

Data Definitions for Core Database and Monthly Updates

We assume no responsibility or liability for any errors or omissions in the data supplied.

Companies only need to declare their turnover if it's more than £36 million.
Accordingly, we only concentrate on extracting key financial data from the balance sheets.
However, where P&L data is available we do try to extract it.

XBRL, which stands for eXtensible Business Reporting Language, created in 1998, is a format that defines or 'tags' individual data items in financial statements in a standardized format. XBRL was developed specifically to communicate information between businesses and other users of financial information.

In 2005, the Securities and Exchange Commission started accepting voluntary XBRL submissions, and since 2009, all financial reporting is submitted using XBRL.

In 2008, UK companies were able to file accounts at Companies House using iXBRL or inline eXtensible Business Reporting Language, an open standard that enables a single document to provide both human-readable and structured, machine-readable data. HMRC requires that all Company Tax Returns for periods ended after 31 March 2010 to be submitted online using iXBRL.

A taxonomy is a dictionary that links each tag to the accounting or tax concept that it identifies. The iXBRL taxonomies correspond with accounting standards such as FRS, IFRS and GAAP.

There are many taxonomies, all using different tags; so cash can be tagged as 'CashBank', 'CashInHand', 'CashOnHand', 'CashBankOnHand', 'CashBankInHand' or 'CashCashEquivalents' depending on the taxonomy. Likewise debtors can be 'Debtors', 'TradeDebtors', 'DebtorsDueWithinOneYear', 'TradeOtherReceivables' or 'TradeDebtorsTradeReceivables'.

In addition, the tags do not always denote a unique data item; thus, the tag 'Equity' can refer to called up share capital, share premium account, revaluation reserve, capital redemption reserve, other reserves, retained earnings or shareholders funds or most often, all of them.

Although there are thousands of different tags used by the various taxonomies, as far as filings at Companies House are concerned, just over 4,300 iXBRL tags are actually in use.

Notwithstanding, the UK does not follow a structured format for Companies House filing; you can file accounts data virtually in any order that is reasonably sensible. The French, on the other hand, use the 'PCG' (Plan Comptable Général) where there are well defined boxes, in a well defined order, in which balance sheet data is inserted.

There are inherent problems in extracting key data from UK balance sheets where sub totals are not always tagged, some tags are missing or where tags can have several meanings.

<ul style="list-style-type: none"> Property Plant and Equipment Fixed Assets Creditors Debtors Cash Current Assets Net Current Assets Long Term Creditors Capital Employed Total Assets 	<p>The 10 key data tags extracted should cover most uses of the financial data. Certainly, enough to establish the company's health. We only extract balance sheet data where the tag is present. To be fair to all companies, we only aggregate data where necessary otherwise we do not manipulate the data. Sometimes, there may be gaps in the years and some tags may be missing in successive years. We do not squirrel around in the 'Notes' for additional data; so do not attempt to fine tune the data.</p>
---	---

The core database is just over 27 million records. Each company's financial data is held as one record in a CSV file. Each record consists of 68 items separated by commas.

The core database consists of separate files for years 2008 to 2022 and then separate monthly files. All files are zipped and sent by e-mail. Zipped or compressed files take up less storage space and can be transferred to other computers more quickly than uncompressed files.

In Microsoft, to unzip files, right-click the zip file and select 'Extract all'; for Linux, use the unzip package; for MacOS, right-click the zip file and select 'Open'.

A comma-separated values (CSV) file is a delimited text file that uses commas to separate values. Each line of the file is a data record. Thus, each record consists of one or more fields, separated by commas. The use of the comma as a field separator is the source of the name for this file format.

Our CSV files cannot be opened directly by spreadsheet programs like Excel as they are too large. If you open a large CSV in a spreadsheet program, then close it, you will likely lose that data that could not be loaded into memory. You can use a CSV splitter program to make large files smaller. However, most users will be using a SQL database:

Microsoft SQL Server	use either SQL Server Integration Services (SSIS) package or PowerShell script.
Oracle SQL Server	for a new table, on the Database menu, click Import Data; for an existing table, in Database Explorer, right-click a table and select Import Data.
MySQL or MariaDB	use LOAD DATA INFILE.

