Udemy Project

February 26, 2025

```
[2]: import pandas as pd
     import numpy as np
     import matplotlib.pyplot as plt
     import seaborn as sns
```

```
[3]: df=pd.read_csv('udemy_courses_clean.csv')
```

```
[4]: df
```

```
[4]:       Unnamed: 0  course_id  \
     0              0    1070968
     1              1    1113822
     2              2    1006314
     3              3    1210588
     4              4    1011058
     ...          ...        ...
     3678        3678     775618
     3679        3679    1088178
     3680        3680     635248
     3681        3681     905096
     3682        3682     297602


                                           course_title  \
     0                Ultimate Investment Banking Course
     1       Complete GST Course & Certification - Grow You…
     2       Financial Modeling for Business Analysts and C…
     3       Beginner to Pro - Financial Analysis in Excel …
     4            How To Maximize Your Profits Trading Options
     ...                                                ...
     3678  Learn jQuery from Scratch - Master of JavaScri…
     3679  How To Design A WordPress Website With No Codi…
     3680                    Learn and Build using Polymer
     3681  CSS Animations: Create Amazing Effects on Your…
     3682  Using MODX CMS to Build Websites: A Beginner's…


                                                   url is_paid price  \
     0     https://www.udemy.com/ultimate-investment-bank…    True   200
     1          https://www.udemy.com/goods-and-services-tax/    True    75
```

```
2      https://www.udemy.com/financial-modeling-for-b…      True     45
3      https://www.udemy.com/complete-excel-finance-c…      True     95
4      https://www.udemy.com/how-to-maximize-your-pro…      True    200
…                                                   …       …    …
3678   https://www.udemy.com/easy-jquery-for-beginner…      True    100
3679   https://www.udemy.com/how-to-make-a-wordpress-…      True     25
3680   https://www.udemy.com/learn-and-build-using-po…      True     40
3681   https://www.udemy.com/css-animations-create-am…      True     50
3682   https://www.udemy.com/using-modx-cms-to-build-…      True     45

       num_subscribers   num_reviews   num_lectures                 level  \
0                 2147            23             51            All Levels
1                 2792           923            274            All Levels
2                 2174            74             51     Intermediate Level
3                 2451            11             36            All Levels
4                 1276            45             26     Intermediate Level
…                    …             …              …                     …
3678              1040            14             21            All Levels
3679               306             3             42        Beginner Level
3680               513           169             48            All Levels
3681               300            31             38            All Levels
3682               901            36             20        Beginner Level

       content_duration   published_timestamp              subject  \
0               1.5 hours  2017-01-18T20:58:58Z   Business Finance
1                39 hours  2017-03-09T16:34:20Z   Business Finance
2               2.5 hours  2016-12-19T19:26:30Z   Business Finance
3                 3 hours  2017-05-30T20:07:24Z   Business Finance
4                 2 hours  2016-12-13T14:57:18Z   Business Finance
…                      …                     …                   …
3678              2 hours  2016-06-14T17:36:46Z   Web Development
3679            3.5 hours  2017-03-10T22:24:30Z   Web Development
3680            3.5 hours  2015-12-30T16:41:42Z   Web Development
3681              3 hours  2016-08-11T19:06:15Z   Web Development
3682              2 hours  2014-09-28T19:51:11Z   Web Development

                                    clean_course_title
0                   Ultimate Investment Banking Course
1           Complete GST Course  Certification  Grow Practice
2         Financial Modeling Business Analysts Consultants
3              Beginner Pro  Financial Analysis Excel 2017
4                       Maximize Profits Trading Options
…                                                   …
3678       Learn jQuery Scratch  Master JavaScript library
3679                    Design WordPress Website Coding
3680                              Learn Build Polymer
3681        CSS Animations Create Amazing Effects Website
```

```
     3682              MODX CMS Build Websites Beginners Guide

     [3683 rows x 14 columns]
```

[4]: `df.columns`

[4]:
```
Index(['Unnamed: 0', 'course_id', 'course_title', 'url', 'is_paid', 'price',
       'num_subscribers', 'num_reviews', 'num_lectures', 'level',
       'content_duration', 'published_timestamp', 'subject',
       'clean_course_title'],
      dtype='object')
```

[5]: `df.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3683 entries, 0 to 3682
Data columns (total 14 columns):
 #   Column               Non-Null Count  Dtype
---  ------               --------------  -----
 0   Unnamed: 0           3683 non-null   int64
 1   course_id            3683 non-null   int64
 2   course_title         3683 non-null   object
 3   url                  3683 non-null   object
 4   is_paid              3683 non-null   object
 5   price                3683 non-null   object
 6   num_subscribers      3683 non-null   int64
 7   num_reviews          3683 non-null   int64
 8   num_lectures         3683 non-null   int64
 9   level                3683 non-null   object
 10  content_duration     3683 non-null   object
 11  published_timestamp  3683 non-null   object
 12  subject              3683 non-null   object
 13  clean_course_title   3663 non-null   object
dtypes: int64(5), object(9)
memory usage: 403.0+ KB
```

[6]: `df.isnull().sum()`

[6]:
```
Unnamed: 0            0
course_id            0
course_title         0
url                  0
is_paid              0
price                0
num_subscribers      0
num_reviews          0
num_lectures         0
level                0
```

```
content_duration        0
published_timestamp     0
subject                 0
clean_course_title     20
dtype: int64
```

1 subject

