**CHAPTER 1 – BIVARIATE DATA**

**Bivariate data** analysis will determine if a relationship or association exists between two variables.

The **explanatory variable (independent)** is the one that explains or influences change in the response variable.

The **response variable** **(dependent)** measures an outcome or study. It depends on the explanatory variable.
- A significant change in proportion of percentages in the data indicates there is an association.

A **scatter plot** can be used to graph two numerical variables. The **explanatory** variable goes on the x-axis and the **response** variable goes on the y-axis.

When **interpreting the scatter plot**, discuss FORM (Linear, non linear, no relationship), DIRECTION (Positive, negative) , STRENGTH (Strong, moderate, weak)
- If there is a linear relationship between two variables, we can calculate the **correlation coefficient (*r*)**If r = 1 or -1, there is a perfect positive or perfect negative relationship
If r is between +- 0.75 and 1, there is a strong relationship
If r is between +- 0.5 and 0.75, there is a moderate relationship
If r is between +-0.25 and 0.5, there is a weak relationship
If r is between -0.25 and +0.25 there is no relationship

When describing using r, state the type of relationship (F,D,S) and then say that the y variable should \_\_ as the x variable \_\_

\*If two variables are correlated it doesn’t necessarily mean that a change in one **causes** a change in the other. There may be **external factors** causing the relationship.

**CHAPTER 2 – BIVARIATE DATA – FURTHER ANALYSIS**

The **least squares regression line** is y=ax + b.
The equations should be written in terms of the variable names.

- To interpret the y-intercept in the LSR line, we state “The y variable is “\_” units when the x variable is zero units.”

To interpret the gradient in the LSR line, we state “The y-variable increases/decreases by “\_” units for every one unit increase in the x-variable.”

To **plot the least squares regression line**, plot the y-intercept and then select an x value and substitute it into the LSR line equation to find a corresponding y value.

The **coefficient of determination (*r2*)** tells us the percentage of variation in the dependent variable that can be explained by the independent variable, therefore determining how useful or appropriate a linear model is.

When **interpreting the coefficient of determination**, state “(r2 x 100)% of the variation in the response variable can be explained by the variation in the explanatory variable.”

When using the least squares regression line to **make predictions**, it is either interpolation or extrapolation.

**Interpolation:** prediction made within the original data

**Extrapolation:** prediction made outside the original data therefore less reliable.

**Residual value** is the difference between the actual y value - predicted y value. Residuals can be graphed with the explanatory variable on the x axis and the residual values on the y axis. It will have 2 quadrants (positive and negative) **Randomly scattered residuals = probably linear data**

**CHAPTER 3 – SEQUENCES – RECURSION**

**Recursive rules**

Rules can be described in words, for example the sequence:
5,8,11,14… - Each term is obtained by adding 3 to the previous term.

\*When writing a recursive rule (numerically) you must include the first term!!!!

So the recursive rule for the sequence above is:
Tn+1 = Tn + 3 where T1 = 5

This reads: To find any term, take the previous term and +3, starting at 5

**Eg.** Write a recursive rule for the sequence 7,1,-5,-11
Tn+1 = Tn – 6 where T1 = 7

**Eg.** Find the first 5 terms of the sequence defined recursively as Tn+1 = 2n x Tn where T1 = 5

5,10,40,320
**Tn will always equal the ‘previous’ term. But single ‘n’ needs to be the number you must add to 1 to get the number of the term you’re trying to find**

**CHAPTER 4 – SEQUENCES – SOME SPECIFIC TYPES**

**Arithmetic –** common difference (+ or -)

a = first term, d = common difference
Recursively, Tn+1 = Tn + d where T1 = a

The general formula for finding any term in an **arithmetic** sequence is: **Tn = a + (n-1)d**

**Geometric –** common ratio (times)

a = first term, r = common ratio

Recursively, Tn+1 = Tn x r where T1 = a

The general formula for finding any term in a **geometric** sequence is: **Tn = ar(n-1)**

**First order** linear recurrence equation:
Tn+1 = b x Tn + c where T1 = a

**Second order** linear recurrence equation:
Tn+2 = b x Tn+1 + c x Tn + d where T1 = a

**Examples:**

**1.** The 6th term of an arithmetic progression is double the 4th term. The first term is 20. Write a recursive formula.

