

# Lab Practical Assignment II

## The Analysis of Bivariate Data: Linear Regression

**Ordinary Least Squares Regression: Depth (ft.)-H2O Content**

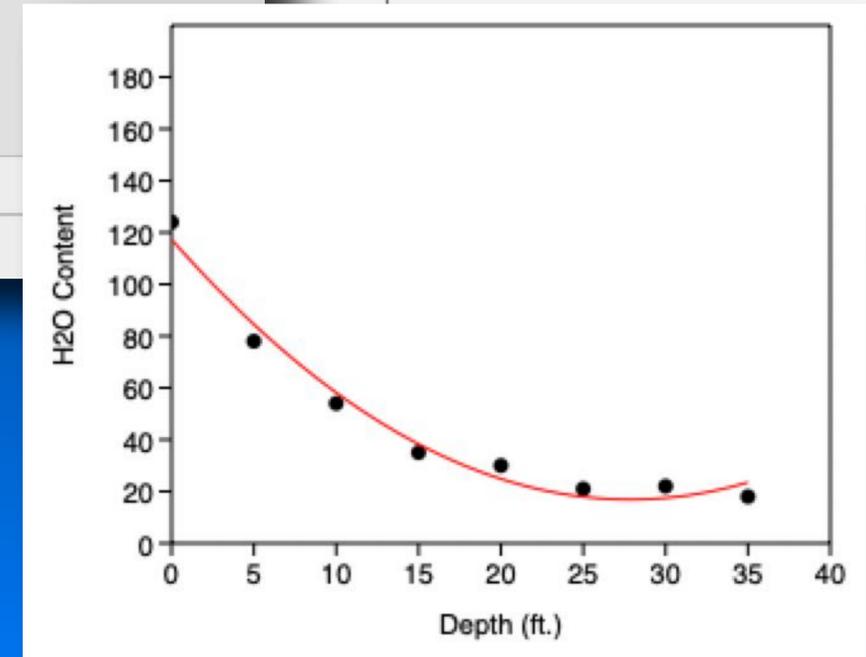
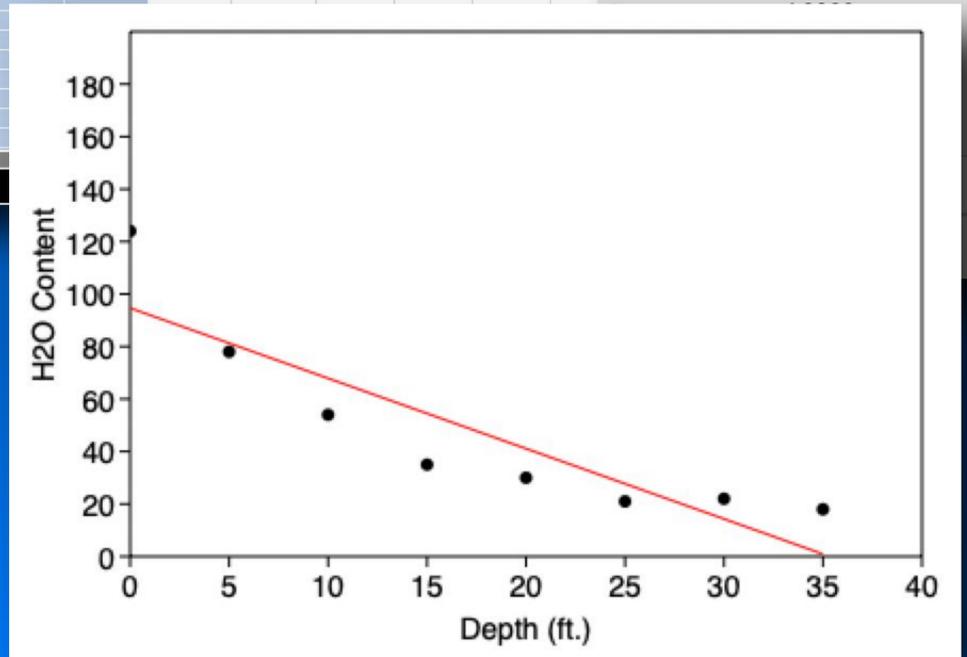
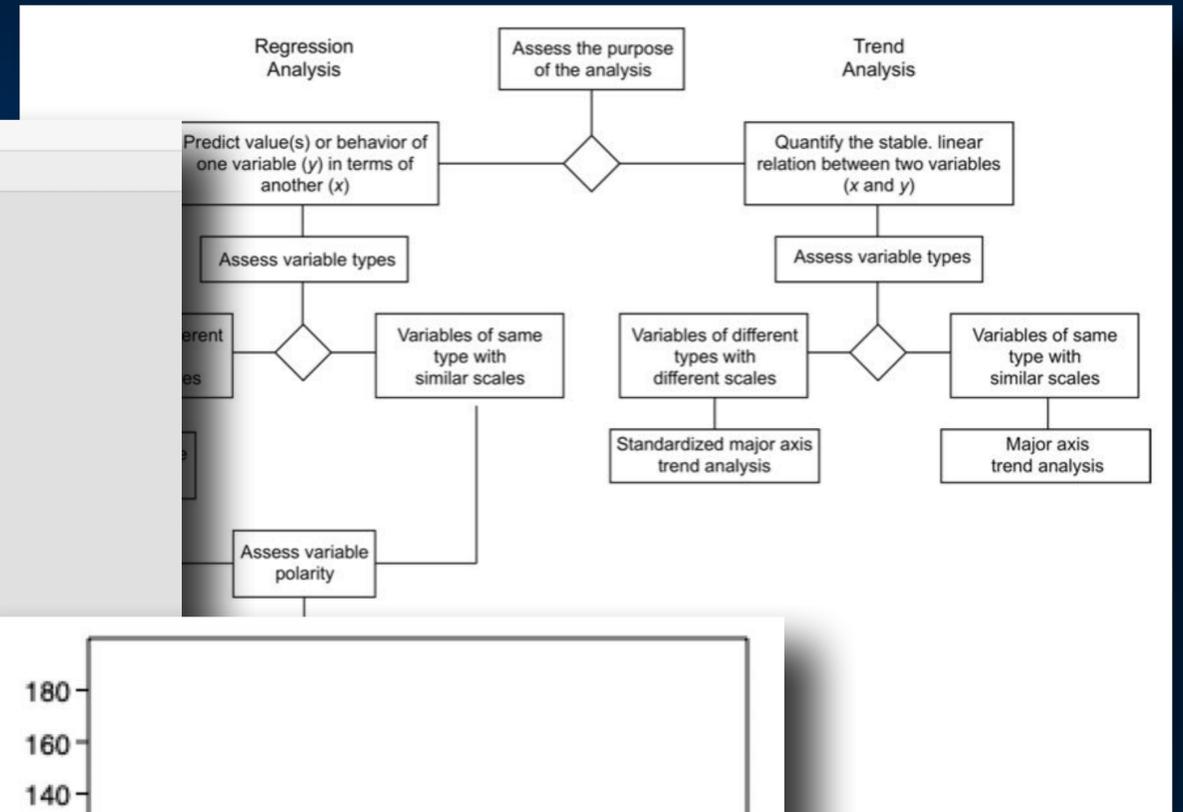
<b>Slope a:</b>	-2.681	<b>Std. error a:</b>	0.55815
<b>t:</b>	4.8033	<b>p (slope):</b>	0.0029906
<b>Intercept b:</b>	94.667	<b>Std. error b:</b>	11.674

**95% bootstrapped confidence intervals (N=1999):**

<b>Slope a:</b>	(-4.156, -1.071)
<b>Intercept b:</b>	(72.832, 132.32)

**Correlation:**

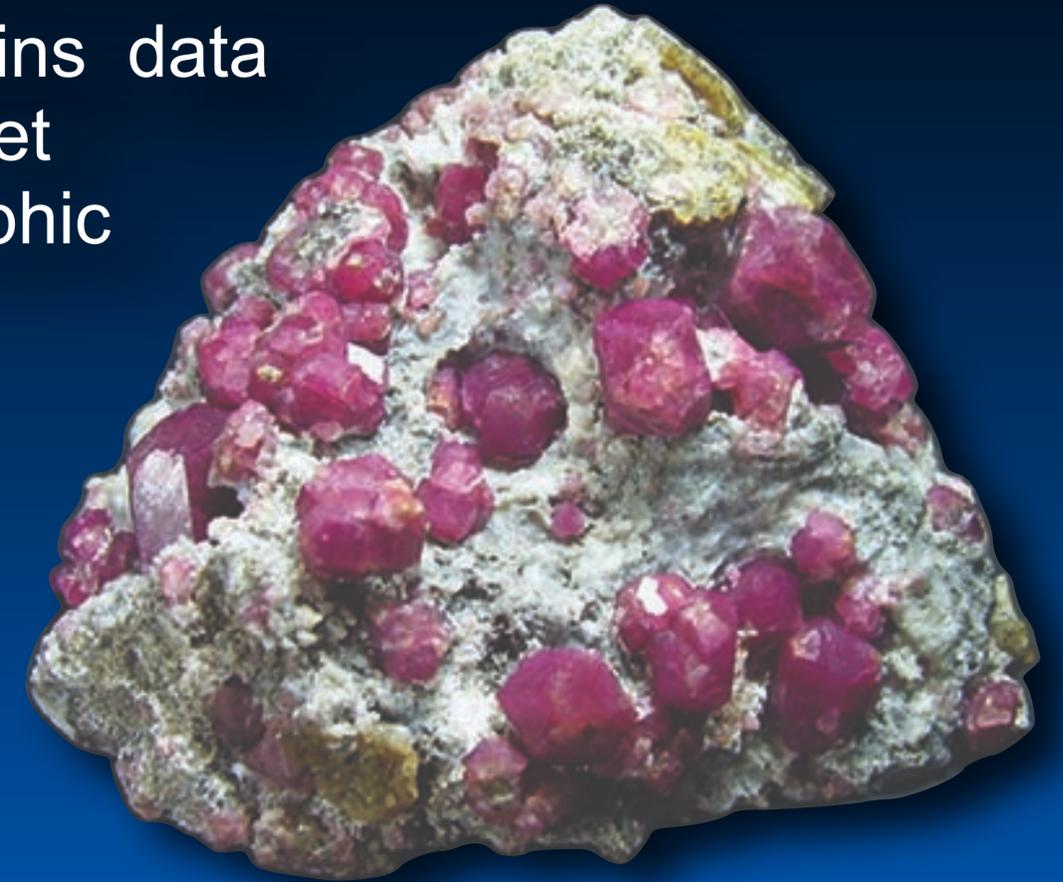
<b>r:</b>	-0.89085
<b>r<sup>2</sup>:</b>	0.79361



# Lab Practical Assignment II

---

1. The Garnets dataset (Garnets.dat, Garnets.csv) contains data that quantify changes in the abundance of iron in garnet crystals collected from a core drilled into the metamorphic halo surrounding an igneous intrusion.
  - Plot these data. (10 points)
  - Select a linear regression model that will allow the concentration of iron (Fe) to be predicted as a function of depth into metamorphic halo. (10 points)
    - a. Justify your selection. (20 points)
  - List the equation of the regression line for the model you have selected. (10 points)
    - a. Use this equation to predict the Fe concentration at a depth of 300 meters. (20 points)



# Lab Practical Assignment II

---

- Use an ANOVA  $F$  test to estimate the significance of the regression model.
  - a. To two decimal places state the probability value associated with the ANOVA test result. (10 points)
  - b. Provide an interpretation of the ANOVA test result in terms of the degree to which the regression model can be regarded as constituting an accurate prediction. (20 points)
  - c. Estimate the 95% confidence interval for the regression result. (20 points)
    1. Estimate the range of variation in Fe concentration values that would be expected at depth of 300 meters. (30 points)
  
- Do these data conform to the assumptions of an ANOVA test?
  - a. Justify your answer. (20 points)



# Lab Practical Assignment II

---

- Perform any additional test(s) you deem appropriate in order to confirm the validity of the regression model.
  - a. Show all plots, secondary statistical tests, and results associated with these additional tests (if any are warranted). (20 points)
  - b. Describe how the results of any additional tests (if any are warranted) either increased or decreased your confidence in your regression analysis. (20 points)



# Lab Practical Assignment II

## Garnets Dataset

Kentucky

Sample	Group	Depth (m)	Fe (%)
1	Garnet	0	14.21
2	Garnet	3	19.35
3	Garnet	10	17.22
4	Garnet	14	15.87
5	Garnet	23	13.62
6	Garnet	30	16.31
7	Garnet	36	14.13
8	Garnet	48	13.95
9	Garnet	59	15.00
10	Garnet	66	14.23
11	Garnet	68	16.81
12	Garnet	81	15.93
13	Garnet	94	16.02
14	Garnet	96	17.85
15	Garnet	102	17.02

Kentucky-1

Sample	Group	Depth (m)	Fe (%)
16	Garnet	115	15.87
17	Garnet	121	19.84
18	Garnet	130	16.94
19	Garnet	163	16.72
20	Garnet	168	19.20
21	Garnet	205	20.41
22	Garnet	239	16.88
23	Garnet	251	18.74
24	Garnet	283	16.67
25	Garnet	297	18.56
26	Garnet	322	18.87
27	Garnet	335	20.81
28	Garnet	351	24.52
29	Garnet	370	25.03
30	Garnet	408	25.11

