

Inside and Beyond the Capital: Latent Profiles of Gay Men in Peru

Alejandro Aybar-Flores¹ and Elizabeth Espinoza-Portilla²

¹Universidad del Pacifico, Lima, Peru

²Universidad Continental, Lima, Peru

a.aybarf@alum.up.edu.pe

eespinozap@continental.edu.pe

Abstract: Research on sexual minority health in Peru has largely focused on the capital city, overlooking how geographic inequalities shape heterogeneity among gay men living outside Lima. This study addresses this gap by comparing latent sociodemographic and risk-related profiles of gay men residing in Lima and in Peru's inland regions, where structural constraints and social exclusion may operate differently. Using data from the 2017 First Virtual Survey for LGBTIQ+ People in Peru, we analyze two subsamples: gay men living in inland regions (n=657) and those residing in Lima (n=1,528). The study applies a two-stage unsupervised analytical framework across both settings. Behavioral, social, and structural indicators are summarized through dimensionality reduction techniques, followed by the evaluation of multiple clustering algorithms. Optimal solutions are selected based on internal validity metrics, and feature relevance is examined using Random Forest-based importance measures to support interpretability while preserving cross-regional comparability. Results reveal marked differences in both the number and composition of latent profiles across geographic contexts. In inland regions, the optimal solution identifies two clusters primarily differentiated by age and employment status, with secondary differences in access to STI/HIV information and health services. In contrast, the Lima subsample yields three distinct clusters, indicating greater internal heterogeneity. The largest Lima cluster is dominated by young adults with lower employment levels and higher prevalence of mental health problems and suicidal ideation, while older clusters display greater socioeconomic stability and lower psychosocial vulnerability. From a gender research perspective, these findings demonstrate that sexual minority inequalities are shaped less by individual risk behaviors than by gendered structural and spatial contexts. Geographic location influences not only levels of vulnerability but also the internal configuration of risk and protection among gay men in Peru. The study underscores the need for territorially differentiated research and interventions that move beyond capital-centered and behavior-focused frameworks.

Keywords: Unsupervised learning, Dimensionality reduction, Clustering, Regional inequality, Gay men, Peru

1. Introduction

Research on sexual minority health consistently shows that gay men experience disproportionate burdens in mental health, sexual and reproductive health, and social well-being compared to heterosexual populations (Frost and Svensson, 2018). These inequalities are largely driven by stigma, discrimination, and structural exclusion operating across multiple social levels (Hatzenbuehler, 2016). Yet, particularly in low- and middle-income countries, much of the literature has treated gay men as a relatively homogeneous group, focusing on isolated behavioral risks while neglecting the social and geographic contexts through which vulnerability is produced (Goldhammer and Mayer, 2011).

In Latin America, and especially in Peru, evidence on the health and social conditions of gay men remains limited and uneven (Abdo, 2020). Research has focused primarily on Lima, where greater visibility of sexual minorities and stronger institutional and civil society infrastructure facilitate data collection (INEI, 2017; Más Igualdad Perú, 2021). While these studies have yielded important insights into sexual behavior, HIV prevention, and mental health, their capital-centered focus may obscure the experiences of gay men living outside Lima, where social norms, institutional capacity, and access to resources often differ substantially (Romani et al., 2021). Consequently, geographic inequalities within the country remain insufficiently examined (Perez-Brumer et al., 2013).

Peru is characterized by marked regional disparities in economic development, health services, and social inclusion (Jaramillo, 2024). Inland regions often combine weaker health systems, fewer culturally competent services, and more conservative social environments, potentially intensifying stigma and limiting access to prevention and care for sexual minorities (Suppes et al., 2021). At the same time, vulnerability cannot be assumed to manifest uniformly across regions. Differences in age structure, employment opportunities, and migration trajectories may generate distinct configurations of risk and protection among gay men in Lima and inland areas (Perez-Brumer et al., 2013).

Despite these dynamics, few studies have examined whether the internal organization of gay men's social and health profiles varies by geographic context within Peru. Findings from Lima-based samples are often implicitly treated as generalizable, potentially obscuring important intra-group differences (Jaramillo, 2024). Recent

advances in unsupervised learning provide tools to address this gap by identifying latent, multidimensional profiles without imposing predefined categories (Scheer et al., 2021).

Building on this approach, the present study examines sociodemographic and risk-related profiles of gay men in Peru, comparing Lima and inland regions using data from the 2017 First Virtual Survey for LGBTIQ+ People in Peru. By applying a two-stage unsupervised analytical strategy separately in each context, the study shows how geography shapes not only levels of vulnerability but also the internal organization of risk and protection, contributing to a more context-sensitive understanding of sexual minority health and to gender research on the intersections of sexuality, age, and place.

2. Theoretical Framing

This study assumes that gay men are not a homogeneous population and that vulnerability and protection emerge from specific configurations of social, demographic, and contextual factors (Carnes, 2016). Rather than conceptualizing risk as isolated behaviors, the framework emphasizes how multiple dimensions combine to produce differentiated profiles shaped by place, age, and access to resources (Halkitis et al., 2015).

