Lab Manual

CS614 – Data Warehousing



Week	Lab Topic	Page No.
No.		
1	Installation of Microsoft SQL Server 2000	3
2	Normalization	18
3	De-Normalization	27
4	Process of Cube Creation	31
5	Dimension Modelling and Star Schema	50
6	Data Extraction	52
7	Basic Sorted Neighborhood Method (BSN)	63
8	Data Quality Rules	65
9	Key Range Partitioning	67
10	Indexing Technique	74
11	Nested Loop, Sort Merge, and Hash Join using SQL Server Query Analyzer	77
12	Kimbal Approach Part I	82
13	Kimbal Approach Part II	84
14	Dimension Model for DWH	85

Installation of Microsoft SQL Server 2000

Important Instruction

- ✓ User must have Administrators permissions in order to install Windows SQL Server 2000.
- ✓ Windows Server 2000 with SP3 or higher must be installed on the system.
- ✓ Close all programs running on computer before installation of SQL Server 2000.

Installation of SQL Server 2000 Step by Step:

1. Open the folder/Drive containing set up of SQL Server 2000 and double click on "autorun".

🈂 SQL2000_Ent (D:)				
<u>Eile E</u> dit <u>V</u> iew F <u>a</u> vorites	Tools Help			<u>.</u>
🚱 Back 🔹 🕥 - 🏂	Search 🏀 Folders			
Address 🛃 D:\				💌 🄁 Go
Folders	×	0	0	
Desktop	books	devtools	html	
My Computer My Computer Josephine Josephi	INSTALL	MSDE	MSEQ	
books books devtools html	msolap	×86	autorun	
INSTALL I INSTALL MSDE I INSTALL MSDE	autorun Setup Information 1 KB	Configuration Settings 1 KB	license Ten Cocument 20 KB	
msolap msolap msolap msola msola	msolap.iss ISS File 2 KB	Text Document 47 KB	redist Text Document 29 KB	
Gond Grand G	MS-DOS Batch File 1 KB	MS-DOS Batch File 1 KB	MS-DOS Batch File 1 KB	
🥑 Recycle Bin	smssql.pdf PDF File 1 KB	sqlcli M5-DOS Batch File 1 KB	sqldLiss ISS File 5 KB	
	sqlcst MS-DOS Batch File 1 KB	sqlcst.iss ISS File 8 KB	MS-DOS Batch File 1 KB	
	sqlins.iss ISS File 2 KB	MS-DOS Batch File 1 KB	sqlsms.iss ISS File 2 KB	

2. At the Microsoft SQL Server 2000 Enterprise Edition screen, press the "SQL Server 2000 Components" button.

Lab 1



3. At the Microsoft SQL Server 2000 Enterprise Edition screen, press the "Install Database Server" button



4. If Windows Server 2003 is installed you will see the SQL Server 2000 window. Press the Continue button.

🗘 sql s	Server 2000	×
<u>.</u>	SQL Server 2000 SQL Server 2000 sp2 and below is not supported by this versio Windows. Please apply sp3 after installation of SQL Server 200	n of 10.
To run t	the program, click Continue. For more information, click Details.	—
🗖 Don	n't display this <u>m</u> essage again	
	Continue Cancel Details	

5. You will be brought to the Welcome window to the Microsoft SQL Server Installation Wizard. Here press the Next button.



6. You will be brought to the Computer Name window. Since you are installing SQL 2000 Server on the Server computer, you will select the Local Computer and press the Next button.

	Enter the name of the computer on which you wan create a new instance of SQL Server or modify an existing instance of SQL Server.	t to
H III	Or, enter the name of a new or existing Virtual SQL Server to manage, SYSTEM7	
	Local Computer Bemote Computer Virtual Server Browse	
	Help < <u>B</u> ack <u>N</u> ext >	Cancel

7. You will be brought to the Installation Selection window. Select the "Create a new instance of SQL Server, or install Client Tools" radio button and press Next button.



8. User will be brought to the User Information window. Enter the computer Name and Company for your system, press Next button.

User Information	the second s	×
	Enter your name below. It is not necessary to enter a company name.	
	N <u>a</u> me:	
	Company:	
	< Back Next > Cancel	-
		-

9. You will be brought to the Software License Agreement window. After reading the legal agreement, press the Yes button.

oftware Li	icense Agr	eement			3
2	Please read the rest of th	I the following Li ne agreement.	icense Agreemen	t. Press the PA(GE DOWN key to see
END-USI SERVER	ER LICENSE LICENSE F	AGREEMENT	FT SQL SERVER	PRODUCTS	<u> </u>
IMPORT. agreement for the M and may document the Produ INSTALL AGREE, YOUR PI The Prod	ANT-READ I nt between y icrosoft softw include asso itation ("Proc act. YOU AC ING, COPYI DO NOT INS LACE OF PU luct may con	CAREFULLY: T iou (either an in vare product ide iciated media, p fuct"). An amer REE TO BE B NG, OR OTHE STALL OR USE IRCHASE FOR itain the followir	This End-User Lick dividual or a single entified above, wh printed materials, a ndment or addend OUND BY THE T RWISE USING T E THE PRODUCT A FULL REFUND ng software:	ense Agreement e entity) and Mic ich includes cor nd "online" or e lum to this EUL4 ERMS OF THIS HE PRODUCT. ; YOU MAY RE).	("EULA") is a legal crosoft Corporation mputer software lectronic A may accompany EULA BY IF YOU DO NOT TURN IT TO
Do you ao will close.	ccept all the To install M	terms of the pre licrosoft SQL Se	eceding License A erver 2000, you m	greement? If yo ust accept this a	ou choose No, Setup agreement.
			< <u>B</u> ack	<u>Y</u> es	No

10. You will be brought to the registration key window. Enter the 25 digit registration Key and press the Next button.



11. You will be brought to the Installation Definition window. Choose the "Serve and Client Tools" radio button and press the Next button.

Installation Definition	You can select one of the following types of installations.	
	C <u>C</u> lient Tools Only	
H m		
	C Connectivity Only	
	This option allows you to install a server and the client tools. Use this option if you want to set up a server with administration capabilities.	
	Help < <u>B</u> ack <u>N</u> ext > Cancel	

12. You will be brought to the Instance Name window. Her you have to press the Next button.

Instance Name		×
	Default For a default installation, leave Default checked and click Next. To install or maintain a named instance of SQL Server on this computer clear the Default checkbox and type or select an instance name. A new name must be 16 characters or less and should start with a letter or other acceptable character. For more information, click Help. Instance name:	
	Help < <u>B</u> ack <u>N</u> ext > Can	cel

13. You will be brought to the Setup Type window. Of the type of Setup preferred, select Typical. And press the Next button.

Setup Type				×
Click the type of 9	Setup you prefer, then c	lick Next.		
• Typical	Installed with the m	ost common option	s. Recommended f	or most users.
C <u>M</u> inimum	Installed with minim	um required options	s.	
○ <u>C</u> ustom	You may choose th advanced users.	e options you want	to install. Recomm	ended for
- Destination Fold	der			
Program Files	B <u>r</u> owse	C:\Program File	es/Microsoft SQL S	erver
Data Files	Bro <u>w</u> se	C:\Program File	es\Microsoft SQL S	erver
Canada an araar	an Glas diiyay		Required:	Available:
Space on suster	ani nies unve. m drive:		182917 K	47708568 K
Space on data I	iles drive:		34432 K	47708568 K
	Help	< <u>B</u> ack	Next >	Cancel

14. You will be brought to the Service Accounts window. Choose the 'Use the same account for each service Auto start SQL Server Service.' radio button. Enter Administrator's Password and press the Next button.

C Customize the settings for	each service.	
Services	Service Settings	
 SUL Server Agent 	Use a Domain User account Username: Administrator	
	Domain: SYSTEM7	_

15. You will be brought to the Authentication Mode window. Select the 'Mixed Mode (Windows Authentication and SQL Server Authentication) radio button. Enter your password in the Enter Password and Confirm Password fields and press the Next button.

then	itication Mode	
	Choose the authentication mode.	
	C Windows Authentication Mode	
	Mixed Mode (Windows Authentication and SQL Server Authentication)	
	Add pressured for the endoring	
	Add password for the sa login.	
	Enter password:	
	Confirm password:	
	Blank Password (not recommended)	
	Help < <u>B</u> ack <u>N</u> ext> Cance	el

16. You will be brought to the Start Copying Files window and press the Next button.



17. Once the file installation is complete, you will be brought to the Choose Licensing Mode. Select either Per Seat or Processor License. Press the Continue button.

Choose Licensing Mode	<
SQL Server 2000 supports two Client Licensing Modes	
Per Seat - Each device that accesses Microsoft SQL Server 2000 requires a separate Client Access License.	
Processor License - Each processor on this server requires a separate Processor License.	
Please click Help to fully understand the implications of the licensing mode you choose.	
Licensing Mode	
Per Seat for O evices	
C Processor License for D processors	
<u>C</u> ontinue <u>Ex</u> it Setup <u>H</u> elp	

18. The System will display this window when loading the system.



19. The system will display this window when installing the Microsoft Full-Text Search Engine.



20. SQL Server 2000 progress screen will display.

🛃 Enterprise Edition		_
Microsoft SQL Serve	r 2000	
	c:\pragram files\microsoft sql server\mssql\binn\opends60.dll 4 %	

21. Setup is preparing to configure the server and setup started the server and installation of selected configuration. After that setup updating window will display.



22. Once the setup is complete, you will be brought to the following window. Press the **Finish** button to complete the installation.



Installation of SQL Server 2000 Client:

1. Open the Drive/Folder and click on "autorun".



2. At the Microsoft SQL Server 2000 Enterprise Edition screen, press the SQL Server 2000 Components button.



3. At the Microsoft SQL Server 2000 – Enterprise Edition – Install Components screen, press the Install Database Server button.



4. If Windows Server 2003 is installed, you will see the Setup window. Press the OK button.



5. You will be brought to the Welcome window to the Microsoft SQL Server Installation Wizard. Press the Next button.



6. You will be brought to the Computer Name window. Since you are installing SQL 2000 Client, the Local Computer option will be the only selection. Press the Next button.

Computer Name	
	Enter the name of the computer on which you want to create a new instance of SQL Server or modify an existing instance of SQL Server.
T T T	Or, enter the name of a new or existing Virtual SQL Server to manage.
	Local Computer Bemote Computer Virtual Server
	Help < Back Next > Cancel

7. You will be brought to the Installation Selection window. The "Create a new instance of SQL Server, or install Client Tools" will be the only option available. Press the Next button.



8. You will brought to the User Information window. Enter the computer Name and Company for your system. Press the Next button.

User Information	
	Enter your name below. It is not necessary to enter a company name.
	N <u>a</u> me:
	Company:
=	
	< <u>B</u> ack <u>N</u> ext > Cancel

9. You will be brought to the Software License Agreement. After reading the legal agreement, press the Yes button.

Software License Agreement
Please read the following License Agreement. Press the PAGE DOWN key to see the rest of the agreement.
END-USER LICENSE AGREEMENT SERVER LICENSE FOR MICROSOFT SQL SERVER PRODUCTS
IMPORTANT-READ CAREFULLY: This End-User License Agreement ("EULA") is a legal agreement between you (either an individual or a single entity) and Microsoft Corporation for the Microsoft software product identified above, which includes computer software and may include associated media, printed materials, and "online" or electronic documentation ("Product"). An amendment or addendum to this EULA may accompany the Product. YOU AGREE TO BE BOUND BY THE TERMS OF THIS EULA BY INSTALLING, COPYING, OR OTHERWISE USING THE PRODUCT. IF YOU DO NOT AGREE, DO NOT INSTALL OR USE THE PRODUCT; YOU MAY RETURN IT TO YOUR PLACE OF PURCHASE FOR A FULL REFUND.
The Product may contain the following software:
Do you accept all the terms of the preceding License Agreement? If you choose No, Setup will close. To install Microsoft SQL Server 2000, you must accept this agreement.
< <u>B</u> ack <u>Y</u> es <u>N</u> o

10. You will be brought to the Registration Key window. Enter the 25 digit CD-Key. Click on the Next button.

CD-Key	<
Enter your 25-digit CD-Key	
< <u>B</u> ack <u>N</u> ext > Cancel]

11. You will be brought to the Installation Definition window. Choose the "Client Tools Only" and then press the Next button.



