EDA on Udemy Finance Data

```python
import seaborn as sns
import pandas as pd
import numpy as np
import matplotlib.pyplot  as plt

br = pd.read_csv(r"D:\
udemy_output_All_Finance__Accounting_p1_p626.csv")
br
```

```
              id                                              title  \
0         762616  The Complete SQL Bootcamp 2020: Go from Zero t...
1         937678  Tableau 2020 A-Z: Hands-On Tableau Training fo...
2        1361790              PMP Exam Prep Seminar -  PMBOK Guide 6
3         648826            The Complete Financial Analyst Course 2020
4         637930  An Entire MBA in 1 Course:Award Winning Busine...
...          ...                                                ...
13603    3005408                       מושגים בסיסיים באופציות חלק ב
13604    3467284                             Programa de Integridade
13605    3368222  Goodwill and Method of Calculating Goodwill (A...
13606    3211345                                  Poderoso Investidor
13607    2935720            Acabou a Previdência e agora? -  Volume 03

                                               url  is_paid  \
0                   /course/the-complete-sql-bootcamp/     True
1                                /course/tableau10/     True
2                        /course/pmp-pmbok6-35-pdus/     True
3          /course/the-complete-financial-analyst-course/     True
4       /course/an-entire-mba-in-1-courseaward-winning...     True
...                                              ...      ...
13603                             /course/1-ndgzhu/    False
13604            /course/iaudit_programa_de_integridade/    False
13605   /course/goodwill-and-method-of-calculating-goo...    False
13606                       /course/poderoso-investidor/    False
13607     /course/acabou-a-previdencia-e-agora-volume-03/     True

       num_subscribers  avg_rating  avg_rating_recent   rating
num_reviews  \
0                295509     4.66019            4.67874  4.67874
78006
1                209070     4.58956            4.60015  4.60015
54581
2                155282     4.59491            4.59326  4.59326
52653
3                245860     4.54407            4.53772  4.53772
46447
4                374836     4.47080            4.47173  4.47173
41630
...                 ...         ...                ...      ...
```

```
...
13603                   47      0.00000              0.00000  0.00000
0
13604                   19      0.00000              0.00000  0.00000
0
13605                   47      0.00000              0.00000  0.00000
0
13606                   48      0.00000              0.00000  0.00000
0
13607                    0      0.00000              0.00000  0.00000
0

       is_wishlisted  num_published_lectures
num_published_practice_tests  \
0              False                       84
0
1              False                       78
0
2              False                      292
2
3              False                      338
0
4              False                       83
0
...              ...                      ...
...
13603          False                        6
0
13604          False                        5
0
13605          False                        5
0
13606          False                       13
0
13607          False                       14
0

                    created        published_time
discount_price__amount  \
0      2016-02-14T22:57:48Z  2016-04-06T05:16:11Z
455.0
1      2016-08-22T12:10:18Z  2016-08-23T16:59:49Z
455.0
2      2017-09-26T16:32:48Z  2017-11-14T23:58:14Z
455.0
3      2015-10-23T13:34:35Z  2016-01-21T01:38:48Z
455.0
4      2015-10-12T06:39:46Z  2016-01-11T21:39:33Z
455.0
```

```
...                  ...                     ...
...
13603  2020-04-15T10:20:17Z   2020-04-23T02:06:55Z
NaN
13604  2020-09-01T17:57:15Z   2020-09-03T15:26:47Z
NaN
13605  2020-07-27T02:29:18Z   2020-07-31T12:14:40Z
NaN
13606  2020-06-07T01:29:53Z   2020-07-15T10:35:38Z
NaN
13607  2020-03-30T19:10:58Z   2020-04-02T16:33:32Z
NaN

       discount_price__currency discount_price__price_string  \
0                           INR                         ₹455
1                           INR                         ₹455
2                           INR                         ₹455
3                           INR                         ₹455
4                           INR                         ₹455
...                         ...                          ...
13603                       NaN                          NaN
13604                       NaN                          NaN
13605                       NaN                          NaN
13606                       NaN                          NaN
13607                       NaN                          NaN

       price_detail__amount price_detail__currency
price_detail__price_string
0                    8640.0                    INR
₹8,640
1                    8640.0                    INR
₹8,640
2                    8640.0                    INR
₹8,640
3                    8640.0                    INR
₹8,640
4                    8640.0                    INR
₹8,640
...                     ...                    ...
...
13603                   NaN                    NaN
NaN
13604                   NaN                    NaN
NaN
13605                   NaN                    NaN
NaN
13606                   NaN                    NaN
NaN
13607                   NaN                    NaN
```

```
NaN

[13608 rows x 20 columns]
```

Question no3

IN this given data of udemy's finance accounting data, we are going to do the analysis of no of courses,subscriber,prices etc it have we will achieve some experience how to work with big data of finance and accounts.

Question no 4

br.isnull().sum()

Here we see the overall null value percentage in this data

```
br.isnull().mean()

id                             0.000000
title                          0.000000
url                            0.000000
is_paid                        0.000000
num_subscribers                0.000000
avg_rating                     0.000000
avg_rating_recent              0.000000
rating                         0.000000
num_reviews                    0.000000
is_wishlisted                  0.000000
num_published_lectures         0.000000
num_published_practice_tests   0.000000
created                        0.000000
published_time                 0.000000
discount_price__amount         0.103101
discount_price__currency       0.103101
discount_price__price_string   0.103101
price_detail__amount           0.036523
price_detail__currency         0.036523
price_detail__price_string     0.036523
dtype: float64
```

sir i have made a special column to fill the null value

Handle Missing Value

1.discount_price__amount column

```
br["discount_price__amount"].unique()
```

```
array([ 455.,   500.,    nan,  540.,   546.,   533.,   481.,   488.,   474.,
        462., 2880.,   494., 1600.,   800., 2080.,   520.,   468.,   507.,
        700.,   514.,   552.,   640., 3200.,   960., 2720., 1920., 1680.,
        611., 2400.,   720., 1280., 2000., 2240., 3040., 2160., 1360.,
       1760., 2480., 1120.,   760., 1520., 1440.,   480., 2320., 1200.,
       2560., 1840., 2800., 2960.,   560., 2640., 1040.,   880., 3120.])
```

```
br['discount_price__amount'].median()
```

```
455.0
```

```
br["discount_price__amount"].isnull().sum()
```

```
1403
```

```
sns.heatmap(br.isnull())
```

```
<Axes: >
```
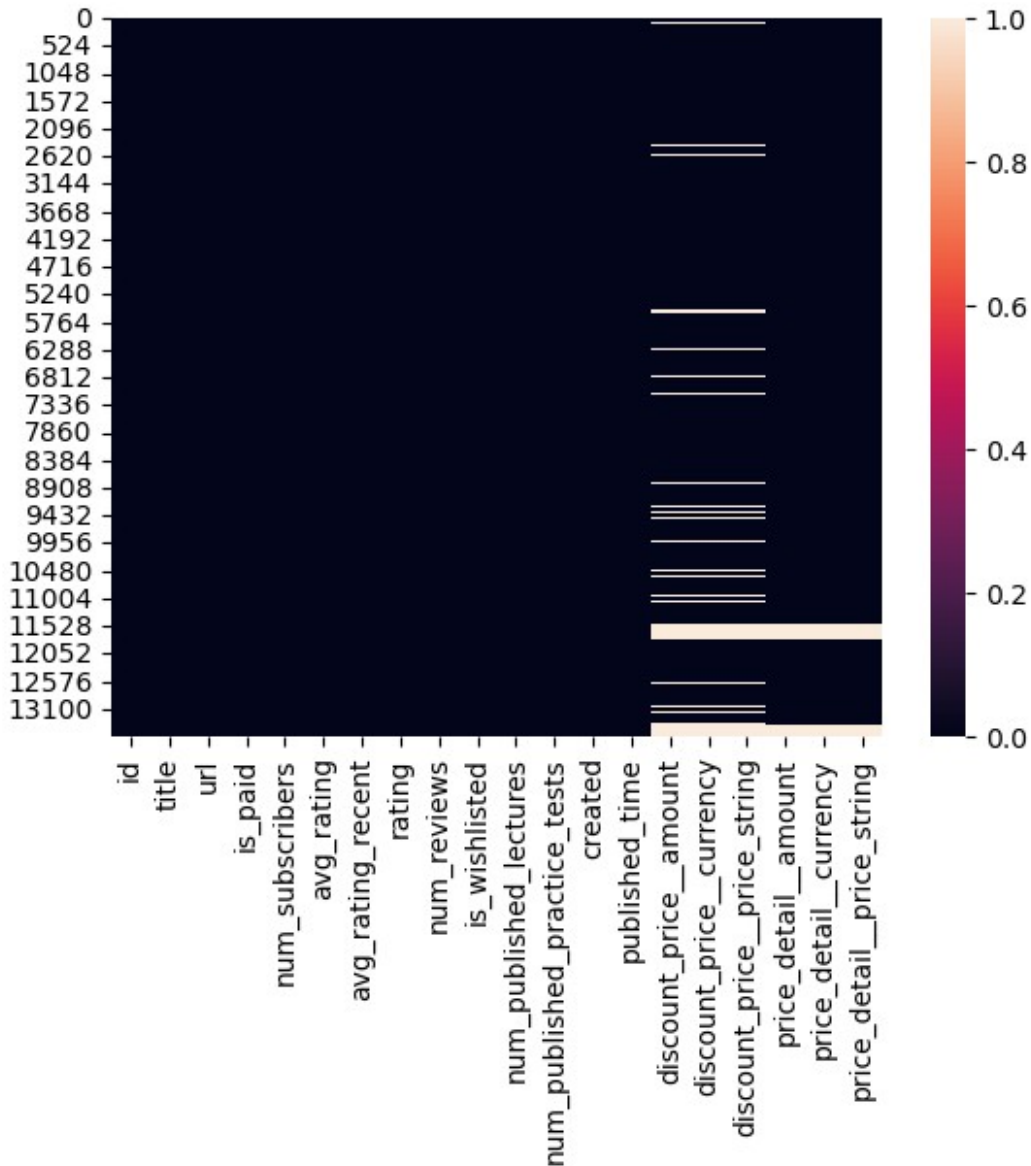
```
type(br['discount_price__amount'].median())

float

br["discount_price__amount_notnull"] =
br["discount_price__amount"].fillna(br["discount_price__amount"].media
n())# jo bhi column ka median ka wo null value me fill hojaega

br# here i make discount_price__amount_notnullcolumn
```

	id	title \
0	762616	The Complete SQL Bootcamp 2020: Go from Zero t...
1	937678	Tableau 2020 A-Z: Hands-On Tableau Training fo...
2	1361790	PMP Exam Prep Seminar - PMBOK Guide 6

```
3         648826              The Complete Financial Analyst Course 2020
4         637930       An Entire MBA in 1 Course:Award Winning Busine...
...          ...                                                    ...
13603    3005408                         מושגים בסיסיים באופציות חלק ב
13604    3467284                                   Programa de Integridade
13605    3368222       Goodwill and Method of Calculating Goodwill (A...
13606    3211345                                        Poderoso Investidor
13607    2935720           Acabou a Previdência e agora? -  Volume 03

                                                  url  is_paid  \
0                  /course/the-complete-sql-bootcamp/     True
1                                   /course/tableau10/     True
2                             /course/pmp-pmbok6-35-pdus/     True
3          /course/the-complete-financial-analyst-course/     True
4       /course/an-entire-mba-in-1-courseaward-winning...     True
...                                                  ...      ...
13603                                   /course/1-ndgzhu/    False
13604              /course/iaudit_programa_de_integridade/    False
13605   /course/goodwill-and-method-of-calculating-goo...    False
13606                         /course/poderoso-investidor/    False
13607     /course/acabou-a-previdencia-e-agora-volume-03/     True

       num_subscribers  avg_rating  avg_rating_recent   rating
num_reviews  \
0              295509     4.66019            4.67874  4.67874
78006
1              209070     4.58956            4.60015  4.60015
54581
2              155282     4.59491            4.59326  4.59326
52653
3              245860     4.54407            4.53772  4.53772
46447
4              374836     4.47080            4.47173  4.47173
41630
...               ...         ...                ...      ...
...
13603              47     0.00000            0.00000  0.00000
0
13604              19     0.00000            0.00000  0.00000
0
13605              47     0.00000            0.00000  0.00000
0
13606              48     0.00000            0.00000  0.00000
0
13607               0     0.00000            0.00000  0.00000
0

       is_wishlisted  ...  num_published_practice_tests
created  \
```

```
0              False  ...                         0  2016-02-
14T22:57:48Z
1              False  ...                         0  2016-08-
22T12:10:18Z
2              False  ...                         2  2017-09-
26T16:32:48Z
3              False  ...                         0  2015-10-
23T13:34:35Z
4              False  ...                         0  2015-10-
12T06:39:46Z
...              ...  ...                       ...
...
13603          False  ...                         0  2020-04-
15T10:20:17Z
13604          False  ...                         0  2020-09-
01T17:57:15Z
13605          False  ...                         0  2020-07-
27T02:29:18Z
13606          False  ...                         0  2020-06-
07T01:29:53Z
13607          False  ...                         0  2020-03-
30T19:10:58Z

            published_time discount_price__amount
discount_price__currency  \
0       2016-04-06T05:16:11Z                   455.0
INR
1       2016-08-23T16:59:49Z                   455.0
INR
2       2017-11-14T23:58:14Z                   455.0
INR
3       2016-01-21T01:38:48Z                   455.0
INR
4       2016-01-11T21:39:33Z                   455.0
INR
...                      ...                     ...
...
13603   2020-04-23T02:06:55Z                     NaN
NaN
13604   2020-09-03T15:26:47Z                     NaN
NaN
13605   2020-07-31T12:14:40Z                     NaN
NaN
13606   2020-07-15T10:35:38Z                     NaN
NaN
13607   2020-04-02T16:33:32Z                     NaN
NaN

      discount_price__price_string price_detail__amount  \
```

```
0                          ₹455                    8640.0
1                          ₹455                    8640.0
2                          ₹455                    8640.0
3                          ₹455                    8640.0
4                          ₹455                    8640.0
...                         ...                       ...
13603                       NaN                       NaN
13604                       NaN                       NaN
13605                       NaN                       NaN
13606                       NaN                       NaN
13607                       NaN                       NaN

       price_detail__currency price_detail__price_string  \
0                          INR                     ₹8,640
1                          INR                     ₹8,640
2                          INR                     ₹8,640
3                          INR                     ₹8,640
4                          INR                     ₹8,640
...                        ...                        ...
13603                      NaN                        NaN
13604                      NaN                        NaN
13605                      NaN                        NaN
13606                      NaN                        NaN
13607                      NaN                        NaN

       discount_price__amount_notnull
0                               455.0
1                               455.0
2                               455.0
3                               455.0
4                               455.0
...                               ...
13603                           455.0
13604                           455.0
13605                           455.0
13606                           455.0
13607                           455.0

[13608 rows x 21 columns]
```
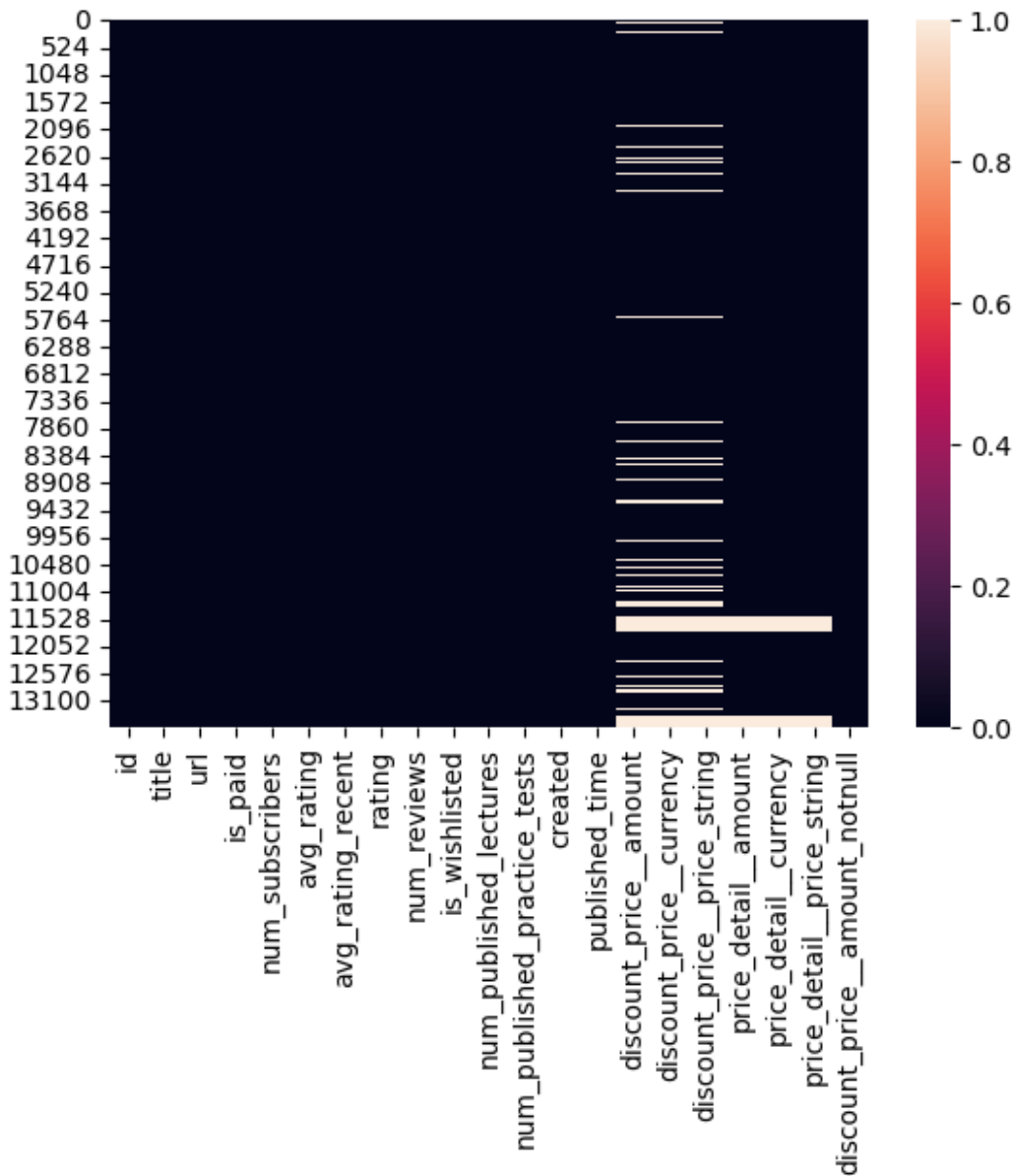
We are going to plot the data with the help of Heatmap to show the null value

```
sns.heatmap(br.isnull())

<Axes: >
```