Geographic context is central to this approach. In Peru, resources relevant to sexual minority health are concentrated in Lima, while inland regions often face more limited health infrastructure, employment opportunities, and access to preventive information (Jaramillo, 2024). These conditions are treated as structuring forces that shape how risk-related characteristics cluster within populations, influencing both overall vulnerability and the internal composition of latent profiles among gay men (Hatzenbuehler, 2016; Hill et al., 2025). Age and socioeconomic position further condition these patterns. Employment, access to information, and engagement with health services vary across the life course, generating distinct trajectories of risk and stability (Fredriksen-Goldsen et al., 2017). These trajectories are mediated by geographic context, as regional environments shape opportunities for integration, information, and care (Lewis, 2009).

Analytically, the study adopts a configuration-based perspective, using dimensionality reduction and clustering techniques to capture vulnerability as the co-occurrence of social, behavioral, and structural factors within specific contexts, rather than as isolated determinants or fixed categories (Blondeel et al., 2021; Scheer et al., 2021).

3. Methodology

3.1 Study Sample

This study uses data from the 2017 First Virtual Survey for LGBTIQ+ People in Peru, a nationwide online survey conducted by the National Institute of Statistics and Informatics (INEI) across all 24 departments and the Constitutional Province of Callao (INEI, 2017). The survey was designed to collect information on sociodemographic characteristics, health conditions, sexual behaviors, discrimination experiences, and access to services among sexual and gender minorities. The analytic sample was restricted to respondents who self-identified as gay men. To enable geographic comparison, the data were divided into two subsamples based on place of residence: individuals living in inland regions of Peru and those residing in Lima, the capital of Peru. After data cleaning and filtering for complete cases on all selected variables, the final sample consisted of 657 gay men living in inland regions and 1,528 gay men residing in Lima.

3.2 Analytical Strategy

The study employs a two-stage unsupervised analytical framework designed to identify latent profiles without imposing predefined risk categories. Given the exploratory nature of the research question and the absence of established typologies of vulnerability among gay men across geographic contexts in Peru, an unsupervised learning approach was considered appropriate. Unlike supervised or regression-based models that require predefined outcomes or fixed categories, unsupervised methods allow patterns to emerge inductively from the data. This is particularly relevant where heterogeneity is expected but its internal structure remains unclear, and where a priori classifications may obscure meaningful configurations of social, structural, and behavioral characteristics. By identifying latent profiles without prespecifying risk groups or isolating single variables, the approach aligns with configuration-based and intersectional perspectives, enabling a more context-sensitive analysis of how vulnerability and protection are jointly organized across territorial settings.

In the first stage, dimensionality reduction techniques (including Principal Component Analysis [PCA], t-distributed Stochastic Neighbor Embedding [t-SNE], Isometric Mapping [Isomap], Uniform Manifold Approximation and Projection [UMAP], and Locally Linear Embedding [LLE]) were used to summarize the

information contained in the original set of variables into a smaller number of dimensions. This step serves to reduce noise, address redundancy among correlated indicators, and facilitate the identification of underlying structure in the data. Importantly, dimensionality reduction does not eliminate information but reorganizes it into a more compact representation that preserves the most relevant patterns.

In the second stage, clustering algorithms (including K-means, K-means++, hierarchical clustering, and a hybrid hierarchical–K-means approach) were applied to identify groups of individuals with similar profiles. To ensure robustness, clustering was performed both on the reduced representations and on the original high-dimensional data. This allowed for a direct comparison of solutions and helped assess whether dimensionality reduction improved cluster quality and interpretability. Across both stages, multiple configurations were tested, varying the number of clusters from two to ten.

3.3 Model Evaluation and Selection

Cluster solutions were evaluated using three complementary internal validation metrics: the Silhouette coefficient, the Calinski–Harabasz index, and the Davies–Bouldin index. Each metric captures a different aspect of clustering quality. The Silhouette coefficient reflects how well observations fit within their assigned clusters relative to others, the Calinski–Harabasz index assesses the balance between within-cluster compactness and between-cluster separation, and the Davies–Bouldin index measures cluster similarity, with lower values indicating better separation.

To avoid reliance on a single indicator, all metrics were normalized to a common scale and combined into an aggregated performance score. The normalized Silhouette and Calinski–Harabasz values were averaged together with the inverse of the Davies–Bouldin index, ensuring that higher aggregated scores consistently represented better-performing solutions. The optimal clustering configuration for each subsample was selected as the one that maximized this aggregated score across all tested combinations.

3.4 Cluster Characterization and Interpretation

Once the optimal clustering solution was identified, clusters were characterized by examining the distribution of all study variables across groups. Furthermore, chi-square tests of independence were conducted for each variable, comparing their distributions across clusters within each subsample. To further support interpretability, a feature importance analysis was conducted using a one-versus-rest strategy. In this approach, each cluster is compared against all others to identify the variables that most strongly differentiate it from the rest of the sample.

4. Results and Discussion

4.1 Univariate Analysis

The univariate analysis, as shown in Table 1, reveals both shared patterns and notable contrasts between gay men living in inland regions and those residing in Lima.