Kentucky-1-1

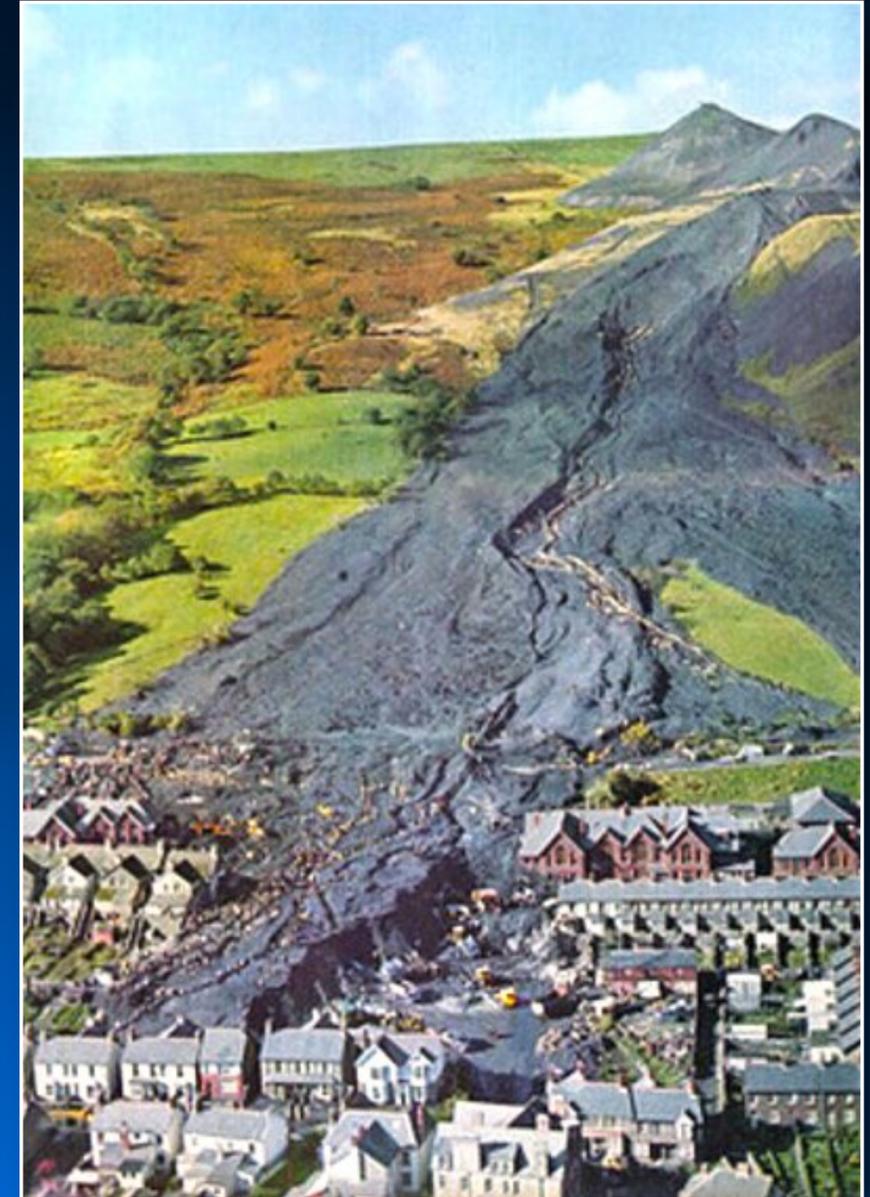
Sample	Group	Depth (m)	Fe (%)
31	Garnet	416	23.28
32	Garnet	419	22.56
33	Garnet	425	19.00
34	Garnet	429	20.53
35	Garnet	443	19.08
36	Garnet	447	22.83
37	Garnet	465	21.06
38	Garnet	474	24.96
39	Garnet	493	19.12
40	Garnet	502	22.24
41	Garnet	522	26.88
42	Garnet	550	21.15
43	Garnet	558	28.92
44	Garnet	571	27.96
45	Garnet	586	25.03

# Lab Practical Assignment II

---

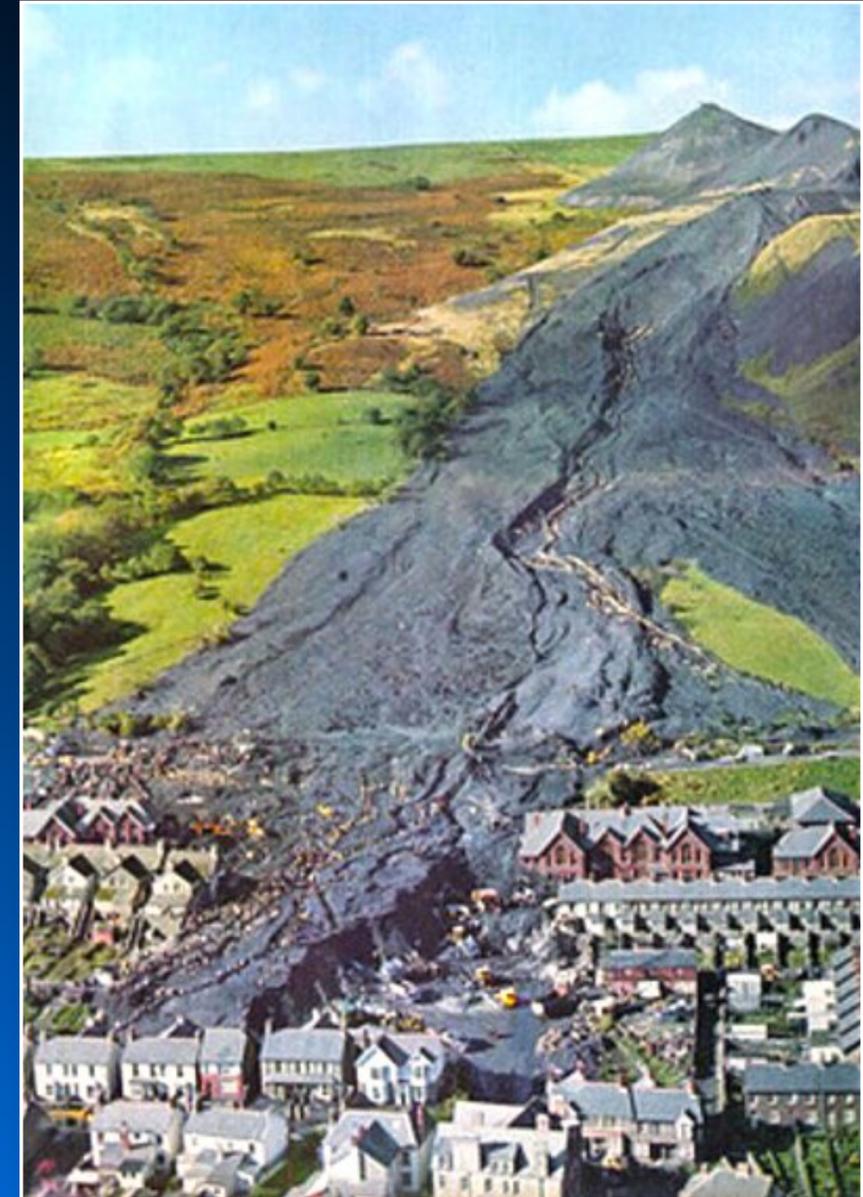
2. On October 21 1966 a colliery spoil heap in the village of Aberfan, Wales collapsed after a period of locally heavy rainfall, burying part of the village that was located immediately downslope from the spoil heap. One-hundred and forty-four villagers died in this incident, including 116 children who attended the Pantglas Junior School in Aberfan that day. As a result, the physical and chemical conditions of colliery spoil heaps throughout Britain were investigated.

Obviously, this raised a serious question regarding the stability of other colliery spoil heaps across the UK. As a result the UK government funded extensive research into the physics of colliery spoil heaps, some of which over 100 years old.



# Lab Practical Assignment II

2. Shear strength measurements were made on the Brancepeth Colliery spoil heap in County Durham (Taylor, 1973). The largest and smallest principal stresses collected from 36 spoil heap samples are listed in the Brancepeth Colliery dataset (Brancepeth Colliery.dat, Brancepeth Colliery.csv). Estimation of the principal stress angle is critical for the purpose of determining whether the spoil heap is prone to collapse under conditions like those that caused the collapse at Aberfan. The sine of the principle stress angle can be estimated as the slope of the “best fit” regression between these two variables.



# Lab Practical Assignment II

## Brancepeth Colliery Dataset

Kentucky

Sample	Group	Largest Shear Stress	Smallest Shear Stress
1	Brancepeth Colliery	66.79	39.40
2	Brancepeth Colliery	71.56	47.28
3	Brancepeth Colliery	95.41	50.43
4	Brancepeth Colliery	136.75	66.19
5	Brancepeth Colliery	157.42	92.98
6	Brancepeth Colliery	182.86	113.47
7	Brancepeth Colliery	189.22	130.80
8	Brancepeth Colliery	197.17	126.07
9	Brancepeth Colliery	221.02	107.16
10	Brancepeth Colliery	252.83	118.19
11	Brancepeth Colliery	219.43	146.56
12	Brancepeth Colliery	232.16	162.32
13	Brancepeth Colliery	310.07	185.96
14	Brancepeth Colliery	316.43	176.50
15	Brancepeth Colliery	321.20	181.23

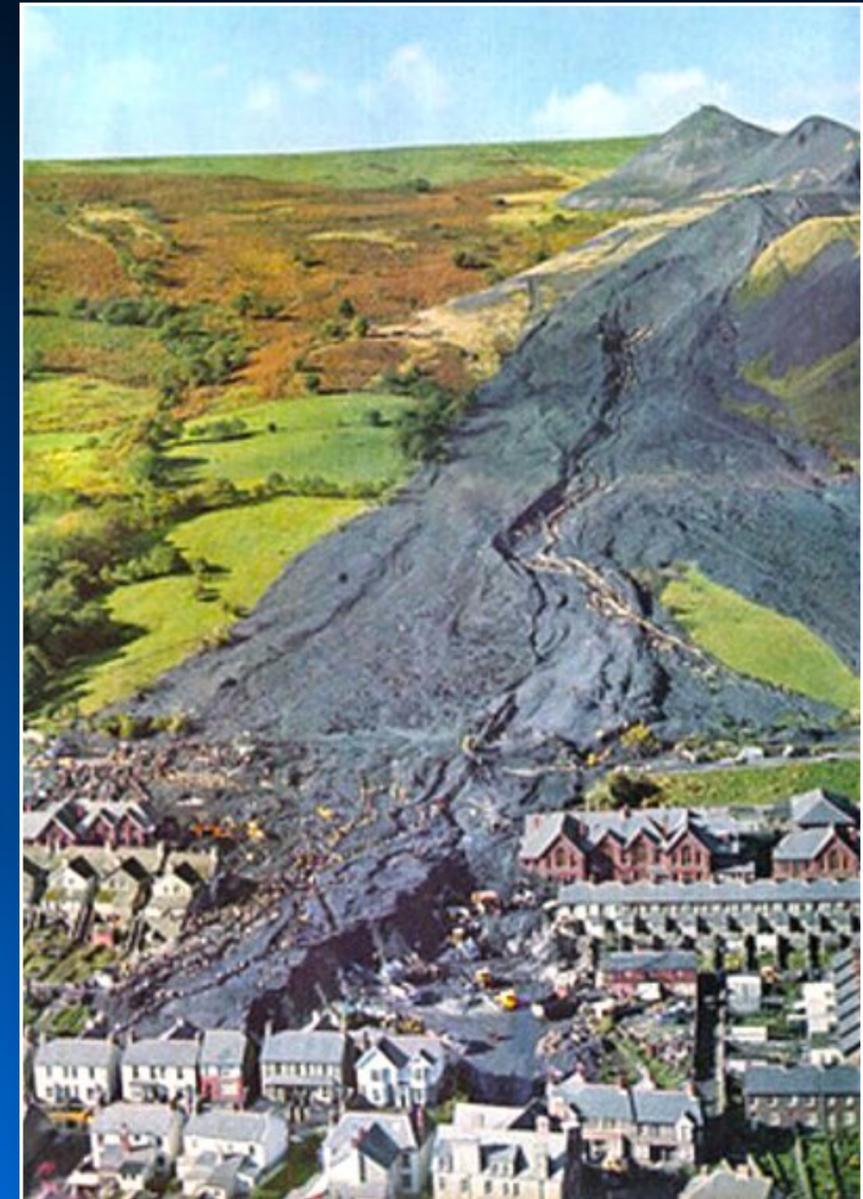
Kentucky-1

Sample	Group	Largest Shear Stress	Smallest Shear Stress
16	Brancepeth Colliery	325.97	193.84
17	Brancepeth Colliery	353.00	217.48
18	Brancepeth Colliery	356.18	225.36
19	Brancepeth Colliery	359.36	231.66
20	Brancepeth Colliery	403.89	165.47
21	Brancepeth Colliery	397.53	192.26
22	Brancepeth Colliery	403.89	209.60
23	Brancepeth Colliery	456.36	250.57
24	Brancepeth Colliery	480.21	296.28
25	Brancepeth Colliery	480.21	310.46
26	Brancepeth Colliery	462.72	330.95
27	Brancepeth Colliery	550.18	330.95
28	Brancepeth Colliery	578.80	378.22
29	Brancepeth Colliery	618.55	308.88
30	Brancepeth Colliery	632.86	307.31

