Read data in your R file:

```r
Country <- c('United States', 'Canada', 'Australia', 'New Zealand', 'United Kingdom', 'Switzerland', 'Ireland', 'Iceland', 'Finland', 'West Germany', 'Netherlands', 'Greece', 'Austria', 'Belgium', 'Mexico', 'Italy', 'Denmark', 'France', 'Sweden', 'Spain', 'Norway')
Cigarette <- c(3900, 3350, 3220, 3220, 2790, 2780, 2770, 2290, 2160, 1890, 1810, 1800, 1770, 1700, 1680, 1510, 1500, 1410, 1270, 1200, 1090)
Coronary <- c(259.9, 211.6, 238.1, 211.8, 194.1, 124.5, 187.3, 110.5, 233.1, 150.3, 124.7, 41.2, 182.1, 118.1, 31.9, 114.3, 144.9, 144.9, 126.9, 43.9, 136.3)
```

Question 1

```r
plot(Cigarette, Coronary)
```

Based on the scatter plot, we could see there is an obvious positive correlation between Cigarette and Coronary.

Question 2

```r
fit <- lm(Coronary ~ Cigarette)
fit
```

```r
## Call:
## lm(formula = Coronary ~ Cigarette)
##
## ##
```
The fitted equation is Coronary = 29.45259 + 0.05568 * Cigarette

The coefficient of determination is the multiple R-squared which is 0.4957

The correlation between Cigarette and Coronary is 0.704081, the determination is the square of correlation coefficient.
The variance of residuals seem to be smaller as the fitted values increase, it is better to use weighted least square method.

Question 5

It is same with check if the slope of regression greater than 0. We could use the result of regression with the assumption the residual following a normal distribution.

```r
shapiro.test(resid(fit))
```

##

```r
plot(fitted(fit), resid(fit))
```
The p value is 0.2608 greater than 0.05, we cannot reject the null hypothesis and we could assume that the residuals follows a normal distribution.

```
summary(fit)$coefficients
```

From the table, we could see the p value for ttest is 0.0003676282, but this is a two side p value, we need divid it by 2 to get one side p value

```
summary(fit)$coefficients['Cigarette','Pr(|t|)']/2
```

So the p value is 0.0001838141, less than 0.01, reject the null hypothesis, the correlation is significantly positive. You can also get same result with following code.

```
cor.test(Cigarette, Coronary, alternative = 'greater')
```

The p value is 0.0001838, less than 0.01, so reject the null hypothesis, and we can conclude they have significant positive relationship.

Question 6

```
resid(fit)[Country == 'Greece']
```

```
12
## -88.48341
```

```
(Coronary - fitted(fit))[Country == 'Greece']
```

```
12
## -88.48341
```

```
```

```
## (Intercept)
## -88.48341
```

Question 7

```
plot(Cigarette, Coronary)
lines(Cigarette, fitted(fit))
```