Vivekanand Education Society's Institute of Technology
Department of Computer Engineering

Subject: -Data Warehouse and Mining Lab

Class:- T.E. (D12) **Semester:- VI** **Div:- A**

Roll No: 21	Name: AMIT JOSHI		
Exp. No: 1	Title: Assignment-1		
DOP:	04/03/2021	DOS:	04/03/2021
GRADE:	LAB OUTCOMES : LO1,LO2	SIGNATURE:	

Amit Joshi 21

CLASSMATE
Date :
Page :

DWM Assignment-1.

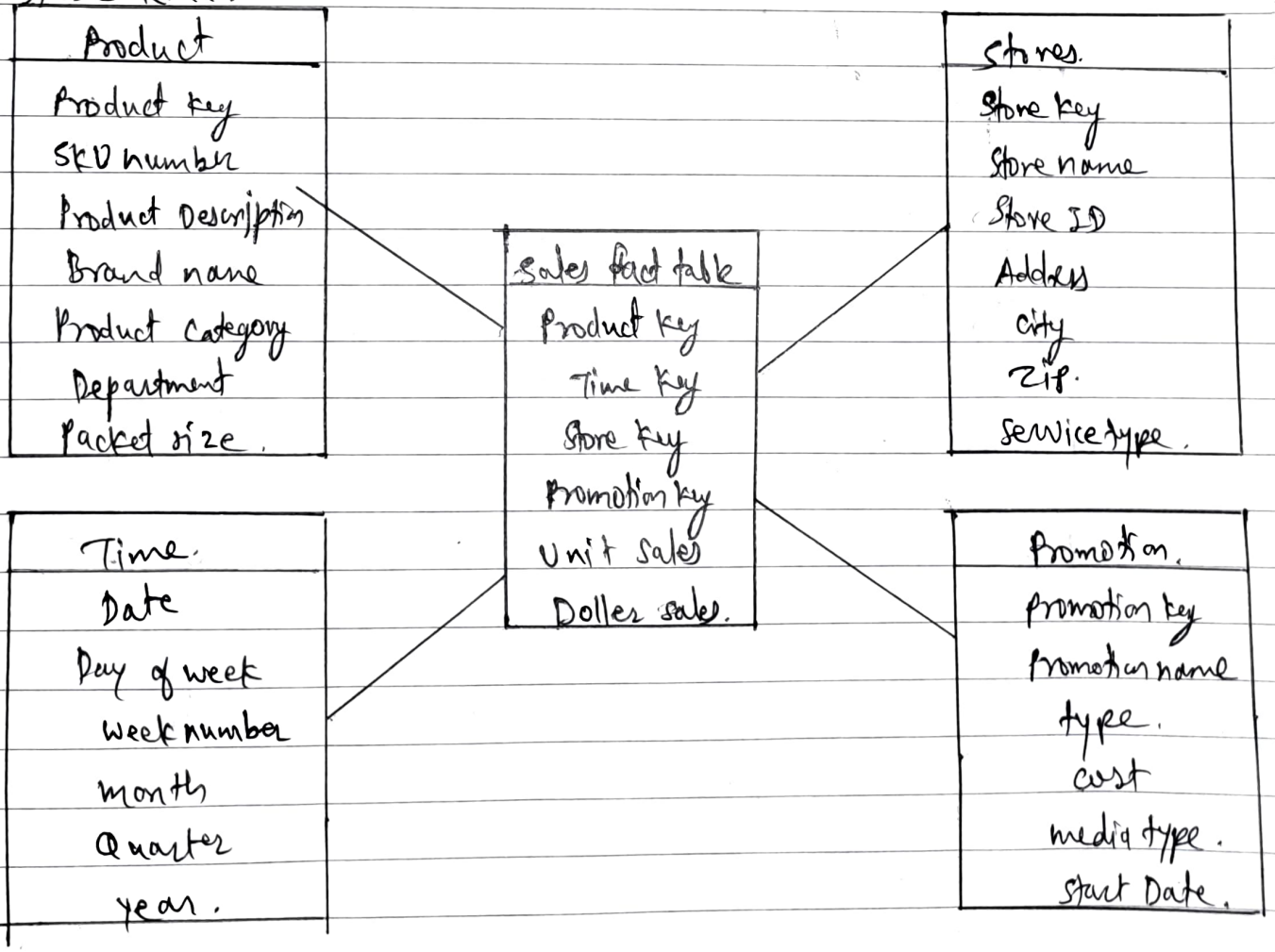
Q1).

Ans: Information Package.

Information subject: sales.

