H&S Startup Insights

Startup Landscape Across Countries and Industries: Exploratory Data Analysis with Jupyter Notebook

Prepared by: Hafiz Sham

Introduction:

1. Business Context
2. Sources of Data
3. Dataset Overview
4. Dataset Variable Definitions

Business Context

- **Startups** are critical drivers of innovation and economic growth.
  - Success and valuation are often influenced by external factors beyond the business model or product.
  - In today's competitive funding landscape, understanding these broader influences is essential for investors, policy makers, and startup founders.

- **The relationship** between macroeconomic indicators and startup performance across multiple countries is analyzed.
  - By combining company-level startup data with country-level economic metrics, we aim to uncover which macroeconomic indicators correlate most with high startup valuations and funding activity.

- **Our goal:**
  - Provide stakeholders with data-driven insights into how macroeconomic factors, may affect the potential and scalability of startups in different regions.

- **The findings** are intended to support more informed decision-making in areas such as:
  - Market expansion,
  - Investment prioritization,
  - Government policy support.

Sources of Data

1. **Kaggle (Link)**
   - `startups_data.csv`

2. **The World Bank**
   - `countries_macroeconomics_data.csv` (with some precleaning)
   - Latest as of 2023

3. **Visual Capitalist (Link)**
   - For imputing rows with missing values such as R&D

4. **StartupBlink (Link)**
   - Inform rationale for selecting chosen countries

Dataset Overview

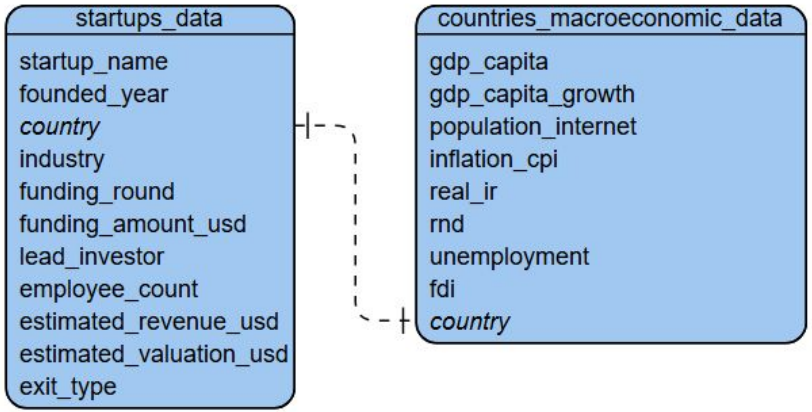
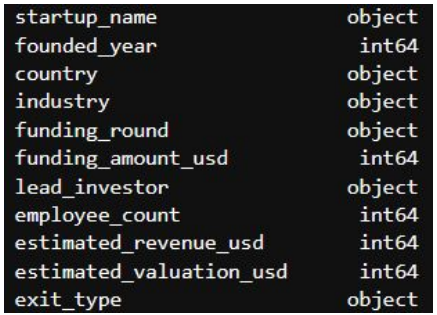
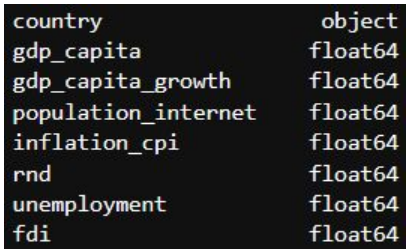