```
[7]: df['subject'].unique()
```

```
[7]: array(['Business Finance', 'Graphic Design', 'Musical Instruments',
            'Web Development'], dtype=object)
```
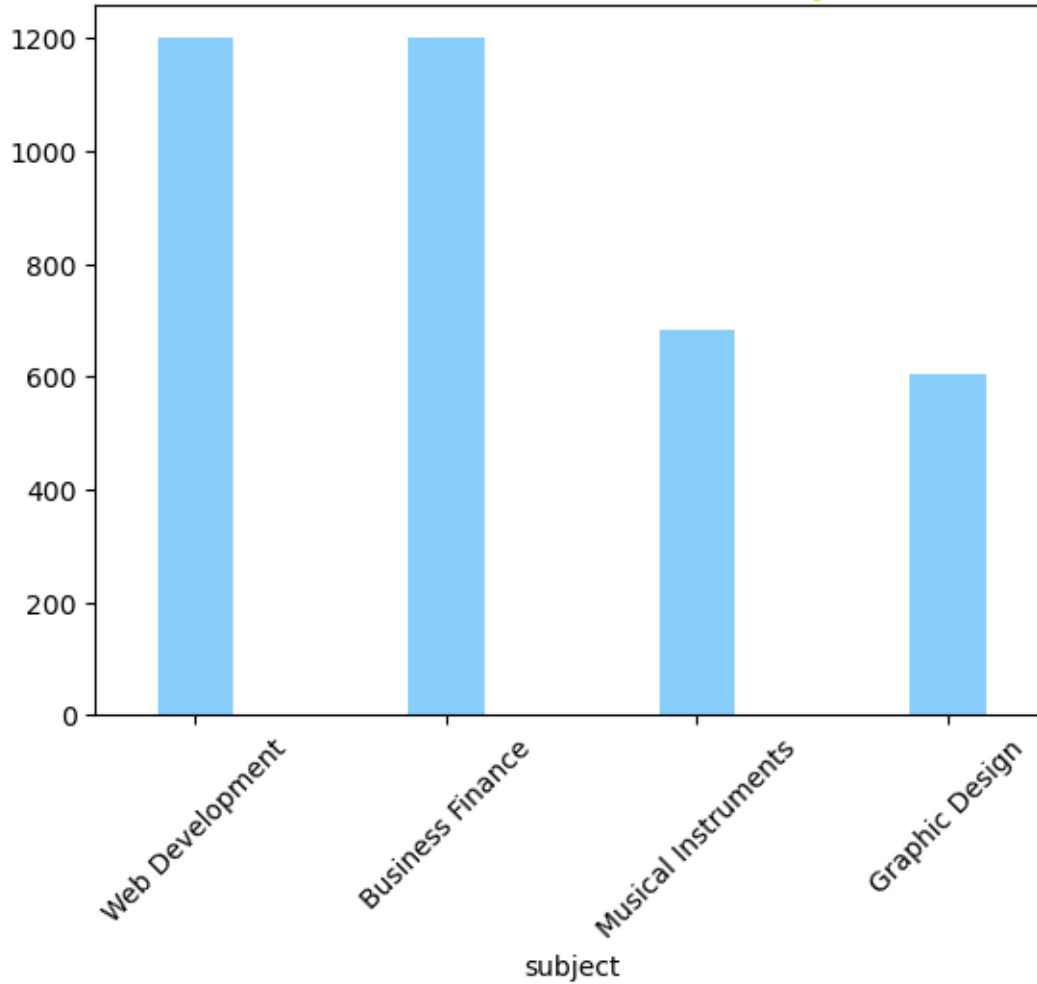
```
[8]: df['subject'].value_counts()
```

```
[8]: subject
     Web Development        1200
     Business Finance       1199
     Musical Instruments     681
     Graphic Design          603
     Name: count, dtype: int64
```

```
[9]: df['subject'].value_counts().plot(kind='bar',fontsize=10,␣
      ↪color='lightskyblue',width=0.3 )
     plt.title('The Distribution of Subjects', fontsize=20 ,c='gold')
     plt.xticks(rotation=45)

     plt.show()
```

The Distribution of Subjects

2 num_subscribers

```
[10]: df.groupby('subject')['num_subscribers'].sum()
```

```
[10]: subject
      Business Finance      1870747
      Graphic Design        1063148
      Musical Instruments    846719
      Web Development       7980572
      Name: num_subscribers, dtype: int64
```

```
[11]: df['num_subscribers'].describe()
```

```
[11]: count        3683.000000
      mean         3193.371165
      std          9498.231406
      min             0.000000
      25%           110.000000
      50%           911.000000
      75%          2537.500000
      max        268923.000000
      Name: num_subscribers, dtype: float64
```

```
[12]: df['num_subscribers'].idxmax()
```

```
[12]: 2832
```

```
[13]: df.iloc[2832]
```

```
[13]: Unnamed: 0                                              2832
      course_id                                              41295
      course_title             Learn HTML5 Programming From Scratch
      url                   https://www.udemy.com/learn-html5-programming-…
      is_paid                                                False
      price                                                   Free
      num_subscribers                                       268923
      num_reviews                                             8629
      num_lectures                                              45
      level                                             All Levels
      content_duration                                   10.5 hours
      published_timestamp                      2013-02-14T07:03:41Z
      subject                                       Web Development
      clean_course_title          Learn HTML5 Programming Scratch
      Name: 2832, dtype: object
```

3 level

```
[14]: df
```

```
[14]:       Unnamed: 0  course_id  \
      0              0    1070968
      1              1    1113822
      2              2    1006314
      3              3    1210588
      4              4    1011058
      …            …         …
      3678        3678     775618
      3679        3679    1088178
      3680        3680     635248
      3681        3681     905096
```

```
3682          3682       297602

                                              course_title  \
0                    Ultimate Investment Banking Course
1        Complete GST Course & Certification - Grow You…
2        Financial Modeling for Business Analysts and C…
3        Beginner to Pro - Financial Analysis in Excel …
4            How To Maximize Your Profits Trading Options
…                                                       …
3678  Learn jQuery from Scratch - Master of JavaScri…
3679  How To Design A WordPress Website With No Codi…
3680                         Learn and Build using Polymer
3681  CSS Animations: Create Amazing Effects on Your…
3682  Using MODX CMS to Build Websites: A Beginner's…


                                              url is_paid price  \
0     https://www.udemy.com/ultimate-investment-bank…    True   200
1          https://www.udemy.com/goods-and-services-tax/   True    75
2     https://www.udemy.com/financial-modeling-for-b…    True    45
3     https://www.udemy.com/complete-excel-finance-c…    True    95
4     https://www.udemy.com/how-to-maximize-your-pro…    True   200
…                                                   …     …    …
3678  https://www.udemy.com/easy-jquery-for-beginner…    True   100
3679  https://www.udemy.com/how-to-make-a-wordpress-…    True    25
3680  https://www.udemy.com/learn-and-build-using-po…    True    40
3681  https://www.udemy.com/css-animations-create-am…    True    50
3682  https://www.udemy.com/using-modx-cms-to-build-…    True    45


      num_subscribers  num_reviews  num_lectures              level  \
0                2147           23            51          All Levels
1                2792          923           274          All Levels
2                2174           74            51   Intermediate Level
3                2451           11            36          All Levels
4                1276           45            26   Intermediate Level
…                   …            …             …                   …
3678             1040           14            21          All Levels
3679              306            3            42      Beginner Level
3680              513          169            48          All Levels
3681              300           31            38          All Levels
3682              901           36            20      Beginner Level


      content_duration  published_timestamp             subject  \
0            1.5 hours  2017-01-18T20:58:58Z  Business Finance
1             39 hours  2017-03-09T16:34:20Z  Business Finance
2            2.5 hours  2016-12-19T19:26:30Z  Business Finance
3              3 hours  2017-05-30T20:07:24Z  Business Finance
4              2 hours  2016-12-13T14:57:18Z  Business Finance
```

