

```
In [1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

```
In [2]: data=pd.read_csv('dataset/udemy.csv')
```

```
In [43]: data=pd.read_csv('dataset/udemy.csv',parse_dates=['published_timestamp'])
```

```
In [44]: data.dtypes
```

```
Out[44]: course_id          int64
course_title        object
url                 object
is_paid             bool
price              int64
num_subscribers     int64
num_reviews         int64
num_lectures        int64
level              object
content_duration    float64
published_timestamp  datetime64[ns, UTC]
subject            object
dtype: object
```

Display Top 10 Rows of The Dataset?

```
In [3]: data.head(10)
```

Out[3]:	course_id	course_title	url	is_paid	price	num_subscribers	num_reviews	num_lectures	lev
0	1070968	Ultimate Investment Banking Course	https://www.udemy.com/ultimate-investment-bank...	True	200	2147	23	51	All Leve
1	1113822	Complete GST Course & Certification - Grow You...	https://www.udemy.com/goods-and-services-tax/	True	75	2792	923	274	All Leve
2	1006314	Financial Modeling for Business Analysts and C...	https://www.udemy.com/financial-modeling-for-b...	True	45	2174	74	51	Intermedia Lev
3	1210588	Beginner to Pro - Financial Analysis in Excel ...	https://www.udemy.com/complete-excel-finance-c...	True	95	2451	11	36	All Leve
4	1011058	How To Maximize Your Profits Trading Options	https://www.udemy.com/how-to-maximize-your-pro...	True	200	1276	45	26	Intermedia Lev
5	192870	Trading Penny Stocks: A Guide for All Levels I...	https://www.udemy.com/trading-penny-stocks-a-g...	True	150	9221	138	25	All Leve
6	739964	Investing And Trading For Beginners: Mastering...	https://www.udemy.com/investing-and-trading-fo...	True	65	1540	178	26	Beginn Lev
7	403100	Trading Stock Chart Patterns For Immediate, Ex...	https://www.udemy.com/trading-chart-patterns-f...	True	95	2917	148	23	All Leve
8	476268	Options Trading 3 : Advanced Stock Profit and ...	https://www.udemy.com/day-trading-stock-option...	True	195	5172	34	38	Expert Lev
9	1167710	The Only Investment Strategy You Need For Your...	https://www.udemy.com/the-only-investment-str...	True	200	827	14	15	All Leve

Check Last 5 Rows of The Dataset?

In [4]: `data.tail()`

Out[4]:	course_id	course_title	url	is_paid	price	num_subscribers	num_reviews	num_lectures	level
3673	775618	Learn jQuery from Scratch - Master of JavaScri...	https://www.udemy.com/easy-jquery-for-beginner...	True	100	1040	14	21	All Levels
3674	1088178	How To Design A WordPress Website With No Codi...	https://www.udemy.com/how-to-make-a-wordpress-...	True	25	306	3	42	Beginner Level
3675	635248	Learn and Build using Polymer	https://www.udemy.com/learn-and-build-using-po...	True	40	513	169	48	All Levels
3676	905096	CSS Animations: Create Amazing Effects on Your...	https://www.udemy.com/css-animations-create-am...	True	50	300	31	38	All Levels
3677	297602	Using MODX CMS to Build Websites: A Beginner's...	https://www.udemy.com/using-modx-cms-to-build-...	True	45	901	36	20	Beginner Level

Find Shape of Our Dataset (Number of Rows And Number of Columns)?

```
In [5]: data.shape
```

```
Out[5]: (3678, 12)
```

Getting Information About Our Dataset Like Total Number Rows, Total Number of Columns, Datatypes of Each Column And Memory Requirement?

