Shri Vile Parle Kelavani Mandal's
DWARKADAS J. SANGHVI COLLEGE OF ENGINEERING
(Autonomous College Affiliated to the University of Mumbai)
NAAC Accredited with "A" Grade (CGPA : 3.18)

A.Y. 2022-2023

DATA MINING AND WAREHOUSE

AYUSH JAIN

COMPUTER ENGINEERING | TE – B2 | 60004200132

EXPERIMENT – 7

Aim: Perform OLAP operations such as Roll up, Drill down, Slice and Dice, Pivot on Datawarehouse.

Theory:

OLAP is an acronym for On Line Analytical Processing**.** Online Analytical Processing: An OLAP system manages large amount of historical data, provides Facilities for summarization and aggregation, and stores and manages information at different levels of granularity.

OLAP operations:

Slice: A slice is a subset of a multi-dimensional array corresponding to a single value for one or more members of the dimensions not in the subset.

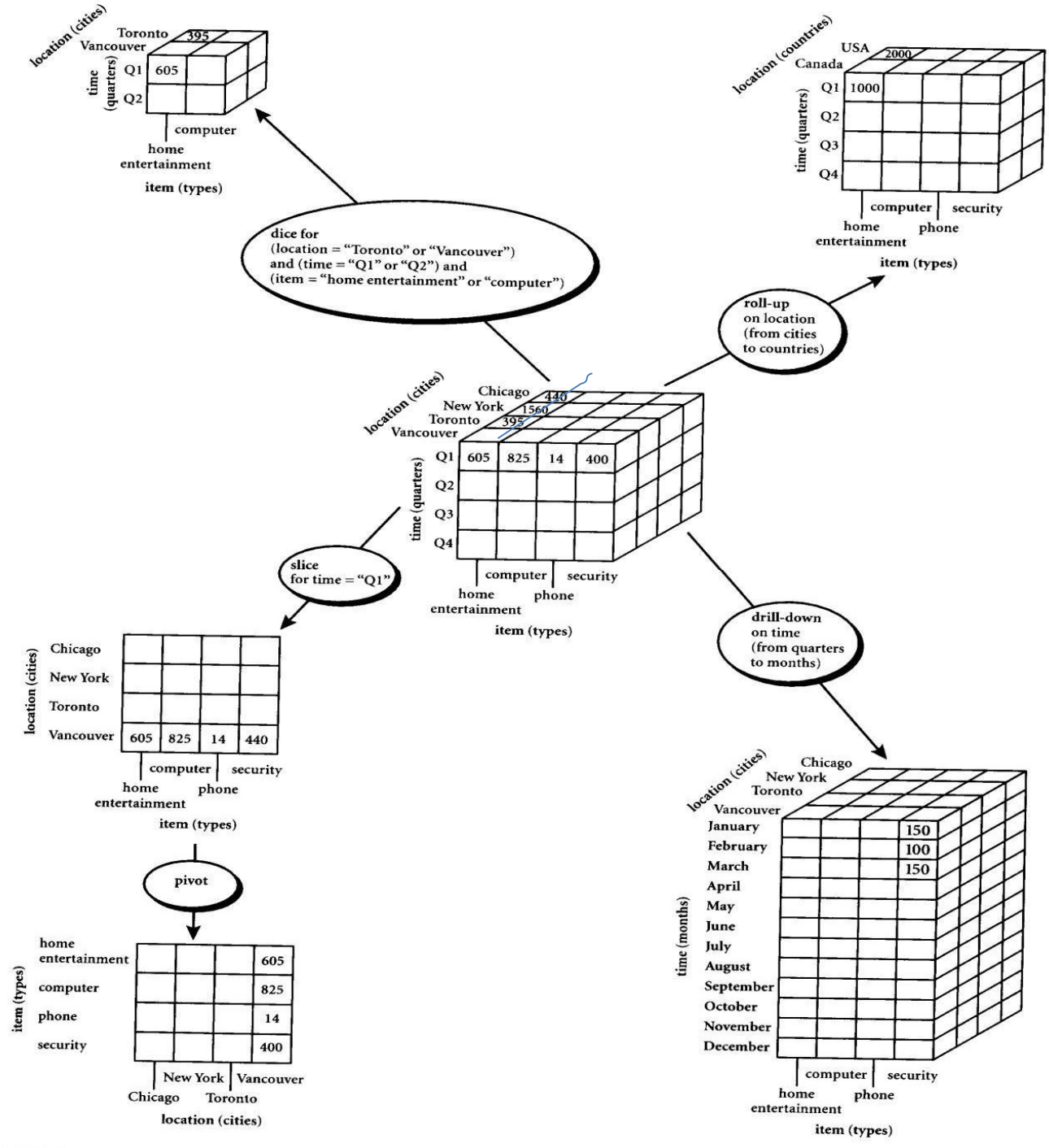
Dice: The dice operation is a slice on more than two dimensions of a data cube (or more than two consecutive slices).

