

Workforce Diversity Analysis Report

Exploring Demographic Patterns Across Sectors and Industries

Corporate Data Science Team

November 2024

Contents

Executive Summary 4

1 Which sectors have similar demographic structures across gender and ethnicity? 5

1.1 Analysis Method 5

1.2 Findings 5

1.2.1 Cluster 1: Service & Manufacturing Sectors 5

1.2.2 Cluster 2: Traditional Industries 5

1.3 Key Insights 5

1.4 Visualizations 6

2 Which sectors are most gender-imbalanced and how does that relate to total workforce size? 7

2.1 Most Gender-Imbalanced Sectors 7

2.2 Gender Balance Leaders 7

2.3 Relationship with Workforce Size 7

2.4 Analysis of Gender Patterns 7

3 Which sectors have demonstrated the most notable shifts in workforce diversity between 2021 and 2023? 9

3.1 Top 5 Sectors by Total Diversity Shift 9

3.2 Detailed Analysis of Top Shifters 9

3.2.1 Mining, Quarrying, and Oil & Gas Extraction 9

3.2.2 Leisure & Hospitality 9

3.2.3 Public Administration 9

3.2.4 Overall Trends 10

4 Which sectors are dominated by a single ethnic group and how extreme is this imbalance? 11

4.1 Ethnic Dominance Ranking 11

4.2 Sectors with EXTREME Ethnic Dominance 11

4.3 Concentration Severity Assessment 11

4.4 Critical Analysis 11

4.5 Policy Implications 12

5 Which sectors show the highest ethnic diversity and how does this relate to sector size? 13

5.1 Ethnic Diversity Rankings 13

5.2	Least Diverse Sectors	13
5.3	Relationship with Sector Size	13
5.4	Most Diverse Sector Profile	14
5.5	Analysis of Diversity Patterns	14
6	Can we group industries based on similar ethnic compositions?	15
6.1	Ethnic Composition Clustering	15
6.1.1	Ethnic Cluster 1: Mainstream Diversity Pattern	15
6.1.2	Ethnic Cluster 2: High Concentration Pattern	15
6.2	Hierarchical Clustering Insights	15
6.3	Synthesis of Ethnic Patterns	15
7	Synthesis and Strategic Recommendations	17
7.1	Priority Matrix	17
7.1.1	Sectors with Critical Diversity Challenges	17
7.1.2	Sectors with Moderate Diversity Challenges	17
7.1.3	Sectors with Strong Diversity Trends	17
7.2	Strategic Recommendations	18
7.2.1	Launch “Construction Diversity Initiative”	18
7.2.2	Establish Agriculture & Mining Task Force	18
7.2.3	Replicate Success Models	18
7.2.4	Address the “Pipeline Myth”	18
7.2.5	Set Ethnic Diversity Targets	18
7.2.6	Implement Intersectional Analysis	18
7.2.7	Investigate Mining Sector Volatility	19
7.2.8	Hispanic/Latino Strategic Focus	19
7.2.9	Extend Analysis to Leadership	19
7.2.10	Create Accountability Mechanisms	19
7.3	Measuring Success	19
8	Methodology Appendix	20
8.1	Data Sources	20
8.2	Statistical Methods	20
8.2.1	Clustering Analysis	20
8.2.2	Diversity Metrics	20
8.2.3	Temporal Analysis	20
8.2.4	Correlation Analysis	20
8.3	Limitations	20

Conclusion

22

Executive Summary

This report analyzes workforce demographic patterns across sectors to identify diversity trends, clustering patterns, and areas requiring policy attention. The analysis reveals significant gender and ethnic imbalances in specific sectors, with notable shifts occurring between 2021 and 2023. Our analysis identified two distinct demographic clusters across the thirteen sectors examined, with Construction exhibiting the most extreme gender imbalance at 89% male. Five sectors demonstrate extreme ethnic concentration exceeding 80% White representation, while the Mining sector experienced the largest diversity shift between 2021 and 2023. Transportation and utilities emerged as the sector with the highest ethnic diversity, setting a positive example for other industries.

1 Which sectors have similar demographic structures across gender and ethnicity?

1.1 Analysis Method

Using hierarchical clustering and K-means analysis on standardized demographic data encompassing both gender and ethnicity, we identified natural groupings of sectors with similar diversity compositions. This multivariate approach reveals patterns that transcend simple categorical comparisons.

1.2 Findings

The analysis yielded an optimal clustering solution of two distinct groups with a silhouette score of 0.472, indicating moderate cluster cohesion. These clusters reflect fundamental differences in industry characteristics and workforce composition patterns.

1.2.1 Cluster 1: Service & Manufacturing Sectors

This cluster encompasses ten sectors including Education and health services, Financial activities, Information, Leisure and hospitality, Manufacturing, Other services, Professional and business services, Public administration, Transportation and utilities, and Wholesale and retail trade. The cluster demonstrates an average composition of 44.7% women with ethnic representation at 76.9% White, 12.7% Black/African American, 6.6% Asian, and 16.3% Hispanic/Latino. With an average workforce size of approximately 2,012k employees, these sectors exhibit more gender-balanced and ethnically diverse characteristics compared to traditional industries.

1.2.2 Cluster 2: Traditional Industries

The second cluster comprises three sectors: Agriculture, forestry, fishing and hunting; Construction; and Mining, quarrying, and oil and gas extraction. This group exhibits markedly different demographics with only 17.7% women on average, representing severe male dominance. Ethnic composition shows 89.4% White, 5.0% Black/African American, 3.3% Asian, and 23.3% Hispanic/Latino representation. Despite an average workforce size of approximately 4,167k employees, these sectors maintain extreme gender imbalance and high ethnic concentration.

