HR EMPLOYEE DISTRIBUTION ANALYSIS REPORT

PREPARED BY: Benjamin Onyatta

DATE: 4/17/2025

Abstract

This analysis project focused on exploring employee data to uncover key workforce insights and support data-driven HR decision-making. The dataset comprised detailed employee records, including demographic attributes (gender, race, age), employment details (job title, department, location), and lifecycle events (hire date, termination date). During data preprocessing in MySQL, an additional variable, "age," was derived from birthdate to facilitate age-based analysis.

The primary goals of the analysis were to examine employee distribution across departments and locations, assess diversity metrics, and identify workforce trends such as hiring patterns, terminations, and tenure. Key findings revealed uneven departmental representation, potential gaps in gender and racial diversity, and regional staffing trends across different states and cities. Age and tenure distributions also provided insights into generational workforce dynamics and potential areas for succession planning or retention strategies.

Overall, the project demonstrated the value of HR analytics in understanding workforce composition, improving employee lifecycle management, and guiding strategic HR interventions. The results were visualized using interactive dashboards to enable stakeholders to monitor key metrics in real-time.

Introduction

Context

In today’s competitive business environment, organizations are increasingly turning to datadriven approaches to manage their human resources more effectively. Human Resource (HR) analytics, also known as people analytics, enables organizations to gain actionable insights into workforce trends, employee behavior, and operational efficiency. This analysis project was undertaken to explore and interpret a comprehensive employee dataset, providing a deeper understanding of the company’s human capital landscape.

The dataset included key employee attributes such as ID, name, gender, race, department, job title, location, hire and termination dates, and more. An additional "age" column was created by deriving it from the birthdate field to support age-specific insights. The analysis aimed to uncover patterns related to employee distribution, diversity, hiring and termination trends, departmental and regional workforce structures, and tenure metrics.

This project plays a vital role in supporting HR strategy, from improving diversity and inclusion initiatives to enhancing talent acquisition and retention. By transforming raw HR data into meaningful insights, this analysis empowers stakeholders to make informed, strategic decisions that align with organizational goals

Objective

The primary objective of this HR and employee data analysis project was to derive meaningful insights from workforce data to support strategic decision-making within the organization. Specifically, the analysis aimed to:

1. **Understand Employee Demographics**
   Analyze the workforce composition by age, gender, race, and location to assess diversity and inclusion levels across the organization.
2. **Evaluate Departmental and Job Role Distribution**
   Identify the distribution of employees across departments and job titles to understand staffing structure and departmental capacity.
3. **Examine Employment Lifecycle Trends**
   Explore hiring and termination patterns over time to assess workforce stability, turnover rates, and tenure trends.
4. **Identify Location-Based Workforce Insights**
   Assess how employee distribution varies across cities and states to support regional workforce planning.
5. **Support Data-Driven HR Decision-Making**
   Generate visual and statistical insights that can inform strategic initiatives such as recruitment planning, diversity programs, and employee retention strategies.

Scope and Data Sources

This report focuses on the analysis of internal employee data to uncover patterns and trends that can inform human resource strategies within the organization. The scope is limited to exploring the workforce distribution, demographic diversity, job roles, departmental structure, employment lifecycle (hire and termination trends), and location-based insights.

The data was cleaned, transformed, and analyzed using **MySQL**, with key metrics and patterns visualized through interactive dashboards in **Power BI**. The dataset reflects a snapshot of employee records over a defined period and is assumed to be representative of the organization's workforce during that time.

This report does not include performance data, compensation, training records, or employee feedback. Instead, it focuses solely on static and time-bound employee information for structural and demographic analysis.

Methodology

The analysis followed a structured approach that combined data extraction, cleaning, transformation, and visualization to ensure accuracy and relevance of the insights. The methodology involved the following key steps:

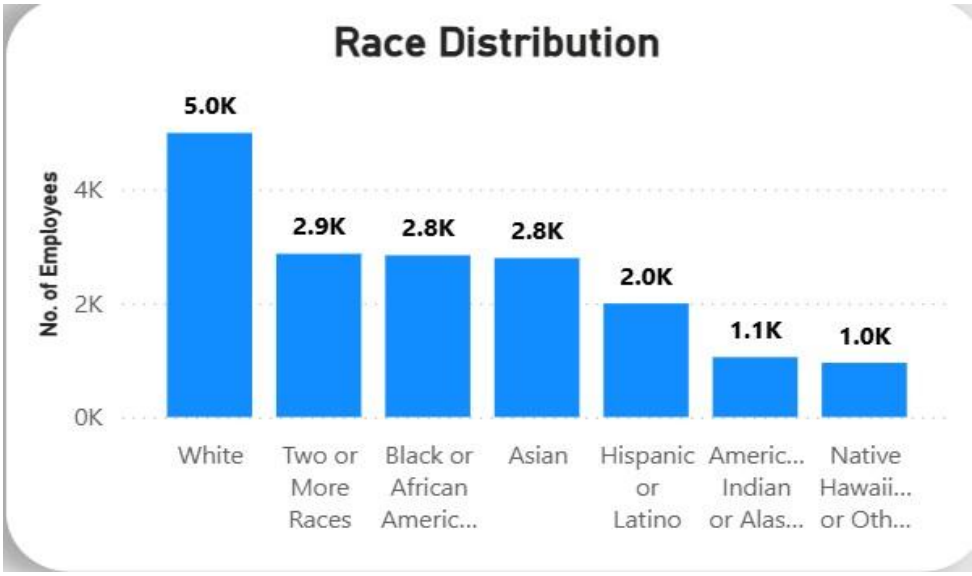
1. **Data Acquisition and Preparation**
   The raw dataset was acquired from the organization's HR information system and imported into a **MySQL database** for preprocessing. The data included employee demographic details, job-related information, and employment dates.
2. **Data Cleaning and Transformation**
   Using **MySQL**, the data was cleaned to handle missing values, format inconsistencies, and irrelevant fields. Key transformations included:
   - Standardizing date formats (birthdate, hire_date, term_date)
   - Calculating a new column, age, by subtracting birth year from the current year
   - Filtering out incomplete or invalid records (e.g., missing hire dates)
   - Deriving tenure-related metrics for analysis of employment duration 3.

