Solve the following problems. Create a web portfolio in presenting the answers. Use Full Stack developing the programs for these:

1. Create a scatterplot for each combinatorics of the metadata of the IRIS Data set.
    a. Compute for each pair's Coefficient of Regression
    b. Compute for the following Regression Curves for each pair
        i. Linear
        ii. Logarithmic
        iii. Exponential
        iv. Hyperbolic
    c. Separate the species by color.

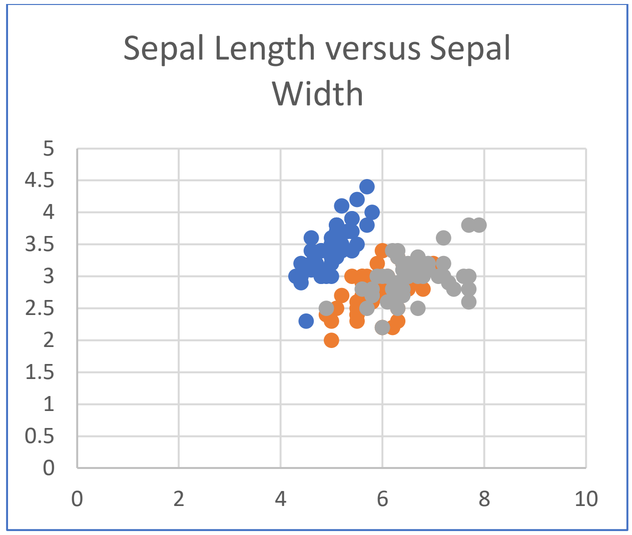
2. The following data sets are needed to answer the following questions:

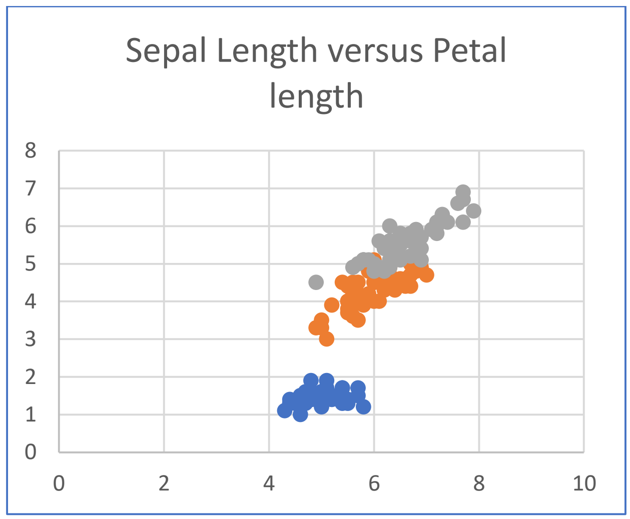
    Energy Production
    https://drive.google.com/open?id=11eVuROefnuKNLpOXh1_w5vXo3wWnnCHQ

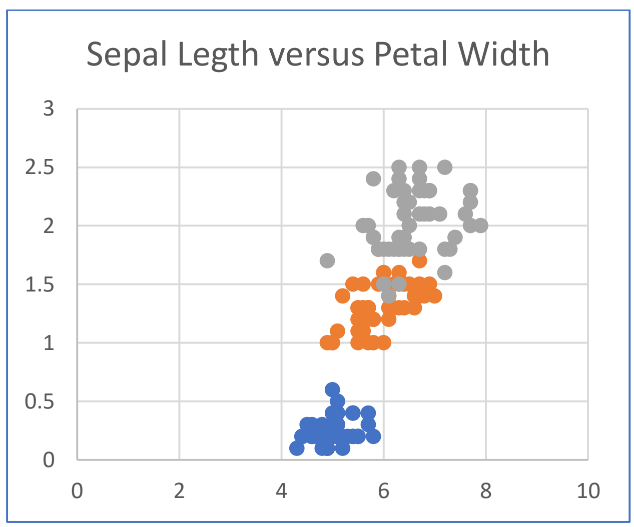
    Carbon Dioxide
    https://drive.google.com/open?id=1rVtGYzeXNj9O0VWlHSXc7Ka_rBBRFO0T

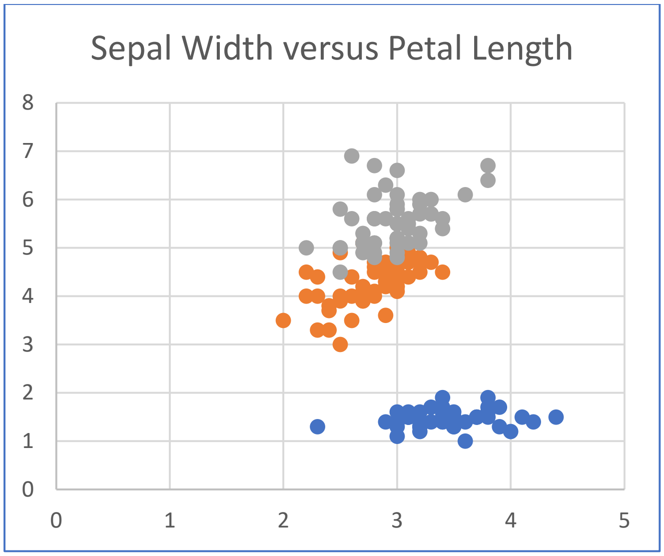
    Create a scatterplot that will correlate the Energy Production with the Carbon Dioxide Emission.
    a. Find the Coefficient of Regression
    b. Compute for the following Regression Curves for each pair
        a. Linear
        b. Logarithmic
        c. Exponential
        d. Hyperbolic

