A Geostatistical Approach to the Characteristic Values

Horatiu V. Corbeanu
Halcrow/CH2M HILL

07/11/12 – 12:15

Pro Statistics

- *"Statistical methods . . . constitute the science of collecting, analysing and interpreting data **in the best possible way**"* C. Chatfield (*Statistics of Technology, 3rd ed. 1983*)

- *"The product of an arithmetical computation is the answer to an equation; it is not the solution to a problem"* - G.O. Ashley

- *"Years ago a statistician might have claimed that statistics deals with the processing of data. . . today's statistician will be more likely to say that statistics is concerned with decision making in the face of uncertainty"* – H.Chernonff & L.E. Moses (*Elementary Decision Theory, 1959*)

Geotechnica ME equip6

Halcrow
A CH2M HILL COMPANY

Abuse of Statistics

- "*Attempts by statisticians to tackle geotechnical design have often ended in ridicule, and it is very difficult for one person to have sufficient grasp of both disciplines that he can use them sensibly*", Decoding Eurocode 7, A.Bond, A.Harris, 2008

- "*There are lies, damned lies, and statistics*", Benjamin Disraeli

Twisted Example

$$a = b$$

$$a^2 = b^2$$

$$a^2 - b^2 = 0$$

$$(a-b)*(a+b) = 0$$

$$(a-b)*(a+b) = (a-b)$$

$$\frac{(a-b)*(a+b)}{(a-b)} = 1$$

$$a+b = 1$$

$$a = 1-b$$

Geotechnica ME equip

Halcrow
A CH2M HILL COMPANY

Objective

- The Problem
- The Dataset
- General presentation of geotechnical data
- Distribution of geotechnical data
  - Descriptive statistical techniques
  - Transformations to a normal distribution
- Conclusions

Geotechnica ME equip

Halcrow
A CH2M HILL COMPANY

The Problem

1) Provide a reliable characteristic value
2) Identify the mathematical tool that can provide confidence to the engineering judgement.

Geotechnica ME equip6

Halcrow
A CH2M HILL COMPANY

The Dataset

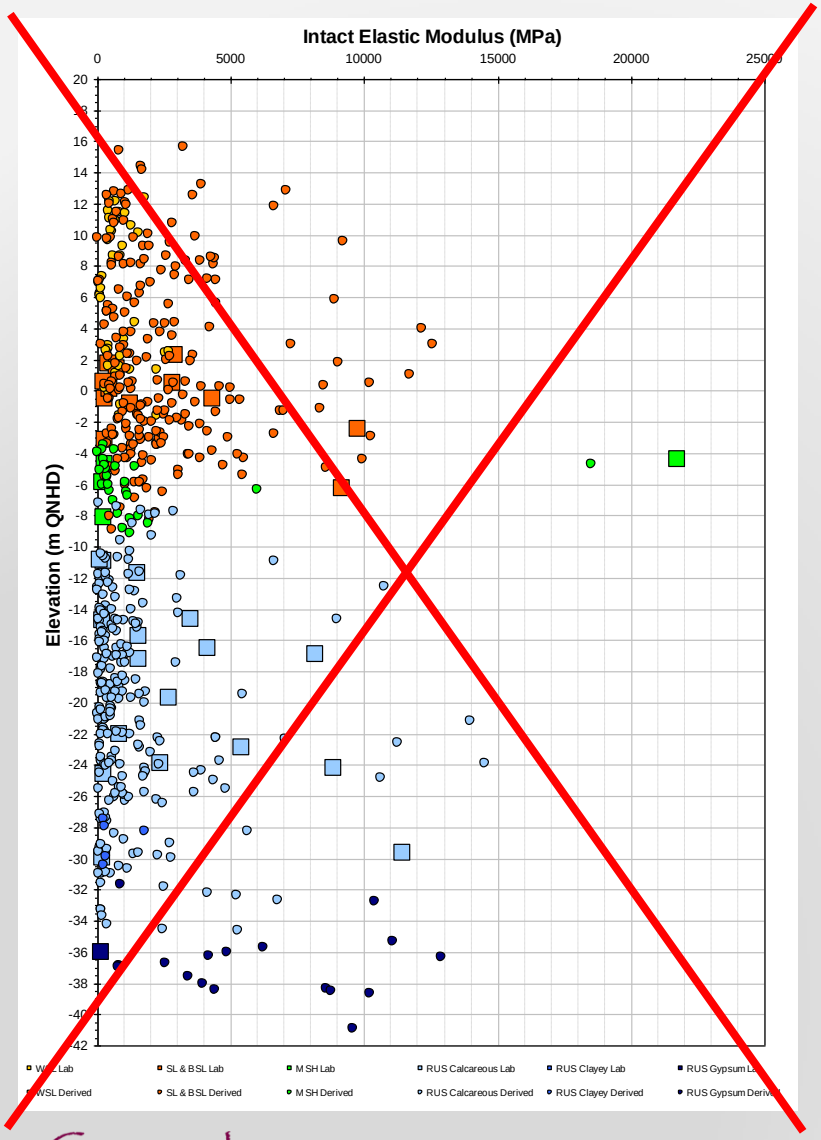
HOLE	EASTING	NORTHING	GEOL_GEOL	ELEVATION	ROCK_UCS	ROCK_E	ROCK_MU	ROCK_MC	ROCK_BDEN
01.BH-353	228,131	393,303	SL	-6.09	13.851	13100	0.265	1.7	21
01.BH-353	228,131	393,303	RUS	-19.59	5.678	8300		1.4	17.6
01.BH-536	229,228	393,111	RUS	-15.52	7.18			1.8	18.9
01.BH-536	229,228	393,111	RUS	-21.52	41.884	26900	0.172	1.2	26.2
01.BH-570	225,071	394,022	MSH	-3.80	25.599	22800		0.6	18
01.BH-570	225,071	394,022	RUS	-12.80	18.3	17600		1.1	17.4
01.BH-570	225,071	394,022	RUS	-23.82	12.85	11700	0.206	1.3	20.2
01.BH-572	226,451	393,594	RUS	-10.49	7.957			1.6	20.4
01.BH-572	226,451	393,594	RUS	-13.49	8.156	7700	0.212	1.9	20.5
01.BH-572	226,451	393,594	RUS	-22.49	12.174	11800	0.238	2.1	21.6
01.BH-755	224,397	394,360	SL	-6.49	11.736			2	22.9
01.BH-755	224,397	394,360	MSH	-14.54	40.022			1.4	24
01.BH-755	224,397	394,360	RUS	-17.54	15.158			2.8	21.9
01.BH-756	220,978	395,899	RUS	0.81	15.197	19500		2.6	17
01.BH-756	220,978	395,899	RUS	-5.19	12.452	16400		2.1	17.1
01.BH-756	220,978	395,899	RUS	-12.69	10.523	14100	0.215	2.5	25.1
01.BH-757	221,265	395,957	MSH	1.38	16.331	17000	0.200	3.9	22.6

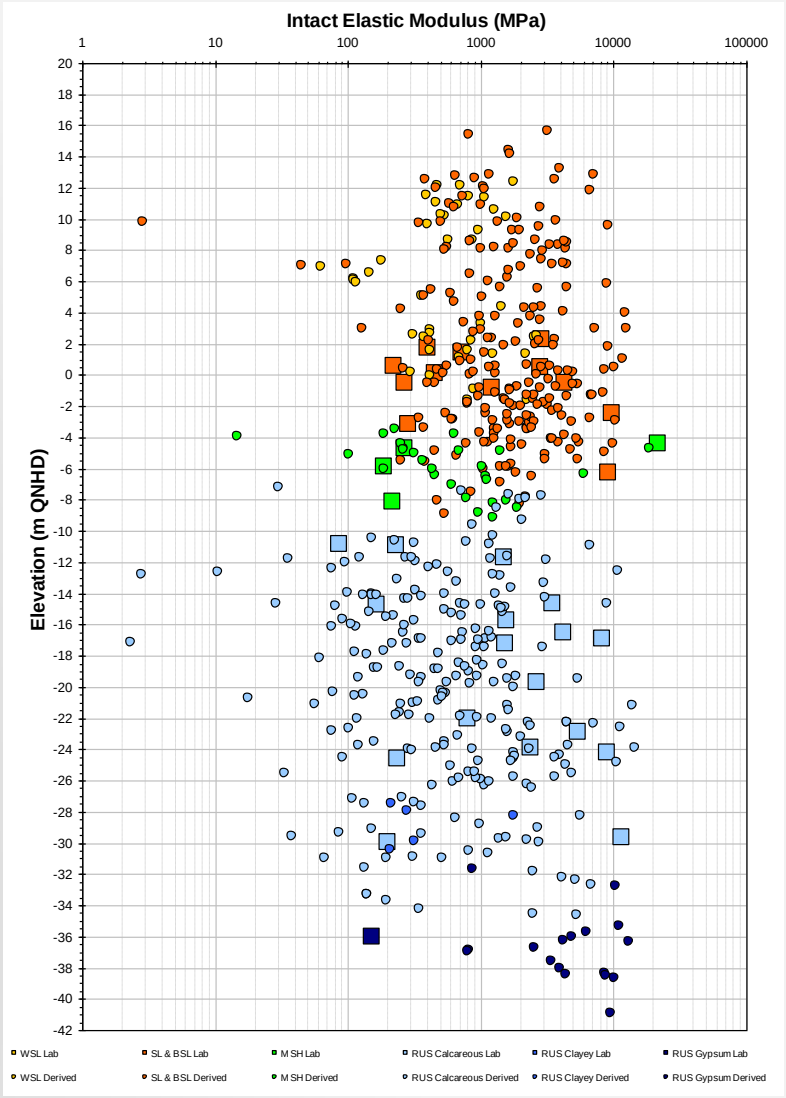
Count : 226
Minimum : 2.83 MPa
Maximum : 12,567 MPa
Mean : 2,531 MPa
5% Fractile : 240 MPa

Geotechnica ME equip

Halcrow
A CH2M HILL COMPANY

General Presentation of Data

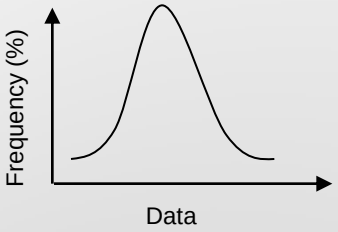




Geotechnica ME equip

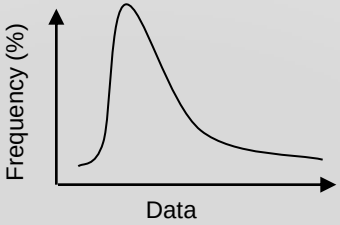
Halcrow
A CH2M HILL COMPANY

Shapes of Data Distribution

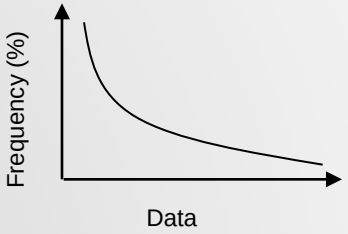


Normal

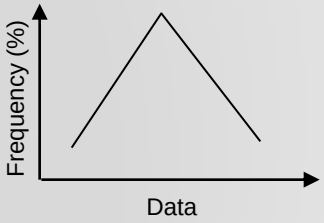
Uniform



Lognormal



Exponential



Triangular

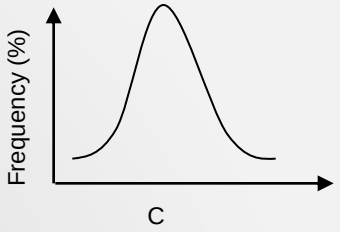
Geotechnica ME equip

Halcrow
A CH2M HILL COMPANY

Fact for Log Normal Distribution

Two geological processes: a & b (i.e. rate of deposition, duration of deposition)