Table 1: Characteristics of gay men in inland regions and Lima

Variable	Label	Gay inland inhabitants		Gay inhabitants in Lima	
		n	%	n	%
Health insurance	Yes	447	68	1164	76.2
	No	210	32	364	23.8
Infectious disease history	No	438	66.7	880	57.6
	Yes	219	33.3	648	42.4
Mental health problems	No	383	58.3	956	62.6
	Yes	274	41.7	572	37.4
Condom use	Yes	599	91.2	1393	91.2
	No	58	8.8	135	8.8
Orientation expression	No	389	59.2	836	54.7
	Yes	268	40.8	692	45.3

Variable	Label	Gay inland inhabitants		Gay inhabitants in Lima	
		n	%	n	%
Employment	Yes	398	60.6	1013	66.3
	No	259	39.4	515	33.7
Sex work	No	527	80.2	1357	88.8
	Yes	130	19.8	171	11.2
Health assistance	Yes	545	83	1379	90.2
	No	112	17	149	9.8
Age	18–24	288	43.8	667	43.7
	25–29	269	40.9	517	33.8
	30–34	64	9.7	236	15.4
	35+	36	5.5	108	7.1
Discrimination	Yes	526	80.1	1232	80.6
	No	131	19.9	296	19.4
Education level	Higher education	334	50.8	842	55.1
	Secondary education	308	46.9	677	44.3
	Primary education	11	1.7	8	0.5
	No education	4	0.6	1	0.1
STI/HIV information	Yes	510	77.6	1285	84.1
	No	147	22.4	243	15.9
Number of sexual partners	None	392	59.7	894	58.5
	One	250	38.1	620	40.6
	More than one	15	2.3	14	0.9
Ethnicity	Mixed	567	86.3	1395	91.3
	Indigenous	66	10	69	4.5
	White	24	3.7	64	4.2
Sexual violence	No	529	80.5	1237	81
	Yes	128	19.5	291	19
Suicide ideation	No	357	54.3	933	61.1
	Yes	300	45.7	595	38.9
Psychological support	No	418	63.6	918	60.1
	Yes	239	36.4	610	39.9
Substance abuse	No	474	72.1	1134	74.2
	Yes	183	27.9	394	25.8

4.2 Framework Experiments

Each dimensionality reduction technique evaluated in Stage 1 involved its own set of parameters, which were systematically combined with the clustering algorithms and configurations assessed in Stage 2. By exploring all relevant combinations of dimensionality reduction methods, clustering algorithms, and numbers of clusters, the analytical framework generated a large set of candidate solutions. In total, 13,716 clustering experiments were conducted for the subsample of gay men living in inland regions, while 13,680 experiments were performed for the subsample residing in Lima.

Comparison of Clustering Techniques through Dimensionality Reduction

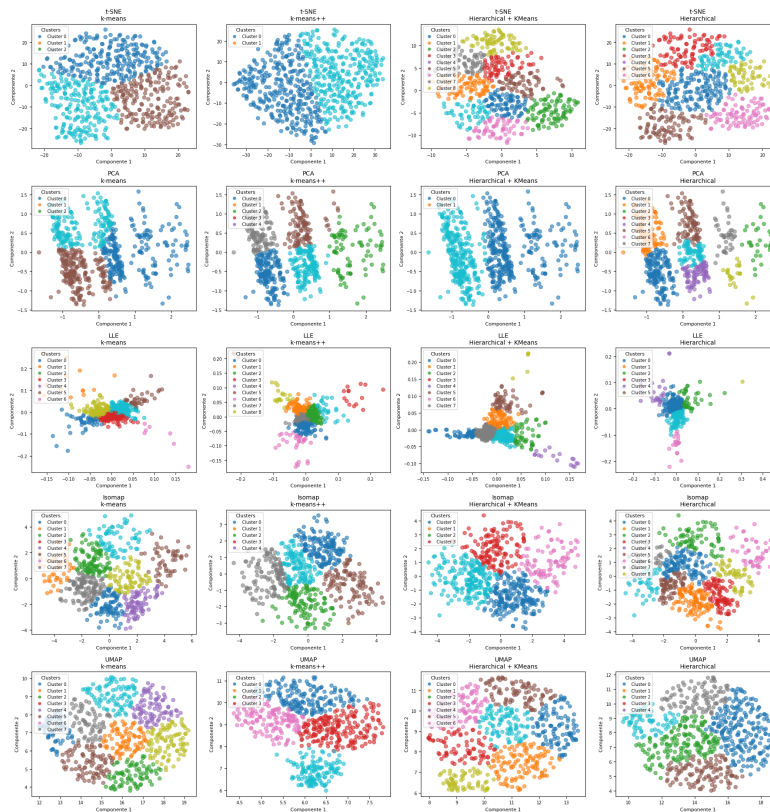


Figure 1: Visualization of a subset of experiments

In that sense, Figure 1 shows visualizations of a subset of experiments using the specified dimensionality reduction and clustering techniques. Linear methods such as PCA tend to produce more overlapping groupings, whereas non-linear approaches like UMAP, t-SNE, and Isomap reveal more compact and visually separable clusters. Moreover, to assess the relative performance of different parameter combinations, aggregated scores from all experiments were calculated. The top 5 configurations are shown in Table 2.

