2° Investing Initiative

PACTA for Banks Training Webinar:
Matching a loan book to physical assets in the real economy

Supported by:

Federal Ministry
for the Environment, Nature Conservation
and Nuclear Safety

based on a decision of the German Bundestag

About our funders: *This project is part of the International Climate Initiative (IKI). The Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (BMU) supports this initiative on the basis of a decision adopted by the German Bundestag. This project has also received funding from the European Union LIFE program. The views expressed here are the sole responsibility of the authors and do not necessarily reflect the views of the funders.*

What we'll cover

Matching a bank's loan book to physical assets in the real economy, using the Asset-level Data set provided by Asset Resolution and the r2dii.match software package

Agenda

- Recap of methodology & the matching process
- User resources
- Matching work flow with examples and code (r2dii.match)
  - Step 1 – Import files
    - Sector Classification Codes
  - Step 2 – match_name function
    - String Matching algorithms
  - Step 3 – Manual matching
  - Step 4 – Overwrite file
  - Step 5 – Prioritize function
- Next steps
- Q&A

Methodology Recap

Corporate lending portfolios

Alignment of loan books is benchmarked against **climate change scenarios** and the **market**

Loans are mapped to the **physical assets** in **the real economy** and their corresponding **production values**

Climate Change Scenarios

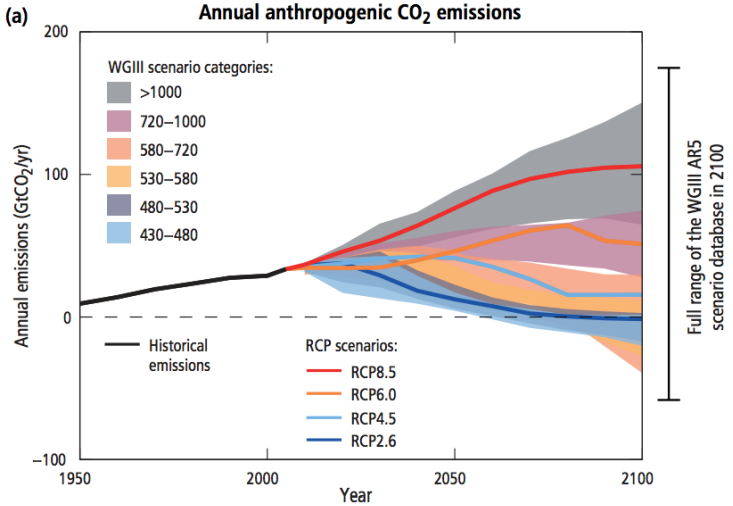
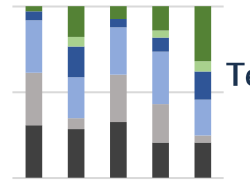
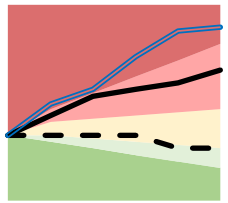


Figure SPM.5 from page 9 of the IPCC AR5 Summary for Policymakers:

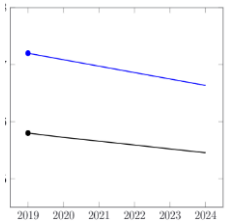
Metrics



Technology mix



Volume trajectory



Emission Intensity

Physical Assets in the Real Economy

Mapping physical assets to a bank's exposures – in practical terms

Corporate lending portfolios

Physical assets in the real economy

r2dii.match

Loan book (LBK)

	A	B	C	D	E	F
1	id_loan	id_direct_	name_dire	id_interm	name_inte	id_ultimat
2	L1	c301	BP plc	NA	NA	UP101
3	L2	c301	BP plc	NA	NA	UP101
4	L3	c301	BP	NA	NA	UP101
5	L4	c302	duke energ	NA	NA	UP102
6	L5	c303	duke energ	NA	NA	UP102

Asset Level Data (ALD)

Asset Resolution

	A	B	C	D	E	F
1	name_cor	sector	technolog	year	productio	productio
2	bp plc	oil and gas	gas	2019		GJ
3	bp plc	oil and gas	gas	2019		GJ
4	bp plc	oil and gas	gas	2019		GJ
5	bp plc	oil and gas	gas	2019		GJ
6	bp plc	oil and gas	gas	2019		GJ

Terminology / Disclaimer

- Examples given in this presentation are illustrative only
- The names of companies may be represented differently in the asset-level data set provided by Asset Resolution

Term	Abbreviation	Description
Asset Level Data	ALD	The data set to which the loan book is being matched
Loan book	LBK	The bank's corporate lending book
Direct Loan Taker	DL	The counter party that receives the loan
Ultimate Parent	UP	The owner of the counter party receiving the loan
2 Degrees Investing Initiative	2DII	The think tank behind the PACTA methodology
Asset Resolution	AR	The data provider proving the free PACTA for banks data set

