decarbonisation-strats

User Guide

Welcome to NatWest Decarbonisation Dashboard user guide! This is a read only dashboard, which means there will be no *create*, *update* or *delete* functionality.

This user guide is designed to provide documentation on the usage of the data pipeline, dashboard as well as containing information to any upcoming features.

1. Quick start

1. Launch the dashboard. You should see the home page of the dashboard which looks something like this.

2. Under the `Client Selector` heading, search and select a company.

3. Under the `Client Selector` heading, seleted a year and click `Submit` .

4. Click on one of the button (i.e `Current Emissions` , `Emission Reduction Target` or `Transition Plans`). For example, the image below is when `Current Emissions` is

clicked, which appears as a modal.

5. Refer to the Features below for details of the dashboard.

2. Data Pipeline

ℹ️ **Notes about the data pipeline:**

- This section provides documentation on the data that are to be rendered on the dashboard.
- To understand and use the data pipeline, basic familiarity with `Python` and *Machine Learning (ML)* are **highly recommended**.
- For first time users, kindly ensure that you have the relevant packages installed before executing the code.
- The files mentioned in the following sections **must** be a *csv* file and the naming must strictly adhere to the header of the section.

2.1. The Training Set

This is the input to the pipeline script in order to train the model. The training set includes the following two files namely `Processed_Corp_List.csv` and `Labelled_Answers.csv`.

2.1.1. Processed_Corp_List.csv

This file contains information about the company such as the company name, sector as well as their Sustainability Report URLs. Columns include:

- *IssuerNamer*
- *ISIN*
- *Ticker*

- *CountryOfIncorporation*
- *GICSSector*
- *GICSSubIndustry*
- *Year*
- *Report URLs*

For the **current version**, the training dataset in this case would be the 2021 Sustainability Reports of 61 Energy Sector companies.

	IssuerName	ISIN	Ticker	CountryOfIncorporation	GICSSector	GICSSubIndustry	Year	Report URL
0	Repsol SA	ES0173516115	REP	Spain	Energy	Integrated Oil & Gas	2021	https://www.repsol.com/content/dam/repsol-corp...
1	OMV AG	AT0000743059	OMV	Austria	Energy	Integrated Oil & Gas	2021	https://www.omv.com/services/downloads/00/omv....
2	TotalEnergies SE	FR0000120271	TTE	France	Energy	Integrated Oil & Gas	2021	https://totalenergies.com/system/files/documen...
3	Energean Plc	GB00BG12Y042	ENOG	United Kingdom	Energy	Oil & Gas Exploration & Production	2021	https://www.energean.com/media/5201/2021-susta...
4	Eni SpA	IT0003132476	ENI	Italy	Energy	Integrated Oil & Gas	2021	https://www.eni.com/assets/documents/eng/just-...
...	...	...	...	...	...	...	...	...
62	Crescent Point Energy Corp.	CA22576C1014	CPG	Canada	Energy	Oil & Gas Exploration & Production	2021	https://www.crescentpointenergy.com/sites/defa...
63	Tethys Oil AB	SE0017859259	TETY	Sweden	Energy	Oil & Gas Exploration & Production	2021	https://www.tethysoil.com/en/wp-content/upload...
64	KOSMOS ENERGY LTD.	US5006881065	KOS	USA	Energy	Oil & Gas Exploration & Production	2021	https://www.kosmosenergy.com/KosmosSustainabil...
65	CNOOC Limited	HK0883013259	883	Hong Kong	Energy	Oil & Gas Exploration & Production	2021	https://www.cnoocltd.com/attach/0/220413233504...
66	Frontera Energy Corporation	CA35905B1076	FEC	Canada	Energy	Oil & Gas Exploration & Production	2021	http://www.fronteraenergy.ca/content/uploads/2...