```
...          ...               ...                    ...
3678        2 hours   2016-06-14T17:36:46Z    Web Development
3679        3.5 hours 2017-03-10T22:24:30Z    Web Development
3680        3.5 hours 2015-12-30T16:41:42Z    Web Development
3681        3 hours   2016-08-11T19:06:15Z    Web Development
3682        2 hours   2014-09-28T19:51:11Z    Web Development

                                  clean_course_title
0                      Ultimate Investment Banking Course
1          Complete GST Course  Certification  Grow Practice
2         Financial Modeling Business Analysts Consultants
3              Beginner Pro  Financial Analysis Excel 2017
4                          Maximize Profits Trading Options
...                                     ...
3678      Learn jQuery Scratch  Master JavaScript library
3679                     Design WordPress Website Coding
3680                               Learn Build Polymer
3681      CSS Animations Create Amazing Effects Website
3682            MODX CMS Build Websites Beginners Guide

[3683 rows x 14 columns]
```

[15]: `df['level'].unique()`

[15]:
```
array(['All Levels', 'Intermediate Level', 'Beginner Level',
       'Expert Level', '52'], dtype=object)
```

[16]:
```
df1=df['level']=='52'
df1
```

[16]:
```
0       False
1       False
2       False
3       False
4       False
        ...
3678    False
3679    False
3680    False
3681    False
3682    False
Name: level, Length: 3683, dtype: bool
```
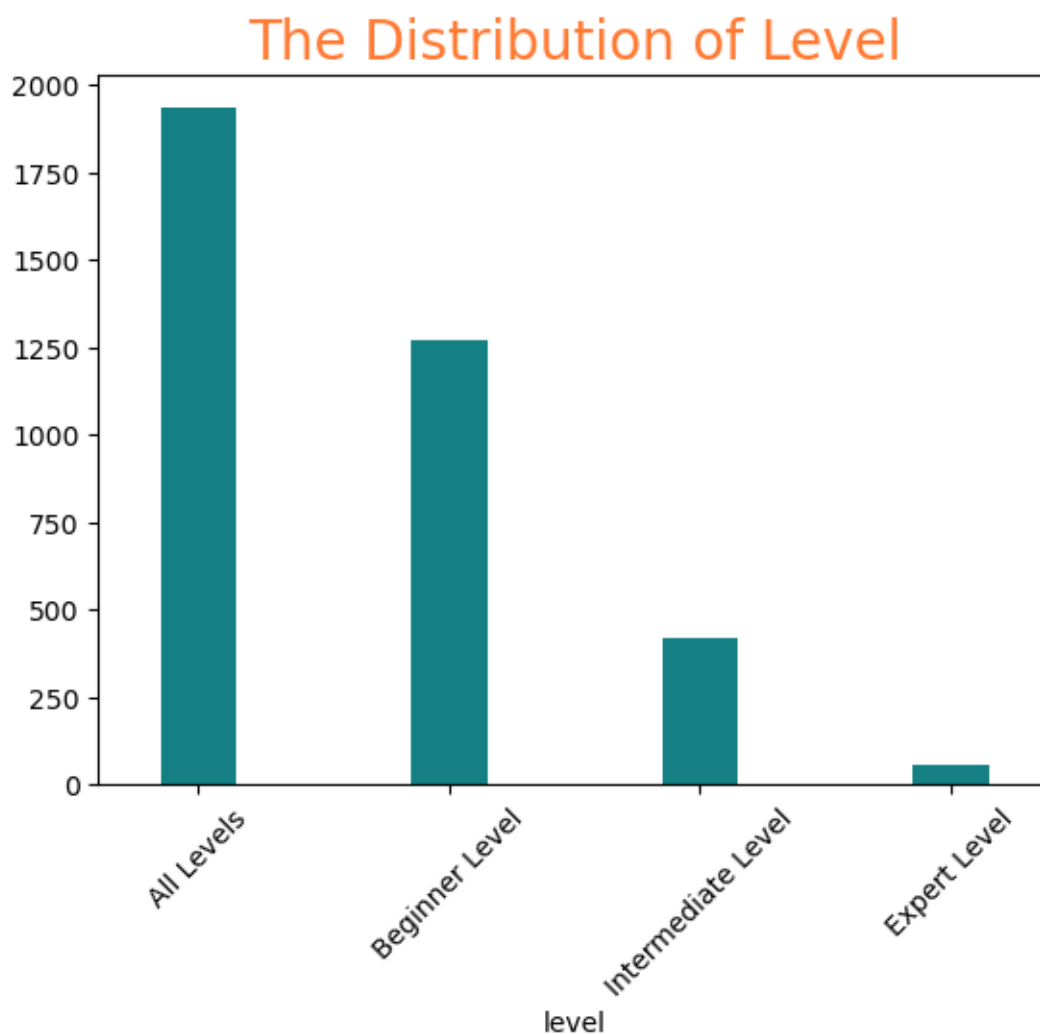
[17]: `df=df[df['level']!='52']`

[18]: `df['level'].value_counts()`
```

```
[18]: level
      All Levels           1932
      Beginner Level       1271
      Intermediate Level    421
      Expert Level           58
      Name: count, dtype: int64
```

```
[19]: df['level'].value_counts().plot(kind='bar',fontsize=10, color='#158083',width=0.
      ⤷3 )