we are going to check null value with in alpha numerical form

```
br["discount_price__amount_notnull"].isnull().sum()

0

br["discount_price__amount"].isnull().sum()

1403
```

```
br.groupby('discount_price__amount_notnull').size()

discount_price__amount_notnull
455.0      12813
462.0         47
468.0        135
474.0         30
480.0         17
481.0         44
488.0         27
494.0         22
500.0          7
507.0          5
514.0          6
520.0         34
533.0          2
540.0          2
546.0          2
552.0          2
560.0          5
611.0          1
640.0         24
700.0         86
720.0         12
760.0          1
800.0         39
880.0          3
960.0         13
1040.0         5
1120.0         7
1200.0         8
1280.0         6
1360.0         5
1440.0        11
1520.0         8
1600.0        27
1680.0         3
1760.0         4
1840.0         1
1920.0         8
2000.0         5
2080.0         9
2160.0         5
2240.0         7
2320.0         2
2400.0        16
2480.0         1
2560.0         2
2640.0         2
2720.0         6
```

```
2800.0         3
2880.0        10
2960.0         2
3040.0         3
3120.0         2
3200.0        61
dtype: int64

br.groupby('discount_price__amount_notnull').size()

discount_price__amount_notnull
455.0     12813
462.0        47
468.0       135
474.0        30
480.0        17
481.0        44
488.0        27
494.0        22
500.0         7
507.0         5
514.0         6
520.0        34
533.0         2
540.0         2
546.0         2
552.0         2
560.0         5
611.0         1
640.0        24
700.0        86
720.0        12
760.0         1
800.0        39
880.0         3
960.0        13
1040.0        5
1120.0        7
1200.0        8
1280.0        6
1360.0        5
1440.0       11
1520.0        8
1600.0       27
1680.0        3
1760.0        4
1840.0        1
1920.0        8
2000.0        5
2080.0        9
```

```
2160.0        5
2240.0        7
2320.0        2
2400.0       16
2480.0        1
2560.0        2
2640.0        2
2720.0        6
2800.0        3
2880.0       10
2960.0        2
3040.0        3
3120.0        2
3200.0       61
dtype: int64
```
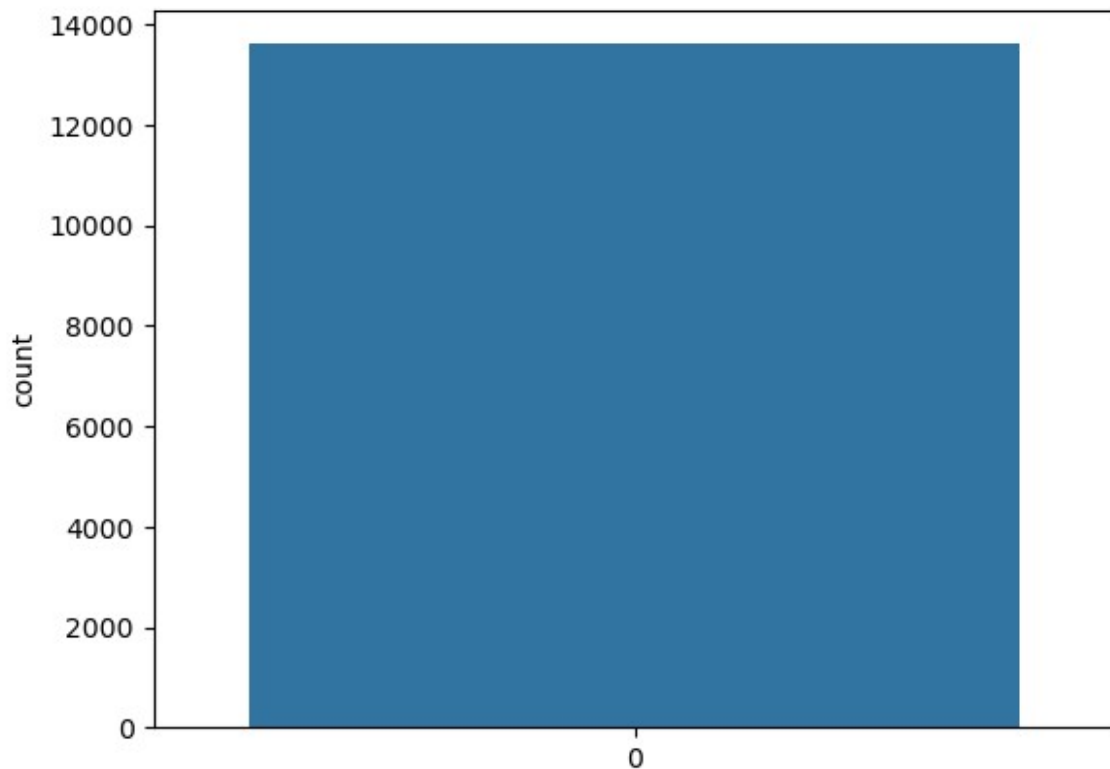
```
sns.countplot(br["discount_price__amount_notnull"])
```

```
<Axes: ylabel='count'>
```



```
br["discount_price__amount_notnull"].describe()
```

```
count    13608.000000
mean       489.928645
std        253.920382
```

```
min          455.000000
25%          455.000000
50%          455.000000
75%          455.000000
max         3200.000000
Name: discount_price__amount_notnull, dtype: float64

sns.countplot(br["discount_price__amount_notnull"])

<Axes: ylabel='count'>
```
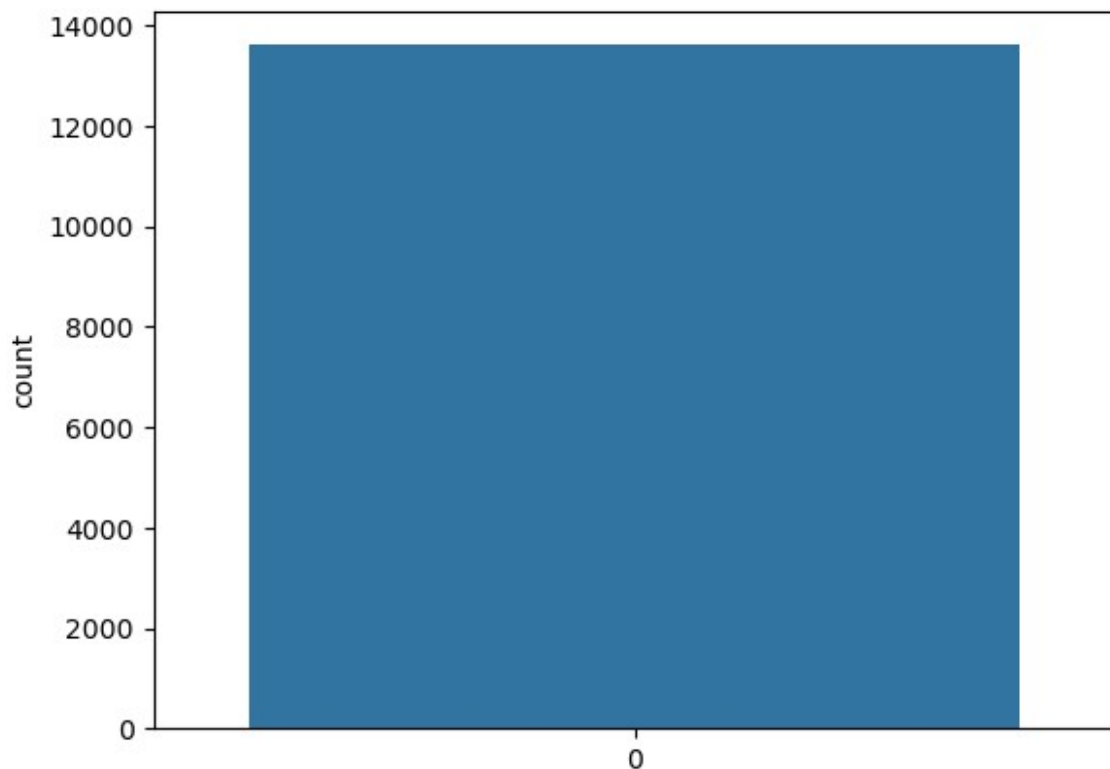

2) In discount_price__currency column

This is the categorical column

```
br["discount_price__currency"].isnull().sum()

1403
```

There are 1403 null value are there in discount_price__currency column

```
br["discount_price__currency"].unique()

array(['INR', nan], dtype=object)

br.groupby('discount_price__currency').size()

discount_price__currency
INR     12205
dtype: int64
```

We have used ' INR ' to fill all the null values in this discount_price__currency column.

Because we are tackling with the Indian currency (INR) data in this column

```
br["discount_price__currency_notnull"] =
br["discount_price__currency"].fillna("INR")
```

we are going to check null value with in alpha numerical form

o null value are there

```
br['discount_price__currency_notnull'].isnull().sum()

0
```

We are going to plot the data with the help of Heatmap to show the null value

```
sns.heatmap(br.isnull())

<Axes: >
```
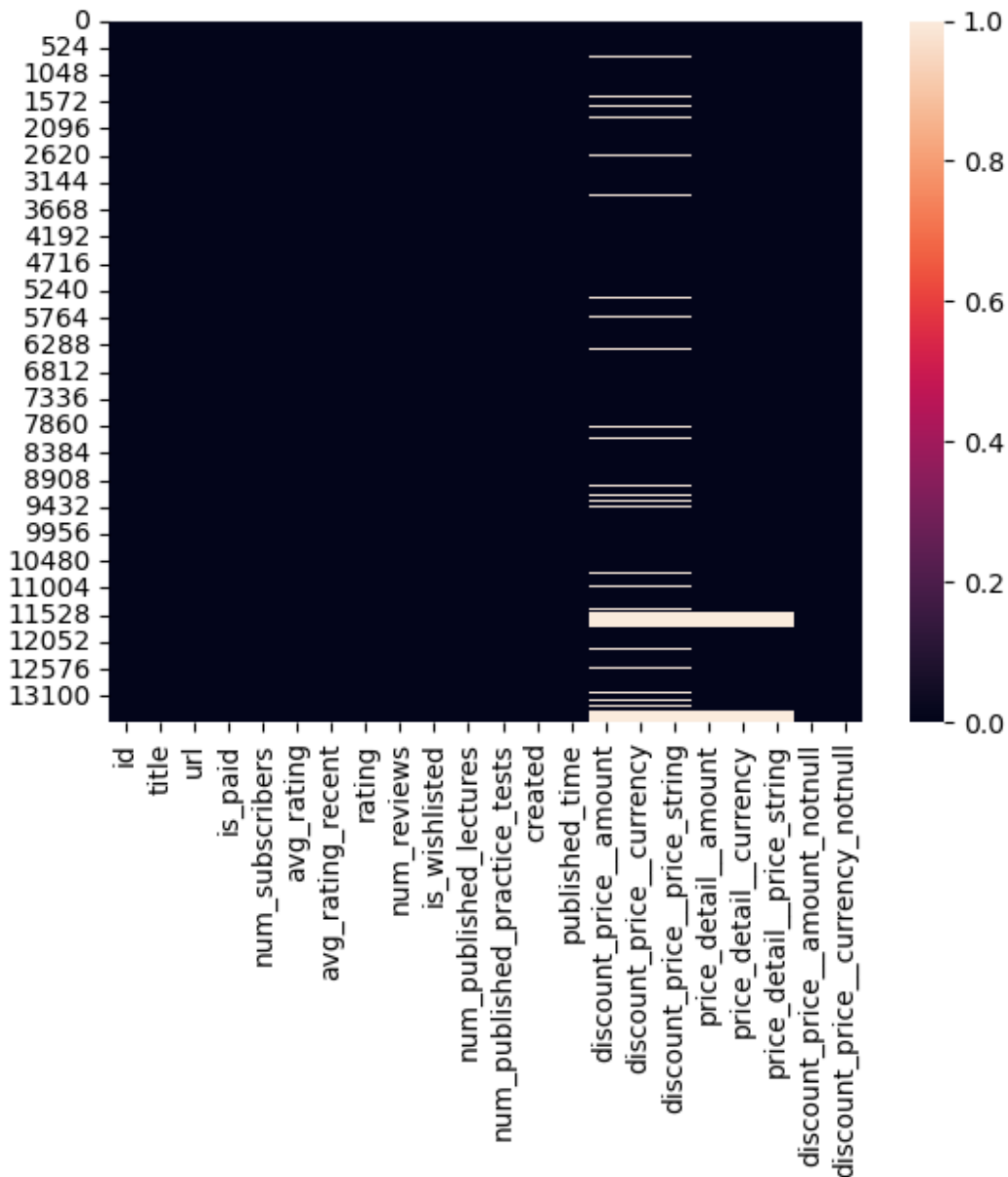
Now you see only INR data is there in COUNTPLOT

```
sns.countplot(x = "discount_price__currency_notnull",data = br)
```

```
<Axes: xlabel='discount_price__currency_notnull', ylabel='count'>
```
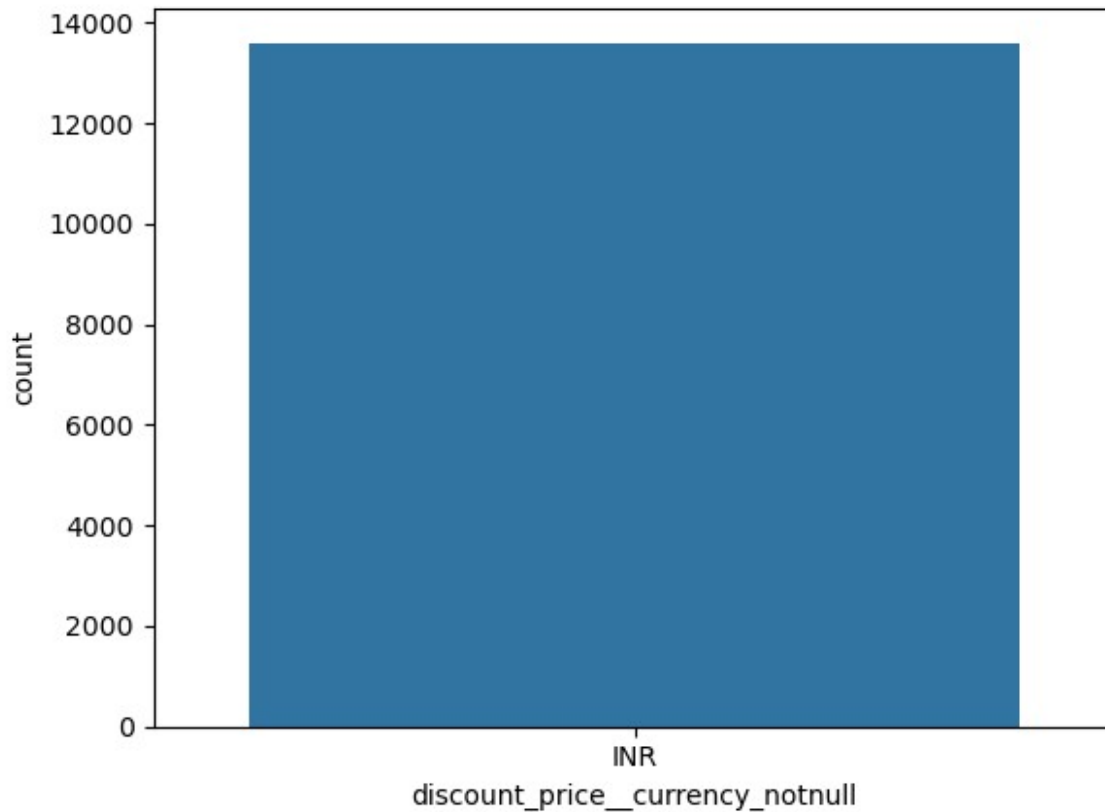
```
br.groupby('discount_price__currency_notnull').size()

discount_price__currency_notnull
INR     13608
dtype: int64
```

There is four null value column still left in udemy finanace accounting dataset

```
br.isnull().sum()
```

```
id                          0
title                       0
url                         0
is_paid                     0
num_subscribers             0
avg_rating                  0
avg_rating_recent           0
rating                      0
num_reviews                 0
is_wishlisted               0
num_published_lectures      0
```

```
num_published_practice_tests          0
created                                0
published_time                         0
discount_price__amount              1403
discount_price__currency            1403
discount_price__price_string        1403
price_detail__amount                 497
price_detail__currency               497
price_detail__price_string           497
discount_price__amount_notnull         0
discount_price__currency_notnull       0
dtype: int64
```

3rd.discount_price__price_stringColumn (Where the null values count is 1403)

here we can see string values are present in this discount_price__price_string column

```
br["discount_price__price_string"].info

<bound method Series.info of 0            ₹455
1           ₹455
2           ₹455
3           ₹455
4           ₹455
            ...
13603        NaN
13604        NaN
13605        NaN
13606        NaN
13607        NaN
Name: discount_price__price_string, Length: 13608, dtype: object>

type(br["discount_price__price_string"])

pandas.core.series.Series
```