1.3 Key Insights

The clustering analysis reveals that demographic patterns correlate more strongly with industry type than workforce size. Traditional blue-collar industries form a distinct demographic pattern characterized by male dominance and limited ethnic diversity. In contrast, service and knowledge-based sectors demonstrate more similar diversity profiles despite significant variations in their workforce sizes, suggesting that industry culture and job characteristics play a more determinative role than organizational scale.

1.4 Visualizations

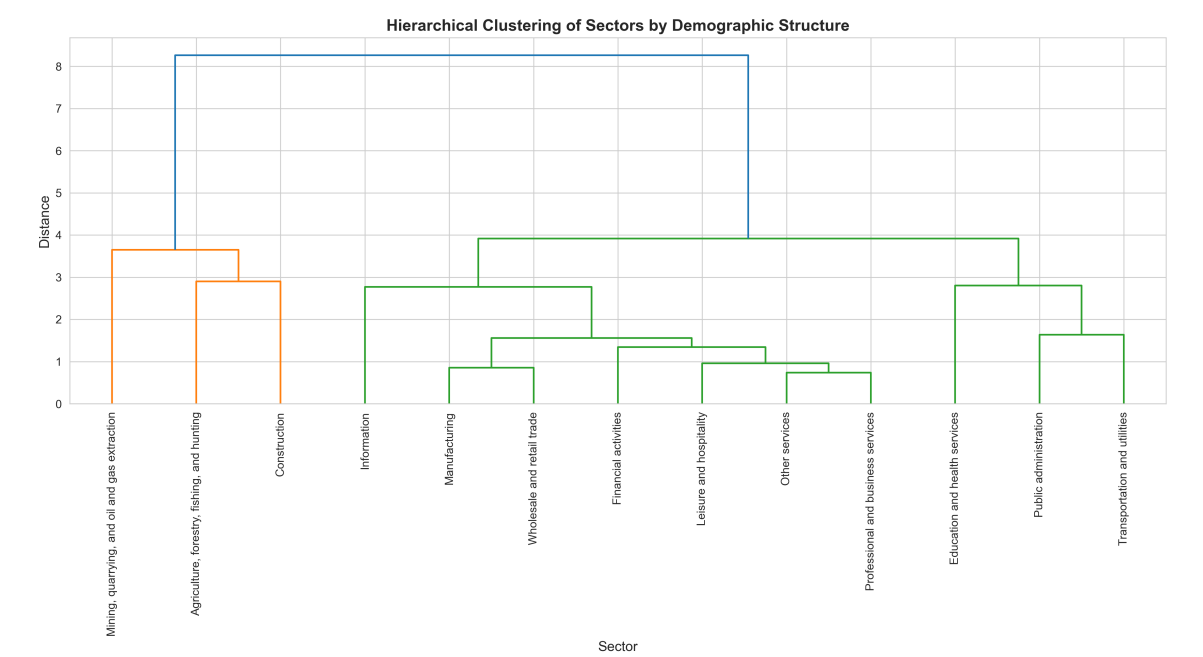


Figure 1: Hierarchical Clustering of Sectors by Demographic Structure

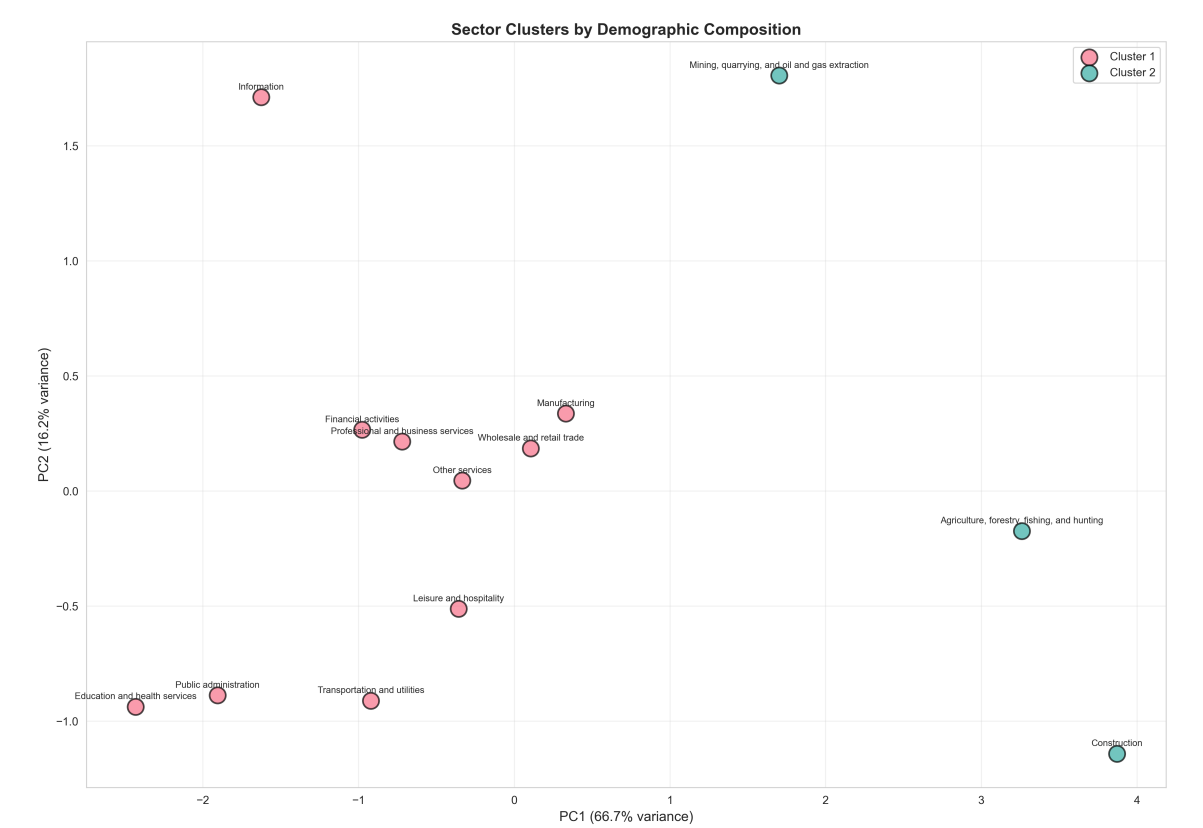


Figure 2: Sector Clusters by Demographic Composition (PCA Visualization)