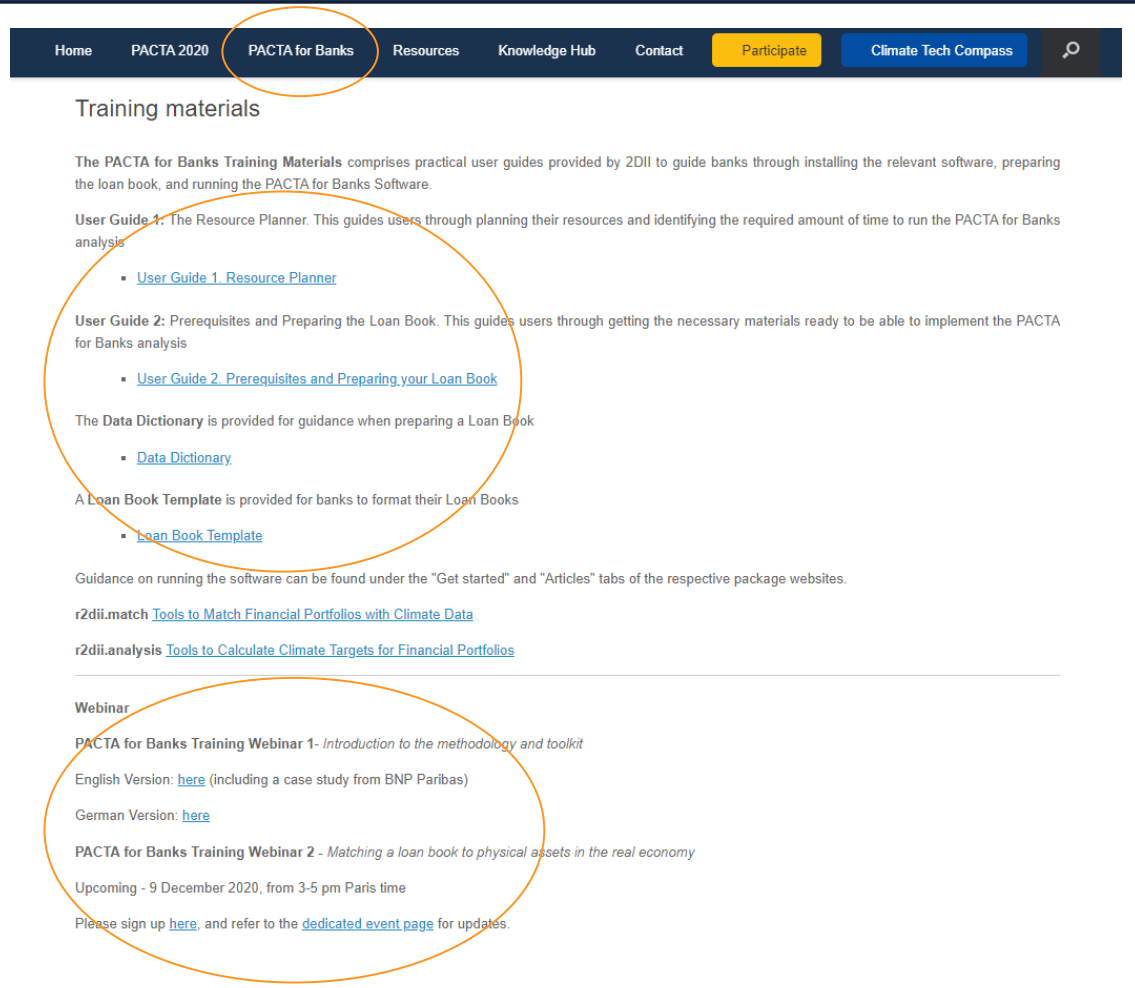
User resources: Transition Monitor website

User Guides and templates

User Guide 1 – Resource Planner
User Guide 2 - Prerequisites and Preparing your loan book
Loan Book Template

Webinars

Webinar 1 Introduction to the Methodology and tool kit
Webinar 2 (today) Matching a loan book to physical assets in the real economy
Webinar 3 (TBC) Analysis and Visualisation



User resources: r2dii.match website

Get started – instructions for using the code with reproducible examples and mock data sets

