Data for Impact Primer

LinkedIn Data Available through Data for Impact

Fall 2023

LinkedIn data can foster more resilient, sustainable, and equitable labor markets

As part of our vision to create economic opportunity, Data for Impact leverages data to:

- Amplify the impact of public investments in industrial and workforce development
- Inform public policy decisions on economic development strategy
- Enable cutting-edge research on efforts to improve labor market function

WORLD BANK GROUP

OECD

UNDP

European Bank for Reconstruction and Development

DESTATIS Statistisches Bundesamt

HAI Stanford University Human-Centered Artificial Intelligence

IYF

IAB INSTITUTE FOR EMPLOYMENT RESEARCH The Research Institute of the Federal Employment Agency

CAF DEVELOPMENT BANK OF LATIN AMERICA

CENIA CENTRO NACIONAL DE INTELIGENCIA ARTIFICIAL

INTERNATIONAL MONETARY FUND

IDB

Federal Ministry of Labour and Social Affairs

Federal Reserve Bank of San Francisco

DFI efforts have advanced LinkedIn's thought leadership, relevance among global institutions, and integration with government services

Innovate

- DFI facilitates EG's efforts to help IAB assess labor market trends
- By contributing standardized datasets, DFI frees up EG to innovate new metrics alongside IAB

Inform

- DFI data informed a $1.7BN World Bank strategy for Argentina
- DFI's data on gender inequities and industry hiring helps WB economists craft better labor market interventions

Integrate

- The German Statistical Authority relies on monthly data from DFI
- DFI positions LinkedIn as a critical partner by integrating our data into government services

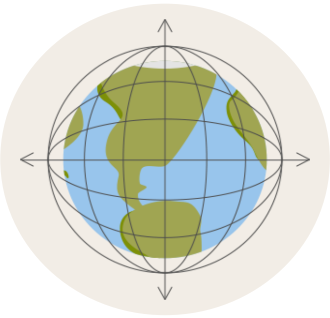
Relevant Datasets

As of Fall 2023, LinkedIn offers 6 indicators on hiring, career pathways, and skills, which can be cut by 4 dimensions

Indicators

Indicator	Answers the question:
LinkedIn Hiring Rate	Which industries are experiencing the most job transition growth right now in my market?
Labor Market Tightness	How do active job openings compare to active applicants in any given industry?
Female Representation	What does female representation look like, by industry and seniority level?
Career Transitions	What is the ratio of LinkedIn members moving from one industry to another?
Skills Genome	What skills are most representative for a given occupation?
Skills Penetration	What share of the representative skills for a labor market segment are tech, green, or soft skills?

Dimensions



National[1]

Industry[2]

Skill Groups[3]

Gender

1 – Subnational cuts available for Germany, India, and US for some indicators.
2 – LinkedIn's industry taxonomy includes 20 industries but they do not align to ISIC.
3 – Skills groups described in section 3.

LinkedIn data—like most privately-sourced data—has limitations, so it should always be used to supplement, not replace, publicly-sourced data

Privacy

- Member privacy is our first priority
  - Geography based on member reported data, not physical location / IP address
  - Consent mechanisms for outcome tracking not enabled
  - Minimum thresholds prevent the sharing of data for granular labor market segments

Representativeness

- LinkedIn members more likely to:
  - have regular internet access
  - live in higher-income places
  - work in tech-enabled industries and occupations
- Data still largely informative:
  - High coverage in US and Europe, with highest platform growth among first-line professionals
  - World Bank has validated data for use for 80+ countries

Supply vs. Demand

- Our data is largely supply-based:
  - pulls from our 950+M members
  - Data on job postings (demand) not as strong or easily standardized
- Demand metrics coming online:
  - Labor Market Tightness metric has been used by partners seeking to test validity

Sections

LinkedIn Hiring Rate (YOY)