2 Which sectors are most gender-imbalanced and how does that relate to total workforce size?

2.1 Most Gender-Imbalanced Sectors

Rank	Sector	% Women	Gender Imbalance	Workforce Size (k)
1	Construction	10.9%	39.1%	11,436
2	Mining, quarrying, oil & gas	15.1%	34.9%	276
3	Transportation & utilities	24.6%	25.4%	1,534
4	Education & health services	74.2%	24.2%	4,880
5	Agriculture, forestry, fishing	27.0%	23.0%	790

Table 1: Most Gender-Imbalanced Sectors

2.2 Gender Balance Leaders

Sector	% Women	Imbalance
Leisure & hospitality	48.5%	1.5%
Other services	47.1%	2.9%
Financial activities	46.8%	3.2%

Table 2: Sectors with Best Gender Balance

2.3 Relationship with Workforce Size

The correlation coefficient of 0.412 indicates a moderate positive relationship between sector size and gender imbalance, though the association is not particularly strong. This suggests that larger sectors tend to have slightly higher gender imbalance, yet the relationship’s weakness indicates that scale alone does not determine gender balance outcomes. Notably, some of the most imbalanced sectors, particularly Construction and Education & health services, rank among the largest by workforce size, demonstrating that organizational scale does not automatically drive gender balance improvements.

2.4 Analysis of Gender Patterns

Construction stands out with the most extreme imbalance, employing nearly nine in ten male workers. This represents not only the highest gender disparity but also affects the largest workforce among imbalanced sectors at 11.4 million employees. Conversely, Education & health services represents the only large female-dominated sector, with three in four workers being women. Only Leisure & hospitality approaches true gender parity at 48.5% women, demonstrating that balance is achievable but rare. The data strongly suggests that workforce size is not a reliable predictor of gender balance, with cultural and historical factors appearing far more influential in determining sectoral gender composition.