# Lab Practical Assignment II

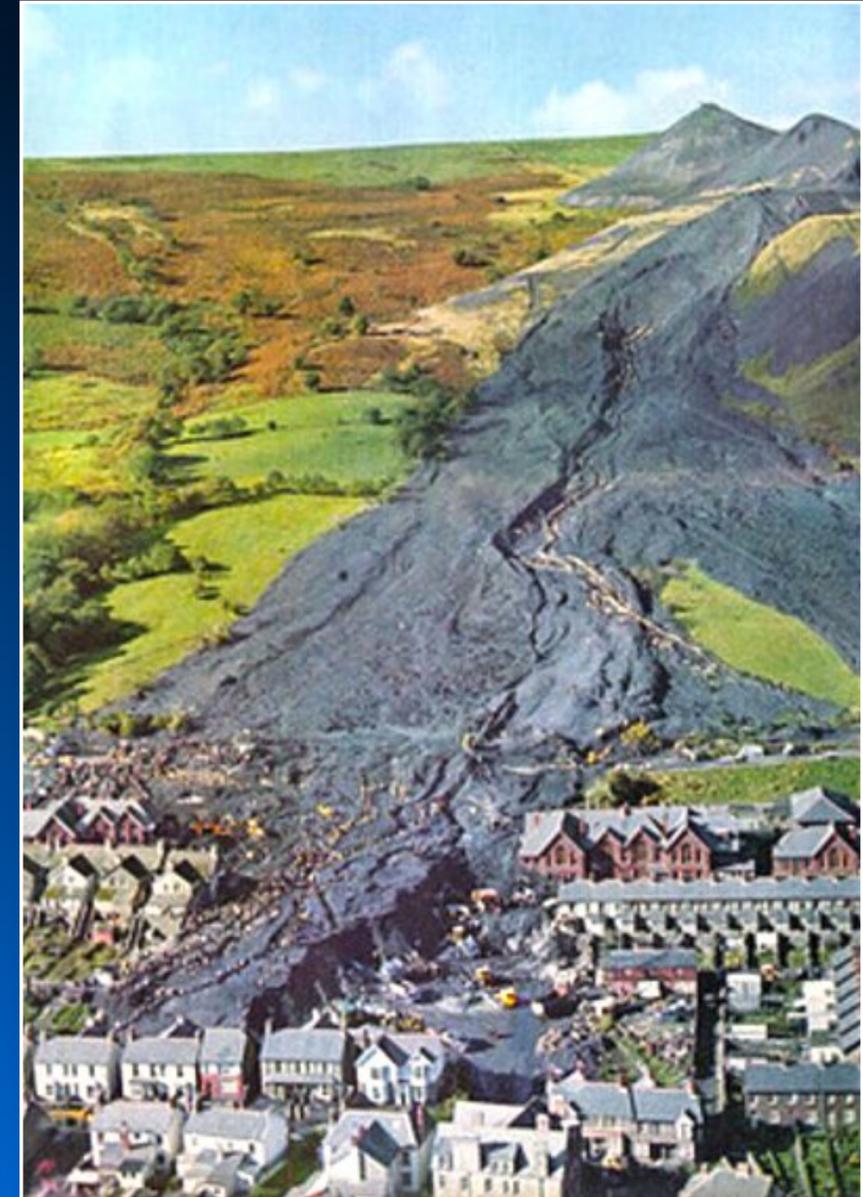
---

- Plot these data. (10 points)
- Select a linear regression model that will allow the principle stress for this colliery spoil heap to be predicted. (10 points)
  - a. Justify your selection. (20 points)
- List the equation of the regression line for the model you have selected. (10 points)
  - a. Use this equation to predict the angle of the principle stress estimate (in degrees). (20 points)



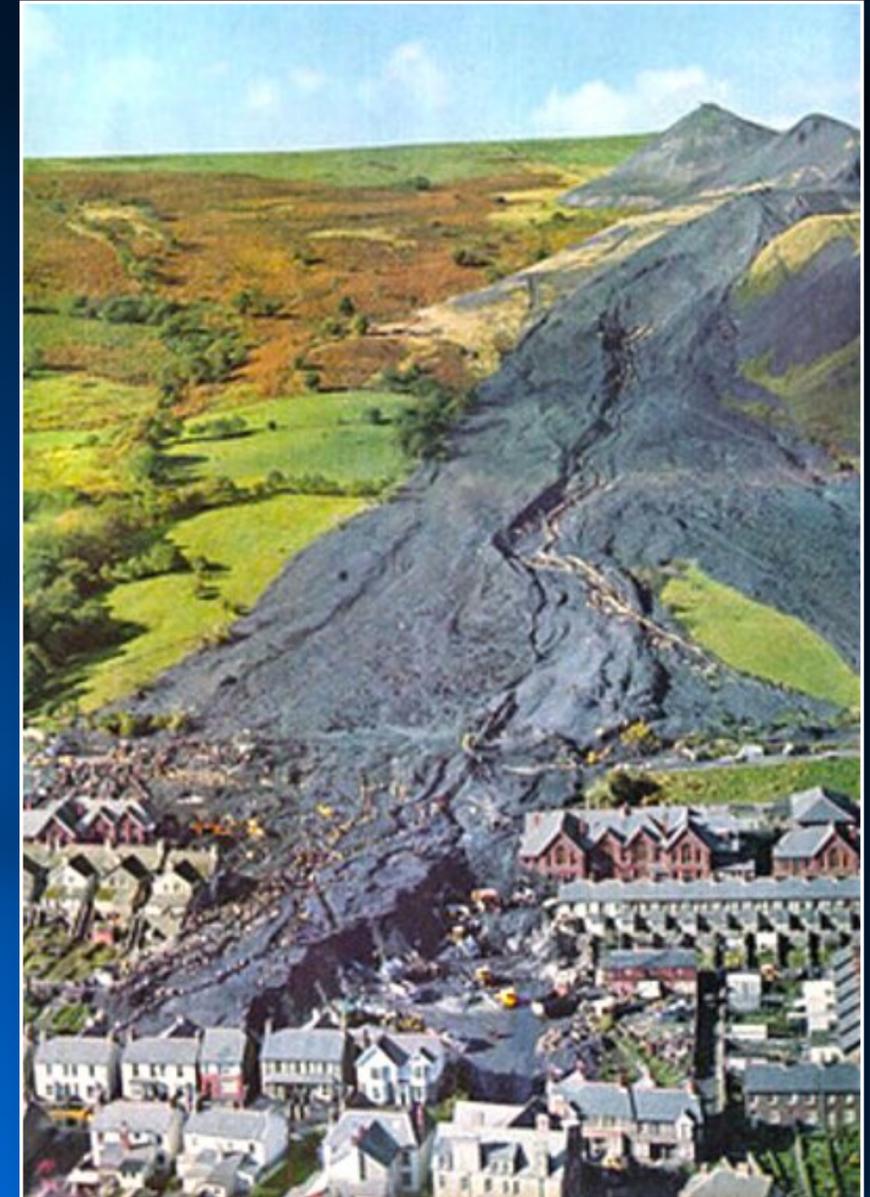
# Lab Practical Assignment II

- Use an ANOVA  $F$  test to estimate the significance of the regression model.
  - a. Based on the purpose of this analysis decide what, in your opinion, an appropriate probability value would be for use in reporting your results.
    - 1. Justify your reasoning. (30 points)
  - b. To two decimal places state the probability value associated with the ANOVA test result. (10 points)
  - c. Provide an interpretation of the ANOVA test result in terms of the degree to which the regression model can be regarded as constituting an accurate prediction. (20 points)
  - d. Estimate the 95% confidence interval for the regression result. (20 points)
    - 1. Estimate the range of principle stress angle variation that would be expected at this colliery. (30 points)



# Lab Practical Assignment II

- Do these data conform to the assumptions of an ANOVA test?
  - a. Justify your reasoning. (20 points)
- Perform any additional test(s) you deem appropriate in order to confirm the validity of the regression model.
  - a. Show all plots, secondary statistical tests, and results associated with these additional tests (if any are warranted). (20 points)
  - b. Describe how the results of any additional tests (if any are warranted) either increased or decreased your confidence in your regression analysis. (20 points)



# Lab Practical Assignment II

3. The Mowry Shale is a formation occurring in the US states of Colorado, Wyoming and Montana. It is a thick unit of black shale beds interspersed with numerous minor beds of commercially exploitable bentonite. Bentonite is almost entirely composed of the clay mineral montmorillonite, which is an alteration product of rhyolitic or andesitic volcanic ash.

The Mowry Shale dataset (Mowry Shale.dat, Mowry Shale.csv) contains measurements of the thicknesses (in cm) and positions (in m) above the base of the Mowry Shale of a sequence of 26 bentonite beds.



# Lab Practical Assignment II

---

3. If it can be assumed that the marine shales of the Mowry Formation accumulated at a constant rate it may be possible to determine the frequency of local volcanic eruptions. Test this hypothesis using linear regression analysis.



# Lab Practical Assignment II

## Mowry Shale Dataset

Kentucky

Sample	Group	Thickness	Height above Base
1	Mowry Shale	0	0.00
2	Mowry Shale	4	4.00
3	Mowry Shale	26	30.00
4	Mowry Shale	4	34.00
5	Mowry Shale	5	39.00
6	Mowry Shale	4	43.00
7	Mowry Shale	17	60.00
8	Mowry Shale	3	63.00
9	Mowry Shale	6	69.00
10	Mowry Shale	4	73.00
11	Mowry Shale	35	108.00
12	Mowry Shale	2	110.00
13	Mowry Shale	15	125.00
14	Mowry Shale	10	135.00
15	Mowry Shale	23	158.00

Kentucky-1

Sample	Group	Thickness	Height above Base
16	Mowry Shale	8	166.00
17	Mowry Shale	7	173.00
18	Mowry Shale	47	220.00
19	Mowry Shale	14	234.00
20	Mowry Shale	17	251.00
21	Mowry Shale	5	256.00
22	Mowry Shale	10	266.00
23	Mowry Shale	5	271.00
24	Mowry Shale	6	277.00
25	Mowry Shale	11	288.00
26	Mowry Shale	29	317.00

# Lab Practical Assignment II

- Plot these data. (10 points)
- Select a linear regression model that will test the data for use in the estimation of eruption frequency. (10 points)
  1. Justify your selection. (20 points)
- List the equation of the regression line for the model you have selected. (10 points)
  1. Use this equation to predict the frequency of eruptions (if appropriate) assuming a rock accumulation rate of between 0.23 cm/yr and 4.6 cm/yr. (20 points)



# Lab Practical Assignment II

---

- Use an ANOVA  $F$  test to estimate the significance of the regression model.
  - a. To two decimal places state the probability value associated with the ANOVA test result. (10 points)
  - b. Provide an interpretation of the ANOVA test result in terms of the degree to which the regression model can be regarded as constituting an accurate prediction. (20 points)
  - c. Estimate the 95% confidence interval for the regression result. (20 points)
    - 1. Estimate the range of variation in eruption frequency estimates that would be expected based on your analysis. (30 points)



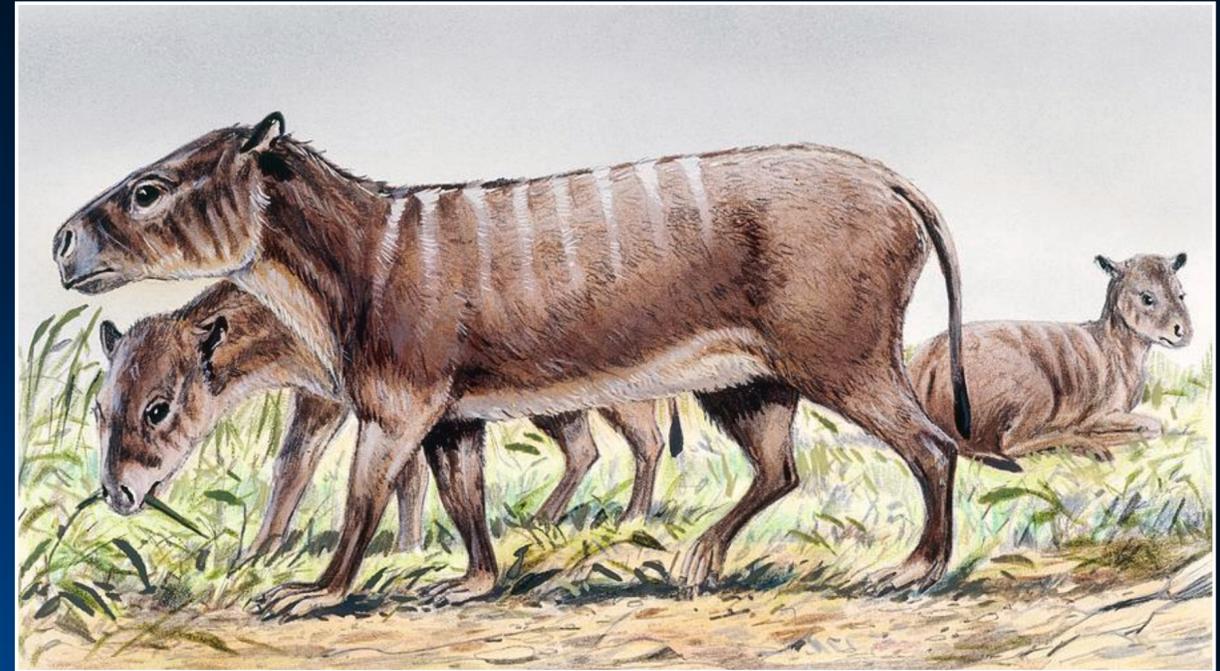
# Lab Practical Assignment II

- Do these data conform to the assumptions of an ANOVA test?
  - a. Justify your reasoning. (20 points)
- Perform any additional test(s) you deem appropriate in order to confirm the validity of the regression model.
  - a. Show all plots, secondary statistical tests, and results associated with these additional tests (if any are warranted). (20 points)
  - b. Describe how the results of any additional tests (if any are warranted) either increased or decreased your confidence in your regression analysis. (20 points)
- If no significant linear trend is found list a few factors that may be responsible for this negative finding. (10 points)



# Lab Practical Assignment II

4. Merycoidodontoids (known previously as “oreodonts” are an extinct group of Cenozoic pig-like mammals whose skulls are commonly found vertebrate fossils in the American west. They were a primitive group of cud-chewing artiodactyls with short faces and long, fang-like canine teeth. The Oreodont skulls dataset (Oreodont Skulls.dat, Oreodont Skulls.csv) contains a list of skull measurements collected from 72 specimens representing seven species. Fit a regression model to these data such that the groups are arranged in a systematic manner (with the groups as well-separated as possible) using any bivariate combination of variables and any appropriate linear regression model. Feel free to transform the variables if this improves the model fit.



