

### **HW3INF552 - Time Series Classification:**

#### **1. ISLR 3.7.4:**

**Answer a.** Without knowing more details about the training data, it is difficult to know which training RSS is lower between linear or cubic. However, as the true relationship between  $x$  and  $y$  is linear, we may expect the least squares line to be close to the true regression line and consequently the RSS for the linear regression may be lower than the cubic regression.

**Answer b.** In this case the test RSS depends upon the test data, so we have not enough information to conclude. However, we may assume that polynomial regression will have a higher test RSS as the overfit from training would have more error than the linear regression.

**Answer c.** Polynomial regression has lower train RSS than the linear fit because of higher flexibility: no matter what the underlying true relationship is the more flexible model will closer follow points and reduce train RSS.

**Answer d.** There is not enough information to tell which test RSS would be lower for either regression given the problem statement is defined as not knowing "how far it is from linear". If it is closer to linear than cubic, the linear regression test RSS could be lower than the cubic regression test RSS. Or, if it is closer to cubic than linear, the cubic regression test RSS could be lower than the linear regression test RSS. It is due to bias-variance tradeoff: it is not clear what level of flexibility will fit data better.