Q1

$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$ ① $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$ ②

Plug $\bar{x}$ to ① $\hat{\beta}_0 + \hat{\beta}_1 \bar{x}$

Plug ② $\bar{y} - \hat{\beta}_1 \bar{x} + \hat{\beta}_1 \bar{x} = \bar{y}$

So $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$ go through $(\bar{x}, \bar{y})$

Q2 $\mathrm{Corr}(X,Y) = \dfrac{SS_{xy}}{\sqrt{SS_{xx}} \cdot \sqrt{SS_{yy}}}$

a) $RSS = SS_{YY} - \hat{\beta}_1^2 SS_{XX}$

$$= SS_{YY} - \frac{(SS_{XY})^2}{(SS_{XX})^2} \cdot SS_{XX}$$

$$= SS_{YY} - \frac{(SS_{XY})^2}{SS_{XX}}$$

$$= SS_{YY} - \frac{SS_{yy}}{SS_{yy}} \cdot \frac{(SS_{xy})^2}{SS_{xx}}$$

$$= SS_{yy}\left(1 - \frac{(SS_{xy})^2}{SS_{yy} \cdot SS_{xx}}\right)$$

$$= SS_{yy}\left(1 - \mathrm{Corr}(X,Y)^2\right)$$

b) $$\frac{\hat{\beta}_1}{\sqrt{\widehat{\mathrm{Var}}(\hat{\beta}_1 | X)}} = \frac{SS_{xy}}{SS_{xx}} \cdot \frac{1}{\sqrt{\frac{RSS}{n-2} \cdot \frac{1}{SS_{XX}}}}$$

$$= \frac{SS_{xy}}{SS_{xx}} \cdot \sqrt{n-2} \cdot \sqrt{SS_{XX}} \cdot \frac{1}{\sqrt{SS_{yy} \cdot (1 - \mathrm{Corr}(X,Y)^2)}}$$

$$= \frac{SS_{xy}}{SS_{xx}} \cdot \sqrt{SS_{xx}} \cdot \frac{1}{\sqrt{SS_{yy}}} \cdot \sqrt{n-2} \cdot \frac{1}{\sqrt{1 - \mathrm{Corr}(X,Y)^2}}$$

$$= \frac{SS_{XY}}{\sqrt{SS_{XX}} \cdot \sqrt{SS_{yy}}} \cdot \sqrt{n-2} \cdot \frac{1}{\sqrt{1 - \mathrm{Corr}(X,Y)^2}}$$

$$= \sqrt{n-2} \cdot \frac{\mathrm{Corr}(X,Y)}{\sqrt{1 - \mathrm{Corr}(X,Y)^2}}$$

Q3 $H_0: \beta_1 + 2\beta_2 = 4\beta_3$

$H_1: \beta_1 + 2\beta_2 \neq 4\beta_3$

$$t = \frac{\hat{\beta}_1 + 2\hat{\beta}_2 - 4\hat{\beta}_3}{\sqrt{\widehat{\mathrm{Var}}(\hat{\beta}_1 + 2\hat{\beta}_2 - 4\hat{\beta}_3 | X)}}$$

where $\widehat{\mathrm{Var}}(\hat{\beta}_1 + 2\hat{\beta}_2 - 4\hat{\beta}_3 | X)$

$= \widehat{\mathrm{Var}}(\hat{\beta}_1 | X) + 4\widehat{\mathrm{Var}}(\hat{\beta}_2 | X) + 16(\hat{\beta}_3 | X) + 4\mathrm{Cov}(\hat{\beta}_1, \hat{\beta}_2 | X) - 8\mathrm{Cov}(\hat{\beta}_1, \hat{\beta}_3 | X) - 16\mathrm{Cov}(\hat{\beta}_2, \hat{\beta}_3 | X)$

Q4

a) $E(\text{blood pressure} \mid X) = \beta_0 + \beta_1 \cdot X_{Age} + \beta_2 \cdot X_{weight} + \beta_3 \cdot X_{Diabetes\ Status}$

b) No. If we are going to plot the rolation between Diabetes Status and blood pressure, we won't see a quadratic trend since Diabetes Status is a categorical variable

Q5

a) $H_0: \beta_1 = 0$

b) No. To be able to if X_1 is quadratically related to Y, one must include X_1^2 in the model. So I would add $\beta_7 X_1^2$ term to the model

c) $H_0: \beta_5 = 0$

Q6

a) $n-(4+1)=42 \Rightarrow n=47$

b) male

c) female and income

d) When changing from male group to female group, we estime an average decrease of 22.11833 pounds expenditure on gambling by keeping other cofficients unchanged.

e) 52.67%

f) It's same as the unit of gamble variable, which is pounds

g) i $H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$

$H_a:$ At least one $\beta_i \neq 0$ $i = 1, 2, 3, 4$

ii test statistic $= 11.69$, p-value $= 1.815 \times 10^{-6}$

iii F distribution, $F_{4,42}$

iv Since p-value is less than 5%, we reject the null hypothesis. We can conclude that overall model is statistically significant.