# Lab Practical Assignment II

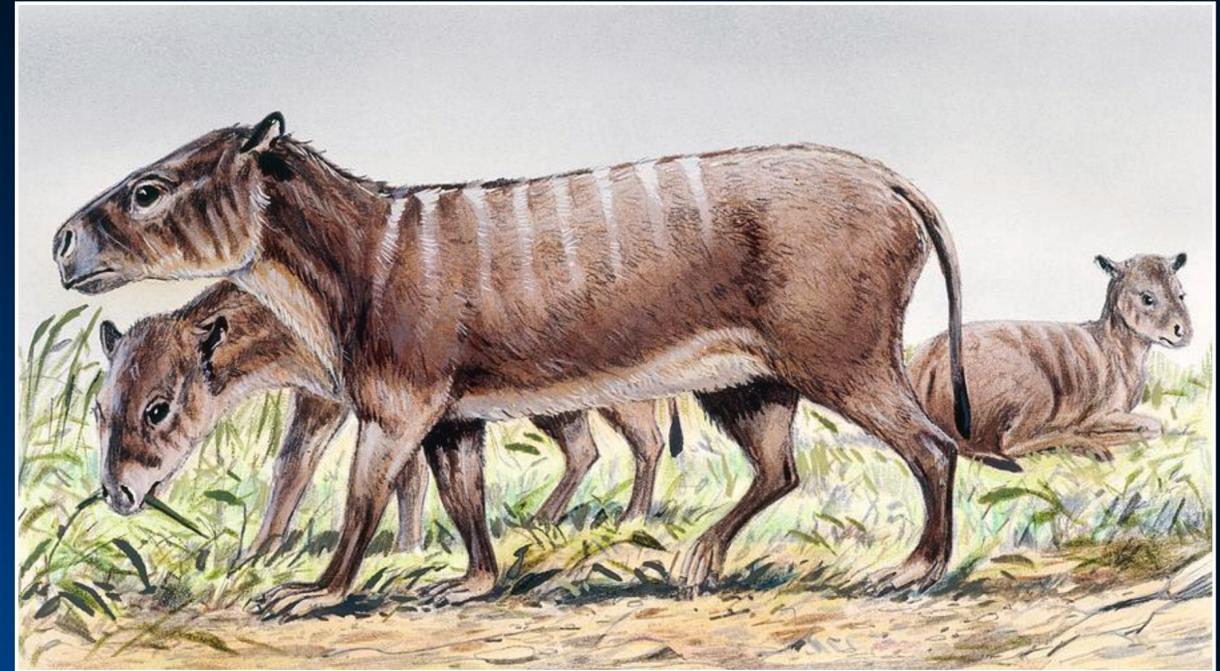
## Oreodont Dataset

Kentucky

No	Species	Skull Width	Cheek Tooth Row Length	Bulla Length	Bulla Depth
1	<i>Desmatochoerus hatcheri</i>	58	129.00	26.00	16.00
2	<i>Desmatochoerus hatcheri</i>	52	126.00	27.00	18.00
3	<i>Desmatochoerus hatcheri</i>	50	122.00	28.00	22.00
4	<i>Desmatochoerus hatcheri</i>	52	123.00	29.00	18.00
5	<i>Desmatochoerus hatcheri</i>	60	138.00	33.00	17.00
6	<i>Desmatochoerus hatcheri</i>	61	122.00	28.00	17.00
7	<i>Desmatochoerus hatcheri</i>	54	132.00	30.00	17.00
8	<i>Desmatochoerus hatcheri</i>	65	131.00	32.00	18.00
9	<i>Desmatochoerus hatcheri</i>	55	130.00	32.00	17.00
10	<i>Desmatochoerus hatcheri</i>	64	125.00	26.00	16.00
11	<i>Desmatochoerus hatcheri</i>	56	124.00	28.00	16.00
12	<i>Merychoidodon culbertsoni</i>	45	91.00	16.00	7.50
13	<i>Merychoidodon culbertsoni</i>	46	93.00	17.00	6.50
14	<i>Merychoidodon culbertsoni</i>	48	92.00	19.00	5.00
15	<i>Merychoidodon culbertsoni</i>	46	91.00	19.00	6.00

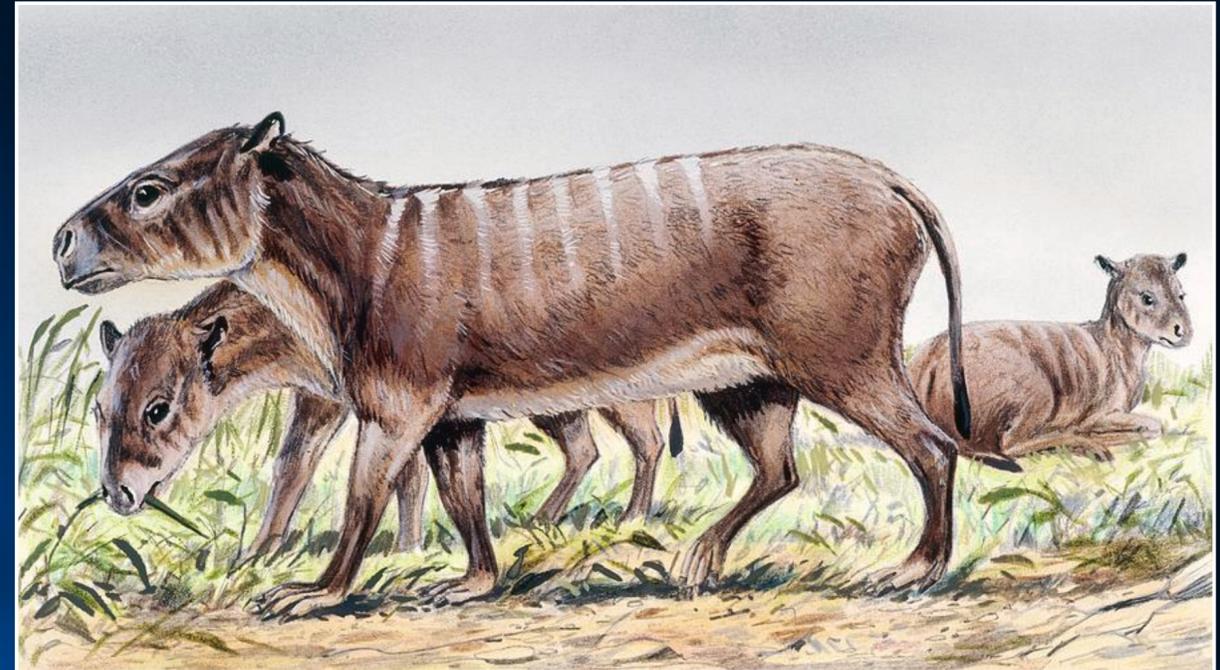
# Lab Practical Assignment II

- Plot these data. (10 points)
- Select a set of linear regression variables that will allow these to be distinguished from one another to the maximum extent possible. (10 points)
  - a. Justify your selection. (20 points)
- List the equation of the regression line for the model you have selected. (10 points)
  - a. Use this equation to predict the taxonomic identities of the unknown skulls. (20 points)



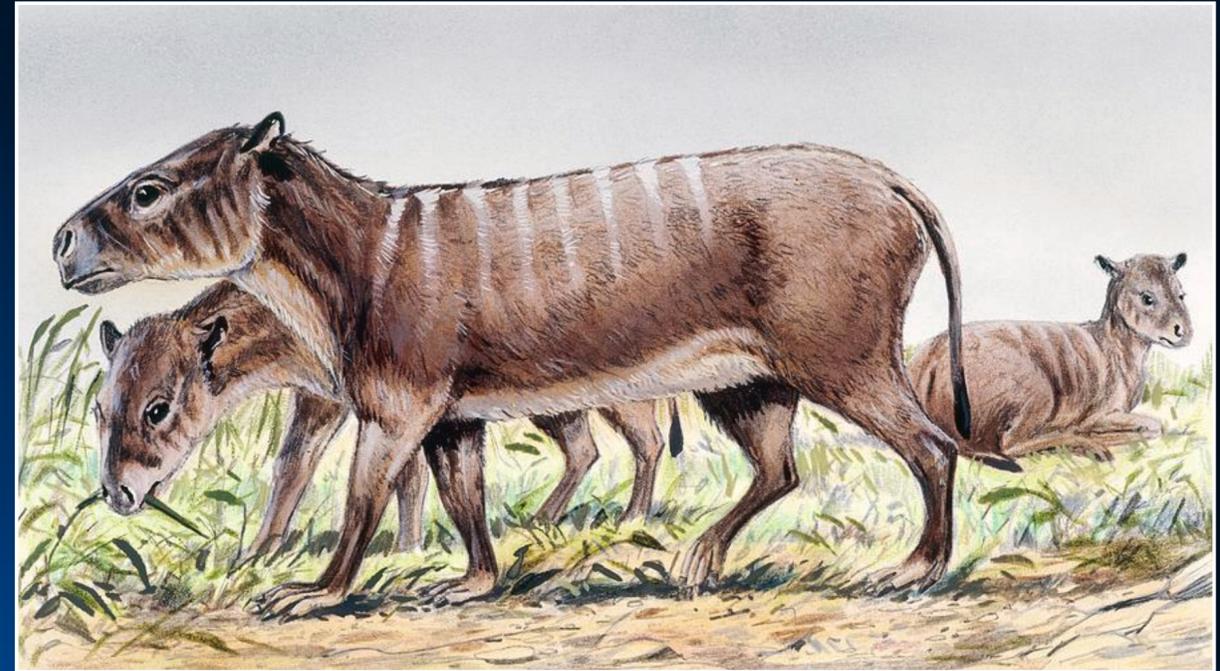
# Lab Practical Assignment II

- Use an ANOVA  $F$  test to estimate the significance of the regression model.
  - a. To two decimal places state the probability value associated with the ANOVA test result. (10 points)
  - b. Provide an interpretation of the ANOVA test result in terms of the degree to which the regression model can be regarded as constituting an accurate prediction. (20 points)
  - c. Estimate the 95% confidence interval for the regression result. (20 points)
    1. Determine with the range of variation predicted values has an effect of the reliability of your identifications. (30 points)