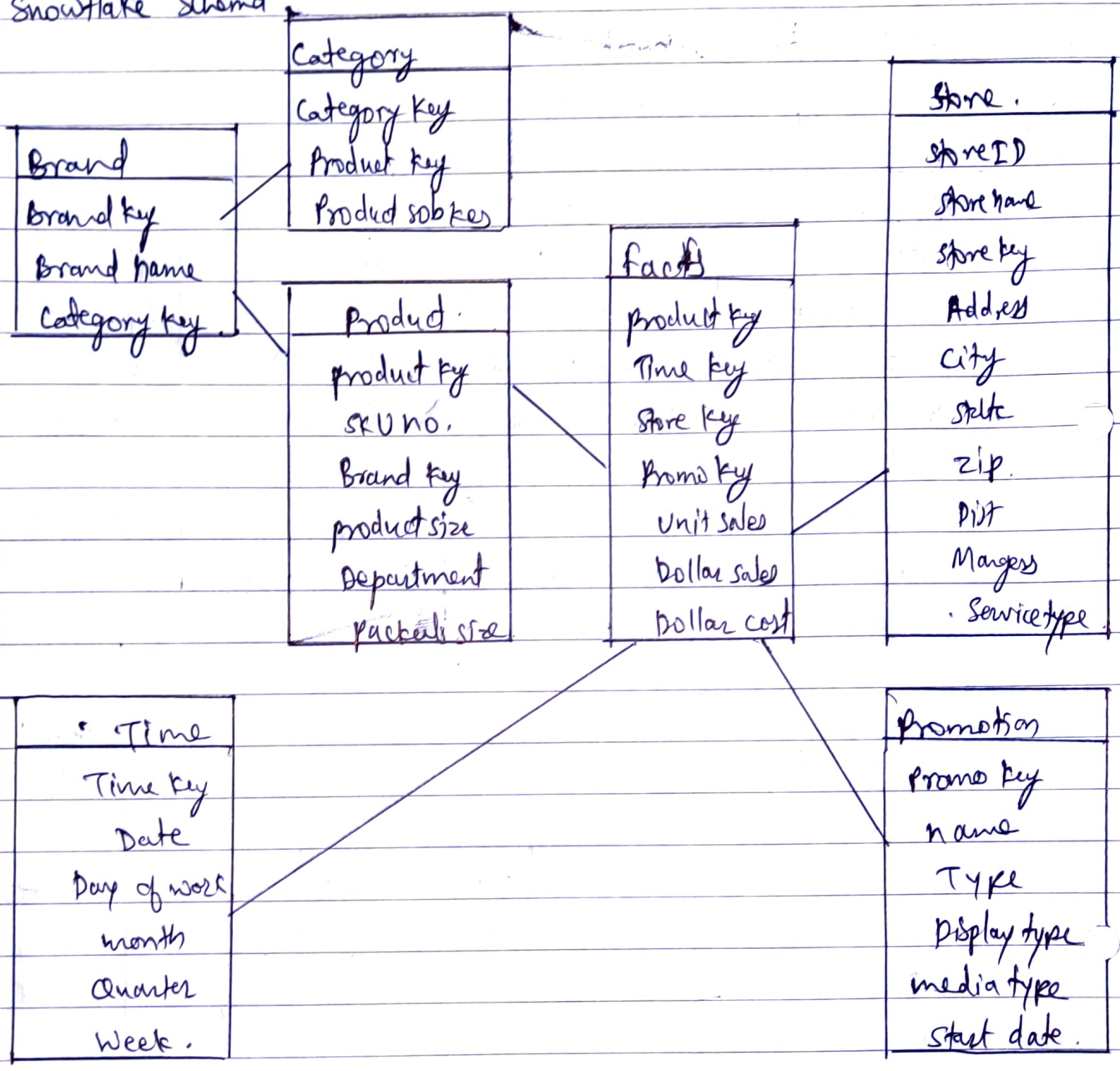
Time.	Product.	store.	Promotion.
Date	Product key	store key	Promotion key
Day of week	SKU number	Store name	promotion type
week number	Product Description	Store ID	promotion Display
month	Brand name	Address	media type
month number	Product category	City	promotion cost
Quarter	Department	Zip	start Date
Year.	Packet spe	Service Type	Responsible Manager.