12. You will be brought to the Select Components window– press on the Next button for default values.

Select Components			X
Select or clear the compon	ents to install	/ remove.	
Components:		Sub-Components:	
🖌 Management Tools	28672 K	Enterprise Manager	25024 K
Client Connectivity	288 K	Profiler	640 K
🖌 Books Online	32512 K	Query Analyzer	1696 K
Development Tools	288 K	DTC Client Support	0 K
Code Samples	0 K	Conflict Viewer	1024 K
Installs interactive manag	ement tools to	help run SQL Server.	
Space on program files dr	ive :	Required: 208 K	Available: 37207484 K
Space on system drive:		151248 K	37207484 K
Space on data files drive:		0 K	37207484 K
	<u>H</u> elp	<a>Back Next >	Cancel

13. At the "Start Copying Files" window, press the Next button.

Start Copying Files	
	Setup has enough information to start copying the program files. If you want to review or change any settings, click Back. If you are satisfied with the settings, click Next to begin copying files.
	< <u>B</u> ack Next> Cancel

14. The system will display this window when loading the system.



15. SQL Server 2000 progress screen will display.



16. Once the setup is done, you will see the following window. Press the Finish button to complete the installation.



Mechanism to Conduct Lab:

Students and teacher communicate through Adobe Connect.

NORMALIZATION

Consider the following table

Project ID	Project Name	Project Budget	EmplD	Emp Name	Hourly Rate
1001	Pakistan International Airlines Database	1 billion	101 102 103	Sana Ali Hasan	60000 80000 45000
1002	NADRA database	20 million	111 112	Amir Umer	90000 80000

Following functional dependences exist.

- 1. ProjectID->ProjectName, Project Budget
- 2. EmpID->EmpoyeeName, HourlyRate
- 3. ProjectID, EmpID->ProjectName, Project Budget , EmpName, HourlyRate

Normalize above given table into first and second normal form.

Solution:

First Normal Form: Remove repeating groups

Project	Project Name	Project Budget	EmpID	Emp Name	Hourly
ID					Rate
1001	Pakistan International	1 billion	101	Sana	60000
	Airlines Database				
1001	Pakistan International	1 billion	102	Ali	80000
	Airlines Database				
1001	Pakistan International	1 billion	103	Hasan	45000
	Airlines Database				
1002	NADRA database	20 million	111	Amir	90000
1002	NADRA database	20 million	112	Umer	80000

Second Normal Form: Remove Partial Dependencies

The above table is not in the second normal form since there exists the partial dependency through the FDs 1, 2 and 3. To bring it into second normal form, we will decompose the table into the following tables:

Lab 2

Employee Table

EmpID	Emp Name	Hourly Rate
101	Sana	60000
102	Ali	80000
103	Hasan	45000
111	Amir	90000
112	Umer	80000

Project Table

Project ID	Project Name	Project
		Budget
1001	Pakistan International Airlines Database	1 billion
1001	Pakistan International Airlines Database	1 billion
1001	Pakistan International Airlines Database	1 billion
1002	NADRA database	20 million
1002	NADRA database	20 million

Employee-Project Table

Project ID	EmpID
1001	101
1001	102
1001	103
1002	111
1002	112

Now, above tables are in second normal form. Create above all tables in SQL server enterprise manager in normalized form. The procedure of table creation is given below.

1. Open SQL Server Enterprise Manager by clicking on Start menu->Programs->Microsoft SQL Server->Enterprise Manager.



2. Click server node



3. Right Click on Databases node and click on New Database...

🚡 Console Root\Microsoft S	QL Servers	SQL Server Gr	oup\(LOCAL	.) (Windows N	T)\Databases	
Console Root		Databases 7	Items			
□ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □						
E 🔂 (LOCAL) (Window	vs NT)					
🕀 🔂 Databases	Mau Daha	123	master I	model	msdb	Northwind
E Data Trans	New Data	<u>o</u> ase				
	All Tas <u>k</u> s	• •				
🕀 🧰 Security	⊻iew	•	temodh			
E Support Se	New <u>W</u> ind	ow from Here				
	Refresh					
	Export <u>L</u> is	t				
	Help					
_			1			
						1

4. Type MyFirstDB in **Name** as Database Name and click oK.

Database Properties - MyFirstDB		×
General Data Files Transaction Log		
<u>N</u> ame: MyFirstDB		
Database		
Status:	(Unknown)	
Owner:	(Unknown)	
Date created:	(Unknown)	
Size:	(Unknown)	
Space available:	(Unknown)	
Number of users:	(Unknown)	
Backup		
Last database backup:	None	
Last transaction log backup:	None	
Maintenance		
Maintenance plan:	None	
Collation name:	(Server default)	
	OK Cancel Help	

5. Expand Databases node, right click on your newly created databases MyFirstDB, click on New and then click on Table.

🚡 Console Root\Microsoft	SQL Server	s\SQL Server (Group\(LOCAL	.) (Windows N	T)\Databases	<u> </u>
Console Root	-	Databases	8 Items			
Console Root Co	New Data New All Tasks New Wind Delete Refresh	Northwind base	B Items master pubs Database Database Table Yiew Stored Pro	User Diagram	msdb	MyFirstDB
	Help		R <u>u</u> le			
	Пеф		De <u>f</u> ault User Defin User Defin New Full-T	ed Data Type ed Function ext ⊆atalog		
			Pu <u>b</u> lication Pull <u>S</u> ubscr	n ription		

6. The following window will open.

New Table in 'MyFirstDE	' on '(LOCAL)'				
Column Name	Data Type	Length	Allow Nulls		
1					
Columns					
				1	
Description					
Default Value					
Precision					
Scale					
Identity Sood					
Identity Jacrement					
Is RowGuid					
Formula					
Collation					
				_	

7. Type Column Name and select data type. See following figure in which we have created variable EmpID.

Column Name Data Type Length Allow Nulls EmpID int 4 ✓ Columns	New Table in 'N	MyFirstDB' (on (LUCAL)		
EmpID int 4 Columns Description	Column Na	ame	Data Type	Length	Allow Nulls
Columns Description Default Value Precision 10 Scale 0 Identity No Identity Seed Identity Increment Is RowGuid Formula Collation	▶ EmpID	ir	nt	4	 V
Description	Columns	1			
Description Image: Constraint of the sector of the secto	Columno	·			
Default Value Precision 10 Scale 0 Identity No Identity Seed 1 Identity Increment 1 Is RowGuid No Formula Collation	Description	1			
Precision 10 Scale 0 Identity No Identity Seed Identity Increment Is RowGuid No Formula Collation	Default Value				
Scale 0 Identity No Identity Seed Identity Increment Is RowGuid No Formula Collation	Precision	1	0		
Identity No Identity Seed	Scale	0			
Identity Seed Identity Increment Is RowGuid No Formula Collation	Identity	N	lo		
Is RowGuid No Formula Collation	Identity Seed	-			
Collation	Is RowGuid	anc N	lo		
Collation	Formula				
	Collation				

8. Right Click on column row and select option Insert Column.

New Table in 'MyFirstDB'	on '(LOCAL)'			
Column Name	Data Type	Length	Allow Nulls	
EmpID	int	4	¥	
Set Primar <u>y</u> Key				
Insert Colu <u>m</u> n				
<u>D</u> elete Column				
Tock				
Таж				
Select <u>A</u> ll	10			
Terdennedikanna	P No			
Indexes/Keys				
Relationships				
Check Constraints	No			
Properties				
]			

9. Similarly create other variables, EmpName and Hourly Rate

7	New Table in 'MyFirstDB	' on '(LOCAL)'			
	Column Name	Data Type	Length	Allow Nulls	<u> </u>
•	EmpName	char	10	~	•
ſ	Columns				1
	Description				
	Default Value				
	Precision	0			
	Scale	0			
	Identity	No			
	Identity Seed				
	Identity Increment				
	Is RowGuid	No			
	Formula				
	Collation	<database default=""></database>			
-					

	New Table in MyFirst	DB ON (LULAL)		
	Column Name	Data Type	Length	Allow Nulls
[Hourly Rate]	int	4	V
i	Coloren 1			
	Columns			
	Description			
	Default Value			
	Precision	10		
	Scale	0		
	Identity	No		
	Identity Seed			
	Identity Increment			
	Is RowGuid	No		
	Formula			
	Collation			
_				
-				
_				

10. To set a column as primary key, right click on column name and click **Set Primary Key**.

7	Design Table 'Employee	e' in 'MyFirstDB' on '	(LOCAL)'		<u>_0×</u>
	Column Name	Data Type	Length	Allow Nulls	•
Ŀ	Set Primary Key	char	10	√	•
	Insert Colu <u>m</u> n				
	<u>D</u> elete Column				1
	Task				
	Select <u>A</u> ll	0			
	Inde <u>x</u> es/Keys	No			
	Relations <u>h</u> ips				
	Check Co <u>n</u> straints				
	Properties	No			
	Collation	<database default=""></database>			
_					

11. Click close icon, the **SQL Server Enterprise Manager** will open. Click **Yes**.



12. Type Employee in variable name. Click OK.

Choose Nam	e			×
<u>E</u> nter a name	e for the table:			
Employee				
	ОК	Cancel	Help	

13. See in the following figure the new created table Employee. It is added in Tables list.

🚡 Console Root\Microsoft SQL Server:	s\SQL Server Group\(LOCAI	.) (Windows NT)\	Databases\My	Fir 💶 🖂 🔀
🖻 🔂 (LOCAL) (Windows NT) 🔺	Tables 21 Items			
Databases	Name 🛆	Owner	Туре	Create [🔺
	📰 dtproperties	dbo	System	10/12/2
ti of model	🔲 Employee	dbo	User	10/12/2
	📰 syscolumns	dbo	System	8/6/200
🖃 🧻 MyFirstDB	syscomments	dbo	System	8/6/200
ara Diagrams	📰 sysdepends	dbo	System	8/6/200
	sysfilegroups	dbo	System	8/6/200
- 60 Views	📰 sysfiles	dbo	System	8/6/200
Stored Proce	📰 sysfiles1	dbo	System	8/6/200
Users	📰 sysforeignkeys	dbo	System	8/6/200
Roles	📰 sysfulltextcatalogs	dbo	System	8/6/200
Rules	📰 sysfulltextnotify	dbo	System	8/6/200
Deraults	📰 sysindexes	dbo	System	8/6/200
	📰 sysindexkeys	dbo	System	8/6/200
	📰 sysmembers	dbo	System	8/6/200
	sysobierts	dbo	System	8/6/200

14. Right click on table name, select **Open Table** options and then **Return all rows** option.

🚡 Console Root\Microsoft SQL Servers	s∖SQL Ser	ver Group\(LOCAL) (Wi	ndows NT)\Da	tabases\MyF	ir 💶 🗵	
🖻 🔂 (LOCAL) (Windows NT) 🔺	Tables	21 Items				
Databases	Name 🕖	Δ	Owner	Туре	Create [🔺	
H master	📰 dtpro	perties	dbo	System	10/12/2	
	🔲 Emplo	wee	dbo	User	10/12/2	
	📰 sys	New Table	dbo	System	8/6/200	
🖃 🔋 MyFirstDB	🔲 sys	Design Table	ldbo	System	8/6/200	
- 🖙 🔤 Diagrams	🛅 sys	Open l'able	Return <u>a</u> ll ro	ws m	8/6/200	
Tables	🔲 sys	Eull-Text Index Table 🔸	Return <u>T</u> op.,	• m	8/6/200	
- 660' Views	🔲 sys –	All Table A	Query	m	8/6/200	
Stored Proce	sys _	All Tas <u>k</u> s •	dbo	System	8/6/200	
	🔲 sys	Cu <u>t</u>	dbo	System	8/6/200	
Roles	sys 📰	⊆ору	dbo	System	8/6/200	
	sys 📰	💷 sys	<u>D</u> elete	dbo	System	8/6/200
. User Defined	sys	Rena <u>m</u> e	dbo	System	8/6/200	
- Se User Defined	sys –	Properties		System	8/6/200	
📕 📔 Full-Text Cat 🚽	sys ===================================	riopercies		System	8/6/200	
		<u>H</u> elp		-500 Pff		

15. Enter data in table manually or using some application.

7	Data in Table 'Em	ployee' in 'MyFir	stDB' on '(LOCAL)'		
	EmpName	EmpID	Hourly Rate		
.1	Sana	101	60000		
*					
L .					
L .					
L .					
L .					
L .					
L .					
L .					
L .					
L .					
L .					
L .					
L .					
L .					
L .					