1403 null values inside this discount_price__price_string column

```
br["discount_price__price_string"].isnull().sum()

1403
```

Here we see the percentage of null values in the discount_price__price_string column

There are 10 % null values inside the discount_price__price_string column against overall null values.

```
br.isnull().mean()

id                                  0.000000
title                               0.000000
url                                 0.000000
is_paid                             0.000000
num_subscribers                     0.000000
avg_rating                          0.000000
avg_rating_recent                   0.000000
rating                              0.000000
num_reviews                         0.000000
is_wishlisted                       0.000000
num_published_lectures              0.000000
num_published_practice_tests        0.000000
created                             0.000000
published_time                      0.000000
discount_price__amount              0.103101
discount_price__currency            0.103101
discount_price__price_string        0.103101
price_detail__amount                0.036523
price_detail__currency              0.036523
price_detail__price_string          0.036523
discount_price__amount_notnull      0.000000
discount_price__currency_notnull    0.000000
dtype: float64

br["discount_price__price_string"].unique()
```

```
array(['₹455', '₹500', nan, '₹540', '₹546', '₹533', '₹481', '₹488',
       '₹474', '₹462', '₹2,880', '₹494', '₹1,600', '₹800', '₹2,080',
       '₹520', '₹468', '₹507', '₹700', '₹514', '₹552', '₹640',
'₹3,200',
       '₹960', '₹2,720', '₹1,920', '₹1,680', '₹611', '₹2,400', '₹720',
       '₹1,280', '₹2,000', '₹2,240', '₹3,040', '₹2,160', '₹1,360',
       '₹1,760', '₹2,480', '₹1,120', '₹760', '₹1,520', '₹1,440',
'₹480',
       '₹2,320', '₹1,200', '₹2,560', '₹1,840', '₹2,800', '₹2,960',
'₹560',
       '₹2,640', '₹1,040', '₹880', '₹3,120'], dtype=object)
```

Now we have to check the data type

```
type(br["discount_price__price_string"][0])#we are not able to get any
relevant information from this value

str
```

In this datset there are Alphanumerical values, so we are going use fillna to fill the null values with the mode.

```
br["discount_price__price_string"].isnull().sum()

1403
```

Finding the Mode

```
br["discount_price__price_string"].mode().values[0]

'₹455'
```

Replacing the Null values with mode of this column i.e.₹455

```
br["discount_price__price_string_notnull"] =
br["discount_price__price_string"].fillna(455)# jo bhi column ka
median ka wo null value me fill hojaega
br
```

```
              id                                              title  \
0         762616  The Complete SQL Bootcamp 2020: Go from Zero t...
1         937678  Tableau 2020 A-Z: Hands-On Tableau Training fo...
2        1361790                   PMP Exam Prep Seminar -  PMBOK Guide 6
3         648826               The Complete Financial Analyst Course 2020
4         637930  An Entire MBA in 1 Course:Award Winning Busine...
...          ...                                                ...
13603    3005408                           מושגים בסיסיים באופציות חלק ב
13604    3467284                               Programa de Integridade
13605    3368222  Goodwill and Method of Calculating Goodwill (A...
13606    3211345                                     Poderoso Investidor
13607    2935720                Acabou a Previdência e agora? -  Volume 03

                                                     url  is_paid  \
0                        /course/the-complete-sql-bootcamp/     True
1                                     /course/tableau10/     True
2                              /course/pmp-pmbok6-35-pdus/     True
3                /course/the-complete-financial-analyst-course/     True
4            /course/an-entire-mba-in-1-courseaward-winning...     True
...                                                  ...      ...
13603                              /course/1-ndgzhu/     False
13604              /course/iaudit_programa_de_integridade/     False
13605  /course/goodwill-and-method-of-calculating-goo...     False
13606                      /course/poderoso-investidor/     False
13607    /course/acabou-a-previdencia-e-agora-volume-03/     True

       num_subscribers  avg_rating  avg_rating_recent   rating
num_reviews  \
0               295509     4.66019            4.67874  4.67874
78006
1               209070     4.58956            4.60015  4.60015
54581
2               155282     4.59491            4.59326  4.59326
52653
3               245860     4.54407            4.53772  4.53772
46447
4               374836     4.47080            4.47173  4.47173
41630
...                ...         ...                ...      ...
...
13603               47     0.00000            0.00000  0.00000
0
13604               19     0.00000            0.00000  0.00000
0
13605               47     0.00000            0.00000  0.00000
0
13606               48     0.00000            0.00000  0.00000
0
13607                0     0.00000            0.00000  0.00000
0
```

```
       is_wishlisted  ...         published_time  discount_price__amount  \
0              False  ...   2016-04-06T05:16:11Z  455.0
1              False  ...   2016-08-23T16:59:49Z  455.0
2              False  ...   2017-11-14T23:58:14Z  455.0
3              False  ...   2016-01-21T01:38:48Z  455.0
4              False  ...   2016-01-11T21:39:33Z  455.0
...              ...  ...                    ...                     ...
13603          False  ...   2020-04-23T02:06:55Z  NaN
13604          False  ...   2020-09-03T15:26:47Z  NaN
13605          False  ...   2020-07-31T12:14:40Z  NaN
13606          False  ...   2020-07-15T10:35:38Z  NaN
13607          False  ...   2020-04-02T16:33:32Z  NaN

       discount_price__currency discount_price__price_string  \
0                           INR                         ₹455
1                           INR                         ₹455
2                           INR                         ₹455
3                           INR                         ₹455
4                           INR                         ₹455
...                         ...                          ...
13603                       NaN                          NaN
13604                       NaN                          NaN
13605                       NaN                          NaN
13606                       NaN                          NaN
13607                       NaN                          NaN

       price_detail__amount price_detail__currency price_detail__price_string  \
0                    8640.0                    INR                     ₹8,640
1                    8640.0                    INR                     ₹8,640
2                    8640.0                    INR                     ₹8,640
3                    8640.0                    INR                     ₹8,640
4                    8640.0                    INR
```

```
₹8,640
...                              ...                              ...
...
13603                      NaN                      NaN
NaN
13604                      NaN                      NaN
NaN
13605                      NaN                      NaN
NaN
13606                      NaN                      NaN
NaN
13607                      NaN                      NaN
NaN

       discount_price__amount_notnull discount_price__currency_notnull
\
0                                455.0                             INR

1                                455.0                             INR

2                                455.0                             INR

3                                455.0                             INR

4                                455.0                             INR

...                                ...                             ...

13603                            455.0                             INR

13604                            455.0                             INR

13605                            455.0                             INR

13606                            455.0                             INR

13607                            455.0                             INR


       discount_price__price_string_notnull
0                                       ₹455
1                                       ₹455
2                                       ₹455
3                                       ₹455
4                                       ₹455
...                                      ...
13603                                    455
13604                                    455
13605                                    455
13606                                    455
13607                                    455
```

```
[13608 rows x 23 columns]

br["discount_price__price_string_notnull"].fillna(br["discount_price__
price_string"].mode(), inplace=True)# jo bhi column ka median ka wo
null value me fill hojaega

br["discount_price__price_string_notnull"].isnull().sum()

0
```

Now there are two type of elements are present inside discount_price__price_string column, String conceded i.e. (₹ and ,) and 2nd is Numeric value i.e. 455

```
br["discount_price__price_string"].info

<bound method Series.info of 0          ₹455
1          ₹455
2          ₹455
3          ₹455
4          ₹455
           ...
13603      NaN
13604      NaN
13605      NaN
13606      NaN
13607      NaN
Name: discount_price__price_string, Length: 13608, dtype: object>
```

This is not a proper way of Data cleaning we should maintain same type elements in particular column

so we are going to remove ₹ and , from elements from discount_price__price_string

```
br['discount_price__price_string_notnull'] =
br['discount_price__price_string'].str.replace(r'\D', '')

C:\Users\Subham Ranjan\AppData\Local\Temp\
ipykernel_8544\3040057425.py:1: FutureWarning: The default value of
regex will change from True to False in a future version.
  br['discount_price__price_string_notnull'] =
br['discount_price__price_string'].str.replace(r'\D', '')

br['discount_price__price_string_notnull'].info

<bound method Series.info of 0           455
1           455
2           455
3           455
4           455
         ...
13603     NaN
13604     NaN
13605     NaN
13606     NaN
13607     NaN
Name: discount_price__price_string_notnull, Length: 13608, dtype:
object>
```

now we are going to fill 0 where null value beacuse we are not able to covert float to int directly because of null value not convert into int

```
br['discount_price__price_string_notnull'] =
br['discount_price__price_string_notnull'].fillna(0)
```

now we can convert float into int

```
br['discount_price__price_string_notnull'] =
br['discount_price__price_string_notnull'].astype(int)

br['discount_price__price_string_notnull'].info

<bound method Series.info of 0          455
1          455
2          455
3          455
4          455
           ...
13603        0
13604        0
13605        0
13606        0
13607        0
Name: discount_price__price_string_notnull, Length: 13608, dtype:
int32>
```

now we can see same type of data inside the discount_price__currency column

```
br["discount_price__price_string_notnull"].info

<bound method Series.info of 0          455
1          455
2          455
3          455
4          455
           ...
13603        0
13604        0
13605        0
13606        0
13607        0
Name: discount_price__price_string_notnull, Length: 13608, dtype:
int32>

type(br["discount_price__price_string_notnull"])

pandas.core.series.Series

br["discount_price__price_string_notnull"].mode()

0     455
Name: discount_price__price_string_notnull, dtype: int32
```

now we are going to fill null value with mode

```python
br["discount_price__price_string_notnull"] =
br["discount_price__price_string_notnull"].fillna(br["discount_price__
price_string_notnull"].mode())

br["discount_price__price_string_notnull"].isnull().sum()

0


br.isnull().sum()
```

```
id                                         0
title                                      0
url                                        0
is_paid                                    0
num_subscribers                            0
avg_rating                                 0
avg_rating_recent                          0
rating                                     0
num_reviews                                0
is_wishlisted                              0
num_published_lectures                     0
num_published_practice_tests               0
created                                    0
published_time                             0
discount_price__amount                   1403
discount_price__currency                 1403
discount_price__price_string             1403
price_detail__amount                      497
price_detail__currency                    497
price_detail__price_string                497
discount_price__amount_notnull             0
discount_price__currency_notnull           0
discount_price__price_string_notnull       0
dtype: int64
```

4 column ("price_detail__amount")

```python
br["price_detail__amount_notnull"] =br["price_detail__amount"].bfill()
```

we tried backward filling but 215 null value still left

```python
br["price_detail__amount_notnull"].isnull().sum()
```

so we are going to fill remaining null values with front fill

```
br["price_detail__amount_notnull"] =br["price_detail__amount"].ffill()
br.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 13608 entries, 0 to 13607
Data columns (total 24 columns):
 #   Column                                 Non-Null Count  Dtype
---  ------                                 --------------  -----
 0   id                                     13608 non-null  int64
 1   title                                  13608 non-null  object
 2   url                                    13608 non-null  object
 3   is_paid                                13608 non-null  bool
 4   num_subscribers                        13608 non-null  int64
 5   avg_rating                             13608 non-null  float64
 6   avg_rating_recent                      13608 non-null  float64
 7   rating                                 13608 non-null  float64
 8   num_reviews                            13608 non-null  int64
 9   is_wishlisted                          13608 non-null  bool
 10  num_published_lectures                 13608 non-null  int64
 11  num_published_practice_tests           13608 non-null  int64
 12  created                                13608 non-null  object
 13  published_time                         13608 non-null  object
 14  discount_price__amount                 12205 non-null  float64
 15  discount_price__currency               12205 non-null  object
 16  discount_price__price_string           12205 non-null  object
 17  price_detail__amount                   13111 non-null  float64
 18  price_detail__currency                 13111 non-null  object
 19  price_detail__price_string             13111 non-null  object
 20  discount_price__amount_notnull         13608 non-null  float64
 21  discount_price__currency_notnull       13608 non-null  object
 22  discount_price__price_string_notnull   13608 non-null  int32
 23  price_detail__amount_notnull           13608 non-null  float64
dtypes: bool(2), float64(7), int32(1), int64(5), object(9)
memory usage: 2.3+ MB
```

now price_detail__amount_notnull has 0 null value

```
br["price_detail__amount_notnull"].isnull().sum()
```

```
0

br.isnull().sum()

id                                     0
title                                  0
url                                    0
is_paid                                0
num_subscribers                        0
avg_rating                             0
avg_rating_recent                      0
rating                                 0
num_reviews                            0
is_wishlisted                          0
num_published_lectures                 0
num_published_practice_tests           0
created                                0
published_time                         0
discount_price__amount              1403
discount_price__currency            1403
discount_price__price_string        1403
price_detail__amount                 497
price_detail__currency               497
price_detail__price_string           497
discount_price__amount_notnull         0
discount_price__currency_notnull       0
discount_price__price_string_notnull   0
price_detail__amount_notnull           0
dtype: int64

br["price_detail__amount_notnull"].unique()

array([ 8640.,  5760.,  1280.,  9280.,  2560.,  4800.,  6720.,  2240.,
        1920., 10240.,  3200.,  8960.,  6400.,  7040.,  9600., 12800.,
        7680.,  2880.,  4480.,  5440.,  1600.,  6080.,  3520.,  8320.,
        7360.,  4160.,  8000.,  3840.,  5120., 12160., 11520.,  9920.,
       11840., 12480., 10560., 11200., 10880.])
```

5 price_detail__currency column

```
br["price_detail__currency"].unique()

array(['INR', nan], dtype=object)
```

There are 497 null value in price_detail__currency column

```
br["price_detail__currency"].isnull().sum()

497

br["price_detail__currency_notnull"] =
br["price_detail__currency"].fillna("INR")

br["price_detail__currency_notnull"].isnull().sum()

0
```

After handling with null values there is nan name element in this column. So we are going to convert this into INR

Now we are going Change nan to INR. it will be relevent type of way to fill null value in this particular column

```
br["price_detail__currency_notnull"] =
br["price_detail__currency_notnull"] .replace(to_replace="nan",
value="INR")

br["price_detail__currency_notnull"].unique()

array(['INR'], dtype=object)

br.groupby('price_detail__currency').size()

price_detail__currency
INR     13111
dtype: int64

br.groupby('price_detail__currency_notnull').size()

price_detail__currency_notnull
INR     13608
dtype: int64
```

previous size = 13111

Present SIZE 0F INR = 13608

```
sns.countplot(x = "price_detail__currency_notnull",data  = br)

<Axes: xlabel='price_detail__currency_notnull', ylabel='count'>
```
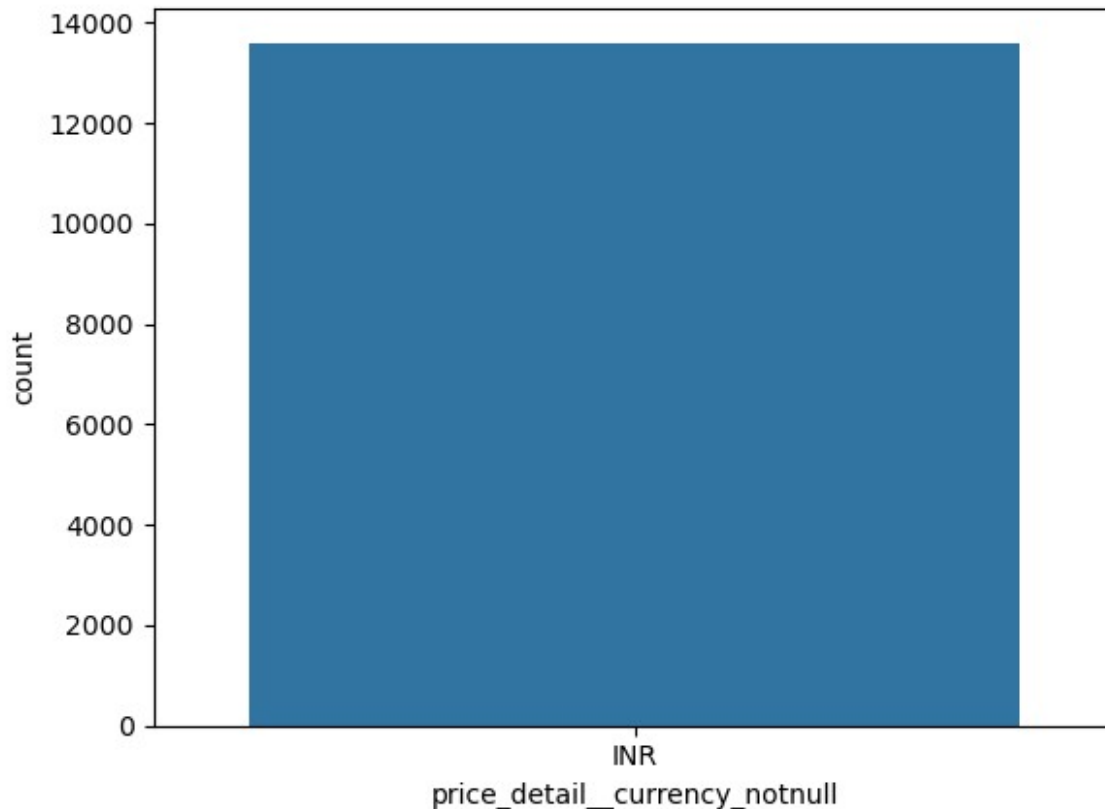

