Multiple Choice:

1. A and D
2. C

Short-Answer Questions:

Question 5:

5a. For each additional drink per day, the predicted average effect on average life expectancy is -3.27, holding smoker status fixed.

5c.

Smoker: 93.68 - 3.27alc - 23.44

Non-smoker: 93.68 - 3.27alc

The average difference in lifespan between smokers and non-smokers is 23.44 years. After taking into account the effects of alcohol consumption, non-smokers on average live 23.44 years longer.

Question 7:

7b. Reliability in this case is determined by meeting all the assumptions necessary for inference, which are zero mean, linearity, uniform spread, independence, normality, and randomness. In 6b, we used an overall ANOVA F-test. The full model in this case would be $\beta_0 + \beta_1 + \beta_2 + \beta_3$ and the reduced model would be just β_0. The assumptions that are already met in the full model would be zero mean, independence, and randomness. Looking at the residual plot, the data points look randomly scattered and generally have more uniform spread, so the linearity and uniform spread assumptions are met. Looking at the normal quantile plot, the data points are generally along the line, therefore, we can check off normality. By default, the assumptions would be met for the reduced model with just β_0. In 7a, we used a nested F-test where the full model was model 3 ($\beta_0 + \beta_1 + \beta_2 + \beta_3$) and the reduced model was model 1 ($\beta_0 + \beta_1$). As stated above, the assumptions were met for model 3, however, for model 1, not all the assumptions are met. If you look at the residual plot for model 1, there's a slight curvature to the data points, which violates the linearity assumption. Therefore, since all the model assumptions were not met for model 1, we can say that the test in 6b was more reliable.