2.1.2. Labelled_Answers.csv

This file contains the data, that is the answers to the question in order to train the ML model. Columns include:

- IssuerName
- Q'X' (where X is the question number. E.g Q1, Q11)

	IssuerName	Q3	Q4	Q7	Q9	Q10	Q11	Q12	Q13	Q14
0	Pharos Energy Plc	NaN	Yes	NaN	Yes	NaN	NaN	NA	NaN	NaN
1	PT Adaro Energy Indonesia Tbk	NaN	Yes	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	Ecopetrol SA	Yes	Yes	Yes	Yes	NaN	Yes	Yes	NaN	NaN
3	Enerplus Corporation	NaN	NaN	NaN	Yes	NaN	NaN	NaN	NaN	NaN
4	Baytex Energy Corp.	Yes	NaN	NaN	Yes	NaN	NaN	NaN	NaN	Yes
...	...	...	...	...	...	...	...	...	...	...
72	Southwestern Energy Company	NaN	Yes	NaN	NaN	NaN	NaN	NaN	NaN	Yes
73	Tourmaline Oil Corp.	NA_2021	NA_2021	NA_2021	NA_2021	NA_2021	NA_2021	NA_2021	NA_2021	NA_2021
74	Oil Co. LUKOIL PJSC	NaN	Yes	Yes	Yes	NaN	Yes	Yes	NaN	NaN
75	Exxon Mobil Corporation	NaN	NaN	Yes	Yes	NaN	NaN	NaN	NaN	NaN
76	Marathon Oil Corporation	NA_2021	NA_2021	NA_2021	NA_2021	NA_2021	NA_2021	NA_2021	NA_2021	NA_2021

77 rows × 10 columns

❗ Notes about the training set:

- If the *file name* or the *column name* differs from above, this will not allow you to execute the data pipeline script!
- If you want to introduce more training data into `Labelled_Answers.csv`, manual labelling has to be done for these data.
- Once you have trained an ML model, you do not need to train it again unless there are changes to the training set.

2.2. Output Answers to Inputs

This section provides a general guide to output the answers to the questionnaires for the comapny's Sustainability Report. This is also applicable to the case where you have a list of new company and it's Sustainability Reports to be included.

2.2.1. New_URLs.csv

This file contains the details to the list of company of interest. Columns include:

- *IssuerNamer*
- *ISIN*
- *Ticker*
- *CountryOfIncorporation*
- *GICSSector*
- *GICSSubIndustry*
- *Year*
- *Report URLs*

	IssuerName	ISIN	Ticker	CountryOfIncorporation	GICSSector	GICSSubIndustry	Year	Report URL
0	Kilroy Realty Corporation	US49427F1084	KRC	USA	Real Estate	Office REITs	2021	https://kilroyrealty.com/wp-content/uploads/20...
1	Highwoods Properties, Inc.	US4312841087	HIW	USA	Real Estate	Office REITs	2021	https://www.responsibilityreports.com/Click/1335
2	Prologis, Inc.	US74340W1036	PLD	USA	Real Estate	Industrial REITs	2021	https://prologis.getbynder.com/m/40de827ea131c...
3	Host Hotels & Resorts, Inc.	US44107P1049	HST	USA	Real Estate	Hotel & Resort REITs	2021	https://www.hosthotels.com/-/media/HostHotels/...
4	DiamondRock Hospitality Company	US2527843013	DRH	USA	Real Estate	Hotel & Resort REITs	2021	https://investor.drhc.com/static-files/52de73f...

2.3. Execute the data pipeline script

Place the above-mentioned files in the same directory as the data pipeline script named `Final_Pipeline.ipynb` and kindly execute `Final_Pipeline.ipynb`. The output of the execution is a *csv* file titled `Consolidated_Output.csv` which has to be **uploaded** onto Dataiku to *pipeline* it into the dashboard.

2.3.1. Consolidated_Output.csv

The output file `Consolidated_Output.csv` It comprises the data in the training dataset and the ML results of the new URLs.

