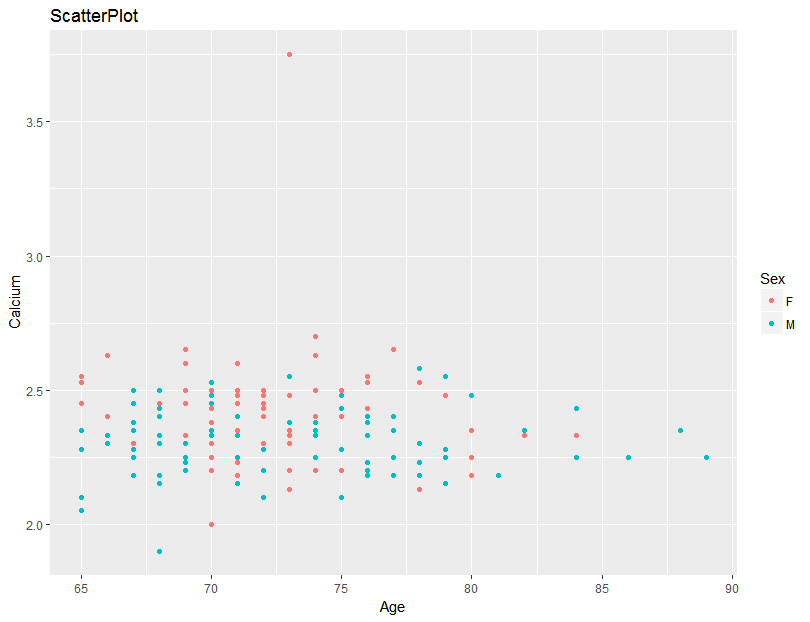
Question 2

1. Scatterplot of Calcium versus Age



Summary

1. This plot shows that blood calcium is negatively associated with age, because values of blood calcium decrease slightly while age increases.
2. Blood calcium tends to have a roughly linear relation with age.
3. Female tends to have a slightly higher blood calcium and male tends to have higher age.
4. There is a distinct outlying point with blood calcium value of 3.75.
5. Correlation Coefficient
6. Total sample: r = -0.0294977
7. Male: r = 0.02895113
8. Female r = -0.0659832

For all the correlation coefficient values above, no matter positive or negative, they are close to zero. This implies that there are no linear associations between age and blood calcium level for both males and females.

1. Simple Linear Regression Line

Summary:

Residuals:

Min 1Q Median 3Q Max

-0.41527 -0.07612 0.01281 0.08154 0.25835

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 2.2718930 0.1697104 13.387 <2e-16 \*\*\*

Age 0.0006379 0.0023347 0.273 0.785

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1224 on 89 degrees of freedom

Multiple R-squared: 0.0008382, Adjusted R-squared: -0.01039

F-statistic: 0.07466 on 1 and 89 DF, p-value: 0.7853

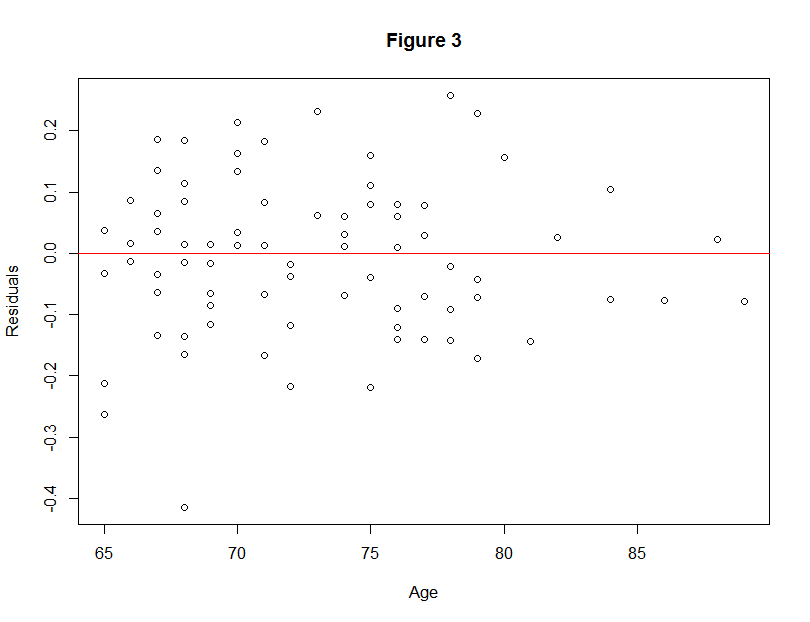
Regression equation:

Calcium: y

Age: x

**y = 0.0006379 \* x + 2.2718930**

1. Residual Plot



Residuals bounce around the zero line, and the variability is uniform, except one outlier with residual value of negative zero point four. This plot indicates that the relation between blood calcium level and age is roughly linear.

1. Coefficient of Determination

According to the summary table, coefficient of determination r^2 = 0.0008382

This implies only 0.08 % variability of values of blood calcium level can be explained by knowing their age. It also means that it’s a poor estimation of y’s values by incorporating x, as opposed to the mean value of y.

1. The estimated blood calcium level for an 83-year-old man is approximately 2.325
2. I don’t recommend using this calcium-age equation to estimate blood calcium level for male by age. There are several reasons.
3. Based on the scatterplot, blood calcium level for man doesn’t show obvious variance while age varies, which indicates there is no association between blood calcium level and age for male.
4. The correlation coefficient between male’s blood calcium level and age is close to zero, moreover, if subtract the outlier, the correlation coefficient will be even close to zero. All these values imply that there is no linear relationship between these two variables.
5. The slope of regression line is 0.0006, in another interpretation, the line is nearly a horizontal line. This implies there is no association between blood calcium level and age for male.
6. Although residuals bounce around the residual line evenly, the variability of these residual is very large, which indicates the sample values vary largely from estimated values.
7. Most importantly, the coefficient of determination shows that only 0.008% y values can be explained by knowing x values. Therefore, it is a poor estimation of males’ blood calcium level by incorporating their ages.

AgeGroup

Sex 65-69 70-74 75-79 80-84 85-89 Sum

F 21 46 15 5 0 87

M 35 24 23 6 3 91

Sum 56 70 38 11 3 178

Marginal distribution:

Male: 51.12%

Female: 48.88%