   Exploratory Data Analysis (EDA)
   SQL queries were used to extract relevant statistics and uncover patterns such as:
   - Distribution of employees by department, gender, race, and age
   - Hiring and termination trends over time
   - Geographic distribution of employees across cities and states
   - Job title concentration and departmental composition
4. **Data Visualization**
   The cleaned and structured data was imported into **Power BI** for report and dashboard development. A series of interactive charts, graphs, and slicers were designed to enable dynamic filtering and deeper exploration of key metrics.
5. **Insight Generation**
   Insights were derived by interpreting visual patterns and relationships across variables. These findings form the basis for actionable recommendations aimed at enhancing workforce planning, diversity management, and HR policy development.

This methodological approach ensured the integrity of the data, the clarity of the visuals, and the relevance of the insights to organizational decision-making.

Data Analysis and Results

A. Race Distribution



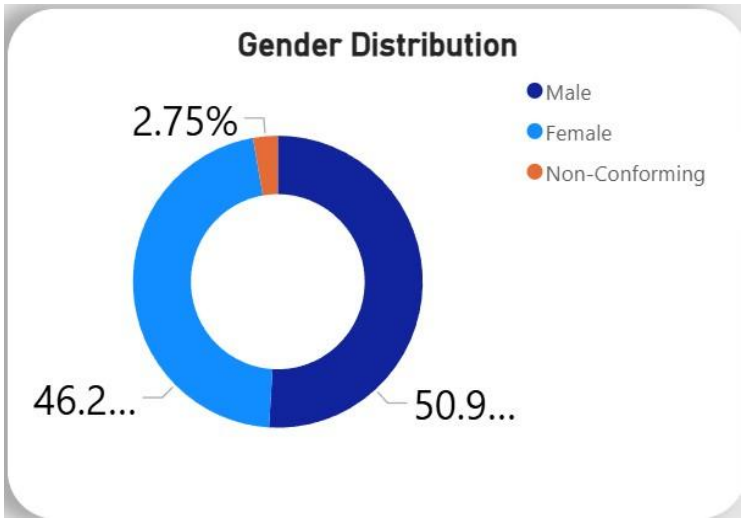
Key Insights:

1. **Predominance of White Employees**
   The bar chart shows that **White** employees make up the largest racial group within the organization. This indicates a potential racial imbalance, depending on the geographical context and organizational diversity goals.
2. **Underrepresentation of Other Races**
   - **Hispanic**, **Asian**, and **Black or African American** employees follow, but their counts are noticeably lower.
   - **Native Hawaiian or Other Pacific Islander**, **American Indian or Alaska Native**, and those marked as **Two or More Races** are represented in much smaller numbers, indicating limited diversity across these groups.
3. **Possibility of Diversity Gaps**
   The sharp skew toward one dominant group suggests the need for reviewing recruitment and retention strategies to promote racial and ethnic diversity.

Insight Summary:

- The racial composition of the organization is heavily skewed toward White employees, with minority groups significantly underrepresented.
- There may be systemic or cultural barriers that limit the inclusion or hiring of racially diverse talent.

B. Gender Distribution



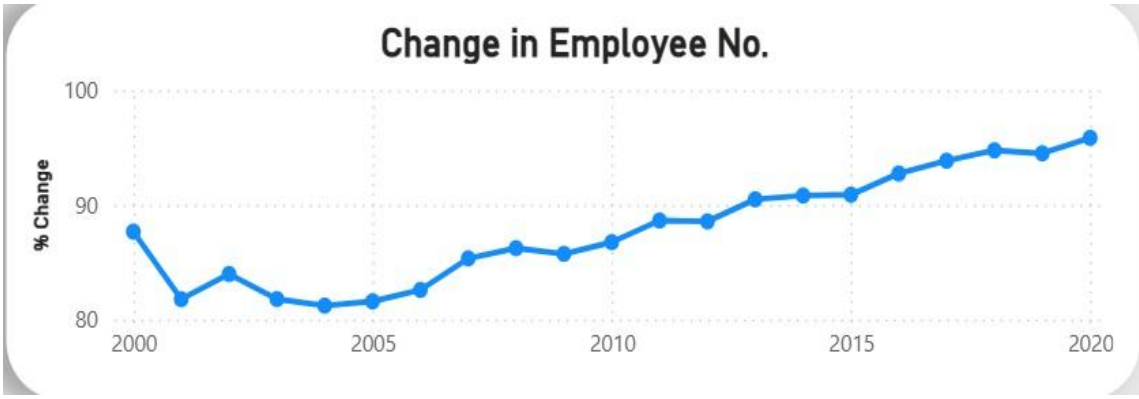
Key Insights:

1. **Male-Dominated Workforce**
   The donut chart shows that **Male employees make up the majority** of the organization's workforce making up for approximately **50%.** This indicates a gender imbalance that may be worth investigating further, especially in roles or departments traditionally dominated by one gender.
2. **Female Representation is Lower**
   While **Female employees are present in significant numbers**, they are noticeably fewer than males making up for approximately **46%.** This may point to disparities in either:
   - Hiring practices,
   - Retention trends, or
   - Departmental concentration (e.g., more males in technical or leadership roles).
3. **Small Representation of Non-Conforming or Other Genders**
   There is little visible representation of **non-binary or gender-diverse individuals**, who only make up for approximately **2.7%** which may suggest:
   - Lack of inclusive gender options in employee records,
   - Or underreporting due to cultural or organizational limitations.

Insight Summary :

- The organization exhibits a **gender imbalance**, with **Male employees being more dominant** in the workforce.
- Although **Female representation is present**, the gap suggests potential challenges in achieving gender equity—especially if males are overrepresented in leadership, technical, or high-paying roles.

C. Employee Count Change Over Time



Key Insights :

1. **2000–2005: Low and Stable Headcount**
    - From the year **2000**, the % change clocks at approximately **89%** and drops all the way to **83%** by the year **2001.**
    - It slightly spikes to **85%** within a year and later drops gradually to **81%** up to the year **2005**
2. **2005–2010: Slight Uptick Begins**
    - Throughout this 5-Year period, the % change has increased considerably and reaching **87%** by the end of the year **2010**
    - It suggests the organization began **scaling up slowly**, possibly setting the stage for expansion in the next phase.
3. **2010–2015: Peak Growth and Rapid Expansion**
    - During this period the % change begins at **87%** and by the year **2015** it reaches a peak of **91%**
    - This period shows the **highest surge in employee count**, marking the company's **peak growth phase**.
    - A significant number of employees were hired during this time, likely due to **business growth, new projects, or strategic expansion**.

