Exercises Solutions

Mastering Python

الدرس #8

Pandas and Data Analysis تحليل البيانات

By:

Hussam Hourani

V1.0 - NOV 2019

Exercises

1. Write a Pandas program to create and print a one-dimensional array-like object containing an array of data([2, 4, 6, 8, 10]).

Exercises

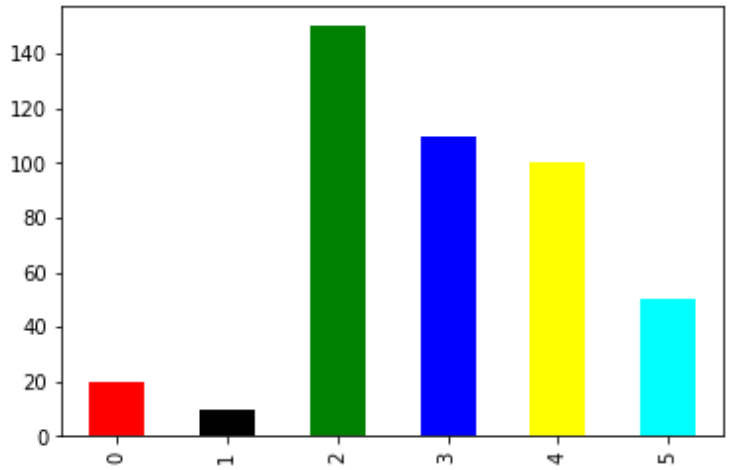
2. Write a Pandas program to convert the following dictionary to a Pandas series and print the series.

d1 = {'a': 100, 'b': 200, 'c':300, 'd':400, 'e':800}

Exercises

3. Write a Pandas program to describe and draw data as per the following graph:

data = [20, 10, 150, 110, 100, 50]

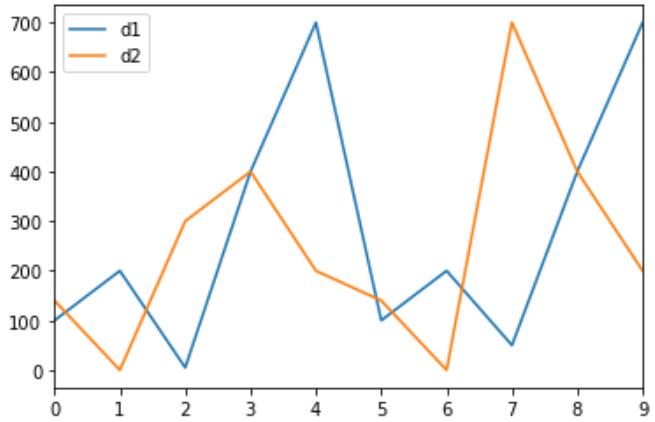


Exercises

4. Write a Pandas program to describe and draw data as per the following graph:

Data = {'d1':[100,200,5,400,700,100,200,50,400,700],

'd2':[140,0,300,400,200,140,0,700,400,200]



Exercises

5. Write a Pandas program to get the 3 arrays values and print them

Note: Sample data: {

'X':[78,85,96,80,86], 'Y':[84,94,89,83,86],'Z':[86,97,96,72,83]}

Expected Output:

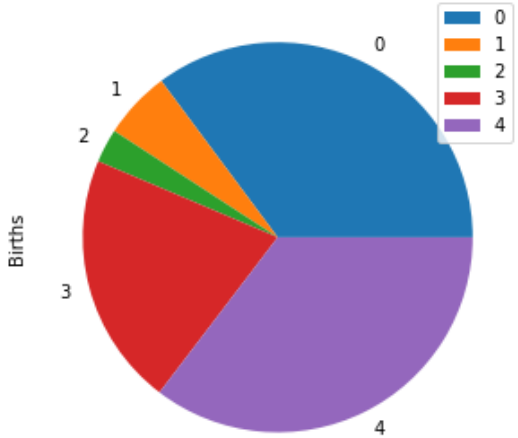
```
    X  Y   Z
0  78 84  86
1  85 94  97
2  96 89  96
3  80 83  72
4  86 86  83
```

Exercises

6. Write a Pandas program print the following dataset as Datafram and to drware the chart as following (use plot = df.plot.pie(y='births', figsize=(5, 5))

names = ['Bob','Jessica','Mary','John','Mel']

births = [968, 155, 77, 578, 973]



Exercises

7. Write a Pandas program to read from csv file the following data (tab separator) and print and describe the data, then write to another csv file with comma(,) separator

Index	Number
0	100
1	2200
2	300
3	400
4	500
5	600
6	700
7	800
8	900

Exercises

7. Write a Pandas program to creat the following :

dates = pd.date_range('20000101', periods=4)

Generate dataframe as np.random.randn(4, 2)

Set the columns as columns=['A','B']

index=dates

Then find the following :

- print (df)
- print(df.head(2))
- print(df[['A']])
- print(df[0:1])
- print(df['20000102':'20000104'])
- print(df.loc['20000102':'20000104', ['A']])
- print(df.iloc[:, 1:2])
- print(df[df > 0])
- print(df[df.B > 0])
- df['A'] = [100,200,300,100]
- print (df)
- print(df[df['A'].isin([200, 300])])

Master in Software Engineering

Hussam Hourani has over 25 years of Organizations Transformation, VROs, PMO, Large Scale and Enterprise Programs Global Delivery, Leadership, Business Development and Management Consulting. His client experience is wide ranging across many sectors but focuses on Performance Enhancement, Transformation, Enterprise Program Management, Artificial Intelligence and Data Science.