Follow above method for all tables.

Mechanism to Conduct Lab:

Students and teacher communicate through Adobe Connect.

Lab 3

Scenario

"Pak Airline" is an airliner reservation company, which is operating in more than 10 countries. They have developed the airline reservation system to avoid the errors faced in manual system. The staff of the airline use airline reservation system form the tasks such as flight scheduling, ticket reservation, announcements in automated way. Similarly, users/passengers can search for flight schedule according to date and time and fare details. The staff of the airline can manage the reservation systems by flight rout, runway details, flight scheduling and reservation.

Ticket reservation system of the Pak Airline provides the information about schedule of flights, availability of seats, flight number and destination. For reservation of ticket user have to provide its personal information such as name, age, address etc. For payment purpose user will provide credit card number and bank details. Moreover, information about flight number, date of departure, no. of tickets to be booked is also required for confirmation of ticket. Following is the ERD of above airline reservation system.



Question Statement:

You are required to de-normalize above tables using table pre-joining technique according the relationship between entities. Carefully identify all tables with such relationships on which Pre-joining technique can be applied.

Solution:

There are One-to-Many relationships in following entities:

- 1. Staff-Announcement
- 2. Airport-Flight
- 3. Flight-Passenger
- 4. Staff-Flight
- 5. Flight-Ticket

As Pre-joining De-normalization technique is based on 1-many relationship so the De-normalization will be performed on following tables.

Airport Table

Airport Id	Name	City

Flight Table

<u>Flight Id</u>	Arrival	Arrival	Depart	Depart	Destination	<u>Airport Id</u>	Staff Id
	uale	ume	Date	nme			

Staff Table

Staff Id	Name	Designation	Contact No		

Announcement Table

Announcement Id	Announcement Text	Start Date	Expire Date	Staff Id

Passenger Table

Passenger Id	Name	Contact No	Address	Email	Credit Cr.No	<u>Flight Id</u>

Ticket Table

Ticket Id	Туре	<u>Flight Id</u>

Merge Staff and announcement table as there is one to many relationship which is the requirement of Pre-joining.

Staff-Announcement

<u>Staff</u>	Name	Designation	Contact	<u>Announcement</u>	Text	Start	Expire
<u>Id</u>			No	<u>Id</u>		Date	Date

Airport-Flight Table

<u>Airport Id</u>	Name	City	<u>Flight Id</u>	Arrival date	Arrival time	Depart Date	Depart Time	Destination

Flight –Passenger Table

<u>Flight</u> Id	Arrival date	Arrival time	Depart Date	Depart Time	Destin ation	Passen ger Id	Name	Contact No	Addr ess	Email	Credit Cr.No	<u>Airport</u> Id	<u>Staff</u> Id

Staff-Flight Table

<u>Staff</u> <u>Id</u>	Name	Designation	Contact No	<u>Flight Id</u>	Arrival date	Arrival time	Depart Date	Depart Time	Destinat ion

Flight-Ticket Table

<u>Flight</u> Id	Arrival date	Arrival time	Depart Date	Depart Time	Destination	<u>Airport</u> Id	<u>Staff</u> Id	<u>Ticket</u> Id	Туре

Mechanism to Conduct Lab:

Students and teacher communicate through Adobe Connect.

Lab 4 Part I

Installation Guide for MS SQL Server 2000 Analysis Services

1. Follow all instructions in Lab lecture 1 section 3.2 "Installing Microsoft SQL Server 2000" and then Double Click on "AutoRun.Exe" icon.



2. The Microsoft SQL Server 2000 Enterprise Edition screen will appear. Now, press the SQL Server 2000 Components button.



3. Now, Press "Install Analysis Services" button.



4. The following window will appear. Click "Continue" button.

sql s	erver 2000 Analysis Services	×	
1	SQL Server 2000 Analysis Services SQL Server 2000 Analysis Services sp2 and below is not supported by this version of Windows. Please apply sp3 after installation of SQL Server 2000 Analysis Services.		
To run the program, click Continue. For more information, click Details.			
Don't display this <u>m</u> essage again			
	<u>C</u> ontinue Cancel <u>D</u> etails		

5. From Microsoft SQL Server 2000 Analysis Services screen, Click Next button.

Microsoft SQL Server Analysis Services	2000
Welcome	X
	Welcome to the Microsoft SQL Server 2000 Analysis Services Setup program.
	It is strongly recommended that you exit all Windows programs before running this Setup program. Click Cancel to quit Setup and then close any programs you have running. Click Next to continue with the Setup program.
	WARNING: This program is protected by copyright law and international treaties.
	Unauthorized reproduction or distribution of this program, or any portion of it, may result in severe civil and criminal penalties, and will be prosecuted to the maximum extent possible under law.
	<u>N</u> ext > Cancel

6. Click "Yes" from following "Software License Agreement" window.

Software License Agreement	×			
Please read the following License Agreement. Press the PAGE DOWN key to the rest of the agreement.	see			
END-USER LICENSE AGREEMENT SERVER LICENSE FOR MICROSOFT SQL SERVER PRODUCTS				
IMPORTANT-READ CAREFULLY: This End-User License Agreement ("EULA") is a legal agreement between you (either an individual or a single entity) and Microsoft Corporation for the Microsoft software product identified above, which includes computer software and may include associated media, printed materials, and "online" or electronic documentation ("Product"). An amendment or addendum to this EULA may accompany the Product. YOU AGREE TO BE BOUND BY THE TERMS OF THIS EULA BY INSTALLING, COPYING, OR OTHERWISE USING THE PRODUCT. IF YOU DO NOT AGREE, DO NOT INSTALL OR USE THE PRODUCT; YOU MAY RETURN IT TO YOUR PLACE OF PURCHASE FOR A FULL REFUND.				
Do you accept all the terms of the preceding License Agreement? If you choose No, Setup will close. To install Microsoft SQL Server 2000 Analysis Services, you must accept this agreement.				
< <u>B</u> ack <u>Y</u> es <u>N</u> o				
7. Select Component as selected in following "Select Components" window.

Select Components		×
	Select the components you want to install, cle you do not want to install. <u>C</u> omponents	ear the components
	🗸 Analysis server	34368 K
= = -	🖌 Analysis Manager	48960 K
	Decision Support Objects	7200 K
	Client components	9984 K
	Sample applications	4672 K
	🖌 Books Online	31424 K
	Destination Folder D:\\Microsoft Analysis Services	B <u>r</u> owse
	Space Required:140992 KSpace Available:339740 K	Disk <u>Space</u>
	< <u>B</u> ack <u>N</u> ext>	Cancel

8. Select Program Folder and click "Next" button.

Select Program Folder		×
	Setup will add program icons to the Program Folder listed below. You may type a new folder name, or select one from the existing Folders list. Click Next to continue.	
	Program Folders:	
	Microsoft SQL Server\Analysis Services	
	Existing Folders:	
	Accessories Administrative Tools	
	Microsoft Office Microsoft SQL Server	
	Microsoft SQL Server - Switch Microsoft SQL Server 2008 B2	
	NeoSmart Technologies	1
	UltraISO	·
		_
	< <u>B</u> ack <u>N</u> ext > Cancel	

9. The following window will appear.



10. Click "Finish" button to complete setup.



Working with Analysis Manager

 Open Analysis Manager by clicking Programs->Microsoft SQL Server->Analysis Services->Analysis Manager.



2. The following Analysis Manager Window will open. To get help on working with Analysis Manager, click on Analysis Manager Concepts & Tutorial.



3. In the Analysis Manager tree view, expand Analysis Servers.



- 4. In above screen, **FoodMart2000** under the server node is the sample database available in Analysis Manager.
- 5. To create a new database:
- 1. Click the name of your server (e.g. EXAMRWP2), a connection with the Analysis server will be established.
- 2. Right-click your server's name, and then click **New Database**.



3. The following database dialog box will open.

atabase							×
Databa:	se name:						
Descript							
Descript	tion:					_	
						-	
			_				
		OK		Cancel	<u>H</u> elp		

4. In the **Database** dialog box, in the **Database** name box, enter "MYFirstDB", and in **Description** box, enter "This is my Fisrt database.", then click **OK**.

Data	base	×
Ē	2atabase name:	
F	MYFirstDB	
D	Description:	
ľ	This is my First database.	
	*	
	_	
	OK Cancel Help	

- 5. In the Analysis Manager tree pane, expand the server, and then expand the **MYFirstDB** database that you just created.
- 6. Your new **MYFirstDB** database contains the following items:
- Data Sources
- Cubes
- Shared Dimensions
- Mining Models
- Database Roles

Lab Exercise: Explore sample database FoodMart2000, all of its tables, tables' schema and browse their data.

Mechanism to Conduct Lab:

Students and teacher communicate through Skype/Adobe Connect. Students perform the task using the following simulator:

Lab 4 Part II

CREATING AND PROCESSING A CUBE

1. Click on newly created database "MYFirstDB", from its drop down list, select New Cube, then select Wizard option.



- 2. The following **Data Link Properties** window will open. Select option **Use data source name** and then select **FoodMart2000** from its dropdown list.
- 3. Write Admin as User name and check Blank Password option.
- 4. After this, Click Test Connection button.

🗒 Data Link Properties 🛛 🔀
Provider Connection Advanced All
Specify the following to connect to ODBC data: 1. Specify the source of data:
FoodMart 2000
Use connection string Connection string:
Build
2. Enter information to log on to the server
User <u>n</u> ame: admin
Password:
☑ Blank password □ Allow saving password
3. Enter the initial catalog to use:
[Test Connection]
OK Cancel Help

5. You will get message "Test Connection Succeeded". Click OK, then again click OK from Data Link Properties.

🖶 Data Link Properties 🛛 🗙
Provider Connection Advanced All
Specify the following to connect to ODBC data: 1. Specify the source of data: Use data source name FoodMart 2000 Befresh Use connection string Connection string Connection string Connection string Connection string Connection string Connection string Connection string Connection succeeded. User Pass CK
3. Enter the initial catalog to use:
<u>I</u> est Connection
OK Cancel Help

6. Cube Wizard window will show different tables which are created in FoodMart2000 database. Select **inventory_fact_1997** table.