Drill Down/Up: Drilling down or up is a specific analytical technique whereby the user navigates among levels of data ranging from the most summarized (up) to the most detailed (down).

Roll-up: A roll-up involves computing all of the data relationships for one or more dimensions.
To do this, a computational relationship or formula might be defined.

Pivot: To change the dimensional orientation of a report or page display.

A.Y. 2022-2023


location (cities)
Toronto
Vancouver
395
time (quarters)
Q1
605
Q2
computer
home entertainment
item (types)
dice for
(location = "Toronto" or "Vancouver")
and (time = "Q1" or "Q2") and
(item = "home entertainment" or "computer")
location (countries)
USA
2000
Canada
Q1
1000
Q2
Q3
Q4
computer
security
home entertainment
phone
item (types)
roll-up
on location
(from cities
to countries)
location (cities)
Chicago
440
New York
1560
Toronto
395
Vancouver
Q1
605
825
14
400
Q2
Q3
Q4
time (quarters)
computer
security
home entertainment
phone
item (types)
slice
for time = "Q1"
drill-down
on time
(from quarters
to months)
location (cities)
Chicago
New York
Toronto
Vancouver
605
825
14
440
computer
security
home entertainment
phone
item (types)
pivot
item (types)
home entertainment
computer
phone
security
605
825
14
400
Chicago
New York
Toronto
Vancouver
location (cities)
location (cities)
Chicago
New York
Toronto
Vancouver
January
February
March
April
May
June
July
August
September
October
November
December
150
100
150
time (months)
computer
security
home entertainment
phone
item (types)

EXERCISE 1

Consider a data Warehouse for a hospital, where there are three dimensions:

(i) Doctor
(ii) Patient
(iii) Time

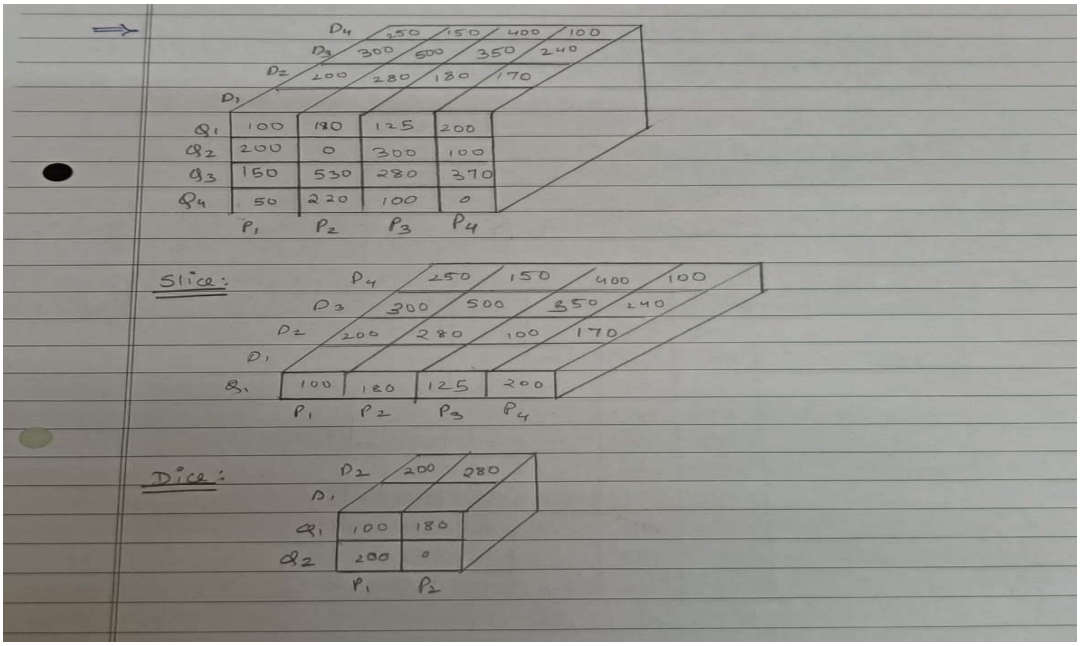
With two measures
(a) Count
(b) Charge

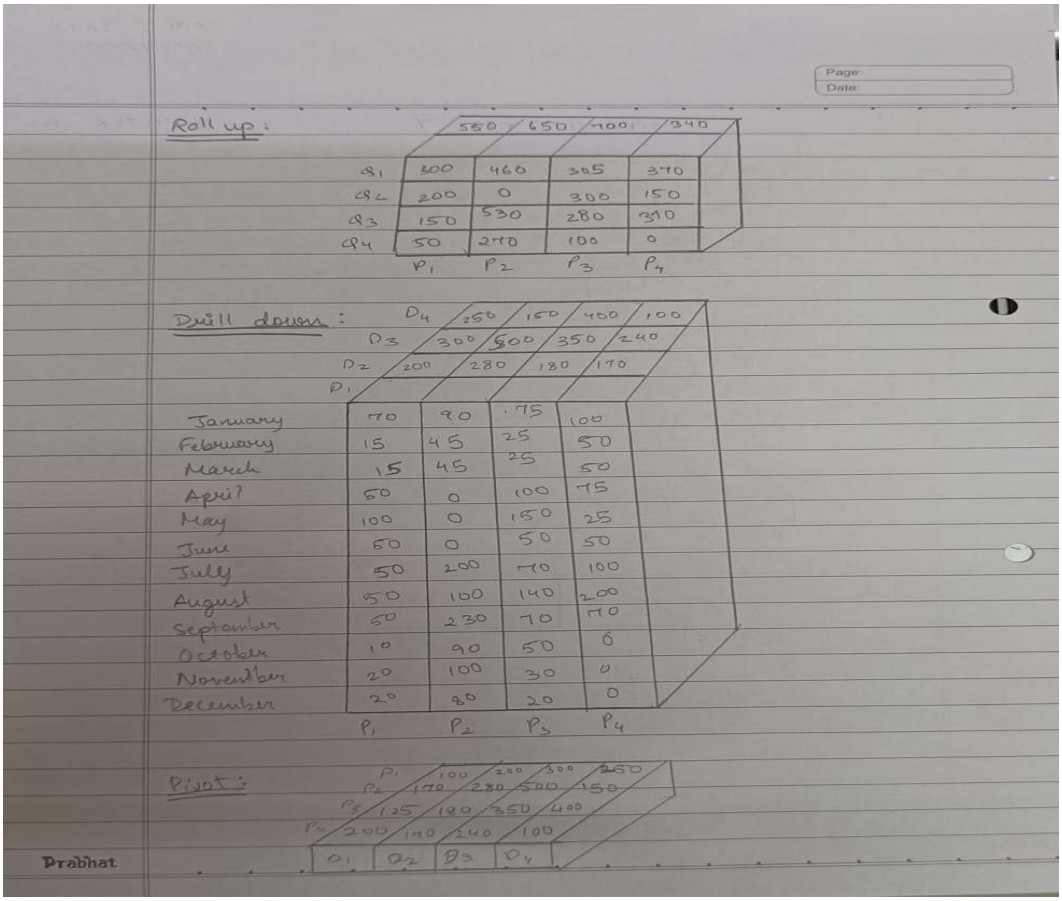
Where Charge is the fee that the Doctor charges a patient for a visit.

