

INTRODUCTION

The Market Intelligence dataset builds on Growth Intelligence by adding workforce structure signals. Each row still represents one company in one year. The additional fields show which function dominates the workforce, how diversified the role mix is, and how the company's organizational shape compares to others in the same market.

HOW TO READ THIS DATASET

This layer uses the same 300 company cohort and the same company year structure as Growth Intelligence. All Growth fields are included. The new fields add a structural lens: you can now see not just how fast a company is growing, but how its workforce is composed.

FIELD REFERENCE

Company and Context

Same fields as Growth Intelligence: company_id, company_name, company_domain, industry, country, company_size_midpoint, year. See the Growth Intelligence Data Dictionary for full descriptions.

Growth Signals

Same fields as Growth Intelligence: headcount_observed, net_headcount_change, growth_rate_yoy, avg_tenure_years, growth_bucket. See the Growth Intelligence Data Dictionary for full descriptions.

Workforce Structure

primary_role_bucket	The dominant functional role category in the company's workforce (e.g. sales, engineering, operations, specialist). Tells you which function has the largest share of headcount.
primary_role_pct	The percentage of total observed headcount that belongs to the primary role bucket. A high value (e.g. 40%+) indicates functional concentration. A lower value suggests a more evenly distributed workforce.
distinct_role_buckets	The number of unique role categories present in the company's workforce. Higher values indicate broader organizational complexity.
role_diversity_bucket	A pre computed classification of how diversified the company's role mix is. Three tiers: 1. Low diversity, 2. Moderate diversity, 3. High diversity. Useful for separating focused operators from broad organizations.

DATA NOTES

- All data notes from the Growth Intelligence layer apply here. The 2020 sample year, tenure coverage limitations, and sample scope remain unchanged.
- Role bucket classifications are derived from standardized role taxonomies applied across the full dataset. They are designed to be comparable across industries and company sizes.
- Primary role percentages reflect the visible workforce. Companies with a higher share of unlabeled roles may show lower concentration in the primary bucket than their actual organizational structure would suggest.