Figure 1: *Entity-Relationship Diagram*





Figures 2a and 2b: `dtypes` *of variables*

- The datasets are merged on the `country` column

Dataset Variable Definitions

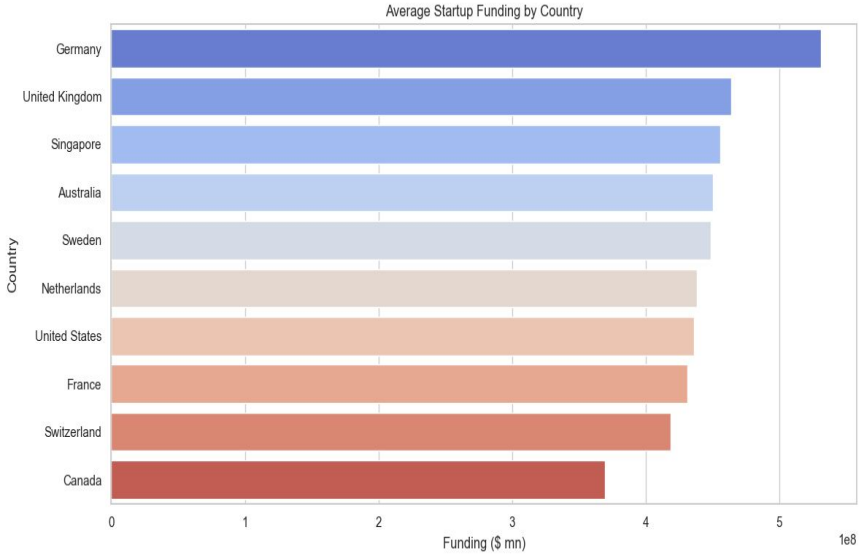
Variable	Definition
startup_name	Official name of the company
founded_year	Year the company was founded
country	Country where the company began operating
industry	Industry which the company operates in
funding_round	Type of funding the company received
funding_amount_usd	Initial funding received ($)
lead_investor	Largest investor in the company
employee_count	Number of employees in the company
estimated_revenue_usd	Estimated revenue in 2023 ($)
estimated_valuation_usd	Estimated valuation in 2023 ($)

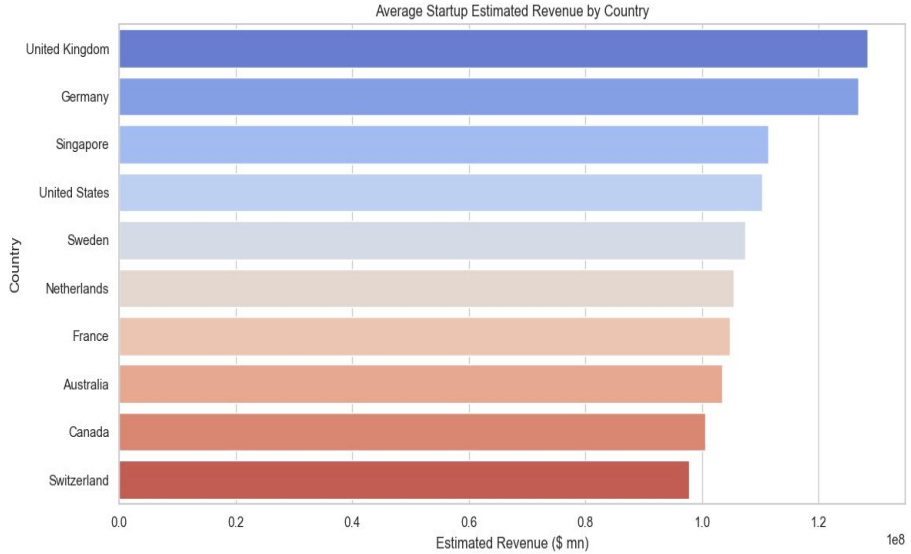
Variable	Definition
estimated_valuation_usd	Estimated valuation in 2023 ($)
exit_type	Company underwent IPO or was acquired
gdp_capita	GDP per capita of the country
gdp_capita_growth	GDP per capita annual growth (%)
population_internet	Proportion of population (%) who have access to the internet
inflation_cpi	Inflation rate (%), measured by changes in Consumer Price Index
rnd	Amount of R&D in technology (as a proportion to GDP, %)
unemployment	National unemployment rate (%)
fdi	Foreign direct investment

Charts and Findings:

1. Bar chart
2. Scatter/bubble plot
3. Line chart
4. Stacked column chart
5. Correlation heatmap

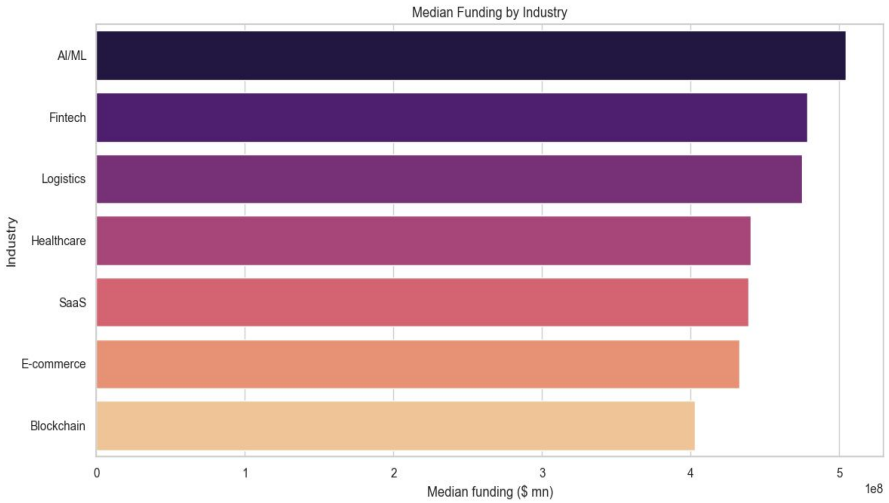
Global trends of funding and revenue

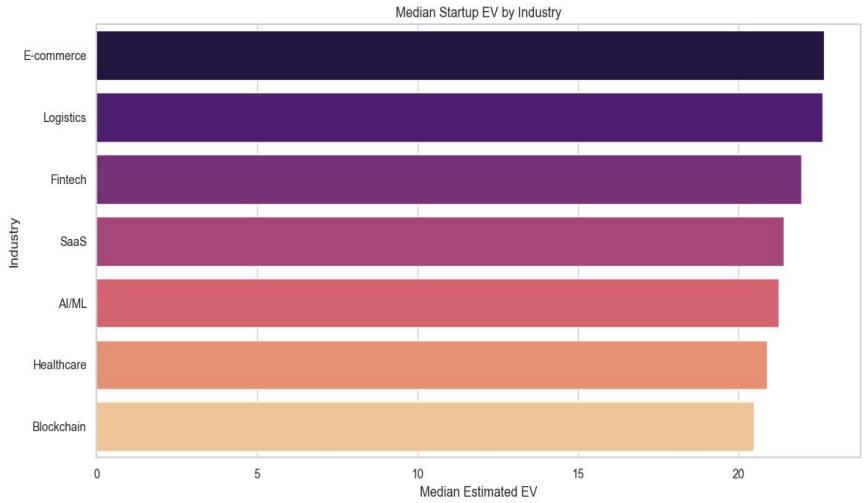




- **Germany, the UK and Singapore** have the highest average startup funding and estimated revenue
- Startups in **France, Switzerland, Canada and Australia** have the lowest average funding and estimated revenue

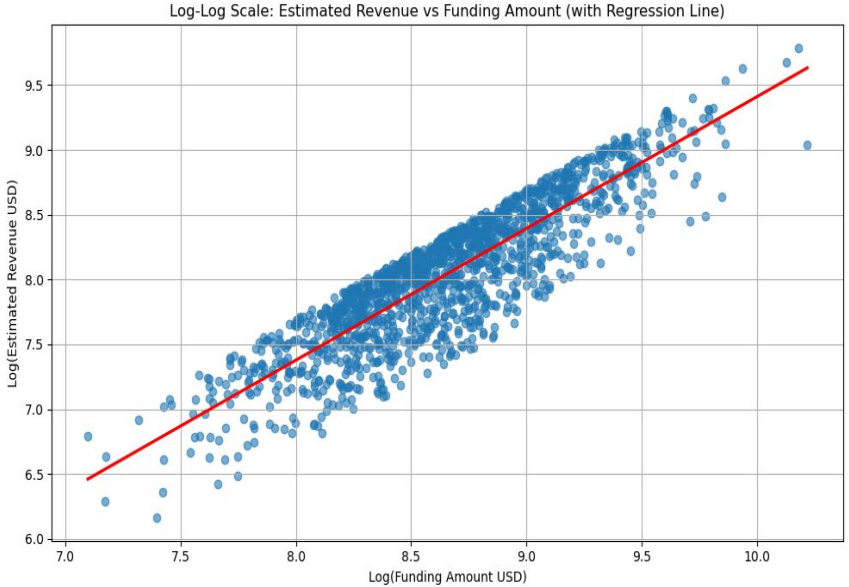
Strong sentiment around AI/ML, fintech, e-commerce and logistics

