Projecting Data using map

Let us go through the details about `map` to project the data.

- We can use `map` on top of `iterable` to return new `iterable` with all the transformed elements based up on the logic.
- It takes transformation logic and iterable as arguments. We can pass transformation logic either as regular function or lambda function.
- `map` returns a special iterable called as `map`. We have to type cast to regular collection such as `list` to preview the data or we can use for loop to iterate and print the data.
- Data from objects such as `filter`, `map` etc will be flushed out once we read from it.
- Number of elements in the `map` object will be same as `iterable` that is passed to it.

```
%run 02_preparing_data_sets.ipynb
```

```
orders[:10]
```

```
['1,2013-07-25 00:00:00.0,11599,CLOSED',
 '2,2013-07-25 00:00:00.0,256,PENDING_PAYMENT',
 '3,2013-07-25 00:00:00.0,12111,COMPLETE',
 '4,2013-07-25 00:00:00.0,8827,CLOSED',
 '5,2013-07-25 00:00:00.0,11318,COMPLETE',
 '6,2013-07-25 00:00:00.0,7130,COMPLETE',
 '7,2013-07-25 00:00:00.0,4530,COMPLETE',
 '8,2013-07-25 00:00:00.0,2911,PROCESSING',
 '9,2013-07-25 00:00:00.0,5657,PENDING_PAYMENT',
 '10,2013-07-25 00:00:00.0,5648,PENDING_PAYMENT']
```

```
len(orders)
```

```
68883
```

```
order_items[:10]
```

```
['1,1,957,1,299.98,299.98',
 '2,2,1073,1,199.99,199.99',
 '3,2,502,5,250.0,50.0',
 '4,2,403,1,129.99,129.99',
 '5,4,897,2,49.98,24.99',
 '6,4,365,5,299.95,59.99',
 '7,4,502,3,150.0,50.0',
 '8,4,1014,4,199.92,49.98',
 '9,5,957,1,299.98,299.98',
 '10,5,365,5,299.95,59.99']
```

```
len(order_items)
```

```
172198
```

- Get order_dates from orders

```
order = '1,2013-07-25 00:00:00.0,11599,CLOSED'
order.split(',')[1]
```

```
'2013-07-25 00:00:00.0'
```

```
order_dates = map(
    lambda order: order.split(',')[1],
    orders
)
```

```
type(order_dates)
```

```
map
```

```
list(order_dates)[:10]
```

```
['2013-07-25 00:00:00.0',
 '2013-07-25 00:00:00.0',
 '2013-07-25 00:00:00.0',
 '2013-07-25 00:00:00.0',
 '2013-07-25 00:00:00.0',
 '2013-07-25 00:00:00.0',
 '2013-07-25 00:00:00.0',
 '2013-07-25 00:00:00.0',
 '2013-07-25 00:00:00.0',
 '2013-07-25 00:00:00.0']
```

```
len(orders)
```

```
68883
```

> **ℹ Note**
>
> This will return 0 as data from map object `order_dates` is flushed out as part of the previous read.

```
len(list(order_dates))
```

```
0
```

> **ℹ Note**
>
> Creating order_dates once again by invoking `map` function to validate the number of elements. Number of elements in order_dates is same as orders.

