import pandas as pd

data=pd.read\_csv(r"C:\Users\HP\OneDrive\Desktop\Datascience\14th\panda\Sample - Superstore\_Orders.csv")

data

Category City Country/Region Customer Name \  
0 Office Supplies Houston United States Darren Powers   
1 Office Supplies Naperville United States Phillina Ober   
2 Office Supplies Naperville United States Phillina Ober   
3 Office Supplies Naperville United States Phillina Ober   
4 Office Supplies Philadelphia United States Mick Brown   
... ... ... ... ...   
10189 Office Supplies New York City United States Patrick O'Donnell   
10190 Office Supplies Fairfield United States Erica Bern   
10191 Office Supplies Loveland United States Jill Matthias   
10192 Technology New York City United States Patrick O'Donnell   
10193 Office Supplies Charlottetown Canada Harry Olson   
  
 Manufacturer Order Date Order ID Postal Code \  
0 Message Book 03-01-2019 US-2019-103800 77095   
1 GBC 04-01-2019 US-2019-112326 60540   
2 Avery 04-01-2019 US-2019-112326 60540   
3 SAFCO 04-01-2019 US-2019-112326 60540   
4 Avery 05-01-2019 US-2019-141817 19143   
... ... ... ... ...   
10189 Wilson Jones 30-12-2022 US-2022-143259 10009   
10190 GBC 30-12-2022 US-2022-115427 94533   
10191 Other 30-12-2022 US-2022-156720 80538   
10192 Other 30-12-2022 US-2022-143259 10009   
10193 Wilson Jones 30-12-2022 CA-2022-143500 C0A   
  
 Product Name Region \  
0 Message Book, Wirebound, Four 5 1/2" X 4" Form... Central   
1 GBC Standard Plastic Binding Systems Combs Central   
2 Avery 508 Central   
3 SAFCO Boltless Steel Shelving Central   
4 Avery Hi-Liter EverBold Pen Style Fluorescent ... East   
... ... ...   
10189 Wilson Jones Legal Size Ring Binders East   
10190 GBC Binding covers West   
10191 Bagged Rubber Bands West   
10192 Gear Head AU3700S Headset East   
10193 Wilson Jones Impact Binders East   
  
 Segment Ship Date Ship Mode State/Province \  
0 Consumer 07-01-2019 Standard Class Texas   
1 Home Office 08-01-2019 Standard Class Illinois   
2 Home Office 08-01-2019 Standard Class Illinois   
3 Home Office 08-01-2019 Standard Class Illinois   
4 Consumer 12-01-2019 Standard Class Pennsylvania   
... ... ... ... ...   
10189 Consumer 03-01-2023 Standard Class New York   
10190 Corporate 03-01-2023 Standard Class California   
10191 Consumer 03-01-2023 Standard Class Colorado   
10192 Consumer 03-01-2023 Standard Class New York   
10193 Consumer 03-01-2023 Standard Class Prince Edward Island   
  
 Sub-Category Discount Profit Quantity Sales   
0 Paper 0.2 5.5512 2 16.448   
1 Binders 0.8 -5.4870 2 3.540   
2 Labels 0.2 4.2717 3 11.784   
3 Storage 0.2 -64.7748 3 272.736   
4 Art 0.2 4.8840 3 19.536   
... ... ... ... ... ...   
10189 Binders 0.2 19.7910 3 52.776   
10190 Binders 0.2 6.4750 2 20.720   
10191 Fasteners 0.2 -0.6048 3 3.024   
10192 Phones 0.0 2.7279 7 90.930   
10193 Binders 0.2 -0.6048 3 3.024   
  
[10194 rows x 19 columns]

data.shape

(10194, 19)

data1=pd.read\_excel(r"C:\Users\HP\OneDrive\Documents\Rajavolu npi ekyc not done.xlsx")  
data1

Proforma-I Unnamed: 1 Unnamed: 2 \  
0 Sl.\nNo. Name of the Mandal Name of the Village   
1 NaN NaN NaN   
2 1 2 3   
3 1 Rajamahendravaram Rural Thorredu   
  
 Unnamed: 3 Unnamed: 4 Unnamed: 5 Unnamed: 6 \  
0 Eligible Applications NaN NaN Pattas Distributed   
1 NPI 90 Days Total NPI   
2 4 NaN NaN 5   
3 253 26 279 245   
  
 Unnamed: 7 Unnamed: 8 Unnamed: 9 ... Unnamed: 14 \  
0 NaN NaN Patta Not Distributed ... NaN   
1 90 Days Total NPI ... Total   
2 NaN NaN 6 ... NaN   
3 15 260 8 ... 248   
  
 Unnamed: 15 Unnamed: 16 Unnamed: 17 \  
0 EKYC Not done out of column(4)\n NaN NaN   
1 NPI 90 Days Total   
2 8 NaN NaN   
3 9 8 17   
  
