CITP
CENTRE FOR INCLUSIVE TRADE POLICY

The Jobs in Trade Database: Methodology and User Guide

James Black,Calum Fox, Mairi Spowage

September 2026

Centre for Inclusive Trade Policy

UKRI Economic and Social Research Council

Centre for Inclusive Trade Policy
https://citp.ac.uk/
info@citp.ac.uk

Established in 2022, the Centre for Inclusive Trade Policy (CITP) is the first research centre dedicated to trade policy to be funded by the Economic and Social Research Council. As a centre of excellence for innovative research on trade policy and its inclusiveness, we aim to equip the UK with the capability to formulate and implement a trade policy tailored to the needs of the whole of the UK, while recognising the importance of the multilateral trading system and the UK's role within it.

Acknowledgements

This work was conducted by members of the Centre for Inclusive Trade Policy, supported by the ESRC [grant number ES/W002434/1].

The analysis was carried out in the Secure Research Service, part of the Office for National Statistics. This work contains statistical data from ONS which is Crown Copyright. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.

We would like to thank all those at the Centre for Inclusive Trade Policy, the Department for Business and Trade, the Scottish Government and the Office for National Statistics who provided advice, feedback and support.

Any questions can be directed to the lead author

James Black: james.black@strath.ac.uk

Contents

1. Introduction 3

2. Source data 3

2.1 Annual Business Survey 4

2.2 International Trade in Services survey 7

2.3 Annual Survey of Hours and Earnings 7

2.4 Annual Purchases Survey 9

2.5 Business Structure Database 10

2.6 Northern Ireland Annual Business Inquiry and Broad Economy Sales & Export Statistics 10

3. Data linkage 11

3.1 Statistical units and the IDBR 11

3.2 Construction of export-status classifications 12

3.3 Export-status classification imputation 14

3.4 Linkage for GB business data outputs 15

3.5 Linkage for GB employee data outputs 15

3.5.1 Balancing employee job outputs 17

3.6 Linkage for Northern Ireland employee data outputs 18

4. Statistical concepts and interpretation 19

4.1 What is an 'exporter' and 'non-exporter'? 19

4.2 What is 'Approximate' Output, GVA and Intermediate Consumption etc? 20

4.3 What is a 'full-time equivalent employee'? 21

4.4 Why do employee counts differ from other data sources? 21

4.4 What are 'skill' groups? 23

4.5 Intended and non-intended uses 23

5. Outputs 24

5.1. Overview of business data outputs 24

5.2. Overview of employee job data outputs 25

5.3 Industry coverage 27

5.4 Public sector coverage 28

5.5 Export coverage 28

5.6 Detailed indicator information 28
5.6.1 Business indicators 28
6. Future extensions 36
7. Conclusion 37
References 38
Appendix 40

1. Introduction

Trade is an important source of opportunity for firms and the wider economy, but its effects are unlikely to be distributed evenly. Exporting may be associated with different types of firms, different types of jobs, different wage structures and different patterns of investment and purchasing. Understanding these differences matters for analysts interested in trade policy, productivity, labour markets and inclusive growth.

However, the published UK evidence base does not currently allow users to examine these questions in much detail. Many economic statistics are available by industry, and trade statistics provide detailed information on exports, but there is little published data showing how exporting and non-exporting firms differ across the wider economy. Existing UK outputs provide some evidence on labour productivity by export status for a small number of broad industry groups. For many economic indicators that are relevant to trade and employment analysis - including output, GVA, intermediate consumption, employment costs, wages, employee characteristics, capital expenditure and purchases - there is no published UK dataset that consistently separates exporters from non-exporters at detailed industry level and across time.

This report documents the construction of a new dataset that addresses part of this gap. The dataset disaggregates selected UK business and employee indicators by industry, year and firm export status. A key feature of the work is that export status is linked to both firm-level economic indicators and employee records. This allows users to examine the characteristics and earnings of employees working for firms classified as exporters or non-exporters. Many outputs are produced for 48 industries and are available as time series, supporting both descriptive analysis and integration with industry-level analytical frameworks that require a consistent industrial structure.

The report focuses on the construction and interpretation of the dataset rather than on substantive findings. Section 2 describes the source datasets. Section 3 explains the linkage and construction of export-status classifications. Section 4 discusses the statistical concepts needed to interpret the outputs. Section 5 describes the published outputs. Section 6 discusses coherence checks and data quality concerns. Section 7 outlines possible future extensions.

2. Source data

We used seven datasets to produce our data.

Data for Great Britain was produced using the Annual Business Survey (ABS) (Office for National Statistics, 2025), International Trade in Services survey (ITIS) (Office for National Statistics, 2025), Annual Survey of Hours and Earnings (ASHE) (Office for

National Statistics, 2025), Annual Purchases Survey (APS) (Office for National Statistics, 2023) and Business Structure Database (BSD) (Office for National Statistics, 2025).

Data for Northern Ireland was produced using NISRA's Annual Business Inquiry (ABI) and Broad Economy Sales & Exports Statistics (BESES).

For GB data, the key sources for determining export classification were ABS and ITIS. The resulting firm-level classifications were linked to ABS, ASHE and APS to disaggregate a range of indicators by firm export status. BSD was used to support the linkage between export classification and ASHE.

For NI data, BESES was the source used for export classification and was linked to ABI to disaggregate employees by export status.

Each dataset comes with its own set of strengths and limitations which we will now outline.

2.1 Annual Business Survey

Type: Survey

Sampling units: The ABS is sampled at the reporting unit level which supports linkage to many other ONS surveys such as ITIS, APS and so on. Units can represent businesses, charitable organisations and public sector.

Sample size: The ONS Annual Business Survey is a large survey with around 50,000 GB respondents annually.

Geographical coverage: The dataset provided to researchers covers reporting units in Great Britain only. NISRA collects comparable NI data separately using the Annual Business Inquiry.

Industry coverage: The ABS samples units from the vast majority of Standard Industrial Classifications ('industries'). Exclusions are laid out in our 'Industry Coverage' section but notable include parts of Agriculture, Finance & Insurance, and Public Administration and Defence.

Public sector coverage: Public sector units are not sampled in 'public sector industries' (SIC 84-88) but are sampled in all other industries. Note that other datasets that use ABS as a source of information may opt to exclude sampled public sector units in favour of administrative data.

More information can be found on the ABS QMI webpage.

Table 1: Variables used from the Annual Business Survey Microdata

Variable	Description	Usage
aweight	Design weight	Used with gwtto or gwtemp to weight data.
gwtto	G-weight derived from IDBR turnover	Used with aweight to weight non-employment-cost data.
gwtemp	G-weight derived from IDBR employment	Used with aweight to weight employment-cost data (wq252, wq446).
frm_type	Form type (i.e. specific questionnaire)	Used to apply correct adjusted Intermediate Consumption derivation by form type.
ruref	Reporting unit reference	Used to link between datasets.
sic07	Standard Industrial Classification 2007	Used to create industry aggregations.
q15	Did your business export goods to individuals, enterprises or other organisations based outside the UK? eg raw materials, semi or finished manufactured goods (Yes/No Question)	Indicates goods export status
wq163	What was your income from services provided to individuals, enterprises or other organisations based outside the UK?	Indicates services export status
wq26	Total turnover from clients resident within the rest of the European Union	Indicates services export status
wq27	Total turnover from clients resident outside the European Union	Indicates services export status
wq600	What was the total value of capital assets? - Value of Acquisitions	Capital expenditure acquisitions
wq699	What was the total value of capital assets? - Value of Proceeds from Disposals	Capital expenditure disposals
wq412	Amounts payable in national non-domestic (business) rates	Part of taxes on production
wq413	Other amounts paid for rates, duties, levies and taxes	Part of taxes on production
wq432	Subsidies received under The Work Programme	Subsidies on production
wq570	Approximate intermediate consumption at purchaser's prices	Use to derive adjusted Intermediate Consumption
wq433	Expenditure on services bought for resale without further processing	Use to derive adjusted Intermediate Consumption
wq479	Expenditure on parts used solely in repair and servicing activities	Use to derive adjusted Intermediate Consumption
wq614	Approximate output at basic prices	
wq345	Value of retail sales of goods direct to the general public for personal or household use	
wq272	Increase in total stocks (year end less year beginning)	
wq252	Total employment cost	
wq446	Gross wages and salaries	
wq613	Approximate GVA at basic prices	
wq499	Total purchases of goods, raw materials, energy and services	

The dataset provided to researchers includes variables for Gross Fixed Capital Formation, Gross Operating Surplus, Gross Value Added in Factor Prices. However, after discussions with ONS we chose not to use these variables due to quality concerns.

We constructed a Net Capital Expenditure indicator using other available variables. Users should be aware that there are some differences between our Net Capital Expenditure indicator and the concept of Gross Fixed Capital Formation used in National Accounts, such as:

- ABS does not capture business research and development expenditure and military weapons systems which is included in National Accounts since Blue Book 2014, reflecting ESA 2010 changes.
- ABS only captures investment in non-residential buildings.
- ABS captures land acquisition, purchases of used buildings and structures, and other non-produced assets such as goodwill. These are not included in the National Accounts GFCF measure. This impact will be somewhat mitigated at the whole economy level as each GB buyer (i.e. acquisition) will often have a corresponding GB seller (i.e. disposal). However, we expect less mitigation at the industry level if firms are likely to transact with firms in other industries.

