



# UDEMY COURSE ANALYSIS USING PYTHON

**This dataset contains all course data for all subjects.**

**We will analyze this data for all subjects. Some of the questions we need to analyze.**

- What are all different subjects for which Udemy is offering courses ?
- Which subject has the maximum number of courses.
- Show all the courses which are Free of Cost.
- Show all the courses which are Paid.
- Which are Top Selling Courses ?
- Which are Least Selling Courses ?
- Show all courses of Graphic Design where the price is below 100 ?
- List out all the courses that are related to 'Python'.
- What are courses that were published in the year 2015 ?

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

In [1]: *# Install all required libraries and import the required data*

```
import numpy as np
import pandas as pd
```

In [2]: `df=pd.read_csv('Udemy Courses.csv')`

In [3]: `df.head(1)`

Out[3]:

	course_id	course_title	is_paid	price	num_subscribers	num_reviews	num_lectures
0	288942	#1 Piano Hand Coordination: Play 10th Ballad i...	True	35	3137	18	68

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

In [4]: `df.subject.unique()`

Out[4]: array(['Musical Instruments', 'Business Finance', 'Graphic Design', 'Web Development'], dtype=object)

## 2. Which subject has the maximum number of courses?

In [6]: `df.subject.value_counts()`

Out[6]:


subject	
Web Development	1200
Business Finance	1199
Musical Instruments	680
Graphic Design	603
Name: count, dtype: int64	

## 3. Show all the courses which are free of cost?

In [7]: `df.head(2)`

Out[7]:

	course_id	course_title	is_paid	price	num_subscribers	num_reviews	num_lectures
0	288942	#1 Piano Hand Coordination: Play 10th Ballad i...	True	35	3137	18	68
1	1170074	#10 Hand Coordination - Transfer Chord Ballad ...	True	75	1593	1	41



In [8]:

```
df[df.is_paid == False]
```

Out[8]:

	course_id	course_title	is_paid	price	num_subscribers	num_reviews	num Lec
<b>41</b>	286070	5 lecciones que todo guitarrista debe tomar	False	Free	4452	263	
<b>49</b>	696630	7 Ways A Beginner Guitarist Can Sound Better, ...	False	Free	4529	193	
<b>57</b>	955914	A beginner's guide to fingerpicking and strumm...	False	Free	3481	29	
<b>67</b>	270976	A how to guide in HTML	False	Free	7318	205	
<b>74</b>	1214144	¡Triunfar en La Bolsa de Valores No Requiere d...	False	Free	338	7	
...	...	...	...	...	...	...	...
<b>3620</b>	552598	WordPress in under an hour - Create/manage you...	False	Free	6474	86	
<b>3638</b>	1068590	Wordpress Website Mastery 2017	False	Free	5181	64	
<b>3652</b>	752706	Write quicker HTML5 and CSS 3; productivity ha...	False	Free	13016	286	
<b>3674</b>	1257976	Your First 10 Guitar Lessons - Learn how to pl...	False	Free	924	4	
<b>3680</b>	1225194	Zoho Books Gestion Financière d'Entreprise pas...	False	Free	229	0	

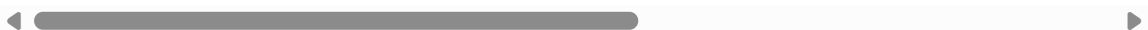
310 rows × 11 columns

In [11]: `df[df.is_paid == False]`

Out[11]:

	course_id	course_title	is_paid	price	num_subscribers	num_reviews	num Lec
<b>41</b>	286070	5 lecciones que todo guitarrista debe tomar	False	Free	4452	263	
<b>49</b>	696630	7 Ways A Beginner Guitarist Can Sound Better, ...	False	Free	4529	193	
<b>57</b>	955914	A beginner's guide to fingerpicking and strumm...	False	Free	3481	29	
<b>67</b>	270976	A how to guide in HTML	False	Free	7318	205	
<b>74</b>	1214144	¡Triunfar en La Bolsa de Valores No Requiere d...	False	Free	338	7	
...	...	...	...	...	...	...	
<b>3620</b>	552598	WordPress in under an hour - Create/manage you...	False	Free	6474	86	
<b>3638</b>	1068590	Wordpress Website Mastery 2017	False	Free	5181	64	
<b>3652</b>	752706	Write quicker HTML5 and CSS 3; productivity ha...	False	Free	13016	286	
<b>3674</b>	1257976	Your First 10 Guitar Lessons - Learn how to pl...	False	Free	924	4	
<b>3680</b>	1225194	Zoho Books Gestion Financière d'Entreprise pas...	False	Free	229	0	

