ST 503 Homework 1

Eric Warren

2024-01-18

Contents

1 Problem 1

1.1 Part A

Consider the analysis of covariance (ANCOVA) model: $y_{ij} = \mu + \alpha_i + x_{ij}\beta + e_{ij}$ for i = 1, 2, 3 and j = 1, . . . , n. Write the model in matrix form, clearly specifying all model components.

- We know that the **y** is a (ij x 1) matrix. Since i goes to 3 and j goes to n then **y** is a (3n x 1) matrix, where we start with y_{11} to y_{1n} then go y_{21} to y_{2n} and then go y_{31} to y_{3n}. Therefore, we can set up

$\mathbf{y} = \begin{pmatrix} y_{11} \\ ... \\ y_{1n} \\ y_{21} \\ ... \\ y_{2n} \\ y_{31} \\ ... \\ y_{3n} \end{pmatrix}$.

- We know that the **e** is a (ij x 1) matrix, as it is very similar to **y** (just this time showing the errors for each value). Since i goes to 3 and j goes to n then **e** is a (3n x 1) matrix, where we start with e_{11} to e_{1n} then go e_{21} to e_{2n} and then go e_{31} to e_{3n}. Therefore, we can set up $\mathbf{e} = \begin{pmatrix} e_{11} \\ ... \\ e_{1n} \\ e_{21} \\ ... \\ e_{2n} \\ e_{31} \\ ... \\ e_{3n} \end{pmatrix}$.
- We also know that β is a (p x 1) matrix where p is the number of predictor variables in the model. In this case we have $\mu, \alpha_1, \alpha_2, \alpha_3, \beta$, which is a total of 5 predictors. Thus, we should have a (5 x 1) matrix with the predictors in order. It is set up as $\beta = \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \beta \end{pmatrix}$.
- Lastly we know that **X** is a (ij x p) matrix where where p is the number of predictor variables in the model. Since i goes to 3 and j goes to n then ij is 3n and we have $\mu, \alpha_1, \alpha_2, \alpha_3, \beta$, which is a total of 5 predictors. Thus, we should have a (3n x 5) matrix. Our matrix will be all 1's for the first column as all of the y values use μ as a predictor. 0's and 1's will be possible values for the next 3 columns (0 if it does not use the α_i predictor and 1 if we do use the α_i predictor which is know by matching the index i values). The last column to represent the predictor β is going to be the x_{ij} value it has. For example if we want to represent y_{11} then we should use x_{11} for that last column value in the matrix. So based on all of the things we have seen we can model the matrix $\mathbf{X} = \begin{pmatrix} 1 & 1 & 0 & 0 & x_{11} \\ ... & & & & \\ 1 & 1 & 0 & 0 & x_{1n} \\ 1 & 0 & 1 & 0 & x_{21} \\ ... & & & & \\ 1 & 0 & 1 & 0 & x_{2n} \\ 1 & 0 & 0 & 1 & x_{31} \\ ... & & & & \\ 1 & 0 & 0 & 1 & x_{3n} \end{pmatrix}$.

Therefore our final equation in matrix format $\mathbf{y} = \mathbf{X}\beta + \mathbf{e}$ where **y** is a (3n x 1) matrix, **X** is a (3n x 5) matrix, β is a (5 x 1) matrix, and **e** is a (3n x 1) matrix is

$$\begin{pmatrix} y_{11} \\ \dots \\ y_{1n} \\ y_{21} \\ \dots \\ y_{2n} \\ y_{31} \\ \dots \\ y_{3n} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 & x_{11} \\ \dots & & & & \\ 1 & 1 & 0 & 0 & x_{1n} \\ 1 & 0 & 1 & 0 & x_{21} \\ \dots & & & & \\ 1 & 0 & 1 & 0 & x_{2n} \\ 1 & 0 & 0 & 1 & x_{31} \\ \dots & & & & \\ 1 & 0 & 0 & 1 & x_{3n} \end{pmatrix} \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \beta \end{pmatrix} + \begin{pmatrix} e_{11} \\ \dots \\ e_{1n} \\ e_{21} \\ \dots \\ e_{2n} \\ e_{31} \\ \dots \\ e_{3n} \end{pmatrix}$$

1.2 Part B

Is the model matrix **X** full column rank? Explain.

We know that matrix **X** does **NOT** have full column rank. We know this because there is linear dependence between the columns. We can say that the first column is equal to the second column plus the third column plus the fourth column. We can double check this and see this is true. We either get $1 = 1 + 0 + 0$, $1 = 0 + 1 + 0$, or $1 = 0 + 0 + 1$; which means linear dependence is present. For this reason the matrix **X** does **NOT** have full column rank. The rank of this matrix is actually just 4 instead of the full of 5 because of the linear combination we can make with the first column is equal to the second column plus the third column plus the fourth column.

2 Problem 2

Consider the teen gambling data, `teengamb`, in the R package `faraway`.

2.1 Read in data

Here we are going to read in the data to allow manipulation and solving this problem. We are also going to see the first couple of entries to get an idea of the data.

```
# install.packages(c("faraway", "tidyverse")) # Uncomment if you do not have packages
library(faraway)
library(tidyverse)
(
  gambling <- as_tibble(teengamb) %>%
    mutate(sex = ifelse(sex == 0, "Male",
                        ifelse(sex == 1, "Female", "Error")))
)
```

```
## # A tibble: 47 x 5
##    sex    status income verbal gamble
##    <chr>   <int>  <dbl>  <int>  <dbl>
##  1 Female     51   2          8   0
##  2 Female     28   2.5        8   0
##  3 Female     37   2          6   0
##  4 Female     28   7          4   7.3
##  5 Female     65   2          8  19.6
##  6 Female     61   3.47       6   0.1
##  7 Female     28   5.5        7   1.45
```

```
##  8 Female     27   6.42      5   6.6
##  9 Female     43   2         6   1.7
## 10 Female     18   6         7   0.1
## # i 37 more rows
```

```
nrow(gambling %>% filter(sex == "Error")) == 0 # Check to make sure the number of "Error" values is 0; i
```

```
## [1] TRUE
```

2.2 Part A

Write a brief description of the dataset. Produce some numerical and graphical summaries of the dataset.

