

MATH 204: PRINCIPLES OF STATISTICS 2

WINTER 2026

EXERCISES ON SIMPLE LINEAR REGRESSION

For each of the following data sets with the given summary statistics, compute the estimates for β_0 , β_1 and σ from a simple linear regression analysis

1. <https://www.math.mcgill.ca/dstephens/204/Exercises/Data/SLR/Data01.csv>

$$n = 24 \quad \bar{x} = 9.979 \quad \bar{y} = 39.029$$

$$SS_{xx} = 827.465 \quad SS_{xy} = 3266.194 \quad SS_{yy} = 13011.362$$

2. <https://www.math.mcgill.ca/dstephens/204/Exercises/Data/SLR/Data02.csv>

$$n = 28 \quad \bar{x} = 14.505 \quad \bar{y} = -23.897$$

$$SS_{xx} = 2545.655 \quad SS_{xy} = -4200.397 \quad SS_{yy} = 7154.351$$

3. <https://www.math.mcgill.ca/dstephens/204/Exercises/Data/SLR/Data03.csv>

$$n = 28 \quad \bar{x} = 19.935 \quad \bar{y} = -29.175$$

$$SS_{xx} = 2102.598 \quad SS_{xy} = -3398.745 \quad SS_{yy} = 5541.064$$

4. <https://www.math.mcgill.ca/dstephens/204/Exercises/Data/SLR/Data04.csv>

$$n = 19 \quad \bar{x} = 6.098 \quad \bar{y} = 50.091$$

$$SS_{xx} = 338.736 \quad SS_{xy} = 2796.262 \quad SS_{yy} = 23131.858$$

5. <https://www.math.mcgill.ca/dstephens/204/Exercises/Data/SLR/Data05.csv>

$$n = 26 \quad \bar{x} = 16.696 \quad \bar{y} = -70.521$$

$$SS_{xx} = 2238.890 \quad SS_{xy} = -9550.420 \quad SS_{yy} = 40958.105$$

6. <https://www.math.mcgill.ca/dstephens/204/Exercises/Data/SLR/Data06.csv>

$$n = 20 \quad \bar{x} = 13.029 \quad \bar{y} = 28.519$$

$$SS_{xx} = 1451.534 \quad SS_{xy} = 3167.811 \quad SS_{yy} = 6921.734$$

7. <https://www.math.mcgill.ca/dstephens/204/Exercises/Data/SLR/Data07.csv>

$$n = 23 \quad \bar{x} = 18.680 \quad \bar{y} = -5.975$$

$$SS_{xx} = 1720.768 \quad SS_{xy} = -327.365 \quad SS_{yy} = 114.189$$

8. <https://www.math.mcgill.ca/dstephens/204/Exercises/Data/SLR/Data08.csv>

$$n = 22 \quad \bar{x} = 23.005 \quad \bar{y} = -114.976$$

$$SS_{xx} = 3446.135 \quad SS_{xy} = -16861.911 \quad SS_{yy} = 82519.905$$

9. <https://www.math.mcgill.ca/dstephens/204/Exercises/Data/SLR/Data09.csv>

$$n = 25 \quad \bar{x} = 11.857 \quad \bar{y} = 35.718$$

$$SS_{xx} = 1601.102 \quad SS_{xy} = 4902.375 \quad SS_{yy} = 15119.839$$

10. <https://www.math.mcgill.ca/dstephens/204/Exercises/Data/SLR/Data10.csv>

$$n = 30 \quad \bar{x} = 8.059 \quad \bar{y} = 97.096$$

$$SS_{xx} = 663.714 \quad SS_{xy} = 8174.566 \quad SS_{yy} = 100811.183$$

Solutions:

1. Results

- $(\hat{\beta}_0, \hat{\beta}_1) = (-0.360, 3.947), \hat{\sigma} = 2.325$

2. Results

- $(\hat{\beta}_0, \hat{\beta}_1) = (0.037, -1.650), \hat{\sigma} = 2.933$

3. Results

- $(\hat{\beta}_0, \hat{\beta}_1) = (3.050, -1.616), \hat{\sigma} = 1.347$

4. Results

- $(\hat{\beta}_0, \hat{\beta}_1) = (-0.251, 8.255), \hat{\sigma} = 1.693$

5. Results

- $(\hat{\beta}_0, \hat{\beta}_1) = (0.699, -4.266), \hat{\sigma} = 3.020$

6. Results

- $(\hat{\beta}_0, \hat{\beta}_1) = (0.085, 2.182), \hat{\sigma} = 0.681$

7. Results

- $(\hat{\beta}_0, \hat{\beta}_1) = (-2.421, -0.190), \hat{\sigma} = 1.572$

8. Results

- $(\hat{\beta}_0, \hat{\beta}_1) = (-2.411, -4.893), \hat{\sigma} = 0.858$

9. Results

- $(\hat{\beta}_0, \hat{\beta}_1) = (-0.586, 3.062), \hat{\sigma} = 2.181$

10. Results

- $(\hat{\beta}_0, \hat{\beta}_1) = (-2.157, 12.316), \hat{\sigma} = 2.155$

Output:

```
+ [1] "Data set 1"
+
+ Call:
+ lm(formula = y ~ x, data = df)
+
+ Residuals:
+   Min       1Q   Median       3Q      Max
+ -3.159 -1.959 -0.133  1.330  5.079
+
+ Coefficients:
+               Estimate Std. Error t value Pr(>|t|)
+ (Intercept)  -0.3598     0.9359   -0.38    0.7
+ x              3.9472     0.0808  48.83   <2e-16 ***
+ ---
+ Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
+
+ Residual standard error: 2.33 on 22 degrees of freedom
+ Multiple R-squared:  0.991, Adjusted R-squared:  0.99
+ F-statistic: 2.38e+03 on 1 and 22 DF,  p-value: <2e-16
+
+ [1] "Data set 2"
+
+ Call:
+ lm(formula = y ~ x, data = df)
+
+ Residuals:
+   Min       1Q   Median       3Q      Max
+  -6.83   -1.51    0.27    1.67    5.91
+
+ Coefficients:
+               Estimate Std. Error t value Pr(>|t|)
+ (Intercept)   0.0371     1.0089    0.04    0.97
+ x             -1.6500     0.0581  -28.39   <2e-16 ***
+ ---
+ Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
+
+ Residual standard error: 2.93 on 26 degrees of freedom
+ Multiple R-squared:  0.969, Adjusted R-squared:  0.968
+ F-statistic: 806 on 1 and 26 DF,  p-value: <2e-16
+
+ [1] "Data set 3"
+
+ Call:
+ lm(formula = y ~ x, data = df)
+
+ Residuals:
+   Min       1Q   Median       3Q      Max
+ -3.821 -0.862  0.218  0.992  2.536
+
+ Coefficients:
```

```

+           Estimate Std. Error t value Pr(>|t|)
+ (Intercept)   3.0498     0.6385    4.78 6.1e-05 ***
+ x             -1.6165     0.0294   -55.03 < 2e-16 ***
+ ---
+ Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
+
+ Residual standard error: 1.35 on 26 degrees of freedom
+ Multiple R-squared:  0.991, Adjusted R-squared:  0.991
+ F-statistic: 3.03e+03 on 1 and 26 DF, p-value: <2e-16
+
+ [1] "Data set 4"
+
+ Call:
+ lm(formula = y ~ x, data = df)
+
+ Residuals:
+   Min       1Q   Median       3Q      Max
+ -3.432 -1.108  0.207  1.238  2.328
+
+ Coefficients:
+           Estimate Std. Error t value Pr(>|t|)
+ (Intercept)   -0.251     0.682   -0.37    0.72
+ x              8.255     0.092   89.72 <2e-16 ***
+ ---
+ Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
+
+ Residual standard error: 1.69 on 17 degrees of freedom
+ Multiple R-squared:  0.998, Adjusted R-squared:  0.998
+ F-statistic: 8.05e+03 on 1 and 17 DF, p-value: <2e-16
+
+ [1] "Data set 5"
+
+ Call:
+ lm(formula = y ~ x, data = df)
+
+ Residuals:
+   Min       1Q   Median       3Q      Max
+  -6.57  -1.80  -0.25   1.31   4.66
+
+ Coefficients:
+           Estimate Std. Error t value Pr(>|t|)
+ (Intercept)   0.6991     1.2193    0.57    0.57
+ x            -4.2657     0.0638  -66.83 <2e-16 ***
+ ---
+ Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
+
+ Residual standard error: 3.02 on 24 degrees of freedom
+ Multiple R-squared:  0.995, Adjusted R-squared:  0.994
+ F-statistic: 4.47e+03 on 1 and 24 DF, p-value: <2e-16
+
+ [1] "Data set 6"
+