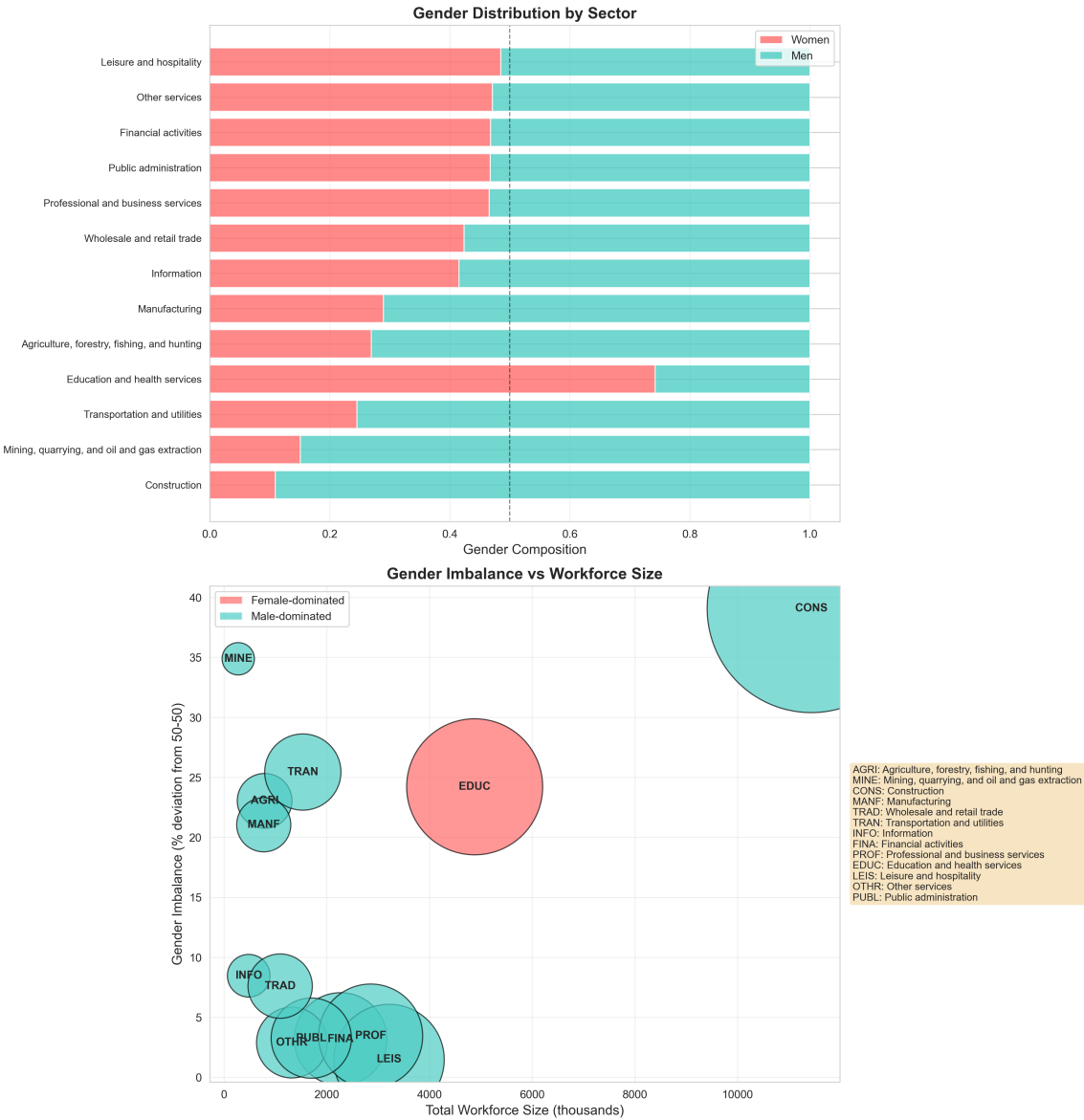


Figure 3: Gender Distribution and Imbalance Analysis

3 Which sectors have demonstrated the most notable shifts in workforce diversity between 2021 and 2023?

3.1 Top 5 Sectors by Total Diversity Shift

Rank	Sector	Total Shift Magnitude
1	Mining, quarrying, oil & gas extraction	0.169
2	Leisure & hospitality	0.040
3	Public administration	0.036
4	Financial activities	0.034
5	Agriculture, forestry, fishing	0.033

Table 3: Sectors with Largest Diversity Shifts (2021-2023)

3.2 Detailed Analysis of Top Shifters

3.2.1 Mining, Quarrying, and Oil & Gas Extraction

The mining sector experienced dramatic demographic volatility over the two-year period. While women’s representation decreased slightly by 0.13 percentage points, ethnic composition showed wild swings, with White representation increasing by 0.88pp and Black/African American by 0.45pp. Most notably, Asian representation plummeted by 15.72 percentage points, suggesting either a data quality issue or major workforce restructuring. Hispanic/Latino representation grew substantially by 6.22pp. This volatility likely stems from the sector’s small workforce size of 276k, where individual hiring or departure events create outsized statistical impacts.

3.2.2 Leisure & Hospitality

This sector demonstrated post-pandemic workforce shifts consistent with the industry’s recovery trajectory. Women’s representation declined by 1.01pp while White representation fell by 1.79pp, offset by gains in Black/African American representation (+2.01pp) and Hispanic/Latino representation (+2.83pp). These patterns reflect increasing racial diversity, likely driven by pandemic-related labor market disruptions and subsequent recovery hiring patterns that drew from more diverse talent pools.

3.2.3 Public Administration

Government diversity initiatives appear to be bearing measurable fruit in the public administration sector. While women’s representation decreased by 1.00pp, White representation fell significantly by 1.77pp. The sector achieved the largest gain in Black/African American representation at 2.73pp, demonstrating substantial progress in Black representation within government roles. Hispanic/Latino representation declined by 1.26pp, suggesting targeted efforts may be needed for this demographic.

3.2.4 Overall Trends

Most sectors exhibited modest changes across the two-year period, with fewer than 2% shifts in any single dimension. The mining sector stands as a clear outlier requiring further investigation to understand the drivers of its volatility. Black/African American representation increased across multiple sectors, particularly in public administration, leisure, and financial services. Hispanic/Latino representation showed mixed results, growing in some sectors while declining in others. Gender progress remains slow across the board, with most changes under one percentage point, suggesting that achieving gender balance will require sustained, long-term efforts.



Figure 4: Workforce Diversity Shifts by Demographic Category (2021-2023)

4 Which sectors are dominated by a single ethnic group and how extreme is this imbalance?

4.1 Ethnic Dominance Ranking

Rank	Sector	Dominant Group	Dominance Level	HHI*	Size (k)
1	Agriculture, forestry, fishing	White	92.3%	0.892	790
2	Mining, quarrying, oil & gas	White	87.9%	0.810	276
3	Construction	White	87.9%	0.885	11,436
4	Wholesale & retail trade	White	80.7%	0.697	1,093
5	Manufacturing	White	80.3%	0.692	773

Table 4: Sectors with Highest Ethnic Concentration (*HHI = Herfindahl-Hirschman Index: >0.5 = High concentration)

4.2 Sectors with EXTREME Ethnic Dominance

Five sectors exceed the 80% threshold for single-group ethnic dominance: Agriculture, forestry, fishing, and hunting leads with 92.3% White representation, followed by Mining, quarrying, and oil and gas extraction and Construction both at 87.9% White. Wholesale and retail trade and Manufacturing round out this group at 80.7% and 80.3% White respectively.

4.3 Concentration Severity Assessment

Using the Herfindahl-Hirschman Index (HHI) to measure concentration, Agriculture (0.892), Construction (0.885), and Mining (0.810) exhibit high concentration with HHI values exceeding 0.5. All remaining sectors fall within the moderate concentration range of 0.25-0.5, while notably, no sector achieves low concentration status below 0.25. This universal presence of at least moderate ethnic concentration underscores the systemic nature of diversity challenges across the economy.

4.4 Critical Analysis

The data reveals that all analyzed sectors maintain White-dominated workforces, with no sector achieving non-White majority representation. Traditional industries—Agriculture, Mining, and Construction—form what might be termed a “concentration triangle” of extreme ethnic homogeneity. Remarkably, workforce size provides no protection against concentration; Construction employs 11.4 million workers yet maintains 88% White dominance, demonstrating that scale and diversity operate independently. Even the most “diverse” sector, Public administration at 72.9% White, falls far short of representing general population demographics. This pattern strongly suggests systematic barriers exist in recruiting and retaining non-White workers, particularly in traditional industries where physical requirements, location constraints, and cultural factors may compound access challenges.

4.5 Policy Implications

These five sectors demand priority targeting for comprehensive diversity interventions. Effective strategies must include expanded outreach to diverse communities, apprenticeship programs specifically targeting underrepresented groups, thorough review of hiring practices and barriers to entry, robust mentorship and retention programs for minorities, and establishment of industry-specific diversity councils with real authority and accountability.

