Package ‘r2dii.data’

September 3, 2021

Title Datasets to Measure the Alignment of Corporate Loan Books with Climate Goals

Version 0.2.2

Description These datasets support the implementation in R of the software ‘PACTA’ (Paris Agreement Capital Transition Assessment), which is a free tool that calculates the alignment between corporate lending portfolios and climate scenarios (<https://2degrees-investing.org/>). Financial institutions use ‘PACTA’ to study how their capital allocation decisions align with climate change mitigation goals. Because both financial institutions and market data providers keep their data private, this package provides fake, public data to enable the development and use of ‘PACTA’ in R.

License CC0


BugReports https://github.com/2DegreesInvesting/r2dii.data/issues

Depends R (>= 3.4)

Imports stats, utils

Suggests covr, rlang, rmarkdown, testthat (>= 2.1.0)

Config/testthat/edition 3

Encoding UTF-8

LazyData true

RoxygenNote 7.1.1

NeedsCompilation no

Author Mauro Lepore [aut, cre, ctr] (<https://orcid.org/0000-0002-1986-7988>), Jackson Hoffart [aut, dtc] (<https://orcid.org/0000-0002-8600-5042>), Klaus Hagedorn [aut], 2 Degrees Investing Initiative [cph, fnd]

Maintainer Mauro Lepore <maurolepore@gmail.com>

Repository CRAN

Date/Publication 2021-09-03 17:50:02 UTC
### Description

Fake data about physical assets (e.g. wind turbine power plant capacities) used to assess the climate alignment of financial portfolios. It imitates data from market-intelligence databases.

Demo datasets are synthetic because most financial data is strictly private; they help to demonstrate and test the implementation in R of 'PACTA' ([https://2degrees-investing.org/resource/pacta/](https://2degrees-investing.org/resource/pacta/)).

### Usage

```r
ald_demo
```

### Format

An object of class `tbl_df` (inherits from `tbl, data.frame`) with 17668 rows and 14 columns.
Definitions

- ald_timestamp (character): Date at which asset data was sourced from the data provider.
- company_id (character): The id of the company owning the asset created by the data provider.
- country_of_domicile (character): Country where company is registered.
- emission_factor (double): Company level emission factor of the technology.
- emission_factor_unit (character): The units that the emission factor is measured in.
- is_ultimate_owner (logical): Flag if company is the ultimate parent in our database.
- lei (character): The legal entity identifier of the company owning the asset.
- name_company (character): The name of the company owning the asset.
- plant_location (character): Country where asset is located.
- production (double): Company level production of the technology.
- production_unit (character): The units that production is measured in.
- sector (character): Sector to which the asset belongs.
- technology (character): Technology implemented by the asset.
- year (integer): Year at which the production value is predicted.

See Also

data_dictionary

Other demo datasets: co2_intensity_scenario_demo, loanbook_demo, overwrite_demo, region_isos_demo, scenario_demo_2020

Examples

head(ald_demo)

---

cnb_classification  Dataset to bridge (translate) common sector-classification codes

Description

This dataset serves as a translation key between common sector-classification systems and sectors relevant to the 'PACTA' tool (https://2degrees-investing.org/resource/pacta/).

Usage

cnb_classification

Format

An object of class tbl_df (inherits from tbl, data.frame) with 220 rows and 5 columns.
Definitions

- borderline (logical): Flag indicating if 2dii sector and classification code are a borderline match. The value TRUE indicates that the match is uncertain between the 2dii sector and the classification. The value FALSE indicates that the match is certainly perfect or the classification is certainly out of 2dii’s scope.
- code (character): Formatted CNB code.
- code_level (double): Level of granularity of CNB code.
- original_code (character): Original CNB sector name.
- sector (character): Associated 2dii sector.

Details

Classification datasets help to standardize sector classification codes from the wild to a relevant subset including 'power', 'oil and gas', 'coal', 'automotive', 'aviation', 'concrete', 'steel', and 'shipping'.

See Also

data_dictionary.