There are 497 null value in price_detail__currency column

```
br

          id                                               title  \
0     762616   The Complete SQL Bootcamp 2020: Go from Zero t...
1     937678   Tableau 2020 A-Z: Hands-On Tableau Training fo...
2    1361790             PMP Exam Prep Seminar -   PMBOK Guide 6
```

```
3         648826            The Complete Financial Analyst Course 2020
4         637930    An Entire MBA in 1 Course:Award Winning Busine...
...          ...                                                   ...
13603    3005408                              מושגים בסיסיים באופציות חלק ב
13604    3467284                                   Programa de Integridade
13605    3368222    Goodwill and Method of Calculating Goodwill (A...
13606    3211345                                       Poderoso Investidor
13607    2935720               Acabou a Previdência e agora? -  Volume 03

                                                     url  is_paid  \
0                     /course/the-complete-sql-bootcamp/     True
1                                     /course/tableau10/     True
2                               /course/pmp-pmbok6-35-pdus/   True
3            /course/the-complete-financial-analyst-course/   True
4        /course/an-entire-mba-in-1-courseaward-winning...   True
...                                                    ...      ...
13603                                   /course/1-ndgzhu/    False
13604              /course/iaudit_programa_de_integridade/   False
13605    /course/goodwill-and-method-of-calculating-goo...   False
13606                        /course/poderoso-investidor/   False
13607       /course/acabou-a-previdencia-e-agora-volume-03/   True

        num_subscribers  avg_rating  avg_rating_recent   rating
num_reviews  \
0                295509     4.66019            4.67874  4.67874
78006
1                209070     4.58956            4.60015  4.60015
54581
2                155282     4.59491            4.59326  4.59326
52653
3                245860     4.54407            4.53772  4.53772
46447
4                374836     4.47080            4.47173  4.47173
41630
...                 ...         ...                ...      ...
...
13603                47     0.00000            0.00000  0.00000
0
13604                19     0.00000            0.00000  0.00000
0
13605                47     0.00000            0.00000  0.00000
0
13606                48     0.00000            0.00000  0.00000
0
13607                 0     0.00000            0.00000  0.00000
0

        is_wishlisted  ...  discount_price__currency  \
0               False  ...                       INR
```

```
1                False  ...                              INR
2                False  ...                              INR
3                False  ...                              INR
4                False  ...                              INR
...                ...  ...                              ...
13603            False  ...                              NaN
13604            False  ...                              NaN
13605            False  ...                              NaN
13606            False  ...                              NaN
13607            False  ...                              NaN

       discount_price__price_string price_detail__amount  \
0                              ₹455               8640.0
1                              ₹455               8640.0
2                              ₹455               8640.0
3                              ₹455               8640.0
4                              ₹455               8640.0
...                             ...                  ...
13603                           NaN                  NaN
13604                           NaN                  NaN
13605                           NaN                  NaN
13606                           NaN                  NaN
13607                           NaN                  NaN

       price_detail__currency price_detail__price_string  \
0                         INR                     ₹8,640
1                         INR                     ₹8,640
2                         INR                     ₹8,640
3                         INR                     ₹8,640
4                         INR                     ₹8,640
...                       ...                        ...
13603                     NaN                        NaN
13604                     NaN                        NaN
13605                     NaN                        NaN
13606                     NaN                        NaN
13607                     NaN                        NaN

       discount_price__amount_notnull discount_price__currency_notnull
\
0                               455.0                              INR

1                               455.0                              INR

2                               455.0                              INR

3                               455.0                              INR

4                               455.0                              INR

...                               ...                              ...
```

13603	455.0	INR
13604	455.0	INR
13605	455.0	INR
13606	455.0	INR
13607	455.0	INR

```
       discount_price__price_string_notnull
price_detail__amount_notnull  \
0                                           455
8640.0
1                                           455
8640.0
2                                           455
8640.0
3                                           455
8640.0
4                                           455
8640.0
...                                         ...                                    .
..
13603                                         0
1280.0
13604                                         0
1280.0
13605                                         0
1280.0
13606                                         0
1280.0
13607                                         0
1280.0

       price_detail__currency_notnull
0                                 INR
1                                 INR
2                                 INR
3                                 INR
4                                 INR
...                               ...
13603                             INR
13604                             INR
13605                             INR
13606                             INR
13607                             INR
```

```
[13608 rows x 25 columns]

br["price_detail__currency_notnull"].isnull().sum()

0
```

Now there is 0 null value is there inside the price_detail__currency column now

6 price_detail__price_string column

```
br["price_detail__price_string"].isnull().sum()

497
```

This is the string data type inside the price_detail__price_string

```
br["price_detail__price_string_notnull"] =
br["price_detail__price_string"].ffill()
br.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 13608 entries, 0 to 13607
Data columns (total 26 columns):
 #   Column                        Non-Null Count  Dtype
---  ------                        --------------  -----
 0   id                            13608 non-null  int64
 1   title                         13608 non-null  object
 2   url                           13608 non-null  object
 3   is_paid                       13608 non-null  bool
 4   num_subscribers               13608 non-null  int64
 5   avg_rating                    13608 non-null  float64
 6   avg_rating_recent             13608 non-null  float64
 7   rating                        13608 non-null  float64
 8   num_reviews                   13608 non-null  int64
 9   is_wishlisted                 13608 non-null  bool
 10  num_published_lectures        13608 non-null  int64
 11  num_published_practice_tests  13608 non-null  int64
 12  created                       13608 non-null  object
 13  published_time                13608 non-null  object
 14  discount_price__amount        12205 non-null  float64
```

```
 15   discount_price__currency                12205 non-null   object
 16   discount_price__price_string            12205 non-null   object
 17   price_detail__amount                    13111 non-null   float64
 18   price_detail__currency                  13111 non-null   object
 19   price_detail__price_string              13111 non-null   object
 20   discount_price__amount_notnull          13608 non-null   float64
 21   discount_price__currency_notnull        13608 non-null   object
 22   discount_price__price_string_notnull    13608 non-null   int32
 23   price_detail__amount_notnull            13608 non-null   float64
 24   price_detail__currency_notnull          13608 non-null   object
 25   price_detail__price_string_notnull      13608 non-null   object
dtypes: bool(2), float64(7), int32(1), int64(5), object(11)
memory usage: 2.5+ MB
```

There is 0 null value inside the price_detail__price_string_notnull column

```
br["price_detail__price_string_notnull"].isnull().sum()

0

br.isnull().sum()
```

```
id                                        0
title                                     0
url                                       0
is_paid                                   0
num_subscribers                           0
avg_rating                                0
avg_rating_recent                         0
rating                                    0
num_reviews                               0
is_wishlisted                             0
num_published_lectures                    0
num_published_practice_tests             0
created                                   0
published_time                            0
discount_price__amount                 1403
discount_price__currency               1403
discount_price__price_string           1403
price_detail__amount                    497
price_detail__currency                  497
price_detail__price_string              497
discount_price__amount_notnull            0
discount_price__currency_notnull          0
discount_price__price_string_notnull      0
price_detail__amount_notnull              0
```

```
price_detail__currency_notnull              0
price_detail__price_string_notnull          0
dtype: int64

br

           id                                              title  \
0      762616  The Complete SQL Bootcamp 2020: Go from Zero t...
1      937678  Tableau 2020 A-Z: Hands-On Tableau Training fo...
2     1361790              PMP Exam Prep Seminar -  PMBOK Guide 6
3      648826          The Complete Financial Analyst Course 2020
4      637930  An Entire MBA in 1 Course:Award Winning Busine...
...       ...                                                ...
13603  3005408                    מושגים בסיסיים באופציות חלק ב
13604  3467284                              Programa de Integridade
13605  3368222  Goodwill and Method of Calculating Goodwill (A...
13606  3211345                                  Poderoso Investidor
13607  2935720         Acabou a Previdência e agora? -  Volume 03

                                                   url  is_paid  \
0                     /course/the-complete-sql-bootcamp/     True
1                                    /course/tableau10/     True
2                              /course/pmp-pmbok6-35-pdus/     True
3         /course/the-complete-financial-analyst-course/     True
4      /course/an-entire-mba-in-1-courseaward-winning...     True
...                                                  ...      ...
13603                               /course/1-ndgzhu/    False
13604            /course/iaudit_programa_de_integridade/    False
13605  /course/goodwill-and-method-of-calculating-goo...    False
13606                      /course/poderoso-investidor/    False
13607    /course/acabou-a-previdencia-e-agora-volume-03/     True

       num_subscribers  avg_rating  avg_rating_recent   rating
num_reviews  \
0              295509     4.66019            4.67874  4.67874
78006
1              209070     4.58956            4.60015  4.60015
54581
2              155282     4.59491            4.59326  4.59326
52653
3              245860     4.54407            4.53772  4.53772
46447
4              374836     4.47080            4.47173  4.47173
41630
...               ...         ...                ...      ...
...
13603              47     0.00000            0.00000  0.00000
0
13604              19     0.00000            0.00000  0.00000
0
```

```
13605                  47       0.00000          0.00000  0.00000
0
13606                  48       0.00000          0.00000  0.00000
0
13607                   0       0.00000          0.00000  0.00000
0

       is_wishlisted  ...  discount_price__price_string
price_detail__amount  \
0              False  ...                           ₹455
8640.0
1              False  ...                           ₹455
8640.0
2              False  ...                           ₹455
8640.0
3              False  ...                           ₹455
8640.0
4              False  ...                           ₹455
8640.0
...              ...  ...                            ...
...
13603          False  ...                            NaN
NaN
13604          False  ...                            NaN
NaN
13605          False  ...                            NaN
NaN
13606          False  ...                            NaN
NaN
13607          False  ...                            NaN
NaN

       price_detail__currency price_detail__price_string  \
0                         INR                     ₹8,640
1                         INR                     ₹8,640
2                         INR                     ₹8,640
3                         INR                     ₹8,640
4                         INR                     ₹8,640
...                       ...                        ...
13603                     NaN                        NaN
13604                     NaN                        NaN
13605                     NaN                        NaN
13606                     NaN                        NaN
13607                     NaN                        NaN

       discount_price__amount_notnull discount_price__currency_notnull
\
0                               455.0                              INR

1                               455.0                              INR
```

2	455.0	INR
3	455.0	INR
4	455.0	INR
...	...	...
13603	455.0	INR
13604	455.0	INR
13605	455.0	INR
13606	455.0	INR
13607	455.0	INR

```
       discount_price__price_string_notnull
price_detail__amount_notnull  \
0                                          455
8640.0
1                                          455
8640.0
2                                          455
8640.0
3                                          455
8640.0
4                                          455
8640.0
...                                        ...                            .
..
13603                                        0
1280.0
13604                                        0
1280.0
13605                                        0
1280.0
13606                                        0
1280.0
13607                                        0
1280.0

       price_detail__currency_notnull
price_detail__price_string_notnull
0                              INR
₹8,640
1                              INR
```

```
₹8,640
2                                    INR
₹8,640
3                                    INR
₹8,640
4                                    INR
₹8,640
...                                  ...                                    ..
.
13603                                INR
₹1,280
13604                                INR
₹1,280
13605                                INR
₹1,280
13606                                INR
₹1,280
13607                                INR
₹1,280

[13608 rows x 26 columns]
```

INT572

```
br

             id                                              title  \
0        762616   The Complete SQL Bootcamp 2020: Go from Zero t...
1        937678   Tableau 2020 A-Z: Hands-On Tableau Training fo...
2       1361790            PMP Exam Prep Seminar -  PMBOK Guide 6
3        648826        The Complete Financial Analyst Course 2020
4        637930   An Entire MBA in 1 Course:Award Winning Busine...
...         ...                                                 ...
13603   3005408                        מושגים בסיסיים באופציות חלק ב
13604   3467284                             Programa de Integridade
13605   3368222   Goodwill and Method of Calculating Goodwill (A...
13606   3211345                                  Poderoso Investidor
13607   2935720            Acabou a Previdência e agora? -  Volume 03


                                                 url  is_paid  \
0                    /course/the-complete-sql-bootcamp/     True
1                                  /course/tableau10/     True
2                           /course/pmp-pmbok6-35-pdus/     True
3         /course/the-complete-financial-analyst-course/     True
4      /course/an-entire-mba-in-1-courseaward-winning...     True
...                                                ...      ...
13603                                 /course/1-ndgzhu/    False
13604           /course/iaudit_programa_de_integridade/    False
```

```
13605   /course/goodwill-and-method-of-calculating-goo...      False
13606                        /course/poderoso-investidor/      False
13607     /course/acabou-a-previdencia-e-agora-volume-03/       True

       num_subscribers  avg_rating  avg_rating_recent    rating
num_reviews  \
0               295509     4.66019            4.67874   4.67874
78006
1               209070     4.58956            4.60015   4.60015
54581
2               155282     4.59491            4.59326   4.59326
52653
3               245860     4.54407            4.53772   4.53772
46447
4               374836     4.47080            4.47173   4.47173
41630
...                ...         ...                ...       ...
...
13603               47     0.00000            0.00000   0.00000
0
13604               19     0.00000            0.00000   0.00000
0
13605               47     0.00000            0.00000   0.00000
0
13606               48     0.00000            0.00000   0.00000
0
13607                0     0.00000            0.00000   0.00000
0

       is_wishlisted  ...  discount_price__price_string
price_detail__amount  \
0              False  ...                          ₹455
8640.0
1              False  ...                          ₹455
8640.0
2              False  ...                          ₹455
8640.0
3              False  ...                          ₹455
8640.0
4              False  ...                          ₹455
8640.0
...              ...  ...                           ...
...
13603          False  ...                           NaN
NaN
13604          False  ...                           NaN
NaN
13605          False  ...                           NaN
NaN
```

```
13606          False  ...                          NaN
NaN
13607          False  ...                          NaN
NaN

      price_detail__currency price_detail__price_string  \
0                        INR                     ₹8,640
1                        INR                     ₹8,640
2                        INR                     ₹8,640
3                        INR                     ₹8,640
4                        INR                     ₹8,640
...                      ...                        ...
13603                    NaN                        NaN
13604                    NaN                        NaN
13605                    NaN                        NaN
13606                    NaN                        NaN
13607                    NaN                        NaN

       discount_price__amount_notnull discount_price__currency_notnull  \
0                               455.0                              INR

1                               455.0                              INR

2                               455.0                              INR

3                               455.0                              INR

4                               455.0                              INR

...                               ...                              ...

13603                           455.0                              INR

13604                           455.0                              INR

13605                           455.0                              INR

13606                           455.0                              INR

13607                           455.0                              INR


      discount_price__price_string_notnull
price_detail__amount_notnull  \
0                                      455
8640.0
1                                      455
8640.0
2                                      455
8640.0
```

```
3                                              455
8640.0
4                                              455
8640.0
...                                            ...                            .
..
13603                                            0
1280.0
13604                                            0
1280.0
13605                                            0
1280.0
13606                                            0
1280.0
13607                                            0
1280.0

       price_detail__currency_notnull
price_detail__price_string_notnull
0                             INR
₹8,640
1                             INR
₹8,640
2                             INR
₹8,640
3                             INR
₹8,640
4                             INR
₹8,640
...                           ...                             ..
.
13603                         INR
₹1,280
13604                         INR
₹1,280
13605                         INR
₹1,280
13606                         INR
₹1,280
13607                         INR
₹1,280

[13608 rows x 26 columns]
```

Q1

How many paid and unpaid courses are there in udemy finance data?

Here i make specific column and mention is_paid true = 1,else 0

```
br['is_paid_plot'] = np.where(br['is_paid']=="True",1,0)
br
```

	id	title \
0	762616	The Complete SQL Bootcamp 2020: Go from Zero t...
1	937678	Tableau 2020 A-Z: Hands-On Tableau Training fo...
2	1361790	PMP Exam Prep Seminar - PMBOK Guide 6
3	648826	The Complete Financial Analyst Course 2020
4	637930	An Entire MBA in 1 Course:Award Winning Busine...
...	...	...
13603	3005408	מושגים בסיסיים באופציות חלק ב
13604	3467284	Programa de Integridade
13605	3368222	Goodwill and Method of Calculating Goodwill (A...
13606	3211345	Poderoso Investidor
13607	2935720	Acabou a Previdência e agora? - Volume 03

	url	is_paid \
0	/course/the-complete-sql-bootcamp/	True
1	/course/tableau10/	True
2	/course/pmp-pmbok6-35-pdus/	True
3	/course/the-complete-financial-analyst-course/	True
4	/course/an-entire-mba-in-1-courseaward-winning...	True
...	...	...
13603	/course/1-ndgzhu/	False
13604	/course/iaudit_programa_de_integridade/	False
13605	/course/goodwill-and-method-of-calculating-goo...	False
13606	/course/poderoso-investidor/	False
13607	/course/acabou-a-previdencia-e-agora-volume-03/	True

	num_subscribers	avg_rating	avg_rating_recent	rating num_reviews \
0	295509	4.66019	4.67874	4.67874 78006
1	209070	4.58956	4.60015	4.60015 54581
2	155282	4.59491	4.59326	4.59326

```
52653
3              245860    4.54407          4.53772  4.53772
46447
4              374836    4.47080          4.47173  4.47173
41630
...                ...        ...              ...      ...
...
13603              47    0.00000          0.00000  0.00000
0
13604              19    0.00000          0.00000  0.00000
0
13605              47    0.00000          0.00000  0.00000
0
13606              48    0.00000          0.00000  0.00000
0
13607               0    0.00000          0.00000  0.00000
0

       is_wishlisted  ...  price_detail__amount
price_detail__currency  \
0              False  ...                8640.0
INR
1              False  ...                8640.0
INR
2              False  ...                8640.0
INR
3              False  ...                8640.0
INR
4              False  ...                8640.0
INR
...              ...  ...                   ...                                ..
.
13603          False  ...                   NaN
NaN
13604          False  ...                   NaN
NaN
13605          False  ...                   NaN
NaN
13606          False  ...                   NaN
NaN
13607          False  ...                   NaN
NaN

      price_detail__price_string discount_price__amount_notnull  \
0                           ₹8,640                          455.0
1                           ₹8,640                          455.0
2                           ₹8,640                          455.0
3                           ₹8,640                          455.0
4                           ₹8,640                          455.0
```

```
...                       ...                                   ...
13603                     NaN                                 455.0
13604                     NaN                                 455.0
13605                     NaN                                 455.0
13606                     NaN                                 455.0
13607                     NaN                                 455.0

       discount_price__currency_notnull
discount_price__price_string_notnull  \
0                                   INR
455
1                                   INR
455
2                                   INR
455
3                                   INR
455
4                                   INR
455
...                                 ...
...
13603                               INR
0
13604                               INR
0
13605                               INR
0
13606                               INR
0
13607                               INR
0

       price_detail__amount_notnull price_detail__currency_notnull  \
0                           8640.0                              INR
1                           8640.0                              INR
2                           8640.0                              INR
3                           8640.0                              INR
4                           8640.0                              INR
...                            ...                              ...
13603                       1280.0                              INR
13604                       1280.0                              INR
13605                       1280.0                              INR
13606                       1280.0                              INR
13607                       1280.0                              INR

       price_detail__price_string_notnull is_paid_plot
0                                   ₹8,640            0
1                                   ₹8,640            0
2                                   ₹8,640            0
3                                   ₹8,640            0
```

```
4                                        ₹8,640              0
...                                        ...            ...
13603                                    ₹1,280              0
13604                                    ₹1,280              0
13605                                    ₹1,280              0
13606                                    ₹1,280              0
13607                                    ₹1,280              0

[13608 rows x 27 columns]
```