Star schema :-



Amit Joshi 2)

Snowflake schema



fact Table size: a) Time Dimension: 5 years x 365 days = 1825

b) Store Dimension: 300 store.

c) Product Dimension: 40,000 product

d) Promotion Dimension: sold Item may be in one prom

Minimum no. of Base fact Tables Records = 1825 x 300 x 40000 x 1

= 2 billion

Amit Joshi 21

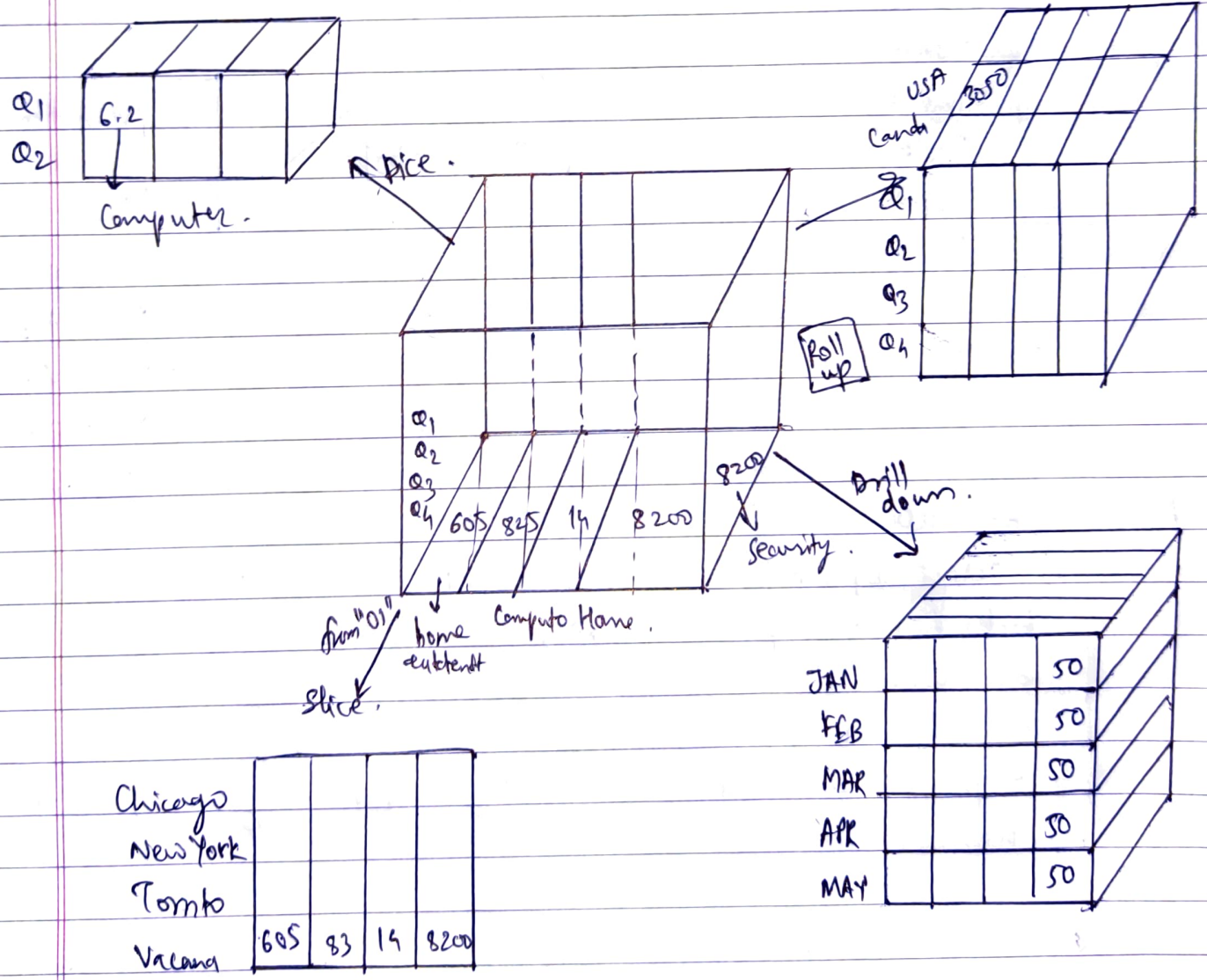
CLASSMATE
Date :
Page :

Q2)

Ans: a) Roll up : The roll up operation (Drill up operation) perform aggregation on data cube either by climbing up a concept heierchy for a dimension or by dimension reduction.

b) Drill down : Reverse of roll up . It goes from less detailed data to more detailed data can be performed by doing down operation on central cube. eg. listing each data into < day, month >