Other datasets for bridging sector classification codes: gics_classification, isic_classification, nace_classification, naics_classification, psic_classification, sector_classifications, sic_classification

Examples

head(cnb_classification)

co2_intensity_scenario_demo

A prepared co2 intensity climate scenario dataset for demonstration

Description

Fake co2 intensity climate scenario dataset, prepared for the software PACTA (Paris Agreement Capital Transition Assessment). It imitates climate scenario data (e.g. from the International Energy Agency (IEA)) including the change through time in production across industrial sectors (calculated by 2DII).

Demo datasets are synthetic because most financial data is strictly private; they help to demonstrate and test the implementation in R of 'PACTA' (https://2degrees-investing.org/resource/pacta/).

Usage

c02_intensity_scenario_demo
Format

An object of class tbl_df (inherits from tbl_data.frame) with 22 rows and 7 columns.

Definitions

- emission_factor (double): The target sector level emissions factor that the scenario prescribes.
- emission_factor_unit (character): The units that the emissions factor is measured in.
- region (character): The region to which the pathway is relevant.
- scenario (character): The name of the scenario.
- scenario_source (character): The source publication from which the scenario was taken.
- sector (character): The sector to which the scenario prescribes a pathway.
- year (integer): The year at which the pathway value is prescribed.

See Also

data_dictionary

Other demo datasets: ald_demo, loanbook_demo, overwrite_demo, region_isos_demo, scenario_demo_2020

Examples

head(co2_intensity_scenario_demo)

---

data_dictionary  Column definitions of all datasets

Description

This dataset provides metadata about all datasets in the package r2dii.data.

Usage

data_dictionary

Format

An object of class tbl_df (inherits from tbl_data.frame) with 104 rows and 4 columns.

Definitions

- column (character): The name of a dataset-column.
- dataset (character): The name of a dataset.
- definition (character): The definition of a dataset-column.
- typeof (character): The result of typeof(), one of double, integer, logical, or character.  

Examples

head(data_dictionary)

gics_classification

Dataset to bridge (translate) common sector-classification codes

Description

This dataset serves as a translation key between common sector-classification systems and sectors relevant to the 'PACTA' tool (https://2degrees-investing.org/resource/pacta/).

Usage

gics_classification

Format

An object of class tbl_df (inherits from tbl, data.frame) with 263 rows and 5 columns.

Definitions

• borderline (logical): Flag indicating if 2dii sector and classification code are a borderline match. The value TRUE indicates that the match is uncertain between the 2dii sector and the classification. The value FALSE indicates that the match is certainly perfect or the classification is certainly out of 2dii’s scope.
• code (character): Original GICS code.
• code_level (double): Level of granularity of GICS code.
• description (character): Original GICS description.
• sector (character): Associated 2dii sector.

Details


See Also

data_dictionary.

Other datasets for bridging sector classification codes: cnb_classification, isic_classification, nace_classification, naics_classification, psic_classification, sector_classifications, sic_classification

Examples

head(gics_classification)
green_or_brown

Determine if a technology is green or brown

Description
This dataset provides a simple lookup table to determine if a technology is meant to increase (green) or decrease (brown) to align with a future global warming of less than 2 degrees.

Usage
green_or_brown

Format
An object of class tbl_df (inherits from tbl, data.frame) with 20 rows and 3 columns.

Definitions
- green_or_brown (character): If the technology is green (increasing) or brown (decreasing), as defined by the Paris-aligned IEA scenarios.
- sector (character): The sector to which the technology belongs.
- technology (character): The technology sub-category within the sector.

See Also
data_dictionary

Examples
head(green_or_brown)

isic_classification

Dataset to bridge (translate) common sector-classification codes

Description
This dataset serves as a translation key between common sector-classification systems and sectors relevant to the 'PACTA' tool (https://2degrees-investing.org/resource/pacta/).

Usage
isic_classification

Format
An object of class tbl_df (inherits from tbl, data.frame) with 768 rows and 4 columns.
Definitions

- borderline (logical): Flag indicating if 2dii sector and classification code are a borderline match. The value TRUE indicates that the match is uncertain between the 2dii sector and the classification. The value FALSE indicates that the match is certainly perfect or the classification is certainly out of 2dii’s scope.
- code (character): Original ISIC code.
- code_level (double): Level of granularity of ISIC code.
- sector (character): Associated 2dii sector.

