

# Programación Actuarial III

## Caso: Reconocimiento de Actividad Humana con Celulares

### Introducción

Una de las áreas más populares del análisis de datos proviene de las computadoras de vestimenta. Para más detalles puedes consultar este artículo:

<http://www.insideactivitytracking.com/datascienceactivitytrackingandthebattlefortheworldstopportsbrand/>

Compañías como Fitbit, Nike y Jawbone Up compiten para desarrollar los algoritmos más avanzados para atraer a nuevos usuarios. Los datos de este caso provienen de los acelerómetros ubicados en celulares Samsung Galaxy S.

Una mayor descripción de la base de datos y su recopilación puede ser hallada en la siguiente liga:

<http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones>

### Datos

Los datos pueden ser descargados de la siguiente liga:

<https://www.dropbox.com/s/j26ksrw1jzqz2a1/getdata-projectfiles-UCI%20HAR%20Dataset.zip?dl=0>

### Instrucciones

El propósito de este proyecto es demostrar tu habilidad para recolectar, trabajar y limpiar base de datos. El objetivo es preparar un conjunto ordenado de información que pueda ser trabajado en análisis posteriores.

Considera incluir:

- 1) La base de datos ordenada
- 2) Dirección de GitHub para el repositorio
- 3) Un libro de códigos para describir las variables, la base de datos y las transformaciones que realices llamado "CodeBook.md".
- 4) Un archivo "README.md" en el repositorio con tus códigos. Este archivo explica el funcionamiento de tus scripts y como se conectan entre ellos.

Deberás crear un código en R llamado "correr\_analisis.R" que haga lo siguiente:

- 1) Une los datos de test con los de training, para crear un solo conjunto de datos.
- 2) Extrae únicamente las medidas de media y desviación estándar de cada medición.

- 3) Usa nombres de actividad para describir los nombres de las actividades en la base de datos.
- 4) Coloca etiquetas apropiadas en la base de datos con nombres de variables que las describan.
- 5) Con los datos del paso 4, crea una segunda base de datos independiente con el promedio de cada variable para cada actividad y cada sujeto.