      plt.title('The Distribution of Level', fontsize=20 ,c='#ff7a33')
      plt.xticks(rotation=45)
      plt.show()
```



```
[20]: df.groupby('level')['num_subscribers'].sum()
```

```
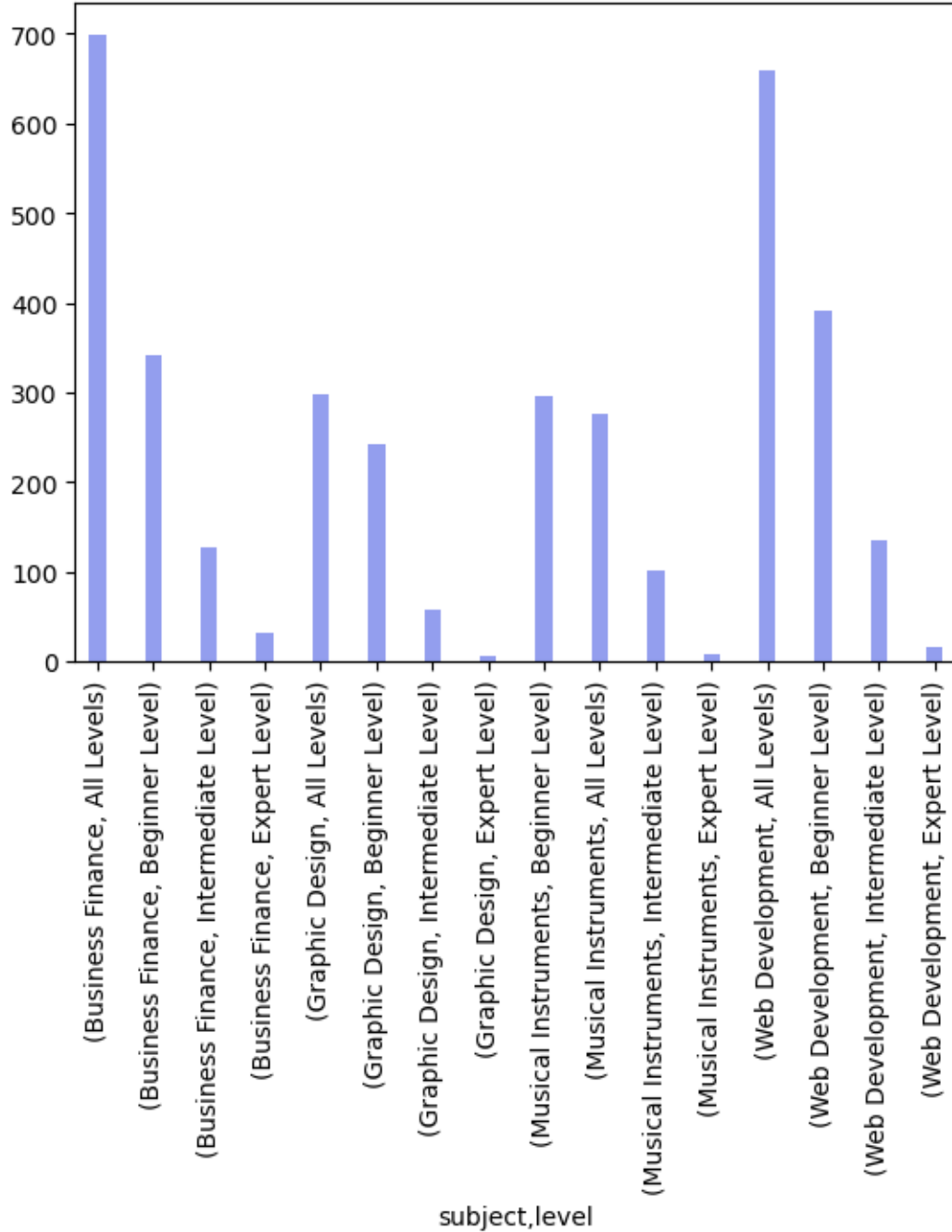[20]: level
      All Levels          6916112
      Beginner Level       4052843
      Expert Level           50196
      Intermediate Level    742005
      Name: num_subscribers, dtype: int64
```

```
[21]: df.groupby('subject')['level'].value_counts()
```

```
[21]: subject              level
      Business Finance     All Levels          699
                           Beginner Level      341
                           Intermediate Level  128
                           Expert Level         31
      Graphic Design       All Levels          298
                           Beginner Level      243
                           Intermediate Level   57
                           Expert Level          5
      Musical Instruments  Beginner Level      296
                           All Levels          276
                           Intermediate Level  101
                           Expert Level          7
      Web Development      All Levels          659
                           Beginner Level      391
                           Intermediate Level  135
                           Expert Level         15
      Name: count, dtype: int64
```

```
[22]: df.groupby('subject')['level'].value_counts().plot(kind='bar',fontsize=10,␣
      ↪color='#939eee',width=0.3 )
      plt.title('The Distribution of Level', fontsize=20 ,c='#6f5b25')
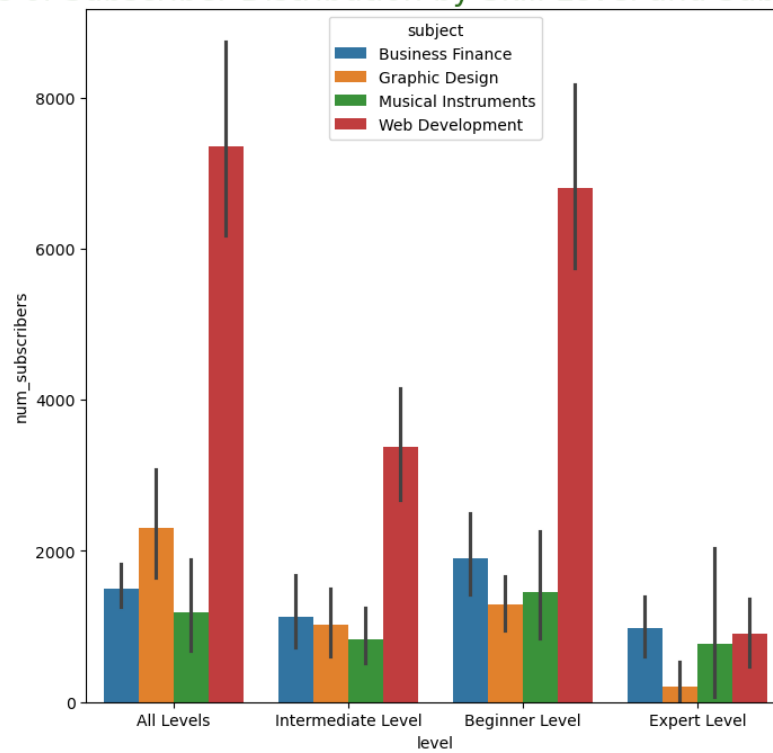
      plt.show()
```

The Distribution of Level


```
[23]: plt.figure(figsize=(8,8))
      sns.barplot(x='level',y='num_subscribers',data=df,hue='subject')
      plt.title('Analysis of Subscriber Distribution by Skill Level and Subject␣
       ↪Category',fontsize=20, c='#2b6c1c')
```

```
plt.show()
```

