Le Thi Phuong Thao

Student: 413707007

HW0224

<u>Question 2.1</u>

a) Complete the entries in the table. Put the sums in the last row. What are the sample means x and y?

X meas = x_bar = 1, Y mean = y_bar = 2

```
> print(data)
   x y x_bar y_bar x_minus_xbar x_minus_xbar_sq y_minus_ybar x_y_diff_prod
1  3 4     1     2            2               4            2             4
2  2 2     1     2            1               1            0             0
3  1 3     1     2            0               0            1             0
4 -1 1     1     2           -2               4           -1             2
5  0 0     1     2           -1               1           -2             2
> print(answers$a_sums)
  sum_x sum_y sum_x_minus_xbar sum_x_minus_xbar_sq sum_y_minus_ybar
1     5    10                0                  10                0
  sum_x_y_diff_prod
1                 8
```

b) Calculate b1 and b2 using (2.7) and (2.8) and state their interpretation

```
> cat("b2 (Slope):", b2, "\n")
b2 (Slope): 0.8
> cat("b1 (Intercept):", b1, "\n")
b1 (Intercept): 1.2
>
```

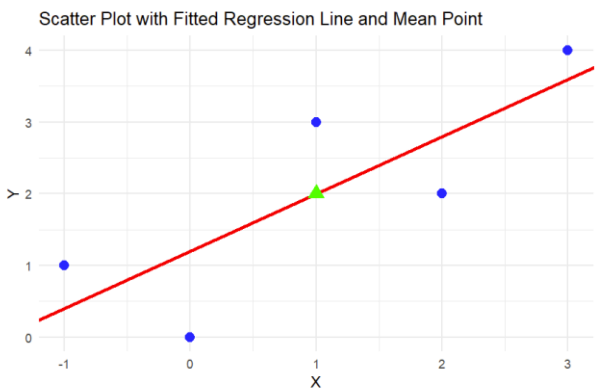
c) Checked

```
> # Verify equation
> # equation 1: Σ(xi - x̄)^2 = Σxi^2 - N*x̄^2
> sum_x_minus_xbar_sq_check <- sum_x_sq - N * data$x_bar[1]^2
> should_match_x = sums_a$sum_x_minus_xbar_sq
>
>
> # Verify equation 2: Σ(xi - x̄)(yi - ȳ) = Σxi*yi - N*x̄*ȳ
> sum_x_y_diff_prod_check <- sum_x_y - N * data$x_bar[1] * data$y_bar[1]
> should_match_xy = sums_a$sum_x_y_diff_prod
> # Print results
> cat("LHS of equation 1:", sum_x_minus_xbar_sq_check, " | RHS of equation 1:", s
hould_match_x, "\n")
LHS of equation 1: 10  | RHS of equation 1: 10
> cat("LHS of equation 2:", sum_x_y_diff_prod_check, " | RHS of equation 2:", sho
uld_match_xy, "\n")
LHS of equation 2: 8  | RHS of equation 2: 8
```

d) Results

```
> print(answers$d_table)
   x y y_hat    e e_sq  x_e
1  3 4   3.6  0.4 0.16  1.2
2  2 2   2.8 -0.8 0.64 -1.6
3  1 3   2.0  1.0 1.00  1.0
4 -1 1   0.4  0.6 0.36 -0.6
5  0 0   1.2 -1.2 1.44  0.0
> print(answers$d_sums)
  sum_x sum_y sum_y_hat         sum_e sum_e_sq         sum_x_e
1     5    10        10 -2.220446e-16      3.6 -1.332268e-15
> print(answers$d_stats)
  s_y_sq   s_x_sq     s_xy     r_xy     CV_x median_x
  2.5000   2.5000   2.0000   0.8000 158.1139   1.0000
> print(median_x)
[1] 1
```

e) Results



f) Yes, the regression line passes through the mean point

g) h)

```
> check_g <- b1 + b2 * x_bar
> check_h <- sums_d$sum_y_hat/N
> y_bar == check_g
[1] TRUE
> y_bar == check_h
[1] TRUE
```

i) Result

```
> sigma_sq_hat <- sums_d$sum_e_sq / (N - 2)
> answers$i_sigma_sq <- sigma_sq_hat
> print(answers$i_sigma_sq)
[1] 1.2
```

j) Results

```
> # j. Compute var(b2 | x) and se(b2)
> var_b2 <- sigma_sq_hat / sums_a$sum_x_minus_xbar_sq
> se_b2 <- sqrt(var_b2)
> answers$j_stats <- c(var_b2 = var_b2, se_b2 = se_b2)
> print(answers$j_stats)
   var_b2     se_b2
0.1200000 0.3464102
```