Details

Classification datasets help to standardize sector classification codes from the wild to a relevant subset including 'power', 'oil and gas', 'coal', 'automotive', 'aviation', 'concrete', 'steel', and 'shipping'.

See Also

data_dictionary.  
Other datasets for bridging sector classification codes: cnb_classification, gics_classification, nace_classification, naics_classification, psic_classification, sector_classifications, sic_classification

Examples

head(isic_classification)

---

iso_codes  Countries and codes

Description

This dataset maps countries to codes.

For information about the ISO standard for country codes see https://www.iso.org/iso-3166-country-codes.html.

Usage

iso_codes

Format

An object of class tbl_df (inherits from tbl, data.frame) with 286 rows and 2 columns.

Definitions

- country (character): Country name.
- country_iso (character): Corresponding ISO code.
See Also

data_dictionary

Other iso codes: region_isos_demo, region_isos

Examples

head(iso_codes)

loanbook_demo

A loanbook dataset for demonstration

Description

Fake financial portfolio.

Demo datasets are synthetic because most financial data is strictly private; they help to demonstrate and test the implementation in R of 'PACTA' (https://2degrees-investing.org/resource/pacta/).

Usage

loanbook_demo

Format

An object of class tbl_df (inherits from tbl, data.frame) with 321 rows and 19 columns.

Definitions

- **fi_type** (character): Financial instrument name or asset class.
- **flag_project_finance_loan** (character): Project finance flag denoting whether a loan is given out to a particular asset or not.
- **id_direct_loantaker** (character): Borrower identifier unique to each borrower/sector combination in loanbook.
- **id_intermediate_parent_n** (character): Optional input: id of the n-th intermediate parent company within the company structure that can be used for more granular mapping than the ultimate parent. Smaller values of n are closer to the direct_loantaker.
- **id_loan** (character): Unique loan identifier.
- **id_ultimate_parent** (character): Ultimate parent identifier unique to each ultimate parent/sector combination.
- **isin_direct_loantaker** (logical): Optional input: providing the isin identifier of the direct loan taker to improve the matching coverage.
- **lei_direct_loantaker** (logical): Optional input: providing the lei (legal entity identifier) of the direct loan taker to improve the matching coverage.
- **loan_size_credit_limit** (double): Total credit limit or exposure at default.
• loan_size_credit_limit_currency (character): Currency corresponding to credit limit.
• loan_size_outstanding (double): Amount drawn by borrower from total credit limit.
• loan_size_outstanding_currency (character): Currency corresponding to outstandings.
• name_direct_loantaker (character): Name of the company directly taking the loan.
• name_intermediate_parent_n (character): Optional input: name of intermediate parent company within the company structure that can be used for more granular mapping than the ultimate parent. Smaller values of n are closer to the direct_loantaker.
• name_project (logical): Required input for loans with the flag_project_finance_loan = TRUE: Name of the project being financed.
• name_ultimate_parent (character): Name of the ultimate parent company to which the borrower belongs. Can be the same as borrower.
• sector_classification_direct_loantaker (double): Sector classification code of the direct_loantaker.
• sector_classification_input_type (character): Flag identifying if the sector classification code or character description is used.
• sector_classification_system (character): Name of the sector classification standard being used.

See Also
data_dictionary

Other demo datasets: ald_demo, co2_intensity_scenario_demo, overwrite_demo, region_isos_demo, scenario_demo_2020

Examples

head(loanbook_demo)

---
nace_classification

Dataset to bridge (translate) common sector-classification codes

Description

This dataset serves as a translation key between common sector-classification systems and sectors relevant to the 'PACTA' tool (https://2degrees-investing.org/resource/pacta/).

