

Want to learn machine learning? Use my **machine learning flashcards**.

Convert A Categorical Variable Into Dummy Variables

20 Dec 2017

```
# import modules  
import pandas as pd
```

```
# Create a dataframe  
raw_data = {'first_name': ['Jason', 'Molly', 'Tina', 'Jake',  
                           'Amy'],  
            'last_name': ['Miller', 'Jacobson', 'Ali', 'Milner',  
                          'Cooze'],  
            'sex': ['male', 'female', 'male', 'female', 'female']}  
df = pd.DataFrame(raw_data, columns = ['first_name', 'last_name',  
                                       'sex'])  
df
```

	first_name	last_name	sex
0	Jason	Miller	male
1	Molly	Jacobson	female
2	Tina	Ali	male
3	Jake	Milner	female
4	Amy	Cooze	female

```
# Create a set of dummy variables from the sex variable  
df_sex = pd.get_dummies(df['sex'])
```

```
# Join the dummy variables to the main dataframe  
df_new = pd.concat([df, df_sex], axis=1)  
df_new
```

CHRIS ALBON

[ARTICLES](#) [ML/AI NOTES](#) [ABOUT](#)