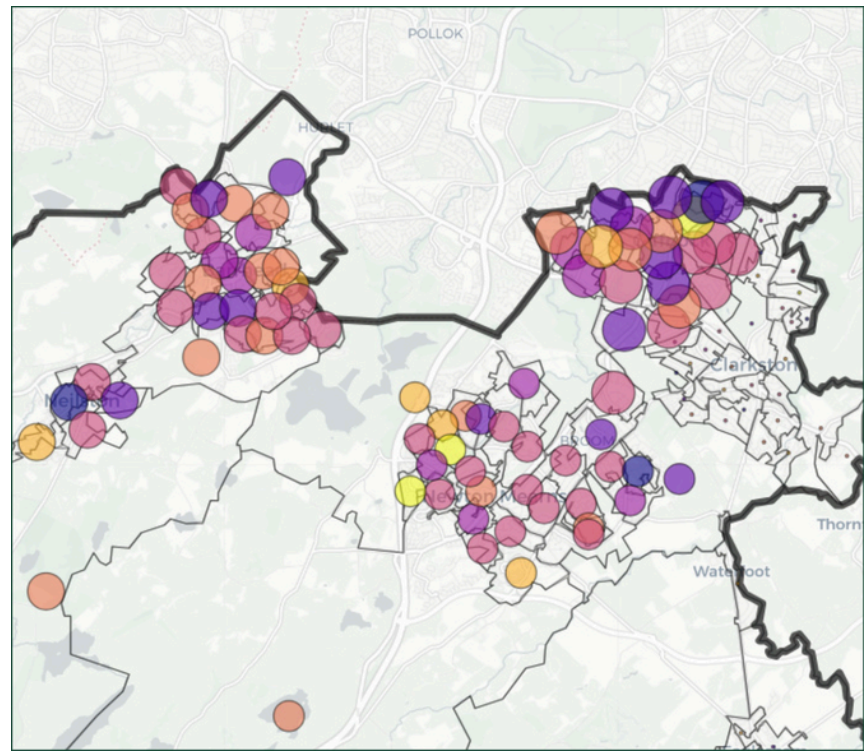


Projects that have utilised this data to date have often taken advantage of its ability to see the effects of macroeconomic shocks in real time, such as the effects of interest rate rises on people’s financial resilience. The nature of the data allows us to go as far as testing the efficacy of interventions, such as seeing if habits changed in areas of emerging distress when given further financial education support by a local authority.



Smart Data Foundry has undertaken a lengthy data validation process, comparing our data to census, Indices of Multiple Deprivation, and other administrative data to judge its representativeness over a variety of characteristics. We found that the data was broadly representative across GB. Additionally, the large sample size aids in the creation of statistics as it accounts for variance.