Usage

nace_classification

Format

An object of class tbl_df (inherits from tbl, data.frame) with 996 rows and 5 columns.
Definitions

- borderline (logical): Flag indicating if 2dii sector and classification code are a borderline match. The value TRUE indicates that the match is uncertain between the 2dii sector and the classification. The value FALSE indicates that the match is certainly perfect or the classification is certainly out of 2dii's scope.
- code (character): Formatted NACE code with periods removed.
- code_level (double): Level of granularity of NACE code.
- original_code (character): Original NACE code.
- sector (character): Associated 2dii sector.

Details

Classification datasets help to standardize sector classification codes from the wild to a relevant subset including 'power', 'oil and gas', 'coal', 'automotive', 'aviation', 'concrete', 'steel', and 'shipping'.

See Also

data_dictionary.

Other datasets for bridging sector classification codes: cnb_classification, gics_classification, isic_classification, naics_classification, psic_classification, sector_classifications, sic_classification

Examples

head(nace_classification)

---

**naics_classification**  
*Dataset to bridge (translate) common sector-classification codes*

Description

This dataset serves as a translation key between common sector-classification systems and sectors relevant to the 'PACTA' tool ([https://2degrees-investing.org/resource/pacta/](https://2degrees-investing.org/resource/pacta/)).

Usage

naics_classification

Format

An object of class tbl_df (inherits from tbl, data.frame) with 1057 rows and 4 columns.
Definitions

- **borderline** (logical): Flag indicating if 2dii sector and classification code are a borderline match. The value TRUE indicates that the match is uncertain between the 2dii sector and the classification. The value FALSE indicates that the match is certainly perfect or the classification is certainly out of 2dii’s scope..
- **code** (character): Six-digit NAICS code.
- **naics_title** (character): Original NAICS sector title.
- **sector** (character): Associated 2dii sector.

Details

Classification datasets help to standardize sector classification codes from the wild to a relevant subset including 'power', 'oil and gas', 'coal', 'automotive', 'aviation', 'concrete', 'steel', and 'shipping'.

See Also

- `data_dictionary`

Other datasets for bridging sector classification codes: `cnb_classification`, `gics_classification`, `isic_classification`, `nace_classification`, `psic_classification`, `sector_classifications`, `sic_classification`

Examples

```r
head(naics_classification)
```

<table>
<thead>
<tr>
<th>overwrite_demo</th>
<th>A demonstration dataset used to overwrite specific entity names or sectors</th>
</tr>
</thead>
</table>

Description

Fake dataset used to manually link loanbook entities to mismatched asset level entities.

Demo datasets are synthetic because most financial data is strictly private; they help to demonstrate and test the implementation in R of 'PACTA' (https://2degrees-investing.org/resource/pacta/).

Usage

`overwrite_demo`

Format

An object of class tbl_df (inherits from tbl, data.frame) with 2 rows and 5 columns.
Definitions

- id_2dii (character): IDs of the entities to overwrite.
- level (character): Which level should be overwritten (e.g. direct_loantaker or ultimate_parent).
- name (character): Overwrite name (if only overwriting sector, type NA).
- sector (character): Overwrite sector (if only overwriting name, type NA).
- source (character): What is the source of this information (leave as "manual" for now, may remove this flag later).

See Also

data_dictionary

Other demo datasets: ald_demo, co2_intensity_scenario_demo, loanbook_demo, region_isos_demo, scenario_demo_2020

Examples

head(overwrite_demo)

---

psic_classification  

Dataset to bridge (translate) common sector-classification codes

Description

This dataset serves as a translation key between common sector-classification systems and sectors relevant to the 'PACTA' tool (https://2degrees-investing.org/resource/pacta/).

Usage

psic_classification

Format

An object of class tbl_df (inherits from tbl, data.frame) with 1271 rows and 5 columns.

Definitions

- borderline (logical): Flag indicating if 2dii sector and classification code are a borderline match. The value TRUE indicates that the match is uncertain between the 2dii sector and the classification. The value FALSE indicates that the match is certainly perfect or the classification is certainly out of 2dii's scope.
- code (character): Formatted psic_classification code.
- code_level (double): Level of granularity of psic_classification code.
- original_code (character): Original psic_classification sector name.
- sector (character): Associated 2dii sector.
Details
Classification datasets help to standardize sector classification codes from the wild to a relevant subset including 'power', 'oil and gas', 'coal', 'automotive', 'aviation', 'concrete', 'steel', and 'shipping'.

