# **Project 5**

Data: “Product\_dim” “Product\_list” “shipped”  
Aim: produce reports in pdf  
Operations: Import the datasets in SAS. Create a unique data set using an appropriate by variable (order the datasets in SAS by using a macro function) for the match. Include the following variables:

* PRODUCT\_ID: identifier of the product
* PRODUCT\_LINE: describes the product line. Create a format that codes the variable as follows:

|  |  |
| --- | --- |
| Children | CH |
| Clothes & Shoes | CL\_SH |
| Outdoors | OU |
| Sports | SP |

* PRODUCT\_NAME: indicates the product name
* SUPPLIER\_COUNTRY: indicates the supplier country
* SHIP\_DATE: date in which the product is shipped. Format this variable as follows: 09-08-2019
* QUANTITY: number of items bought
* PRICE: price of each item
* Create PRICE\_CAT: variable with 4 categories, obtained from the quartiles of PRICE

Create a report with all the previously listed variables, grouped by PRODUCT\_LINE, adding an appropriate title.

Create a table with mean, min, max and sd of the variable PRICE, grouped by country.

Compute a second table in which only the Total expenditure related to the first purchase (by shipping date) of the same product appear in one record. Report the ordered product\_id, quantity, unit price and total, as follows:

|  |  |  |  |
| --- | --- | --- | --- |
| Product\_ID | Quantity | Price | Total expenditure |
| 240800200021 | 2 | $20.50 | $41.00 |
| 240800200035 | 4 | $75.20 | $150.40 |
| 240200100225 | 1 | $60.00 | $120.00 |
| 240200200091 | 3 | $10.25 | $20.50 |
| ... | ... |  | ... |

Compute an appropriate test statistic the potential impact of the variable country on the ranges of expenditure defined as “Low” if the total is lower than 50.00$, “Medium” if the total lays between 50.00$ - 100.00$ and “High” if is greater than 100.00$.