# Lab Practical Assignment II

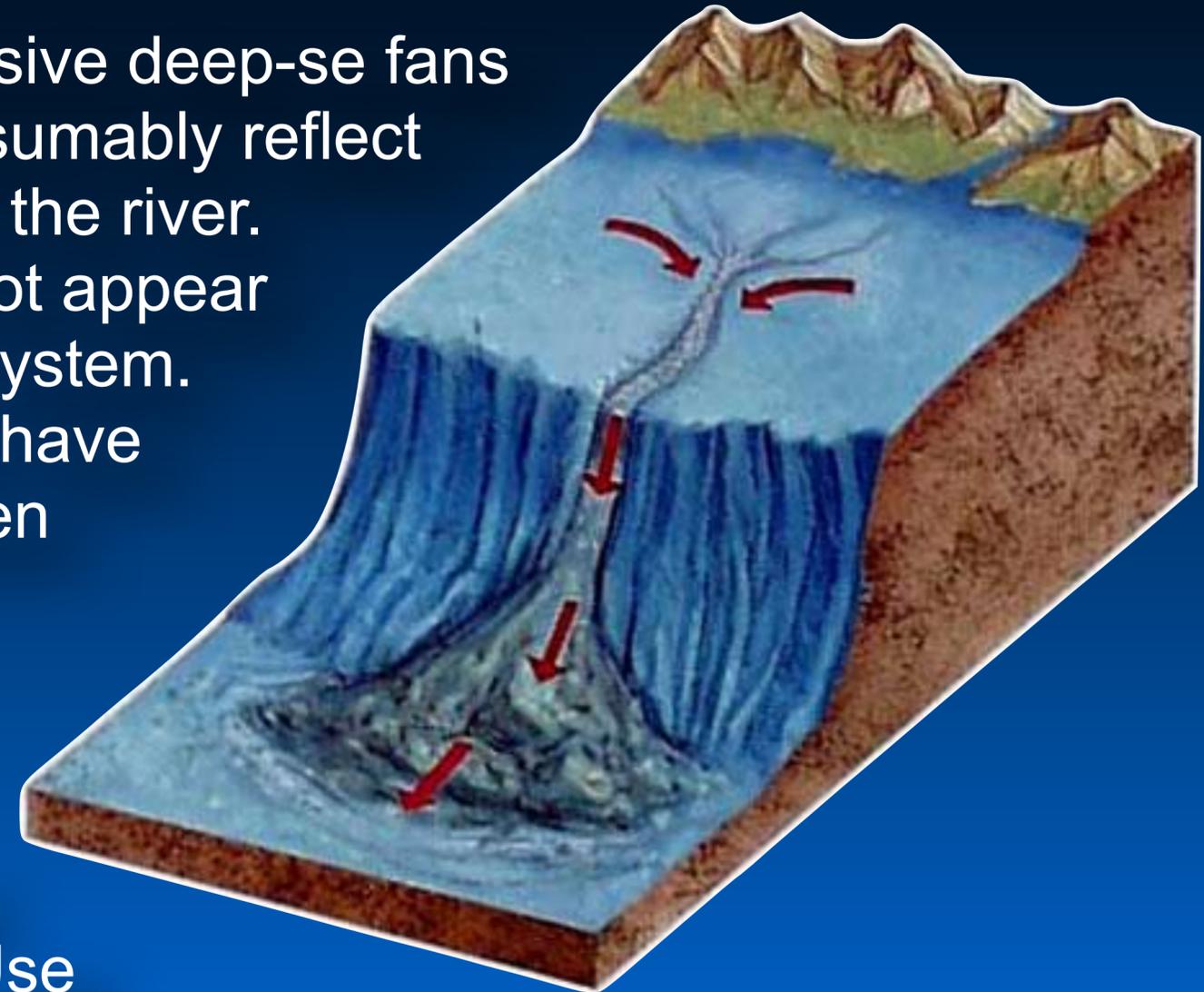
- Do these data conform to the assumptions of an ANOVA test?
  - a. Justify your reasoning. (20 points)
- Perform any additional test(s) you deem appropriate in order to confirm the validity of the regression model.
  - a. Show all plots, secondary statistical tests, and results associated with these additional tests (if any are warranted). (20 points)
  - b. Describe how the results of any additional tests (if any are warranted) either increased or decreased your confidence in your regression analysis. (20 points)



# Lab Practical Assignment II

5. Many major rivers are contiguous with extensive deep-sea fans at their river mouths, the sizes of which presumably reflect the sediments loads that are transported by the river. However, submarine fans do exist that do not appear to be associated with a major current river system. These are probably relict fans whose rivers have either dried up due to climate change or been captured by a different drainage system.

The submarine fans dataset (Submarine Fans.dat, Submarine Fans.csv) contains data for 12 submarine fans associated with extant drainage systems and 5 relict fans. Use multiple regression analysis to predict the relation between the river discharge variable and the other relevant physical variables included in these data.



# Lab Practical Assignment II

## Submarine Fan Dataset

Kentucky

No.	Fan	Length, km	Volume, 1000 km <sup>3</sup>	Age, Ma	Deposition rate, 10 <sup>6</sup> cu m/yr	River discharge 10 <sup>6</sup> MT/yr
1	Amazon	700	700.00	16.50	42.00	900.00
2	Columbia	310	27.00	1.00	27.00	13.00
3	Ganges	2800	4000.00	21.00	190.00	2180.00
4	Crati	16	0.00	0.01	0.15	2.00
5	Indus	1500	1000.00	16.00	66.00	480.00
6	St. Lawrence	1200	100.00	2.00	50.00	15.00
7	Magdalena	270	180.00	21.00	11.10	220.00
8	Mississippi	570	290.00	2.40	121.00	400.00
9	Nile	280	140.00	5.20	27.00	142.00
10	Fraser	260	9.00	0.60	15.00	20.00
11	Rhone	440	40.00	5.20	7.70	10.00
12	Valencia	160	6.00	5.20	1.30	7.00
14	Cap Ferret	75	13.00	2.00	6.50	
15	Delgada	350	40.00	11.20	3.60	
16	La Jolla	40	12	2.0	0.6	

# Lab Practical Assignment II

- Plot these data. (10 points)
- Select a linear regression model that will allow the river discharge rate to be predicted as a function of the other relevant physical variables. (10 points)
  - a. Justify your selection. (20 points)
- List the equation of the regression line for the model you have selected. (10 points)
  - a. Use this equation to estimate the discharge rates for the missing rivers based on their relict fans. (50 points)
- Use an ANOVA  $F$  test to estimate the significance of your regression model.
  - a. To two decimal places state the probability value associated with the ANOVA test result. (10 points)



# Lab Practical Assignment II

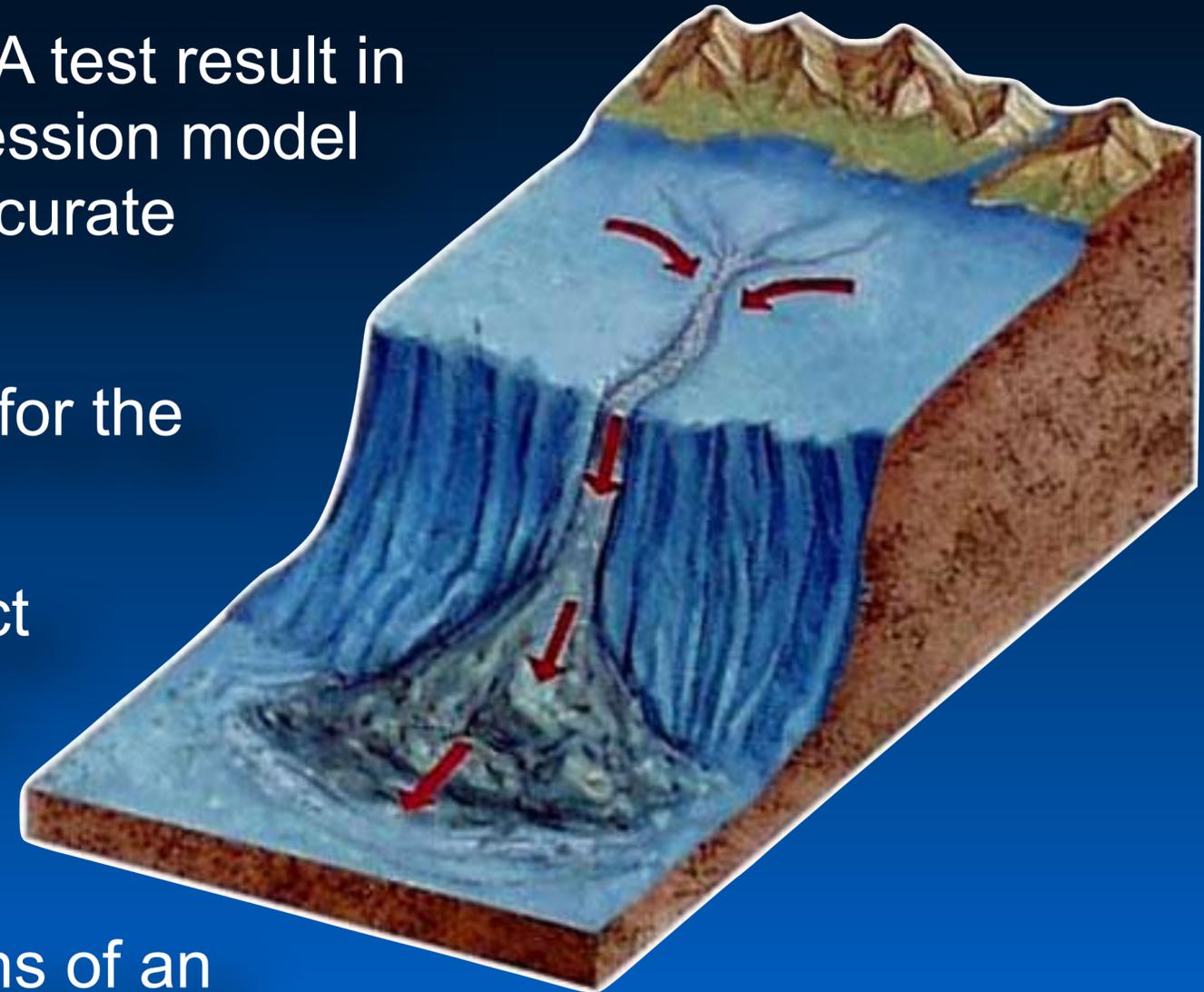
b. Provide an interpretation of the ANOVA test result in terms of the degree to which the regression model can be regarded as constituting an accurate prediction. (20 points)

c. Estimate the 95% confidence interval for the regression result. (20 points)

1. Estimate the range of variation relict river discharge rate values values that would be expected. (30 points)

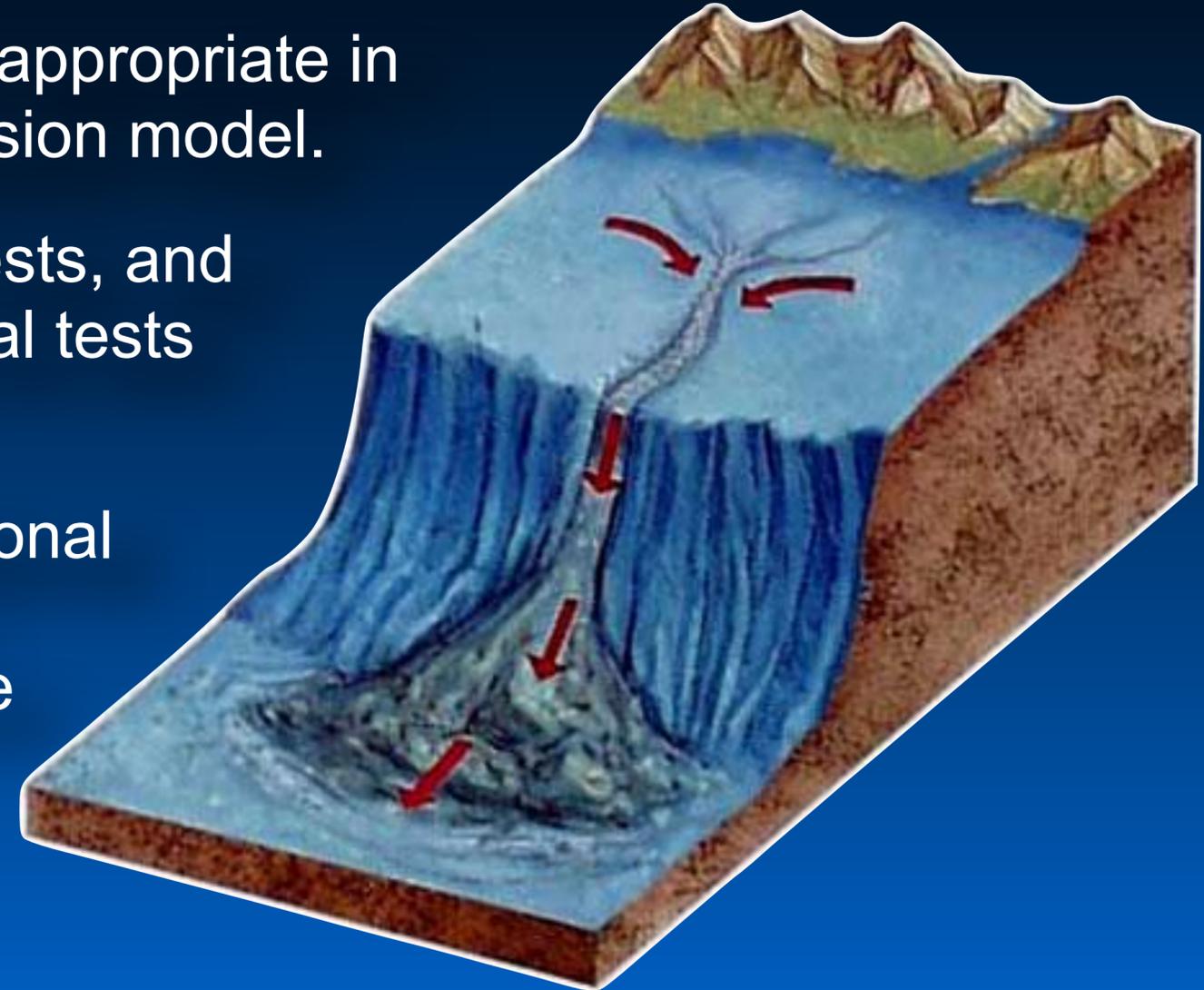
● Do these data conform to the assumptions of an ANOVA test?

a. Justify your reasoning. (20 points)



# Lab Practical Assignment II

- Perform any additional test(s) you deem appropriate in order to confirm the validity of the regression model.
  - a. Show all plots, secondary statistical tests, and results associated with these additional tests (if any are warranted). (20 points)
  - b. Describe how the results of any additional tests (if any are warranted) either increased or decreased your confidence in your regression analysis. (20 points)