```
In [6]: data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3678 entries, 0 to 3677
Data columns (total 12 columns):
#   Column                Non-Null Count  Dtype
---  -
0   course_id             3678 non-null   int64
1   course_title          3678 non-null   object
2   url                   3678 non-null   object
3   is_paid               3678 non-null   bool
4   price                 3678 non-null   int64
5   num_subscribers       3678 non-null   int64
6   num_reviews           3678 non-null   int64
7   num_lectures          3678 non-null   int64
8   level                 3678 non-null   object
9   content_duration      3678 non-null   float64
10  published_timestamp    3678 non-null   object
11  subject                3678 non-null   object
dtypes: bool(1), float64(1), int64(5), object(5)
memory usage: 319.8+ KB
```

Check Null Values In The Dataset?

```
In [47]: data.isnull().values.any()
```

```
Out[47]: False
```

```
In [7]: data.isnull().sum()
```

```
Out[7]: 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
        dtype: int64
```

Check For Duplicate Data and Drop Them?

```
In [48]: data.duplicated().any()
```

```
Out[48]: True
```

```
In [8]: data.duplicated().sum()
```

```
Out[8]: 6
```

```
In [9]: data.drop_duplicates(inplace=True)
```

```
In [10]: data.shape
```

```
Out[10]: (3672, 12)
```

What are all different subjects for which Udemy is offering courses ?

```
In [11]: data.head()
```

```
Out[11]:
```

	course_id	course_title	url	is_paid	price	num_subscribers	num_reviews	num_lectures	lev
0	1070968	Ultimate Investment Banking Course	https://www.udemy.com/ultimate-investment-bank...	True	200	2147	23	51	All Leve
1	1113822	Complete GST Course & Certification - Grow You...	https://www.udemy.com/goods-and-services-tax/	True	75	2792	923	274	All Leve
2	1006314	Financial Modeling for Business Analysts and C...	https://www.udemy.com/financial-modeling-for-b...	True	45	2174	74	51	Intermedia Lev
3	1210588	Beginner to Pro - Financial Analysis in Excel ...	https://www.udemy.com/complete-excel-finance-c...	True	95	2451	11	36	All Leve
4	1011058	How To Maximize Your Profits Trading Options	https://www.udemy.com/how-to-maximize-your-pro...	True	200	1276	45	26	Intermedia Lev

```
In [12]: data.subject.value_counts()
```

```
Out[12]: subject
Web Development      1199
Business Finance     1191
Musical Instruments   680
Graphic Design        602
Name: count, dtype: int64
```

Find Out Number of Courses Per Subjects?

```
In [13]: data.subject.value_counts()
```

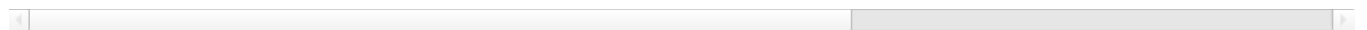
```
Out[13]: subject
Web Development      1199
Business Finance     1191
Musical Instruments   680
Graphic Design        602
Name: count, dtype: int64
```

Show all the courses which are Free of Cost?

```
In [14]: data[data['price']==0]
```

	course_id	course_title	url	is_paid	price	num_subscribers	num_reviews	num_lectures	
95	1148774	Options Trading 101: The Basics	https://www.udemy.com/options-trading-101-the-...	False	0	1514	66	11	E
103	133536	Stock Market Investing for Beginners	https://www.udemy.com/the-beginners-guide-to-t...	False	0	50855	2698	15	E
106	265960	Fundamentals of Forex Trading	https://www.udemy.com/fundamentals-of-forex-tr...	False	0	17160	620	23	
108	923616	Website Investing 101 - Buying & Selling Onlin...	https://www.udemy.com/cash-flow-website-invest...	False	0	6811	151	51	
112	191854	Stock Market Foundations	https://www.udemy.com/how-to-invest-in-the-sto...	False	0	19339	794	9	E
...	
3638	155640	Building a Search Engine in PHP & MySQL	https://www.udemy.com/creating-a-search-engine...	False	0	10110	379	12	
3643	366720	CSS Image filters - The modern web images colo...	https://www.udemy.com/super-awesome-images-wit...	False	0	6315	53	16	
3651	1079078	Drupal 8 Site Building	https://www.udemy.com/drupal-8-site-building/	False	0	1942	23	48	
3665	21386	Beginner Photoshop to HTML5 and CSS3	https://www.udemy.com/psd-html5-css3/	False	0	73110	1716	22	
3668	270976	A how to guide in HTML	https://www.udemy.com/a-how-to-guide-in-html/	False	0	7318	205	8	E

310 rows × 12 columns