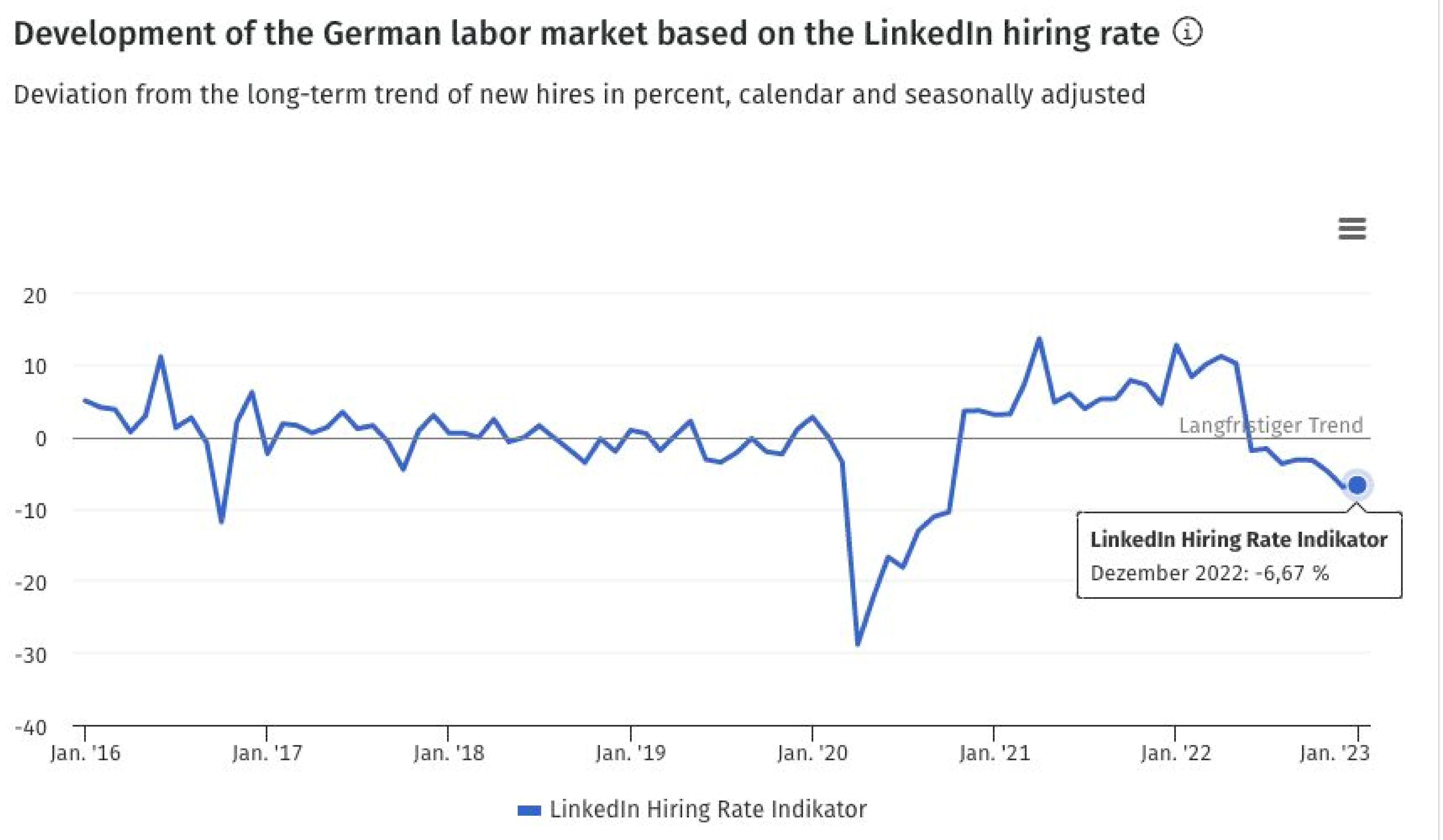
Monthly Refreshes Available

Indicator Description

The year over year change in the number of LinkedIn members who added a new employer to their profile in the same month or year the new job began, divided by the total number of LinkedIn members in that entity (e.g. country, industry, country-industry pair, green workers, etc.), indexed to the average month in 2016.

