NOVA IMS Information Management School

MDSAA

Master Degree Program in

Data Science and Advanced Analytics

Analysis and Evaluation of the Model for Impact Assessment, for a social impact NGO

Validity Testing in a Real World Context

Ana Leonor Gonçalves Vital

Master Thesis

presented as partial requirement for obtaining a Master's Degree in Data Science and Advanced Analytics

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

Analysis and Evaluation of the Model for Impact Assessment, for a social impact NGO

Validity Testing in a Real World Context

by

Ana Leonor Gonçalves Vital

Master Thesis as partial requirement for obtaining the Master's degree in Data Science and Advanced Analytics, with a specialization in Business Analytics

Supervised by

Professor Maria Manuela Simões Aparício da Costa, PhD in Information Management, NOVA Information Management School

December, 2025

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration. I further declare that I have fully acknowledged the Rules of Conduct and Code of Honor from the NOVA Information Management School.

Lisbon, 01st of December 2025

Ana Leonor Gonçalves Vital

DEDICATION

Para o meu Pai,
que mesmo não estando
faz-se presente.

ABSTRACT

This dissertation evaluates the validity and robustness of the Entity B model for measuring social impact in non-profit housing interventions and examines its ability to capture multidimensional change. Guided by two research questions, this study applies a mixed-methods approach structured under the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework. Quantitative analysis was conducted using a multiple linear regression on survey data, while qualitative insights were integrated through sentiment analysis of beneficiaries' open-ended responses. Results confirm strong internal coherence and explanatory power (R-squared = 0.90), with physical indicators emerging as the dominant predictors of perceived impact. Yet, the social and financial dimensions exhibit weak explanatory capacity, and the professional dimension had to be excluded due to data constraints - indicating that the model only partially achieves its multidimensional ambition. The sentiment analysis corroborates these findings, revealing predominantly positive perceptions associated with housing improvements but limited evidence of relational or economic transformation. Theoretically, this research advances methodological innovation by validating a hybrid framework that combines data mining processes with computational text analysis, offering a replicable approach for multidimensional impact assessment. Practically, it provides actionable recommendations for non-profit organisations: baseline data collection, longitudinal follow-up, and stakeholder triangulation, to strengthen accountability and learning mechanisms. Limitations include the absence of baseline data, reliance on self-reported perceptions, short observation windows, and exclusion of the professional dimension. Future research should address these gaps through longitudinal designs, reintegration of omitted dimensions, refinement of relational indicators, and benchmarking against established frameworks such as Social Return on Investment (SROI) and Most Significant Change (MSC).

KEYWORDS

Social Impact Measurement; Nonprofit Housing Interventions; Mixed-Methods Approach; CRISP-DM Framework; Multiple Linear Regression; Sentiment Analysis

Sustainable Development Goals (SDG):

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

CRISP-DM	Cross Industry Standard Process for Data Mining
GRI	Global Report Initiative
MSC	Most Significant Change
NPOs	Non-Profit Organisations
SDGs	Sustainable Development Goals
SIA	Social Impact Assessment
ROI	Social Return on Investment
TOC	Theory of Change
UIC	Units of Complete Impact
VIF	Variance Inflation Factor

NOTE: For confidentiality reasons, the names of the organizations have been anonymized. Organization A refers to the implementing NGO, and Entity B refers to the external methodological developer.

1. INTRODUCTION

The growing demand for transparency and accountability in the social sector has placed increased emphasis on the need for robust, evidence-based tools to assess the impact of interventions (Alomoto et al., 2022; Salido-Andres et al., 2022). Non-profit organisations face the need to demonstrate the effectiveness of their work, not only to secure funding but also to support strategic decision-making and improve programme delivery (Schröer et al., 2021; Epstein & Roy, 2001). Despite this, many social impact assessment models remain conceptually underdeveloped and empirically untested, limiting their reliability and relevance in real-world contexts (Costa & Pesci, 2016; Tuan, 2008).

One of the most visible expressions of social exclusion is housing poverty, a multidimensional problem affecting thousands of families in Portugal (Dunn, 2000; Clair & Hughes, 2019; Stone & Hulse, 2007). Organization A, a Portuguese non-profit organization, is a prominent actor committed to mitigating housing poverty (Social Data Lab, 2024). Operating under the motto "rebuilding houses, restoring lives," the organization has renovated hundreds of homes across the country and mobilized thousands of volunteers (Mullins et al., 2012). As it continues to grow and seeks long-term sustainability, the ability to measure the outcomes of its work has become essential (Costa & Pesci, 2016).

In response to this need, Organization A, in collaboration with the Entity B, developed a statistical model to quantify the social impact of its housing interventions. This model is based on a previous model made by the Entity B to ensure and measure impact in the social sector (Social Data Lab, 2024). However, the effectiveness of this model in capturing multidimensional social impact has not yet been empirically evaluated, creating a methodological and practical gap that this dissertation aims to address (Funnell & Rogers, 2011).

The central research question of this dissertation then is: How effectively does the statistical model developed by the Entity B evaluate the social impact of Organization A*'s* housing interventions? This question not only underscores the need for methodological rigor but also highlights a critical gap in applied evaluations within the non-profit sector. Accordingly, the dissertation is guided by two overarching objectives that shape its analytical trajectory. First, it seeks to assess the conceptual foundations of the Entity B model, examining its alignment with best practices in social impact assessment. Second, it aims to determine the model's effectiveness in capturing meaningful changes in beneficiaries' lives, thereby demonstrating its practical applicability. These objectives collectively reveal the dual ambition of the study: to contribute to theoretical debates while generating actionable insights for practice (Funnell & Rogers, 2011).

This study employs a mixed-methods approach, integrating both quantitative and qualitative data to evaluate the model's effectiveness and applicability. The methodology follows the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework a structured, iterative process widely used in data science projects (Schröer et al., 2021; Daw & Hatfield, 2018). Quantitative analysis is primarily conducted through the construction and validation of a multiple linear regression model (Su et al., 2012), using structured survey data collected from beneficiaries. In parallel, qualitative insights are drawn from interviews and open-ended responses (Davies & Dart, 2005; Tuan, 2008), offering a more holistic view of the intervention's perceived effects. This research benefits from a partnership with two key stakeholders: Organization A, which provides contextual information and access to data; and the Entity B, the external entity responsible for developing the initial version of the model.

The primary contribution of this dissertation lies in its empirical evaluation of a statistical model for social impact assessment within a non-profit context. By evaluating the model's capacity to capture meaningful outcomes, the study aims to support Organization A in strengthening its impact measurement practices. More broadly, it contributes to the academic and practical discourse on social impact assessment by 1) demonstrating how a structured data science methodology (CRISP-DM) can be adapted for use in the social sector, and 2) offering insights and recommendations that may inform similar organisations operating in the same fields (Gómez et al., 2022; Alomoto et al., 2022).

In summary, the introduction establishes the conceptual and practical gap that this dissertation seeks to address - namely, the absence of empirically validated models for social impact assessment in non-profit housing interventions. This gap not only limits organizational learning but also constrains accountability mechanisms. Therefore, the next chapter reviews the theoretical and empirical foundations that inform the design and operationalization of the impact model (Funnell & Rogers, 2011).

2. LITERATURE REVIEW

This chapter provides a comprehensive review of the relevant theoretical and empirical foundations to the assessment of housing-related social interventions, thereby establishing the conceptual scaffolding for the empirical analysis that follows in the next chapters. Notably, it synthesizes evidence across four domains: 1) non-profit roles, 2) housing poverty, 3) intervention strategies, and 4) impact measurement.

2.1 Social Impact Assessment and The Role of Non-Profit Organisations

Non-profit organisations (NPOs) play a pivotal role in addressing unmet social needs, particularly in contexts where state intervention and market mechanisms fall short. Operating in environments marked by vulnerability and exclusion, NPOs mobilize community resources and foster social innovation to mitigate structural inequalities (Gómez et al., 2022). Their contribution extends beyond service delivery, encompassing broader goals such as empowerment, participation, and long-term social transformation (Epstein & Roy, 2001).

In this context, tools such as the Social Impact Assessment (SIA) framework have gained increasing relevance. Historically, the non-profit sector relied heavily on qualitative narratives and subjective evidence to demonstrate effectiveness (Davies & Dart, 2005; Tuan, 2008). However, growing demands for transparency and accountability have led organisations to adopt more systematic and evidence-based approaches (Salido-Andres et al., 2022). SIA thus serves a dual purpose: it enhances organizational legitimacy and sustainability while fulfilling external reporting requirements (Alomoto et al., 2022).

Beyond accountability, SIA contributes to strategic learning and continuous improvement. However, its implementation remains uneven across contexts. This contrast underscores the tension between normative frameworks and operational realities, a challenge widely discussed in the literature (Kročil, 2018; Alomoto et al., 2022). By systematically evaluating outcomes, organisations can identify strengths and weaknesses, refine their interventions, and adapt to evolving social contexts, SIA's role as both a reporting tool and a mechanism for organizational learning (Salido-Andres et al., 2022) is reinforced.

Literature also highlights the unique challenges of measuring social impact in the non-profit sector. Unlike in corporate environments, where financial indicators dominate, NPOs pursue multidimensional goals that are not easily quantifiable. Attempts to monetize intangible outcomes, such as psychological well-being, empowerment, or social cohesion, risk oversimplifying complex realities (Tuan, 2008). Conversely, purely qualitative approaches, such as participatory evaluations or the Most Significant Change (MSC) method, may be criticized for their subjectivity and lack of comparability (Davies & Dart, 2005).

To address these tensions and to capture both the scale and the richness of social change (Alomoto et al., 2022), hybrid approaches that integrate quantitative rigor with qualitative depth have emerged. For example, (Kročil, 2018) advocates for the combination of three tools - Social Return on Investment (SROI), Local Multiplier 3 (LM3), and the Social Enterprise Balanced Scorecard (SEBS) - to achieve a multidimensional perspective. Similarly, (Epstein & Roy, 2001) propose adapting the Balanced Scorecard to the social sector, enabling organisations to monitor both financial sustainability and social value creation.

Participatory and stakeholder-driven approaches have also gained prominence, enhanced the legitimacy of impact assessments and ensured that indicators reflect the lived experiences of beneficiaries (Gómez et al., 2022; Kročil, 2018; Svejenova & Christiansen, 2017). These methods underscore the importance of SIA not only as a retrospective evaluation tool but also as a forward-looking mechanism that informs strategic innovation and stakeholder engagement (Svejenova & Christiansen, 2017).

Despite these advances, significant gaps remain. Many existing models lack empirical validation, raising concerns about their robustness in real-world applications (Alomoto et al., 2022; Salido-Andres et al., 2022). Moreover, there is limited consensus on how to reconcile context-specific indicators with standardized frameworks such as the Sustainable Development Goals (SDGs) or the Global Reporting Initiative (GRI). This fragmentation reflects an ongoing tension between the pursuit of comparability and the need for nuanced, context-sensitive evaluations.

In summary, Social Impact Assessment has become an indispensable practice for non-profit organisations. It strengthens accountability, supports funding strategies, and fosters organizational learning. However, persistent challenges regarding methodological rigor, comparability, and stakeholder inclusion call for innovative, hybrid approaches that bridge the gap between scientific robustness and practical applicability.

2.2 The Concept of Housing Poverty

Housing poverty is widely conceptualized in the academic literature as a multidimensional form of deprivation that transcends the mere absence of shelter. It encompasses situations in which individuals or households lack access to adequate, safe, and affordable housing that meets both physical and social needs (Gómez et al., 2022; Chadee, Martin, Mwasha, & Otuloge, 2022). This definition includes various dimensions such as affordability constraints, overcrowding, poor structural quality, insecurity of tenure, and limited access to essential services.

Over time, the understanding of housing poverty has evolved from a narrow focus on economic indicators, such as rent-to-income ratios to a broader framework that incorporates habitability, dignity, health, and social participation (Dunn, 2000; Stone & Hulse, 2007).

Early contributions to the field emphasized the role of housing as a determinant of health. (Dunn, 2000), for instance, argued that substandard housing conditions can negatively affect both physical and mental well-being. This perspective has been reinforced by empirical studies demonstrating strong correlations between inadequate housing and respiratory illnesses, psychological stress, and reduced quality of life (Clair & Hughes, 2019).

From a human rights standpoint, adequate housing is considered a fundamental entitlement and a cornerstone of social justice. Deprivation in this domain not only reflects material poverty but also perpetuates inequality by restricting access to education, employment, and healthcare (Stone & Hulse, 2007). Housing poverty thus operates both as a consequence of broader socio-economic disadvantage and as a driver of long-term exclusion.

Another critical dimension of housing poverty is relational. Housing is not merely a physical space but also a social one, where individuals build relationships and participate in community life (Gómez et al., 2022; Stone & Hulse, 2007). Inadequate or insecure housing can hinder social interactions, diminish self-esteem, and exacerbate isolation (Gómez et al., 2022). For vulnerable groups, such as the single-parent, elderly, or low-income households, housing insecurity can intensify cycles of exclusion by weakening social networks and reducing resilience to economic shocks.

To operationalize this concept, researchers have developed diverse measurement frameworks. Common indicators include affordability ratios, overcrowding rates, structural adequacy (e.g., presence of damp, poor insulation), security of tenure, and access to basic services such as water, sanitation, and energy. These metrics enable cross-country comparisons and provide policymakers with actionable benchmarks. However, reliance on quantitative indicators alone risks overlooking the subjective and experiential aspects of housing deprivation, such as perceived safety, dignity, and psychological well-being (Chadee, Martin, Mwasha, & Otuloge, 2022). In response, mixed-methods approaches have gained traction, combining quantitative surveys with qualitative accounts of lived experience. These methodologies offer both the comparability required for policy analysis and the depth needed to understand the nuanced realities of affected households. This shift reflects a growing recognition of housing poverty as a multifaceted and experiential condition, not reducible to economic constraints alone.

First, there is no universal consensus on the thresholds that define housing poverty, with definitions and indicators varying across cultural and institutional contexts (Stone & Hulse, 2007). Second, empirical research has disproportionately focused on urban housing deprivation, leaving rural contexts underexplored despite their distinct characteristics (Clair & Hughes, 2019). Third, housing poverty is often intertwined with related phenomena such as energy poverty, neighbourhood deprivation, and environmental hazards, necessitating more interdisciplinary and integrative research approaches.

Housing poverty emerges as a multidimensional phenomenon deeply intertwined with structural inequalities and relational dynamics. This complexity not only challenges simplistic evaluation models but also reveals the necessity of hybrid approaches that integrate quantitative rigor with qualitative depth. Therefore, the methodological choices outlined in the next chapter respond directly to these gaps, operationalizing a mixed-methods design aligned with contemporary best practices.

2.3 HOUSING INTERVENTIONS AND THE ROLE OF NGOS

As established in the previous chapter, housing poverty is a multidimensional phenomenon that affects not only the physical condition of dwellings but also impacts health, dignity, and social participation. Addressing this form of deprivation requires interventions that go beyond technical solutions, engaging with the structural and relational dimensions of poverty (Clair & Hughes, 2019; Dunn, 2000; Stone & Hulse, 2007). Housing interventions must therefore be conceived as social responses that promote empowerment and integration (Gómez et al., 2022; Keans et al., 2010).

It is important to distinguish between housing policies and housing interventions. Housing policies are typically large-scale, state-led initiatives such as social housing programs, rental subsidies, or tenure regulations. These policies shape national housing systems and influence long-term supply. In contrast, housing interventions are often localized efforts targeting specific households or communities. They include the rehabilitation of deteriorated dwellings, energy retrofits, and supportive housing schemes. These interventions are frequently led or co-delivered by organizations NGOs and community-based entities, which operate with greater proximity and responsiveness to local needs (Mullins et al., 2012).

The social implications of housing interventions are well documented. Poor housing conditions are associated with increased rates of chronic illness, mental health challenges, and lower educational attainment (Clair & Hughes, 2019; Dunn, 2000). Conversely, targeted interventions such as those addressing dampness, insulation, or overcrowding, have been shown to improve both physical health and psychological well-being (Keans et al., 2010). These outcomes underscore the role of housing interventions as de facto social policies, generating benefits that extend to health, education, employability, and community cohesion (Mullins et al., 2012). NGOs play a particularly significant role in this landscape. Their proximity to vulnerable populations enables them to build trust, mobilize volunteers and donors, and tailor interventions to specific local contexts.

They are often recognized as drivers of social innovation in the housing sector, developing flexible models that respond to gaps left by public provision (Mullins et al., 2012; Svejenova & Christiansen, 2017).

