

## Ques 12 : How do I change the data type of a pandas Series ?

This is useful when a column's data type is string but actually holds numbers. So, we can change the type from string to number for doing mathematical calculation.

```
In [1]: import pandas as pd
drinks_ds = pd.read_csv('http://bit.ly/drinksbycountry')
drinks_ds.head(3)
```

```
Out[1]:
```

	country	beer_servings	spirit_servings	wine_servings	total_litres_of_pure_alcohol	continent
0	Afghanistan	0	0	0	0.0	Asia
1	Albania	89	132	54	4.9	Europe
2	Algeria	25	0	14	0.7	Africa

```
In [2]: drinks_ds.dtypes    # 'object' means string
```

```
Out[2]: country                object
beer_servings                int64
spirit_servings              int64
wine_servings                int64
total_litres_of_pure_alcohol float64
continent                    object
dtype: object
```

```
In [3]: drinks_ds['beer_servings'] = drinks_ds.beer_servings.astype(float)
```

```
In [4]: drinks_ds.dtypes
```

```
Out[4]: country                object
beer_servings                float64
spirit_servings              int64
wine_servings                int64
total_litres_of_pure_alcohol float64
continent                    object
dtype: object
```

```
In [5]: drinks = pd.read_csv('http://bit.ly/drinksbycountry', dtype={'beer_servings':float})
drinks.dtypes
```

```
Out[5]: country                object
beer_servings                float64
spirit_servings              int64
wine_servings                int64
total_litres_of_pure_alcohol float64
continent                    object
dtype: object
```

```
In [12]: orders_ds = pd.read_table('http://bit.ly/chiporders')
orders_ds.head(3)
```

```
Out[12]:
```

	order_id	quantity	item_name	choice_description	item_price
0	1	1	Chips and Fresh Tomato Salsa	NaN	\$2.39
1	1	1	Izze	[Clementine]	\$3.39
2	1	1	Nantucket Nectar	[Apple]	\$3.39

```
In [13]: orders_ds.dtypes
```

```
Out[13]: order_id          int64
quantity          int64
item_name         object
choice_description object
item_price        object
dtype: object
```

**Datatype of the column 'item\_price' is 'object'. We will cast it as float and then find mean of the column**

```
In [14]: orders_ds.item_price.str.replace('$', '').head()
```

```
Out[14]: 0    2.39
         1    3.39
         2    3.39
         3    2.39
         4   16.98
         Name: item_price, dtype: object
```

```
In [15]: orders_ds.item_price.str.replace('$', '').astype(float).mean()
```

```
Out[15]: 7.464335785374397
```

```
In [16]: # converting True and False to 1 and 0 . This is important in ML
orders_ds.item_name.str.contains('Chicken').astype(int).head()
```

```
Out[16]: 0    0
         1    0
         2    0
         3    0
         4    1
         Name: item_name, dtype: int32
```