Sample Data Description

This sample shows the YOY LinkedIn Hiring Rate for Germany, as posted by the German Statistical Authority, Destatis.



Labor Market Tightness

Monthly Refreshes Available

Indicator Description

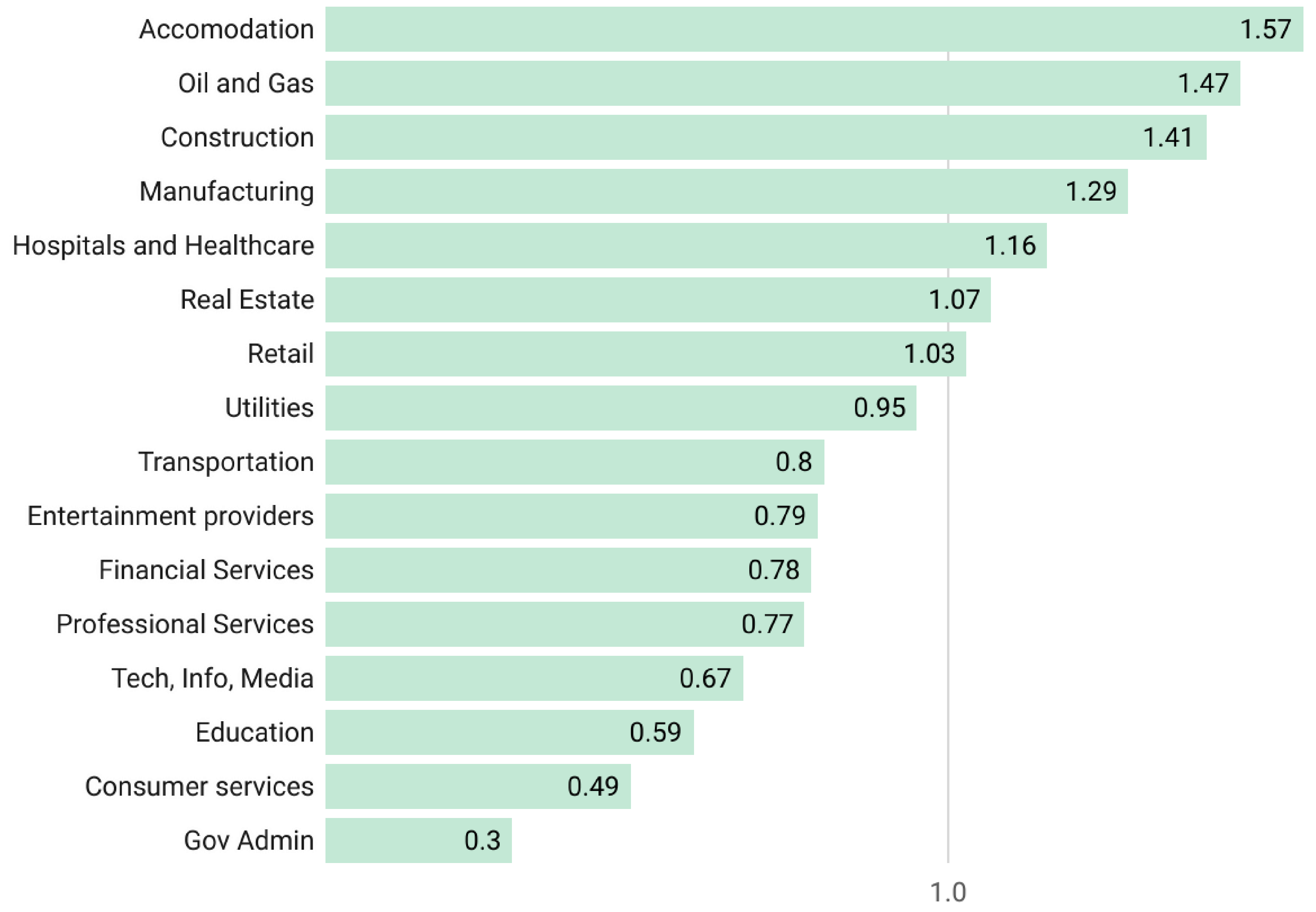
The stock of “active” job openings divided by the number of active applicants.

