Interaction Terms in Linear Regression

$$Y = \alpha + \underline{\beta_1 \, \textsf{primary2004} + \beta_2 \, \textsf{Neighbors}} + \beta_3 \, (\textsf{primary2004} \times \textsf{Neighbors}) + \epsilon$$

- Hierarchy principle for interaction: all low level effect terms must be included
  - Why? (primary2004, Neighbors)

	(0,0)	(1,0)	(0,1)	(1,1)
$\hat{Y}$	$\hat{\alpha}$	$\hat{\alpha} + \hat{\beta}_1$	$\hat{\alpha} + \hat{\beta}_2$	$\hat{\alpha} + \hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3$

  - Now think of a case where you drop the term $\beta_1 \, \textsf{primary2004}$

	(0,0)	(1,0)	(0,1)	(1,1)
$\hat{Y}$	$\hat{\alpha}$	$\hat{\alpha}$	$\hat{\alpha} + \hat{\beta}_2$	$\hat{\alpha} + \hat{\beta}_2 + \hat{\beta}_3$

    - Expected outcomes for (0,0) and (1,0) become identical!
      - Inappropriate assumption!
      - All lower level effect terms $\beta_1 \, \textsf{primary2004} + \beta_2 \, \textsf{Neighbors}$ must be included

Interaction Terms in Linear Regression

- Interaction model with linear age effect

$$Y = \alpha + \beta_1 \text{ age} + \beta_2 \text{ Neighbors} + \beta_3 (\text{age} \times \text{Neighbors}) + \epsilon$$

  - Modeling assumption: (Heterogeneous treatment effect by age)
    - Treatment effect of Neighbors message is a function of age
      - Neighbor message treatment effect for age *x* population

$$(\hat{\alpha} + \hat{\beta}_1 x + \hat{\beta}_2 + \hat{\beta}_3 x) - (\hat{\alpha} + \hat{\beta}_1 x) = \hat{\beta}_2 + \hat{\beta}_3 x$$

        - e.g. age x = 20: 0.0486 + 0.0006 x 20 = 0.0606
        - e.g. age x = 50: 0.0486 + 0.0006 x 50 = 0.0786
  - In case when you do not have interaction term $\beta_3 (\text{age} \times \text{Neighbors})$
    - Treatment effect of Neighbors does not become a function of age

$$(\hat{\alpha} + \hat{\beta}_1 x + \hat{\beta}_2) - (\hat{\alpha} + \hat{\beta}_1 x) = \hat{\beta}_2$$

        - e.g. age x = 20: 0.0486; e.g. age x = 50: 0.0486
        - Heterogeneous treatment effect does not get captured!