See Also

data_dictionary.
Other datasets for bridging sector classification codes: cnb_classification, gics_classification, isic_classification, nace_classification, naics_classification, sector_classifications, sic_classification

Examples
head(psic_classification)

region_isos
A dataset outlining various region definitions

Description
This dataset maps codes representing countries to regions.
For information about the ISO standard for country codes see https://www.iso.org/iso-3166-country-codes.html.

Usage
region_isos

Format
An object of class tbl_df (inherits from tbl, data.frame) with 3722 rows and 3 columns.

Definitions
- isos (character): Countries in region, defined by iso code.
- region (character): Benchmark region name.
- source (character): Source publication from which the regions are defined.

See Also
data_dictionary
Other iso codes: iso_codes, region_isos_demo

Examples
head(region_isos)
A dataset outlining various region definitions

### Description

This dataset maps codes representing countries to regions. It is similar to but smaller than region_isos.

Demo datasets are synthetic because most financial data is strictly private; they help to demonstrate and test the implementation in R of 'PACTA' ([https://2degrees-investing.org/resource/pacta/](https://2degrees-investing.org/resource/pacta/)).

For information about the ISO standard for country codes see [https://www.iso.org/iso-3166-country-codes.html](https://www.iso.org/iso-3166-country-codes.html).

### Usage

```r
region_isos_demo
```

### Format

An object of class `tbl_df` (inherits from `tbl, data.frame`) with 358 rows and 3 columns.

### Definitions

- `isos` (character): Countries in region, defined by iso code.
- `region` (character): Benchmark region name.
- `source` (character): Source publication from which the regions are defined.

### See Also

Other iso codes: `iso_codes, region_isos`

Other demo datasets: `ald_demo, co2_intensity_scenario_demo, loanbook_demo, overwrite_demo, scenario_demo_2020`

### Examples

```r
region_isos_demo
```
A prepared climate scenario dataset for demonstration

Description
Fake climate scenario dataset, prepared for the software PACTA (Paris Agreement Capital Transition Assessment). It imitates climate scenario data (e.g. from the International Energy Agency (IEA)) including the change through time in production across industrial sectors (calculated by 2DII).
Demo datasets are synthetic because most financial data is strictly private; they help to demonstrate and test the implementation in R of 'PACTA' (https://2degrees-investing.org/resource/pacta/).

Usage
scenario_demo_2020

Format
An object of class tbl_df (inherits from tbl.data.frame) with 1512 rows and 8 columns.

Definitions
- region (character): The region to which the pathway is relevant.
- scenario (character): The name of the scenario.
- scenario_source (character): The source publication from which the scenario was taken.
- sector (character): The sector to which the scenario prescribes a pathway.
- smsp (double): Sector market share percentage of the pathway calculated in 2020.
- technology (character): The technology within the sector to which the scenario prescribes a pathway.
- tmsr (double): Technology market share ratio of the pathway calculated in 2020.
- year (integer): The year at which the pathway value is prescribed.

See Also
data_dictionary
Other demo datasets: ald_demo, co2_intensity_scenario_demo, loanbook_demo, overwrite_demo, region_isos_demo

Examples
head(scenario_demo_2020)
sector_classifications

A view of available sector classification datasets

Description

This dataset lists all sector classification code standards used by 'PACTA' (https://2degrees-investing.org/resource/pacta/).

Usage

sector_classifications

Format

An object of class tbl_df (inherits from tbl, data.frame) with 4700 rows and 4 columns.

Definitions

- borderline (character): Flag indicating if 2dii sector and classification code are a borderline match. The value TRUE indicates that the match is uncertain between the 2dii sector and the classification. The value FALSE indicates that the match is certainly perfect or the classification is certainly out of 2dii’s scope.
- code (character): Formatted code.
- code_system (character): Code system.
- sector (character): Associated 2dii sector.

