

Analysis of companies in the UK

Data analysis exercise

Diogo Cunha

A section of the data was chosen to be analyzed for the exercise and loaded into a PostgreSQL database.

Source

- [BasicCompanyData-2024-07-01-part1_7.zip \(69Mb\)](#)
- [BasicCompanyData-2024-07-01-part2_7.zip \(70Mb\)](#)
- [BasicCompanyData-2024-07-01-part3_7.zip \(69Mb\)](#)
- [BasicCompanyData-2024-07-01-part4_7.zip \(70Mb\)](#)
- [BasicCompanyData-2024-07-01-part5_7.zip \(70Mb\)](#)
- [BasicCompanyData-2024-07-01-part6_7.zip \(70Mb\)](#)
- [BasicCompanyData-2024-07-01-part7_7.zip \(42Mb\)](#)



Database



- 850k companies created in the UK
- 55 columns (high dimensionality!)

Some columns suggested more potential for insights, resulting in a final dataset with 16 columns that can be used to produce useful dashboards and statistics.

Column name	Potential Uses
company_name	Link specific companies by name
company_number	Unique identifier
county	Locate geographically and identify possible trends with that area.
country	Geography (...)
country_of_origin	Geography (...)
incorporation_date	Essential time component for analysis.
company_category	Type of entity such as “Private Limited Company” can be used to differentiate companies by size or type of management.
sic_codes (x4 columns)	Industry/area of activity of companies are very relevant to understand trends like new markets rising or falling, profitability, market barriers etc...
company_status	Company current state (active, liquidation) can give good insights on industries that are behaving well or poorly.
mortgages (x4 columns)	Can reflect need to leverage debt, which is more common on certain industries. Can also imply financial trouble (e.g. failure to pay mortgages).

Taking a look at data quality for the selected columns, only 2 (non-crucial) columns of the 16 columns lack information.

Column name	Rows missing values	% Total missing
county	594,595	69.95%
country	170,990	20.11%

- ~80% rows have country specified which is alright, considering the column 'country' allows for more global insights than county, even if county only has ~30% of non-missing rows.

123 sic_1_null_or_empty ▼	123 sic_2_null_or_empty ▼	123 sic_3_null_or_empty ▼	123 sic_4_null_or_empty ▼
0	703,847	787,599	821,338

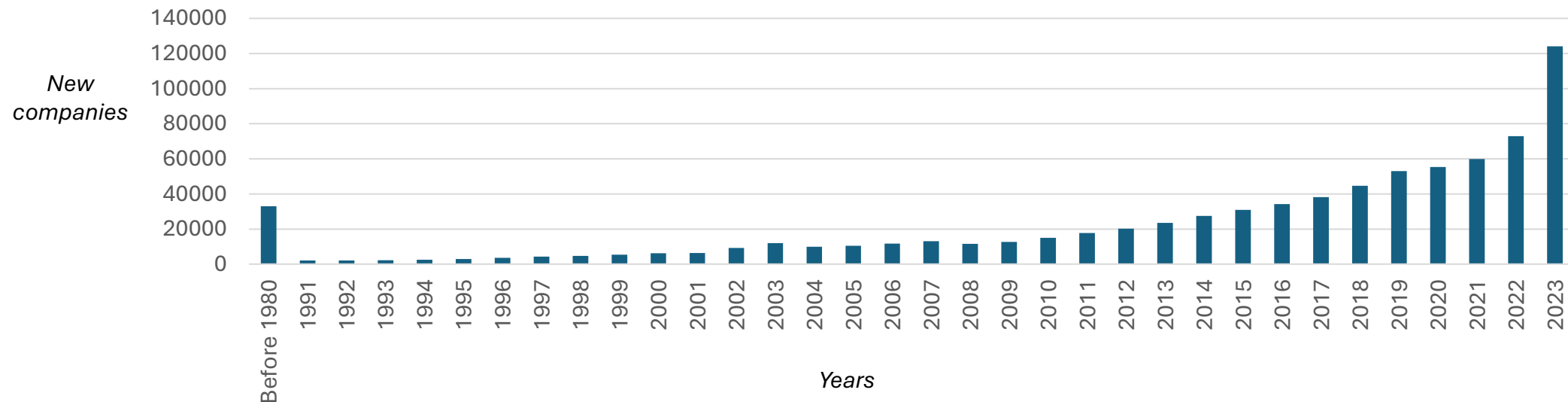
- SIC represents the industry or industries of the companies so it is normal that many companies don't have more than one SIC code specified! It is not as if information is missing. But if it were we also wouldn't be able to tell...
- In fact we can quickly calculate that only ~3.3% of companies have four different areas of activity specified and that ~82.8% of companies have only 1 area of activity!