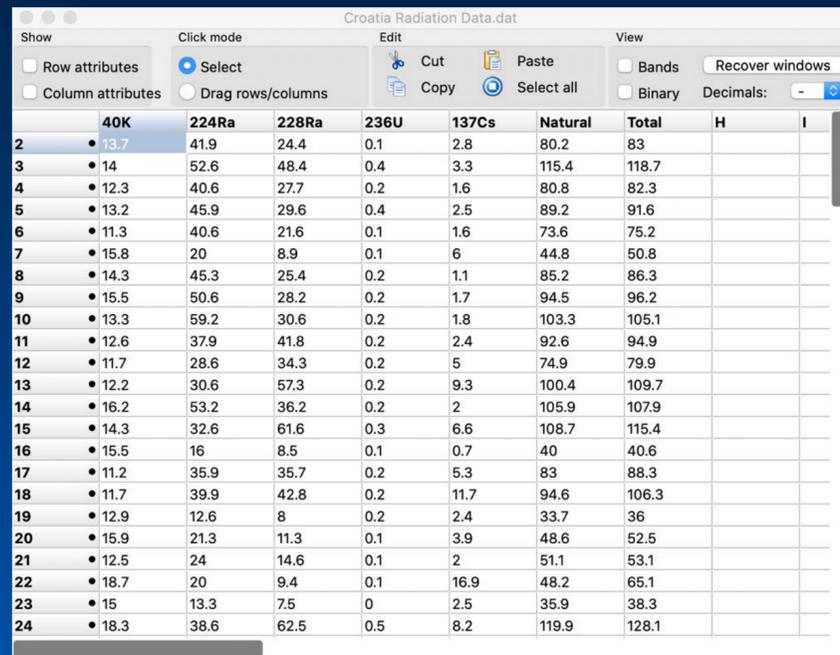


Total = 1150 points

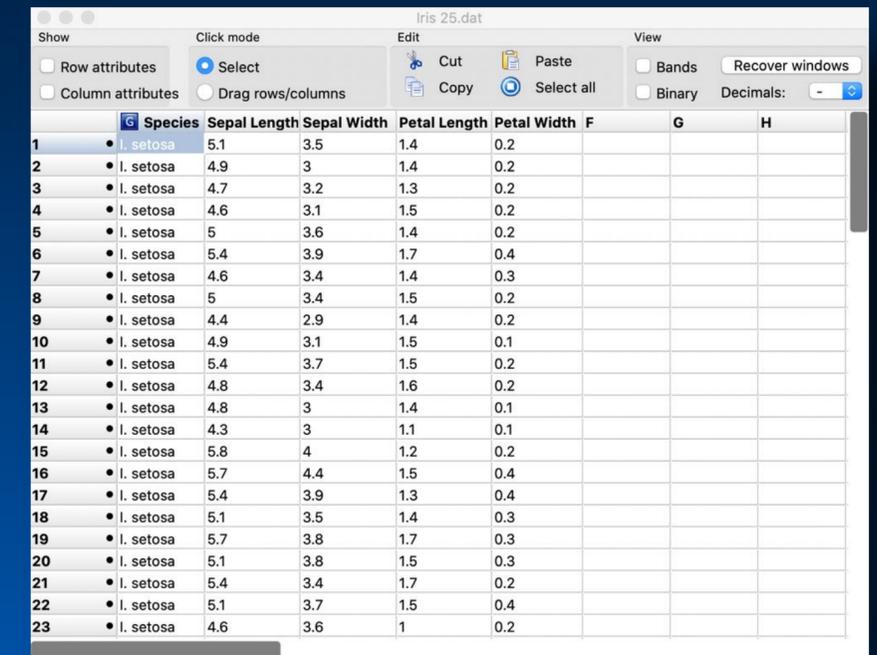
# Introduction to Statistics & Data Analysis in *PAST*

Prof. Norman MacLeod

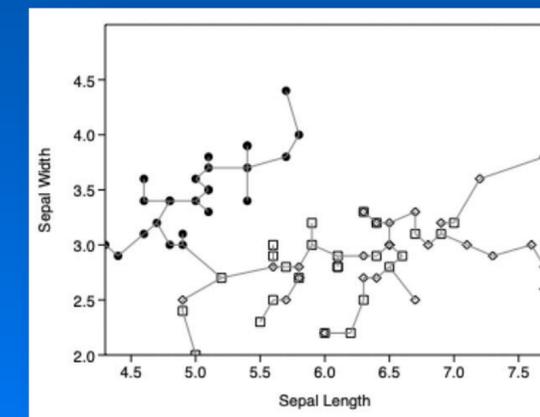
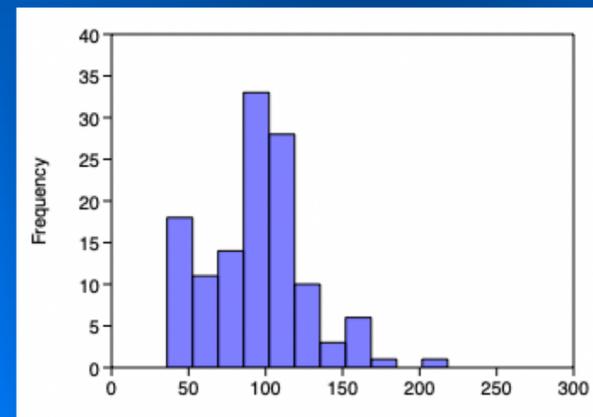
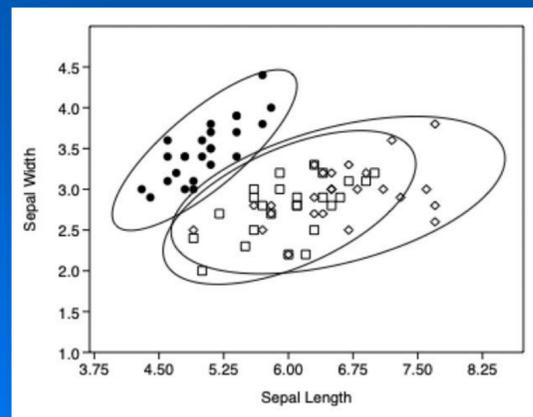
School of Earth Sciences & Engineering, Nanjing University



	40K	224Ra	228Ra	236U	137Cs	Natural	Total	H	I
2	13.7	41.9	24.4	0.1	2.8	80.2	83		
3	14	52.6	48.4	0.4	3.3	115.4	118.7		
4	12.3	40.6	27.7	0.2	1.6	80.8	82.3		
5	13.2	45.9	29.6	0.4	2.5	89.2	91.6		
6	11.3	40.6	21.6	0.1	1.6	73.6	75.2		
7	15.8	20	8.9	0.1	6	44.8	50.8		
8	14.3	45.3	25.4	0.2	1.1	85.2	86.3		
9	15.5	50.6	28.2	0.2	1.7	94.5	96.2		
10	13.3	59.2	30.6	0.2	1.8	103.3	105.1		
11	12.6	37.9	41.8	0.2	2.4	92.6	94.9		
12	11.7	28.6	34.3	0.2	5	74.9	79.9		
13	12.2	30.6	57.3	0.2	9.3	100.4	109.7		
14	16.2	53.2	36.2	0.2	2	105.9	107.9		
15	14.3	32.6	61.6	0.3	6.6	108.7	115.4		
16	15.5	16	8.5	0.1	0.7	40	40.6		
17	11.2	35.9	35.7	0.2	5.3	83	88.3		
18	11.7	39.9	42.8	0.2	11.7	94.6	106.3		
19	12.9	12.6	8	0.2	2.4	33.7	36		
20	15.9	21.3	11.3	0.1	3.9	48.6	52.5		
21	12.5	24	14.6	0.1	2	51.1	53.1		
22	18.7	20	9.4	0.1	16.9	48.2	65.1		
23	15	13.3	7.5	0	2.5	35.9	38.3		
24	18.3	38.6	62.5	0.5	8.2	119.9	128.1		

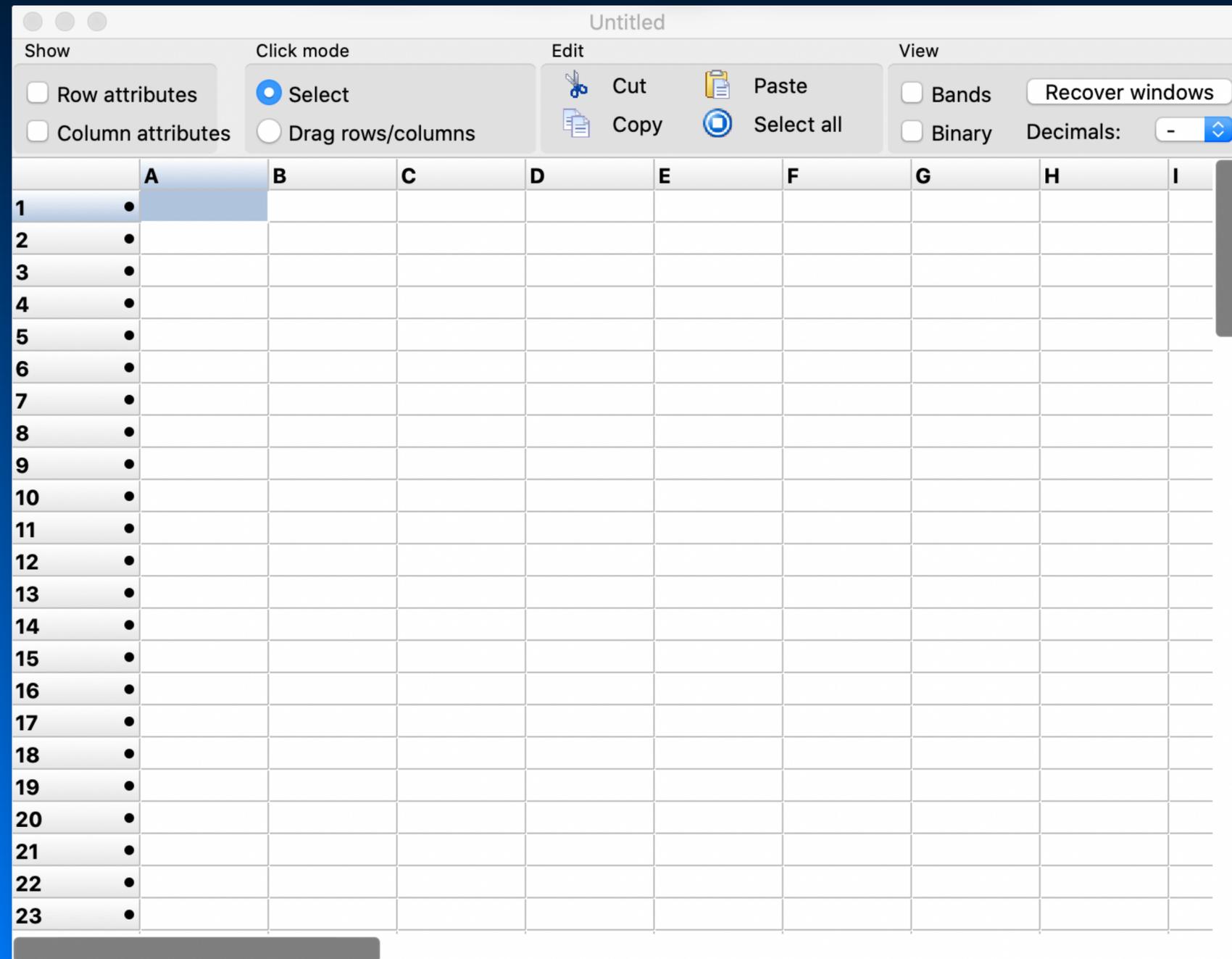


	Species	Sepal Length	Sepal Width	Petal Length	Petal Width	F	G	H
1	i. setosa	5.1	3.5	1.4	0.2			
2	i. setosa	4.9	3	1.4	0.2			
3	i. setosa	4.7	3.2	1.3	0.2			
4	i. setosa	4.6	3.1	1.5	0.2			
5	i. setosa	5	3.6	1.4	0.2			
6	i. setosa	5.4	3.9	1.7	0.4			
7	i. setosa	4.6	3.4	1.4	0.3			
8	i. setosa	5	3.4	1.5	0.2			
9	i. setosa	4.4	2.9	1.4	0.2			
10	i. setosa	4.9	3.1	1.5	0.1			
11	i. setosa	5.4	3.7	1.5	0.2			
12	i. setosa	4.8	3.4	1.6	0.2			
13	i. setosa	4.8	3	1.4	0.1			
14	i. setosa	4.3	3	1.1	0.1			
15	i. setosa	5.8	4	1.2	0.2			
16	i. setosa	5.7	4.4	1.5	0.4			
17	i. setosa	5.4	3.9	1.3	0.4			
18	i. setosa	5.1	3.5	1.4	0.3			
19	i. setosa	5.7	3.8	1.7	0.3			
20	i. setosa	5.1	3.8	1.5	0.3			
21	i. setosa	5.4	3.4	1.7	0.2			
22	i. setosa	5.1	3.7	1.5	0.4			
23	i. setosa	4.6	3.6	1	0.2			