ABS Metadata Information (Office for National Statistics, 2015) states that the ABS indicators for 'Approximate output at basic prices' can be calculated as the sum of 'Approximate GVA at basic prices' plus 'Approximate intermediate consumption at purchaser's prices'. We found that this calculation did not hold for firms responding to the following ABS forms: Motor trades, Wholesale, Retail, Catering and Property. Removing 'Services bought for resale without further processing' from the intermediate consumption of firms associations with these forms and removing 'Expenditure on parts used solely in repair and servicing activities' for Motor Trades resulted in the calculation holding. While the ABS intermediate consumption indicator is not provided for services industries prior to 2013, we could use the other components of the formula to calculate the expected values.

Taxes less subsidies on production covers many but not all taxes and subsidies. Notable exceptions include rail subsidies and research and development tax credit subsidies. These exceptions contribute to differences when compared with National Accounts data, along with the other differences listed in 4.2 What is 'Approximate' Output, GVA and Intermediate Consumption etc?. The magnitude of these differences will vary by industry.

Total purchases of goods, raw materials, energy and services relates to intermediate consumption of goods and services only. E.g. it does not include employment costs or capital investment.

Our main indicator for whether a firm exports goods, q15, is only available from 2011 onwards. As this indicator is used to construct our firm export status classifications, this limits all of our outputs to 2011 onwards.

2.2 International Trade in Services survey

Type: Survey

Sampling units: ITIS is sampled at the reporting unit level.

Sample size: ITIS is a large survey of around 27,000 reporting units annually.

Geographical coverage: UK

Industry coverage: ITIS samples units from the vast majority of industries. Exclusions beyond ABS include 'Distribution of gaseous fuels through mains' (SIC 35.220), all of 'Remediation activities and other waste management services' (SIC 39), most of 'Transport' (SIC 49-52), parts of 'Accommodation' (SIC 55), parts of 'Rental and leasing activities '(SIC 77), parts of 'Travel agency and other reservation services' (SIC79), and parts of education, health, care and social work (SIC 85-88).

Public sector coverage: Yes

More information can be found on the ITIS QMI webpage.

Table 2: Variables used from the International Trade in Services Microdata

Variable	Description	Usage
ru_ref	Reporting unit reference	Used to link between datasets.
j_weight	Product-country imputation weights	
a_weight	Design weights	
o_weight	Outlier weights	
q204	Value of exports	

We are not particularly concerned about the ITIS industry exclusions laid out above as the ABS services export questions provides industrial coverage where ITIS does not. However, both surveys do not capture non-resident expenditure as exports and we would refer users to our discussion in 'Export Coverage'.

2.3 Annual Survey of Hours and Earnings

Type: Survey

Sampling units: The Annual Survey of Hours and Earnings uses a 1% random sample of employee jobs from the HMRC Pay As You Earn (PAYE) system. Employers are legally required use PAYE if any of their employees earn more than the National Insurance Lower Earnings Limit (£129 per week in 2026/27). Self-employed jobs and jobs which do not earn over the reference period are excluded.

Sample size: Around 180,000 employee jobs are sampled.

Geographical coverage: The dataset provided to researchers covers GB only.

Industry coverage: All industries.

Public sector coverage: Yes

Weighting: Results are weighted to match the Labour Force Survey employee job totals for key attributes (age, sex, region, occupation).

More information can be found on the ASHE QMI webpage.

Table 3: Variables used from the Annual Survey of Hours and Earnings Microdata

Variable	Description	Usage
wcoa	Census Output Area of employees work location	Used to map enterprise to reporting unit before linking to export classifications.
wlsoa	Lower Super Output Area of employees work location	Used to map enterprise to reporting unit before linking to export classifications.
wmsoa	Middle Super Output Area of employees work location	Used to map enterprise to reporting unit before linking to export classifications.
wgor	Region of employees work location	Used to map enterprise to reporting unit before linking to export classifications & filter hourly earnings.
entref	Enterprise unit reference	Used to link to BSD
ft	Full / part time marker	Used to calculate full-time equivalent employee jobs
occ10	Standard Occupation Classification 2010	Used to determine skill group
occ20	Standard Occupation Classification 2020	Used to determine skill group
sex	Employee sex	Results by sex
age	Employee age	Used to determine age group & filter hourly earnings.
he	Hourly earnings for the reference period	Used to calculate median hourly earnings
adr	Adult rate marker	Used to filter hourly earnings.
lop/lopf	Loss of pay/loss of pay excluding furlough marker	Used to filter hourly earnings.

For median hourly earnings, we apply the recommended filter in the ASHE documentation. This excludes the following individuals: those who are younger than sixteen; are not paid adult rates; are without a valid work region; whose pay in the reference period was affected by absence; and whose hourly earnings are zero.

ASHE was not intentionally designed to produce job counts and job count results are not treated as definitive. However, ASHE is used for job count analysis in some contexts. ASHE is the official source of estimates for the number of jobs paid below the national minimum wage. We believe the job counts are robust enough for purposes of apportionment but caution that they will not have the rigour of official employee count statistics. Industry totals will not match with other sources such as ONS Business Register and Employment Survey. More discussion can be found in '4.4 Why do employee counts differ from other data sources?'.

Research suggests that smaller and younger organisations may be under-represented in ASHE (Forth, et al., 2025) which may have implications for our results.

2.4 Annual Purchases Survey

Type: Survey

Sampling units: Reporting unit, sampled from the IDBR.

Sample size: Around 33,000 businesses.

Geographical coverage: UK

Industry coverage: Exclusions are parts of 'Agriculture', parts of 'Financial and insurance activities', all of 'Public administration and defence; compulsory social security' (SIC 84), all of 'Activities of households as employers; undifferentiated goods and services producing activities of households for own use' (SIC 97-98), all of 'Activities of extraterritorial organisations and bodies' (SIC 99). These exclusions are a subset of the ABS exclusions and therefore do not further affect our outputs.

Public sector coverage: Private sector only – public and charitable organisations are not covered. Unlike our outputs which use ABS as the source, our outputs of APS-sourced data do not contain any public or charitable organisations.

More information can be found on the APS QMI webpage.

Table 4: Variables used from the Annual Purchases Survey Microdata

Variable	Description	Usage
rureference	Reporting unit reference	Used to link between datasets.
currentsic	Standard Industrial Classification 2007	Used to create industry aggregations.
q4010_p*	Expenditure on intermediate consumption of energy, water and waste	Contributes to total expenditure on intermediate consumption
q6010_p*	Expenditure on intermediate consumption of services	Contributes to total expenditure on intermediate consumption
q8010_p*	Expenditure on intermediate consumption of goods, materials and related services	Contributes to total expenditure on intermediate consumption
q5010	Total expenditure on energy, water and waste in the UK	Used to apportion UK/international values
q5020	Total expenditure on energy, water and waste outside the UK	Used to apportion UK/international values
q7010	Total expenditure on services in the UK	Used to apportion UK/international values
q7020	Total expenditure on services outside the UK	Used to apportion UK/international values
q9010	Total expenditure on goods, materials and related services in the UK	Used to apportion UK/international values
q9020	Total expenditure on goods, materials and related services outside the UK	Used to apportion UK/international values

Annual Purchases Survey captures the value of purchases of energy, services, goods and materials used in intermediate consumption. It therefore excludes expenditure on capital investment, staff costs, and goods and services bought for resale without further processing.

The APS was not collected for the years 2019 and 2020.

2.5 Business Structure Database

Type: Administrative

Source: BSD is a point-in-time snapshot of the Inter-Departmental Business Register (IDBR), a register of all businesses in the UK. The IDBR primarily sources its information from HMRC Value Added Tax records, HMRC Pay As you Earn records and Companies House.

Size: Around 4 million enterprises and 5.5 million local units.

Geographical coverage: UK

Industry coverage: All industries

Public sector coverage: Public sector units are included in the IDBR.

More information can be found on the UK Data Service BSD Documentation.

Table 5: Variables used from the Business Structure Database Microdata

Variable	Description	Usage
death_date	Date of local unit death	Used to remove dead local units.
entref	Enterprise unit reference	Used to link with ASHE.
ruref	Reporting unit reference	Used to link with ABS.
luref	Local unit reference	Used to detect approximate geography and thus reporting unit for mapping enterprise to reporting unit in ASHE.
coa_2011	Local unit Census Output Area (2011)	Used for ASHE enterprise to reporting unit mapping
soa_lower_2011	Local unit 2011 Census Lower SOA code	Used for ASHE enterprise to reporting unit mapping
soa_middle_2011	Local unit 2011 Census Middle SOA code	Used for ASHE enterprise to reporting unit mapping
main_gor	Local unit Government Office Region	Used for ASHE enterprise to reporting unit mapping

2.6 Northern Ireland Annual Business Inquiry and Broad Economy Sales & Export Statistics

Type: Survey

Sampling units: Reporting unit, sampled from the IDBR.

Sample size: Around 11,000 reporting units.

Geographical coverage: Northern Ireland.

Industry coverage: In line with the Annual Business Survey (NISRA, 2026).

Public sector coverage: In line with the Annual Business Survey.

More information can be found on the Annual Business Inquiry webpage.

Table 6: Variables used from the Northern Ireland Annual Business Inquiry Microdata

Variable	Description	Usage
psic	Standard Industrial Classification 2007	Used to create industry aggregations.
calwght	Survey calibration weights	Weighting
wq051	Total number of part-time employees	Used to calculate FTE employees
wq052	Total number of full-time employees	Used to calculate FTE employees
wq050	Total employees	Used to calculate FTE employees
export	Total export value	Used to classify export status.