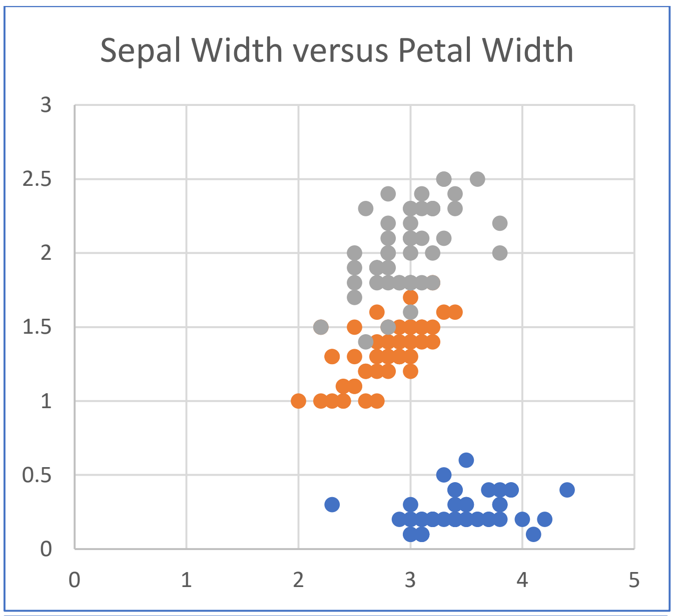
a. This will be presented in the first day of class.
b. Create a group of 4
c. Codes will be inspected
d. A random member will be given another question to be programmed impromptu.
e. This will be counted as quiz 1 with 100 items.

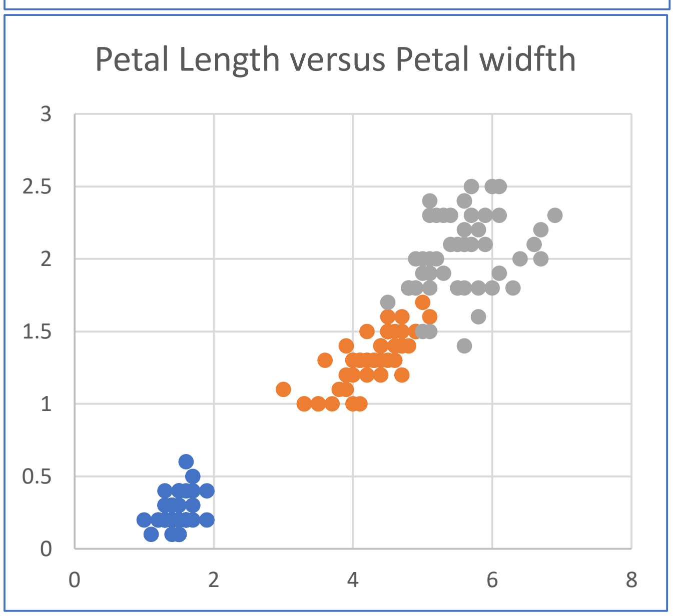










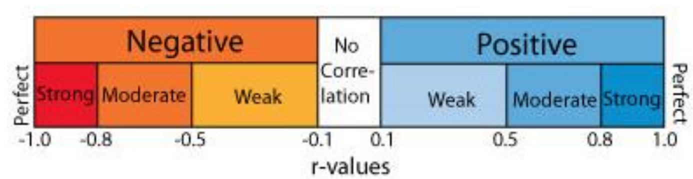


Pearson Product-Moment Correlation Coefficient

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{\left[n\sum x^2 - \left(\sum x\right)^2\right]\left[n\sum y^2 - \left(\sum y\right)^2\right]}}$$

- Has the range of values: $-1 \le r \le 1$
- If r is negative, there is an inverse relationship between x and y, i.e., if x is increasing then y is decreasing or vice versa.
- If r is positive, there is a direct relationship between x and y, i.e., if x is increasing then y is increasing or vice versa.
- If $r = 0$, then the two sets of data are uncorrelated (No Correlation).