# Paleontological Statistics (PAST)

## The PAST Spreadsheet





# Paleontological Statistics (PAST)

## Linear Regression Analysis in PAST

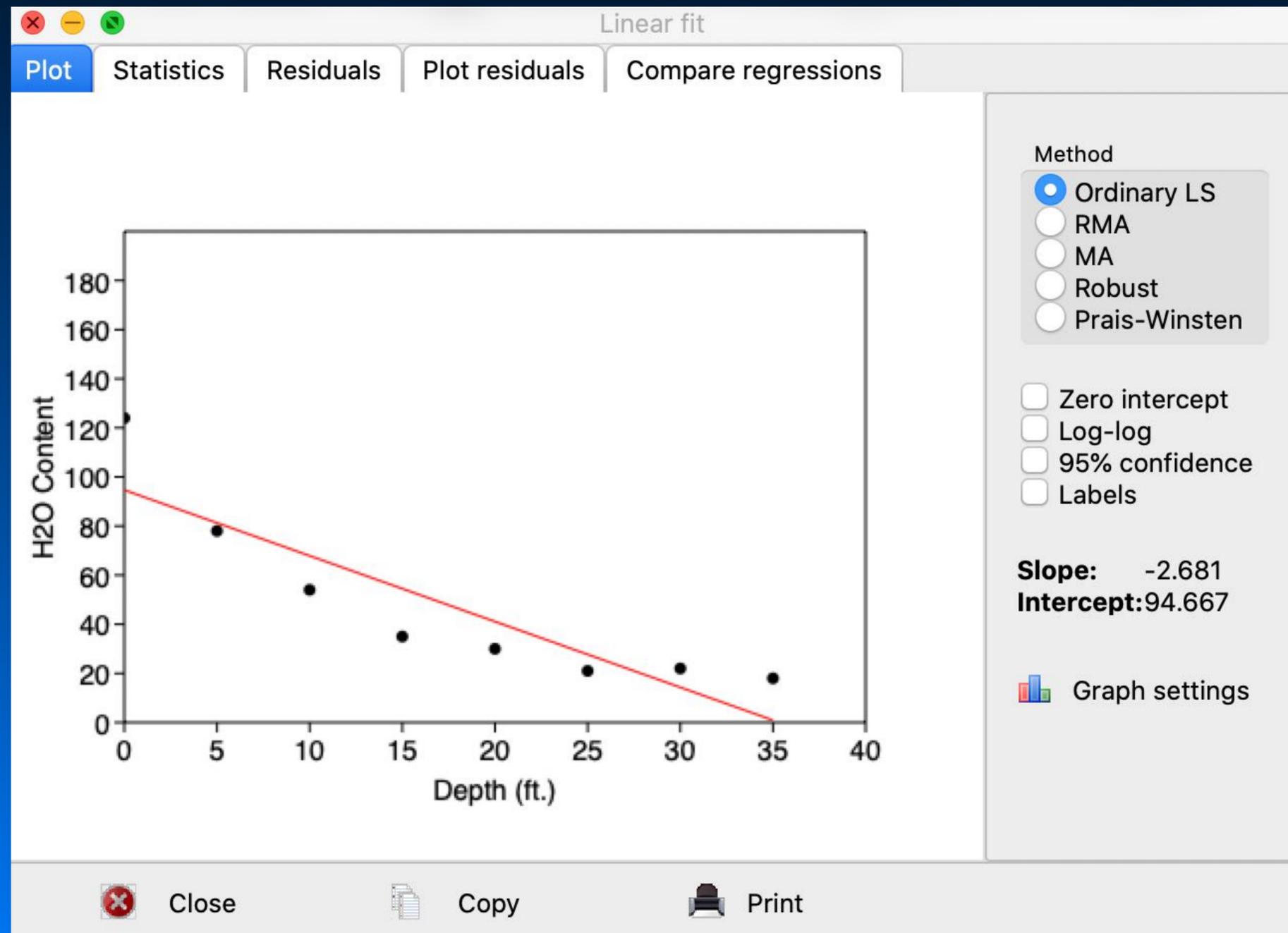
The screenshot displays the PAST software interface. The main window shows a data table with the following columns: Depth (ft.), H2O Content C, D, E, F, G, H, and I. The data is as follows:

	Depth (ft.)	H2O Content C	D	E	F	G	H	I
1	0	124						
2	5	78						
3	10	54						
4	15	35						
5	20	30						
6	25	21						
7	30	22						
8	35	18						
9								
10								
11								
12								
13								
14								
15								
16								
17								
18								
19								
20								
21								
22								
23								

The 'Model' menu is open, showing options for Linear, Generalized Linear Model, Polynomial, Nonlinear fit, Sum-of-sinusoids, Smoothing spline, LOESS smoothing, Abundance distribution, Species packing (Gaussian), Mixture analysis, Logarithmic spiral, and Change-point. The 'Bivariate' submenu is also open, showing options for Multivariate (1 independent, n dependent), Multiple (1 dependent, n independent), and Multivariate multiple (m indep., n dep.).

# Paleontological Statistics (PAST)

## Linear Regression Analysis in PAST



# Paleontological Statistics (PAST)

## Linear Regression Analysis in PAST

Linear fit

Plot Statistics **Residuals** Plot residuals Compare regressions

Depth (ft.)	H2O Content	Regress.	Residual
0	124	94.667	29.333
5	78	81.262	-3.2619
10	54	67.857	-13.857
15	35	54.452	-19.452
20	30	41.048	-11.048
25	21	27.643	-6.6429
30	22	14.238	7.7619
35	18	0.83333	17.167

Std. error of estimate:  
15.663

Durbin-Watson  
statistic:  
0.81116  
**p (no pos. autocorr.):**  
0.0070934

Breusch-Pagan  
statistic:  
1.6967  
**p (homoskedastic):**  
0.19272

Close Copy Print

# Paleontological Statistics (PAST)

---

## Evaluating a Linear Regression Analysis

- Take a look at the regression plot. Does the regression line fit the data as it should?
- Take a look at the regression equation. Do the slope and intercept look correct when compared to the plot?
- Look at the regression statistics. PAST calculates a  $t$ -test rather than an ANOVA to test the regression's significance. Look at the  $p$ -value. Compare the  $p$ -value to the level of significance you think is appropriate or that was requested.
- Look at the plot of residual values. A good regression should have residuals that (1) exhibit no slope and (2) are distributed randomly about the regression line.
- Ask yourself is this the best fit you could achieve for your data.



# Paleontological Statistics (PAST)

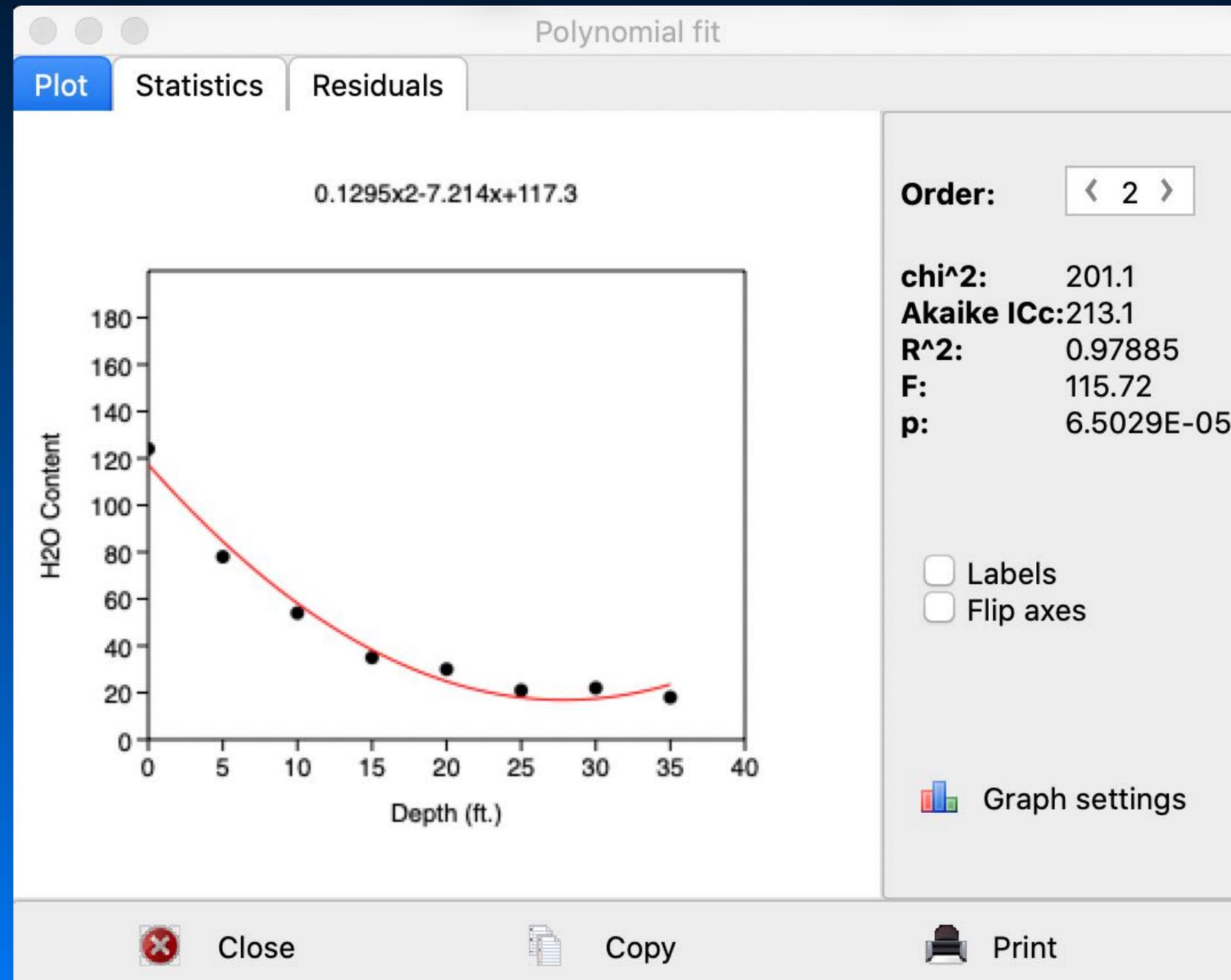
## Polynomial Regression Analysis in PAST

The screenshot displays the PAST software interface. The 'Model' menu is open, showing various statistical options. The 'Polynomial' option is highlighted. Below the menu, a data table is visible with columns for 'Depth (ft.)' and 'H2O Content C'. The data points are as follows:

Row	Depth (ft.)	H2O Content C
1	0	124
2	5	78
3	10	54
4	15	35
5	20	30
6	25	21
7	30	22
8	35	18
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		
21		
22		
23		

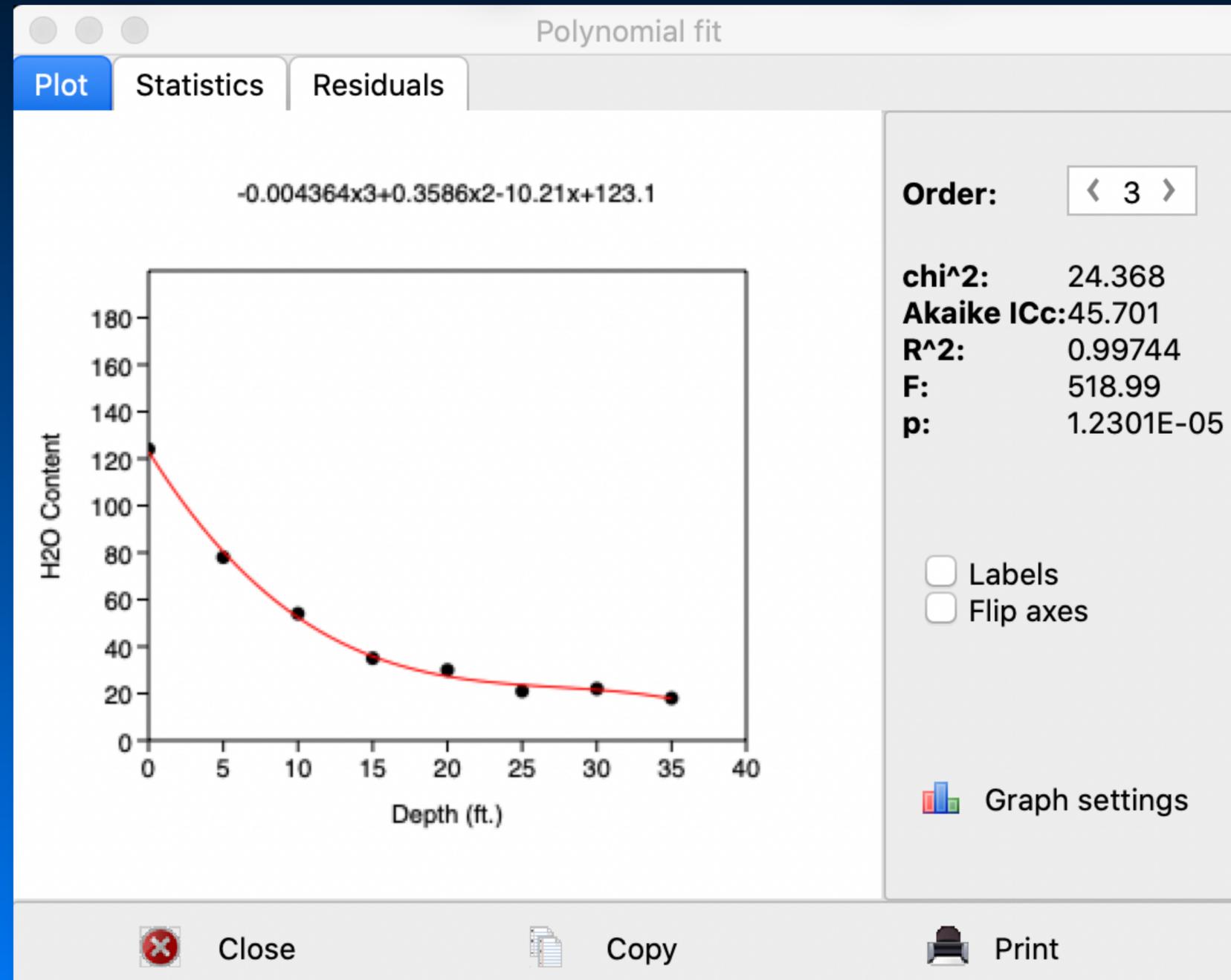
# Paleontological Statistics (PAST)

