

Curso de Pandas (se compara con Excel)

1. Introducción

A) Dos herramientas Excel y Pandas

1. Aprender comparando
2. Vamos a Construir en Excel y Python

2. Excel

1. Revolución de las hojas de calculo
2. Fácil de aprender
3. Muchos de los procesos son gráficos con el mouse

2.1. Crear una tabla Excel para trabajar

1. Nombre,Apellido, edad, peso, telefono, RRSS

3. Pandas

3.1. Que es un DataFrame ?

B) DataFrame es una estructura que contiene datos 2D

1. Similar a las tablas de Excel
2. Las filas y columnas tienen etiquetas
3. Para comenzar a trabajar
 - 3.1 `import pandas as pd`
4. instalar xlrd, openpyxl

3.2. Crear un DataFrame a partir un diccionario de listas

1. `category-map-dic.py`

3.3. Estructura del objeto Dataframe

1. `data`
2. `index`
3. `columns`
4. `dtype`



4. Acceso a la data

C) Piense en el DataFrame como un diccionario de columnas, o Serie-Pandas

1. `s1=df1['Date']`
2. `s2=df1.Date`
3. Puede obtener un solo elemento de una serie Pandas utilizando el index
4. `s1[14]`

4.1. Dos formas de selección de data `loc()` y `iloc()`

1. puede acceder a una fila completa con `df1.loc [14]`
2. Por posición numérica con método `.iloc()`
3. Selección del método mixto esta deprecada
4. Selección de; DataFrame por expresiones booleanas

4.2. Métodos útiles para el acceso a la data

1. Método `concat()`
2. Método `drop_duplicates()`
3. Método versátil `merge()`
4. Método `isin()`
5. `set_dataframe.py`

5. Escribir o Leer hacia y desde una fuente externa

D) Puede cargar y Guardar los datos y las etiquetas hacia y desde varios tipos de archivos

5.1. Leyendo la tabla de Excel en pandas

1. método `read_excel()`
 - 1.1 `index_col=0`
 - 1.2 `skiprows=4`
 - 1.3 `header=None`

5.2. Explorar el DataFrame leído

1. Usar `head()` y `tail()`
2. usar método `describe()` [11.3](#)
 - 2.1 Para datos numéricos
 - 2.2 Para datos no -numericos



5.3. Por que el éxito de pandas ?

1. pandas concentro varias funcionalidades dispersas
2. Carga de datos: read_*
3. Escritura de datos: to_*
4. Limpieza de los datos

5.4. Escribir dataframe en un libro de Excel

1. método `to_excel()`
2. método `to_csv()`
3. Reportar los resultados de python en archivo Excel
4. Crear un data frame para crear un archivo Excel

5.5. Ejemplo de lookup

1. Leer dos hojas de calculo
2. Explicar la funcion "lookup"
3. `exel_lookup.py`
4. Escribir una hoja de calculo nueva
5. Pandas cambio la fecha a `datetime64[ns]` [8](#)

6. Utilizando el Dataframe

6.1. Varias formas de construir un dataframe

1. Construir listas
2. Distribuciones aleatorias de numpy
3. Construir un diccionario [11.4](#)

6.2. Eliminando columnas o filas

1. `df.drop([3,4], axis=0)`
2. `df.drop('Country', axis=1)`
3. `treasures_pandas.py`
4. Hay una sintaxis alternativa para ambos

6.3. Gran cantidad de Operaciones

1. Operaciones aritméticas
2. Comparaciones
3. Filtro booleano



4. Sumar dos columnas y producir una tercera columna
5. Sumar a lo largo de la fila y crear columna de sub-totales
6. Sumar una columna
7. Sumar cada columna y crear fila de sub-totales

6.4. Operaciones de análisis

1. Tabla de frecuencia entre dos columna-factor
 - 1.1 titanic_xt.py

6.5. Dos operaciones de Excel en Pandas

1. Funcion IF de Excel en Pandas
 - 1.1 List comprehensions
 - 1.2 excel_easy1.py
2. Funcion VLOOKUP de Excel en Pandas
 - 2.1 Funcion merge() de Pandas
 - 2.2 excel_easy2.py

6.6. operaciones pivot

1. Método pivot()
2. Método pivot_table()
3. excel_easy3.py
4. small-pivot.py
5. conditional-prob.py

7. Preparación de la data

7.1. Tratar con los datos faltantes

1. Contar el numero de NAN [11.3](#)
2. Agregar un valor predeterminado para los datos faltantes
3. Eliminar columnas que tienen datos faltantes
4. Eliminar filas que tienen datos faltantes

7.2. Limpiar campos

1. Limpiar un DataFrame con caracteres no deseables
- 2.
- 3.
- 4.



8. Fechas y Horas

E) ¿Por qué programar con fechas y horas es un desafío?

1. Pandas cambio la fecha a datetime64[ns]
2. Como Excel maneja las fechas y horas ?
3. Como Pandas maneja las fechas y horas ?
4. Usos horarios
5. formatos de fechas

8.1. leer data de los sensores de tu celular

- 1.
- 2.

9. Análisis exploratorio de la data

1. columnas con data categóricas no numéricas
 - 1.1 .describe() [11.3](#)
 - 1.2 .counts()
 - 1.3 .unique()
 - 1.4 .top()
 - 1.5 .freq()
2. describe() series de tiempo

10. Visualización del dataframe

F) Gráficos en Excel con click y gráficos con código

1. En Excel hay una galería de opciones
2. En código los procesos son mas replicables y reproducibles

10.1. columnas con data numérica

F)1 .boxplot

F)2 Análisis Multivariable

10.2. Otros gráficos

1. .FacetGrid
2. .jointplot



10.3. Tablas dinamicas

1. .heatmap

10.4. Cual es la idea de seaborn ?

1. .countplot
2. Análisis bivariante

11. Ejemplos prácticos con librería pandas

G) Convertir archivo a DataFrame rápido y directo

11.1. pagos-inter.py

- 0.1 Convertir campos fechas
- 0.2 Convertir campos a punto flotante
- 0.3 plt.xticks(rotation=25)

11.2. sales-dataframe.py

- 0.1 Reducir a tres columnas
- 0.2 Agrupar por cliente
- 0.3 Sumar las ventas
- 0.4 Grafico de barras de ventas totales

11.3. aq.py

- 0.1 Método describe()
- 0.2 Contar numero de NAN
- 0.3 Cambiar columna Month a tipo category

11.4. create_dataframe.py

- 0.1 Crear un dataframe usando funciones numpy
- 0.2 Crear una serie de tiempo entre dos fechas
- 0.3 Convertir indice en datetime