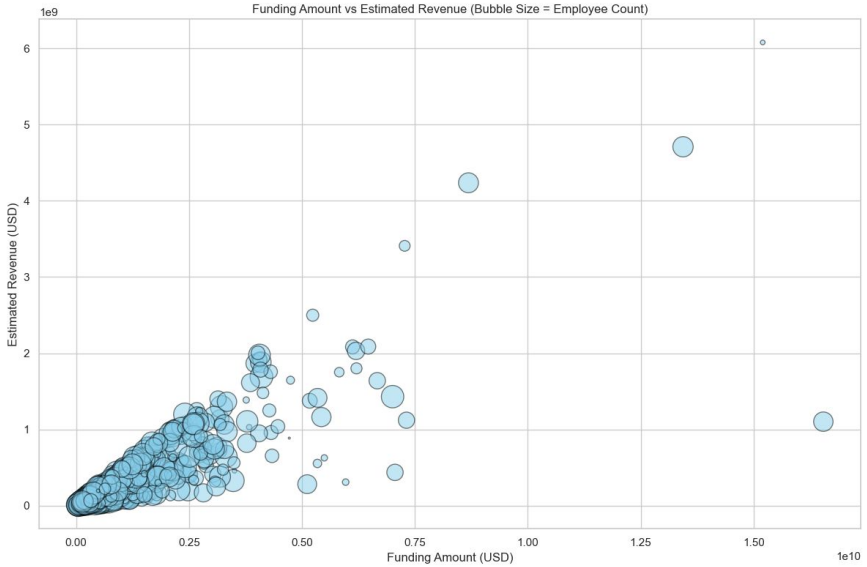




- **AI/ML, fintech, and logistics** are the top funded industries, indicating investor confidence towards these industries
- **E-commerce, logistics and fintech** have the highest EV multiples, indicating high expected future growth

Positive correlation between funding and estimated revenue (with a few outliers)

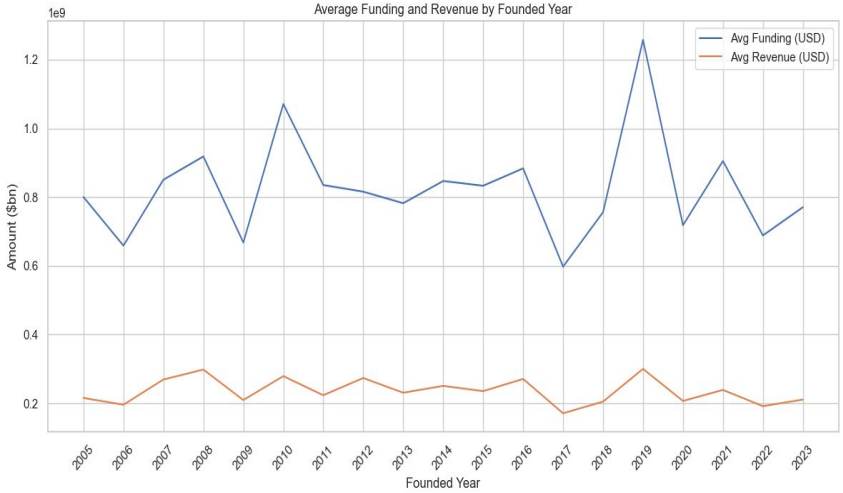


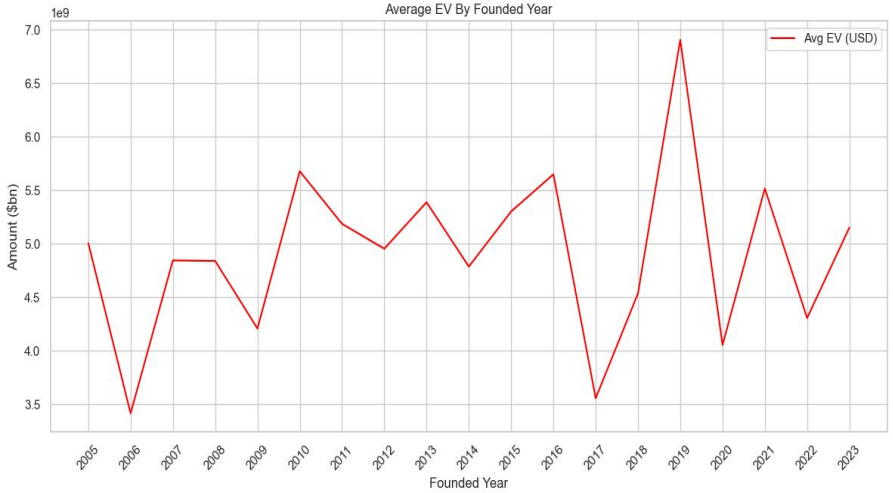


- There is a **linear relationship** between funding amount and estimated revenue

- The funding of most companies is concentrated around $2.5bn and below
- Most companies have large numbers of employees although some relatively lean outliers still have high valuations