\*When you see the 6th term, think a+5d

a+5d = 2(a+3d)
20+5d = 40+6d
d = -20
Therefore, recursively: Tn+1 = Tn – 20 where T1 = 20

**2.** Tn = 2n – Tn-1 where T1 = 4. Find T2 and T3
T2 = 2x2 – 4 = 0

T3 = 2x3 – 0 – 6

**CHAPTER 5 – NETWORKS**

Made up of vertices, edges and faces. \*Outer area = 1 face

- 2 vertices connected by one or more edges are **adjacent**

- When 2+ edges connect the same pair of vertices, the graph is said to have **multiple edges**

- A vertex not connected to any other vertex is **isolated**

TYPES OF GRAPHS

**Simple:** No loops or multiple edges

**Connected:** No isolated, every vertex reachable from other

**Complete:** Every vertex connected to every other vertex

**Directed:** Edges are directed (arrows)

**Bipartite:** Vertices split into groups so that each edge connects each vertex in the first group with one in the second

DEGREE AND ADJACENCY

**Degree** is the number of edges coming out of a vertex with loops counted twice. Degree sum = 2 x no. of edges

**Adjacency matrix:** Loops only count as 1

PLANAR GRAPHS AND EULER’S FORMULA

**Planar graph:** Can be redrawn so that no edges cross over

**For a planar graph**, **Euler’s formula applies**:
V = E – F + 2

F = E – V + 2

E = V + F – 2

WALKS, PATHS & TRAILS

**Walk:** Can include repeated edges and vertices. **Open walk** = start and finish at different vertices. **Closed walk** = start and finish same vertex

**Path:** All edges and vertices are different. **Open path** = starts and finishes at different vertices, **closed path/cycle** = starts and finishes same vertices

**Trail:** You can only travel along an edge once. Open/closed

**Bridge:** Keeps graph connected

EULERIAN & SEMI-EULERIAN

**Eulerian:** All vertices must be of even degree, start and finish same vertex

**Semi-Eulerian:** Only 2 vertices can be of odd degree, start and finish at the vertices of odd degree

*Both can include repeated vertices not edges***Traversable:** If a network is either Eulerian or semi-Eulerian, it is traversable.

HAMILTONIAN PATHS AND CYCLES

**Hamiltonian path:** Includes every vertex once only, starts and finishes at different vertices.

**Hamiltonian cycle:** Includes every vertex once only, starts and finishes at the same vertex

**UNIT 4 CHAPTER 1 – TIMES SERIES DATA**

In times series, the explanatory variable is **time** and it is plotted on the x axis.

- Time series plots aim to investigate how a variable changes with time

To plot a time series graph on the classpad:
Calc 🡪 2 variables. Time in list 1 (numbers 1…)

Setgraph: xyLine

- The line of best fit can be found using linear regression

**Describing time series plots**

**Trend:**
- Long term increase or decrease in the dependent variable as time

passes

POSITIVE/NEGATIVE:
Overall positive movement = positive secular trend

Overall negative movement = negative secular trend

SEASONAL: The actual seasons of the year affect the dependent variable. There will be a pattern of regular peaks and troughs at corresponding months

CYCLIC: Significant peaks and troughs occur at irregular intervals and are unpredictable

RANDOM: Unpredictable and no significant peaks. There are random fluctuations at a relatively stable mean

**CHAPTER 2 – MOVING AVERAGES AND SEASONAL EFFECTS**Smoothing times series is used to level the fluctuations and show a clearer picture of an overall trend.

