mean = 72.5

sd = 12

max = 91.5

≥90 = 6

≥80 = 22

A —> 88

A- ... > 80

B+ > 72.5

B or B- > 55

C .....

ANOVA table: a break-down of squares (variation)

TABLE 3.4 The Overall Analysis of Variance Table

Source	df	SS	MS	F	p-value
Regression	p	$SSreg$	$SSreg/p$	$MSreg/\hat{\sigma}^2$	
Residual	$n-(p+1)$	RSS	$\hat{\sigma}^2 = RSS/(n-(p+1))$		
Total	$n-1$	SYY			

$$\sum_{i=1}^{n}[y_i - \bar{y}]^2 = \sum_{i=1}^{n}[y_i - \hat{y}_i]^2 + \sum_{i=1}^{n}[\hat{y}_i - \bar{y}]^2$$

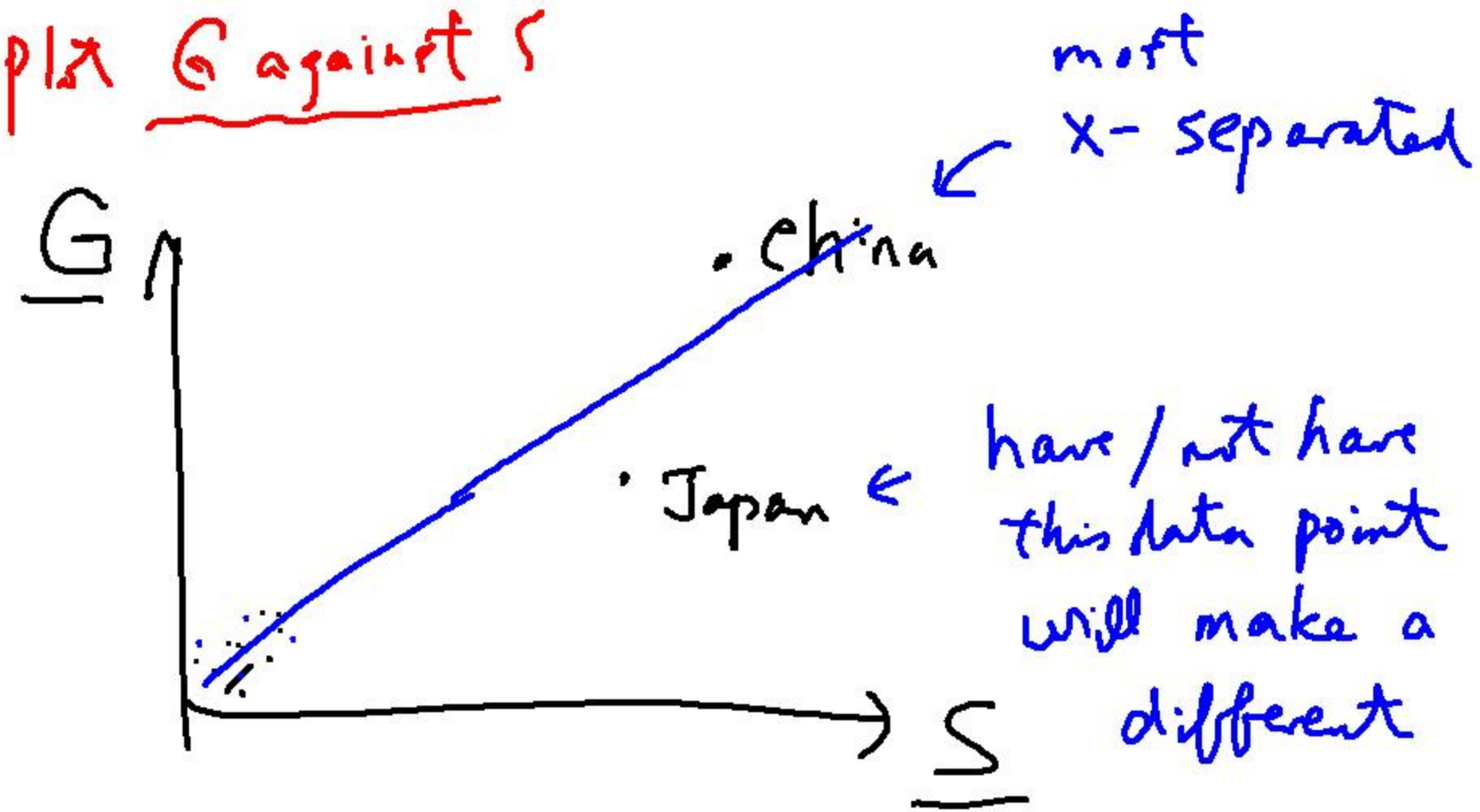
$$TSS = SYY = RSS + SSreg$$

Variation of the data — $Y'(I-J)Y$

Variation not explained by regression — $Y'(I-H)Y$

Variation explained by regression — $Y'(H-J)Y$

plot G against S



most X-separated

have / not have this data point will make a different

4.1 Understanding parameter estimates parameters=($\beta_0, \beta_1, \ldots \beta_p, \sigma^2$)

$$E(Y|X) = \beta_0 X_0 + \beta_1 X_1 + \cdots + \beta_p X_p$$

$$E(Fuel|X) = 154.19 - 4.23\,Tax + 0.47\,Dlic - 6.14\,Income + 18.54\log(Miles)$$

$Y = X\beta + \varepsilon$

$Var(\varepsilon) = \sigma^2$

- Unit of βs:
  - unit of y/unit of x. (e.g. gallon/$1000 for -6.14)
- Unit of σ^2:
  - (unit of y)^2
- Meaning of β_i:
  - Rate of change of y on x_i , after adjusting for other variables
  - e.g.
    - Fuel decreased by 6.14 gallon when Income increased by $1000
    - Fuel increased by 18.54 gallon when Miles is doubled ($\log_2 2x = 1+\log(x)$)
- Meaning of σ^2:
  - Variability that can't be explained by the regression line.

$Y = \beta + \beta_1 X_1 \cdots + \beta_i X_i + \cdots + \varepsilon$

$Y = \alpha_0 + \beta_i X_i + \varepsilon$

not adjust for other variables