

# Package: onlineretail (via r-universe)

March 11, 2025

**Type** Package

**Title** Online Retail Dataset

**Version** 0.1.2

**Date** 2021-05-13

**Description** Transactions occurring for a UK-based and registered, non-store online retail between 01/12/2010 and 09/12/2011 (Chen et. al., 2012, <[doi:10.1145/1835804.1835882](https://doi.org/10.1145/1835804.1835882)>). This dataset is included in this package with the donor's permission, Dr. Daqing Chen.

**Depends** R (>= 3.5.0)

**License** CC0

**URL** <https://github.com/allanvc/onlineretail/>,  
<https://doi.org/10.1057/dbm.2012.17>,  
<https://www.researchgate.net/profile/Daqing-Chen>

**BugReports** <https://github.com/allanvc/onlineretail/issues/>

**Encoding** UTF-8

**LazyData** true

**RoxygenNote** 7.1.1

**LazyDataCompression** xz

**Repository** <https://allanvc.r-universe.dev>

**RemoteUrl** <https://github.com/allanvc/onlineretail>

**RemoteRef** HEAD

**RemoteSha** 3a77afe42f2650c54221573f9537e45cde646834

## Contents

onlineretail . . . . .	2
<b>Index</b>	<b>3</b>

---

onlinetail

*Online Retail Data Set*

---

### **Description**

This Online Retail dataset contains all the transactions occurring for a UK-based and registered, non-store online retail between 01/12/2010 and 09/12/2011. The company mainly sells unique all-occasion gift-ware. Many customers of the company are wholesalers.

### **Usage**

`data(onlinetail)`

### **Format**

A data frame with eight variables:

**InvoiceNo** A character indicating the invoice number, which is a 6-digit integral number uniquely assigned to each transaction. If this code starts with the letter 'c', it indicates a cancellation.

**StockCode** A character indicating the product (item) code, which is a 5-digit integral number uniquely assigned to each distinct product. It can be accompanied by a trailing uppercase letter.

**Description** A character indicating the Product (item) name.

**Quantity** A numeric indicating the quantities of each product (item) per transaction.

**InvoiceDate** A POSIXct indicating the invoice day and time when a transaction was generated.

**UnitPrice** A numeric indicating the product price per unit in sterling (£)

**CustomerID** A numeric indicating the customer number, which is a 5-digit integral number uniquely assigned to each customer.

**Country** A character indicating the name of the country where a customer resides.

### **Source**

[UCI Machine Learning Repository](#)

### **References**

Daqing Chen, Sai Liang Sain, and Kun Guo (2012), Data mining for the online retail industry: A case study of RFM model-based customer segmentation using data mining. *Journal of Database Marketing and Customer Strategy Management*, Vol. 19, No. 3, pp. 197-208, 2012 (Published online before print: 27 August 2012. doi: 10.1057/dbm.2012.17).

### **Examples**

`data(onlinetail)`

# Index

\* **datasets**

onlinetail, 2

onlinetail, 2