	Unnamed: 0	IssuerName	ISIN	Ticker	CountryOfIncorporation	GICSSector	GICSSubIndustry	2021 Report URL	Processed_Sentences	Q3	Q4	Q7	Q9	Q10	Q11	Q13	Q14	Q12
0	0	Kilroy Realty Corporation	US49427F1084	KRC	USA	Real Estate	Office REITs	https://kilroyrealty.com/wp-content/uploads/20...	['S', 'EP4 GREEN POWER PARTNERSHIP in 2021,	1.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
1	1	Highwoods Properties, Inc.	US4312841087	HIW	USA	Real Estate	Office REITs	https://www.responsibilityreports.com/Click/1335	[' Despite the continued challenges of the pan...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
2	2	Prologis, Inc.	US74340W1036	PLD	USA	Real Estate	Industrial REITs	https://prologis.getbynder.com/m/40de827ea131c...	['2021-22 ESG Report Prologis Eindhoven Distri...	0.0	1.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0
3	3	Host Hotels & Resorts, Inc.	US44107P1049	HST	USA	Real Estate	Hotel & Resort REITs	https://www.hosthotels.com/-/media/HostHotels/...	[' None of the owners of these trademarks has ...	0.0	1.0	0.0	0.0	0.0	1.0	0.0	0.0	1.0
4	4	DiamondRock Hospitality Company	US2527843013	DRH	USA	Real Estate	Hotel & Resort REITs	https://investor.drhc.com/static-files/52de73f...	['2021 CORPORATE RESPONSIBILITY REPORT ##PAGE_...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0

2.4. The model used for each question - FYI Only

The following questions requires machine learning models and are dependent on using the `Labelled_Answers.csv` in the training dataset. The following ML models would be ran on the

dataset after conducting the relevant *data cleaning* and *feature engineering*.

Questions	Feature Engineering	Machine Learning Model
Question 3	Count Vectorizer + Vader Sentiment Analysis	DecisionTree
Question 4	Count Vectorizer + Vader Sentiment Analysis	RandomForest
Question 7	Count Vectorizer + Vader Sentiment Analysis	ExtraTrees
Question 9	TFIDF Vectorizer	ExtraTrees
Question 10	Count Vectorizer + Vader Sentiment Analysis	RandomForest
Question 11	Count Vectorizer	GradientBoosting
Question 12	Count Vectorizer + Vader Sentiment Analysis	GradientBoosting
Question 13	Count Vectorizer + Vader Sentiment Analysis	LogisticRegression
Question 14	Count Vectorizer	GradientBoosting

2.5. Questions that do not need a machine learning model - FYI Only

The following questions do not require using the `Labelled_Answers.csv` in the training dataset. However *Computer Vision* techniques, *Heuristics* as well as *Cosine Similarity* are used to obtain answes to the following questions.

Questions	Techniques	Output
Question 1	Computer Vision + Heuristics	Page number + cropped table image
Question 2	Computer Vision + Heuristics	Page number + cropped table image
Question 5	Cosine Similarity	Up to 3 sentences
Question 6	Cosine Similarity	Up to 3 sentences

3. Features

ℹ️ **Notes about the dashboard:**

- The default definition of `peer` in this dashboard refers to **all the companies** that have released their Sustainability Report in the **current year**.
  For example, if `Year` selected is `2021`, peer refers to all the companies that have released their Sustainability Report in `2021`.
- The term *company* and *client* are used synonymously.

3.1. Client Selector dropdown

Provides filtering function which filters for the desired company's sustainability performance as well as peer's performance. Types of filters as follows:

3.1.1. Company dropdown

The `company` dropdown allows users to filter for their desired company. This dropdown is **compulsory**.

3.1.2. Year

The `year` dropdown allows user to filter for the sustainability performance of their desired company in the specified year. This dropdown is **compulsory**.