Details

Classification datasets help to standardize sector classification codes from the wild to a relevant subset including 'power', 'oil and gas', 'coal', 'automotive', 'aviation', 'concrete', 'steel', and 'shipping'.

See Also

data_dictionary.

Other datasets for bridging sector classification codes: cnb_classification, gics_classification, isic_classification, nace_classification, naics_classification, psic_classification, sic_classification

Examples

head(sector_classifications)
sic_classification

Dataset to bridge (translate) common sector-classification codes

Description
This dataset serves as a translation key between common sector-classification systems and sectors relevant to the 'PACTA' tool (https://2degrees-investing.org/resource/pacta/).

Usage
sic_classification

Format
An object of class tbl_df (inherits from tbl.data.frame) with 256 rows and 4 columns.

Definitions
- borderline (character): Flag indicating if 2dii sector and classification code are a borderline match. The value TRUE indicates that the match is uncertain between the 2dii sector and the classification. The value FALSE indicates that the match is certainly perfect or the classification is certainly out of 2dii's scope.
- code (character): Original SIC code.
- description (character): Original SIC description.
- sector (character): Associated 2dii sector.

Details
Classification datasets help to standardize sector classification codes from the wild to a relevant subset including 'power', 'oil and gas', 'coal', 'automotive', 'aviation', 'concrete', 'steel', and 'shipping'.

See Also
data_dictionary.
Other datasets for bridging sector classification codes: cnb_classification, gics_classification, isic_classification, nace_classification, naics_classification, psic_classification, sector_classifications

Examples
head(sic_classification)
Index

* datasets for bridging sector classification codes
  cnb_classification, 3
  gics_classification, 6
  isic_classification, 7
  nace_classification, 10
  naics_classification, 11
  psic_classification, 13
  sector_classifications, 17
  sic_classification, 18

* datasets
  ald_demo, 2
  cnb_classification, 3
  co2_intensity_scenario_demo, 4
  data_dictionary, 5
  gics_classification, 6
  green_or_brown, 7
  isic_classification, 7
  iso_codes, 8
  loanbook_demo, 9
  nace_classification, 10
  naics_classification, 11
  overwrite_demo, 12
  psic_classification, 13
  region_isos, 14
  region_isos_demo, 15
  scenario_demo_2020, 16
  sector_classifications, 17
  sic_classification, 18

* demo datasets
  ald_demo, 2
  co2_intensity_scenario_demo, 4
  loanbook_demo, 9
  overwrite_demo, 12
  region_isos_demo, 15
  scenario_demo_2020, 16

* iso codes
  iso_codes, 8
  region_isos, 14
  region_isos_demo, 15

* meta
  data_dictionary, 5

* miscellaneous datasets
  green_or_brown, 7
  ald_demo, 2, 5, 10, 13, 15, 16
  cnb_classification, 3, 6, 8, 11, 12, 14, 17, 18
  co2_intensity_scenario_demo, 3, 4, 10, 13, 15, 16
  data_dictionary, 3–5, 6–14, 16–18
  gics_classification, 4, 6, 8, 11, 12, 14, 17, 18
  green_or_brown, 7
  isic_classification, 4, 6, 7, 11, 12, 14, 17, 18
  iso_codes, 8, 14, 15
  loanbook_demo, 3, 5, 9, 13, 15, 16
  nace_classification, 4, 6, 8, 10, 12, 14, 17, 18
  naics_classification, 4, 6, 8, 11, 14, 17, 18
  overwrite_demo, 3, 5, 10, 12, 15, 16
  psic_classification, 4, 6, 8, 11, 12, 13, 17, 18
  region_isos, 9, 14, 15
  region_isos_demo, 3, 5, 9, 10, 13, 14, 15, 16
  scenario_demo_2020, 3, 5, 10, 13, 15, 16
  sector_classifications, 4, 6, 8, 11, 12, 14, 17, 18
  sic_classification, 4, 6, 8, 11, 12, 14, 17, 18