Now i am going to change True = 1 and False = 0

```
br.is_paid = br.is_paid.replace({True: 1, False: 0})

br

            id                                           title  \
0        762616  The Complete SQL Bootcamp 2020: Go from Zero t...
1        937678  Tableau 2020 A-Z: Hands-On Tableau Training fo...
2       1361790          PMP Exam Prep Seminar -  PMBOK Guide 6
3        648826         The Complete Financial Analyst Course 2020
4        637930  An Entire MBA in 1 Course:Award Winning Busine...
...         ...                                             ...
13603   3005408                    מושגים בסיסיים באופציות חלק ב
13604   3467284                           Programa de Integridade
13605   3368222  Goodwill and Method of Calculating Goodwill (A...
13606   3211345                               Poderoso Investidor
13607   2935720         Acabou a Previdência e agora? -  Volume 03

                                                  url  is_paid  \
0                   /course/the-complete-sql-bootcamp/        1
1                                   /course/tableau10/        1
2                             /course/pmp-pmbok6-35-pdus/        1
3          /course/the-complete-financial-analyst-course/        1
4        /course/an-entire-mba-in-1-courseaward-winning...        1
...                                               ...      ...
13603                             /course/1-ndgzhu/        0
13604              /course/iaudit_programa_de_integridade/        0
13605  /course/goodwill-and-method-of-calculating-goo...        0
13606                       /course/poderoso-investidor/        0
13607    /course/acabou-a-previdencia-e-agora-volume-03/        1

        num_subscribers  avg_rating  avg_rating_recent   rating
num_reviews  \
0                295509     4.66019            4.67874  4.67874
78006
```

```
1              209070       4.58956                4.60015   4.60015
54581
2              155282       4.59491                4.59326   4.59326
52653
3              245860       4.54407                4.53772   4.53772
46447
4              374836       4.47080                4.47173   4.47173
41630
...              ...           ...                    ...       ...
...
13603             47       0.00000                0.00000   0.00000
0
13604             19       0.00000                0.00000   0.00000
0
13605             47       0.00000                0.00000   0.00000
0
13606             48       0.00000                0.00000   0.00000
0
13607              0       0.00000                0.00000   0.00000
0

       is_wishlisted  ...  price_detail__amount
price_detail__currency  \
0              False  ...                8640.0
INR
1              False  ...                8640.0
INR
2              False  ...                8640.0
INR
3              False  ...                8640.0
INR
4              False  ...                8640.0
INR
...              ...  ...                   ...                       ..
.
13603          False  ...                   NaN
NaN
13604          False  ...                   NaN
NaN
13605          False  ...                   NaN
NaN
13606          False  ...                   NaN
NaN
13607          False  ...                   NaN
NaN

     price_detail__price_string discount_price__amount_notnull  \
0                         ₹8,640                          455.0
1                         ₹8,640                          455.0
```

```
2                              ₹8,640                              455.0
3                              ₹8,640                              455.0
4                              ₹8,640                              455.0
...                               ...                                ...
13603                             NaN                              455.0
13604                             NaN                              455.0
13605                             NaN                              455.0
13606                             NaN                              455.0
13607                             NaN                              455.0

       discount_price__currency_notnull
discount_price__price_string_notnull  \
0                                   INR
455
1                                   INR
455
2                                   INR
455
3                                   INR
455
4                                   INR
455
...                                 ...
...
13603                               INR
0
13604                               INR
0
13605                               INR
0
13606                               INR
0
13607                               INR
0

       price_detail__amount_notnull price_detail__currency_notnull  \
0                            8640.0                              INR
1                            8640.0                              INR
2                            8640.0                              INR
3                            8640.0                              INR
4                            8640.0                              INR
...                             ...                              ...
13603                        1280.0                              INR
13604                        1280.0                              INR
13605                        1280.0                              INR
13606                        1280.0                              INR
13607                        1280.0                              INR

       price_detail__price_string_notnull is_paid_plot
0                              ₹8,640                 0
```

```
1                                        ₹8,640              0
2                                        ₹8,640              0
3                                        ₹8,640              0
4                                        ₹8,640              0
...                                        ...            ...
13603                                    ₹1,280              0
13604                                    ₹1,280              0
13605                                    ₹1,280              0
13606                                    ₹1,280              0
13607                                    ₹1,280              0

[13608 rows x 27 columns]
```

Now we can see i convert True and False into 1 and 0

```
br["is_paid"].unique()

array([1, 0], dtype=int64)
```

Now you can see no of paid courses and unpaid courses

paid courses = 1

unpaid courses = 0

no of paid courses = 13112
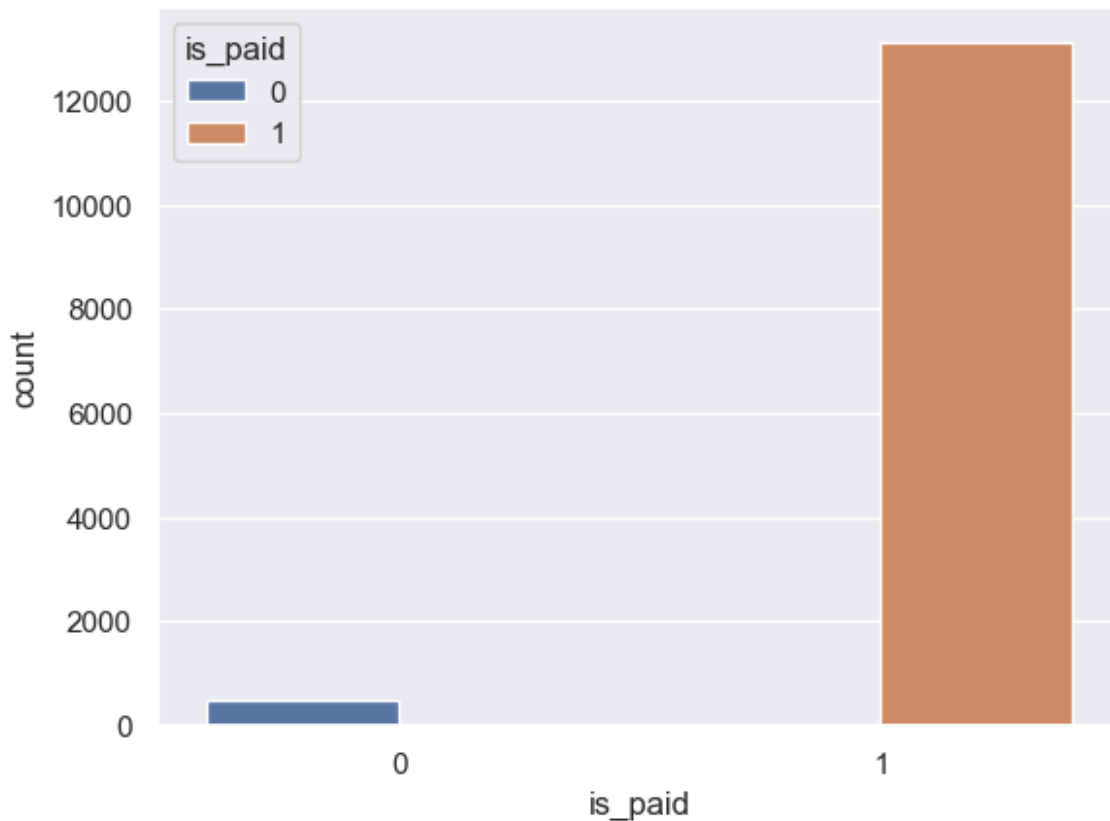
no of unpaid courses = 496

```
br.groupby('is_paid').size()

is_paid
0        496
```

```
1    13112
dtype: int64

sns.set(style = "darkgrid")
sns.countplot(x = "is_paid",data = br,hue = "is_paid")

<Axes: xlabel='is_paid', ylabel='count'>
```


Q NO 2

No of Whislisted courses in udemy finance accounting data?

```
br["is_wishlisted"].unique()
```

```
array([False])
```

Now you can see there is no course are whislisted by the customer

This is finance accounting data, that's why only "False" is mentioned in the wishlist.

Conclusion

So udemy should not include these unnecessary type of column in this finance accounting data which donot have any relevency to this.so it will increase overall size of dataset

```
br["is_wishlisted"].info()

<class 'pandas.core.series.Series'>
RangeIndex: 13608 entries, 0 to 13607
Series name: is_wishlisted
Non-Null Count  Dtype
--------------  -----
13608 non-null  bool
dtypes: bool(1)
memory usage: 13.4 KB
```

So i am going to drop this column

```
br.drop("is_wishlisted",axis = 1)

             id                                            title  \
0        762616  The Complete SQL Bootcamp 2020: Go from Zero t...
1        937678  Tableau 2020 A-Z: Hands-On Tableau Training fo...
2       1361790          PMP Exam Prep Seminar -  PMBOK Guide 6
3        648826        The Complete Financial Analyst Course 2020
4        637930  An Entire MBA in 1 Course:Award Winning Busine...
...         ...                                              ...
13603   3005408                       מושגים בסיסיים באופציות חלק ב
13604   3467284                            Programa de Integridade
```

```
13605  3368222  Goodwill and Method of Calculating Goodwill (A...
13606  3211345                          Poderoso Investidor
13607  2935720       Acabou a Previdência e agora? -  Volume 03

                                                    url  is_paid  \
0                        /course/the-complete-sql-bootcamp/        1
1                                /course/tableau10/        1
2                        /course/pmp-pmbok6-35-pdus/        1
3          /course/the-complete-financial-analyst-course/        1
4      /course/an-entire-mba-in-1-courseaward-winning...        1
...                                              ...      ...
13603                          /course/1-ndgzhu/        0
13604            /course/iaudit_programa_de_integridade/        0
13605  /course/goodwill-and-method-of-calculating-goo...        0
13606                      /course/poderoso-investidor/        0
13607     /course/acabou-a-previdencia-e-agora-volume-03/        1

       num_subscribers  avg_rating  avg_rating_recent   rating
num_reviews  \
0              295509    4.66019              4.67874  4.67874
78006
1              209070    4.58956              4.60015  4.60015
54581
2              155282    4.59491              4.59326  4.59326
52653
3              245860    4.54407              4.53772  4.53772
46447
4              374836    4.47080              4.47173  4.47173
41630
...               ...        ...                  ...      ...
...
13603              47    0.00000              0.00000  0.00000
0
13604              19    0.00000              0.00000  0.00000
0
13605              47    0.00000              0.00000  0.00000
0
13606              48    0.00000              0.00000  0.00000
0
13607               0    0.00000              0.00000  0.00000
0

       num_published_lectures  ...  price_detail__amount  \
0                           84  ...                8640.0
1                           78  ...                8640.0
2                          292  ...                8640.0
3                          338  ...                8640.0
4                           83  ...                8640.0
...                        ...  ...                   ...
13603                        6  ...                   NaN
```

```
13604                   5  ...                        NaN
13605                   5  ...                        NaN
13606                  13  ...                        NaN
13607                  14  ...                        NaN

       price_detail__currency price_detail__price_string  \
0                         INR                      ₹8,640
1                         INR                      ₹8,640
2                         INR                      ₹8,640
3                         INR                      ₹8,640
4                         INR                      ₹8,640
...                       ...                         ...
13603                     NaN                         NaN
13604                     NaN                         NaN
13605                     NaN                         NaN
13606                     NaN                         NaN
13607                     NaN                         NaN

       discount_price__amount_notnull discount_price__currency_notnull
\
0                               455.0                              INR

1                               455.0                              INR

2                               455.0                              INR

3                               455.0                              INR

4                               455.0                              INR

...                               ...                              ...

13603                           455.0                              INR

13604                           455.0                              INR

13605                           455.0                              INR

13606                           455.0                              INR

13607                           455.0                              INR

       discount_price__price_string_notnull
price_detail__amount_notnull  \
0                                        455
8640.0
1                                        455
8640.0
2                                        455
8640.0
```

```
3                                                 455
8640.0
4                                                 455
8640.0
...                                               ...                        .
..
13603                                               0
1280.0
13604                                               0
1280.0
13605                                               0
1280.0
13606                                               0
1280.0
13607                                               0
1280.0

       price_detail__currency_notnull
price_detail__price_string_notnull  \
0                                  INR
₹8,640
1                                  INR
₹8,640
2                                  INR
₹8,640
3                                  INR
₹8,640
4                                  INR
₹8,640
...                                ...                          ..
.
13603                              INR
₹1,280
13604                              INR
₹1,280
13605                              INR
₹1,280
13606                              INR
₹1,280
13607                              INR
₹1,280

       is_paid_plot
0                 0
1                 0
2                 0
3                 0
4                 0
...             ...
```

```
13603                0
13604                0
13605                0
13606                0
13607                0

[13608 rows x 26 columns]
```

