COMP9318 Tutorial 1

Wei WANG

The University of New South Wales

weiw@cse.unsw.edu.au

---

① Data Warehouse and OLAP

Q1

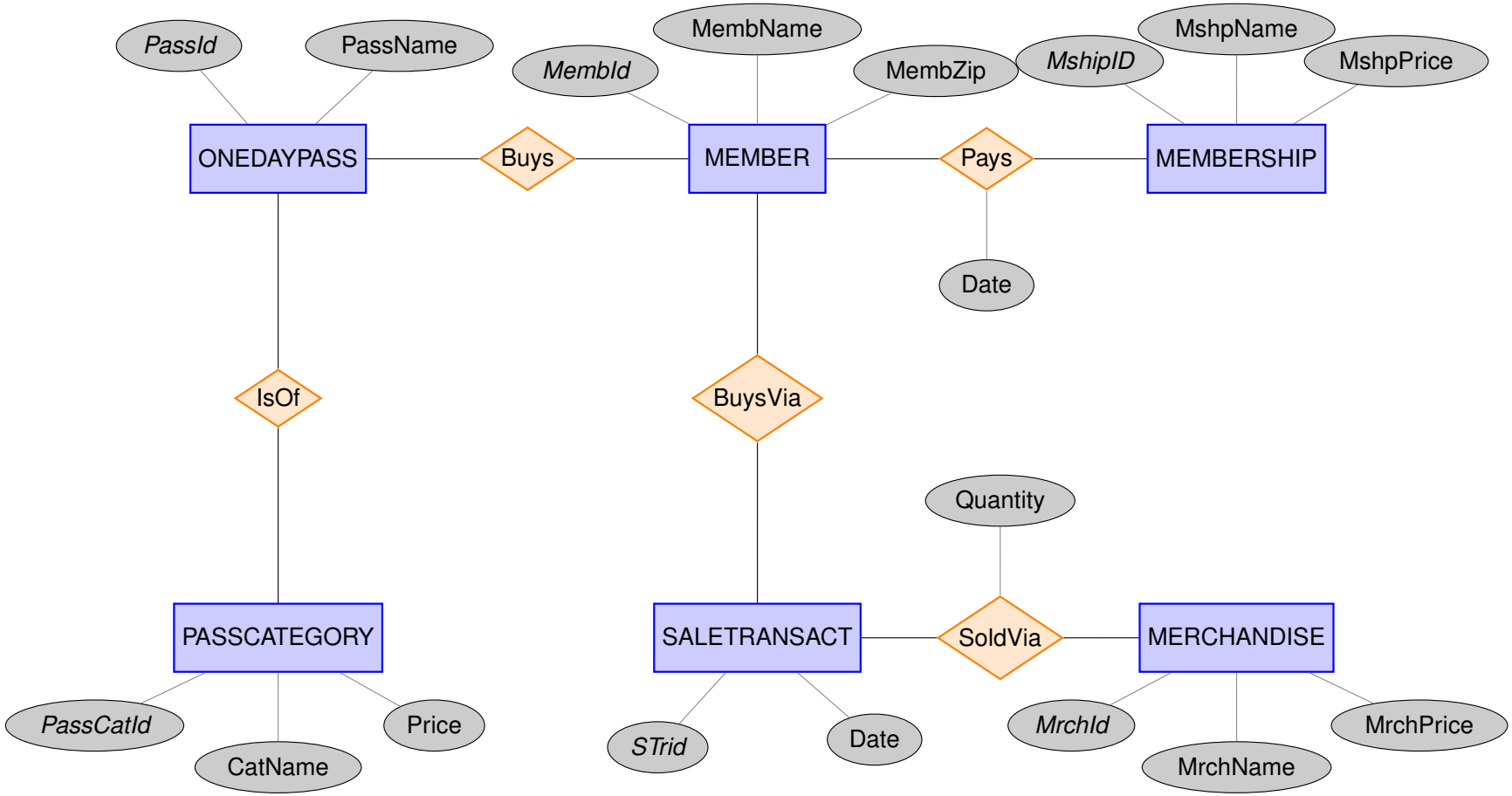
① Create a star schema diagram that will enable FIT-WORLD GYM INC. to analyze their revenue.
   - ➜ The fact table will include — for every instance of revenue taken — attribute(s) useful for analyzing revenue.
   - ➜ The star schema will include all dimensions that can be useful for analyzing revenue
   - ➜ The only two data sources are shown below

② Appreciate the ETL process involved populating the data warehouse.