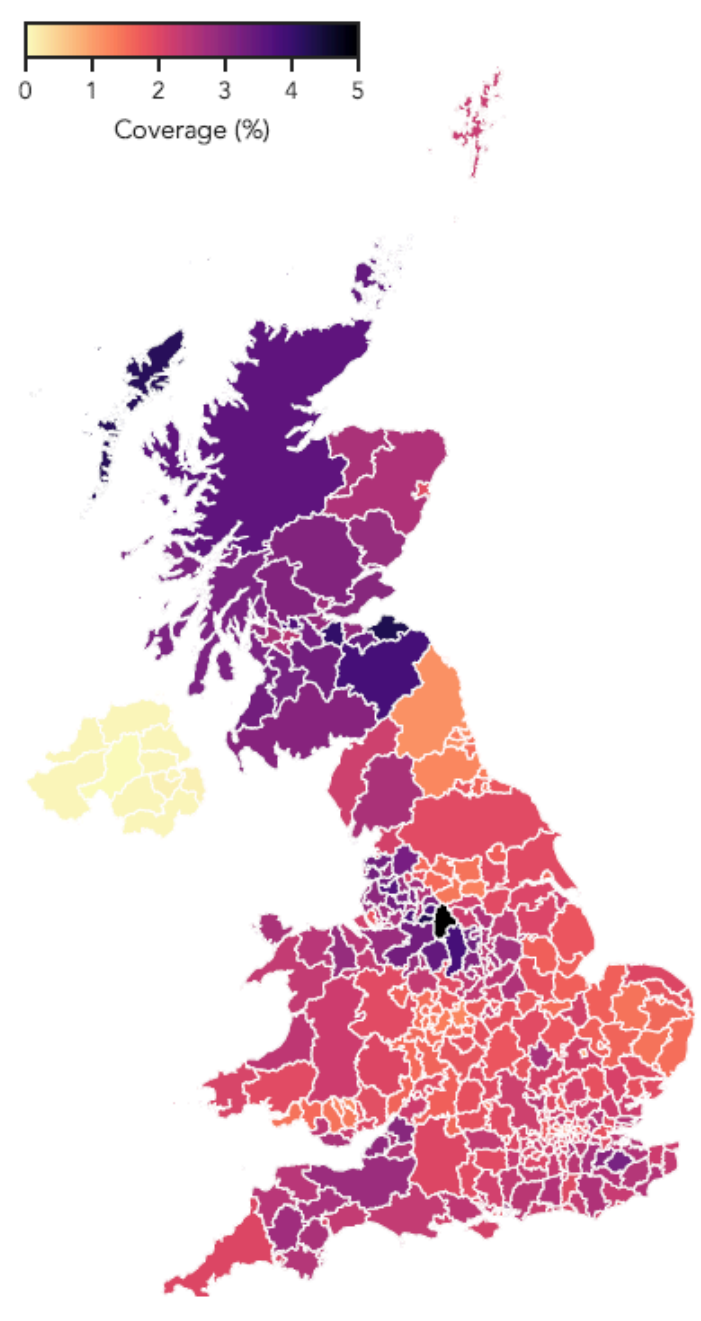
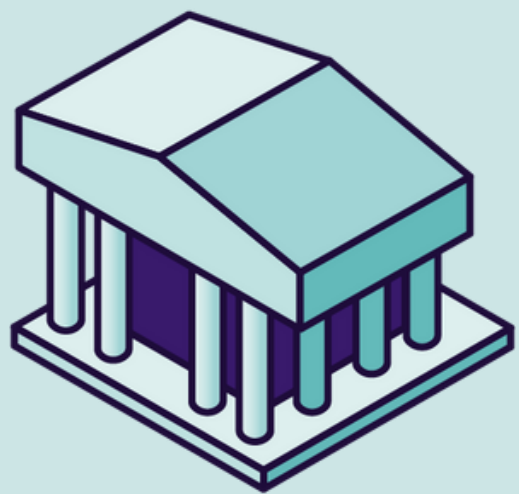


Figure 2: Coverage of the sample by local authority district across the UK (mid-2022).

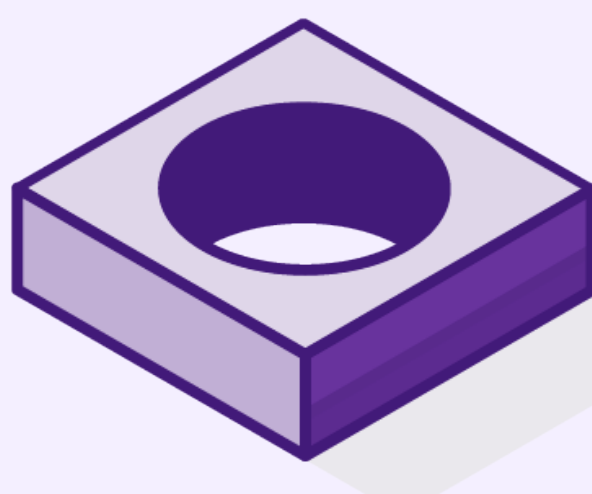
Strengths

People vote with their wallets, allowing data to be much more unfiltered than from a survey. Financial transaction data can therefore be considered rich behavioural data across a range of demographics. The data arrives weekly with persistent identifiers, so is near-real time and usable in longitudinal studies. Once can simultaneously see income and expenditure, and the relationships between the two.



Weaknesses

As with all real data, it possesses certain oddities and anomalous datapoints, especially given the varying ways people manage their finances. The data covers individuals, not households, to ensure the privacy of the data subjects. Data categorisation is a commercially sensitive subject, so it is not always the most transparent. While geographic and basic demographic representation is good, it still comes from only one data provider..