4. **2015–2020: Steady Decline in Workforce**

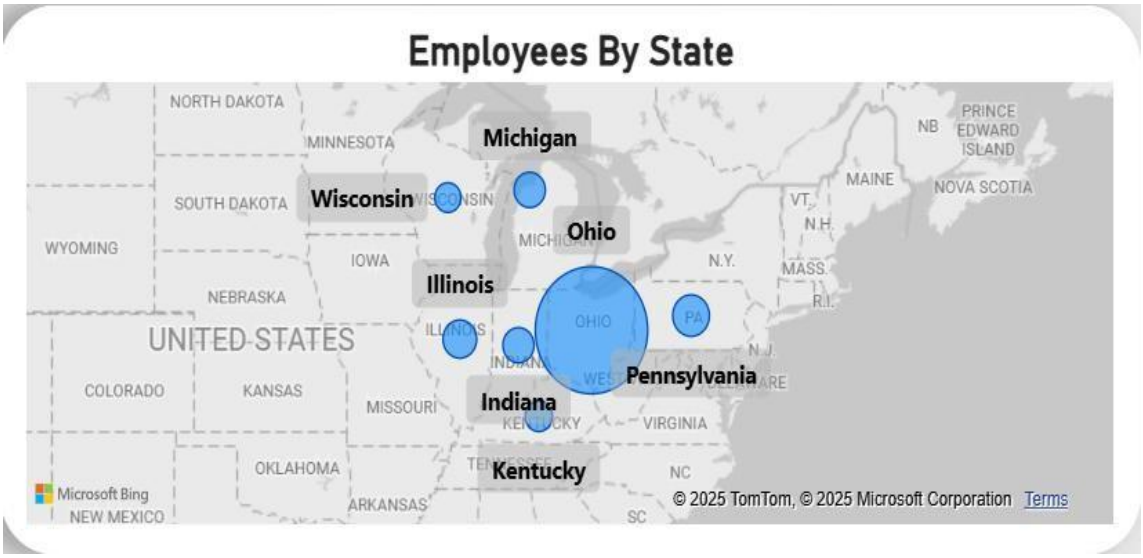
- Following the previous year, 2015, which ended at 91%, it finally peaks at **96%** by the year **2020**.
- After peaking, the chart shows a **consistent incline in employee numbers** throughout this final period.

Insight Summary:

- The company saw **minimal activity from 2000 to 2005**, followed by a **gradual buildup (2005–2010)**.
- **2015–2020 was the prime growth phase**, with the **largest increase in staff count**.

- The **2005–2010 period marks a clear reversal**, with a consistent increase in employee numbers—an important area for further HR analysis and action.

D. Distribution of Employees By State



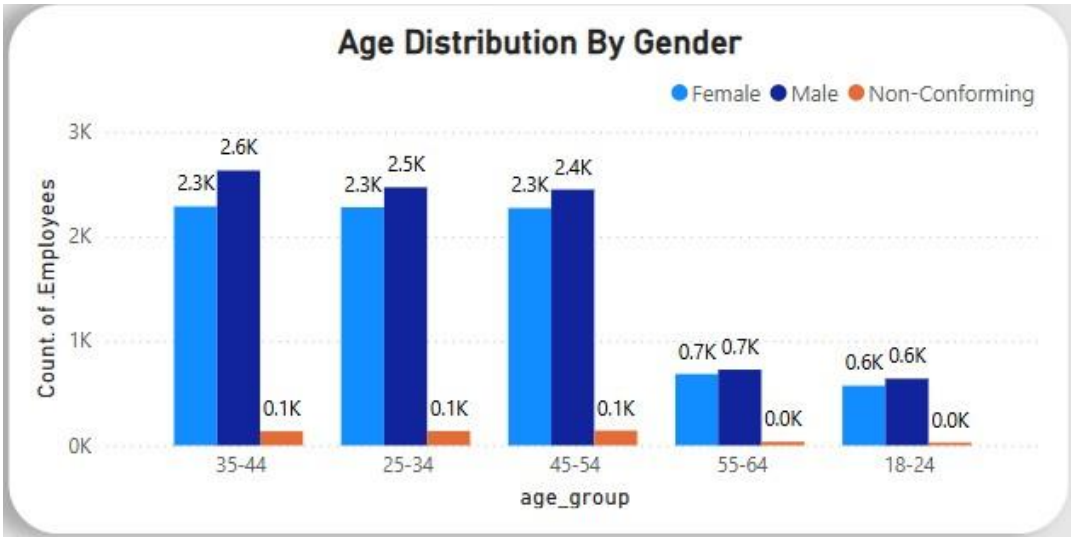
Key Insights:

1. **States with the Largest Bubbles (Highest Employee Count):**
    - **Ohio** clearly has the **largest bubble**, indicating it hosts the **highest number of employees** among all states.
      Likely a major headquarters or central operational hub.
    - **Pennsylvania**, **Illinois,** and **Michigan** follow closely with **large-sized bubbles**, signaling **significant employment presence**.
      These could be major regional offices, tech centers, or strategic market zones.
2. **Moderately Represented States (Medium Bubbles):**
    - **Indiana**, **Kentucky**, **Georgia**, and **Wisconsin** show **medium-sized bubbles**.
      These likely represent **mid-sized offices** or **concentrated teams** supporting core functions like sales, support, or marketing.

Insights Summary:

- **Ohio** leads in employee count, followed by **Pennsylvania, Illinois,** and **Michigan**.
- The **employee distribution is heavily clustered** in a few strategic states — primarily **urban, economically vibrant, and coastal**.
- The **Midwest and rural states** have **limited or no presence**, reflecting either market strategy or infrastructural constraints.