```
In [15]: data[data['is_paid']==False]
```

Out[15]:

	course_id	course_title	url	is_paid	price	num_subscribers	num_reviews	num_lectures	
95	1148774	Options Trading 101: The Basics	https://www.udemy.com/options-trading-101-the-...	False	0	1514	66	11	E
103	133536	Stock Market Investing for Beginners	https://www.udemy.com/the-beginners-guide-to-t...	False	0	50855	2698	15	E
106	265960	Fundamentals of Forex Trading	https://www.udemy.com/fundamentals-of-forex-tr...	False	0	17160	620	23	
108	923616	Website Investing 101 - Buying & Selling Online...	https://www.udemy.com/cash-flow-website-invest...	False	0	6811	151	51	
112	191854	Stock Market Foundations	https://www.udemy.com/how-to-invest-in-the-sto...	False	0	19339	794	9	E
...	
3638	155640	Building a Search Engine in PHP & MySQL	https://www.udemy.com/creating-a-search-engine...	False	0	10110	379	12	
3643	366720	CSS Image filters - The modern web images colo...	https://www.udemy.com/super-awesome-images-wit...	False	0	6315	53	16	
3651	1079078	Drupal 8 Site Building	https://www.udemy.com/drupal-8-site-building/	False	0	1942	23	48	
3665	21386	Beginner Photoshop to HTML5 and CSS3	https://www.udemy.com/psd-html5-css3/	False	0	73110	1716	22	
3668	270976	A how to guide in HTML	https://www.udemy.com/a-how-to-guide-in-html/	False	0	7318	205	8	E

310 rows × 12 columns

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

Show all the courses which are Paid?

In [16]: data[data['price']>0]

Out [16]:

	course_id	course_title	url	is_paid	price	num_subscribers	num_reviews	num_lectures	
0	1070968	Ultimate Investment Banking Course	https://www.udemy.com/ultimate-investment-bank...	True	200	2147	23	51	All L
1	1113822	Complete GST Course & Certification - Grow You...	https://www.udemy.com/goods-and-services-tax/	True	75	2792	923	274	All L
2	1006314	Financial Modeling for Business Analysts and C...	https://www.udemy.com/financial-modeling-for-b...	True	45	2174	74	51	Interme
3	1210588	Beginner to Pro - Financial Analysis in Excel ...	https://www.udemy.com/complete-excel-finance-c...	True	95	2451	11	36	All L
4	1011058	How To Maximize Your Profits Trading Options	https://www.udemy.com/how-to-maximize-your-pro...	True	200	1276	45	26	Interme
...	
3673	775618	Learn jQuery from Scratch - Master of JavaScri...	https://www.udemy.com/easy-jquery-for-beginner...	True	100	1040	14	21	All L
3674	1088178	How To Design A WordPress Website With No Codi...	https://www.udemy.com/how-to-make-a-wordpress-...	True	25	306	3	42	Beq
3675	635248	Learn and Build using Polymer	https://www.udemy.com/learn-and-build-using-po...	True	40	513	169	48	All L
3676	905096	CSS Animations: Create Amazing Effects on Your...	https://www.udemy.com/css-animations-create-am...	True	50	300	31	38	All L
3677	297602	Using MODX CMS to Build Websites: A Beginner's...	https://www.udemy.com/using-modx-cms-to-build-...	True	45	901	36	20	Beq

3362 rows × 12 columns

--	--	--

In [17]:

