



Online Retail

Donated on 11/5/2015

This is a transactional data set which contains all the transactions occurring between 01/12/2010 and 09/12/2011 for a UK-based and registered non-store online retail.

Dataset Characteristics

Multivariate, Sequential, Time-Series

Subject Area

Business

Associated Tasks

Classification, Clustering

Feature Type

Integer, Real

Instances

541909

Features

6

Dataset Information



Additional Information

This is a transactional data set which contains all the transactions occurring between 01/12/2010 and 09/12/2011 for a UK-based and registered non-store online retail. The company mainly sells unique all-occasion gifts. Many customers of the company are wholesalers.

Has Missing Values?

No

Introductory Paper



[Data mining for the online retail industry: A case study of RFM model-based customer segmentation using data mining](#)

By Daqing Chen, Sai Laing Sain, Kun Guo. 2012

Published in Journal of Database Marketing and Customer Strategy Management, Vol. 19, No. 3

Variables Table



Variable Name	Role	Type	Description	Units	Missing Values
InvoiceNo	ID	Categorical	a 6-digit integral number uniquely assigned to each transaction. If this code starts with letter 'c', it indicates a cancellation		no
StockCode	ID	Categorical	a 5-digit integral number uniquely assigned to each distinct product		no
Description	Feature	Categorical	product name		no
Quantity	Feature	Integer	the quantities of each product (item) per transaction		no
InvoiceDate	Feature	Date	the day and time when each transaction was generated		no
UnitPrice	Feature	Continuous	product price per unit	sterling	no
CustomerID	Feature	Categorical	a 5-digit integral number uniquely assigned to each customer		no
Country	Feature	Categorical	the name of the country where each customer resides		no

Rows per page 10 ▾ 0 to 8 of 8 < >

Additional Variable Information ^

InvoiceNo: Invoice number. Nominal, a 6-digit integral number uniquely assigned to each transaction. If this code starts with letter 'c', it indicates a cancellation.

StockCode: Product (item) code. Nominal, a 5-digit integral number uniquely assigned to each ...

SHOW MORE ▾

Dataset Files ^

File	Size
Online Retail.xlsx	22.6 MB

Papers Citing this Dataset

 SORT BY YEAR, DESC

[Finding Robust Itemsets Under Subsampling](#)

By Nikolaj Tatti, Fabian Moerchen, Toon Calders. 2019
Published in ArXiv.

[Moment-Based Quantile Sketches for Efficient High Cardinality Aggregation Queries](#)

By Edward Gan, Jialin Ding, Kai Tai, Vatsal Sharan, Peter Bailis. 2018
Published in ArXiv.

[Efficient and Scalable Multi-task Regression on Massive Number of Tasks](#)

By Xiao He, Francesco Alesiani, Ammar Shaker. 2018
Published in ArXiv.

[Efficient Mining Top-k Regular-Frequent Itemset Using Compressed Tidsets](#)

By Komate Amphawan, Philippe Lenca, Athasit Surarerks. 2011
Published in PAKDD Workshops.

[Layered critical values: a powerful direct-adjustment approach to discovering significant patt...](#)

By Geoffrey Webb. 2008
Published in Machine Learning.

Reviews

5.0  (1 rating)

LOGIN TO WRITE A REVIEW

Sugiyamto Gik



-

12/23/2024

 **DOWNLOAD** (22.6 MB)



IMPORT IN PYTHON

CITE

 **8 citations**

 **284073 views**

Keywords

sales

Creators

Daqing Chen



chend@lsbu.ac.uk

School of Engineering, London South Bank University

DOI

[10.24432/C5BW33](https://doi.org/10.24432/C5BW33)

License

This dataset is licensed under a [Creative Commons Attribution 4.0 International](#) (CC BY 4.0) license.

This allows for the sharing and adaptation of the datasets for any purpose, provided that the appropriate credit is given.



UC Irvine
Machine Learning
Repository

THE PROJECT

About Us

CML

National Science Foundation

NAVIGATION

[Home](#)

[View Datasets](#)

[Donate a Dataset](#)

LOGISTICS

[Contact](#)

[Privacy Notice](#)

[Feature Request or Bug Report](#)