Q no 3

How many courses has 5.0 star rating?

```
top_rating = br['rating'].nlargest(n=500).index
top_rating

Int64Index([ 3085,  4341,  5108,  6000,  6084,  6296,  6300,  6821,
7233,
            7323,
            ...
             8248,  7957, 12954,  8102,  9063, 12731, 12071,  8112,
10583,
            3759],
           dtype='int64', length=500)

top_rating_df = br.iloc[top_rating]

top_rating_df
```

```
            id                                    title  \
3085   1538742   Advanced Upwork Interviews: A Simple Way to Ea...
4341   2350744   Price Action: The complete Price Action Master...
5108   1531238   New! Raise funds for your innovative business ...
6000     60463   Shipping address and Shipping cost for E-comme...
6084   3339914   Emotional Intelligence Training for Increased ...
...        ...                                       ...
12731  3234013                   LOS 4 ELEMENTOS DEL IMPUESTO
12071  3207347             Risk Expert - Aprenda na Prática
8112   3141018   Build a passive business empire with Mini Kind...
10583  2805993       Careers in US Capital Market and PE & VC
3759   1855058   Open a Non-Medical Home Care Agency - Home Bas...


                                                   url  is_paid  \
3085                        /course/upwork-interviews/        1
4341                       /course/daytrading-masterclass/     1
5108          /course/eu-grant-funding-innovative-business/   1
6000                       /course/ecommerce-shipping/      1
6084    /course/emotional-intelligence-to-increase-sales/    1
```

```
...                                                         ...      ...
12731                    /course/los-4-elementos-del-impuesto/       1
12071                                    /course/risk-expert/         1
8112     /course/build-a-passive-business-empire-with-m...          1
10583                       /course/careers-in-finance-1/            1
3759          /course/open-a-non-medical-home-care-agency/          1

        num_subscribers  avg_rating  avg_rating_recent    rating
num_reviews  \
3085                508     5.00000            5.00000   5.00000
71
4341                439     5.00000            5.00000   5.00000
38
5108               2233     5.00000            5.00000   5.00000
28
6000                130     5.00000            5.00000   5.00000
19
6084                 38     5.00000            5.00000   5.00000
18
...                 ...         ...                ...       ...
...
12731                22     4.87500            4.86951   4.86951
4
12071               411     4.88889            4.86944   4.86944
55
8112                 59     4.85714            4.86939   4.86939
8
10583                 8     4.87500            4.86813   4.86813
4
3759                245     4.75000            4.86810   4.86810
50

        is_wishlisted  ...  price_detail__amount
price_detail__currency  \
3085            False  ...                8640.0
INR
4341            False  ...                8640.0
INR
5108            False  ...                9600.0
INR
6000            False  ...                1280.0
INR
6084            False  ...                6400.0
INR
...               ...  ...                   ...                    ..
.
12731           False  ...                1280.0
INR
12071           False  ...                8640.0
```

```
INR
8112        False  ...                 8640.0
INR
10583       False  ...                 1280.0
INR
3759        False  ...                12800.0
INR

      price_detail__price_string discount_price__amount_notnull  \
3085                      ₹8,640                           455.0
4341                      ₹8,640                           455.0
5108                      ₹9,600                           455.0
6000                      ₹1,280                           455.0
6084                      ₹6,400                           455.0
...                          ...                              ...
12731                     ₹1,280                           700.0
12071                     ₹8,640                           455.0
8112                      ₹8,640                           455.0
10583                     ₹1,280                           455.0
3759                     ₹12,800                           455.0

       discount_price__currency_notnull  discount_price__price_string_notnull  \
3085                                INR
455
4341                                INR
455
5108                                INR
0
6000                                INR
0
6084                                INR
455
...                                 ...
...
12731                               INR
700
12071                               INR
455
8112                                INR
455
10583                               INR
0
3759                                INR
0

      price_detail__amount_notnull  price_detail__currency_notnull  \
3085                        8640.0                             INR
4341                        8640.0                             INR
5108                        9600.0                             INR
```
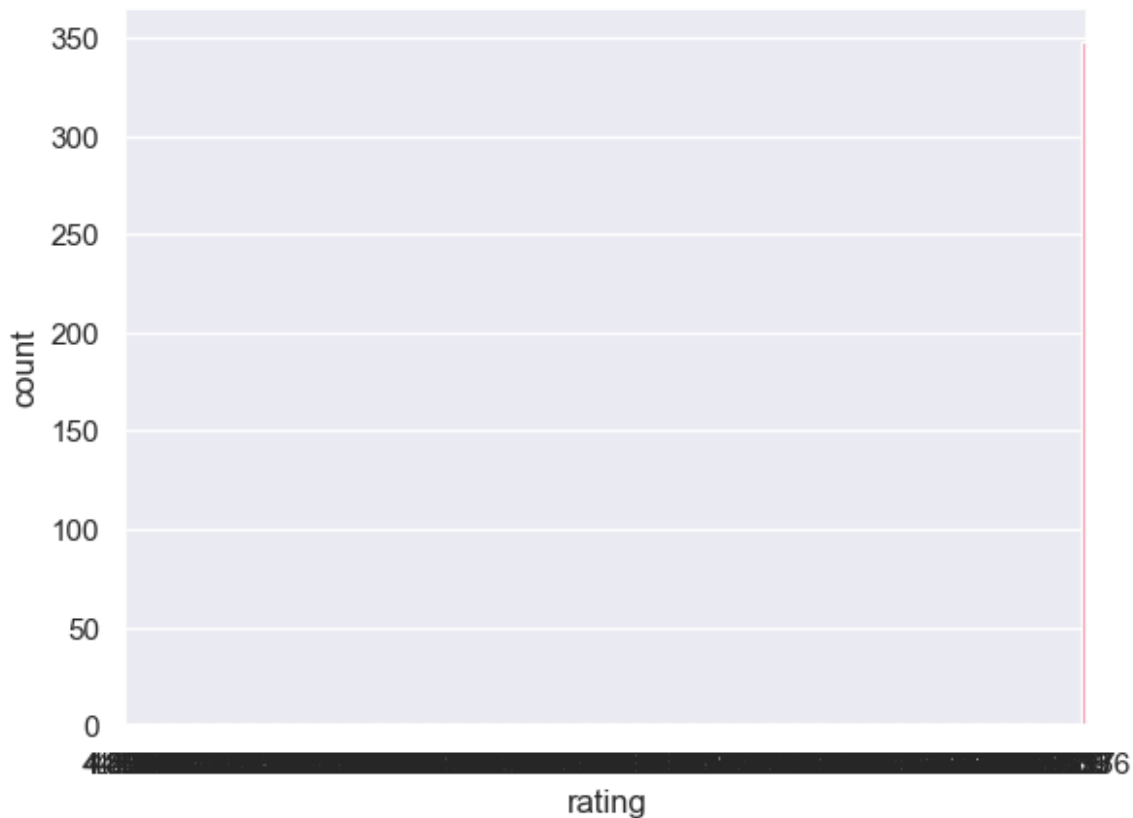
```
6000                                    1280.0                                    INR
6084                                    6400.0                                    INR
...                                        ...                                     ...
12731                                   1280.0                                    INR
12071                                   8640.0                                    INR
8112                                    8640.0                                    INR
10583                                   1280.0                                    INR
3759                                   12800.0                                    INR

       price_detail__price_string_notnull is_paid_plot
3085                                 ₹8,640            0
4341                                 ₹8,640            0
5108                                 ₹9,600            0
6000                                 ₹1,280            0
6084                                 ₹6,400            0
...                                     ...          ...
12731                                ₹1,280            0
12071                                ₹8,640            0
8112                                 ₹8,640            0
10583                                ₹1,280            0
3759                                ₹12,800            0

[500 rows x 27 columns]
```

```python
sns.set(style = "darkgrid")
sns.countplot(data=top_rating_df,x='rating')
```

```
<Axes: xlabel='rating', ylabel='count'>
```

4 Top Five courses name has highest rating?

```
rating_list = [5.0,4.98,4.97,4.8,4.7]
course_lists = ['Advanced Upwork Interviews','Beginners Course on
Warehousing and Distribution','Tools For Inspecting Your Home Or One
You May',"Public Relations: Crisis Communications Oil an.","Learn
About Safety Norms in Schools as a Priority"]
Highest_cd = ['AUI','BCWD',"TI","PR","LP"]
C_H =
pd.DataFrame({"Highest_rating":rating_list,"Highest_courses":course_li
sts,"full_cname1":Highest_cd})
plt.show()

sns.scatterplot(x = "full_cname1",y = "Highest_rating",data =
C_H,style = "full_cname1",hue = "full_cname1" ,sizes
=(500,600,500,500,500),palette = "gist_ncar")

<Axes: xlabel='full_cname1', ylabel='Highest_rating'>
```
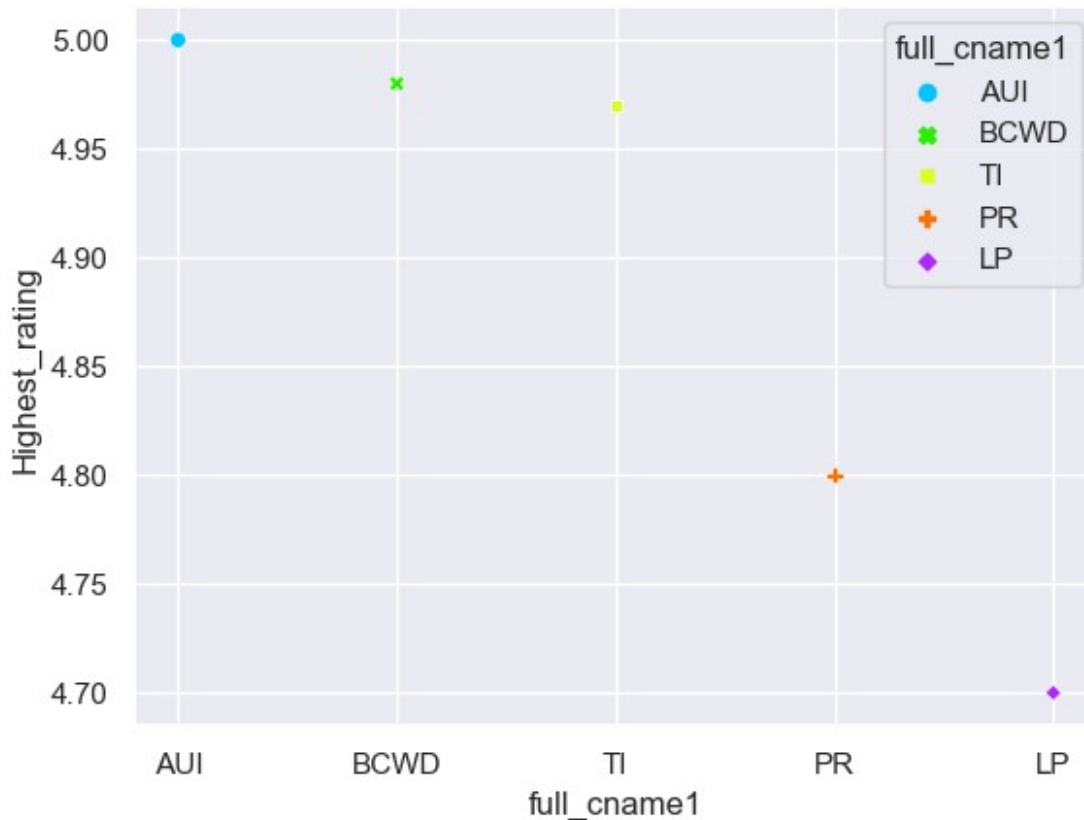
5 Which courses has highest no reviews ?

```
top_nreviews = br['num_reviews'].nlargest(n=5).index
top_nreviews

Int64Index([0, 1, 2, 3, 4], dtype='int64')

top_no_df = br.iloc[top_rating]

top_no_df

            id                                          title  \
3085    1538742  Advanced Upwork Interviews: A Simple Way to Ea...
4341    2350744  Price Action: The complete Price Action Master...
5108    1531238  New! Raise funds for your innovative business ...
6000      60463  Shipping address and Shipping cost for E-comme...
6084    3339914  Emotional Intelligence Training for Increased ...
...         ...                                             ...
12731   3234013                       LOS 4 ELEMENTOS DEL IMPUESTO
12071   3207347                    Risk Expert - Aprenda na Prática
8112    3141018  Build a passive business empire with Mini Kind...
10583   2805993            Careers in US Capital Market and PE & VC
3759    1855058  Open a Non-Medical Home Care Agency - Home Bas...
```

```
                                                    url  is_paid  \
3085                          /course/upwork-interviews/        1
4341                       /course/daytrading-masterclass/        1
5108          /course/eu-grant-funding-innovative-business/        1
6000                         /course/ecommerce-shipping/        1
6084      /course/emotional-intelligence-to-increase-sales/        1
...                                                   ...      ...
12731                 /course/los-4-elementos-del-impuesto/        1
12071                                 /course/risk-expert/        1
8112      /course/build-a-passive-business-empire-with-m...        1
10583                        /course/careers-in-finance-1/        1
3759              /course/open-a-non-medical-home-care-agency/        1

       num_subscribers  avg_rating  avg_rating_recent   rating  \
num_reviews
3085               508     5.00000            5.00000  5.00000
71
4341               439     5.00000            5.00000  5.00000
38
5108              2233     5.00000            5.00000  5.00000
28
6000               130     5.00000            5.00000  5.00000
19
6084                38     5.00000            5.00000  5.00000
18
...                ...         ...                ...      ...
...
12731               22     4.87500            4.86951  4.86951
4
12071              411     4.88889            4.86944  4.86944
55
8112                59     4.85714            4.86939  4.86939
8
10583                8     4.87500            4.86813  4.86813
4
3759               245     4.75000            4.86810  4.86810
50

       is_wishlisted  ...  price_detail__amount
price_detail__currency  \
3085           False  ...                8640.0
INR
4341           False  ...                8640.0
INR
5108           False  ...                9600.0
INR
6000           False  ...                1280.0
INR
6084           False  ...                6400.0
```

```
INR
...              ...  ...                    ...                      ..
.
12731          False  ...                 1280.0
INR
12071          False  ...                 8640.0
INR
8112           False  ...                 8640.0
INR
10583          False  ...                 1280.0
INR
3759           False  ...                12800.0
INR

       price_detail__price_string  discount_price__amount_notnull  \
3085                       ₹8,640                           455.0
4341                       ₹8,640                           455.0
5108                       ₹9,600                           455.0
6000                       ₹1,280                           455.0
6084                       ₹6,400                           455.0
...                           ...                             ...
12731                      ₹1,280                           700.0
12071                      ₹8,640                           455.0
8112                       ₹8,640                           455.0
10583                      ₹1,280                           455.0
3759                      ₹12,800                           455.0

       discount_price__currency_notnull
discount_price__price_string_notnull  \
3085                                INR
455
4341                                INR
455
5108                                INR
0
6000                                INR
0
6084                                INR
455
...                                 ...
...
12731                               INR
700
12071                               INR
455
8112                                INR
455
10583                               INR
0
```

```
3759                                          INR
0

       price_detail__amount_notnull  price_detail__currency_notnull  \
3085                          8640.0                             INR
4341                          8640.0                             INR
5108                          9600.0                             INR
6000                          1280.0                             INR
6084                          6400.0                             INR
...                              ...                             ...
12731                         1280.0                             INR
12071                         8640.0                             INR
8112                          8640.0                             INR
10583                         1280.0                             INR
3759                         12800.0                             INR

       price_detail__price_string_notnull  is_paid_plot
3085                               ₹8,640              0
4341                               ₹8,640              0
5108                               ₹9,600              0
6000                               ₹1,280              0
6084                               ₹6,400              0
...                                   ...            ...
12731                              ₹1,280              0
12071                              ₹8,640              0
8112                               ₹8,640              0
10583                              ₹1,280              0
3759                              ₹12,800              0

[500 rows x 27 columns]

review_list = [71,38,28]
course_lst = ["Advanced Upwork Interviews: A Simple Way to
Ea..","Price Action: The complete Price Action Master...","New! Raise
funds for your innovative business ."]
course_t = ["AUP","PA","NRF"]
reviewplot =
pd.DataFrame({"Highest_reviews":review_list,"course_name":course_lst,"
course_hue":course_t})

sns.set(style = "darkgrid")
sns.barplot(x = "course_hue",y ="Highest_reviews",data
=reviewplot,saturation = 10,errcolor="b")

<Axes: xlabel='course_hue', ylabel='Highest_reviews'>
```
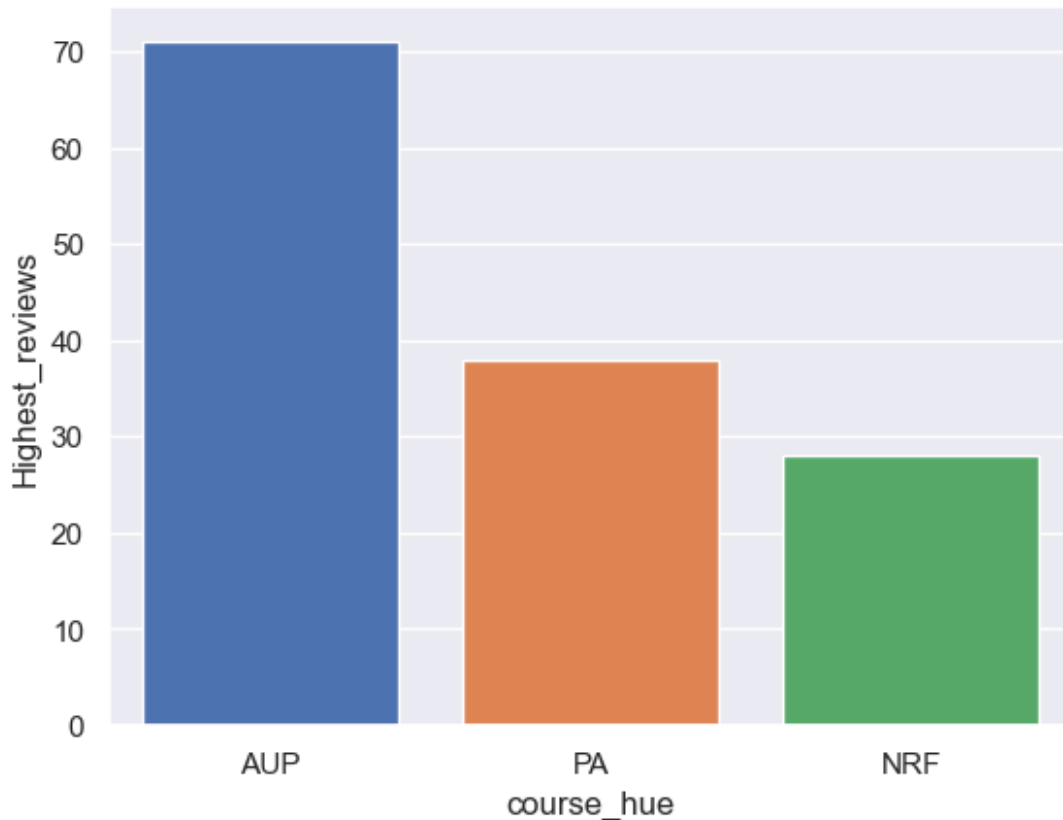
Q6