BESES is collected as part of the NIABI.

An in-depth examination of the Northern Ireland ABI data and the lack of stability of our estimates led to potential concerns about quality. We have opted not to publish business data for Northern Ireland and only to retain the employee data which appears much more in line with FTE employee data from BRES.

3. Data linkage

This section explains how the source datasets were linked to assign export-status classifications to business and employee records, and summarises the main linkage assumptions and limitations.

3.1 Statistical units and the IDBR

The Inter-Departmental Business Register is a register of 2.7 million UK businesses collected for statistical purposes such as acting as the main sampling frame for many business surveys.

IDBR units are categorised as administrative units, statistical units and reporting units:

- Administrative units include Value Added Tax (VAT) traders and Pay As You Earn (PAYE) employers.
- Statistical units include the enterprise, enterprise group, local unit and kind of activity unit (KAU).
- The reporting unit is the unit that businesses reports at.

For our purposes, the three that are most important are:

- Local unit: A single location for the business, e.g. an office, factory, workshop etc.
- Reporting unit: This represents the mailing address for many ONS business surveys. Complex business may have multiple reporting units that reflect

different kinds of activities. The reporting unit provides data reflecting all of its associated local units

- Enterprise unit: Defined as ‘the smallest combination of legal units that is an organizational unit producing goods or services, which benefits from a certain degree of autonomy in decision-making, especially for the allocation of its current resources. An enterprise carries out one or more activities at one or more locations. An enterprise may be a sole legal unit.’ (European Union, 2008)

A simple business which operates out of a single location, can have a single local unit, reporting unit and enterprise unit. For example, an independently owned hairdresser.

A more complex business which operates out of multiple locations can have multiple local units, a single reporting unit and a single enterprise unit. For example, a small hairdresser chain.

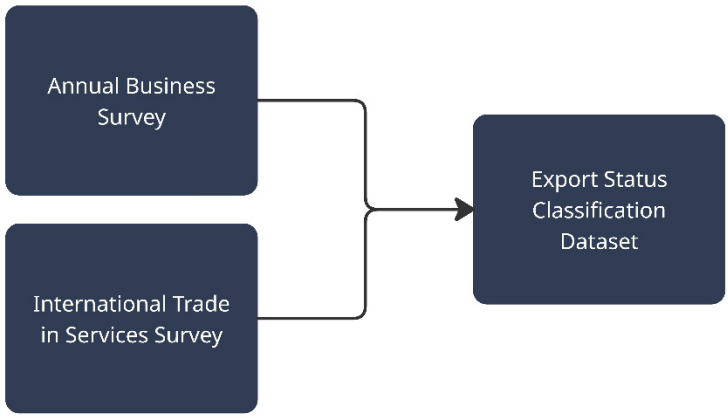
A yet more complex business could operate out of multiple locations performing different activities can have multiple local units, multiple reporting units and one or more enterprise units. For example, a company that owns a hairdresser chain and also manufactures hair products.

Businesses can be more complex still, with many brands, subdivisions and so on. However, for our purposes it is most important to understand that one enterprise can have one or multiple reporting units and each reporting unit can have one or multiple local units.

Most ONS business surveys sample reporting units from the IDBR and therefore collect data at the reporting unit level. This means the data from these surveys can be linked together using the common reporting unit. However, other data sources can be collected at different levels of statistical units. For example, the ONS Annual Survey of Hours and Earnings samples HMRC PAYE units. PAYE units are mapped to enterprise units but researchers then face an issue trying to link data at the enterprise unit to other data at the reporting unit, as large firms can have a one-to-many mapping.

3.2 Construction of export-status classifications

The firm export-status classifications used in our GB outputs were constructed at the reporting unit-year level by combining data from the Annual Business Survey (ABS) and International Trade in Services survey (ITIS) for years 2011 to 2022.



Between 2011 and 2022, the combined ABS-ITIS dataset included, on average, 36,907 reporting units per year that appeared in the ABS only, 11,242 that appeared in ITIS only, and 8,030 that appeared in both surveys. The annual breakdown is provided in the Appendix (Table 11).

Goods export status was taken directly from the ABS goods export question (q15: "Did your business export goods to individuals, enterprises or other organisations based outside the UK?" Yes/No).

Services export status was constructed using ITIS services exports where these were available and falling back to ABS services export variables where ITIS did not provide a classification. A reporting unit was classified as a services exporter where a positive services export value was observed and as a services non-exporter where a zero value was observed.

Total export status was then defined as:

- 'Exporter' if a reporting unit had evidence of *either* goods exports or services exports in a given year.
- 'Non-exporter' if a reporting unit had evidence of *both* no goods exports or services exports in a given year.
- 'Unknown' if a reporting unit had either no goods or services export data in a year or if it had only one of these and that figure suggested non-exporter[1].

In our dataset, 322,463 observations (reporting unit-year pairings) were classified as non-exporters, 249,474 as exporters and 102,207 as unknown (Table 12).

[1] Where only one of goods exporters or services exports is available for a reporting unit in a given year, if this suggests that the unit is an exporter then we can still classify the unit as an exporter. However, if it suggests that the unit is a non-exporter, we do not have sufficient information to claim that it is a non-exporter. There are asymmetric requirements for the classification due to the OR nature of the exporter classification and the AND nature of the non-exporter classification.

3.3 Export-status classification imputation

Statistical disclosure risks limit our ability to publish data with an ‘Unknown’ export status category. The ‘Unknown’ category is often relatively small and would require suppression and, with suppression, we would have to undertake secondary suppression of many other cells to avoid calculation of the suppressed cells. We chose to follow an imputation strategy to estimate missing values. This reduces suppression requirements which enhances the usefulness of our outputs with a trade off in terms of assumptions.

To impute, we first balanced the reporting unit panel so that all reporting units observed at any point between 2011 and 2022 had an observation for every year in that period. Missing goods export status, services export status and industry were then imputed using information from the same reporting unit in other years.

Our first rule used the nearest available observations in time. Where there were two equally near observations with disagreement between export statuses, the rule moved a further year out to seek agreement.

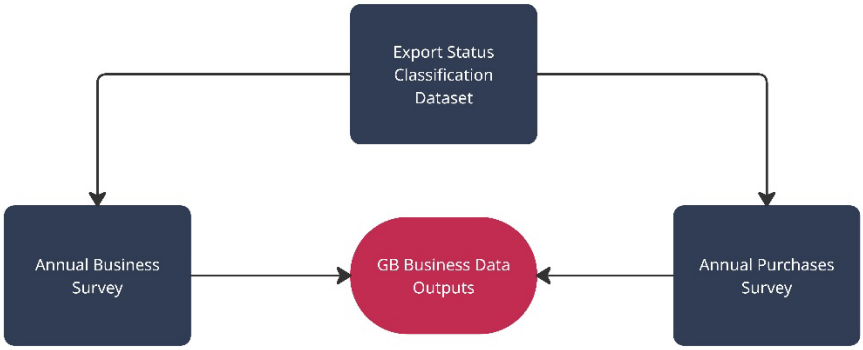
Our second rule applied to any observations still missing a status after the first rule. We assigned the majority observed status for that reporting unit across years where there was a clear majority.

No data was available for some units. These units were typically very small, with many having zero employment, and made up a small proportion of each industry. Most ‘unknown’ classifications are due to missing goods export status and driven by firms neither selecting “yes” or “no” to the respective export question in the Annual Business Survey. Production of a separate ‘unknown’ group in the results would result in statistical disclosure issues that would require suppression of exporter and non-exporter groups. We therefore took the decision to classify these firms as non-exporters. This changes interpretation of our ‘non-exporter’ classification from ‘evidence of not exporting’ to ‘no evidence of exporting’.

For industry, any remaining missing values were assigned to the most frequently observed industry for that reporting unit across years.

The impact of this strategy is to greatly increase the number of reporting unit-year pairs in our dataset by imputing non-surveyed years. Results can be found in Table 13.

3.4 Linkage for GB business data outputs

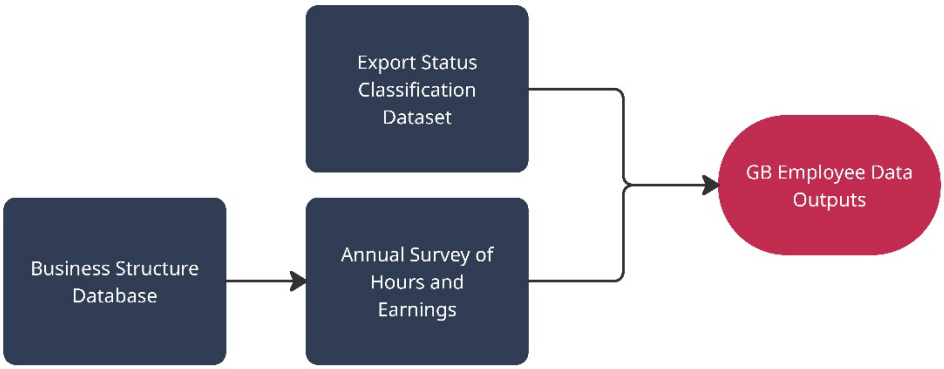


For business data outputs covering Great Britain, our export status classification dataset can be linked via reporting units to the Annual Business Survey and Annual Purchases Survey. This enriches both of these datasets with firm-level export status classifications and enabled us to produce breakdowns of a range of ABS and APS indicators by export status.