User resources: r2dii.match Website

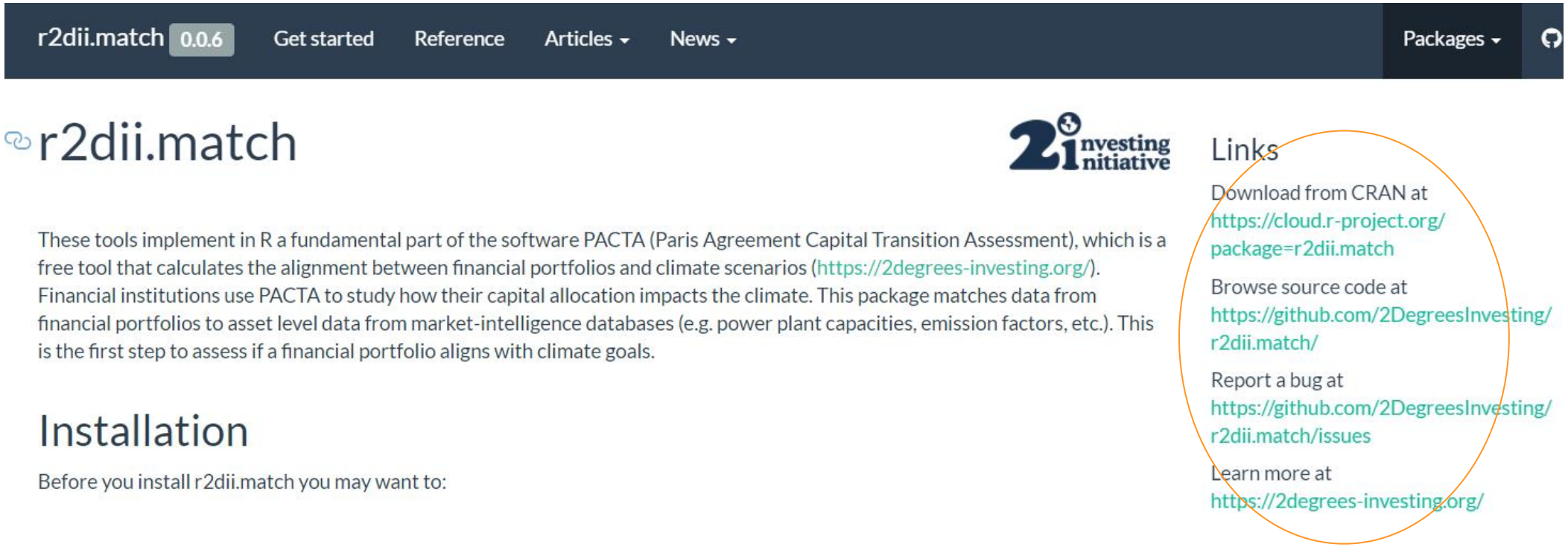
Articles cover certain issues that may arise and other useful topics. For example, matching a large loan book and calculating matching coverage.

User resources: r2dii.match Website

News contains updates on the code. For example, bugs fixes, feature enhancements, new features, etc...

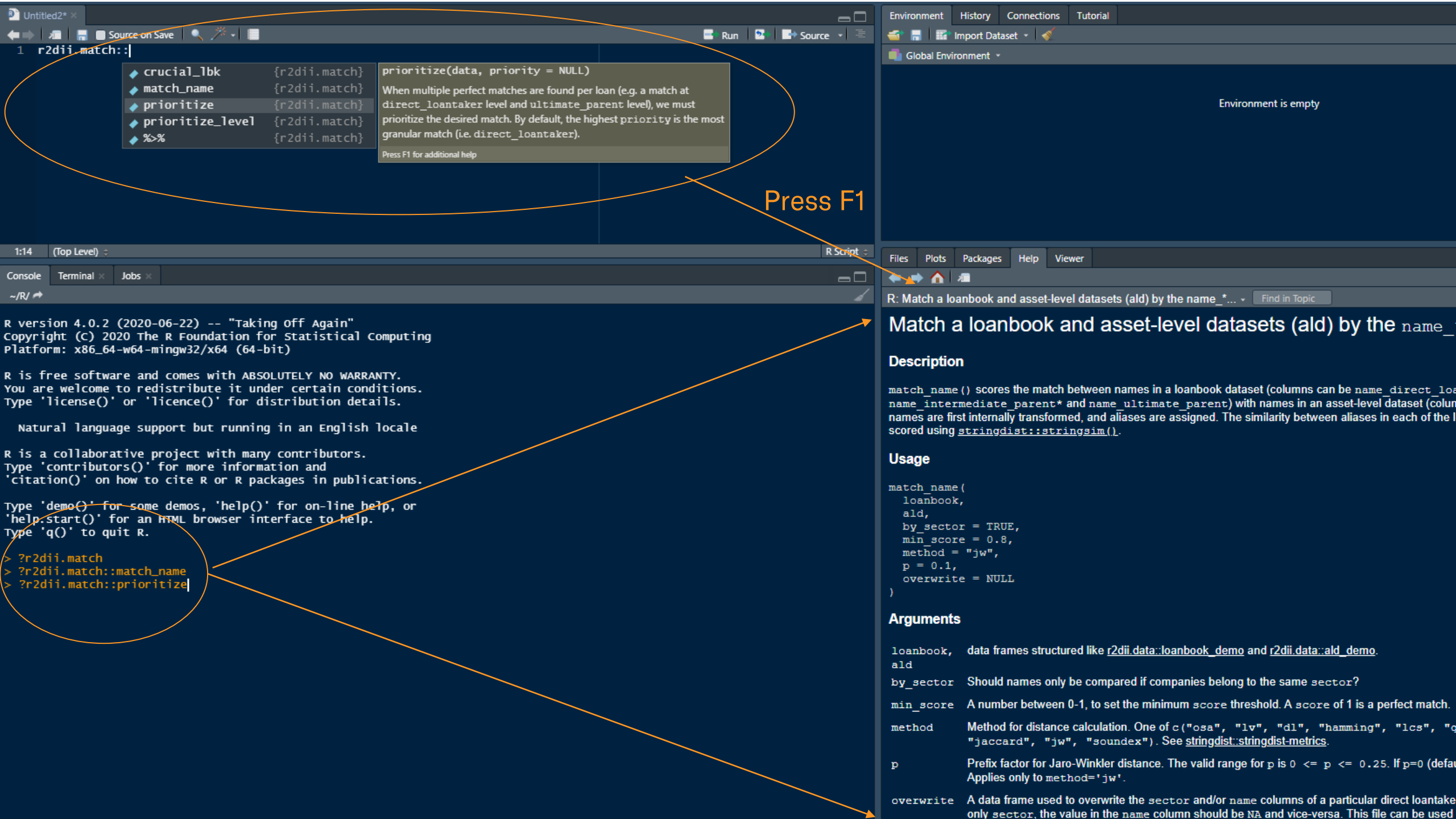
Resources available to a user: r2dii.match Website

