



Diversity Analytics: Understanding and Measuring D&I in Organizations

Effective diversity and inclusion (D&I) practices should be integrated throughout an employee's entire journey, from recruitment to exit. Poor D&I management can manifest in workplace bias, unfair treatment, increased turnover of valuable employees, and a lack of innovation and customer empathy in corporate strategies.

Why Focus on Diversity and Inclusion?

Social Case

Employers have a moral obligation to treat employees with fairness and dignity, ensuring decisions are made without prejudice and stereotypes.

Business Case

Benefits include maximizing people potential, better understanding customers, drawing from a wider candidate pool, developing a more positive company image, increasing employee engagement, and improving retention, innovation and team performance.

While the business case for diversity can be difficult to quantify and may take time to realize benefits, analytics can help break down the data and demystify the link between a diverse organization and its benefits.





Moving Beyond Descriptive Reports

Current HR functions often stop at simply describing workforce characteristics, such as stating "12% of our senior managers are Black or Asian ethnicity and 20% of our senior managers are women."

Without context, these numbers lack meaning. Is 12% what we should expect? What is the likelihood of getting 20% women completely by chance? How do we know if this is due to bias or discrimination?

Descriptive Reports

Current state snapshots that often become checkbox exercises

Diversity Benchmarking

Comparing statistics with other organizations, but with caution about different job structures

Predictive Models

Using statistical tools to bridge the gap and make sound decisions



Gender and Job Grade Analysis: Using Chi-Square

Let's examine a management consulting firm, SlidesRUs, with 1,493 employees (746 female, 745 male, 2 undeclared). If there was no glass ceiling or discrimination, we would expect equal proportions of men and women in each job grade.

The chi-square test helps determine if this pattern is statistically meaningful or could be due to chance. It checks whether the observed gender distribution differs from what we would expect if there was no link between gender and seniority.

The larger the difference between observed and expected patterns, the more our chi-square test will indicate something significant is occurring that cannot be ignored.

Interpreting Chi-Square Results



In our example, the chi-square statistic of 164.7 with a p-value effectively zero ($2.2\text{e-}16$) tells us there is a statistically significant link between gender and grade.

This means we can be certain that the pattern of females being underrepresented in senior grades and over-represented in junior grades is meaningful with a very high level of certainty.

There is almost zero chance that this gender proportion could happen by coincidence, suggesting some other factor has influenced it.

✖ Statistical Interpretation Caveat

It's one thing to say women are underrepresented in senior roles and another to say there is discrimination. The statistical result needs to be carefully combined with rational arguments about why the pattern may be occurring.



Exploring Ethnic Diversity Across Teams

1

Why Analyze Team Diversity?

To investigate whether specific ethnic groups are clustered together and identify patterns that could explain variations in underrepresented groups (UG) across teams.