Of the total 539,238 reporting unit-year pairs in the Annual Business Survey between 2011 and 2022, 337,380 were classified as non-exporters, 172,962 as exporters and 28,896 as unknown (Table 14). As mentioned previously, units classified as 'unknown' are reclassified as 'non-exporters'.

Of the 28,896 'unknown' observations, 1,752 belong to the 'Excluded' industry classification. The remaining observations are disproportionately more likely to be sole traders or partnerships and less likely to be companies or public sector. Excluding public sector and 'Excluded' industry units, exporting units had a median employment of 68, non-exporting units of 15 and 'unknown' units of 2 (mean 411, 201 and 36 respectively). Units with an unknown classification are typically very small firms, with around 37% of this group having employment of zero or one.

3.5 Linkage for GB employee data outputs



For employee data outputs covering Great Britain, an additional linkage step through the Business Structure Database is required because the Annual Survey of Hours and

Earnings survey only reports the enterprise unit and is not enriched with reporting unit[2]. As our export status classification dataset operates at the reporting unit while ASHE operates at the enterprise unit, there were two options:

1. Aggregate our export status classification dataset from reporting unit to enterprise unit and perform the linkage at the enterprise unit level.
2. Enrich ASHE with reporting units and perform the linkage at the reporting unit level.

While the first option would be easiest to implement, it has a key disadvantage. For enterprises with multiple associated reporting units, aggregating the export status classification dataset to enterprise unit would obscure within-enterprise heterogeneity in export status, because reporting units that differ in export behaviour would be assigned a single enterprise-level classification. As enterprises with multiple reporting units tend to be large firms, this could result in many employees being associated with a different firm export status and, in particular, biasing away from a 'non-exporter' status as it would only take a single 'exporter' status among all reporting units in an enterprise unit for the enterprise unit to be classified as an 'exporter'.

The second option has the significant advantage of maintaining reporting unit level export classifications. Enrichment was performed by comparing the reported location of the employees workplace with the location of local units associated with a given reporting unit in an enterprise. This began at the most granular geography provided, census output areas, and successively moved out to less granular geographies when no match was found. Where an employee's work location was matched to locations associated with a single reporting unit in a given enterprise, we listed the reporting unit of the employee's firm as this matched unit. A limitation of using geography was that the output areas related to the 2011 census and were only introduced in the datasets from 2015 onwards, thus limiting our analysis to the same period. ASHE included an average of 55,573 enterprise units between 2015 and 2022. Of these, our method fully matched all employees to reporting units in 51,739 enterprise units, matched some but not all employees in 948 enterprise units and did not match employees in 2,887 enterprise units (Table 15). The 55,573 enterprise units represented the average ASHE sample of 164,473 employees between 2015 and 2022. Of these, an average of 154,460 employees were matched to reporting unit, representing 93.9% of the observations (Table 16). Our enrichment process has therefore successfully estimated reporting unit for a high proportion of employees and enterprises sampled by ASHE.

The remaining unmatched employee observations comprises of two groups. The first group is employees with no match and we assume this is due to dataset timing

[2] Dataset metadata claims that ASHE also includes local unit however, on inspection, we found that these local units had been newly generated for this dataset rather than referring to standard local units used across ONS datasets. This means that these local units cannot be used for linkage.

differences, missing information or industry exclusions. The second group is employees who have been matched to multiple potential reporting units which belong to a single enterprise unit and both operate in the same geography. In this case, we are unable to determine which reporting unit these employees belong to. Of the unmatched employees, around a fifth fall in our 'Excluded' industry and a further 39% fall in heavily public sector industries (SIC 85 & 86). Private sector firms that employ the unmatched employees have a unweighted median employment of 17, compared to 724 and 2178 for non-exporters and exporters respectively.

Once ASHE is enriched with reporting unit, we can link the resulting dataset to our export status classification dataset and results in employee observations in ASHE being association with a firm-level 'exporter' or 'non-exporter' classification.

3.5.1 Balancing employee job outputs

An important part of producing these statistics was ensuring that the outputs do not disclose any information on individual employees or firms. As a result, some of the data has been rounded where necessary. We have then undertaken a constraints method to ensure that column and row totals balance.

The ONS Secure Research Service sets minimum business and employee job counts for each cell of data as part of its Statistical Disclosure Control processes. Where a cell has fallen below this threshold, the data must be obfuscated or suppressed in some way. For any problematic cells in the GB employee jobs data, we opted to round the unweighted counts before weighting, thus obfuscating the true weighted result. These cells are marked in orange in the accompanying spreadsheet and the cell is marked as 'R' for 'rounded' in the associated 'Markers' sheet.

As the employee job data has row and column totals and subtotals, further cells must also be suppressed to ensure that the true values cannot be revealed for the obfuscated cells using the totals. These further suppressed cells have also had their unweighted counts rounded before weighting. They are marked in yellow in the accompanying spreadsheet and the cell is marked 'S' for 'secondary' in the associated 'Markers' sheet. Any cells marked 'S' have more than the minimum number of business and employee job counts and so the rounding of their unweighted counts is likely to lead to less proportionate change relative to the true value than it is for cells marked 'R'.

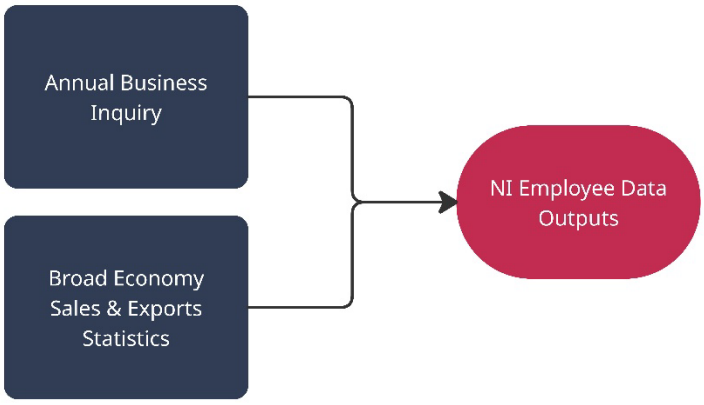
The benefit of this approach to obfuscation over complete suppression is that some cells, particularly those suppressed due to secondary suppression, still provide useful data even if it is somewhat rounded, rather than complete removal of those cells. However, the drawback is that row and column cells no longer round to totals and subtotals which is unhelpful for modelling uses. As a result, we have performed a balancing methodology to ensure that cells sum to the totals and subtotals. The following rules are used:

- Only cells that are rounded, i.e. those marked ‘R’ or ‘S’ are modified. Cells that have not been rounded are not modified.
- Cells cannot be changed by more than the proportionate error indicated by a corresponding table of unweighted counts, which has been rounded to the nearest 20. For example, if a cell is marked ‘S’ and it’s corresponding unweighted count is shown as 100 after having been rounded to the nearest 20, we know that the true unweighted count is greater or equal to 90 and less than 110. This indicates the maximum potential rounding of that cell.
- Cells must round to column subtotals, row totals, and the whole period (2015-2022) values).
- Proportional changes to cells are minimised.
- In some cases, judgements have been made using other relevant datasets to set a new starting value while keeping the original error boundaries. For example, comparing employee job totals by sex to datasets that publish similar information.

While the result is a balanced set of results, we would encourage caution when using yellow shaded cells and would advise against direct use of orange shaded cells if they are a focus of an analysis.

Rebalancing was only applied to GB employee job data due to the complexity and resourcing requirements of the process.

3.6 Linkage for Northern Ireland employee data outputs



The Northern Ireland component of the Interdepartmental Business Register (IDBR) is maintained by NISRA with NISRA also conducting a number of surveys independently from the ONS. Associated NISRA researcher datasets are often not available in the same Trusted Research Environment as ONS data and, even if they were, would not necessarily be readily linkable to ONS data. For example, the ONS collects the Annual Business Survey (ABS) while NISRA collects the Annual Business Inquiry (ABI). These are combined by the ONS for National Accounts but are not available in the same workspace and readily linkable for researchers.

As a result, our treatment of Northern Ireland units has had to differ from GB units. We have drawn on our outputs from a prior exploratory project (Black, Catalano, Crummey, & Spowage, 2023) to present some Northern Ireland data. This research uses the Annual Business Inquiry and Broad Economy Sales & Export Statistics, as laid out earlier in this document. BESES is collected as part of the ABI so no further linking is required. We had also explored linking to the Northern Ireland version of BRES but this plan was abandoned after discovering that the reporting units listed in BRES were not produced on a linkable basis to the ABI.

Due to the inability to link the ABI/BESES to ASHE, we are not able to produce breakdowns of employee jobs by characteristics. The FTE employee results for Northern Ireland are produced using the IDBR-sourced employee data rather than ASHE employee job data, so differs to the GB results.

While we initially produced results for a range of ABI indicators similar to the business indicators produced for GB, lack of stability of the results across time led to concerns about robustness and so we have decided not to publish these indicators. FTE employee data appears more reliable and is our sole Northern Ireland output.

4. Statistical concepts and interpretation

Our dataset introduces firm export status as an additional dimension for a range of economic indicators. This supports many potential uses:

- Investigating the differences between exporter and non-exporter groups of firms at a detailed industry level.
- Exploring trends for exporters and non-exporter groups and their employees across time.
- Studying differences in employee characteristics of employees working for exporters and non-exporter groups and the differential impacts of changes in the workforce.
- Data produced at a level of granularity that supports integration and the extension of many industry-level economic models, such as Input Output, Global Value Chain and Computable General Equilibrium models.