Links – here you can find the source codes and report any bugs, new features or feature enhancement



User resources: R-Documentation

- Within the code, you can access the function documentation
- This explains how to use the functions and all the arguments
- This can be accessed by typing
  - In the script, type r2dii.match::match_name
    r2dii.match::prioritize
  - In the console type ?r2dii.match


Untitled2*
Source on Save
Run
Source
1 r2dii.match::
crucial_lbk {r2dii.match}
match_name {r2dii.match}
prioritize {r2dii.match}
prioritize_level {r2dii.match}
%>% {r2dii.match}
prioritize(data, priority = NULL)
When multiple perfect matches are found per loan (e.g. a match at direct_loantaker level and ultimate_parent level), we must prioritize the desired match. By default, the highest priority is the most granular match (i.e. direct_loantaker).
Press F1 for additional help
Press F1
1:14
(Top Level)
R Script
Environment
History
Connections
Tutorial
Import Dataset
Global Environment
Environment is empty
Console
Terminal
Jobs
~/R/
R version 4.0.2 (2020-06-22) -- "Taking Off Again"
Copyright (C) 2020 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)
R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.
Natural language support but running in an English locale
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.
> ?r2dii.match
> ?r2dii.match::match_name
> ?r2dii.match::prioritize
Files
Plots
Packages
Help
Viewer
R: Match a loanbook and asset-level datasets (ald) by the name_*...
Find in Topic
Match a loanbook and asset-level datasets (ald) by the name_
Description
match_name() scores the match between names in a loanbook dataset (columns can be name_direct_loa
name_intermediate_parent* and name_ultimate_parent) with names in an asset-level dataset (colun
names are first internally transformed, and aliases are assigned. The similarity between aliases in each of the l
scored using stringdist::stringsim().
Usage
match_name(
loanbook,
ald,
by_sector = TRUE,
min_score = 0.8,
method = "jw",
p = 0.1,
overwrite = NULL
)
Arguments
loanbook, ald
data frames structured like r2dii.data::loanbook_demo and r2dii.data::ald_demo.
by_sector
Should names only be compared if companies belong to the same sector?
min_score
A number between 0-1, to set the minimum score threshold. A score of 1 is a perfect match.
method
Method for distance calculation. One of c("osa", "lv", "dl", "hamming", "lcs", "
"jaccard", "jw", "soundex"). See stringdist::stringdist-metrics.
p
Prefix factor for Jaro-Winkler distance. The valid range for p is 0 <= p <= 0.25. If p=0 (defau
Applies only to method='jw'.
overwrite
A data frame used to overwrite the sector and/or name columns of a particular direct loantake
only sector, the value in the name column should be NA and vice-versa. This file can be used

Resources available to a user: External resources

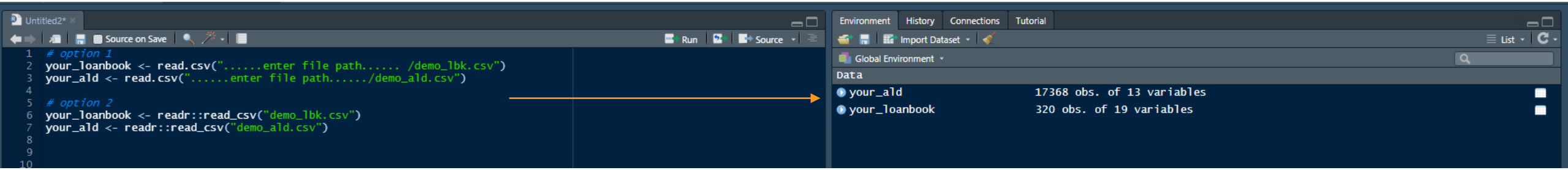
- You may encounter errors that are unrelated to the r2dii packages
- Read the error messages and work out what is going wrong
- These are some commonly used resources to help overcome common errors messages in R

