Generating New Data Insights on Enterprise Characteristics Through Data Integration

by **Cui Hui Min** and **Jung Hwi Leng**
Business Statistics Division
Singapore Department of Statistics

Introduction

As the National Statistical Office, the Singapore Department of Statistics (DOS) produces a wide range of statistics on the Singapore economy, population and households. DOS has developed integrated databases on enterprises and individuals to compile new indicators and conduct statistical studies across subject domains to meet emerging data demands.

To ensure the integrity and security of these integrated databases and maintain the confidentiality of the entities within, DOS manages the data using a robust **data governance framework**. This framework comprises several components, including legislation, data management policies, process safeguards, and Information Technology (IT) systems [1].

This article shares DOS's experience in producing new data insights on enterprise characteristics through data integration. Additionally, it will highlight two examples of how new data insights are derived from the integrated databases, specifically focusing on high growth firms and women-owned companies in Singapore.

Integrated Databases

Integrated Database for Enterprises

DOS develops and manages the Statistical Business Register (SBR), which serves as the foundational statistical database for the compilation of business and economic statistics.

Enterprises in the SBR are identified by their Unique Entity Numbers (UEN), a unique identifier issued by UEN issuance agencies such as the Accounting and Corporate Regulatory Authority (ACRA) to all Singapore-registered enterprises upon their registrations for their transactions with the Government.
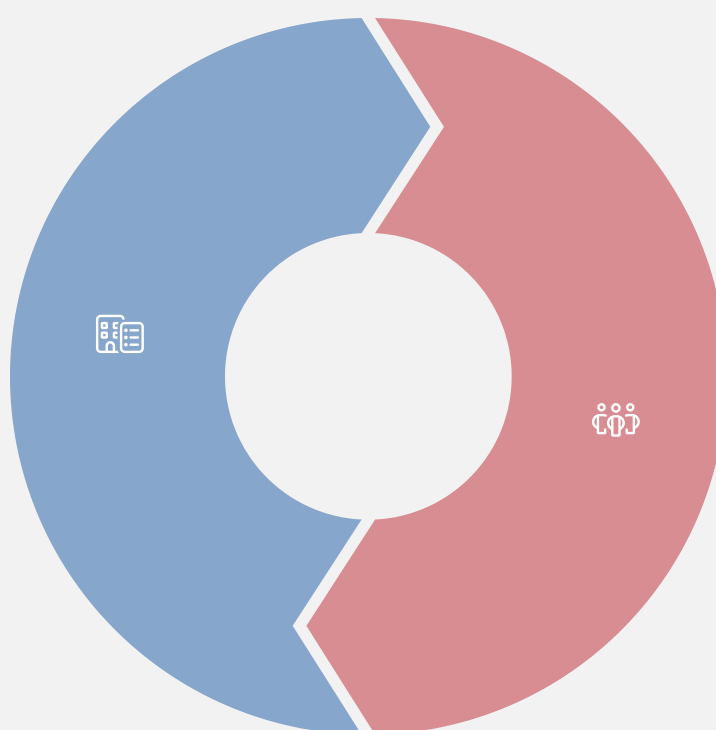
Integrated Database for Individuals

DOS develops and manages the Household Registration Database (HRD), a population statistical database that contains the basic demographic information on Singapore's population. The HRD is used to compile the Singapore population count and profiles, facilitating data integration for in-depth analysis.

Singapore residents in the HRD are identified by their **Unique Identification Numbers** (UIN), a unique identifier issued by the Government to Singapore citizens and permanent residents only.

The availability of UEN and UIN enables DOS to process and integrate both enterprise- and individual-level data across administrative and survey data sources to update the SBR and HRD respectively for the compilation of statistics.

Integrated Longitudinal Databases

With the SBR and HRD as the foundational databases, DOS developed integrated longitudinal databases to support robust, evidence-based policymaking in the Government. The database on enterprises includes data on enterprise characteristics (e.g., economic sector, firm age), firm performance (e.g., revenue, profit), employment & wages, international trade in goods and services, international investment, and business grants. The database on individuals covers demographics (e.g., age and sex), education, employment & income, marriage & family, household, and housing & property (Figure 1).

[1] For detailed information on data integration and data governance in Singapore, please refer to the References appended at the end of the article.

Figure 1: Integrated Longitudinal Database on Enterprises and Individuals

📋 **Enterprise – Individual Data Linking**

Traditionally, data on enterprises and individuals have been treated as separate statistical domains. However, by linking these two types of data on enterprises and individuals, we can uncover new data insights. To support studies that require multi-dimensional analysis cutting across these statistical domains, DOS is able to fuse enterprise (UEN) and individual (UIN) data through a UEN–UIN data link.

Deriving New Data Insights through Data Integration

Integrated Longitudinal Enterprise Data to Profile High Growth Firms

Using the integrated longitudinal database on enterprises, DOS conducted a statistical study to examine the enterprise characteristics and economic contributions of high-growth firms (HGF) between 2016 and 2019. In this study, HGF were defined as firms with average annualised revenue growth exceeding 10% over the three-year period and with at least 10 employees at the start of the period [2].

The study showed that half of the HGF were between 5 and 20 years old (Chart 1). Majority of the HGF were smaller firms with revenue not exceeding S$10 million or with employment size of 10 to 25 workers.

Chart 1: Profile of Firms with High Growth between 2016 and 2019 [3]
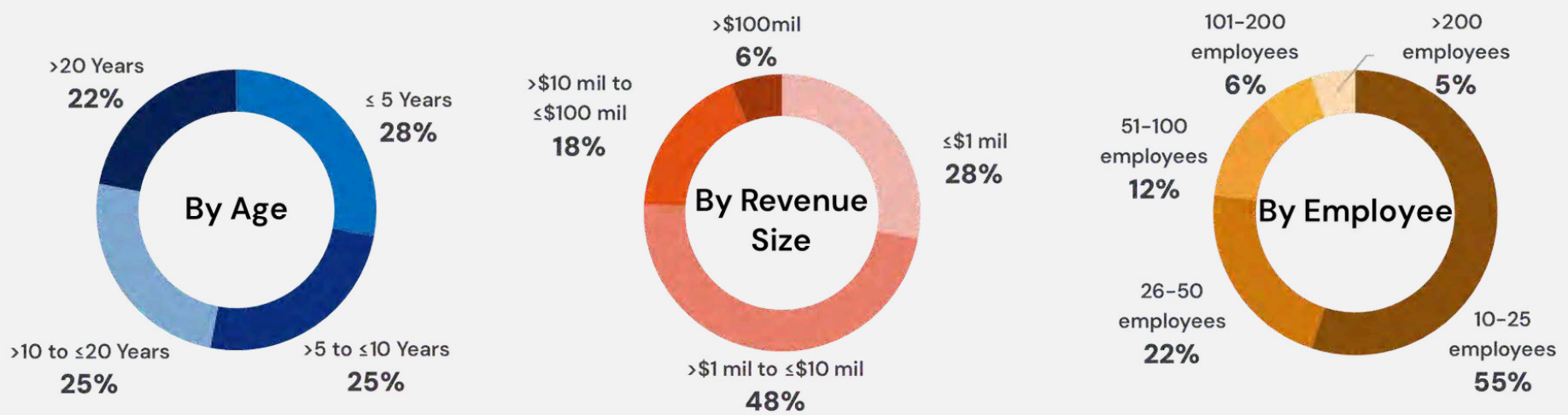
HGF also tended to outperform non-HGF with higher productivity, larger revenue and employment size, and higher average wage per worker (Table 1).

Table 1: Average Revenue, Employees, Productivity, and Wage per Employee of HGF and Non-HGF, 2019

Average Statistics	HGF	Non-HGF
Revenue ($mil)	257	49
Employees (No.)	90	56
Productivity ($'000)	233	105
Annual Wage Per Employee ($'000)	67	52

[2] The threshold of a minimum of 10 employees is meant to mitigate small firm bias, where a small increase in revenue results in disproportionately high revenue growth due to the small denominator. Details of the study are available in the paper '**High Growth Firms in Singapore**'.

[3] The respective categories refer to the revenue size, employment size and age of firms in 2016, which is the start of the revenue growth period between 2016 and 2019.

Integrated Enterprise and Individual Data to Profile Women-Owned Companies

DOS utilised the Enterprise-Individual data linking to study women-owned companies, linking the UEN and UIN across the enterprise and individual databases (Chart 2).

The integrated longitudinal database on enterprises contains information on enterprise characteristics and firm performance, while the SBR stores details about shareholding information of Singapore registered companies. However, both sources do not contain information on the sex of shareholders, which was required to identify women-owned firms.

The integrated longitudinal database on individuals contains demographic information such as sex. By using UIN as the match key to link individual and business data, women-owned companies could be effectively determined.

Chart 2: Integration of Enterprise and Individual Data to Profile Women-Owned Companies
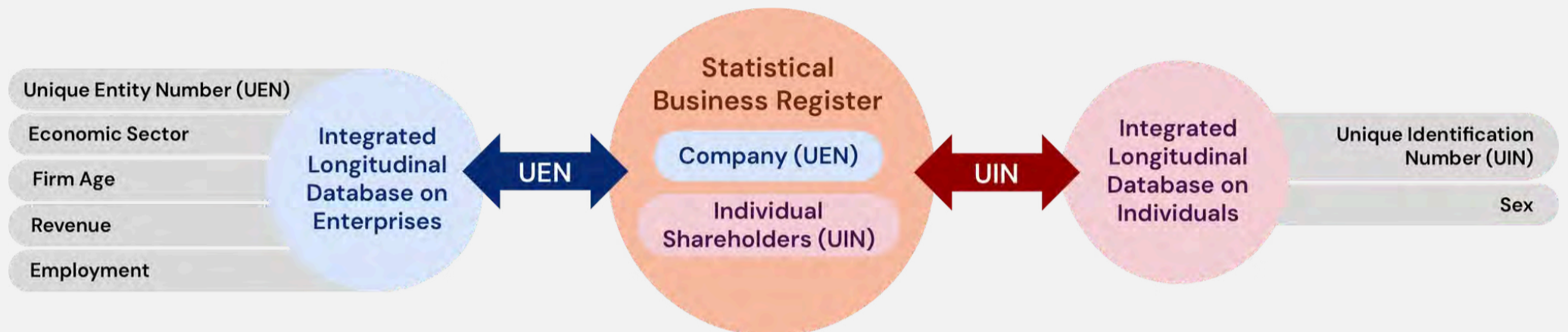
With the fused enterprise-individual data, the study was able to derive the percentage shareholding by male and female shareholders and classify Singapore resident-owned companies [4] into:

Women-Owned Companies

Companies with more than 50% of their ordinary shares held by female Singapore residents.

Men-Owned Companies

Companies with more than 50% of their ordinary shares held by male Singapore residents.

Other Companies

Companies where male and female residents do not separately own more than 50% of their ordinary shares (e.g., male residents own 30% and female residents own 25%, with the remaining owned by company and/ or foreigners).

The study showed that women-owned companies have grown in number and significantly contributed to the Singapore economy. In 2022, women-owned companies generated S$85.3 billion in revenue and hired 202,000 employees, accounting for 21.3% and 18.5% of the total revenue and employment of Singapore resident-owned companies respectively (Chart 3).

The findings provide valuable insights for shaping inclusive growth policies and fostering entrepreneurial talents in Singapore to seize new growth opportunities.

Chart 3: Resident-owned Companies' Revenue and Employment by Type of Ownership, 2022
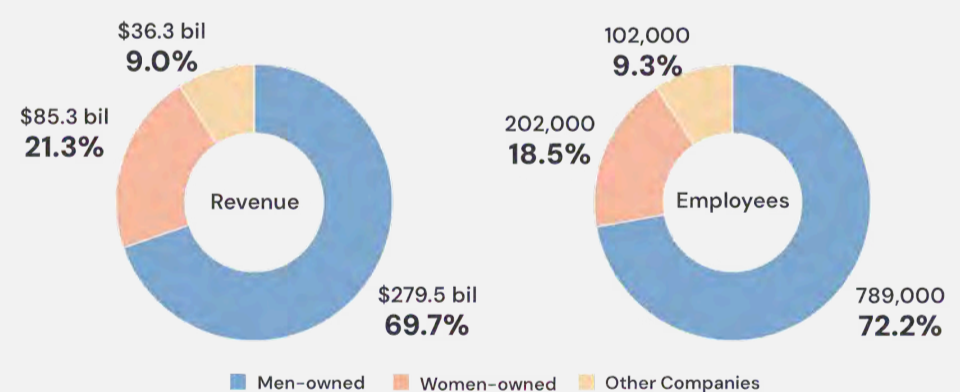
Conclusion

The studies on high-growth firms and women-owned companies have reaped the benefits of data integration across various statistical domains in generating new data insights on emerging topics. By integrating data, DOS can now meet new data demands without increasing the reporting burden of firms. DOS will continue to leverage integrated data to produce new statistics and insights that support evidence-based policymaking.

[4] In the study, resident-owned companies are defined as companies with revenue or employment, and with more than 50% of their ordinary shares held by Singapore residents. For more details of the study, please refer to the References appended at the end of the article.

▼ **References**

- Eurostat (2019). **Micro Data Linking**.
- Luppes M., & Nielsen P. B. (2020). **Micro Data Linking: Addressing New Emerging Topics Without Increasing Reporting Burden**, Statistical Journal of the IAOS Vol 36 Issue 3
- Singapore Department of Statistics (2023). **High Growth Firms in Singapore**
- Singapore Department of Statistics (2024). **Women-owned Companies in Singapore**
- Singapore Department of Statistics (2024). **Integrated Longitudinal Database for Supporting Policy Studies**
- Singapore Department of Statistics (2024). **Data Governance and Data Integration in Singapore**