4.1 What is an ‘exporter’ and ‘non-exporter’?

Export status is classified at the firm level, specifically the reporting unit level.

A firm is classified as an exporter in a given year if we have any evidence that it has exported either goods or services in that year. Evidence can include a firm responding in the affirmative to exporting in ONS ABS, a firm providing a positive non-zero export return in ONS ITIS or, when a firm is not sampled, imputed data from the closest available year of sampled data.

A firm is classified as a non-exporter in a given year if we do not have evidence that it has exported either goods or services in that year. Note that this is subtly different to being a true 'non-exporter' as there are a very small number of firms for which we have no information (typically new or small firms). As introducing a third 'unknown' category would have led to significant statistical disclosure issues, we have classified these firms as non-exporters.

Firms are grouped by their exporter classification and industry. Both export classification and industry can vary across years and so firms can move in and out of their export status-industry pairings. This has the potential to make some data appear more volatile across time if, for example, a relatively large firm in an industry is reclassified to a different export status or a different industry.

As our exporter status is classified at the firm level, our employee data represents employees working for a firm classified as an exporter or non-exporter, rather than representing the specific activities of that employee.

4.2 What is 'Approximate' Output, GVA and Intermediate Consumption etc?

Terms such as output, GVA, intermediate consumption and so on are concepts used in UK National Accounts. These are compiled using a range of data sources before making a number of adjustments to ensure the data lines up with the concept.

The Annual Business Survey is the primary data source for many National Accounts indicators, such as output, intermediate consumption and GVA, with adjustments summarised as (Office for National Statistics, 2014):

- Removal of the partial coverage of non-market activity included in ABS;
- Addition of non-market and own-use data from other sources (e.g. administrative public sector and charitable organisation data, own-use output of households);
- Adjustments to align taxes less subsidies on production;
- Adjustments to align inventories (changes in stocks);
- Coverage adjustments (e.g. for non-sampled industries, businesses which are too small to appear on the sampling frame etc);
- Conceptual adjustments (e.g. addition of illegal activities, benefits in kind to employees etc);
- Coherence adjustments (e.g. balancing between the three approaches to measuring GDP).

ABS indicators are therefore referred to as "approximate" versions of the National Accounts concepts.

When considering only the industrial sections covered by the ABS, approximate Gross Value Added (aGVA) made up between 90% and 94% of GVA in the years 2008 to 2011 (Office for National Statistics, 2014). ABS-sourced approximate indicators can be excellent proxies of their National Accounts counterparts but we caution users that they should be aware of why trends may diverge, such as:

- Some industries are not covered;
- Public sector units are not covered in "public sector" industries;
- ABS does not capture own-use activity of households;
- Even when ABS does capture data, National Accounts indicators may use other data sources. E.g. National Accounts uses administrative data in place of ABS non-market output and ABS-sourced taxes and subsidies;
- ABS does not capture subsidies on rail transport and tax credits on R&D activities;
- Other adjustments made may not scale linearly.

4.3 What is a 'full-time equivalent employee'?

We produce our employee data in 'full-time equivalent' terms which means our employee results will not be comparable to standard headcount data. Full-time equivalence is a simple adjustment intended to support comparisons across industries and time that takes the extent of part-time work into account. Our adjustment follows the common convention to count each full-time employee as one full-time equivalent employee and each part-time employee as half a full-time equivalent employee.

4.4 Why do employee counts differ from other data sources?

The ONS collects several surveys that capture employee counts. Each of these can differ in their purposes, methods, coverage, definitions and timing.

We introduce the following definitions:

- 'Employees' are individuals who are paid via payroll, e.g. the Pay As You Earn (PAYE) system.
- 'Employment' further adds self-employed individuals.
- 'Employee jobs' refers to the total number of payroll jobs. That is, one individual can hold multiple jobs.
- 'Employment jobs' is employee jobs with the addition of self-employed jobs.
- 'People in employment' is the number of individuals who are either employed or self-employed. Importantly, an individual who holds multiple jobs would only count as a single person in employment rather than having each job counted independently like in an 'employment jobs' definition.

The Business Register and Employment Survey (BRES) is dispatched to businesses using the IDBR as the sampling frame. BRES measures both employee jobs and employment jobs. The only form of self-employment captured in the latter is working proprietors, e.g. business owners who are not on payroll. BRES is a point-in-time measurement based on a single reference day in September. It is an excellent source of information on jobs by industry and region.

The ONS Labour Force Survey (LFS) is despatched to households. It primarily measures people in employment, including self-reported self-employed workers. Second job information is collected and reported separately but is not part of the primary reported indicator. The LFS is a regular rolling survey based on individual reference weeks which much more comprehensively covers a whole year once combined and therefore is useful for capturing seasonal employment. It is an excellent source of economy-wide employment and unemployment data.

Our employee data is collected from the Annual Survey of Hours and Earnings, which is a survey dispatched to businesses using the PAYE system as the sampling frame. As it uses the PAYE system as the sampling frame, it measures employees but not self-employment. We have not excluded second jobs from our analysis so our data is an employee jobs measure. ASHE is a point-in-time measurement based on a single reference day (and corresponding pay period) in April.

BRES is therefore the more conceptually similar dataset to our data. However, note that we have adjusted our data to report full-time equivalents rather than simple employee job counts. BRES also counts employees who did not work on the specific reference day whereas ASHE's sampling frame only includes employees who were paid during the reference period (Office for National Statistics, 2022), so employees with zero-hour contracts who did not work during the reference period or who were otherwise not paid during the reference period would be included in BRES statistics but excluded from ASHE statistics. Further differences exist between the datasets in sampling methods, weighting, timing of collection and reference period. These differences mean that users should not expect FTE-adjusted BRES data to align with our FTE employee job data. Constraining to FTE-adjusted BRES industry totals may be appropriate in some situations.

Note that our Northern Ireland data is not sourced from ASHE but is instead based on a point-in-time snapshot of the IDBR. The IDBR, in turn, sources its information from many sources such as BRES, PAYE and so on. This data will also be on an employee jobs basis and, similar to ASHE data, is not designed to produce representative counts of employment for the economy (such as how LFS is used for unemployment statistics) but is sufficient for our purposes.

4.4 What are ‘skill’ groups?

We define skill groups based on the Standard Occupational Code (SOC). Jobs are classified in SOC according to both skill level, the “length of time deemed necessary for a person to become fully competent in the performance of the tasks associated with a job” and skill specialisation, the “field of knowledge required for competent, thorough and efficient conduct of the tasks that comprise a job” (Office for National Statistics, 2026).

There are nine major groups in SOC 2010 and SOC 2020 which ONS has aggregated into four skill levels (Office for National Statistics, 2026). Attempting to produce our results at these four skill levels would have led to excessive suppression to avoid statistical disclosure, so we further aggregated these into two skill groups by combining skill levels 1 & 2 into ‘Lower’ and skill levels 3 & 4 into ‘Higher’.

Table 7: Skill groups mapped to Standard Occupational Classification

‘Skill’ Group	SOC2010/2020	SOC2010 Description
Higher	1	Managers, Directors and Senior Officials
Higher	2	Professional Occupations
Higher	3	Associate Professional and Technical Occupations
Lower	4	Administrative and Secretarial Occupations
Higher	5	Skilled Trades Occupations
Lower	6	Caring, Leisure and Other Service Occupations
Lower	7	Sales and Customer Service Occupations
Lower	8	Process, Plant and Machine Operatives
Lower	9	Elementary Occupations

4.5 Intended and non-intended uses

The dataset provides descriptive data that can be cut in several ways: by industry, export status, time period and, for some outputs, employee characteristics. These breakdowns can be used to examine how economic indicators differ between exporters and non-exporters, how those differences change over time, and how the data might be incorporated into wider analytical frameworks. The outputs should be interpreted as descriptive evidence, not as estimates of the effect of exporting.

The detailed industrial breakdown supports integration with a range of existing modelling approaches. For example, integration with Input Output tables, Computable General Equilibrium models and so on. In our next output, we will produce an Input Output table where industries are disaggregated into exporter and non-exporter groups (see 9. Future extensions).

However, our data does not assess causal relationships. For example, we cannot tell from our data whether exporting firms pay higher median hourly earnings because they

are exporters or whether firms that pay higher median hourly earnings are more likely to export.

Our data are experimental. All datasets come with limitations and linking several datasets leads to further assumptions. While these datasets have been linked in the past, this is the first time they have been linked to create our outputs. The novelty of our outputs means that they have not yet gone through years of extensive peer review. We recommend users acquaint themselves with the limitations laid out in this document as this may affect how they interpret some results, for example to consider whether changes are "real" changes or whether they are simply statistical artifacts.

5. Outputs

Our outputs consist of two spreadsheets:

1. Business data outputs: Indicators from the Annual Business Survey and Annual Purchases Survey disaggregated by industry, firm export status and year.
2. Employee data outputs: Annual Survey of Hours and Earnings based employee count and earnings indicators disaggregated by industry, firm export status and demographic breakdowns.

5.1. Overview of business data outputs

The business data workbook includes fourteen indicator sheets and four additional supporting sheets. All indicators are disaggregated by industry, firm export status and either year or year group. Where years are grouped, the data is presented in terms of annual averages.

Indicators represent Great Britain and do not include Northern Ireland as this data is held in separate datasets.

We have deflated any data which has been aggregated across years to 2022 prices using the GDP deflators provided in the supporting sheet.

