

A Brief Summary of week 3

1. Presentation of datasets in the form of a Table

- Each attribute/field is a column in the table (like: Card No., Gender, Total, etc)
- Each card is a row in the table.
-

Card No	Name	Gender	DOB	---
0	Bhuvanesh	M	7 Nov	
1	Harish	M	3 June	

- Example: Sores dataset, Words dataset.

<p>Name Bhuvanesh 0</p> <p>Gender M</p> <p>Date of Birth 7 Nov</p> <p>Town/City Erode</p> <table> <tr> <td>Mathematics</td> <td>68</td> </tr> <tr> <td>Physics</td> <td>64</td> </tr> <tr> <td>Chemistry</td> <td>78</td> </tr> <tr> <td>TOTAL</td> <td>210</td> </tr> </table>	Mathematics	68	Physics	64	Chemistry	78	TOTAL	210	<p>Name Harish 1</p> <p>Gender M</p> <p>Date of Birth 3 June</p> <p>Town/City Salem</p> <table> <tr> <td>Mathematics</td> <td>62</td> </tr> <tr> <td>Physics</td> <td>45</td> </tr> <tr> <td>Chemistry</td> <td>91</td> </tr> <tr> <td>TOTAL</td> <td>198</td> </tr> </table>	Mathematics	62	Physics	45	Chemistry	91	TOTAL	198
Mathematics	68																
Physics	64																
Chemistry	78																
TOTAL	210																
Mathematics	62																
Physics	45																
Chemistry	91																
TOTAL	198																

- Difficultify when a card has a variable number of attributes
Multiple rows:- duplication of data
Split as separate tables:- need to link via unique attributes
- Example: Shopping bill card

SV Stores		Srivatsan 1			
Item	Category	Qty	Price	Cost	
Carrots	Vegetables/Food	1.5	50	75	
Soap	Toiletries	4	32	128	
Tomatoes	Vegetables/Food	2	40	80	
Bananas	Vegetables/Food	8	8	64	
Socks	Footwear/Apparel	3	56	168	
Curd	Dairy/Food	0.5	32	16	
Milk	Dairy/Food	1.5	24	36	
				567	

Big Bazaar		Sudeep 2			
Item	Category	Qty	Price	Cost	
Baked Beans	Canned/Food	1	125	125	
Chicken Wings	Meat/Food	0.5	600	300	
Cocoa powder	Canned/Food	1	160	160	
Capsicum	Vegetables/Food	0.8	180	144	
Tie	Apparel	2	390	780	
Clips	Household	0.5	32	16	
				1525	

2. Below average students in two iterations (non-nested) and grade allocation

- First we compute average marks of some subject (1st iteration)
- And then in next iteration we compare each student marks with average marks of that particular subject

Eg:

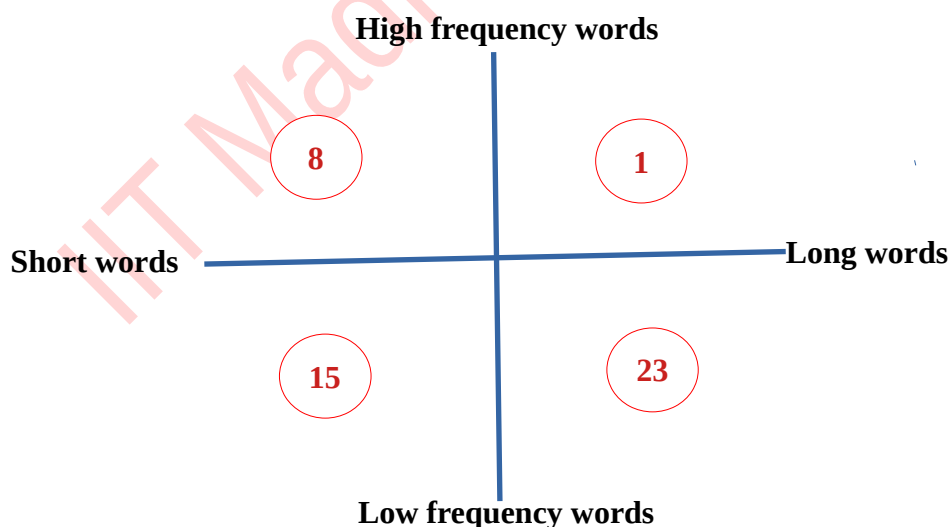
```
while(-----){  
    sum = sum + X.Maths  
    count = count + 1  
}  
avg = sum/count  
while(-----)  
    if(not (X.Maths < avg)) ---- (not True) --- (False)  
        avgS = avgS + 1  
    }  
}
```

Grade allocation: Find maximum and minimum marks of a subject
: Divide into grade bands and
: Count the numbers of students per grade bands

3. Systematic process of hypothesis verification

Example: To find the relation between word length and frequency

- : Assumption – Long words ≥ 6 and Short words ≤ 5
- High frequency > 1 and Low frequency = 1



Hypothesis: (i) High frequency words tend to be short.

: (Words which have high frequency and are short) / (Total high frequency words)
 $= 8 / (8 + 1) = 8 / 9 = 0.89$

(ii) Low frequency words tend to be long?

4. Three prize problem

For getting prizes, it should satisfy three conditions:

- (i) Should be in top three of total marks
- (ii) Should be in a top three in at least one subjects
- (iii) And there should be a representation of boys and girls

5. Introduction to procedures and parameters

- Procedure name: **SumMaths**
- Argument receives value: **gen**
- Calling procedure with a parameter: **SumMaths (F)**
- Argument variable is assigned parameter value
- Procedure call **SumMaths(F)** starts with, **gen = F**
- Procedure returns the value stored in **Sum** using **return(Sum)**

```
Procedure SumMaths(gen)
    Sum = 0
    while (Pile 1 has more cards) {
        Pick a card X from Pile 1
        Move X to Pile 2
        if (X.Gender == gen) {
            Sum = Sum + X.Maths
        }
    }
    return(Sum)
end SumMaths
```

Interface and Implementation:

Interface: Parameters to be passed
: Value to returned
: What side effects are possible

Implementation: Procedure definition

Eg:

```
Procedure pro1 (a, b)
    return( 0.5*(a+b))
End proc1
```

```
Procedure pro2 ( a, b)
    c = a + b
    c = 0.5*c
    return(c)
End proc2
```

Truth Tables:

(i) AND Table

A	B	A AND B
True	True	True
True	False	False
False	True	False
False	False	False

(ii) OR Table

A	B	A OR B
True	True	True
True	False	True
False	True	True
False	False	False