Each record appears as:

```
00178226,2021-07-31,,,,,,,,,,,,,109666,,,,,,,,,26153,,,,,,,,,26153,,135819,,,135819,,,,,135819,26153,0,GBP
00179710,2021-10-31,,,,,,,,,,,,,4500,,,,,,,,,537,,,,,19172,19709,,,,,1263,,,,,18446,,22946,,,,,,,,,18446,0,GBP
00180478,2021-04-30,,,,,,,,,,,,,12810,,,,,,,,,18050,,,,,1233,19283,,,,,1014,,,,,1400,18269,,31079,,,,,29679,,,,,18269,0,GBP
00366930,2021-06-30,[.],,159384,,,,,,,,,6335,,,,,54426,,,,,-48091,7500,111293,,,,,103793,,,,,-48091,1,GBP
00375364,2021-04-30,[.],,735569,,,,,,,,,1898,,,,,86923,88821,,,,,50575,,,,,38246,,773815,,,,,,,,,38246,0,GBP
00376022,2021-06-30,[.],,127568,,,,,86199,,1545933,,,,,132760,1764892,,,,,940215,,,,,824677,3801,952245,,,,,24238,924206,,,,,824677,39,GBP
00380316,2021-05-31,[.],,1800000,,,,,31420,,,,,152536,183956,,,,,35022,,,,,148934,,1948934,,,,,-104038,1844896,,,,,148934,2,GBP
00385534,2021-04-05,[.],,999153,,,,,293111,111,71078,,,,,496020,860320,,,,,128207,,,,,732113,,1731266,,56265,1675001,,,,,1859473,732113,5,GBP
```

Some previous month's records can occasionally be duplicated in later months.

Where a company declares itself as 'dormant' we zeroise all the data to prevent people wasting their time on such companies. Sometimes, where there is large cash figure, we will include all balance sheet data.

Year	Total Balance Sheets	Each year's CSV file is zipped		XBRL accounts filed in the last 12 months			
		Year	Zip Size in bytes	Month	Total	New	Updated
2008	185,187	2008	3,988,425	Jan 2022	249,408	36,851	212,557
2009	344,403	2009	9,0341,85	Feb 2022	223,034	38,304	184,730
2010	585,123	2010	17,024,229	Mar 2022	296,103	51,119	244,984
2011	792,993	2011	23,735,332	Apr 2022	231,997	41,782	190,215
2012	1,032,997	2012	31,252,112	May 2022	255,996	45,776	210,220
2013	1,329,547	2013	39,869,220	Jun 2022	265,760	45,766	219,994
2014	1,586,434	2014	47,281,121	Jul 2022	250,863	43,780	207,083
2015	1,974,565	2015	58,388,979	Aug 2022	248,974	43,678	205,296
2016	2,345,447	2016	69,800,293	Sep 2022	336,535	62,707	273,828
2017	2,342,852	2017	67,074,758	Oct 2022	271,842	55,356	216,486
2018	2,590,818	2018	73,343,030	Nov 2022	285,182	47,772	237,410
2019	2,754,058	2019	78,334,691	Dec 2022	206,076	18,456	187,620
2020	2,597,354	2020	75,604,866				
2021	3,526,911	2021	103,831,273				
2022	3,121,770	2022	92,775,893				
2008-2022	27,110,459						