  - Stackoverflow 

  - rstudio community 

Work flow:
Step 1. Import files

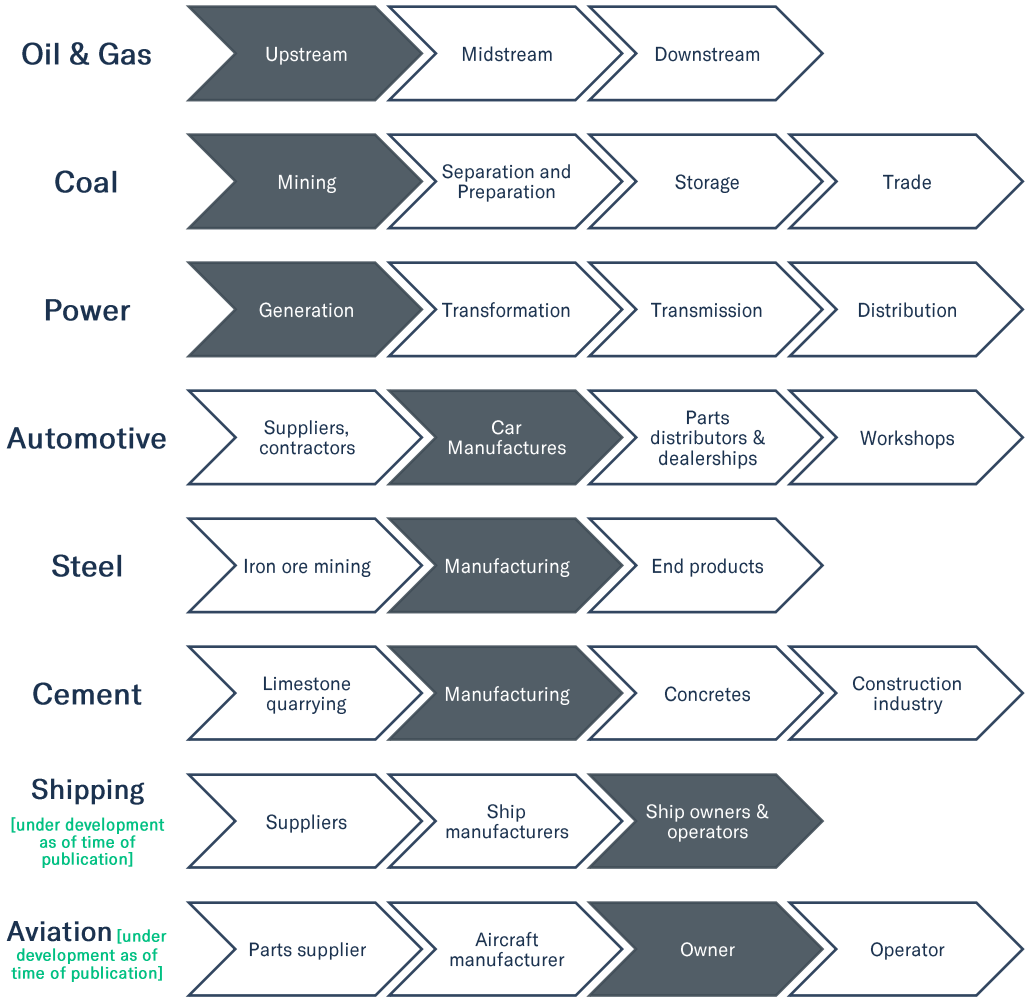
- Import loan book – in the correct template: r2dii.data::loanbook_demo
- Import Asset level Data set
  - No right way to do this
  - Possible functions
    - Readr
    - Read.csv



- Note (optional) – You may want to segment the LBK and ALD by sectors
  - This will make it more manageable to match and prevent potential issues with the size of the files and memory limits

Sector Classification Bridges

- Sector classification bridges between common code systems and 2dii sectors are provided
- Including:
  - **NACE**
  - **GICS**
  - **ISIC**
  - **SIC**
  - **NAICS**
- This can be accessed via r2dii.data::nace_classification
- More codes than are technically in scope are included



Sector Classification Bridges

Example from the NACE bridge

	original_code	code	code_level	sector	borderline
	All	All	All	gas	All
1	6	60	2	oil and gas	FALSE
2	6.1	610	3	oil and gas	FALSE
3	6.1	6100	4	oil and gas	FALSE
4	6.2	620	3	oil and gas	FALSE
5	6.2	6200	4	oil and gas	FALSE
6	9.1	910	3	oil and gas	TRUE
7	9.1	9100	4	oil and gas	TRUE
8	35.2	352	3	oil and gas	TRUE
9	35.21	3521	4	oil and gas	TRUE
10	35.22	3522	4	oil and gas	TRUE
11	35.23	3523	4	oil and gas	TRUE

