

Employment Structure by Education Level in Finland

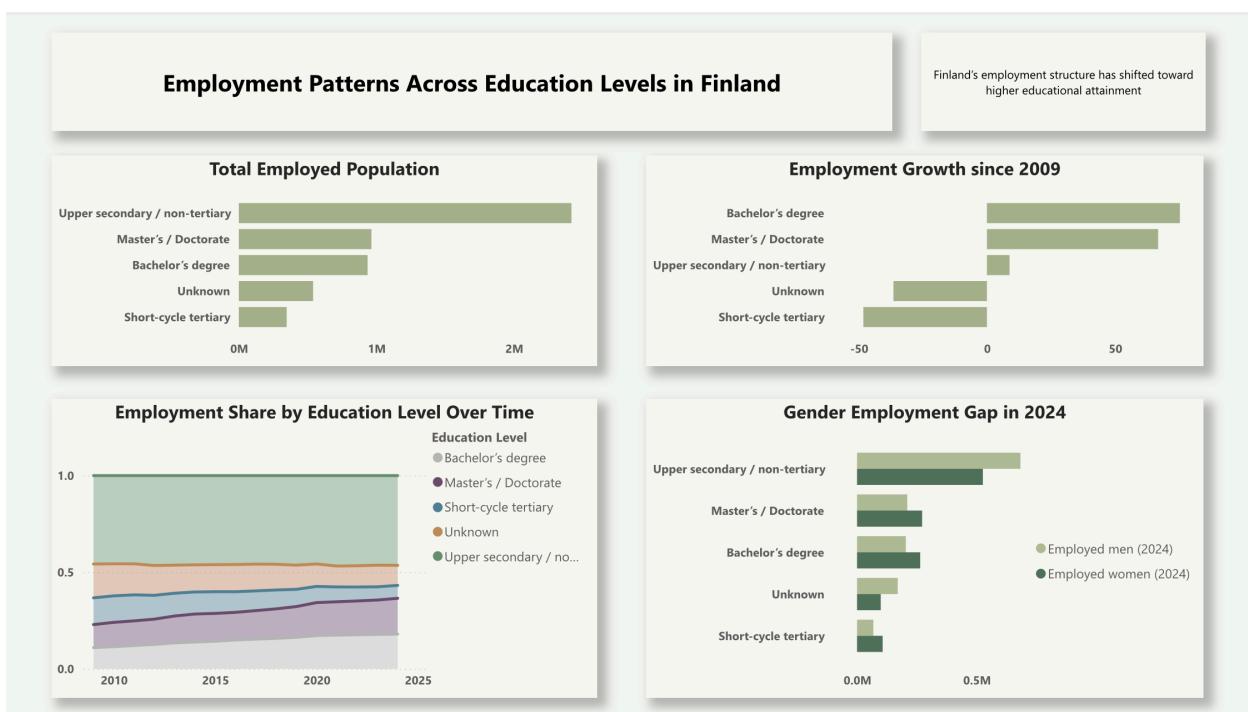
Introduction

This project analyzes how employment in Finland is distributed across education levels, how this structure has evolved since 2009, and how growth patterns and gender differences vary by educational attainment. The goal was to demonstrate an end-to-end analytics workflow: from raw public data to clean SQL analysis and interactive business-style visualization.

The project emphasizes interpretation and communication, not just querying, mirroring how data analysts work with labor-market and socio-economic data in practice.

Key Questions

- How is total employment distributed across education levels?
- How has the composition of employment changed over time?
- Which education groups have contributed most to employment growth since 2009?
- Are gender employment gaps present, and how do they vary by education level?



Skills demonstrated

- SQL-based data cleaning, aggregation, and analytical modeling
- Labor-market and categorical data analysis
- Data validation with **Python** (pandas)
- Interactive dashboard development in **Power BI**
- Clear analytical communication and insight generation

Tools & Methods

SQL, Python, PowerBI

Data

- Statistics Finland (StatFin) - *Employed persons aged 15–74 by educational level and sex, 2009–2024*

Education categories follow the ISCED classification. For clarity and analytical validity, aggregate categories (e.g. Total and Tertiary (all)) were excluded to avoid overlap and double-counting.

Data Preparation (SQL)

- Imported raw StatFin data into SQLite
- Cleaned and standardized education categories
- Created analytical SQL views with:
 - Consistent education labels
 - Removal of overlapping aggregate categories
 - Conversion of employment figures to absolute counts
- Calculated:
 - Employment shares
 - Absolute employment levels
 - Absolute and relative employment growth since 2009
 - Gender-specific employment counts

Analysis & Visualization

- Exported analytical outputs into a single long-format dataset
- Built an interactive **Power BI dashboard** (Power BI Online)

Python (pandas, matplotlib/seaborn) was used for validation and exploratory analysis, while Power BI served as the primary visualization and presentation tool.

Key Insights

- Upper secondary / non-tertiary education remains the largest employment group, accounting for nearly half of all employed persons, reflecting its prevalence in the workforce rather than higher employment prospects.
- Employment has shifted toward higher educational attainment over time, with Bachelor's and Master's/Doctorate groups increasing their share of total employment.
- Employment growth since 2009 has been strongest among tertiary-educated groups, particularly at the Bachelor's and Master's/Doctorate levels.
- Gender employment gaps vary by education level, with larger absolute gaps in more populous education groups and smaller gaps in short-cycle tertiary education.

- Employment shares describe the composition of the employed workforce and should not be interpreted as employment rates.