4.)

(a) Expect cubic RSS to be lower. Since more params, will ~~fit~~ over fit to some noise.

(b) " Linear to be lower, since cubic has overfit to noise and will not generalize well to new data.

(c) ~~Not enough info~~ Poly lower

(d) Not enough info

5.)

$$\hat{y} = \left(\frac{\sum x_i y_i}{\sum x_i^2}\right) \cdot x_i \Rightarrow \hat{y} = x_i \left(\frac{\sum x_i y_i}{\sum x_i^2}\right)$$

$$\hat{y}_i = \left(\frac{x_i}{\sum x_i^2}\right) \sum x_i y_i$$

$$\hat{y}_i = \sum \left(\left(\frac{x_i \sum x_i}{\sum x_i^2}\right)\right) y_i$$

$= a_i'$

$$= \frac{2yx \sum xy - \left(\frac{\sum xy}{\sum x^2}\right)(x^2)}{\sum x^2 \sum y^2}$$

$$= \left[2yx - \left(\frac{x^2}{\sum x^2}\right)\right] \sum xy / \ldots$$

$$= \frac{2yx\sum x^2}{\sum x^2} - \frac{x^2}{\sum x^2} \ldots$$

$$= \left(\frac{2yx\sum x^2 - x^2}{\sum x^2}\right)$$

6.)

When $x_i = \bar{x}$, $\Rightarrow \hat{\beta}_1 = \frac{\sum 0 (\ldots)}{\ }$ $(y-\bar{y})(y-\bar{y})$

$\beta_0 = 0$

$\beta_1 = \frac{\sum xy}{\sum x^2}$

$y = \hat{\beta}_0 + \hat{\beta}_1 \cdot x \Rightarrow$

$Y_{pred} = \hat{\beta}_0 + \hat{\beta}_1 \cdot \bar{x} =$

$= (\bar{y} - \hat{\beta}_1 \bar{x}) + \hat{\beta}_1 \bar{x} = \bar{Y}$

So $x = \bar{x} \Rightarrow y = \bar{y}$.

Assume $\bar{x} = \bar{y} = 0$

7.)

$$cor(x,y) = \frac{\sum xy}{\sqrt{\sum x^2}\sqrt{\sum y^2}}$$

$$c^2 = \frac{\sum xy \cdot \sum xy}{\sum x^2 \sum y^2} \stackrel{?}{=}$$

$$R^2 = \frac{\sum y^2 - \sum (y-\hat{y})^2}{\sum y^2} = 1 - \frac{\sum (y-\hat{y})^2}{\sum y^2}$$

$$= \frac{\sum y^2 - [\sum y^2 - 2y\hat{y} + \hat{y}^2]}{\sum y^2} = \frac{2y\hat{y} - \hat{y}^2}{\sum y^2}$$

$$= \frac{2y\left(\frac{\sum xy}{\sum x^2}\right)x - \left(\frac{\sum xy}{\sum x^2}\right)^2 x^2}{\sum y^2} =$$