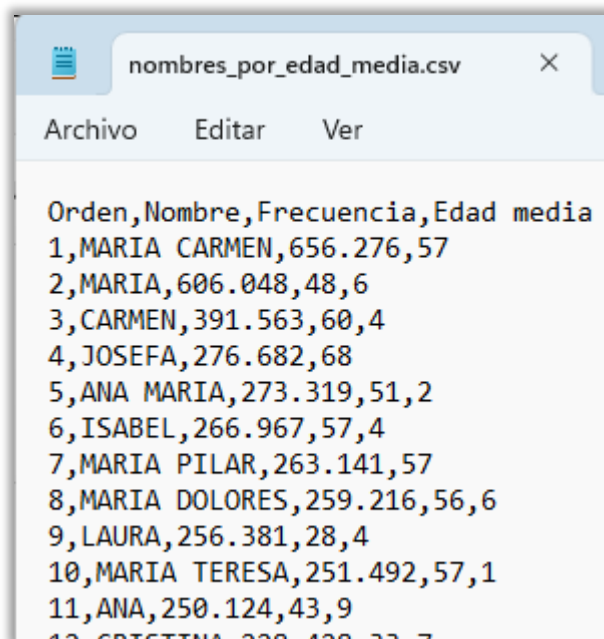




Programación. Python

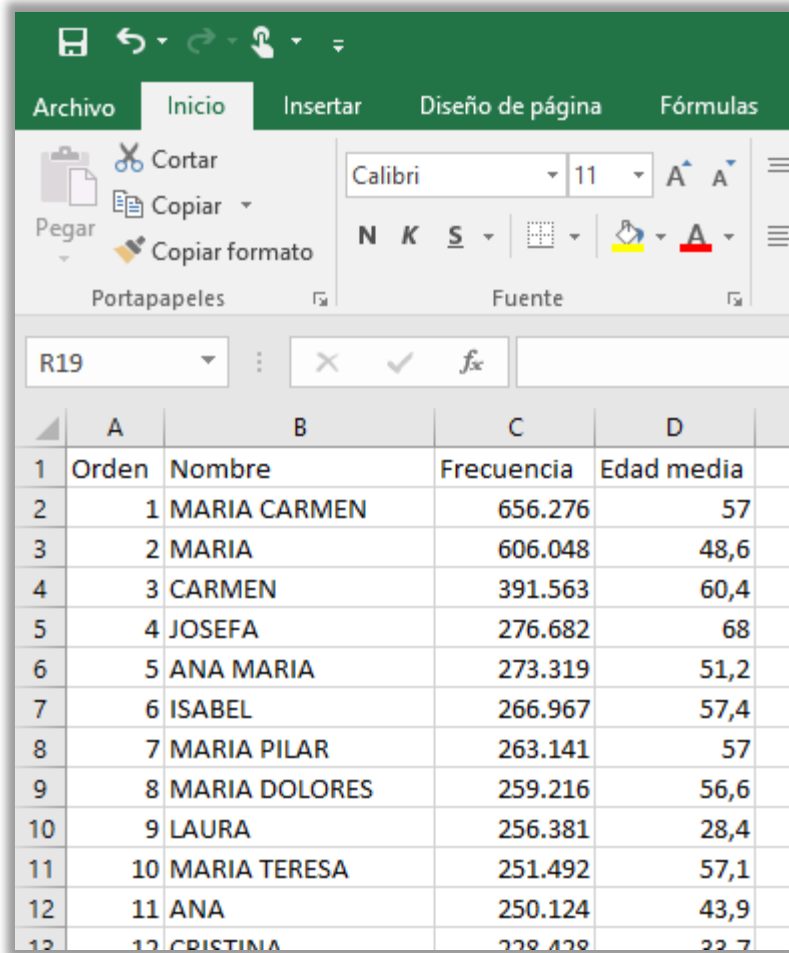
Archivos csv y el formato json

Archivos CSV



Orden	Nombre	Frecuencia	Edad media
1	MARIA CARMEN	656.276	57
2	MARIA	606.048	48,6
3	CARMEN	391.563	60,4
4	JOSEFA	276.682	68
5	ANA MARIA	273.319	51,2
6	ISABEL	266.967	57,4
7	MARIA PILAR	263.141	57
8	MARIA DOLORES	259.216	56,6
9	LAURA	256.381	28,4
10	MARIA TERESA	251.492	57,1
11	ANA	250.124	43,9
12	CRISTINA	228.428	22,7