Participatory approaches are central to this role, as they enhance legitimacy and foster a sense of ownership among beneficiaries. International evidence supports this perspective: community housing associations in the United Kingdom and participatory upgrading initiatives in Latin America have demonstrated how collaborative approaches can simultaneously improve physical living conditions and strengthen social networks (Acolin & Wachter, 2017).

The expanding role of NGOs in housing interventions also raises important questions regarding equity and sustainability. While NGOs ideally complement state action, in contexts of welfare retrenchment they may assume responsibilities traditionally held by public institutions. This dynamic creates opportunities for innovation but also risks fragmentation, as NGO-led provision can be uneven across regions and dependent on unstable funding streams (Mullins et al., 2012). These tensions highlight the need to situate NGOs within coherent housing strategies that balance the roles of the state, market, and third sector.

NGOs have pioneered qualitative and mixed-method approaches that capture intangible outcomes, including dignity, empowerment, and social connectedness (Davies & Dart, 2005; Tuan, 2008). These approaches anticipate broader debates on social impact measurement, emphasizing the need for hybrid frameworks that combine quantitative robustness with sensitivity to lived experiences.

In summary, NGOs are key actors in housing interventions, addressing gaps left by public provision while advancing participatory, innovative, and multidimensional approaches. Their contributions extend beyond material improvements to include social innovation and methodological experimentation. However, their growing prominence also raises critical questions about equity, sustainability, and accountability. Understanding this dual role is essential for situating NGOs within the broader field of housing interventions and for framing their contributions to ongoing debates on social impact evaluation.

2.4 Social Impact Measurement

The measurement of social impact has become a central concern in the evaluation of interventions implemented by non-profit and third-sector organisations. The third sector refers to organisations operating outside the public and private spheres, typically including non-governmental, voluntary, and community-based entities whose primary aim is to create social value rather than profit (Alomoto et al., 2022; Mullins et al., 2012). It enables these entities to assess the effectiveness of their actions, justify the allocation of resources, and continuously improve their practices. Unlike traditional performance assessments, which often emphasize efficiency or productivity, social impact measurement (SIM) seeks to capture the broader, multidimensional outcomes that interventions generate in the lives of individuals, households, and communities (Costa & Pesci, 2016).

SIM is the differentiation between outputs, outcomes, and impacts. Outputs refer to the immediate and tangible products of activities, such as the number of houses rehabilitated. Outcomes denote short and medium-term changes experienced by beneficiaries, such as improved health or social integration. Impacts, in turn, represent long-term and structural transformations, including enhanced well-being, empowerment, and community resilience (Costa & Pesci, 2016).

Three major methodological approaches have emerged in the literature:

2.4.1 Cost-Based Approaches

These approaches integrate financial data into impact assessment, allowing the costs of interventions to be related to the benefits generated. Among the most used methods are:

1. Cost-Effectiveness Analysis (CEA): Evaluates the cost per unit of non-monetary result (e.g., cost per beneficiary with improved housing) (Zinovieff & Rotem, 2008);
2. Cost-Benefit Analysis (CBA): Assigns monetary values to social benefits to calculate the return on investment (Zinovieff & Rotem, 2008; Tuan, 2008);
3. Social Return on Investment (SROI): Expands CBA by including stakeholders' perspectives and social, environmental, and economic impacts, expressed in a ratio (e.g., 1:3, meaning 1 euro invested generates 3 euros of social value (Tuan, 2008; Alomoto et al., 2022));

Despite their utility, cost-based approaches face significant challenges, namely the difficulty of assigning monetary values to intangible benefits, such as dignity or psychological well-being. These methods also rely heavily on subjective assumptions, which may compromise the validity of results (Alomoto et al., 2022).

2.4.2 Outcome-Based Approaches

Outcome-based methodologies focus on the changes observed in beneficiaries' lives, without necessarily translating these results into financial terms. They are particularly useful for capturing complex and qualitative social transformations. Prominent models include:

1. Theory of Change: Maps the causal logic between activities, results, and impact, making underlying assumptions explicit (Guidelines and Principles For Social Impact Assessment the Interorganizational Committee on Guidelines and Principles for Social Impact Assessment, 1994);
2. Logic Model: Structures the impact pathway in five stages: inputs → activities → outputs → outcomes → impact (Funnell & Rogers, 2011)
3. Outcome Mapping: Focuses on changes in stakeholders' behaviour, being useful in community development contexts (Davies & Dart, 2005);

4. Most Significant Change (MSC): Collects qualitative narratives from beneficiaries about the most significant changes in their lives. (Davies & Dart, 2005).

These approaches are especially relevant for organisations such as Organization A*,* which aim to evaluate the impact of housing interventions based on the lived experiences of beneficiaries (Social Data Lab, 2024; Just a Change, n.d.). In this context, the Entity B adopted an outcome-based approach, integrating quantitative data (e.g., pre and post-intervention surveys) with qualitative data (e.g., testimonials), in line with the multidimensional nature of the impact being assessed.

2.4.3 Standardized Models and Frameworks

To promote comparability and transparency, standardized measurement frameworks have been developed, such as:

1. IRIS+ (Global Impact Investing Network): A catalogue of social, environmental, and financial indicators (Hand et al., 2020; Li et al., 2022);
2. Global Reporting Initiative (GRI): Widely used sustainability reporting framework;
3. Sustainable Development Goals (SDGs): Alignment with global goals, such as poverty eradication (SDG 1) or decent housing (SDG 11);
4. Social Value International (SVI): Principles for measuring social value with a focus on stakeholders (Ortega Maldonado, 2015).

While these frameworks have contributed to the consolidation of a common language in the sector, they still face challenges in adapting to local contexts and integrating qualitative data (Alomoto et al., 2022; Tuan, 2008).

2.4.4 Methodological Considerations

The choice of impact measurement methodology should be guided by several factors (Gómez et al., 2022; Kročil, 2018):

- The quality and nature of available data (quantitative, qualitative, or both);
- The organization’s capacity to collect and analyse data;
- The degree of stakeholder involvement in the evaluation process;
- The need for triangulation to validate results.

Triangulation may involve: 1) cross-referencing surveys, interviews, and administrative data (data triangulation), 2) combining methods such as SROI, thematic analysis, and inferential statistics (methodological triangulation), and 3) collecting perspectives from beneficiaries, volunteers, partners, and funders (stakeholder triangulation).

2.5 RELEVANT STUDIES AND EMPIRICAL EVIDENCE

The literature reviewed throughout this chapter underscores the multidimensional nature of housing poverty and its far-reaching implications for health, social integration, and overall well-being. Housing is not merely a physical asset but a critical determinant of social and psychological outcomes (Dunn, 2000; Stone & Hulse, 2007; Clair & Hughes, 2019). Empirical studies have consistently demonstrated that inadequate housing conditions are associated with a range of negative effects, including poor physical health, mental distress, and diminished life satisfaction.

Their research highlights how substandard housing environments, characterized by dampness, overcrowding, and poor insulation can contribute to respiratory illnesses, psychological stress, and reduced quality of life. These findings reinforce the argument that housing deprivation must be understood as a public health issue, not solely a matter of infrastructure. Beyond health, housing also plays a central role in shaping social cohesion. (Stone & Hulse, 2007), emphasize that housing poverty restricts access to essential services such as education and employment, thereby perpetuating cycles of exclusion. (Keans et al., 2010), further demonstrate that improvements in housing conditions can yield psychosocial benefits, including enhanced feelings of security, dignity, and community belonging. These studies collectively affirm that housing interventions have the potential to generate transformative social outcomes. Robust evidence of the relationship between housing conditions and health outcomes has been provided by previous studies (Dunn, 2000; Clair & Hughes, 2019)

Non-governmental organisations (NGOs) and third-sector entities have emerged as key actors in addressing housing poverty through innovative and context-sensitive interventions. According to (Gómez et al., 2022; Mullins, Czischke & van Bortel, 2012), NGOs' proximity to vulnerable communities enables them to design and implement socially innovative responses that are tailored to local needs. However, despite their relevance, evaluations of NGO-led housing interventions often rely on narrow output measures, such as the number of units rehabilitated failing to capture the broader social transformations they produce (Costa & Pesci, 2016; Alomoto et al., 2022).

This limitation underscores the importance of robust social impact measurement (SIM). The literature illustrates how SIM has evolved from accountability-driven frameworks to more participatory and multidimensional models. (Costa & Pesci, 2016), stress the role of stakeholder engagement in ensuring that metrics are meaningful and context-sensitive. (Davies & Dart, 2005; Tuan, 2008), propose qualitative approaches that capture intangible outcomes through narratives and lived experiences. These methods reflect a growing consensus that hybrid models combining quantitative and qualitative data are essential for capturing the full spectrum of social impact.

Despite these methodological advances, several gaps persist. First, empirical studies on housing interventions continue to prioritize measurable outputs, with limited attention to long-term outcomes such as empowerment, social mobility, or community resilience. Second, while the association between housing and health is well established (Dunn, 2000; Clair & Hughes, 2019), the causal mechanisms underlying these relationships remain underexplored, partly due to the scarcity of longitudinal data and robust counterfactual designs. Third, although the integration of qualitative and quantitative approaches is increasingly recommended (Costa & Pesci, 2016; Gómez et al., 2022), its implementation remains uneven, constrained by methodological complexity and limited organizational resources.

Housing poverty emerges as a multidimensional phenomenon deeply intertwined with structural inequalities and relational dynamics (Stone & Hulse, 2007; Clair & Hughes, 2019). This complexity not only challenges simplistic evaluation models but also reveals the necessity of hybrid approaches that integrate quantitative rigor with qualitative depth (Costa & Pesci, 2016; Kročil, 2018). Therefore, the methodological choices outlined in the next chapter respond directly to these gaps, operationalizing a mixed-methods design aligned with contemporary best practices (Alomoto et al., 2022; Gómez et al., 2022).

In sum, prior research underscores the complexity of housing poverty and the methodological challenges of measuring social impact in nonprofit contexts. While hybrid approaches combining quantitative and qualitative data are increasingly advocated, empirical applications remain scarce. These insights justify the mixed-methods design adopted in this study and inform the selection of indicators and analytical strategies presented in the methodology chapter.

3. METHODOLOGY

This dissertation aims to critically assess the methodology developed by the Entity B for measuring the social impact of housing rehabilitation initiatives implemented by the non-profit organization. The evaluation process is grounded in current academic debates on the effectiveness and contextual relevance of impact assessment tools used in the third sector, with particular emphasis on their ability to capture the multidimensional nature of social change (Costa& Pesci, 2026; Alomoto et al., 2022).

Organization A is a non-profit organization operating under the motto "*Reabilitamos casas, reconstruímos vidas*," and is recognized for its dual mission: improving the physical conditions of housing while fostering broader social transformation. As its operations have expanded, the need to demonstrate the effectiveness of its interventions has become increasingly salient not only for internal learning and strategic development, but also to meet the expectations of external stakeholders, including funders, partners, and public institutions (Gómez et al., 2022; Epstein & Roy, 2001). The organization relies significantly on external investment, both public and private, which further underscores the importance of robust impact evaluation (Social Data Lab, 2024).

Rather than proposing a new impact measurement framework, this dissertation focuses on evaluating the effectiveness of an existing model in capturing social impact. The model seeks to quantify changes in beneficiaries' lives across multiple dimensions: physical, psychological, social, and financial, drawing on data collected through structured surveys and interviews. (Social Data Lab, 2024). The evaluation adopts a mixed-methods approach, integrating statistical analysis with qualitative interpretation, in alignment with best practices in social impact assessment (Tuan, 2008; Schröer et al., 2021).

To guide the analytical process, the Cross Industry Standard Process for Data Mining (CRISP-DM) methodology was selected. The following sections detail each phase within the context of this evaluation process. By evaluating this model, the dissertation contributes to the development of more effective and replicable methodologies for social impact assessment in non-profit housing interventions, offering insights that may inform future evaluations.

3.1 Research Hypotheses

Based on the multidimensional structure of the impact model, the following hypotheses guide the empirical analysis:

1. H1 (Physical): Improvements in housing conditions and safety are positively associated with beneficiaries' perceived overall impact;
2. H2 (Personal): Gains in emotional stability and self-confidence are positively associated with perceived overall impact;
3. H3 (Social): Increases in social integration and frequency of interactions are positively associated with perceived overall impact;
4. H4 (Financial): Improvements in perceived economic security are positively associated with perceived overall impact.

3.2 CRISP-DM

The CRISP-DM methodology is one of the most widely used methodologies for data mining and analytics projects. It provides a structured approach to handling data science projects, ensuring consistency and efficiency. The methodology consists of six iterative phases (Schröer et al., 2021): 1) Context Understanding: Understanding the business objectives and defining the problem; 2) Data Understanding: Collecting initial data, exploring it, and assessing its quality; 3) Data Preparation: Cleaning, transforming, and structuring the data for analysis; 4) Modelling: Applying various modelling techniques and tuning parameters; 5) Evaluation: Assessing model performance and determining if it meets business objectives; 6) Deployment: Implementing the model in a real-world scenario.

By leveraging the CRISP-DM framework, this study ensures a structured and systematic approach to impact measurement. The integration of both qualitative and quantitative data enhances the credibility of the results, providing a comprehensive understanding of the Organization A programme's impact (Daw & Hatfield, 2018; Ncr& Clinton, 1999).

3.2.1 Phase 1: Context Understanding

The Context Understanding phase involves clarifying the analytical objectives and aligning them with the organizational context and the goals of the dissertation. In this case, the primary aim is to validate the impact measurement model developed by the Entity B, which is currently used to assess the outcomes of housing rehabilitation interventions carried out by Organization A

As detailed in Chapter above, Organization A operates in contexts of housing deprivation and social exclusion. This section focuses on the methodological implications of evaluating its impact (Dunn, 2000; Clair & Hughes, 2019). The organization's mission is also discussed above.

In this chapter, I emphasize the need for an impact model capable of capturing both tangible and intangible outcomes. Therefore, any impact measurement tool must be capable of capturing both tangible improvements and intangible transformations in beneficiaries' lives (Gómez et al., 2022). The study is guided by two central research questions, which I come back to in Chapter 4 during the empirical analysis. These research questions articulate the dual objective of this dissertation: first, to evaluate the methodological integrity of the selected impact measurement model; and second, to determine its effectiveness in capturing the changes reported by the target population Indicators were collaboratively defined with stakeholders and multiple dimensions of well-being. A detailed operationalization is presented in Chapter 4 (Social Data Lab, 2024). This section reflects a holistic understanding of well-being, as advocated by multidimensional impact frameworks (Alomoto et al., 2022; Gómez et al., 2022).

By establishing a clear understanding of the organizational context and the analytical objectives, this phase lays the foundation for the next stages of the CRISP-DM process. It ensures that the evaluation process remains aligned with both theoretical principles and practical requirements, contributing to the refinement of impact assessment practices in non-profit housing intervention (Schröer et al., 2021). These methodological considerations are guided by two central research questions: (1) How valid and robust is the statistical model developed by the Entity B for measuring the social impact of Organization A*'s* interventions?, and (2) To what extent is this model effective in capturing the changes reported by the target population? These questions structure the empirical analysis presented in the following chapters and ensure that the model evaluation remains aligned with the theoretical and practical objectives of this dissertation.

3.2.2 Phase 2: Data Understanding

The Data Understanding phase involves a systematic exploration of the data sources and their alignment with the analytical objectives of the study (Chapman et al., 2000). In the context of social impact assessment, this phase is critical for ensuring that the data collected are suitable for capturing multidimensional changes in beneficiaries' lives (Schröer et al., 2021).

This study adopts a mixed-methods approach, combining structured quantitative data with qualitative insights. The dependent variable is conceptualised as a measure of perceived impact, reflecting the beneficiary's subjective assessments of change following the intervention. This aligns evaluation models, which prioritise the lived experiences of participants as valid and meaningful indicators of social transformation (Davies & Dart, 2005; Tuan, 2008).

Independent variables are grouped into thematic dimensions: physical, personal, social, and financial - following established frameworks for multidimensional well-being (Costa & Pesci, 2016; Gómez et al., 2022).

These dimensions are then operationalized through indicators that capture both tangible and intangible aspects of change, such as housing conditions, emotional stability, social integration, and economic security.

This phase also involves preliminary data exploration to assess structure, consistency, and relevance. This includes identifying potential sources of bias, evaluating the completeness of the dataset, and ensuring that the selected variables are theoretically grounded and analytically appropriate. Such preparatory work is essential for maintaining the integrity of subsequent modelling efforts and for supporting triangulation across data types (Schröer et al., 2021).