310 rows × 11 columns



## 4. Show all the courses which are paid?

```
In [13]: df[df.is_paid == True]
```

Out[13]:

	course_id	course_title	is_paid	price	num_subscribers	num_reviews	num_lectu
<b>0</b>	288942	#1 Piano Hand Coordination: Play 10th Ballad i...	True	35	3137	18	
<b>1</b>	1170074	#10 Hand Coordination - Transfer Chord Ballad ...	True	75	1593	1	
<b>2</b>	1193886	#12 Hand Coordination: Let your Hands dance wi...	True	75	482	1	
<b>3</b>	1116700	#4 Piano Hand Coordination: Fun Piano Runs in ...	True	75	850	3	
<b>4</b>	1120410	#5 Piano Hand Coordination: Piano Runs in 2 ...	True	75	940	3	
...	...	...	...	...	...	...	...
<b>3676</b>	498488	Your First Successful Forex Trades - With Case...	True	200	1079	34	
<b>3677</b>	328960	Your Own Site in 45 Min: The Complete Wordpres...	True	120	1566	29	
<b>3678</b>	552700	Your Second Course on Piano: Two Handed Playing	True	70	1018	12	
<b>3679</b>	631754	Zend Framework 2: Learn the PHP framework ZF2 ...	True	40	723	130	
<b>3681</b>	964478	Zombie Apocalypse Photoshop Actions	True	50	12	1	

3372 rows × 11 columns

## 5. Which are top selling courses?

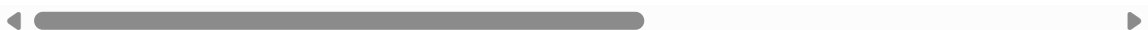
```
In [15]: df.sort_values('num_subscribers', ascending = False)
```



Out[15]:

	course_id	course_title	is_paid	price	num_subscribers	num_reviews	num_lectu
<b>2230</b>	41295	Learn HTML5 Programming From Scratch	False	Free	268923	8629	
<b>776</b>	59014	Coding for Entrepreneurs Basic	False	Free	161029	279	
<b>3385</b>	625204	The Web Developer Bootcamp	True	200	121584	27445	
<b>640</b>	173548	Build Your First Website in 1 Week with HTML5 ...	False	Free	120291	5924	
<b>3316</b>	764164	The Complete Web Developer Course 2.0	True	200	114512	22412	
...	...	...	...	...	...	...	...
<b>1390</b>	1136038	Free Ways to Create Passive Income Online	True	25	0	0	
<b>1909</b>	1157298	Introduction to Forex Trading Business For Beg...	True	20	0	0	
<b>1910</b>	1157298	Introduction to Forex Trading Business For Beg...	True	20	0	0	
<b>1325</b>	1187084	Forex how traders beat the markest with little...	True	100	0	0	
<b>44</b>	1231958	60 Minutes to Fundamental Accounting Skills	True	75	0	0	

3682 rows × 11 columns



## 6. Which are least selling courses?

In [16]: `df.sort_values('num_subscribers', ascending = True)`

Out[16]:

	course_id	course_title	is_paid	price	num_subscribers	num_reviews	num_lectu
--	-----------	--------------	---------	-------	-----------------	-------------	-----------

98	1219820	Accounting - The 6 most important rules you mu...	True	20	0	0	
2350	371900	Learn to create Text fire effect in Adobe Phot...	True	20	0	0	
3221	1057040	Teeter-Totter Accounting	True	90	0	0	
649	1233314	Building a Balanced Scorecard	True	50	0	0	
954	1169196	Credit concepts and checklist for bankers, cre...	True	20	0	0	
...	...	...	...	...	...	...	...
3316	764164	The Complete Web Developer Course 2.0	True	200	114512	22412	
640	173548	Build Your First Website in 1 Week with HTML5 ...	False	Free	120291	5924	
3385	625204	The Web Developer Bootcamp	True	200	121584	27445	
776	59014	Coding for Entrepreneurs Basic	False	Free	161029	279	
2230	41295	Learn HTML5 Programming From Scratch	False	Free	268923	8629	

3682 rows × 11 columns



## 7. Show all courses of graphic design where the price is below 100

```
In [17]: df[(df.subject == 'Graphic Design') & (df.price < '100')]
```

```
Out[17]:
```

course_id	course_title	is_paid	price	num_subscribers	num_reviews	num_lectures	l
-----------	--------------	---------	-------	-----------------	-------------	--------------	---



## 8. Show all the courses of web development where the price is above 50?

```
In [19]: df[(df.subject == 'Web Development') & (df.price > '50')]
```