<https://docs.python.org/3/library/csv.html>



	A	B	C	D
1	Orden	Nombre	Frecuencia	Edad media
2	1	MARIA CARMEN	656.276	57
3	2	MARIA	606.048	48,6
4	3	CARMEN	391.563	60,4
5	4	JOSEFA	276.682	68
6	5	ANA MARIA	273.319	51,2
7	6	ISABEL	266.967	57,4
8	7	MARIA PILAR	263.141	57
9	8	MARIA DOLORES	259.216	56,6
10	9	LAURA	256.381	28,4
11	10	MARIA TERESA	251.492	57,1
12	11	ANA	250.124	43,9
13	12	CRISTINA	228.428	22,7

Archivos csv

```
import csv
```

```
csvFile = csv.reader(open("nombres_por_edad_media.csv", "r"))  
for row in csvFile:  
    print(row)
```

```
['Orden', 'Nombre', 'Frecuencia', 'Edad media']  
['1', 'MARIA CARMEN', '656.276', '57']  
['2', 'MARIA', '606.048', '48', '6']  
['3', 'CARMEN', '391.563', '60', '4']  
['4', 'JOSEFA', '276.682', '68']  
['5', 'ANA MARIA', '273.319', '51', '2']  
['6', 'ISABEL', '266.967', '57', '4']  
['7', 'MARIA PILAR', '263.141', '57']  
['8', 'MARIA DOLORES', '259.216', '56', '6']  
['9', 'LAURA', '256.381', '28', '4']  
['10', 'MARIA TERESA', '251.492', '57', '1']  
['11', 'ANA', '250.124', '42', '0']
```

De otro modo:

```
import csv  
with open('nombres_por_edad_media.csv', 'r') as csvFile:  
    reader = csv.reader(csvFile, delimiter=';')  
    for row in reader:  
        print(row)
```

Archivos CSV

*# Y también, leyendo la cabecera por separado,
seleccionando algunas columnas y realizando una conversión, ya de paso:*

```
with open('nombres_por_edad_media.csv', 'r') as csvFile:  
    reader = csv.reader(csvFile)  
    cab = next(reader)  
    print(cab[1], cab[3])  
    for row in reader:  
        nombre, edad_media = row[1], row[3]  
        print(nombre, int(edad_media))
```

