UC Irvine
Machine Learning
Repository

Datasets   Contribute Dataset   About Us   Sea   → Login

Online Retail
Donated on 11/5/2015

This is a transnational data set which contains all the transactions occurring between 01/12/2010 and 09/12/2011 for...

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 transnational 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 ...

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

DOWNLOAD

IMPORT IN PYTHON

CITE

🔖 **8 citations**
👁 **216603 views**

Keywords

(sales)

Creators

Daqing Chen
chend@lsbu.ac.uk
School of Engineering, London South Bank University

DOI

10.24432/C5BW33

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

Variables Table ⌃

Variable Name	Role	Type	Description	Un
InvoiceNo	ID	Categorical	a 6-digit integral number uniquely assigned to each transaction. If this code starts with letter 'c', it indicates a cancellation	
StockCode	ID	Categorical	a 5-digit integral number uniquely assigned to each distinct product	
Description	Feature	Categorical	product name	
Quantity	Feature	Integer	the quantities of each product (item) per transaction	
			the day and time when	

InvoiceDate	Feature	Date	each transaction was generated	
UnitPrice	Feature	Continuous	product price per unit	ste
CustomerID	Feature	Categorical	a 5-digit integral number uniquely assigned to each customer	
Country	Feature	Categorical	the name of the country where each customer resides	

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. ...

SHOW MORE ⌄

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 Cardin...](#)

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

[Efficient and Scalable Multi-task Regression on Massive Nu...](#)

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

[Efficient Mining Top-k Regular-Frequent Itemset Using Co...](#)

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

[Layered critical values: a powerful direct-adjustment appro...](#)

By Geoffrey Webb. 2008
Published in Machine Learning.

Rows per page (5)   0 to 5 of 8   <   ( )

Reviews ⌃

There are no reviews for this dataset yet.   LOGIN TO WRITE A REVIEW

THE PROJECT   NAVIGATION   LOGISTICS