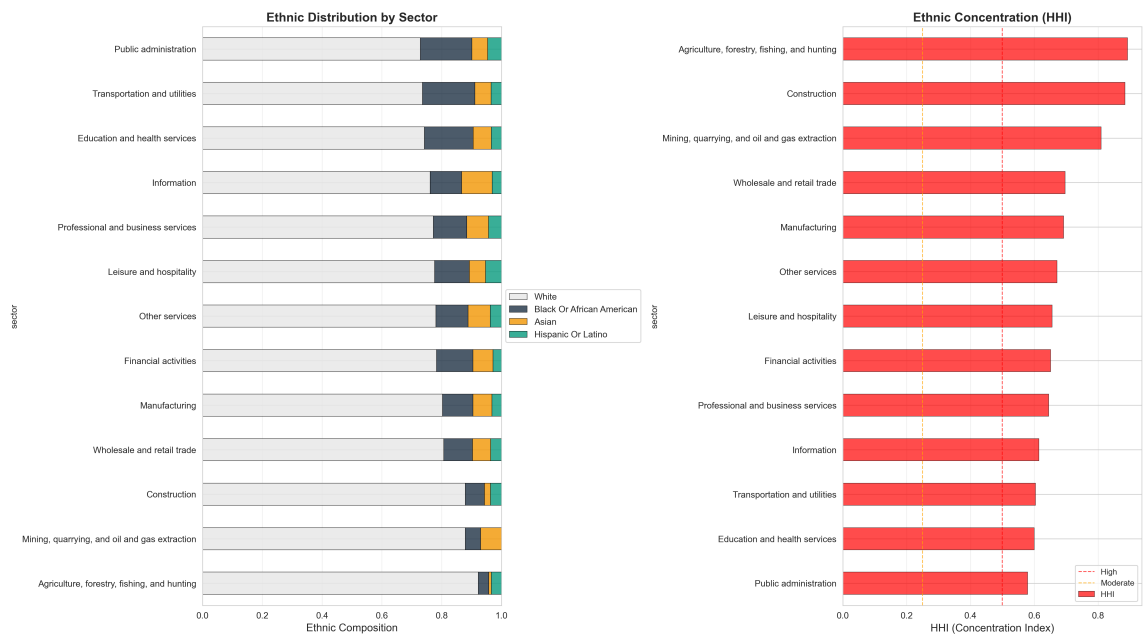


Figure 5: Ethnic Distribution and Concentration Analysis by Sector

5 Which sectors show the highest ethnic diversity and how does this relate to sector size?

5.1 Ethnic Diversity Rankings

Using the Shannon Diversity Index, which assigns higher values to more diverse distributions, we can quantify ethnic diversity across sectors. Transportation & utilities leads with a Shannon Index of 0.990, followed closely by Education & health services at 0.960 and Other services at 0.956. Public administration and Professional & business services complete the top five with indices of 0.949 and 0.940 respectively.

Rank	Sector	Shannon Index	Diversity Score	Size (k)
1	Transportation & utilities	0.990	0.396	1,534
2	Education & health services	0.960	0.400	4,880
3	Other services	0.956	0.328	1,321
4	Public administration	0.949	0.421	1,698
5	Professional & business services	0.940	0.354	2,855

Table 5: Most Ethnically Diverse Sectors

5.2 Least Diverse Sectors

At the opposite end of the spectrum, Agriculture, forestry, fishing ranks last with a Shannon Index of only 0.555, while Construction and Mining score 0.733 and 0.756 respectively. These low scores reflect the extreme ethnic concentration documented earlier in this report.

Rank	Sector	Shannon Index	Size (k)
13	Agriculture, forestry, fishing	0.555	790
12	Construction	0.733	11,436
11	Mining, quarrying, oil & gas	0.756	276

Table 6: Least Ethnically Diverse Sectors

5.3 Relationship with Sector Size

Surprisingly, the correlation between sector size and ethnic diversity proves weakly negative at -0.140, suggesting that larger sectors are slightly less diverse than smaller ones, though the relationship’s weakness indicates this is not a strong pattern. This finding challenges assumptions that organizational scale drives diversity through broader recruitment reach or resources. Instead, it suggests that scale alone does not promote diversity, and larger sectors may actually maintain more entrenched recruitment patterns that perpetuate homogeneity. Industry type emerges as a far more important determinant than workforce size.

5.4 Most Diverse Sector Profile

Transportation & Utilities demonstrates that high ethnic diversity is achievable at scale. The sector's composition includes 73.6% White, 19.4% Black/African American, 5.5% Asian, and 20.1% Hispanic/Latino workers. This sector shows particularly strong representation across all groups, with Black representation at nearly 20% standing out as exceptional among large sectors.

5.5 Analysis of Diversity Patterns

Service sectors consistently demonstrate greater ethnic diversity than goods-producing sectors. Public-facing industries such as transportation, education, and hospitality show better diversity outcomes, likely due to customer-base alignment incentives and broader talent pool access. However, even these “diverse” sectors remain far from population parity, with most still maintaining 70-75% White majorities. Construction emerges as a critical outlier—it is simultaneously the largest sector and possesses the second-lowest diversity, making it a high-impact target for intervention. The data decisively demonstrates that both small and large sectors can achieve diversity, confirming that success is fundamentally about commitment and strategy rather than organizational scale.