E. Employee Age Distribution By Gender

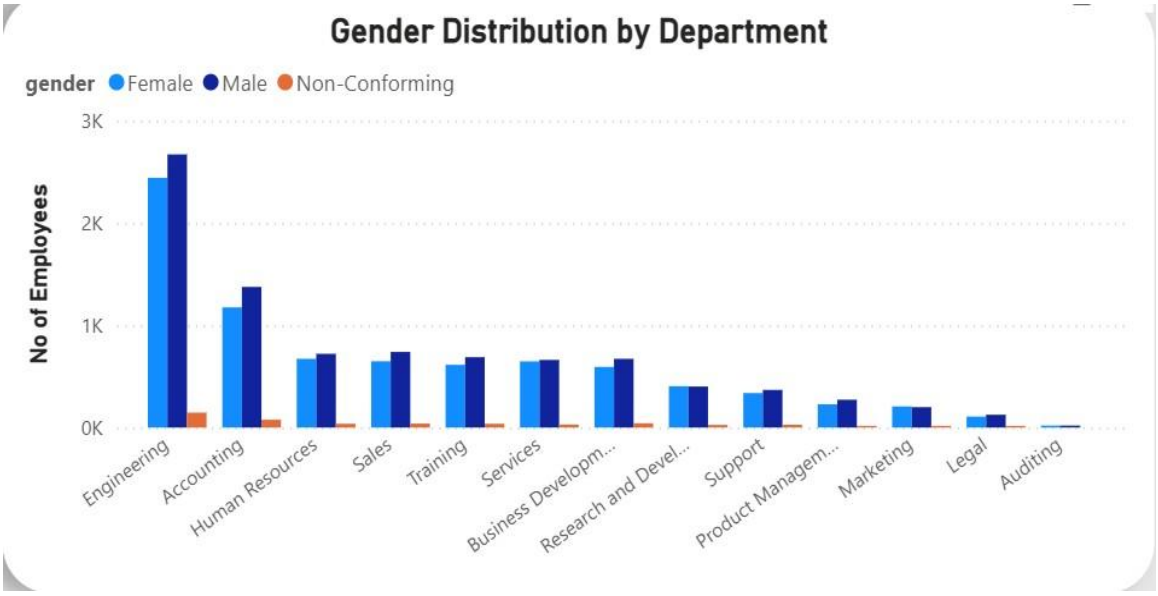


1. **Dominant Age Groups**:
    - The **35-44** age group has the highest employee count (**2.6K**), followed closely by **25-34** and **45-54** (both at **2.3K**).
    - The smallest groups are **18-24** (**0.6K**) and **55-64** (**0.7K**), indicating fewer younger and older employees.
2. **Gender Distribution**:
    - **Female** and **Male** employees are nearly balanced in most age groups, with minor variations (e.g., **45-54** has **2.3K** females vs. **2.4K** males).
    - **Non-Conforming** employees are significantly underrepresented, with only **0.1K** in each age group.
3. **Peak Workforce Age**:
    - Employees aged **25-54** (early career to mid-career) form the core workforce, suggesting the company attracts or retains professionals in this range.

Insights Summary

The workforce is **middle-aged (25-54) dominated**, with balanced binary gender representation but minimal diversity in non-conforming genders**. Younger (18-24)** and **older (55-64**) employees are a small minority, which may reflect hiring trends, retention challenges, or industry norms. The lack of non-conforming representation highlights potential gaps in inclusivity efforts.

F. Employee Gender Distribution By Department



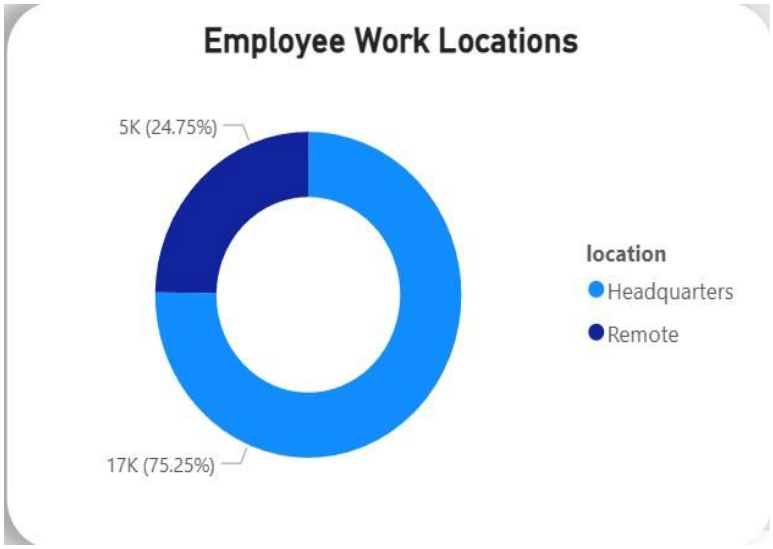
Key Insights

1. **Overall Gender Representation**:
    - **Female** and **Male** employees dominate across all departments, with **Female** representation slightly higher in most cases (e.g., **2K** vs. **1K** in some departments).
    - **Non-Conforming** employees are nearly absent (close to **0K**), indicating a significant lack of gender diversity beyond binary categories.
2. **Department-Specific Trends**:
    - **Female** employees appear to outnumber **Male** employees in certain departments (e.g**., Human Resources, Training**), while others (e.g., **Engineering, Legal**) may show closer parity or male dominance (exact numbers are unclear due to chart resolution).
    - No department shows meaningful representation of **nonconforming** individuals, suggesting systemic underrepresentation.
3. **Potential Imbalances**:
    - If certain departments (e.g., **Engineering, Legal**) skew heavily toward one gender, this could indicate hiring biases, cultural barriers, or retention issues for underrepresented genders.

Insight Summary

The workforce exhibits **binary gender dominance (Female/Male)**, with **Non-Conforming employees marginalized** across all departments. While some departments (e.g., **HR**) may lean female, others (e.g., **Engineering**) may reflect traditional gender imbalances. The near-zero representation of **Non-Conforming** individuals highlights a critical gap in inclusivity.