Nombre Edad media

MARIA CARMEN 57

MARIA 48

CARMEN 60

JOSEFA 68

ANA MARIA 51

ISABEL 57

MARIA PILAR 57

MARIA DOLORES 56

LAURA 28

MARIA TERESA 57

ANA 43

CRISTINA 33

MARIA ANGELES 55

MARTA 29

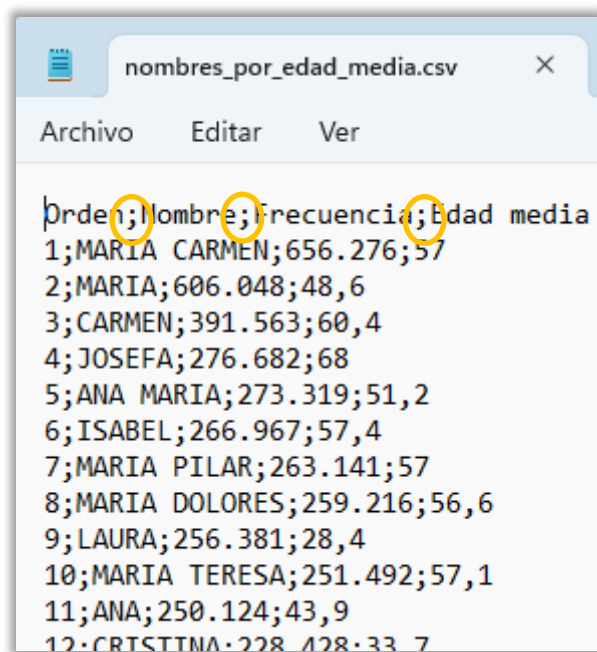
FRANCISCA 64

ANTONIA 64

MARIA ISABEL 52

MARIA JOSE 46

Archivos csv



Orden	Nombre	Frecuencia	Edad media
1	MARIA CARMEN	656.276	57
2	MARIA	606.048	48,6
3	CARMEN	391.563	60,4
4	JOSEFA	276.682	68
5	ANA MARIA	273.319	51,2
6	ISABEL	266.967	57,4
7	MARIA PILAR	263.141	57
8	MARIA DOLORES	259.216	56,6
9	LAURA	256.381	28,4
10	MARIA TERESA	251.492	57,1
11	ANA	250.124	43,9
12	CRISTINA	228.428	33,7

```
csv_file = csv.reader(open("nombres_por_edad_media_pc.csv", "r", delimiter=';'))  
reader = csv.reader(csv_file, delimiter=';')
```

Archivos CSV

```
# Escritura:


ids_columnas = ['Nombre', 'Matemáticas', 'Lengua', 'Historia']

filas = [ ['Juan', '5.7', '2.5', '9.0'],
          ['Sara', '9.5', '6.7', '9.3'],
          ['Alberto', '7.5', '7.5', '7.5'],
          ['Sara', '4.9', '5.2', '8.0']]

with open("calificaciones_2019.csv", 'w', newline='') as csv_archivo:
    csvwriter = csv.writer(csv_archivo, delimiter=";")
    csvwriter.writerow(ids_columnas)
    csvwriter.writerows(filas)

print("Hecho")
```

Hecho

 nombres_por_edad_media.csv

Archivo Editar Ver

Nombre;Matemáticas;Lengua;Historia
Juan;5.7;2.5;9.0
Sara;9.5;6.7;9.3
Alberto;7.5;7.5;7.5
Sara;4.9;5.2;8.0

Transformación de datos

```
# Enteros:
```

```
print(int("7"),int("123.000".replace('.', '')))
```

```
# Reales:
```

```
print(float("4.5"), float("4,5".replace(",", ".")))  
print(float("123.000,75".replace('.', '').replace(',', '.')))
```

```
# Fechas:
```

```
from datetime import datetime  
fecha_str = '10-24-2019'  
  
fecha_objeto = datetime.strptime(fecha_str, '%m-%d-%Y').date()  
print(type(fecha_objeto))  
print(fecha_objeto)
```

```
7 123000
```

```
4.5 4.5
```

```
123000.75
```

```
<class 'datetime.date'>
```

```
2019-10-24
```

El formato JSON

```
example_1.json: Bloc de notas
Archivo Edición Formato Ver Ayuda
{
  "fruit": "Apple",
  "size": "Large",
  "color": "Red"
}
```

```
import json

with open("example_1.json") as archivo:
    datos = json.loads(archivo.read())

print(datos)
print(type(datos))
```

```
{'fruit': 'Apple', 'size': 'Large', 'color': 'Red'}
<class 'dict'>
```

Impresión con sangrado:

```
print(json.dumps(datos, indent=4))
```

```
{
    "fruit": "Apple",
    "size": "Large",
    "color": "Red"
}
```

*# Obsérvese que el archivo `example_1.json`
no contiene un diccionario,
sino únicamente cadenas de caracteres:*

```
with open("example_1.json") as archivo:
    for row in archivo:
        print(row,end="")
```

```
{
    "fruit": "Apple",
    "size": "Large",
    "color": "Red"
}
```

str

dict

str

El formato JSON

```
with open("example_2.json") as archivo:  
    datos = json.loads(archivo.read())  
datos
```

```
{'quiz': {'sport': {'q1': {'question': 'Which one is correct team name in NBA?',  
    'options': ['New York Bulls',  
    'Los Angeles Kings',  
    'Golden State Warriros',  
    'Huston Rocket'],  
    'answer': 'Huston Rocket'}}},  
'maths': {'q1': {'question': '5 + 7 = ?',  
    'options': ['10', '11', '12', '13'],  
    'answer': '12'},  
    'q2': {'question': '12 - 8 = ?',  
    'options': ['1', '2', '3', '4'],  
    'answer': '4'}}}}
```