Which courses has highest no subscribers ?

```
top_subscriber = br['num_subscribers'].nlargest(n=5).index
top_subscriber

Int64Index([4, 0, 11484, 3, 19], dtype='int64')

top_subscriber_df = br.iloc[top_subscriber]
top_subscriber_df

           id                                        title  \
4      637930  An Entire MBA in 1 Course:Award Winning Busine...
0      762616  The Complete SQL Bootcamp 2020: Go from Zero t...
11484  133536              Stock Market Investing for Beginners
3      648826        The Complete Financial Analyst Course 2020
19     980086  Deep Learning Prerequisites: The Numpy Stack i...


                                              url  is_paid  \
4      /course/an-entire-mba-in-1-courseaward-winning...        1
```

```
0                       /course/the-complete-sql-bootcamp/          1
11484    /course/the-beginners-guide-to-the-stock-market/          0
3              /course/the-complete-financial-analyst-course/       1
19        /course/deep-learning-prerequisites-the-numpy-...       1

       num_subscribers  avg_rating  avg_rating_recent    rating
num_reviews  \
4                 374836     4.47080            4.47173  4.47173
41630
0                 295509     4.66019            4.67874  4.67874
78006
11484             259888     4.34928            4.35902  4.35902
23635
3                 245860     4.54407            4.53772  4.53772
46447
19                239309     4.47381            4.54005  4.54005
18322

       is_wishlisted  ...  price_detail__amount
price_detail__currency  \
4              False  ...                8640.0
INR
0              False  ...                8640.0
INR
11484          False  ...                   NaN
NaN
3              False  ...                8640.0
INR
19             False  ...                1280.0
INR

       price_detail__price_string  discount_price__amount_notnull  \
4                           ₹8,640                           455.0
0                           ₹8,640                           455.0
11484                          NaN                           455.0
3                           ₹8,640                           455.0
19                          ₹1,280                           455.0

       discount_price__currency_notnull
discount_price__price_string_notnull  \
4                                   INR
455
0                                   INR
455
11484                               INR
0
3                                   INR
455
19                                  INR
455
```

```
      price_detail__amount_notnull  price_detail__currency_notnull  \
4                          8640.0                             INR
0                          8640.0                             INR
11484                     12800.0                             INR
3                          8640.0                             INR
19                         1280.0                             INR

      price_detail__price_string_notnull  is_paid_plot
4                                  ₹8,640             0
0                                  ₹8,640             0
11484                             ₹12,800             0
3                                  ₹8,640             0
19                                 ₹1,280             0

[5 rows x 27 columns]
subs_list = [374836,295509,25988,245860,239309]
course_s = ["An Entire MBA in 1 Course:Award Winning Busine.","The
Complete SQL Bootcamp 2020: Go from Zero t..","Stock Market Investing
for Beginners","The Complete Financial Analyst Course 2020","Deep
Learning Prerequisites: The Numpy Stack i."]
course_sname = ["EMBA","SQL","STM","FIAN","NUMPY"]
C_subs =
pd.DataFrame({"subs_listn":subs_list,"course_name":course_sname,"cours
e_ns":course_s})

sns.set(style = "darkgrid")
sns.barplot(x ="course_name",y ="subs_listn",data =C_subs,saturation =
100,errcolor="b",palette = "icefire")

<Axes: xlabel='course_name', ylabel='subs_listn'>
```
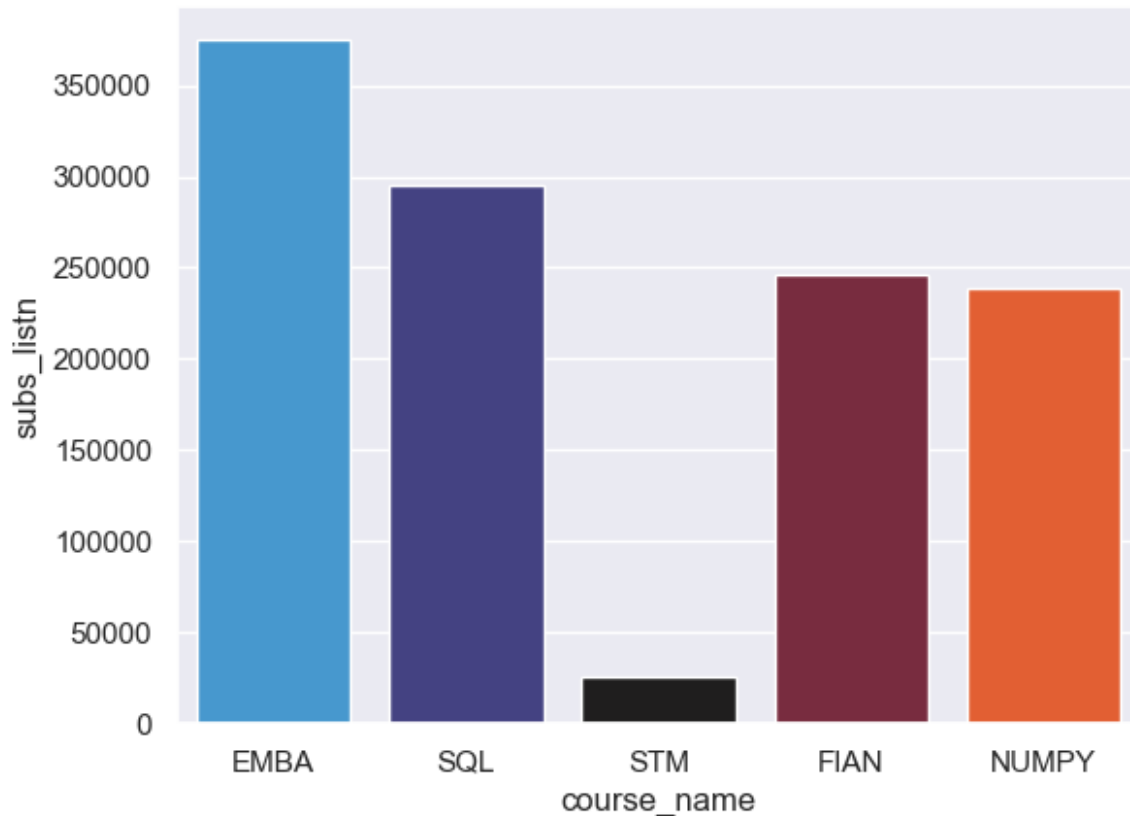
Q7

Which courses having highest no of discount?

```
top_discount =
br['discount_price__amount_notnull'].nlargest(n=120).index
top_discount

Int64Index([  365,    491,    677,    789,  1172,  1202,  1252,  1453,
2307,
               2924,
             ...
               1044,   1500,   3380,   4195, 10004, 12129, 13260,  1115,
2102,
               6268],
           dtype='int64', length=120)

top_discount_df = br.iloc[top_discount]
top_discount_df
```

```
               id                                              title  \
365        936754              Certified Six Sigma Black Belt Training
491       1464064                    QlikSense Data Architect Masterclass
677       1673566              Certified Quality Engineer Training
789       2103578          Certified Six Sigma Green Belt Training
1172       387514   Professional Sales Training - Compelling Conve...
...           ...                                                ...
12129     2122342                    Fundamentos da Análise de Crédito
13260     2580040       O PASSO A PASSO DO ISS — Imposto sobre Serviços
1115      1618964    Lead Auditor Tools for a Professional Internal...
2102      1739996    How to Write Effortless Quality Procedures & S...
6268      1595148          How to Become a Freelance Business Consultant

                                                    url  is_paid  \
365                         /course/six-sigma-black-belt/        1
491          /course/qliksense-data-architect-masterclass/        1
677               /course/quality-engineer-certification/        1
789           /course/six-sigma-green-belt-with-minitab/        1
1172             /course/compelling-sales-conversations/        1
...                                                   ...      ...
12129           /course/fundamentos-da-analise-de-credito/        1
13260   /course/o-passo-a-passo-do-iss-imposto-sobre-s...        1
1115             /course/be-a-pro-lead-auditor-from-scratch/        1
2102       /course/how-to-write-effortless-procedures-sops/        1
6268       /course/how-to-start-a-business-as-a-freelance...        1

        num_subscribers  avg_rating  avg_rating_recent   rating
num_reviews  \
365                7138     4.55333            4.58557  4.58557
1321
491                3472     4.37162            4.34758  4.34758
962
677                2680     4.48333            4.45706  4.45706
694
789                2600     4.62879            4.60111  4.60111
562
1172               2619     4.48148            4.49195  4.49195
341
...                 ...         ...                ...      ...
...
12129                94     4.33333            4.31730  4.31730
40
13260                11     0.00000            0.00000  0.00000
0
1115               1054     4.38462            4.38047  4.38047
365
2102                583     4.11111            4.09607  4.09607
134
6268               1426     3.95000            3.90116  3.90116
17
```

```
       is_wishlisted  ...  price_detail__amount
price_detail__currency  \
365            False  ...                8640.0
INR
491            False  ...                8640.0
INR
677            False  ...                8960.0
INR
789            False  ...                8640.0
INR
1172           False  ...                8640.0
INR
...              ...  ...                   ...                              ..
.
12129          False  ...                8640.0
INR
13260          False  ...                8640.0
INR
1115           False  ...                8640.0
INR
2102           False  ...                8640.0
INR
6268           False  ...                8640.0
INR

       price_detail__price_string discount_price__amount_notnull  \
365                         ₹8,640                         3200.0
491                         ₹8,640                         3200.0
677                         ₹8,960                         3200.0
789                         ₹8,640                         3200.0
1172                        ₹8,640                         3200.0
...                            ...                            ...
12129                       ₹8,640                         2240.0
13260                       ₹8,640                         2240.0
1115                        ₹8,640                         2160.0
2102                        ₹8,640                         2160.0
6268                        ₹8,640                         2160.0

       discount_price__currency_notnull
discount_price__price_string_notnull  \
365                                 INR
3200
491                                 INR
3200
677                                 INR
3200
789                                 INR
3200
1172                                INR
```

```
3200
...                                                 ...
...
12129                                               INR
2240
13260                                               INR
2240
1115                                                INR
2160
2102                                                INR
2160
6268                                                INR
2160

       price_detail__amount_notnull  price_detail__currency_notnull  \
365                          8640.0                              INR
491                          8640.0                              INR
677                          8960.0                              INR
789                          8640.0                              INR
1172                         8640.0                              INR
...                             ...                              ...
12129                        8640.0                              INR
13260                        8640.0                              INR
1115                         8640.0                              INR
2102                         8640.0                              INR
6268                         8640.0                              INR

       price_detail__price_string_notnull  is_paid_plot
365                                 ₹8,640             0
491                                 ₹8,640             0
677                                 ₹8,960             0
789                                 ₹8,640             0
1172                                ₹8,640             0
...                                    ...           ...
12129                               ₹8,640             0
13260                               ₹8,640             0
1115                                ₹8,640             0
2102                                ₹8,640             0
6268                                ₹8,640             0

[120 rows x 27 columns]

t_discount_price = [3200,2400,2240,2160,1760]
discount_c= ["Certified Six Sigma Black Belt Training","Exceptional
Leadership: Leading at a Higher Level","Fundamentos da Análise de
Crédito","Lead Auditor Tools for a Professional Internal.","ISO
22000:2018 Food Safety Management System "]
discount_cd = ["CT","EL","FD","LPI","ISO"]
dis_var =
```
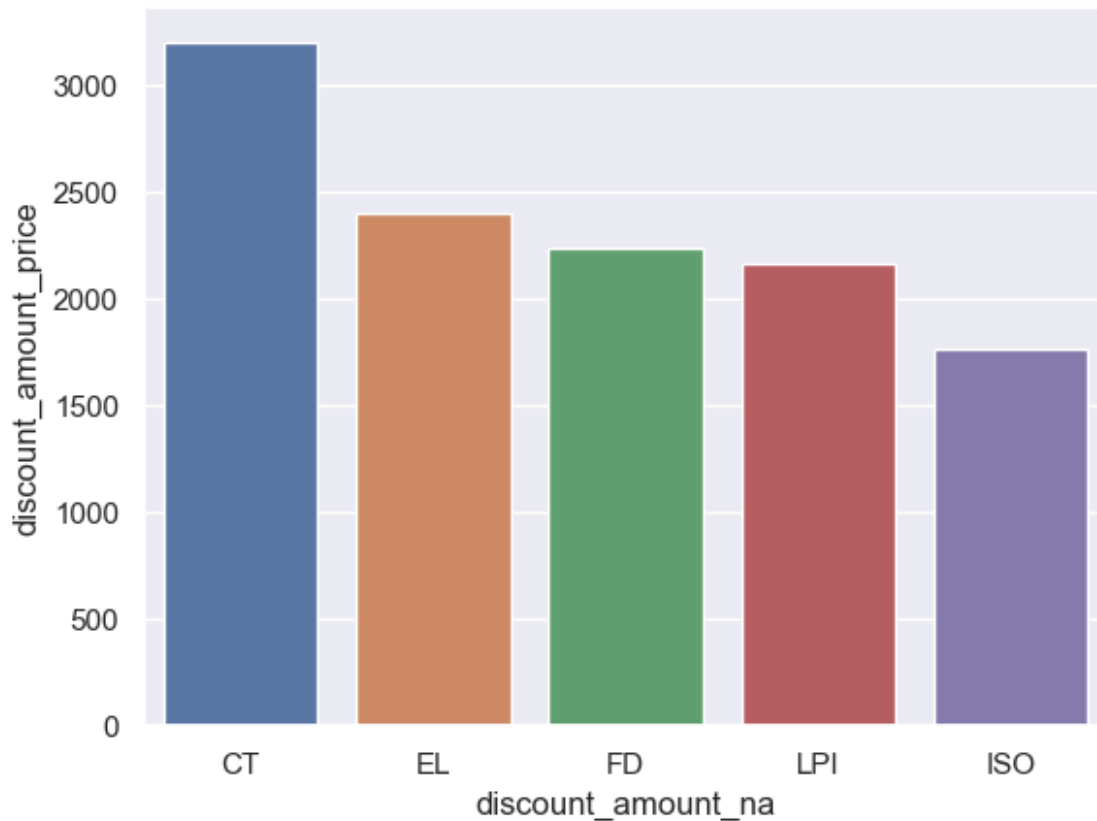
```
pd.DataFrame({"discount_amount_price":t_discount_price,"discount_amoun
t_na":discount_cd})

sns.barplot(x ="discount_amount_na",y = "discount_amount_price",data =
dis_var )

<Axes: xlabel='discount_amount_na', ylabel='discount_amount_price'>
```


8 which is the latest published courses of udemy?

```
br['published_time'] = pd.to_datetime(br['published_time'])

least_recent_date = br['published_time'].min()
most_recent_date = br['published_time'].max()

least_recent_date

Timestamp('2010-04-14 16:32:46+0000', tz='UTC')

most_recent_date
```

```
Timestamp('2020-09-10 06:31:11+0000', tz='UTC')

least_recent_date = br['published_time'].nlargest(n=5).index

least_recent_date

Int64Index([11269, 11761, 11080, 11203, 13293], dtype='int64')

top_df = br.loc[[11269,11761,11080,11203,13293],"title"]
top_df

11269    CBBF (BTA - CERTIFIED BLOCKCHAIN BUSINESS FOUN...
11761    Trading Trend Like A Pro : Perfect Strategy fo...
11080    Forex Online Course - Beginners / Intermediate...
11203    Forex Beginners Course - Complete Guide to For...
13293                                 Financial Accounting
Name: title, dtype: object

top_dfd = br.loc[[11269,11761,11080,11203,13293],"published_time"]
top_dfd

11269    2020-09-10 06:31:11+00:00
11761    2020-09-09 04:21:55+00:00
11080    2020-09-09 02:24:43+00:00
11203    2020-09-09 02:11:42+00:00
13293    2020-09-08 22:52:22+00:00
Name: published_time, dtype: datetime64[ns, UTC]

published_date_list =
[2020_09_10 ,2020_09_09,2020_09_09 ,2020_09_09,2020_09_08]
title_l = ["Trading Trend Like A Pro : Perfect Strategy
fo...","Trading Trend Like A Pro : Perfect Strategy fo...","Forex
Online Course - Beginners / Intermediate...","Forex Beginners Course -
Complete Guide to For...","Financial Accounting"]
title_s = ["PS","TT","FO","FB","FA"]
co_data =
pd.DataFrame({"published_data":published_date_list,"title_name":title_
l,"TITLE":title_s})

sns.lineplot(x = "TITLE",y = "published_data",data = co_data)

<Axes: xlabel='TITLE', ylabel='published_data'>
```
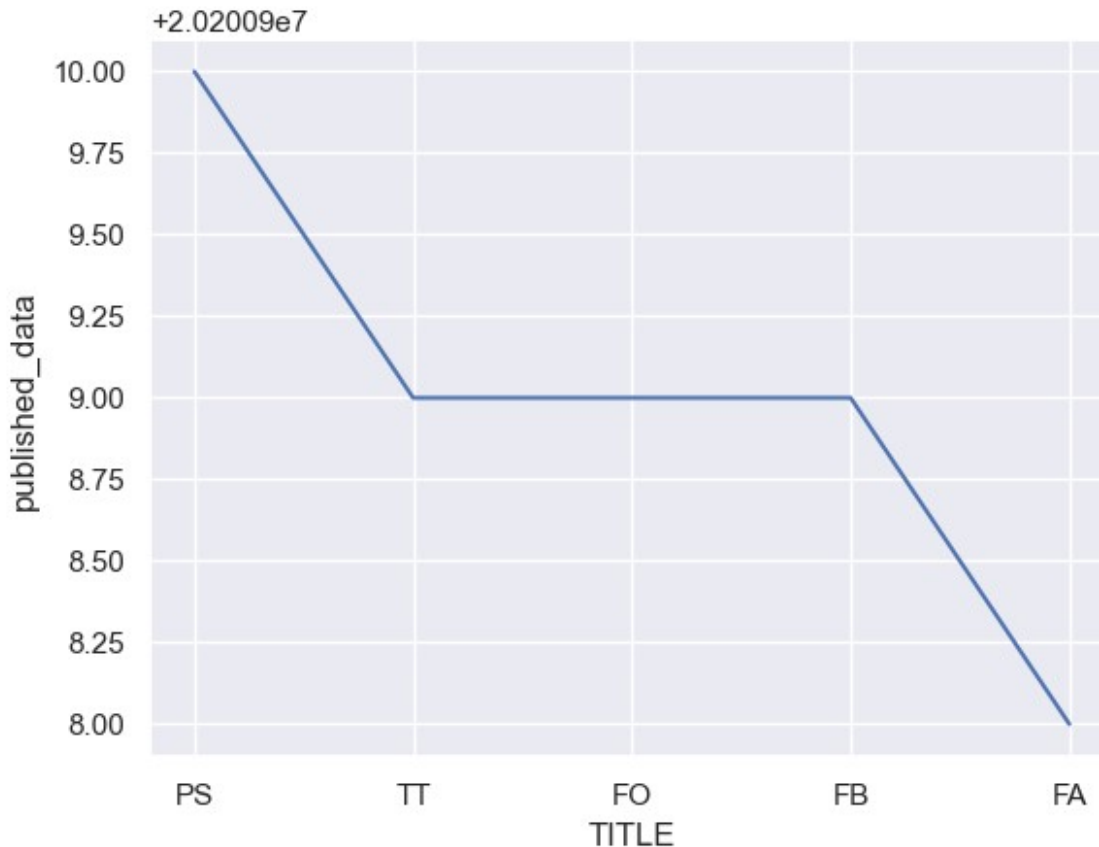
+2.02009e7