Let's gather some super quick insights from the data!

- The 'Ashford Cattle Market Company' is the oldest company registered in the dataset. Its origin dates back to 1856! 3 other companies also date back to 1856.

ABC company_name	ABC sic_1
ASHFORD CATTLE MARKET COMPANY LIMITED(THE)	46110 - Agents selling agricultural raw materials, livestock, textile raw materials and semi-finished goods
B.L.HOLDINGS LIMITED	41100 - Development of building projects
BRITISH INDIA STEAM NAVIGATION COMPANY LIMITED	50200 - Sea and coastal freight water transport
00000258 LIMITED	7499 - Non-trading company

- As expected the rate of new companies appearing every year increased over time. The number of companies registered in 2023 is **1.8x** the number in 2022 and **3.76x** the number of companies registered before 1980!



Let's gather some super quick insights from the data!

- That big jump in companies registered between 2022 and 2023 is mostly due to companies dedicated to “Retail sale via mail order houses or via Internet”, which account for **8.91%** of companies from 2023.
- However, between the years of 2017 and 2022, the main sector responsible for more companies registered was related to **“Operating of own or leased real estate”**.
- In terms of geography, around **50%** of all companies are registered in England, as it would be expected. In terms of counties, the distribution is not so skewed, with Essex, Kent and Surrey occupying the top 3 counties with a clear tendency for more companies to be registered around London nearby counties.



Let's gather some super quick insights from the data!

- In terms of **company health**, at the time of this snapshot, **91.94%** of the companies are active and only **1.90%** went through liquidation process.
- We can also use mortgages as a proxy for estimating company/sector health, although they can also be part of more complex debt management strategies (e.g. real estate). **13.08%** of the companies have taken at least 1 mortgage and only **2.37%** we were able to fully pay their debt.

¹²³ have_taken_mortgages ▼	¹²³ avoided_mortgages ▼	¹²³ paid_mortgages ▼	¹²³ still_paying_mortgages ▼
111,193	738,806	20,168	91,025

Let's gather some super quick insights from the data!

ABC sic	123 total_mortgages	123 company_count	123 avg_mortgages	123 percent_companies_with_mortgages
65202 - Non-life reinsurance	6,140	149	41.21	59.73
65201 - Life reinsurance	1,842	58	31.76	48.28
65120 - Non-life insurance	9,397	679	13.84	44.33
7012 - Buying & sell own real estate	267	26	10.27	96.15
7020 - Letting of own property	489	49	9.98	87.76

- It is also clear and expectable that mortgages are related to companies connected to the real estate sector.
- We can check that **insurance related companies tend to have more mortgages on average**. This makes sense since they tend to use real estate as part of their portfolio diversification strategy as these can generate stable, long-term returns that align well with the long-term nature of their liabilities (e.g., life insurance payouts).
- However, contrarily to companies that operate directly in real estate, insurance related companies do not always have mortgages going on. While real-estate companies have **87%-96%** with at least 1 mortgage, the insurance related companies value falls into **44%-60%** range.
- So, as a conclusion, opting for mortgages is a very common strategy for real estate companies, despite, on average, going for less mortgages than in the insurance sector. However, the companies in the insurance sector that do opt for mortgages tend to go all-in in that direction!

Thanks!