Table 8: Structure and source of our business data outputs

Sheet	Industries	Years	Prices	Source
Output at Basic Prices	48	2011-2022	Current	ABS
Total Purchases	48	2011-2022	Current	ABS
Intermediate Consumption at Purchaser's Prices	48	2011-2022	Current	ABS
Net Capital Expenditure	48	2011-2022	Current	ABS
GVA at Basic Prices	48	2011-2022	Current	ABS
Direct Sales to Households	1	2011-2022	Current	ABS
Change in Stocks	36	2011-2022	Current	ABS
Employment Costs	48	2011-2022	Current	ABS
Gross Wages	48	2011-2022	Current	ABS
Taxes less Subsidies on Production	48	2011-2022 [1]	Current	ABS
Purchases from UK Sources	48	2015-2022 [2] (3 year groups)	2022 Prices	APS
Purchases from International Sources	48	2015-2022 [2] (3 year groups)	2022 Prices	APS
Purchases by Product 2015-18	48	2015-2018 (1 year group)	2022 Prices	APS
Purchases by Product 2021-22	48	2021-2022 (1 year group)	2022 Prices	APS

ABS = ONS Annual Business Survey; APS = ONS Annual Purchases Survey.

[1] Data for 2020-2021 has been suppressed due to reliability concerns.

[2] The Annual Purchases Survey was not collected for years 2019-2020.

The four supporting sheets are:

- Unweighted Counts for Annual Business Survey sourced data
- Unweighted Counts for Annual Purchases Survey sourced data
- GDP Deflators at Market Prices

5.2. Overview of employee job data outputs

The employee data workbook includes nine indicator sheets and seven additional supporting sheets. There are two core indicators: full-time equivalent employee counts and real median hourly earnings and these are disaggregated by employee characteristics to produce nine indicator sheets in total. All indicator sheets are disaggregated by industry, firm export status and either year or year group.

Eight of the indicator sheets represent Great Britain, excluding Northern Ireland, while one sheet represents Northern Ireland.

All GB indicator sheets are broken down by individual year whereas the NI indicator sheet is disaggregated into three year-groups. Where years are grouped, the data is presented in terms of annual averages. Each sheet also includes a whole period annual average (e.g. 2015-2022 annual average).

The business data workbook includes fourteen indicators and four additional supporting sheets. All indicators are disaggregated by industry, firm export status and either year or year group. Indicators represent Great Britain and do not include Northern Ireland as this data is held in separate datasets.

All hourly earnings data has been deflated using the CPIH deflators provided in the supporting sheet to support cross-year comparisons.

Table 9: Structure and source of our employee job data outputs

Sheet	Industries	Geog.	Years	Prices	Source
Full-Time Equivalent Employees	48	GB	2011-2022	n/a	ASHE
Full-Time Equivalent Employees in Northern Ireland	48	NI	2014-2020 (3 year groups)	n/a	ABI
Full-Time Equivalent Employees by Sex	48	GB	2011-2022	n/a	ASHE
Full-Time Equivalent Employees by Age Group	48	GB	2011-2022	n/a	ASHE
Full-Time Equivalent Employees by Skill Group	48	GB	2011-2022	n/a	ASHE
Real Median Hourly Earnings	48	GB	2011-2022	2022	ASHE
Real Median Hourly Earnings by Sex	48	GB	2011-2022	2022	ASHE
Real Median Hourly Earnings by Age Group	48	GB	2011-2022	2022	ASHE
Real Median Hourly Earnings by Skill Group	48	GB	2011-2022	2022	ASHE

ASHE = ONS Annual Survey of Hours and Earnings; ABI = NISRA Annual Business Inquiry.
GB = Great Britain excluding Northern Ireland; NI = Northern Ireland only.

The seven supporting sheets are:

1. Employees Markers: Flags indicating adjustments made to cells in the 'Full-Time Equivalent Employees' GB dataset.
2. Employees Markers – Sex: Flags for 'Full-Time Equivalent Employees by Sex'
3. Employees Markers - Age Group: Flags for 'Full-Time Equivalent Employees by Age Group'
4. Employees Markers - Skill Group: Flags for 'Full-Time Equivalent Employees by Skill Group'
5. Unweighted Employee Counts
6. Unweighted Firm Counts
7. CPIH Deflators

5.3 Industry coverage

Table 10: Industry coverage of business and employee job outputs

SIC Division(s)	Description
02-03	Forestry and logging; Fishing and aquaculture
05-09	Mining and quarrying of coal, oil and gas, metal ores, and other support activities
10	Manufacture of food products
11-12	Manufacture of beverages and tobacco products
13-15	Manufacture of textiles, apparel, leather and related products
16-17	Manufacture of wood, paper and related products
18	Printing and reproduction of recorded media
19-21	Manufacture of coke, refined petroleum, chemicals and pharmaceuticals
22	Manufacture of rubber and plastic products
23	Manufacture of other non-metallic mineral products
24-25	Manufacture of basic metals and fabricated metal products
26-27	Manufacture of electronic, optical and electrical equipment
28	Manufacture of machinery and equipment
29-30	Manufacture of motor vehicles and other transport equipment
31-32	Manufacture of furniture and other manufacturing
33	Repair and installation of machinery and equipment
35	Electricity, gas, steam and air conditioning supply
36-39	Water supply, sewerage, waste management and remediation
41-43	Construction of buildings, civil engineering and specialised construction
45	Sale and repair of motor vehicles and motorcycles
46	Wholesale trade (excluding motor vehicles)
47	Retail trade (excluding motor vehicles)
49	Land transport and transport via pipelines
50-53	Water and air transport, and support transport services
55	Accommodation services
56	Food and beverage service activities
58	Publishing activities
59-60	Motion picture, video, TV production and broadcasting
61-63	Telecommunications, IT and information services
68	Real estate activities
69.1	Legal activities
69.2	Accounting, bookkeeping and tax consultancy
70	Head offices and management consultancy activities
71	Architectural, engineering and technical testing activities
72	Scientific research and development
73	Advertising and market research
74-75	Design, photography, translation, consulting and veterinary services
77	Renting and leasing of equipment and assets
78	Employment placement and recruitment activities
79	Travel agency and tour operator services
80	Security and investigation activities
81	Building services and landscape activities
82	Business support service activities
85	Education
86	Human health activities
87-88	Residential care and social work activities
90-95	Arts, entertainment and recreation
96	Other personal service activities
EXCLUDED	Data included in ABS or ASHE but excluded from our outputs.

Our industry inclusions and exclusions broadly mirror the coverage of the ONS Annual Business Survey (ABS).

Parts of 'Agriculture' (SIC 01.6-01.7) and parts of 'Financial and insurance activities' (SIC 65.1-65.2) are surveyed by the ABS but aggregated into our 'Excluded' category due to low sample sizes and/or volatility of the data. This category is retained to support replicability of our results.

ABS does not cover own-use output by households, such as the concept of 'imputed rent'. Users examining the real estate industry using National Accounts data should note that our indicators will not include the imputed rent component, sometimes labelled as industry 68IMP or similar.

ABS does not sample 'Public administration and defence' (SIC 84), 'Medical and dental practice activities' (SIC 86.2), 'Activities of households as employers; undifferentiated goods and services producing activities of households for own use' (SIC 97-98), Activities of extraterritorial organisations and bodies (SIC 99), and the remaining parts of 'Agriculture' and 'Financial and insurance activities'.

SIC 01, 64-66, 85 and 97-99 are thus either entirely excluded from our dataset or embodied within the industry labelled 'Excluded', whereas the 'Human health activities' industry (SIC 86) has partial coverage only.

5.4 Public sector coverage

The ABS samples public sector units for all 'non-public sector industries'. That is, public sector units are not sampled in industries 85-88, while industry 84 is wholly excluded. This means that industries 85-88 represent non-public sector units only in our dataset.

5.5 Export coverage

We use the ONS Annual Business Survey and ONS International Trade in Services survey to classify firms as either exporters or non-exporters in a given year. However, these data sources do not capture all forms of exports. Notably, non-resident expenditure (e.g. spending by tourists) is not captured. Caution should therefore be taken when interpreting industries with high levels of non-resident expenditure, e.g. Air transport, Accommodation services etc.

5.6 Detailed indicator information

5.6.1 Business indicators

Output at Basic Prices

Description: Approximate Output at Basic Prices

Units: £ millions

Source variable(s): Annual Business Survey, wq614. Derived variable – details available via "ABS Variable Details, 2017-2022" spreadsheet on the UK Data Service.

Notes:

- Output is not the same as turnover, although they are closely related.
- Approximate output is not identical to output. See chapter 4.2 What is 'Approximate' Output, GVA and Intermediate Consumption etc? for more information.
- Within the structure of the ABS survey and derived variables, ABS documentation states:

$$Approximate\ Output\ at\ Basic\ Prices \approx Approximate\ Intermediate\ Consumption\ at\ Purchaser's\ Prices + Approximate\ Gross\ Value\ Added\ at\ Basic\ Prices$$

- Basic prices measure the value of output received by the producer, excluding any taxes on products (e.g. VAT) and including any subsidies on products.

Total Purchases

Description: Total purchases of energy, goods, materials and services.

Units: £ millions

Source variable(s): Annual Business Survey, wq499.

Notes:

- Excludes purchases charged to the capital account (see Net Capital Expenditure).

Intermediate Consumption at Purchaser's Prices

Description: Approximate Intermediate Consumption at Purchaser's Prices

Units: £ millions

Source variable(s): Annual Business Survey, various.

$wq570 - wq433 - wq479$ for Motor Trades forms.