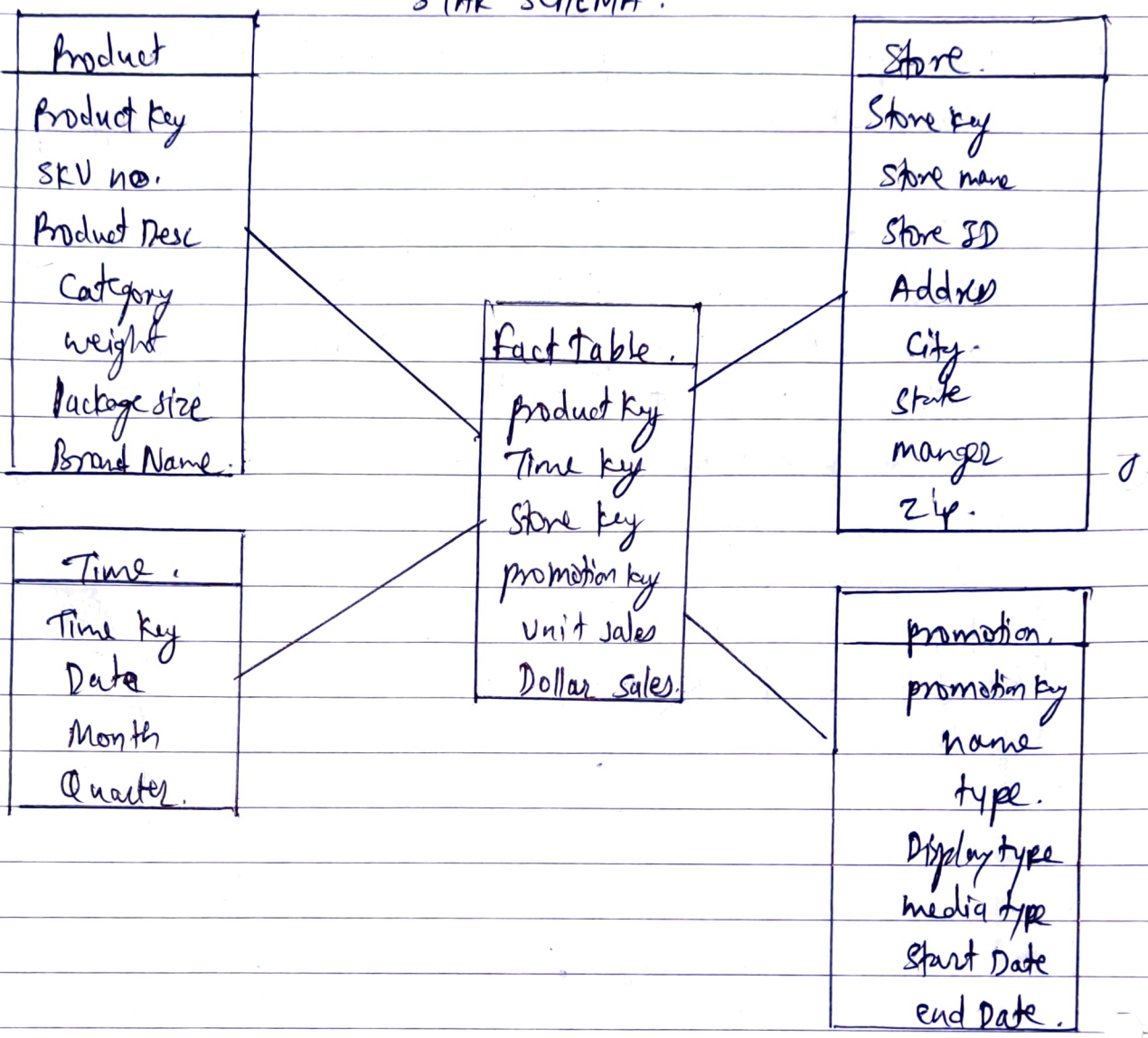
c) Slice and Dice : Slice operation performs a selection one one dimension of the given cube resulting in a sub-cube.



Amit Joshi2)

Q3)

STAR SCHEMA.



Q4)

Ans OLAP Queries.

a) select bp. Pub_name, count(bp.pub_name) from Comic books, fact cbf, publisha Rbsp, time t, distributer d, books category bc where cbf distributer_key = d. distribution key and t time key = cbf time key and cbf books_category bc books_category and t year = 2014 and cbf. pub_key = bp. pub_key group by publisha_name.

b) select sum (cost_sales), bp. pub_name from comic books_fact cf time t publisha pb where cf. time_key = t. time key and bp. p key = cf. pub_key and t. years = 2015 group by bp. pub_no

④ total sales of marvel publisher over each month in 2005

Ans select bp. pub-name, t.month, t.years, sum(cst.sst.sale)) from Comik book-fact ct, timet, book-Pub-key = bp. pub key and ct. time-key = bp.time-key and bp pub-name like marvel. t. and t. year IN(2005) group by month.

Q5)

Ans: when you rise to higher level in hierarchies of two dimensions and keep the level at the lowest in the other dimension, you create 2 way aggrate tables eg.

a) product category by terrority by Date
b) product category by Regions by Date
c) product category by all stores by Date
d) product category by month by store
e) product category by quarter by store
f) product category by terrority by date
g) product category by Region by date.
h) product department by all stores by date
i) product department by month by store
j) product department by Quarter by store
k) product. department by year by store.
l) All products by territory by date
m) All products by Region by date
n) District by Quarter by Product
o) Territory by month by product
p) Region by year by product.