G. Employee Work Locations



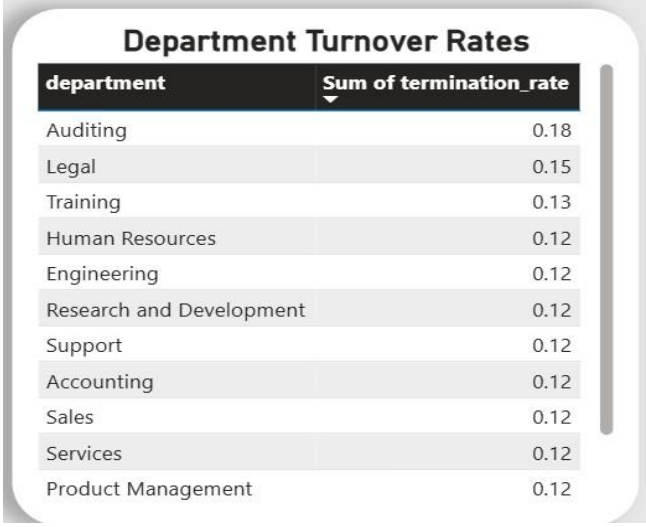
Key Insights

1. **Dominant Work Location**:
    - **Headquarters (HQ)** employs the vast majority of workers at **75.23% (17K employees)**.
    - **Remote** work accounts for only **24.73% (5K employees)**, indicating a primarily office-centric workforce.
2. **Implications of the Split**:
    - The company heavily favors **on-site work**, which may reflect organizational culture, role requirements, or leadership preferences.
    - The relatively small remote workforce (**~25%**) suggests either **limited flexibility policies** or that most roles are location-dependent (e.g., manufacturing, labbased).

Insight Summary

The workforce is **overwhelmingly centralized at Headquarters**, with remote work being a secondary option. This could impact **talent acquisition** (limiting geographic diversity), **employee satisfaction** (if flexibility is expected), and **resilience** (e.g., adaptability to disruptions).

I. Department turnover Rates



Department Turnover Rates

department	Sum of termination_rate
Auditing	0.18
Legal	0.15
Training	0.13
Human Resources	0.12
Engineering	0.12
Research and Development	0.12
Support	0.12
Accounting	0.12
Sales	0.12
Services	0.12
Product Management	0.12

Key Insights

1. **Highest Turnover Departments**
   Auditing (0.18) and **Legal (0.15)** have the **highest termination rates**, significantly above the company average (**0.12**).
   - *Potential Causes*: High stress, workload imbalances, or misaligned role expectations in these high-compliance areas.
2. **Moderate Turnover Departments**
   Training (0.13) and **Human Resources (0.12)** follow, suggesting challenges in employee development or internal support roles.
3. **Stable Departments**
   Most departments (e.g., **Engineering, Sales, Accounting**) cluster around **0.12**, indicating consistent but manageable attrition.
   - *Implication*: This could reflect industry norms, but still warrants monitoring for trends.
4. **Uniformity in Mid-Range Turnover**
   10 of 11 departments fall within a narrow band (**0.12–0.13**), except Auditing/Legal.
   - *Takeaway*: The company may have systemic retention issues (e.g., culture, compensation) affecting most teams.

Insight Summary

Turnover is **highest in compliance-driven roles (Auditing/Legal)**, likely due to job pressures, while other departments face uniform but moderate attrition. The consistency across most teams suggests **company-wide retention challenges**, whereas outliers (Auditing/Legal) need targeted interventions

H. Employee Distribution By City

Key Insights

1. Geographic Concentration

- Employees are **heavily concentrated in North America and Europe**, with minimal presence in **Africa and South America**.
- **Mansard Bay** (likely a corporate HQ or major office) is explicitly highlighted, suggesting it's a central hub.

2. Regional Imbalances

- **Africa and South America** show sparse representation, which may reflect:
  - ✓ Limited business operations in these regions.
  - ✓ Missed talent pool opportunities in emerging markets.

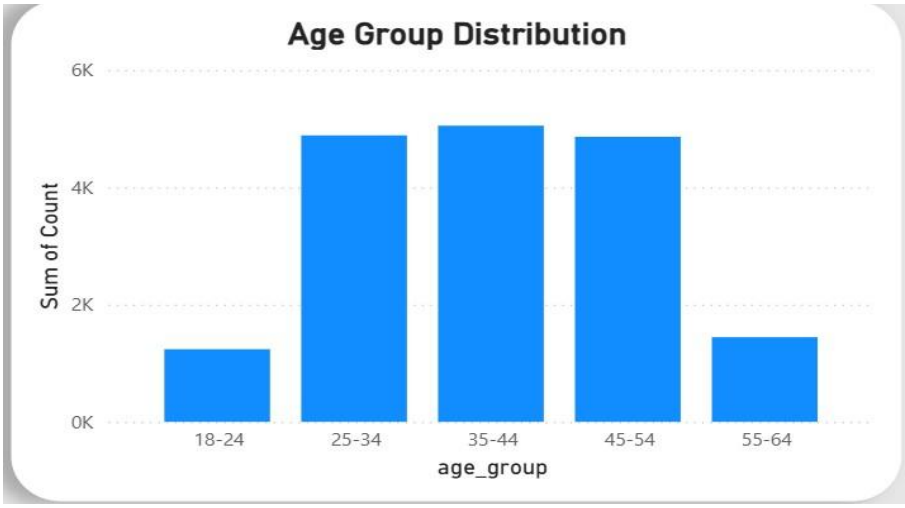
3. Implications of the Distribution

- **Pros**: Focused infrastructure in key markets (**NA/EU**) may streamline operations.
- **Cons**: Lack of geographic diversity could limit cultural perspectives, market adaptability, and global talent acquisition.

Insight Summary

The workforce is **overwhelmingly centralized in North America and Europe**, with negligible representation in other continents. This may align with the company's business model. but it could pose risks for global scalability and diversity.

I. Age Group Distribution



Key Insights

1. Workforce Demographics by Age

- **25–34** is the **largest age group (4K employees)**, representing early to mid-career professionals.
- **35–44** follows as the second-largest group (**2K employees**), indicating a strong midcareer presence.
- Younger (**18–24**) and older (**55–64**) employees are the smallest cohorts (**≤1K each**), suggesting limited representation of entry-level and pre-retirement workers.