$wq570 - wq433$ for Wholesale, Retail, Catering, and Property forms.

$wq570$ for remaining forms.

Where $wq570$ is 'Approximate intermediate consumption at purchaser's prices', $wq433$ is 'Services bought for resale without further processing', and $wq479$ is ' Expenditure on parts used solely in repair and servicing activities'.

Notes:

- Approximate intermediate consumption is not the same as intermediate consumption. See chapter 4.2 What is 'Approximate' Output, GVA and Intermediate Consumption etc? for more information.
- Within the structure of the ABS survey and derived variables, ABS documentation states:

$$Approximate\ Output\ at\ Basic\ Prices \approx Approximate\ Intermediate\ Consumption\ at\ Purchaser's\ Prices + Approximate\ Gross\ Value\ Added\ at\ Basic\ Prices$$

  However, this equation was found not to hold for industries relating to the Motor Trades, Wholesale, Retail, Catering and Property forms. We have adjusted ABS variable $wq570$ to handle the resale of services without further processing for various industries and expenditure on parts used solely in repair and servicing and motor trades. This avoids otherwise overstating estimates of Approximate Intermediate Consumption.

- Data for services industries for years 2011-2013 are not included in the ABS dataset provided to researchers and so are calculated using the equation above.
- Purchaser's prices measure the amount paid by the purchaser to take delivery of a good or service, excluding deductible VAT. Unlike basic prices, which reflect the amount receivable by the producer, purchaser's prices include taxes on products and trade and transport margins.

Net Capital Expenditure

Description: Net Capital Expenditure, excluding Research & Development

Units: £ millions

Source variable(s): Annual Business Survey, $wq600 - wq699$ where $wq600$ is the total value of capital asset acquisitions and $wq699$ is the total value of capital asset disposals. These are derived variables – details available via "ABS Variable Details, 2017-2022" spreadsheet on the UK Data Service.

Notes:

- Following discussions with the ONS, we chose not to use the GFCF variable provided in the researcher dataset due to reliability concerns.

- We instead constructed an indicator we named 'Net Capital Expenditure'. This shares many similarities to what would be otherwise named 'Approximate Gross Fixed Capital Formation' but with some differences to the National Accounts concept. Further discussion can be found in 2.1 Annual Business Survey.
- Notable exclusions compared to an 'Approximate Gross Fixed Capital Formation' concept are:
    - Expenditure on business R&D and military weapons systems are not captured. These have been part of the National Accounts concept since Blue Book 2014.
    - Investment in residential buildings.
- Notable inclusions compared to an 'Approximate GFCF' concept are:
    - Land acquisition
    - Purchases of used buildings and structures
    - Other non-produced assets such as goodwill etc.
- Industry breakdowns are on the basis of the firm making the acquisition or disposal. This can differ to GFCF sources such as Supply and Use Tables where the industry breakdown is on the basis of the industry undertaking the capital formation. For example, consider that the industry of a firm spending money to upgrade a non-residential building may differ from the industry of the construction firm that performs the activity.

GVA at Basic Prices

Description: Approximate Gross Value Added at Basic Prices

Units: £ millions

Source variable(s): Annual Business Survey, wq613. Derived variable – details available via "ABS Variable Details, 2017-2022" spreadsheet on the UK Data Service.

Notes:

- Approximate output is not identical to output. See chapter 4.2 What is 'Approximate' Output, GVA and Intermediate Consumption etc? for more information.
- Within the structure of the ABS survey and derived variables, ABS documentation states:

$$\begin{aligned} &Approximate\ Output\ at\ Basic\ Prices \\ &\quad \approx\ Approximate\ Intermediate\ Consumption\ at\ Purchaser's\ Prices \\ &\quad + Approximate\ Gross\ Value\ Added\ at\ Basic\ Prices \end{aligned}$$

- Basic prices value GVA excluding taxes on products and including subsidies on products.

Direct Sales to Households

Description: Value of sales of goods direct to the general public for personal or household use, including VAT.

Units: £ millions

Source variable(s): Annual Business Survey, wq345

Notes:

- This question is only included in the Retail and Catering surveys which relate to industries 47 and 55-56 respectively. However, we have omitted results for industries 55-56 to avoid misleading users. The Catering form asks respondents to exclude activities such as 'income received from the letting of accommodation', 'Sales of food and drink as a catering activity (including staff canteens, take-away food and bar sales)', and 'The sale of food and drink for consumption on the premises'. As these exclusions represents some of the primary activities of industries 55-56, we have omitted results for this industry to avoid confusion.
- Notable exclusions are 'Sales, repair and maintenance of motor vehicles, motorcycles, their parts and accessories' and 'Sales of services (eg holidays, cinema and other tickets, membership fees)'.

Change in Stocks

Description: Change in stocks over the year

Units: £ millions

Source variable(s): Annual Business Survey, wq272. Derived variable – details available via "ABS Variable Details, 2017-2022" spreadsheet on the UK Data Service.

Notes:

- A positive value indicates an increase in stocks over the year. A negative value indicates a reduction in stocks over the year.
- The National Accounts concept uses data from the Quarterly Stocks Inquiry (QSI) constrained to annual totals from the ABS. QSI data is also adjusted to remove the impact of price differences on end year stocks compared to start year stocks whereas no such adjustments are made to ABS data. Differences will therefore exist between our indicator and National Accounts data.
- We removed data for the year 2020 due to very large differences between the industry totals of our estimates and published ABS data. We are unsure what has caused this issue but we are aware that smoothing can sometimes occur in National Accounts when there are concerns about the reliability of survey

returns. It is possible that the combination of the pandemic and the UK leaving the European Union resulted in smoothing of the published series, but we cannot provide any evidence for or against this hypothesis.

Employment Costs

Description: Total employment cost

Units: £ millions

Source variable(s): Annual Business Survey, wq252. Derived variable – details available via "ABS Variable Details, 2017-2022" spreadsheet on the UK Data Service.

Notes:

- Differs from 'Gross Wages' primarily by capturing redundancy payments, employers' National Insurance contributions, employers' contributions to pension funds.
- Excludes employment costs that were reimbursed from government sources due to the pandemic (e.g. furlough payments).

Gross Wages

Description: Expenditure on gross wages and salaries in cash or kind to all employees.

Units: £ millions

Source variable(s): Annual Business Survey, wq446.

Notes:

- Including payments in kind, overtime, bonuses, commissions, payments to those temporarily absent (e.g. holiday, sick or parental leave.
- Excluding payments to working proprietors/partners/executive directors not in receipt of a regular salary, fee or commission, payments to subcontractors, payments to employment agencies for the services of agency staff, National Insurance contributions and contributions to other pension and welfare schemes.

Taxes less Subsidies on Production

Description: Approximate Taxes less Subsidies on Production

Units: £ millions

Source variable(s): Annual Business Survey, various. Taxes on production is the sum of $wq412$ ('Amounts payable in national non-domestic (business) rates') and $wq431$ ('Amounts payable for vehicle excise duty'). Subsidies on production is equal to $wq432$ ('Subsidies received under The Work Programme; subsidies related to the coronavirus pandemic eg the Coronavirus Job Retention Scheme (CJRS), Job Support Scheme (JSS)').

Notes:

- National Accounts measures use administrative tax and subsidy data in place of ABS-sourced data. The ABS survey also does not cover subsidies on rail transport and tax credits on research and development activities. ABS collects information on the amount of tax paid during the year whereas National Accounts tax data is on an accrual basis.
- Includes subsidies related to the pandemic, e.g. Coronavirus Job Retention Scheme etc.
- We have removed data for years 2020 and 2021 due to large differences between the industry totals of our outputs and those of the ONS Supply and Use Tables. We note that 2022 data also appears to be lower quality than pre-pandemic years but we have not deemed the quality to be sufficiently low to warrant removal.

Purchases from UK Sources

Description: Purchases from UK Sources

Units: £ millions

Source variable(s): Annual Purchases Survey, sum of q5010 ('Total expenditure on energy, water and waste in the UK'), q7010 ('Total expenditure on services in the UK'), q9010 ('Total expenditure on goods, materials and related services in the UK'). After combining with the related international variables, these have been constrained to q5000, q7000 and q9000.

Notes:

- Covers private sector only.
- Excludes expenditure on staff and capital investment.
- This indicator refers to the location of purchases rather than considering whether there has been a change of economic ownership between residents and non-residents.

Purchases from International Sources

Description: Purchases from International Sources

Units: £ millions

Source variable(s): Annual Purchases Survey, sum of q5020 ('Total expenditure on energy, water and waste outside the UK'), q7020 ('Total expenditure on services outside the UK'), q9020 ('Total expenditure on goods, materials and related services outside the UK'). After combining with the related UK variables, these have been constrained to q5000, q7000 and q9000.

Notes:

- Covers private sector only.
- Excludes expenditure on staff and capital investment.
- There are conceptual differences between 'purchases from international sources' and 'imports'. Imports are concerned with the change of economic ownership between residents and non-residents, rather than simply cross-border transactions. Purchases also do not capture spending by residents abroad, among other conceptual differences.

Purchases by Product 2015-18

Description: Expenditure on intermediate consumption by product (CPA code)

Units: £ millions

Source variable(s): Annual Purchases Survey, sum of variables named in pattern q*010_p****.

Notes:

- Covers private sector only.
- Excludes expenditure on staff and capital investment.

Purchases by Product 2021-22

Description: Expenditure on intermediate consumption by product (CPA code)

Units: £ millions

Source variable(s): Annual Purchases Survey, sum of variables named in pattern q*010_p****.