Cube Wizard	×
Select a fact table from a data source	*
Data gources and tables:	Details: Column product_id time_id warehouse_id store_id units_ordered warehouse_sales warehouse_sales supply_time store_invoice Browse Data
< <u>B</u> ack <u>N</u> ex	t > Cancel <u>H</u> elp

7. You can browse data of selected table by clicking on Browse Data button (See above window).

8. Now you will define your cube measures. Select table numeric columns to define cube measures.

Fact table n <u>u</u> meric colun	nns:		<u>C</u> ube measures:	
Column	Туре		Measure name	Source column
product_id	Integer			
🥫 time_id	Integer	>		
🧾 warehouse_id	Integer			
🧮 store_id	Integer	<		
🧾 units_ordered	Double			
🧾 units_shipped	Integer			
🧾 warehouse_sales	Numeric			
🧾 warehouse_cost	Numeric	>>		
🧾 supply_time	Small Integer	201		
🧾 store_invoice	Numeric	~~~		

9. Select e.g. Units Ordered, Links Shipped,, Warehouse Sales, and Warehouse Cost. Then click on Next button.

Cube Wizard				×
Select the numeric co	lumns that define ye	our measu	ires	*
Fact table n <u>u</u> meric colu	umns:		<u>C</u> ube measures:	
Column	Туре		Measure name	Source column
📴 product_id	Integer		#≰¥ Units Ordered	units_ordered
🧾 time_id	Integer	>	#≰¥ Units Shipped	units_shipped
🔲 warehouse_id	Integer		#≰¥ Warehouse Sales	warehouse_sales
📑 store_id	Integer		#≰¥ Warehouse Cost	warehouse_cost
supply_time	Small Integer			
📑 store_invoice	Numeric			
		>> <<		
•			•	
	< <u>B</u> ack	<u>N</u> ext >	Cancel	

10. Now, you have to select cube dimension. From **Cube Wizard** window, click on **New Dimension**.

Cube Wizard				×
Select the dimensions for your cube				(
Shared dimensions:	Ω	ube dimens	ions:	
Name		Name		
	> <			
	>> <<			
New Dimension]			
< Back	<u>N</u> ext >		Cancel	Help

11. Click on Next button.

Dimension Wizard	×
	Welcome to the Dimension Wizard
	This wizard helps you create dimensions for your cubes.
	Dimensions are derived from the relational tables and columns containing the categorical data you want to analyze.
	Categories such as time, geography, or product line are typical examples of dimensions.
VED	Dimensions move beyond the structure of a relational database by allowing you to define hierarchical relationships in your data.
	These data hierarchies are grouped into levels that combine to form the values for the next highest level in the hierarchy.
	\Box Skip this screen in the future
	< Back Cancel Help

13. Select **Star Schema: A single dimension tables** option. You can use some other option too. Click **Next**.

Dimension Wizard				
Choose how you want to create the dimension				
Create a dimension and define its hierarchy from:				
Star Schema: A single dimension table				
O Sno <u>w</u> flake Schema: Multiple, related dimension tables				
O Parent-Child: Two related columns in a single dimension table				
O Virtual Dimension: The member properties of another dimension				
C Mining Model: A predictable column of an OLAP mining model				
Description: In a later step, select one or more columns from a single table. Each column contributes a level to the dimension (or a single date column can be parsed to create multiple levels.) If you select multiple columns, they often contain increasingly detailed information. For example, the columns Country, Province, City, and Store create a Stores dimension. This option creates a regular dimension.				
< Back Next > Cancel Help				

14. Now, you have to select Dimension table. In this lab, we have selected **product** table. You can choose any table of your choice. Click **Next**.

Dimension Wizard	×
Select the dimension table	(
Available tables:	Details: Column product_class_id product_id brand_name product_name SKU SRP gross_weight net_weight
→ isales_fact_1997 → isales_fact_1998 → isales_fact_dec_1998	Browse Data
< <u>B</u> ack	xxt > Cancel <u>H</u> elp

15. Now, select **level of your dimension**. Here in this example, we have selected BrandName, ProductId and ProductName as shown in following Figure. You can also select any other columns. Then, click **Next**.

Dimension Wizard		×
Select the levels for your dimension		
Available columns:	Dimension levels:	Source column brand_name product_id product_name
< <u>B</u> ack	Next > Canc	el <u>H</u> elp

16. Now, select member key columns and Click Next.

Dimension Wizard		×
Specify the member k	ey columns	Ð
The member key column o as the member name colu	ontains the keys for your level members. It may or may not be the same column nn. The key and the name column come from the same table.	
Name	Member Key Colump	
Brand Name	"product","brand name"	
Product Id	"product","product_id"	
Product Name	"product"."product_name"	
	< Back Next > Cancel Help	

17. Click Next.

Dimension Wizard	X
Select advanced options	
Select the advanced options you want to set for your new dimension:	
Options:	
Changing dimension	Select <u>A</u> ll
Storage mode and member groups	Deselect All
Description Storage mode determines the storage location for the dimension members. Grouping can be associated with a large level to facilitate navigation within the dimension.	
< <u>B</u> ack <u>N</u> ext > Cance	el <u>H</u> elp

18. Click Next.

Dimension Wizard			×
	Finish the	Dimension Wizard	
	Dimension name:	Product	
	₽review: I	Create a hierarchy of a dimension Unable to browse the dimension 'New Dimension'. Unspecified error	
	< <u>B</u> ack	Einish Cancel <u>H</u> elp	

19. Type Product in **Dimension Name** and click **Finish**.

20. By following same steps, add two more dimensions Product, Region and time.

Cube Wizard				×
Select the dimensions for your cube				(
Shared dimensions:		⊆ube dimens	ions:	
Name	> < ×	Name	:	
New Dimension				
< <u>B</u> ack	<u>N</u> ext >		Cancel	Help

21. Click Next. Then, Cube Wizard window will open.

Cube Wizard	×
Select the dimensions for your cube	X
Shared dimensions:	Cube dimensions:
Fact Table Row Count	
Counting the fact table rows may take	a while to complete. Do you want to count fact table rows?
	<<
New Dimension	
< <u>B</u> ack	Next > Cancel Help

22. Click **Yes** in **Fact Table Row Count** window. Cube Wizard window will open. Type MyFirstCube in **Cube Name** and then Click **Finish**.



23. Cube Editor Window will open which shows Star Schema of your cube.



24. To process cube, click on cube name MyFirstCube->Process.



25. Select the processing method and click Next.

Process a Cube	×
Select the processing method	
Select the processing method:	
O Incremental update	
C Refresh data	
Euli Process	
Description	
You must process your cube when you change its structure. Use this option to completely rebuild your entire cube. Users can continue to query the cube during this operation, but they must reconnect after processing to see new data.	
Incrementally update the shared dimensions used in this cube.	
Specify processing settings for the cube:	
Se <u>t</u> tings	
OK Cancel <u>H</u> elp	

26. Cube processing completed.



Lab Exercise: Right Click on cube name and select Browse Data option to view cube data. Perform different cube operation on this data e.g. drill down, roll up etc.

Mechanism to Conduct Lab:

Students and teacher communicate through Adobe Connect.

STAR SCHEMA

Question Statement:

Following is an **Entity Relationship Diagram (ERD)** of Airline Reservation System, you have to design **Star Schema** using all steps of dimension modelling in any drawing tool e.g. MS Visio etc.



Lab 5

Solution:

Following are steps of dimensional modeling.

Step-1: Choose the Business Process

The business process is "Air Ticket Reservation System".

Step-2: Choose the Grain

Grain represents the atomic level of information required from business process and it is termed as unit of analyses. The grain statement is **"Total number of passengers arrives/depart in each flight by the airline."**

Step-3: Choose the Facts

Facts are numeric, continuously valued and additive. The fact in our case is "**Total number of passengers** in an airplane".

Step-4: Choose the Dimensions

The dimensions will be "Airport", "Flight", "Staff", "Announcement", "Passenger", "Airplane", and "Tickets".

Star Schema:



Mechanism to Conduct Lab:

Students and teacher communicate through Adobe Connect Session.

DATA EXTRACTION

1. Open SQL Server Enterprise Manager and click on SQL Server Group.



2. Right Click on Data Transformation Services node and click New Package.

🚡 Console Root\Microsoft SQL Servers\SQL	Server Group		<u> </u>
Console Root Console Root SQL Server Group CLOCAL (Windows NT) CLOCAL (Windows NT) CLOCA	SQL Server Group 1 Ite	m	
			11

1. The following window will open.

🚡 DTS Pack	age: <	New Packag	e>								<u>- 0 ×</u>
Package	<u>E</u> dit	Connection	<u>T</u> ask	Workflow	i	¥ 🖻	6	alb	• =	면영 🤅	2
Connection	-										_
S 🙆 🕖											
🖪 🗟 🗊											
5 🔊											
Task	1										
3 🚳 🏵											
🗟 😨 📓											
🗐 🖆 🎁											
i 🕹 🙆 🦉											_

2. Click on **Package** menu and click **Save** to save the package.



3. Enter MyFirstPackage in Package name and Click OK.

Save DTS Package	x					
You can save structured sto	a a DTS package to SQL Server, Meta Data Services, a rage file, or a Visual Basic file.					
Package <u>n</u> ame:	MyFirstPackage					
Owner password:	Use <u>r</u> password:					
Location:	SQL Server					
To save to Micro username, and p	osoft SQL Server (Local), you must enter the server, bassword.					
<u>S</u> erver:	Server: (LOCAL)					
○ Use <u>W</u> indows Authentication						
Use SQL Se	erver Authentication					
<u>U</u> sername:	sa					
Password:	жикихихи					
	OK Cancel Help					

4. Right click on MyFirstPackage and select execute package option.

🚡 Console Root\Microsoft SQL Servers\SQL	Server Gro	up\(LOCAL) (Window	s NT)\Data Tra	ansforma 📕	
Console Root	Local Pac	kages 3 Items			
Microsoft SQL Servers SQL Server Group SQL Server Group Data Transformation Services Cocal Packages Meta Data Services Packag Meta Data Meta Data Amagement Replication Security Support Services Cocal Package Meta Data	Name A	Design Package Execute Package Schedule Package Versions Package Logs All Tasks Delete Help	Description	Owner VRWP01\A VRWP01\A VRWP01\A	Crea 10/1 10/4, 10/4,
	<u> </u>				

5. Click OK. Then Click Done.

Executing Package
E Contraction of the second
Status: Step Package Execution Results Step Successfully completed execution of package.
▲ Done Help

6. Click on Tools Menu->Data Transformation Services->Import Data.



7. Click Next.



8. Select **Text File** as data source and click Next.

Choose a Data Source From where do you wa following sources.	and ant to copy data? You can copy data from one of the	×
Data Source:	Microsoft OLE DB Provider for SQL Server Microsoft Text-Treiber (*.txt; *.csv) Microsoft Visual FoxPro-Treiber Microsoft Visual FoxPro Driver Other (DDBC Data Source) Paradox 3.X Paradox 3.X Paradox 5.X Text File Kefresh Advanced	*
	< Back Next > Cancel	Help

9. Select text file which you want to import and then click Open.

🍓 DTS Import/Export Wizard	×
Choose a Data Source From where do you want to copy data? You can copy data from one of the following sources.	Q
Data Source: 📄 Text File	•
Text files can be delimited or fixed field. To connect, you must selec	at a file.
<u>File name:</u>	
< <u>B</u> ack <u>N</u> ext > Cancel	Help

10. You can create a text on your system containing sample data. You can enter data about customers or students etc. Select the file and click Open. Then click **Next.**

😽 DTS Import/E	xport Wizard			×	
Select File					<u>?</u> ×
Look jn:	🞯 Desktop		🖬 🗗 🖬	•	
My Recent Documents Desktop My Documents My Computer	My Documents My Computer Creating cube Sql2ksp3 SQL Server 200 UtraISO.PE.v9 VUTES 7.0_Pac Web Extension Win Srv 2K3 En Etst.txt	ces)0 .3.6.2750.Retail.Silent.Install kage s t-VU LHR			
My Network	File <u>n</u> ame:	SampleFile.txt	-	<u>0</u> F	ien
Flaces	Files of type:	Text Files (*.txt,*.csv)	•	Car	ncel

