
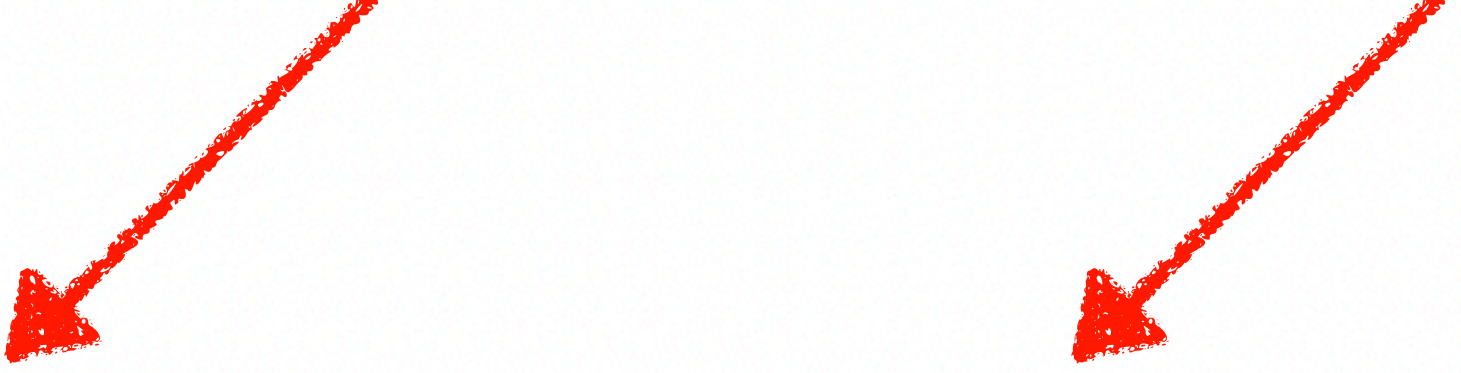


WHEN ARE PREDICTORS CHOSEN?

```
1 lm_spec = LinearRegression()  
2  
3 lm_spec.fit(X = penguins[["bill_length_mm", "bill_depth_mm"]],  
4             y = penguins["body_mass_g"])
```



```
1 lm_spec <- linear_reg()  
2  
3 lm_spec %>%  
4   fit(body_mass_g ~ bill_length_mm + bill_depth_mm,  
5       data = penguins)
```



WHAT IF WE ADD PREPROCESSING?

```
1 lm_spec = LinearRegression()  
2 log_tf = FunctionTransformer(log)  
3  
4 pipeline = make_pipeline(  
5     log_tf,  
6     lm_spec  
7 )  
8  
9 pipeline.fit(X = penguins[["bill_length_mm", "bill_depth_mm"]],  
10             y = penguins["body_mass_g"])
```