(Q1) $K=2 \Rightarrow$ 迴歸式 $y = b_1 + b_2 x$

$$b_1 = \bar{y} - b_2\bar{x}, \quad b_2 = \frac{\Sigma(x_i-\bar{x})(y_i-\bar{y})}{\Sigma(x_i-\bar{x})^2}$$

$$b = (X'X)^{-1}X'Y$$

其中 $X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}$ $Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}$

① $X'X = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & 1 \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} = \begin{bmatrix} n & \Sigma x_i \\ \Sigma x_i & \Sigma x_i^2 \end{bmatrix}$

$X'Y = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} \Sigma Y_i \\ \Sigma x_i Y_i \end{bmatrix}$

$$(X'X)^{-1} = \frac{1}{n\Sigma x_i^2 - (\Sigma x_i)^2} \begin{bmatrix} \Sigma x_i^2 & -\Sigma x_i \\ -\Sigma x_i & n \end{bmatrix}$$

故 $b = \dfrac{1}{n\Sigma x_i^2 - (\Sigma x_i)^2} \begin{bmatrix} \Sigma x_i^2 & -\Sigma x_i \\ -\Sigma x_i & n \end{bmatrix} \begin{bmatrix} \Sigma Y_i \\ \Sigma x_i Y_i \end{bmatrix}$

$$= \frac{1}{n\Sigma x_i^2 - (\Sigma x_i)^2} \begin{bmatrix} \Sigma x_i^2 \Sigma Y_i - \Sigma x_i \Sigma x_i Y_i \\ -\Sigma x_i \Sigma Y_i + n\Sigma x_i Y_i \end{bmatrix}$$

故 $b_1 = \dfrac{\Sigma x_i^2 \Sigma Y_i - \Sigma x_i \Sigma x_i Y_i}{n\Sigma x_i^2 - (\Sigma x_i)^2} = \bar{y} - b_2\bar{x}$

$b_2 = \dfrac{n\Sigma x_i Y_i - \Sigma x_i \Sigma Y_i}{n\Sigma x_i^2 - (\Sigma x_i)^2}$ 得證

(Q2) (2.14) $Var(b_1|X)=\sigma^2\left[\frac{\sum x_i^2}{N\sum(x_i-\bar{x})^2}\right]$

(2.15) $Var(b_2|X)=\frac{\sigma^2}{\sum(x_i-\bar{x})^2}$

(2.16) $cov(b_1,b_2|X)=\sigma^2\left[\frac{-\bar{x}}{\sum(x_i-\bar{x})^2}\right]$

$Var(b)=\sigma^2(X'X)^{-1}$

令 $X=\begin{bmatrix}1 & x_1\\ 1 & x_2\\ \vdots & \vdots\\ 1 & x_n\end{bmatrix}$ $X'X=\begin{bmatrix}1 & 1 & \cdots & 1\\ x_1 & x_2 & \cdots & x_n\end{bmatrix}\begin{bmatrix}1 & x_1\\ 1 & x_2\\ \vdots & \vdots\\ 1 & x_n\end{bmatrix}=\begin{bmatrix}n & \sum x_i\\ \sum x_i & \sum x_i^2\end{bmatrix}$

$\sum(x_i-\bar{x})^2$
$=\sum(x_i^2-2x_i\bar{x}+\bar{x}^2)$

$(X'X)^{-1}=\frac{1}{n\sum x_i^2-(\sum x_i)^2}\begin{bmatrix}\sum x^2 & -\sum x_i\\ -\sum x_i & n\end{bmatrix}$

$Var(b)=\sigma^2(X'X)^{-1}=\sigma^2\frac{1}{n\sum x_i^2-(\sum x_i)^2}\begin{bmatrix}\sum x_i^2 & -\sum x_i\\ -\sum x_i & n\end{bmatrix}=\begin{bmatrix}\sigma_1 & \sigma_{12}\\ \sigma_{21} & \sigma_2\end{bmatrix}$.

