EXPERIMENT NO. 07

Aim: To perform various OLAP operations such as slice, dice, drilldown, rollup, pivot

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:

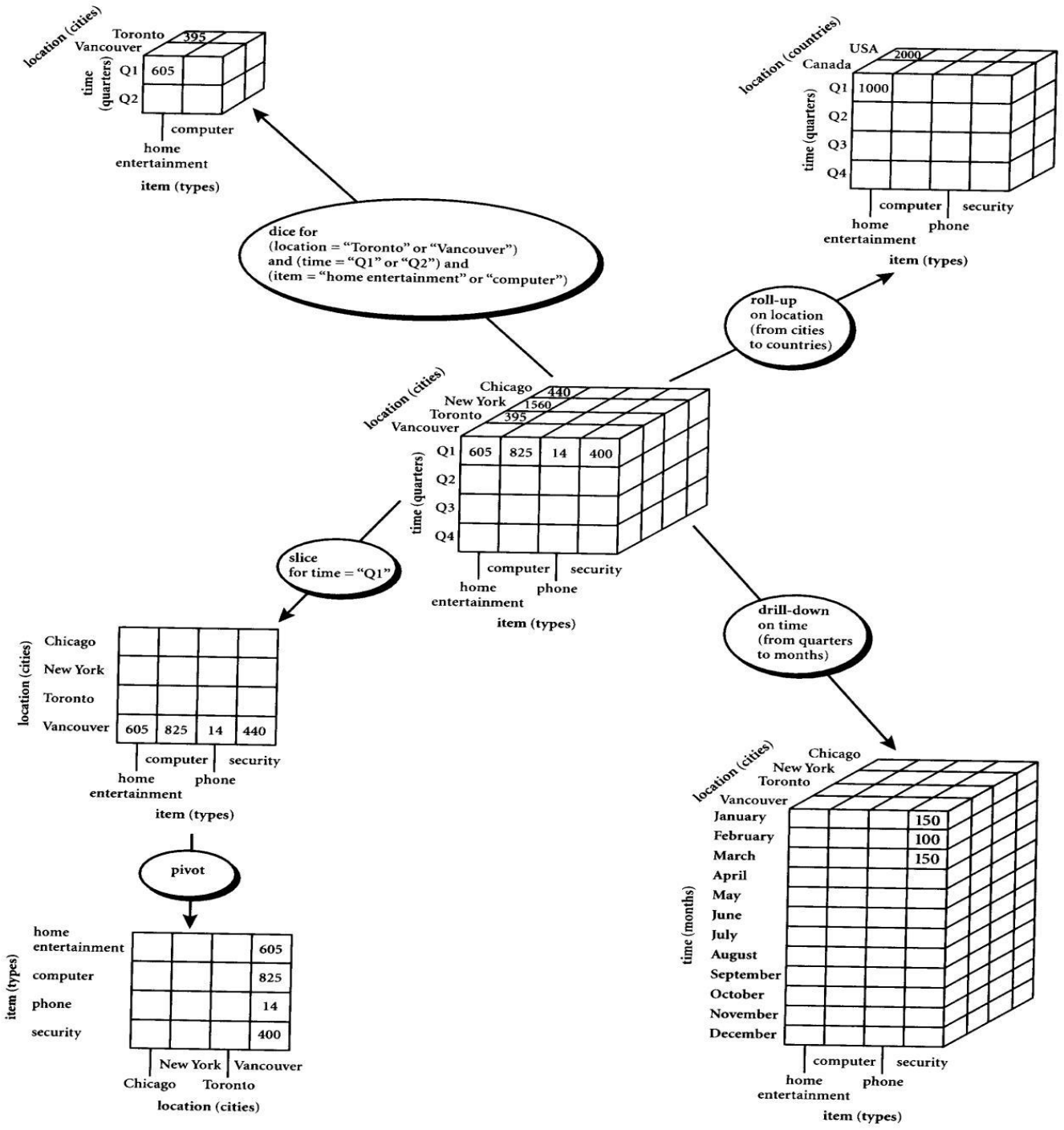
Slice: A slice is a subset of a multi-dimensional array corresponding to a single value for one or more members of the dimensions not in the subset.

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.


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
New York
Vancouver
Chicago
Toronto
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

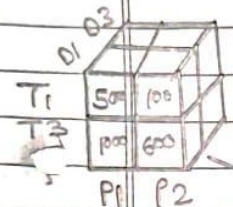
NOTE: Assume data according to the dimensions and measures and explore individual task diagrammatically.