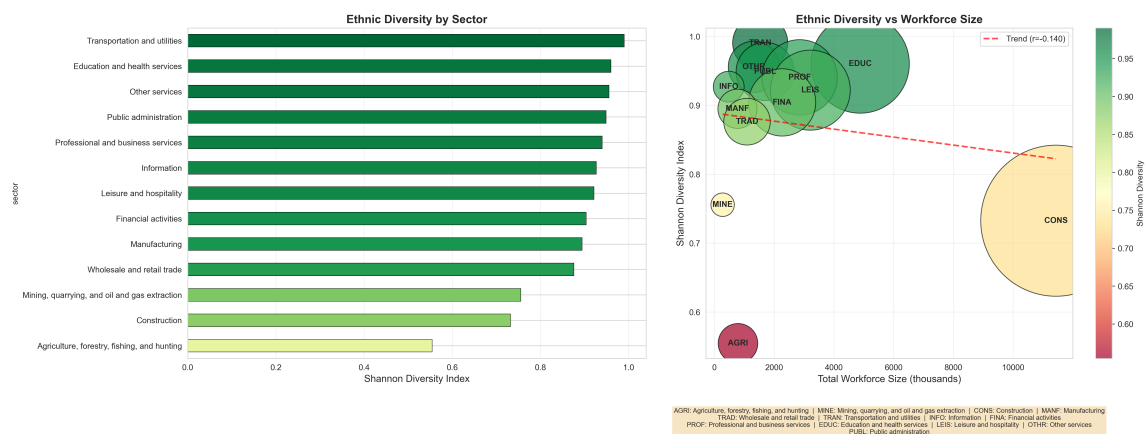


Figure 6: Ethnic Diversity Rankings and Relationship with Sector Size

6 Can we group industries based on similar ethnic compositions?

6.1 Ethnic Composition Clustering

Applying hierarchical and K-means clustering focused specifically on ethnic proportions, the analysis identified two distinct ethnic composition patterns that transcend the broader demographic clustering discussed earlier. This focused analysis reveals that ethnic composition alone creates meaningful distinctions among sectors.

6.1.1 Ethnic Cluster 1: Mainstream Diversity Pattern

Eleven sectors comprise this cluster: Education and health services, Financial activities, Information, Leisure and hospitality, Manufacturing, Mining, Other services, Professional and business services, Public administration, Transportation and utilities, and Wholesale and retail trade. The average ethnic composition shows 77.9% White, 12.0% Black/African American, 6.7% Asian, and 16.4% Hispanic/Latino representation. This cluster exhibits more balanced ethnic distribution with representation approaching national demographics, and encompasses most service and knowledge-based industries.

6.1.2 Ethnic Cluster 2: High Concentration Pattern

Only two sectors form this cluster: Agriculture, forestry, fishing and hunting, and Construction. Their average composition reveals extreme White dominance at 90.1%, with 5.0% Black/African American, merely 1.4% Asian, and 26.4% Hispanic/Latino representation. These traditional manual labor industries demonstrate very low Black and Asian representation despite somewhat higher Hispanic/Latino presence driven by agricultural labor patterns.

6.2 Hierarchical Clustering Insights

The dendrogram analysis reveals clear separation between traditional industries and others, confirming the distinctiveness of the high-concentration cluster. Public administration and Education/health services cluster closely together, both being public sector dominated. Financial and Professional services show similar patterns, reflecting their shared knowledge work characteristics. Agriculture and Construction prove ethnically similar despite vastly different work types, suggesting that factors beyond job duties drive ethnic composition.

6.3 Synthesis of Ethnic Patterns

Industry type predicts ethnic composition better than any other factor examined in this analysis. The data reveals “Two Americas” in the workforce: mainstream diverse sectors making slow progress toward representativeness versus concentrated traditional industries maintaining near-homogeneity. Construction and Agriculture form a distinct “low diversity” cluster despite different work characteristics, united by high physical demands, decentralized employment patterns, and traditional recruitment methods. Hispanic representation proves complex—high in some low-diversity sectors due to specific labor market dynamics, but not translating into overall diversity given the near-absence of other minority groups. The service economy demonstrates

substantially more ethnic integration than goods-producing sectors, though even service leaders remain far from demographic parity.

