

- ① Missing values (however R `lm()` ignores them)
- ② Spurious regression  
(explain GDP with number of penguins)
- ③ Usage of values of categorical factors as numbers (conversion as `factor()` is required)
- ④ Everytime think about really important factor values: instead of  $\{1 \dots 20\} \Rightarrow \{"Bad", "Avg", "Good"\}$  group them!

## Regression Traps