Aims

This project aims to determine which characteristics will most likely allow a firm to succeed.

I will initially analyse opinions from firms on what their largest obstacles to success are, this is split into small, medium, and large firms. Geographical location of high growth firms and different industry growth will then be considered. I will then focus on emerging economies to determine whether the size of a firm and the innovation index of a country are related to high growth firms.

All charts are interactive with tooltips encoded for more detailed information.

Data

I used a variety of data sources to collect my data- these sources can be viewed by clicking the ‘Data Source’ button under each chart. My methods to clean them can be viewed by clicking the ‘colab’ button

For the first chart I built a scraper using python in google collab to retrieve a table from a pdf produced by the World Bank which gave the opinions of firms of different sizes on what their major obstacles for success were.

I used data downloaded from the ONS as a csv on the number of high growth firms per region and then cross referenced this with a Geojson file to produce a the heatmap.

Initially I used an API from Trading Economics to download the growth of 6 different industries as this would ensure data was kept up to date. However, this site only offered a 7 day free trial so instead I had to download each file separately. I used python to merge the datasets as doing this manually could have caused errors.

I retrieved data from a book produced by the World Bank using a scraper in colab to plot how the size of firm in emerging economies is related to their growth potential. This data source was also used alongside an innovation index downloaded as a csv from theglobaleconomy.com to produce a regression chart. I downloaded a csv from growjo on funding and employment growth to produce my last chart.

Challenges

Ideally for my map heatmap I would have used regions such as ‘North East’ instead of Local Authority Districts, however there were no publicly available Geojson files that were of small enough size. To show the data better, I worked out the number of high growth firms as a percentage of all firms in that area as just using the count of high growth firms did not show the distribution accurately as Local Authority Districts are different sizes.

I found problems when embedding chart 4 into my website as the compiled vega restacked the groups on top of each other despite it reverting back to the intended chart when you click ‘Open in Vega Editor’. Whilst this is not ideal, it still shows how high growth firms tend to be larger. I have added a tooltip so you can see the value and firm type of each bar to be extra clear.

Many attempts were made to scrape data from tables or pictures within PDFs using python, however, often parts of tables were heavily cut off and therefore could not be used. I was able to find other data sources instead that I was able to use by using the tabula package in python.

Conclusion

There are a variety of factors affecting a firm’s success. It seems that companies deem financing, inflation and political instability to be their largest obstacles to growth, however, these factors are less of an obstacle to larger firms. We see in chart 4 that in emerging economies, companies that are larger (determined by average number of workers) are more likely to be high growth firms.

Location also creates an environment for high growth firms. Chart 2 shows that high growth firms appear in clusters, such as very central London, South Wales, and Edinburgh. This could be explained by their being more innovation or competition in these areas which would stimulate growth. More innovation leads to a higher share of high growth firms as shown in chart 5 with an R-squared value of 0.34.

The type of industry a business is in, is likely to affect its success. Manufacturing, services, and public administration prove to be growing industries, whilst mining is significantly declining, and agriculture is particularly volatile as shown in chart 3. Therefore, the sector a business is in will affect its probability of success.

Intuitively, businesses with more funding are more likely to have higher employment growth which is shown in chart 6 with an R-squared value of 0.35, suggesting a significant relationship.

Location, industry type, innovation, funding and the size of a firm are all determinants in whether a company will be successful. Future work could test other characteristics such as number of women on the board, level of education of managers or comparisons between the public and private sector.