3.2.3 Phase 3: Data Preparation

The Data Preparation phase focuses on refining and structuring the dataset to ensure its analytical appropriateness. This involves a series of methodological procedures aimed at enhancing data quality and coherence, thereby enabling robust statistical and interpretative analysis (Schröer et al., 2021). Quantitative data are typically subject to standard preprocessing steps, including verification of scale consistency, normalization of variables, and formatting adjustments to ensure compatibility with the chosen modelling techniques. These procedures are guided by the requirements of inferential statistical methods, such as linear regression, which assume specific data properties for valid estimation (Su et al., 2012).

Qualitative data, often derived from semi-structured interviews or open-ended survey responses, are organised thematically according to the analytical dimensions defined in the study. While formal coding may not be applied in all cases, thematic grouping facilitates integration with quantitative findings and supports methodological triangulation (Costa & Pesci, 2016).

In studies adopting composite indicators such as Depth, Perpetuity, Peculiarity, and Severity (Social Data Lab, 2024). Data preparation also involves aligning raw inputs with the conceptual structure of the impact model. This requires ensuring that each variable corresponds accurately to its intended construct and that the dataset supports the calculation of multidimensional metrics. Such alignment is crucial for preserving the validity of the impact assessment and for enabling meaningful interpretation of results.

3.2.3.1 Qualitative Data Preprocessing

All open-ended responses originally collected in Portuguese were translated into English using an automated pipeline based on OpenAI's language models. This step ensured compatibility with the NRC Emotion Lexicon, which is optimized for English and widely validated in sentiment analysis research (Mohammad & Turney, 2013). The choice was driven by two factors: (1) higher coverage and accuracy compared to Portuguese equivalents, and, (2) alignment with international standards for replicability. To reduce semantic distortion, a random sample of translations was manually validated and emotion-related terms were checked for consistency.

The preprocessing pipeline included three main stages:

1. Compilation and Cleaning of Textual Data: All responses were consolidated into a unified dataset. Preprocessing involved normalization of text, tokenization, and removal of stop words to eliminate non-informative elements. These steps are essential to reduce noise and improve the accuracy of subsequent classification tasks, as recommended by text mining best practices (Silge & Robinson, 2017).

2. Lemmatization: To preserve semantic meaning while reducing morphological variations, lemmatization was applied to the text. This process ensures that words are analysed in their base form, facilitating accurate matching with sentiment lexicons. The decision to apply lemmatization follows recommendations in Natural Language Processing literature, which highlight its role in improving the reliability of sentiment detection (Mohammad & Turney, 2013; González Canché & Zheng, 2025).

3. Lexicon-Based Mapping: After normalization, the NRC Emotion Lexicon was used to identify polarity (positive, negative) and emotional categories within the text. This lexicon-based approach provides a replicable and transparent method for sentiment classification, aligning with established practices in mixed-methods research (Mohammad & Turney, 2013; Suzara, Peña, & Botham, 2025).

These procedures were essential for ensuring that qualitative data could be systematically analysed using computational techniques, supporting the integration of sentiment analysis into the empirical study.

3.2.4 Phase 4: Modelling

The Modelling phase focuses on selecting and applying analytical techniques that allow for the validation of the impact assessment framework (Chapman et al., 2000). In this study, the modelling approach is primarily inferential and explanatory. Rather than aiming to predict outcomes, the objective is to understand how different dimensions of the intervention contribute to the perceived impact reported by beneficiaries (Alomoto et al.,2022). This approach supports the validation of the internal logic of the impact model and enables the identification of the most influential variables.

The modelling process is grounded in supervised learning, as the dependent variable "perceived impact" is known and measured through structured data collection (Su et al., 2012). Among the available techniques, a multiple linear regression is selected due to its interpretability, robustness, and widespread use in social impact research (Su et al., 2012). This method allows for the estimation of the relationship between a continuous outcome and multiple explanatory variables, facilitating the analysis of multidimensional effects. To ensure methodological rigour, standard diagnostic procedures are considered prior to model estimation. These include the verification of assumptions such as linearity, homoscedasticity, and absence of multicollinearity. Additionally, complementary techniques such as sentiment analysis may be employed to enrich the interpretation of qualitative data, particularly in capturing emotional and perceptual dimensions of change (Schröer et al., 2021; Alomoto et al., 2022)

The table below summarises the alignment between the study's sub-objectives and the analytical techniques selected for each:

Table 1 - Alignment between the study's sub-objectives and the analytical techniques

Sub-Objective	Analytical Technique	Purpose
Assess the relationship between intervention dimensions and perceived impact	Multiple Linear Regression (using Python with sklearn lib)	Inference and explainability
Explore emotional and perceptual responses to the intervention	Sentiment Analysis (using Python + nltk lib+ NRC file)	Complementary qualitative insight
Evaluate the internal coherence of the impact model	Model diagnostics and triangulation	Validation of methodological logic

Alternative modelling strategies, such as Difference-in-Differences, may be considered in contexts where baseline data or control groups are available (Daw & Hatfield, 2018). However, the viability of such approaches depends on the structure and completeness of the dataset, and their applicability is discussed in the empirical chapter.

The modelling phase, as outlined here, provides the analytical foundation for the evaluation of the impact model. Its operationalisation and empirical application are presented in the following chapter, where the techniques described are applied to the case of Organization A.

3.2.5 Phase 5: Evaluation

The Evaluation phase is dedicated to assessing the methodological soundness and contextual relevance of the analytical models applied (Chapman et al., 2000). In the context of social impact assessment, this phase does not merely verify technical adequacy but also ensures that the chosen modelling strategies align with the theoretical framework and the objectives of the intervention under study (Schröer et al.,2021).

This phase is particularly relevant to the overarching aim of the dissertation: to validate the impact measurement model developed by the Entity B in the context of housing rehabilitation interventions carried out by Organization A. As such, evaluation serves as a mechanism to determine whether the model is capable of capturing multidimensional social change in a manner that is both empirically grounded and theoretically coherent.

From a methodological standpoint, evaluation involves a critical appraisal of the model's internal logic, its capacity to reflect stakeholder-defined indicators, and its alignment with best practices in social impact measurement. This process is especially important in third-sector contexts, where impact is often intangible, subjective, and distributed across various domains of well-being (Alomoto et al., 2022; Costa & Pesci, 2016).

To ensure robustness, the evaluation phase integrates several complementary strategies:

1. Theoretical Validation: The model is assessed against established frameworks such as Theory of Change, Logic Models, and multidimensional impact assessment approaches. This ensures that the analytical logic reflects causal pathways and captures the complexity of social transformation (Tuan, 2008; Funnell & Rogers, 2011).

2. Triangulation of Methods and Sources: Methodological triangulation is employed to enhance credibility, combining quantitative modelling techniques with qualitative insights derived from interviews and testimonials. This approach is consistent with participatory evaluation principles, which emphasize the importance of integrating lived experiences into impact measurement (Schröer et al., 2021; Davies & Dart, 2005).

3. Benchmarking: Where feasible, the model is compared with similar impact assessment tools used in the third sector. Although direct benchmarking may be limited by contextual differences, structural comparisons with models such as Social Return on Investment (SROI), Most Significant Change (MSC), and SDG-aligned frameworks allow for the identification of strengths and limitations in the model's design and application (Gómez et al., 2022; Tuan, 2008).

4. Assessment of Limitations and Mitigation Strategies: A critical component of this phase involves acknowledging and addressing methodological constraints. One of the most significant limitations is the absence of baseline data collected prior to intervention, which restricts the use of designs such as Difference-in-Differences (Daw & Hatfield, 2018). In this study, impact is assessed through post-intervention data only, based on beneficiaries' self-reported perceptions and qualitative accounts. While this approach does not allow for direct comparison with pre-intervention conditions, it enables a multidimensional understanding of perceived change, supported by triangulation across data sources. Additionally, the reliance on self-reported data introduces potential biases, particularly in emotionally sensitive contexts. These are addressed through careful questionnaire design, interviewer training, and the integration of multiple data sources to enhance validity.

At last, the evaluation phase serves as a bridge between modelling and deployment, ensuring that the findings generated are not only statistically valid but also meaningful and actionable. By grounding the assessment in both theoretical rigour and contextual sensitivity, this phase contributes to the refinement of impact measurement practices and supports the development of more robust methodologies for evaluating social interventions.

3.2.6 Phase 6: Deployment

The Deployment phase refers to the translation of analytical outputs into actionable insights and communicable results (Chapman et al., 2000). In the context of social impact assessment, deployment does not imply the implementation of automated systems or predictive algorithms, but rather the structured dissemination of findings to relevant stakeholders, including funders, public institutions, and community partners (Schröer et al., 2021).

The primary objective of this phase is to ensure that the validated results of the modelling and evaluation stages are effectively communicated and utilised (Chapman et al., 2000). This involves the preparation of comprehensive impact reports that integrate both quantitative indicators and qualitative insights, allowing for a multidimensional interpretation of the intervention's effects. These reports are designed to be accessible to diverse audiences and may include data visualisations, thematic summaries, and evidence-based recommendations. Deployment also serves a strategic function within the organisation.

By providing a clear picture of the outcomes achieved, the findings support internal decision-making processes, inform programmatic adjustments, and contribute to long-term planning. Moreover, they enhance external accountability, offering funders and partners a transparent account of the value generated through the intervention. This aligns with broader trends in the third sector, where impact reporting is increasingly seen as a tool for legitimacy, learning, and stakeholder engagement (Salido-Andres et al., 2022; Epstein & Roy, 2001).

By documenting the methodological process and the structure of the impact model, the study facilitates future applications in similar contexts. This is particularly relevant for organisations seeking to expand their operations or adapt their evaluation frameworks to new settings. The deployment of findings thus extends beyond the immediate scope of the intervention, offering a foundation for methodological innovation and sector-wide learning.

It is important to note, however, that the interpretation and use of results must be approached with caution. As discussed in the evaluation phase, the absence of baseline data and the reliance on self-reported perceptions introduce limitations that constrain causal inference. Deployment must therefore include appropriate caveats and contextualisation, ensuring that conclusions are presented transparently and ethically.

As noted by (Tuan ,2008), results should be communicated in ways that match stakeholders' needs and capacities, ensuring they are both technically accurate and socially meaningful (Gómez et al., 2022). Returning findings to those directly affected reinforces co-responsibility and mutual learning. By presenting results in an accessible and context-sensitive format, this phase promotes a more democratic and accountable approach to social impact assessment. By presenting findings in a format that is both accessible and context-sensitive, the deployment phase contributes to a more democratic and accountable approach to social impact assessment.

Taken together, the methodological choices outlined in this chapter provide the foundation for a rigorous and context-aware evaluation of the impact measurement model. The following chapter builds on this framework by applying it to the empirical case of Organization A*,* detailing the data collected, the analytical procedures undertaken, and the insights generated through the evaluation of its housing rehabilitation interventions.

4. EMPIRICAL STUDY

4.1 DESCRIPTION OF THE SLI2 MODEL

This chapter presents the empirical application of the impact assessment model, detailing the data collection process, sample characteristics, and analytical procedures. The aim is to provide a comprehensive description of the empirical study that underpins the evaluation of the Entity B Lab Impact Index (SLI2) model's effectiveness (Social Data Lab, 2024).

The SLI2 is a multidimensional framework designed to measure direct social impact at the individual level. It integrates principles from Theory of Change (Funnell & Rogers, 2011), and Logic Models (Knowlton & Phillips, 2012) to establish causal pathways between interventions and outcomes. Additionally, it applies corrective factors such as: Perpetuity, Peculiarity, and Severity, to account for sustainability, uniqueness, and initial vulnerability, reflecting best practices in social impact measurement (Costa & Pesci, 2016).

Impact is assessed across four dimensions: physical, personal, social, and financial. Each dimension includes indicators capturing tangible and intangible changes, such as housing conditions, emotional well-being, social integration, and economic security (Social Data Lab, 2024).

The model expresses results through two complementary metrics:

1. Relative Impact: The percentage change perceived by beneficiaries
2. Units of Complete Impact (UIC): An estimate of the number of individuals whose lives were fully transformed.

Data collection combines structured questionnaires and qualitative interviews, supporting a mixed-methods approach recommended for robust social impact assessment (Costa & Pesci, 2016). This empirical study applies the SLI2 model exclusively to direct effects on beneficiaries, providing the basis for the statistical evaluation presented in this chapter.

4.1.1 Structure and Dimensions of the SLI2 Index

The impact measurement is structured around several key dimensions that quantify changes in different aspects of the beneficiaries' and volunteers' lives. These dimensions include

1. **Depth**

Depth measures the extent to which the intervention has influenced the lives of beneficiaries and volunteers. This is expressed as a percentage (ranging from 0% to 100%), representing the estimated change in an individual's life due to the program (Social Data Lab, 2024).

To ensure a comprehensive evaluation, depth is divided into specific components, which vary between beneficiaries and volunteers:

For Beneficiaries:

- General Well-being: Overall perception of life improvements.
- Physical Impact: Includes health and safety improvements, changes in housing conditions, better hygiene practices, and enhanced thermal comfort.
- Psychological Impact: Measures changes in self-confidence, motivation, happiness, and emotional stability.
- Social Integration: Assesses improvements in family relationships, frequency of social interactions, and feelings of social inclusion.
- Professional Development: Evaluates changes in employability, motivation to work or study, and acquisition of useful skills.

The depth value is calculated as the arithmetic mean of responses collected through interviews and questionnaires.

2. **Perpetuity**

This indicator estimates how long the impact of the intervention is expected to last in the lives of beneficiaries. It is measured through responses to questions like, "Do you think these changes will still be present in your life in five years?" The responses provide an estimate of how sustainable the benefits of the intervention are over time (Social Data Lab, 2024).

Based on a broader study conducted by Entity B, this indicator uses benchmarks from prior research on social impact durability to improve the reliability of estimates.

3. **Peculiarity**

This factor assesses the uniqueness of the support provided by Organization A, particularly in comparison to other available aid programs. A higher peculiarity score indicates that the services provided by Organization A are difficult to find elsewhere, which strengthens the program's significance (Social Data Lab, 2024)

4. **Severity**

This metric considers the initial conditions of beneficiaries before the intervention took place. Individuals who were in more severe conditions prior to receiving support are expected to experience a greater perceived impact, which is taken into account when analysing overall impact scores (Social Data Lab, 2024).

Additionally, the methodology incorporates a “Deve e Haver” (Balance of Give and Take) component for volunteers, ensuring that the perceived balance between what they contribute and what they receive is factored into the analysis.

4.1.2 Impact Measure Access

These dimensions are calculated using a combination of qualitative and quantitative data, collected through structured interviews and questionnaires (Social Data Lab, 2024). The model also introduces the concept of Units of Complete Impact (UIC), which estimates the number of individuals whose lives were fully transformed by the intervention. This metric facilitates benchmarking across programs and institutions (Social Data Lab, 2024).

The SLI2 model aligns with international standards for social impact evaluation, including those proposed by the United Nations Development Programme (UNDP) and the European Commission. Its emphasis on direct, individual-level impact ensures clarity in attribution and avoids the dilution of results through indirect or systemic effects.

The impact measurement process follows a structured calculation method. First the relative impact is measured and the absolute impact. The Relative Impact represents the percentage change that beneficiaries and volunteers feel has occurred in their lives due to the program. The Absolute Impact value is calculated to allow comparison with other projects and institutions. This is done using the concept of Units of Complete Impact (UIC), which estimates the number of individuals whose lives were completely transformed by the program.

The Units of Complete Impact (UIC) metric estimates the number of individuals whose lives were fully transformed by the intervention. To enhance precision, the raw UIC value is adjusted using three corrective factors: Perpetuity (P), Peculiarity (Q), and Severity (S).

Definitions:

1. Perpetuity (P): Expected duration of the impact (scale 0–1), reflecting sustainability over time.
2. Peculiarity (Q): Degree of uniqueness of the support provided by Organization A compared to other programs (scale 0–1).
3. Severity (S): Initial vulnerability level of beneficiaries before intervention (scale 0–1).

1. Complete Formula - *UIC_adjusted = (Impact Relative×N) × (P×Q×S)*

Example Calculation, with: Impact Relative = 0.61; N = 165; and P = 0.85, Q = 0.90, S = 0.95

2. *UIC_adjusted = (0.61×165) × (0.85×0.90×0.95)*

3. *UIC_adjusted = 100,65×0.72675 ≈ 73.15*

This adjustment ensures that the UIC metric accounts for both the intensity and sustainability of change, as well as the uniqueness of the intervention and the vulnerability of the target population. Sensitivity Note: Small variations in corrective factors can significantly influence the adjusted UIC value. Therefore, benchmarks for P, Q, and S should be based on empirical evidence and periodically reviewed to maintain methodological robustness.