$c = a + b$

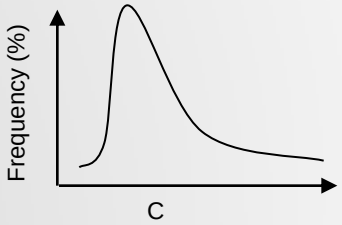


Normal

$c = a * b$

$\text{Log}(c) = \text{Log}(a*b)$

$\text{Log}(c) = \text{Log}(a) + \text{Log}(b)$

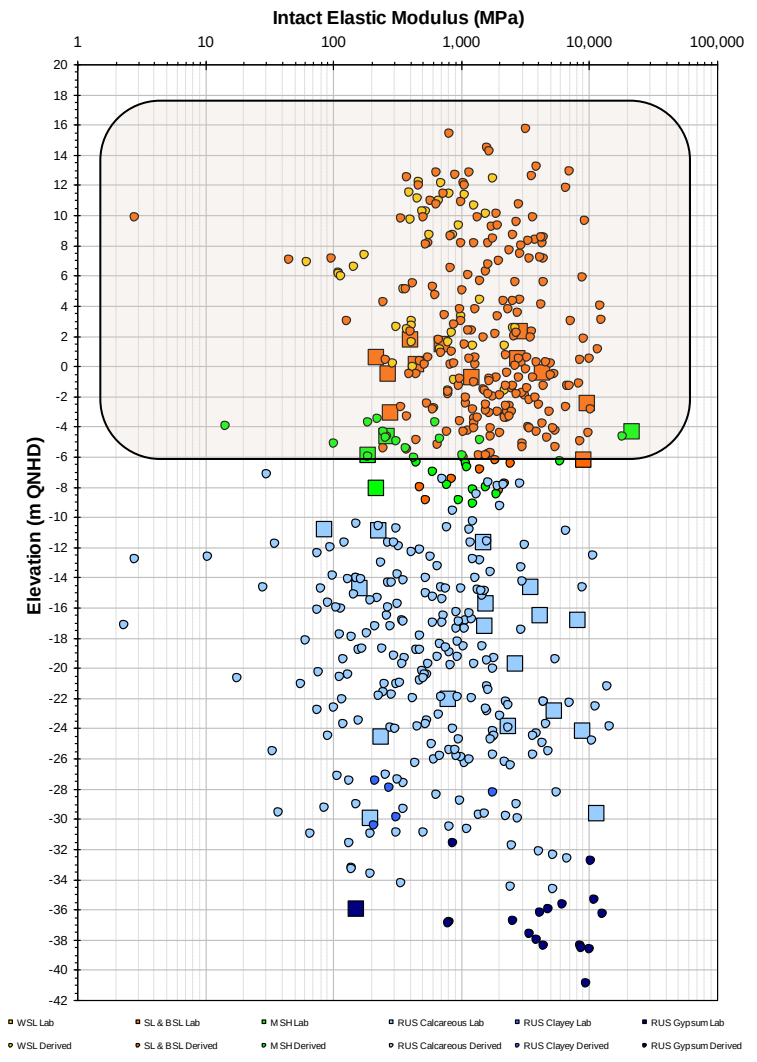


Lognormal

Geotechnica ME equip6

Halcrow
A CH2M HILL COMPANY

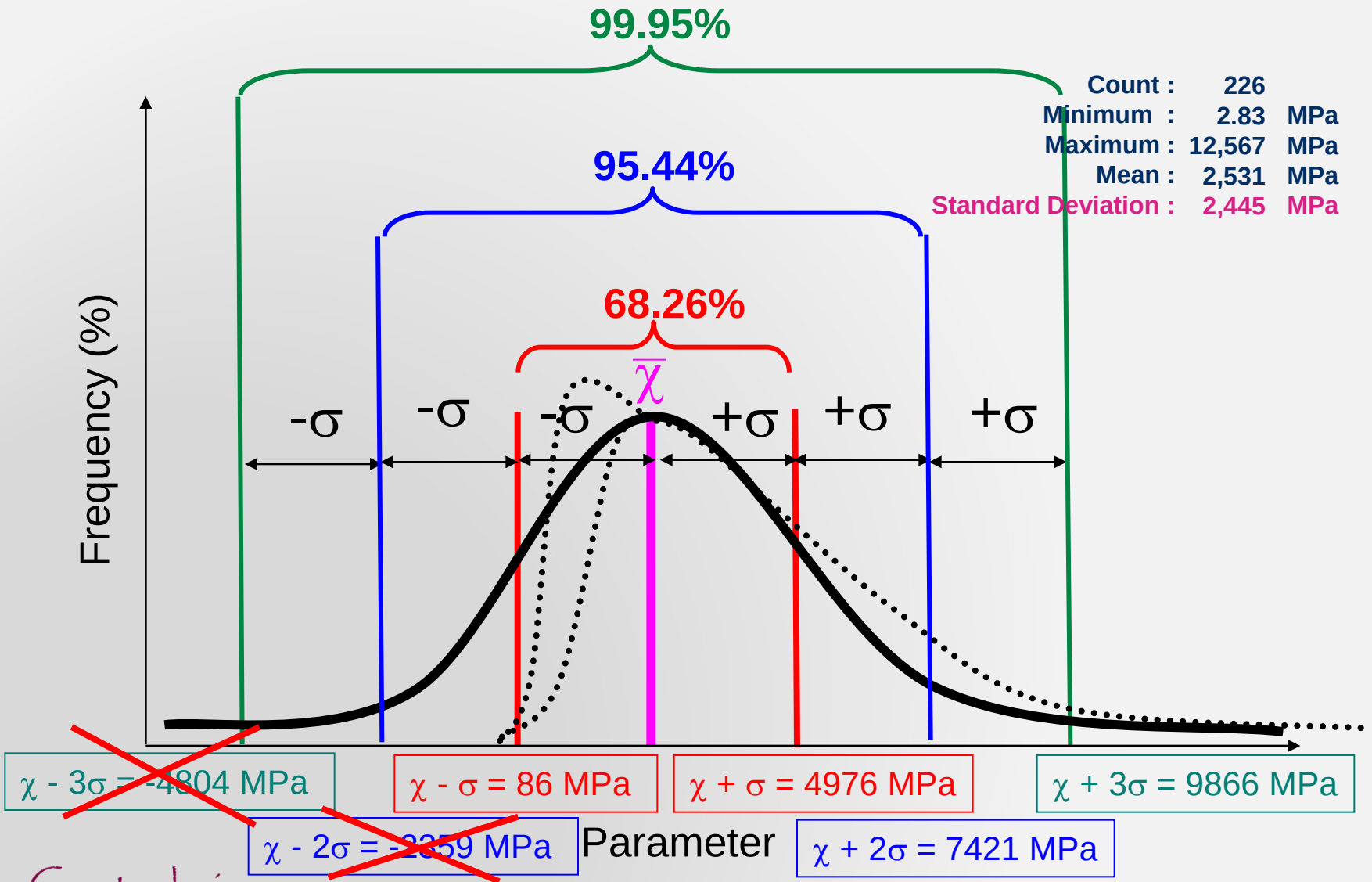
How does it look like?



Geotechnica ME equip

Halcrow
A CH2M HILL COMPANY

The distribution of Data

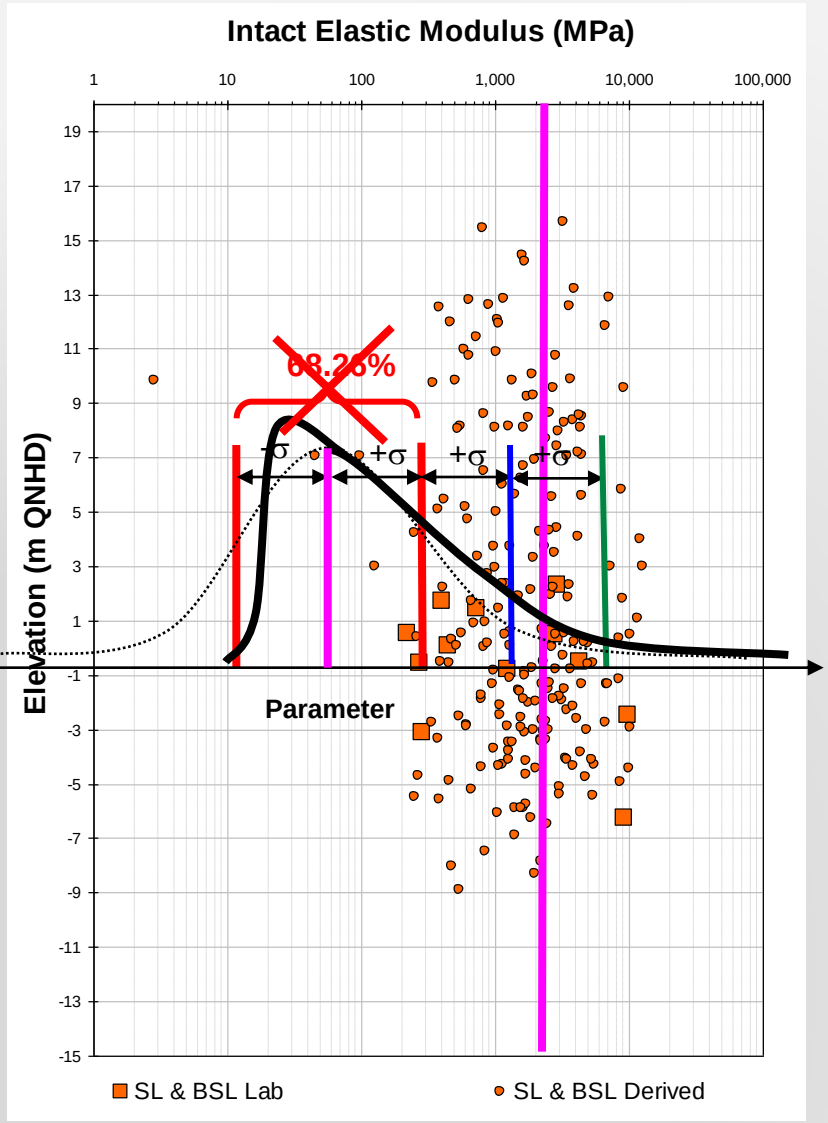


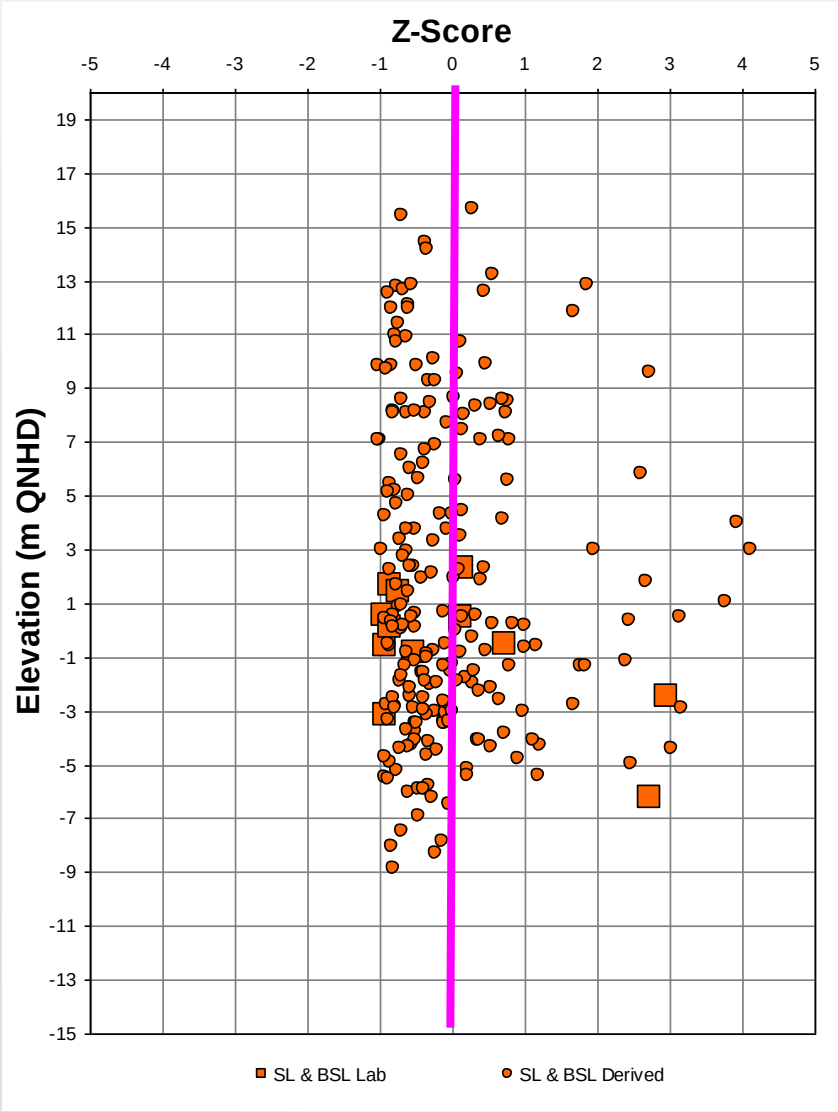
Geotechnica ME equip

Halcrow
A CH2M HILL COMPANY

Z-Score

$Z = (x - \bar{\chi}) / \sigma$

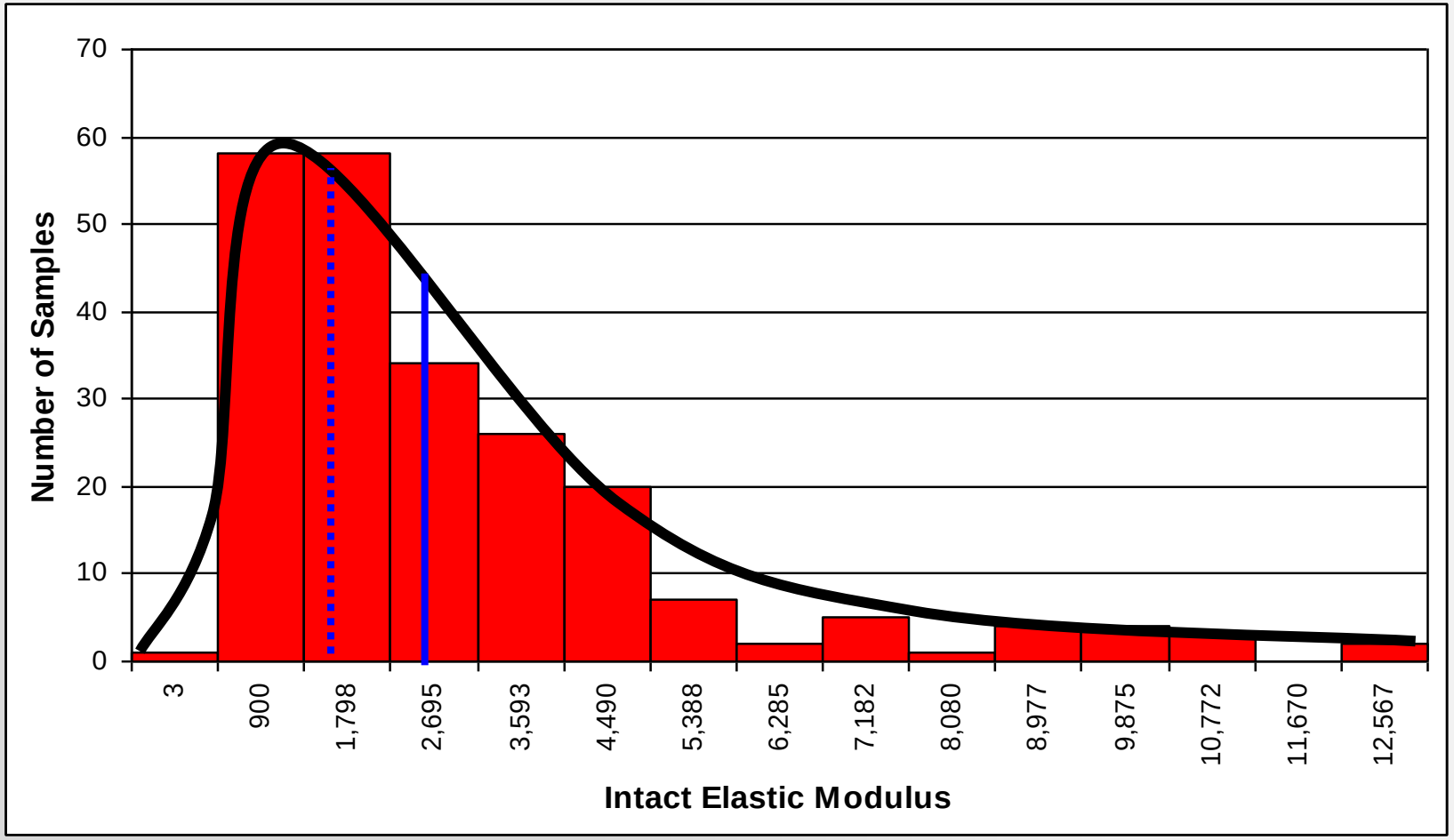




Geotechnica ME

Halcrow
A CH2M HILL COMPANY

Log-normal Distribution



Geotechnica ME equip

Halcrow
A CH2M HILL COMPANY

Ways Forward

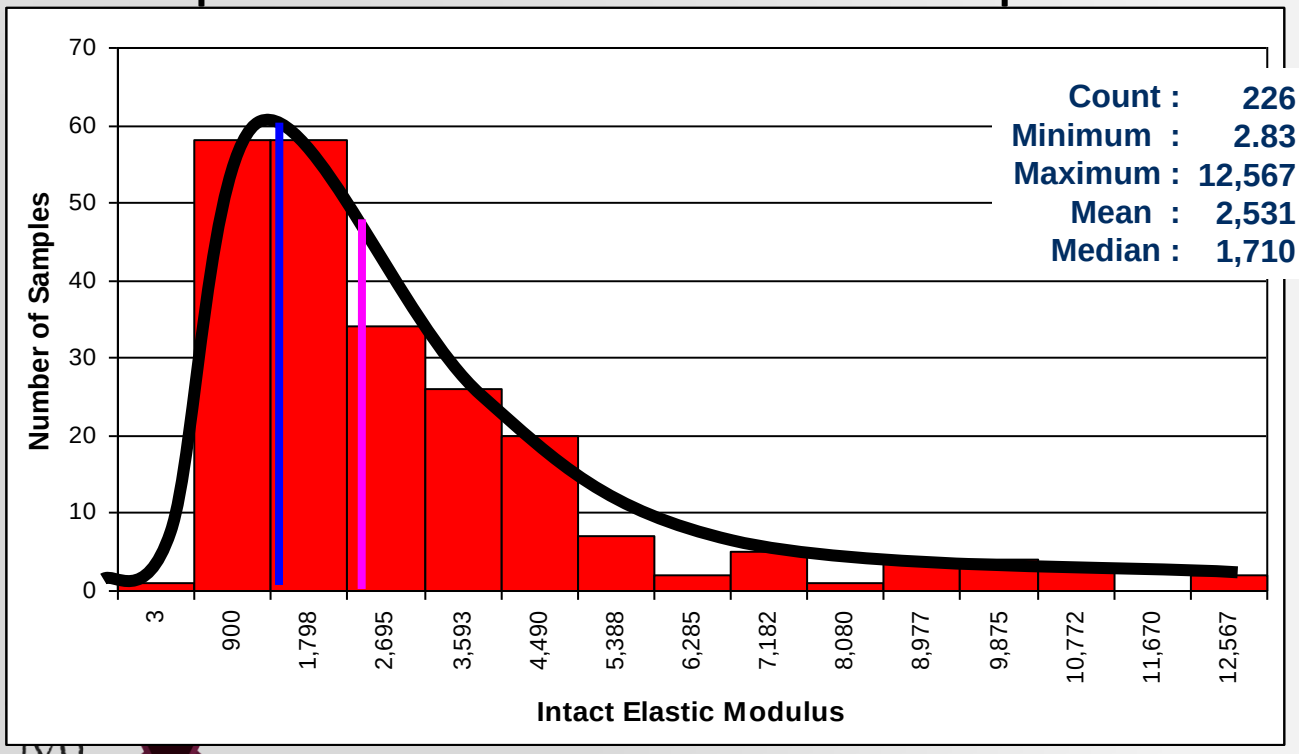
- Ignoring the mathematical tools, introducing further descriptive statistical parameters (median, skewness, kurtosis, coefficient of variation, etc.)

- Applying transformations (rotations, translations, multiplications, etc.) to the existing data in such a way to be represented as a normal distribution

Descriptive Statistics Approach

- Median

  The midpoint of the observed values ***if they are arranged in increasing order*** (ex. No of Samples = 226 – median = sample no 113)



Halcrow
A CH2M HILL COMPANY

Descriptive Statistics Approach

- Median
- Deciles, Percentile, Quartiles
- Interquartile Range (IQR)

Splitting the data into tenth (deciles), hundredths (percentile) or any other fraction (quartiles)

IQR = difference between $q_{0.75}$ and $q_{0.25}$

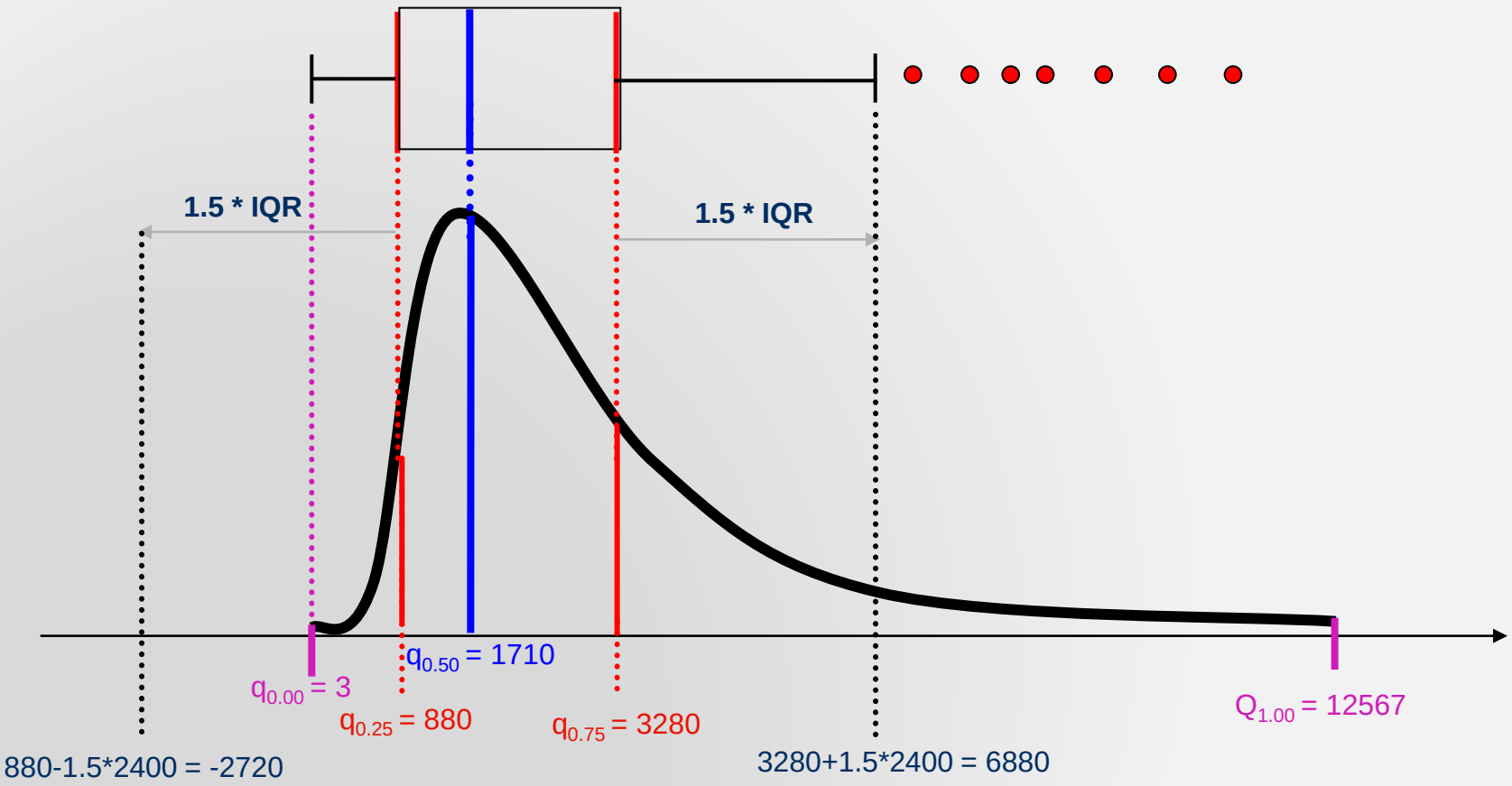
Attention: *Excel works a bit different – percentile for any fraction, quartile for min, 25%, median, 75%, max*

Geotechnica ME equip

Halcrow
A CH2M HILL COMPANY

Box Plot

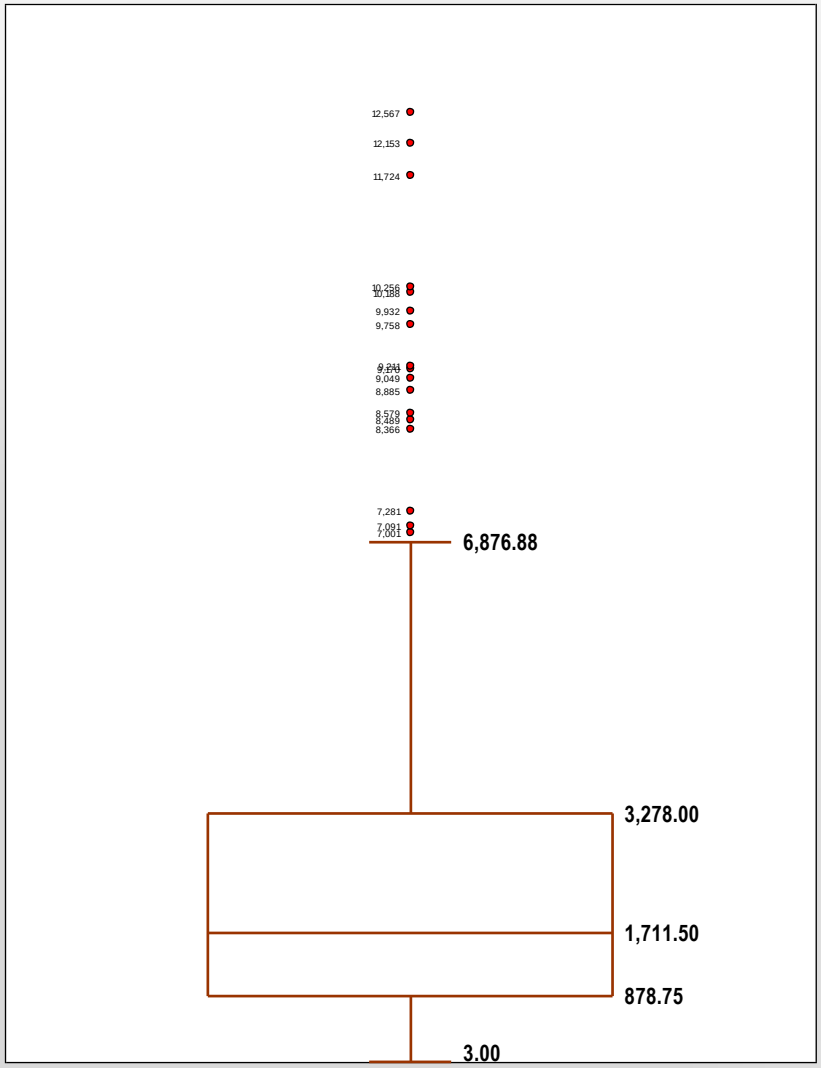
$IQR = q_{0.75} - q_{0.25} = 2400$

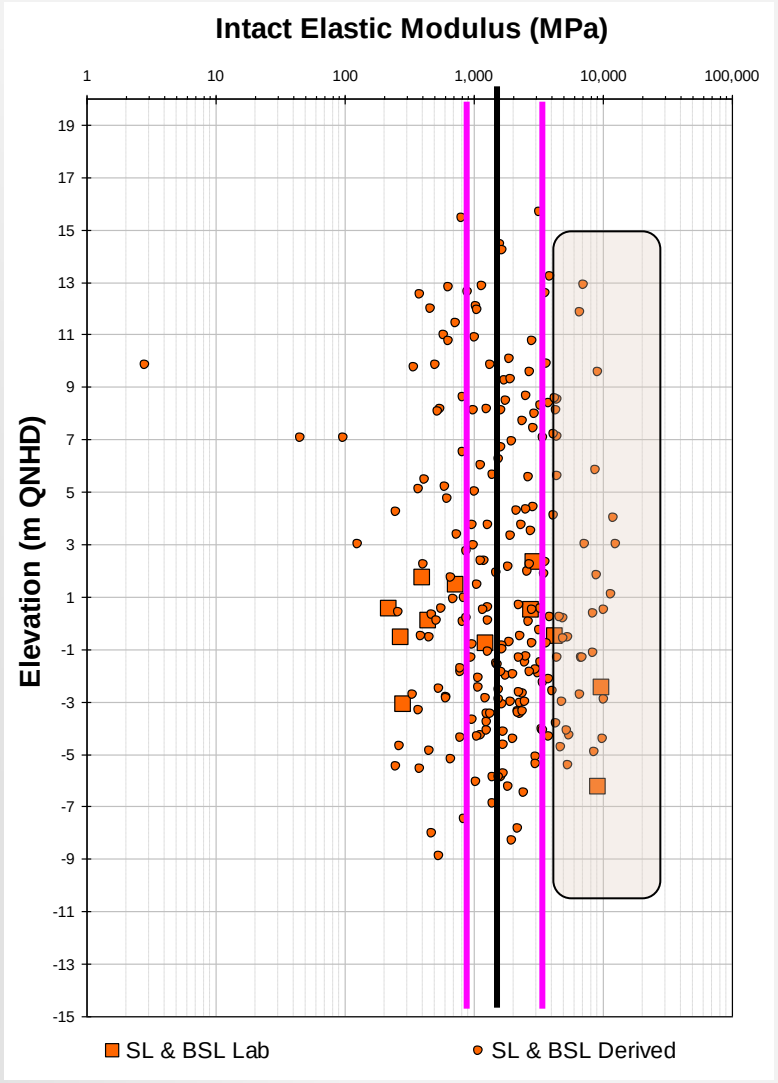


Geotechnica ME equip6

Halcrow
A CH2M HILL COMPANY

Distribution of Data

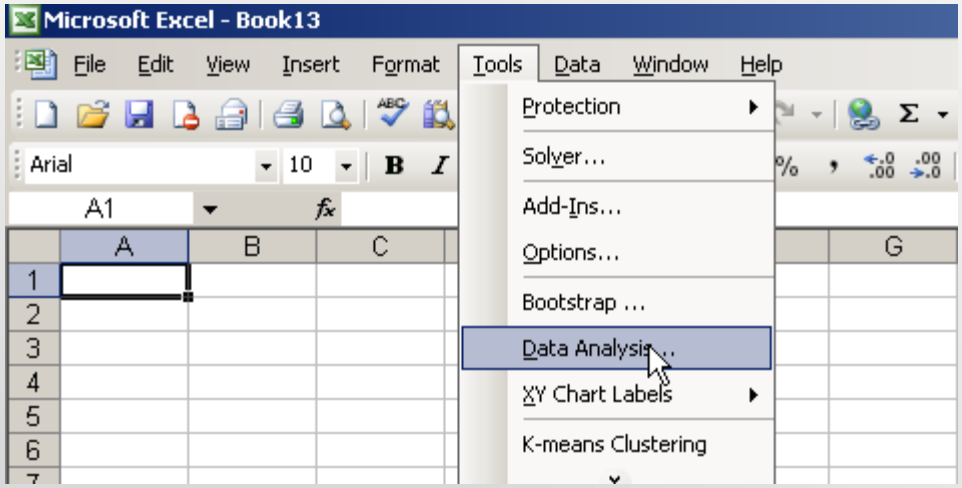


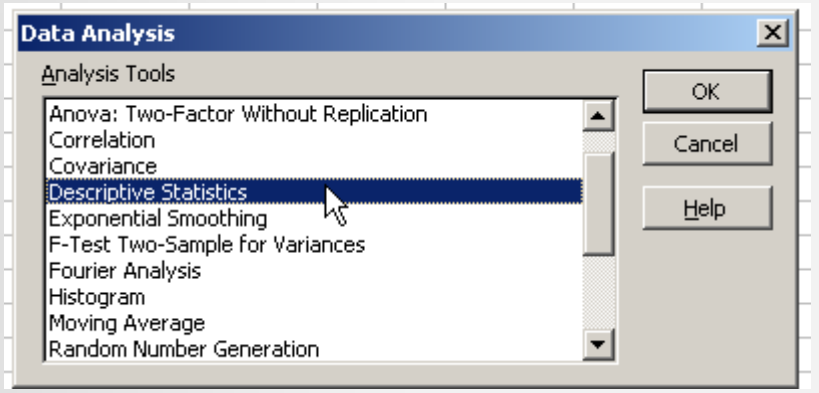


Geotechnica ME

Halcrow
A CH2M HILL COMPANY

Descriptive Statistics in Excel



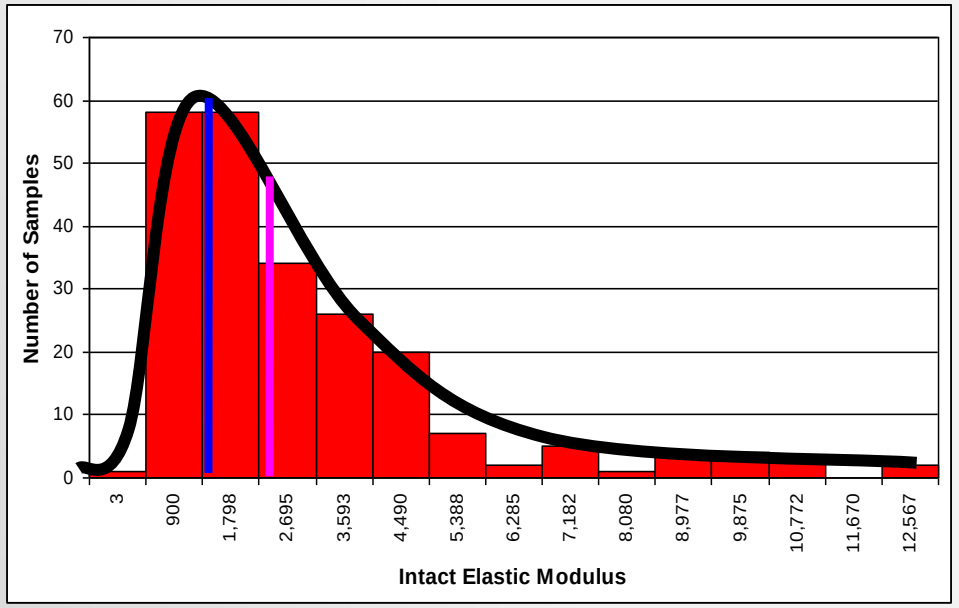


Geotechnica ME equip6

Halcrow
A CH2M HILL COMPANY

Descriptive Statistics Approach

	Column1
Mean	2531
Standard Error	163
Median	1712
Mode	2221
Standard Deviation	2445
Sample Variance	5975961
Kurtosis	4
Skewness	2
Range	12564
Minimum	3
Maximum	12567
Sum	572008
Count	226
Largest(1)	12567
Smallest(1)	3
Confidence Level(95.0%)	320



Geotechnica ME equip6

Halcrow
A CH2M HILL COMPANY

Transformations to a Normal Distribution

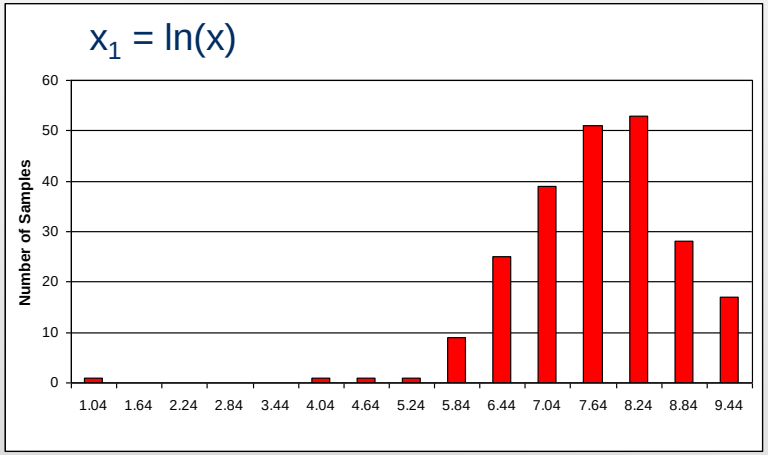
- Applying any mathematical calculation over the existing dataset (square, log, exponential, etc.)
- Applying rotation and translation matrices

Geotechnica ME equip

Halcrow
A CH2M HILL COMPANY

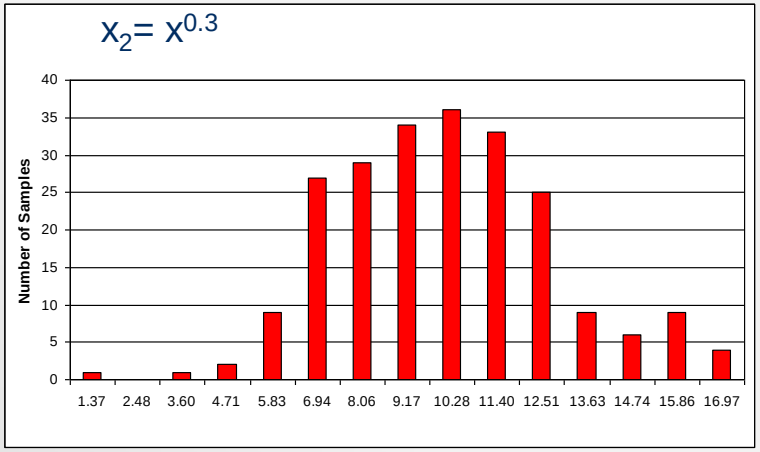
Transformations

$x_1 = \ln(x)$



Descriptive Statistics	
Count	226.000
Minimum	1.041
Maximum	9.439
Range	8.398
Mean	7.392
Median	7.445
Mode	7.706
Standard Deviation	1.060
Sample Variance	1.123
Kurtosis	5.192
Skewness	-1.184
Confidence Level (95%)	0.004

$x_2 = x^{0.3}$



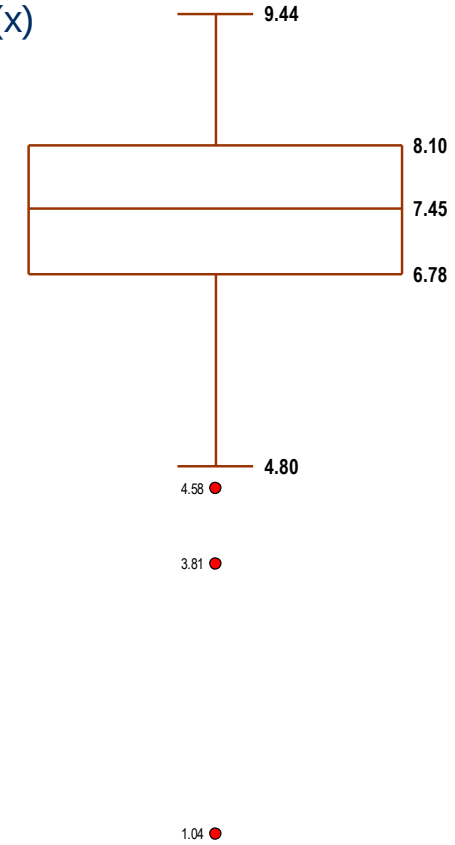
Descriptive Statistics	
Count	226.000
Minimum	1.366
Maximum	16.973
Range	15.607
Mean	9.613
Median	9.333
Mode	10.092
Standard Deviation	2.777
Sample Variance	7.709
Kurtosis	0.098
Skewness	0.336
Confidence Level (95%)	0.012

Geotechnica ME equip

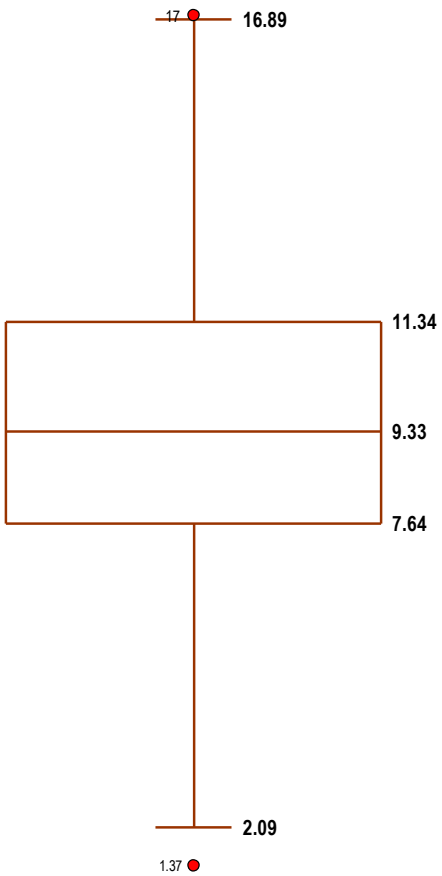
Halcrow
A CH2M HILL COMPANY

Box Plots

$x_1 = \ln(x)$



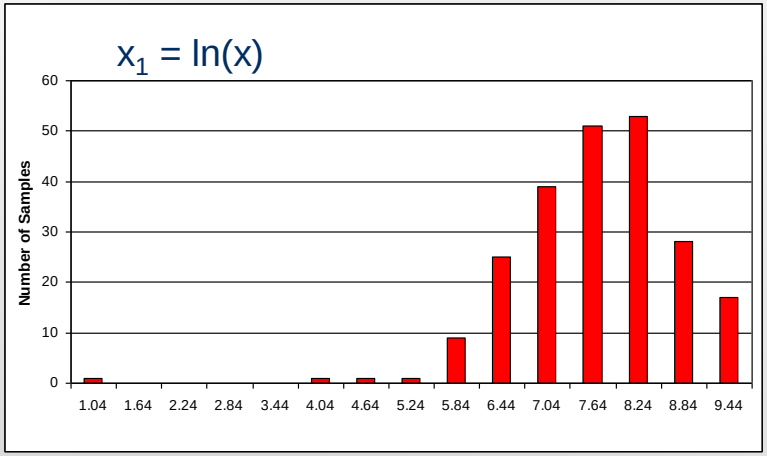
$x_2 = x^{0.3}$

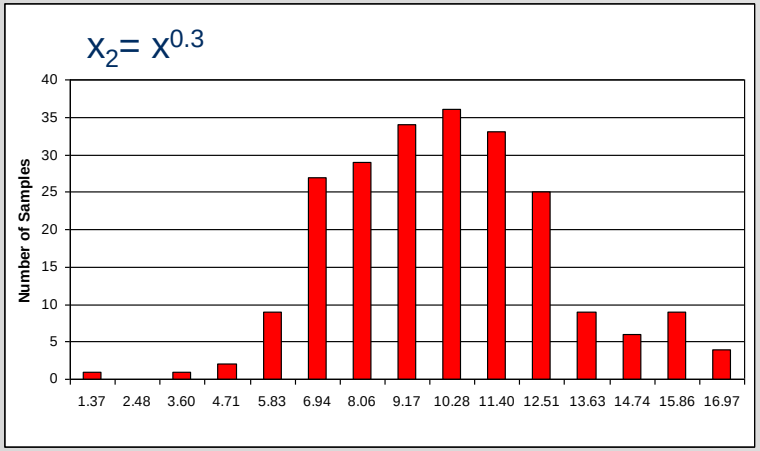


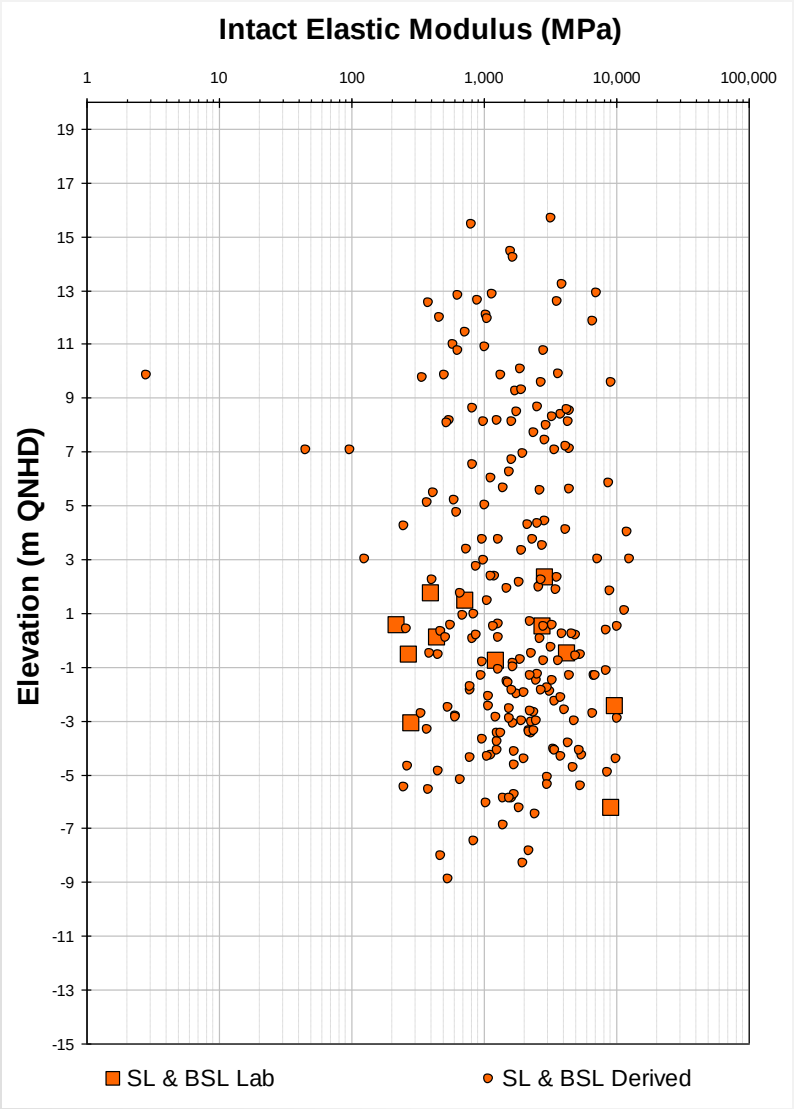
Geotechnica ME equip

Halcrow
A CH2M HILL COMPANY

Histograms



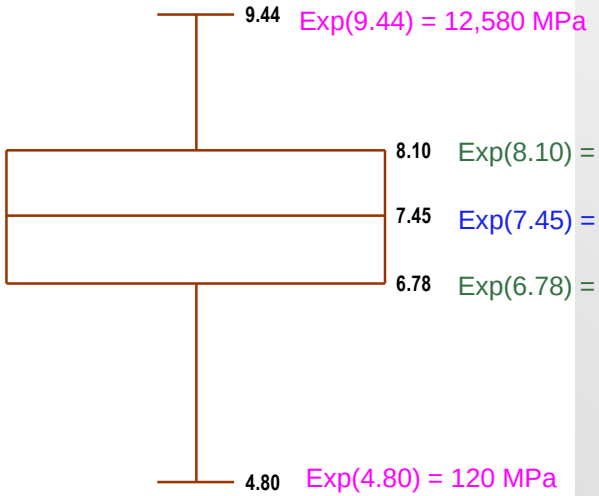




Geotechnica ME equip

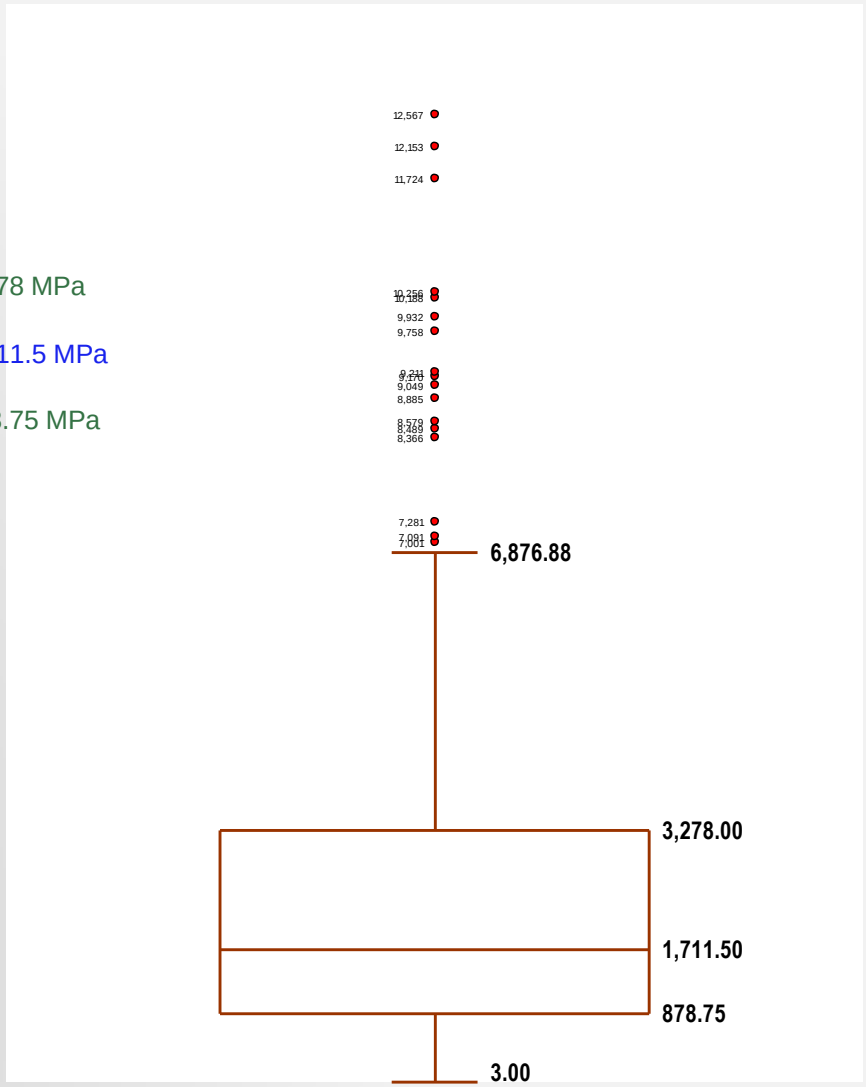
Halcrow
A CH2M HILL COMPANY

Back calculation



Median Ln : 7.45

Back calculated median : Exp(7.45) = 1,711.5 MPa



Conclusion

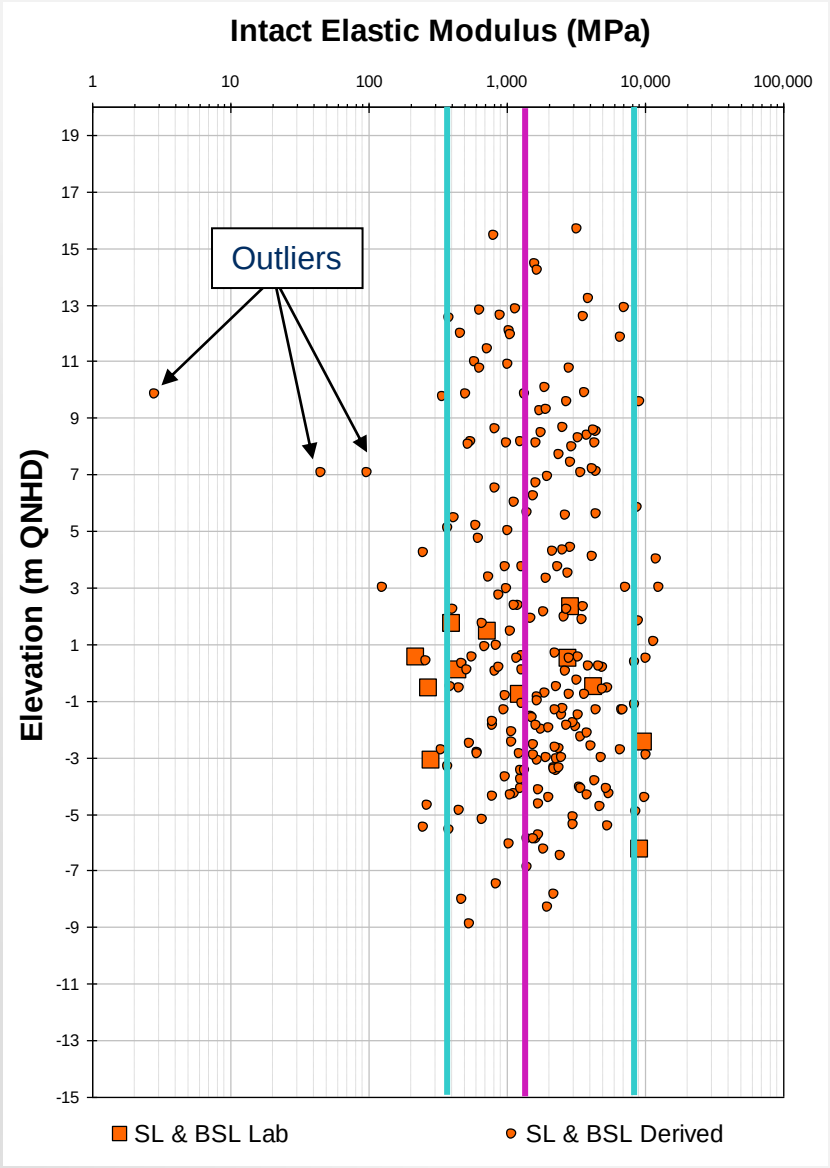
Simsima Limestone

Original assessment:

Minimum	=	2.83 MPa
Maximum	=	12,567 MPa
Characteristic value	=	2,531 MPa
5% Fractile	=	240 MPa
95% Fractile	=	8,570 MPa

Statistical assessment:

Minimum	=	120 MPa
Maximum	=	12,567 MPa
Characteristic value	=	1,600 – 1700 MPa
5% Fractile	=	380 MPa
95% Fractile	=	8,570 MPa



Geotechnica ME equip6

Halcrow
A CH2M HILL COMPANY

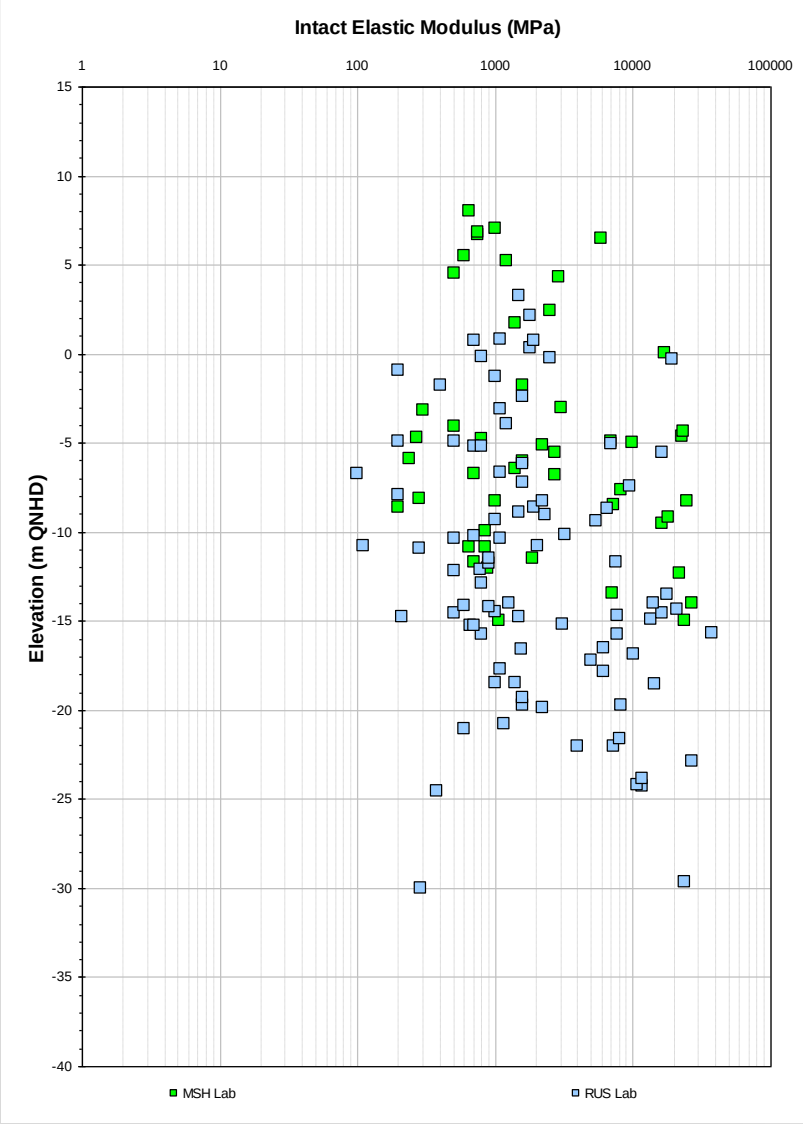
Conclusion

- Earth data are generally characterized by log normal distributions
- The basic statistical tools available are used to quantify our engineering judgment and to provide confidence in our assessment
- Applying blind mathematical tools without the understanding of our dataset can result in misleading answers.

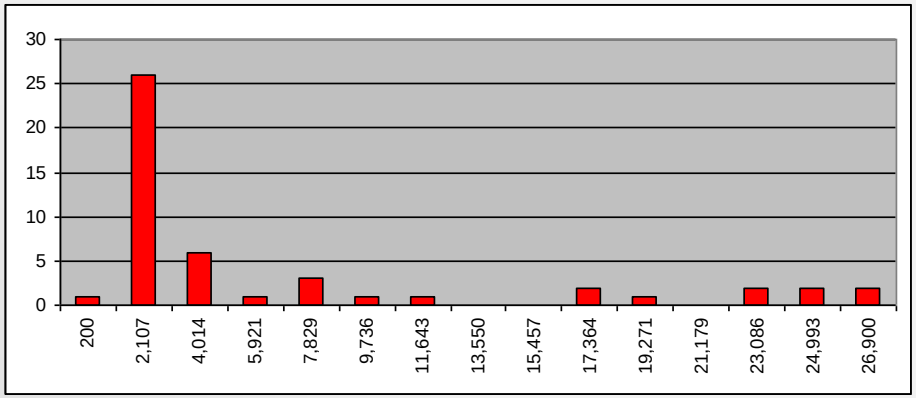
Geotechnica ME equip6

Halcrow
A CH2M HILL COMPANY

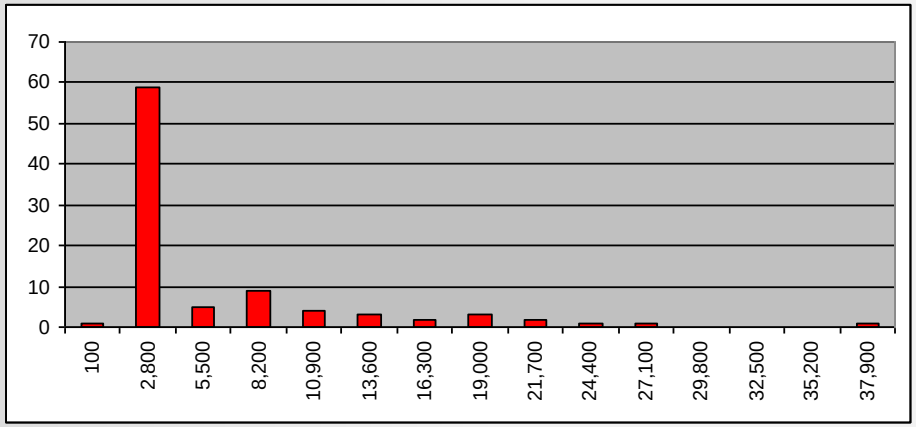
Midra Shale and Rus Formation



Midra Shale



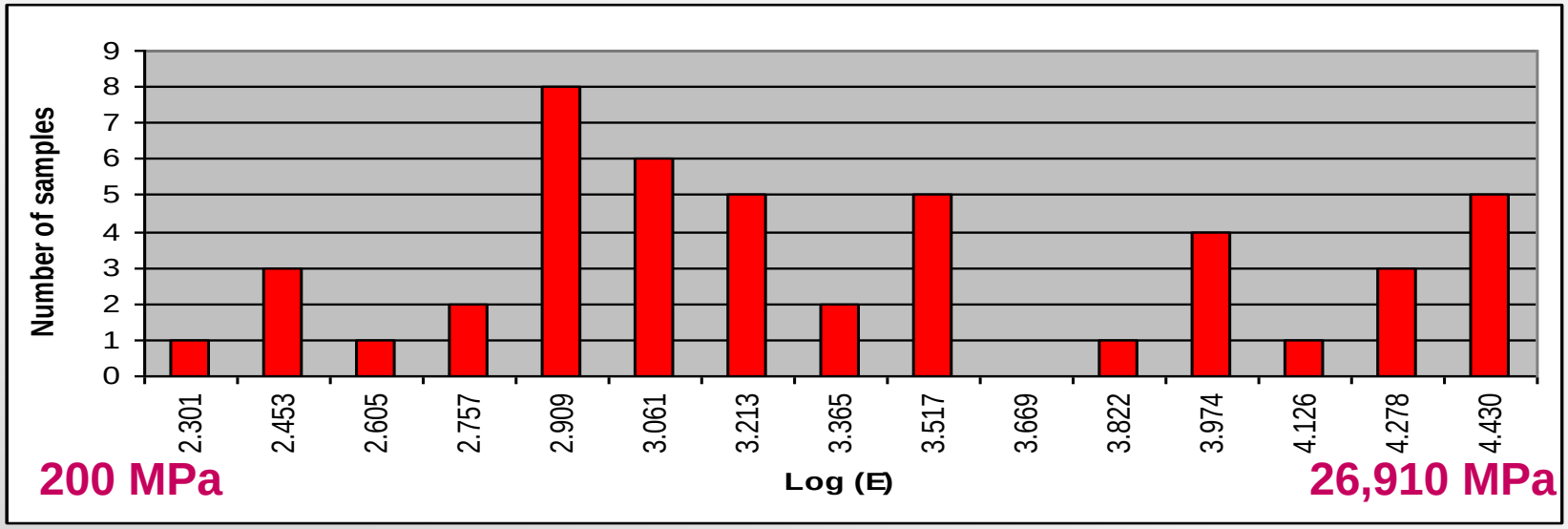
Rus Formation



Geotechnica ME equip

Halcrow
A CH2M HILL COMPANY

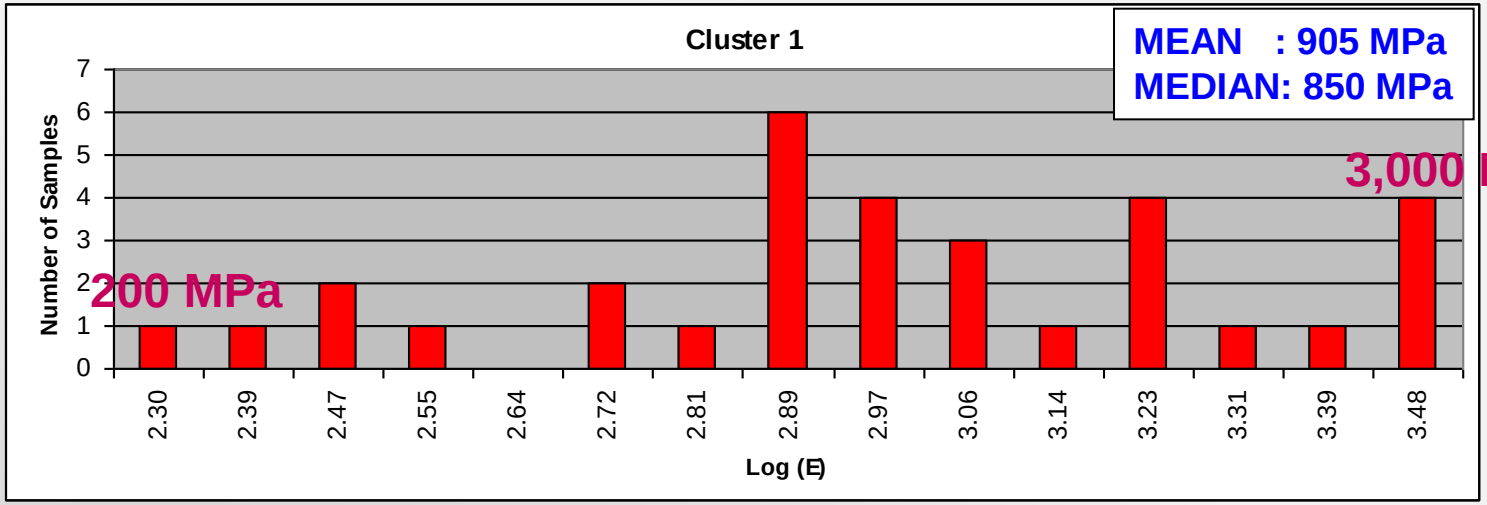
Midra Shale

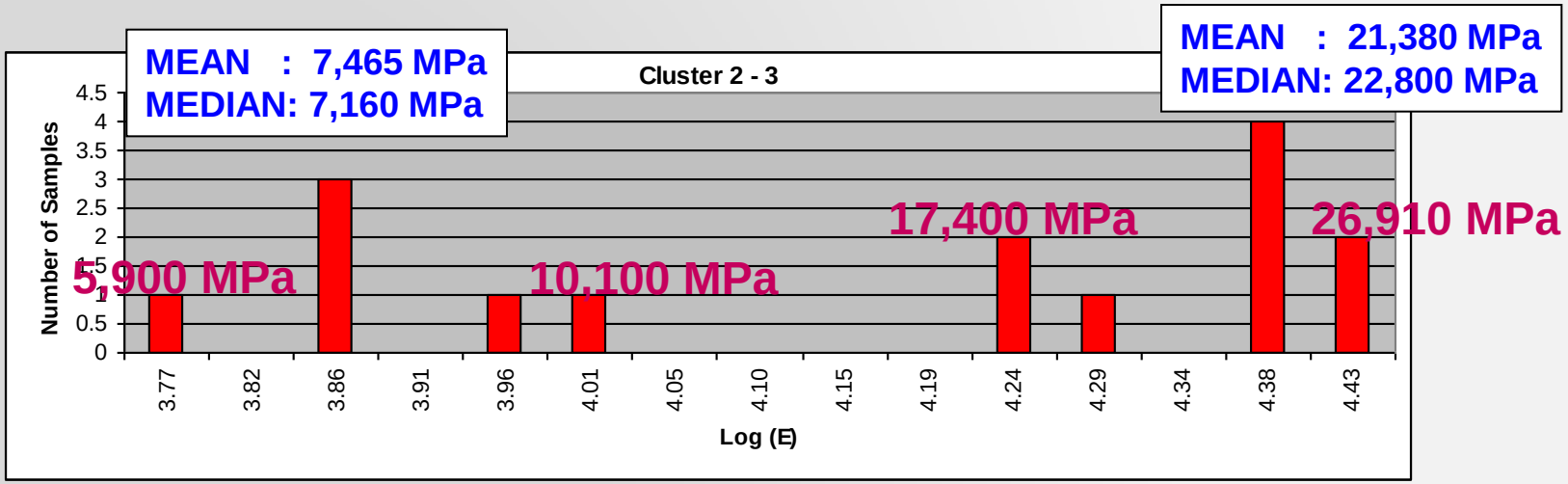


Geotechnica ME

Halcrow
A CH2M HILL COMPANY

Midra Shale

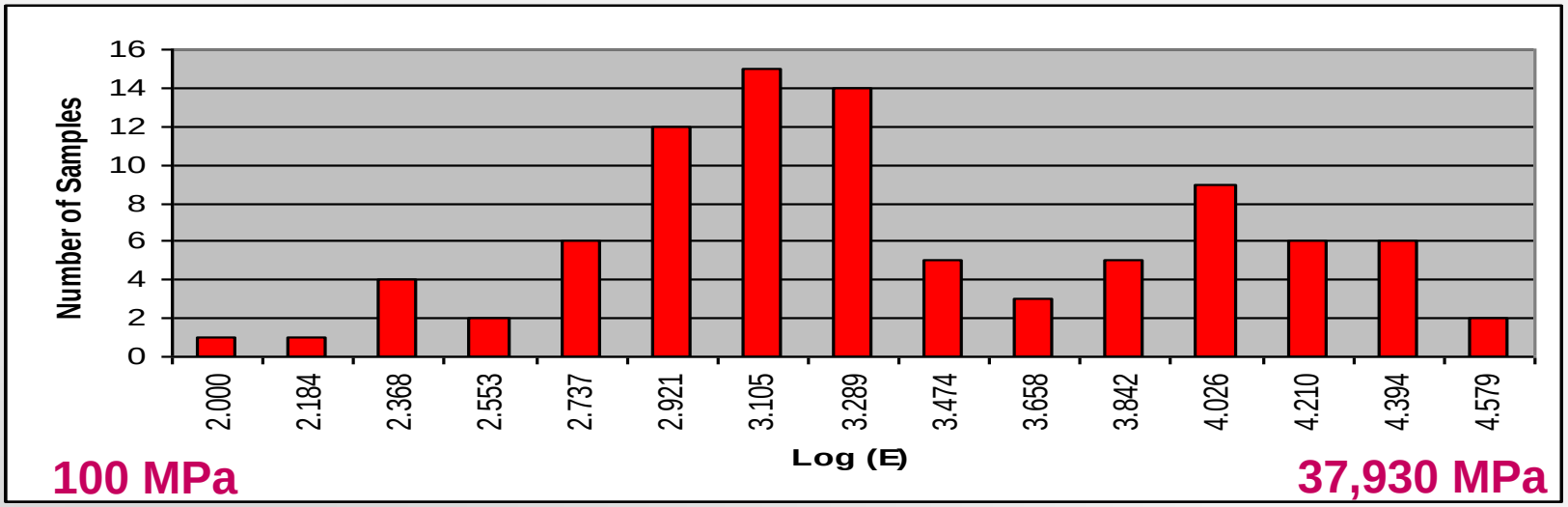




Geotechnica ME equip6

Halcrow
A CH2M HILL COMPANY

Rus Formation



Geotechnica ME equip6

Halcrow
A CH2M HILL COMPANY

Conclusion

Midra Shale

Original assessment:

Minimum	=	200 MPa
Maximum	=	26,900 MPa
Characteristic value	=	5.813 MPa
5% Fractile	=	275 MPa
95% Fractile	=	23,555 MPa
Confidence Level (95)	=	2,370 MPa

Statistical assessment:

Lithology 1 (Possible Shale)

Minimum	=	200 MPa
Maximum	=	3,000 MPa
Characteristic value	=	850 - 1170 MPa
5% Fractile	=	260 MPa
95% Fractile	=	2,810 MPa
Confidence Level (95)	=	295 MPa