The dataset contains 47 rows and 5 columns. This was gathered from a survey done to study teenage gambling in Britain. The data consists of the following columns:

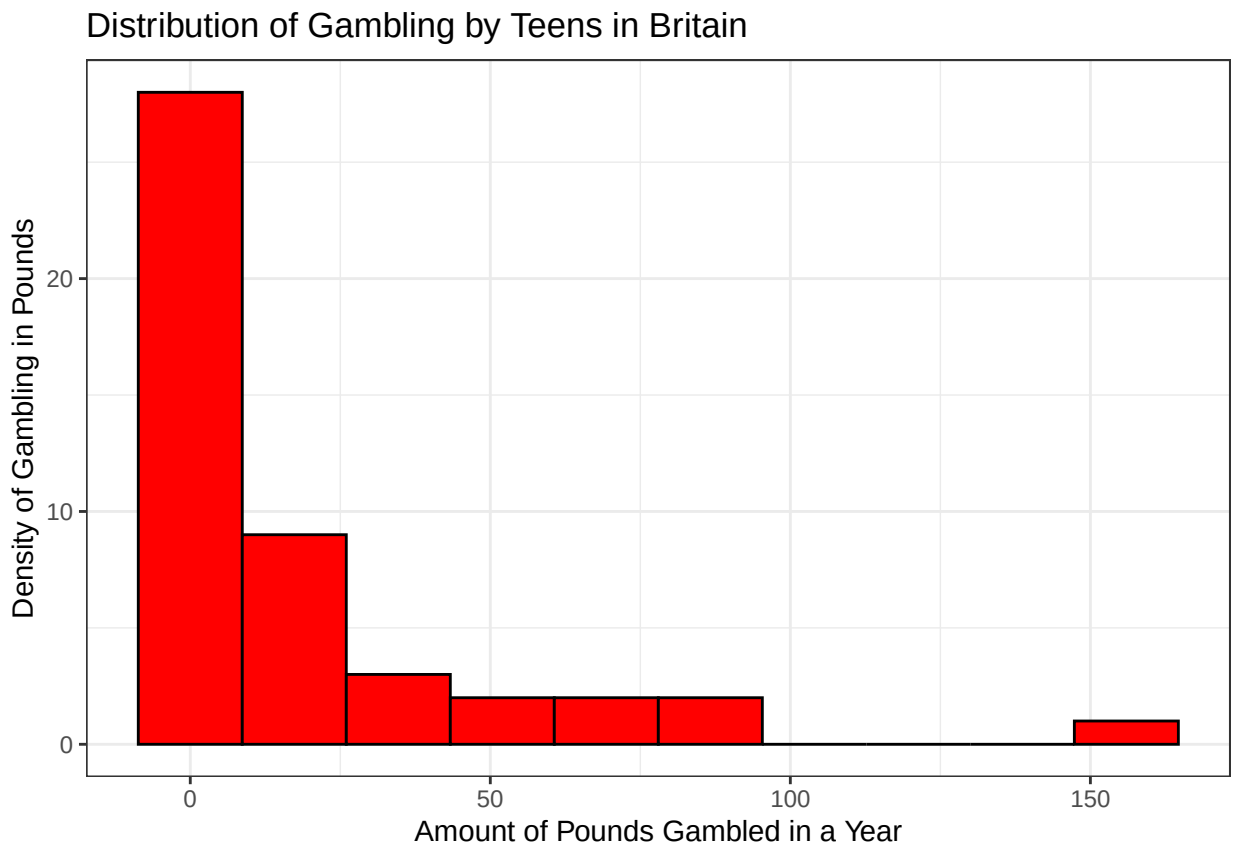
- `sex` (binary): Male (shown as 0 before manipulation) and Female (shown as 1 before manipulation)
- `status` (number): Socioeconomic status score based on parents' occupation
- `income` (number): Income in pounds per week
- `verbal` (number): Verbal score in words out of 12 correctly defined
- `gamble` (number): Expenditure on gambling in pounds per year

Let us first take a look at the numeric summary of gambling amounts per year and make a histogram showing this distribution.

```
# Numeric summary
summary(gambling$gamble)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     0.0     1.1     6.0    19.3    19.4   156.0
```

```
# Histogram showing distribution
gambling %>%
  ggplot(aes(x = gamble)) +
  geom_histogram(bins = 10, fill = "red", color = "black") +
  labs(title = "Distribution of Gambling by Teens in Britain",
       x = "Amount of Pounds Gambled in a Year",
       y = "Density of Gambling in Pounds") +
  theme_bw()
```