Banking on Change
A practitioner’s SWOT analysis
of banking transaction data for
social good

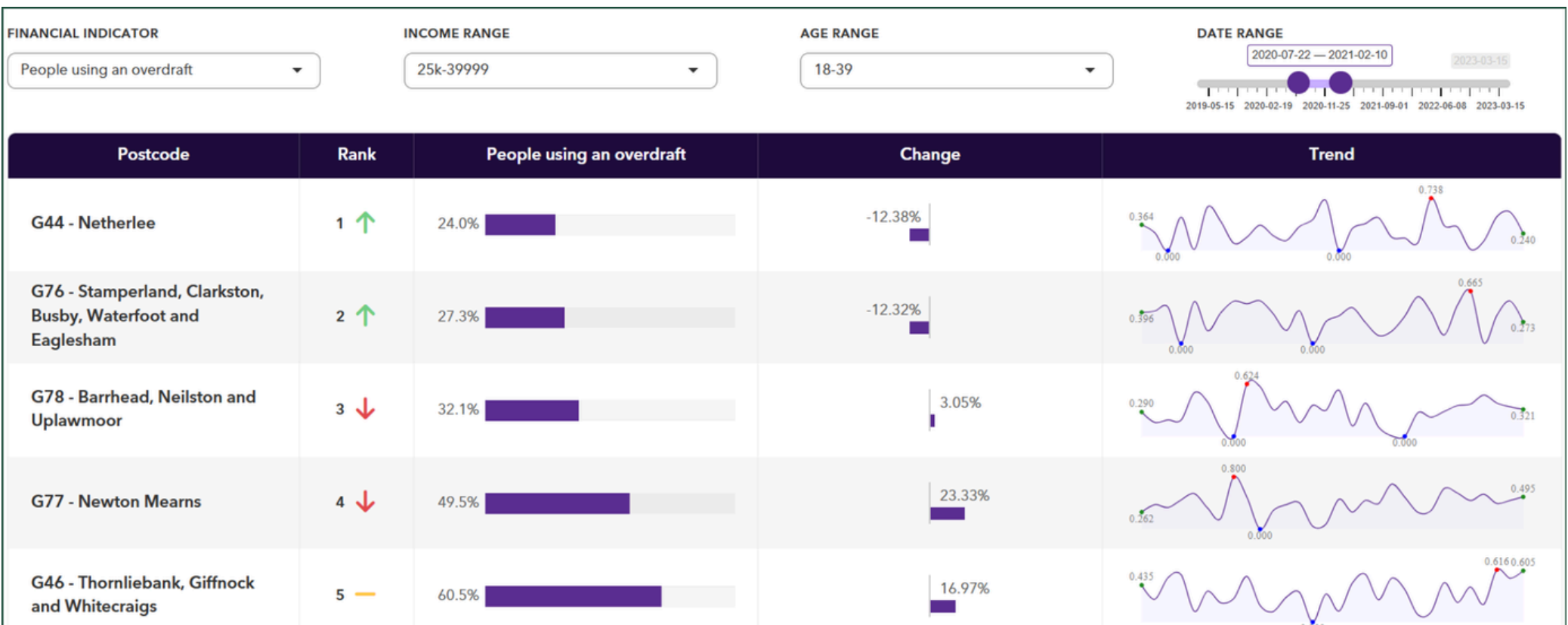
Utilising de-identified consumer transaction data for 5.3m individuals aggregated over ~60 income and expenditure categories on a weekly basis.



Valuable data can provide new insight into research areas, adding a strong quantitative basis which can direct or support further qualitative and quant work. This can support the data-driven strategies of local and national governments, providing detailed and regular insight beyond what is usually available. The ability to connect academia and the public sector is one core advantage of this type of data, helping to revolutionise policy decision-making.

Opportunities

Additional contextual data such as council, census, health, and education data can allow for the creation of powerful research and insights tools, such as our Economic Wellbeing Explorer, highlighting changing patterns in financial wellbeing over time.



This is highly sensitive data, requiring appropriate safeguards to protect it from being misused. Despite being de-identified, such data can still form a target for malicious actors due to its perceived value. We must also ensure that those using the data are well-informed as to risks and associated regulations, enabling them to safely use the data without fear of re-identification or other privacy breaches, especially when in novel trusted research environments.

Threats

Smart Data Foundry closely follows the 5 Safes framework developed by the ONS, also utilising a trusted research environment provided by EPCC. Strong security protocols are in place through the use of airgapping and virtual machines to make sure that threats are minimised.



SDF also has an in-house Information Governance team that closely manage data access in line with ours and regulatory policies. We ensure that all data use is in line with our missions and remit as defined in collaboration with the ICO.

