**Examples of data analysis**

This notebook shows how to analyse the questionnaire data example from the ABCE Open Data Project.

View the dataset here: https://figshare.com/articles/dataset/BoE\_Opendata/20364429

The raw dataset (MS Excel document) of the BoE can be download here: <https://github.com/building-energy/ABCE_Open_Data_Project/blob/main/BoE%20Opendata%20by%20S%20Li/BoE%20NMG%20household%20survey%20data%20-%202004-11.xlsx>

Please be aware that this is the official data set from the official website of the Bank of England (BoE). For more information please visit: <https://www.bankofengland.co.uk/statistics/research-datasets>

View this notebook on GitHub here: <https://github.com/building-energy/ABCE_Open_Data_Project/blob/main/BoE%20Opendata%20by%20S%20Li/read-me.txt>

Instead, please click: <https://figshare.com/articles/dataset/BoE_Opendata/20364429?file=36403065>

**Setup**

In [1]:

**import** requests

**import** json

**import** pandas **as** pd

**from** matplotlib **import** pyplot **as** plt

**import** io

**Get CSV file from figshare**

This downloads the Annual Questionnaire Survey Data output CSV file directly from the Figshare data repository. Here is an example of downloading the survey data from the year of 2004.

In [2]:

csv\_download\_url**=**'https://figshare.com/articles/dataset/BoE\_Opendata/20364429?file=36403068'

response**=**requests**.**get(csv\_download\_url)

csv\_text**=**response**.**text

csv\_text**.**split('\n')[:6]

Out[2]:

['id,weight, year, sex, age, age\_grp, region, gor, mastat, hushop, hhsize, lstage, nkids, parent, ch03, ch45, ch69, ch1015, qual, jbstat , jbrgsc, hoh, jbstathoh, fihhyr2, fihhyr\_a, dfihhyr, dfihhyr\_a, dfihhyrchg, dfihhyrchg\_s, us, ustot, ustot\_a, ustot\_grp5, ustoty\_a, ustotdy\_a, ustoty\_grp5, ustot, y\_grp5, xpus, xpus\_a, xpusy\_a, xpusy\_grp5, usn, usa, usb, usc, usd, use, usg, ush, usi, usj, tenure, tenure\_grp2, tenure\_grp3, huresp, hsval, hsval\_a, reshsfall, mghave, intmg, mg1tot, mg1tot\_a, mg1tot\_grp5, mg1toty\_a, mg1totdy\_a, nequity, xpmg, xpmg\_a, xpmgy\_a, xpmgdy\_a, xpmgy\_grp5, xpmgdy\_grp5, fixmg, fixmg09, fixpmg\_a, fnxpmg\_a, fixpmg09, fnxpmg09, changepmg, changepmg\_a, fxuse1, fxuse2, fxuse3, fxuse4, fxuse5, fxuse6, pmguse1, pmguse2, pmguse3, pmguse4, pmguse5, pmguse6, pmguse7, pmguse8, xphsdf, xphpdf, billscc, xphdr1, xphdr2, xphdr3, xphdr4, xphdr5, xphdr6, xphdr7, xphdr8, xphdr9, xphdr10, xphdr11, xphdr12, xphdr13, xphdr14, xphdr15, xphdd1, xphdd2, xphdd3, xphdd4, xphdd5, xphdd6, xphdd7, xphdd8, xphdd9, xphdd10, xphdd11, xphdd12, xphdd13, xphdd14, hscntcr1, hscntcr2, hscntcr, hscrchg, saveamount, saveam, unt\_a, saving, saving11, savebcs, savebcs1, savebcs2, savebcs3, savebcs4, savebcs5, savebcs6, savebcs7, savebcs8, savebcs9, savebcs10, savebcs11, sav, ebcs12, savebcs13, desave11bcs1, desave11bcs2, desave11bcs3, desave11bcs4, desave11bcs5, desave11bcs6, save11bcs1, save11bcs2, save11bcs3, save11bcs4, save11bcs5, save11bcs6, save11bcs7, save11bcs8, save11bcs9, save11, bcs10, save11, bcs11, save11, bcs12, spend, nvesttot, nvesttot\_a, othassets, debtconc, debtconc\_act1, debtconc\_act2, debtconc\_act3, debtconc\_act4, debtconc\_act5, debtconc\_act6, debtconc\_act7, chdebtconc, pubsecty, fisc\_conc1, fisc\_conc2, fisc\_conc3, fisc\_conc4, fisc\_conc5, fisc\_conc6, fisc\_conc7, fisc\_conc8, fisc\_conc9, fisc\_act1, fisc\_act2, fisc\_act3, fisc\_act4 , fisc\_act5, fisc\_act6, fisc\_act7, fisc\_act8, fisc\_act9, fisc\_impact1, fisc\_impact2, fisc\_impact3, fisc\_impact4, fisc\_impact5, fisc\_impact6, fisc\_impact7, fisc\_impact8, fisc11\_act1, fisc11\_act2, fisc11\_act3, fisc11\_act4, fisc11\_act5, fisc11\_act6, fisc11\_conc1, fisc11\_conc2, fisc11\_conc3, fisc11\_conc4, fisc11\_conc5, fisc11\_conc6, fisc11\_conc7, fisc11\_conc8, fisc\_likact1,fisc\_likact2, fisc\_likact3, fisc\_likact4, fisc\_likact5, fisc\_likact6, forebearsec1, forebearsec2, forebearsec3, forebearsec4, forebearsec5, forebearsec6, forebearsec7, forebearsec8, forebeareffsec, brrise, forebearunsec1, forebearunsec2, forebearunsec3, forebearunsec4, forebearunsec5, forebeareffunsec, uncert,uncertch']