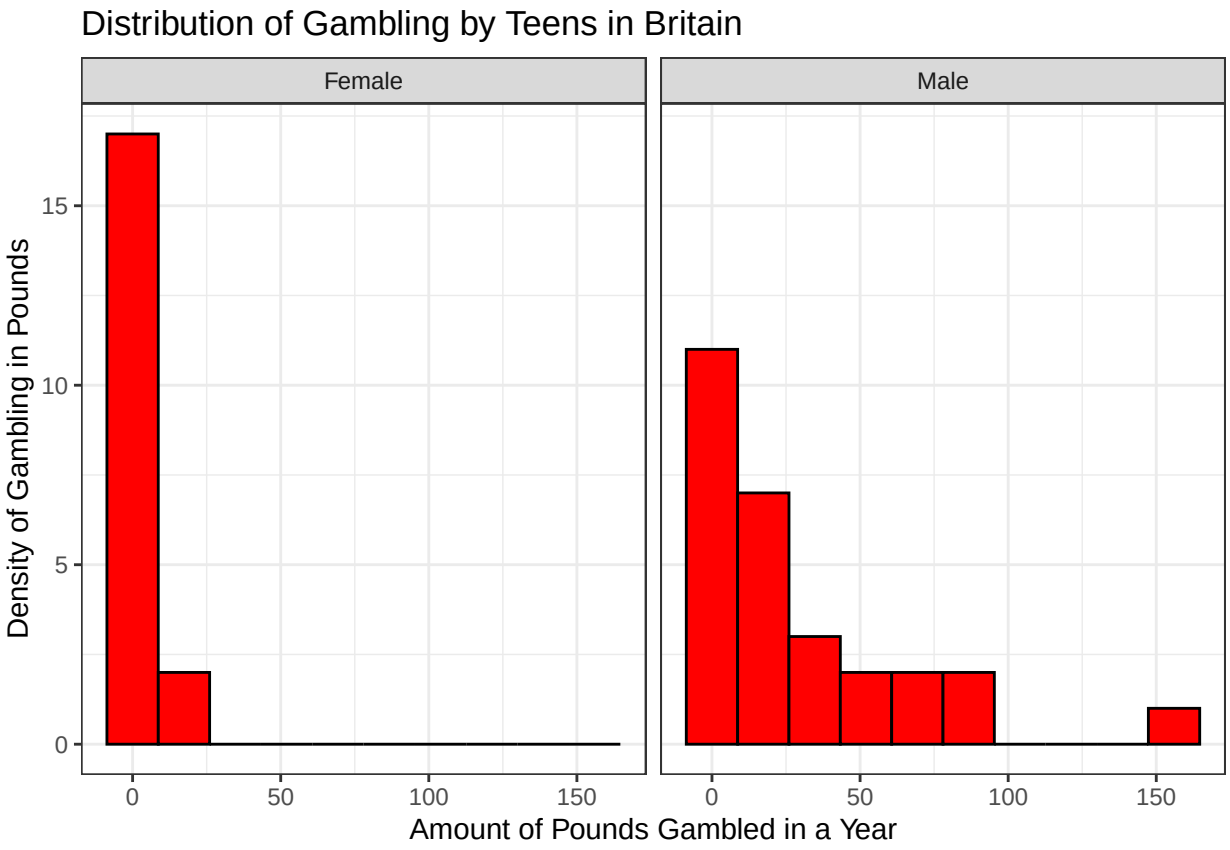
Our summary and histogram both support that gambling in Britain is represented by a skewed right distribution. This means we have some high outliers in the amount of pounds gambled by teens, but in all reality most teens do not gamble much at all. These high outliers support why the mean is so much higher than the median.

Let us see if there is a pattern in gambling between females and males? Does one sex look to gamble more than the others?

```
# Numeric summary
gambling %>%
  group_by(sex) %>%
  summarize(min = min(gamble),
            q1 = quantile(gamble, 0.25),
            median = median(gamble),
            mean = mean(gamble),
            q3 = quantile(gamble, 0.75),
            max = max(gamble)
            )
```

```
## # A tibble: 2 x 7
##   sex      min    q1 median  mean    q3   max
##   <chr>  <dbl> <dbl>  <dbl> <dbl> <dbl> <dbl>
## 1 Female     0  0.1     1.7  3.87   6    19.6
## 2 Male       0  2.78   14.2 29.8   42.2 156
```

```r
# Histogram showing distribution
gambling %>%
  ggplot(aes(x = gamble)) +
  geom_histogram(bins = 10, fill = "red", color = "black") +
  labs(title = "Distribution of Gambling by Teens in Britain",
       x = "Amount of Pounds Gambled in a Year",
       y = "Density of Gambling in Pounds") +
  facet_wrap(~ sex) +
  theme_bw()
```

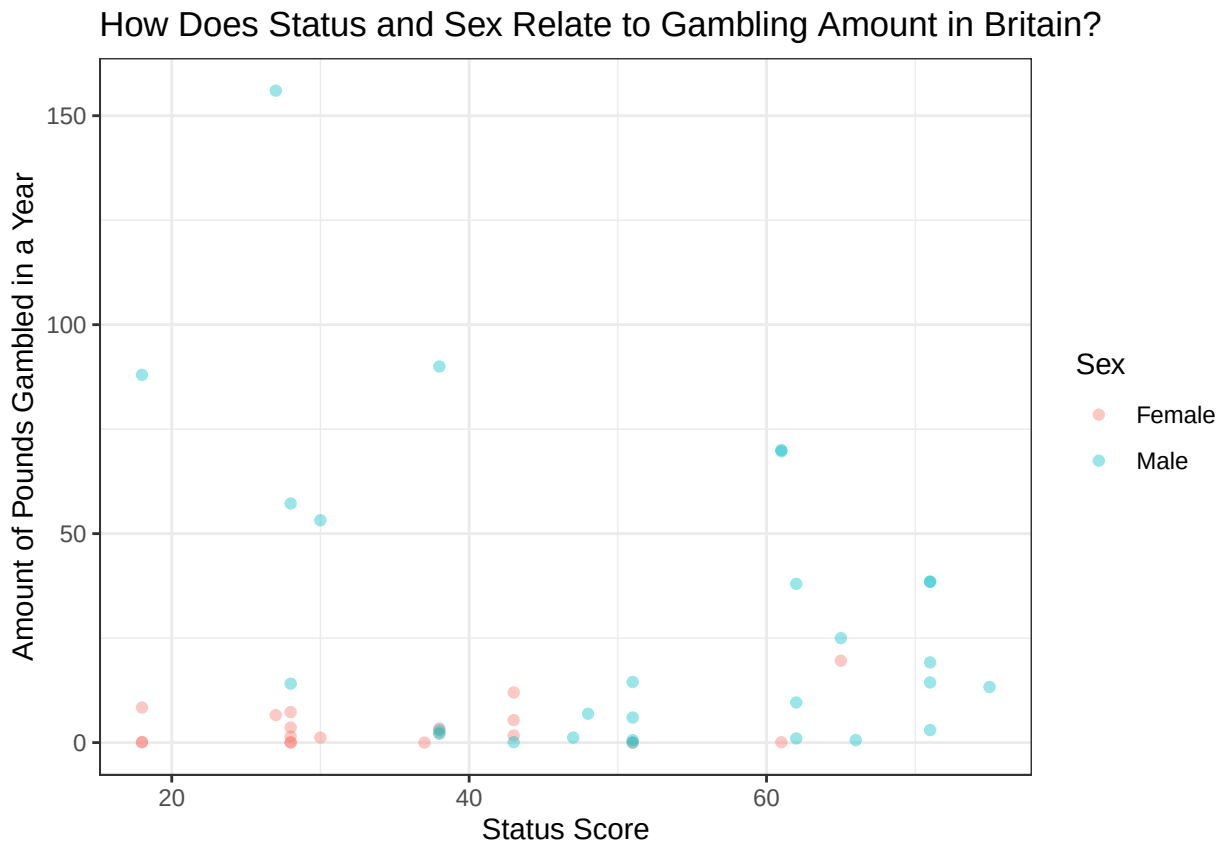


From this, we can see that males tend to gamble more than their female counterparts. We can hypothesis that `sex` is an important predictor for gambling amounts.