Analysis of Subscriber Distribution by Skill Level and Subject Category

[24]: `df`

[24]:
```
      Unnamed: 0  course_id  \
0              0    1070968
1              1    1113822
2              2    1006314
3              3    1210588
4              4    1011058
...          ...        ...
3678        3678     775618
3679        3679    1088178
3680        3680     635248
3681        3681     905096
3682        3682     297602


                                    course_title  \
0                   Ultimate Investment Banking Course
1        Complete GST Course & Certification - Grow You…
2        Financial Modeling for Business Analysts and C…
3        Beginner to Pro - Financial Analysis in Excel …
```

```
4                How To Maximize Your Profits Trading Options
…                                                              …
3678   Learn jQuery from Scratch - Master of JavaScri…
3679   How To Design A WordPress Website With No Codi…
3680                          Learn and Build using Polymer
3681   CSS Animations: Create Amazing Effects on Your…
3682   Using MODX CMS to Build Websites: A Beginner's…


                                              url is_paid price  \
0      https://www.udemy.com/ultimate-investment-bank…    True   200
1          https://www.udemy.com/goods-and-services-tax/    True    75
2      https://www.udemy.com/financial-modeling-for-b…    True    45
3      https://www.udemy.com/complete-excel-finance-c…    True    95
4      https://www.udemy.com/how-to-maximize-your-pro…    True   200
…                                                     …   …    …
3678   https://www.udemy.com/easy-jquery-for-beginner…    True   100
3679   https://www.udemy.com/how-to-make-a-wordpress-…    True    25
3680   https://www.udemy.com/learn-and-build-using-po…    True    40
3681   https://www.udemy.com/css-animations-create-am…    True    50
3682   https://www.udemy.com/using-modx-cms-to-build-…    True    45


       num_subscribers  num_reviews  num_lectures               level  \
0                 2147           23            51          All Levels
1                 2792          923           274          All Levels
2                 2174           74            51  Intermediate Level
3                 2451           11            36          All Levels
4                 1276           45            26  Intermediate Level
…                    …            …             …                   …
3678              1040           14            21          All Levels
3679               306            3            42      Beginner Level
3680               513          169            48          All Levels
3681               300           31            38          All Levels
3682               901           36            20      Beginner Level


       content_duration   published_timestamp           subject  \
0             1.5 hours  2017-01-18T20:58:58Z  Business Finance
1              39 hours  2017-03-09T16:34:20Z  Business Finance
2             2.5 hours  2016-12-19T19:26:30Z  Business Finance
3               3 hours  2017-05-30T20:07:24Z  Business Finance
4               2 hours  2016-12-13T14:57:18Z  Business Finance
…                     …                     …                 …
3678            2 hours  2016-06-14T17:36:46Z   Web Development
3679          3.5 hours  2017-03-10T22:24:30Z   Web Development
3680          3.5 hours  2015-12-30T16:41:42Z   Web Development
3681            3 hours  2016-08-11T19:06:15Z   Web Development
3682            2 hours  2014-09-28T19:51:11Z   Web Development
```

```
                                    clean_course_title
0                    Ultimate Investment Banking Course
1        Complete GST Course  Certification  Grow Practice
2        Financial Modeling Business Analysts Consultants
3            Beginner Pro  Financial Analysis Excel 2017
4                        Maximize Profits Trading Options
...                                                    ...
3678       Learn jQuery Scratch  Master JavaScript library
3679                      Design WordPress Website Coding
3680                                   Learn Build Polymer
3681       CSS Animations Create Amazing Effects Website
3682             MODX CMS Build Websites Beginners Guide

[3682 rows x 14 columns]
```

[25]: `df['price']`

[25]:
```
0          200
1           75
2           45
3           95
4          200

          ...
3678       100
3679        25
3680        40
3681        50
3682        45
Name: price, Length: 3682, dtype: object
```

[26]: `df['price'].unique()`

[26]:
```
array(['200', '75', '45', '95', '150', '65', '195', '30', '20', '50',
       '175', '140', '115', '190', '125', '60', '145', '105', '155',
       '185', '180', '120', '25', '160', '40', 'Free', '100', '90', '35',
       '80', '70', '55', '165', '130', '85', '170', '110', '135'],
      dtype=object)
```

[27]: `df['price']=df['price'].str.replace('Free','0',regex=True)`

```
C:\Users\bios\AppData\Local\Temp\ipykernel_12752\3660464292.py:1:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  df['price']=df['price'].str.replace('Free','0',regex=True)
```

```
[28]: df['price'].unique()
```

```
[28]: array(['200', '75', '45', '95', '150', '65', '195', '30', '20', '50',
             '175', '140', '115', '190', '125', '60', '145', '105', '155',
             '185', '180', '120', '25', '160', '40', '0', '100', '90', '35',
             '80', '70', '55', '165', '130', '85', '170', '110', '135'],
            dtype=object)
```

```
[29]: df['price']=df['price'].astype(float)
```

C:\Users\bios\AppData\Local\Temp\ipykernel_12752\3617968478.py:1:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  df['price']=df['price'].astype(float)

```
[30]: df['price'].mean()
```

```
[30]: 66.01032047800109
```

```
[31]: df['price'].describe()
```

```
[31]: count    3682.000000
      mean       66.010320
      std        60.984167
      min         0.000000
      25%        20.000000
      50%        45.000000
      75%        95.000000
      max       200.000000
      Name: price, dtype: float64
```

```
[32]: df['price'].idxmax()
```

```
[32]: 0
```

```
[33]: df.iloc[0]
```

```
[33]: Unnamed: 0                                                    0
      course_id                                               1070968
      course_title                    Ultimate Investment Banking Course
      url                   https://www.udemy.com/ultimate-investment-bank…
      is_paid                                                    True
      price                                                     200.0
      num_subscribers                                            2147
```

```
num_reviews                                                     23
num_lectures                                                    51
level                                                   All Levels
content_duration                                         1.5 hours
published_timestamp                           2017-01-18T20:58:58Z
subject                                            Business Finance
clean_course_title             Ultimate Investment Banking Course
Name: 0, dtype: object
```

[34]: `df['profit']=df['price']*df['num_subscribers']`

```
C:\Users\bios\AppData\Local\Temp\ipykernel_12752\2819405905.py:1:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  df['profit']=df['price']*df['num_subscribers']
```