2

Team-Level Analysis Benefits

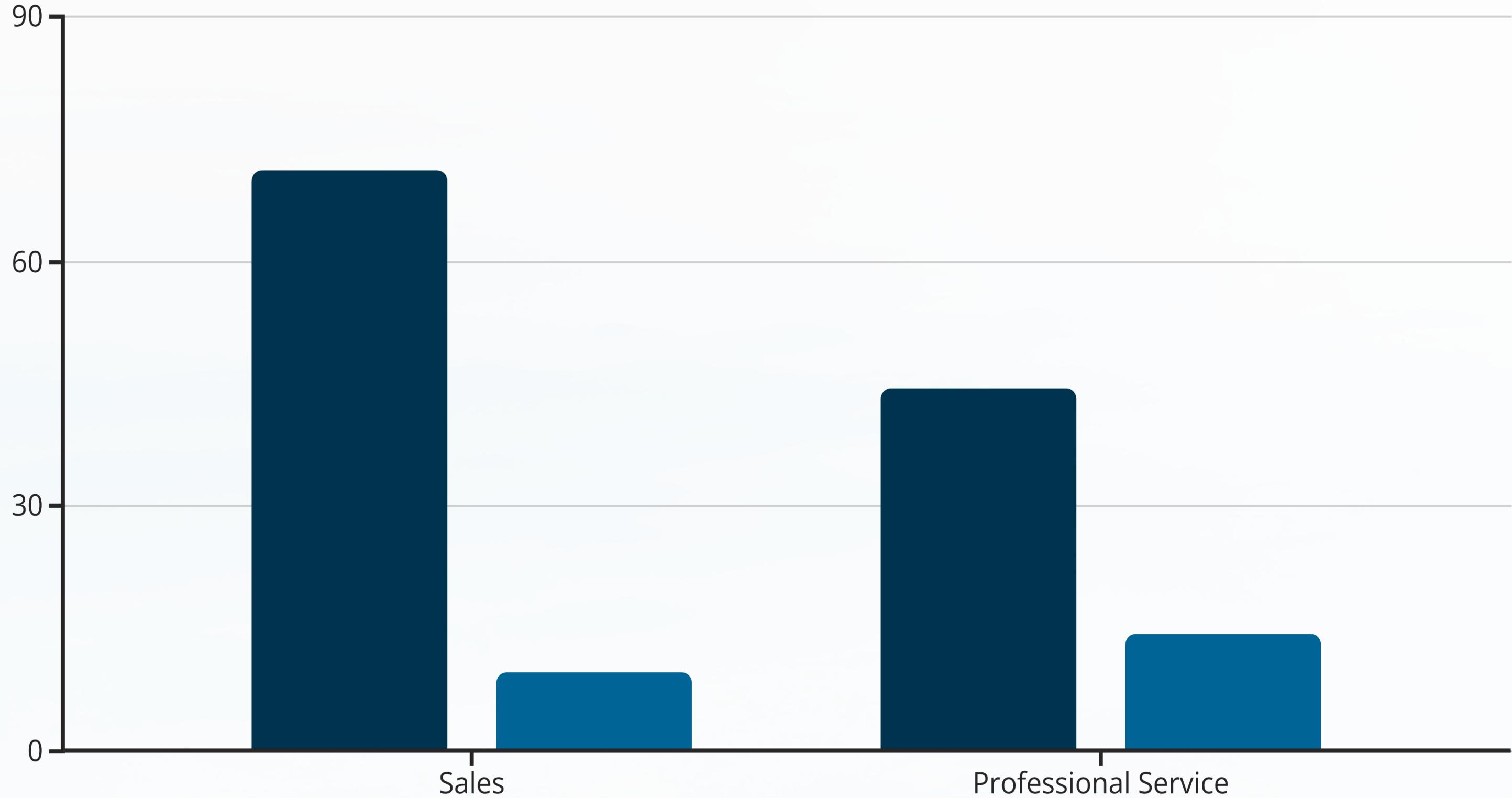
Allows linking team-level UG statistics with team-level scores from employee engagement surveys, providing richer insights than individual-level data alone.

3

Predictive Modeling Approach

Identifying patterns that show where people from underrepresented groups are found or not found can help predict levels of UG representation and inform targeted interventions.

Comparing Functions Using Independent Samples T-Test



Multiple Linear Regression: Predicting Ethnic Diversity

What is Multiple Regression?

Multiple regression helps determine what factors predict a low prevalence of UG individuals in teams by examining the relationship between UG percentage (dependent variable) and multiple predictor variables.

It shows how much each predictor variable impacts the dependent variable on its own and which has the greatest impact, helping identify problem areas.

Variables in Our Model

- Dependent: UG percentage
- Independent: Location (London or not), Function (sales or professional service), Group size, Number of female team leads, Percentage male





Interpreting Regression Results

Model Significance

The model accounts for 17.4% of variance in UG prevalence across teams with a significance level of $p < 0.001$, meaning there's less than 1 in 1,000 chance this result occurred by chance.

Significant Predictors

Two variables had significant impact: Location (Beta = -0.361) and Function (Beta = 0.166). Teams in London have higher UG prevalence, and professional service teams have higher UG prevalence than sales teams.

Geographic Context

Even accounting for geographical differences (London has ~30% UG residents vs. only 2.4% in areas like Northumberland), there's still significantly lower UG prevalence in sales compared to professional service.

Beyond Basic Diversity Metrics

Comprehensive D&I Measurement

While we've focused on ethnicity and gender, comprehensive D&I programs should measure many dimensions:

- Primary characteristics: gender, ethnicity, disabilities, age, marital status, sexual orientation
- Secondary characteristics: parental status, social status, nationality, language, accent, communication style, appearance, education, religion

Measuring Inclusion

Many organizations measure how inclusive their culture is through survey questions, creating an "inclusion index" similar to engagement indices.

The ability to quantify the benefits of having a diverse and inclusive organization remains the key challenge in making the business case for diversity.

