Question 1

On voit clairement sur le nuage de points (circonference/hauteur) que cela suit une droite. On essaye de trouver les valeurs de la droites qui minimisent le risque quadratique.

Question 2

Pour minimiser la fonction $\varphi(\beta_1, \beta_2)$ il faut trouver dériver la fonction par rapport à $beta_1$ et $beta_2$ et trouver les valeurs qui annulent les 2 dérivées.

$$\frac{\partial\varphi(\beta_1,\beta_2)}{\beta_1} = \frac{\sum_{i=1}^{n}(Y_i-\beta_1-\beta_2 x_i)^2}{\beta_1} = \frac{\sum_{i=1}^{n} Y_i^2-\beta_1 Y_i-\beta_2 x_i Y_i-\beta_1 Y_i+\beta_1^2+\beta_1\beta_2 x_i-\beta_2 x_i Y_i+\beta_2 x_i\beta_1+\beta_2^2 x_i^2}{\beta_1}$$

$$= \sum_{i=1}^{n} -Y_i - Y_i + 2\beta_1 + \beta_2 x_i + \beta_2 x_i = 2\sum_{i=1}^{n} -Y_i + \beta_2 x_i + \beta_1$$

$$\frac{\partial\varphi(\beta_1,\beta_2)}{\beta_2} = \frac{\sum_{i=1}^{n}(Y_i-\beta_1-\beta_2 x_i)^2}{\beta_2} = \frac{\sum_{i=1}^{n} Y_i^2-\beta_1 Y_i-\beta_2 x_i Y_i-\beta_1 Y_i+\beta_1^2+\beta_1\beta_2 x_i-\beta_2 x_i Y_i+\beta_2 x_i\beta_1+\beta_2^2 x_i^2}{\beta_2}$$

$$= \sum_{i=1}^{n} -x_i Y_i + \beta_1 x_i - x_i Y_i + \beta_1 x_i + 2\beta_2 x_i^2 = 2\sum_{i=1}^{n} x_i(-Y_i + \beta_2 x_i + \beta_1)$$

On cherche $\hat{\beta_1}$ et $\hat{\beta_2}$ les valeurs qui annulent le système

$$\begin{cases} \sum_{i=1}^{n} x_i(-Y_i+\beta_2 x_i+\beta_1) = 0 \\ \sum_{i=1}^{n} -Y_i+\beta_2 x_i+\beta_1 = 0 \end{cases}$$

$$\begin{cases} \sum_{i=1}^{n} -Y_i x_i + \sum_{i=1}^{n} \beta_2 x_i^2 + \sum_{i=1}^{n} \beta_1 x_i = 0 & (1) \\ \sum_{i=1}^{n} -Y_i + \sum_{i=1}^{n} \beta_2 x_i + \sum_{i=1}^{n} \beta_1 = 0 & (2) \end{cases}$$

$$\begin{cases} \sum_{i=1}^{n} -Y_i x_i + \beta_2\sum_{i=1}^{n} x_i^2 + \beta_1\sum_{i=1}^{n} x_i = 0 & (1) \\ \sum_{i=1}^{n} -Y_i + \beta_2\sum_{i=1}^{n} x_i + n\beta_1 = 0 & (2) \end{cases}$$

On fait $(3) = n(1) - (2)\sum_{i=1}^{n} x_i$

$$\begin{cases} n\sum_{i=1}^{n} -Y_i x_i + n\beta_2\sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} Y_i \sum_{i=1}^{n} x_i - \beta_2\left(\sum_{i=1}^{n} x_i\right)^2 = 0 & (3) \\ \sum_{i=1}^{n} -Y_i + n\beta_2\sum_{i=1}^{n} x_i + n\beta_1 = 0 & (2) \end{cases}$$

$$\begin{cases} -n\sum_{i=1}^{n} Y_i x_i + \sum_{i=1}^{n} Y_i \sum_{i=1}^{n} x_i = \beta_2\left(\sum_{i=1}^{n} x_i\right)^2 - n\beta_2\sum_{i=1}^{n} x_i^2 & (3) \\ \sum_{i=1}^{n} -Y_i + n\beta_2\sum_{i=1}^{n} x_i + n\beta_1 = 0 & (2) \end{cases}$$

$$\begin{cases} \beta_2 = \frac{\sum_{i=1}^{n} Y_i \sum_{i=1}^{n} x_i - n\sum_{i=1}^{n} Y_i x_i}{\left(\sum_{i=1}^{n} x_i\right)^2 - n\sum_{i=1}^{n} x_i^2} & (3) \\ \beta_1 = \frac{1}{n}\left(\sum_{i=1}^{n} Y_i - n\beta_2\sum_{i=1}^{n} x_i\right) & (2) \end{cases}$$

$$\begin{cases} \beta_2 = \frac{\sum_{i=1}^{n} Y_i \sum_{i=1}^{n} x_i - n\sum_{i=1}^{n} Y_i x_i}{\left(\sum_{i=1}^{n} x_i\right)^2 - n\sum_{i=1}^{n} x_i^2} & (3) \\ \beta_1 = \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i Y_i - \sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} Y_i}{\left(\sum_{i=1}^{n} x_i\right)^2 - n\sum_{i=1}^{n} x_i^2} & (2) \end{cases}$$

Question 3

Voir Python On obtient $\beta_1 = 9.037475668452768$ et $\beta_2 = 0.257137855007109$.