

# Package ‘ouladFormat’

September 17, 2024

**Type** Package

**Title** Loads and Formats the Open University Learning Analytics Dataset for Data Analysis

**Version** 1.1.2

**Description** The Open University Learning Analytics Dataset (OULAD) is available from Kuzilek et al. (2017) <[doi:10.1038/sdata.2017.171](https://doi.org/10.1038/sdata.2017.171)>. The 'ouladFormat' package loads, cleans and formats the OULAD for data analysis (each row of the returned data set is an individual student). The package's main function, `combined_dataset()`, allows the user to choose whether the returned data set includes assessment, demographics, virtual learning environment (VLE), or registration variables etc.

**License** GPL (>= 2)

**Encoding** UTF-8

**RoxygenNote** 7.3.2

**Imports** dplyr, httr, magrittr, stats, tidyr

**NeedsCompilation** no

**Author** Emma Howard [aut, cre] (<<https://orcid.org/0000-0002-8104-1351>>)

**Maintainer** Emma Howard <emhoward@tcd.ie>

**Repository** CRAN

**Date/Publication** 2024-09-17 11:50:04 UTC

## Contents

<code>combined_dataset</code> . . . . .	2
<code>convert_VLE</code> . . . . .	4
<code>dataset_assessment</code> . . . . .	5
<code>dataset_demographics</code> . . . . .	8
<code>dataset_registration</code> . . . . .	9
<code>dataset_VLE_activity</code> . . . . .	10
<code>dataset_VLE_time</code> . . . . .	12
<code>load_github_modified</code> . . . . .	14
<code>path_to_file</code> . . . . .	15
<code>VLE_learning_classification</code> . . . . .	15

---

combined_dataset	Returns the combined formatted OULAD
------------------	--------------------------------------

---

### Description

Combines multiple OULAD .csv files into one tibble that is formatted for data analysis and where each row represents a unique student.

### Usage

```
combined_dataset(
  module = c("AAA", "BBB", "CCC", "DDD", "EEE", "FFF", "GGG"),
  presentation = c("2013J", "2014J", "2013B", "2014B", "All"),
  repeat_students = c("remove", "keep"),
  withdrawn_students = c("remove", "keep"),
  demographics = FALSE,
  registration = FALSE,
  VLE = c("omit", "daily", "weekly", "activity", "FSLM", "FSLSM", "OLS", "VARK"),
  VLE_clicks = c("total", "binary", "standardise1", "standardise2", "logarithmic"),
  week_begin = -4,
  week_end = 39,
  assessment = FALSE,
  na.rm = FALSE,
  example_data = FALSE
)
```

### Arguments

module	name of the module to be included, either "AAA", "BBB", "CCC", "DDD", "EEE", "FFF" or "GGG".
presentation	name of the semester of the module to be included, either "2013B", "2014B", "2013J", "2014J", or "All". "B" indicates a February start time whereas "J" indicates an October start time. "All" indicates that all presentations of the module will be included in the returned data.
repeat_students	indicator of whether students who had previous attempts at the module should be removed, either "remove" or "keep". When presentation is set to "All", this value is set to "remove".
withdrawn_students	indicator of whether students who had withdrawn from the module should be removed, either "remove" or "keep". Students are removed based on whether they have withdrawn by the "week_end" value.
demographics	logical. Indicates whether demographic data is included in the outputted data set.
registration	logical. Indicates whether registration data is included in the outputted data set.

VLE	indicates whether VLE data is included in the outputted data set (default is "omit" - VLE data not to be included). If included, the VLE data can be returned based on "weekly" VLE interactions, "daily" VLE interactions or interactions according to Open University activities ("activity"). The VLE data can also be returned as classified under a specific learning model, either "FSLM", "FSLSM", "OLS", or "VARK".
VLE_clicks	indicates the format that the VLE data should be returned as, either "total" views (default), "binary" views, "standardise1" (standardises each variable individually by using the scale function), "standardise2" (performs global standardisation of the data set), or "logarithmic" values.
week_begin	the first semester week of VLE and assessment data to be included in formatted data. Depending on the module presentation, students started to view activities four weeks prior to the initial module start date. Weeks prior to the initial module start are indicated by a negative integer.
week_end	the last semester week of VLE and assessment data to be included in the formatted data. Week 39 is the last week material was viewed (and earlier in some module presentations). This parameter is also used to remove withdrawn students.
assessment	logical. Indicates whether assessment data is included in the outputted data set.
na.rm	logical. Indicates whether NAs should be omitted from the average continuous assessment calculations or treated as zeroes (default). This calculation only includes continuous assessment that was due between the period set by "week_begin" and "week_end" inclusive, and only occurs when a specific module presentation is requested (e.g., 'BBB 2013J').
example_data	logical. Indicates whether to run a subset of the VLE data as an example.