Sample Data Description

This sample shows the US Labor Market Tightness metric by industry for March 2023.

Labor Market Conditions Vary Across Industries

March 2023 Labor market tightness relative to pre-pandemic average



The category equals 1.0 when it returns to its pre-pandemic level. Pre-pandemic level is the average monthly labor market tightness ratio from Dec 2019 to Feb 2020.

Source: LinkedIn Economic Graph • Created with Datawrapper

Female Representation

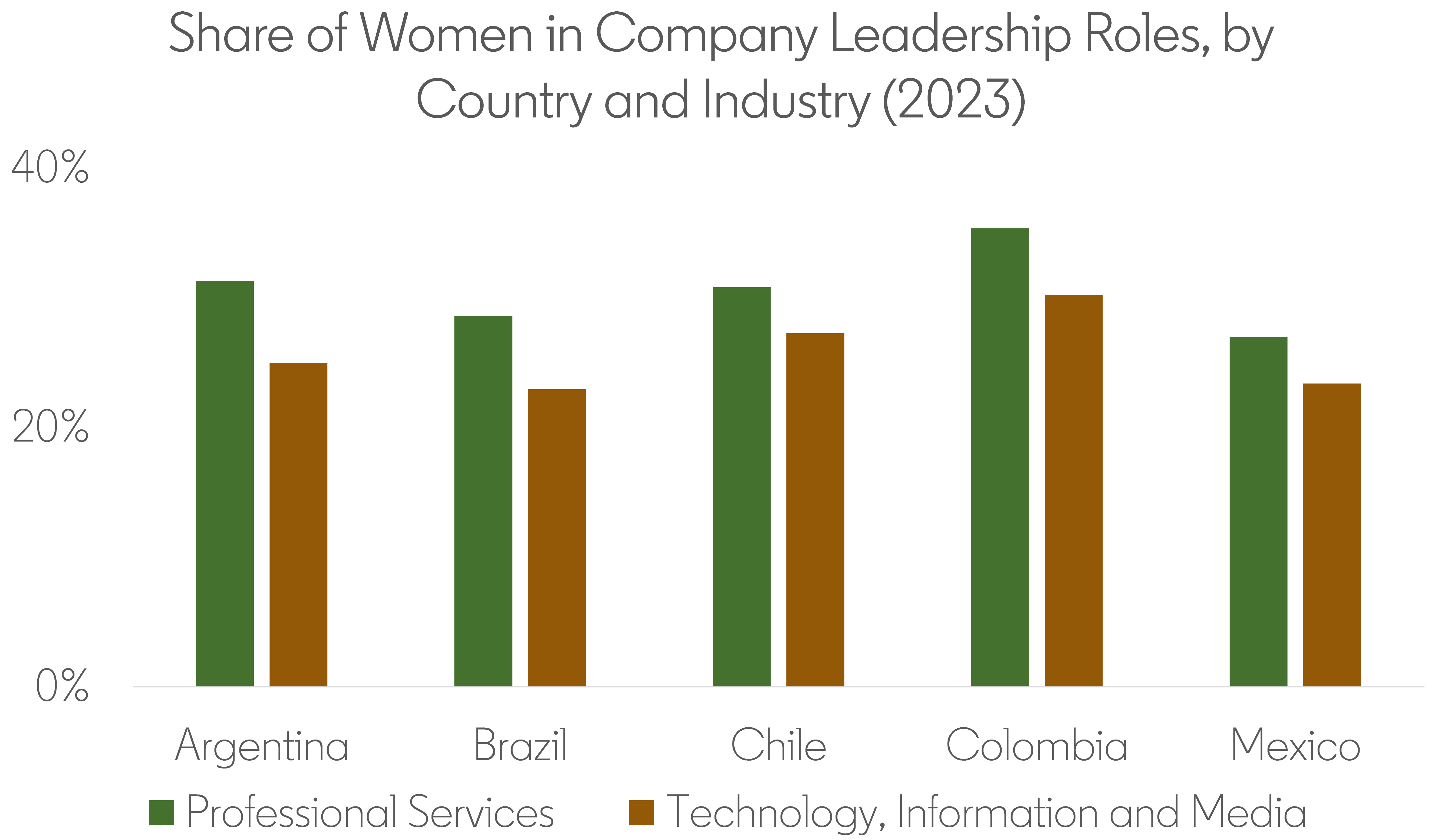
Annual Refreshes Available

Indicator Description

The share of a country's labor force comprised of women, by seniority, or the share of an industry's leadership roles occupied by women, by country.

Sample Data Description

This sample shows the share of women in leadership roles by country and industry, as received by the Inter-American Development Bank for their 2023 Flagship Report on Gender.



Note: Gender identity isn't binary and we recognize that some LinkedIn members identify beyond the traditional gender constructs of "men" and "women." If not explicitly self-identified, we have inferred the gender of members included in this analysis either by the pronouns used on their LinkedIn profiles, or inferred on the basis of first name. Members whose gender could not be inferred as either man or women were excluded from this analysis.

Career Transitions

Monthly Refreshes Available

Indicator Description

Career Transitions shows share of job to job moves between entities (e.g. industries). This can include transitions to jobs within the same companies (internal) or to another company (external).

Sample Data Description

This sample shows German career transitions for the month of January 2021.

- The “ratio” column shows [transitions from A to B / transitions from B to A].
- The “from” column shows all transitions *from* industry A by share of destination industry B.
- The “to” column shows all transitions *to* industry A by share of source industry B.

Industry A	Industry B	Ratio	From	To
Manufacturing	Accommodation and Food Services	0.65	0.8	1.2
Manufacturing	Administrative and Support Services	0.81	1.6	2.1
Manufacturing	Construction	1.54	2.1	1.4
Manufacturing	Consumer Services	1.59	0.9	0.6
Manufacturing	Education	0.63	1.6	2.6
Manufacturing	Entertainment Providers	0.92	0.7	0.8
Manufacturing	Financial Services	1.26	2.6	2.2
Manufacturing	Government Administration	1.69	1.7	1
Manufacturing	Hospitals and Health Care	0.96	1.5	1.7
Manufacturing	Manufacturing	1	51.9	53.5
Manufacturing	Oil, Gas, and Mining	0.82	0.9	1.2
Manufacturing	Professional Services	1.02	16.2	16.3
Manufacturing	Real Estate and Equipment Rental Services	2	0.8	0.4
Manufacturing	Retail	1.12	3.2	2.9
Manufacturing	Technology, Information and Media	1.23	6	5
Manufacturing	Transportation, Logistics, Supply Chain and Storage	0.73	1.8	2.6
Manufacturing	Utilities	1.79	2.2	1.2
Manufacturing	Wholesale	1.1	3.5	3.3

Skill Genome

Yearly Refreshes Available

Indicator Description

For any entity (occupation or job, country, industry, etc), the skill genome is an ordered list (a vector) of the 50 'most characteristic skills' of that entity. These most characteristic skills are identified using a TF-IDF algorithm to identify the most representative skills of the target entity, while down-ranking ubiquitous skills that add little information about that specific entity (e.g. Microsoft Word).

Sample Data Description

This sample forms the basis of a larger dataset which aims to understand the unique skills of the Software & IT Services industry group in various countries.

10 Most Characteristic Skills of the Software and IT Services Industry, by Country *(2020)*

Rank	France	India	UAE	USA
1	Agile Methodologies	Core Java	Requirements Analysis	Software Development Life Cycle (SDLC)
2	Git	Software Development Life Cycle (SDLC)	Pre-sales	Software as a Service (SaaS)
3	Cloud Computing	Requirements Analysis	Software Development Life Cycle (SDLC)	Enterprise Software
4	SQL	SQL	Vendor Management	Agile Methodologies
5	Integration	C (Programming Language)	Cloud Computing	Salesforce.com
6	Pre-sales	Java	SQL	Cloud Computing
7	Scrum	Manual Testing	Business Analysis	Requirements Analysis
8	Linux	Agile Methodologies	Integration	Integration
9	Software as a Service (SaaS)	JavaScript	Enterprise Software	Amazon Web Services (AWS)
10	Java	Unix	Solution Architecture	Vendor Management