Question 2.14

a) Elasticity of wages with respect to education at the "point of the means.": 1.247214

b) Standard error of the elasticity of wages with respect to education at the "point of the means.": 0.0956108

c) The predicted wage for an individual with 12 years of education in each area, With 16 years of education

```
> print("Predicted Wages for 12 and 16 Years of Education:")
[1] "Predicted Wages for 12 and 16 Years of Education:"
> print(predicted_wages)
  Education_Years Urban_Wage Rural_Wage
1              12      18.76      16.72
2              16      28.60      23.92
```

Question 2.16

a) CAPM is a simple regression model because it models asset returns as a linear function of a single independent variable (market returns). It follows the standard OLS regression framework, making it a fundamental tool in finance for estimating risk and expected returns.

b) Estimate the CAPM model for each firm, and comment on their estimated beta values

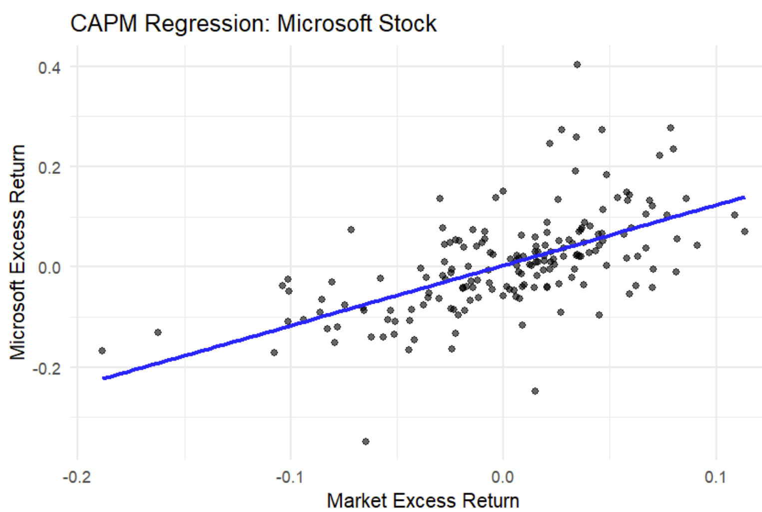
```
> print(capm_results)
# A tibble: 12 × 6
   stock      term          estimate std.error statistic  p.value
   <chr>      <chr>            <dbl>     <dbl>     <dbl>    <dbl>
 1 GE         (Intercept)  -0.000959   0.00442    -0.217 8.29e- 1
 2 GE         market_excess  1.15      0.0895     12.8   4.47e-27
 3 IBM        (Intercept)   0.00605    0.00483     1.25  2.12e- 1
 4 IBM        market_excess 0.977      0.0978      9.98  6.37e-19
 5 Ford       (Intercept)   0.00378    0.0102      0.370 7.12e- 1
 6 Ford       market_excess 1.66       0.207       8.03  1.27e-13
 7 Microsoft  (Intercept)   0.00325    0.00604     0.538 5.91e- 1
 8 Microsoft  market_excess 1.20       0.122       9.84  1.63e-18
 9 Disney     (Intercept)   0.00105    0.00468     0.224 8.23e- 1
10 Disney     market_excess 1.01       0.0946     10.7   6.50e-21
11 ExxonMobil (Intercept)   0.00528    0.00354     1.49  1.37e- 1
12 ExxonMobil market_excess 0.457      0.0716      6.38  1.48e- 9
```

Ford company has the most aggressive beta, while ExxonMobil has the most defensive beta

c) All alpha is insignificant (alpha is closed to 0) supports CAPM theory

```
> alpha_values <- capm_results %>%
+   filter(term == "(Intercept)") %>%
+   select(stock, estimate, std.error, statistic, p.value)
> print(alpha_values)
# A tibble: 6 × 5
  stock       estimate std.error statistic p.value
  <chr>          <dbl>     <dbl>     <dbl>   <dbl>
1 GE         -0.000959   0.00442    -0.217   0.829
2 IBM         0.00605    0.00483     1.25    0.212
3 Ford        0.00378    0.0102      0.370   0.712
4 Microsoft   0.00325    0.00604     0.538   0.591
5 Disney      0.00105    0.00468     0.224   0.823
6 ExxonMobil  0.00528    0.00354     1.49    0.137
```

Microsoft stock plot



d) Beta is quite similar regardless of whether the model includes or excludes alpha (intercept).

	stock	beta_with_intercept	beta_without_intercept
	<chr>	<dbl>	<dbl>
1	GE	1.15	1.15
2	IBM	0.977	0.984
3	Ford	1.66	1.67
4	Microsoft	1.20	1.21
5	Disney	1.01	1.01
6	ExxonMobil	0.457	0.463