Table 2: Top-ranked tested configurations

	Rank	Stage 1: Reduction		Stage 2: Clustering		Metrics Evaluation			
		Technique	Dimension	Method	k	Silhouette	Calinski-Harabasz	Davies-Bouldin	Final Score
Lima	1	UMAP	3	Hierarchical + KMeans	3	1	1	1	1
	2	UMAP	2	KMeans ++	3	1	1	1	1
	3	UMAP	2	KMeans	3	1	1	1	1
	4	UMAP	3	Hierarchical	3	1	1	0.998	0.99933
	5	UMAP	3	KMeans ++	3	0.987	0.991	0.992	0.99030
Inland	1	UMAP	2	Hierarchical	2	1	1	1	1
	2	UMAP	3	Hierarchical + KMeans	2	1	1	1	1
	3	Isomap	2	Hierarchical + KMeans	4	1	1	0.996	0.99873
	4	Isomap	2	Hierarchical + KMeans	2	1	1	0.996	0.99873
	5	Isomap	2	Hierarchical + KMeans	2	1	1	0.996	0.99873

For the Lima subsample, the highest-ranked configurations consistently identified a three-cluster solution, with UMAP emerging as the dominant dimensionality reduction technique across top-performing models. In inland regions, optimal configurations consistently favored a two-cluster structure, with UMAP and Isomap yielding the strongest performance. Figure 2 displays the optimal clustering configurations identified for each subsample: for gay men living in inland regions, UMAP with a two-dimensional representation combined with

hierarchical clustering and a two-cluster solution; and for gay men residing in Lima, UMAP with a three-dimensional representation combined with a hybrid hierarchical–K-means approach yielding three clusters.

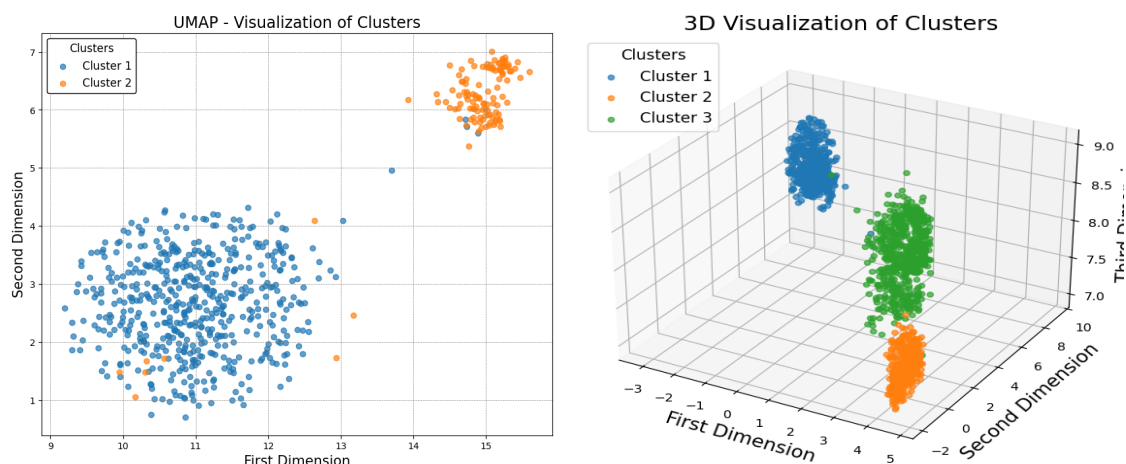


Figure 2: Distribution of optimal clusters for (left to right): (a) inland and (b) main city gay inhabitants

4.3 Clustering Results

Table 3 illustrates each identified cluster’s risk behavior and socio-demographic profiles for the study sample.