```
data[data['is_paid']==True]
```

Out[17]:		course_id	course_title	url	is_paid	price	num_subscribers	num_reviews	num_lectures	
	0	1070968	Ultimate Investment Banking Course	https://www.udemy.com/ultimate-investment-bank...	True	200	2147	23	51	All L
	1	1113822	Complete GST Course & Certification - Grow You...	https://www.udemy.com/goods-and-services-tax/	True	75	2792	923	274	All L
	2	1006314	Financial Modeling for Business Analysts and C...	https://www.udemy.com/financial-modeling-for-b...	True	45	2174	74	51	Interme
	3	1210588	Beginner to Pro - Financial Analysis in Excel ...	https://www.udemy.com/complete-excel-finance-c...	True	95	2451	11	36	All L
	4	1011058	How To Maximize Your Profits Trading Options	https://www.udemy.com/how-to-maximize-your-pro...	True	200	1276	45	26	Interme
	
	3673	775618	Learn jQuery from Scratch - Master of JavaScri...	https://www.udemy.com/easy-jquery-for-beginner...	True	100	1040	14	21	All L
	3674	1088178	How To Design A WordPress Website With No Codi...	https://www.udemy.com/how-to-make-a-wordpress-...	True	25	306	3	42	Beg
	3675	635248	Learn and Build using Polymer	https://www.udemy.com/learn-and-build-using-po...	True	40	513	169	48	All L
	3676	905096	CSS Animations: Create Amazing Effects on Your...	https://www.udemy.com/css-animations-create-am...	True	50	300	31	38	All L
	3677	297602	Using MODX CMS to Build Websites: A Beginner's...	https://www.udemy.com/using-modx-cms-to-build-...	True	45	901	36	20	Beg
3362 rows × 12 columns										
<div> <div></div> <div></div> </div>										

Display The Count of Paid and Free Courses?

```
In [18]: data[data['price']>0].shape[0]
```

```
Out[18]: 3362
```

```
In [19]: data[data['price']==0].shape[0]
```

```
Out[19]: 310
```

For Which Levels, Udemy Courses Providing The Courses?

```
In [20]: data.level.value_counts()
```



```
Out[20]: level
All Levels      1925
Beginner Level  1268
Intermediate Level  421
Expert Level    58
Name: count, dtype: int64
```

Which are Top Selling Courses ?

```
In [21]: data.num_subscribers.max()
```

```
Out[21]: 268923
```

```
In [22]: data.sort_values(by='num_subscribers',ascending=False).head(5)
```

```
Out[22]:
```

	course_id	course_title	url	is_paid	price	num_subscribers	num_reviews	num_lectures	level
	2827	41295 Learn HTML5 Programming From Scratch	https://www.udemy.com/learn-html5-programming-...	False	0	268923	8629	45	All Levels
	3032	59014 Coding for Entrepreneurs Basic	https://www.udemy.com/coding-for-entrepreneurs...	False	0	161029	279	27	Beginner Level
	3230	625204 The Web Developer Bootcamp	https://www.udemy.com/the-web-developer-bootcamp/	True	200	121584	27445	342	All Levels
	2783	173548 Build Your First Website in 1 Week with HTML5 ...	https://www.udemy.com/build-your-first-website...	False	0	120291	5924	30	Beginner Level
	3232	764164 The Complete Web Developer Course 2.0	https://www.udemy.com/the-complete-web-develop...	True	200	114512	22412	304	All Levels

Which are Least Selling Courses ?

```
In [23]: data.sort_values(by='num_subscribers').head()
```

```
Out[23]:
```

	course_id	course_title	url	is_paid	price	num_subscribers	num_reviews	num_lectures	level
	458	1258666 Financial Statement Auditing Cycles	https://www.udemy.com/financial-statement-audi...	True	50	0	0	9	Intermediate
	2400	1026762 Coaching Course: Rhythm Instruction for Music ...	https://www.udemy.com/coaching-bundle-rhythm-i...	True	125	0	0	7	All Levels
	2401	1234656 4 Week Rhythm Mastery	https://www.udemy.com/4-week-rhythm-mastery/	True	20	0	0	28	Beginner
	1236	1177156 Como Criar Logos de Sucesso - Start Designer	https://www.udemy.com/como-criar-logos-de-suce...	True	25	0	0	29	Beginner
	1237	177614 Super Textos con Photoshop	https://www.udemy.com/super-textos-con-photoshop/	True	20	0	0	8	Beginner

```
In [24]: data[data['num_subscribers']==0].shape[0]
```

```
Out[24]: 65
```

Show all courses of Graphic Design where the price is below 100 ?