### Value

Returns the inputs specified for whether assessment, demographics, registration and VLE variables are to be included, as well as the other inputs. Also, one tibble (object of class `tbl_df`) is returned, `dataset_combined`.

### dataset\_combined tibble

A tibble where each row represents a unique student. Depending on the inputs specified, the tibble includes assessment, demographics, registration and VLE data for each student.

### References

Kuzilek, J., Hlosta, M., & Zdrahal, Z. (2017). Open university learning analytics dataset. *Scientific Data* volume 4, (pp. 1–8). <https://doi.org/10.1038/sdata.2017.171>.

### See Also

For more information on different inputs and variables in the `dataset_combined` tibble, see:

- [dataset\\_assessment\(\)](#) for information on the assessment performance data,
- [dataset\\_demographics\(\)](#) for information on the demographics data,

- `dataset_registration()` for information on the registration data,
- `dataset_VLE_time()` for information on the VLE daily or weekly data,
- `dataset_VLE_activity()` for information on the VLE activity data,
- `VLE_learning_classification()` for information on the VLE activities classified under a learning model, and
- `convert_VLE()` for information on transforming the data type of the VLE data.

### Examples

```
# Uses subset of the VLE data set for example
combined_dataset(module = "AAA", presentation = "2013J",
  repeat_students = "remove", withdrawn_students = "remove",
  demographics = TRUE, registration = TRUE,
  assessment = TRUE, na.rm = FALSE,
  VLE = "weekly", VLE_clicks = "total",
  example_data = TRUE)
```

```
# Slow to run as it loads the full VLE data set
combined_dataset(module = "BBB", presentation = "2013J",
  repeat_students = "remove", withdrawn_students = "remove",
  demographics = TRUE, registration = FALSE,
  assessment = TRUE, na.rm = FALSE,
  VLE = "activity", VLE_clicks = "standardise1",
  week_begin = -4, week_end = 14, example_data = FALSE)
```

---

convert\_VLE

*Converts the VLE data format*

---

### Description

Converts the data format of the Virtual Learning Environment (VLE) data set from total view counts to binary, standardised or logarithmic view count data.

### Usage

```
convert_VLE(
  data,
  conversion = c("binary", "standardise1", "standardise2", "logarithmic")
)
```

### Arguments

<code>data</code>	VLE data set to be converted. The VLE data set should consist of <code>student_id</code> as the first column followed by total view counts per time period/activity in each of the remaining columns.
<code>conversion</code>	type of conversion to implement, either "binary", "standardise1" (standardises each variable individually by using the <code>scale</code> function), "standardise2" (performs global standardisation of the data set), or "logarithmic" (where 1 is added first to the data).

**Value**

Two tibbles are returned: 1) original\_data, and 2) converted\_data.

**original\_data tibble**

A tibble of the original data inputted into the function.

**converted\_data tibble**

A tibble where all columns, except the first column, have been changed from total view counts to either binary, standardised, globally standardised or logarithmic view count data.

**See Also**

[dataset\\_VLE\\_time\(\)](#) or [dataset\\_VLE\\_activity\(\)](#) for obtaining the VLE data set needed for the input data.

**Examples**

```
VLE_data = dataset_VLE_activity(example_data = TRUE)$resource_data
convert_VLE(VLE_data, conversion = "standardise1")
```

```
VLE_data = dataset_VLE_time(example_data = TRUE)$weekly_data
convert_VLE(VLE_data, conversion = "logarithmic")
```

---

dataset_assessment	<i>Returns the formatted assessment data sets</i>
--------------------	---

---

**Description**

Load, combines and formats the assessment and student assessment data sets from the OULAD for data analysis.