Escritura:

```
with open("example_3.json", "w") as archivo:  
    archivo.write(json.dumps(datos))
```

example_2.json: Bloc de notas

Archivo Edición Formato Ver Ayuda

```
{  
  "quiz": {  
    "sport": {  
      "q1": {  
        "question": "Which one is correct team name in NBA?",  
        "options": [  
          "New York Bulls",  
          "Los Angeles Kings",  
          "Golden State Warriros",  
          "Huston Rocket"  
        ],  
        "answer": "Huston Rocket"  
      }  
    },  
    "maths": {  
      "q1": {  
        "question": "5 + 7 = ?",  
        "options": [  
          "10",  
          "11",  
          "12",  
          "13"  
        ],  
        "answer": "12"  
      },  
      "q2": {  
        "question": "12 - 8 = ?",  
        "options": [  
          "1",  
          "2",  
          "3",  
          "4"  
        ],  
        "answer": "4"  
      }  
    }  
  }  
}
```

El formato JSON

```
estaciones.json
Archivo  Editar  Ver

[ {
  "latitud" : "431825N",
  "provincia" : "A CORUÑA",
  "altitud" : "98",
  "indicativo" : "1387E",
  "nombre" : "A CORUÑA AEROPUERTO",
  "indsinop" : "08002",
  "longitud" : "082219W"
}, {
  "latitud" : "432157N",
  "provincia" : "A CORUÑA",
  "altitud" : "58",
  "indicativo" : "1387",
  "nombre" : "A CORUÑA",
  "indsinop" : "08001",
  "longitud" : "082517W"
}, {
  "latitud" : "430938N",
  "provincia" : "A CORUÑA",
  "altitud" : "50",
  "indicativo" : "1393",
  "nombre" : "CABO VILAN",
  "indsinop" : "08006",
  "longitud" : "091239W"
}, {
  "latitud" : "434710N",
```

```
import json

f = open("estaciones.json")
estaciones_dicc = json.load(f)
estaciones_dicc

[{'latitud': '431825N',
  'provincia': 'A CORUÑA',
  'altitud': '98',
  'indicativo': '1387E',
  'nombre': 'A CORUÑA AEROPUERTO',
  'indsinop': '08002',
  'longitud': '082219W'},
 {'latitud': '432157N',
  'provincia': 'A CORUÑA',
  'altitud': '58',
  'indicativo': '1387',
  'nombre': 'A CORUÑA',
  'indsinop': '08001',
  'longitud': '082517W'},
 {'latitud': '430938N',
  'provincia': 'A CORUÑA',
  'altitud': '50',
  'indicativo': '1393',
  'nombre': 'CABO VILAN',
  'indsinop': '08006',
  'longitud': '091239W'},
 {'latitud': '434710N',
  'provincia': 'A CORUÑA',
  'altitud': '100',
  'indicativo': '1393',
  'nombre': 'CABO VILAN',
  'indsinop': '08006',
  'longitud': '091239W'}]
```