The dashboard system used by Entity B ensures dynamic and interactive exploration of data, allowing for real-time updates and comparisons. By using this structured approach, the methodology ensures that the impact of the Organization A program is assessed in a rigorous and transparent manner, providing reliable insights into its effectiveness and sustainability.

The SLI2 model aligns with international standards for social impact evaluation, including those proposed by the United Nations Development Programme (2020) and the European Commission (2014). Its emphasis on direct, individual-level impact ensures clarity in attribution and avoids the dilution of results through indirect or systemic effects.

Furthermore, the model's reliance on self-reported data is supported by participatory evaluation literature, which highlights the importance of capturing beneficiaries' perspectives to understand meaningful change (Davies & Dart, 2005; Schröer et al., 2021).

4.2 METHOD OF DATA COLLECTION

To assess the social impact of the Organization A program using the SLI2, data collection focused exclusively on the beneficiaries of the intervention. This decision reflects the primary objective of the study. Understand the direct effects of housing rehabilitation on individuals living in vulnerable conditions.

Data were collected through semi-structured face-to-face interviews, conducted in the homes of the beneficiaries. This method was chosen for its capacity to generate rich, contextualized insights into the lived experiences of participants. Semi-structured interviews allow for a balance between consistency across respondents and flexibility to explore individual narratives (Cunha et al., 2018; Zinovieff & Rotem, 2008). In previous years, these interviews were conducted by volunteers from Organization A, who visited beneficiaries' homes to collect data. While this approach fostered proximity and trust, it also carried the risk of response bias, as beneficiaries might have felt inhibited in expressing criticism or sensitive feedback.

In the current study, interviews were conducted by external professionals, independent from Organization A. This methodological shift is aligned with best practices in social impact evaluation and aims to enhance the neutrality and credibility of the data. The presence of an external interviewer helps ensure that beneficiaries feel free to express their views openly, without concern for how their responses might be perceived by the organization.

Maintaining the interview setting in the beneficiaries' homes remains a deliberate choice. It provides a familiar and safe environment, which facilitates openness and comfort during the conversation (Svejenova & Christiansen, 2017). This setting also allows interviewers to capture non-verbal cues and contextual elements that enrich the qualitative data, ensuring that beneficiaries' voices are fully heard and respected.

This approach is particularly effective when working with socially vulnerable populations, as they help build rapport and trust, reduce misunderstandings, and allow for the observation of non-verbal cues (Davies & Dart, 2005; Svejenova & Christiansen, 2017). These advantages are crucial when assessing subjective dimensions such as emotional well-being, perceived safety, and social integration. The exclusive use of interviews with beneficiaries aligns with best practices in participatory evaluation, which emphasize the importance of capturing the perspectives of those directly affected by social interventions (Tuan, 2008; Schröer et al., 2021). By focusing on beneficiaries, the study ensures that the impact assessment reflects the voices of those who experienced the intervention firsthand.

Moreover, qualitative methods are particularly suited to multidimensional impact models such as SLI2, which require nuanced data across various life domains, such as physical, personal, social, and financial (Alomoto et al., 2022; Gómez et al., 2022). The depth and richness of interview data support the accurate calculation of indicators such as Depth, Perpetuity, are central to the SLI2 framework.

4.3 Data Description and Characterisation of the Sample

Organization A intervenes in the housing conditions of its beneficiaries, meaning that the impact extends not only to the initially identified individual but also to their entire household (Just a Change, n.d.). The target population of the program consists of individuals and families living in situations of housing poverty and social vulnerability, often facing multiple and overlapping disadvantages such as low income, poor health, limited access to services, and social isolation (Gómez et al., 2022). This population is characterised by structural fragilities that compromise their ability to participate fully in society, making them particularly susceptible to exclusion and marginalisation (Gómez et al., 2022; Stone & Hulse, 2007).

The sample universe comprises 85 primary beneficiaries with whom Organization A worked during the year 2023. However, these 85 individuals represent only the main points of contact and are not the sole recipients of the intervention within their households. In reality, these 85 primary contacts correspond to a total of 165 beneficiaries.

The sampling method adopted involved the random selection of 45 beneficiaries, based solely on the 85 individuals identified as the primary contacts. Nevertheless, due to external circumstances, not all selected individuals responded to the questionnaire, either because they were unavailable or had passed away. Ultimately, considering both the primary contacts and the household members present at the time of the interview, the final sample comprises 53 beneficiaries.

It is important to emphasise that the primary contact is not always the most vulnerable individual within the household. This introduces a bias in the assessment of actual vulnerability in relation to other characteristics, which is not accounted for in the statistical model employed.

The beneficiary population is relatively homogeneous, particularly in socioeconomic terms. The vast majority live in poverty and under precarious living conditions. This homogeneity facilitates the analysis, as the behaviours of these individuals tend to be similar, thereby reducing statistical error.

Prior to analysis, the data were stratified based on five variables: (1) date of birth; (2) district and municipality of residence; (3) gender; (4) complexity of the intervention; and (5) social vulnerability. The term, Complexity of Intervention refers to the scope and intensity of the rehabilitation works carried out in each household.

According to the Entity B framework, interventions are classified as medium or high complexity based on factors such as the number of structural elements addressed (e.g., roofing, plumbing, insulation), the duration of the works, and the logistical challenges involved (Social Data Lab, 2024).

Social vulnerability is understood here as the degree to which individuals or groups are exposed to conditions that limit their access to rights, resources, and opportunities, increasing their susceptibility to harm and exclusion. This includes economic deprivation, lack of social support, and limited access to housing, education, and healthcare (Gómez et al., 2022; Costa & Pesci, 2016).

The sample is characterised as follows: 56% of the population is female, while 44% is male. Ages range from 18 to 84 years, with the average age of beneficiaries being 56. In terms of social vulnerability, based on corrected values from the Entity B report, 20% are considered to have low vulnerability, 40% are classified as vulnerable, and 40% as highly vulnerable. Regarding the complexity of the intervention, 47% of the beneficiaries' homes underwent modifications of medium complexity, while 53% were subject to highly complex interventions (Social Data Lab, 2024).

The sample includes individuals from various regions of the country, with the majority residing in the Lisboa and Vale do Tejo, particularly in the municipalities of Lisboa, Torres Vedras, Óbidos, Almeirim, Golegã, Chamusca, and Ferreira do Zêzere. Of the respondents, 74% were the primary contact, while the remaining 26% were household members who had directly benefited from Organization A's intervention (Social Data Lab, 2024).

The average age distribution indicates a relatively young population, with the majority falling within the 35 to 54 age brackets. Additionally, 21% of beneficiaries lived alone, approximately 43% lived with one other person, and the remainder lived in households comprising two or more individuals (Social Data Lab, 2024).

Figure 1 - Gender distribution of beneficiaries

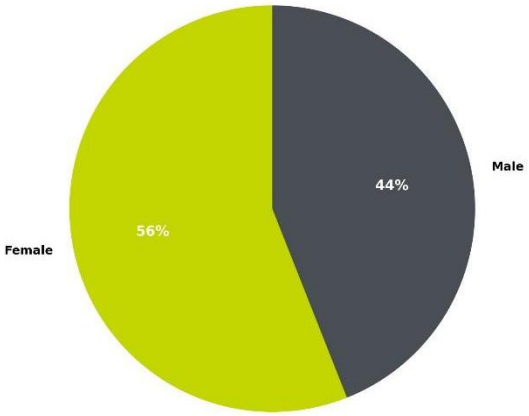


Figure 2 - Age distribution

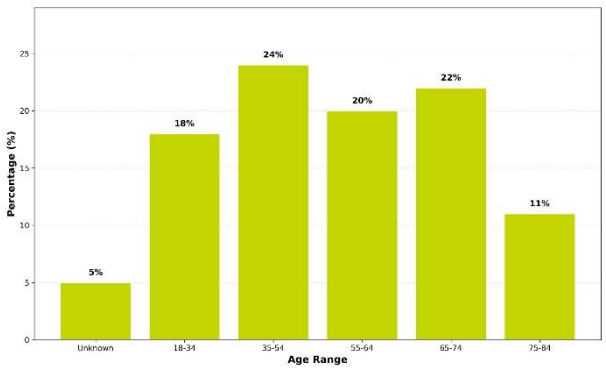


Figure 3 - Social vulnerability levels

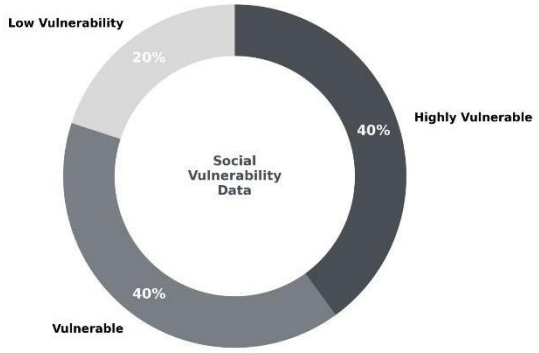


Figure 4 - Intervention complexity

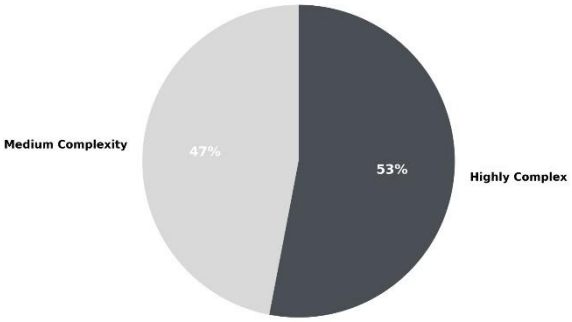
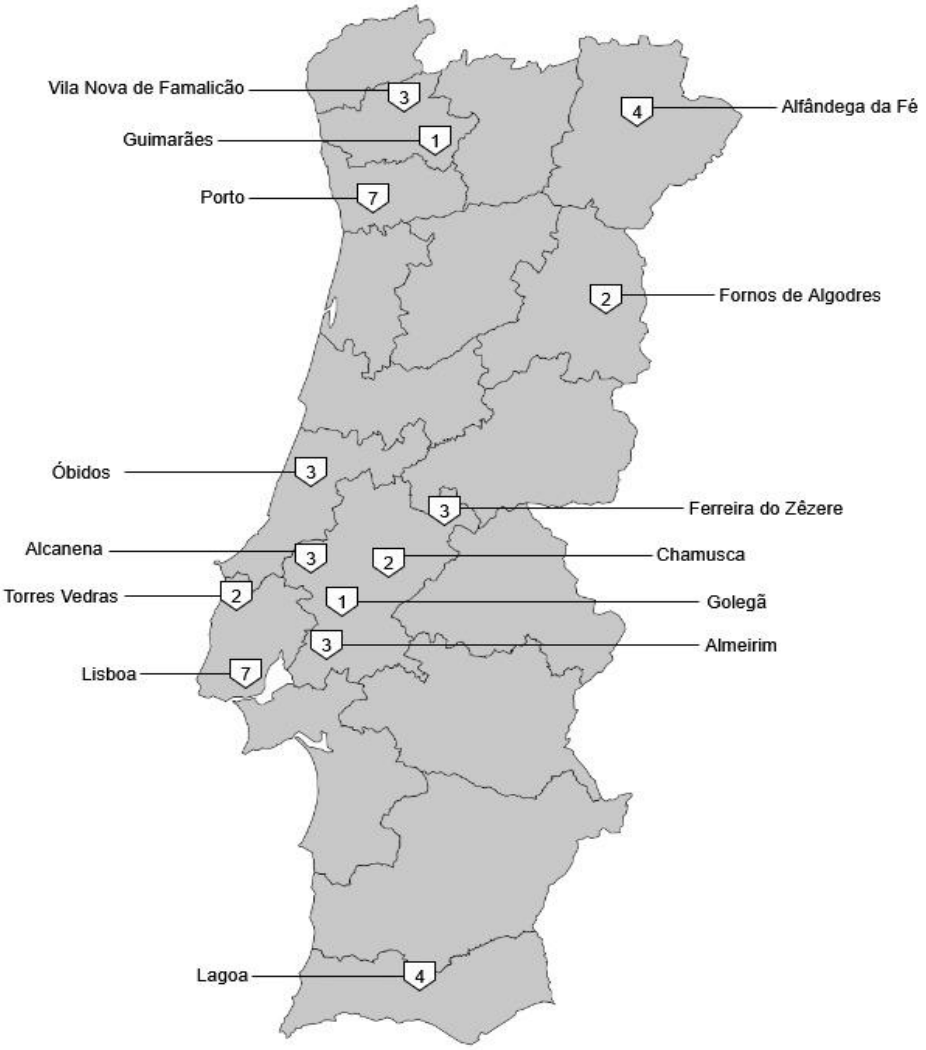


Figure 5 - Beneficiaries' geographic distribution



4.4 DATA PROCESSING

For the purposes of this analysis, I chose to group the 31 questions from the questionnaires into five distinct dimensions, based on the aspects of life they relate to in the context of the beneficiaries. The questions are categorised as follows: (1) Physical Dimension; (2) Personal Dimension; (3) Social Dimension; (4) Professional Dimension and (5) Financial Dimension.

However, for this thesis, only four dimensions were considered, with the Professional Dimension excluded. This decision was made because not all respondents were working or studying at the time of the interview, and the questions within this dimension were directed exclusively at individuals engaged in employment or education. To avoid introducing bias into the analysis, this final dimension was therefore omitted from the present study.

The data processing phase involved both quantitative and qualitative preparation steps. While the detailed procedures for qualitative preprocessing are described in Section 3.1.3.1, it is important to note that these steps ensured the consistency and reliability of the sentiment analysis presented later in this chapter (Schröer et al., 2021). This integration of structured and unstructured data supports the mixed-methods approach adopted in the empirical study (Gómez et al.,2022).

Table 2 – Variables per dimension

Physical Dimension	House conditions; Health; Hygiene; Comfort in the house; House temperature; Security; Protection from the outside
Personal Dimension	Self-Confident; Proud; Confiance in others; Life meaning; Happiness; Willing to work; Motivation
Social Dimension	Family relation; Interaction with others; Capable of inviting people home; Self ware;
Financial Dimension	Personal finances management; utility Expenses (water, electricity or gas)

4.5 Application of the Model

4.5.1 Multi Linear Regression

To analyse the data collected from beneficiaries, a multiple linear regression model was applied. This model was chosen because it allows us to understand how different dimensions of the intervention contribute to the overall perceived impact (Su et al., 2012). Rather than predicting future outcomes, the goal is to explain the variation in responses based on the physical, personal, social and financial aspects of change (Alomoto et al., 2022).

This choice is aligned with the literature reviewed in Chapter 2, which highlights the importance of combining quantitative and qualitative methods to assess social impact. Multiple linear regression is a well-established technique in social research and supports the methodological triangulation discussed earlier (Su et al., 2012).

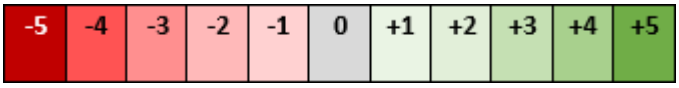
It also fits within the broader logic of the Theory of Change and the multidimensional frameworks presented in the Entity B report (Funnell & Rogers, 2011).

Finally, the use of this model is coherent with the overall objective of the dissertation: to validate the internal structure of an existing impact measurement model. By applying a transparent and interpretable statistical method, it becomes possible to assess whether the model reflects the reality reported by beneficiaries and to identify which dimensions have the greatest influence on perceived change.

The first question in the questionnaires is: “To what extent do you believe that Organization A’s intervention in your home has changed your life overall?” In this context, impact is understood as the perceived change in one’s life, a factor assessed by the beneficiaries themselves, as previously mentioned. This question was employed as the dependent variable, to be explained by the remaining variables in a multiple linear regression model.

Each respondent provided a numerical answer ranging from -5 to 5, whereby the impact could be negative (indicated in red), neutral (0), or positive (indicated in green)

Figure 6 – Scale used to measure the perceived impact



To assess the margin of error, we compared the result of the regression, which represents the estimated actual change in the beneficiaries’ lives, with the value provided by the beneficiaries themselves in response to the first question of the questionnaire: “To what extent do you believe that Organization A’s intervention in your home has changed your life overall?”

To ensure methodological rigor, the regression model was subjected to standard diagnostic procedures:

1. Linearity and Homoscedasticity: Verified through residual plots to confirm that errors are randomly distributed and variance remains constant across fitted values.
2. Normality of Residuals: Assessed using Q–Q plots and the Shapiro–Wilk test to validate assumptions for inference.
3. Multicollinearity: Examined via Variance Inflation Factor (VIF), with all predictors presenting VIF values below the critical threshold of 5, indicating acceptable independence among variables.
4. Influential Observations: Identified using Cook's Distance and leverage statistics; no cases exceeded conventional cut-off points, suggesting that results are not driven by outliers.
5. Robustness Checks: Complementary Analysis included bootstrapping of coefficients (1,000 resamples) and robust regression estimation to confirm the stability of parameter estimates under alternative specifications.