```
In [25]: data[(data['subject']=='Graphic Design') & (data['price']<100)]
```

Out[25]:

	course_id	course_title	url	is_paid	price	num_subscribers	num_reviews	num_lect
1196	1197206	Illustrator CC MasterClass	https://www.udemy.com/illustrator-cc-masterclass/	True	95	462	50	
1198	1219520	Adobe Illustrator T-Shirt Design for Merch by ...	https://www.udemy.com/merchbyamazondesign/	True	20	390	44	
1202	28556	Discover How to Draw and paint Comics	https://www.udemy.com/learn-to-draw-and-paint/	True	65	8901	424	
1203	317278	Typographic Logos: Typography and Lettering fo...	https://www.udemy.com/typographic-logos-typogr...	True	25	4235	427	
1204	573064	Photoshop in Ease: Create World Amazing Graph...	https://www.udemy.com/photoshop-knights-become...	True	20	14440	182	
...
1793	456388	Adobe Photoshop CC The Essential Guide	https://www.udemy.com/adobe-photoshop-cc-the-e...	True	20	254	5	
1794	955242	Learn to Composite a 2D Action Shot in Photoshop	https://www.udemy.com/3dmotive-learn-to-compos...	True	20	40	2	
1795	496430	Infographic Design: How To Create Your Own Inf...	https://www.udemy.com/infographic-design-how-t...	True	20	84	8	
1796	793246	Autodesk Inventor 2016 : Complete Guide	https://www.udemy.com/learn-autodesk-inventor-...	True	20	23	4	
1797	547190	Amazing Graphic Design for Beginners	https://www.udemy.com/amazing-graphic-design-f...	True	20	32	10	

484 rows × 12 columns

Find Most Popular Course Title?

In [26]:

data.sort_values(by='num_subscribers',ascending=False).head(1)

Out[26]:

	course_id	course_title	url	is_paid	price	num_subscribers	num_reviews	num_lectures	level	c
2827	41295	Learn HTML5 Programming From Scratch	https://www.udemy.com/learn-html5-programming-...	False	0	268923	8629	45	All Levels	

Display 10 Most Popular Courses As Per Number of Subscribers?

In [27]:

data.sort_values(by='num_subscribers',ascending=False).head(10)

Out[27]:

	course_id	course_title	url	is_paid	price	num_subscribers	num_reviews	num_lectures	
	2827	41295	Learn HTML5 Programming From Scratch	https://www.udemy.com/learn-html5-programming-...	False	0	268923	8629	45
	3032	59014	Coding for Entrepreneurs Basic	https://www.udemy.com/coding-for-entrepreneurs...	False	0	161029	279	27
	3230	625204	The Web Developer Bootcamp	https://www.udemy.com/the-web-developer-bootcamp/	True	200	121584	27445	342
	2783	173548	Build Your First Website in 1 Week with HTML5 ...	https://www.udemy.com/build-your-first-website...	False	0	120291	5924	30
	3232	764164	The Complete Web Developer Course 2.0	https://www.udemy.com/the-complete-web-develop...	True	200	114512	22412	304
	1896	19421	Free Beginner Electric Guitar Lessons	https://www.udemy.com/beginner-electric-guitar...	False	0	101154	1042	95
	2589	473160	Web Design for Web Developers: Build Beautiful...	https://www.udemy.com/web-design-secrets/	False	0	98867	6512	20
	2619	94430	Learn Javascript & JQuery From Scratch	https://www.udemy.com/learn-javascript-jquery-...	True	30	84897	2685	10
	3289	130064	Practical PHP: Master the Basics and Code Dyna...	https://www.udemy.com/code-dynamic-websites/	False	0	83737	4598	45
	3247	364426	JavaScript: Understanding the Weird Parts	https://www.udemy.com/understand-javascript/	True	175	79612	16976	85

Find Total Number of Courses Related To Python?