El formato JSON

```
import json
```

```
f = open("estaciones.json")
estaciones_dicc = json.load(f)
estaciones_dicc
```

```
[{'latitud': '431825N',
  'provincia': 'A CORUÑA',
  'altitud': '98',
  'indicativo': '1387E',
  'nombre': 'A CORUÑA AEROPUERTO',
  'indsinop': '08002',
  'longitud': '082219W'},
 {'latitud': '432157N',
  'provincia': 'A CORUÑA',
  'altitud': '58',
  'indicativo': '1387',
  'nombre': 'A CORUÑA',
  'indsinop': '08001',
  'longitud': '082517W'},
 {'latitud': '430938N',
  'provincia': 'A CORUÑA',
  'altitud': '50',
  'indicativo': '1393',
  'nombre': 'CABO VILAN',
  'indsinop': '08006',
  'longitud': '091239W'},
 {'latitud': '434710N',
  'provincia': 'A CORUÑA',
  'altitud': '80',
  'indicativo': '1351',
  'nombre': 'ESTACA DE BARES',
  'indsinop': '08004',
  'longitud': '074105W'},
 {'latitud': '425529N',
  'provincia': 'A CORUÑA',
  'altitud': '230',
  'indicativo': '1400',
  'nombre': 'FISTERRA',
  'indsinop': '08040',
  'longitud': '091729W'},
 ...
 {'latitud': '411952N',
  'provincia': 'ZARAGOZA',
  'altitud': '600',
  'indicativo': '9394X',
  'nombre': 'CALATAYUD',
  'indsinop': '08156',
  'longitud': '013843W'},
 {'latitud': '410652N',
  'provincia': 'ZARAGOZA',
  'altitud': '779',
  'indicativo': '9390',
  'nombre': 'DAROCA',
  'indsinop': '08157',
  'longitud': '012436W'},
 {'latitud': '422927N',
  'provincia': 'ZARAGOZA',
  'altitud': '626',
  'indicativo': '9244X',
  'nombre': 'SOS DEL REY CATÓLICO',
  'indsinop': '08090',
  'longitud': '011249W'},
 {'latitud': '413938N',
  'provincia': 'ZARAGOZA',
  'altitud': '249',
  'indicativo': '9434',
  'nombre': 'ZARAGOZA AEROPUERTO',
  'indsinop': '08160',
  'longitud': '010015W'},
 {'latitud': '413715N',
  'provincia': 'ZARAGOZA',
  'altitud': '254',
  'indicativo': '9434P',
  'nombre': 'ZARAGOZA, VALDESPARTERA',
  'indsinop': '08159',
  'longitud': '005606W'}]
```

```
import pandas
```

```
estaciones = pandas.DataFrame(estaciones_dicc)
estaciones
```

	latitud	provincia	altitud	indicativo	nombre	indsinop	longitud
0	431825N	A CORUÑA	98	1387E	A CORUÑA AEROPUERTO	08002	082219W
1	432157N	A CORUÑA	58	1387	A CORUÑA	08001	082517W
2	430938N	A CORUÑA	50	1393	CABO VILAN	08006	091239W
3	434710N	A CORUÑA	80	1351	ESTACA DE BARES	08004	074105W
4	425529N	A CORUÑA	230	1400	FISTERRA	08040	091729W
...
286	411952N	ZARAGOZA	600	9394X	CALATAYUD	08156	013843W
287	410652N	ZARAGOZA	779	9390	DAROCA	08157	012436W
288	422927N	ZARAGOZA	626	9244X	SOS DEL REY CATÓLICO	08090	011249W
289	413938N	ZARAGOZA	249	9434	ZARAGOZA AEROPUERTO	08160	010015W
290	413715N	ZARAGOZA	254	9434P	ZARAGOZA, VALDESPARTERA	08159	005606W

