classtask-may24

July 18, 2024

```python
[7]: import pandas as pd
     import numpy as np
```

```python
[15]: #duplicates

      std=pd.DataFrame({'id':[1000,1001,1002,1003,1004,1000,1001,1002,1008,1001],
                        'name':
       ↪['jayaram','sai','pradeep','lebin','dhivya','nithish','rajesh','sai','rani','lebin'],
                        'age':[22,23,25,22,26,28,24,24,23,25],
                        'city':
       ↪['salem','erode','coimbatore','chennai','trichy','thindukal','madurai','chennai','chennai',

      std
```

```
[15]:      id     name  age        city
      0  1000  jayaram   22       salem
      1  1001      sai   23       erode
      2  1002  pradeep   25  coimbatore
      3  1003    lebin   22     chennai
      4  1004   dhivya   26      trichy
      5  1000  nithish   28   thindukal
      6  1001   rajesh   24     madurai
      7  1002      sai   24     chennai
      8  1008     rani   23     chennai
      9  1001    lebin   25       erode
```

```python
[16]: std.duplicated('name')
```

```
[16]: 0    False
      1    False
      2    False
      3    False
      4    False
      5    False
      6    False
      7     True
      8    False
```

```
9     True
dtype: bool
```

[17]: `std.duplicated('id')`

```
[17]: 0    False
      1    False
      2    False
      3    False
      4    False
      5     True
      6     True
      7     True
      8    False
      9     True
      dtype: bool
```

[18]: `std.duplicated('city')`

```
[18]: 0    False
      1    False
      2    False
      3    False
      4    False
      5    False
      6    False
      7     True
      8     True
      9     True
      dtype: bool
```

[19]: `std.drop_duplicates('id')`

```
[19]:      id     name  age         city
      0  1000  jayaram   22        salem
      1  1001      sai   23        erode
      2  1002  pradeep   25  coimbatore
      3  1003    lebin   22      chennai
      4  1004   dhivya   26       trichy
      8  1008     rani   23      chennai
```

[22]: `std.drop_duplicates('name')`

```
[22]:      id     name  age         city
      0  1000  jayaram   22        salem
      1  1001      sai   23        erode
      2  1002  pradeep   25  coimbatore
```

2

```
3   1003    lebin   22     chennai
4   1004   dhivya   26      trichy
5   1000  nithish   28   thindukal
6   1001   rajesh   24     madurai
8   1008     rani   23     chennai
```

[23]:
```python
#concatenation

std1=pd.DataFrame({'id':[1000,1001,1002,1003,1004,1000,1001,1002,1008,1001],
                  'name':
  ↪['jayaram','sai','pradeep','lebin','dhivya','nithish','rajesh','sai','rani','lebin'],
                  'age':[22,23,25,22,26,28,24,24,23,25],
                  'city':
  ↪['salem','erode','coimbatore','chennai','trichy','thindukal','madurai','chennai','chennai',
print(std1)

std2=pd.DataFrame({'id':[1011,1012,1013,1014,1015,1016,1017,1018,1019,1020],
                  'name':
  ↪['guna','yazhini','preethi','stephen','prince','gobi','kingsley','suseela','ponraj','angel']
                  'age':[29,23,28,23,20,28,23,27,27,30],
                  'city':
  ↪['erode','coimbatore','tripur','salem','madurai','dhindukal','chennai','gobi','vilupuram','
print(std2)
```

```
     id      name  age         city
0  1000   jayaram   22        salem
1  1001       sai   23        erode
2  1002   pradeep   25   coimbatore
3  1003     lebin   22      chennai
4  1004    dhivya   26       trichy
5  1000   nithish   28    thindukal
6  1001    rajesh   24      madurai
7  1002       sai   24      chennai
8  1008      rani   23      chennai
9  1001     lebin   25        erode
     id      name  age         city
0  1011      guna   29        erode
1  1012   yazhini   23   coimbatore
2  1013   preethi   28       tripur
3  1014   stephen   23        salem
4  1015    prince   20      madurai
5  1016      gobi   28    dhindukal
6  1017  kingsley   23      chennai
7  1018   suseela   27         gobi
8  1019    ponraj   27    vilupuram
9  1020     angel   30        hosur
```

```
[25]: pd.concat([std1,std2])
```

```
[25]:        id      name  age        city
      0  1000   jayaram   22       salem
      1  1001       sai   23       erode
      2  1002   pradeep   25  coimbatore
      3  1003     lebin   22     chennai
      4  1004    dhivya   26      trichy
      5  1000   nithish   28   thindukal
      6  1001    rajesh   24     madurai
      7  1002       sai   24     chennai
      8  1008      rani   23     chennai
      9  1001     lebin   25       erode
      0  1011      guna   29       erode
      1  1012   yazhini   23  coimbatore
      2  1013   preethi   28      tripur
      3  1014   stephen   23       salem
      4  1015    prince   20     madurai
      5  1016      gobi   28   dhindukal
      6  1017  kingsley   23     chennai
      7  1018   suseela   27        gobi
      8  1019    ponraj   27   vilupuram
      9  1020     angel   30       hosur
```

```
[28]: pd.concat([std1,std2],ignore_index=2)
```

```
[28]:         id      name  age        city
      0   1000   jayaram   22       salem
      1   1001       sai   23       erode
      2   1002   pradeep   25  coimbatore
      3   1003     lebin   22     chennai
      4   1004    dhivya   26      trichy
      5   1000   nithish   28   thindukal
      6   1001    rajesh   24     madurai
      7   1002       sai   24     chennai
      8   1008      rani   23     chennai
      9   1001     lebin   25       erode
      10  1011      guna   29       erode
      11  1012   yazhini   23  coimbatore
      12  1013   preethi   28      tripur
      13  1014   stephen   23       salem
      14  1015    prince   20     madurai
      15  1016      gobi   28   dhindukal
      16  1017  kingsley   23     chennai
      17  1018   suseela   27        gobi
      18  1019    ponraj   27   vilupuram
      19  1020     angel   30       hosur
```

```
[ ]:
```