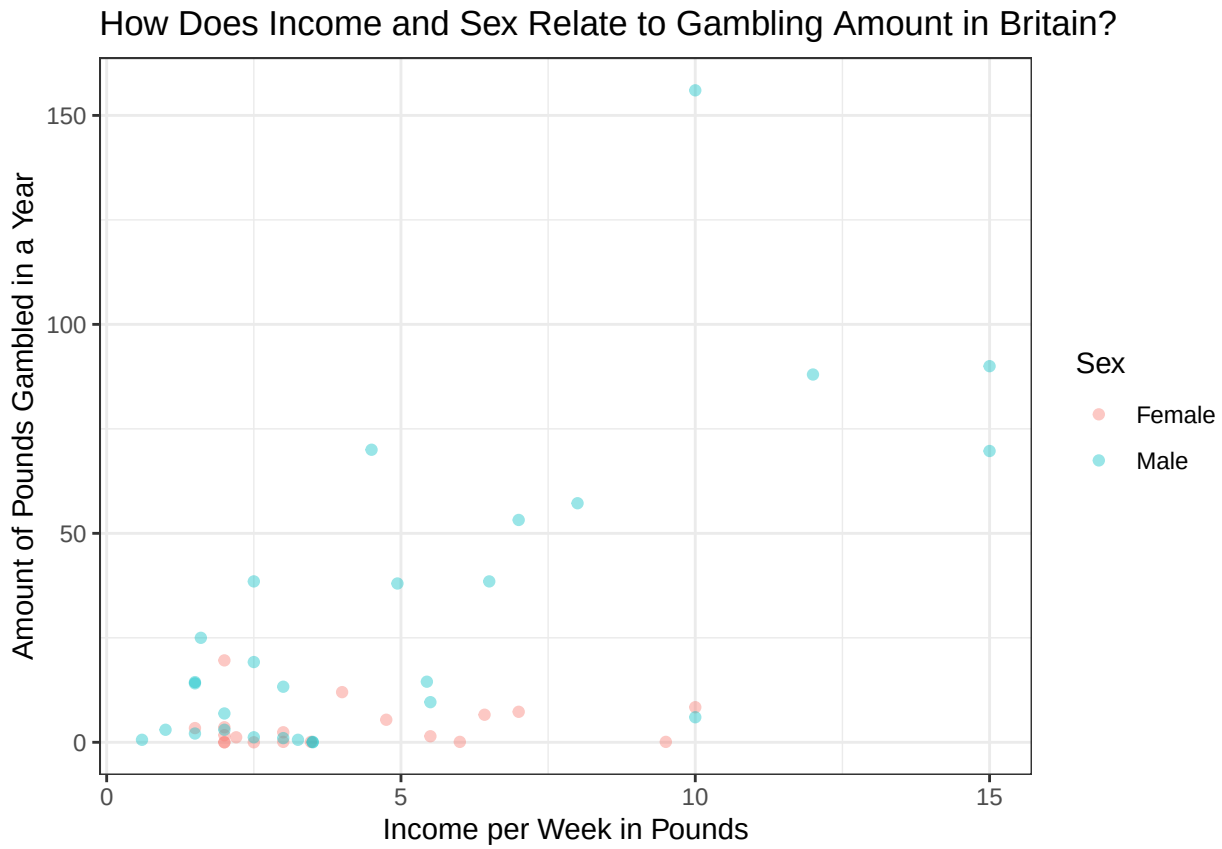
We are also going to view some scatterplots looking at `status` compared to `gamble`, `income` compared to `gamble`, and `verbal` compared to `gamble` to see if there are any trends with that. With all 3 we are going to add a grouping variable of males and females since we have already seen that `sex` seems to be a strong predictor.

```r
# Scatterplot of status and gamble
gambling %>%
  ggplot(aes(x = status, y = gamble, color = sex)) +
  geom_point(alpha = 0.4) +
  labs(title = "How Does Status and Sex Relate to Gambling Amount in Britain?",
       x = "Status Score",
       y = "Amount of Pounds Gambled in a Year",
       color = "Sex") +
```