Correlation Scale



Linear Regression

$$\hat{y} = A + Bx$$

Where:

$$B = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - \left(\sum x\right)^2}$$

$$A = \bar{y} - B\bar{x}$$

---

1. Exponential Model

$$y = a_1 e^{b_1 x}$$

Linearization:

$$\ln y = \ln a_1 + b_1 x$$

Equivalence:

$$\hat{y} = \ln y \qquad A = \ln a_1 \qquad B = b_1 \qquad x = x$$

2. Simple Power Equation

$$y = a_1 x^{b_1}$$

Linearization:

$$\log y = \log a_1 + b_1 \log x$$

Equivalence:

$$\hat{y} = \log y \qquad A = \log a_1 \qquad x = \log x \qquad B = b_1$$

3. Saturation-growth-rate/hyperbolic equation

$$y = a_1 \frac{x}{b_1 + x}$$

Linearization: $$\frac{1}{y} = \frac{1}{a_1} + \frac{b_1}{a_1}\frac{1}{x}$$

Equivalence:

$$\hat{y} = \frac{1}{y} \qquad A = \frac{1}{a_1} \qquad B = \frac{b_1}{a_1} \qquad x = \frac{1}{x}$$

Example:

Given the following sets of data:

x	y
0.50	1.90
1.00	1.50
1.30	1.20
1.60	1.00
2.00	0.80
2.20	0.78
2.50	0.65
3.10	0.46
3.90	0.30
4.40	0.23

Solve for the best-fit mathematical model.

Linear

x	y	x^2	y^2	xy
0.50	1.90	0.25	3.61	0.95
1.00	1.50	1.00	2.25	1.5
1.30	1.20	1.69	1.44	1.56
1.60	1.00	2.56	1.00	1.6
2.00	0.80	4.00	0.64	1.6
2.20	0.78	4.84	0.61	1.716
2.50	0.65	6.25	0.42	1.625
3.10	0.46	9.61	0.21	1.426
3.90	0.30	15.21	0.09	1.17
4.40	0.23	19.36	0.05	1.012
22.50	8.82	64.77	10.33	14.16

$$\hat{y} = 1.79 - 0.40x \qquad r = -0.95$$

Power

		X	Y	X^2	Y^2	XY
x	y	logx	logy	(logx)^2	(logy)^2	logxlogy
0.50	1.90	-0.30103	0.278754	0.090619	0.077704	-0.08391
1.00	1.50	0	0.176091	0	0.031008	0
1.30	1.20	0.113943	0.079181	0.012983	0.00627	0.009022
1.60	1.00	0.20412	0	0.041665	0	0
2.00	0.80	0.30103	-0.09691	0.090619	0.009392	-0.02917
2.20	0.78	0.342423	-0.10791	0.117253	0.011644	-0.03695
2.50	0.65	0.39794	-0.18709	0.158356	0.035001	-0.07445
3.10	0.46	0.491362	-0.33724	0.241436	0.113732	-0.16571
3.90	0.30	0.591065	-0.52288	0.349357	0.273402	-0.30906
4.40	0.23	0.643453	-0.63827	0.414031	0.407391	-0.4107
22.50	8.82	2.78	-1.36	1.52	0.97	-1.10

$$\hat{y} = 1.368x^{-0.98} \qquad r = -0.95$$

Exponential

X		Y	X^2	Y^2	XY
x	y	lny	x^2	(lny)^2	xlny
0.50	1.90	0.641854	0.25	0.41	0.320927
1.00	1.50	0.405465	1.00	0.16	0.405465
1.30	1.20	0.182322	1.69	0.03	0.237018
1.60	1.00	0	2.56	0.00	0
2.00	0.80	-0.22314	4.00	0.05	-0.44629
2.20	0.78	-0.24846	4.84	0.06	-0.54661
2.50	0.65	-0.43078	6.25	0.19	-1.07696
3.10	0.46	-0.77653	9.61	0.60	-2.40724
3.90	0.30	-1.20397	15.21	1.45	-4.69549
4.40	0.23	-1.46968	19.36	2.16	-6.46657
22.50	8.82	-3.12	64.77	5.12	-14.68

$$\hat{y} = 2.47e^{-0.54x} \qquad r = -1.00$$

Hyperbolic

		X	Y	X^2	Y^2	XY
x	y	1/x	1/y	(1/x)^2	(1/y)^2	(1/x)(1/y)
0.50	1.90	2	0.526316	4	0.277008	1.052632
1.00	1.50	1	0.666667	1	0.444444	0.666667
1.30	1.20	0.769231	0.833333	0.591716	0.694444	0.641026
1.60	1.00	0.625	1	0.390625	1	0.625
2.00	0.80	0.5	1.25	0.25	1.5625	0.625
2.20	0.78	0.454545	1.282051	0.206612	1.643655	0.582751
2.50	0.65	0.4	1.538462	0.16	2.366864	0.615385
3.10	0.46	0.322581	2.173913	0.104058	4.725898	0.701262
3.90	0.30	0.25641	3.333333	0.065746	11.11111	0.854701
4.40	0.23	0.227273	4.347826	0.051653	18.90359	0.988142
22.50	8.82	6.56	16.95	6.82	42.73	7.35

$$\hat{y} = \frac{0.374x}{x - 0.5576} \qquad r = -0.63$$

Interpretation

Since exponential model's r-value has the nearest absolute value to 1 among the other mathematical models, then it is considered to be the best-fit curve for the system of data.