These procedures enhance confidence in the validity of the model and mitigate risks associated with assumption violations, reinforcing the credibility of the empirical finding.

4.5.2 Sentiment Analysis

The analytical procedures presented in this section are designed to test the explanatory power of the SLI2 model, ensuring that empirical findings are directly linked to its theoretical foundations. This section operationalizes the second overarching objective of the dissertation: to assess the effectiveness of the impact measurement model developed by Entity B.

While multiple linear regression (MLR) provides robust statistical insights into the relationship between intervention dimensions and perceived impact, its reliance on structured quantitative data introduces limitations when assessing intangible outcomes such as emotional well-being or perceived dignity. To mitigate these limitations and validate the consistency of findings, this study adopts a mixed-methods approach by integrating sentiment analysis as a complementary technique. This integration enables cross-validation between self-reported perceptions and model-based estimates, strengthening the reliability of the empirical evidence and addressing potential biases inherent in single-method designs (Costa & Pesci, 2016; Schröer, Kruse, & Gómez, 2021).

Sentiment analysis offers a systematic way to capture the emotional and perceptual dimensions embedded in beneficiaries' qualitative responses.

Unlike traditional thematic coding, which is often criticized for subjectivity and lack of comparability, sentiment analysis applies computational techniques to classify text into positive, negative, or neutral categories, thereby introducing rigor and replicability into qualitative interpretation (Mohammad & Turney, 2013; González Canché & Zheng, 2025).

Its inclusion in this study serves two purposes:

1. To complement the explanatory power of MLR by revealing patterns in beneficiaries' narratives that may not emerge from quantitative indicators.
2. To operationalize a hybrid evaluation model that reflects both measurable and experiential dimensions of social impact, as recommended by recent mixed-methods frameworks (Suzara, Peña, & Botham, 2025; Alomoto, Niñerola, & Pié, 2022).

The choice of sentiment analysis is grounded in its capacity to bridge the gap between quantitative rigor and qualitative depth. Recent advances in Natural Language Processing (NLP) and lexicon-based approaches, such as the NRC Emotion Lexicon, allow for accurate polarity detection and emotion mapping across large datasets without compromising contextual meaning (Mohammad & Turney, 2013). Furthermore, frameworks like Sentiment and Emotion Network Analysis (SENA) demonstrate how sentiment analysis can be integrated with mixed-methods research to enhance interpretative validity and enable hypothesis testing on perceptual data (González Canché & Zheng, 2025). Empirical evidence from (Suzara et al., 2025) reinforces its applicability in social contexts, showing how sentiment-based metrics can quantify shifts in attitudes and perceptions following interventions.

All qualitative responses were originally collected in Portuguese. To ensure compatibility with the NRC Emotion Lexicon and maintain methodological consistency, responses were translated into English using an automated translation pipeline based on OpenAI's language models. This choice was motivated by two factors: 1) The NRC Emotion Lexicon and most validated sentiment dictionaries are optimized for English, offering higher coverage and accuracy compared to Portuguese equivalents; and 2) Using English-based lexicons facilitates alignment with international studies and ensures that results can be compared across contexts. To mitigate potential semantic distortions, the translation process included manual validation of a random sample of responses and consistency checks for emotion-related terms. These steps aimed to preserve the original meaning and reduce bias introduced by automated translation.

By incorporating sentiment analysis as a cross-validation mechanism, this study advances methodological innovation in social impact assessment. The integration of computational text analysis with regression-based modelling ensures that findings are not only statistically robust but also contextually meaningful.

This approach responds to calls for multidimensional evaluation strategies in the third sector (Gómez, de León, Sánchez-Plaza, & Soler-Gallart, 2022) and contributes to the development of replicable frameworks that combine quantitative precision with qualitative insight. Ultimately, the inclusion of sentiment analysis strengthens the credibility of the conclusions drawn and provides a richer understanding of how housing interventions influence beneficiaries' lives beyond physical improvements.

4.6 DATA ANALYSIS

The analytical procedures presented in this section are designed to test the explanatory power of the SLI2 model, ensuring that empirical findings are directly linked to its theoretical foundations. This section operationalizes the second overarching objective of the dissertation: to assess the effectiveness of the impact measurement model developed by Entity B.

To achieve this, the analysis combines quantitative and qualitative techniques. The multiple linear regression model, described in the previous section, is used to examine how each dimension of the intervention contributes to the perceived impact. These results are complemented by qualitative insights from beneficiary interviews, allowing for a more complete understanding of the changes reported.

The interpretation of the statistical indicators used in the multiple linear regression model is essential to assess the explanatory capacity of each dimension in relation to the perceived impact.

The coefficient of determination (R-squared) quantifies the proportion of variance in the dependent variable that is explained by the independent variables, offering a direct measure of model fit (Su et al., 2012). The adjusted R-squared refines this estimate by accounting for the number of predictors, thus providing a more reliable metric in models with multiple explanatory variables (Schröer et al., 2021). Additionally, the Mean Absolute Error (MAE) captures the average magnitude of prediction errors, with lower values indicating greater precision in estimating beneficiaries' responses. These indicators, widely used in social impact research, align with best practices in quantitative evaluation and contribute to the empirical evaluation of the model's internal logic (Costa & Pesci, 2016; Tuan, 2008). Their combined use ensures that the statistical analysis remains robust and theoretically grounded, reinforcing the credibility of the findings presented in this study.

By analysing all dimensions as a whole, both the data from multi linear regression and the sentiment Analysis, it is possible to observe the perceived change in beneficiaries' lives. By incorporating sentiment analysis, I will be doing an assessment of the margin of error by comparing the regression results (which estimate the actual change in beneficiaries' lives) with the self-reported responses to the first questionnaire item: "To what extent do you believe that Organization A's intervention in your home has changed your life overall?"

Taken together, these results provide critical evidence regarding the explanatory power of the model. The strong influence of physical variables confirms the model's capacity to capture tangible improvements in housing conditions, which aligns with the core mission of Organization A. However, the limited contribution of social and financial dimensions suggests that the model may not fully account for relational and economic aspects of well-being.

This finding highlights the need for methodological refinements, such as incorporating additional indicators or mixed-method approaches, to enhance the comprehensiveness of future impact assessments.

This approach ensures alignment between the empirical findings and the theoretical assumptions outlined in the literature review and in the Entity B framework. It also supports the evaluation of the model's internal logic, as discussed throughout Chapter 4.

The empirical study has described the application of the SLI2 model to real-world data, including sample characteristics, data preparation, and modelling strategies. This operationalization enables the evaluation of the model's capacity to capture multidimensional change, setting the stage for the interpretation of results and their theoretical and practical implications in the next chapter.

5. RESULTS AND DISCUSSION

This chapter will focus on translating analytical outputs into actionable insights. Firstly, by examining the explanatory power of the model and the relative contribution of each dimension to perceived impact. Secondly, by exploring how the integration of quantitative and qualitative evidence enhances the understanding of beneficiaries' experiences. And f finally, by contextualizing the results through connections to the literature and compared with key studies on social impact assessment in the third sector, to ensure that conclusions are both theoretically grounded and practically meaningful.

5.1 INTERPRETATION AND DISCUSSION

This section presents the empirical results of the impact assessment, disaggregated by dimension, to identify which aspects of the intervention most strongly influence perceived change. The analysis focuses on the explanatory power of each dimension, without anticipating the interpretative discussion.

5.1.1 Analysis per dimension

Based solely on the physical variables, we observed that if these were considered in isolation, the results would be as follows:

Table 3 – Statistical analysis per dimension

	R-squared	Adjusted R-squared	Mean Absolute Error
Physical Dim	0.816	0.449	0.669
Personal Dim	0.513	0.397	0.744
Social Dim	0.226	0.117	0.935
Financial Dim	0.220	0.182	1.025

The regression Analysis provide robust evidence of the differential explanatory power across dimensions. The physical dimension demonstrates the strongest performance (R-squared = 0.816; Adjusted R-squared = 0.787), confirming that tangible housing improvements are the primary drivers of perceived impact. Within this dimension, P5 (β = 0.7017, $p < .001$) emerges as a highly significant predictor, while P6 (β = 0.1263, p = 0.024) also contributes meaningfully, reinforcing the predominance of structural variables in short-term evaluations.

The personal dimension exhibits moderate explanatory capacity (Adjusted R-squared = 0.397), with P12 (β = 0.3253, p = 0.026) as the only significant contributor, suggesting that emotional stability and self-confidence exert a secondary but relevant effect. Conversely, the social and financial dimensions display limited explanatory power (Adjusted R-squared = 0.117 and 0.182 respectively), and none of their predictors reach conventional significance thresholds ($p > 0.05$), corroborating the interpretation that relational and economic changes require longer timeframes to materialize (Stone & Hulse, 2007; Clair & Hughes, 2019).

5.1.2 General Analysis

Table 4 – General statistical analysis

Coefficient	Value
R-squared	0.900
Adjusted R-squared	0.802
Mean Absolute Error	0.396

The empirical findings allow for the assessment of the proposed hypotheses. H1 is strongly supported, as physical indicators explain 81% of the variance in perceived impact (adjusted R-squared = 0.79). H2 receives moderate support, with personal indicators accounting for 40% of the variance. Conversely, H3 and H4 are only weakly supported, as social and financial dimensions exhibit limited explanatory power (adjusted R-squared < 0.20). These results not only confirm the predominance of physical improvements in short-term impact assessments but also align with prior evidence emphasizing the immediate effects of housing quality on well-being (Dunn, 2000; Clair & Hughes, 2019). In contrast, relational and economic changes often require longer timeframes to materialize, as highlighted by (Stone & Hulse, 2007; Gómez et al., 2022).

The full model, incorporating all dimensions, achieves an Adjusted R-squared of 0.802, confirming strong overall coherence; however, multicollinearity and the dominance of physical indicators highlight the need for methodological refinements to better capture intangible outcomes (Costa & Pesci, 2016; Alomoto et al., 2022). These findings align with prior literature emphasizing the immediacy of physical improvements and the delayed manifestation of social and financial effects, reinforcing calls for hybrid approaches that integrate quantitative rigor with qualitative depth (Dunn, 2000; Gómez et al., 2022).

5.1.2.1 Model Diagnostics

To ensure that the assumptions of multiple linear regression were reasonably satisfied, a set of diagnostic checks was conducted. Figure 10 presents the residuals versus fitted values plot, which suggests that the assumptions of linearity and homoscedasticity are broadly met, although some curvature and heteroscedasticity are observed. The histogram and Q-Q plot of residuals (Figures 11 and 12) indicate that residuals approximate a normal distribution, with minor deviations in the tails (Su et al., 2012).

Variance Inflation Factor (VIF) values were computed to assess multicollinearity. Several predictors exhibit high VIF values (greater than 10), indicating strong multicollinearity. While this may affect the stability of individual coefficients, it does not compromise the overall purpose of the analysis, which is descriptive rather than inferential (Su et al., 2012). Finally, a robustness check using heteroskedasticity-consistent standard errors (HC3) was performed (Daw & Hatfield, 2018). Results confirm that coefficient estimates remain stable, although many predictors are not statistically significant. Overall, albeit acknowledging limitations in multicollinearity and heteroscedasticity, these diagnostics indicate that the model assumptions are reasonably satisfied for descriptive purposes. Multicollinearity was not addressed because the primary objective of this study is to validate the social impact measurement methodology rather than to optimize predictive accuracy or infer causal relationships.

5.1.3 Sentiment Analysis

The inclusion of sentiment analysis at this stage aims to complement quantitative findings with perceptual data, as it was conducted a sentiment analysis to capture emotional and perceptual dimensions of impact that are not fully explained by numerical indicators. This qualitative component complements the multiple linear regression findings and strengthens the triangulation strategy adopted in this study. The sentiment analysis presented here builds on the qualitative preprocessing procedures described in the Methodology chapter 3.

After manual validation, the final distribution of sentiment polarity, shows that 72% of responses expressed positive sentiment, 23% were neutral, and only 6% were negative. This predominance of positive sentiment strongly confirms the overall effectiveness of housing intervention, while also aligning with the strong explanatory power of the physical dimension observed in the regression analysis. Emotion mapping further revealed dominant positive categories, matched with relatively high occurrences for positive sentiments — trust (49), anticipation (37), joy (37), and surprise (17) — while negative emotions were far less frequent fear (8), sadness (5), disgust (3), and anger (1). These patterns strongly suggest that the intervention not only improved housing conditions but also fostered feelings of security, optimism, and satisfaction among beneficiaries. non-physical domains, such as financial or relational aspects, which is consistent with the weaker explanatory power of these dimensions

The presence of neutral responses and occasional negative emotions indicates that some beneficiaries perceived limited change in the quantitative analysis. Taken together, these findings underscore the explanatory strength of the physical dimension while revealing methodological blind spots in capturing intangible outcomes. This imbalance reflects a broader challenge in social impact measurement: the difficulty of operationalizing constructs such as empowerment and social cohesion within short observation windows (Costa & Pesci, 2016; Tuan, 2008). Notably, the integration of sentiment analysis mitigates this limitation by providing perceptual insights that complement quantitative indicators, thereby reinforcing calls for hybrid evaluation frameworks (Alomoto et al., 2022; Suzara et al., 2025).

Figure 7 – Sentiment polarity distribution

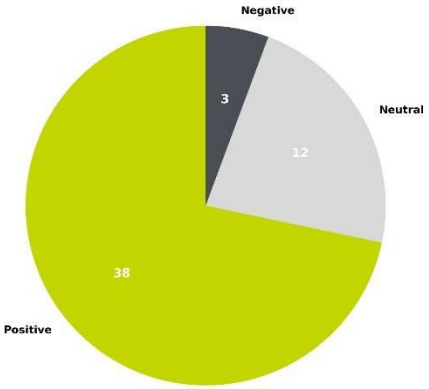
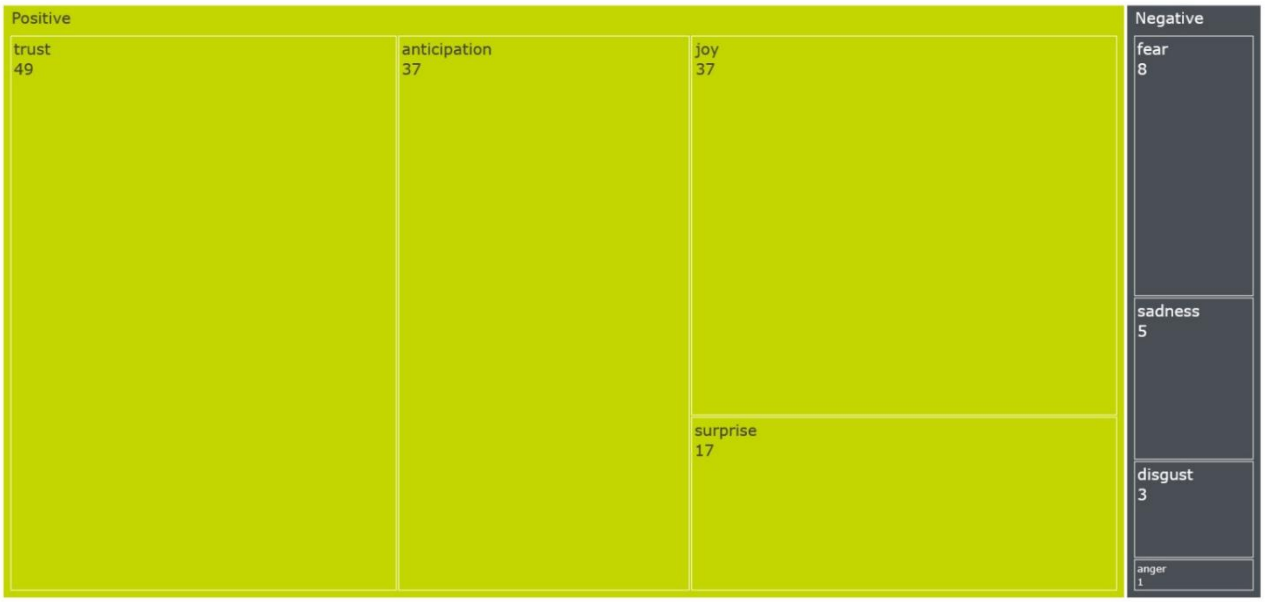


Figure 8 – Word cloud

Figure 9 – Emotion frequencies



5.2 Interpretation and Discussion

The statistical analysis presented above demonstrates that the model exhibits strong overall explanatory power, with a combined R-squared of 0.90 and an adjusted R-squared of 0.80 (Su et al., 2012). These values indicate that the dimensions included in the model account for a substantial proportion of the variance in perceived impact, confirming its internal coherence. However, the contribution of individual dimensions reveals a marked imbalance. The Physical dimension alone explains approximately 82% of the variance. The evidence confirms that the housing intervention primarily drives tangible improvements in physical living conditions, while relational and economic changes remain marginal. This imbalance underscores the need for methodological refinements, such as integrating additional indicators or complementary qualitative approaches, to better capture the complexity of social transformation (Alomoto et al., 2022).