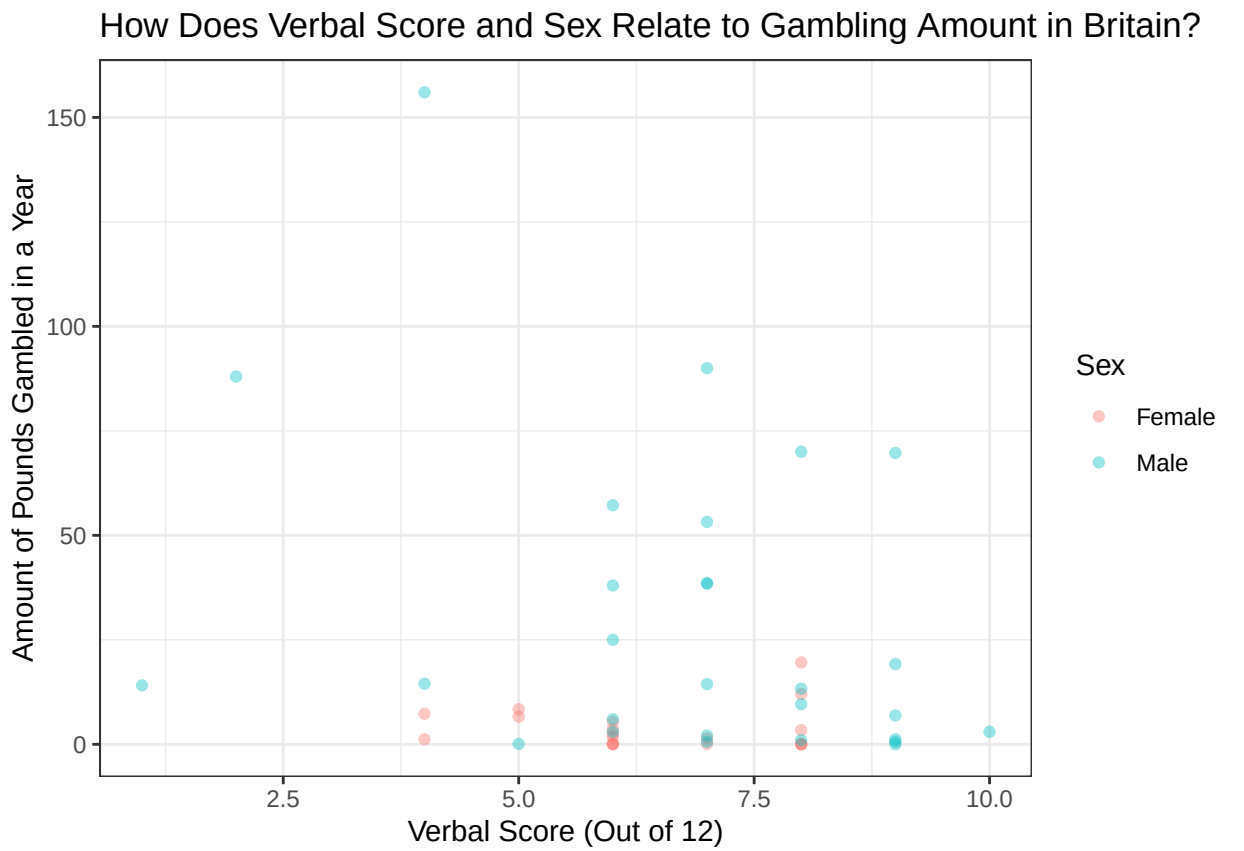
```
  theme_bw()
```




```
# Scatterplot of income and gamble
gambling %>%
  ggplot(aes(x = income, y = gamble, color = sex)) +
  geom_point(alpha = 0.4) +
  labs(title = "How Does Income and Sex Relate to Gambling Amount in Britain?",
       x = "Income per Week in Pounds",
       y = "Amount of Pounds Gambled in a Year",
       color = "Sex") +
  theme_bw()
```




```
# Scatterplot of verbal score and gamble
gambling %>%
  ggplot(aes(x = verbal, y = gamble, color = sex)) +
  geom_point(alpha = 0.4) +
  labs(title = "How Does Verbal Score and Sex Relate to Gambling Amount in Britain?",
       x = "Verbal Score (Out of 12)",
       y = "Amount of Pounds Gambled in a Year",
       color = "Sex") +
  theme_bw()
```



So what are some of the major takeaways?

- The males tend to live in families with higher statuses. This is interesting to note. Is there a `sex` bias (or discrimination involved or just the "random sample" collected)?
- We can also see that males are gambling more than females in all the charts.
- Is `income` a good predictor? There seems to be a similarity that a higher income might tend to gamble more (which makes sense since they have more money to gamble with), but we can also see this trend is caused by the males. The females seem to not have this at all. So do more rich males make this trend? Does it actually exist? Are males much wealthier than females in Britain? (This last point seems to be shown in the `income` plot.) All questions we hope to answer later.
- We can also see that `status` and `verbal` scores do not seem to be good indicators of one's gambling habits. We expect to not use these variables in our modeling.

2.3 Part B

Fit a linear model using the `lm()` function with the `gamble` variable as the response, the `income` variable as the predictor, and report the regression coefficients.

Here we are going to fit a simple linear regression model where the `gamble` variable is the response and the `income` variable is the predictor.

```
# Fit the model
gambleIncomeLR <- lm(gamble ~ income, gambling)

# Get the coefficient values
summary(gambleIncomeLR)
```

```
##
## Call:
## lm(formula = gamble ~ income, data = gambling)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -46.020 -11.874  -3.757  11.934 107.120
##
## Coefficients:
##             Estimate Std. Error t value   Pr(>|t|)
## (Intercept)   -6.325      6.030  -1.049        0.3
## income         5.520      1.036   5.330 0.00000305 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 24.95 on 45 degrees of freedom
## Multiple R-squared:  0.387,  Adjusted R-squared:  0.3734
## F-statistic: 28.41 on 1 and 45 DF,  p-value: 0.000003045
```

Here we get the regression coefficients to be -6.324559 as our intercept and 5.5204853 as our slope value for the `income` predictor variable. This means we expect someone with no income to gamble -6.32 pounds per year (which makes no sense – since you cannot gamble negative money) and for every 1 more pound in weekly income we expect them to gamble an additional 5.52 pounds per year.

2.4 Part C

Write the mathematical form of the model you fit in **Part B**. Clearly define each component in your model.

We know from **Part B** β_0 = -6.324559 and β_1 = 5.5204853.

- We also know that simple linear regression models for making predictions are in the form $\hat{y}_i = \beta_0 + \beta_1 x_i$ where x_i is the income amount in pounds per week, $\hat{y}_i$ is the predicted amount the British teen gambles in a year, β_0 is the intercept value (amount someone gambles with no income), and β_1 is the additional amount of pounds one gambles per year for each additional pound in weekly income. Thus, we can plug in our β_0 and β_1 values to get for prediction purposes we can write our simple linear regression model as $\hat{y}_i$ = -6.324559 + 5.5204853 x_i where x_i is the income amount in pounds per week and $\hat{y}_i$ is the predicted amount the British teen gambles in a year.
- For simple linear regression models in general form (with the error term) $y_i = \beta_0 + \beta_1 x_i + e_i$ where x_i is the income amount in pounds per week, y_i is the actual amount the British teen gambles in a year, e_i is the error term for how much the predicted gambling amount was off from the actual amount, β_0 is the intercept value (amount someone gambles with no income), and β_1 is the additional amount of pounds one gambles per year for each additional pound in weekly income. Thus, we can plug in our β_0 and β_1 values to get our general form for our model (with the error term) we can say y_i = -6.324559 + 5.5204853 $x_i + e_i$ where where x_i is the income amount in pounds per week, y_i is the actual amount the British teen gambles in a year, and e_i is the error term for how much the predicted gambling amount was off from the actual amount.

