Q1

- (a) subset - best model selection has the smallest RSS
- (b) subset - best model selection has the smallest RSS
- (c)
    - i. true
    - ii. true
    - iii. false
    - iv. false
    - v. false

Q2

- (a) iii. Less flexible and hence will give improved prediction accu racy when its increase in bias is less than its decrease in variance.
- (b) Same as lasso.
- (c) ii. More flexible and hence will give improved prediction accu racy when its increase in variance is less than its decrease in bias.

Q3

- (a) iv. Steadily decrease. When s is large enougth, $\displaystyle\sum_{j=1}^p|b_j|$ is equal to RSS, thus regression is optimal. When s = 0, regression line is not option, thus RSS is large at the beginning.
- (b) ii. Decrease initially, and then eventually start increasing in a U shape. Regression line will be poor at the beginning, then is will find an optimal point, then it will move away.
- (c) iii. Steadily increase. Flexibility of the model grows, so the variance grows too
- (d) iv. Steadily decrease. Flexibility of the model grows, so the squared bias decreases.
- (e) v. Remain constant. It's a gap between squired bias, variance, and the total error.

Q4

- (a) Steadily increase. We are starting from the point where training RSS is minimum, and moving towords less complexity.
- (b) ii. Decrease initially, and then eventually start increasing in a U shape. We are starting near optimal points, and moving towords less complexity. Up to the optimal point test RSS decreases, and after then increases.
- (c) iv. Steadily decrease. Explanation as in (a)
- (d) iii. Steadily increase. Explanation as in (a)
- (e) v. Remain constant. It's a gap between squired bias, variance, and the total error.