**Usage**

```
dataset_assessment(
  module = c("All", "AAA", "BBB", "CCC", "DDD", "EEE", "FFF", "GGG"),
  presentation = c("2013B", "2014B", "2013J", "2014J", "All", "Summer", "Winter"),
  repeat_students = c("remove", "keep"),
  week_begin = 1,
  week_end = 39,
  na.rm = FALSE
)
```

## Arguments

module	name of the module to be included, either "All", "AAA", "BBB", "CCC", "DDD", "EEE", "FFF" or "GGG".
presentation	name of the semester of the module to be included, either "2013B", "2014B", "2013J", "2014J", "All", "Summer" or "Winter". "B" indicates a February start time whereas "J" indicates an October start time. "All" indicates that all presentations of the module will be included in the returned data. Where possible, "Summer" returns both "2013B" and "2014B", and "Winter" returns both "2013J" and "2014J".
repeat_students	indicator of whether students who had previous attempts at the module should be removed, either "remove" or "keep".
week_begin	the first semester week of the assessment data to be included in formatted data.
week_end	the last semester week of the assessment data to be included in the formatted data. To ensure all continuous assessment is included regardless of the module presentation, set this to 39, the last week of data in the OULAD.
na.rm	logical. Indicates whether NAs should be omitted from the average continuous assessment calculations or treated as zeroes (default). This calculation only includes continuous assessment that were due between the period set by "week_begin" and "week_end" inclusive, and only occurs when a specific module presentation is requested (e.g., 'BBB 2013J').

## Value

Returns four tibbles (objects of class `tbl_df`) based on the OULAD `studentAssessment.csv` and `assessments.csv` files and the specified inputs (module, presentation, and repeat\_students). The four tibbles are: 1) `assessment_data`, 2) `assessments`, 3) `assessment_performance`, and 4) `assessment_reactivity`.

### assessment\_data tibble

A tibble based on the combined OULAD files of `studentAssessment.csv` and `assessments.csv`, and the inputs. The tibble consists of (Kuzilek et al., 2017):

- `id_student` - unique student identification number.
- `code_module` - module identification code.
- `code_presentation` - module presentation identification code.
- `id_assessment` - assessment identification number.
- `assessment_type` - type of assessment.
- `date` - information about the final submission date of the assessment calculated as the number of days since the start of the module-presentation.
- `weight` - weight of the assessment in %.
- `date_submitted` - day of the assessment submission by student.
- `is_banked` - logical indicator whether assessment result has been transferred from a previous presentation.
- `score` - student's score for the specific assessment (range 0-100).
- `reactivity` - calculated using the `date` variable minus `date_submitted` variable.

**assessments tibble**

A tibble of seven columns which details for each assessment item for a module presentation, the type of assessment it is, the due date and week of the assessment, and the weight of the assessment. The assessment types are Tutor Marked Assessment (TMA), Computer Marked Assessment (CMA) and Final Exam (Exam). Exams are generally treated separately to other assessments and have weight equal to 100%. The sum of all other assessments is also 100% (Kuzilek et al., 2017).

**assessment\_performance tibble**

A tibble where each row represents a unique student and their scores in the range of 0-100 for different assessment items. When a specific module (not 'All') and presentation (not 'All' or 'Winter' or 'Summer') is selected, the final column states each student's average continuous assessment score based on their continuous assessment scores and the assessments weights. For this calculation, the NAs may be not included or replaced by 0 (default; see na.rm argument). The average assessment score will not be calculated if the only variable outputted is exam score. Not all module presentations have exam scores available. Details of the assessment items can be seen in the assessments tibble.

**assessment\_reactivity tibble**

A tibble where each row represents a unique student and their reactivity for different assessment items. Treuillier and Boyer (2021) define reactivity for the OULAD as the delay between the date the assessment is returned and the deadline (in days). Negative numbers indicate overdue assessments. Details of the assessment items can be seen in the assessments tibble.

**References**

Treuillier, C., & Boyer, A. (2021). Identification of class-representative learner personas. In LA4SLE 2021 - Learning Analytics for Smart Learning Environments (pp. 38-45). Bolzano, Italy.

Kuzilek, J., Hlosta, M., & Zdrahal, Z. (2017). Open university learning analytics dataset. Scientific Data volume 4, (pp. 1–8). <https://doi.org/10.1038/sdata.2017.171>.

**See Also**

[combined\\_dataset\(\)](#)

**Examples**

```
dataset_assessment(module = "BBB", presentation = "2013J",  
repeat_students = "remove", week_begin = 1, week_end=39, na.rm = FALSE)
```

---

dataset\_demographics *Returns the formatted demographic data set*

---

## Description

Load and formats the student demographic data set from the OULAD for data analysis.