Borderline = FALSE
6100 = Extraction of crude petroleum
6200 = Extraction of natural gas

Perfectly in scope

Oil & Gas

Close to scope – there is a chance it is mislabelled

Borderline = TRUE
3521 = Manufacture of gas
3522 = Distribution of gaseous fuels through mains

Work flow: Step 2. Matching

- Match the LBK and the ALD using r2dii.match::match_name

```
# match lbk to ald
match_file <- match_name(your_loanbook, your_ald)
```

```
match_file      404 obs. of 28 variables
your_ald        17368 obs. of 13 variables
your_loanbook   320 obs. of 19 variables
```

- Optional advanced matching
- Within the match_name function there are additional functions allowing a user to customise this stage

```
match_file <- match_name(your_loanbook, your_ald, by_sector = TRUE,
                         min_score = 0.8, method = "jw", p = 0.1, overwrite = NULL)
```

min_score allows you to set the matching score threshold

by_sector allows you to match any names irrespective of the sector classification

- More on overwrite later

String matching

- String matching assesses the characters and the order of the characters between two names
- It then outputs a measure of the similarity of the two words
- For example, there is a 80% similarity between “Royal Dutch Shell” and “R Dutch Shell” based on the order and letters involved

Name in LBK	Name in ALD	Score
BP plc	BP plc	1
BP	BP plc	0.85
British Petroleum	BP plc	0.3
Shell	BP	0

Illustrative example of string matching BP

String matching

There are different algorithms that you can choose from

```
match_file <- match_name(your_loanbook, your_ald, by_sector = TRUE,
                         min_score = 0.8, method = "jw", p = 0.1, overwrite = NULL)
```

The different methods use slightly different algorithms to determine the score in the previous slide

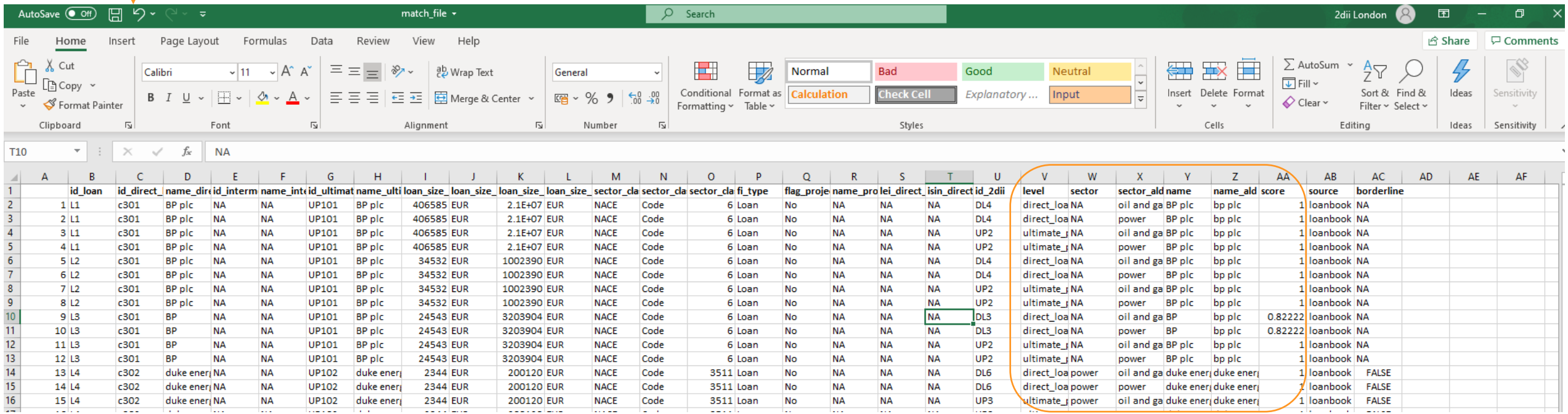
Options:
"osa", "lv", "dl", "hamming", "lcs", "qgram", "cosine", "jaccard", "jw", "soundex".

See stringdist::stringdist-metrics.