$\sigma_1=Var(b_1)=\sigma^2\frac{\sum x_i^2}{n\sum x_i^2-(\sum x_i)^2}=\sigma\frac{\sum x_i^2}{n\sum(x_i-\bar{x})^2}$

$\sigma_2=Var(b_2)=\sigma^2\frac{1}{n\sum(x_i-\bar{x})^2}\cdot n=\frac{\sigma^2}{\sum(x_i-\bar{x})^2}$

$\sigma_{12}=Cov(b_1,b_2)=\sigma^2\frac{-n\bar{x}}{n\sum(x_i-\bar{x})^2}=\sigma^2\frac{-\bar{x}}{\sum(x_i-\bar{x})^2}$ 得證

5.3

(a) (i) $t=\frac{\hat{\beta}_1}{SE(\hat{\beta}_1)}=\frac{1.4515}{2.2019}=0.6592$

(ii) $se(b_2)=\frac{\hat{\beta}_2}{t}=\frac{2.7648}{5.7103}=0.4842$

(iii) $b_3=t\cdot SE(\hat{\beta}_3)=0.3695\cdot(-3.9376)=-1.4549$

(iv) $SD=\sqrt{\frac{TSS}{n-1}}\Rightarrow TSS=SD^2(n-1)=(6.39547)^2(1200-1)\doteqdot 49041.54$

$R^2=1-\frac{RSS}{TSS}=1-\frac{46221.62}{49041.54}=0.0575$

(v) $\hat{\sigma}=\sqrt{\frac{RSS}{n-k}}=\sqrt{\frac{46221.62}{1200-4}}=6.2167$

(b)

$$y = 1.4515 + 2.7648 \ln(TOTEXP) - 1.4549NK - 0.1503AGE$$

b_2：當家庭支出 +1%，預算中 alcohol + 2.7648%。
代表支出越高的家庭花在 alcohol 的支出越多。

b_3：當家庭多1個小孩，酒精支出會下降 0.3695%，可能因責任考量

b_4：當家庭主要負責者年紀大1歲，酒精支出 -0.1503%
可能因健康意識而減少飲用酒。

(c) $C.I.\ 95\%$ of $\beta_4 = b_4 \pm t_{0.025, df=1196} SE(b_4)$

$$= -0.1503 \pm 1.96 \times 0.0235$$

$$= (-0.1964, -0.1042)$$

區間內皆為負，代表在95%的信心水準下，
年紀較大的家庭負責者在酒精的支出

(d) 只有 b_1 的 t-statistic = 0.6592，其它皆顯著 $\Rightarrow t \geq |1.96|$

(e)
$$\begin{cases} H_0: \beta_3 = -2 \\ H_1: \beta_3 \neq -2 \end{cases}$$

$$t = \frac{\hat{\beta}_3 - (-2)}{SE(\beta_3)} = \frac{-1.4549 - (-2)}{0.3695} = 1.4752$$

$\because 1.4752 < 1.96$ $\therefore$ We can't reject $H_0: \beta_3 = -2$ in 95% C.I.

(5.23)

(a) 預期 b_1 quant 為負，因買越多折扣越多

b_2 qual 為正，越純成本越高

b_3 trend 不確定

(b)

```
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 90.84669    8.58025  10.588 1.39e-14 ***
quant       -0.05997    0.01018  -5.892 2.85e-07 ***
qual         0.11621    0.20326   0.572   0.5700
trend       -2.35458    1.38612  -1.699   0.0954 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

$b_1 < 0$，說明非法藥物也有折扣現象

$b_2 > 0$ 正影響但不顯著，說明品質差異對price影響有限。

$b_3 < 0$ 隨時間過去 1984~1991，價格↓，但只有90% C.I.下成立

(c) $R^2 = 0.5097$，說有51%的價格的變異，可以被 quant, qual, time 解釋

(d) $\begin{cases} H_0: \beta_2 = 0 \\ H_1: \beta_2 < 0 \end{cases}$

根據(b)結果，$b_1 = -0.05997$，且 $p = 2.85 \times 10^{-7} < 0.01$

故 reject H_0，quantity↑，price↓，賣方為了規避被抓的risk，故希望降價換取大宗交易。

(e) $\begin{cases} H_0: \beta_3 = 0 \\ H_1: \beta_3 > 0 \end{cases}$

$\because \beta_3$ p-value $= \frac{0.57}{2} = 0.285 > 0.05$ $\therefore$ we can't reject H_0

$\therefore$ 在 95 C.I. 下，no significant evidence 可說明 high purity 的 cocaine 賣比較貴。

(f) $\hat{\beta}_3 = -2.35458$，說明price每年平均下降2.35458美元，

可能是法規等因素限制毒品交易的發展。