## Usage

```
dataset_demographics(  
  module = c("All", "AAA", "BBB", "CCC", "DDD", "EEE", "FFF", "GGG"),  
  presentation = c("2013B", "2014B", "2013J", "2014J", "All", "Summer", "Winter"),  
  repeat_students = c("remove", "keep")  
)
```

## Arguments

module	name of the module to be included, either "All", "AAA", "BBB", "CCC", "DDD", "EEE", "FFF" or "GGG".
presentation	name of the semester of the module to be included, either "2013B", "2014B", "2013J", "2014J", "All", "Summer" or "Winter". "B" indicates a February start time whereas "J" indicates an October start time. "All" indicates that all presentations of the module will be included in the returned data. "Summer" returns both "2013B" and "2014B". "Winter" returns both "2013J" and "2014J".
repeat_students	indicator of whether students who had previous attempts at the module should be removed, either "remove" or "keep".

## Value

Returns one tibble (object of class `tbl_df`), called 'studentInfo', based on the OULAD `studentInfo.csv` file and the specified inputs (module, presentation, and repeat\_students).

The tibble consists of 12 columns (Kuzilek et al., 2017):

- `code_module` - module identification code.
- `code_presentation` - module presentation identification code.
- `id_student` - the unique student identification number.
- `gender` - student's gender, either Male or Female.
- `region` - the geographic region where the student lived while taking the module-presentation.
- `highest_education` - the highest student education level on entry to the module presentation.
- `imd_band` - the index of multiple deprivation band of the place where the student lived during the module-presentation.
- `age_band` - a band of student's age.

- num\_of\_prev\_attempts - the number of times the student has attempted this module previously.
- studied\_credits - the total number of credits for the modules the student is currently studying.
- disability - indicates whether the student has declared a disability.
- final\_result - student's final result in the module-presentation.

## References

Kuzilek, J., Hlosta, M., & Zdrahal, Z. (2017). Open university learning analytics dataset. Scientific Data volume 4, (pp. 1–8). <https://doi.org/10.1038/sdata.2017.171>.

## See Also

[combined\\_dataset\(\)](#)

## Examples

```
dataset_demographics(module = "BBB", presentation = "2013J", repeat_students = "remove")
```

---

`dataset_registration` *Returns the formatted registration data set*

---

## Description

Load and formats the student registration data set from the OULAD for data analysis.

## Usage

```
dataset_registration(
  module = c("All", "AAA", "BBB", "CCC", "DDD", "EEE", "FFF", "GGG"),
  presentation = c("2013B", "2014B", "2013J", "2014J", "All", "Summer", "Winter"),
  repeat_students = c("remove", "keep")
)
```

## Arguments

module	name of the module to be included, either "All", "AAA", "BBB", "CCC", "DDD", "EEE", "FFF" or "GGG".
presentation	name of the semester of the module to be included, either "2013B", "2014B", "2013J", "2014J", "All", "Summer" or "Winter". "B" indicates a February start time whereas "J" indicates an October start time. "All" indicates that all presentations of the module will be included in the returned data. Where possible, "Summer" returns both "2013B" and "2014B", and "Winter" returns both "2013J" and "2014J".
repeat_students	indicator of whether students who had previous attempts at the module should be removed, either "remove" or "keep".

**Value**

Returns one tibble (object of class `tbl_df`), called 'studentRegistration', based on the OULAD studentRegistration.csv file and the specified inputs (module, presentation, and repeat\_students).

The tibble consists of five columns (Kuzilek et al., 2017):

- `id_student` - the unique student identification number.
- `code_module` - module identification code.
- `code_presentation` - module presentation identification code.
- `date_registration` - the day of student's registration for the module presentation. Modules start on day 0.
- `date_unregistration` - the day of student unregistration from the module presentation. This is NA if the student completed the module presentation.

**References**

Kuzilek, J., Hlosta, M., & Zdrahal, Z. (2017). Open university learning analytics data set. Scientific Data volume 4 , (pp. 1–8). <https://doi.org/10.1038/sdata.2017.171>.

**See Also**

[combined\\_dataset\(\)](#)

**Examples**

```
dataset_registration(module = "BBB", presentation = "2013J", repeat_students = "remove")
```

---

`dataset_VLE_activity` *Returns the formatted VLE data set based on activity types*

---

**Description**

Load and formats the student Virtual Learning Environment (VLE) data set from the OULAD for data analysis.