Work flow:
Step 3. Manual matching

- Export the “match_file” file from R to Excel

```
#export match_file to excel
write.csv(match_file, ".....filepath...../match_file.csv")
```



	id_loan	id_direct_	name_dir	id_interm	name_int	id_ultimat	name_ulti	loan_size_	loan_size_	loan_size_	loan_size_	sector_cla	sector_cla	sector_cla	fi_type	flag_proje	name_pro	lei_direct_	isin_direct	id_2dii	level	sector	sector_ald	name	name_ald	score	source	borderline
1	L1	c301	BP plc	NA	NA	UP101	BP plc	406585	EUR	2.1E+07	EUR	NACE	Code	6	Loan	No	NA	NA	NA	DL4	direct_loa	NA	oil and ga	BP plc	bp plc	1	loanbook	NA
2	L1	c301	BP plc	NA	NA	UP101	BP plc	406585	EUR	2.1E+07	EUR	NACE	Code	6	Loan	No	NA	NA	NA	DL4	direct_loa	NA	power	BP plc	bp plc	1	loanbook	NA
3	L1	c301	BP plc	NA	NA	UP101	BP plc	406585	EUR	2.1E+07	EUR	NACE	Code	6	Loan	No	NA	NA	NA	UP2	ultimate_	NA	oil and ga	BP plc	bp plc	1	loanbook	NA
4	L1	c301	BP plc	NA	NA	UP101	BP plc	406585	EUR	2.1E+07	EUR	NACE	Code	6	Loan	No	NA	NA	NA	UP2	ultimate_	NA	power	BP plc	bp plc	1	loanbook	NA
5	L2	c301	BP plc	NA	NA	UP101	BP plc	34532	EUR	1002390	EUR	NACE	Code	6	Loan	No	NA	NA	NA	DL4	direct_loa	NA	oil and ga	BP plc	bp plc	1	loanbook	NA
6	L2	c301	BP plc	NA	NA	UP101	BP plc	34532	EUR	1002390	EUR	NACE	Code	6	Loan	No	NA	NA	NA	DL4	direct_loa	NA	power	BP plc	bp plc	1	loanbook	NA
7	L2	c301	BP plc	NA	NA	UP101	BP plc	34532	EUR	1002390	EUR	NACE	Code	6	Loan	No	NA	NA	NA	UP2	ultimate_	NA	oil and ga	BP plc	bp plc	1	loanbook	NA
8	L2	c301	BP plc	NA	NA	UP101	BP plc	34532	EUR	1002390	EUR	NACE	Code	6	Loan	No	NA	NA	NA	UP2	ultimate_	NA	power	BP plc	bp plc	1	loanbook	NA
9	L3	c301	BP	NA	NA	UP101	BP plc	24543	EUR	3203904	EUR	NACE	Code	6	Loan	No	NA	NA	NA	DL3	direct_loa	NA	oil and ga	BP	bp plc	0.82222	loanbook	NA
10	L3	c301	BP	NA	NA	UP101	BP plc	24543	EUR	3203904	EUR	NACE	Code	6	Loan	No	NA	NA	NA	DL3	direct_loa	NA	power	BP	bp plc	0.82222	loanbook	NA
11	L3	c301	BP	NA	NA	UP101	BP plc	24543	EUR	3203904	EUR	NACE	Code	6	Loan	No	NA	NA	NA	UP2	ultimate_	NA	oil and ga	BP plc	bp plc	1	loanbook	NA
12	L3	c301	BP	NA	NA	UP101	BP plc	24543	EUR	3203904	EUR	NACE	Code	6	Loan	No	NA	NA	NA	UP2	ultimate_	NA	power	BP plc	bp plc	1	loanbook	NA
13	L4	c302	duke ener	NA	NA	UP102	duke ener	2344	EUR	200120	EUR	NACE	Code	3511	Loan	No	NA	NA	NA	DL6	direct_loa	power	oil and ga	duke ener	duke ener	1	loanbook	FALSE
14	L4	c302	duke ener	NA	NA	UP102	duke ener	2344	EUR	200120	EUR	NACE	Code	3511	Loan	No	NA	NA	NA	DL6	direct_loa	power	power	duke ener	duke ener	1	loanbook	FALSE
15	L4	c302	duke ener	NA	NA	UP102	duke ener	2344	EUR	200120	EUR	NACE	Code	3511	Loan	No	NA	NA	NA	UP3	ultimate_	power	oil and ga	duke ener	duke ener	1	loanbook	FALSE

Work flow:
Step 3. Manual match

- Allocate 1 = match and 0 = no match - to each loan in the score column
- Tip – allocate 1 or 0 in a separate column – once you are confident of your matches, you can replace the 1/0 in the “Score” column

- **Certain rules apply:**
  1. Never match 1 name in the loan book to 2 names in the ALD
  2. You can however match 2 names in the loan book to 1 name in the ALD

level	sector	sector_ald	name	name_ald	score	manual match		
direct_loantaker	automotive	automotive	Aston Martin	aston martin	1	1		rule 1
direct_loantaker	automotive	automotive	Aston Martin	Aston Martin UK	0.9	0		
direct_loantaker	oil&gas	oil&gas	BP plc	BP plc	1	1		rule 2
direct_loantaker	oil&gas	oil&gas	BP	BP plc	0.9	1		

Work flow:
Step 3. Manual match

- **Rules continued:**
  3. Do not match a direct loan taker to an ultimate parent – see column “level”

level	sector	sector_ald	name	name_ald	score	manual match		
direct_loantaker	power	power	duke florida	duke	0.8	0		rule 3
ultimate_parent	power	power	duke	duke	1	1		
direct_loantaker	power	power	duke florida	duke florida	1	1		

