**Grouping and Aggregating**

## **Objective**

The object of this lab is to give you some practical experience of wrangling data and deriving data from existing datasets.

Specifically by allowing you to demonstrate your understanding of the following concepts

* Merging DataFrames
* Creating Multi Part Indexes
* Creating Pivot Tables
* Grouping and Aggregating Data
* Creating Cross Sections

## **Create a new notebook**

Import the libraries

pandas

numpy

matplotlib.pyplot

## Import the sales and Stores datasets for Adidas

Using the files ***../Data/Adidas\_Sales.csv*** and ***../Data/Adidas\_Stores.csv***

Merge both datasets into a single DataFrame

## **Build a Multi Index**

Use the columns

- Asset ID

- Location

- Province

- Cit

- Asset Type

- Brand

- Financials

## **Create some Pivot Tables**

1) Produce pivots for the following

- by Asset ID

- by Location

- by Province

- by City

- by Asset Type

- by Brand

- by Financials

2) Create a pivot of the following

- Location, Province, Asset Type

- For each quarter between Mar-16 and Dec-17 (inclusive)

## **Create some Groups**

Using the same Multi Index created for part 2, use the pandas.groupby method to produce groups of the following

1) The sum of all Financials for each Location for the quarters Mar-16 through to Dec-17 inclusive

2) The total number of employees for the entire enterprise grouped by Location, Brand and City

## **Create some Cross Sections**

Use the pandas.xs method to produce the following cross sections

1) Total quarterly Financials for the USA for the year 2018

2) Total quarterly Operating Income for each country for the year 2017