Notes:

- Covers private sector only.

- Excludes expenditure on staff and capital investment.

6. Future extensions

There are several natural extensions to this work. The first is to embed the dataset more directly within existing modelling and analytical frameworks. The industry-by-export-status structure is designed to support this kind of use, particularly in input-output analysis, computable general equilibrium modelling and related approaches that require consistent industry-level economic indicators. In a future report, we will publish results from input-output modelling in which UK industries are disaggregated into exporter and non-exporter groups.

A second extension would be to develop richer trader classifications. The current dataset distinguishes firms by export status, but the same broad approach could be used to investigate import status, or to distinguish between goods exporters and services exporters. For services trade, future work could also examine whether firms exporting to different major trading partners, such as the EU or the United States, differ in their economic characteristics.

Further refinements could also consider alternative definitions of exporter status. In this report, firms are classified as exporters where there is evidence of any goods or services exports. Future work could test threshold-based definitions for services exporters, such as classifying firms as exporters only where services export values exceed a specified level or account for a given share of turnover. This would allow users to distinguish firms with incidental export activity from those with a more substantial propensity to export.

There is also scope to link the trader classification dataset to additional ONS microdata sources. For example, linking to datasets such as UK Manufacturers' Sales by Product (PRODCOM) or the Annual Survey of Goods and Services (ASGS) would support analysis of whether exporters and non-exporters differ in the goods and services they produce. These extensions would move beyond aggregate industry indicators and provide a richer picture of production patterns within exporter and non-exporter groups.

Finally, future work could refine the linkage and imputation methods used in the current dataset. One possible improvement would be to further increase the match rate when mapping ASHE enterprise references to reporting units by assigning unresolved one-to-many cases to the reporting unit with the largest employment. This would improve coverage, but it would also introduce an additional assumption, so the effect on results would need to be assessed carefully before adoption.

7. Conclusion

This report has documented the construction of a new dataset that disaggregates selected UK business and employee indicators by industry, year and firm export status. The dataset brings together several secure microdata sources to classify firms as exporters or non-exporters and to link these classifications to economic and employee indicators.

The dataset fills a clear gap in published UK economic statistics. While many indicators are available by industry, and some information is available on trade and productivity by firm export status, there is limited published data showing how exporters and non-exporters differ across a broader range of economic and employee indicators. By providing time series outputs for many indicators at a detailed industry level, the dataset gives users a new basis for descriptive analysis and for incorporating exporter/non-exporter distinctions into industry-level analytical frameworks.

The outputs should be treated as descriptive and experimental. They should not be interpreted as estimates of the causal effect of exporting, and they should not be expected to align exactly with National Accounts or other official statistics. Users should therefore read the data alongside the definitions, coverage notes and interpretation guidance set out in this report.

Future work could build on the dataset in several ways. The industry-by-export-status structure is designed to support integration with input-output analysis, computable general equilibrium modelling and related approaches. Further extensions could also consider import status, separate goods and services exporter classifications, trade-partner groupings, or links to additional microdata sources. These extensions would help users examine the role of internationally trading firms in the UK economy in more detail.

References

Black, J., Catalano, A., Crummey, C., & Spowage, M. (2023, 10 19). *Characteristics of exporting firms in Northern Ireland*. Retrieved from ADR Northern Ireland: https://www.adruk.org/news-publications/publications-reports/publication/data-insight-characteristics-of-exporting-firms-in-northern-ireland/

European Union. (2008, 12 11). *Council Regulation (EEC) No 696/93 of 15 March 1993 on the statistical units for the observation and analysis of the production system in the Community*. Retrieved from European Union: http://data.europa.eu/eli/reg/1993/696/oj

Forth, J., Bryson, A., Phan, V., Ritchie, F., Singleton, C., Stokes, L., & Whittard, D. (2025). The Representativeness of the Annual Survey of Hours and Earnings and Its Implications for UK Wage Policy. *British Journal of Industrial Relations, 64*(1), 21-36. doi:https://doi.org/10.1111/bjir.70010

NISRA. (2026). *ABI sample coverage*. Retrieved from Northern Ireland Statistics and Research Agency: https://www.nisra.gov.uk/statistics/annual-business-inquiry/abi-sample-coverage

Office for National Statistics. (2014). *A Comparison between Annual Business Survey and National Accounts Measures of Value Added*. Retrieved from Office for National Statistics.

Office for National Statistics. (2015, April). Annual Business Survey (ABS) Metadata. *3*. Retrieved from https://datacatalogue.ukdataservice.ac.uk/studies/study/7451#documentation

Office for National Statistics. (2022, 04 27). *Comparison of labour market data sources*. Retrieved from Office for National Statistics: https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/methodologies/comparisonoflabourmarketdatasources

Office for National Statistics. (2023, 06 13). *Annual Purchases Survey - UK*. Retrieved from ONS SRS Metadata Catalogue: https://doi.org/10.57906/zd3q-d088

Office for National Statistics. (2025). *Annual Business Survey - GB*. Retrieved from ONS SRS Metadata Catalogue: https://doi.org/10.57906/ks2s-qx24

Office for National Statistics. (2025, 03 19). *Annual International Trade in Services - UK*. Retrieved from ONS SRS Metadata Catalogue: https://doi.org/10.57906/d42y-c117

Office for National Statistics. (2025). *Annual Survey of Hours and Earnings - GB*. Retrieved from ONS SRS Metadata Catalogue: https://doi.org/10.57906/x25d-4j96

Office for National Statistics. (2025). *Business Structure Database - UK*. Retrieved from ONS SRS Metadata Catalogue: https://doi.org/10.57906/7kh0-0910

Office for National Statistics. (2026, 05 26). *Non-financial business economy, UK: Sections A to S*. Retrieved from Office for National Statistics: https://www.ons.gov.uk/businessindustryandtrade/business/businessservices/datasets/uknonfinancialbusinesseconomyannualbusinesssurveysectionsas

Office for National Statistics. (2026). *SOC 2020 Volume 1: structure and descriptions of unit groups*. Retrieved from Office for National Statistics: https://www.ons.gov.uk/methodology/classificationsandstandards/standardoccupationalclassificationsoc/soc2020/soc2020volume1structureanddescriptionsofunitgroups#main-areas-of-revision-from-soc-2010-to-soc-2020

Appendix

Table 11: Count of unweighted reporting units matched between the Annual Business Survey and International Trade in Services Survey, 2011-2022

Result	ABS only	ITIS only	ABS & ITIS
2011	45,627	4,320	2,152
2012	45,982	4,527	2,265
2013	37,117	6,833	10,067
2014	37,887	14,191	11,120
2015	34,790	14,656	10,862
2016	27,774	15,603	17,707
2017	30,908	14,995	15,817
2018	41,873	11,499	4,620
2019	31,577	12,401	3,975
2020	32,249	13,663	4,282
2021	37,436	10,477	5,934
2022	39,661	11,741	7,556
Average	**36,907**	**11,242**	**8,030**

Table 12: Count of unweighted reporting unit-year pairings classified by export status, 2011-2022

Classification	Goods Export Status	Services Export Status	Overall Export Status
Non-exporter	367,917	498,183	322,463
Exporter	104,136	174,191	249,474
Unknown	202,091	1,770	102,207
Total	**674,144**	**674,144**	**674,144**

Count of goods export status and services export status will not sum to overall export status as information for overall status is only available for non-exporters if both goods and services information exists for that year. 'Unknown' represents reporting units which have not answered the corresponding questions.

Table 13: Count of unweighted reporting unit-year pairings classified by export status after imputation, 2011-2022

Classification	Goods Export Status	Services Export Status	Overall Export Status
Non-exporter	2,841,123	3,404,882	2,542,218
Exporter	450,417	638,986	960,024
Unknown	758,268	5,940	547,566
Total	**4,049,808**	**4,049,808**	**4,049,808**

Table 14: Count of unweighted reporting unit-year pairings in the Annual Business Survey classified by post-imputation export status, 2011-2022

Classification	Goods Export Status	Services Export Status	Overall Export Status
Non-exporter	394,606	441,025	337,380
Exporter	111,580	96,588	172,962
Unknown	33,052	1,625	28,896
Total	**539,238**	**539,238**	**539,238**

Table 15: Count of enterprise units with all employees matched to reporting unit, some employees, or no employees, Annual Survey of Hours and Earnings, 2015-2022

Year	Full match	Partial match	No match	Total enterprises
2015	55,473	1,087	3,032	**59,592**
2016	54,243	1,018	2,895	**58,156**
2017	54,345	1,048	3,005	**58,398**
2018	59,077	1,140	3,391	**63,608**
2019	54,857	1,088	3,007	**58,952**
2020	46,589	728	2,584	**49,901**
2021	44,757	704	2,543	**48,004**
2022	44,568	770	2,637	**47,975**
Average	**51,739**	**948**	**2,887**	**55,573**

Table 16: Count of unweighted employee jobs matched between the Annual Survey of Hours and Earnings and Business Structure Database, 2015-2022

Year	Matched Employees	Total Employees	Match rate
2015	171,742	183,475	93.6%
2016	167,644	179,022	93.6%
2017	168,143	178,943	94.0%
2018	168,998	180,185	93.8%
2019	166,744	177,930	93.7%
2020	131,073	138,385	94.7%
2021	128,533	136,176	94.4%
2022	132,804	141,675	93.7%
Average	**154,460**	**164,474**	**93.9%**