```
In [28]: data[data.course_title.str.contains('python',case=False)].shape[0]
```

Out[28]: 29

Display 10 Most Popular Python Courses As Per Number of Subscribers?

```
In [29]: a=data[data.course_title.str.contains('python',case=False)]
```

```
In [30]: a.sort_values(by='num_subscribers',ascending=False)[:10]
```

Out[30]:	course_id	course_title	url	is_paid	price	num_subscribers	num_reviews	num_lectures	
	2497	16646 Web Programming with Python	https://www.udemy.com/web-programming-with-pyt...	True	50	35267	217	53	
	2570	47963 Coding for Entrepreneurs: Learn Python, Django...	https://www.udemy.com/coding-for-entrepreneurs/	True	195	23412	799	251	
	2528	391546 Learn Python and Django: Payment Processing	https://www.udemy.com/learn-django-code-accept...	True	70	17714	198	23	
	3200	822444 Python and Django Full Stack Web Developer Boo...	https://www.udemy.com/python-and-django-full-s...	True	200	11832	1883	191	
	3507	394832 Fun and creative web engineering with Python a...	https://www.udemy.com/web-engineering-with-pyt...	False	0	10917	319	25	
	2553	938560 The Complete Ethical Hacking Course 2.0: Pytho...	https://www.udemy.com/penetration-testing-ethi...	True	195	7827	268	66	
	3138	631128 Complete Python Web Course: Build 8 Python Web...	https://www.udemy.com/the-complete-python-web-...	True	110	7489	941	173	
	3197	76052 Try Django 1.9 Build a Blog and Learn Python...	https://www.udemy.com/try-django/	True	50	7407	172	147	
	2681	477702 Python for Beginners: Python Programming Langu...	https://www.udemy.com/python-course/	True	150	6153	125	84	
	3203	970600 REST APIs with Flask and Python	https://www.udemy.com/rest-api-flask-and-python/	True	110	5151	737	115	Int

What are courses that published in year 2015 ?

```
In [31]: b=data.published_timestamp.str.split('T')

In [32]: data[['Date','Time']]=data.published_timestamp.str.split('T',expand=True)

In [33]: data['NEW_TIME']=data.Time.str.rstrip('Z')

In [34]: data.drop(columns='Time',inplace=True)

In [35]: data.rename(columns={'NEW_TIME':'Time'},inplace=True)

In [36]: data['Only_Year']=pd.to_datetime(data['Date'],format='mixed')

In [37]: data.head(2)
```

```
Out[37]:
```

	course_id	course_title	url	is_paid	price	num_subscribers	num_reviews	num_lectures	level	co
0	1070968	Ultimate Investment Banking Course	https://www.udemy.com/ultimate-investment-bank...	True	200	2147	23	51	All Levels	
1	1113822	Complete GST Course & Certification - Grow You...	https://www.udemy.com/goods-and-services-tax/	True	75	2792	923	274	All Levels	

```
In [38]: data['Only_Time']=pd.to_datetime(data['Time'],format='mixed')
```

```
In [39]: data['Only_Year'].dt.year.value_counts()
```

```
Out[39]: Only_Year
2016    1204
2015    1014
2017     713
2014     490
2013     201
2012      45
2011       5
Name: count, dtype: int64
```

What are the Max. Number of Subscribers for Each Level of courses ?

```
In [41]: data.groupby('level')['num_subscribers'].max()
```

```
Out[41]: level
All Levels          268923
Beginner Level      161029
Expert Level         5172
Intermediate Level   29167
Name: num_subscribers, dtype: int64
```

```
In [ ]:
```

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js