 Unnamed: 18 Unnamed: 19 Unnamed: 20 \  
0 Pattas Given Possision Shown out of colum(5) NaN NaN   
1 NPI 90 Days Total   
2 9 NaN NaN   
3 245 15 260   
  
 Unnamed: 21 Unnamed: 22 Unnamed: 23   
0 Pattas Given Possision Not Shownout of colum(5) NaN NaN   
1 NPI 90 Days Total   
2 10 NaN NaN   
3 0 0 0   
  
[4 rows x 24 columns]

data1.shape

(4, 24)

data2=pd.read\_html(r"C:\Users\HP\OneDrive\Documents\Demo.html")  
data2

[ Month Savings  
 0 January $100  
 1 feb $200  
 2 march $300  
 3 april $400  
 4 may $500]

data2.shape

---------------------------------------------------------------------------  
AttributeError Traceback (most recent call last)  
~\AppData\Local\Temp\ipykernel\_4888\2942304966.py in <module>  
----> 1 data2.shape  
  
AttributeError: 'list' object has no attribute 'shape'

h=data2[0]

h

Month Savings  
0 January $100  
1 feb $200  
2 march $300  
3 april $400  
4 may $500

h.shape

(5, 2)

data3=pd.read\_json(r"C:\Users\HP\Downloads\sample2.json")

---------------------------------------------------------------------------  
ValueError Traceback (most recent call last)  
~\AppData\Local\Temp\ipykernel\_4888\2057421949.py in <module>  
----> 1 data3=pd.read\_json(r"C:\Users\HP\Downloads\sample2.json")  
  
~\anaconda3\lib\site-packages\pandas\util\\_decorators.py in wrapper(\*args, \*\*kwargs)  
 205 else:  
 206 kwargs[new\_arg\_name] = new\_arg\_value  
--> 207 return func(\*args, \*\*kwargs)  
 208   
 209 return cast(F, wrapper)  
  
~\anaconda3\lib\site-packages\pandas\util\\_decorators.py in wrapper(\*args, \*\*kwargs)  
 309 stacklevel=stacklevel,  
 310 )  
--> 311 return func(\*args, \*\*kwargs)  
 312   
 313 return wrapper  
  
~\anaconda3\lib\site-packages\pandas\io\json\\_json.py in read\_json(path\_or\_buf, orient, typ, dtype, convert\_axes, convert\_dates, keep\_default\_dates, numpy, precise\_float, date\_unit, encoding, encoding\_errors, lines, chunksize, compression, nrows, storage\_options)  
 610   
 611 with json\_reader:  
--> 612 return json\_reader.read()  
 613   
 614   
  
~\anaconda3\lib\site-packages\pandas\io\json\\_json.py in read(self)  
 744 obj = self.\_get\_object\_parser(self.\_combine\_lines(data\_lines))  
 745 else:  
--> 746 obj = self.\_get\_object\_parser(self.data)  
 747 self.close()  
 748 return obj  
  
~\anaconda3\lib\site-packages\pandas\io\json\\_json.py in \_get\_object\_parser(self, json)  
 766 obj = None  
 767 if typ == "frame":  
--> 768 obj = FrameParser(json, \*\*kwargs).parse()  
 769   
 770 if typ == "series" or obj is None:  
  
~\anaconda3\lib\site-packages\pandas\io\json\\_json.py in parse(self)  
 878 self.\_parse\_numpy()  
 879 else:  
--> 880 self.\_parse\_no\_numpy()  
 881   
 882 if self.obj is None:  
  
~\anaconda3\lib\site-packages\pandas\io\json\\_json.py in \_parse\_no\_numpy(self)  
 1130   
 1131 if orient == "columns":  
-> 1132 self.obj = DataFrame(  
 1133 loads(json, precise\_float=self.precise\_float), dtype=None  
 1134 )  
  
~\anaconda3\lib\site-packages\pandas\core\frame.py in \_\_init\_\_(self, data, index, columns, dtype, copy)  
 634 elif isinstance(data, dict):  
 635 # GH#38939 de facto copy defaults to False only in non-dict cases  
--> 636 mgr = dict\_to\_mgr(data, index, columns, dtype=dtype, copy=copy, typ=manager)  
 637 elif isinstance(data, ma.MaskedArray):  
 638 import numpy.ma.mrecords as mrecords  
  
~\anaconda3\lib\site-packages\pandas\core\internals\construction.py in dict\_to\_mgr(data, index, columns, dtype, typ, copy)  
 500 # TODO: can we get rid of the dt64tz special case above?  
 501   
--> 502 return arrays\_to\_mgr(arrays, columns, index, dtype=dtype, typ=typ, consolidate=copy)  
 503   
 504   
  