```
order_dates = map(
    lambda order: order.split(',')[1],
    orders
)
```

```
len(list(order_dates))
```

```
68883
```

```
order_dates = map(
    lambda order: order.split(',')[1],
    orders
)
```

```
set(order_dates)
```

```
{'2013-07-25 00:00:00.0',
 '2013-07-26 00:00:00.0',
 '2013-07-27 00:00:00.0',
 '2013-07-28 00:00:00.0',
 '2013-07-29 00:00:00.0',
 '2013-07-30 00:00:00.0',
 '2013-07-31 00:00:00.0',
 '2013-08-01 00:00:00.0',
 '2013-08-02 00:00:00.0',
 '2013-08-03 00:00:00.0',
 '2013-08-04 00:00:00.0',
 '2013-08-05 00:00:00.0',
 '2013-08-06 00:00:00.0',
 '2013-08-07 00:00:00.0',
 '2013-08-08 00:00:00.0',
 '2013-08-09 00:00:00.0',
 '2013-08-10 00:00:00.0',
 '2013-08-11 00:00:00.0',
 '2013-08-12 00:00:00.0',
 '2013-08-13 00:00:00.0',
 '2013-08-14 00:00:00.0',
 '2013-08-15 00:00:00.0',
 '2013-08-16 00:00:00.0',
 '2013-08-17 00:00:00.0',
 '2013-08-18 00:00:00.0',
 '2013-08-19 00:00:00.0',
 '2013-08-20 00:00:00.0',
 '2013-08-21 00:00:00.0',
 '2013-08-22 00:00:00.0',
 '2013-08-23 00:00:00.0',
 '2013-08-24 00:00:00.0',
 '2013-08-25 00:00:00.0',
 '2013-08-26 00:00:00.0',
 '2013-08-27 00:00:00.0',
 '2013-08-28 00:00:00.0',
 '2013-08-29 00:00:00.0',
 '2013-08-30 00:00:00.0',
 '2013-08-31 00:00:00.0',
 '2013-09-01 00:00:00.0',
 '2013-09-02 00:00:00.0',
 '2013-09-03 00:00:00.0',
 '2013-09-04 00:00:00.0',
 '2013-09-05 00:00:00.0',
 '2013-09-06 00:00:00.0',
 '2013-09-07 00:00:00.0',
 '2013-09-08 00:00:00.0',
 '2013-09-09 00:00:00.0',
 '2013-09-10 00:00:00.0',
 '2013-09-11 00:00:00.0',
 '2013-09-12 00:00:00.0',
 '2013-09-13 00:00:00.0',
 '2013-09-14 00:00:00.0',
 '2013-09-15 00:00:00.0',
 '2013-09-16 00:00:00.0',
 '2013-09-17 00:00:00.0',
 '2013-09-18 00:00:00.0',
 '2013-09-19 00:00:00.0',
 '2013-09-20 00:00:00.0',
 '2013-09-21 00:00:00.0',
 '2013-09-22 00:00:00.0',
 '2013-09-23 00:00:00.0',
 '2013-09-24 00:00:00.0',
 '2013-09-25 00:00:00.0',
 '2013-09-26 00:00:00.0',
 '2013-09-27 00:00:00.0',
 '2013-09-28 00:00:00.0',
 '2013-09-29 00:00:00.0',
 '2013-09-30 00:00:00.0',
 '2013-10-01 00:00:00.0',
 '2013-10-02 00:00:00.0',
 '2013-10-03 00:00:00.0',
 '2013-10-04 00:00:00.0',
 '2013-10-05 00:00:00.0',
 '2013-10-06 00:00:00.0',
 '2013-10-07 00:00:00.0',
 '2013-10-08 00:00:00.0',
 '2013-10-09 00:00:00.0',
 '2013-10-10 00:00:00.0',
 '2013-10-11 00:00:00.0',
 '2013-10-12 00:00:00.0',
 '2013-10-13 00:00:00.0',
 '2013-10-14 00:00:00.0',
 '2013-10-15 00:00:00.0',
 '2013-10-16 00:00:00.0',
 '2013-10-17 00:00:00.0',
 '2013-10-18 00:00:00.0',
 '2013-10-19 00:00:00.0',
 '2013-10-20 00:00:00.0',
```

```
'2013-10-21 00:00:00.0',
'2013-10-22 00:00:00.0',
'2013-10-23 00:00:00.0',
'2013-10-24 00:00:00.0',
'2013-10-25 00:00:00.0',
'2013-10-26 00:00:00.0',
'2013-10-27 00:00:00.0',
'2013-10-28 00:00:00.0',
