What is Contingency Table?

1. It is a table, when shows count of elements in dataframe
2. It is a tally of counts between two or more categorical variables.
3. It is a table of metrics
4. It is a table of unic row of original data

---------------------------------------------------------------------------------------

What function and what library provide to create Contingency Table?

1. numpy.contingency\_table()
2. pandas.contingency\_table()
3. pandas.crosstab()
4. statmodels.contingency\_table()

---------------------------------------------------------------------------------------

What will be printed after this code execution (dataframe on picture)

  
pd.crosstab(index=df1['Y'], columns=df1['X2'])

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | | **X2** | **0** | **1** | | --- | --- | --- | | **Y** |  |  | | **0** | 2 | 3 | | **1** | 5 | 2 | |
| 2 | | **X2** | **0** | **1** | | --- | --- | --- | | **Y** |  |  | | **0** | 2 | 3 | | **1** | 1 | 3 | |
| 3 | | **X2** | **0** | **1** | | --- | --- | --- | | **Y** |  |  | | **0** | 1 | 3 | | **1** | 1 | 2 | |
| 4 | | **X2** | **0** | **1** | | --- | --- | --- | | **Y** |  |  | | **0** | 2 | 3 | | **1** | 1 | 2 | |

---------------------------------------------------------------------------------------

we have the following code

pd.crosstab(index=df1['Y'], columns=df1['X1'], margins=True)

and output in dataframe format

| **X1** | **0** | **1** | **All** |
| --- | --- | --- | --- |
| **Y** |  |  |  |
| **0** | 3 | 2 | 5 |
| **1** | 1 | 2 | 3 |
| **All** | 4 | 4 | 8 |

What is All (row and coll) mean?

1. it is norm of classification algorithm
2. it is margins (subtotals)
3. it is sum by col and by row
4. it is difference of row and col

---------------------------------------------------------------------------------------

What will be printed after this code execution(dataset in third questtion)

pd.crosstab(index=df1['Y'], columns=df1['X1'],normalize="all", margins=True)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | | **X1** | **0** | **1** | **All** | | --- | --- | --- | --- | | **Y** |  |  |  | | **0** | 3 | 2 | 5 | | **1** | 1 | 2 | 3 | | **All** | 4 | 4 | 8 | |
| 2 | | **X1** | **0** | **1** | | --- | --- | --- | | **Y** |  |  | | **0** | 0.600000 | 0.400000 | | **1** | 0.333333 | 0.666667 | | **All** | 0.500000 | 0.500000 | |
| 3 | | **X1** | **0** | **1** | **All** | | --- | --- | --- | --- | | **Y** |  |  |  | | **0** | 0.375 | 0.25 | 0.625 | | **1** | 0.125 | 0.25 | 0.375 | | **All** | 0.500 | 0.50 | 1.000 | |
| 4 | | **X1** | **0** | **1** | **All** | | --- | --- | --- | --- | | **Y** |  |  |  | | **0** | 0.375 | 0.35 | 0.625 | | **1** | 0.125 | 0.35 | 0.375 | | **All** | 0.500 | 0.50 | 1.000 | |

---------------------------------------------------------------------------------------

We have following code

**pd.crosstab(index=df['Country'], columns=df['Product'], margins=True)**

and crosstab result

**Product Comp Radio TV All**

**Country**

**A 1 0 3 4**

**B 3 2 3 8**

**C 2 3 3 8**

**All 6 5 9 20**

Select incorrect answers (many variants)

1. A total of 4 orders were made from country A.
2. A total of 8 orders were made from country B.
3. A total of 3 orders were made from country B.
4. A total of 8 orders were made from country C.
5. A total of 20 orders were made from country C.
6. A total of 6 computers were purchased.
7. A total of 5 radios were purchased.
8. A total of 6 radios were purchased.
9. A total of 9 TV’s were purchased.