Skill Penetration

Yearly Refreshes Available

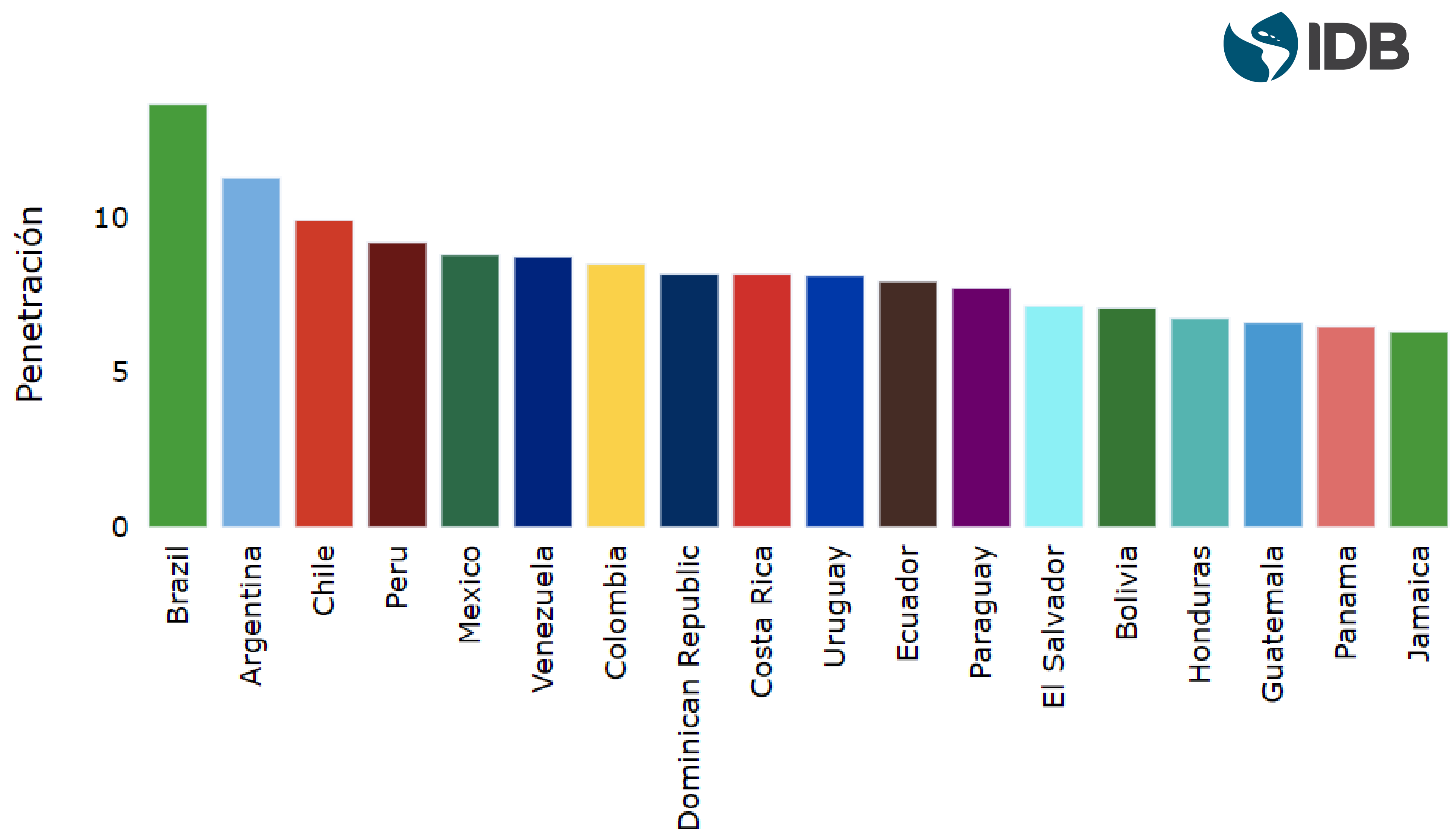
Indicator Description

The Skill Penetration metric calculates the share of an occupation's, industry's, or country's top 50 skills that come from a singular skills group. For example, if 5 of 50 skills for Data Scientists in the Information Services industry fall into the Artificial Intelligence skill group, Artificial Intelligence has a 10% penetration for Data Scientists in Information Services. See the appendix for more on Skill Penetration methodology.

Sample Data Description

This sample compares the technical skills of Latin American country members as part of a Labor Observatory hosted by the Inter-American Development Bank here.

Technical Skills Penetration, 2021



Sections

Geographic Coverage

Data availability varies by country

81 Available Countries

Currently, LinkedIn can provide at least some data for 81 countries. Cutting datasets by more (and more granular) dimensions will reduce data availability. We are NOT able to provide data by region.