The triangulation between perceived impact (P1) and sentiment analysis indicates a strong alignment between beneficiaries' self-reported change and their emotional valence. Overall, positive emotions dominate with 72%, neutral 23% and negative 6% sentiment substantially less frequent. On average, beneficiaries with positive sentiment report higher perceived impact (mean P.1 = 4.29), compared to neutral (mean P1 = 3.50) and negative (mean P1 = - 0,33) sentiments. Furthermore, the correspondence check between sentiment and P1 shows high agreement overall at approximately 81%. These patterns suggest that the immediate, tangible improvements captured by the physical dimension are not only reflected in quantitative scores but are also echoed by beneficiaries' emotional narratives.

Table 5 – Cross Validation Sentiment Analysis and P1

Sentiment		Average value of P1 according to the sentiment
Positive	72%	4.29
Neutral	23%	3.5
Negative	6%	- 0.33

Table 6 – Cross Validation Sentiment Analysis and P1

Correspondence between sentiment and P1	
yes	19%
no	81%
Positive correlation	**70%**

Of the 81% correspondence between sentiment and P1 responses, 70% indicate a positive alignment, reflecting a shared perception of positive impact.

The relationship between short-term perceived impact and expectations of durability is likewise consistent. Among high perceived impact cases (P1 ≥ 4), the vast majority anticipate that changes will persist ("De certeza que sim" or "Em princípio, sim"). In aggregate, cases combining positive sentiment, high P1 (≥ 4), and favourable sustainability expectations represent 62% of the sample, reinforcing a convergent profile between immediate impact and long-term confidence.

The distribution of P14 responses is also skewed toward sustained change (43% "De certeza que sim" and 40% "Em princípio, sim"), with limited uncertainty ("NS/NR" of 8%) and few unfavourable expectations ("Em princípio, não" of - 9%).

Importantly, no statistically meaningful geographical differences emerge in these patterns. Positive sentiment and high perceived impact are broadly observed across districts, indicating that the intervention's short-term effectiveness and expected durability are not contingent on location.

Although the model demonstrates strong explanatory power overall, its limited capacity to capture social and financial dimensions raises significant methodological concerns. These domains are inherently complex and often manifest through gradual, intangible changes that are not easily quantifiable in short-term evaluations.

As highlighted by both studies (Stone & Hulse, 2007; Gómez et al., 2022), relational improvements, such as strengthened family ties or enhanced social networks require sustained interaction and broader structural conditions, which are beyond the immediate scope of housing interventions. Similarly, financial empowerment depends on long-term factors such as employability and income stability, which cannot be fully assessed through post-intervention self-reports conducted within a few months.

The integration of sentiment analysis provides complementary insights that enrich this interpretation. While quantitative results highlight the primacy of physical factors, qualitative data reveal strong emotional responses associated with trust, joy, and anticipation, confirming the positive impact of the intervention beyond material improvements.

At the same time, the presence of neutral and occasional negative sentiments indicates that changes in social and financial domains were perceived as limited, reinforcing the findings from the regression analysis. Together, these results illustrate the value of combining quantitative and qualitative approaches to achieve a more comprehensive understanding of social impact.

The findings presented in this study align with existing evidence on the relationship between housing conditions and well-being. As highlighted by both studies (Dunn, 2000; Clair & Hughes, 2019), improvements in physical housing quality are strongly associated with better health outcomes and enhanced quality of life. The predominance of the Physical dimension in explaining perceived impact is therefore consistent with these theoretical expectations, confirming that tangible improvements in living conditions remain the most immediate and visible drivers of social change.

This result also reflects the nature of Organization A's mission, present in its motto "rebuilding houses, restoring lives." By providing safe and adequate housing to vulnerable populations, the organisation addresses an urgent and concrete need: the absence of shelter and security. In the short term, beneficiaries experience the most significant changes in physical conditions, which explains why this dimension dominates the model. Data collection occurs only a few months after the intervention, a timeframe that captures immediate benefits.

Social and personal improvements, such as strengthened relationships or enhanced emotional stability, require longer periods to materialise and therefore remain underrepresented in this analysis. The short observation window limits the ability to capture these changes, a challenge widely acknowledged in the literature (Stone & Hulse, 2007), argue that relational and economic transformations are more complex, often requiring sustained interventions and broader structural changes. These findings are consistent with the broader context, as the majority of Organization A's beneficiaries tend to experience social isolation, often living alone and having limited contact with family members. Since the organisation's intervention focuses on housing improvements, it does not directly address the issue of loneliness.

Consequently, the inclusion of variables that depend on the beneficiaries' social context introduces bias into the measurement of perceived impact. This reinforces the need for multidimensional approaches that go beyond physical indicators to capture intangible outcomes such as empowerment and social cohesion.

Finally, the integration of sentiment analysis responds to calls for hybrid methodologies in social impact assessment. (Costa, 2016; Alomoto et al., 2022), emphasize the importance of combining quantitative rigor with qualitative depth to ensure that evaluations reflect both measurable and experiential dimensions of change. By triangulating regression results with emotional patterns in beneficiaries' narratives, this study demonstrates the practical value of such approaches and contributes to ongoing debates on methodological innovation in the third sector.

Taken together, these findings confirm the model's capacity to measure physical improvements with high reliability, partially meeting the first research objective. However, the limited sensitivity to relational and economic transformations indicates that the second objective, capturing multidimensional change, was only partially achieved.

5.3 Limitations of the Model

Despite the model's strong explanatory power, several limitations constrain its validity. The main one is the absence of baseline data, which precludes counterfactual designs such as Difference-in-Differences (Daw & Hatfield, 2018). Additionally, reliance on self-reported perceptions introduces potential bias, even when mitigated by external interviewers (Davies & Dart, 2005; Tuan, 2008). Finally, the exclusion of the professional dimension, although justified by data constraints, reduces the comprehensiveness of the analysis, particularly regarding economic empowerment (Costa, 2016).

The absence of baseline data and the exclusive reliance on post-intervention perceptions constrain the validity of causal inferences. Without pre-intervention measurements, it is impossible to determine the magnitude of change attributable to the intervention versus external factors (Daw & Hatfield, 2018). Furthermore, self-reported data are susceptible to social desirability bias, particularly in contexts where beneficiaries feel gratitude toward the organization (Davies & Dart, 2005; Tuan, 2008). These limitations have practical consequences: they restrict the organization's ability to demonstrate impact credibly to funders and policymakers and hinder strategic learning for program improvement.

The reliance on self-reported data introduces potential response bias. Although interviews were conducted by external professionals to mitigate this risk, beneficiaries may still overstate positive outcomes due to gratitude or social desirability. This challenge is widely acknowledged in participatory evaluation literature (Davies & Dart, 2005; Tuan, 2008). Also, it is important to evidence the exclusion of the professional dimension, which reduces the comprehensiveness of the analysis.

Questions related to employability and skills development were omitted because they applied only to a subset of respondents, but this decision limits the model's capacity to assess economic empowerment, an important aspect of long-term social transformation (Costa & Pesci, 2016).

It is also important to emphasise that the primary contact is not always the most vulnerable individual within the household. This introduces a bias in the assessment of actual vulnerability in relation to other characteristics, which is not accounted for in the statistical model employed. Similarly, the inclusion of variables that depend on the beneficiaries' social context introduces additional bias. These findings are consistent with the broader context, as the majority of Organization A´s beneficiaries tend to experience social isolation, often living alone and having limited contact with family members.

Since the organisation's intervention focuses on housing improvements, it does not directly address the issue of loneliness. Consequently, indicators related to social interaction or family relationships may underestimate or misrepresent the impact of the intervention (Gómez et al., 2022). These limitations have direct implications for the interpretation of results. They constrain the external validity of findings and highlight the need for complementary qualitative approaches and longitudinal data collection to strengthen future evaluations (Alomoto et al., 2022)

5.3.1 Strategies for Mitigation

To strengthen the robustness of future evaluations, several low-cost strategies can be implemented without imposing excessive operational burdens on Organization A. These recommendations align with best practices in social impact assessment, which emphasize the importance of longitudinal measurement, triangulation, and stakeholder engagement (Costa, 2016; Alomoto et al., 2022; Gómez et al., 2022).

Baseline Data Collection: Literature highlights the critical role of baseline data in enabling causal inference and improving validity (Daw & Hatfield, 2018). Introducing a simplified pre-intervention survey would allow for a more accurate comparison between initial and post-intervention conditions. This can be conducted during the initial home visit, by volunteers a few days before the intervention, through short telephone interviews, or in partnership with external factors such as social workers or local organisations. These approaches reflect the participatory principles advocated by (Davies & Dart, 2005; Gómez et al., 2022), ensuring that evaluations capture context-specific realities.

Longitudinal Follow-Up: Several authors argue that social and financial changes often require extended timeframes to materialize (Stone & Hulse, 2007; Clair & Hughes, 2019). Incorporating a third interview approximately one year after the intervention, and ideally a fourth at the two-year mark, would enable the measurement of delayed effects. Leveraging existing contact points, such as volunteer visits or mentorship programs, minimizes additional costs while enhancing methodological rigor.

Hybrid Approaches: Combining quantitative indicators with qualitative narratives and computational techniques such as sentiment analysis responds to calls for multidimensional evaluation frameworks (Suzara et al., 2025; González Canché & Zheng, 2025). This integration improves sensitivity to intangible outcomes, such as empowerment and emotional well-being, which are often overlooked in purely quantitative models.

Broader Stakeholder Engagement: Beyond interviewing only the direct beneficiaries, including perspectives from partners who accompany these individuals, such as social workers, local NGOs, food banks, and church groups aligns with participatory evaluation principles (Davies & Dart, 2005; Gómez et al., 2022). This triangulation enriches the understanding of impact and validates self-reported data, addressing concerns about bias and enhancing credibility

These strategies aim to enhance the comprehensiveness and validity of impact measurement, improving both internal learning and external accountability. By adopting these practices, Organization A can move toward a more robust and context-sensitive evaluation framework, consistent with contemporary debates on methodological innovation in the third sector.

5.4 Comparative Analysis with the Literature

The approach adopted in this dissertation shares important similarities with established frameworks for social impact assessment in the third sector, while also presenting distinctive features (Costa, 2016) emphasizes the need for multidimensional models that integrate both tangible and intangible outcomes. The SLI2 model aligns with this principle by incorporating physical, personal, social, and financial dimensions, although the empirical results reveal an imbalance that favours physical indicators. This contrasts with Costa's recommendation for stronger emphasis on relational and empowerment-related metrics

(Alomoto et al., 2022), advocate for hybrid methodologies that combine quantitative rigor with qualitative depth, a perspective fully reflected in this study through the integration of regression analysis and sentiment analysis. Similarly, (Gómez et al., 2022) highlight the importance of participatory approaches and stakeholder-defined indicators. While the present model incorporates beneficiary perceptions, its reliance on structured surveys limits the degree of participatory engagement compared to the iterative processes described by (Gómez et al., 2022). Finally, (Tuan, 2008) discusses the challenges of monetizing social value and the limitations of cost-based approaches, proposing alternative strategies that prioritize experiential outcomes. Although this dissertation does not adopt a monetization framework, its focus on perceived impact and emotional responses resonates with Tuan's call for methods that capture subjective dimensions of change.

In summary, the study is consistent with contemporary debates on methodological innovation in social impact assessment, particularly in its use of mixed methods. However, it diverges from best practices in areas such as longitudinal measurement and participatory design, highlighting opportunities for refinement in future evaluations. For instance, while the integration of sentiment analysis represents an innovative step compared to traditional frameworks, the predominance of physical indicators challenges the multidimensional perspective advocated by (Costa, 2016).

This tension underscores the importance of refining indicator design to better capture intangible outcomes such as empowerment and social cohesion. Building on these insights, the next chapter consolidates the main contributions of this research and proposes methodological refinements and practical recommendations to enhance the measurement of multidimensional social impact. The results demonstrate the model's strong overall explanatory power while revealing imbalances across dimensions, with physical indicators dominating short-term impact assessments. These findings highlight both the strengths and limitations of the current approach and underscore the need for methodological refinements.

6. CONCLUSIONS AND FUTURE WORK

This chapter synthesizes the main findings by linking objectives, methods, and results. The first research question received strong support: the Entity B model exhibited high internal coherence and explanatory power (R-squared = 0.90), effectively capturing tangible improvements in housing conditions and aligning with theoretical expectations and the organization's mission. These results are consistent with prior evidence on the primacy of physical housing quality in shaping well-being, particularly in the short term (Dunn, 2000; Clair & Hughes, 2019), and with broader calls in social impact measurement to privilege empirically grounded indicators for direct outcomes (Costa & Pesci, 2016; Alomoto, Niñerola, & Pié, 2022).

The second research question, whether the model can fully capture multidimensional change, was only partially confirmed. While the physical domain accounts for most of the short-term variance in perceived impact, the social and financial domains contribute more modestly, a pattern that reflects the difficulty of operationalizing relational cohesion and economic security within brief observation windows and the exclusion of the professional dimension due to data constraints (Stone & Hulse, 2007; Costa & Pesci, 2016).

Methodologically, this study introduces a hybrid evaluation framework that combines CRISP-DM with sentiment analysis, illustrating how structured data-science processes can enhance rigor and reproducibility in non-profit contexts (Chapman et al., 2000; Schröer, Kruse, & Gómez, 2021). Computational text analytics complement quantitative findings by analysing beneficiaries' narratives, using lexicon-based methods for transparent and replicable classification of polarity and emotion (Mohammad & Turney, 2013; Silge & Robinson, 2017). Recent mixed-methods applications further demonstrate the value of these techniques for triangulation (Suzara, Peña, & Botham, 2025; González Canché & Zheng, 2025).

Practically, the results offer actionable guidance for accountability and organizational learning by clarifying where the model is strongest, capturing direct, material effects, and where sensitivity should be improved to detect intangible, delayed outcomes.

At the same time, several limitations temper their external validity. The absence of baseline data precludes counterfactual inference and constrains causal claims, even when interviews are conducted by external professionals to reduce bias (Daw & Hatfield, 2018; Davies & Dart, 2005).

Short time horizons underrepresent longer-term social and financial effects, and omitting the professional dimension reduces comprehensiveness. Addressing these constraints points to a coherent agenda for future work that remains feasible within realistic operational boundaries. First, baseline data collection prior to intervention would enable Difference-in-Differences designs and strengthen internal validity (Daw & Hatfield, 2018).

Second, longitudinal follow-ups at one- and two-year intervals would capture the temporal unfolding of empowerment, relational cohesion, and household economic stability (Stone & Hulse, 2007; Clair & Hughes, 2019). Third, reintegrating professional indicators such as employability, skills acquisition, and perceived economic security, would restore the intended multidimensionality of the model and align measurement with the broader objectives of social transformation (Costa & Pesci, 2016; Alomoto et al., 2022). Fourth, advancing qualitative sensitivity beyond lexicon-based sentiment requires modern NLP, including transformer-based approaches and network frameworks capable of extracting nuanced emotion and relational structure (González Canché & Zheng, 2025; Mohammad & Turney, 2013). Fifth, stakeholder triangulation should systematically incorporate the perspectives of volunteers, social workers, and community partners, thereby enriching the understanding of impact pathways and validating self-reports (Davies & Dart, 2005; Gómez et al., 2022).

Finally, benchmarking against established frameworks such as Social Return on Investment and Most Significant Change while maintaining contextual nuance, would improve comparability and external validity and enhance alignment with sector standards and European guidance on impact measurement (Tuan, 2008; Ortega Maldonado & Corbey, 2015; European Commission, 2014). Complementary reference taxonomies and indicator sets (e.g., IRIS+ / GIIN) could further support harmonization for reporting and learning across organisations (Hand et al., 2020; Li et al., 2022).

Taken together, these recommendations chart a feasible pathway for improving robustness, inclusivity, and multidimensionality in social impact assessment. They preserve the clarity and interpretability of the existing statistical core while expanding the evidentiary base required to apprehend change that is gradual, relational, and sometimes indirect. By situating the current findings within this forward-looking methodological trajectory, the thesis moves beyond documenting short-term physical improvements to articulate how impact measurement can evolve toward genuinely multidimensional assessment, responsive to the lived realities of vulnerable populations and to the accountability demands faced by nonprofit organisations (Costa & Pesci, 2016; Gómez et al., 2022).