  4. Do not match across sectors. (This is only relevant when setting the by_sector argument to equal FALSE)

level	sector	sector_ald	name	name_ald	score	manual match		
direct_loantaker	power	power	duke florida	duke florida	1	1		rule 4
direct_loantaker	power	automotive	duke florida	duke florida	1	0		

Work flow:
Step 3. Manual match

- **Rules continued:**

5. Do not assume a match to be the same company if it looks “close enough”

level	sector	sector_ald	name	name_ald	score	manual match		
direct_loantaker	cement	cement	cement tex	cementex	0.9	0		rule 5

6. If you cannot find a match you should look it up in the ALD

level	sector	sector_ald	name	name_ald	score	manual match		
direct_loantaker	oil&gas	oil&gas	pjsc gazprom	wazpromotion	0.8	0		rule 6
direct_loantaker	oil&gas	oil&gas	pjsc gazprom	gazprom pjsc	0.7	1		

Work flow:
Step 3. Manual match

- **Rules continued:**
  7. Change in ownership of a company not recorded in either the ALD or the LBK – investigate and use the overwrite file (next slide) to change the LBK so the company can be matched to the ALD

  8. If all of the above fail, then it is possible that a company does not exist in the ALD --> Contact the data provider

Work flow:
Step 4. Overwrite file

- In certain cases you may wish to change the name or sector of the company in the loan book to match that in the ALD
- To do this, open the overwrite file in excel – this can be found in r2dii.data::overwrite_demo

- Populate it with the name / sector changes you wish to carry out

	A	B	C	D	E	F
1		level	id_2dii	name	sector	source
2	1	direct_loantaker	DL294	bee handshoe	coal	manual
3	2	ultimate_parent	UP15	alpine india	power	manual
4						
5						

- Load this file in as an argument in the r2dii.match::match_name function

```
matched_file_overwrite <- match_name(your_loanbook, your_ald, overwrite = overwite_file)
```

Work flow:
Step 5.1. Combine all your matches

- All the matches must be combined into one file
- e.g. “validated_matches” file
- This can be done in Excel
- Read the “validated_matches” file back into R

```
validated_matches <- read.csv(".....file path.../validated_matches.csv")
```

Work flow:
Step 5.2. Prioritize Matches

- The prioritize function selects the best match for the loan
- By default, this is set at the **Direct Loan taker level.** When the Direct loan taker can not be matched, the Ultimate Parent is used

```
lbk_ready <- prioritize(validated_matches)
```

- If instead you would like the production values of the Ultimate Parent to be selected you can do so using the following code

```
lbk_ready <- prioritize(validated_matches, priority = rev)
```

rev – reverses the order of priority

- Note – This is an essential step – do not try to skip it

Next steps

- **Analysis / Visualisation webinar (date TBC) – www.2degrees-investing.org/events**
- **Follow the instructions on the r2dii.analysis website**
- To access the previous webinar on the methodology and further training materials and user guides, please visit **www.TransitionMonitor.com and the PACTA for Banks Tab**

Q&A

Contacts

- PACTA for Banks related queries: banks@2degrees-investing.org
- General 2DII queries: contact@2degress-investing.org
- www.2degrees-investing.org
- www.transitionmonitor.com

Supported by:

Federal Ministry
for the Environment, Nature Conservation
and Nuclear Safety

based on a decision of the German Bundestag

About our funders: *This project is part of the International Climate Initiative (IKI). The Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (BMU) supports this initiative on the basis of a decision adopted by the German Bundestag. This project has also received funding from the European Union LIFE program. The views expressed here are the sole responsibility of the authors and do not necessarily reflect the views of the funders.*

Annex
Rules

V	W	X	Y	Z	AA	AB	AC	AD
level	**sector**	**sector_ald**	**name**	**name_ald**	**score**	**manual match**		
direct_loantaker	automotive	automotive	Aston Martin	aston martin	1	1		rule 1
direct_loantaker	automotive	automotive	Aston Martin	Aston Martin UK	0.9	0		
direct_loantaker	oil&gas	oil&gas	BP plc	BP plc	1	1		rule 2
direct_loantaker	oil&gas	oil&gas	BP	BP plc	0.9	1		
direct_loantaker	power	power	duke florida	duke	0.8	0		rule 3
ultimate_parent	power	power	duke	duke	1	1		
direct_loantaker	power	power	duke florida	duke florida	1	1		
direct_loantaker	power	power	duke florida	duke florida	1	1		rule 4
direct_loantaker	power	automotive	duke florida	duke florida	1	0		
direct_loantaker	cement	cement	cement tex	cementex	0.9	0		rule 5
direct_loantaker	oil&gas	oil&gas	pjsc gazprom	wazpromotion	0.8	0		rule 6
direct_loantaker	oil&gas	oil&gas	pjsc gazprom	gazprom pjsc	0.7	1		