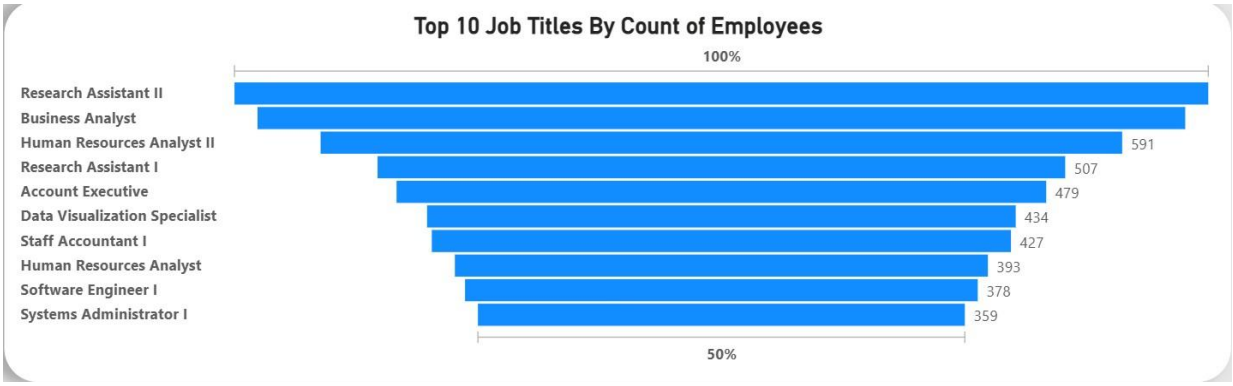
2. Implications of the Distribution

- **Early/Mid-Career Dominance**: The workforce skews toward employees in their 20s–40s, which may reflect:
    - ✓ A focus on hiring for roles requiring moderate experience.
    - ✓ Potential gaps in early talent pipelines (interns/apprentices) or retention of older employees.
- **Low Representation of Younger/Older Workers**:
    - ✓ **18–24**: Could indicate limited campus recruiting or barriers to entry-level roles.
    - ✓ **55–64**: May signal early retirements, age bias in hiring, or lack of phased retirement options.

Insight Summary

The workforce is **youthful and mid-career heavy**, with employees aged **25–44** making up the majority. This suggests a focus on experienced hires but potential vulnerabilities in knowledge transfer (due to fewer older workers) and pipeline development (due to fewer young hires).

J. Top 10 Job Titles By Count of Employees



Key Insights

1. Dominant Job Titles

- **Research Assistant II (591 employees)** is the most common role, followed by **Business Analyst (507)** and **Human Resources Analyst II (479)**.
- **Technical roles** like **Software Engineer I (393)** and **Systems Administrator I (359)** appear lower on the list, suggesting a stronger focus on research, business, and HR functions.

2. Workforce Composition Trends

- **Support & Operational Roles Dominate**: The prevalence of **Research Assistants, HR Analysts, and Account Executives** indicates a workforce heavily skewed toward **support functions** rather than core technical or leadership roles.
- **Entry-Level Titles**: Many titles include **"I" or "II"** (e.g., *Research Assistant I/II*, *Staff Accountant I*), implying a large **early-career or mid-level workforce** with fewer senior positions listed.

3. Potential Organizational Insights

- **Research-Driven Culture**: High numbers of **Research Assistants** suggest the company may prioritize R&D or data-driven projects.
- **Possible Talent Pipeline Gaps**: Fewer senior/leadership titles (e.g., no "Manager" or "Director" roles in the top 10) could indicate a **flat hierarchy** or challenges in career progression.

Insight Summary

The workforce is **heavily concentrated in research, business analysis, and HR support roles**, with fewer employees in technical or leadership positions. This may reflect the company's **operational priorities** (e.g., a focus on research services) or **structural gaps** (e.g., lack of upward mobility for junior staff).

Discussion

1. What the Results Mean in Context

The **predominance of White employees** and **male dominance in the workforce** suggest a potential lack of diversity, particularly in racial and gender representation. Although these trends may partially reflect broader societal or regional labor market dynamics, they raise important questions about the organization's **inclusivity, recruitment policies**, and **internal workplace culture**.

Geographically, states like **Ohio, Pennsylvania, Illinois, and Michigan** show a significant concentration of employees, indicating key operational hubs or regional headquarters. This spatial clustering can have implications for talent pipeline development, regional diversity, and targeted HR strategies.

The **temporal change in employee count** shows clear phases in the company's growth. A **low and stable headcount (2000–2005)** transitions into **gradual buildup (2005–2010)**, **rapid expansion (2010–2015)**, and **further incline into 2020**, suggesting a company that has matured and grown considerably over the past two decades. These periods may correspond with strategic growth initiatives, business expansion, or organizational restructuring efforts.

The data reveals a **middle-aged (25-44), HQ-centric workforce** with strong representation in research and support roles but notable gaps in:

- **Geographic diversity** (concentrated in **NA/EU**).
- **Gender diversity** (non-conforming employees nearly absent).
- **Age diversity** (few employees under 25 or over 55).
- **Role diversity** (few senior titles in top job listings).

This suggests a **traditional, office-driven culture** with potential blind spots in inclusion, global talent acquisition, and career progression.

2. How the Data Answers Key Research Questions

- **"Are we retaining talent effectively?"**
    - ✓ Turnover is highest in **Auditing (0.18)** and **Legal (0.15)**, signaling burnout or misalignment in high-pressure departments.
    - ✓ Most departments hover at **0.12**, indicating systemic retention challenges.
- **"Is our workforce diverse?"**
    - ✓ **No**: Gender diversity is binary (Female/Male dominate; non-conforming employees excluded).
    - ✓ **No**: Geographic diversity is limited to NA/EU.
    - ✓ **Partial**: Age diversity skews mid-career, lacking younger/older representation.

- **"Where should we focus hiring/development?"**
    - ✓ **Remote work**: Only 24.73% of employees are remote, potentially limiting talent pools.
    - ✓ **Technical roles**: Underrepresented in top job titles (e.g., fewer engineers than analysts).

3. Limitations of the Analysis

- **Data Granularity**: Missing breakdowns (e.g., gender by department, remote work by role) limit deeper insights.
- **Cultural Context**: Reasons for disparities (e.g., low non-conforming representation) aren't explained by the data alone.
- **Industry Benchmarks**: Without external comparisons, it's unclear if turnover/age skews are atypical.
- **Lack of Contextual Benchmarking:** The analysis does not compare the company's diversity figures with industry or regional labor market benchmarks. This limits the ability to determine if representation is disproportionately low.
- **Potential Underreporting of Gender Diversity:** The small percentage of non-binary or gender-diverse employees may not reflect reality, as this could be influenced by lack of disclosure options or comfort in self-identification.

Recommendations

1. Promote Racial Diversity and Inclusion

Insight Link: Predominance of White employees and underrepresentation of other races.