Ultimately, this research not only validates a hybrid framework for social impact assessment but also sets a foundation for future methodological innovation, contributing to more rigorous and socially meaningful evaluations in the non-profit sector.

BIBLIOGRAPHICAL REFERENCES

Acolin, A., & Wachter, S. (2017). Opportunity and Housing Access. Cityscape, 19(1), 135–149.

Alomoto, W., Niñerola, A., & Pié, L. (2022). *Social Impact Assessment: A Systematic Review of Literature*. Social Indicators Research, 161(1), 225–250. https://doi.org/10.1007/s11205-021-02809-1

Chadee, A. A., Martin, H. H., Mwasha, A., & Otuloge, F. (2022). *Rationalizing Critical Cost Overrun Factors on Public Sector Housing Programmes.* Emerging Science Journal, 6(3), 647–666. https://doi.org/10.28991/ESJ-2022-06-03-016

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). *CRISP-DM 1.0: Step-by-step data mining guide* (Technical Report). The CRISP-DM Consortium / SPSS Inc.

Clair, A., & Hughes, A. (2019). Housing and health: New evidence using biomarker data. Journal of Epidemiology and Community Health, 73(3), 256–262. https://doi.org/10.1136/jech-2018-211431

Costa, E., & Pesci, C. (2016). Social impact measurement: Why do stakeholders matter? Sustainability Accounting, Management and Policy Journal, 7(1), 99–124. https://doi.org/10.1108/SAMPJ-12-2014-0092

Cunha, J., & Benneworth, P. (2020). How to measure the impact of social innovation initiatives? International Review on Public and Nonprofit Marketing, 17(1), 59–75. https://doi.org/10.1007/s12208-019-00240-4

Davies, R., & Dart, J. (2005). *The "Most Significant Change" (MSC). Technique. CARE* International. http://www.mande.co.uk/wp-content/uploads/2005/MSCGuide.pdf

Daw, J. R. , & Hatfield, l. A. (2018*). Matching and Regression to the Mean in Difference-in-Differences Analysis. Health Services Research*, 53(6), 4138–4156. https://doi.org/10.1111/1475-6773.12993

Dunn, J. R. (2000). Housing and Health Inequalities: Review and Prospects for Research. Housing Studies, 15(3), 341–366. https://doi.org/10.1080/02673030050009221

Epstein, M. J., & Roy, M.J. (2001). Sustainability in Action: Identifying and Measuring the Key Performance Drivers. Long Range Planning, 34(5), 585–604. https://doi.org/10.1016/S0024-6301(01)00084-X

European Commission. (2014). *Proposed approaches to social impact measurement in European Commission legislation and in practice relating to EuSEFs and the EaSI*. Publications Office of the European Union. https://ec.europa.eu

Eurostat. (n. d.). Eurostat: The home of high-quality statistics and data on Europe. Retrieved December 1, 2025, https://ec.europa.eu/eurostat

Funell, S. C., & Rogers, P. J. (2011). Purposeful Program Theory: Effective Use of Theories of Change and Logic Models. Jossey-Bass, Wiley.

Gómez, A., de León, M. , Sánchez-plaza, a., & Soler-Gallart, M. (2022*). Social Impact Ex-Post Evaluation Protocol. International Journal of Qualitative Methods*, 21. https://doi.org/10.1177/16094069221133231

González Canché, M. S., & Zheng, K. (2025). Sentiment and Emotion Network Analysis—SENA: An Analytic Framework and No-Code Software to Integrate Natural Language Processing, Linguistic Analysis, and Data Science with Qualitative and Mixed-Methods Research. International Journal of Qualitative Methods, 24, 16094069251346929.

Hand, D., Dithrich, H. , Sunderji, S., & Nova, N. (2020, June 11). *2020 Annual Impact Investor Survey.* Global Impact Investing Network. https://thegiin.org/publication/research/impinv-survey-2020/

Instituto Nacional de Estatística. (2024, março 15*). Rendimento e condições de vida – Habitação, dificuldades habitacionais e eficiência energética dos alojamentos* – 2023. INE. https://www.ine.pt

The Interorganizational Committee on Guidelines and Principles for Social Impact Assessment. (1994). *Guidelines and principles for social impact assessment.* U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Marine Fisheries Service.

Joy, V., Jain, V. K., Agarwal, S., & Kumar, P. (2021). Revisiting Determinants of Social Entrepreneurial Intentions among Indian Students using Interpretive Structural Modelling. *Elementary Education Online*, 20(5), 4299-4299.

Just a Change. (n.d.). Just a Change: *Renovating homes, rebuilding lives*. Retrieved December 1, 2025, from https://justachange.pt/

Just a Change. (2023). *Impact report 2023* (English version). https://justachange.pt/wpcontent/uploads/2024/05/Relatorio-Impacto-2023_ENG.pd

Just a Change. (2022). *Impact report 2022* (English version). https://justachange.pt/wpcontent/uploads/2023/05/relatorio-2022_eng-1.pdf

Kearns, A., Hiscock, R., Ellaway, A., & MaCintyre, S. (2000). “Beyond Four Walls”. The Psycho-social Benefits of Home: Evidence from West Central Scotland. Housing Studies, 15(3), 387–410. https://doi.org/10.1080/02673030050009249

Knowlton, L. H., & Phillips, C. C. (2012). *The logic model guidebook: Better strategies for great results* (2nd ed.). Jossey-Bass.

Kročil, O. (2018). Pospíšil: *comprehensive measurement of social enterprise impact ekonomski pregled* (vol. 69, Issue 5).

Li, S., Gray, E. , Dennis, m., Hand, D., & Avi, S. (2022). *Unlocking Early-Stage Financing for SDG Partnerships.* World Resources Institute. https://doi.org/10.46830/wrirpt.21.00041

Mohammad, S. M., & Turney, P. D. (2013). *Crowdsourcing a word–emotion association lexicon.* Computational Intelligence, 29(3), 436–465. https://doi.org/10.1111/j.1467-8640.2012.00460.x

Mullins, D., Czischke, D. , & Van Bortel, g. (2012). *Exploring the meaning of hybridity and social enterprise in housing organisations.* In Housing Studies (Vol. 27, Issue 4, pp. 405–417). Routledge. https://doi.org/10.1080/02673037.2012.689171

Ortega Maldonado, M., & Corbey, M. (2015). Social Return on Investment (SROI): *A review of the technique. Management Accounting*, 90(E), 80–87

Salido-Andres, N., Garcia-Rodriguez, N., & Cachero-Martinez, S. (2022). *Connecting Social Enterprises and Sustainable Consumption: Systematic Review, Bibliometric Analysis, and Conceptual Framework*. Sustainability, 14(20), 13428. https://doi.org/10.3390/su142013428

Schröer, C., Kruse, F. , & Gómez, J. M. (2021). *A systematic literature review on applying CRISP-DM process mode*l. Procedia Computer Science, 181, 526–534. https://doi.org/10.1016/j.procs.2021.01.199

Silge, J., & Robinson, D. (2017). *Text mining with R: A tidy approach.* O'Reilly Media.

Social Data Lab. (2024, Julho). *Relatório de medição de impacto social: Programa Just a Change*. Social Data Lab.

Stone, W., and Hulse, K. (2007) *Housing and social cohesion: an empirical exploration, AHURI Final Report No. 100, Australian Housing and Urban Research Institute Limited, Melbourne*, https://www.ahuri.edu.au/research/final-reports/100

Suzara, M., Peña, C., & Botham, C. (2025). *An exploratory mixed-methods sentiment analysis of a social-emotional learning (SEL)-informed grant writing workshop in the biosciences*. Journal of Microbiology & Biology Education, 26(2). https://doi.org/10.1128/jmbe.00219-24

Su, Y., Gao, X. , li, x., & Tao, D. (2012). *Multivariate multilinear regression. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 42(6), 1560–1573. https://doi.org/10.1109/TSMCB.2012.2195171

Svejenova, S. (2018). Creative leadership for social impact. *In Frontiers of Creative Industries: Exploring Structural and Categorical Dynamics* (Vol. 55, pp. 47-72). Emerald Publishing Limited.

Tuan, M. T. (2008). *Measuring and/or estimating social value creation: Insights into eight integrated cost approaches.* Bill & Melinda Gates Foundation Impact Planning and Improvement.

United Nations Development Programme. (n.d.). *Selecting indicators for impact evaluation*. UNDP. Retrieved December 1, 2025 https://www.undp.org

United Nations Human Settlements Programme. (n. d.). UN-Habitat: A better urban future. Retrieved December 1, 2025, https://unhabitat.org/

Zinovieff, M., & Rotem, A. (2008). Review and analysis of training impact evaluation methods, and proposed measures to support a United Nations system fellowships evaluation framework. World Health Organization. https://studylib.net/doc/8208760/review-and-analysis-of-training-impact-evaluation-methods

APPENDIX A

Ethics Committee Report

Dear Leonor Vital,
Dear professor Manuela Aparício,

Thank you for filling out the Research Ethics Checklist. After reviewing your request, you can proceed with the study as we do not foresee any major ethical concerns with the project.

Project No.: **STAT2025-5-258332**

Project Title: **Just a Change Analysis and Evaluation of the Statistical Model for Impact Assessment**

Principal Researcher: **Ana Leonor Vital**

According to the regulations of the Ethics Committee of NOVA IMS and the MagIC Research Center, this project was considered to meet the requirements of the NOVA IMS Internal Review Board and is hereby considered **APPROVED** on **02/06/2025**.

It is the Principal Researcher's responsibility to ensure that all researchers and stakeholders associated with this project are aware of the conditions of approval and which documents have been approved.
The Principal Researcher is required to notify the Ethics Committee, via amendment or progress report, of:

- Any significant change to the project and the reason for that change;
- Any unforeseen events or unexpected developments that merit notification;
- The inability of the Principal Researcher to continue in that role or any other change in research personnel involved in the project.

Lisbon, 02/06/2025
NOVA IMS Ethics Committee
ethicscommittee@novaims.unl.pt

APPENDIX B

BENEFICIÁRIOS

Nesta conversa, vamos falar sobre o trabalho desenvolvido pelo **Just a Change**. Todas as perguntas dizem respeito a este projeto específico.

P.0.1. O(a) senhor(a) foi a pessoa que manteve o principal contacto com a equipa do **Just a Change**?

SIM	**1**
NÃO	**2**

P.0.2. Qual a sua idade? |_||_|

P.0.3. Quantas pessoas moravam na sua casa durante a intervenção do **Just a Change** na sua casa? |_||_|

Vamos agora fazer-lhe um conjunto de perguntas para perceber até que ponto a **intervenção do Just a Change na sua casa mudou ou não mudou aspetos da sua vida**. Para responder a cada uma destas perguntas, vamos pedir-lhe para escolher um número a partir de todos estes que lhe vou mostrar. (**MOSTRAR CARTÃO 1)**

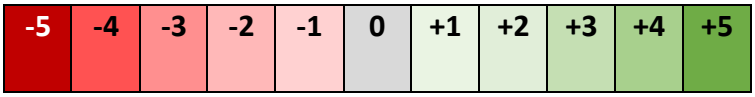
-5	**-4**	**-3**	**-2**	**-1**	**0**	**+1**	**+2**	**+3**	**+4**	**+5**

Neste cartão, tem números **positivos**, que estão a **verde**, e significam que acha que a sua vida **melhorou**. Também tem números **negativos**, que estão a **vermelho**, e significam que acha que a sua vida **piorou**. Também tem o 0, que está a cinzento, e significa que acha que a sua vida **não mudou**, ou seja, não melhorou nem piorou.

Portanto, a primeira coisa que lhe peço é para **pensar se a sua vida**, para cada ponto que lhe vou dizer, **mudou** para **melhor**, mudou para **pior** ou **não mudou**. Depois, se acha que mudou, tem de escolher um número entre 1 e 5. Quanto **maior for o número**, **maior foi a mudança**. Assim:

- **1** significa que mudou muito pouco;
- **2** significa que mudou pouco;
- **3** significa que mudou médio;
- **4** significa que mudou muito;
- **5** significa que mudou muitíssimo.

Por exemplo, **- 4** significa que mudou **muito**, **para** pior. Ou seja, **piorou muito**. **+ 2** significa que mudou **pouco**, mas para **melhor**. Ou seja, **melhorou pouco**. De acordo? (**SE NECESSÁRIO, REPETIR)**



P.1. Então, vamos começar: até que ponto acha que a intervenção do **Just a Change** na sua casa **mudou a sua vida em geral**? (**MOSTRAR CARTÃO 1)**

NÃO SABE 9

SE A PESSOA RESPONDEU 0 – NÃO MUDOU NADA – SIGA PARA P.14

P.2. E até que ponto acha que esta intervenção do **Just a Change** mudou as **condições em que vive na sua casa**? Qual é o número que escolhe? (**MOSTRAR SEMPRE CARTÃO 1)**

P.3. Usando os mesmos números, vou pedir-lhe agora que me diga até que ponto a **intervenção do Just a Change na sua casa** trouxe mudanças ao nível…

3.1. … da sua saúde (por exemplo, ao nível de necessidade de ir ao médico ou a frequência com que esteve doente) |___|

3.2. … da melhoria dos seus hábitos de higiene (por exemplo, tomar banho, utilização da sanita, etc.)

3.3. … do seu bem-estar físico (sentir mais conforto dentro de casa, ter mais facilidade em

descansar, etc.) |___|

3.4 ... ao nível do seu conforto térmico (não sentir frio, não ter dependência de aquecedores, não sentir calor excessivo)?

3.5. … da segurança com que utiliza certos locais dentro da casa ou equipamentos específicos (por exemplo, o fogão ou a banheira da casa de banho) |___|

3.6. … de sentir mais proteção em casa em relação ao exterior |___|

P.4. Usando os mesmos números, vou pedir-lhe agora que me diga até que ponto a **intervenção do Just a Change na sua casa** trouxe mudanças noutros aspetos, por exemplo ao nível…

4.1. … de se sentir mais confiante consigo mesmo(a) |___|

4.2. … de passar a cuidar mais de si |___|

4.3. … de acordar de manhã com mais vontade de fazer coisas (por exemplo, atividades de trabalho ou de lazer) |_|

4.4. … de sentir que as pessoas se preocupam consigo e gostam de si |___|

4.5. … de sentir que consegue fazer as coisas sozinho(a) |___|

4.6. … de se sentir mais feliz, menos frustrado(a), menos irritável |___|

4.7. … de sentir que a sua vida tem mais sentido |___|

4.8. … de confiar mais nos outros |___|

4.9. … de se sentir mais orgulhoso(a) com a sua casa e de ter mais vontade de cuidar dela |___|

4.10. … de se sentir mais orgulhoso(a) com a sua situação na vida |___|

P.5. Vou pedir-lhe agora que me diga até que ponto a **intervenção do Just a Change na sua casa** trouxe outro tipo de mudanças, por exemplo a nível…

5.1. … da melhoria da sua relação com a sua família |___|

5.2. … da sua vontade em estar com outras pessoas |___|

5.3. … de convidar pessoas para ir à sua casa (familiares, amigos, vizinhos, etc.) |___|

5.4. … de ser uma pessoa mais disponível para os outros (familiares, amigos, vizinhos) |__|

5.5. … do número de momentos de convívio com outras pessoas |___|

5.6. … de passar a recorrer mais vezes a ajuda dos outros |___|

5.7. … da maneira como a sua família, os seus amigos e os seus vizinhos o(a) veem como pessoa |___|

P.6. Encontrava-se a trabalhar/estudar ou à procura de trabalho/estudo, no momento da intervenção do **Just a Change** na sua casa?

SIM	**1**	□ PASSAR PARA P.7
NÃO	**2**	□ PASSAR PARA P.8
NS/NR	**9**	□ PASSAR PARA P.8

P.7. (**MOSTRAR CARTÃO 1)** Vou pedir-lhe agora que me diga até que ponto a **intervenção do Just a Change na sua casa** trouxe mudanças a nível…

7.1. … da sua motivação e confiança para trabalhar/estudar |___|

7.2. … de sentir que o seu trabalho/estudo ou a sua procura está a correr melhor |___|

7.3. … de ganhar conhecimentos importantes para a sua vida profissional ou de estudo |___|

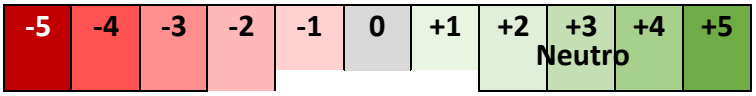
7.4. … ter melhores condições para trabalhar ou estudar|___|

P.8. (**MOSTRAR CARTÃO 1)** Até que ponto a **intervenção do Just a Change** mudou a sua capacidade de gerir o seu dinheiro pessoal? |___|

P.9. (**MANTER CARTÃO 1)** E até que ponto a **intervenção** do **Just a Change** alterou as despesas da casa como água, eletricidade ou gás? |_|

P.10. Como se sente no que diz respeito à relação com os voluntários? (**MOSTRAR CARTÃO 3)**

Muitíssimo feliz



Muitíssimo triste

s do **Just a Change** no último ano?

