

Chuleta para Python (abril 2021)

MCAF - UCM

Instalar una librería	<pre>C:/> pip install jupyter</pre>
Abrir jupyter	<pre>C:/> jupyter notebook</pre>
Consultar librerías instaladas	<pre>C:/> pip list</pre>
Crear función	<pre>def sumauno(x): return x + 1</pre>
Importar una librería	<pre>import numpy</pre>
Importar librería con alias	<pre>import numpy as np</pre>
Importar x funciones de una librería	<pre>from numpy import mean, var, std</pre>
Consultar versión instalada	<pre>print(numpy.__version__)</pre>
Consultar funciones que incluye la librería	<pre>dir(numpy)</pre>
Numpy	
Convención para importar	<pre>import numpy as np</pre>
Genera puntos espaciados linealmente	<pre>np.linspace(inicio, fin, n-pasos)</pre>
Fijar una semilla	<pre>np.random.seed(n)</pre>
Genera n·m aleatorios según dist. normal	<pre>np.random.randn(n,m)</pre>
Elegir distribución de probabilidad	<pre>np.random.normal(loc, scale, size)</pre>
Pandas	
Convención para importar	<pre>import pandas as pd</pre>
Crear una serie	<pre>s = pd.Series(datos)</pre>
Crear un DataFrame	<pre>df = pd.DataFrame(datos, columns=['Col1'])</pre>
Importar datos:	
Importar un csv	<pre>pd.read_csv('fichero.csv', sep=';', decimal='.')</pre>
Importar un excel	<pre>pd.read_excel('fichero.xls')</pre>
Importar el cuadro n de una web	<pre>pd.read_url(url)[n]</pre>
Limpieza:	
Reemplazar 'M' por 1	<pre>.replace('M', 1)</pre>
Cambiar '1' por 'M' en un campo	<pre>df['Sexo'].apply(lambda x: 1 if x=='M' else 0)</pre>
Rellenar valores vacíos o N/A con 0	<pre>df.fillna(0)</pre>
Borrar registros con valores N/A	<pre>df.dropna(inplace=True)</pre>
Borrar la fila n	<pre>df.drop(n, inplace=True)</pre>
Borrar una columna	<pre>df.drop('Col', axis=1, inplace=True)</pre>
Muestras los principales estadísticos	<pre>df.describe()</pre>
Muestra información sobre el df	<pre>df.info()</pre>
Fórmulas:	<i>axis=0 (index) axis=1 (columns)</i>
Suma	<pre>.sum()</pre>
Producto	<pre>.prod()</pre>
Media, Mínimo, Máximo	<pre>.mean() .min() .max()</pre>
q-quantil (entre 0 y 1)	<pre>.quantile(q)</pre>
Diferencias entre celdas	<pre>.diff()</pre>
Suma acumulada	<pre>.cumsum()</pre>
Producto acumulado	<pre>.cumprod()</pre>
Redondear a n-decimales	<pre>.round(n)</pre>
Muestreo (con reemplazamiento)	<pre>.sample(n) .sample(n, replace=True)</pre>
Convertir a entero	<pre>.astype(int) (int, float, str)</pre>
Crear gráfica	<pre>.plot()</pre>
Agrupar y pivotar:	
Agrupar por un campo y sumar el resto	<pre>df.groupby('Col1').sum()</pre>
Agrupar y elegir resultados	<pre>df.groupby('Col1', 'Col2').agg({'col3': 'sum', 'col4': 'mean'})</pre>
Resetear el index	<pre>.reset_index()</pre>
Tabla pivot	<pre>pivot=df.pivot(index='Fecha', columns='Dev', values='Importe')</pre>
Manejo de fechas:	
Formatear una fecha	<pre>pd.to_datetime('31/12/2020', format='%d/%m/%Y')</pre>
Crear una Serie de fechas	<pre>pd.date_range(start, end, period, freq)</pre>
Convertir fecha a periodo anual	<pre>df['Fecha'].dt.to_period(freq='Y')</pre>