Lithology 2 (Possible Dolomitic Limestone)

Minimum	=	16,400 MPa
Maximum	=	26,900 MPa
Characteristic value	=	21,655 – 22,800 MPa
5% Fractile	=	16,640 MPa
95% Fractile	=	26,140 MPa
Confidence Level (95)	=	2,833 MPa

Rus Formation

Original assessment:

Minimum	=	100 MPa
Maximum	=	37,900 MPa
Characteristic value	=	4,750 MPa
5% Fractile	=	205 MPa
95% Fractile	=	18,550 MPa
Confidence Level (95)	=	1,430 MPa

Statistical assessment:

Lithology 1 (Possible Chalk/Calcisiltite)

Minimum	=	100 MPa
Maximum	=	2,500 MPa
Characteristic value	=	1,000 – 1,050 MPa
5% Fractile	=	200 MPa
95% Fractile	=	2,200 MPa
Confidence Level (95)	=	160 MPa

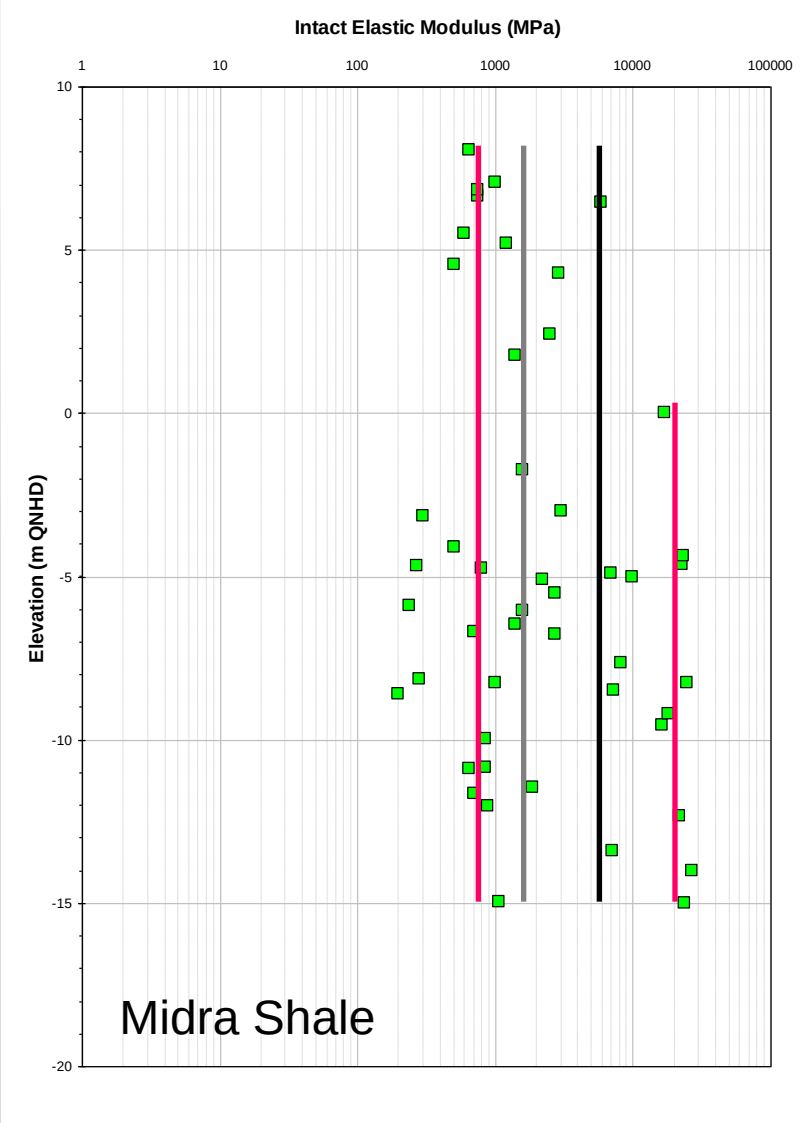
Lithology 2 (Possible Limestone)

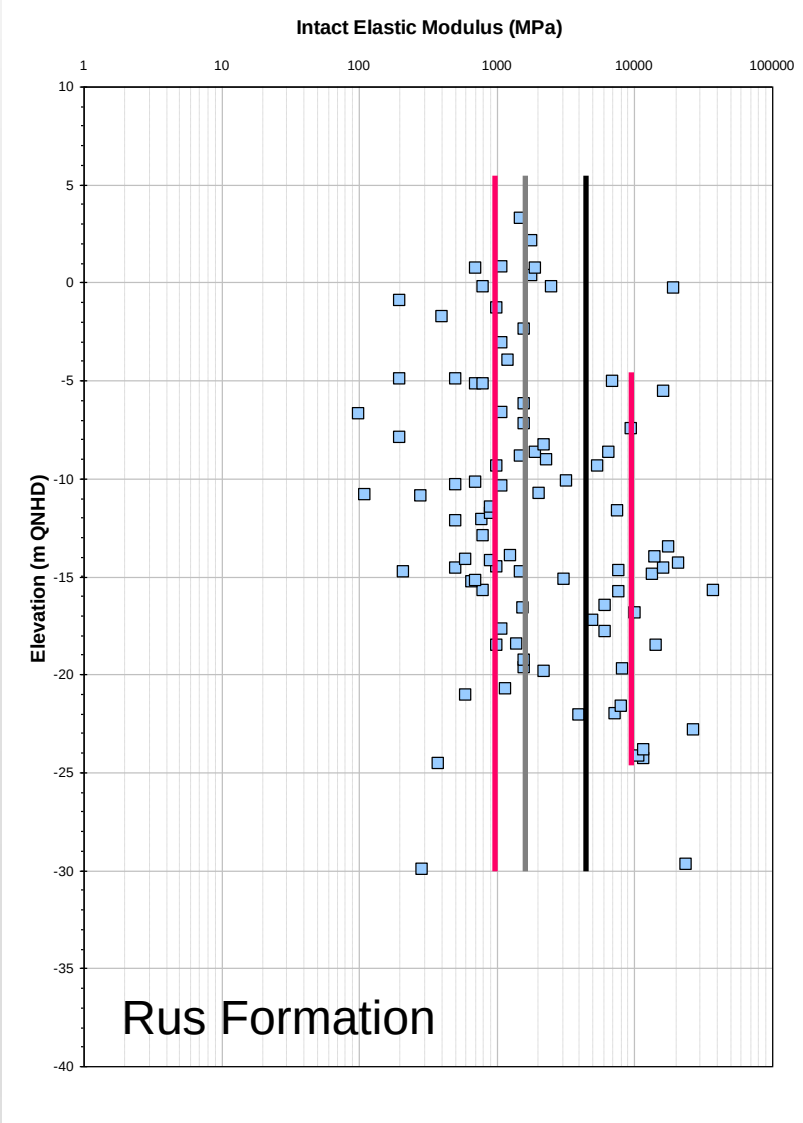
Minimum	=	3,100 MPa
Maximum	=	37,900 MPa
Characteristic value	=	9,600 – 11,900 MPa
5% Fractile	=	3,600 MPa
95% Fractile	=	25,450 MPa
Confidence Level (95)	=	2,848 MPa

Geotechnica ME equip6

Halcrow
A CH2M HILL COMPANY

Conclusion





Geotechnica ME equip6

Halcrow
A CH2M HILL COMPANY