

This advanced statistical analysis confirms that Cornwall's housing system reflects significant structural inequalities, many of which are not immediately visible in descriptive data. Newer housing is not uniformly occupied by high-income or working households. Rather, it accommodates both economically active younger households and older, wealthier in-migrants, two very different groups. Meanwhile, renters, particularly in social housing, continue to face affordability constraints and report lower satisfaction with their homes and surroundings. The findings from the BLR and OLR models demonstrate that property tenure, household age, ethnicity, and location are significant predictors of both material outcomes and subjective experiences. Cluster analysis enriches this picture by showing how affordability, transport mode, and housing form combine to shape distinct resident experiences. Although newer housing may offer physical improvements, it does not automatically resolve structural inequalities. This report offers clear implications for housing policy and planning in Cornwall. To be effective, future interventions must go beyond simply increasing housing stock. Instead, they should ensure that affordability, location, and access to amenities are embedded into new developments, and that longstanding barriers faced by local residents, particularly those identifying as Cornish, are addressed directly. The evidence gathered here supports targeted, equitable housing strategies grounded in the lived realities of Cornwall's diverse population.