③ Appreciate the difference of formulating queries: “Find the percentage of revenue generated by members in the last year”.

④ How many cuboids are there in the complete data cube?

ER Diagram


PassId
PassName
MembId
MembName
MembZip
MshipID
MshpName
MshpPrice
ONEDAYPASS
Buys
MEMBER
Pays
MEMBERSHIP
Date
IsOf
BuysVia
Quantity
PASSCATEGORY
SALETRANSACT
SoldVia
MERCHANDISE
PassCatId
CatName
Price
STrid
Date
MrchId
MrchName
MrchPrice

Data Instances

MEMBER

Membid	*MembName*	*MembZip*	*MshpID*	*MsDatePayed*
111	Joe	60611	M1	1-Jan-04
222	Mary	60640	M3	1-Jan-04
333	Sue	60611	M3	1-Jan-04

MEMBERSHIP

MshpID	*MshpName*	*MshpPrice*
M1	Platinum	$1,000
M2	Gold	$800
M3	Value	$300

ONEDAYPASS

PassID	*PassDate*	*PassCatID*	*Membid*
1-001	1-Jan-04	PSA	111
1-002	1-Jan-04	PSA	333
1-003	2-Jan-04	PSK	333

MERCHANDISE

MrchID	*MrchName*	*MrchPrice*
AP1	T-shirt	$11
AP2	Hat	$9
EQ1	Jump Rope	$12

PASSCATEGORY

PassCatId	*CatName*	*Price*
PSA	Adult	$20
PSS	Senior	$10
PSK	Kid	$3

SOLDVIA

STrid	*MrchID*	*Quatity*
11111	AP1	1
11112	AP2	1
11112	AP2	1
11113	EQ1	3

SALESTRANSACT

STrid	*Date*	*Membid*
11111	1-Jan-04	333
11112	2-Jan-04	222
11113	3-Jan-04	111

Note: MEMEBERS can bring in non-member guests. For each non-member guest, a member buys a one-day-guest-pass of a certain pass category.

