




[About](#)
[Citation Policy](#)
[Donate a Data Set](#)
[Contact](#)

☐ Repository
 ☒ Web
 

**Machine Learning Repository**
[View ALL Data Sets](#)

[Center for Machine Learning and Intelligent Systems](#)

Check out the [beta version](#) of the new UCI Machine Learning Repository we are currently testing! ✕  
[Contact us](#) if you have any issues, questions, or concerns. [Click here to try out the new site.](#)

## Wholesale customers Data Set

Download: [Data Folder](#), [Data Set Description](#)

**Abstract:** The data set refers to clients of a wholesale distributor. It includes the annual spending in monetary units (m.u.) on diverse product categories

<b>Data Set Characteristics:</b>	Multivariate	<b>Number of Instances:</b>	440	<b>Area:</b>	Business
<b>Attribute Characteristics:</b>	Integer	<b>Number of Attributes:</b>	8	<b>Date Donated</b>	2014-03-31
<b>Associated Tasks:</b>	Classification, Clustering	<b>Missing Values?</b>	N/A	<b>Number of Web Hits:</b>	430383

### Source:

Margarida G. M. S. Cardoso, [margarida.cardoso '@' iscte.pt](mailto:margarida.cardoso '@' iscte.pt), ISCTE-IUL, Lisbon, Portugal

### Data Set Information:

Provide all relevant information about your data set.

### Attribute Information:

- 1) FRESH: annual spending (m.u.) on fresh products (Continuous);
- 2) MILK: annual spending (m.u.) on milk products (Continuous);
- 3) GROCERY: annual spending (m.u.) on grocery products (Continuous);
- 4) FROZEN: annual spending (m.u.) on frozen products (Continuous)
- 5) DETERGENTS\_PAPER: annual spending (m.u.) on detergents and paper products (Continuous)
- 6) DELICATESSEN: annual spending (m.u.) on delicatessen products (Continuous);
- 7) CHANNEL: customers' Channel - Horeca (Hotel/Restaurant/Café) or Retail channel (Nominal)
- 8) REGION: customers' Region - Lisbon, Oporto or Other (Nominal)

Descriptive Statistics:

(Minimum, Maximum, Mean, Std. Deviation)

FRESH ( 3, 112151, 12000.30, 12647.329)

MILK (55, 73498, 5796.27, 7380.377)

GROCERY (3, 92780, 7951.28, 9503.163)

FROZEN (25, 60869, 3071.93, 4854.673)

DETERGENTS\_PAPER (3, 40827, 2881.49, 4767.854)

DELICATESSEN (3, 47943, 1524.87, 2820.106)

**REGION Frequency**

Lisbon 77

Oporto 47

Other Region 316

Total 440

**CHANNEL Frequency**

Horeca 298

Retail 142

Total 440

**Relevant Papers:**

Cardoso, Margarida G.M.S. (2013). Logical discriminant models – Chapter 8 in Quantitative Modeling in Marketing and Management Edited by Luiz Moutinho and Kun-Huang Huarng. World Scientific. p. 223-253. ISBN 978-9814407717

Jean-Patrick Baudry, Margarida Cardoso, Gilles Celeux, Maria Jos  Amorim, Ana Sousa Ferreira (2012). Enhancing the selection of a model-based clustering with external qualitative variables. RESEARCH REPORT N  8124, October 2012, Project-Team SELECT. INRIA Saclay -  le-de-France, Projet select, Universit  Paris-Sud 11

**Citation Request:**

The data set is originated from a larger database referred on:

Abreu, N. (2011). Analise do perfil do cliente Recheio e desenvolvimento de um sistema promocional. Mestrado em Marketing, ISCTE-IUL, Lisbon

Supported By:



In Collaboration With:



[About](#) || [Citation Policy](#) || [Donation Policy](#) || [Contact](#) || [CML](#)