**Usage**

```
dataset_VLE_activity(  
  module = c("AAA", "BBB", "CCC", "DDD", "EEE", "FFF", "GGG"),  
  presentation = c("2013B", "2014B", "2013J", "2014J", "All"),  
  repeat_students = c("remove", "keep"),  
  week_begin = -4,  
  week_end = 39,  
  example_data = FALSE  
)
```

**Arguments**

module	name of the module to be included, either "AAA", "BBB", "CCC", "DDD", "EEE", "FFF" or "GGG".
presentation	name of the semester of the module to be included, either "2013B", "2014B", "2013J", "2014J", or "All". "B" indicates a February start time whereas "J" indicates an October start time. "All" indicates that all presentations of the module will be included in the returned data.
repeat_students	indicator of whether students who had previous attempts at the module should be removed, either "remove" or "keep". If presentation is set to "All", this is automatically set to "remove".
week_begin	the first semester week of VLE data to be included in the formatted data. Depending on the module presentation, students started to view activities up to four weeks prior to the initial module start date. Weeks prior to the initial module start are indicated by a negative integer.
week_end	the last semester week of VLE data to be included in the outputted data. Week 39 is the last week material was viewed (and earlier in some module presentations).
example_data	logical. Indicates whether to run a subset of the data as an example.

**Value**

Returns two tibbles (objects of class `tbl_df`) based on the OULAD `studentVle.csv` and `vle.csv` files, the specified inputs (module, presentation, and repeat\_students), and the range of the weeks included in the tibbles. `week_begin` and `week_end` indicates the first and last semester week respectively that is included in the output tibbles. These may be different to the corresponding input parameters. Weeks prior to the initial module start day are indicated by a negative integer. The two tibbles returned are: 1) `filtered_data`, and 2) `resource_data`.

**filtered\_data tibble**

A tibble based on the combined oulad files of `studentVLE.csv` and `vle.csv`, and the inputs. The tibble consists of (Kuzilek et al., 2017):

- `code_module` - the module identification code.
- `code_presentation` - the presentation identification code.
- `id_site` - the VLE material identification number.
- `activity_type` - the role associated with the module material.
- `week_from` - the week from which the material is planned to be used.
- `week_to` - the week until which the material is planned to be used.
- `id_student` - the unique student identification number.
- `date` - the day of student's interaction with the material (module starts on day 0).
- `sum_click` - the number of times the student interacted with the material for the date.

**resource\_data tibble**

A tibble where each row represents a unique student and each column their number of interactions with different VLE activity types for the weeks and module presentation inputted.

**References**

Kuzilek, J., Hlosta, M., & Zdrahal, Z. (2017). Open university learning analytics dataset. Scientific Data volume 4, (pp. 1–8). <https://doi.org/10.1038/sdata.2017.171>.

**See Also**

[convert\\_VLE\(\)](#), [dataset\\_VLE\\_time\(\)](#), [VLE\\_learning\\_classification\(\)](#) and [combined\\_dataset\(\)](#)

**Examples**

```
# Uses subset of the VLE data set for example
dataset_VLE_activity(example_data = TRUE)

# Slow to run as it loads the full VLE data set
dataset_VLE_activity(module = "BBB", presentation = "2013J",
  repeat_students = "remove", week_begin = 1, week_end = 13,
  example_data=FALSE)
```

---

dataset_VLE_time	<i>Returns the formatted VLE data set for daily and weekly counts</i>
------------------	---

---

**Description**

Load and formats the student Virtual Learning Environment (VLE) data set from the OULAD for data analysis.

**Usage**

```
dataset_VLE_time(
  module = c("AAA", "BBB", "CCC", "DDD", "EEE", "FFF", "GGG"),
  presentation = c("2013B", "2014B", "2013J", "2014J", "A11"),
  repeat_students = c("remove", "keep"),
  week_begin = -4,
  week_end = 39,
  example_data = FALSE
)
```

**Arguments**

module	name of the module to be included, either "AAA", "BBB", "CCC", "DDD", "EEE", "FFF" or "GGG".
--------	--

presentation	name of the semester of the module to be included, either "2013B", "2014B", "2013J", "2014J", or "All". "B" indicates a February start time whereas "J" indicates an October start time. "All" indicates that all presentations of the module will be included in the returned data.
repeat_students	indicator of whether students who had previous attempts at the module should be removed, either "remove" or "keep". If presentation is set to "All", this is automatically set to "remove".
week_begin	the first semester week of VLE data to be included in formatted data. Depending on the module presentation, students started to view activities four weeks prior to the initial module start date. Weeks prior to the initial module start are indicated by a negative integer.
week_end	the last semester week of VLE data to be included in the formatted data. Week 39 is the last week material was viewed (and earlier in some module presentations).
example_data	logical. Indicates whether to run a subset of the data as an example.