Table 3: Characteristics across clusters among gay men

		Inland			Lima			
		Clusters		X ²	Clusters			X ²
Variables	Labels	C1 (n = 536)	C2 (n = 121)	p-value	C1 (n = 541)	C2 (n = 357)	C3 (n = 630)	p-value
Health insurance	Yes	182 (34%)	28 (23%)	< 0.01	159 (29%)	74 (21%)	131 (21%)	< 0.01
	No	354 (66%)	93 (77%)		382 (71%)	283 (79%)	499 (79%)	
Infectious disease history	No	361 (67%)	77 (64%)	< 0.01	344 (64%)	204 (57%)	332 (53%)	< 0.01
	Yes	175 (33%)	44 (36%)		197 (36%)	153 (43%)	298 (47%)	
Mental health problems	No	308 (57%)	75 (62%)	< 0.01	276 (51%)	269 (75%)	411 (65%)	< 0.01
	Yes	228 (43%)	46 (38%)		265 (49%)	88 (25%)	219 (35%)	
Condom use	Yes	45 (8%)	13 (11%)	< 0.01	49 (9%)	35 (10%)	51 (8%)	< 0.01
	No	491 (92%)	108 (89%)		492 (91%)	322 (90%)	579 (92%)	
Orientation expression	No	326 (61%)	63 (52%)	< 0.01	300 (55%)	190 (53%)	346 (55%)	< 0.01
	Yes	210 (39%)	58 (48%)		241 (45%)	167 (47%)	284 (45%)	
Employment	Yes	226 (42%)	33 (27%)	< 0.01	291 (54%)	75 (21%)	149 (24%)	< 0.01
	No	310 (58%)	88 (73%)		250 (46%)	282 (79%)	481 (76%)	
Sex work	No	434 (81%)	93 (77%)	< 0.01	466 (86%)	322 (90%)	569 (90%)	< 0.01
	Yes	102 (19%)	28 (23%)		75 (14%)	35 (10%)	61 (10%)	
Health assistance	Yes	94 (18%)	18 (15%)	< 0.01	83 (15%)	17 (5%)	49 (8%)	< 0.01
	No	442 (82%)	103 (85%)		458 (85%)	340 (95%)	581 (92%)	
Age	18–24	288 (54%)	0 (0%)	< 0.01	500 (92%)	0 (0%)	17 (3%)	< 0.01
	25–29	243 (45%)	26 (21%)		41 (8%)	26 (7%)	600 (95%)	
	30–34	5 (1%)	59 (49%)		0 (0%)	223 (62%)	13 (2%)	
	35+	0 (0%)	36 (30%)		0 (0%)	108 (30%)	0 (0%)	
Discrimination	Yes	106 (20%)	25 (21%)	< 0.01	95 (18%)	94 (26%)	107 (17%)	< 0.01
	No	430 (80%)	96 (79%)		446 (82%)	263 (74%)	523 (83%)	

Variables	Labels	Inland			Lima			
		Clusters		X ²	Clusters			X ²
		C1 (n = 536)	C2 (n = 121)	p-value	C1 (n = 541)	C2 (n = 357)	C3 (n = 630)	p-value
Education level	Higher education	3 (1%)	1 (1%)	< 0.01	1 (0%)	0 (0%)	0 (0%)	< 0.01
	Secondary education	8 (1%)	3 (2%)		6 (1%)	2 (1%)	0 (0%)	
	Primary education	289 (54%)	45 (37%)		390 (72%)	97 (27%)	190 (30%)	
	No education	236 (44%)	72 (60%)		144 (27%)	258 (72%)	440 (70%)	
STI/HIV information	Yes	130 (24%)	17 (14%)	< 0.01	110 (20%)	39 (11%)	94 (15%)	< 0.01
	No	406 (76%)	104 (86%)		431 (80%)	318 (89%)	536 (85%)	
Number of sexual partners	None	324 (60%)	68 (56%)	< 0.01	354 (65%)	180 (50%)	360 (57%)	< 0.01
	One	204 (38%)	46 (38%)		182 (34%)	173 (48%)	265 (42%)	
	More than one	8 (1%)	7 (6%)		5 (1%)	4 (1%)	5 (1%)	
Ethnicity	Mixed	57 (11%)	9 (7%)	< 0.01	18 (3%)	11 (3%)	35 (6%)	< 0.01
	Indigenous	462 (86%)	105 (87%)		500 (92%)	332 (93%)	563 (89%)	
	White	17 (3%)	7 (6%)		23 (4%)	14 (4%)	32 (5%)	
Sexual violence	No	428 (80%)	101 (83%)	< 0.01	438 (81%)	281 (79%)	518 (82%)	< 0.01
	Yes	108 (20%)	20 (17%)		103 (19%)	76 (21%)	112 (18%)	
Suicide ideation	No	284 (53%)	73 (60%)	< 0.01	293 (54%)	243 (68%)	397 (63%)	< 0.01
	Yes	252 (47%)	48 (40%)		248 (46%)	114 (32%)	233 (37%)	
Psychological support	No	341 (64%)	77 (64%)	< 0.01	330 (61%)	215 (60%)	373 (59%)	< 0.01
	Yes	195 (36%)	44 (36%)		211 (39%)	142 (40%)	257 (41%)	
Substance abuse	No	390 (73%)	84 (69%)	< 0.01	408 (75%)	272 (76%)	454 (72%)	< 0.01
	Yes	146 (27%)	37 (31%)		133 (25%)	85 (24%)	176 (28%)	

Among gay men living in inland regions, the clustering solution reveals a relatively simple structure composed of two uneven groups. The larger cluster consists mainly of younger, economically active individuals with greater access to STI/HIV information and health services, while the smaller cluster is older, less likely to be employed, and faces more limited access to health information and care. In Lima, the clustering solution identifies greater internal heterogeneity, with three distinct profiles. The largest cluster is dominated by young adults and characterized by elevated levels of mental health problems and suicidal ideation, alongside lower employment rates, despite high access to condom use and STI/HIV information. A second cluster comprises older, socioeconomically stable individuals with higher employment and substantially better mental health outcomes. The third cluster occupies an intermediate position, combining moderate employment levels with persistent mental health challenges and higher reported histories of infectious disease. Moreover, the feature importance scores, presented in Figures 3 and 4, were calculated by developing a Random Forest model for each identified cluster. Variables are ordered by importance from top to bottom.