P.11. Como descreveria a sua rela ção com os voluntário (**MOSTRAR CARTÃO 2)**

Sinto que são meus amigos e estiveram sempre disponíveis quando precisei	**1**
Sinto que são meus amigos, mas pouco presentes na minha vida	**2**
Foram meus amigos durante as obras, mas depois não houve mais contacto	**3**
Não sinto que chegámos a construir uma amizade	**4**
NS/NR	**9**

P.12. Durante este ano, após a **intervenção do Just a Change na sua casa**, que tipo de acompanhamento sentiu por parte de voluntários ou assistentes sociais? (**MOSTRAR CARTÃO 4)**

Nenhum acompanhamento	**1**
Pouco acompanhamento	**2**
Acompanhamento razoável	**3**
Muito acompanhamento	**4**
NS/NR	**9**

P.13. Gostava de ter sido mais acompanhado/a ou apoiado/a?

Sim, pelos voluntários do Just a Change	1
Sim, pelos assistentes sociais	2
Sim, por ambos	3
Não, não foi preciso mais apoio	4
NS/NR	9

P.14. Daqui a 5 anos, acha que estas mudanças na sua vida de que temos vindo a falar se vão, de uma maneira geral, manter? (**MOSTRAR CARTÃO 5)**

De certeza que sim	1
Em princípio, sim	2
Em princípio, não	3
De certeza que não	4
NS/NR	9

P.15. Se o **Just a Change** não existisse, acha que seria fácil ou difícil encontrar quem lhe desse este tipo de ajuda? (**MOSTRAR CARTÃO 6)**

Fácil	1
Difícil	2
Impossível	3
NS/NR	9

P.16. Quer deixar-nos algumas palavras sobre esta ajuda do **Just a Change** na sua vida?

ANNEXS

Figure 10 - Residuals vs Fitted

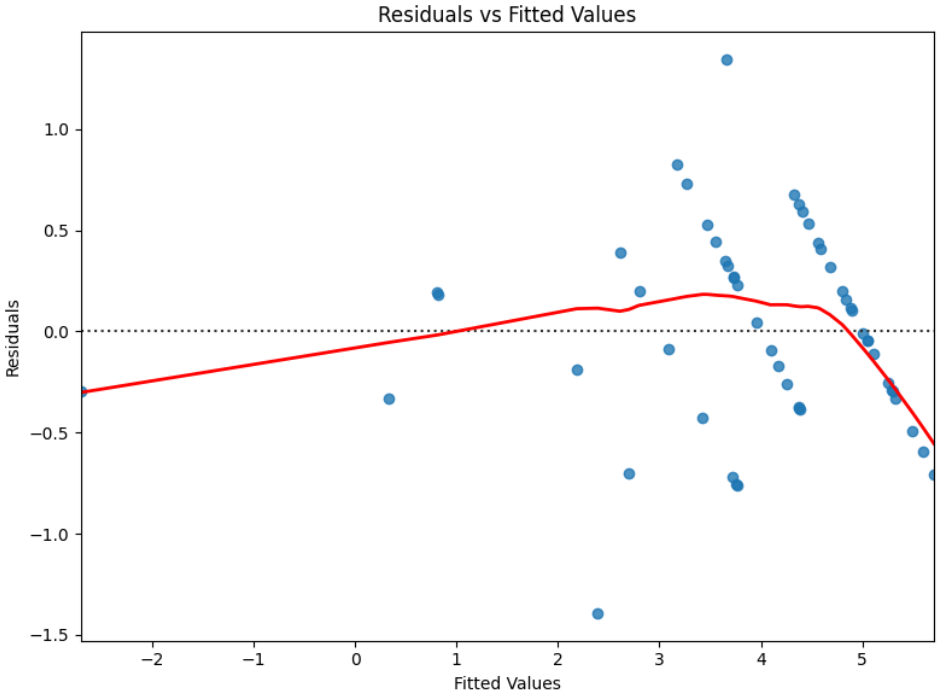


Figure 11 – Histogram

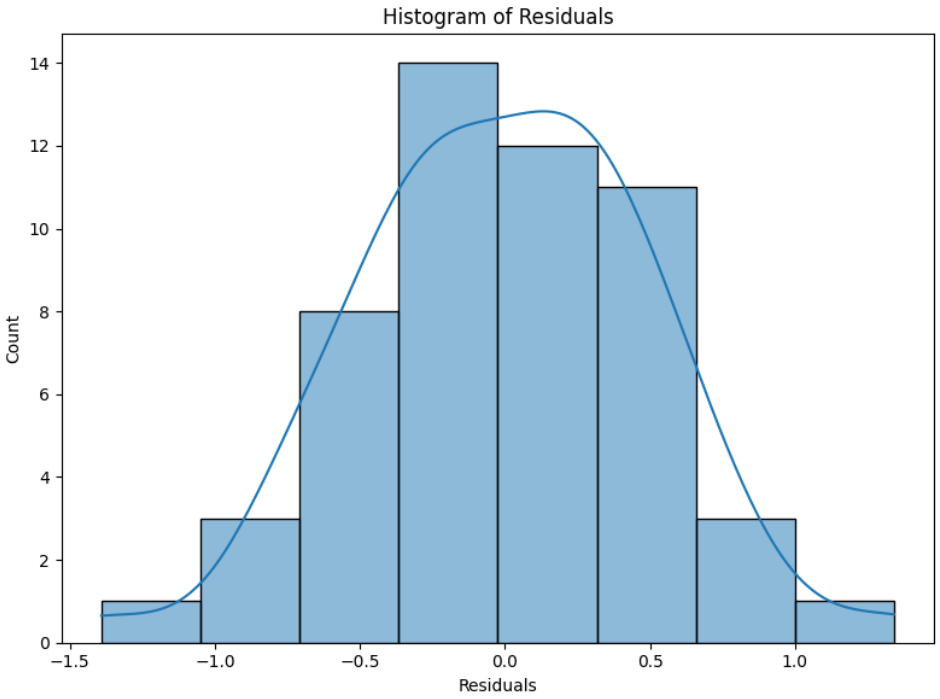


Figure 12 – Q-Q Plot

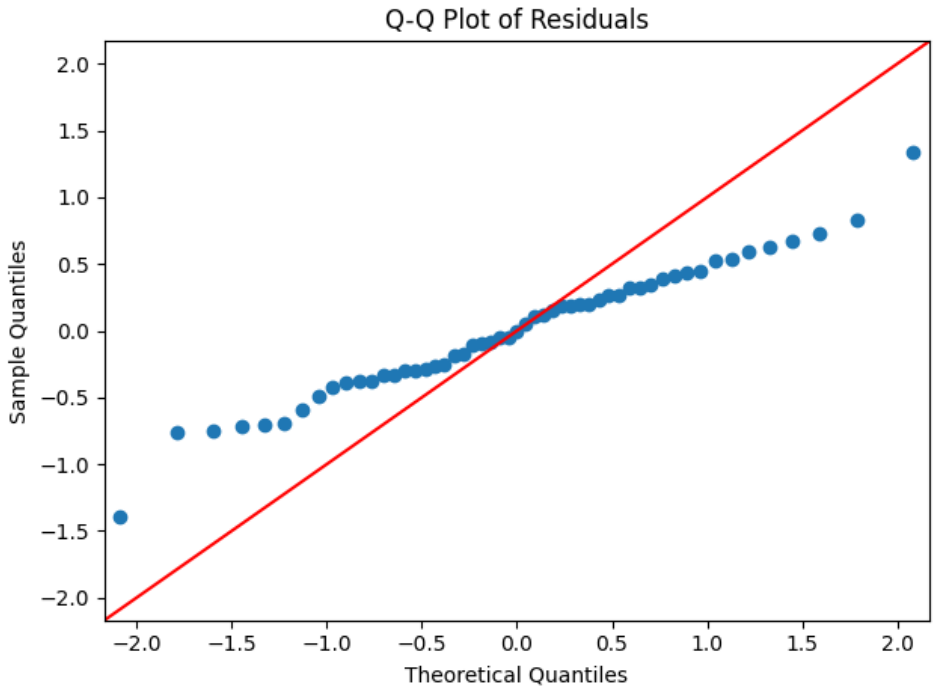


Figure 13 – OLS Regression Results Physical Dimension

Physical Dimension

OLS Regression Results

Dep. Variable:	P4	R-squared:	0.816
Model:	OLS	Adj. R-squared:	0.787
Method:	Least Squares	F-statistic:	28.46
Date:	Thu, 27 Nov 2025	Prob (F-statistic):	1.59e-14
Time:	21:36:55	Log-Likelihood:	-53.911
No. Observations:	53	AIC:	123.8
Df Residuals:	45	BIC:	139.6
Df Model:	7		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	0.5599	0.278	2.011	0.050	-0.001	1.121
P5	0.7017	0.104	6.759	0.000	0.493	0.911
P6	0.1263	0.054	2.333	0.024	0.017	0.235
P7	-0.0058	0.082	-0.070	0.944	-0.172	0.160
P8	0.0922	0.083	1.108	0.274	-0.075	0.260
P9	0.0692	0.049	1.427	0.161	-0.029	0.167
P10	0.0027	0.074	0.036	0.972	-0.147	0.152
P11	-0.0631	0.076	-0.833	0.409	-0.216	0.090

Omnibus:	37.959	Durbin-Watson:	1.773
Prob(Omnibus):	0.000	Jarque-Bera (JB):	168.889
Skew:	-1.744	Prob(JB):	2.12e-37
Kurtosis:	11.019	Cond. No.	27.0

Figure 14 – OLS Regression Results Personal Dimension

Personal Dimension

OLS Regression Results

Dep. Variable:	P4	R-squared:	0.513
Model:	OLS	Adj. R-squared:	0.397
Method:	Least Squares	F-statistic:	4.419
Date:	Thu, 27 Nov 2025	Prob (F-statistic):	0.000290
Time:	21:23:55	Log-Likelihood:	-79.681
No. Observations:	53	AIC:	181.4
Df Residuals:	42	BIC:	203.0
Df Model:	10		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	1.7566	0.412	4.267	0.000	0.926	2.587
P12	0.3253	0.141	2.305	0.026	0.040	0.610
P13	-0.2397	0.168	-1.429	0.161	-0.578	0.099
P14	-0.1770	0.188	-0.940	0.353	-0.557	0.203
P15	-0.0741	0.113	-0.656	0.516	-0.302	0.154
P16	-0.1639	0.136	-1.209	0.233	-0.437	0.110
P17	0.4548	0.155	2.940	0.005	0.143	0.767
P18	0.1834	0.210	0.873	0.388	-0.241	0.608
P19	-0.0431	0.169	-0.255	0.800	-0.384	0.298
P20	-0.0008	0.177	-0.004	0.996	-0.358	0.357
P21	0.2697	0.212	1.271	0.211	-0.159	0.698

Omnibus:	25.344	Durbin-Watson:	1.670
Prob(Omnibus):	0.000	Jarque-Bera (JB):	73.469
Skew:	-1.205	Prob(JB):	1.11e-16
Kurtosis:	8.240	Cond. No.	26.8

Figure 15 – OLS Regression Results Social Dim

Social Dimension

OLS Regression Results

Dep. Variable:	P4	R-squared:	0.236
Model:	OLS	Adj. R-squared:	0.117
Method:	Least Squares	F-statistic:	1.981
Date:	Thu, 27 Nov 2025	Prob (F-statistic):	0.0788
Time:	21:23:55	Log-Likelihood:	-91.614
No. Observations:	53	AIC:	199.2
Df Residuals:	45	BIC:	215.0
Df Model:	7		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	3.0350	0.338	8.982	0.000	2.354	3.716
P22	-0.0569	0.149	-0.381	0.705	-0.358	0.244
P23	-0.0440	0.217	-0.203	0.840	-0.481	0.393
P24	0.2792	0.179	1.563	0.125	-0.081	0.639
P25	-0.0042	0.187	-0.023	0.982	-0.381	0.373
P26	0.0940	0.276	0.341	0.735	-0.461	0.649
P27	-0.1474	0.242	-0.608	0.546	-0.636	0.341
P28	0.2616	0.166	1.574	0.122	-0.073	0.596

Omnibus:	27.823	Durbin-Watson:	1.497
Prob(Omnibus):	0.000	Jarque-Bera (JB):	62.975
Skew:	-1.509	Prob(JB):	2.11e-14
Kurtosis:	7.406	Cond. No.	13.4

Figure 16 – OLS Regression Results Financial Dim

Financial Dimension

OLS Regression Results

Dep. Variable:	P4	R-squared:	0.220
Model:	OLS	Adj. R-squared:	0.189
Method:	Least Squares	F-statistic:	7.062
Date:	Thu, 27 Nov 2025	Prob (F-statistic):	0.00199
Time:	21:23:55	Log-Likelihood:	-92.140
No. Observations:	53	AIC:	190.3
Df Residuals:	50	BIC:	196.2
Df Model:	2		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	3.4540	0.228	15.157	0.000	2.996	3.912
P34	0.3358	0.123	2.739	0.009	0.090	0.582
P35	0.0845	0.114	0.738	0.464	-0.145	0.314

Omnibus:	23.394	Durbin-Watson:	1.584
Prob(Omnibus):	0.000	Jarque-Bera (JB):	39.264
Skew:	-1.423	Prob(JB):	2.98e-09
Kurtosis:	6.111	Cond. No.	3.41

Figure 17 – OLS Regression Results

All Dimensions

OLS Regression Results

Dep. Variable:	P4	R-squared:	0.901
Model:	OLS	Adj. R-squared:	0.802
Method:	Least Squares	F-statistic:	9.112
Date:	Thu, 27 Nov 2025	Prob (F-statistic):	1.42e-07
Time:	21:36:56	Log-Likelihood:	-37.419
No. Observations:	53	AIC:	128.8
Df Residuals:	26	BIC:	182.0
Df Model:	26		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	0.3804	0.374	1.018	0.318	-0.388	1.149
P5	0.8154	0.148	5.499	0.000	0.511	1.120
P6	0.1066	0.092	1.159	0.257	-0.083	0.296
P7	-0.0990	0.119	-0.834	0.412	-0.343	0.145
P8	0.0979	0.112	0.874	0.390	-0.132	0.328
P9	0.0457	0.066	0.697	0.492	-0.089	0.180
P10	0.0026	0.118	0.022	0.982	-0.239	0.245
P11	0.0014	0.105	0.013	0.990	-0.214	0.217
P12	0.1249	0.116	1.078	0.291	-0.113	0.363
P13	-0.0812	0.128	-0.633	0.532	-0.345	0.183
P14	-0.0868	0.154	-0.564	0.577	-0.403	0.229
P15	-0.0092	0.098	-0.093	0.926	-0.211	0.193
P16	-0.2670	0.113	-2.363	0.026	-0.499	-0.035
P17	0.1968	0.123	1.596	0.123	-0.057	0.450
P18	-0.1777	0.181	-0.980	0.336	-0.550	0.195
P19	-0.0487	0.128	-0.380	0.707	-0.312	0.214
P20	0.0392	0.153	0.256	0.800	-0.276	0.355
P21	-0.0449	0.155	-0.289	0.775	-0.364	0.275
P22	0.0393	0.093	0.421	0.677	-0.153	0.231
P23	-0.0271	0.175	-0.155	0.878	-0.386	0.332
P24	0.0522	0.136	0.384	0.704	-0.227	0.331
P25	0.0380	0.119	0.320	0.752	-0.207	0.283
P26	-0.1496	0.216	-0.691	0.496	-0.595	0.295
P27	0.1389	0.197	0.703	0.488	-0.267	0.545
P28	0.1955	0.112	1.746	0.093	-0.035	0.426
P34	0.0370	0.093	0.400	0.693	-0.153	0.227
P35	0.1027	0.095	1.086	0.288	-0.092	0.297

Omnibus:	1.005	Durbin-Watson:	2.102
Prob(Omnibus):	0.605	Jarque-Bera (JB):	0.361
Skew:	-0.075	Prob(JB):	0.835
Kurtosis:	3.376	Cond. No.	68.4

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação

Universidade Nova de Lisboa