

به نام خدا

پاکسازی داده ها

از قسمت دیتا سه فایل زیر را دانلود کرده و پردازش لازم انجام شد و خروجی ها ذخیره و آپلود شدند.

Asia - Export Value و Asia - Export Quantity و Asia – Agriculture

هر کدام از فایل ها شامل دیتاهایی بودند که کاربردی نداشتند پس با دستور **drop** آنها را از فایل ها حذف کردم. یعنی سه فایل ما در ابتدا ۱۲ ستون بودند که هشت ستون حذف کردم.

ستون ها قبل از دستور دراپ :

AsiaEV.head()												
	Area Code	Area	Item Code	Item	Element Code	Element	Year Code	Year	Unit	Value	Flag	continent
0	2	Afghanistan	4004	Calcium ammonium nitrate (CAN) and other mixtu...	5922	Export Value	2003	2003	1000 US\$	1.51	R	Asia
1	2	Afghanistan	4030	Fertilizers n.e.c.	5922	Export Value	2011	2011	1000 US\$	10.00	R	Asia
2	2	Afghanistan	4030	Fertilizers n.e.c.	5922	Export Value	2015	2015	1000 US\$	16.00	R	Asia
3	2	Afghanistan	4030	Fertilizers n.e.c.	5922	Export Value	2017	2017	1000 US\$	2.02	R	Asia
4	2	Afghanistan	4023	Monoammonium phosphate (MAP)	5922	Export Value	2003	2003	1000 US\$	58.81	R	Asia

و بعد ستون های اضافی را حذف میکنیم:

```
AsiaEV = AsiaEV.drop(['Area Code','Item Code','Element Code','Element','Year Code','Unit','Flag','continent'],axis=1)
AsiaEQ = AsiaEQ.drop(['Area Code','Item Code','Element Code','Element','Year Code','Unit','Flag','continent'],axis=1)
AsiaA = AsiaA.drop(['Area Code','Item Code','Element Code','Element','Year Code','Unit','Flag','continent'],axis=1)
```

و فایل های جدید :

```
[ ] AsiaA.head()
```

	Area	Item	Year	Value
0	Afghanistan	Calcium ammonium nitrate (CAN) and other mixtu...	2003	0.03
1	Afghanistan	Fertilizers n.e.c.	2011	21.00
2	Afghanistan	Fertilizers n.e.c.	2015	9.15
3	Afghanistan	Fertilizers n.e.c.	2017	34.00
4	Afghanistan	Monoammonium phosphate (MAP)	2003	200.00

و چون داده های ما سری زمانی هست ستون زمان را بعنوان اندیس دیتا تعیین میکنیم.

و با دستور `AsiaA[AsiaA['Value'] == 0]` داده هایی که مقدار آنها برابر صفر است رو بررسی میکنیم.