2.5 Part D

Further numerical investigation: compute the mean and standard deviation of `gamble` and `income` for males (`sex` = 0) and females (`sex` = 1) separately. Comment on the results.

Here we are going to show the mean and standard deviations as asked above.

```
gambling %>%
  group_by(sex) %>%
  summarize(mean_gamble = mean(gamble),
            sd_gamble = sd(gamble),
            mean_income = mean(income),
            sd_income = sd(income)
            )
```

```
## # A tibble: 2 x 5
##   sex    mean_gamble sd_gamble mean_income sd_income
##   <chr>        <dbl>     <dbl>       <dbl>     <dbl>
## 1 Female        3.87      5.15        4.15      2.60
## 2 Male         29.8      37.3         4.98      4.09
```

As we can see, the mean and standard deviations of the `income` are fairly close and we could say that males and females might make about the same. While the average of income is higher for males, there is a much higher standard deviation showing there might be more outliers who is making more. Now when we look at the `gamble` means and standard deviations, this is wildly drastic. While females tend to gamble less, more tend to be around the same with the smaller standard deviation (way less outliers). The male average gambling amount is a lot higher, but this is due to some pretty high outliers in gambling habits (since the standard deviation is much higher too). This shows that if you have a "high roller" (or large gambler), it is more likely to be a male since the high outliers are males. While the inclination is to say males gamble significantly more, we would have to do some more testing on that to see if that is true or if it is some high gambling males who make it seem that way.

2.6 Part E

Fit the same linear regression as in **Part B**, but separately for male and females. Report the regression coefficients.

First, we are going to separate the data into males and females.

```
# Get male data
(
  males <- gambling %>%
    filter(sex == "Male")
)
```

```
## # A tibble: 28 x 5
##    sex   status income verbal gamble
##    <chr>  <int>  <dbl>  <int>  <dbl>
##  1 Male      51   3.5        9    0
##  2 Male      62   3          8    1
##  3 Male      47   2.5        9    1.2
##  4 Male      43   3.5        5    0.1
##  5 Male      27  10          4  156
##  6 Male      71   6.5        7   38.5
##  7 Male      38   1.5        7    2.1
##  8 Male      51   5.44       4   14.5
##  9 Male      38   1          6    3
## 10 Male      51   0.6        7    0.6
## # i 18 more rows
```

```r
# Get female data
(
  females <- gambling %>%
    filter(sex == "Female")
)
```

```
## # A tibble: 19 x 5
##    sex    status income verbal gamble
##    <chr>   <int>  <dbl>  <int>  <dbl>
##  1 Female     51   2         8   0
##  2 Female     28   2.5       8   0
##  3 Female     37   2         6   0
##  4 Female     28   7         4   7.3
##  5 Female     65   2         8  19.6
##  6 Female     61   3.47      6   0.1
##  7 Female     28   5.5       7   1.45
##  8 Female     27   6.42      5   6.6
##  9 Female     43   2         6   1.7
## 10 Female     18   6         7   0.1
## 11 Female     18   3         6   0.1
## 12 Female     43   4.75      6   5.4
## 13 Female     30   2.2       4   1.2
## 14 Female     28   2         6   3.6
## 15 Female     38   3         6   2.4
## 16 Female     38   1.5       8   3.4
## 17 Female     28   9.5       8   0.1
## 18 Female     18  10         5   8.4
## 19 Female     43   4         8  12
```

Here we are going to fit a simple linear regression model where the `gamble` variable is the response and the `income` variable is the predictor for the **male** data first.

```r
# Fit the model
gambleIncomeMalesLR <- lm(gamble ~ income, males)

# Get the coefficient values
summary(gambleIncomeMalesLR)
```

```
##
## Call:
## lm(formula = gamble ~ income, data = males)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -56.522 -16.402  -2.342   7.901  93.478
##
## Coefficients:
##             Estimate Std. Error t value  Pr(>|t|)
## (Intercept)   -2.660      8.021  -0.332     0.743
## income         6.518      1.255   5.195 0.0000201 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 26.64 on 26 degrees of freedom
## Multiple R-squared:  0.5093, Adjusted R-squared:  0.4905
## F-statistic: 26.99 on 1 and 26 DF,  p-value: 0.0000201
```

Here we get the regression coefficients to be -2.6596291 as our intercept and 6.5181197 as our slope value for the `income` predictor variable. This means we expect males with no income to gamble -2.66 pounds per year (which makes no sense – since you cannot gamble negative money) and for every 1 more pound in weekly income we expect males to gamble an additional 6.52 pounds per year.

Now we are going to fit a simple linear regression model where the `gamble` variable is the response and the `income` variable is the predictor for the **female** data.

```
# Fit the model
gambleIncomeFemalesLR <- lm(gamble ~ income, females)

# Get the coefficient values
summary(gambleIncomeFemalesLR)
```

```
##
## Call:
## lm(formula = gamble ~ income, data = females)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -4.702 -3.527 -1.790  1.883 16.110
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.1400     2.3273   1.349    0.195
## income        0.1749     0.4789   0.365    0.719
##
## Residual standard error: 5.279 on 17 degrees of freedom
## Multiple R-squared:  0.007786,   Adjusted R-squared:  -0.05058
## F-statistic: 0.1334 on 1 and 17 DF,  p-value: 0.7194
```

Here we get the regression coefficients to be 3.1399737 as our intercept and 0.1749176 as our slope value for the `income` predictor variable. This means we expect females with no income to gamble 3.14 pounds per year and for every 1 more pound in weekly income we expect females to gamble an additional 0.17 pounds per year.

