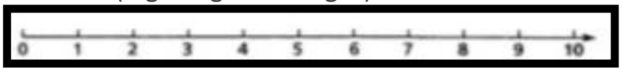
Geologic data Analysis

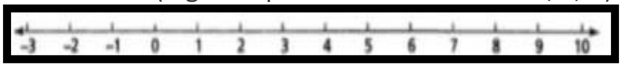
Shaas Hamdan

DATA COLLECTION

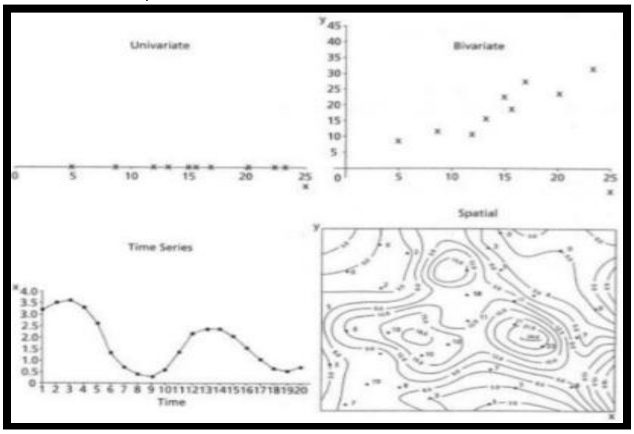
- **Randomness in data analytical results may indicate**:
  1. Natural randomness in geological setting
  2. Failure of the measurement procedure
- In bulk of typical data set we should make sure that compare “like with like” , & Numbers must be measured under the same strict definition (Units must be the same, Same data retrieval procedure used throughout)
- **Missing Data**: the easiest option is to eliminate affected data from further considerations, if not acceptable the missing data should be represented by (* or -99.99) that any future analyst will recognize that it isn’t a real value
- **Discrete Data:** have certain specific values & special frequency distributions, integer, most common is data of counts of objects (e.g. number of fossils in 1m²)
- **Continuous scales (all points used):**
  1. **Ratio or Ordinary Scale [0, ∞]**: best quality & most versatile (e.g. length & weight)



  2. **Interval Scale [-∞, +∞]:** 0 not fundamental termination (e.g. Temperature measured in C, F, K)

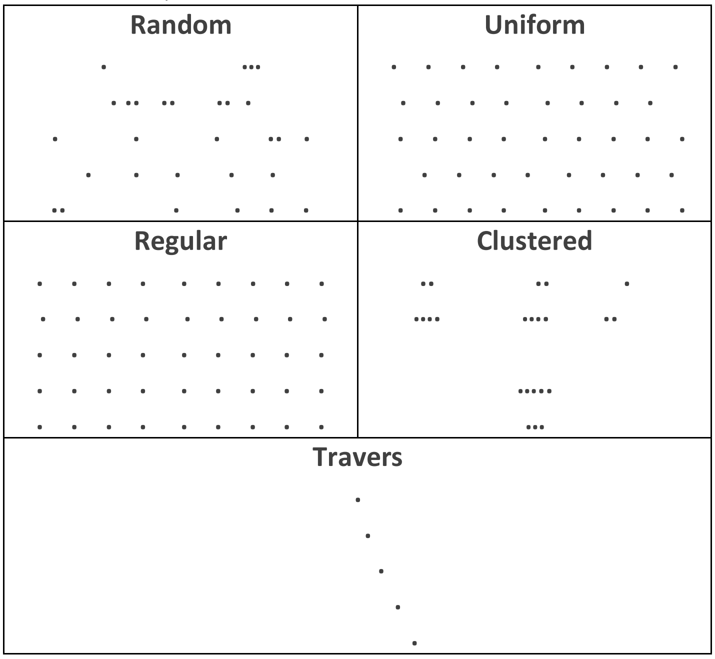


  3. **Closed Scale**: Proportion (i.e. %, ppm)
  4. **Directional Scale**: expressed in Angles (e.g. strike)
  5. **Ordinal Scale**: Lower quality, & Scale is not regular (e.g. Moh’s scale or Richter scale)
- **Statistics & Data Analytical techniques:**
  1. **Univariate**: Each variable analysed in isolation
  2. **Bivariate**: 2 variables, 2D coordinates (scatter)
  3. **Multivariate**: Number of variables analysed is ≥ 3
  4. **Spatial Analysis**: 3-4 variables, 2 to 3 grid references, & the other is measurement of interest



- **The Population:** is the total set of measurements
- **The Sample**: collection of objects (sample or specimen) or measurements & taken to represent the population
- **The entity** is geological object “or sample”
- **Bias:** Samples should be unbiased subset of population
- **Precision (repeated)** measurement is precise if repeated measurements of the same geological entity are similar
  - Precision = measurements
- **Accuracy**: The measurement is accurate if it is close to the true value (In geology true value is unknown)

- **Spatial Sampling scheme**: Random, Uniform, Regular, Clustered, & Traverse



PROBLEMS

- Based on the following data

Mineral	Concentration [in ppm]
Plagioclase	524
K Feldspar	334
Quartz	728
Hornblende	123
Augite	100
Mica	224
Volatiles	*

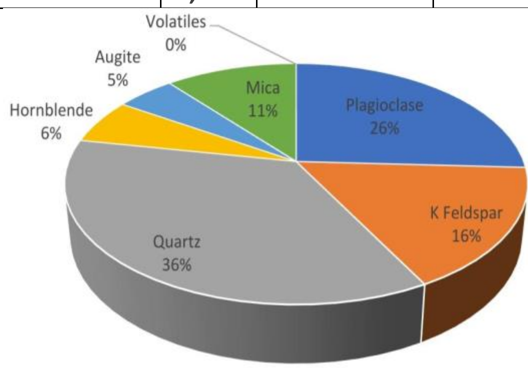
1. What does the star symbol mean? (of volatiles)

   Missing Data
2. What type of data is this? (Hint. geochemical data)

   Is a closed data
3. Are these data Univariate, Bivariate, or Multivariate

   Univariate
4. Draw the data on the Pi-chart

Mineral	ppm	Frequency%	Pi-data
Plagioclase	**524**	**25.8**	**92.9**
K Feldspar	**334**	**16.4**	**59.0**
Quartz	**728**	**35.8**	**129.0**
Hornblende	**123**	**6.1**	**22.0**
Augite	**100**	**4.9**	**17.6**
Mica	**224**	**11.0**	**39.6**
Volatiles	*	**00.0**	**00.0**
Total	**2,033**	**100.0**	**360**



UNIVARIATE DATA

- **Frequency**: is the number of counts in each class
- **Classes:** dividing scale to number of equal intervals
- **Frequency density f(x):** is a function described smooth curve that approximate the outline of the histogram
- **Parameters**: constant indicate properties of population

LOCATION PARAMETERS

- **Mean:** Arithmetic average of the data values

$$\text{Arthmetic Mean } (\bar{X}) = \frac{\Sigma X}{n}$$

$$\text{Weighted Mean } (\bar{X}) = \frac{\Sigma(Xf)}{f}$$

$$\text{Geometric Mean } (\bar{X}) = (X_1 X_2 \ldots)^{\frac{1}{n}}$$

- **Median:** Midpoint of observed values

$$Q = X^{th}_{\left[\frac{n+1}{2}\right]} \text{if n odd} = \left(\frac{X_{\left[\frac{n}{2}\right]} + X_{\left[\frac{n}{2}+1\right]}}{2}\right)^{th} \text{ if n even}$$

- **Mode:** most frequently value

DISPERSION PARAMETER

- **Range:** is the differences between Max. & Min. values

$$R = X_{Max} - X_{Min}$$

- **Interquartile Deviation or Range:** range is overcome by ignoring top 25% & bottom 25%

$$IQR = Q_3 - Q_1$$

$$Q_3 = (X_{3/4 n})^{th}, Q_1 = (X_{n/4})^{th}$$

- **Variance:** the scatter of values about the mean (average square difference of observed values from their mean)

$$\sigma^2 = \frac{\Sigma(X_i - \mu)^2}{n}$$

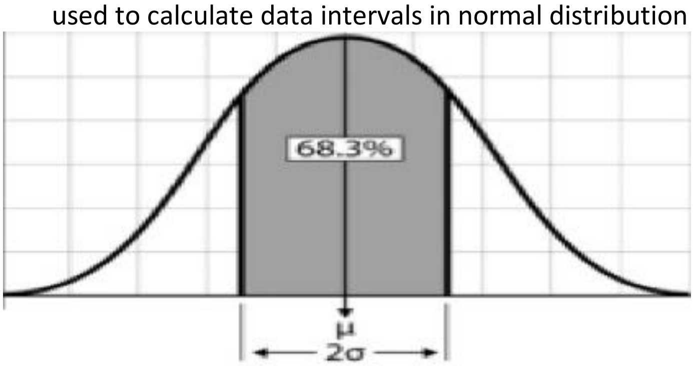
σ^2: variance, μ: mean, Xi: ith value

- **Standard Deviation:** The square root of the variance used to measure spread or dispersion around the mean
  - Is very sensitive to outliers

$$\sigma = root(\sigma^2) = \left(\frac{\Sigma(X_i - \mu)^2}{n}\right)^{½}$$

Interpretation of Standard Division
The higher the σ, the lower the dispersion or spread
The lower the σ, the higher the reliability & sorting

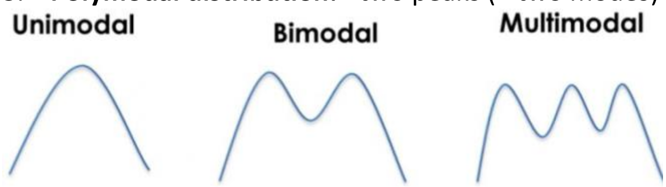
- **Distribution diagram:** standard division (σ) & mean (μ) used to calculate data intervals in normal distribution



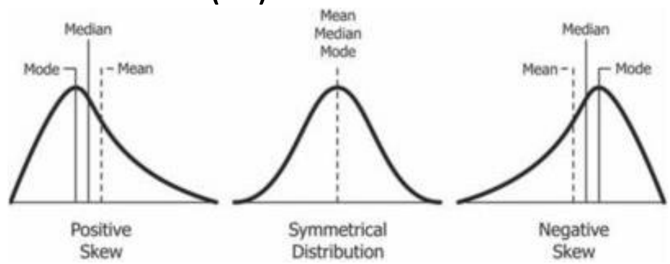
68% of the data lie in the interval [μ-σ, μ+σ]
95% of the data lie in the interval [μ-2σ, μ+2σ]
99% of the data lie in the interval [μ-3σ, μ+3σ]

SHAPE PARAMETER

- **Number of Modes:**
  1. **Unimodal distribution:** one peak (one mode)
  2. **Bimodal distribution:** two peaks (two modes)
  3. **Polymodal distribution:** >two peaks (> two modes)



- **Skewness**: Frequency distribution may be symmetrical or skew, which could be to the right or to the left
  - **Symmetrical population:** Mode = Mean = Median
  - **Asymmetrical population:** skew population
    Right-skewed (+ve): Mean > Median > Mode
    Left-skewed (-ve): Mode > Median > Mean



$$\text{Pearson Sk} = \frac{\mu - \text{mode}}{\sigma} = \frac{3(\mu - Q_{50})}{\sigma}$$

$$\text{Fisher Sk} = \frac{1}{n}\Sigma\frac{(X_i - \bar{X})^3}{\sigma^3} = Avg\frac{(value - mean)^3}{(standard\ diviation)^3}$$

Interpretation of Skewness
If Skewness = 0, the distribution is normal
If Skewness ≠ 0, the distribution is log-normal
+ve Skewness: when the diagram skewed to the right
-ve Skewness: when the diagram skewed to the left

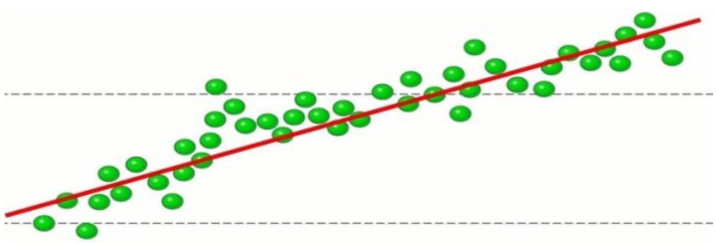
- **Kurtosis**: The "peakedness" of the distribution
  1. **High Kurtosis**: attenuated modal peak frequency
  2. **Low Kurtosis:** describe a plateau- like distribution

Interpretation of kurtosis
If kurtosis >3 → heavier tails than a normal distribution
If kurtosis <3 → lighter tails than a normal distribution

- **Coefficient of Variation (CV):** for +ve skewed data, indicates erratic high values that have impact on estimate

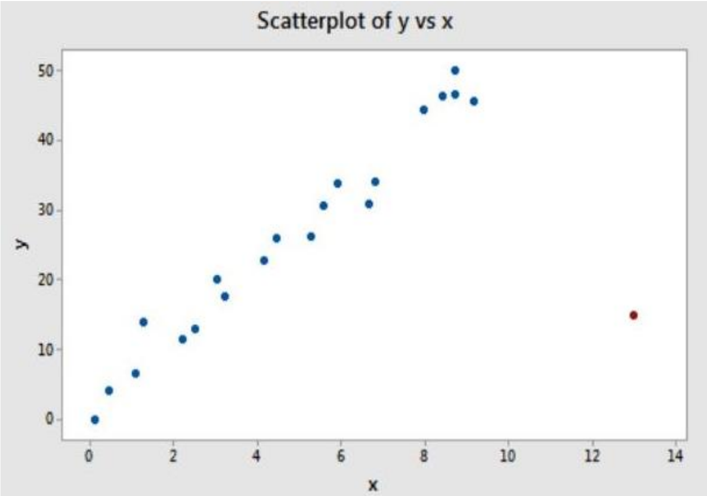
$$CV = \frac{\sigma}{\bar{X}} = \frac{standard\ diviation}{mean}$$

Interpretation of Coefficient of Variation
The higher the coefficient of variation, the greater the level of dispersion around the mean. & The lower the coefficient of variation, the more precise the estimate

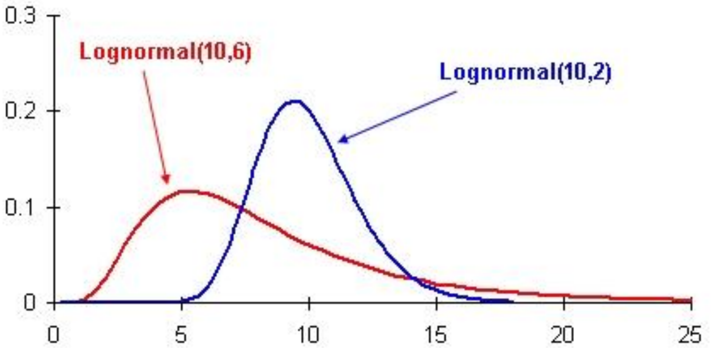


HISTOGRAM SHAPE

- **Histogram**: is rough approximation to the shape of the source population frequency distribution
- **Normal or Gaussian Distribution:** special unimodal, & symmetrical, are the most used probability distribution
- **Outliers:** Data points having high anomalous values



- **Testing for distributions:** chi-squared, Kolmogorov-Smirnov, Normal score, & Normal probability graphs
- **Log Normal probability distribution:** if data aligned in straight line so its log normal distribution



GRAIN SIZE DISTRIBUTION

- **Grain size distribution**: follow normal probability distribution, & quantified in phi (ϕ) scale

$$\boldsymbol{\Phi = -\log_2[\mathrm{X\ in\ mm}] = -\left(\frac{\ln[X]}{\ln 2}\right)}$$

the –ve mean higher +ve values are associated with finer grain size

Step.1: calculate %wt from data & cumulative wt%

$$\mathbf{wt\% = \frac{X}{\Sigma X}}, \quad \mathbf{Cwt\% = \left(X_i + \Sigma^{i-1} X\right)}$$

Step. 2: drow cumulative frequency curve using cumulative wt% (on Y-axis) & grain size in Φ-unit (on X-axis) & determine Φ5, Φ16, Φ25, Φ50, Φ75, Φ84, & Φ95

Step. 3: Calculate the following parameters

$$\mathbf{Median\ (Q) = \Phi_{50}}$$

$$\mathbf{Mean\ (\mu) = \frac{\Phi_{84} + \Phi_{50} + \Phi_{16}}{3}}$$

$$\mathbf{Kurtosis = \frac{\Phi_{95} - \Phi_5}{2.44(\Phi_{75} - \Phi_{25})}}$$

$$\boldsymbol{\beta = \left(\frac{Q^2 - f_1.f_2}{(f_1 + f_2) - 2Q}\right)}$$

$f_1 = 15\%$, $f_2 = 85\%$, Q = 50%

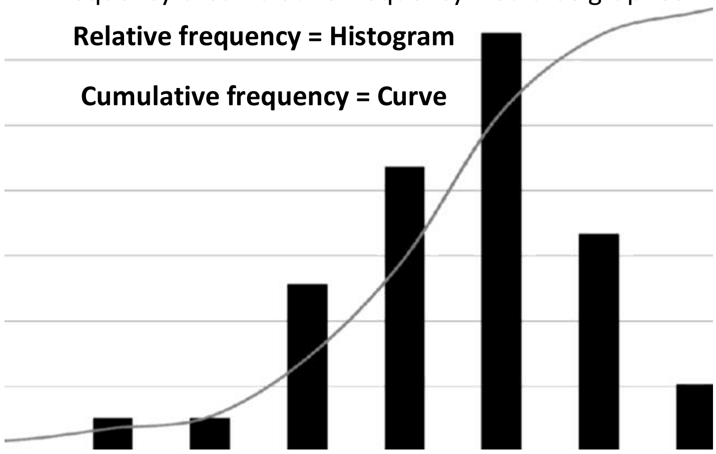
β: deflection or depletion factor

PROBLEMS

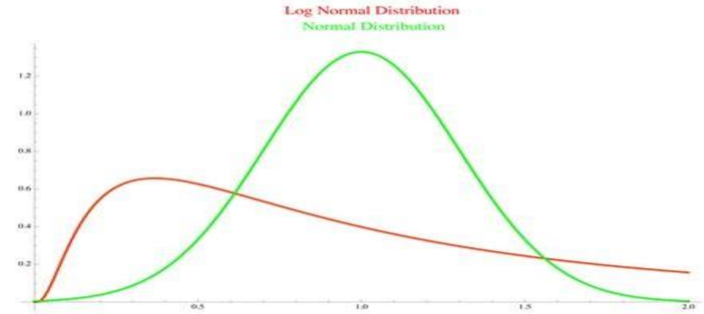
- Explain the difference between relative frequency & cumulative frequency

The relative frequency is the percentage of the data that falls in each class. the cumulative frequency is the sum of the frequencies of that class & all previous classes

- Sketch a generalized diagram showing how relative frequency & cumulative frequency would be graphed



- Sketch a generalized diagram showing a normal distribution & log-normal distribution, in each case where do the most frequently data occurs? & what is the relationship between mean, median, & mode?



In the normal distribution the most frequently data (i.e. the mode) occur at the center of a graph (hear at 1.0), in which mode = mean = median = 1.0

In log-normal distribution (that skewed to the right or to the left) the most frequently data (i.e. the mode) occur at the peak of a graph (0.25), mode ≠ median ≠ mean

- Defined the following terms

Weighted Mean: is the arithmetic mean in which some data points contribute more than others

Mean: is the arithmetic average of the dataset

Mode: is the most frequently value

Median: is the mid point of the dataset

Sample: part of population that represents the population, & taken randomly

Population: is the total set of measurements

Dispersion of data: is the extent to which a distribution is stretched or squeezed

Range: is the differences between Max. & Min. values

Variance: is the scatter of values about the mean (Avg. square difference of observed values from their mean)

Standard Division: is the square root of the variance used to measure spread or dispersion around the mean & very sensitive to outliers

WORKED EXAMPLE: LAB 1

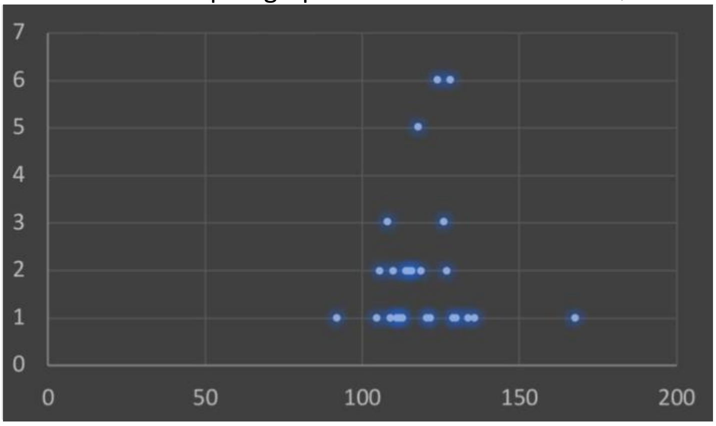
Below are recorded temperature (in °F) for 50 stations

168	122	108	130	126	110	108	116	126	128
112	118	128	124	127	124	109	124	134	126
118	124	119	118	128	113	115	106	121	118
124	118	110	114	116	129	108	136	114	124
111	115	119	105	106	92	128	127	128	128

1. Construct a frequency distribution for the data

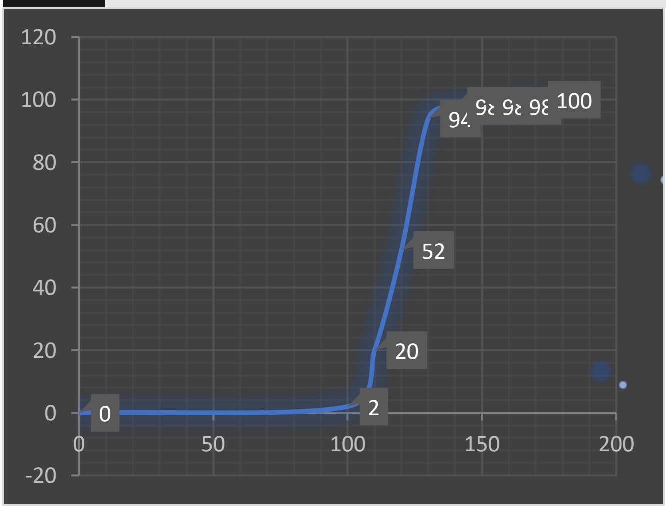
X	92	105	106	108	109	110	111	112	113
F	1	1	2	3	1	2	1	1	1
F%	2	2	4	6	2	4	2	2	2
X	114	115	116	118	119	121	122	124	126
F	2	2	2	5	2	1	1	6	3
F%	4	4	4	10	4	2	2	12	6
X	127	128	129	130	134	136	168	Sum	
F	2	6	1	1	1	1	1	50	
F%	4	12	2	2	2	2	2	100	

2. Draw scatter plot graph for the data results in Q 1



3. Using the data above draw cumulative frequency curve if the intervals (bins) is 10

bins	Frequency	Frequency %	Cumulative frequency
90-100	1	2	2
100-110	9	18	20
110-120	16	32	52
120-130	21	42	94
130-140	2	4	98
140-150	0	0	98
150-160	0	0	98
160-170	1	2	100



Cumulative Frequency Curve

4. For the frequency curve results in question 3
   A. Which interval contains the most recorded T? & what is the mid point of the interval?

   The interval that contains higher frequency which is 120-130 (21 value = 42% of whole data)

   B. Which temperature was exceeded by 50%?

   Is the T above 50% (median) = 119°F

   C. Which temperature was exceeded by 25%?

   Below 75% of data Q_3 = 127°F

   D. Which temperature was exceeded by 75%?

   Below 25% of data Q_1 = 113°F

   E. What percentage of temperature exceeded 118°F?

   50% of data

   F. What temperature marks 25th percentile?

   113°F

5. Calculate the range (R)

   R = 168 – 92 = 76

6. Calculate the Mean (µ)

   Mean = 6000/50 = 120

7. Calculate the Median (Q)

   Q = 119

8. Calculate the Mode (M)

   M = 128 & M = 124 (Bimodal)

9. Calculate the Variance (σ^2)

   $\Sigma(X-\mu)^2 = 6125 \rightarrow \sigma^2 = 123.04$

10. Calculate the Standard Division (σ)

    σ = 11.09

11. What are the variance & standard division values tell us?

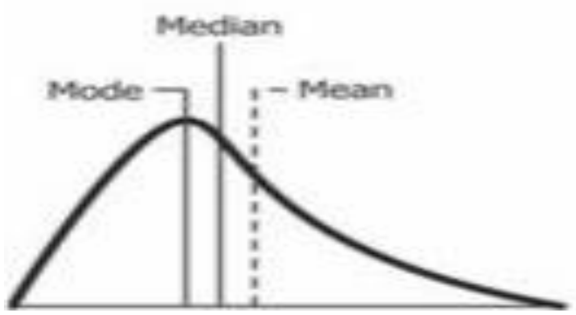
    The standard division (& so variance) is very high & the higher the σ, the lower the dispersion or spread, the lower the reliability & the sorting

12. Calculate the Skewness (Sk)

    $Sk = 3(\mu-Q)/\sigma = 3/11.09 = 0.3$

13. What is the Skewness value tell us?

    +ve skewness, so the distribution is abnormal (log-normal) & skewed to the right in which mean > median



14. Calculate the Interquartile Deviation or Range (IQR)

    $IQR = Q_3 - Q_1 = 127 – 113 = 14$

15. Calculate the Coefficient of Variation (CV)

    $CV = \sigma/\mu = 11.09/120 = 0.092$

16. What is the Coefficient of Variation value tell us?

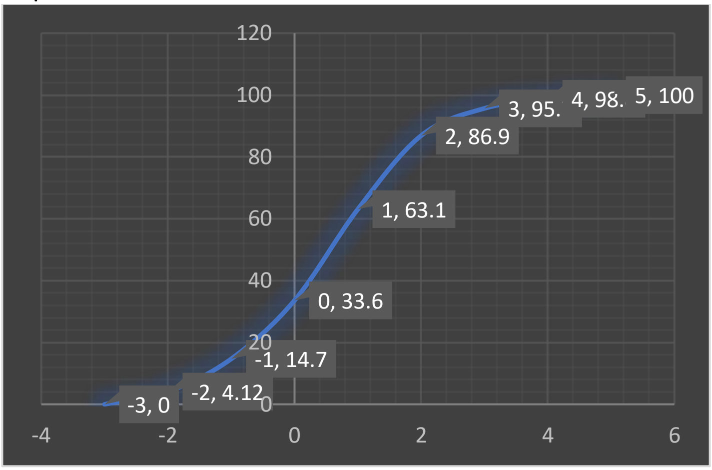
    The Coefficient of variation is low, then there's low level of dispersion around the mean, & the estimation is more precise

17. Is this data normally distributed? How do you know?

    Is not normally distributed data, but is a log-normal distributed data because mean ≠ mode ≠ median & skewed by value of 0.3 to the right (+ve Skewness)

WORKED EXAMPLE: GRAIN SIZE

Use the following cumulative frequency curve that represent the particle size distribution of sediments in a channel to calculate the mean, the median, the sorting, the kurtosis, the skewness, the interquartile range, & the depletion factor



Φ_5 = -1.875
Φ_{16} = -0.938
Φ_{25} = -0.438
Φ_{50} = 0.563
Φ_{75} = 1.500
Φ_{84} = 1.875
Φ_{95} = 2.938

$$\textbf{Median (Q)} = \Phi_{50} = 0.563$$

$$\textbf{Mean} (\mu) = \frac{\Phi_{84} + \Phi_{50} + \Phi_{16}}{3}$$

$$\mu = \frac{1.875 + 0.563 - 0.938}{3} = 0.500$$

$$\textbf{Kurtosis (Kr)} = \frac{\Phi_{95} - \Phi_5}{2.44(\Phi_{75} - \Phi_{25})}$$

$$Kr = \frac{2.938 + 1.875}{2.44(1.500 + 0.438)} = 1.018$$

$$\textbf{Skewness} = \frac{\Phi_{16} + \Phi_{84} - 2Q}{2(\Phi_{84} - \Phi_{16})} + \frac{\Phi_5 + \Phi_{95} - 2Q}{2(\Phi_{95} - \Phi_5)} =$$

$$\frac{-0.938 + 1.875 - 1.226}{2(1.875 + 0.938)} + \frac{-1.875 + 2.938 - 1.875}{2(2.938 + 1.875)} = -0.040$$

$$Sorting\ (St) = \frac{\Phi_{84} - \Phi_{16}}{4} + \frac{\Phi_{95} - \Phi_5}{6.6}$$

$$St = \frac{1.875 + 0.938}{4} + \frac{2.938 + 1.875}{6.6} = 1.430$$

$$Interquartile\ Range\ (IQR) = \Phi_{75} - \Phi_{25}$$

$$IQR = 1.500 + 0.438 = 1.938$$

$$\textbf{Depletion Factor} (\beta) = \left(\frac{Q^2 - \Phi_{15}.\Phi_{85}}{(\Phi_{15} + \Phi_{85}) - 2Q}\right)$$

$$\beta = \left(\frac{0.563^2 + 0.920x1.880}{(1.880 - 0.920) - 2x0.563}\right) = -12.329$$

WORKED EXAMPLE: LAB 2

Use the following data to answer the following questions (USING MICROSOFT EXCEL SOFTWARE)

15	12	24	27	30	1	2	18	18	20
16	17	34	36	29	9	4	18	21	15
16	19	22	24	25	10	7	19	19	22
21	18	27	27	32	4	10	15	17	19
21	8	20	27	29	19	9	16	19	22
15	16	17	23	24	25	7	15	21	20
14	15	13	16	17	18	14	5	28	25
14	15	16	15	16	17	13	12	40	38
16	17	11	29	38	55	11	13	34	39
22	26	4	32	36	20	1	14	31	35

- Calculate the mean, the median, the variance, the standard division, the skewness, & quartile of Q1, & Q3

Excel Function	Values
=MEAN(selected data)	19.57
=MEDIAN(selected data)	18
=VAR.S(selected data)	89.86
=STDEV(selected data)	9.48
=SKEW(selected data)	0.68
=QUARTILE(selected data,1)	14.75
=QUARTILE(selected data,3)	25.00

- Calculate the Interquartile Range

IQR = Q3 - Q1 = 25.00 = 14.75 = 10.25

- Calculate the standard division manually using excel

$\Sigma(X-\mu)^2$ = 8896.51
$\sigma^2 = \Sigma(X-\mu)^2/n$ = 8896.51/100 = 88.965
σ = 9.432

- Calculate the skewness manually & explain the result

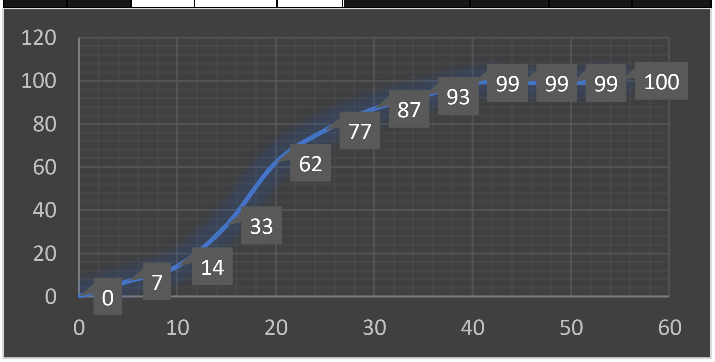
Sk = 3(μ-Q)/σ = 3(19.57-18)/9.432 = 0.499
+ve skewed mean the distribution is log-normal & skewed to the right in which μ>Q

- Calculate the Coefficient of Variation (CV)

CV = σ/μ = 9.432/19.57 = 0.482

- Draw the cumulative frequency curve (assume bias = 5)

PINS		F	F%	Cf	PINS		F	F%	Cf
0	5	7	7	7	30	35	6	6	93
5	10	7	7	14	35	40	6	6	99
10	15	19	19	33	40	45	0	0	99
15	20	29	29	62	45	50	0	0	99
20	25	15	15	77	50	55	1	1	100
25	30	10	10	87	Sum		100	100	-



- Calculate the mode based on the results

Using Excel: =MODE(selected values) = 15
Using Data: (20-15)/2 = 17.5

BIVARIATE DATA

- **Bivariate:** 2 variables relate to the same object
- **Bivariate Scatter Diagram:** simplest graphical concept of univariate data, series of points along a scaled line
- **Correlation Coefficient (ρ):** linear correlation degree, used to summarize the relationship between 2 variables
  - A measure of how close the to falling on a straight line at 45°(independent of magnitude) ranging -1–1
  - As ρ increases the slope of regression line increases & the depression around regression line decrease

$$\rho = \frac{1}{n}\frac{\Sigma_{n=1}^{n}\left((X_i - \mu_x)(Y_i - \mu_y)\right)}{\sigma_x \sigma_y} = \frac{\text{Covariance}}{\sigma_x \sigma_y}$$

Interpretation of Correlation Coefficient
Indicators the linear relationship between two variables,
>0 indicates a direct linear relationship
<0 indicates a inverse linear relationship

- **Corrected Sum of Products CSP**

$$\text{CSP}\ (xy) = \Sigma_{i=1}^{n}\left[(X_i - \mu_x)(Y_i - \mu_y)\right]$$

$$\text{Covariance: COV}\ (xy) = \frac{\text{CPS(xy)}}{n-1} = \mu_{xy} - \mu_x \mu_y$$

Interpretation of Covariance
Indicates the relationship of 2 variables if changes in one variable results in changes in the other variable
>0: as one variable increase the other variable increase

- **Linear Regression:** A strong relationship between 2 variables, help predicting one variable if other is known
  - The simplest type of prediction is Linear Regression

$$Y = aX + b = b_0 + b_1 X$$

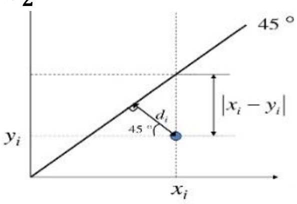
$$a = \rho \frac{\sigma_y}{\sigma_x}\ \&\ b = \mu_y - a\mu_x$$

- **Polynomial regression:** If data scatter shows certain types of more complex curvature (e. g. Parabolic)
  - Independent variables are Powers of X

$$Y = b_0 + b_1 X + b_2 X^2$$

- **The degree of dependence** between X-Y characterized by the spread of the scattergram around the 45° line

$$d_i^2 = \frac{(X_i - Y_i)^2}{2}$$



- **Variogram:** inertia moment of scattergram around 45°
  - moment of inertia is called semi-variogram for pairs
  - The variogram $2\gamma_{xy}$ is the average squared difference between the 2 components of each other

$$\textit{Simivariogram}\ (\gamma_{xy}) = \frac{1}{2n}\Sigma_{i=1}^{n}(X_i - y_i)^2$$

$$\text{Variogram (VAR}, 2\gamma_{xy}) = \frac{1}{n}\Sigma_{i=1}^{n} d_i^2$$

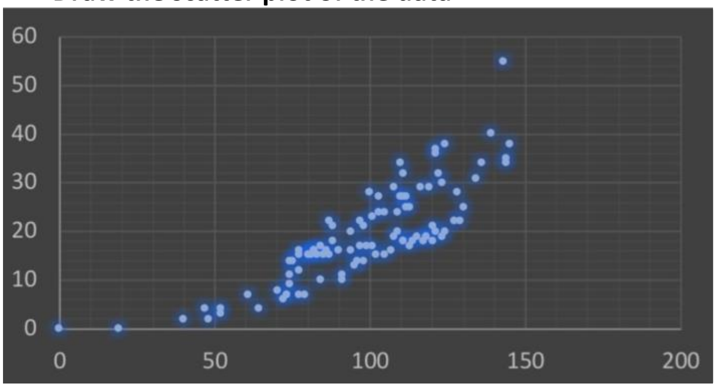
Interpretation of Variogram
The greater the variogram value, the greater the scatter between 2 variables (X-Y), the less the relationship (both direct & Inverse) between them

WORKED EXAMPLE: LAB 3

The following table represents the locations of 100 selected Au & Hg samples (Ag above '+' & Hg below) use these data to answer the following questions (USING EXCEL)

81	77	103	112	123	19	40	111	114	120
+	+	+	+	+	+	+	+	+	+
15	12	24	27	30	0	2	18	18	18
82	81	110	121	119	77	52	111	117	124
+	+	+	+	+	+	+	+	+	+
16	7	34	36	29	7	4	18	18	20
82	74	97	105	112	91	73	115	118	129
+	+	+	+	+	+	+	+	+	+
16	9	22	24	25	10	7	19	19	22
88	70	103	111	122	54	84	105	113	123
+	+	+	+	+	+	+	+	+	+
21	8	27	27	32	4	10	15	17	19
89	88	94	110	116	108	73	107	118	127
+	+	+	+	+	+	+	+	+	+
21	18	20	27	29	19	7	16	19	22
77	82	86	101	109	113	79	102	120	121
+	+	+	+	+	+	+	+	+	+
15	16	16	23	24	25	7	15	21	20
74	80	85	90	97	101	96	72	128	130
+	+	+	+	+	+	+	+	+	+
14	15	15	16	17	18	14	6	28	25
75	80	83	87	94	99	95	48	139	145
+	+	+	+	+	+	+	+	+	+
14	15	15	15	16	17	13	2	40	38
77	84	74	108	121	143	91	52	136	144
+	+	+	+	+	+	+	+	+	+
16	17	11	29	37	55	11	3	34	35
87	100	47	111	124	109	0	98	134	144
+	+	+	+	+	+	+	+	+	+
22	28	4	32	38	20	0	14	31	34

- **Draw the scatter plot of the data**



- Calculate the Covariance

Corrected Sum of products (CSP) = Sum (Z) = 21623.2
Covarince = CSP/(n-1) = 218.417

- Calculate the correlation coefficient

Standard Diviation Y = root of Σ(Y-μy)²/n = 9.81
Standard Diviation X = root Σ(X-μx)²/n = 26.2
Correlation Coefficient (ρ) = 1/n Σ[Z/(σx σy)] = 0.8487

- **Calculate the variogram**

Variogram = 2Simi-variogram = 6521.61

- **Comment on the results**

NON-PARAMETRIC STATISTICS

- If the relationship between 2 variables is not linear, the correlation coefficient may be a very poor
- the rank correlation coefficient calculated by Equation of rank & test are based on median rather than mean
- **Non Parametric Statistics Uses**
  1. The measurement scale of data is ordinal rather than interval or ratio, so the units along the scale are not constant (e.g. difference between 9 & 10 is not the same as the difference between 1 & 2)
  2. The measurements are on interval or ratio scales but investigation of frequency distribution shows a marked departure from the normal distribution
  3. if the sample size is small (<20), Non parametric method is valid regardless of the size & distribution
- **Rank**: is the position of a data value in the ordered sequence of highest to lowest data values.
  - For a data value xi, the rank is symbolised by Rxi (i.e. The lowest value having the rank Rx1 = 1)
  - The change of a scale (as multiplying by 1000) cannot be change the rank
  - If there are 2 or more identical values, ranks are tied, & the rank allocated is average of the rank for the pair of group

Example									
Xi	**56**	**42**	**61**	**42**	**55**	**35**	**42**	**39**	**65**

اولا نرتب القيم من الاقل للاكثر ثم نعطيها ترتيب والقيم المتشابهة نعطينها نفس الترتيب ثم نأخذ المتوسط للقيم المتشابهة، وكل ما يطلب بهذه الاسئلة)مثلا لو كان مطلوب حساب المتوسط (نحسبه عن طريق الترتيب وليس القيم الاصلية

Xi	**56**	**42**	**61**	**42**	**55**	**35**	**42**	**39**	**65**
نرتب القيم من الاقل للاكثر									
Xi	**35**	**39**	**42**	**42**	**42**	**55**	**56**	**61**	**65**
نكتب ترتيب القيم والقيم المتشابهة نعطيها نفس الترتيب									
Rxi	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
نأخذ متوسط القيم المتشابهة ونعطيهم نفس الترتيب									
Rxi	**1**	**2**		**4**		**6**	**7**	**8**	**9**
نحسب المطلوب من السؤال من الرانك وليس القيم الاصلية									

- **Ranked Correlation coefficient or Spearman's rank correlation coefficient:** calculated by the ranks of the data values rather than to the original sample values

$$\rho_{rank} = \frac{1}{n} x \frac{\sum_{i=1}^{n}(Rx_i - \mu_{R_x})(Ry_i - \mu_{R_y})}{\sigma_{R_x}\sigma_{R_y}}$$

The lowest of the z values would appear first on a sorted list & receive a rank of 1; the highest z would appear last on the list & receive a rank of n

Ry: the rank of y; among all the other y values

MRx: is the mean of all of the ranks (Rx1… Rxn)

σRx: standard deviation

- p rank & p aren't the same often quite revealing about the location of extreme pairs on the scatterplot
- rank coefficient is not influenced by extreme pairs
- difference between traditional & rank coefficients arise from the location of extreme pairs on scatterplot
- If it is p that is quite high while prank is quite low, then it is likely that the high value of p is due largely to the influence of a few extreme pairs

WORKED EXAMPLE: QUIZ

Using the following data that represent surface runoff, calculate the rank correlation coefficient

Station (#)	Forest (X)	Surface Runoff (Y)
1	0.28	-8.75
2	-0.33	3.60
3	0.00	-5.44
4	-1.48	210.00
5	-0.50	210.00
6	-2.66	106.06
7	-1.52	150.00
8	-3.04	150.00
9	0.00	0.00
10	-3.46	150.00
11	-0.50	116.85
12	-1.72	210.00
13	-0.33	29.94
14	-1.85	210.00
15	-1.08	136.31

First, We need to arrange the values to find the rank

#	X	Rank (Rx)			#	Y	Rank (Ry)	
		I	F				I	F
10	-3.46	1	1		1	-8.75	1	1
8	-3.04	2	2		3	-5.44	2	2
6	-2.66	3	3		9	0.00	3	3
14	-1.85	4	4		2	3.60	4	4
12	-1.72	5	5		13	29.94	5	5
7	-1.52	6	6		6	106.06	6	6
4	-1.48	7	7		11	116.85	7	7
15	-1.08	8	8		15	136.31	8	8
5	-0.50	9	9.5		7	150.00	9	10
11	-0.50	10	9.5		8	150.00	10	10
2	-0.33	11	11.5		10	150.00	11	10
13	-0.33	12	11.5		4	210.00	12	13.5
9	0.00	13	13.5		5	210.00	13	13.5
3	0.00	14	13.5		12	210.00	14	13.5
1	0.28	15	15		14	210.00	15	13.5

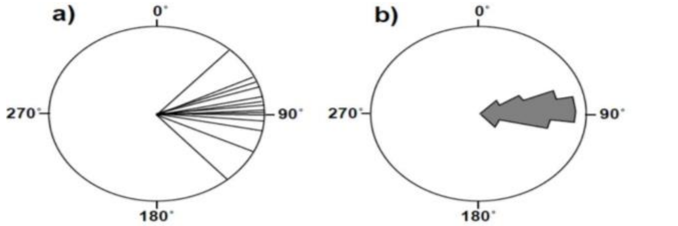
Second, Rearrange the data to its original order, & calculate the ranked mean for both X & Y, the ranked standard division for both X & Y, then ranked correlation coefficient

#	Rx	Ry	(Rx-Rμ)	(Ry-Rμ)	d	(Rx-Rμ)²	(Ry-Rμ) ²
1	15.0	01.0	7.00	-7.00	-49.00	49.00	49.00
2	11.5	04.0	3.50	-4.00	-14.00	12.25	16.00
3	13.5	02.0	5.50	-6.00	-33.00	30.25	36.00
4	07.0	13.5	-1.00	5.50	-05.50	01.00	30.25
5	09.5	13.5	1.50	5.50	08.25	02.25	30.25
6	03.0	06.0	-5.00	-2.00	10.00	25.00	04.00
7	06.0	10.0	-2.00	2.00	-04.00	04.00	04.00
8	02.0	10.0	-6.00	2.00	-12.00	36.00	04.00
9	13.5	03.0	5.50	-5.00	-27.50	30.25	25.00
10	01.0	10.0	-7.00	2.00	-14.00	49.00	04.00
11	09.5	07.0	1.50	-1.00	-01.50	02.25	01.00
12	05.0	13.5	-3.00	5.50	-16.50	09.00	30.25
13	11.5	05.0	3.50	-3.00	-10.50	12.25	09.00
14	04.0	13.5	-4.00	5.50	-22.00	16.00	30.25
15	08.0	08.0	0.00	0.00	00.00	00.00	00.00
Mean	8.0	8.0	0.00	0.00	-12.75	18.57	18.20
Standard Division						04.31	04.27

$$\rho_{rank} = \frac{-12.75}{4.31x4.27} = -0.693$$

DIRECTIONAL DATA ANALYSIS

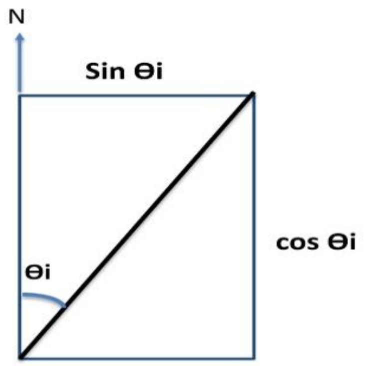
- **Directional data**: take on unique values in 0-360° range & can be displayed in circular diagrams (e.g. Rose)
- Types of graphical presentations of directional data: Rose diagram & Sector diagram



- **Rose diagram** not acceptable as unbiased representation
  - ➢ Visual impression of segment is proportional to the area, & the area is proportional to the r^2, so Rose diagram overemphasis high frequencies & under-emphasises low frequencies. This can lead to false impressions, that can be avoided by
    1. Plot r proportional to square root of frequencies
    2. means of Kite diagrams
- **Sector diagram** biased in way it presents the data, but still the most commonly used type of display for directional data, consider the area of a sector of Δα

$$\mathbf{A} = \frac{\pi r^2 \Delta\alpha}{360°}\ \alpha\ r^2$$

- Why we need spherical methods for circular data? Imagine we have 2 directional measurements 1° & 359° both are due N, but the simple arithmetic mean is (1+359)/2= 180° which is due south, so we need a method which treats 1° & 359°as similar numbers with 0° - 360° as identical numbers, & this problem solved using **trigonometrical function** since cosθ & tanθ repeat every 180°, so sin0 = sin360 & the same for cos & tan
- Dispersion is probably much more important in directional data than in ordinary univariate, it is used to diagnose braided ***alluvial system (Low dispersion)*** & ***meandering systems (high dispersion)***



Xi = sinθ
Yi = cosθ
Xr = Σsinθi
Yr = Σcosθi
The mean direction θ is the direction of the hypotenuse

$$\tan^{-1}\left(\frac{x_r}{y_r}\right) = tan^{-1}\left(\frac{\Sigma sin\theta_i}{\Sigma cos\theta_i}\right)$$

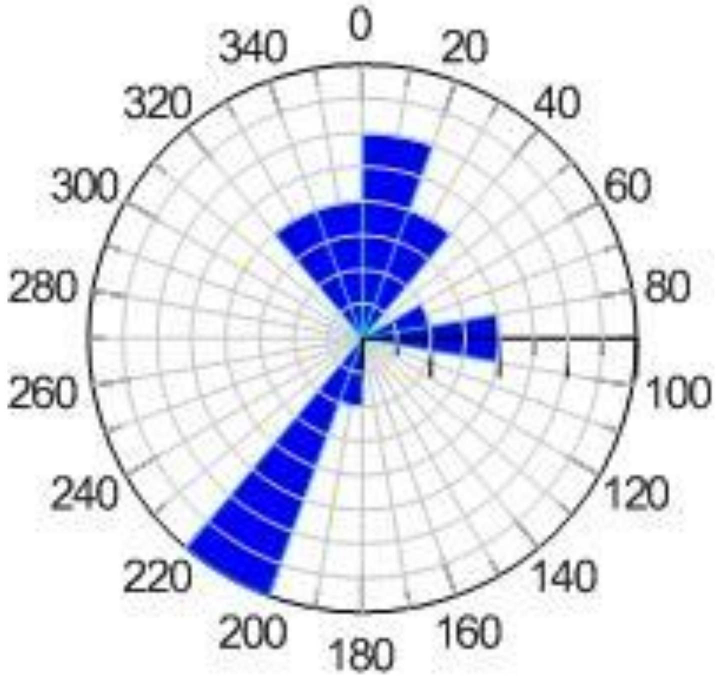
$$mean\ (\overline{R}) = \frac{R}{n} = \frac{\left(X_r^2 + Y_r^2\right)^{½}}{n}$$

$$\sigma_0 = S_0^2 = 1 - \overline{R} = \frac{n - R}{n}$$

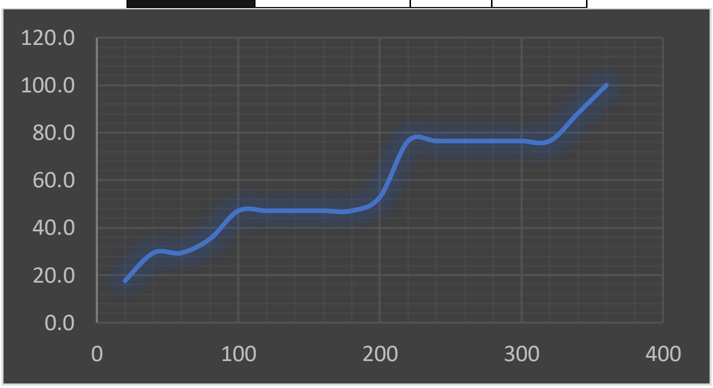
High mean R = less variance, circular variance (range between 1-0 & 0 indicate that all directions are identical)

WORKED EXAMPLE: LAB 5

Examine the following rose diagram & fill in the data table, & draw a cumulative frequency curve for these data (if every 1 unit on diagram = 2)



Bins	Frequency	f%	Cf
0-20	12	17.6	17.6
20-40	8	11.8	29.4
40-60	0	0.0	29.4
60-80	4	5.9	35.3
80-100	8	11.8	47.1
100-120	0	0.0	47.1
120-140	0	0.0	47.1
140-160	0	0.0	47.1
160-180	0	0.0	47.1
180-200	4	5.9	52.9
200-220	16	23.5	76.5
220-240	0	0.0	76.5
240-260	0	0.0	76.5
260-280	0	0.0	76.5
280-300	0	0.0	76.5
300-320	0	0.0	76.5
320-340	8	11.8	88.2
340-360	8	11.8	100.0
Total	68		



WORKED EXAMPLE: LAB 6

Strike	Dip	Dip Direction	Strike	Dip	Dip Direction
435	65	NE	60	83	SE
329	83	SW	52	84	SE
328	75	NE	68	84	SE
347	74	NE	73	86	NW
330	32	SW	57	70	SE
346	81	NE	70	84	NW
277	74	NE	47	80	NW
330	81	NE	18	85	NW
347	78	NE	38	40	SE
349	48	NE	77	82	SE
292	84	SW	11	27	NE
350	77	NE	78	80	NW
351	15	NE	48	85	NW
350	84	SW	6	56	SE
349	83	NE	39	66	SE
358	80	SW	63	82	SE
300	82	NE	68	80	NW
338	55	NE	53	88	SE
343	78	NE	17	83	NW
353	38	NE	55	82	NW
348	83	SW	72	80	NW
348	81	NE	42	66	SE
287	82	NE	83	75	SE
303	78	NE	52	69	NW
328	75	NE	1	88	SE
293	85	SW	37	84	NW
348	76	NE	68	75	NW
307	69	NE	51	62	NW
316	43	SW	4	88	SE
273	83	SW	55	84	SE
298	18	SW	18	82	NW
303	83	NE	18	80	SE
342	76	SW	54	82	SE
340	22	SW	68	75	SE
280	38	SW	32	85	NW
306	38	SW	52	87	NW
285	82	NE	72	50	NW
312	70	SW	41	64	SE
298	18	SW	84	90	NW
310	37	SW	3	53	NW
313	83	SW	53	70	NW
347	51	NE	71	72	NW
20	78	NW	15	68	SE
60	83	SE	45	70	NW

Use data on the previous table which represents strike & dip for different strata to calculate the mean direction & standard divisions of strike & dip, & to draw a rose diagram for strikes if bias = 10° increments & determine the trending direction

First step we need to convert dip direction into real numbers (by adding 90° for dip that extended in SE, 180° for dip that extended in SW direction, 270° for dip that extended in NW, we don't add anything to dip that extended in NE direction) then we need to calculate sinθ & cosθ for both strike & dip

Strike	radian	sinθ	cosθ	Dip	radians	sinθ	cosθ
435	7.6	1.0	0.3	65	1.1	0.9	0.4
329	5.7	-0.5	0.9	263	4.6	-1.0	-0.1
328	5.7	-0.5	0.8	75	1.3	1.0	0.3
347	6.1	-0.2	1.0	74	1.3	1.0	0.3
330	5.8	-0.5	0.9	212	3.7	-0.5	-0.8
346	6.0	-0.2	1.0	81	1.4	1.0	0.2
277	4.8	-1.0	0.1	74	1.3	1.0	0.3
330	5.8	-0.5	0.9	81	1.4	1.0	0.2
347	6.1	-0.2	1.0	78	1.4	1.0	0.2
349	6.1	-0.2	1.0	48	0.8	0.7	0.7
292	5.1	-0.9	0.4	264	4.6	-1.0	-0.1
350	6.1	-0.2	1.0	77	1.3	1.0	0.2
351	6.1	-0.2	1.0	15	0.3	0.3	1.0
350	6.1	-0.2	1.0	264	4.6	-1.0	-0.1
349	6.1	-0.2	1.0	83	1.4	1.0	0.1
358	6.2	-0.0	1.0	260	4.5	-1.0	-0.2
300	5.2	-0.9	0.5	82	1.4	1.0	0.1
338	5.9	-0.4	0.9	55	1.0	0.8	0.6
343	6.0	-0.3	1.0	78	1.4	1.0	0.2
353	6.2	-0.1	1.0	38	0.7	0.6	0.8
348	6.1	-0.2	1.0	263	4.6	-1.0	-0.1
348	6.1	-0.2	1.0	81	1.4	1.0	0.2
287	5.0	-1.0	0.3	82	1.4	1.0	0.1
303	5.3	-0.8	0.5	78	1.4	1.0	0.2
328	5.3	-0.5	0.8	75	1.3	1.0	0.3
293	5.1	-0.9	0.4	265	4.6	-1.0	-0.1
348	6.1	-0.2	1.0	76	1.3	1.0	0.2
307	5.4	-0.8	0.6	69	1.2	0.9	0.4
316	5.5	-0.7	0.7	223	3.9	-0.7	-0.7
273	4.8	-1.0	0.1	263	4.6	-1.0	-0.1
298	5.2	-0.9	0.5	198	3.5	-0.3	-1.0
303	5.3	-0.8	0.5	83	1.4	1.0	0.1
342	6.0	-0.3	1.0	256	4.5	-1.0	-0.2
340	5.9	-0.3	0.9	202	3.5	-0.4	-0.9
280	4.9	-1.0	0.2	218	3.8	-0.6	-0.8
306	5.3	-0.8	0.6	218	3.8	-0.6	-0.8
285	5.0	-1.0	0.3	82	1.4	1.0	0.1
312	5.0	-0.7	0.7	197	3.4	-0.3	-1.0
298	5.2	-0.9	0.5	198	3.5	-0.3	-1.0
310	5.4	-0.8	0.6	217	3.8	-0.6	-0.8
313	5.5	-0.7	0.7	218	3.8	-0.6	-0.8
347	6.1	-0.2	0.9	51	0.9	0.8	0.6

20	0.3	0.3	0.9	348	6.1	-0.2	1.0
60	1.0	0.9	0.5	163	2.8	0.3	-1.0
60	1.0	0.9	0.5	163	2.8	0.3	-1.0
52	0.9	0.8	0.6	164	2.9	0.3	-1.0
68	1.2	0.9	0.4	164	2.9	0.3	-1.0
73	1.3	1.0	0.3	356	6.0	-0.1	1.0
57	1.0	0.8	0.5	160	2.8	0.3	-0.9
70	1.2	0.9	0.3	354	6.2	-0.1	1.0
47	0.8	0.7	0.7	350	6.1	-0.2	1.0
18	0.3	0.3	1.0	355	6.2	-0.1	1.0
38	0.7	0.6	0.8	130	2.3	0.8	-0.6
77	1.3	1.0	0.2	172	3.0	0.1	-1.0
11	0.2	0.2	1.0	27	0.5	0.5	1.0
78	1.4	1.0	0.2	350	6.1	-0.2	1.0
48	0.8	0.7	0.7	355	6.2	-0.1	1.0
6	0.1	0.1	1.0	146	2.5	0.6	-0.8
39	0.7	0.6	0.8	156	2.7	0.4	-0.9
63	1.1	0.9	0.5	172	3.0	0.1	-1.0
68	1.2	0.9	0.4	350	6.1	-0.2	1.0
53	0.9	0.8	0.6	178	3.1	0.0	-1.0
17	0.3	0.3	1.0	353	6.2	-0.1	1.0
55	1.0	0.8	0.6	352	6.1	-0.1	1.0
72	1.3	1.0	0.3	350	6.1	-0.2	1.0
42	0.7	0.7	0.7	156	2.7	0.4	-0.9
83	1.4	1.0	0.1	165	2.9	0.3	-1.0
52	0.9	0.8	0.6	339	5.9	-0.4	0.9
1	0.0	0.0	1.0	178	3.1	0.0	-1.0
37	0.6	0.6	0.8	354	6.2	-0.1	1.0
68	1.2	0.9	0.4	345	6.0	-0.3	1.0
51	0.9	0.8	0.6	332	5.8	-0.5	0.9
4	0.1	0.1	1.0	178	3.1	0.0	-1.0
55	1.0	0.8	0.6	174	3.0	0.1	-1.0
18	0.3	0.3	1.0	352	6.1	-0.1	1.0
18	0.3	0.3	1.0	170	3.0	0.2	-1.0
54	0.9	0.8	0.6	172	3.0	0.1	-1.0
68	1.2	0.9	0.4	165	2.9	0.3	-1.0
32	0.6	0.5	0.8	355	6.2	-0.1	1.0
52	0.9	0.8	0.6	357	6.2	-0.1	1.0
72	1.3	1.0	0.3	320	5.6	-0.6	0.8
41	0.7	0.7	0.8	154	2.7	0.4	-0.9
84	1.5	1.0	0.1	360	6.3	0.00	1
3	0.1	0.1	1.0	323	5.6	-0.6	0.8
53	0.9	0.8	0.6	340	5.9	-0.3	0.9
71	1.3	0.9	0.3	342	6.0	-0.3	1.0
15	0.3	0.3	1.0	158	2.8	0.4	-0.9
45	0.8	0.7	0.7	340	5.9	-0.3	0.9
Sum		10.1	58.8			9.8	2.2

For Strike

$$mean\ (\overline{R}) = \frac{R}{n} = \frac{\left(10.1^2 + 58.8^2\right)^{½}}{88} = 0.68$$

$$mean\ direction = \tan^{-1}\left(\frac{10.1}{58.8}\right) = 9.7$$

$$S_0^2 = 1 - \overline{R} = 1 - 0.68 = 0.32$$

Environments : Alluvial

For Dip

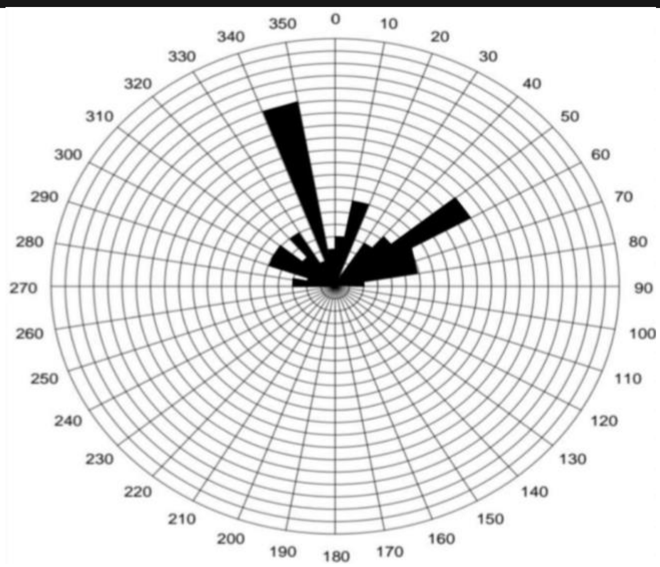
$$mean\ (\overline{R}) = \frac{R}{n} = \frac{\left(9.8^2 + 2.2^2\right)^{½}}{88} = 0.11$$

$$mean\ direction = \tan^{-1}\left(\frac{9.8}{2.2}\right) = 7.7$$

$$S_0^2 = 1 - \overline{R} = 1 - 0.68 = 0.89$$

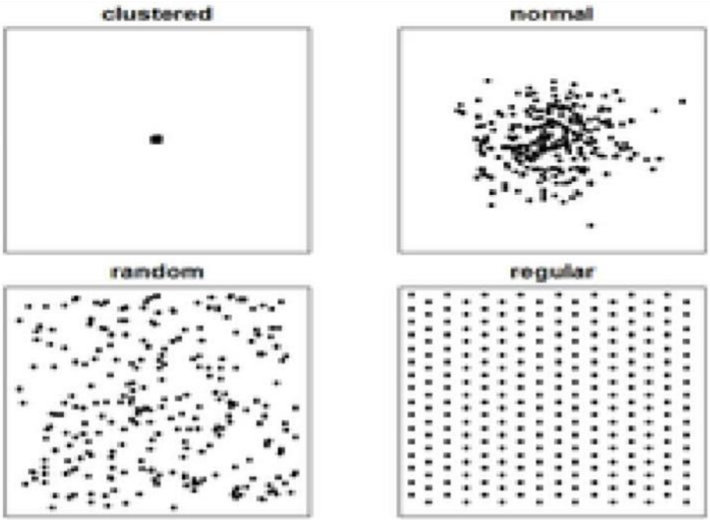
Rose Diagram For strikes

Bin	f	Bin	F
0-10	40	180-190	0
10-20	70	190-200	0
20-30	0	200-210	0
30-40	40	210-220	0
40-50	50	220-230	0
50-60	110	230-240	0
60-70	60	240-250	0
70-80	60	250-260	0
80-90	20	260-270	0
90-100	0	270-280	30
100-110	0	280-290	20
110-120	0	290-300	50
120-130	0	300-310	50
130-140	0	310-320	30
140-150	0	320-330	50
150-160	0	330-340	20
160-170	0	340-350	140
170-180	0	350-360	30
SUM		360	870

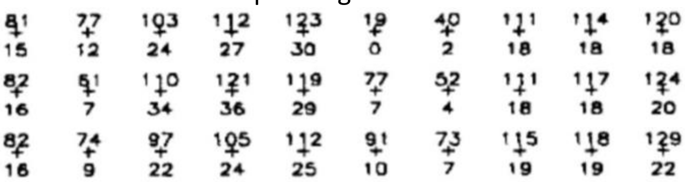


GEOGRAPHIC DISTRIBUTION

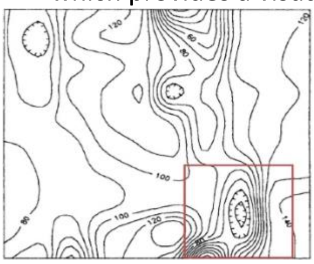
- **Randomness**: when the position of point has operated independently of positions of other points & involved equal probability of occurrence in equal subdivision
- **Non Randomness**
  1. **Uniform**: distribution tending to more homogeneity
  2. **Clustered**: distribution tend to more heterogeneous
  3. **Isotopic**: no directional relationships (no fabric)
  4. **Anisotropic**: points forming lineations or trending



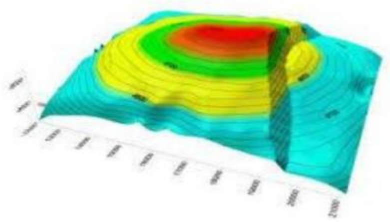
- **Postings**: important initial step in analyzing spatial data
  - The simplest display of spatial data is a data posting, a map on which each data location is plotted along with its corresponding data value



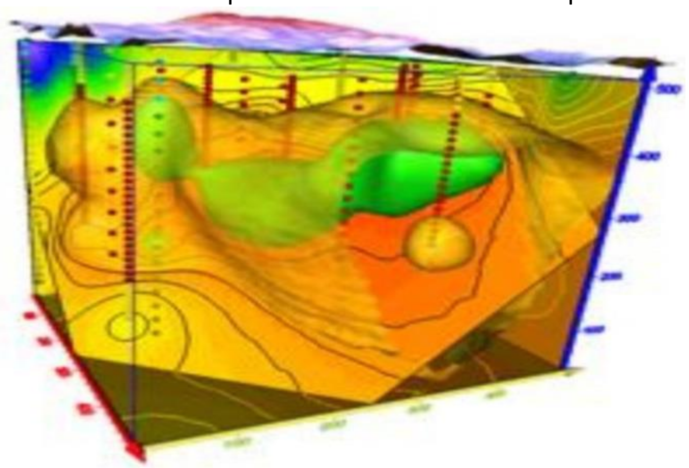
- **Contour Maps:** overall trends revealed by contour map, which provides a visual display & contouring by hand



2D Contour Map

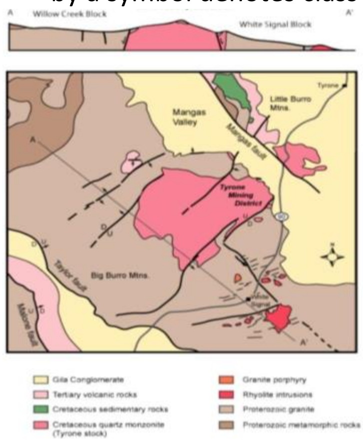


3D Surface map

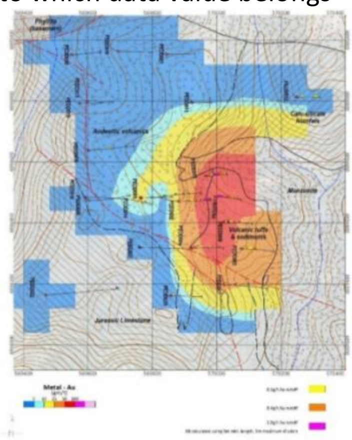


Surface & Subsurface 3D map

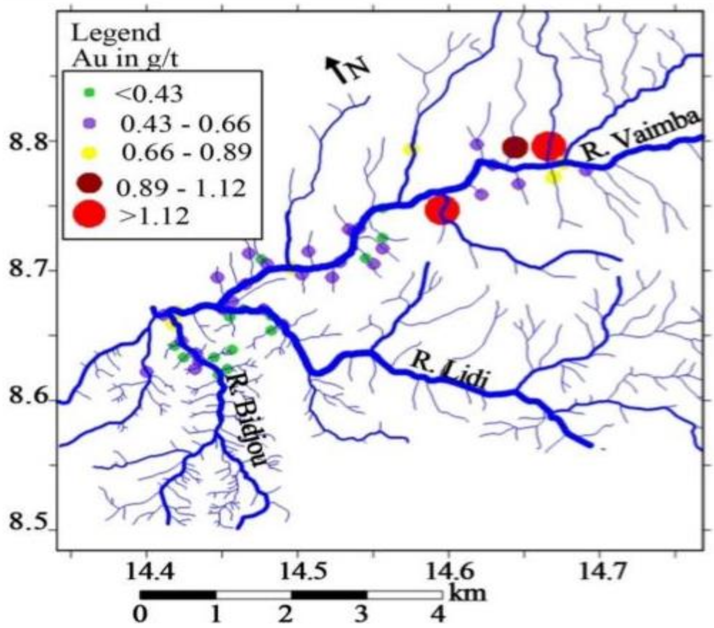
- For many very large regularly gridded data sets, a posting of all the data values may not be feasible, & a contour map may mask many of the interesting local details so we use symbol map
- **Symbol maps:** data posting with each location replaced by a symbol denotes class to which data value belongs



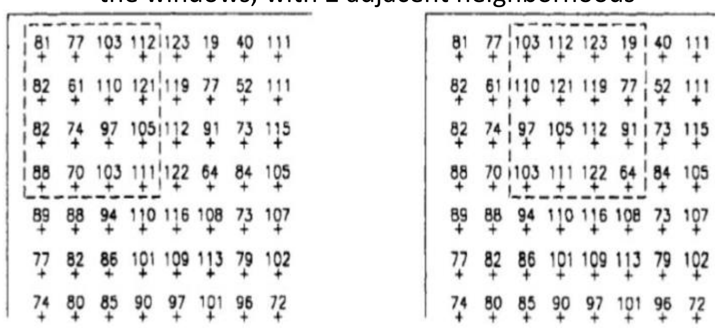
Symbol Map



Symbolic Colored Mab



- The geologic data sets are most interested in the anomalies or the impermeable layers that condition flow in a reservoir
- **Moving Window**: calculation of few summary statistics within moving windows is frequently used to investigate anomalies both in the average value & in variability
  - A good compromise is often found in overlapping the windows, with 2 adjacent neighborhoods



Overlapping moving window to calculate moving average statistics

We have chosen to use a 4x4m² window so that we will have 16 data in each window. By moving the window only 2m so that it overlaps half of the previous window, we can fit 16 such windows into our 10x10m² area

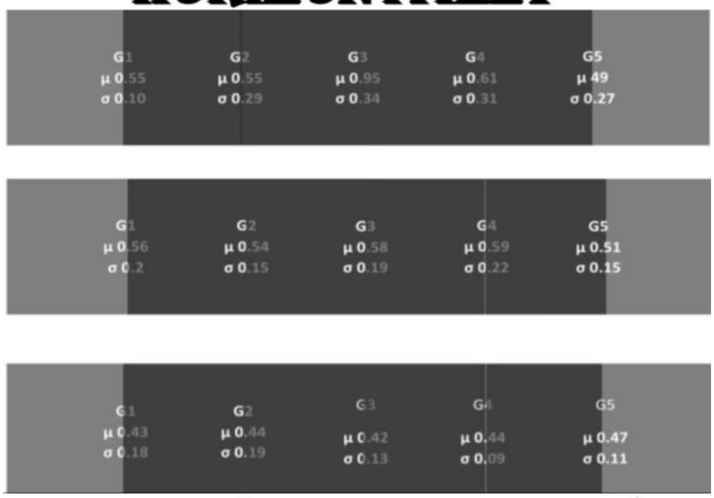
WORKED EXAMPLE: LAB 6

In the following table of 300m*300m total grid size

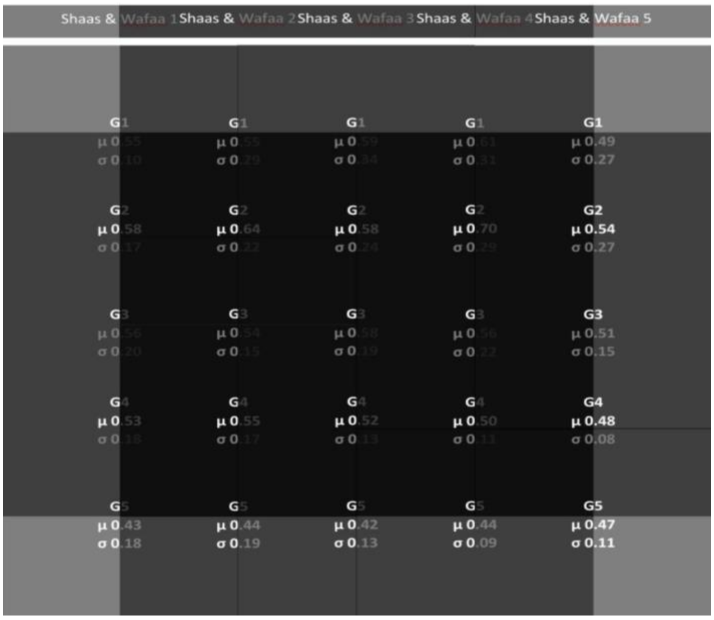
1. Use the moving window techniques to calculate the average grad of 100*100m² grid block. Move in 50m increments along the direction of vertical axis then in horizontal axis
2. Determine standard division of 100m*100m grid block
3. Draw a scatterplot of the local means versus the local standard divisions from 16 local neighborhoods
4. Draw a sketch represent all mean & standard division
5. Comment on the result

5.8	5.8	4.9	3.3	2.2	2.1	3.4	2.8	4.6	3.8	3.5	3.5
6.0	6.4	6.6	4.4	2.8	1.6	2.8	4.9	6.8	9.4	4.5	3.4
5.6	6.2	5.9	5.6	11.1	9.8	6.6	8.4	2.6	4.9	2.8	3.8
6.6	4.0	4.4	6.9	6.1	9.6	9.9	9.8	12.7	4.7	4.1	2.9
9.1	9.1	5.8	4.5	4.9	8.1	11	6.9	5.5	10.2	6.3	6.2
4.2	5.8	3.0	6.3	5.4	4.9	5.0	4.5	4.5	5.0	5.0	4.5
4.2	3.9	5.0	5.6	4.2	4.6	6.1	4.5	3.9	4.9	3.9	4.0
3.4	4.4	7.1	5.4	4.9	4.4	8.3	4.9	4.4	4.0	4.9	4.6
4.6	7.4	9.2	5.6	7.8	5.8	4.4	4.1	4.4	4.9	5.5	6.5
4.0	4.1	4.6	3.6	4.0	3.6	6.0	5.1	4.8	4.8	6.3	4.4
3.3	2.7	3.1	3.4	3.6	3.5	4.2	3.0	3.9	3.9	6.2	3.2
4.2	3.5	2.4	3.6	2.8	3.3	3.2	3.7	3.4	6.4	3.8	3.5
	50		100		150		200		250		

HORIZONTALLY



VERTICALLY



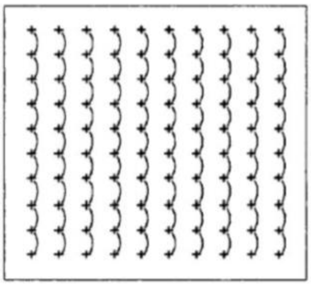
Solution: Click Here

ESTIMATION

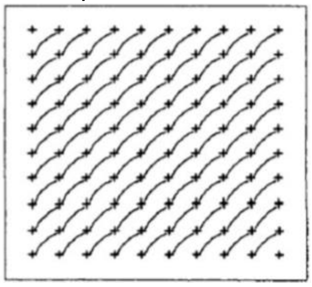
- Anomalies in the local variability will have an impact on the accuracy of our estimates
- A scatterplot of the local means & the local standard deviations from our moving window calculations is a good way to check for a relationship between the two. If it exists, such a relationship is generally referred to as a proportional effect
- **For lognormally distributed values**, a scatterplot of local means versus local standard deviations will show a **linear relationship between the two**

H-SCATTERPLOTS

- **h-scatterplot** shows all possible pairs whose locations separated by certain distance in a particular direction
- On h-scatterplots, the x-axis is labeled V(t) & the y-axis is labeled V(t + h), The x-coordinate of a point corresponds to V at a particular location & y-coordinate to V value distance & direction h away



H (0,1)



H (1,1)

$$\mathrm{COR} = C(h) = \frac{1}{N(h)} \Sigma_{(i,j)|h_{ij}=h} (v_i . v_j - m_{-h} . m_{+h})$$

$$m_{-h} = \frac{1}{N(h)} \Sigma_{i|h_{ij}=h} V_i \ \& \ m_{+h} = \frac{1}{N(h)} \Sigma_{i|h_{ij}=h} V_j$$

$$\mathrm{CORRELOGRAM} = \rho(h) = \frac{C(h)}{\sigma_{-h} . \sigma_{+h}}$$

$$\sigma^2_{-h} = \frac{1}{N(h)} \Sigma_{i|h_{ij}=h} (V_i^2 - m^2_{-h})$$

$$\sigma^2_{+h} = \frac{1}{N(h)} \Sigma_{i|h_{ij}=h} (V_j^2 - m^2_{+h})$$

WORKED EXAMPLE: LAB 4

For the data in the table below

1. Draw the h-scatter plot for h(0,1), h(0,2), h(0,3), & h(0,4)
2. Draw the h-scatter plot for h(1,0), h(2,0), h(3,0), & h(4,0)
3. Calculate COV, COR, & VAR for each h-scatter plot
4. Plot the vertical distance Vs VAR, COV, & COR

81	77	103	112	123	19	40	111	114	120
82	61	110	121	119	77	52	111	117	124
82	74	97	105	112	91	73	115	118	129
88	70	103	111	122	64	84	105	113	123
89	88	94	110	116	105	73	107	118	127
77	82	85	101	109	113	79	102	120	121
74	80	85	90	97	101	95	72	128	130
75	80	83	87	94	99	95	48	139	145
77	84	74	108	121	143	91	52	136	144
87	100	47	111	124	109	0	98	134	144

Solution: Click Here

TRICKY EXAMPLE

Use the following table to calculate the correlation coefficient of h=(0, 1), h=(0, 2), h=(0, 3), h=(0, 4), h=(1, 0), h=(2, 0), h=(3, 0), & h=(4, 0), & draw the curve represents the correlation coefficient Vs vertical & horizontal distance

6	2		2
	7	5	
			2
2	5	4	8
3		2	
	2		6
2		6	

h=(0, 1) → COR = -0.67936622

X	3	7	2	8
Y	2	2	4	2

h=(0, 2) → COR = 0.436367693

X	2	2	5	6	4	6	2
Y	3	5	7	2	5	8	2

h=(0, 3) → COR = -0.555420384

X	2	2	5	6	2	6	8
Y	2	6	2	4	5	2	2

h=(0, 4) → COR = -1

X	3	2
Y	6	7

h=(1, 0) → COR = -0.371690882

X	2	5	4	7	6
Y	5	4	8	5	2

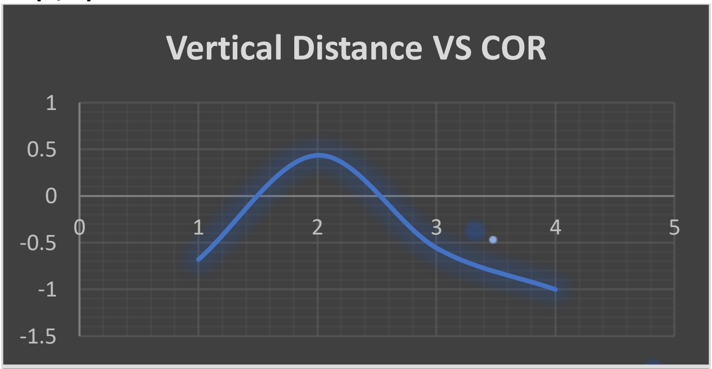
h=(2, 0) → COR = 0.5

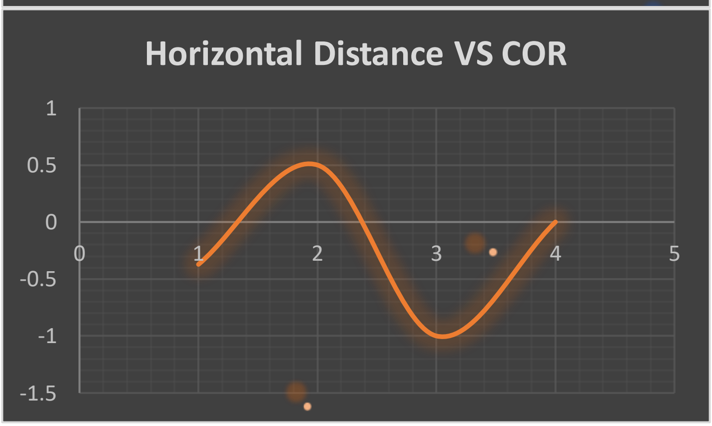
X	2	2	3	2	5	2
Y	6	6	2	4	8	2

h=(3, 0) → COR = 0.5

X	2	6
Y	8	2

h=(4, 0) → COR = 0

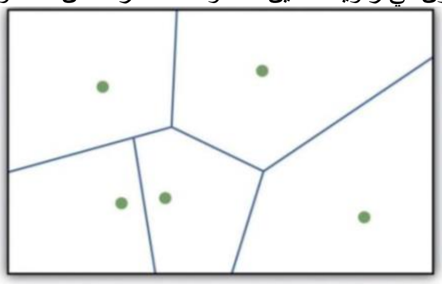


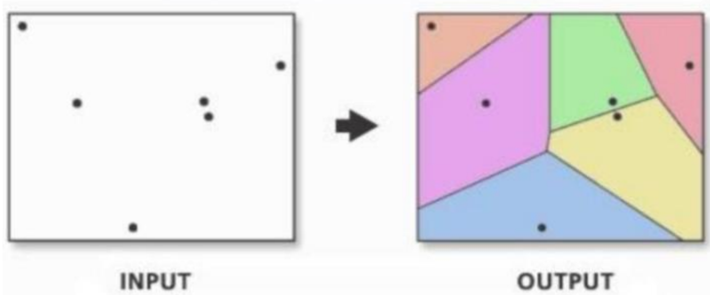


ESTIMATION METHODS

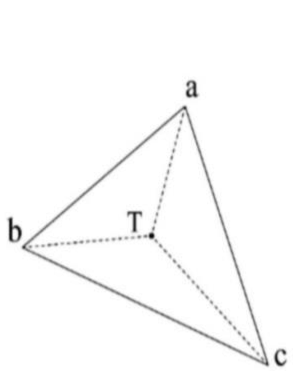
- **Polygonal method**: Connect between points
  - Draw a line perpendicular to the connection line (pass trough the 2 points) & connect lines to create a polygon around each point
  - All the area in the polygon have the same value

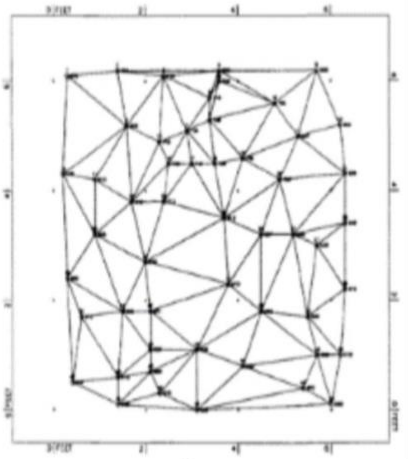
نرسم خطوط بين كل نقطتين ونمد خط عامودي عليها بحيث ينصف الخط بين النقطتين ويكون في زاوية تسعين معه ونحذف الزائد من الخطوط

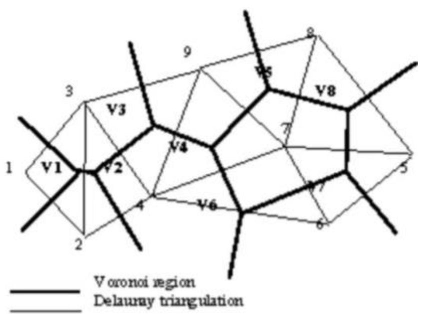




- **Triangulation**: estimation of a new grid point value can be done using points at apices of the triangle which it lies or more commonly & for a more accurate result, these 3 + 3 extra points at the apices of the 3 adjoining triangles
  - Result: every point is within or on edge of a triangle
  - **Delauney Triangulation**: standard & best method in which the acuteness of triangles is minimized mostly done by a computer algorithm. Which fortunately available in in spatial analysis packages







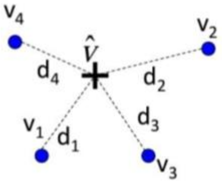
- **Weighted Linear Combinations**: Different approaches to assigning the weights to the data values give rise to many different methodologies

$$\bar{v} = \Sigma w_i v_i$$

- **Inverse distance technique**: simplest interpolation method, give more weight to the closest samples, & less to those that are farthest away.

$$\bar{v} = \Sigma w_i v_i = \Sigma \frac{\left(\frac{1}{d_i}\right)}{\left(\Sigma \frac{1}{d_i^n}\right)} x V_i$$

Inverse Distance Square

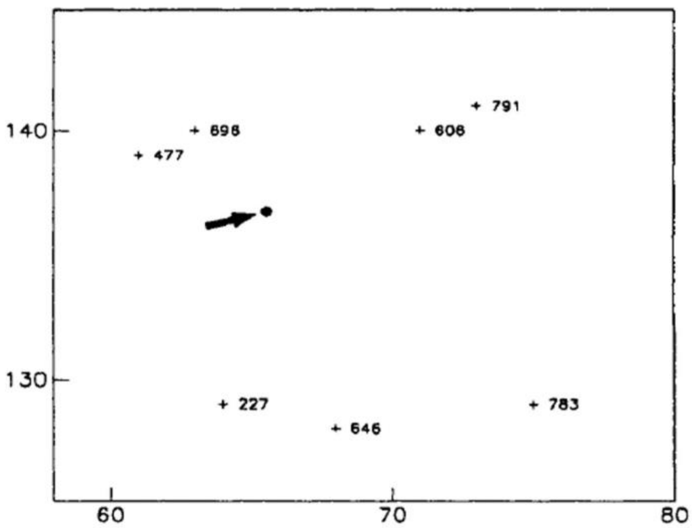


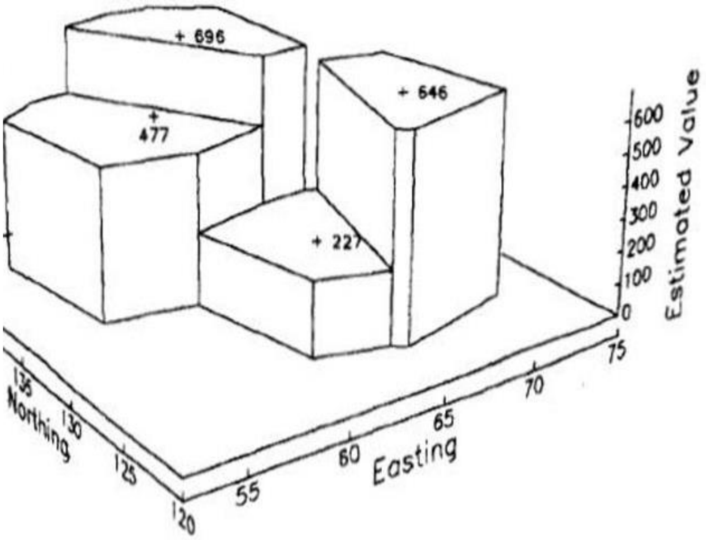
$$\hat{v} = \frac{\frac{1}{d_1^2}}{\sum_{i=1}^{4} \frac{1}{d_i^2}} v_1 + \frac{\frac{1}{d_2^2}}{\sum_{i=1}^{4} \frac{1}{d_i^2}} v_2 + \frac{\frac{1}{d_3^2}}{\sum_{i=1}^{4} \frac{1}{d_i^2}} v_3 + \frac{\frac{1}{d_4^2}}{\sum_{i=1}^{4} \frac{1}{d_i^2}} v_4$$

- **Error**

$$r = \overline{X}_{True} - \overline{X}_{Estimated}$$
$$r = \sigma^2_{True} - \sigma^2_{Estimated}$$
$$r = \overline{V} - V$$

Example: estimate the V value at point 65E-137N located by arrow from the surrounding 7 samples in the figure below (Table shows the distances from the 65E, 137N)



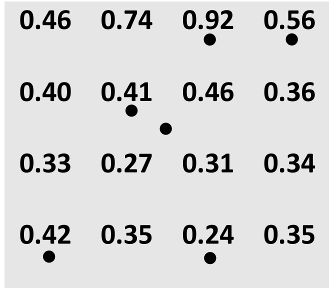


Using Polygonal Estimate: We choose the value that is closest to the point V (63E,140N) so our polygonal estimate (65E, 137N) is 696 ppm. Polygonal estimates of the V value at other points near 65E, 137N will also be 696 ppm

WORKED EXAMPLE

Suppose the area of the whole data below is 100m*100m

1. Estimate the grade of central point by inverse distance techniques & inverse distance squared techniques
2. Calculate mean & standard division of estimated values
3. Calculate the estimation error for μ & σ in each case
4. Draw a diagram represents the estimation errors



Using inverse distance techniques

V	d-X	d-Y	d	1/d	W	$\overline{V}$
0.42	37.5	37.5	53.03	0.019	0.130	0.055
0.24	12.5	37.5	39.53	0.025	0.175	0.042
0.41	12.5	12.5	17.69	0.057	0.390	0.160
0.92	12.5	37.5	39.53	0.025	0.175	0.161
0.56	37.5	37.5	53.03	0.019	0.130	0.073
		Sum		0.145		0.490

$$\mu_{estimated} = \frac{0.42 + 0.24 + 0.41 + 0.92 + 0.56 + 0.49}{6} = 0.507$$

$$\sigma_{estimated} = \left(\frac{(0.42 - 0.507)^2 + \cdots + (0.49 - 0.507)^2}{n-1}\right)^{½} = 0.229$$

$$\mu_{True} = \frac{0.46 + \cdots + 0.35}{16} = 0.433$$

$$\sigma_{True} = \left(\frac{(0.46 - 0.4325)^2 + \cdots + (0.35 - 0.4325)^2}{n-1}\right)^{½} = 0.177$$

$$r = \overline{X}_{True} - \overline{X}_{Estimated} = 0.433 - 0.507 = -0.074$$
$$r = \sigma^2_{True} - \sigma^2_{Estimated} = 0.177 - 0.229 = -0.052$$

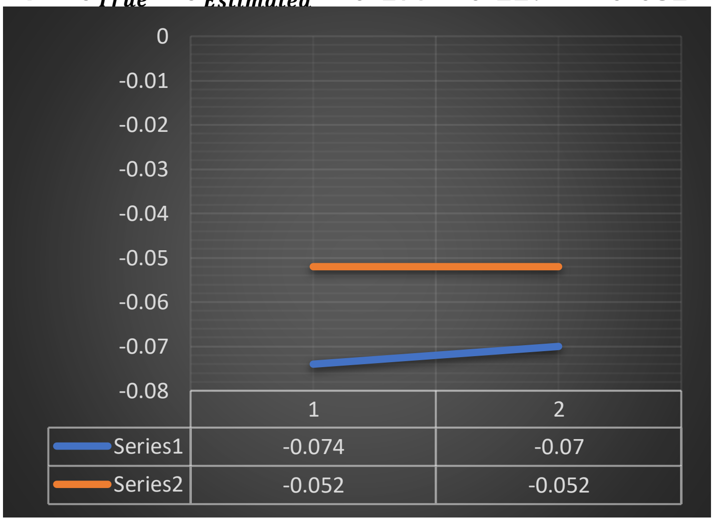
Using inverse distance square techniques

V	d-X	d-Y	d	1/d²	W	$\overline{V}$
0.42	37.5	37.5	53.03	0.00036	0.069	0.029
0.24	12.5	37.5	39.53	0.00064	0.123	0.030
0.41	12.5	12.5	17.69	00.0032	0.620	0.254
0.92	12.5	37.5	39.53	0.00064	0.123	0.113
0.56	37.5	37.5	53.03	0.00036	0.069	0.039
		Sum		0.0052		0.465

$$\mu_{estimated} = \frac{0.42 + 0.24 + 0.41 + 0.92 + 0.56 + 0.465}{6} = 0.503$$

$$\sigma_{esti.} = \left(\frac{(0.42 - 0.507)^2 + \cdots + (0.465 - 0.507)^2}{n-1}\right)^{½} = 0.229$$

$$r = \overline{X}_{True} - \overline{X}_{Estimated} = 0.433 - 0.503 = -0.070$$
$$r = \sigma^2_{True} - \sigma^2_{Estimated} = 0.177 - 0.229 = -0.052$$



TRICKY EXAMPLE

The following data are ***coordination*** of several points, use these data to estimate the point V that located at ***(0.2, 0.2)*** coordination using inverse distance techniques (ID), inverse distance square techniques (IDS), & Inverse distance cube techniques (IDC), then sketch a curve that represent all values calculated (relationships with distance)

	V	X	Y
1	0.3	0.4	0.8
2	0.2	0.5	0.2
3	0.4	0.2	0.4
4	0.2	0.3	0.6
5	0.3	0.5	0.2

First Step: Calculate the vertical & horizontal distance between the V & 1; 2; 3; 4; & 5, then calculate the true distance (triangle chord length) between V & these points

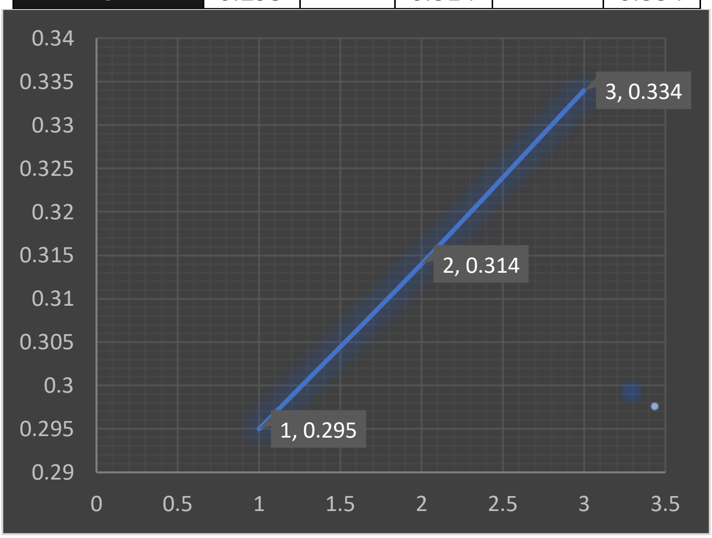
	V	X	Y	d	1/d	$1/d^2$	$1/d^3$
1	0.3	0.2	0.6	0.63	1.58	02.50	03.95
2	0.2	0.3	0.0	0.30	3.33	11.11	37.04
3	0.4	0.0	0.2	0.20	5.00	25.00	125.0
4	0.2	0.1	0.4	0.41	2.43	05.88	14.27
5	0.3	0.3	0.0	0.30	3.33	11.11	37.04

Second Step: Calculate Wi for inverse distance techniques (ID), inverse distance square techniques (IDS), & Inverse distance cube techniques (IDC)

	1/d	W_1	$1/d^2$	W_2	$1/d^3$	W_3
1	1.58	0.10	02.50	0.045	03.95	0.0182
2	3.33	0.21	11.11	0.200	37.04	0.1707
3	5.00	0.32	25.00	0.450	125.0	0.5760
4	2.43	0.15	05.88	0.106	14.27	0.0658
5	3.33	0.21	11.11	0.200	37.04	0.1707

Third step: Calculate Wi for inverse distance techniques (ID), inverse distance square techniques (IDS), & Inverse distance cube techniques (IDC)

	V	W_1	$V*W_1$	W_2	$V*W_2$	W_3	$V*W_3$
1	0.3	0.10	0.030	0.045	0.014	0.0182	0.006
2	0.2	0.21	0.042	0.200	0.034	0.1707	0.034
3	0.4	0.32	0.127	0.450	0.180	0.5760	0.230
4	0.2	0.15	0.031	0.106	0.021	0.0658	0.013
5	0.3	0.21	0.064	0.200	0.060	0.1707	0.051
V			0.295		0.314		0.334



WORKED EXAMPLE: LAB7-8

1. For the deposit assignd below, assume it was modeled by using 100m*100m blocks, estimate the grade of these blocks for the exploration data set by using the inverse distance square (IDS) techniques with rectangular search neighborhood of 50m*50m
2. Using the estimated block values results from question one, determine the overall estimated bench averaging the grades of copper ore at 0.5ppm cutoff
3. Using the estimated block values results from question one, determine the overall estimated bench averaging the grades of 16 blast holes that fall within each block
4. Compare the inverse distance estimated average block grade in Q1 to the block grades determined in Q3
   a) Draw the histogram of estimation errors calculated tracking the difference between the IDS (Q1) & the true blast hole calculated grade (Q3)
   b) Determine the quantitative statistics about the errors by determining the Min. & Max. error & calculate the mean, the variance, the first quartile, the second quartile, the third quartile of the errors & make comments about these statistics, what would be the influence of these errors?
   c) Plot IDS against true grades on X-Y scatter diagram
   d) Determine the following
      I. Number of blocks estimated by inverse distance square to be >100ppm V yet they have true grades (blast hole estimation) <100ppm
      II. Number of blocks estimated by IDS to have average grade <0.5%Cu yet their true grade in >100ppm V
      III. Make comments on the results (Differences)

	5.8	5.6	4.9	3.3	2.2	2.1	3.4	2.8	4.6	3.8	3.5	3.5
	6.0	6.4	6.6	4.4	2.8	1.6	2.8	4.9	6.8	9.4	4.6	3.4
	5.6	6.2	5.9	5.6	11	9.8	6.6	8.4	2.6	4.9	2.8	3.8
	6.6	4.0	4.4	6.9	6.1	9.6	9.9	9.8	13	4.7	4.1	2.9
200	9.1	9.1	5.8	4.5	4.9	8.1	11	6.9	5.5	10	6.3	6.2
	4.2	5.8	3.0	6.3	5.4	4.9	5.0	4.5	4.5	5.0	5.0	4.5
	4.2	3.9	5.0	8.6	4.2	4.6	6.1	4.5	3.9	4.9	3.9	4.0
100	3.4	4.4	7.1	5.4	4.9	4.4	8.3	4.9	4.4	4.0	4.9	4.6
	4.6	7.4	9.2	5.6	7.8	5.6	4.4	4.1	4.4	4.9	5.5	6.5
	4.0	4.1	4.6	3.6	4.0	3.6	6.0	5.1	4.8	4.8	6.3	4.4
	3.3	2.7	3.1	3.4	3.6	3.5	4.2	3.0	3.8	3.9	6.2	3.2
	4.2	3.5	2.4	3.5	2.8	3.3	3.2	3.7	3.4	6.4	3.8	3.5
0				100				200				

TONNAGE CALCULATIONS FOR NORMAL DEPOSITS

EXAMPLE OF CALCULATION OF TONNAGE & AVERAGE GRADE FOR NORMAL DISTRIBUTION

A normal deposits of 20Mton with m_0 (mean) -0.742%Cu, & variance 0.145, calculate the TONNAGE & AVERAGE GRADE ABOVE THE CUTOFF of Z=-0.92% & Z=-0.51%

اولاً : يجب حساب standard normal value

$$X_c = \frac{z - m_0}{\sigma}$$

$$X_c = \frac{-0.92 - -0.742}{0.145^{½}} = -0.47$$

ثانياً : نحسب ال Cumulative Density Function

F(z) = Cdf = P(Z≤ z) = P(Z≤-0.47)

$$F(X_c) = ½\left[1 - \left(1 - e^{-\frac{2X_c{}^2}{\pi}}\right)^{½}\right]$$

او نستخدم جدول التوزيع الطبيعي

$$F(X_c) = ½\left[1 - \left(1 - e^{-0.13911}\right)^{½}\right] = 0.3198$$

Normal Distribution table
The number in table represents all probabilities that less than or equal z P(Z≤z)

z	0.0x	0.07	0.03
x.x sf	Cdf, F(x)		
-0.4		0.3192	
0.5			0.7019

ثالثاً : نحسب ال Inverse Cumulative Density Function

G(z) = P(Z> z) = P(Z>-0.47)

$$G(X_c) = F(X_c)\ if\ X < 0$$
$$G(X_c) = 1 - F(X_c)\ if\ X > 0$$

$$G(X_c) = F(X_c) = 0.3198$$

رابعاً : نحسب ال TONNAGE

$$T_0 = wtxG(X_c)$$

$$T_0 = 20Mton * 0.3198 = 6.396Mton$$

خامساً : نحسب ال AVERAGE GRADE ABOVE THE CUTOFF

$$m(X_c) = m_0 + \frac{e^{-\frac{X^2}{2}} * \sigma}{G(x) * (2\pi)^{½}}$$

$$m(X_c) = -0.742 + \frac{e^{-\frac{(-0.47)^2}{2}} x\, 0.145^{½}}{0.3198x(2\pi)^{½}} = -0.317$$

TONNAGE CALCULATIONS (LOG-NORMAL DEPOSITS)

A log-normal deposits of 20Mton with m_0 (mean) 0.512%Cu, & variance 0.041, calculate the TONNAGE & AVERAGE GRADE ABOVE THE CUTOFF of Z=0.4% & Z=0.6%

اولاً : هذه البيانات ليست طبيعية (log normal) لذا يجب تحويل ال m & σ الى m & σ لبيانات طبيعية (α, & β)

$$\beta^2 = \ln^{\left(\frac{\sigma^2}{m^2}+1\right)}$$

$$\alpha = \ln^m - ½\beta^2 = \ln^m - ½\ln^{\left(\frac{\sigma^2}{m^2}+1\right)}$$

$$\beta^2 = \ln^{\left(\frac{\sigma^2}{m^2}+1\right)} = 0.145 \rightarrow \beta = 0.381$$

$$\alpha = \ln^m - ½\beta^2 = ln^{0.512} - ½ - 0.145 = -0.742$$

ثانياً : نحسب القيمة المعيارية Xc التي تعادل التوزيع الطبيعي ليتسنى لنا استخدام جداول التوزيع الطبيعي والمعادلات

$$X_c = \frac{ln^Z - \alpha}{\beta}$$

$$X_{0.4} = \frac{\ln^{0.4} - -0.742}{0.381} = \frac{\ln^{0.4} - -0.742}{0.381} = -0.46$$

ثالثاً : نحسب ال Cumulative Density Function

F(z) = Cdf = P(Z≤ z) = P(Z≤-0.46)

$$F(X_c) = ½\left[1 - \left(1 - e^{-\frac{2X_c{}^2}{\pi}}\right)^{½}\right]$$

$$F(z) = ½\left[1 - \left(1 - e^{-0.135}\right)^{½}\right] = 0.32232$$

ثالثاً : نحسب ال Inverse Cumulative Density Function

G(z) = P(Z≤ z) = P(Z>-0.46)

$$G(X_c) = F(X_c)\ if\ X < 0$$
$$G(X_c) = 1 - F(X_c)\ if\ X > 0$$

$$G(-0.46) = F(-0.46) = 0.32232$$

رابعاً : نحسب ال TONNAGE

$$T_0 = wtxG(X_c)$$

$$T_0 = 20 * 0.32232 = 6.4464$$

خامساً : نحسب ال AVERAGE GRADE ABOVE THE CUTOFF

$$m(X_0) = \frac{G'(X_c)}{G(X_c)}$$

$$G'(X_c) = F'(X_c{}')\ if\ Xc < 0 = 1 - F'(X'_c)\ if\ X_c > 0$$

$$F'(X_c) = ½\left[1 - \left(1 - e^{-\frac{2X_c'^2}{\pi}}\right)^{½}\right]$$

$$X'_c = X_c - \beta$$

$$X'_c = -0.46 - 0.381 = -0.841$$
$$F'(-0.84) = 0.2005$$
$$G'(-0.84) = F'(-0.84) = 0.2005$$
$$m(X_0) = \frac{G'(X_c)}{G(X_c)} = \frac{0.2005}{0.32232} = 0.6221$$

For Z=0.6 → Tonnage = 5.4Mton & Avg. = 0.777

MULTIVARIATE DATA

	Variable 1	Variable 2	…	Variable n
Item A	X_{1A}	X_{2A}		X_{nA}
Item B	X_{1B}	X_{2B}		X_{nB}
…				
…				
Item n	X_{1n}	X_{2n}		X_{nn}

- **Euclidean Distance (Ed):**

$$\mathbf{Ed_{xn}} = \mathbf{root}[\boldsymbol{\Sigma}(\boldsymbol{X_x} - \boldsymbol{X_n})^2]$$

$$\mathrm{Ed_{An}} = \mathrm{root}[\Sigma(X_{1A} - X_{1n})^2 + \cdots + (X_{nA} - X_{nn})^2]$$

$$\mathrm{Ed_{Bn}} = \mathrm{root}[\Sigma(X_{1B} - X_{1n})^2 + \cdots + (X_{nB} - X_{nn})^2]$$

…

- **Euclidean Distance Coefficient (EdC):**

$$\mathbf{Ed_{xn}} = \frac{\mathbf{root}[\boldsymbol{\Sigma}(\boldsymbol{X_x} - \boldsymbol{X_n})^2]}{\boldsymbol{m}}$$

$$\mathrm{EdC_{An}} = \frac{\mathrm{Ed}}{\mathrm{m}} = \frac{[\Sigma(X_{1A} - X_{1n})^2 + \cdots + (X_{nA} - X_{nn})^2]^{½}}{m}$$

$$\mathrm{EdC_{Bn}} = \frac{\mathrm{Ed}}{\mathrm{m}} = \frac{[\Sigma(X_{1B} - X_{1n})^2 + \cdots + (X_{nB} - X_{nn})^2]^{½}}{m}$$

…

- **Manhattan Distance (Md):**

$$\mathbf{Md_{xn}} = \mathbf{\Sigma}|(\boldsymbol{X_X} - \boldsymbol{X_n})|$$

$$\mathrm{Md_{An}} = |(X_{1A} - X_{1n}) + (X_{2A} - X_{2n}) + (X_{nA} - X_{nn})|$$

$$\mathrm{Md_{Bn}} = |(X_{1B} - X_{1n}) + (X_{2B} - X_{2n}) + (X_{nB} - X_{nn})|$$

…

- **Manhattan Distance Coefficient (MdC):**

$$\mathbf{MdC_{xn}} = \frac{\mathbf{\Sigma}|(\boldsymbol{X_X} - \boldsymbol{X_n})|}{\boldsymbol{m}}$$

$$\mathrm{MdC_{An}} = \frac{\mathrm{Md}}{\mathrm{m}} = \frac{|(X_{1A} - X_{1n}) + (X_{2A} - X_{2n}) + (X_{nA} - X_{nn})|}{m}$$

$$\mathrm{MdC_{Bn}} = \frac{\mathrm{Md}}{\mathrm{m}} = \frac{|(X_{1B} - X_{1n}) + (X_{2B} - X_{2n}) + (X_{nB} - X_{nn})|}{m}$$

- **Correlation Similarity Coefficient (r_{AB})**

$$\mathbf{r_{xn}} = \frac{\mathbf{\Sigma_{i=1}^{i=n}(X_x - \mu_x)(X_n - \mu_n)}}{\mathbf{S_xS_n(m-1)}}$$

$$\mathrm{r_{An}} = \frac{\Sigma_{i=1}^{i=n}(\mathrm{X_{iA}} - \mu_A)(\mathrm{X_{in}} - \mu_n)}{\mathrm{S_AS_n(m-1)}}$$

$$\mathrm{r_{Bn}} = \frac{\Sigma_{i=1}^{i=n}(\mathrm{X_{iB}} - \mu_A)(\mathrm{X_{in}} - \mu_n)}{\mathrm{S_BS_n(m-1)}}$$

- **Jaccard Association Coefficient or Association Coefficient (J_{AB})**

$$\mathbf{J_{AB}} = \frac{\mathbf{a}}{\mathbf{a+b+c+d}}$$

a: is the number of attributes present in both samples
b: is the number of attributes present in only A sample
c: is the number of attributes present in only B sample
c: is the number of attributes absent in both samples

$J_{AB} = \frac{a}{a+b+c}$		Sample B	
		Present	Absent
Sample A	Present	a	c
	Absent	b	d

BOOK EXAMPLE

Use the matrix given below to calculate the following

1. Euclidean Coefficient of AB, AC, & BC
2. Manhattan Coefficient of AB, AC, & BC
3. Correlation Similarity Coefficient of AB, AC, & BC
4. Comment on the results

	X	Y	Z	M	K
A	-0.21	1.12	0.76	0.14	-0.93
B	0.64	3.13	2.58	-0.15	-0.50
C	-0.12	-0.39	0.76	-0.01	-0.07

Euclidean Distance Coefficient Using Excel

EdC AB = 0.577719655

EdC AC = 0.349302161

EdC BC = 0.812032019

Manhattan Distance Coefficient (MdC):

MdC AB = 1.08

MdC AC = 0.522

MdC BC = 1.334

Correlation Similarity Coefficien (r)

r AB = 0.906549698

r AC = 0.143549018

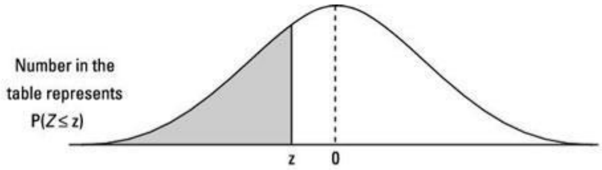
r BC = 0.179947104

FINALLY

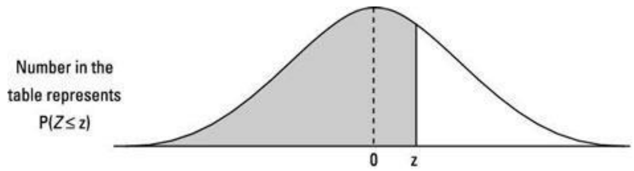
Calculate Jaccard Association Coefficient (J_{AB})

		Sample B	
		Present	Absent
Sample A	Present	10	7
	Absent	1	2

$$\mathbf{J_{AB}} = \frac{\mathbf{10}}{\mathbf{7+1+10}} = \frac{\mathbf{10}}{\mathbf{18}} = \mathbf{0.556}$$



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
–3.6	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
–3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
–3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
–3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
–3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
–3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
–3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
–2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
–2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
–2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
–2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
–2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
–2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
–2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
–2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
–2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
–2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
–1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
–1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
–1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
–1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
–1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
–1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
–1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
–1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
–1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
–1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
–0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
–0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
–0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
–0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
–0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
–0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
–0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
–0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
–0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
–0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.5	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998
3.6	.9998	.9998	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999