11. Click Next.

🍓 DTS Import/Export Wizar	d	×
Select file format To import the data, confir are correctly detected be	m the source file format. Confirm fore proceeding.	that the file properties
© Eelimit © Fixed fi	ed. The columns are separated b ield. Information is aligned into co	y any character(s)
File <u>type</u> :	ANSI Skip	rows: 0 🗧
<u>R</u> ow delimiter:	(CR){LF}	jirst row has column names
Text gualifier:	Double Quote {"}	
Preview of file: D:\Docu	uments and	
1,"Eldon Base 2,"1.7 Cubic 3,"Cardinal S	for stackable sto Foot Compact ""Cul lant-D® Ring Binde	orage shelf, l▲ be"" Office R∉ er, Heavy Gau{y
•		F
	< <u>B</u> ack <u>N</u> ext >	Cancel Help

12. Click Next. Select Column Delimiter which in this case is comma. Click Next.

comma, tab, ser	micolon, or another charac	oter.		
Comma	O I ab O S	emicolon	◯ <u>O</u> ther:	
Preview: Col001	Col002	Col003	Col004	-
1	Eldon Base for st	Muhammed Mac	3	
2	1.7 Cubic Foot C	Barry French	293	
3	Cardinal Slant-D	Barry French	293	
4	R380	Clay Rozendal	483	
5	Holmes HEPA Ai	Carlos Soltero	515	
6	G.E. Longer-Life	Carlos Soltero	515	
7	Angle-D Binders	Carl Jackson	613	
8	SAFCO Mobile D	Carl Jackson	613	
9	SAFCO Commer	Monica Federle	643	-
•				

13. Choose destination and click Next.

🔩 DTS Imp	ort/Export Wizard		×
Choose Towi destir	a destination here do you want to co hations.	py data? You can copy data to one of the following	Q D
Destin	nation:	crosoft OLE DB Provider for SQL Server	•
Ē.	To connect to Mic name, and passwo	crosoft SQL Server, you must specify the server, user ord.	
	<u>S</u> erver:	(LOCAL)	•
	O Use <u>W</u> indows	Authentication	
	Use SQL Serv	ver Authentication	
	<u>U</u> sername:	sa	
	Password:	******	
	<u>D</u> atabase:	Image: Second	
		< Back Next > Cancel	Help

14. Select Create destination table option and create OK.

Destinal	D:\D tion: [mas)ocuments an ster].[dbo].[Sar	d npleFile]				
Column Mappings Transformations							
 Create destination table 			Edit	<u>s</u> ql	1		
Column	the standard state	I.					
	s in destination tab	ne	і <u>р</u> гор а	na recr	eate destina	ation table	•
C Append rows to destination table							
C Append roy	vs to destination ta	able	🔲 Enable	identit	y insert		
C Append rov	vs to destination ta	able	🔲 Enable	identitį	y insert		
C Append rov Mappings:	vs to destination ta	Tupe	Nullable	identity Size	y insert	Scale	
Append rov <u>Mappings:</u> Source Col001	Destination ta	Type	Nullable	identity Size	y insert Precision	Scale	
C Append rov <u>Mappings:</u> Source Col001 Col002	Destination ta	ble Type varchar	Nullable	identity Size 8000	y insert Precision	Scale	
C Append rov Mappings: Source Col001 Col002 Col003	Destination ta	Type varchar varchar varchar	Nullable	identity Size 8000 8000	y insert	Scale	
C Append rov Mappings: Source Col001 Col002 Col003 Col004	Destination ta Destination Col001 Col002 Col003 Col004	Type varchar varchar varchar varchar	Nullable	Size 8000 8000 8000 8000	y insert	Scale	
C Append rov Mappings: Source Col001 Col002 Col003 Col004 Col005	Destination ta Destination Col001 Col002 Col003 Col004 Col005	Type varchar varchar varchar varchar varchar	Nullable	identity Size 8000 8000 8000 8000 8000	Precision	Scale	
C Append rov Mappings: Source Col001 Col002 Col003 Col004 Col005 Col005 Col005	Destination ta Destination Col001 Col002 Col003 Col004 Col005 Col005	Type varchar varchar varchar varchar varchar varchar	Nullable	identity Size 8000 8000 8000 8000 8000	Precision	Scale	
C Append rov Mappings: Source Col001 Col002 Col003 Col004 Col005 Col005 Col005 Col005	Destination ta Destination Col001 Col002 Col003 Col004 Col005 Col005	Type varchar varchar varchar varchar varchar	Nullable	Size 8000 8000 8000 8000 8000 8000	Precision	Scale	
C Append rov Mappings: Source Col001 Col002 Col003 Col004 Col005 Col005 Col005 Source column	bestination ta Destination Col001 Col002 Col003 Col004 Col005 Col005 Col005 Col005	Type varchar varchar varchar varchar varchar varchar varchar varchar	Nullable	Size 8000 8000 8000 8000 8000 8000	Precision	Scale	

🐝 DTS	5 Import/Export Wiza	rd		×
Se : :	lect Source Tables a You can choose one or and data as it appears in ActiveX scripts.	nd Views more tables or views to cop n the source or click () to tr	y. You can copy the schema ransform the data using	B
:	\underline{T} able(s) and View(s):			
-	Source	Destination	Transform	
:	🔽 🔐 D:\Document	s a 🛅 [master].[dbo].[\$	SampleFile]	
r				
i				
	1			
	<u>S</u> elect All	Deselect All Pre	eview	
		< <u>B</u> ack <u>N</u> ext	> Cancel	Help

15. Click **Preview....** to view data.

Source:	D:\Documents and				6
Col001	Col002	Col003	Col004		
1	Eldon Base for st	Muhammed Mac	3		
2	1.7 Cubic Foot C	Barry French	293		h
3	Cardinal Slant-D	Barry French	293		
4	R380	Clay Rozendal	483		
5	Holmes HEPA Ai	Carlos Soltero	515		
6	G.E. Longer-Life	Carlos Soltero	515		
7	Angle-D Binders	Carl Jackson	613		
8	SAFCO Mobile D	Carl Jackson	613		
9	SAFCO Commer	Monica Federle	643		
10	Xerox 198	Dorothy Badders	678	-	
•					
		01		Help	

16. Click Next.

When	
Run immediately	Use replication to publish destination data
🔲 Sched <u>u</u> le DTS package f	for later execution
Occurs every 1 day(s), at	(12:00:00 AM.
Save	
Save DTS Package	I S <u>Ω</u> L Server
2	C SQL Server Meta Data Services
	C Structured Storage Eile

17. Select Run immediately and click Next.

DTS Import/Export Wizard	×
Save, schedule, and replicate p Specify if you want to save this D schedule the package to be exec	Dackage TS package. You may also replicate the data or suited at a later time.
-When	
Run immediately	Use replication to publish destination data
🔲 Sched <u>u</u> le DTS package fo	or later execution
Occurs every 1 day(s), at	12:00:00 AM.
Save	C SOL Server
Save DTS Package	C SQL Server Meta Data Services
	C Structured Storage Elle
	C ⊻isual Basic File
< <u>B</u>	ack <u>N</u> ext> Cancel Help

18. Click Finish.

🙀 DTS Import/Export Wizard	1	X
	Completing the DTS Import/Export Wizard	
	copy, transform, or transfer the data. Review the selections below and click on FINISH to execute	
	<u>S</u> ummary:	
	Source: Flat File Using DTS Flat File Provider Location: D:\Documents and Settings\Administrator.VR' Destination: Microsoft SQL Server Using Microsoft OLE DB Provider for SQL Server Location: (LOCAL) Database: master Tables	
	< Back Finish Cancel Help	

19. Click OK. Then Click Done.

-	Executing Package		
ř	Flat File	(iz.)	Microsoft SQL Server
	Progress:		
	DTS Import/Export Wiz	ard	×
K	Sta	opied 1 table(s) from Flat Fil	e to Microsoft SQL Server.
		OK	
	, <u>, , , , , , , , , , , , , , , , , , </u>	_	Done

20. View your table in master table node.

Console Root	Tables 61 Items			
Microsoft SQL Servers	Name A	Owner	Туре	C
E SQL Server Group	MSreplication_options	dbo	System	8
Databases	🗐 SampleFile	dbo	User	1
	spt_datatype_info	dbo	System	8
master	🔢 📰 spt_datatype_info_ext	dbo	System	8
Gan Tables	🗾 📰 spt_fallback_db	dbo	System	8
	📰 spt_fallback_dev	dbo	System	8
	📕 🧮 spt_fallback_usg	dbo	System	8
	🔲 📰 spt_monitor	dbo	System	8
Users	spt_provider_types	dbo	System	8
Roles	spt_server_info	dbo	System	8
Rules	spt_values	dbo	System	8
Defaults	📰 sysaltfiles	dbo	System	8
Service Defined Data	syscacheobjects	dbo	System	8
	🗾 📰 syscharsets	dbo	System	8
	- svscolumos	dho	System	8

21. To view data, right click on table SampleFile, click on Open Table->Return all rows.

Console Root\Microsoft SQL Servers\SQL	Server Group\(LOCA	L) (Windows NT)\Databa	ses\mast	<u> </u>
Console Root	Tables 61 Items			
Microsoft SQL Servers	Name A	Owner	Туре	C 🔺
E SQL Server Group	MSreplication_optio	ns dbo	System	8,
	🔟 SampleFile	dba	User	1
	📰 spt datatype info	New Ta <u>b</u> le	System	8.
master	spt datatype info.	De <u>s</u> ign Table	Einstein	
	🗐 sot fallback db	Open Table 🔹 🕨	Return <u>a</u> ll r	ows
- Marchael	🗐 spt fallback dev	Full-Text Index Table 🕨	Return <u>T</u> op)
Stored Procedures	sot fallback uso		Query	
Extended Stored Pr	spt monitor	All Tas <u>k</u> s 🕨 🕨	System	8.
Users	spt provider type	Out	System	8.
Roles	spt server info	Cogu	System	8.
Rules	spt values	<u>C</u> opy Delete	System	8.
- Defaults	📰 sysaltfiles	Depere	System	8.
🔤 🖾 User Defined Data 🐪	syscacheobiects	Kena <u>m</u> e	System	8
- 🔩 User Defined Funct	syscharsets	P <u>r</u> operties	System	8.
🔄 📔 Full-Text Catalogs 👻	Syscolumps		System	8.
	•	Help		•

22. This is data of table SampleFile.

7	🖥 Data in Table 'SampleFile' in 'master' on '(LOCAL)'						
	Col001	Col002	Col003	Col004	Col005	Col006	Col007
►	1	Eldon Base for stac	Muhammed MacInt [,]	3	-213.25	38.94	35
	2	1.7 Cubic Foot Corr	Barry French	293	457.81	208.16	68.02
	3	Cardinal Slant-D® I	Barry French	293	46.71	8.69	2.99
	4	R380	Clay Rozendal	483	1198.97	195.99	3.99
	5	Holmes HEPA Air Pu	Carlos Soltero	515	30.94	21.78	5.94
	6	G.E. Longer-Life In	Carlos Soltero	515	4.43	6.64	4.95
	7	Angle-D Binders wit	Carl Jackson	613	-54.04	7.3	7.72
	8	SAFCO Mobile Desk	Carl Jackson	613	127.70	42.76	6.22
	9	SAFCO Commercial	Monica Federle	643	-695.26	138.14	35
	10	Xerox 198	Dorothy Badders	678	-226.36	4.98	8.33
*							
	-						

Mechanism to Conduct Lab:

Students and teacher communicate through Adobe Connect.

BASIC SORTED NEIGHBORHOOD (BSN) METHOD

PatientID	Patient Name	Patient Contact No	Address
F101	Sana	051-123456	Islamabad
M102	Ali	051-123456	Lahore
F103	Aliya		Rawalpindi
F104	Hina		Faisalabad
F103	Aliaa		Rawalpindi
F101	Sanah		Islamabad
M105	Hassan		Karachi

Question:

Consider the above table and apply all three steps (**Create key, Sort the data, Merge)** of Basic Sorted Neighborhood (BSN) method to find out the duplicate records in the table. Records will be considered duplicate if the value of column "<u>PatientID</u>" is same in this table data.

