

9.1 - don't be fanatic

July 12, 2017

```
In [1]: import pandas as pd
import numpy as np
```

```
In [2]: df = pd.DataFrame(['a b c']*10000, columns=['col'])
```

```
In [3]: df
```

```
Out[3]:
```

	col
0	a b c
1	a b c
2	a b c
3	a b c
4	a b c
5	a b c
6	a b c
7	a b c
8	a b c
9	a b c
10	a b c
11	a b c
12	a b c
13	a b c
14	a b c
15	a b c
16	a b c
17	a b c
18	a b c
19	a b c
20	a b c
21	a b c
22	a b c
23	a b c
24	a b c
25	a b c
26	a b c
27	a b c
28	a b c
29	a b c

```
...      ...
9970 a b c
9971 a b c
9972 a b c
9973 a b c
9974 a b c
9975 a b c
9976 a b c
9977 a b c
9978 a b c
9979 a b c
9980 a b c
9981 a b c
9982 a b c
9983 a b c
9984 a b c
9985 a b c
9986 a b c
9987 a b c
9988 a b c
9989 a b c
9990 a b c
9991 a b c
9992 a b c
9993 a b c
9994 a b c
9995 a b c
9996 a b c
9997 a b c
9998 a b c
9999 a b c
```

```
[10000 rows x 1 columns]
```

```
In [4]: pd.DataFrame(df['col'].apply(lambda x : pd.Series(x.split())))
```

```
Out[4]:
```

	0	1	2
0	a	b	c
1	a	b	c
2	a	b	c
3	a	b	c
4	a	b	c
5	a	b	c
6	a	b	c
7	a	b	c
8	a	b	c
9	a	b	c
10	a	b	c

11	a	b	c
12	a	b	c
13	a	b	c
14	a	b	c
15	a	b	c
16	a	b	c
17	a	b	c
18	a	b	c
19	a	b	c
20	a	b	c
21	a	b	c
22	a	b	c
23	a	b	c
24	a	b	c
25	a	b	c
26	a	b	c
27	a	b	c
28	a	b	c
29	a	b	c
...
9970	a	b	c
9971	a	b	c
9972	a	b	c
9973	a	b	c
9974	a	b	c
9975	a	b	c
9976	a	b	c
9977	a	b	c
9978	a	b	c
9979	a	b	c
9980	a	b	c
9981	a	b	c
9982	a	b	c
9983	a	b	c
9984	a	b	c
9985	a	b	c
9986	a	b	c
9987	a	b	c
9988	a	b	c
9989	a	b	c
9990	a	b	c
9991	a	b	c
9992	a	b	c
9993	a	b	c
9994	a	b	c
9995	a	b	c
9996	a	b	c
9997	a	b	c

```
9998 a b c
9999 a b c
```

```
[10000 rows x 3 columns]
```

```
In [5]: %timeit pd.DataFrame(df['col'].apply( lambda x : pd.Series(x.split())))
```

```
1 loop, best of 3: 1.35 s per loop
```

```
In [6]: %timeit pd.DataFrame(df['col'].apply( lambda x : x.split()))
```

```
100 loops, best of 3: 5.21 ms per loop
```