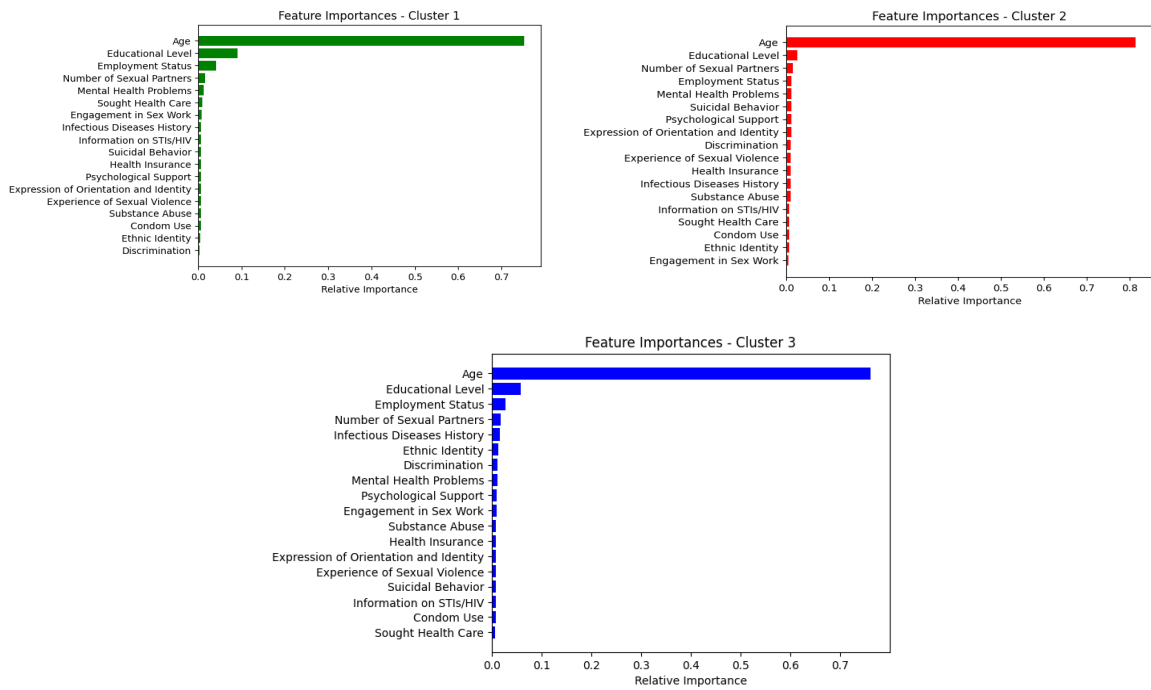


Figure 3: Variable importance plots for main city inhabitants' clusters

Among gay men living in Lima, age emerged as the most influential variable across all three clusters in the Random Forest feature importance analysis, underscoring its central role in differentiating profiles within the urban context. In Cluster 1, age dominates, with employment status and educational level contributing at a secondary level. Cluster 2 is even more clearly structured by age, with minimal influence from other variables, suggesting a profile largely defined by life stage. In Cluster 3, age remains the primary driver, while employment status and educational level show slightly greater relevance, indicating a closer interaction between age and socioeconomic positioning. Across clusters, behavioral and psychosocial factors such as mental health problems, substance use, discrimination, and sexual violence exhibit consistently low importance, reflecting that cluster differentiation is driven primarily by age-related and socioeconomic dimensions rather than risk behaviors.

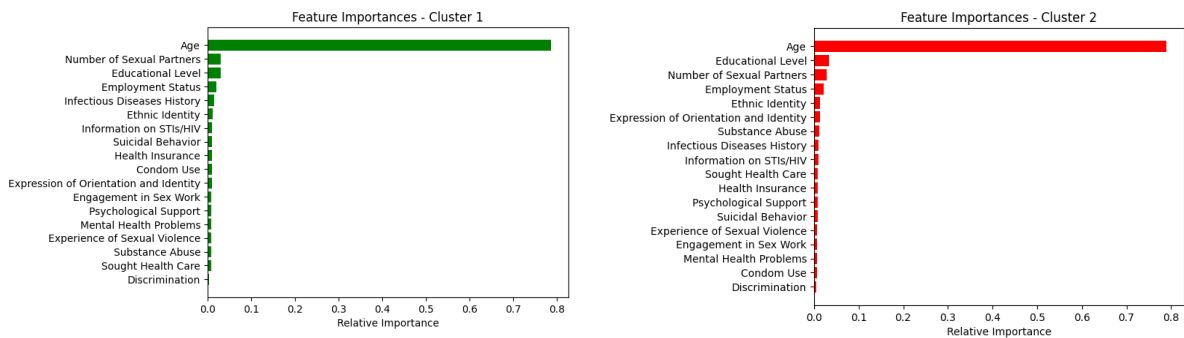


Figure 4: Variable importance plots for inland inhabitants' clusters

Regarding gay inland inhabitants, in both clusters, age emerged as the most influential variable in the Random Forest model, indicating its central role in differentiating risk behaviors and access to resources. Cluster 1 is characterized by a larger share of individuals with stable risk employment, higher access to STI/HIV information, and a greater proportion reporting multiple sexual partners. In contrast, Cluster 2 is composed mainly of younger participants, with a higher representation of those under 25 years old, fewer sexual partners overall, and lower levels of access to STI/HIV information. Although variables like educational level, number of sexual partners, and employment status also contributed moderately to cluster differentiation, most other factors, such as mental health issues, substance abuse, and discrimination, had relatively lower importance.