**Get metadata.json file from figshare**

The CSV data is in CSV on the Web (CSVW) format and so there is an additional accompanying metadata file. The code below downloads this directly from the Figshare data repository.

In [3]:

metadata\_json\_download\_url**=**' https://figshare.com/articles/dataset/BoE\_Opendata/20364429?file=36403071 '

response**=**requests**.**get(metadata\_json\_download\_url)

metadata\_text**=**response**.**text

metadata\_dict**=**json**.**loads(metadata\_text)

metadata\_dict

Out[3]:

{

"@context": ["http://www.w3.org/ns/csvw", {"@language": "en"}],

"tables": [{

"url": "Survey Data 2004.CSV",

"tableSchema": {

"columns": [

{

"name": "id",

"datatype": "integer",

"description": "Each answer has a related response id which is assigned to the specific respondant's response, so that answers for different questions can be tracked between different respondants"

},

{

"name": "weight",

"datatype": "integer",

"description": "Each respondent has the unique weight value for the whole survey"

},

{

"name": "year",

"datatype": "date",

"description": "This explains the year in which the survey was conducted"

},

{

"name": "sex",

"datatype": "string",

"description": "The gender of each respondent"

},

{

"name": "age",

"datatype": "integer",

"description": "The respondent's age of the year during which the survey was conducted"

},

{

"name": "age\_grp",

"datatype": "integer",

"description": "The age group which the respondent belongs to"

},

{

"name": "region",

"datatype": "string",

"description": "The region where the respondent comes from"

},

{

"name": "gor",

"datatype": "string",

"description": "Respondent's government office region"

},

{

"name": "mastat",

"datatype": "string",

"description": "Respondent's marutal status"

},

{

"name": "hushop",

"datatype": "string",

"description": "Check whether the respondent is a main shopper"

},

{

"name": "hhsize",

"datatype": "integer",

"description": "Number of individuals in househould"

},

{

"name": "Istage",

"datatype": "string",

"description": "Respondent's life stage"

},

{

"name": "nkids",

"datatype": "string",

"description": "Number of children in the household"

},

{

"name": "parent",

"datatype": "string",

"description": "Whether the respond is the parent of children"

},

{

"name": "ch03",

"datatype": "string",

"description": "Whether the respond has children between 0-3"

},

{

"name": "ch45",

"datatype": "string",

"description": "Whether the respond has children between 4-5"

},

{

"name": "ch69",

"datatype": "string",

"description": "Whether the respond has children between 6-9"

},

{

"name": "ch1015",

"datatype": "string",

"description": "Whether the respond has children/teenager between 10-15"

},

{

"name": "qual",

"datatype": "string",

"description": "The question about the highest degree/education of each respondent"

},

{

"name": "jbstat",

"datatype": "string",

"description": "Working status of each respondent"

},

{

"name": "jbrgsc",

"datatype": "string",

"description": "Social grade"

},

{

"name": "hoh",

"datatype": "string",

"description": "Whether the respondent is the chief income earner"

},

{

"name": "jbstathoh",

"datatype": "string",

"description": "working status of chief income earner"

},

{

"name": "fihhyr2",

"datatype": "string",

"description": "income (only asked to main shopper or chief income earner)"

},

{

"name": "fihhyr\_a",

"datatype": "string",

"description": "Midpoints of fihhyr2 from 2008 (see notes column for pre-2008)"

},

{

"name": "dfihhyr, dfihhyr\_a, dfihhyrchg, dfihhyrchg\_a",

"datatype": "string",

"description": "There is no data under these columns. For more information please refer to the questionnaire template"

},

{

"name": "us",

"datatype": "string",

"description": "Ask the respondent if the household currently owes any money on the following types of loan or credit agreement (list: usa-usj)"

},

{

"name": "ustot",

"datatype": "string",

"description": "Question about how much money in total the respondent or household currently owes on these loans"

},

{

"name": "answer\_text",

"datatype": "string",

"description": "\*\*Please see previous column descriptions for examples\*\* Each answer column should have a description which provides context for the answer (i.e. 1 is low and 10 is high etc). Also, the .JSON file should have a given datatype relevant to the answer provided (usually text or integer)."