### Value

Returns three tibbles (objects of class `tbl_df`) based on the OULAD `studentVle.csv` file, the specified inputs (module, presentation, and `repeat_students`), and the range of the weeks included in the tibbles. `week_begin` and `week_end` indicates the first and last semester week respectively that is included in the output tibbles. These may be different to the input parameters of `week_begin` and `week_end`. Weeks prior to the initial module start day are indicated by a negative integer. The three tibbles returned are: 1) `filtered_data`, 2) `daily_data`, and 3) `weekly_data`.

### filtered\_data tibble

A tibble based on the students' VLE interactions and the inputs. The tibble consists of (Kuzilek et al., 2017):

- `id_student` - the unique student identification number.
- `code_module` - the module identification code.
- `code_presentation` - the presentation identification code.
- `id_site` - the VLE material identification number.
- `date` - the day of the student's interaction with the material.
- `sum_click` - the number of times the student interacted with the material on a specific day.

### daily\_data tibble

A tibble where each row represents a unique student, and each column their number of interactions with the VLE for different days of the module. The module starts on day 0. Days prior to the initial start day, are indicated by negative integers.

### weekly\_data tibble

A tibble where each row represents a unique student, and each column their number of interactions with the VLE for different weeks of the module. Weeks prior to the initial start week (week 1) are indicated by the term 'pre' in the week name.

**References**

Kuzilek, J., Hlosta, M., & Zdrahal, Z. (2017). Open university learning analytics dataset. Scientific Data volume 4 , (pp. 1–8). <https://doi.org/10.1038/sdata.2017.171>.

**See Also**

[convert\\_VLE\(\)](#), [dataset\\_VLE\\_activity\(\)](#), and [combined\\_dataset\(\)](#)

**Examples**

```
# Uses subset of the VLE data set for example
dataset_VLE_time(example_data = TRUE)

# Slow to run as it loads the full VLE data set
dataset_VLE_time(module = "BBB", presentation = "2013J",
repeat_students = "remove", week_begin = 1, week_end = 13,
example_data=FALSE)
```

---

load\_github\_modified *A modified version of the Rfssa load\_github\_data() function*

---

**Description**

Loads .RData files directly from a GitHub repository. The original Rfssa function (Haghbin et al., 2022) is based on the Stack Overflow post (Hartman, 2016): <https://stackoverflow.com/questions/24846120/importing-data-into-r-rdata-from-github>, and loads the .RData to the global environment. This function allows the user to choose the environment which the .RData files are loaded to.

**Usage**

```
load_github_modified(github_data_url, env)
```

**Arguments**

```
github_data_url      The GitHub url of the .RData data set to be loaded.
env                  Environment where the data is to be loaded into.
```

**Value**

Loads the .RData into the local environment.

**References**

Haghbin, H., Trinkka, J., Najibi, S., M., & Maadooliat, M. (2022). Package 'Rfssa'. <https://cran.r-project.org/web/packages/Rfssa/Rfssa.pdf>.

Hartman, R. (2016). Importing data into R (.rData) from Github. <https://stackoverflow.com/questions/24846120/importing-data-into-r-rdata-from-github>.

**Examples**

```
env = environment()
load_github_modified("https://github.com/ehoward1/oulad_data/blob/main/vle.RData", env)
```

---

path_to_file	<i>Get file path to the sample_studentVLE.csv file</i>
--------------	--

---

**Description**

ouladFormat comes bundled with the sample studentVLE data in its inst/extdata directory (sample\_studentVLE.csv). This .csv file is a subset (rows 10,000-14,999) of the original OULAD studentVLE.csv. This function makes it easy to access.

**Usage**

```
path_to_file(path = NULL)
```

**Arguments**

path                    Name of file in quotes with extension. If NULL, the example file will be listed.

**Value**

No return value, called for side effects.

**Source**

This function is adapted from readxl::readxl\_example().