ℹ️ **Notes about the following dropdown:**

- These dropwdowns are **optional** dropdowns.
- These dropdowns are meant to change the definition of `peer`. For example, if a `Sector` is selected, `peer` will be defined as those companies in sector `Sector` that have released their sustainability report in year `Year`.
- The intended effect of these dropdowns will be *accumulative*. For example, if `Sector` and `Country` dropdowns are selected, `peer` will be defined as those companies in sector `Sector` and country `Country` that have released their sustainability report in year `Year`.
- If `Sub-Sector` is selected, the option chosen in `Sector` will be ignored.

3.1.3. Sector

This feature allows filtering to compare against peers in the specified `Sector`.

3.1.4. Sub-Sector

This feature allows filtering to compare against peers in the specified `Sub-Sector`. Please take note of the side effects as mentioned above.

3.1.5. Country

This feature allows filtering to compare against peers in the specified `Country`.

3.1.6. Comparison

This feature allows side by side comparison of the sustainability performance of 2 different companies.

3.2. Company Details

Shows the details of the selected `Client`. Details include the `Client's` *Company Name*, *ISIN*, *Sector*, *Sub-Sector*, *Country* as well as the link to the Sustainability Report. When `Year` is selected, *Report Year* will also display the selected `Year`.

Company Details
Company Name: Advantage Energy Ltd. (Ticker: AAV)
ISIN: CA00791P1071
Report Year: 2021
Sector: Energy; Sub-Sector: Oil & Gas Exploration & Production
Country: Canada
Sustainability Report: View

3.3. Current Emissions

This appears as a modal (i.e pop-out) to the user. This page contains information such as the tables as well as the page number to scope 1-3 emissions if any. It also contains other binary questions. The following is the list of questions that are answered in this page.

- Question 1. What are your absolute Scope 1-2 emissions?
- Question 2. What are your absolute Scope 3 emissions?
- Question 3. Have your scope 1-3 emissions been verified by a third party?
- Question 4. Does the company have an active program to support increasing green space or promote biodiversity?

3.3.1. Tables for Scope 1-3 emissions

The image above shows when there is a table containing information to scope 1-3 information found in the report. In the case when no such tables are found due to reasons such as the report not containing such information, the dashboard will display the following image:

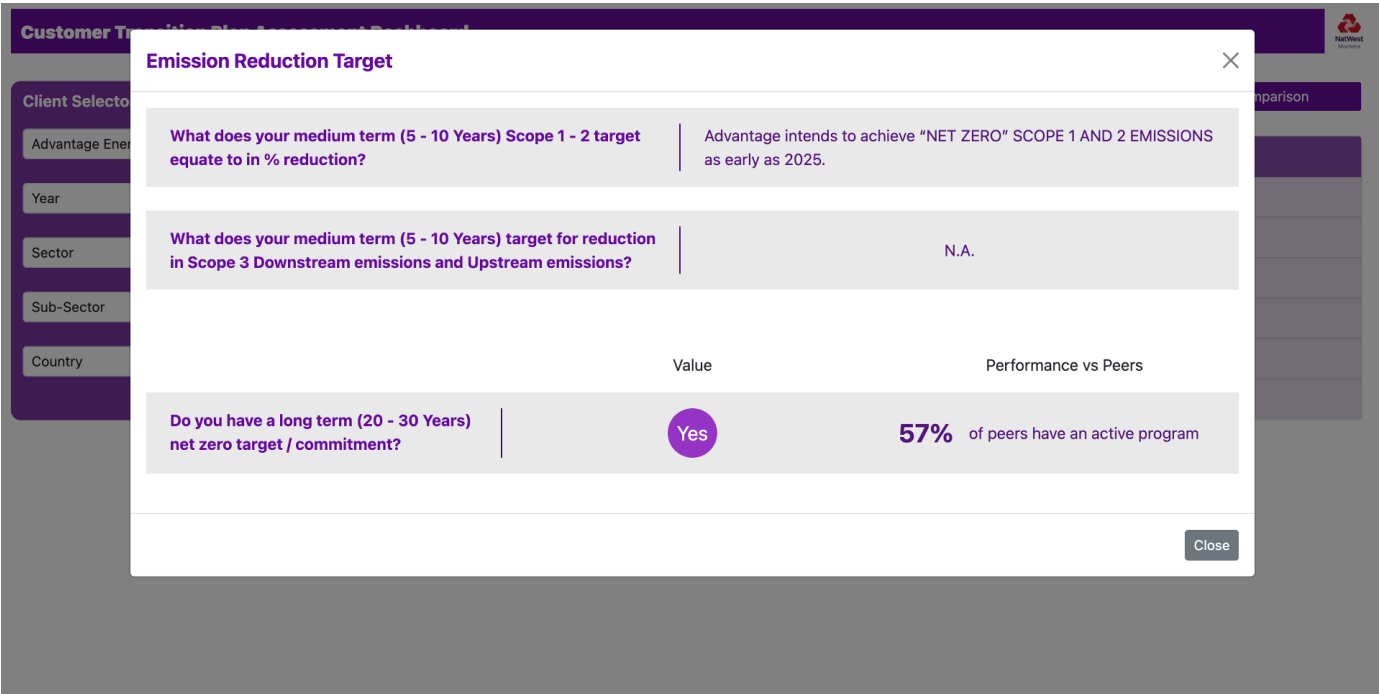To reduce scrolling, users can choose to collapse the pane.