Another Data Source

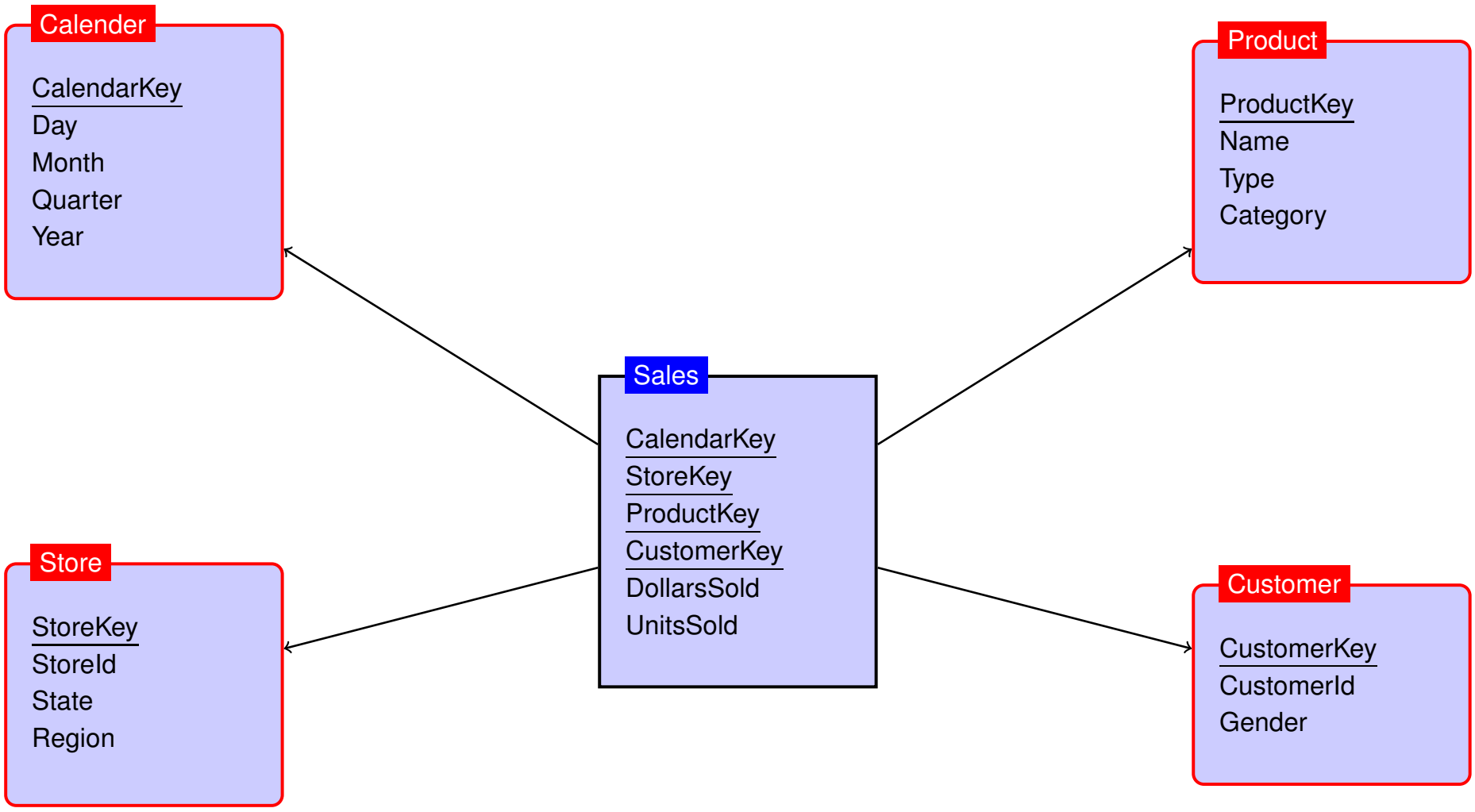
SPECIALEVENT

CorpCustID	*CorpCustNameLoc*	*EventTypeCode*	*EvetType*	*EventDate*	*AmountCharged*
CC1	Sears, Chicago 60640	L-A	All Day Rental,	January 4, 2004	$3500
CC2	Boeing, Chicago 60611	L-H	Half Day Rental,	January 5, 2004	$2200

Q2

Consider the star schema below

- ➜ Write an MDX query that display total DollasSold for each product category and each store in the State 'CA'.
- ➜ Write an MDX query that display total DollasSold for each product category and each store in the State 'CA' in 2007.
- ➜ create a star schema that has **Month** and **Region** as the finest granularity on the corresponding dimensions. Show all tables in the new data model populated with the data based on the data from the original model.


Calender
CalendarKey
Day
Month
Quarter
Year
Product
ProductKey
Name
Type
Category
Sales
CalendarKey
StoreKey
ProductKey
CustomerKey
DollarsSold
UnitsSold
Store
StoreKey
StoreId
State
Region
Customer
CustomerKey
CustomerId
Gender

Populated Tables

CALENDAR

CalendarKey	*Day*	*Month*	*Quarter*	*Year*
1	1	Jan	1	2003
2	2	Jan	1	2003
3	1	Feb	1	2003

PRODUCT

ProKey	*ProName*	*ProType*	*Category*
1	Luvs 50	Diapers	Infant Care
2	Huggies 24	Diapers	Infant Care
3	High C	Vitamin	Dietary Supp

STORE

StoreKey	*StoreID*	*State*	*Region*
1	X1	Maine	East
2	X2	New Jersey	East
3	Y1	Ohio	Midwest

CUSTOMER

CustKey	*CustID*	*Gender*
1	12	Male
2	23	Male
3	34	Female

SALES

CalendarKey	*ProKey*	*StoreKey*	*CustKey*	*$Sold*	*UnitsSold*
1	1	1	1	$15	1
1	2	2	2	$20	1
1	2	1	1	$40	2
1	2	2	1	$20	1
2	2	1	1	$19	1
2	2	2	1	$19	1
3	3	3	1	$9	2
3	3	3	2	$9	1
3	3	3	3	$9	1