```

```

+ Call:
+ lm(formula = y ~ x, data = df)
+
+ Residuals:
+   Min       1Q   Median       3Q      Max
+ -0.991 -0.489 -0.127  0.565  1.086
+
+ Coefficients:
+               Estimate Std. Error t value Pr(>|t|)
+ (Intercept)   0.0854     0.2781    0.31   0.76
+ x             2.1824     0.0179  122.16 <2e-16 ***
+ ---
+ Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
+
+ Residual standard error: 0.681 on 18 degrees of freedom
+ Multiple R-squared:  0.999, Adjusted R-squared:  0.999
+ F-statistic: 1.49e+04 on 1 and 18 DF, p-value: <2e-16
+
+ [1] "Data set 7"
+
+ Call:
+ lm(formula = y ~ x, data = df)
+
+ Residuals:
+   Min       1Q   Median       3Q      Max
+ -2.349 -0.918 -0.388  1.226  3.769
+
+ Coefficients:
+               Estimate Std. Error t value Pr(>|t|)
+ (Intercept)  -2.4211     0.7802   -3.10  0.0054 **
+ x            -0.1902     0.0379   -5.02  5.7e-05 ***
+ ---
+ Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
+
+ Residual standard error: 1.57 on 21 degrees of freedom
+ Multiple R-squared:  0.545, Adjusted R-squared:  0.524
+ F-statistic: 25.2 on 1 and 21 DF, p-value: 5.72e-05
+
+ [1] "Data set 8"
+
+ Call:
+ lm(formula = y ~ x, data = df)
+
+ Residuals:
+   Min       1Q   Median       3Q      Max
+ -1.2425 -0.4746  0.0425  0.4612  2.1130
+
+ Coefficients:
+               Estimate Std. Error t value Pr(>|t|)
+ (Intercept)  -2.4112     0.3827   -6.3  3.8e-06 ***
+ x            -4.8930     0.0146  -334.9 < 2e-16 ***
+ ---

```

```

+ Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
+
+ Residual standard error: 0.858 on 20 degrees of freedom
+ Multiple R-squared:  1, Adjusted R-squared:  1
+ F-statistic: 1.12e+05 on 1 and 20 DF,  p-value: <2e-16
+
+ [1] "Data set 9"
+
+ Call:
+ lm(formula = y ~ x, data = df)
+
+ Residuals:
+   Min       1Q   Median       3Q      Max
+ -3.326 -1.420 -0.323  2.381  3.568
+
+ Coefficients:
+              Estimate Std. Error t value Pr(>|t|)
+ (Intercept)  -0.5863     0.7796   -0.75    0.46
+ x              3.0619     0.0545   56.18 <2e-16 ***
+ ---
+ Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
+
+ Residual standard error: 2.18 on 23 degrees of freedom
+ Multiple R-squared:  0.993, Adjusted R-squared:  0.992
+ F-statistic: 3.16e+03 on 1 and 23 DF,  p-value: <2e-16
+
+ [1] "Data set 10"
+
+ Call:
+ lm(formula = y ~ x, data = df)
+
+ Residuals:
+   Min       1Q   Median       3Q      Max
+ -4.467 -1.306  0.068  1.481  4.928
+
+ Coefficients:
+              Estimate Std. Error t value Pr(>|t|)
+ (Intercept)  -2.1566     0.7804   -2.76    0.01 **
+ x             12.3164     0.0836  147.26 <2e-16 ***
+ ---
+ Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
+
+ Residual standard error: 2.15 on 28 degrees of freedom
+ Multiple R-squared:  0.999, Adjusted R-squared:  0.999
+ F-statistic: 2.17e+04 on 1 and 28 DF,  p-value: <2e-16

```