Using the above example describe the following operations:

(i) Slice
(ii) Dice
(iii) Roll Up
(iv) Drill Down
(v) Pivot

NOTE: Assume data according to the dimensions and measures and explore individual tasks diagrammatically.


Slice:
Dice:


Roll up:
Drill down:
Pivot:
Prabhat

EXERCISE 2

To create Pivot of Table using MS Excel

Follow these

steps ... 1.

Start with M.S

Excel.

2. In excel sheet create 4 columns PRODUCT, ORIGIN, DAY OF SALE, SOLD UNITS (FACT COLUMN).
3. Insert around fifty rows of data.
4. Save the table data.
5. Go to Insert Tab-> click on Pivot Table-> New work sheet-> Ok.
6. Right side you will find pivot table fields.
   It contains all columns of our table that we created.
   Select product in

   rows, Days in

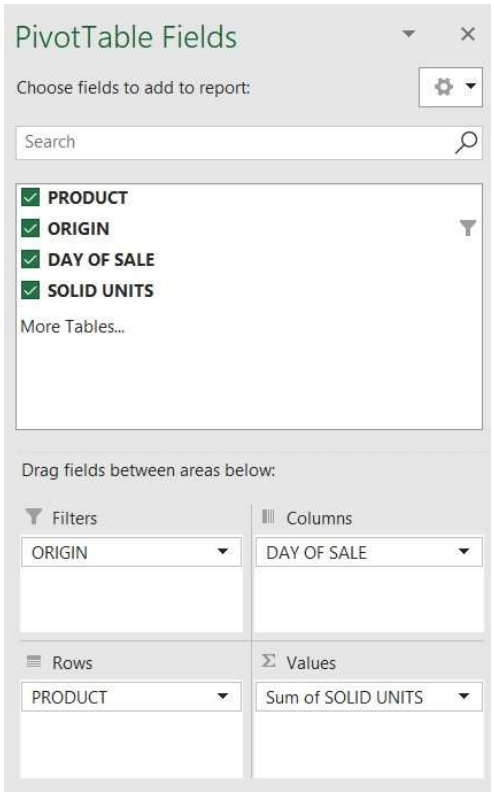
   column,

   Unit sold in ∑ values.
   Later apply filter using Origin.
   Also we can flip the rows & columns or combine together as rows

   only to see different views of same data.

   Dataset:

A	B	C	D
PRODUCT	ORIGIN	DAY OF SALE	SOLID UNITS
A	East	01-01-2023	8
B	Central	02-01-2023	4
C	Central	03-01-2023	2
D	Central	04-01-2023	5
E	West	05-01-2023	6
F	East	06-01-2023	9
G	Central	07-01-2023	9
H	Central	08-01-2023	1
I	West	09-01-2023	1
J	East	01-01-2023	8
K	Central	02-01-2023	4
L	East	03-01-2023	6
M	East	04-01-2023	7
N	East	05-01-2023	9
O	Central	06-01-2023	5
P	East	07-01-2023	2
Q	Central	08-01-2023	1
R	East	09-01-2023	3
S	East	01-01-2023	6
T	Central	02-01-2023	5
U	Central	03-01-2023	8
V	East	04-01-2023	4
X	Central	05-01-2023	8
Y	Central	06-01-2023	9
Z	East	07-01-2023	2



ORIGIN	East							
Sum of SOLID UNITS	**Column Labels**							
Row Labels	**01-01-2023**	**03-01-2023**	**04-01-2023**	**05-01-2023**	**06-01-2023**	**07-01-2023**	**09-01-2023**	**Grand Total**
A	8							8
F					9			9
J	8							8
L		6						6
M			7					7
N				9				9
P						2		2
R							3	3
S	6							6
V			4					4
Z						2		2
Grand Total	**22**	**6**	**11**	**9**	**9**	**4**	**3**	**64**