Etapa 2

a) Os parâmetros β_0, β_1, β_2 são parâmetros das retas ajustadas por meio do método dos mínimos quadrados. Estas retas tem os pontos ^yi calculados sendo que a distância deles até os pontos da nuvem com o mesmo x (xi) seja a menor possível. O cálculo dos parâmetros pode ser feito a partir do próprio método dos mínimos quadrados.

$$L = \sum_{i=1}^{n} (yi - y^{i})^2$$

Que pode ser reescrita na forma:

$$L = \sum_{i=1}^{n} (yi - \beta0 - \beta1 * \mathrm{xi})^2$$

Este somatório é a soma de todas as diferenças entre os pontos reais y e a reta calculada ^y. Para encontrar as relações entre β0 e β1 derivamos o somatório em relação a β0 e β1.

$$\frac{\delta L}{\delta \beta 0}$$

$$\frac{\delta L}{\delta \beta 1}$$

Como queremos saber os pontos onde a distância entre yi e ^yi é a menor possível, igualamos as derivadas a 0.

$$\sum_{i=1}^{n} -2 * (yi - \beta0 - \beta1 * \mathrm{xi}) = 0$$

$$\sum_{i=1}^{n} -2 * (yi - \beta0 - \beta1 * \mathrm{xi}) * xi = 0$$

Assim chegamos em um sistema com duas equações (equações acima) e duas incógnitas (β0 e β1). Desenvolvendo uma das equações chegamos em um valor para β0 em função de β1, substituímos esse valor na outra equação do sistema, chegando em um valor para β1. Sendo n o número de elementos analisados.

$$\beta0 = \frac{\sum_{i=1}^{n} yi}{n} - \frac{\sum_{i=1}^{n} xi}{n}$$

$$\beta 1 = \frac{\sum_{i=1}^{n} yi - n * xmédio * ymédio}{\sum_{i=1}^{n} xi^2 - n * \ xmédio}$$

Assim podemos calcular os parâmetros das retas utilizando os dados da população em questão. Para calcular β2 basicamente teríamos outra reta para ajustar, então adicionaríamos uma equação e uma incógnita ao sistema citado.

b) O teste de hipóteses nesta análise se dá em relação à inclinação da reta ajustada aos dados. H0 seria então um ponto que consideramos que as variáveis explicativas (X1 e X2) não influenciam na variável resposta (Y) e H1 seria uma inclinação suficiente para considerarmos que as variáveis explicativas influenciam no valor de Y.

c) No caso do meu problema, se parâmetros β1 e β2 forem muito grandes, maior será a influência da renda per capita e consumo de energia per capita na emissão de CO_2 per capita, ou seja, para pouca adição em X1 e X2, fará com que Y se altere muito rápido. O contrário aconteceria se β1 e β2 fossem pequenos, indicando que Y não se altera muito quando se alteram X1 e X2. Lembrando que isso não indica correlação, isso indica a influência quantitativa entre X1 e X2 com Y.

d) Para realizar a regressão linear necessitamos fazer algumas suposições. Os erros dos dados coletados são independentes entre si e seguem uma distribuição normal. O ajuste precisa ser linear (a reta citada nos itens anteriores) e que a variância dos erros seja = $sigma^2$.
Poderíamos checar se os erros são independentes entre si, fazendo cálculos para xi escolhidos, e analisar se os erros são próximos do erro médio, ou visualmente, observando se a nuvem comporta-se aproximadamente da mesma forma em torno da reta ajustada. Poderíamos checar se eles formam uma normal realizando a distribuição deles e analisando visualmente. A variância poderia ser apenas calculada e comparada ao desvio padrão da população. E o ajuste linear poderia ser verificado visualmente, observando como a nuvem de observações se comporta.