- **Implement Targeted Recruitment Campaigns:** Focus on universities, professional associations, and job boards that serve underrepresented racial groups.
- **Set Diversity Benchmarks:** Establish racial diversity KPIs aligned with broader DEI goals to track progress in hiring and retention.
- **Create Safe Reporting & Feedback Channels:** Enable anonymous employee feedback and open forums for minority voices to express their experiences.
- **Review Hiring Panels:** Ensure diverse interview panels to reduce unconscious bias during candidate selection.

2. Address Gender Imbalance

Insight Link: Males make up the majority workforce (~50%) while females (~46%) and nonbinary (~2.7%) trail behind.

- **Support Female Career Advancement:** Launch mentorship programs and leadership training for women to bridge any leadership gaps.
- **Audit Role Distribution by Gender:** Assess whether certain roles or departments are gender-skewed (e.g., technical roles heavily male).
- **Inclusive Policies:** Provide gender-neutral parental leave, flexible working arrangements, and anti-discrimination training.
- **Update HR Systems:** Allow for more inclusive gender identification options beyond the binary categories.

3. Leverage Regional Workforce Dynamics

Insight Link: Ohio, Pennsylvania, Illinois, and Michigan hold the largest employee clusters.

- **Regional Workforce Planning:** Consider whether these hubs are at risk of being overconcentrated and explore the feasibility of redistributing talent to underutilized states.
- **Expand Talent Acquisition in Underserved States:** Target recruitment campaigns in low-representation areas (based on the map) to access fresh talent pools.
- **Enhance Remote Work Infrastructure:** Enable distributed teams to reduce overreliance on a few physical locations and promote regional equity.

4. Align Workforce Growth with Strategic Planning

Insight Link: Rapid employee growth observed between 2010–2020, especially 2015–2020.

- **Evaluate Onboarding & Training Programs:** Ensure your rapid hires from growth phases are effectively integrated and upskilled.

- **Conduct Workforce Planning Reviews:** Assess whether past expansions led to longterm productivity or if course corrections are needed.
- **Forecast Future Trends:** Use historical hiring data to model expected growth/contraction periods and align hiring with business cycles.

5. Launch Return-Ship Programs

Insight Link: Middle-aged dominance (25-54), minimal non-conforming representation.

- **Launch "Return-ship" Programs**: Target professionals 55+ to fill experience gaps (e.g., mentorship roles).
- **Non-Conforming Inclusion**: Partner with LGBTQ+ orgs for recruitment and audit HR policies for inclusivity.

6. Regular Auditing

Insight Link: High attrition in Auditing (0.18) and Legal (0.15).

- **Stress Audits**: Anonymous surveys in Auditing/Legal to identify pain points (e.g., workload, lack of tools).
- **Rotation Programs**: Cross-train employees in Legal/Auditing to prevent burnout (e.g., 6-month rotations).

7. Introduce Apprenticeship programs

Insight Link: Few employees under 25 or over 55.

- **Apprenticeships**: Create paid apprenticeships for 18-24-year-olds in high-turnover depts (e.g., Auditing).
- **Phased Retirement**: Offer part-time roles for 55-64 employees to retain institutional knowledge.

8. Introduce Promotional Paths for Employees

Insight Link: Research/HR roles dominate; few senior titles.

- **Leadership Tracks**: Define clear promotion paths for "Analyst II" to "Manager" roles.
- **Skills Mapping**: Audit if technical roles (e.g., Software Engineers) are undercounted due to outsourcing.

Conclusion

This HR Employee Distribution Analysis has provided a comprehensive overview of workforce patterns across key dimensions such as race, gender, geographic location, and temporal changes in employee count. By leveraging MySQL for data analysis and Power BI for dynamic visualization, the project revealed important insights into the organization's demographic composition and growth trends.

This analysis reveals a highly centralized workforce, mid-career dominated, and lacking diversity in gender, age, and geography. Key challenges include:

- **Retention risks** in high-pressure departments (Auditing/Legal).
- **Inclusion gaps**, particularly for non-conforming genders and older/younger employees.
- **Geographic rigidity**, with over 75% of employees based at HQ, limiting global talent acquisition.
- **Role imbalances**, with research and support roles overshadowing technical and leadership positions.

The findings highlight notable imbalances, including a workforce predominantly composed of White and male employees, with minority racial groups and gender-diverse individuals significantly underrepresented. Geographically, certain states—particularly Ohio, Pennsylvania, Illinois, and Michigan—host a large concentration of employees, indicating regional clustering that may influence organizational operations and talent strategy. The temporal analysis further uncovered a major growth phase between 2010 and 2020, which aligns with a possible period of organizational expansion.

These insights underscore the need for data-driven decision-making in HR policy formulation. Recommendations such as enhancing diversity recruitment efforts, promoting gender inclusivity, and optimizing regional workforce planning offer a roadmap for more equitable and strategic talent management. While the analysis was insightful, limitations such as the lack of departmentlevel granularity and potential underreporting in demographic categories suggest areas for further exploration.

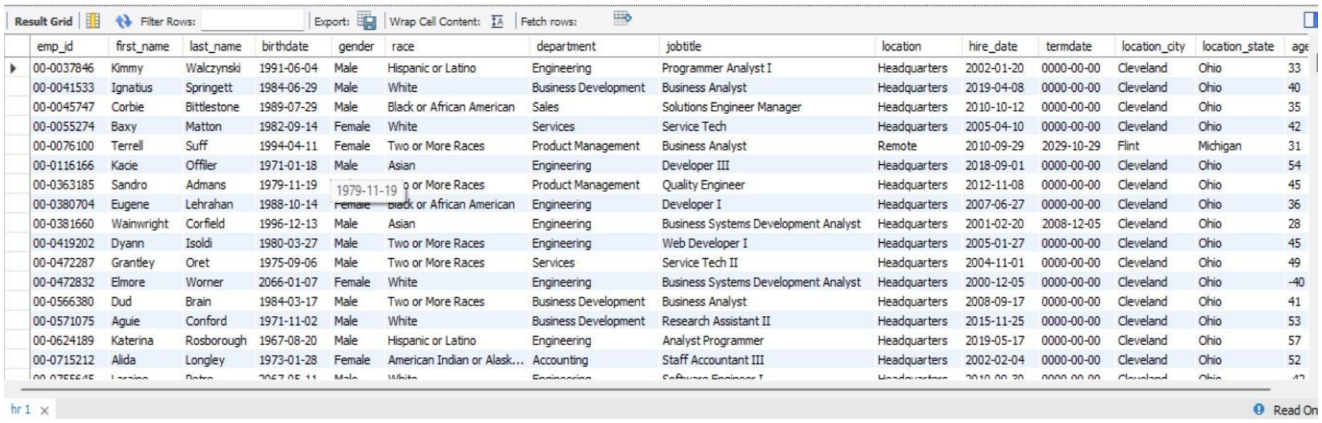
Ultimately, this report serves as a baseline for continuous improvement and strategic workforce planning. By embracing data-informed HR practices, the organization can foster a more inclusive, balanced, and resilient workplace.