Geographic Coverage of LinkedIn's DDP Data *(2022)*

Representativeness

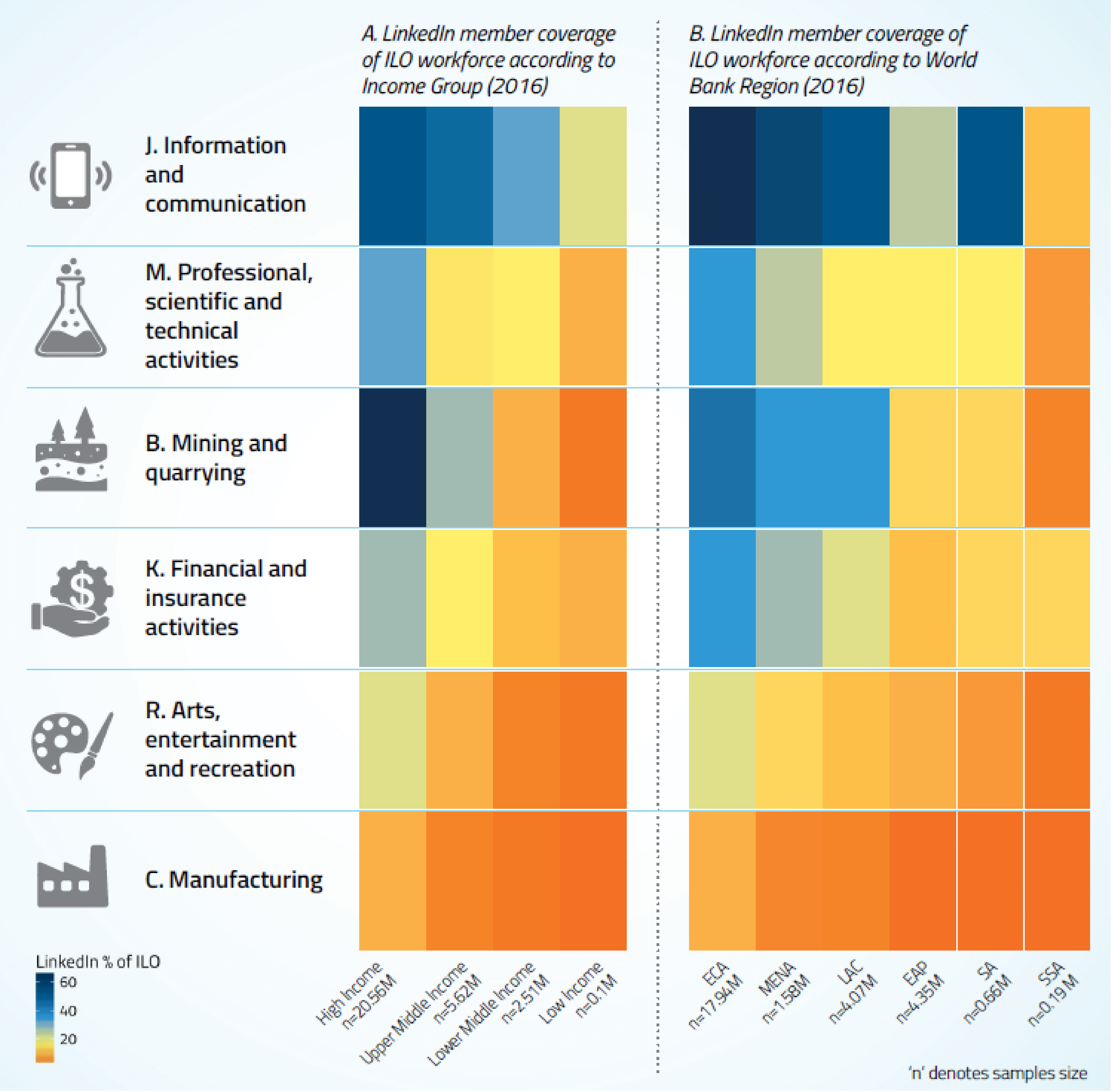
Representativeness varies by industry and region

A Complementary Dataset

LinkedIn data represents some of the most comprehensive, granular sources of labor market and skills data available. It is, however, influenced by how members choose to use the platform, which can vary based on professional, social, and regional culture, as well as overall site availability and accessibility. As such, aggregated LinkedIn datasets shared through the DDP will likely be most useful when paired with traditional sources of labor market and economic data.

**LinkedIn's global membership has grown by 50+% since this analysis, so we anticipate that data is even more representative in 2022 than in 2016; an updated analysis is forthcoming.*

Representativeness of LinkedIn's DDP Data *(2016*)*



Based on calculation using LinkedIn and International Labor Organization (ILO) data in 92 countries

Source: Zhu,Tingting Juni; Fritzler,Alan; Orlowski,Jan Alexander Kazimierz. 2018. *World Bank Group-LinkedIn Data Insights : Jobs, Skills and Migration Trends Methodology and Validation Results* Vol. 1 of *World Bank Group-LinkedIn Data Insights : Jobs, Skills and Migration Trends Methodology and Validation Results*. Washington, D.C. : World Bank Group. https://hubs.worldbank.org/docs/ImageBank/Pages/DocProfile.aspx?nodeid=30610436