3.3.2. Binary questions

If no answers are found in the sustainability report, value will be `N.A.`. Otherwise, the value will be `Yes`.

3.4. Emission Reduction Target

This appears as a modal (i.e pop-out) to the user. This page contains information such as medium term scope 1-3 target which is presented in the form of paragraph if any. It also contains other binary questions. The following is the list of questions that are answered in this page.

- Question 5. What does your medium term (5 - 10 Years) Scope 1 - 2 target equate to in % reduction?
- Question 6. What does your medium term (5 - 10 Years) target for reduction in Scope 3 Downstream emissions and Upstream emissions?
- Question 7. Do you have a long term (20 - 30 Years) net zero target / commitment?

3.4.1. Medium term Scope 1-3 emissions target

If such information is found in the report, the paragraph containing the information will be displayed. Otherwise, will display `N.A.`.
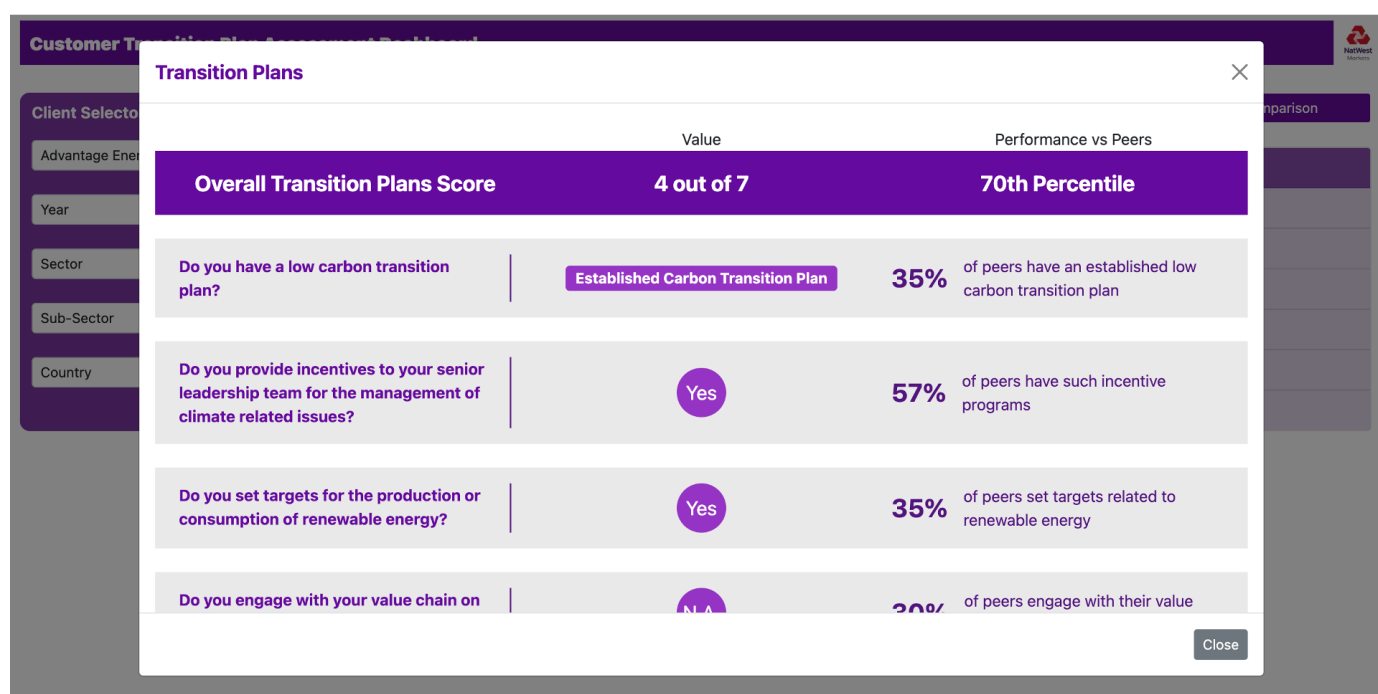
3.4.2. Binary questions