291 rows × 7 columns

El formato JSON

```
estaciones.json
Archivo  Editar  Ver

[ {
  "latitud" : "431825N",
  "provincia" : "A CORUÑA",
  "altitud" : "98",
  "indicativo" : "1387E",
  "nombre" : "A CORUÑA AEROPUERTO",
  "indsinop" : "08002",
  "longitud" : "082219W"
}, {
  "latitud" : "432157N",
  "provincia" : "A CORUÑA",
  "altitud" : "58",
  "indicativo" : "1387",
  "nombre" : "A CORUÑA",
  "indsinop" : "08001",
  "longitud" : "082517W"
}, {
  "latitud" : "430938N",
  "provincia" : "A CORUÑA",
  "altitud" : "50",
  "indicativo" : "1393",
  "nombre" : "CABO VILAN",
  "indsinop" : "08006",
  "longitud" : "091239W"
}, {
  "latitud" : "434710N",
  "provincia" : "A CORUÑA",
  "altitud" : "80",
  "indicativo" : "1351",
  "nombre" : "ESTACA DE BARES",
  "indsinop" : "08004",
  "longitud" : "074105W"
}, {
  "latitud" : "425529N",
  "provincia" : "A CORUÑA",
  "altitud" : "230",
  "indicativo" : "1400",
  "nombre" : "FISTERRA",
  "indsinop" : "08040",
  "longitud" : "091729W"
}, ...
286 411952N ZARAGOZA 600 9394X CALATAYUD 08156 013843W
287 410652N ZARAGOZA 779 9390 DAROCA 08157 012436W
288 422927N ZARAGOZA 626 9244X SOS DEL REY CATÓLICO 08090 011249W
289 413938N ZARAGOZA 249 9434 ZARAGOZA AEROPUERTO 08160 010015W
290 413715N ZARAGOZA 254 9434P ZARAGOZA, VALDESPARTERA 08159 005606W
291 434710N
```

```
datos_pandas = pandas.read_json("estaciones.json", encoding="latin1")
datos_pandas
```

	latitud	provincia	altitud	indicativo	nombre	indsinop	longitud
0	431825N	A CORUÑA	98	1387E	A CORUÑA AEROPUERTO	08002	082219W
1	432157N	A CORUÑA	58	1387	A CORUÑA	08001	082517W
2	430938N	A CORUÑA	50	1393	CABO VILAN	08006	091239W
3	434710N	A CORUÑA	80	1351	ESTACA DE BARES	08004	074105W
4	425529N	A CORUÑA	230	1400	FISTERRA	08040	091729W
...
286	411952N	ZARAGOZA	600	9394X	CALATAYUD	08156	013843W
287	410652N	ZARAGOZA	779	9390	DAROCA	08157	012436W
288	422927N	ZARAGOZA	626	9244X	SOS DEL REY CATÓLICO	08090	011249W
289	413938N	ZARAGOZA	249	9434	ZARAGOZA AEROPUERTO	08160	010015W
290	413715N	ZARAGOZA	254	9434P	ZARAGOZA, VALDESPARTERA	08159	005606W

291 rows × 7 columns

El formato JSON

También se puede manejar la orientación:

```
estaciones = pandas.read_json("estaciones.json", orient="index")
estaciones
```

	0	1	2	3	4	5
altitud	98	58	50	80	230	685
indicativo	1387E	1387	1393	1351	1400	1437O
indsinop	08002	08001	08006	08004	08040	08043
latitud	431825N	432157N	430938N	434710N	425529N	424314N
longitud	082219W	082517W	091239W	074105W	091729W	085524W
nombre	A CORUÑA AEROPUERTO	A CORUÑA	CABO VILAN	ESTACA DE BARES	FISTERRA	MONTE IROITE
provincia	A CORUÑA	A CORUÑA	A CORUÑA	A CORUÑA	A CORUÑA	A CORUÑA

7 rows × 291 columns

```
estaciones = pandas.read_json("estaciones.json", orient="column")
estaciones
```

	altitud	indicativo	indsinop	latitud	longitud	nombre
0	98	1387E	08002	431825N	082219W	A CORUÑA AEROPUERTO
1	58	1387	08001	432157N	082517W	A CORUÑA
2	50	1393	08006	430938N	091239W	CABO VILAN
3	80	1351	08004	434710N	074105W	ESTACA DE BARES
4	230	1400	08040	425529N	091729W	FISTERRA
5	685	1437O	08043	424314N	085524W	MONTE IROITE
6	100	1473A	08039	424418N	083738W	PADRÓN
7	370	1428	08042	425317N	082438W	SANTIAGO DE COMPOSTELA AEROPUERTO
8	240	1475X	08041	425234N	083321W	SANTIAGO DE COMPOSTELA



Programación. Python

Archivos csv y el formato json