},

{

"name": "ustot\_a",

"datatype": "string",

"description": "Midpoints of ustot after 2006, sum of individual balances on different types of unsecured in 2004 and 2005"

},

{

"name": "ustoty\_a",

"datatype": "string",

"description": "Unsecured debt to income ratio"

},

{

"name": "ustotdy\_a",

"datatype": "string",

"description": "Unsecured debt to disposable income ratio"

},

{

"name": "xpus",

"datatype": "string",

"description": "Taking into account of your unsecured debts, how much was your last monthly payment on all of the loans?"

},

{

"name": "xpus\_a",

"datatype": "string",

"description": "Midpoints of xpus. Please refere to last question for more information"

},

{

"name": "xpusy\_a",

"datatype": "string",

"description": "last monthly unsecured debt payment to income ratio"

},

{

"name": "xpusdy\_a",

"datatype": "string",

"description": "last monthly unsecured debt payment to disposable"

},

{

"name": "usn",

"datatype": "string",

"description": "Number of unsecured debt instruments held"

},

{

"name": "usa",

"datatype": "string",

"description": "Unsecured debt instrument: hire purchase"

},

{

"name": "usb",

"datatype": "string",

"description": "Unsecured debt instrument: personal loan"

},

{

"name": "usc",

"datatype": "string",

"description": "Unsecured debt instrument: credit card"

}, {

"name": "usd",

"datatype": "string",

"description": "Unsecured debt instrument: DSS social fund loan"

}, {

"name": "ude",

"datatype": "string",

"description": "Unsecured debt instrument: something else"

},

{

"name": "usg",

"datatype": "string",

"description": "Unsecured debt instrument: personal loan"

},

{

"name": "ush",

"datatype": "string",

"description": "Unsecured debt instrument: overdraft"

},

{

"name": "usi",

"datatype": "string",

"description": "Unsecured debt instrument: student loan"

},

{

"name": "usj",

"datatype": "string",

"description": "Unsecured debt instrument: store card"

},

{

"name": "tenure",

"datatype": "string",

"description": "Housing tenure group"

},

{

"name": "tenure\_grp2",

"datatype": "string",

"description": "Housing tenure group"

},

{

"name": "tenure\_grp3",

"datatype": "string",

"description": "Housing tenure group"

},

{

"name": "hsval",

"datatype": "string",

"description": "Midpoints of hsval"

},

{

"name": "hsval\_a",

"datatype": "string",

"description": "Midpoints of hsval"

},

{

"name": "reshsfall",

"datatype": "string",

"description": "Over the past year, the price of an average home has fallen by about 10%. How would your household spending on items such as clothes, leisure and groceries be affected if house prices were to fall by another 10% in the next year. Would you say.."

},

{

"name": "mghave",

"datatype": "string",

"description": "Question about if there is a mortgage or loan secured on respondent's home"

},

{

"name": "intmg",

"datatype": "string",

"description": "Type of interest rate being paid on the mortage or loan"

},

{

"name": "mg1tot",

"datatype": "string",

"description": "Roughly how much is left to pay on your current mortgage(s) and secured loan(s) on your home? (only asked to adults responsible for financial decision making)"

},

{

"name": "mg1tot\_a",

"datatype": "string",

"description": "Midpoints of mg1tot"

},

{

"name": "mg1toty\_a",

"datatype": "string",

"description": "Secured debt to income ratio"

},

{

"name": "mg1totdy\_a",

"datatype": "string",

"description": "Secured debt to disposable income ratio"

},

{

"name": "nequity",

"datatype": "integer",

"description": "If you were able to sell your home today, do you think it would be worth more than what you owe on your mortgage and any other secured loans?"

},

{

"name": "xpmg",

"datatype": "string",

"descript

This file contains more than 10 kB. To see the remaining content please download the original file.

Switch View

Switch between different file views

Hide files

Previous file

Next file

5/19

Survey Data 2004.csv-metadata.

JSON

(13.31 kB)

**Analysis: Importing data into SPSS and making analysis**

Once gaining the data with its metadata, it can be converted into an Excel document (see Figure 1).

A picture containing text, window

Description automatically generated

Figure 1: Converting data and meta data into Excel

In SPSS, the database downloaded from the Figshare can be opened to start analyse data (Figure 2). Instead, CSV files can also be opened by SPSS directly.

A picture containing graphical user interface

Description automatically generated

Figure 2: Open the downloaded dataset

The dataset presented by SPSS can be found in Figure 3.

Graphical user interface, application, table

Description automatically generated

Figure 3: Data presented by SPSS

To analyse data, for example descriptive analysis, it can be processed by SPSS (Figure 4).

A picture containing text

Description automatically generated

Figure 4: Descriptive analysis

For more information about how to use SPSS, please visit the website of Loughborough University Doctoral College.