SDS Page No.
Date / /

Exercise 1

Dimensions ~~Fact~~

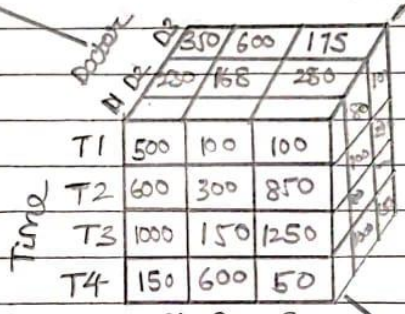
1) Doctor
2) Patient
3) Time



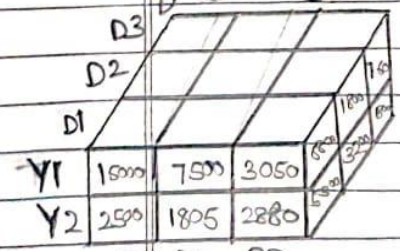
DICE for
(doctor = "D1" or "D3")
and (patient = "P1" or "P2")
and time = "T1" or T3")

	D1	D2	D3
T1	150	230	350
T2	600	830	125
T3	1000	500	375
T4	150	185	450

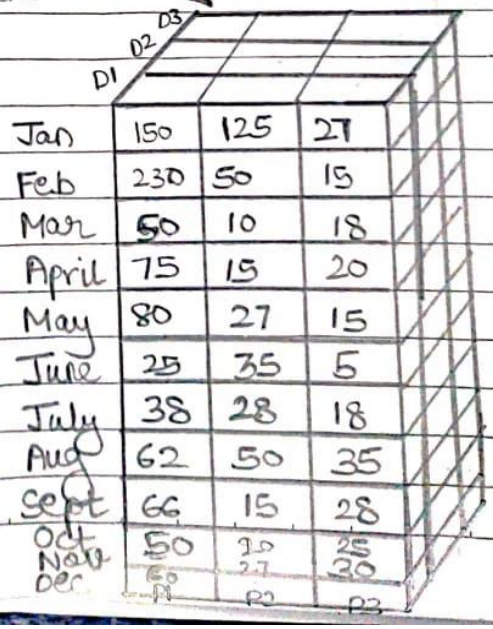
SLICE for
Patient = "P1"



ROLL UP
on time (from quarterly to half yearly)



DRILL DOWN for
on time
(from quarters to months)



