

Temas avanzados en física computacional: Análisis de Datos

Tarea #1

Semestre académico 2016-I

Elaborada por Prof. J. Bazo

Parte 1. (3 ptos)

Con un ejemplo individual escribir los pasos del proceso de análisis de datos según se presentó en la primera clase:

- Definir pregunta
- Definir juego de datos ideal
- Obtener datos accesibles
- Limpiar datos
- Análisis exploratorio
- Modelamiento estadístico
- Graficar resultados
- Interpretar resultados
- Crear código reproducible
- Distribuir resultados

Parte 2. (7 ptos)

Se analizará un archivo de datos (NuAstro_4yr_IceCube_Data.txt que encuentra en el Campus Virtual del curso) correspondiente al experimento IceCube recopilado de <https://icecube.wisc.edu/science/data>.

Este archivo contiene en cada fila los eventos de neutrinos astrofísicos de alta energía de IceCube con vértices de interacción dentro del volumen activo del detector. Los datos cubren un total de 1347 días entre 2010 y 2014.

Las columnas tienen la siguiente información:

- ID : número de identificación
- Dep_Energy_TeV: energía electromagnética equivalente depositada.
- Ene_Err_min: error en la energía por debajo del valor central
- Ene_Err_max: error en la energía por encima del valor central
- Time_MJD: tiempo fraccional de días Julianos modificados
- Declination_deg: ángulo de declinación en coordenadas ecuatoriales
- RA_deg: ángulo de ascensión recta en coordenadas ecuatoriales
- Med_Ang_Res_deg: resolución angular mediana
- Topology: topología del evento

Los pasos a seguir en script que se debe crear son:

- A.) Leer el archivo del path actual.
- B.) Remover los eventos con datos incompletos (0.5ptos)
- C.) Hallar mediante 2 métodos el promedio de la resolución angular para las diferentes topologías. (1pto)
- D.) Crear una función (mjd2posix) que convierta la fecha de MJD (Modified Julian Day) en tiempo POSIX (GMT) (1pto)
- E.) Aplicar la función anterior para añadir una columna Time_GMT que sea Time_MJD convertido a GMT. (0.5ptos)
- F.) Escribir una función (evt_month) que tome como argumentos la tabla de datos y un mes dado (numérico) y retorne el número de eventos en los diversos años ocurridos en dicho mes. Luego usar ésta para imprimir el mes con el mayor número de eventos. (1,5 ptos)
- G.) Hacer un gráfico de la distribución de los eventos en coordenadas ecuatoriales. (0.5ptos)
- H.) Pensar en una cantidad interesante de calcular y escribir la función correspondiente comentando de qué se trata. (1pto)

Guardar paulatinamente el código en un archivo .R que se deberá actualizar en su cuenta de Github. (1pto)

Enviar la parte 1 y el link al código en Github por correo electrónico a jbazo@pucp.edu.pe a más tardar el 13 de abril a las 23:59.

San Miguel, 8 de abril de 2016.