**Moving averages**Odd numbers: the new average value corresponds with the middle t value

- When plotting the moving average data, make sure you start and end with the right t values

- A line of best fit for the moving averages can be found

Even numbers: If we calculate the moving averages of even numbers, the new values don’t correspond to any of the t values. We therefore have to calculate the **centered moving averages.**

- This can be done by calculating the moving averages and then averaging out each pair

- OR “half of the first, half of the last and the middle three divided by 4” (for a 4 point centered moving average)

**Quantifying the seasonal effect**

- Calculating the mean for certain periods of time

- Expressing each t value as a percentage of the mean

**Seasonal index:**

- All corresponding percentages averaged out gives the seasonal index

- The average of the seasonal indices should = 100%

**Explaining seasonal index:**- If above 100, the value tends to be ‘\_%’ above the average

- If below 100, the value tends to be ‘\_%’ below the average

**Deseasonalising or seasonally adjusting data**

- The original data includes a ‘\_%’ seasonal increase or decrease

- Dividing each value from the original data by its corresponding seasonal index as a decimal gives the seasonally adjusted value

**Making predictions**

If we use a line of best fit for the moving averages or deseasonalised values, any prediction made will be a prediction that has been deseasonalised/smoothed

- In order for the predictions to be ‘real life values’ or more like the original data, we need to factor back in the seasonal effect.

- Eg. If we determine a line of best fit using deseasonalised values, we will be able to calculate a predicted value. However, for it to fit in with the original data we should multiply it by its corresponding seasonal index.

\* Deseasonalised data = real data/seasonal index

Therefore, Real data = deseasonalised data x seasonal index

**CHAPTER 3 – FINANCE I – SAVING AND BORROWING**

**SIMPLE INTEREST:** I = PRT

**Eg.** *An investor wishes to see her initial investment grow to $64000 interest over a five-year period from an account that earns 12% p.a simple interest. What was her initial investment?*P + I = 64000

I = 64000 – P
Therefore, 64000 – P = P x 0.12 x 5
P = $40 000

**COMPOUND INTEREST:** A = P(1 + $\frac{R}{n}$ )nt
A = Principal + interest

**Eg.** *At an 8% annual compound interest rate, with compounding every six months, how many years would it take for an initial investment of $2500 to grow to $3700?*P = 2500
r = 0.08

n = 2
Therefore, 3700 = 2500 x (1 + $\frac{0.08}{2}$ )(t x 2), t = approx 5 years
**EFFECTIVE ANNUAL INTEREST RATE**

Effective annual interest rate of ‘i’ per annum (as a decimal) and ‘n’ compounding periods per year,
ieffective = (1 + $\frac{i}{n}$ )n – 1

**COMPOUND INTEREST INVESTMENTS WITH REGULAR DEPOSITS**

**Annuity investment:** Initial deposit + regular payments while earning fixed rate of compound interest

**Eg.** *At the start of each year, Ben deposits $1200 into an account earning 8% p.a. interest compounded annually. Interest is paid at the end of the year before his next deposit.
Write a recurrence relation that gives the balance at the end of year n, Bn.*

Bn = 1.08 x (Bn-1 + 1200), B0 = 1200

**Eg.** *At the beginning of one month, $100 is invested into an account paying interest at 7.5% p.a. compounded monthly, and an extra $100 is invested at the end of that first month and the end of every month thereafter. How much is the account worth at the end of the month two years later, just after the $100 deposit for that month is made?*

- Compounded monthly therefore change rate into monthly (7.5/12)
= 0.625
- Recursively: Bn = (1.00625 x Bn-1) + 100, B0 = 100
- After 2 years = B24, = $2696.80

- Can also be easily done in financial
**DEPRECIATION**

The reduction in the value of an item as it ages over a period of time

A = P(1-R)n

**LOAN REPAYMENTS**

**Eg.** *Ben takes out a loan of $7000. Interest of 9.4% p.a. is added to the loan annually and Ben repays $800 at the end of each one year period. How much will Ben still owe immediately after he makes the $800 repayment at the end of year 8?*

Recursively: Bn = 1.094 x Bn-1 – 800, B0 = 7000
Therefore, after the $800 repayment at the end of year 8, Ben will still owe $5411.09

**CHANGING THE TERMS OF A LOAN**

**Eg.** *Ben took out a reducing balance loan for $60 000 with interest calculated monthly at a rate of 10.25% p.a. He makes monthly repayments of $918.41*

*a) How long will it take him to repay the loan?*

FINANCIAL:

N = 95.999 therefore 96 months = 8 years.