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,

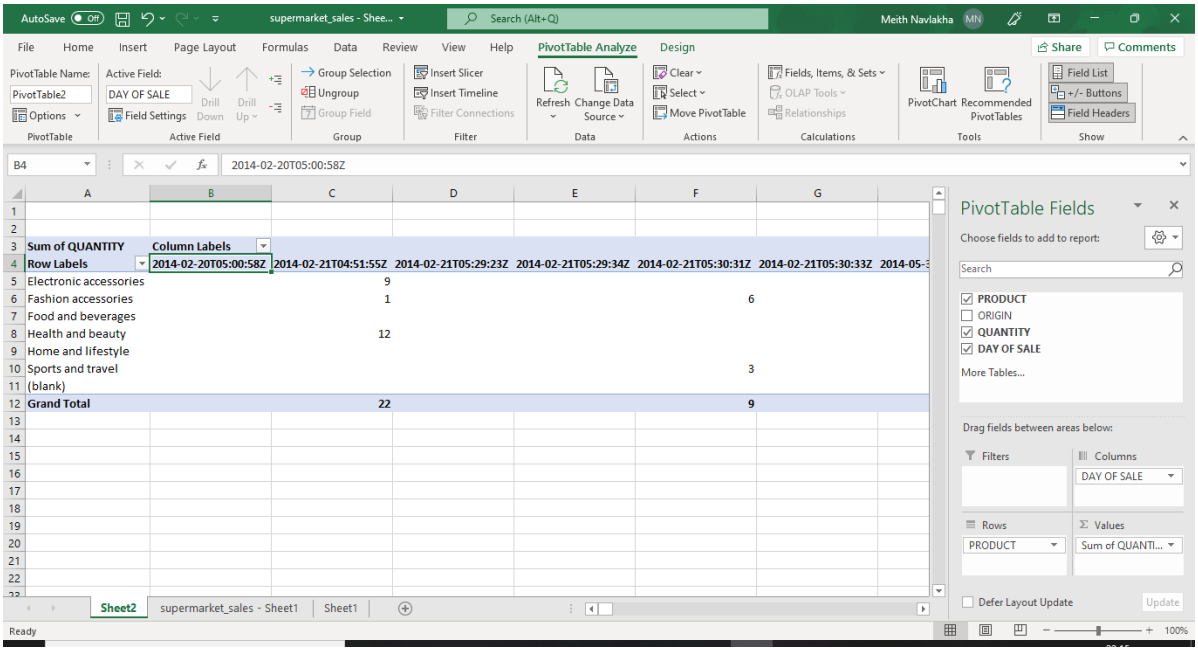
   Days in column,

   Unit sold in ∑ values.

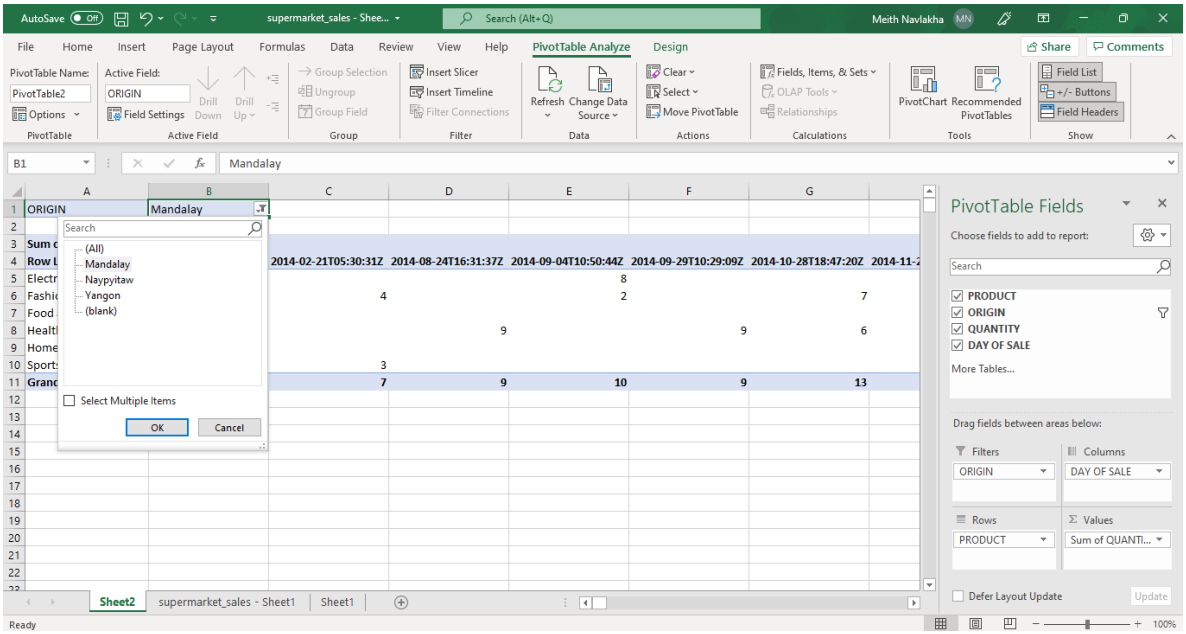
   Later apply filter using Origin.

   Also we can flip the rows & columns or combine together as rows only to see different views of same data.