The steps which you have to follow are:

Key:

Key will consist of one character from "**PatientID**", then first three characters from "**Patient Name**" and then first two characters from "**Address**" column.

Step 1: Create key

In step-1, you will create the key according to the rules as mentioned above against each record. Add an extra column at the end of the table to show the new key created against each record.

Step 2: Sort the data

In step-2, you will sort the record on the basis of key which you created in step-1.

Step 3: Merge

In step-3, consider the window size (w) equal to two (2) and identify the similar records on the basis of sorted key.

Solution:

Step 1: Create key

Key is created in first step as per rules given in question statement.

PatientID	Patient Name	Patient Contact No	Address	Кеу
F01	Sana	051-123456	Islamabad	FSanIs
M01	Ali	051-123456	Lahore	MAliLa
F02	Aliya		Rawalpindi	FAliRa
F03	Hina		Faisalabad	FHinFa
F02	Aliaa		Rawalpindi	FAliRa
F01	Sanah		Islamabad	FSanIs
M02	Hassan		Karachi	MHasKa

Step 2: Sort the data

Now sort the records based on the above created keys

PatientID	Patient Name	Patient	Address	Кеу
		Contact No		
F02	Aaliya		Rawalpindi	FAalRa
F02	Aaliaa		Rawalpindi	FAalRa
F03	Hina		Faisalabad	FHinFa
F01	Sana	051-123456	Islamabad	FSanIs
F01	Sanah		Islamabad	FSanIs
M01	Ali	051-123456	Lahore	MAliLa
M02	Hassan		Karachi	MHasKa

Step 3: Merge

Duplicate/Identical keys

PatientID	Patient Name	Patient Contact No	Address	Кеу
F02	Aaliya		Rawalpindi	FAalRa
F02	Aaliaa		Rawalpindi	FAalRa

F01	Sana	051-123456	Islamabad	FSanIs
F01	Sanah		Islamabad	FSanIs

Mechanism to Conduct Lab:

Students and teacher communicate through Adobe Connect.

Lab 8

DATA QUALITY RULES

"Pak Airline" is running in more than 15 countries and maintaining its data in a database. They have also started using data warehouse. They have different policies for running their business. One of their policies is that, if the airline itself cancels the flight due to weather or any other technical reason, then they have to payback 100% amount to customers and if flight is canceled by passenger then 40% amount will be returned. The Payback codes 'FP' or 'PP' are put in the payment claims to identify whether the claim is for full or partial payment respectively. Another policy is that allowed luggage weight per passenger is 25 kg.

There exists some specific data problems which are linked to business rules, and then generic and specific rule sets are developed for measuring how good the data is within an information system. These rule sets shows the data quality metrics in order to judge conformance of data according to these business rules.

Note: Here, Assume, amount of ticket is Rs. 15,000 per passenger.

Question:

Considering above scenario, apply data quality rules on the historical data problems faced by "Pak Airline". For this, fill the following table (Rule Type, Generic Rule Set and Specific Rule Set) against each historical data problems as given below.

Historical Data Problem	Rule Type	Generic Rule Set	Specific Rule Set
Payback amount is less than 100% when flight is canceled due to weather or other technical reasons.			
The field "reason_cancel" were often left blank.			
The luggage per passenger exceeds 25 kg.			
The payback code in payback claim is not 'FP' or 'PP' sometimes.			

Solution:

Historical Data Problem	Rule Type	Generic Rule Set	Specific Rule Set
Payback amount is less than 100% when flight is canceled due to weather or other technical reasons.	Business Rule	If total payback amount is less than Rs. 15,000 and reason of flight cancelation is weather or technical, then, error.	Select total_payback_amount from Claim where total_payback_amount <15000 and reason_cancel='weather' or reason_cancel='technical';
The field "reason_cancel" were often left blank.	Null Constraints	If reason_cancel is blank or null, then, error	Select reason_cancel from Claim where reason_cancel= ' ' or reason_cancel =NULL;
The luggage per passenger exceeds 25 kg.	Operational Rule	If luggage_weight is greater than 25kg, then, error.	Select luggage_weight from Passenger where luggage_weight>25; Note: Assume that weight in field luggage_weight is input as kg.
The payback code in payback claim is not 'FP' or 'PP' sometimes.	Domain Validation	If payback_code is not 'FP' or 'PP', then, error.	Select payback_code from Claim where NOT (payback_code= 'FP' or 'PP');

Mechanism to Conduct Lab:

Students and teacher communicate through Adobe Connect.

KEY RANGE PARTITIONING

PART I

Consider the subset of dataset data taken from <u>https://community.tableau.com/docs/DOC-1236</u>.

You are required to perform range partitioning on this data into yearly partitions of ship date.

Solution:

Partition 1 of year 2014:

			Customer	Customer			
Order ID	Order Date	Ship Date	ID	Name	State	Region	Product ID
CA-2014-				Brosina			FUR-FU-
115812	6/9/2014	6/14/2014	BH-11710	Hoffman	California	West	10001487
CA-2014-				Brosina			OFF-AR-
115812	6/9/2014	6/14/2014	BH-11710	Hoffman	California	West	10002833
CA-2014-				Brosina			
115812	6/9/2014	6/14/2014	BH-11710	Hoffman	California	West	TEC-PH-10002275
CA-2014-				Brosina			
115812	6/9/2014	6/14/2014	BH-11710	Hoffman	California	West	OFF-BI-10003910
CA-2014-				Brosina			OFF-AP-
115812	6/9/2014	6/14/2014	BH-11710	Hoffman	California	West	10002892
CA-2014-				Brosina			FUR-TA-
115812	6/9/2014	6/14/2014	BH-11710	Hoffman	California	West	10001539
CA-2014-				Brosina			
115812	6/9/2014	6/14/2014	BH-11710	Hoffman	California	West	TEC-PH-10002033
CA-2014-							
105893	11/11/2014	11/18/2014	PK-19075	Pete Kriz	Wisconsin	Central	OFF-ST-10004186
CA-2014-				Alejandro			
167164	5/13/2014	5/15/2014	AG-10270	Grove	Utah	West	OFF-ST-10000107
CA-2014-				Zuschuss			OFF-AR-
143336	8/27/2014	9/1/2014	ZD-21925	Donatelli	California	West	10003056
CA-2014-				Zuschuss			
143336	8/27/2014	9/1/2014	ZD-21925	Donatelli	California	West	TEC-PH-10001949
CA-2014-				Zuschuss			
143336	8/27/2014	9/1/2014	ZD-21925	Donatelli	California	West	OFF-BI-10002215

Lab 9

Order ID	Order Date	Ship Date	Customer ID	Customer Name	State	Region	Product ID
CA-2016-							
152156	11/8/2016	11/11/2016	CG-12520	Claire Gute	Kentucky	South	FUR-BO-10001798
CA-2016-							
152156	11/8/2016	11/11/2016	CG-12520	Claire Gute	Kentucky	South	FUR-CH-10000454
CA-2016-							
138688	6/12/2016	6/16/2016	DV-13045	Darrin Van Huff	California	West	OFF-LA-10000240
US-2015-							
108966	10/11/2015	10/18/2015	SO-20335	Sean O'Donnell	Florida	South	FUR-TA-10000577
US-2015-	40/44/2045	10/10/2015	60 20225		et a data		055 ST 40000760
108966	10/11/2015	10/18/2015	50-20335	Sean O'Donnell	Florida	South	OFF-ST-10000760
CA-2014-	C /0 /201 4	C /1 A /201 A		Dresing Hoffman	California	Mont	
115812	6/9/2014	6/14/2014	BH-11/10	Brosina Horiman	California	west	FUK-FU-10001487
CA-2014-	6/0/2014	6/14/2014	BH-11710	Brosina Hoffman	California	West	OFE-AP-10002822
CA_2014_	0/9/2014	0/14/2014	BII-11/10		California	VVESL	011-AII-10002855
115812	6/9/2014	6/14/2014	BH-11710	Brosina Hoffman	California	West	TEC-PH-10002275
CA-2014-	0/5/2014	0/14/2014	51111/10		Camornia	WCSC	12011110002275
115812	6/9/2014	6/14/2014	BH-11710	Brosina Hoffman	California	West	OFF-BI-10003910
CA-2014-	0,0,2021	0,, _ 0					
115812	6/9/2014	6/14/2014	BH-11710	Brosina Hoffman	California	West	OFF-AP-10002892
CA-2014-	, , ,						
115812	6/9/2014	6/14/2014	BH-11710	Brosina Hoffman	California	West	FUR-TA-10001539
CA-2014-							
115812	6/9/2014	6/14/2014	BH-11710	Brosina Hoffman	California	West	TEC-PH-10002033
CA-2017-					North		
114412	4/15/2017	4/20/2017	AA-10480	Andrew Allen	Carolina	South	OFF-PA-10002365
CA-2016-							
161389	12/5/2016	12/10/2016	IM-15070	Irene Maddox	Washington	West	OFF-BI-10003656
US-2015-							
118983	11/22/2015	11/26/2015	HP-14815	Harold Pawlan	Texas	Central	OFF-AP-10002311
US-2015-							
118983	11/22/2015	11/26/2015	HP-14815	Harold Pawlan	Texas	Central	OFF-BI-10000756
CA-2014-	11/11/2011	11/10/2011	DK 40075	Deterio			055 ST 4000 4405
105893	11/11/2014	11/18/2014	PK-19075	Pete Kriz	Wisconsin	Central	OFF-ST-10004186
CA-2014-	E /12/2014	E /1E /2014	AC 10270	Alajandra Crova	litab	West	OFF ST 10000107
CA 2014	5/15/2014	5/15/2014	AG-10270	Alejanuro Grove	Utan	west	0FF-31-10000107
LA-2014-	8/27/2014	9/1/2014	70-21025	Zuschuss Donatelli	California	Wost	OFF-AR-10003056
CA-2014-	0/2//2014	5/1/2014	20 21525	Zuschuss Donatem	Camornia	WCSC	OTT AN 10005050
143336	8/27/2014	9/1/2014	ZD-21925	Zuschuss Donatelli	California	West	TFC-PH-10001949
CA-2014-	0, _ / , _ 0 _ 1	0, _, _0					
143336	8/27/2014	9/1/2014	ZD-21925	Zuschuss Donatelli	California	West	OFF-BI-10002215
CA-2016-							
137330	12/9/2016	12/13/2016	KB-16585	Ken Black	Nebraska	Central	OFF-AR-10000246
CA-2016-							
137330	12/9/2016	12/13/2016	KB-16585	Ken Black	Nebraska	Central	OFF-AP-10001492
US-2017-							
156909	7/16/2017	7/18/2017	SF-20065	Sandra Flanagan	Pennsylvania	East	FUR-CH-10002774

Partition 2 of year 2015:

Order ID	Order Date	Ship Date	Customer ID	Customer Name	State	Region	Product ID
US-2015- 108966	10/11/2015	10/18/2015	SO-20335	Sean O'Donnell	Florida	South	FUR-TA- 10000577
US-2015- 108966	10/11/2015	10/18/2015	SO-20335	Sean O'Donnell	Florida	South	OFF-ST- 10000760
US-2015- 118983	11/22/2015	11/26/2015	HP-14815	Harold Pawlan	Texas	Central	OFF-AP- 10002311
US-2015- 118983	11/22/2015	11/26/2015	HP-14815	Harold Pawlan	Texas	Central	OFF-BI- 10000756

Partition 3 of year 2016:

	Order		Customer	Customer			
Order ID	Date	Ship Date	ID	Name	State	Region	Product ID
CA-2016-							FUR-BO-
152156	11/8/2016	11/11/2016	CG-12520	Claire Gute	Kentucky	South	10001798
CA-2016-							FUR-CH-
152156	11/8/2016	11/11/2016	CG-12520	Claire Gute	Kentucky	South	10000454
CA-2016-							OFF-LA-
138688	6/12/2016	6/16/2016	DV-13045	Darrin Van Huff	California	West	10000240
CA-2016-							OFF-BI-
161389	12/5/2016	12/10/2016	IM-15070	Irene Maddox	Washington	West	10003656
CA-2016-							OFF-AR-
137330	12/9/2016	12/13/2016	KB-16585	Ken Black	Nebraska	Central	10000246
CA-2016-							OFF-AP-
137330	12/9/2016	12/13/2016	KB-16585	Ken Black	Nebraska	Central	10001492

	Order		Customer	Customer			
Order ID	Date	Ship Date	ID	Name	State	Region	Product ID
CA-2017-					North		OFF-PA-
114412	4/15/2017	4/20/2017	AA-10480	Andrew Allen	Carolina	South	10002365
US-2017-				Sandra			FUR-CH-
156909	7/16/2017	7/18/2017	SF-20065	Flanagan	Pennsylvania	East	10002774
PART II

Perform list portioning by Partitioning by following sales table by region. You can use values of region from table given in **Question 1**.