References

1. Randel, A. E. (2023). Inclusion in the Workplace: A Review and Research Agenda. Group & Organization Management, 50(1), 119-162. https://doi.org/10.1177/10596011231175578 (Original work published 2025)
2. Ng E. S. and Sears G. J., "Walking the Talk on Diversity: CEO Beliefs, Moral Values, and the Implementation of Workplace Diversity Practices," *Journal of Business Ethics volume* 164, p. 437–450, 2020.

Appendices

Data Preview



Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

emp_id	first_name	last_name	birthdate	gender	race	department	jobtitle	location	hire_date	termdate	location_city	location_state	age
00-0037846	Kimmy	Walczynski	1991-06-04	Male	Hispanic or Latino	Engineering	Programmer Analyst I	Headquarters	2002-01-20	0000-00-00	Cleveland	Ohio	33
00-0041533	Ignatius	Springett	1984-06-29	Male	White	Business Development	Business Analyst	Headquarters	2019-04-08	0000-00-00	Cleveland	Ohio	40
00-0045747	Corbie	Bittlestone	1989-07-29	Male	Black or African American	Sales	Solutions Engineer Manager	Headquarters	2010-10-12	0000-00-00	Cleveland	Ohio	35
00-0055274	Baxy	Matton	1982-09-14	Female	White	Services	Service Tech	Headquarters	2005-04-10	0000-00-00	Cleveland	Ohio	42
00-0076100	Terrell	Suff	1994-04-11	Female	Two or More Races	Product Management	Business Analyst	Remote	2010-09-29	2029-10-29	Flint	Michigan	31
00-0116166	Kacie	Offiler	1971-01-18	Male	Asian	Engineering	Developer III	Headquarters	2018-09-01	0000-00-00	Cleveland	Ohio	54
00-0363185	Sandro	Admans	1979-11-19		or More Races	Product Management	Quality Engineer	Headquarters	2012-11-08	0000-00-00	Cleveland	Ohio	45
00-0380704	Eugene	Lehrahan	1988-10-14	Female	Black or African American	Engineering	Developer I	Headquarters	2007-06-27	0000-00-00	Cleveland	Ohio	36
00-0381660	Wainwright	Corfield	1996-12-13	Male	Asian	Engineering	Business Systems Development Analyst	Headquarters	2001-02-20	2008-12-05	Cleveland	Ohio	28
00-0419202	Dyann	Isoldi	1980-03-27	Male	Two or More Races	Engineering	Web Developer I	Headquarters	2005-01-27	0000-00-00	Cleveland	Ohio	45
00-0472287	Grantley	Oret	1975-09-06	Male	Two or More Races	Services	Service Tech II	Headquarters	2004-11-01	0000-00-00	Cleveland	Ohio	49
00-0472832	Elmore	Worner	2066-01-07	Female	White	Engineering	Business Systems Development Analyst	Headquarters	2000-12-05	0000-00-00	Cleveland	Ohio	-40
00-0566380	Dud	Brain	1984-03-17	Male	Two or More Races	Business Development	Business Analyst	Headquarters	2008-09-17	0000-00-00	Cleveland	Ohio	41
00-0571075	Aguie	Conford	1971-11-02	Male	White	Business Development	Research Assistant II	Headquarters	2015-11-25	0000-00-00	Cleveland	Ohio	53
00-0624189	Katerina	Rosborough	1967-08-20	Male	Hispanic or Latino	Engineering	Analyst Programmer	Headquarters	2019-05-17	0000-00-00	Cleveland	Ohio	57
00-0715212	Alida	Longley	1973-01-28	Female	American Indian or Alask...	Accounting	Staff Accountant III	Headquarters	2002-02-04	0000-00-00	Cleveland	Ohio	52

1979-11-19

hr 1 ×

Read On

SQL QUERY SAMPLES
Data Cleaning

```
#Converting hiredate column to date format
UPDATE hr
SET hire_date = CASE
    WHEN hire_date LIKE '%/%' THEN date_format(str_to_date(hire_date,'%m/%d/%Y'), '%Y-%m-%d')
    WHEN hire_date LIKE '%-%' THEN date_format(str_to_date(hire_date,'%m-%d-%Y'), '%Y-%m-%d')
ELSE NULL
END;

ALTER TABLE hr
MODIFY COLUMN hire_date DATE;

#Converting termdate column to date format
UPDATE hr
SET termdate = IF(termdate IS NOT NULL AND termdate != '', date(str_to_date(termdate, '%Y-%m-%d %H:%i:%s UTC')), '0000-00-00')
WHERE true;
```

(Converting the 'Hire-date' Column to MySQL Date Format)

```
SET sql_mode = 'ALLOW_INVALID_DATES';
ALTER TABLE hr
MODIFY COLUMN termdate DATE;

#Adding a new column'age' for easier calculations
ALTER TABLE hr
ADD COLUMN age INT;

# Calculating the age values
UPDATE hr
SET age = timestampdiff(YEAR,birthdate,CURDATE());

SELECT birthdate,age FROM hr;
```

(Introducing New Columns and Calculating its Values)

Data Analysis & Explorations

```
# Age Distribution of Employees
SELECT
    MIN(age) AS youngest,
    MAX(age) AS oldest
    FROM hr
    WHERE age >= 18 AND termdate = '0000-00-00';

SELECT
    CASE
    count(*) AS Count
    FROM hr
    WHERE age >= 18 AND termdate = '0000-00-00'
    GROUP BY age_group
    ORDER BY age_group;

SELECT
    CASE
    count(*) AS Count
    FROM hr
    WHERE age >= 18 AND termdate = '0000-00-00'
    GROUP BY age_group,gender
    ORDER BY age_group,gender;
```

(Calculating the Age Distribution of Employees)