[35]: `df['profit']`

```
[35]: 0        429400.0
      1        209400.0
      2         97830.0
      3        232845.0
      4        255200.0
                 …
      3678     104000.0
      3679       7650.0
      3680      20520.0
      3681      15000.0
      3682      40545.0
Name: profit, Length: 3682, dtype: float64
```

[36]: `df['profit'].max()`

[36]: 24316800.0

[37]: `df['profit'].idxmax()`

[37]: 3235

[38]: `df.iloc[3234]`

```
[38]: Unnamed: 0                                         3235
      course_id                                        625204
      course_title                   The Web Developer Bootcamp
```

```
url                https://www.udemy.com/the-web-developer-bootcamp/
is_paid                                                         True
price                                                          200.0
num_subscribers                                               121584
num_reviews                                                    27445
num_lectures                                                     342
level                                                     All Levels
content_duration                                            43 hours
published_timestamp                             2015-11-02T21:13:27Z
subject                                              Web Development
clean_course_title                           Web Developer Bootcamp
profit                                                    24316800.0
Name: 3235, dtype: object
```

[39]: `df`

[39]:
```
      Unnamed: 0   course_id  \
0              0     1070968
1              1     1113822
2              2     1006314
3              3     1210588
4              4     1011058
...          ...         ...
3678        3678      775618
3679        3679     1088178
3680        3680      635248
3681        3681      905096
3682        3682      297602

                                          course_title  \
0                  Ultimate Investment Banking Course
1        Complete GST Course & Certification - Grow You…
2        Financial Modeling for Business Analysts and C…
3        Beginner to Pro - Financial Analysis in Excel …
4              How To Maximize Your Profits Trading Options
...                                                 …
3678     Learn jQuery from Scratch - Master of JavaScri…
3679     How To Design A WordPress Website With No Codi…
3680                        Learn and Build using Polymer
3681     CSS Animations: Create Amazing Effects on Your…
3682     Using MODX CMS to Build Websites: A Beginner's…

                                          url is_paid  price  \
0        https://www.udemy.com/ultimate-investment-bank…    True  200.0
1            https://www.udemy.com/goods-and-services-tax/    True   75.0
2        https://www.udemy.com/financial-modeling-for-b…    True   45.0
3        https://www.udemy.com/complete-excel-finance-c…    True   95.0
```

17

```
4       https://www.udemy.com/how-to-maximize-your-pro…     True  200.0
…                                              …                 …      …
3678    https://www.udemy.com/easy-jquery-for-beginner…     True  100.0
3679    https://www.udemy.com/how-to-make-a-wordpress-…     True   25.0
3680    https://www.udemy.com/learn-and-build-using-po…     True   40.0
3681    https://www.udemy.com/css-animations-create-am…     True   50.0
3682    https://www.udemy.com/using-modx-cms-to-build-…     True   45.0

      num_subscribers  num_reviews  num_lectures               level  \
0                2147           23            51          All Levels
1                2792          923           274          All Levels
2                2174           74            51  Intermediate Level
3                2451           11            36          All Levels
4                1276           45            26  Intermediate Level
…                   …            …             …                   …
3678             1040           14            21          All Levels
3679              306            3            42      Beginner Level
3680              513          169            48          All Levels
3681              300           31            38          All Levels
3682              901           36            20      Beginner Level

      content_duration    published_timestamp           subject  \
0            1.5 hours   2017-01-18T20:58:58Z  Business Finance
1             39 hours   2017-03-09T16:34:20Z  Business Finance
2            2.5 hours   2016-12-19T19:26:30Z  Business Finance
3              3 hours   2017-05-30T20:07:24Z  Business Finance
4              2 hours   2016-12-13T14:57:18Z  Business Finance
…                    …                     …                 …
3678           2 hours   2016-06-14T17:36:46Z   Web Development
3679         3.5 hours   2017-03-10T22:24:30Z   Web Development
3680         3.5 hours   2015-12-30T16:41:42Z   Web Development
3681           3 hours   2016-08-11T19:06:15Z   Web Development
3682           2 hours   2014-09-28T19:51:11Z   Web Development

                                 clean_course_title    profit
0                   Ultimate Investment Banking Course  429400.0
1      Complete GST Course  Certification  Grow Practice  209400.0
2      Financial Modeling Business Analysts Consultants   97830.0
3           Beginner Pro  Financial Analysis Excel 2017  232845.0
4                       Maximize Profits Trading Options  255200.0
…                                                   …         …
3678     Learn jQuery Scratch  Master JavaScript library  104000.0
3679               Design WordPress Website Coding     7650.0
3680                          Learn Build Polymer    20520.0
3681     CSS Animations Create Amazing Effects Website   15000.0
3682          MODX CMS Build Websites Beginners Guide    40545.0
```

```
[3682 rows x 15 columns]
```

[40]:
```python
df['is_paid'].value_counts()
```

[40]:
```
is_paid
True      2738
TRUE       634
False      264
FALSE       46
Name: count, dtype: int64
```