Solution:

First you will define values for region partitions.

Region partitions will be **East**, **West**, **South** and **Central Partition**.

List of East partition values: Pennsylvania

List of West Partition Values: Utah, California, Washington

List of South Partition Values: Florida, Kentucky, North Carolina

List of Central Partition Values: Wisconsin, Texas, Nebraska

South Partition:

Order ID	Order Date	Ship Date	Customer ID	Customer Name	State	Product ID
CA-2016-						FUR-BO-
152156	11/8/2016	11/11/2016	CG-12520	Claire Gute	Kentucky	10001798
CA-2016-						FUR-CH-
152156	11/8/2016	11/11/2016	CG-12520	Claire Gute	Kentucky	10000454
US-2015-						FUR-TA-
108966	10/11/2015	10/18/2015	SO-20335	Sean O'Donnell	Florida	10000577
US-2015-						OFF-ST-
108966	10/11/2015	10/18/2015	SO-20335	Sean O'Donnell	Florida	10000760
CA-2017-					North	OFF-PA-
114412	4/15/2017	4/20/2017	AA-10480	Andrew Allen	Carolina	10002365

Central Partition:

Order ID	Order Date	Ship Date	Customer ID	Customer Name	State	Region	Product ID
US-2017-							FUR-CH-
156909	7/16/2017	7/18/2017	SF-20065	Sandra Flanagan	Pennsylvania	East	10002774

Order ID	Order Date	Ship Date	Customer ID	Customer Name	State	Product ID
CA-2016- 152156	11/8/2016	11/11/2016	CG-12520	Claire Gute	Kentucky	FUR-BO-10001798
CA-2016- 152156	11/8/2016	11/11/2016	CG-12520	Claire Gute	Kentucky	FUR-CH-10000454
CA-2016- 138688	6/12/2016	6/16/2016	DV-13045	Darrin Van Huff	California	OFF-LA-10000240
US-2015- 108966	10/11/2015	10/18/2015	SO-20335	Sean O'Donnell	Florida	FUR-TA-10000577
US-2015- 108966	10/11/2015	10/18/2015	SO-20335	Sean O'Donnell	Florida	OFF-ST-10000760
CA-2014- 115812	6/9/2014	6/14/2014	BH-11710	Brosina Hoffman	California	FUR-FU-10001487
CA-2014- 115812	6/9/2014	6/14/2014	BH-11710	Brosina Hoffman	California	OFF-AR-10002833
CA-2014- 115812	6/9/2014	6/14/2014	BH-11710	Brosina Hoffman	California	TEC-PH-10002275
CA-2014- 115812	6/9/2014	6/14/2014	BH-11710	Brosina Hoffman	California	OFF-BI-10003910
CA-2014- 115812	6/9/2014	6/14/2014	BH-11710	Brosina Hoffman	California	OFF-AP-10002892
CA-2014- 115812	6/9/2014	6/14/2014	BH-11710	Brosina Hoffman	California	FUR-TA-10001539
CA-2014- 115812	6/9/2014	6/14/2014	BH-11710	Brosina Hoffman	California	TEC-PH-10002033
CA-2017- 114412	4/15/2017	4/20/2017	AA-10480	Andrew Allen	North Carolina	OFF-PA-10002365
CA-2016- 161389	12/5/2016	12/10/2016	IM-15070	Irene Maddox	Washington	OFF-BI-10003656
US-2015- 118983	11/22/2015	11/26/2015	HP-14815	Harold Pawlan	Texas	OFF-AP-10002311

US-2015- 118983	11/22/2015	11/26/2015	HP-14815	Harold Pawlan	Texas	OFF-BI-10000756
CA-2014-		44/40/2044		5		055 67 1000 1106
105893	11/11/2014	11/18/2014	PK-19075	Pete Kriz	Wisconsin	OFF-ST-10004186
CA-2014-						
167164	5/13/2014	5/15/2014	AG-10270	Alejandro Grove	Utah	OFF-ST-10000107
CA-2014-				Zuschuss		
143336	8/27/2014	9/1/2014	ZD-21925	Donatelli	California	OFF-AR-10003056
CA-2014-				Zuschuss		
143336	8/27/2014	9/1/2014	ZD-21925	Donatelli	California	TEC-PH-10001949
CA-2014-				Zuschuss		
143336	8/27/2014	9/1/2014	ZD-21925	Donatelli	California	OFF-BI-10002215
CA-2016-						
137330	12/9/2016	12/13/2016	KB-16585	Ken Black	Nebraska	OFF-AR-10000246
CA-2016-						
137330	12/9/2016	12/13/2016	KB-16585	Ken Black	Nebraska	OFF-AP-10001492
US-2017-						
156909	7/16/2017	7/18/2017	SF-20065	Sandra Flanagan	Pennsylvania	FUR-CH-10002774

West Partition:

Order ID	Order Date	Ship Date	Customer ID	Customer Name	State	Product ID
CA-2016-						
138688	6/12/2016	6/16/2016	DV-13045	Darrin Van Huff	California	OFF-LA-10000240
CA-2014-						
115812	6/9/2014	6/14/2014	BH-11710	Brosina Hoffman	California	FUR-FU-10001487
CA-2014-						
115812	6/9/2014	6/14/2014	BH-11710	Brosina Hoffman	California	OFF-AR-10002833
CA-2014-						
115812	6/9/2014	6/14/2014	BH-11710	Brosina Hoffman	California	TEC-PH-10002275
CA-2014-						
115812	6/9/2014	6/14/2014	BH-11710	Brosina Hoffman	California	OFF-BI-10003910
CA-2014-						
115812	6/9/2014	6/14/2014	BH-11710	Brosina Hoffman	California	OFF-AP-10002892
CA-2014-						
115812	6/9/2014	6/14/2014	BH-11710	Brosina Hoffman	California	FUR-TA-10001539

CA-2014-						
115812	6/9/2014	6/14/2014	BH-11710	Brosina Hoffman	California	TEC-PH-10002033
CA-2016-						
161389	12/5/2016	12/10/2016	IM-15070	Irene Maddox	Washington	OFF-BI-10003656
CA-2014-						
167164	5/13/2014	5/15/2014	AG-10270	Alejandro Grove	Utah	OFF-ST-10000107
CA-2014-						
143336	8/27/2014	9/1/2014	ZD-21925	Zuschuss Donatelli	California	OFF-AR-10003056
CA-2014-						
143336	8/27/2014	9/1/2014	ZD-21925	Zuschuss Donatelli	California	TEC-PH-10001949
CA-2014-						
143336	8/27/2014	9/1/2014	ZD-21925	Zuschuss Donatelli	California	OFF-BI-10002215

East Partition:

Order ID	Order Date	Ship Date	Customer ID	Customer Name	State	Product ID
US-2015-						
118983	11/22/2015	11/26/2015	HP-14815	Harold Pawlan	Texas	OFF-AP-10002311
US-2015-						
118983	11/22/2015	11/26/2015	HP-14815	Harold Pawlan	Texas	OFF-BI-10000756
CA-2014-						
105893	11/11/2014	11/18/2014	PK-19075	Pete Kriz	Wisconsin	OFF-ST-10004186
CA-2016-						
137330	12/9/2016	12/13/2016	KB-16585	Ken Black	Nebraska	OFF-AR-10000246
CA-2016-						
137330	12/9/2016	12/13/2016	KB-16585	Ken Black	Nebraska	OFF-AP-10001492

Mechanism to Conduct Lab:

Students and teacher communicate through Adobe Connect.

Lab 10

CLUSTER INDEX

"Pak Airline" is an airliner reservation company, which is operating in more than 10 countries. They have developed the airline reservation system to avoid the errors faced in manual system. The staff of the airline use airline reservation system form the tasks such as flight scheduling, ticket reservation, announcements in automated way. Similarly, users/passengers can search for flight schedule according to date and time and fare details. The staff of the airline can manage the reservation systems by flight rout, runway details, flight scheduling and reservation.

Ticket reservation system of the Pak Airline provides the information about schedule of flights, availability of seats, flight number and destination. For reservation of ticket user have to provide its personal information such as name, age, address etc. For payment purpose user will provide credit card number and bank details. Moreover, information about flight number, date of departure, no. of tickets to be booked is also required for confirmation of ticket.

Question Statement:

You are required to create Cluster Index based on destination city, Designation of staff, and address of passenger. For this consider the following tables and apply Cluster indexing technique on required columns.

<u>Airport</u>	Flight	Arrival	Arrival	Depart	Depart	Destination	Aiport Name	City
<u>Id</u>	<u>Id</u>	date	time	Date	Time			
AP1	Pk01	15-10-	12:30	16-10-	10:00	United	Allama Iqbal Airport	Lahore
		2018	PM	2018	AM	Kingdom		
AP2	Pk02	19-11-	10:00	19-11-	04:00	Dubai	Baynazir Airport	Rawalpindi
		2018	AM	2018	PM			
AP3	Bg01	16-10-	12:30	16-10-	06:30	KSA	New Islamabad	Islamabad
		2018	PM	2018	PM		Airport	
Ap4	Pk03	17-10-	09:00	17-10-	04:00	United	Jinnah Airport	Karachi
		2018	PM	2018	PM	Kingdom		
Ap5	Bg02	25-10-	08:00	26-10-	12:00	USA	New Lahore Airport	Lahore
		2018	AM	018	PM			
Ap6	Bg03	19-11-	10:00	19-11-	04:00	Dubai	International Quetta	Quetta
		2018	AM	2018	PM		Airport	

Aiport-Flight Table

Staff Table

Staff Id	Name	Designation	Contact No
St01	Micheal	Operation Agent	0092-345-7865439
St02	Jackob	Flight Attendant	0062-876-0987654
St03	Alfard	Avionic Technician	0092-321-9865321
St04	Jackson	Operation Agent	0072-098-7854321
St05	Joseph	Flight Dispatcher	0092-333-9054213

St06	Joliana	Passenger Service	0062-900-6789012
		Agent	
St07	Thomsan	Flight Attendant	0052-321-9084563

Passenger Table

Passenger Id	Name	Contact No	Address	Email	Credit
					Cr.No
Ps01	Julia	0092-345-7865439	Islamabad	julia@gmail.com	123-987
Ps02	Alexandra	0062-876-0987654	Dubai	alex@gmail.com	324-908
Ps03	Robert	0092-321-9865321	London	robert@live.com	457-975
Ps04	Alaf	0072-098-7854321	Islamabad	alaf@yahoo.com	345-075
Ps05	Julia Sanf	0092-333-9054213	New York	jausanf@gmail.com	123-890
Ps06	Charistea	0062-900-6789012	London	charist@live.com	768-054

Solution:

Cluster index on **Destination** column

<u>Airport</u>	<u>Flight</u>	Arrival	Arrival	Depart	Depart	Destination	Aiport Name	City
<u>Id</u>	<u>Id</u>	date	time	Date	Time			
AP2	Pk02	19-11-	10:00	19-11-	4:00 PM	Dubai	Baynazir Airport	Rawalpindi
		2018	AM	2018				
Ap6	Bg03	19-11-	10:00	19-11-	4:00 PM	Dubai	International Quetta	Quetta
		2018	AM	2018			Airport	
AP3	Bg01	16-10-	12:30	16-10-	16-10-	KSA	New Islamabad	Islamabad
		2018	PM	2018	2018		Airport	
AP1	Pk01	15-10-	12:30	16-10-	10:00	United	Allama Iqbal Airport	Lahore
		2018	PM	2018	AM	Kingdom		
Ap4	Pk03	17-10-	9:00	17-10-	4:00 PM	United	Jinnah Airport	Karachi
		2018	PM	2018		Kingdom		
Ap5	Bg02	25-10-	8:00	26-10-	12:00	USA	New Lahore Airport	Lahore
		2018	AM	018	PM			

Cluster index on **Designation** column

Staff Id	Name	Designation	Contact No
St03	Alfard	Avionic Technician	0092-321-9865321
St02	Jackob	Flight Attendant	0062-876-0987654
St07	Thomsan	Flight Attendant	0052-321-9084563
St05	Joseph	Flight Dispatcher	0092-333-9054213
St01	Micheal	Operation Agent	0092-345-7865439
St04	Jackson	Operation Agent	0072-098-7854321
St06	Joliana	Passenger Service Agent	0062-900-6789012

Cluster index on **Address** column

Passenger	Name	Contact No	Address	Email	Credit
<u>Id</u>					Cr.No
Ps01	Julia	0092-345-7865439	Islamabad	julia@gmail.com	123-987
Ps04	Alaf	0072-098-7854321	Islamabad	alaf@yahoo.com	345-075
Ps02	Alexandra	0062-876-0987654	Dubai	alex@gmail.com	324-908
Ps03	Robert	0092-321-9865321	London	robert@live.com	457-975
Ps06	Charistea	0062-900-6789012	London	charist@live.com	768-054
Ps05	Julia Sanf	0092-333-9054213	New York	jausanf@gmail.com	123-890

Mechanism to Conduct Lab:

Students and teacher communicate through Adobe Connect.

Lab 11

Nested Loop, Sort Merge, and Hash Join using SQL Server Query Analyzer

1. Open SQL Query Analyzer by clicking Programs->Microsoft SQL Server and Query Analyzer.



2. Select SQL Server **ISASERVER** and click oK.

Connect	to SQL Server	×
Ø	<u>S</u> QL Server:	ISASERVER
Connect	using: ndows authentica	tion
i su Lo	gin name:	
Pa	ssword:	
	<u>ОК</u>	Cancel Help

3. The following SQL Query Analyzer window will open. In this window, you will write SQL queries.

🝯 SQL Query Analyzer						
<u>File E</u> dit Query <u>T</u> ools <u>V</u>	<u>V</u> indow <u>H</u> elp					
🖆 🕶 🖬 🗐 🐰	• 6 ď #	$\bowtie \blacksquare \bullet \checkmark \flat =$	📋 ma	ster		🚏 🔍 📑 🖥
🛋 Query - ISASERVER.m	aster.VRWP01\Ad	ministrator - Untitled1				
4						
Ready	ISASERVER (8.0)	VRWP01\Administrator (63)	master	0:00:00	0 rows	Ln 1, Col 1

4. First select sample database Northwind.

🔞 SQL Query Analyzer						
<u>File Edit Query T</u> ools <u>W</u>	indow <u>H</u> elp					
] 🖆 🕶 🖨 🖶 🖁 🛛 🐰 ।	6 C M	$\bowtie \blacksquare \bullet \checkmark \flat ~ \blacksquare$		master	•	1 🐕 🔍 🖻
Query - ISASERVER.me	ister.¥R₩P01\Ad	lministrator - Untitled1		master model mothwind pubs SurfControl_We tempdb <details></details>	sbFilter	
4						
Ready	ISASERVER (8.0)	VRWP01\Administrator (63)	maste	r 0:00:00	0 rows	Ln 1, Col 1

5. Write following SQL query in query analyzer window.

select * from customers, orders where Customers.CustomerID=Orders.CustomerID;

Here customers and orders are table of Northwind database.

	🔨 SQL Query Analyzer		
l	Eile Edit Query Tools Window Help		
	🖆 → 🚅 🛃 🗐 ½ 🗈 🎕 🗹 🏘 ∽ 🎟 → ✔ 🕨 🔳 🚺 master	- 🗄 📽 🔍 🖆 🖬	
	Query - ISASERVER.master.VRWP01\Administrator - Untitled1*	_	
	select * from customers, orders where Customers.CustomerID	=Orders.CustomerID;	
	Ready ISASERVER (8.0) VRWP01\Administrator (63) master	er 0:00:00 0 rows Ln 1. Col 7	8 /

6. Execute SQL query by clicking **Query** menu and select **Execute** option. You can also click execute button available on standard toolbar.

🗃 SOL Ouery Analyzer			
File Edit Query Tools W	(indow Help		
📔 🔸 🔓 🚺 Change Datab	iase Ctrl+U	🗸 🕨 🗉 🚺 Northwind 💽 🚭 📽 🕰 🔛	
eelt Current Curren	Ctrl+FS Ing Query Alt+Break ted Execution Plan Ctrl+L Wizard Ctrl+I t Ctrl+T t Ctrl+D	or-Untitled1* ce Customers.CustomerID-Orders.CustomerID;	_
N Results to Elle	Ctrl+Shift+F in Plan Ctrl+K (race Ctrl+Shift+T tatistics Ctrl+Shift+S	_	
Current Conne	sction Properties	1	
•			Þ

7. Click on **Query** menu and select option **Display Estimated Execution Plan** to see execution plan of query. Estimated execution plan can also be viewed by using highlighted options in following figure.

🛒 SQL Query Analyzer	
Eile Edit Query Tools Window Help	
📔 🕶 🚡 📋 C <u>h</u> ange Database Ctrl+U	🗸 🕨 🗉 📋 Notthwind 💽 💮 📽 🔃 📰
Query Zarse Ctrl+F5	or - Untitled1*
Self Execute F5	ce Customers.CustomerID=Orders.CustomerID;
Display Estimated Execution Plan Ctrl+L	
Index Tuning Wizard Ctrl+I	
Quers 📰 Results in Text Ctrl+T	batch): 100.00%
Quers Results in Grid Ctrl+D	orders where Customers.CustomerID=Orders.Customer]
Results to File Ctrl+Shift+F	
Show Execution Plan Ctrl+K	- ॡ ← ♣
L Show Server Trace Ctrl+Shift+T	d Loope/In Customers PK Cu
Cos Ctrl+Shift+S	Cost: 3% Cost: 33%
Current Connection Properties	
	600
	Orders.Customer
	Cost: 10%
•	
Estimated Execution Plan	

8. By default, Nested loop join is performed, you are required to analyze execution plan.

🗃 SOL Query Analyzer	
<u>File Edit Query Tools Window Help</u>	
🖆 • 😅 🖬 🗐 🐰 ங 🛍 📽 📽 🛤 🕫 🎟 • 🗸 🕨 = 🚺 Northwind 💿 🚭 😽 🔍 🗉	a
Query - ISASERVER.Northwind.VRWP01\Administrator - Untitled1*	×
orders where Customers.CustomerID=Orders.CustomerID;	-
<u>د</u> ۲	-
Query 1: Query cost (relative to the batch): 100.00% Query text: select * from customers, orders where Customers CustomerID=Orc	ie:
	<u></u>
Cost: 0% Cost: 54% Cost: 3% Cost: 3%	
5th	
Orders.Customer Cost: 10%	
	F
Estimated Execution Plan D Messages	_

Merge Join Query:

For Merge Join, write following query in SQL Query Analyzer window.

SELECT * FROM customers C

INNER Merge JOIN orders O ON C.CustomerID = O.CustomerID



Hash Join Query:

For Merge Join, write following query in SQL Query Analyzer window.

SELECT * FROM customers C

INNER Hash JOIN orders O ON C.CustomerID = O.CustomerID



Lab Exercise: You are required to analyze Nested Loop Join, Merge Join and Hash Join for the given query in terms of which one is efficient in terms of execution time.

Mechanism to Conduct Lab:

Students and teacher communicate through Adobe Connect.

Lab 12

DWH IMPLEMENTATION: Goal Driven Approach

Ralph Kimball's Approach Part I

In this Lab, you have to perform the second part of the semester project which details have been provided in lecture 36. This lab deals with the DWH implementation life cycle that you have already studied in great detail in lectures 32-35.

In lectures 33and 34, you have studied the **Ralph Kimball's approach for data warehouse implementation**. You will use this approach in this lab. Note that, you are not required to do any development and deployment work. As you have studied in lectures that the DWH lifecycle road-map was divided into three parts, you only have to cover these parts i.e. (i)project planning (ii) user requirement definition and (iii) three parallel tracks. You are NOT required to discuss or do DWH deployment or do any analytics development.

In this lab, you are required to perform following two tasks:

Task 1: Identify Organization

- Do a complete data warehouse implementation life cycle study.
- Identify a large company/organization that is a prime candidate for a DWH.
- Prepare report_1 giving and explaining any four reasons for selecting a company.
- Submit report_1 and get the company/organization selected approved by the instructor before
 proceeding ahead.

Task 2: Project Planning

After identifying organization, perform the following task as part of Project Planning. For this,

- Prepare a questionnaire (at least 15 non-trivial questions).
- Identify and contact a key person who will help you.
- Prepare and Submit report_2

Your report should include following:

• Report No.

- Title of course, semester & submission date
- Your Name and roll no.
- Campus and name of city.
- Table of contents.
- 1-page executive summary of the report.
- Description of task 1 and task2.
- Attach (scanned) hard/soft copies of all related material collected and referenced.

Mechanism to Conduct Lab:

Students and teacher communicate through Adobe Connect.

Lab 13

DWH IMPLEMENTATION: Goal Driven Approach

Ralph Kimball's Approach Part II

User requirement definition

- Set an appointment to meet business users.
- Collect answers to questions from business users to understand business requirements (Use your questionnaire which you have prepared in previous lab).
- Compile report of interview and document it in meaningful way.
- Identify business processes.
- Identify Requirements of key processes.
- Prepare report_3 of interview report, business processes and requirements of key business processes.
- Submit report_3.

Note: For details of above tasks, watch video lecture 33 and 36.

Your report should include following:

- Report No.
- Title of course, semester & submission date
- Your Name and roll no.
- Campus and name of city.
- Table of contents.
- 1-page executive summary of the report.
- Interview report, identified business processes and requirements of key business processes.
- Attach (scanned) hard/soft copies of all related material collected and referenced.

Mechanism to Conduct Lab:

• Students and teacher communicate through Adobe Connect.

Lab 14

DIMENSION MODEL OF AIRLINE DATA WAREHOUSE

Scenario

"Pak Airline" is an airliner reservation company, which is operating in more than 10 countries. They have developed the airline reservation system to avoid the errors faced in manual system. The staff of the airline use airline reservation system form the tasks such as flight scheduling, ticket reservation, announcements in automated way. Similarly, users/passengers can search for flight schedule according to date and time and fare details. The staff of the airline can manage the reservation systems by flight rout, runway details, flight scheduling and reservation.

Ticket reservation system of the Pak Airline provides the information about schedule of flights, availability of seats, flight number and destination. For reservation of ticket user have to provide its personal information such as name, age, address etc. For payment purpose user will provide credit card number and bank details. Moreover, information about flight number, date of departure, no. of tickets to be booked is also required for confirmation of ticket. Following is the ERD of above airline reservation system.

Question: You are required to draw the Dimension model for the airline data warehouse following above given scenario.

Solution:



Mechanism to Conduct Lab:

Students and teacher communicate through Adobe Connect.