01	CompanyNumber	‡Company Number
02	BalanceSheetDate	Balance Sheet Date yyyy-mm-dd
03	Turnover	Turnover
04	Costsales	Cost of Sales
05	Grossprofitloss	Gross Profit or Loss
06	Distributioncosts	Distribution Costs
07	Administrativeexpenses	Administrative Expenses
08	Increasedecreaseinstocksinventoriesfinishedgoodsworkingprogress	Change in Stock
09	Otheroperatingincomeformat2	Other Operating Income
10	Rawmaterialsconsumablesused	Raw Materials
11	Otherexternalcharges	Other External Charges
12	Staffcostsemployeebenefitexpense	Staff Costs
13	Depreciationamortisationimpairmentexpense	Depreciation
14	Furtheritemoperatingincomegainincomestatementitemcomponentoperatingprofitloss	Other Operating Income
15	Operatingprofitloss	Operating Profit or Loss
16	Totaloperatingprofitloss	Total Operating Profit or Loss
17	Incomefromotherfixedassetinvestments	Income from Other Investments
18	Otherinterestreceivablesimilarincomefinanceincome	Interest Receivable
19	Interestpayablesimilarchargesfinancecosts	Interest Payable
20	Gainlossduetoforeignexchangedifferencesrecognisedinprofitforloss	Foreign Exchange Gains or Losses
21	Otherfinanceincome	Other Financial Income
22	Profitlossonordinaryactivitiesbeforetax	Profit or Loss before Taxation
23	Taxtaxcreditonprofitforlossonordinaryactivities	Tax Credit
24	Taxrelatedtoprofitforlossonordinaryactivities	Corporation Tax on Ordinary Activities
25	Totalukcurrentcorporationtax	Total Corporation Tax
26	Profitlossonordinaryactivitiesaftertax	Profit or Loss after Taxation
27	Profitloss	Profit or Loss
28	FixedAssetsInvestments	Investments
29	InvestmentProperty	Property
30	IntangibleFixedAssets	Intangibles
31	Propertyplantequipment	Property Plant and Equipment
32	Othertangiblefixedasset	Other Tangible Fixed Assets
33	Fixedassets	Fixed Assets
34	Tangiblefixedassets	Tangible Fixed Assets
35	Noncurrentassets	Non Current Assets
36	Stock	Stock
37	CurrentAssetsInvestments	Current Assets Investments
38	Debtors	Debtors
39	Otherdebtors	Other Debtors
40	Debtors2	Additional Debtors
41	Prepayments	Prepayments
42	Cash	Cash
43	Currentassets	Current Assets
44	OtherCreditors	Other Creditors
45	Bankborrowingoverdrafts	Bank Borrowings
46	Bankloanoverdraftswithinoneyear	Short Term Bank Loans
47	Corporationtaxpayable	Corporation Tax Payable
48	Creditors	Creditors
49	CreditorsDueAfterOneYear	Long Term Creditors
50	Bankborrowingoverdrafts2	Long Term Bank Borrowings
51	Othercreditorsafteroneyear	Other Long Term Creditors
52	Accruedliabilitiesnotexpressedwithincreditorssubtotal	Accruals
53	Netcurrentassetsliabilities	Net Current Assets
54	LongTermCreditors	Long Term Creditors
55	CapitalEmployed	Capital Employed
56	Bankloanoverdraftsafteroneyear	Long Term Bank Loans
57	DeferredTax	Deferred Tax
58	Provisions	Provisions
59	NetAssets	Net Assets
60	Shareholderfunds	Shareholder Funds
61	Calledupsharecapitalnotpaidnotexpressedascurrentasset	Unpaid Share Capital
65	Totalassets	‡Total Assets
63	Totalliabilities	‡Total Liabilities
64	Propertyplantequipment2	Other Property Plant and Equipment
65	TotalAssets	‡Total Assets
66	WorkingCapital	Working Capital
67	Employees	No of Employees
68	Currency	§Currency

‡ 01 Company Number 8 digits or NC NI OC R0 SC SO plus 6 digits

§ 68 Currency ARS AUD CHF DKK EUR GBP ISK JPY NOK PKR QAR SEK USD

‡ The following three are extracted and are not computed and they normally refer to dormant or non-trading companies:
62 Total Assets 63 Total Liabilities 65 Total Assets

¶ Net Current Assets and Working Capital is a sop to those older clients using different structured data; they are, of course, the same.

Each month, we produce two profile files: active.csv and dissolved.csv. They hold approximately 3.2 million and 2.4 million records respectively. Their structure is as follows:

CompanyNumber	Company Number
check_accents	EUR=European characters present OK=otherwise
CompanyName	Full Company Name
IncorporationDate	Date of Incorporation yyyy-mm-dd
DissolutionDate	Date of Dissolution yyyy-mm-dd or null
SICCode1	first SIC code
SICCode2	second SIC code
SICCode3	third SIC code
SICCode4	fourth SIC code
PreviousName	Previous Name
CareOf	Care of
POBox	Post Office Box
AddressLine1	First Line of Address
AddressLine2	Second Line of Address
PostTown	Post Town
County	County
Country	Country
PostCode	Postcode

Companies House may still categorise a dissolved company as 'active'. Some companies that are dissolved may be administratively restored; in which case, they will move from the dissolved_companies.csv to the active_companies.csv

Some address data by registration agents or by company promoters themselves can be placed in the wrong field, commas separating parts of the address can be missing, fields can be concatenated or duplicated; and users should be aware that some addresses can be dirty and require cleaning due to sloppy input.

A company's size must satisfy at least two from annual turnover, balance sheet total and average number of employees.

Company Type	Parameters	Filing Requirements
micro-entity	turnover \leq £632,000 balance sheet \leq £316,000 employees \leq 10	micro balance sheet
small company	turnover \leq £10.2 million balance sheet \leq £5.1 million employees \leq 50	abridged balance sheet
medium-sized company	turnover \leq £36 million balance sheet \leq £18 million employees \leq 250	full balance sheet and P&L (special exemption provisions)
large company	turnover $>$ £36 million balance sheet $>$ £18 million employees $>$ 250	full balance sheet and P&L

A company qualifies for audit exemption if it has at least 2 of the following:
annual turnover \leq £10.2 million; balance sheet \leq £5.1 million and employees \leq 50.