**Examples**

```
path_to_file()
read.csv(path_to_file("sample_studentVLE.csv"))
```

---

VLE_learning_classification	<i>Classify VLE activities</i>
-----------------------------	--------------------------------

---

**Description**

Transform Open University VLE activities to classifications under either the:

- Filder-Selverman Learning Style Model (FSLM) mapped by Balti et al. (2023),
- Felder–Silverman Learning Style Model (FSLSM) (see Felder & Silverman, 1988) mapped by Nazempour and Darabi (2023),
- VARK Learning Style Model (see Fleming and Colleen, 1992) mapped by Balti et al. (2023),  
or
- Corresponding to the Online Learning Style Characteristics (OLS) mapped by Yan et al. (2021) and drawing on Li and Yin (2017).

**Usage**

```
VLE_learning_classification(
  activity_data,
  classification = c("FSLM", "FSLSM", "OLS", "VARK")
)
```

**Arguments**

`activity_data` VLE activity data set to be converted.

`classification` classification system to be used to categorise the different VLE activities under.

**Value**

Returns the classification specified, the mapping for the VLE activities to each category in the specified learning model, and one tibble, `VLE_classified_data`.

**VLE\_classified\_data**

A tibble where each row represents a unique student and their number of interactions with different categories of the specified learning model.

**References**

- Balti, R., Hedhili, A., Chaari, W.L. & Abed, M. (2023). Hybrid analysis of the learner's online behavior based on learning style. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-023-11595-x>
- Felder, R. M., & Silverman, L. K. (1988). Learning and Teaching Styles in Engineering Education. *Engineering Education*, 78(7), 674–681. <https://www.engr.ncsu.edu/wp-content/uploads/drive/1QP6kBIiQmpQbTXL-08HSI0PwJ5BYnZW/1988-LS-plus-note.pdf>
- Fleming, N. D., & Mills, C. (1992). Not Another Inventory, Rather a Catalyst for Reflection. *To Improve the Academy*. 246. <https://digitalcommons.unl.edu/podimproveacad/246>
- Kuzilek, J., Hlosta, M., & Zdrahal, Z. (2017). Open university learning analytics dataset. *Scientific Data* volume 4 , (pp. 1–8). <https://doi.org/10.1038/sdata.2017.171>.
- Li, R., & Yin, C. (2017). Analysis of Online Learning Style Model Based on K-means Algorithm. *Proceedings of the 3rd International Conference on Economics, Management, Law and Education (EMLE 2017)*. (pp. 692-697). Atlantis Press. <https://doi.org/10.2991/emle-17.2017.148>
- Nazempour, R., & Darabi, H. (2023). Personalized Learning in Virtual Learning Environments Using Students' Behavior Analysis. *Education Sciences*. 2023; 13(5), 457. <https://doi.org/10.3390/educsci13050457>
- Yan, L., Yin, C., Chen, H., Rong, W., Xiong, Z., & David, B. (2021). Learning Resource Recommendation in E-Learning Systems Based on Online Learning Style. In: Qiu, H., Zhang, C., Fei, Z., Qiu, M., Kung, SY. (eds) *Knowledge Science, Engineering and Management. KSEM 2021. Lecture Notes in Computer Science*, 12817. Springer, Cham. [https://doi.org/10.1007/978-3-030-82153-1\\_31](https://doi.org/10.1007/978-3-030-82153-1_31)

**See Also**

[dataset\\_VLE\\_activity\(\)](#) for obtaining the VLE data set needed for input data

**Examples**

```
VLE_data = dataset_VLE_activity(example_data = TRUE)$resource_data  
VLE_learning_classification(VLE_data, classification = "FSLSM")
```

# Index

`combined_dataset`, [2](#)  
`combined_dataset()`, [7](#), [9](#), [10](#), [12](#), [14](#)  
`convert_VLE`, [4](#)  
`convert_VLE()`, [4](#), [12](#), [14](#)

`dataset_assessment`, [5](#)  
`dataset_assessment()`, [3](#)  
`dataset_demographics`, [8](#)  
`dataset_demographics()`, [3](#)  
`dataset_registration`, [9](#)  
`dataset_registration()`, [4](#)  
`dataset_VLE_activity`, [10](#)  
`dataset_VLE_activity()`, [4](#), [5](#), [14](#), [16](#)  
`dataset_VLE_time`, [12](#)  
`dataset_VLE_time()`, [4](#), [5](#), [12](#)

`load_github_modified`, [14](#)

`path_to_file`, [15](#)

`VLE_learning_classification`, [15](#)  
`VLE_learning_classification()`, [4](#), [12](#)