## Polynomial Regression Analysis in PAST



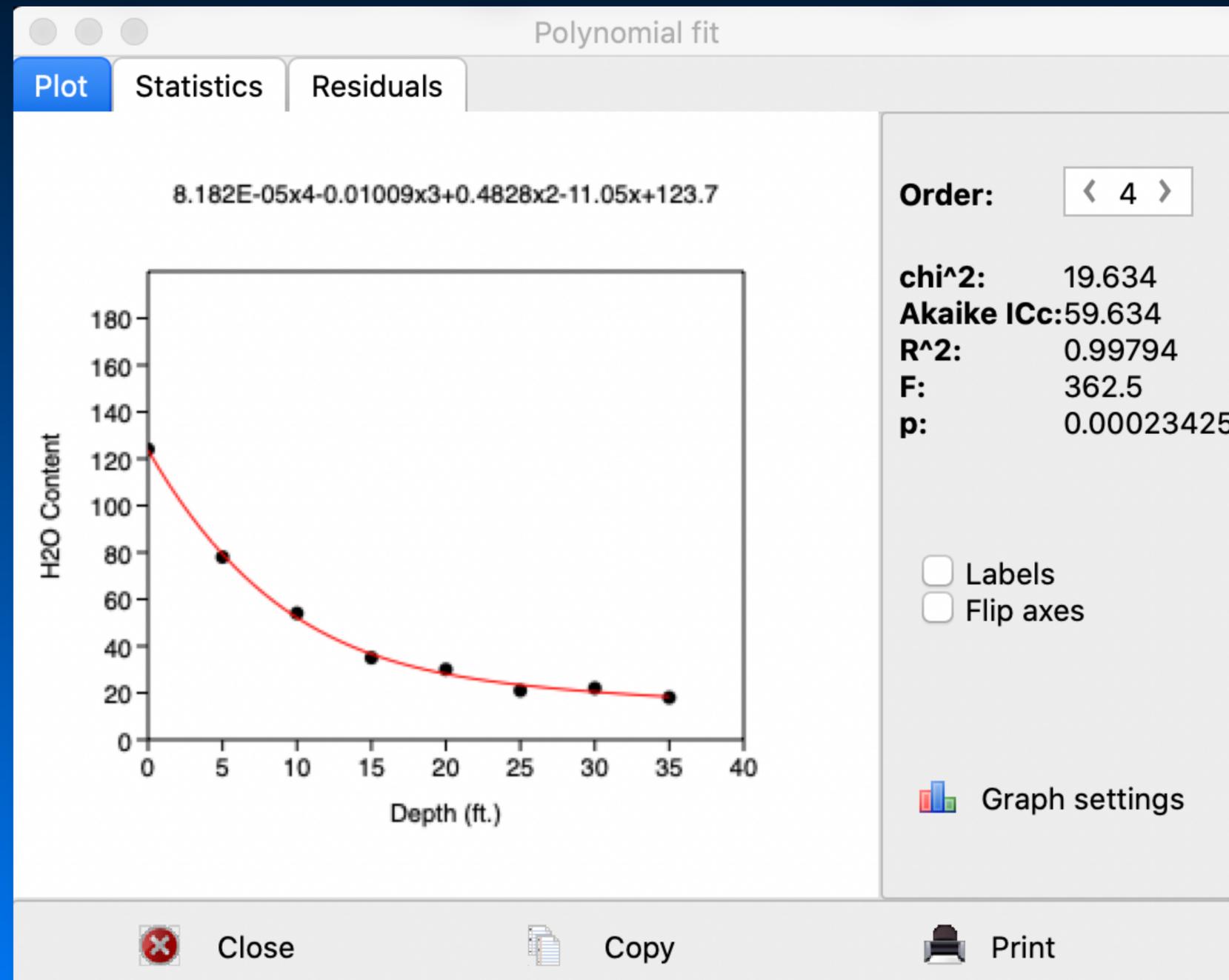
# Paleontological Statistics (PAST)

## Polynomial Regression Analysis in PAST



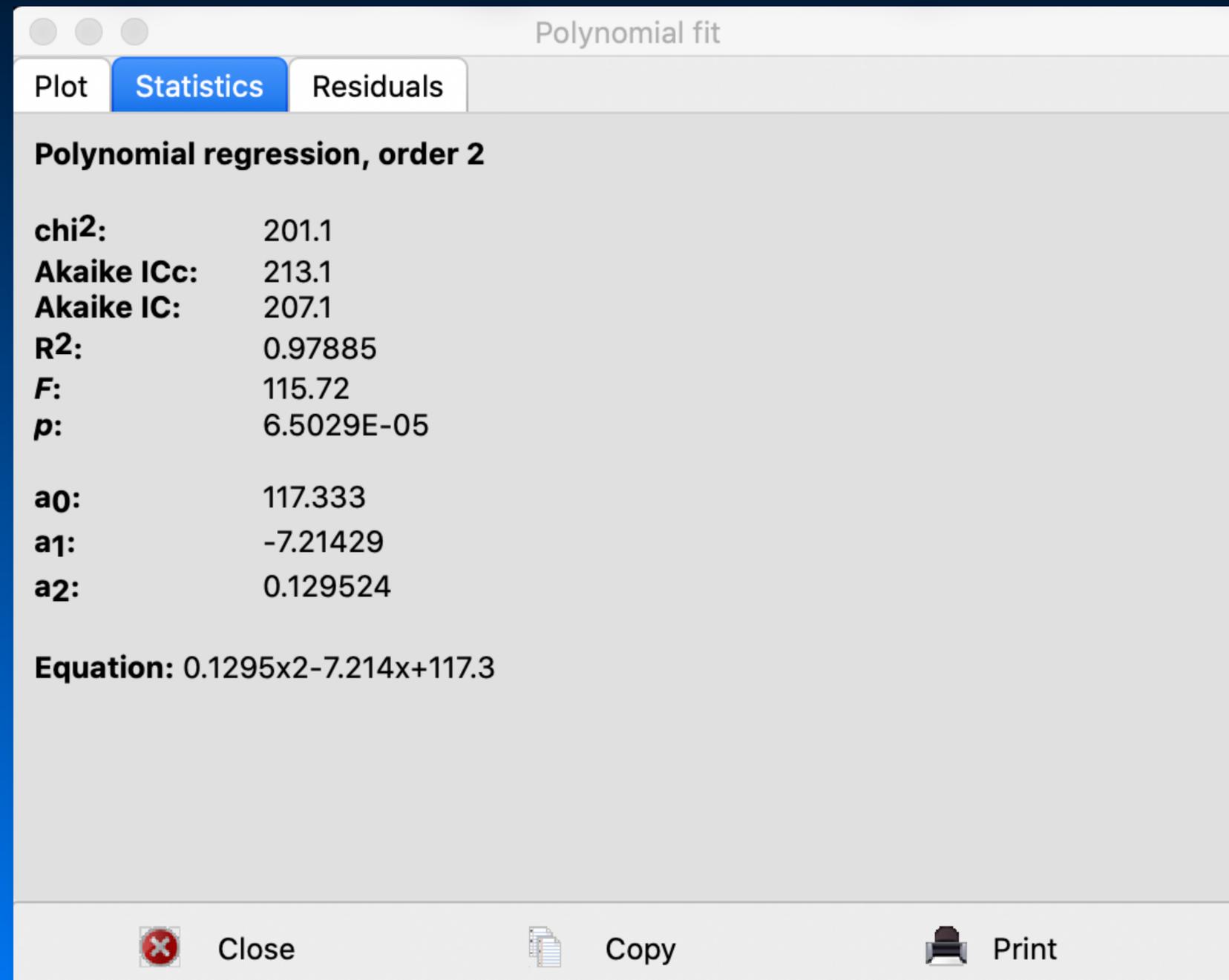
# Paleontological Statistics (PAST)

## Polynomial Regression Analysis in PAST



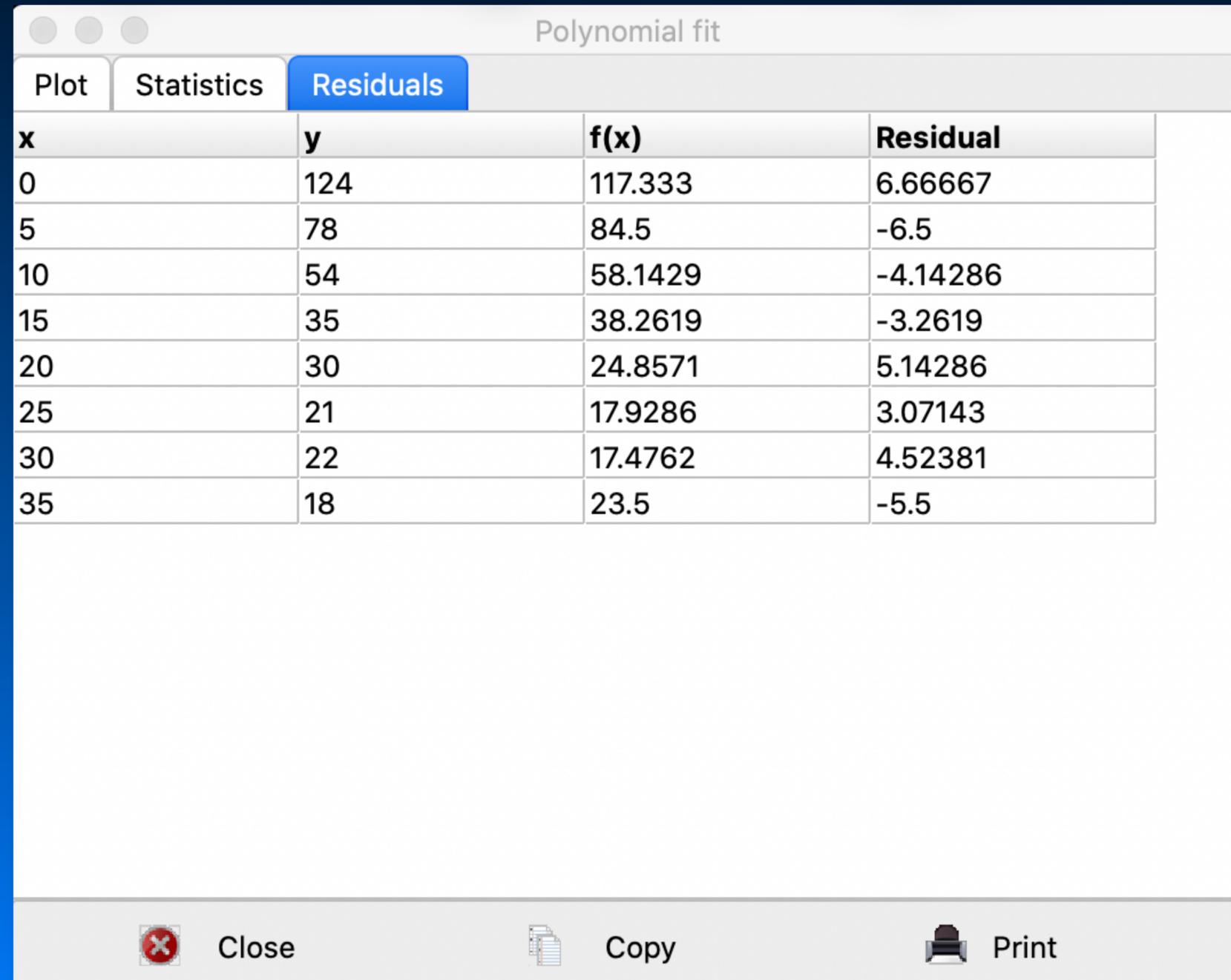
# Paleontological Statistics (PAST)

## Polynomial Regression Analysis in PAST



# Paleontological Statistics (PAST)

## Polynomial Regression Analysis in PAST



The screenshot shows a window titled "Polynomial fit" with three tabs: "Plot", "Statistics", and "Residuals". The "Residuals" tab is active, displaying a table with four columns: "x", "y", "f(x)", and "Residual". The data points are as follows:

x	y	f(x)	Residual
0	124	117.333	6.66667
5	78	84.5	-6.5
10	54	58.1429	-4.14286
15	35	38.2619	-3.2619
20	30	24.8571	5.14286
25	21	17.9286	3.07143
30	22	17.4762	4.52381
35	18	23.5	-5.5

At the bottom of the window, there are three buttons: "Close" (with a red X icon), "Copy" (with a document icon), and "Print" (with a printer icon).

# Lab Practical Assignment II

## The Analysis of Bivariate Data: Linear Regression

**Ordinary Least Squares Regression: Depth (ft.)-H2O Content**

<b>Slope a:</b>	-2.681	<b>Std. error a:</b>	0.55815
<b>t:</b>	4.8033	<b>p (slope):</b>	0.0029906
<b>Intercept b:</b>	94.667	<b>Std. error b:</b>	11.674

**95% bootstrapped confidence intervals (N=1999):**

<b>Slope a:</b>	(-4.156, -1.071)
<b>Intercept b:</b>	(72.832, 132.32)

**Correlation:**

<b>r:</b>	-0.89085
<b>r<sup>2</sup>:</b>	0.79361

