

Guía de Estudio Extracción de Datos

Análisis de Datos

Máster Universitario en
Tecnologías del Sector Financiero:
Fintech

Objetivos

Extraer datos de cotizaciones de Yahoo Finance y realizar transformaciones básicas para explorar y hacer gráficas con los datos.

Materiales

En Aula Global está disponible un fichero .zip que contiene varios scripts de R que se irán ampliando según las necesidades del curso. Estos scripts tienen un conjunto de funciones que nos permiten manipular los datos obtenidos desde Yahoo Finance. Podemos hacer transformaciones básicas para extraer un subconjunto de la serie o calcular por ejemplo las rentabilidades diarias o totales de las series.

Las funciones en *yahoodatatools.R* leerán los ficheros .csv descargados desde Yahoo Finance en el subdirectorio *basedata* y luego de completar las series a todos los días laborales, escribirá el fichero procesado en el subdirectorio *data*. La función *stockGetData* buscará si existe una serie limpia en *data* para evitar el pre-procesado. De lo contrario buscará un fichero nuevo en *basedata*.

Desarrollo

1) Ejecutar el script "ExtraccionDatos.R" para ver cómo utilizar las distintas funciones. Las actividades que llevamos a cabo son:

- Lectura y pre-procesado de las series de cotizaciones con la función *stockGetData*.
- Extraer parte de la serie de cotizaciones por rango de fechas con la función *stockSubset*.
- Crear gráficas con la librería *ggplot2*.

2) Analizar el código en los scripts de R, para comprender las funciones implementadas

Ejercicios

1. Descargar las cotizaciones históricas de BBVA y Mapfre desde el 2010.
2. Extender el script de R para incluir en la gráfica de comparación al BBVA.
3. Crear un script de R en el que generemos una gráfica para visualizar la correlación de los rendimientos diarios entre el IBEX35 y el Santander, y el IBEX35 y Mapfre, ambos desde el 2011. Para esto es necesario:
 - Obtener y procesar las cotizaciones correspondientes y generar los dataframes con los datos desde 2011.
 - Generar los vectores de rendimientos diarios con la función *stepReturn* programada en los scripts de R.

- Generar una gráfica de puntos con la librería ggplot2.