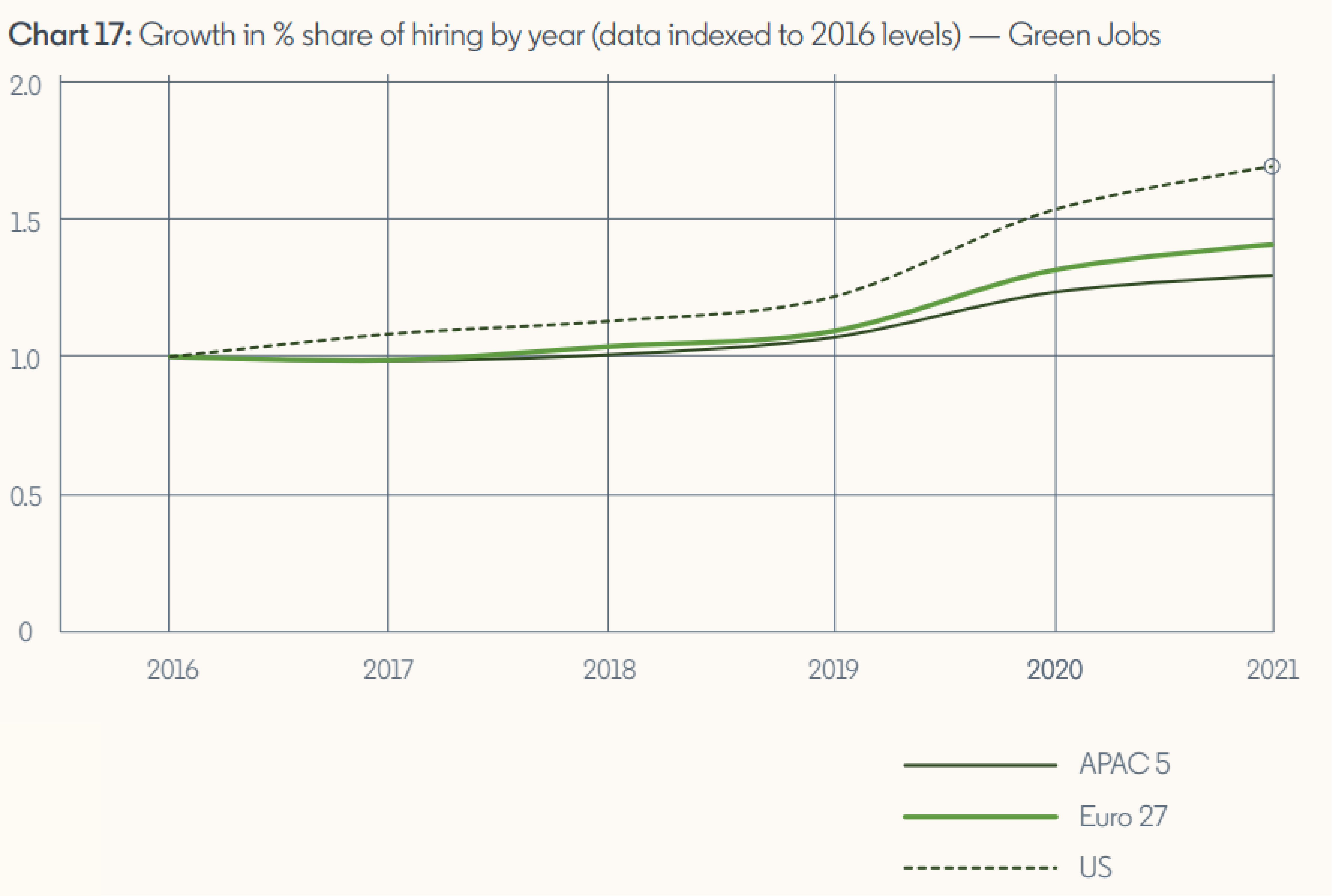
Sections

Green Data

See 2023 Green Global Green Skills Report

Green Dimension	Definition
Green skills	Skills that enable the environmental sustainability of economic activities
Green jobs	Jobs that cannot be performed without extensive knowledge of green skills
Greening jobs	Jobs that can be performed without green skills, but typically require some green skills
Greening potential jobs	Jobs that be performed without green skills, but occasionally require some level of green skills
Non-green jobs	Jobs that do not require green skills to be performed
Green talent	A LinkedIn member who has explicitly added green skills to their profile and/or are working in a green or greening job

Sample Green Data Analysis



Specialized Skill Categories

Skills are annotated as pertaining to specific skill categories, which can be applied as indicator dimensions

Skill Group Category	Description
Soft Skills	Non-cognitive skills or personality traits valued in the labor market but not assessed by achievement tests. IQ or achievement tests cannot predict these skills.
Business Skills	Knowledge and skills required to start or operate an enterprise. Examples include Business Management, Project Management, Entrepreneurship.
Tech Skills	Defined as a range of abilities to use digital devices, communication applications, and networks to access and manage information. They enable people to create and share digital content, communicate and collaborate, and solve problems.
Disruptive Tech Skills	Skills associated with developing new technologies that are expected to impact labor markets in the coming years. Examples include Robotics, Genetic Engineering, and Artificial Intelligence. Artificial Intelligence can be isolated as a skill group category itself (see OECD.AI).