[41]:
```python
paid_dic={'TRUE':True,'FALSE':False}
```

[42]:
```python
df['is_paid']=df['is_paid'].map(paid_dic)
df['is_paid'].value_counts()
```

```
C:\Users\bios\AppData\Local\Temp\ipykernel_12752\1306579327.py:1:
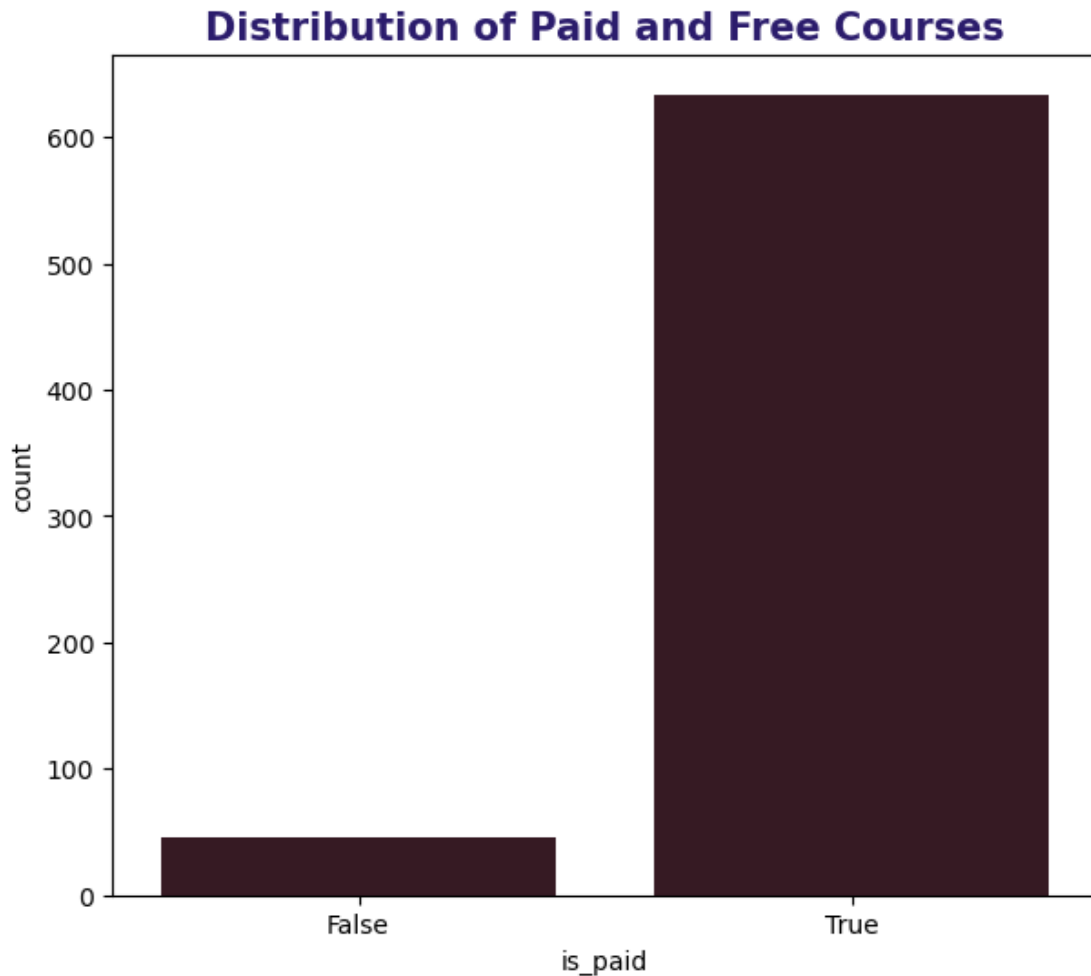SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  df['is_paid']=df['is_paid'].map(paid_dic)
```

[42]:
```
is_paid
True      634
False      46
Name: count, dtype: int64
```

[43]:
```python
plt.figure(figsize=(7,6))
sns.countplot(x=df['is_paid'],color='#3b1521')
plt.title('Distribution of Paid and Free␣
  ↪Courses',fontsize=15,fontweight='bold',c='#2b1c6c')
plt.show()
```

Distribution of Paid and Free Courses

[44]:
```
df['course_title_len']=df['course_title'].str.len()
df['course_title_len'].min()
```

C:\Users\bios\AppData\Local\Temp\ipykernel_12752\3924821331.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  df['course_title_len']=df['course_title'].str.len()

[44]: 6

[45]:
```
df['course_title_len'].max()
```

[45]: 243

```
[46]: df['course_title_len'].idxmax()
```

```
[46]: 2191
```

```
[47]: df.iloc[2190]
```

```
[47]: Unnamed: 0                                          2191
      course_id                                         286898
      course_title          Aprende a tocar el Acordeón 'de oído y con téc…
      url                   https://www.udemy.com/aprende-los-secretos-de-…
      is_paid                                             True
      price                                               75.0
      num_subscribers                                      158
      num_reviews                                           17
      num_lectures                                          59
      level                                           All Levels
      content_duration                                 2.5 hours
      published_timestamp                     2014-07-28T13:55:43Z
      subject                                 Musical Instruments
      clean_course_title    Aprende tocar el Acorden de odo con tcnicahttp…
      profit                                            11850.0
      course_title_len                                     243
      Name: 2191, dtype: object
```

```
[48]: df['published_timestamp']
```

```
[48]: 0       2017-01-18T20:58:58Z
      1       2017-03-09T16:34:20Z
      2       2016-12-19T19:26:30Z
      3       2017-05-30T20:07:24Z
      4       2016-12-13T14:57:18Z
                      …
      3678    2016-06-14T17:36:46Z
      3679    2017-03-10T22:24:30Z
      3680    2015-12-30T16:41:42Z
      3681    2016-08-11T19:06:15Z
      3682    2014-09-28T19:51:11Z
      Name: published_timestamp, Length: 3682, dtype: object
```

```
[49]: df['datetime']=df['published_timestamp'].str.split('T')
```

```
C:\Users\bios\AppData\Local\Temp\ipykernel_12752\963559493.py:1:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
```

```
       df['datetime']=df['published_timestamp'].str.split('T')
```

```
[50]:  df['datetime']=df['datetime'].str.get(0)
```

C:\Users\bios\AppData\Local\Temp\ipykernel_12752\4083654934.py:1:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  df['datetime']=df['datetime'].str.get(0)

```
[51]:  df['datetime']
```

```
[51]:  0          2017-01-18
       1          2017-03-09
       2          2016-12-19
       3          2017-05-30
       4          2016-12-13
                     …
       3678       2016-06-14
       3679       2017-03-10
       3680       2015-12-30
       3681       2016-08-11
       3682       2014-09-28
       Name: datetime, Length: 3682, dtype: object
```

```
[52]:  df['datetime']=pd.to_datetime(df['datetime'],format='%Y-%m-%d')
```

C:\Users\bios\AppData\Local\Temp\ipykernel_12752\3142870180.py:1:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  df['datetime']=pd.to_datetime(df['datetime'],format='%Y-%m-%d')

```
[53]:  df['datetime']
```

```
[53]:  0          2017-01-18
       1          2017-03-09
       2          2016-12-19
       3          2017-05-30
       4          2016-12-13
                     …
       3678       2016-06-14
       3679       2017-03-10
```

```
3680    2015-12-30
3681    2016-08-11
3682    2014-09-28
Name: datetime, Length: 3682, dtype: datetime64[ns]
```

[54]:
```python
df['Year']=df['datetime'].dt.year
df['Year']
```

```
C:\Users\bios\AppData\Local\Temp\ipykernel_12752\3085403916.py:1:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  df['Year']=df['datetime'].dt.year
```