~\anaconda3\lib\site-packages\pandas\core\internals\construction.py in arrays\_to\_mgr(arrays, columns, index, dtype, verify\_integrity, typ, consolidate)  
 118 # figure out the index, if necessary  
 119 if index is None:  
--> 120 index = \_extract\_index(arrays)  
 121 else:  
 122 index = ensure\_index(index)  
  
~\anaconda3\lib\site-packages\pandas\core\internals\construction.py in \_extract\_index(data)  
 675   
 676 if have\_dicts:  
--> 677 raise ValueError(  
 678 "Mixing dicts with non-Series may lead to ambiguous ordering."  
 679 )  
  
ValueError: Mixing dicts with non-Series may lead to ambiguous ordering.

data4=pd.read\_json(r"C:\Users\HP\Downloads\Sample - Superstore\_Orders.json")

data4

Category City Country \  
0 Office Supplies Houston {'Region': 'United States'}   
1 Office Supplies Naperville {'Region': 'United States'}   
2 Office Supplies Naperville {'Region': 'United States'}   
3 Office Supplies Naperville {'Region': 'United States'}   
4 Office Supplies Philadelphia {'Region': 'United States'}   
... ... ... ...   
10189 Office Supplies New York City {'Region': 'United States'}   
10190 Office Supplies Fairfield {'Region': 'United States'}   
10191 Office Supplies Loveland {'Region': 'United States'}   
10192 Technology New York City {'Region': 'United States'}   
10193 Office Supplies Charlottetown {'Region': 'Canada'}   
  
 Customer Name Manufacturer Order Date Order ID \  
0 Darren Powers Message Book 03-01-2019 US-2019-103800   
1 Phillina Ober GBC 04-01-2019 US-2019-112326   
2 Phillina Ober Avery 04-01-2019 US-2019-112326   
3 Phillina Ober SAFCO 04-01-2019 US-2019-112326   
4 Mick Brown Avery 05-01-2019 US-2019-141817   
... ... ... ... ...   
10189 Patrick O'Donnell Wilson Jones 30-12-2022 US-2022-143259   
10190 Erica Bern GBC 30-12-2022 US-2022-115427   
10191 Jill Matthias Other 30-12-2022 US-2022-156720   
10192 Patrick O'Donnell Other 30-12-2022 US-2022-143259   
10193 Harry Olson Wilson Jones 30-12-2022 CA-2022-143500   
  
 Postal Code Product Name Region \  
0 77095 Message Book, Wirebound, Four 5 1/2" X 4" Form... Central   
1 60540 GBC Standard Plastic Binding Systems Combs Central   
2 60540 Avery 508 Central   
3 60540 SAFCO Boltless Steel Shelving Central   
4 19143 Avery Hi-Liter EverBold Pen Style Fluorescent ... East   
... ... ... ...   
10189 10009 Wilson Jones Legal Size Ring Binders East   
10190 94533 GBC Binding covers West   
10191 80538 Bagged Rubber Bands West   
10192 10009 Gear Head AU3700S Headset East   
10193 C0A Wilson Jones Impact Binders East   
  
 Segment Ship Date Ship Mode \  
0 Consumer 07-01-2019 Standard Class   
1 Home Office 08-01-2019 Standard Class   
2 Home Office 08-01-2019 Standard Class   
3 Home Office 08-01-2019 Standard Class   
4 Consumer 12-01-2019 Standard Class   
... ... ... ...   
10189 Consumer 03-01-2023 Standard Class   
10190 Corporate 03-01-2023 Standard Class   
10191 Consumer 03-01-2023 Standard Class   
10192 Consumer 03-01-2023 Standard Class   
10193 Consumer 03-01-2023 Standard Class   
  
 State Sub-Category Discount Profit \  
0 {'Province': 'Texas'} Paper 0.2 5.5512   
1 {'Province': 'Illinois'} Binders 0.8 -5.4870   
2 {'Province': 'Illinois'} Labels 0.2 4.2717   
3 {'Province': 'Illinois'} Storage 0.2 -64.7748   
4 {'Province': 'Pennsylvania'} Art 0.2 4.8840   
... ... ... ... ...   
10189 {'Province': 'New York'} Binders 0.2 19.7910   
10190 {'Province': 'California'} Binders 0.2 6.4750   
10191 {'Province': 'Colorado'} Fasteners 0.2 -0.6048   
10192 {'Province': 'New York'} Phones 0.0 2.7279   
10193 {'Province': 'Prince Edward Island'} Binders 0.2 -0.6048   
  
 Quantity Sales   
0 2 16.448   
1 2 3.540   
2 3 11.784   
3 3 272.736   
4 3 19.536   
... ... ...   
10189 3 52.776   
10190 2 20.720   
10191 3 3.024   
10192 7 90.930   
10193 3 3.024   
  
[10194 rows x 19 columns]

data4.shape

(10194, 19)