

EU Startup Monitor 2018

Methodological note

Table of contents

- 1. Introduction..... 1
- 2. Technical background..... 1
- 3. Coding, programming and data cleansing..... 2

1. Introduction

This document is intended to give an overview of the data sources and methodologies used to compile the dataset on start-ups and scale-ups for the EU Startup Monitor and parts of the Annual Report on European SMEs 2017/2018 and of the 2018 SBA Fact Sheets, commissioned by the European Commission (DG GROW). The key objective of the Work Package is to add to the existing data on start-ups and scale-ups and deliver real time information on the state of startup ecosystems in Europe. Geographically the data covers 18 of the EU28 members states as well as Switzerland, Turkey and Israel. The data has been collected with a 7-category online questionnaire (demographics, internationalisation, innovation, leadership, markets, start-up and strategy) between February and May 2018. All inputs were voluntarily provided by founders and c-level (top-tier) employees of start-ups and scale-ups. The communication of the data collection and call for participation has been supported by a large number of European stakeholders ranging from startup associations to Incubators, Accelerators and other important ecosystem representatives. The individually presented sections have been chosen in accordance with the SME Performance Review team.

2. Technical background

The online survey has been delivered using the latest PHP version 7.1 and MySQL database. The tool used to collect the data was Limesurvey which is secure and highly customisable with PHP and JavaScript code, allows responsive frontend design and offers live previews to constantly monitor the quality of data.

3. Coding, programming and data cleansing

The questionnaire was programmed in English, based on the final questionnaire that had been scientifically designed by the ESCP researchers Prof. Dr. René Mauer and MSc. Lisa Steigertahl. Each category draws on existing and academically acknowledged scales from international journals of excellence. Due to the lack of historical data on start-ups and scale-ups and without a common European definition on the terms the researchers also matched their work with existing studies on innovation and start-ups and available literature to ensure to be in line with the work of practitioners as well as academic scholars. 140 unique tokens to allow clustering and monitor the sources of data were given out, all inputs were anonymous.

The following question types were used:

- (Uneven) 5-Point Likert scales with additional “Prefer not to say” option
- Single or multiple input boxes
- Single- and multi-answer boxes

Additionally, some questions needed individual programming in PHP and JavaScript / JQuery for the specific purposes and user friendliness, offering previously given inputs as answer options for upcoming questions.

The data was cleaned and presented as is and without any estimations or other secondary inputs to ensure consistency among results three cleansing criteria were established: Considered companies had to be younger than ten years, have an innovative product/ service or business model and show clear intent to grow (in number of employees and/or markets operated in). Besides content related cleansing, all inputs that were made in under five minutes, using a clear answering scheme or providing conflicting inputs had been removed. Since the total population of start-ups or scale-ups in Europe is unknown, the researchers could not estimate the representativeness of the data, but found to be in line with other relevant studies in the field.

4. Limitations

There are limitations to the study which need to be considered when drawing conclusions from the findings. First, there are no reliable data about the overall population of start-ups in Europe.

Hence it is unclear what relative sample size our sample emerged from. Second, in terms of representativeness there is variation in the response behaviour among countries, which often is dependent on the level of sophistication within each country's startup ecosystem. Third, there is a clear bias towards start-ups with digital business models. Although we can assume that the percentage of digital startups is a lot higher these days than for non-digital startups, we do not have information about the actual spread, which may have implications for the representativeness of the sample. Finally, the dataset is cross-sectional and hence only captures a momentary situation in 2018 not allowing for conclusions about dynamics. Most respondents did not fill out the survey in their mother tongue as the survey language was English.

*For the SME Performance Review, specific country and industry expertise by Author Lisa Steigertahl (former Head of Research German Startups Association and CEO European Startup Network) was applied to provide a layer of qualitative information with the quantitative results and to ensure the data has been presented in a well reflected manner. More information on the researchers and partners of the EU Startup Monitor can be found online: www.startupmonitor.eu.