

The Sales Trends Dataset

This anonymized data set consists of the following files:

- daily_sales.csv – daily historical sales data from January 2013 to October 2015.
- item_categories.csv – supplemental information about the item categories.
- items.csv – supplemental information about the items/products.
- shops.csv – supplemental information about the shops.
- interested_sales_trends – the [shop,item] pairs to be investigated for the sales trends.

For your group project, you may investigate the monthly sales trends specified in the interested_sales_trends.csv file, and you may simply choose a subset (e.g., fewer than 20) of the [shop, item] pairs for your investigation and business case report.

The descriptions of the fields/attributes in the data set are as follows.

- ID – a unique identifier that represents a [shop, item] pair in the interested_sales_trends.csv file
- shop_id – a unique identifier of a shop
- item_id – a unique identifier of an item/product
- item_category_id – a unique identifier of item category
- item_cnt_day – the number of items sold
- item_price – the price of an item sold
- date – the date in the format dd/mm/yyyy
- date_block_num - a consecutive month number, used for convenience. January 2013 is 0, February 2013 is 1, ..., October 2015 is 33
- item_name – the name of item (deleted)
- shop_name – the name of shop (deleted)
- item_category_name – the name of item category (deleted)