'2013-10-29 00:00:00.0',
'2013-10-30 00:00:00.0',
'2013-10-31 00:00:00.0',
'2013-11-01 00:00:00.0',
'2013-11-02 00:00:00.0',
'2013-11-03 00:00:00.0',
'2013-11-04 00:00:00.0',
'2013-11-05 00:00:00.0',
'2013-11-06 00:00:00.0',
'2013-11-07 00:00:00.0',
'2013-11-08 00:00:00.0',
'2013-11-09 00:00:00.0',
'2013-11-10 00:00:00.0',
'2013-11-11 00:00:00.0',
'2013-11-12 00:00:00.0',
'2013-11-13 00:00:00.0',
'2013-11-14 00:00:00.0',
'2013-11-15 00:00:00.0',
'2013-11-16 00:00:00.0',
'2013-11-17 00:00:00.0',
'2013-11-18 00:00:00.0',
'2013-11-19 00:00:00.0',
'2013-11-20 00:00:00.0',
'2013-11-21 00:00:00.0',
'2013-11-22 00:00:00.0',
'2013-11-23 00:00:00.0',
'2013-11-24 00:00:00.0',
'2013-11-25 00:00:00.0',
'2013-11-26 00:00:00.0',
'2013-11-27 00:00:00.0',
'2013-11-28 00:00:00.0',
'2013-11-29 00:00:00.0',
'2013-11-30 00:00:00.0',
'2013-12-01 00:00:00.0',
'2013-12-02 00:00:00.0',
'2013-12-03 00:00:00.0',
'2013-12-04 00:00:00.0',
'2013-12-05 00:00:00.0',
'2013-12-06 00:00:00.0',
'2013-12-07 00:00:00.0',
'2013-12-08 00:00:00.0',
'2013-12-09 00:00:00.0',
'2013-12-10 00:00:00.0',
'2013-12-11 00:00:00.0',
'2013-12-12 00:00:00.0',
'2013-12-13 00:00:00.0',
'2013-12-14 00:00:00.0',
'2013-12-15 00:00:00.0',
'2013-12-16 00:00:00.0',
'2013-12-17 00:00:00.0',
'2013-12-18 00:00:00.0',
'2013-12-19 00:00:00.0',
'2013-12-20 00:00:00.0',
'2013-12-21 00:00:00.0',
'2013-12-22 00:00:00.0',
'2013-12-23 00:00:00.0',
'2013-12-24 00:00:00.0',
'2013-12-25 00:00:00.0',
'2013-12-26 00:00:00.0',
'2013-12-27 00:00:00.0',
'2013-12-28 00:00:00.0',
'2013-12-29 00:00:00.0',
'2013-12-30 00:00:00.0',
'2013-12-31 00:00:00.0',
'2014-01-01 00:00:00.0',
'2014-01-02 00:00:00.0',
'2014-01-03 00:00:00.0',
'2014-01-04 00:00:00.0',
'2014-01-05 00:00:00.0',
'2014-01-06 00:00:00.0',
'2014-01-07 00:00:00.0',
'2014-01-08 00:00:00.0',
'2014-01-09 00:00:00.0',
'2014-01-10 00:00:00.0',
'2014-01-11 00:00:00.0',
'2014-01-12 00:00:00.0',
'2014-01-13 00:00:00.0',
'2014-01-14 00:00:00.0',
'2014-01-15 00:00:00.0',
'2014-01-16 00:00:00.0',
'2014-01-17 00:00:00.0',
```

```
'2014-01-18 00:00:00.0',
'2014-01-19 00:00:00.0',
'2014-01-20 00:00:00.0',
'2014-01-21 00:00:00.0',
'2014-01-22 00:00:00.0',
'2014-01-23 00:00:00.0',
'2014-01-24 00:00:00.0',
'2014-01-25 00:00:00.0',
'2014-01-26 00:00:00.0',
'2014-01-27 00:00:00.0',
'2014-01-28 00:00:00.0',
'2014-01-29 00:00:00.0',
'2014-01-30 00:00:00.0',
'2014-01-31 00:00:00.0',
'2014-02-01 00:00:00.0',
'2014-02-02 00:00:00.0',
'2014-02-03 00:00:00.0',
'2014-02-04 00:00:00.0',
'2014-02-05 00:00:00.0',
'2014-02-06 00:00:00.0',
'2014-02-07 00:00:00.0',
'2014-02-08 00:00:00.0',
'2014-02-09 00:00:00.0',
'2014-02-10 00:00:00.0',
'2014-02-11 00:00:00.0',
'2014-02-12 00:00:00.0',
'2014-02-13 00:00:00.0',
'2014-02-14 00:00:00.0',
