Introducción a la Ciencia de los Datos

Tarea 6

Nota: entregar las respuestas el próximo miércoles al correo de Daniela Peña ([daniela.pena@alumnos.cide.edu](mailto:daniela.pena@alumnos.cide.edu)) en un archivo: nombre\_apellido\_tarea6. py (o en archivo texto)

Realizar un Script en Spyder con las siguientes características:

(i) Lee el archivo tips en formato csv a través de un Dataframe de Pandas

(ii) Muestra los primeros cinco renglones

(iii) Hacer una visualización de dos histogramas con 10 bins de las variables 'total\_bill'

y 'tip'. Deben incluir títulos y etiquetas para los ejes de las Xs y de las Ys. Ponerlos

de distinto color.

(iv) Hacer visualización de un diagrama de dispersión con variables 'total\_bill' y 'tip',

agregar título y etiquetas en los ejes.

(v) Hacer una visualización con boxplot (mediana, cuartiles y outliers) en un mismo panel

para la variable 'tip' en donde se divide por fumadores ('smoker') y no fumadores. Poner

una etiqueta a cada categoría, y etiquetas en los ejes Xs y Ys así como el título

Para este ejercicio --no usar seaborn—

(vi) Hacer un diagrama de dispersión con variables 'total\_bill' y 'tip', en el que el color

del punto se elige a partir de la variable 'time' y el tamaño del punto en función

de la variable 'size'. Agregar título y etiquetas en los ejes.