

*Project PEACH is a large scale open source community driven Data Science project by the Computer Science Department at University College London. The goal of Project PEACH is to provide a data science tool for medical professionals and researchers that can aid them in diagnostic and analytics processes through the use of big data, machine learning and data visualisation. <https://code-4-health.org/peach>*

Project Supervisors: Dr. Dean Mohamedally, Dr. Atia Rafiq, Marios Constantinides, George Spithourakis

Team Members: Nikolaos Georgiadis, Narmin Huseynli, Efthymia Kazakou, Johann Kristthorsson, Mahalakshmi Sabapathi, Siri Vinay, Stephen Whelan

Project Contact: Stephen Whelan (Lead Architect, Template 1), [stephen.whelan.15@ucl.ac.uk](mailto:stephen.whelan.15@ucl.ac.uk)

The following document is a description of the file structure for the Retail Shopping data set.

## **PEACH\_retailShopping.csv** - (comma delimited)

This file is a running output of users' shopping from various stores including supermarkets, pharmacies, athletic apparel and outdoor gear. This information could be tracked through store loyalty programs and OCR receipt scanning applications.

### *File Structure*

<b>Field</b>	<b>Values</b>	<b>Description</b>
User Id	Integer	This field is a user's unique Id. Each user maintains their unique Id throughout all data sets in Project PEACH.
Date	Date (YYYYMMDD)	This field is the date the item was purchased by the user.
Item	String	This value is the item purchased. The items fall into the following broad categories {cosmetics, outdoor/athletic gear, social/alcohol related, tobacco, medicine, fast food, cleaning supplies, organic foods, dairy products}.