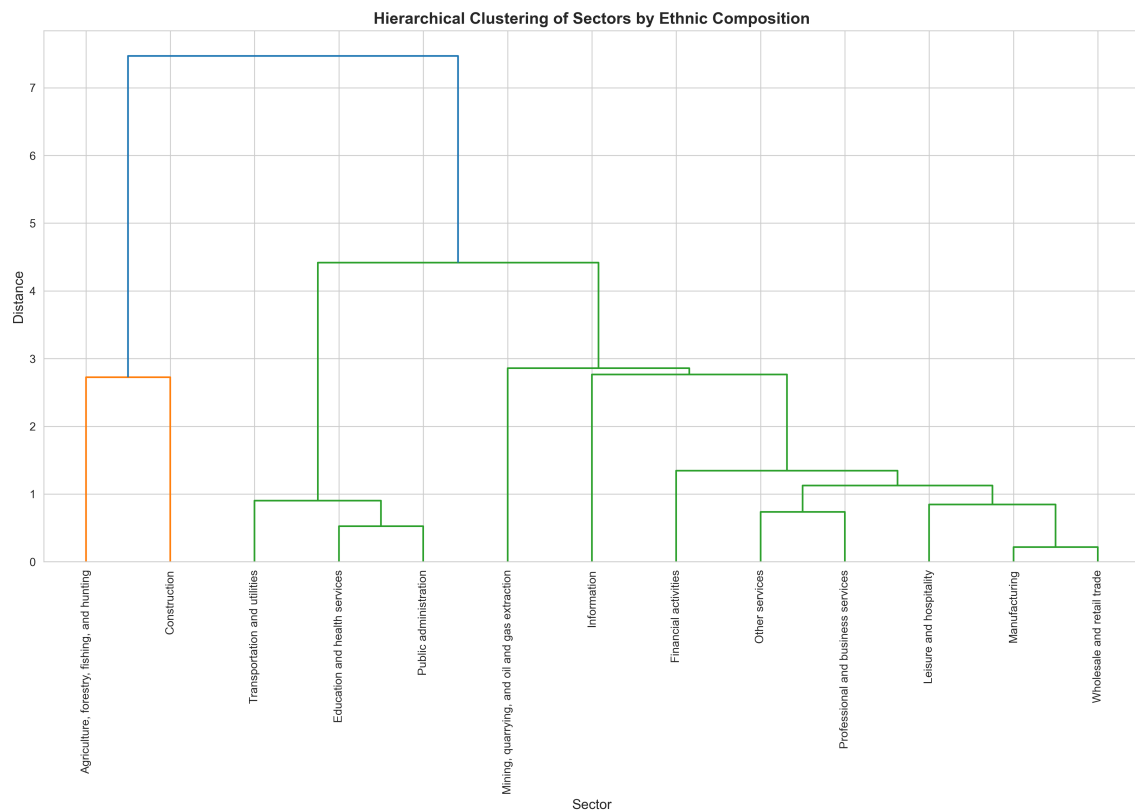


Figure 7: Hierarchical Clustering of Sectors by Ethnic Composition

7 Synthesis and Strategic Recommendations

7.1 Priority Matrix

Based on the comprehensive analysis, sectors require differentiated intervention strategies corresponding to the severity and nature of their diversity challenges.

7.1.1 Sectors with Critical Diversity Challenges

Construction (11.4M employees): This sector demands immediate, comprehensive intervention as it combines the most extreme gender imbalance (89% male) with the third-highest ethnic concentration (88% White). As the largest sector exhibiting severe imbalances on both dimensions, successful intervention here offers maximum impact potential, potentially affecting workforce demographics at a national scale.

Mining, Quarrying, Oil & Gas Extraction: This sector ranks second in both gender imbalance (85% male) and ethnic concentration (88% White) while simultaneously exhibiting the highest diversity volatility. The unstable workforce composition suggests underlying structural issues requiring investigation and stabilization strategies.

Agriculture, Forestry, Fishing: With extreme ethnic concentration (92% White), severe gender imbalance (73% male), and ranking as the least ethnically diverse overall, this sector represents the most homogeneous workforce examined. Its smaller size should not diminish urgency, as it serves as a critical case study in diversity barriers.

7.1.2 Sectors with Moderate Diversity Challenges

Transportation & Utilities: Despite leading in ethnic diversity, this sector maintains significant gender imbalance (75% male). The priority here is maintaining diversity leadership while addressing gender gaps, potentially serving as a model for achieving multidimensional diversity.

Manufacturing: With high ethnic concentration (80% White) and significant gender imbalance (71% male), this large sector requires targeted interventions. Its size offers substantial impact potential if effective strategies can be identified and implemented.

7.1.3 Sectors with Strong Diversity Trends

Education & Health Services: This female-dominated sector (74% women) demonstrates that extreme gender imbalance is not inevitable, offering important lessons about pipeline development and workplace culture. Combined with above-average ethnic diversity and status as the largest sector showing strong diversity, it provides a valuable case study for best practices.

Leisure & Hospitality: Approaching true gender parity (49% women) while maintaining strong ethnic diversity and showing positive two-year trends, this sector exemplifies achievable balance. Understanding its success factors could inform interventions elsewhere.

Public Administration: Government's strongest Black representation growth and above-average ethnic diversity demonstrate that public sector diversity initiatives can succeed. As the government can set standards through policy and contracting requirements, documenting and propagating these practices is essential.

7.2 Strategic Recommendations

7.2.1 Launch “Construction Diversity Initiative”

Given Construction’s scale (11.4M) and severe imbalances across both gender and ethnicity, a sector-specific federal/state initiative could transform workforce demographics at national scale. This initiative should mandate diversity reporting for all government contracts, establish apprenticeship quotas for women and minorities with enforcement mechanisms, and create partnerships with trade schools in diverse communities to build sustainable pipelines. Success here would represent the single highest-impact diversity intervention possible.

7.2.2 Establish Agriculture & Mining Task Force

Traditional industries face unique challenges requiring specialized approaches. A cross-sector working group should study barriers to entry including physical requirements, geographic isolation, and cultural factors. The task force must develop industry-specific recruitment strategies accounting for dispersed worksites and small firm structures, while directly addressing potential discrimination and hostile work environments that may deter diverse candidates.

7.2.3 Replicate Success Models

Transportation & Utilities achieved high ethnic diversity while maintaining large scale—understanding why and how is crucial. Systematic documentation of best practices in recruitment, retention, and culture-building should be compiled into transferable case studies. Creating a knowledge transfer mechanism to share these practices with similar sectors could accelerate progress industry-wide.

7.2.4 Address the “Pipeline Myth”

Education & Health Services proves that sectors can become female-dominated, demonstrating that gender imbalance is not inevitable or natural. Research should analyze what made E&H attractive to women, examining educational pathways, workplace flexibility, cultural messaging, and career progression opportunities. These lessons should be applied to male-dominated sectors, particularly challenging assumptions about “natural” gender distributions in technical and physical work.

7.2.5 Set Ethnic Diversity Targets

Ambitious but realistic targets grounded in local demographics should replace vague aspirations. Currently, most sectors remain 75-80% White. By 2030, concentrated sectors should target 65-70% White representation, with a 2035 goal of approaching population demographics (approximately 60% White). These targets should be disaggregated by ethnic group to ensure broad-based progress rather than token representation.

7.2.6 Implement Intersectional Analysis

Current data’s lack of gender-by-ethnicity intersections obscures compound disadvantages facing women of color. Organizations must collect intersectional data enabling analysis of Women of Color representation specifically. This enhanced data will identify which groups face the most

severe exclusion, enabling development of targeted interventions addressing unique barriers faced at these intersections.