Sections

LinkedIn Hiring Rate

Methodological Notes

Indicator	Equation	In a sentence
LinkedIn Hiring Rate (LHR)	$LHR_{2018} = \dfrac{\left(\dfrac{new\ hires_{2018}}{total\ employees_{2018}}\right)}{\dfrac{new\ hires_{2016}}{total\ employees_{2016}}}$	The share of Manufacturing workers changing jobs was much X% greater in 2018 than in 2016.
LHR YOY	$LHR\ YOY_{2018} = \dfrac{LHR_{2018} - LHR_{2017}}{LHR_{2017}} = \dfrac{\left(\left(\dfrac{new\ hires_{2018}}{total\ employees_{2018}}\right) - \left(\dfrac{new\ hires_{2017}}{total\ employees_{2017}}\right)\right)}{\left(\dfrac{new\ hires_{2017}}{total\ employees_{2017}}\right)}$	The share of Manufacturing workers changing jobs increased by X% from 2017 to 2018.

Skills Genome

Methodological Notes

Goal

Identify the most relevant skills for groups of members holding specific occupations in specific countries and specific geographies.

Challenge

A set of generic skills (such as Microsoft Office) often occupies the top spots in the skills vector for many occupations.

Solution

TF-IDF downweighs skills that are common across other occupations. The method is a combination of skill popularity and uniqueness.

Data Analyst Skills in Spain:

Popular *Vs.* Popular AND Unique (Skills Genome)

Popular	Popular AND Unique
Data Analysis	SQL
SQL	Tableau
Python	Python
Tableau	Data Analysis
Microsoft Excel	Hive
Microsoft Office	R
R	Machine Learning
Microsoft PowerPoint	Data Mining
Microsoft Word	Post-GRE SQL
Hive	Data Visualization

Interpretation

The most distinct skills claimed by LinkedIn members working as Data Analysts in Spain include SQL, Machine Learning, and Data Visualization.

Skills Penetration

Methodological Notes

Goal

Determine how many of the most distinct 50 skills, for each occupation in each country, belong to a given skill group (e.g. Tech Skills, Soft Skills, etc.).

Challenge

LinkedIn's occupational distribution and coverage varies by country, even within the same industry.

Solution

Compare each country's skill penetration to a global composite benchmark based on that country's occupational distribution.

Tech Skills in Healthcare Industry

Ctry	Industry	AVG SP	GLOBAL SP	RELATIVE SP
X	Healthcare	0.1	0.2875	0.3478
Y	Healthcare	0.3	0.3625	0.8276
Z	Healthcare	0.5	0.333	1.5015

Interpretation

On average, tech skills comprise 10% of the occupational skill genomes in Country X's Healthcare Industry. Holding constant for differences in occupational distributions, this penetration of tech skills is ~35% of the global average for this industry.

Skill Penetration (SP) Calculations

Methodological Notes

1. Start with Skills Penetration data at Country/Industry/Occupation level

*Tech Skills Penetration by Ctry/Ind/Occ**

Country	Industry	Occupation	Skill Penetration (SP)
X	Healthcare	Nurse	0.05
X	Healthcare	Accountant	0.15
Y	Healthcare	Nurse	0.25
Y	Healthcare	Surgeon	0.35
Z	Healthcare	Nurse	0.60
Z	Healthcare	Surgeon	0.50
Z	Healthcare	Accountant	0.40

2A. AVERAGE skill penetration based on occupations at country-industry level

AVG Tech Skills Penetration by Ctry/Ind

Country	Industry	Calculation	AVG SP
X	Healthcare	(0.05+0.15)/2	0.1
Y	Healthcare	(0.25+0.35)/2	0.3
Z	Healthcare	(0.60+0.50+0.40)/3	0.5

2B. Calculate GLOBAL industry skill penetration from occupational mix

Tech Skills Penetration by Ind/Occ (Global)

Country	Occupation	Calculation	Skill Penetration
-	Nurse	(0.05+0.25+0.60)/3	0.3
-	Accountant	(0.15+0.40)/2	0.275
-	Surgeon	(0.35+0.50)/2	0.425

GLOBAL Tech Skills Penetration by Ctry/Ind

Country	Industry	Calculation	GLOBAL SP
X	Healthcare	(0.3+0.275)/2	0.2875
Y	Healthcare	(0.3+0.425)/2	0.3625
Z	Healthcare	(0.3+0.275+0.425)/3	0.333

3A. Calculate RELATIVE industry skill penetration based on global SP

RELATIVE Tech Skills Penetration by Ctry/Ind

Country	Industry	Calculation	RELATIVE SP
X	Healthcare	0.1 / 0.2875	0.3478
Y	Healthcare	0.3 / 0.3625	0.8276
Z	Healthcare	0.5 / 0.333	1.5015

3B. Final Outputs are AVG, GLOBAL, and RELATIVE Skills Penetration by Ctry/Ind

All Tech Skills Penetration by Ctry/Ind

Ctry	Ind	AVG SP	GLOBAL SP	RELATIVE SP
X	Healthcare	0.1	0.2875	0.3478
Y	Healthcare	0.3	0.3625	0.8276
Z	Healthcare	0.5	0.333	1.5015

**Ctry is Country. Ind is Industry. Occ is Occupation.*