[54]:
```
0       2017
1       2017
2       2016
3       2017
4       2016
        ...
3678    2016
3679    2017
3680    2015
3681    2016
3682    2014
Name: Year, Length: 3682, dtype: int32
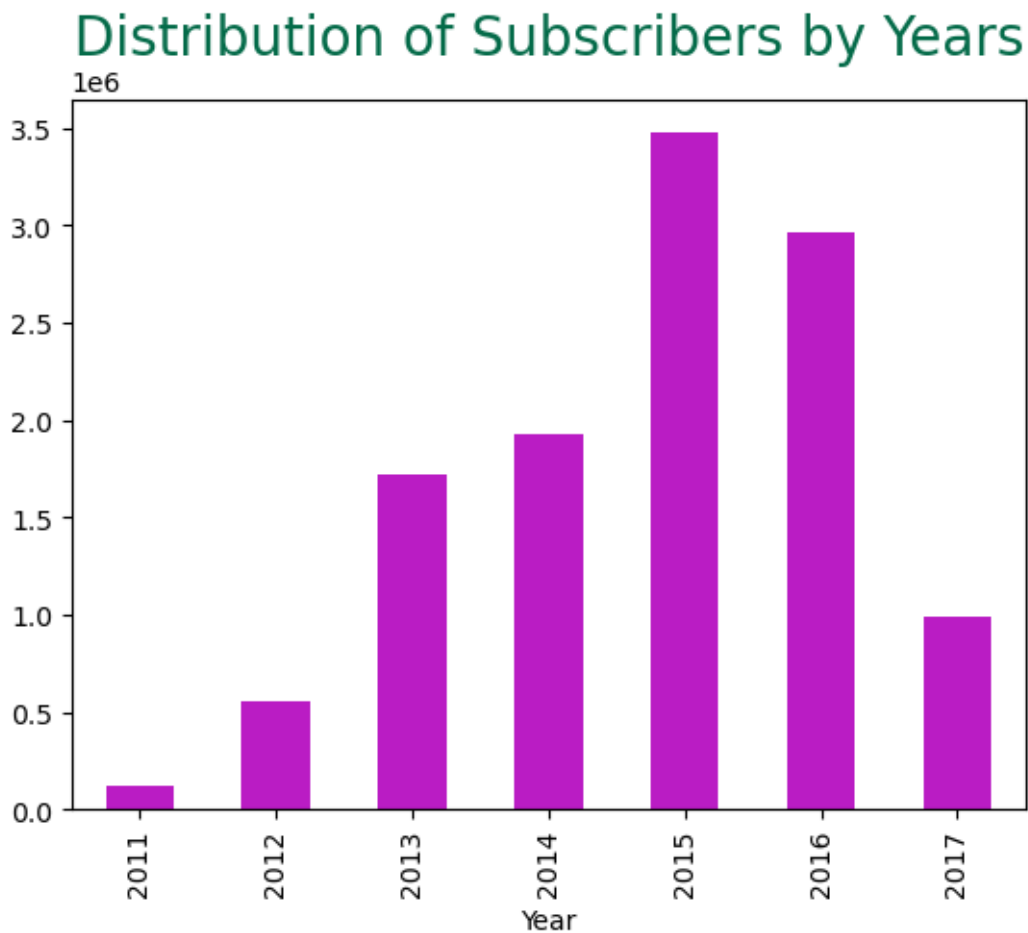```

[55]:
```python
df['Year'].unique()
```

[55]:
```
array([2017, 2016, 2014, 2015, 2013, 2012, 2011])
```

[56]:
```python
df.groupby('Year')['num_subscribers'].sum()
```

[56]:
```
Year
2011     119028
2012     555339
2013    1723438
2014    1930406
2015    3475324
2016    2966644
2017     990977
Name: num_subscribers, dtype: int64
```

[57]:
```python
df.groupby('Year')['num_subscribers'].sum().plot(kind='bar',color='#ba1cc4')
plt.title('Distribution of Subscribers by Years',fontsize=20,color='#046b49')
```

```
plt.show()
```


Distribution of Subscribers by Years

```
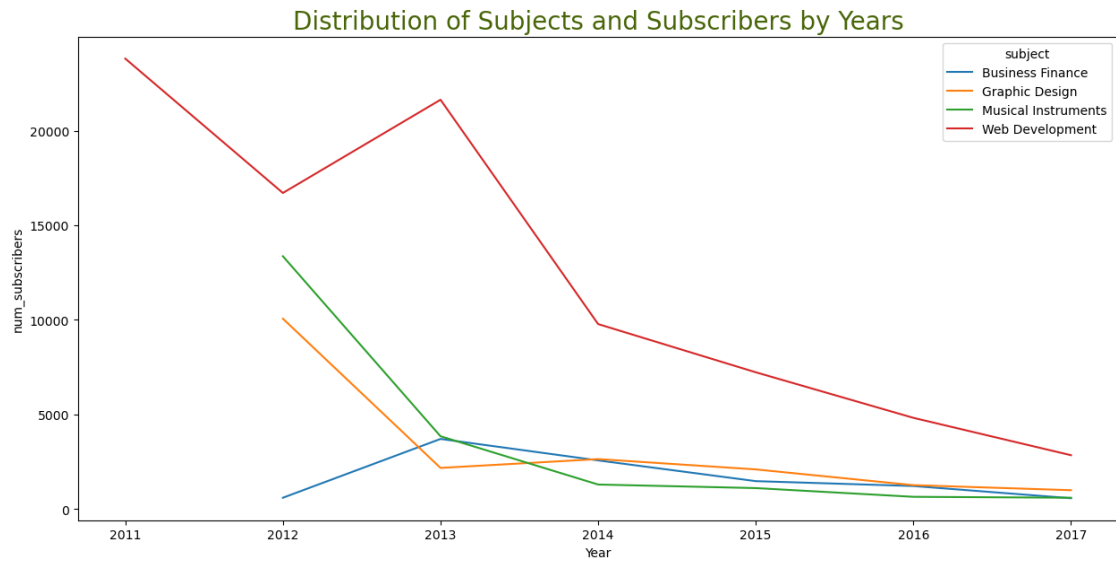[58]: plt.figure(figsize=(15,7))
      sns.lineplot(x='Year',y='num_subscribers',data=df,hue='subject',ci=None)
      plt.title('Distribution of Subjects and Subscribers by␣
       ↪Years',fontsize=20,color='#486608')
      plt.show()
```

C:\Users\bios\AppData\Local\Temp\ipykernel_12752\3664452510.py:2: FutureWarning:

The `ci` parameter is deprecated. Use `errorbar=None` for the same effect.

  sns.lineplot(x='Year',y='num_subscribers',data=df,hue='subject',ci=None)

Distribution of Subjects and Subscribers by Years

[ ]: