# How about more categories?

## One-hot

- There are a few ways of modeling predictors with many categories:
  - **Contrasts**: Each category is compared to a **baseline**, and the coefficients are comparisons between baseline and levels
  - **One-hot**: coefficients are means of each level of the predictor
  - **Residuals**: an overall mean is measured, and coefficients are differences between each level and the global mean

```
> x
[1] "B" "A" "A" "B" "C" "B" "B" "C" "B"
> onehot
  xA xB xC
1  0  1  0
2  1  0  0
3  1  0  0
4  0  1  0
5  0  0  1
6  0  1  0
7  0  1  0
8  0  0  1
9  0  1  0
```

# How about more categories?

## One-hot

- There are a few ways of modeling predictors with many categories:
  - **Contrasts**: Each category is compared to a **baseline**, and the coefficients are comparisons between baseline and levels
  - **One-hot**: coefficients are means of each level of the predictor
  - **Residuals**: an overall mean is measured, and coefficients are differences between each level and the global mean

```
> onehot = model.matrix(~0+x, data = df)
> fit_onehot = stan_glm(y ~ 0 + onehot, data =
df, cores = 4)
> summary(fit_onehot)[1:3, 1:3]
          mean mcse   sd
onehotxA 0.72    0 0.13
onehotxB 2.04    0 0.14
onehotxC 2.93    0 0.12
```