Out[19]:

	course_id	course_title	is_paid	price	num_subscribers	num_reviews	num_lectu
<b>67</b>	270976	A how to guide in HTML	False	Free	7318	205	
<b>145</b>	425084	Administering Microsoft SQL Server 2012 Databa...	True	60	586	65	
<b>193</b>	990532	Advance Unix : File Processing: AWK   SED   GR...	True	75	578	40	
<b>209</b>	212398	Advanced Node.js Development	True	55	2565	32	
<b>222</b>	866384	Advanced WordPress Topics: Post Types Taxonomi...	True	95	301	44	
...	...	...	...	...	...	...	...
<b>3631</b>	339086	WordPress Security - The Ultimate Security Guide	True	60	2267	17	
<b>3638</b>	1068590	Wordpress Website Mastery 2017	False	Free	5181	64	
<b>3651</b>	424864	Write PHP Like a Pro: Build a PHP MVC Framewor...	True	70	1703	435	
<b>3652</b>	752706	Write quicker HTML5 and CSS 3; productivity ha...	False	Free	13016	286	
<b>3661</b>	688872	XML For Absolute Beginners: Learn to create yo...	True	75	1945	58	

314 rows × 11 columns



## 9. List out all the courses that are related with Python?

```
In [20]: len(df[df.course_title.str.contains('Python', case = True)])
```

```
Out[20]: 29
```

## 10. What are courses that published in year 2015?

```
In [21]: df.dtypes
```

```
Out[21]: course_id          int64
course_title      object
is_paid           bool
price            object
num_subscribers   int64
num_reviews       int64
num_lectures      int64
level            object
content_duration  object
published_timestamp object
subject          object
dtype: object
```

- We need to convert "published\_timestamp" to datetime

```
In [22]: df['published_timestamp'] = pd.to_datetime(df.published_timestamp) '''this func
```

```
In [23]: df.dtypes
```

```
Out[23]: course_id          int64
course_title      object
is_paid           bool
price            object
num_subscribers   int64
num_reviews       int64
num_lectures      int64
level            object
content_duration  object
published_timestamp datetime64[ns, UTC]
subject          object
dtype: object
```

```
In [24]: df['Year'] = df['published_timestamp'].dt.year
```

```
In [25]: df[df.Year == 2015]
```

Out[25]:

	course_id	course_title	is_paid	price	num_subscribers	num_reviews	num_lectures
<b>10</b>	591880	1 - Concepts of Statistics For Beginners Step ...	True	200	273	4	1
<b>16</b>	302450	10 Numbers Every Business Owner Should Know	True	20	13	1	1
<b>19</b>	384928	101 Blues riffs - learn how the harmonica supe...	True	200	1350	65	1
<b>23</b>	550842	16 Guitar Chords to Jam With - (Beginner - Int...	True	20	1224	19	2
<b>26</b>	486240	2 Easy Steps To Investment And Avoiding Traps	True	20	828	1	2
...	...	...	...	...	...	...	...
<b>3675</b>	551170	Your First Course on Piano	True	70	2002	34	3
<b>3676</b>	498488	Your First Successful Forex Trades - With Case...	True	200	1079	34	1
<b>3677</b>	328960	Your Own Site in 45 Min: The Complete Wordpres...	True	120	1566	29	3
<b>3678</b>	552700	Your Second Course on Piano: Two Handed Playing	True	70	1018	12	2

	course_id	course_title	is_paid	price	num_subscribers	num_reviews	num_lectur
		Zend Framework 2: Learn the PHP framework ZF2 ...	True	40	723	130	:

1014 rows × 12 columns

- There are 1014 courses which are published in 2015

## 10. What are the maximum number of subscribers for each level of courses?

```
In [26]: df.level.unique()
```

```
Out[26]: array(['All Levels', 'Intermediate Level', 'Beginner Level',
               'Expert Level'], dtype=object)
```

```
In [27]: # df.groupby('level') - this is the syntax to create a group in any column.
         # ['num_subscribers'] - this is the syntax to work on second column.
```

```
df.groupby('level')['num_subscribers'].max()
```

```
Out[27]: level
All Levels      268923
Beginner Level  161029
Expert Level     5172
Intermediate Level  29167
Name: num_subscribers, dtype: int64
```

```
In [28]: df.groupby('level').max()
```

Out[28]:

	course_id	course_title	is_paid	price	num_subscribers	num_reviews	nur
level							
All Levels	1277924	6時間でインターバンク市場を攻略! 最短距離でトレード基礎力	True	Free	268923	27445	
Beginner Level	1282064	7日でマスター ビギナー向け Adobe Illustrator ときどきセミナー	True	Free	161029	5924	
Expert Level	1275790	[Value Investing] Where White People Keep Their Money	True	95	5172	249	
Intermediate Level	1276182	株式投資で本当のテクニカル分析ができるようになる	True	Free	29167	3326	



## reference -

<https://www.youtube.com/watch?v=AO5uhxa1R84&list=PLy3lFw0OTlutzXFVwttrtaRGEEyLEdnp&index=9>

In [ ]: