


```
(47): df.age = df.age.astype('int64')

In [48]: df

Out [48]:
```

	age	sex	chronic_disease_binary	outcome	asymptomatic	cough	fatigue	fever	headache	malaise	pneumonia	respiratory problems	runny nose	sore throat
476	60	1		0	1	0	0	0	1	0	0	1	0	0
504	45	1		0	1	0	0	0	1	0	0	0	0	0
756	25	1		0	1	0	1	0	1	0	0	0	0	0
765	45	1		0	1	0	1	0	1	0	0	0	0	0
767	85	1		0	1	0	0	0	1	0	0	0	0	0
...
658596	65	1		0	1	0	1	0	0	0	0	0	0	0
671470	45	1		0	1	0	0	0	0	0	0	0	0	0
672418	35	1		1	1	0	1	0	1	0	0	0	0	0
672428	65	1		0	1	0	1	0	0	0	0	0	0	0
674650	39	0		0	1	0	0	0	0	0	0	0	0	0

1577 rows × 16 columns

```
In [49]: df.dtypes
```

```
Out [49]: age                int64
sex                int32
chronic_disease_binary  int32
outcome            int32
asymptomatic       int64
cough              int64
fatigue            int64
fever              int64
headache           int64
malaise            int64
pneumonia          int64
respiratory problems  int64
runny nose         int64
sore throat        int64
diabetes           int64
hypertension       int64
dtype: object
```

Réorganiser les colonnes

```
In [50]: cols = df.columns.tolist()
cols = cols[0:3] + cols[4:16] + cols[3:4]

In [51]: cols

Out [51]: ['age',
'sex',
'chronic_disease_binary',
'asymptomatic',
'cough',
'fatigue',
'fever',
'headache',
'malaise',
'pneumonia',
'respiratory problems',
'runny nose',
'sore throat',
'diabetes',
'hypertension',
'outcome']

In [53]: df = df[cols] # OR df = df.ix[:, cols]
```

Notre jeu de données après le nettoyage

```
In [54]: df

Out [54]:
```

	age	sex	chronic_disease_binary	asymptomatic	cough	fatigue	fever	headache	malaise	pneumonia	respiratory problems	runny nose	sore throat
476	60	1		0	0	0	1	0	0	1	0	0	0
504	45	1		0	0	0	1	0	0	0	0	0	0
756	25	1		0	0	1	0	1	0	0	0	0	1
765	45	1		0	0	1	0	1	0	0	0	0	0
767	85	1		0	0	0	1	0	0	0	0	0	0
...
658596	65	1		0	0	1	0	0	0	0	0	0	0
671470	45	1		0	0	0	0	0	0	0	0	0	0
672418	35	1		1	0	1	0	1	0	0	0	0	0
672428	65	1		0	0	1	0	0	0	0	0	0	0
674650	39	0		0	0	1	0	0	0	0	0	0	0

1577 rows × 16 columns

Enregistrer ce jeu données sous format CSV

```
In [55]: df.to_csv('covid_data_cleaned.csv', index=False)
```

Réalisé par : AABBOU ZAKARIA

- [E-mail](mailto:zakaria.abbou199434@gmail.com) : zakaria.abbou199434@gmail.com
- [GitHub](https://github.com/ZakariaAABBOU) : github.com/ZakariaAABBOU
- [Linkedin](https://www.linkedin.com/in/zakaria-aabbou/) : linkedin.com/in/zakaria-aabbou/

```
In [ ] :
```