Geographic context emerges in this study as a key factor shaping not only levels of vulnerability among gay men in Peru but also the ways in which sociodemographic, behavioral, and psychosocial dimensions are

internally configured (Hatzenbuehler, 2016). The concentration of resources, services, and visibility in Lima appears to generate greater internal differentiation, producing distinct life-stage-based profiles in which younger gay men experience heightened psychosocial vulnerability despite relatively high access to preventive information (Krueger et al., 2016). This pattern aligns with broader evidence linking early adulthood, labor market precarity, and minority stress to adverse mental health outcomes among sexual minorities (Fredriksen-Goldsen et al., 2017). In contrast, the simpler and more polarized structure observed in inland regions suggests that structural constraints related to employment, access to health services, and information availability may play a more dominant role in shaping vulnerability, limiting the range of distinct configurations that emerge (Jaramillo, 2024). Across both contexts, age consistently functions as a central organizing axis, reinforcing the relevance of life-course perspectives for understanding sexual minority health (Carnes, 2016). Importantly, the limited role of behavioral and psychosocial variables in differentiating clusters, as indicated by feature importance analyses, suggests that vulnerability among gay men in Peru may be driven less by individual risk behaviors than by broader structural and socioeconomic conditions (Hatzenbuehler, 2016). These findings support calls to move beyond behavior-centered and capital-focused frameworks and to adopt more context-sensitive approaches that account for how place, age, and structural position interact to shape lived experiences among sexual minorities (Abdo, 2020).

4.3.1 Implications for gender research: Sexuality, place, and structural inequality

Beyond identifying latent profiles, the results contribute to gender research by showing that vulnerability among gay men is socially and spatially produced rather than primarily behavior-driven. Differences in clustering structures between Lima and inland regions indicate that geographic context operates as a gendered social structure, shaping both access to resources and internal organization of risk.

In Lima, greater institutional density, sexual minority visibility, and labor market segmentation generate differentiated life-course trajectories. The presence of a large cluster of young gay men with elevated mental health problems and suicidal ideation, despite high levels of HIV/STI information and condom use, reveals a gendered paradox: informational access alone does not ensure psychosocial well-being under conditions of structural precarity and minority stress. Conversely, inland regions display a more polarized and less differentiated clustering structure. Vulnerability is primarily organized around structural constraints related to employment and access to health services and information, suggesting that weaker institutional environments limit the range of social positions available to gay men and reinforce divides between protected and structurally marginalized groups.

Across both contexts, behavioral indicators play a secondary role in profile differentiation, while age, employment, and access to institutional resources emerge as the main organizing features. This challenges behavior-centered paradigms in sexual minority health research and supports structural and intersectional approaches that frame vulnerability as a product of gendered social positioning within specific territorial contexts.

5. Conclusion

This study shows that gay men in Peru are not a homogeneous population and that geographic context decisively shapes patterns of vulnerability and protection. Unsupervised analyses conducted separately for Lima and inland regions reveal differences not only in overall conditions but also in the internal structure of heterogeneity.

In inland regions, vulnerability is organized in a relatively simple and polarized manner, largely shaped by age, employment status, and unequal access to health information and services. Lima, by contrast, displays greater complexity, with profiles reflecting more diverse life-course trajectories. Within this setting, younger gay men emerge as particularly vulnerable, marked by weak labor market integration and elevated mental health problems and suicidal ideation despite broad access to preventive information. Across both contexts, age consistently operates as a central axis of differentiation, highlighting the importance of life-course dynamics. From a gender research perspective, these findings underscore that sexual minority inequalities cannot be understood without attention to spatial and structural contexts. Policies targeting gay men in Peru should therefore be territorially differentiated and sensitive to life-course processes, moving beyond behavior-centered approaches. Gender-sensitive public health and social policies must explicitly account for how place and structural exclusion shape lived experiences.

As the analysis is based on a nationwide online survey, potential selection biases related to unequal internet access should be acknowledged. Despite its broad geographic reach, the survey may underrepresent gay men with limited digital connectivity, particularly in inland regions where infrastructural inequalities are more pronounced. As a result, the identified clusters likely reflect the experiences of individuals who are more connected to online networks and information channels. The profiles should be interpreted as context-specific configurations within the surveyed population rather than as fully generalisable representations of all gay men in each region. Nonetheless, the comparative design remains informative for examining how heterogeneity is structured under different territorial conditions.

Ethics declaration: Ethical clearance wasn't required for the research.

AI declaration: No AI tools were used in the creation of this paper.