'2014-02-15 00:00:00.0',
'2014-02-16 00:00:00.0',
'2014-02-17 00:00:00.0',
'2014-02-18 00:00:00.0',
'2014-02-19 00:00:00.0',
'2014-02-20 00:00:00.0',
'2014-02-21 00:00:00.0',
'2014-02-22 00:00:00.0',
'2014-02-23 00:00:00.0',
'2014-02-24 00:00:00.0',
'2014-02-25 00:00:00.0',
'2014-02-26 00:00:00.0',
'2014-02-27 00:00:00.0',
'2014-02-28 00:00:00.0',
'2014-03-01 00:00:00.0',
'2014-03-02 00:00:00.0',
'2014-03-03 00:00:00.0',
'2014-03-04 00:00:00.0',
'2014-03-05 00:00:00.0',
'2014-03-06 00:00:00.0',
'2014-03-07 00:00:00.0',
'2014-03-08 00:00:00.0',
'2014-03-10 00:00:00.0',
'2014-03-11 00:00:00.0',
'2014-03-12 00:00:00.0',
'2014-03-13 00:00:00.0',
'2014-03-14 00:00:00.0',
'2014-03-15 00:00:00.0',
'2014-03-16 00:00:00.0',
'2014-03-17 00:00:00.0',
'2014-03-18 00:00:00.0',
'2014-03-19 00:00:00.0',
'2014-03-20 00:00:00.0',
'2014-03-21 00:00:00.0',
'2014-03-22 00:00:00.0',
'2014-03-23 00:00:00.0',
'2014-03-24 00:00:00.0',
'2014-03-25 00:00:00.0',
'2014-03-26 00:00:00.0',
'2014-03-27 00:00:00.0',
'2014-03-28 00:00:00.0',
'2014-03-29 00:00:00.0',
'2014-03-30 00:00:00.0',
'2014-03-31 00:00:00.0',
'2014-04-01 00:00:00.0',
'2014-04-02 00:00:00.0',
'2014-04-03 00:00:00.0',
'2014-04-04 00:00:00.0',
'2014-04-05 00:00:00.0',
'2014-04-06 00:00:00.0',
'2014-04-07 00:00:00.0',
'2014-04-08 00:00:00.0',
'2014-04-09 00:00:00.0',
'2014-04-10 00:00:00.0',
'2014-04-11 00:00:00.0',
'2014-04-12 00:00:00.0',
'2014-04-13 00:00:00.0',
'2014-04-14 00:00:00.0',
'2014-04-15 00:00:00.0',
'2014-04-16 00:00:00.0',
'2014-04-17 00:00:00.0',
```

```
'2014-04-18 00:00:00.0',
'2014-04-19 00:00:00.0',
'2014-04-20 00:00:00.0',
'2014-04-21 00:00:00.0',
'2014-04-22 00:00:00.0',
'2014-04-23 00:00:00.0',
'2014-04-24 00:00:00.0',
'2014-04-25 00:00:00.0',
'2014-04-26 00:00:00.0',
'2014-04-27 00:00:00.0',
'2014-04-28 00:00:00.0',
'2014-04-29 00:00:00.0',
'2014-04-30 00:00:00.0',
'2014-05-01 00:00:00.0',
'2014-05-02 00:00:00.0',
'2014-05-03 00:00:00.0',
'2014-05-04 00:00:00.0',
'2014-05-05 00:00:00.0',
'2014-05-06 00:00:00.0',
'2014-05-07 00:00:00.0',
'2014-05-08 00:00:00.0',
'2014-05-09 00:00:00.0',
'2014-05-10 00:00:00.0',
'2014-05-11 00:00:00.0',
'2014-05-12 00:00:00.0',
'2014-05-13 00:00:00.0',
'2014-05-14 00:00:00.0',
'2014-05-15 00:00:00.0',
'2014-05-16 00:00:00.0',
'2014-05-17 00:00:00.0',
'2014-05-18 00:00:00.0',
'2014-05-19 00:00:00.0',
'2014-05-20 00:00:00.0',
'2014-05-21 00:00:00.0',
'2014-05-22 00:00:00.0',
'2014-05-23 00:00:00.0',
'2014-05-24 00:00:00.0',
'2014-05-25 00:00:00.0',
'2014-05-26 00:00:00.0',
'2014-05-27 00:00:00.0',
'2014-05-28 00:00:00.0',
'2014-05-29 00:00:00.0',
'2014-05-30 00:00:00.0',
'2014-05-31 00:00:00.0',
'2014-06-01 00:00:00.0',
'2014-06-02 00:00:00.0',
'2014-06-03 00:00:00.0',
'2014-06-04 00:00:00.0',
'2014-06-05 00:00:00.0',