*b) After 3 years, he increases his monthly repayment to $1200. By how many months is the length of the loan reduced?*- 3 years (36 months) of $918.41 repayments, FV (still owes) $42 975.73
- That then becomes the new PV, and payments change to $1200. FV = 0, and N now = 43 months.
- Therefore, life of loan is now (36 + 43 = 79) instead of 96 months
- Therefore it reduced by 17 months

*c) Find the amount saved in interest by increasing the size of the monthly repayment*

Amount originally paid: 96 x 918.41 = 88 167.36

Amount paid after increasing payments: (36 x 918.41) + (43 x 1200) = 84 662.76
Therefore amount in interest saved = 88 167.36 – 84 662.76
= $3504.60

**CHAPTER 4 – FINANCE II – DRAWING DOWN THE INVESTMENT**

**SUPERANNUATION** allows people to accumulate funds to provide them with income after retirement

**Eg.** *Ben wishes to deposit 200 000 into an account that will earn interest if 8% p.a. compounded annually, withdrawing 25 000 at the end of the first year, 25 500 at the end of the second year, 26 010 at the end of the third year etc, with each annual withdrawal being a 2% increase on the last one.*

Recursively: Tn+1 = 1.08 x Tn – 1.02n x 25 000, T0 = 200 000

**PERPETUITIES** Means a constant amount is taken out each year, leaving the balance at the same amount as well

**Eg.** *What initial investment is required into an account that pays annual interest of 6% compounded monthly, if the interest earned each year is to pay for an annual perpetual award of $15 000?*- Change interest rate to monthly: 6/12 = 0.5

- 1.00512 = 1.061677812
- Therefore there will be an increase of 6.1677812% on the initial loan, A

- A x 0.061677812 = 15 000

- A = 243 199.29

\* To the power of n \*

\* If the compounding occurs quarterly, monthly etc, change annual rate to a different rate\*
\* If payments occur more frequently than compounding, adjust the compounding period to match the payment period. **Eg.** Payments occurring quarterly but compounding occurring annually, interest rate of 8% p.a.

Therefore, x4 = 1.08, so fourth root of 1.08 = 1.019426547

**Eg. Compounding occurs more frequently than payments**

*Suppose that $250 000 is invested into an account paying 8% interest p.a. which compounds quarterly, with $40 000 withdrawn from the account at the end of every year. Determine the balance of the account immediately after the 6th withdrawal*

**Recursively:** Tn+1 = Tn x 1.024 – 40 000, T0 = 250 000 and then find T6

**Financial:**
N = 6
I = 8
PV = 250 000

PMT = - 40 000

**FV (ANS) = $106 866.65**
P/Y = 1

C/Y = 4

**Eg. Payments occur more frequently than compoundings**

*Suppose that $250 000 is invested into an account paying 8% interest p.a, compounded annually, with $10 000 withdrawn from the account every 3 months (quarterly)*

*Determine the balance of the account after 3 years*

**Recursively:** Rate = 8% per annum

Ratio = 1.08

Fourth root of 1.08 = 1.019426547

Therefore, Tn+1 = Tn x 1.019426547 – 10 000, so T12 = $181 238.78

**Financial:**N = 12 (3 years x 4 withdrawals)
I = 8

PV = 250 000

PMT = - 10 000

**FV** **(ANS) = $181 238.78**
P/Y = 4

C/Y = 1

**CHAPTER 5 – MINIMUM SPANNING TREES**

- Minimum spanning trees to not need to connect every vertex to every other vertex
- Every vertex must be reachable in some way from others

****- No loops

**Eg.**

Minimum spanning tree = 188 units

- Prim’s algorithm

- Kruskal’s algorithm (shortest edges)

**Minimum spanning tree from distance table**

Using Prim’s algorithm

- Select a vertex, delete its row and mark its column.
- Scan all marked columns for the lowest non-zero entry and circle that entry. If there
 is a tie, pick an entry at random.
- Delete the row containing the circled entry and mark its column.
- Repeat steps 2 and 3until all rows in the matrix are deleted.