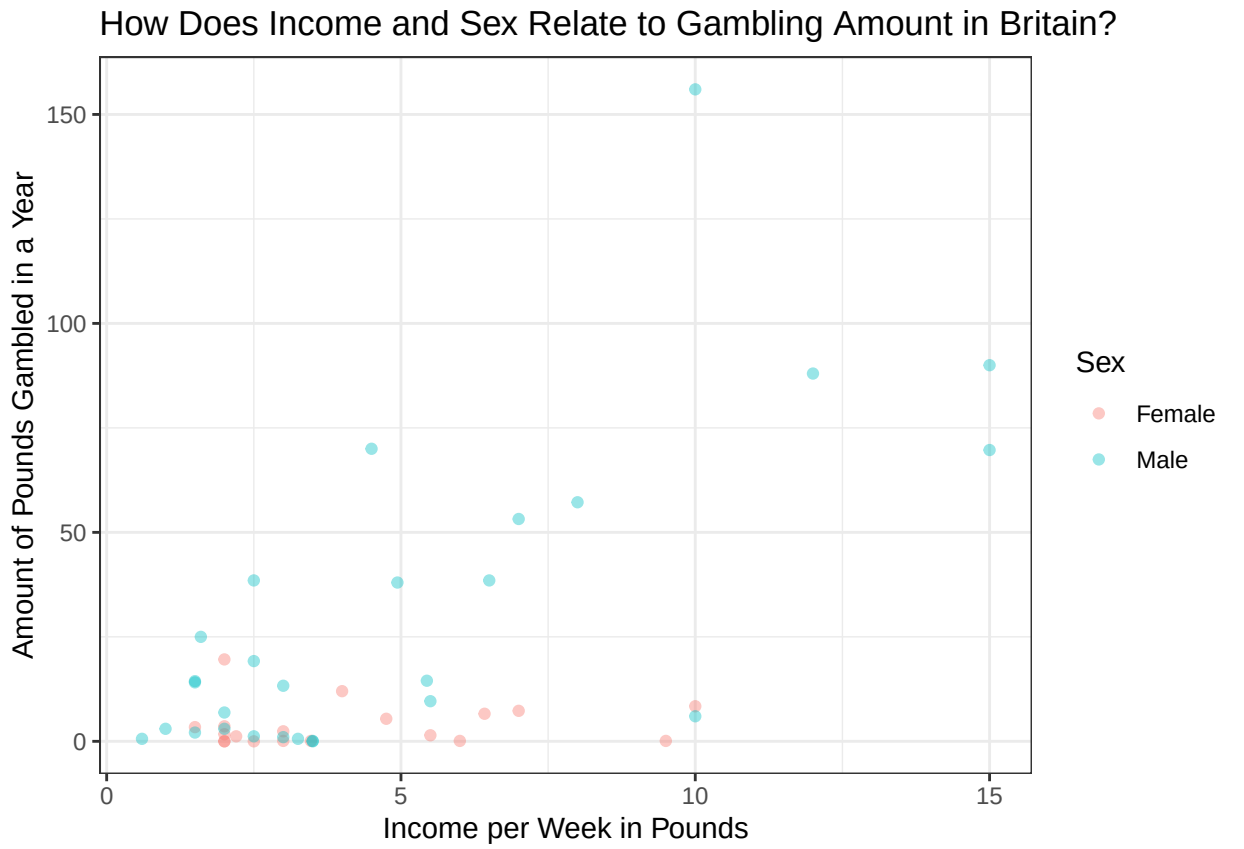
Something to note from this analysis: the `income` variable was significant for the males but not the females. This means that we can say just from this model that males with more money tend to gamble more, while females with more money tend to have similar gambling habits (in amount) as females with not as much money. So we can say `income` is a good predictor or gambling amounts for males but not for females.

2.7 Part F

Create a scatterplot between `gamble` (in y axis) and `income` (x axis), and color the points by `sex`. Then add two fitted regression lines from **Part E** to the plot. Comment on the results.

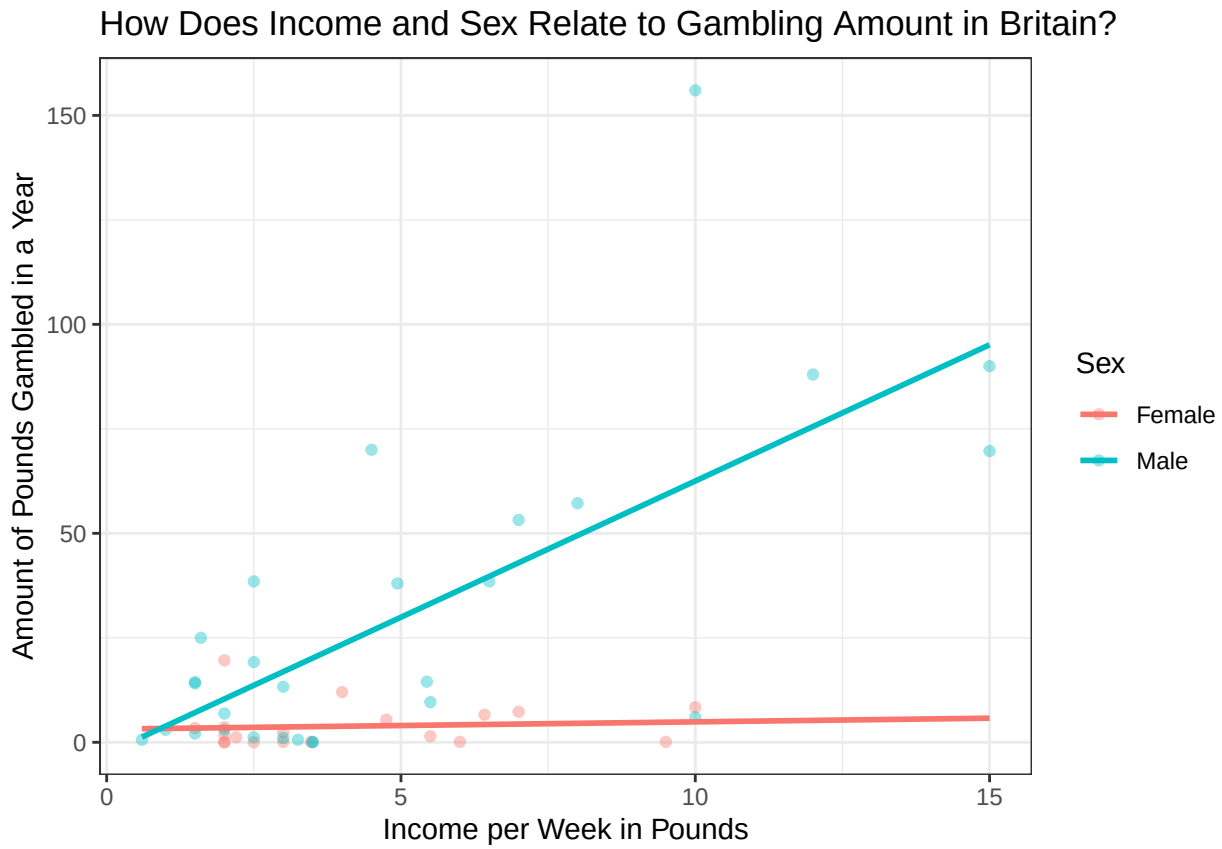
We are going to make our scatterplot that shows the relationship between `gamble` (in y axis) and `income` (x axis) with the grouping variable `sex` (as color).

```r
# Scatterplot of income and gamble
gambling %>%
  ggplot(aes(x = income, y = gamble, color = sex)) +
  geom_point(alpha = 0.4) +
  labs(title = "How Does Income and Sex Relate to Gambling Amount in Britain?",
       x = "Income per Week in Pounds",
       y = "Amount of Pounds Gambled in a Year",
       color = "Sex") +
  theme_bw()
```