'2014-06-06 00:00:00.0',
'2014-06-07 00:00:00.0',
'2014-06-08 00:00:00.0',
'2014-06-09 00:00:00.0',
'2014-06-10 00:00:00.0',
'2014-06-11 00:00:00.0',
'2014-06-12 00:00:00.0',
'2014-06-13 00:00:00.0',
'2014-06-14 00:00:00.0',
'2014-06-15 00:00:00.0',
'2014-06-16 00:00:00.0',
'2014-06-17 00:00:00.0',
'2014-06-18 00:00:00.0',
'2014-06-19 00:00:00.0',
'2014-06-20 00:00:00.0',
'2014-06-21 00:00:00.0',
'2014-06-22 00:00:00.0',
'2014-06-23 00:00:00.0',
'2014-06-24 00:00:00.0',
'2014-06-25 00:00:00.0',
'2014-06-26 00:00:00.0',
'2014-06-27 00:00:00.0',
'2014-06-28 00:00:00.0',
'2014-06-29 00:00:00.0',
'2014-06-30 00:00:00.0',
'2014-07-01 00:00:00.0',
'2014-07-02 00:00:00.0',
'2014-07-03 00:00:00.0',
'2014-07-04 00:00:00.0',
'2014-07-05 00:00:00.0',
'2014-07-06 00:00:00.0',
'2014-07-07 00:00:00.0',
'2014-07-08 00:00:00.0',
'2014-07-09 00:00:00.0',
'2014-07-10 00:00:00.0',
'2014-07-11 00:00:00.0',
'2014-07-12 00:00:00.0',
'2014-07-13 00:00:00.0',
'2014-07-14 00:00:00.0',
'2014-07-15 00:00:00.0',
```

```
    '2014-07-16 00:00:00.0',
    '2014-07-17 00:00:00.0',
    '2014-07-18 00:00:00.0',
    '2014-07-19 00:00:00.0',
    '2014-07-20 00:00:00.0',
    '2014-07-21 00:00:00.0',
    '2014-07-22 00:00:00.0',
    '2014-07-23 00:00:00.0',
    '2014-07-24 00:00:00.0'}
```

```
order_dates = map(
    lambda order: order.split(',')[1],
    orders
)
```

```
len(set(order_dates))
```

```
364
```

- Use orders and extract order_id as well as order_date from each element in the form of a tuple. Make sure that order_id is of type int.

```
orders[:10]
```

```
['1,2013-07-25 00:00:00.0,11599,CLOSED',
 '2,2013-07-25 00:00:00.0,256,PENDING_PAYMENT',
 '3,2013-07-25 00:00:00.0,12111,COMPLETE',
 '4,2013-07-25 00:00:00.0,8827,CLOSED',
 '5,2013-07-25 00:00:00.0,11318,COMPLETE',
 '6,2013-07-25 00:00:00.0,7130,COMPLETE',
 '7,2013-07-25 00:00:00.0,4530,COMPLETE',
 '8,2013-07-25 00:00:00.0,2911,PROCESSING',
 '9,2013-07-25 00:00:00.0,5657,PENDING_PAYMENT',
 '10,2013-07-25 00:00:00.0,5648,PENDING_PAYMENT']
```

```
[(1, '2013-07-25 00:00:00.0'), (2, '2013-07-25 00:00:00.0')]
```

```
[(1, '2013-07-25 00:00:00.0'), (2, '2013-07-25 00:00:00.0')]
```

```
order = orders[0]
```

```
(int(order.split(',')[0]), order.split(',')[1])
```

```
(1, '2013-07-25 00:00:00.0')
```

```
order_tuples = map(
    lambda order: (int(order.split(',')[0]), order.split(',')[1]),
    orders
)
```

```
list(order_tuples)[:10]
```

```
[(1, '2013-07-25 00:00:00.0'),
 (2, '2013-07-25 00:00:00.0'),
 (3, '2013-07-25 00:00:00.0'),
 (4, '2013-07-25 00:00:00.0'),
 (5, '2013-07-25 00:00:00.0'),
 (6, '2013-07-25 00:00:00.0'),
 (7, '2013-07-25 00:00:00.0'),
 (8, '2013-07-25 00:00:00.0'),
 (9, '2013-07-25 00:00:00.0'),
 (10, '2013-07-25 00:00:00.0')]
```