**CHAPTER 6 – MAXIMUM FLOW
Eg.**

*Determine the maximum flow possible with point A as the source and point F as the sink*

- Work through all combinations from top to bottom, crossing out the smallest value and subtracting it from all other values in the combination

- Write down all of the initial minimum values, and at the end, add them all together to find the maximum flow

*Using the systematic approach…*

ABEF: 20

ABCEF: 10

ACEF: 10

ACF: 10

ACDF: 5

ADF: 25

*Re-draw the network showing the true flow*

- Go through working: wherever there is a 0, write the original

 value, wherever there is left over, subtract it from original and

 write the answer

 **Maximum flow = Minimum cut**

We can check that we have found the maximum flow by considering ‘cuts’

- Draw vertical lines through the source and the sink

- Draw cuts that go through arcs from above the network to below, working through left to right

- The smallest cut = maximum flow

\***Cuts that must be horizontal (towards the source) aren’t included in the value of the cut**

**CHAPTER 7 – PROJECT NETWORKS**

***Any task must be completed before another can commence***EXAMPLE 1



*Find the minimum completion time and the critical path for the project network shown below…*

**Minimum completion time** = 32 hours

**Critical path =** P – Q – V

**Slack time for S:**-Task S is ready to start after
6 hours and doesn’t need to
be finished until after 22
hours.

- So, we can allow
up to 16 hours for
the task, which means
there is room for (16-12)

= 4 hours’ slack time

EXAMPLE 2



**EARLIEST STARTING TIME:**
- EST for the starting activities is 0
- EST for the other activities: add the EST from the previous activity to the activity duration (e.g. EST of activity H is 12 as EST of activity E is 9 and duration for activity E is 3, hence 9 + 3 = 12).
- If there are multiple activities feeding into another activity (such as Activity E whose predecessors are activities B, C and D), choose the highest duration of those activities (i.e. Activity D with duration of 9).
**LATEST STARTING TIME:**
Backwards scanning:
- Set LST equal to the EST of the finishing time.
- Using the LST of the finish time, work backwards through the network by subtracting the activity duration from the LST of the previous activity (e.g. duration of Activity J is 3 and LST of Activity J is 26 − 3 = 23). If there are multiple activities feeding into another activity (such as Activity E whose predecessors are activities G, H and I), choose the lowest LST’s of those activities to subtract from (i.e. Activity G with LST of 12).
\* If there are 2 ways, choose the smallest number to be the LST
**SLACK TIME = LST – EST**- If slack time = 0, it is on the critical path
- *By how much can activity H be lengthened without changing critical path?*ANS = 16 – 12 = 4 hours

**CHAPTER 8 – ASSIGNMENT PROBLEMS**- Can be represented using a bipartite graph or a table/matrix- Can be asked to find minimum or maximum cost**Systematic approach: Hungarian Algorithm
Minimum:**- Put all values into a matrix
- Work across the rows, finding the smallest value and subtracting it from every other value in the row
- Once you have done all the rows, do the same with all the columns: find the smallest and subtract from every other value in the column
- Find a combination of 0s that work
**Maximum:**- Find the largest value out of them all and subtract each value from it to create a new matrix
- Start with rows and work through the same process as before **ADJUSTING:**- If number of lines drawn through zeros does not equal the number of options, you need to adjust.
- Find the smallest uncovered number and subtract it from each other uncovered number and add it to where two lines intersect

\* If Kn is a complete graph with n vertices, the total number of edges is:
 $\frac{n(n-1)}{2}$

\* Don’t connect the dots when plotting sequences