References

- Abdo, C.H.N. (2020) "Latin American and Latina/Latino issues in sexual health", in Rowland, D. and Jannini, E. (eds.), *Cultural differences and the practice of sexual medicine*, Springer, Cham, pp. 183–205.
- Blondeel, K., Dias, S., Furegato, M., Seuc, A., Gama, A., Fuertes, R., Mendão, L., Temmerman, M. and Toskin, I. (2021) "Sexual behaviour patterns and STI risk: Results of a cluster analysis among men who have sex with men in Portugal", *BMJ Open*, Vol. 11, No. 2, e033290.
- Carnes, N. (2016) "Gay men and men who have sex with men: Intersectionality and syndemics", in Wright, E. and Carnes, N. (eds.), *Understanding the HIV/AIDS epidemic in the United States: Social disparities in health and health care*, Springer, New York, pp. 43–69.
- Fredriksen-Goldsen, K.I., Bryan, A.E.B., Jen, S., Goldsen, J., Kim, H.-J. and Muraco, A. (2017) "The unfolding of LGBT lives: Key events associated with health and well-being in later life", *The Gerontologist*, Vol. 57, No. 1, pp. S15–S29.
- Frost, D.M. and Svensson, M. (2018) "Sexual minority health", in Revenson, T.A. and Gurung, R. (eds.), *Handbook of health psychology*, Routledge, New York, pp. 312–324.
- Goldhammer, H. and Mayer, K.H. (2011) "Focusing on sexual health promotion to enhance preventive behaviors among gay men and other men who have sex with men: Report from a state-of-the-art conference", *AIDS and Behavior*, Vol. 15, No. 1, pp. S1–S8.
- Halkitis, P.N., Kapadia, F., Ompad, D.C. and Perez-Figueroa, R. (2015) "Moving toward a holistic conceptual framework for understanding healthy aging among gay men", *Journal of Homosexuality*, Vol. 62, No. 5, pp. 571–587.
- Hatzenbuehler, M.L. (2016) "Structural stigma and health inequalities: Research evidence and implications for psychological science", *American Psychologist*, Vol. 71, No. 8, pp. 742–751.
- Hill, A.O., Ruppanner, L., Armstrong, G., Lafferty, L., Bavinton, B.R., Gilmour, S., Bourne, A., Amos, N. and Kaneko, N. (2025) "Social correlates of social capital access among gay and bisexual men in Japan", *Japanese Journal of Sociology*.
- INEI (2017) *Primera Encuesta Virtual para Personas LGBTI 2017: Principales resultados*, Instituto Nacional de Estadística e Informática, Lima.
- Jaramillo Lizana, J. (2024) "Regional inequality in Peru: Causes, effects, and strategies for equitable development", SSRN Working Paper, Lima.
- Krueger, E.A., Chiu, C.J., Menacho, L.A. and Young, S.D. (2016) "HIV testing among social media-using Peruvian men who have sex with men: Correlates and social context", *AIDS Care*, Vol. 28, No. 10, pp. 1301–1305.
- Lewis, N.M. (2009) "Mental health in sexual minorities: Recent indicators, trends, and their relationships to place in North America and Europe", *Health & Place*, Vol. 15, No. 4, pp. 1029–1045.
- Li, W., Zhang, Q. and Wang, X. (2025) "Interpretable clustering with adaptive heterogeneous causal structure learning in mixed observational data", *ArXiv Working Paper*, arXiv:2509.04415.
- Mas Igualdad Perú (2021) *Salud mental de personas LGBTQ+ en Perú: Informe final*, Mas Igualdad Perú, Lima.
- Perez-Brumer, A.G., Konda, K.A., Salvatierra, H.J., Segura, E.R. and Hall, E.R. (2013) "Prevalence of HIV, STIs, and risk behaviors in a cross-sectional community- and clinic-based sample of men who have sex with men in Lima, Peru", *PLOS ONE*, Vol. 8, No. 4, e59072.
- Romani, L., Ladera-Porta, K., Quiñones-Laveriano, D.M., Rios-Garcia, W., Juarez-Ubillus, A. and Vilchez-Cornejo, J. (2021) "Factors associated with the non-use of health services in LGBTI people from Peru", *Revista Peruana de Medicina Experimental y Salud Pública*, Vol. 38, No. 2, pp. 240–247.
- Salazar-Vizcaya, L., Nicca, D., Christinet, V., Kouyos, R.D., Vock, F., Andresen, S., Lehner, A., Haerry, D., Günthard, H.F., Schmidt, A.J. and Rauch, A. (2025) "Machine learning identifies sexual behavior subgroups among men who have sex with men in Switzerland", *Archives of Sexual Behavior*, Vol. 54, No. 6, pp. 2043–2055.
- Scheer, J.R., Clark, K.A., Maiolatesi, A.J. and Pachankis, J.E. (2021) "Syndemic profiles and sexual minority men's HIV-risk behavior: A latent class analysis", *Archives of Sexual Behavior*, Vol. 50, No. 8, pp. 2825–2841.
- Suppes, A., van der Toorn, J. and Begeny, C.T. (2021) "Unhealthy closets, discriminatory dwellings: The mental health benefits and costs of being open about one's sexual minority status", *Social Science & Medicine*, Vol. 285, 114286.