



Python/Pandas

Map the headers to a column

Map the headers to a column with pandas?

Data set: Stack Over Flow 2018 insights

- <https://insights.stackoverflow.com/survey>
- <https://insights.stackoverflow.com/survey/2018#technology>

Topics

- map a headers based on a value to a new column

Bonus

- pandas dot method - matrix multiplication
- understand np.where
- map single column of dataframe
- map all columns of a dataframe
- map and NaN
- check all distinct values in dataframe
- Optimize big data frames:
 - Columns have mixed types. Specify dtype option on import or set low_memory=False.

0	1	0	0
1	1	1	0
2	1	1	0
3	0	0	0
4	1	0	0

```
In [8]: # map headers to columns way 1
df_answers['answer'] = np.where(df_answers, df_answers.columns, '').sum(axis=1)
df_answers.head()
```

	Hobby	OpenSource	Student	answer
0	1	0	0	Hobby
1	1	1	0	HobbyOpenSource
2	1	1	0	HobbyOpenSource
3	0	0	0	
4	1	0	0	Hobby

```
In [9]:
```