

CS112 : Introduction to Python Programming

Week 13: lab practice

Practice 1

- Read the *sales_data.csv* file
- Group the data by region, calculate the total sales and total revenue for each region, and sort the results by total sales in descending order.

	total_sales	total_revenue
region		
South	550	11000
West	550	11000
North	400	8000
East	190	3800

Practice 2

- Read the *marketing_data.csv* file.
- Merge the grouped data from Practice 1 (sales_summary) with the marketing_data.csv file on the region column to create a larger table that includes total sales, total revenue, and marketing budget for each region.

	region	total_sales	total_revenue	marketing_budget
0	South	550	11000	40000
1	West	550	11000	60000
2	North	400	8000	50000
3	East	190	3800	30000

Practice 3

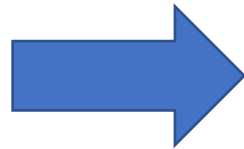
- Calculate the mean device price for each company:

Hint: `groupby()` can not only group by a single column.

The DataFrame can be generated by copy and paste the following:

```
df = pd.DataFrame({"price": [200, 300, 400, 500, 300, 600, 700, 900],  
                  "brand": ["apple", "google", "apple", "apple", "google", "apple", "google", "google"],  
                  "device": ["phone", "phone", "computer", "phone", "computer", "computer", "phone", "computer"]})
```

	price	brand	device
0	200	apple	phone
1	300	google	phone
2	400	apple	computer
3	500	apple	phone
4	300	google	computer
5	600	apple	computer
6	700	google	phone
7	900	google	computer

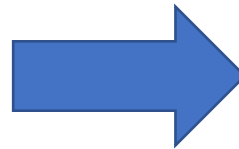


	brand	device	mean_price
0	apple	computer	500.0
1	apple	phone	350.0
2	google	computer	600.0
3	google	phone	500.0

Practice 4

- Read the *stocks_data.csv* as a DataFrame. Pivot the table using *date* as index, *symbol* as columns and *close* as values.

	date	symbol	open	high	low	close	volume
0	2019-03-01	AMZN	1655.13	1674.26	1651.00	1671.73	4974877
1	2019-03-04	AMZN	1685.00	1709.43	1674.36	1696.17	6167358
2	2019-03-05	AMZN	1702.95	1707.80	1689.01	1692.43	3681522
3	2019-03-06	AMZN	1695.97	1697.75	1668.28	1668.95	3996001
4	2019-03-07	AMZN	1667.37	1669.75	1620.51	1625.95	4957017
5	2019-03-01	AAPL	174.28	175.15	172.89	174.97	25886167
6	2019-03-04	AAPL	175.69	177.75	173.97	175.85	27436203
7	2019-03-05	AAPL	175.94	176.00	174.54	175.53	19737419
8	2019-03-06	AAPL	174.67	175.49	173.94	174.52	20810384
9	2019-03-07	AAPL	173.87	174.44	172.02	172.50	24796374



symbol	AAPL	AMZN	GOOG
date			
2019-03-01	174.97	1671.73	1140.99
2019-03-04	175.85	1696.17	1147.80
2019-03-05	175.53	1692.43	1162.03
2019-03-06	174.52	1668.95	1157.86
2019-03-07	172.50	1625.95	1143.30

Practice 5

- Concat the following DataFrames:

	col1	col2	col3
0	1	a	a1
1	2	b	b2
2	3	c	c3

	col1	col2	col3
0	4	d	d4
1	5	e	e5
2	6	f	f6

	col1	col2	col3
0	7	g	g7
1	8	h	h2
2	9	i	i3



	col1	col2	col3
0	1	a	a1
1	2	b	b2
2	3	c	c3
3	4	d	d4
4	5	e	e5
5	6	f	f6
6	7	g	g7
7	8	h	h2
8	9	i	i3

```
df1 = pd.DataFrame({'col1': [1, 2, 3], 'col2': ['a', 'b', 'c'], 'col3': ['a1', 'b2', 'c3']})
df2 = pd.DataFrame({'col1': [4, 5, 6], 'col2': ['d', 'e', 'f'], 'col3': ['d4', 'e5', 'f6']})
df3 = pd.DataFrame({'col1': [7, 8, 9], 'col2': ['g', 'h', 'i'], 'col3': ['g7', 'h2', 'i3']})
```