Startup funding and revenue peaked in 2019

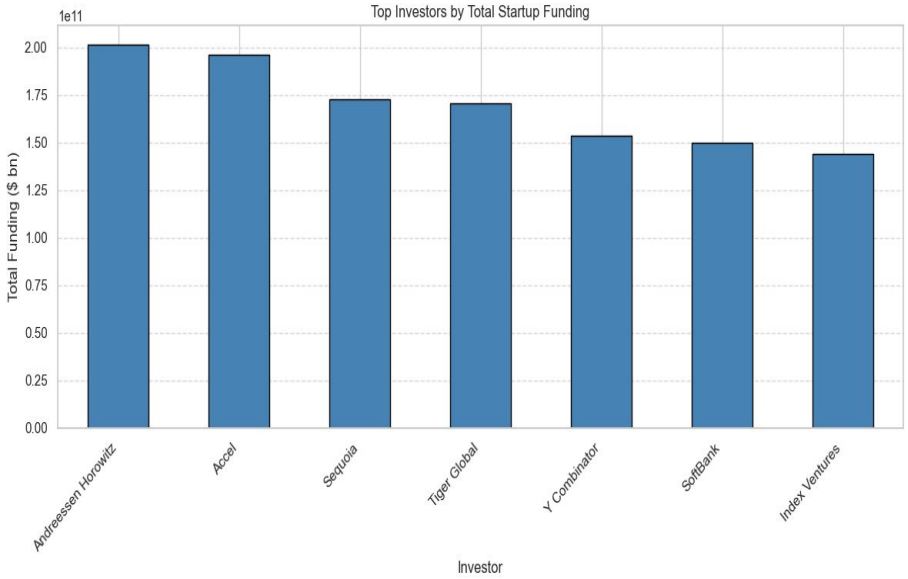


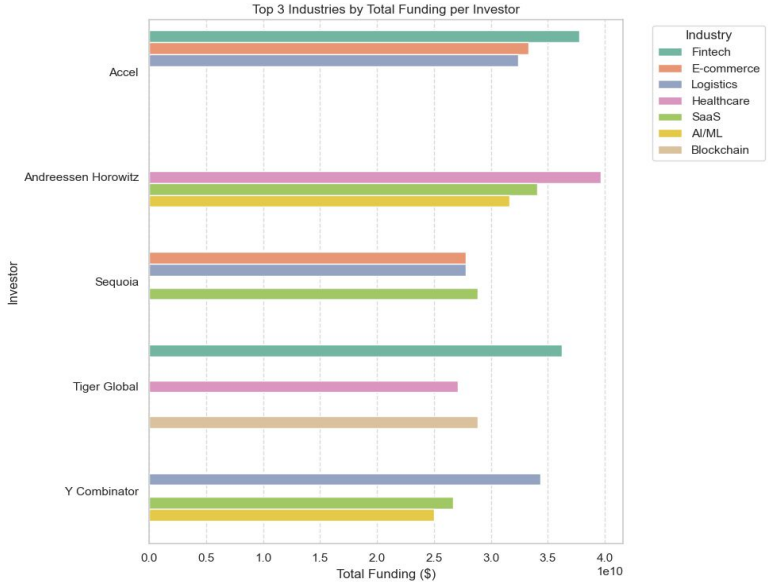


- Funding and revenue seem to increase slowly for companies established after the **mid-2010s**
- Companies founded in 2019 receive the highest funding and revenue, before dropping sharply the following year

- A similar trend is seen for EV as compared to the graph on the left

Investments strategy vary among investors

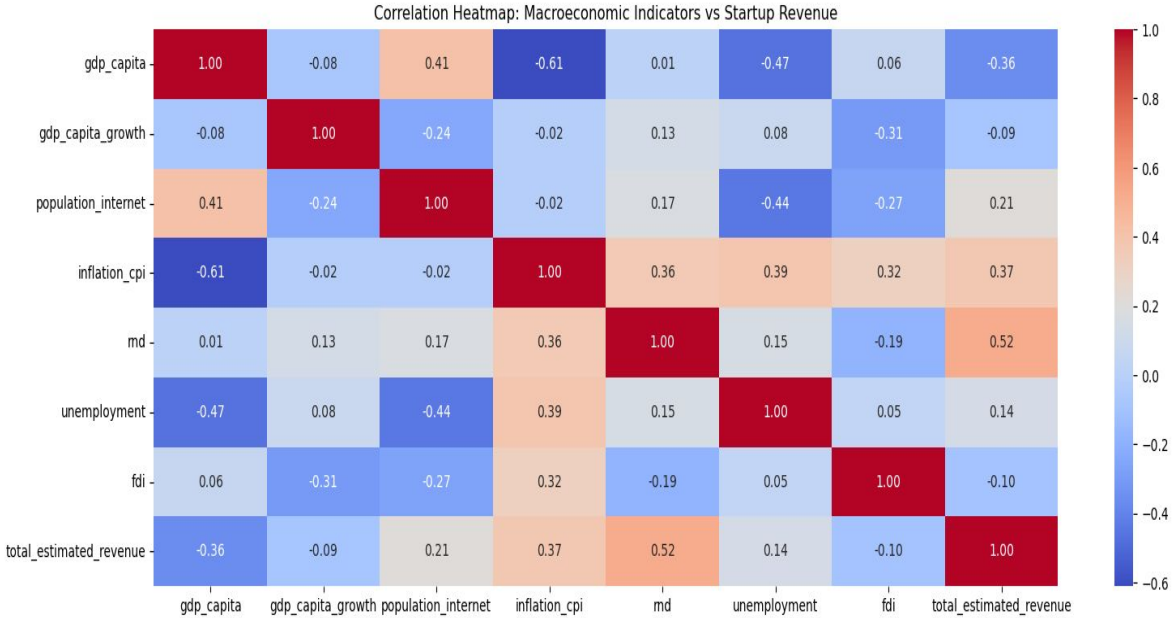




- **Andreessen Horowitz, Accel** and **Sequoia** invest the most in startup funding

- Each company's top three most invested industries are different, indicating varying investment strategies
- For example, Accel focuses on fintech, Andreessen Horowitz focuses on healthcare, and Sequoia focuses on SaaS.

Relationship between macroeconomic indicators with startup industry performance



Positive Correlation

- Mortality Rate (md) = 0.52, Countries with higher mortality rates have significantly higher startup revenue
- Inflation Rate = 0.37, Higher inflation environments generate more revenue

Negative Correlation

- GDP per capita = -0.36, Wealthier countries show lower total startup revenue. This suggests startups in emerging markets may be capturing larger market share or scaling faster

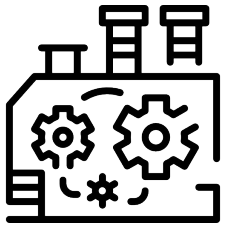
Main Takeaways:

1. Conclusion
2. Key Recommendations

Conclusion

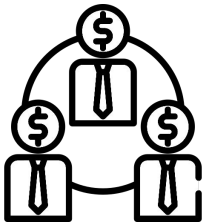
Macroeconomic conditions:

- Startup valuations vary significantly across industries and geographies, but are often higher in countries with **stronger macroeconomic fundamentals** (higher R&D spending, GDP, and FDI)
- Countries with more robust macroeconomic conditions tend to foster better-funded startup ecosystems, even if the number of startups is smaller.



Type of Industry:

- Solid sectors such as **AI/ML, Fintech and Logistics** consistently attract higher average funding across regions, indicating that these are established industries.
- Founding year shows that **younger startups** are still receiving high valuations, indicating a sustained appetite for innovation in recent years.



Venture Capital:

- Venture capitalists ascribe high enterprise value towards **AI/ML, E-commerce and Fintech** despite lower revenues, indicating that they forecasted increased growth in the future.
- **Blockchain and Healthcare** have the lowest enterprise values, indicating that they are either face higher regulatory or adoption barriers, or currently attract less investor confidence compared to more mature or commercially proven sectors.

Key Recommendations

1. **Investors**:
   a. Focus on industries and countries with favorable macroeconomic indicators, particularly R&D
   b. Understand investment strategies of varying investors and use them as a guideline for investing in certain industries

2. **Policymakers**:
   a. Enhance national innovation ecosystems by investing in research and development and maintaining economic stability to attract startup activity

3. **Startup entrepreneurs**:
   a. Start in high valuation sectors (AI/ML, e-commerce and fintech) and countries (such as Germany, UK and Singapore) to increase your chances of receiving significant funding and higher valuations
   b. Tailor company strategy to attract the largest investors (e.g. Andreessen Horowitz, Accel, Sequoia**)**

4. **Future Analysis**:
   a. Improve investor granularity
   b. Integrate additional startup growth metrics such as year-on-year Profit & Loss (P&L), Return On Investment (ROI), and Compounded Growth Annual Rate (CAGR)
   c. Incorporate more longitudinal data to capture business lifecycle effects (e.g Launch, Growth, Maturity, Decline)

End of presentation