7.2.7 Investigate Mining Sector Volatility

The 15.7% Asian representation drop in mining demands immediate investigation. If data quality issues exist, methodology must be corrected immediately. If accurate, urgent inquiry into causes—whether mass layoffs, discriminatory practices, or sector exit—is essential. Beyond investigation, strategies must be developed to stabilize diverse representation and prevent similar volatility.

7.2.8 Hispanic/Latino Strategic Focus

Hispanic representation is growing but unevenly distributed, often concentrated in manual labor roles. Strategy must understand why Hispanic representation remains high in Agriculture and Construction but low elsewhere, then systematically expand opportunities beyond manual labor sectors. Addressing language barriers, credential recognition, and networks will be essential to enabling mobility into knowledge work.

7.2.9 Extend Analysis to Leadership

Current workforce composition analysis likely masks more severe leadership representation gaps. Collection of C-suite and management diversity data by sector should become standard practice. Setting leadership diversity targets separate from workforce targets, combined with executive development programs specifically for underrepresented groups, addresses the critical “diversity at the top” challenge.

7.2.10 Create Accountability Mechanisms

Tracking without accountability enables continued stagnation. Federal contracts should be explicitly tied to diversity progress with meaningful penalties for non-compliance. Public sector diversity scorecards should enable comparative performance assessment and create reputational incentives. Industry-specific diversity certifications could create market differentiation, while annual progress reports naming specific sectors and large employers would leverage transparency for change.

7.3 Measuring Success

2025 Milestones: No sector should exceed 85% single ethnic group representation, 80% single gender representation, or fail to show measurable improvement in HHI for large sectors (exceeding 1 million workers).

2028 Milestones: The ceiling should drop to 80% for ethnic group representation and 75% for gender representation, with all sectors meeting the “moderate diversity” threshold of HHI below 0.5.

2030 Vision: Sector demographics should roughly mirror regional population composition, gender imbalance should fall below 20% in all sectors, and sustained year-over-year progress should become the norm rather than the exception. This vision is ambitious but achievable with sustained commitment and evidence-based intervention.

8 Methodology Appendix

8.1 Data Sources

The analysis drew on company workforce demographic data spanning 2020-2023, encompassing 1,272 records across thirteen sectors after excluding the aggregate “Total” category to focus on specific sector patterns.

8.2 Statistical Methods

8.2.1 Clustering Analysis

Sector grouping employed K-means clustering with StandardScaler normalization to ensure equal weighting across demographic variables. Silhouette analysis determined optimal cluster numbers by maximizing between-cluster separation while maintaining within-cluster cohesion. Hierarchical clustering using Ward linkage provided alternative perspectives and validated K-means results. Principal Component Analysis (PCA) enabled dimensionality reduction for visualization while preserving maximum variance.

8.2.2 Diversity Metrics

Gender Imbalance quantification used the absolute deviation from parity. The Herfindahl-Hirschman Index measured ethnic concentration through the summation of the squared proportions, with proportion values above 0.5 indicating high concentration. Shannon Diversity Index captured ethnic diversity using the sum of the negative product of the proportion and the natural logarithm of the proportion, with higher values indicating greater diversity.

8.2.3 Temporal Analysis

Year-over-year comparison (2021 vs. 2023) captured diversity trajectory. Total shift magnitude employed vector distance to quantify overall change magnitude. Percentage point changes provided intuitive interpretability over ratio-based metrics.

8.2.4 Correlation Analysis

Pearson correlation coefficients quantified linear relationships between sector size and diversity metrics. Scatter plots with fitted trend lines visualized these relationships while revealing potential non-linear patterns or outliers.

8.3 Limitations

The analysis faces several important limitations. Industry and subsector data proved sparse, necessitating sector-level focus that obscures within-sector heterogeneity. Gender categories include only male/female, missing non-binary and other identities. Ethnic categories’ limitation to four groups misses important diversity nuance within broad categories. Intersectional analysis remains impossible without gender-by-ethnicity cross-tabulations, obscuring compound disadvantages. Leadership representation data’s absence means composition analysis captures

only overall workforce, likely underestimating diversity gaps at senior levels. The four-year time period prevents assessment of longer-term trends or cyclical patterns. Finally, company-specific data may not generalize to other organizations or the broader economy, though sectoral patterns likely reflect industry-wide dynamics.

Conclusion

This analysis reveals a workforce fundamentally divided along both gender and ethnic dimensions, with traditional industries—Construction, Agriculture, and Mining—forming a distinct cluster of extreme demographic concentration. While select service sectors demonstrate encouraging diversity progress, no sector examined approaches true demographic parity with the general population. The glacial pace of change observed between 2021 and 2023 suggests that voluntary diversity efforts prove insufficient to drive meaningful transformation.

The Construction sector emerges unambiguously as the single highest-priority intervention target. Its unique combination of massive scale (11.4 million workers), extreme gender imbalance (89% male), and high ethnic concentration (88% White) means that progress here could dramatically improve overall U.S. workforce demographics. Success in Construction would affect more workers and send stronger market signals than interventions in any other sector.

However, Construction's challenges exemplify broader systemic issues. Policy interventions, robust accountability mechanisms, and sector-specific initiatives will be needed to accelerate progress across all industries. By systematically studying and replicating success stories—Leisure & Hospitality's approach to gender parity and Transportation & Utilities' ethnic diversity strategies—while applying evidence-based interventions to problem sectors, meaningful progress becomes achievable within a 5-10 year timeframe.

The path forward demands sustained commitment, adequate resources, and willingness to implement strong accountability measures. The choice facing policymakers and industry leaders is clear: continue incremental voluntary efforts yielding minimal results, or embrace comprehensive intervention strategies capable of transforming workforce demographics at scale.