Same as here.

3.5 Transition Plans

This appears as a modal (i.e pop-out) to the user. This page contains information to the overall transition plan score and ranking amongst peer. Apart from binary questions, it also contains a scale question. The following is the list of questions that are answered in this page.

- Question 8. Do you have a low carbon transition plan?
- Question 9. Do you provide incentives to your senior leadership team for the management of climate related issues?
- Question 10. Do you set targets for the production or consumption of renewable energy?
- Question 11. Do you engage with your value chain on climate related issues?
- Question 12. Does you set targets for production of more energy efficient or environmentally friendly products (e.g. recyclable)?
- Question 13. Does your transition plan include direct engagement with suppliers to drive them to reduce their emissions, or even switching to suppliers producing low carbon materials?
- Question 14. Do you engage in activities that influence public policy on cimate related issues to support net zero aligned transition through any of the following?

3.5.1. Scale question

Scale question contains `Established Carbon Transition Plan` , `Plans to Transtion to Low Carbon Environment` if there are some information to low carbon transition plan found. Otherwise, `N.A.`

3.5.2. Binary questions

Same as here.

3.5.3. Score calculations

Each `Yes` is awarded 1 point. For the scale question, `Established Carbon Transition Plan` is awarded 1 point, `Plans to Transtion to Low Carbon Environment` is awarded 0.5 point. 0 point otherwise.

❗ **Notes about the Transition Plans:**

- The percentile calculated is based on the ranking of the score calculated.
- While peer's performance value can change when any of the optional dropdown is selected, the **percentile** will always remain the same.

3.6. Comparison

This appears as a modal (i.e pop-out) to the user. This page serves to provide side by side comparison of the sustainability performance between 2 companies. Data displayed here will be all the questions except for questions 1 & 2.

FAQ

Q1:

A1:

Details Coming Soon...

Data Pipeline Summary

Important files:

- `Processed_Corp_List.csv`
- `Labelled_Answers.csv`
- `New_URLs.csv`
- `Consolidated_Output.csv`

Script to execute: `Final_Pipeline.ipynb`

Feature Summary

Feature	Feature Description
Client Selector	Dropdown selections for filtering and analysis purpose
Company Details	Display selected client
Current Emissions	Contains current emissions details which includes tables containing information to scope 1-3 emission values
Emission Reduction Target	Contains emission reduction target details.
Transition Plans	Contains transition plan details and the client's score in this aspect.
Comparison	Allows for side by side comparison of 2 company's sustainability performance