Question 3

OLS multiple regression output from Excel is presented in the table below. The estimated regression equation is:

$\hat{Y} = -6420 + 9881.8$(Male) + 11645(Over 65 Years) + 1265(High Cost User)

```
Call:
lm(formula = Cost ~ Male + `Over 65 Years` + `High Cost User`,
    data = Question3_Data_High_Cost_Users)

Residuals:
    Min      1Q  Median      3Q     Max
-3467.8 -2646.8  -782.5   -15.0  9192.2

Coefficients:
                Estimate Std. Error t value       Pr(>|t|)
(Intercept)         -6420       1647  -3.899       0.000526 ***
Male                 9882       1490   6.633 0.000000285299 ***
`Over 65 Years`     11645       1458   7.989 0.000000008236 ***
`High Cost User`    12625       1419   8.900 0.000000000867 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4027 on 29 degrees of freedom
Multiple R-squared:  0.8405,    Adjusted R-squared:  0.824
F-statistic: 50.93 on 3 and 29 DF,  p-value: 0.00000000001121
```

The multiple regression output suggests that all three predictors (gender, age, and last year's cost) are statistically significant predictors of next year's cost.

Interpretation of partial slope coefficients is:

Male: The average next year's cost for males is $9,882 higher than the average for females.

Over 65 Years: The average next year's cost for individuals who are over 65 years of age is $11,645 higher than the average for individuals who are 65 years or younger.

High Cost user: The average next year's for high cost users is $12,625 higher than the average for other users.