which is oldest courses of udemy?

```
most_recent_date = br['published_time'].min()

most_recent_date

Timestamp('2010-04-14 16:32:46+0000', tz='UTC')

least_recent_date = br['published_time'].nsmallest(n=5).index

least_recent_date

Int64Index([686, 1191, 856, 885, 3677], dtype='int64')

top_dfo = br.loc[[686,1191,856,885,3677],"published_time"]
top_dfo

686     2010-04-14 16:32:46+00:00
1191    2011-07-06 14:06:34+00:00
856     2011-07-11 06:29:02+00:00
885     2011-08-22 01:55:25+00:00
3677    2011-10-12 17:21:40+00:00
Name: published_time, dtype: datetime64[ns, UTC]
```

```python
top_dfon = br.loc[[686,1191,856,885,3677],"title"]
top_dfon
```

```
686      Simple Strategy for Swing Trading the Stock Ma...
1191     How to Create an Awesome Demo Video for Your B...
856                                      The Lean Startup
885                      Speak Up! Stand Out and be HAPPY
3677                            How to Start a Business
Name: title, dtype: object
```

```python
oldest_cor = [2010_04_14,2011_07_06,2011_07_11,2011_08_22,2011-10-12]
old_course = ["Simple Strategy for Swing Trading the Stock Ma...","How
to Create an Awesome Demo Video for Your B..","The Lean
Startup","Speak Up! Stand Out and be HAPPY","How to Start a Business"]
old_sh = ["ST","HB","ST","SH","HB"]
C_OD =
pd.DataFrame({"oldest_pub_data":oldest_cor,"old_pub_cname":old_sh})

sns.lineplot(x = "old_pub_cname",y = "oldest_pub_data",data =
C_OD ,palette = 'Accent',markers = ["o",">"])
```

```
C:\Users\Subham Ranjan\AppData\Local\Temp\
ipykernel_8544\3837670618.py:1: UserWarning: Ignoring `palette`
because no `hue` variable has been assigned.
  sns.lineplot(x = "old_pub_cname",y = "oldest_pub_data",data =
C_OD ,palette = 'Accent',markers = ["o",">"])
```

```
<Axes: xlabel='old_pub_cname', ylabel='oldest_pub_data'>
```
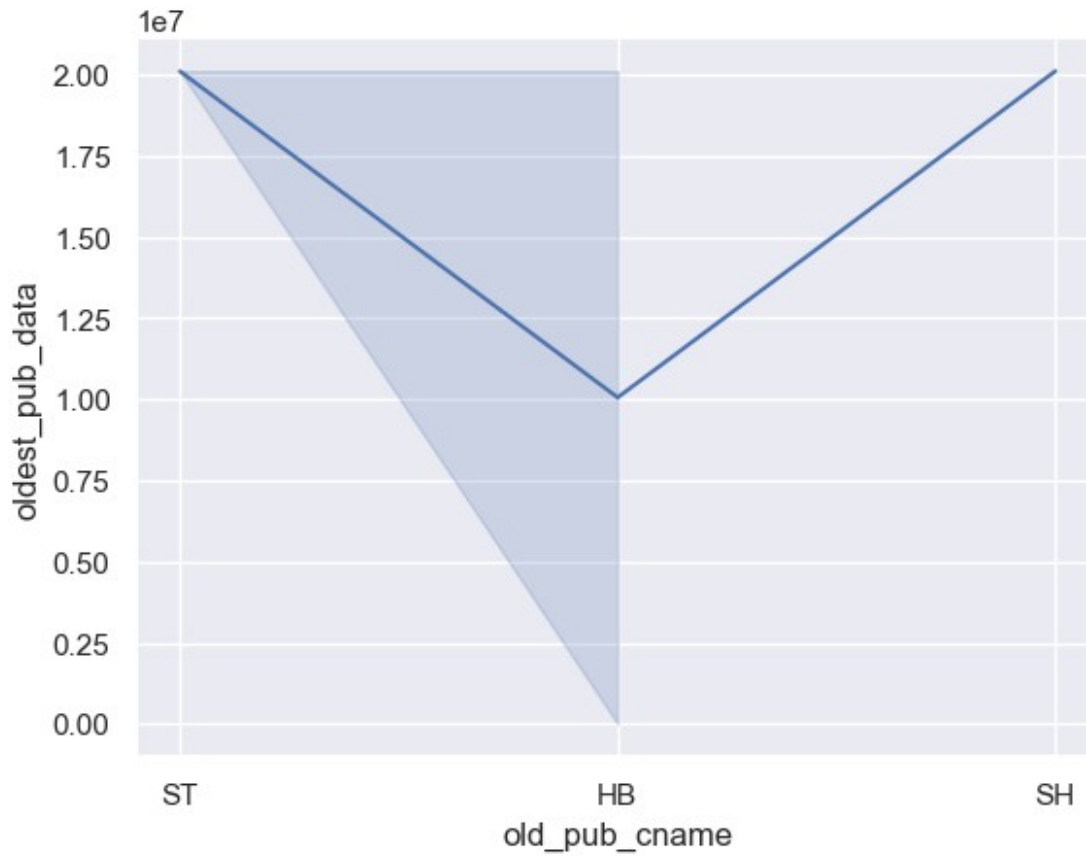
```
rating_list = [5.0,4.98,4.97,4.8,4.7]

expensive_course =
br['price_detail__amount_notnull'].nlargest(n=550).index
expensive_course

Int64Index([ 90,  95, 102, 142, 213, 358, 460, 499, 566, 663,
            ...
            335, 371, 409, 440, 544, 550, 557, 568, 673, 677],
           dtype='int64', length=550)

top_expensive_df = br.iloc[expensive_course]
top_expensive_df

           id                                          title  \
90     328266   Project Management Professional (PMP)® - 35 Co...
95     475170                     Seth Godin's Freelancer Course
102    665152            Mastering Agile Scrum Project Management
142   2296344   Speak Like a Pro: Public Speaking for Professi...
213   1123940                 Think Like a Leader with Brian Tracy
..        ...                                                ...
550   2441460   PMI-PMP-Project Management Professional - 35 C...
557    402836   Accounting 101: How to read an Accounting Bala...
568    455794   HR 101: Compliance Made Easy  -  U.S. Employer...
```

```
673    1336300              Manager's Guide to Difficult Conversations
677    1673566                      Certified Quality Engineer Training

                                                      url  is_paid  \
90                      /course/pmp-certification-complete/        1
95                      /course/seth-godin-freelancer-course/        1
102    /course/mastering-agile-scrum-project-manageme...        1
142    /course/speak-like-a-pro-public-speaking-for-p...        1
213          /course/think-like-a-leader-with-brian-tracy/        1
..                                                     ...      ...
550    /course/pmi-pmp-project-management-professiona...        1
557    /course/financial-statements-101-how-to-read-a...        1
568                  /course/hr-101-compliance-made-easy/        1
673    /course/managers-guide-to-difficult-conversati...        1
677              /course/quality-engineer-certification/        1

     num_subscribers  avg_rating  avg_rating_recent    rating  num_reviews  \
90             27653     4.40769            4.41472   4.41472         4579
95             24449     4.69697            4.69071   4.69071         4361
102            26047     4.27826            4.28871   4.28871         4163
142             7556     4.34922            4.36594   4.36594         3270
213            12113     4.49656            4.52477   4.52477         2268
..               ...         ...                ...       ...          ...
550             2910     4.41991            4.39563   4.39563          848
557             5684     4.52163            4.54807   4.54807          837
568             3443     4.51869            4.54344   4.54344          826
673             2810     4.31379            4.30476   4.30476          700
677             2680     4.48333            4.45706   4.45706          694

     is_wishlisted  ...  price_detail__amount  price_detail__currency  \
90           False  ...               12800.0                     INR
95           False  ...               12800.0                     INR
102          False  ...               12800.0                     INR
142          False  ...               12800.0                     INR
```

```
213          False  ...              12800.0                    INR

..             ...  ...                  ...                    ...

550          False  ...               8960.0                    INR

557          False  ...               8960.0                    INR

568          False  ...               8960.0                    INR

673          False  ...               8960.0                    INR

677          False  ...               8960.0                    INR


     price_detail__price_string  discount_price__amount_notnull  \
90                       ₹12,800                           455.0
95                       ₹12,800                           455.0
102                      ₹12,800                           455.0
142                      ₹12,800                           455.0
213                      ₹12,800                           455.0
..                           ...                             ...
550                       ₹8,960                           481.0
557                       ₹8,960                           474.0
568                       ₹8,960                           474.0
673                       ₹8,960                           481.0
677                       ₹8,960                          3200.0

     discount_price__currency_notnull
discount_price__price_string_notnull  \
90                                INR
0
95                                INR
0
102                               INR
0
142                               INR
0
213                               INR
0
..                                ...
...
550                               INR
481
557                               INR
474
568                               INR
474
673                               INR
```

```
481
677                                               INR
3200

     price_detail__amount_notnull  price_detail__currency_notnull  \
90                         12800.0                             INR
95                         12800.0                             INR
102                        12800.0                             INR
142                        12800.0                             INR
213                        12800.0                             INR
..                             ...                             ...
550                         8960.0                             INR
557                         8960.0                             INR
568                         8960.0                             INR
673                         8960.0                             INR
677                         8960.0                             INR

     price_detail__price_string_notnull  is_paid_plot
90                                ₹12,800             0
95                                ₹12,800             0
102                               ₹12,800             0
142                               ₹12,800             0
213                               ₹12,800             0
..                                    ...           ...
550                                ₹8,960             0
557                                ₹8,960             0
568                                ₹8,960             0
673                                ₹8,960             0
677                                ₹8,960             0

[550 rows x 27 columns]

top_exo =
br.loc[[365,491,677,789,1172],"discount_price__amount_notnull"]
top_exo

365     3200.0
491     3200.0
677     3200.0
789     3200.0
1172    3200.0
Name: discount_price__amount_notnull, dtype: float64

my_list = [12800,9600,8960,8640]
my_cname = ["Project Management Professional (PMP)® _35
Co...""",Introduction to Finance, Accounting,Certified Quality Engineer
Training, Modeling" ,"Udemy Affiliate  Social Media marketing cours"]
mys_cname = ["PP","IF","ACM","AFM"]
C_NAME = pd.DataFrame({"exp_amt":my_list,"exp_cname":mys_cname})
```
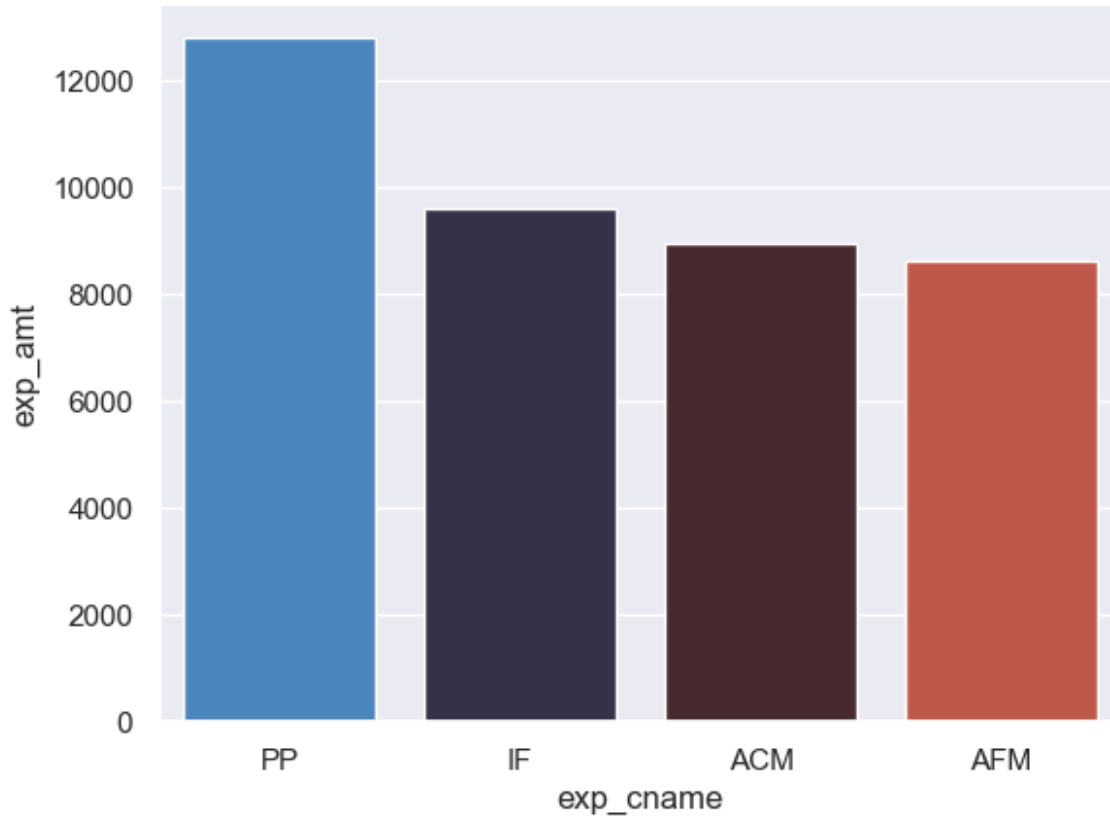
```
sns.barplot(x = "exp_cname",y = "exp_amt",data = C_NAME,palette =
"icefire")
```

```
<Axes: xlabel='exp_cname', ylabel='exp_amt'>
```



```
cheapest_course =
br['price_detail__amount_notnull'].nsmallest(n=3500).index
cheapest_course
```

```
Int64Index([  19,   46,   59,   64,   78, 139, 148, 149, 156,
157,
             ...
           5183, 5195, 5198, 5235, 5242, 5254, 5264, 5302, 5349,
5380],
           dtype='int64', length=3500)
```

```
top_cheapest_df = br.iloc[cheapest_course]
top_cheapest_df
```

```
        id                                       title  \
19    980086  Deep Learning Prerequisites: The Numpy Stack i...
46    536482                    Security Awareness Training
59    672290  Brilliant Customer Service: How to Impress you...
64   1758098                      How to tell a story with data
```

```
78      129118                                Product Management 101
...         ...                                                  ...
5254    288964    Consulting Strategies: Succeed as a Niche Cons...
5264   2112078                           Visualizing data using Stata
5302     49009                     How I pay my rent selling balloons
5349   2270114              Managing commercial website projects
5380   1074696                              Learn business dress code

                                                    url  is_paid  \
19      /course/deep-learning-prerequisites-the-numpy-...        1
46                    /course/security-awareness-training/        1
59            /course/brilliant-customer-service-lecturio/        1
64                        /course/tell-a-story-with-data/        1
78                         /course/productmanagement101/        1
...                                                   ...      ...
5254    /course/consulting-strategies-succeed-as-a-nic...        1
5264               /course/visualizing-data-using-stata/        1
5302    /course/steps-to-start-your-own-balloon-distri...        1
5349       /course/managing-commercial-website-projects/        1
5380                  /course/learn-business-dress-code/        1

      num_subscribers  avg_rating  avg_rating_recent   rating
num_reviews  \
19             239309     4.47381            4.54005  4.54005
18322
46              16511     4.51703            4.55482  4.55482
8915
59              12728     4.34038            4.34485  4.34485
6401
64              15059     4.36463            4.34164  4.34164
6075
78              40950     4.46622            4.46130  4.46130
4960
...               ...         ...                ...      ...
...
5254             2449     3.90000            4.14263  4.14263
26
5264             2729     4.55000            4.51303  4.51303
26
5302              463     4.25000            3.99639  3.99639
25
5349               66     4.05000            4.02094  4.02094
25
5380              765     4.40000            4.52771  4.52771
25

      is_wishlisted  ...  price_detail__amount  price_detail__currency
\
19            False  ...                1280.0                     INR
```

```
46            False  ...              1280.0                      INR

59            False  ...              1280.0                      INR

64            False  ...              1280.0                      INR

78            False  ...              1280.0                      INR

...             ...  ...                 ...                      ...

5254          False  ...              1600.0                      INR

5264          False  ...              1600.0                      INR

5302          False  ...              1600.0                      INR

5349          False  ...              1600.0                      INR

5380          False  ...              1600.0                      INR


      price_detail__price_string  discount_price__amount_notnull  \
19                        ₹1,280                           455.0
46                        ₹1,280                           455.0
59                        ₹1,280                           455.0
64                        ₹1,280                           455.0
78                        ₹1,280                           462.0
...                          ...                             ...
5254                      ₹1,600                           455.0
5264                      ₹1,600                           455.0
5302                      ₹1,600                           455.0
5349                      ₹1,600                           455.0
5380                      ₹1,600                           455.0

      discount_price__currency_notnull discount_price__price_string_notnull  \
19                                 INR                                  455
46                                 INR                                  455
59                                 INR                                  455
64                                 INR                                  455
78                                 INR                                  462
...                                ...                                  ...
5254                               INR                                  455
```

```
5264                                    INR
455
5302                                    INR
455
5349                                    INR
455
5380                                    INR
455

      price_detail__amount_notnull  price_detail__currency_notnull  \
19                          1280.0                             INR
46                          1280.0                             INR
59                          1280.0                             INR
64                          1280.0                             INR
78                          1280.0                             INR
...                            ...                             ...
5254                        1600.0                             INR
5264                        1600.0                             INR
5302                        1600.0                             INR
5349                        1600.0                             INR
5380                        1600.0                             INR

      price_detail__price_string_notnull is_paid_plot
19                               ₹1,280            0
46                               ₹1,280            0
59                               ₹1,280            0
64                               ₹1,280            0
78                               ₹1,280            0
...                                 ...          ...
5254                             ₹1,600            0
5264                             ₹1,600            0
5302                             ₹1,600            0
5349                             ₹1,600            0
5380                             ₹1,600            0

[3500 rows x 27 columns]

my_ch = [1280,1600,2240,3840,]
mych_cname = ["Deep Learning Prerequisites: The Numpy Stack
i.","Microsoft Excel Aplicado a las Finanzas / Exce.","SAP
Accounting","Master Public Policy Analysis: Advanced"]
mych_scname =["NUMPY_s","EXCEL","S_AC","MPPA"]
CH = pd.DataFrame({"cheap_price":my_ch,"mycourse_scname":mych_scname})

sns.barplot(x = "mycourse_scname",y = "cheap_price",data = CH)

<Axes: xlabel='mycourse_scname', ylabel='cheap_price'>
```