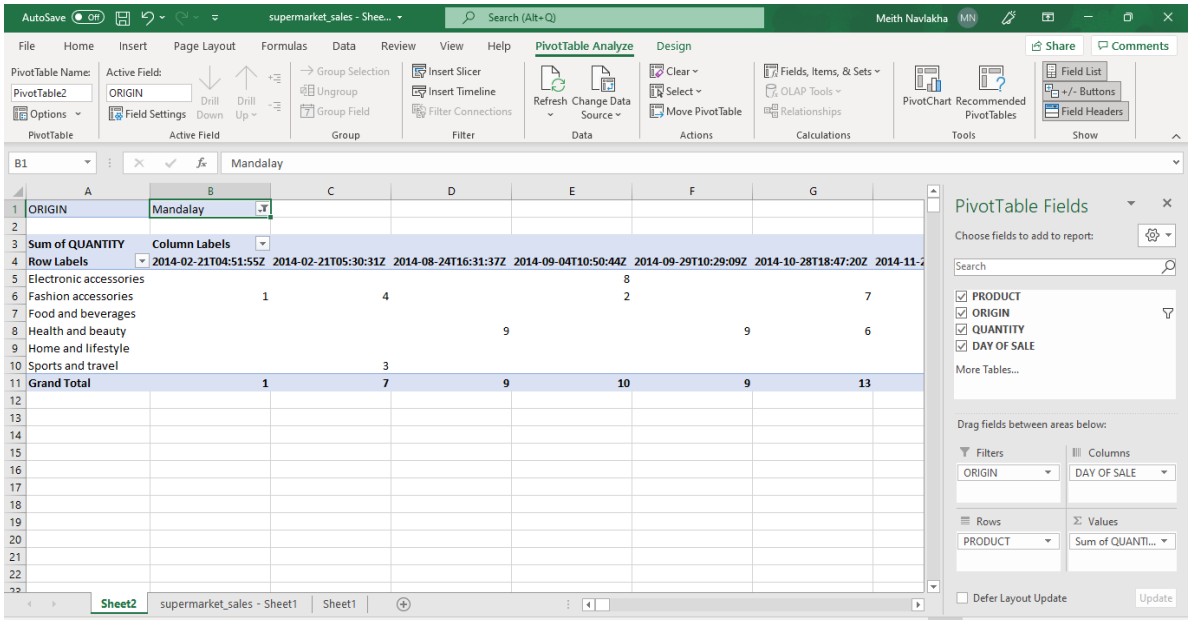
OUTPUT: PIVOT TABLE



Filter by Origin

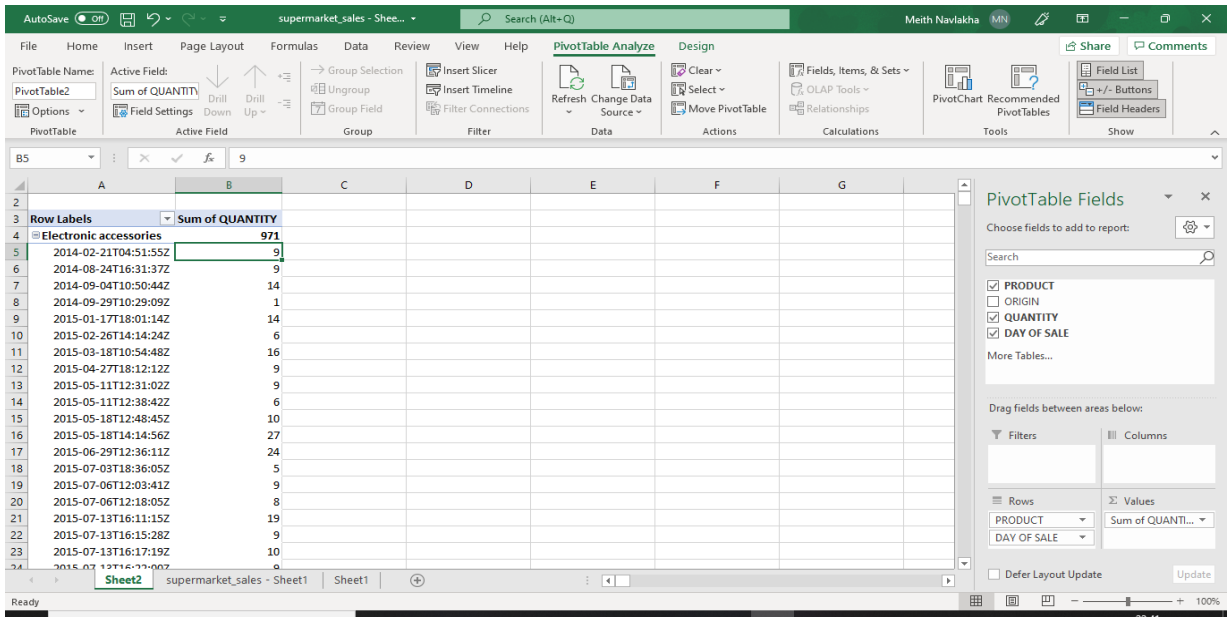


After Applying Filter by origin



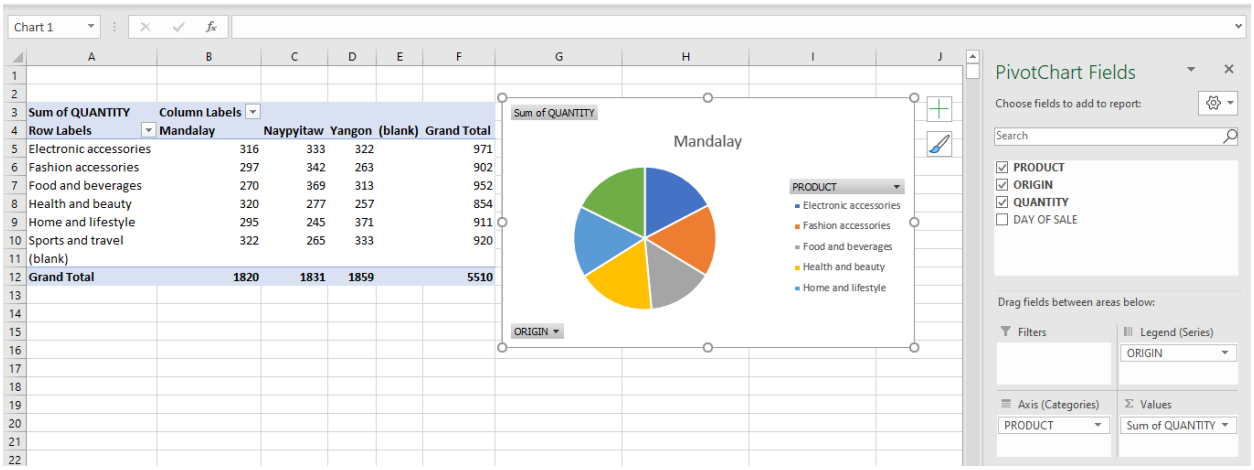
ORIGIN	Mandalay
Sum of QUANTITY	Column Labels
Row Labels	2014-02-21T04:51:55Z	2014-02-21T05:30:31Z	2014-08-24T16:31:37Z	2014-09-04T10:50:44Z	2014-09-29T10:29:09Z	2014-10-28T18:47:20Z
Electronic accessories				8		
Fashion accessories	1	4		2		7
Food and beverages						
Health and beauty			9		9	6
Home and lifestyle						
Sports and travel		3				
Grand Total	1	7	9	10	9	13

Products and Day of Sales on the rows with summation over Quantity sold along with filter by Origin.



Row Labels	Sum of QUANTITY
Electronic accessories	971
2014-02-21T04:51:55Z	9
2014-08-24T16:31:37Z	9
2014-09-04T10:50:44Z	14
2014-09-29T10:29:09Z	1
2015-01-17T18:01:14Z	14
2015-02-26T14:14:24Z	6
2015-03-18T10:54:48Z	16
2015-04-27T18:12:12Z	9
2015-05-11T12:31:02Z	9
2015-05-11T12:38:42Z	6
2015-05-18T12:48:45Z	10
2015-05-18T14:14:56Z	27
2015-06-29T12:36:11Z	24
2015-07-03T18:36:05Z	5
2015-07-06T12:03:41Z	9
2015-07-06T12:18:05Z	8
2015-07-13T16:11:15Z	19
2015-07-13T16:15:28Z	9
2015-07-13T16:17:19Z	10

Distribution of products sold by category with respect to the origin place.



Sum of QUANTITY	Column Labels				
Row Labels	Mandalay	Naypyitaw	Yangon	(blank)	Grand Total
Electronic accessories	316	333	322		971
Fashion accessories	297	342	263		902
Food and beverages	270	369	313		952
Health and beauty	320	277	257		854
Home and lifestyle	295	245	371		911
Sports and travel	322	265	333		920
(blank)					
Grand Total	1820	1831	1859		5510

The above table summarizes the number of units sold by a particular category at different locations as well as the total sales for that category. It also gives insights about the category of the products which are most sold in the particular location. Pie chart for the sales distribution in Mandalay is also being displayed.

CONCLUSION:

In this experiment, I designed the data warehouse for a Hospital and performed OLAP operations such as Roll up by aggregating the quarterly time to yearly, Drill down by expanding to a more granular level i.e. months, Slicing the warehouse with respect to the patient (P1) and Dice operation on the warehouse. In the 2nd exercise I implemented Pivot on M.S Excel on the data warehouse. Through the experiment, I learnt the methods of data visualizing and analysis, how a data warehouse can be expanded or aggregated with respect to different fields and from that, strategic decisions can be formulated.