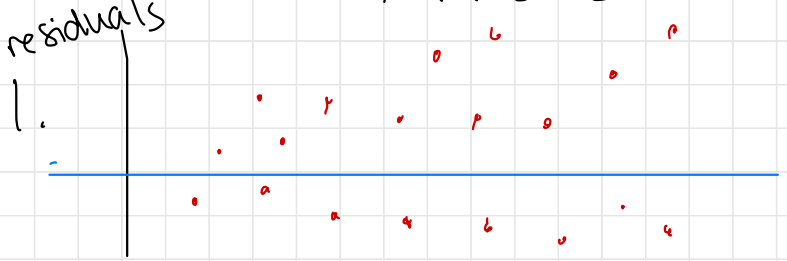
Three General Points Based on Presentations:



If it's hard to see evidence of nonconstant variance from a plot like this, try plotting absolute value of residuals against fitted values

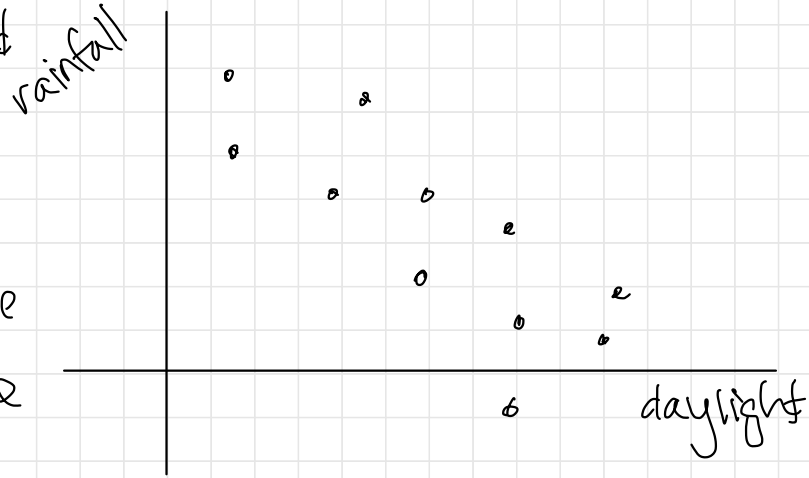
2. Why do we care about correlated predictors?

Example: Measured a response, student happiness

Covariates: Daily sunlight where student lives
Daily rainfall where the student lives

Wanted to ask - is daily sunlight associated with student happiness?

2. why do we care about correlated predictors?



Example: Measured a response, student happiness

Covariates: Daily sunlight where student lives
Daily rainfall where the student lives

Wanted to ask - is daily sunlight associated with student happiness?

Imagine fit $Y_i = \beta_0 + \beta_1 X_{i1} + \varepsilon_i$, $\varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$
X_{i1} daily sunlight

get estimate $b_1 = 1$ average change in student happiness given 1 more hour of sunlight, reject null $\beta_1 = 0$ w/ a t-test

problem - 1 unit increase in sunlight in our data is associated w/ less rainfall

If we want to isolate relationship between student happiness and sunlight, need to include rainfall too in our model, $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$, $\varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$, x_{i2} is rainfall

$\Longrightarrow$ This allows us to interpret the sunlight regression coefficient as the average happiness change given a one unit increase in sunlight holding rainfall constant

3. Think carefully about interpretation of indicators