Again we can see this relationship that males with more income tend to gamble more, while this trend does not seem to apply for females very much. Now we are going to add the two fitted regression lines (one for male the other for female) from **Part E** to the plot.

```r
# Scatterplot of income and gamble
gambling %>%
  ggplot(aes(x = income, y = gamble, color = sex)) +
  geom_point(alpha = 0.4) +
  geom_smooth(method = "lm",
              se = FALSE,
              fullrange = TRUE) +
  labs(title = "How Does Income and Sex Relate to Gambling Amount in Britain?",
       x = "Income per Week in Pounds",
       y = "Amount of Pounds Gambled in a Year",
       color = "Sex") +
  theme_bw()
```



Again adding this line makes us strongly believe that females tend to gamble the same amount regardless of `income` as their regression line seems to be very flat, while males tend to gamble more based on getting more `income`. For males, the simple linear regression model might not be the best model as there are many points away from the line and I wonder if other techniques (like cubic) would be better. Also with more data since we only have 28 males this might be too small of amount to make accurate determinations on models and instead find trends like already expressed.

3 Problem 3

Consider the simple linear regression model: $y_i = \beta_0 + x_i\beta_1 + e_i$ where $i = 1, ..., n$ and the x variable has been centered and scaled so that $\sum x_i = 0$ and $\sum x_i^2 = 1$.

3.1 Part A

Write the model matrix, $\mathbf{X}$.

We know the matrix $\mathbf{X}$ is a (i x p) matrix where i is the number of observations and p is the number of predictors. In this case $i = 1, ..., n$ so $i = n$ and we have two predictors (β_0 and β_1) so $p = 2$. So matrix $\mathbf{X}$ is a (n x 2) matrix. Now note the first column is all 1's since all values will use the β_0 value and in the second column all the values will be its corresponding x_i value (for example, if we are looking at y_1 then the corresponding x_i is x_1). So from this we know our model matrix,

$$\mathbf{X} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ ...(1) & ...(x_i) \\ 1 & x_n \end{pmatrix}$$

3.2 Part B

Write the expression for $\mathbf{X^T X}$, and solve the normal equations.

We know $\mathbf{X^T}$ is just flipping the X matrix on its side by flipping the indices. So now the second row and first column value in the $\mathbf{X}$ matrix become the first row and second column in the $\mathbf{X^T}$ matrix. Thus, we can see that

$$\mathbf{X} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ ...(1) & ...(x_i) \\ 1 & x_n \end{pmatrix}$$

now becomes our matrix we are trying to solve which is

$$\mathbf{X^T} = \begin{pmatrix} 1 & 1 & ...(1) & 1 \\ x_1 & x_2 & ...(x_i) & x_n \end{pmatrix}$$

Therefore we can now get our expression of

$$\mathbf{X^T X} = \begin{pmatrix} 1 & 1 & ...(1) & 1 \\ x_1 & x_2 & ...(x_i) & x_n \end{pmatrix} \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ ...(1) & ...(x_i) \\ 1 & x_n \end{pmatrix} = \begin{pmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{pmatrix} = \begin{pmatrix} n & 0 \\ 0 & 1 \end{pmatrix}$$

Now we know the solution to the normal equations is found by saying that $\hat{\beta} = (\mathbf{X^T X})^{-1} \mathbf{X^T y}$ as long as $\mathbf{X^T X}$ is invertible, which we know is true because the inverse of a diagonal matrix is a diagonal matrix where the elements are the reciprocals of the corresponding elements in the original matrix. Therefore, if $\mathbf{X^T X} = \begin{pmatrix} n & 0 \\ 0 & 1 \end{pmatrix}$ then $(\mathbf{X^T X})^{-1} = \begin{pmatrix} \frac{1}{n} & 0 \\ 0 & 1 \end{pmatrix}$. Now,

$$\mathbf{X^T y} = \begin{pmatrix} 1 & 1 & ...(1) & 1 \\ x_1 & x_2 & ...(x_i) & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ ...(y_i) \\ y_n \end{pmatrix} = \begin{pmatrix} \sum y_i \\ \sum x_i y_i \end{pmatrix} = \begin{pmatrix} n\bar{y} \\ \sum x_i y_i \end{pmatrix}$$

So now we can rewrite solving our normal equations as

$$\hat{\beta} = (\mathbf{X^T X})^{-1} \mathbf{X^T y} = \begin{pmatrix} \frac{1}{n} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} n\bar{y} \\ \sum x_i y_i \end{pmatrix} = \begin{pmatrix} \bar{y} \\ \sum x_i y_i \end{pmatrix}$$

So we have found our solutions to the normal equations being that $\hat{\beta} = \begin{pmatrix} \bar{y} \\ \sum x_i y_i \end{pmatrix}$.