# ETL Step by Step Exam Help ETL Step by Step Example

#### DB1

#### Customer Table 1

CustID	CustName	Gender
1	Mark	Male
2	John	Male
3	Mila	Female
4	Ann	Female
5	Philip	Male

#### DB2

#### Customer Table 2

CustID	CustName	Gender
1	Mark	1
2	Joe	1
3	Martha	0
4	Paula	0
5	Philip	1

#### Order Table 1

ItemID	CustID	Date	Qty
1	1	1/1/10	8
1	2	1/10/10	6
2	2	4/5/11	1
2	4	2/2/09	5
2	5	1/1/12	9
s <sup>3</sup> onn	něnt Pi	-2/ <del>2/1</del> 3+ 1	F <sup>1</sup> var

# As signment Project Exam Help

4https://tutofcs:com

#### Order Table 2

Itewice	Chat:	csautor	<b>C</b> \$y
1	1	5/1/10	8
1	5	4/10/10	6
2	2	6/5/11	1
2	4	7/2/12	5
2	4	3/1/12	9
3	5	21/2/13	1
3	3	13/1/14	3
4	4	15/7/12	5

#### Items Table 1

ItemID	Description	Price
1	iPad	899
2	Laptop	688
3	Lemon	1.5
4	Pen	5
5	Table	59

#### Items Table 2

ItemID	Description	Price
1	chair	56
2	Laptop	688
3	Apple	1.5
4	Spoon	0.7
5	Table	59

#### TARGET STAR SCHEMA

#### Customer DIM

<u>CustSK</u>	CustName	Gender
		Male

#### DATE Dim

<u>DateSK</u>	Day	Month	Year

#### Item DIM

#### **FACT TABLE**

<u>ItemSK</u>	<u>CustSK</u>	<u>DateSK</u>	Total	Qty
lssigni	ment F	roject	Exa	m H
htt	ps://tu	tores.c	om	
We	Chat:	cstuto	rcs	
	ssign	ssignment F	ssignment Project https://tutorcs.c	ItemSK       CustSK       DateSK       Total         Ssignment Project       Exa         https://tutorcs.com         WeChat:       cstutorcs

<u>ItemSK</u>	Description	Price
n		

# Assignment Project Exam Help

# CUSTOPS/EPPONDENSION

#### DB1

#### Customer Table 1

CustID	CustName	Gender
1	Mark	Male
2	John	Male
3	Mila	Female
4	Ann	Female
5	Philip	Male

#### DB2

#### Customer Table 2

CustID	CustName	Gender
1	Mark	1
2	Joe	1
3	Martha	0
4	Paula	0
5	Philip	1

#### **Customer Staging Table**

	CustID	CustName	Gender	SourceDB	SK
	1	Mark	Male	1	
	2	John	Male	1	
	3	Mila	Female	1	
A agianma	4 nt Dag	Ann	Female	1	
Assignme LOAD	11t P10	Philip EXA	Male	1	
https:	/\tuto	res.com	1	2	
	2	Joe	1	2	
	hat: cs	stutorcs	0	2	
	4	Paula	0	2	
	5	Philip	1	2	

Consolidation /
Entity Matching.
 Assume same name =
same customer. Assign
same surrogate key



Consolidation /
 Entity Matching.
 Assume same name =
 same customer. Assign
 same surrogate key

2. Assign Surrogate Keys. This information is used for future load of new data, it is not temporary!

	CustID	CustName	Gender	SourceDB	CustSK
	1	Mark	Male	1	1
	2	John	Male	1	2
	3	Mila	Female	1	3
aaianma	4 Dro	Ann	Female	1	4
ssignme	<b>III F I</b> C	Philip EXA	Male	1	5
<u>https:</u>	/\tuto	Mark CS.COM	Male	2	1
	2	Joe	Male	2	6
WeC	hat: cs	stutores	Female	2	7
	4	Paula	Female	2	8
	5	Philip	Male	2	5



Standardize data

#### **Customer Dimension**

CustID	CustName	Gender	SourceD B	CustSK		CustSK	CustName	Gender
			_			1	Mark	Male
1	Mark	Male	1	1		2	John	Male
2	John	Male	1	2				IVIAIC
3	Mila	Female	1	3		3	Mila	Female
3	IVIIIa	Telliale		_		4	Ann	Female
4	Ann	Female	1 A	ssignr	nent Project E	<del>x</del> am l	Help Philip	Mala
5	Philip	Male	1	5		3	Philip	Male
1	N. A. o. w.l.	1	2	1 http	s://tutorcs.com	6	Joe	Male
1	Mark	1	2	1 11001		7	Martha	Female
2	Joe	1	2	6	Chate actutora	Co		
3	Martha	0	2	7	Chat: cstutorc	8	Paula	Female
	D. I.	0	2	0				
4	Paula	0	2	8				
5	Philip	1	2	5				

LOAD DATA INTO THE DIMENSION CUSTOMER

## Assignment Project Exam Help

# DIMPENSOCIONEM

DB1

#### Items Table 1

ItemID	Description	Price
1	iPad	899
2	Laptop	688
3	Lemon	1.5
4	Pen	5
5	Table	59

#### DB2 Items Table 2

ItemID	Description	Price
1	chair	56
2	Laptop	688
3	Apple	1.5
4	Spoon	0.7
5	Table	59

#### Item Staging Table

	ItemID	CustName	Price	SourceDB	SK
	1	iPad	899	1	
	2	Laptop	688	1	
	3	Lemon	1.5	1	
A sai anna a	4 24 Dags	Pen Evro	5 II	1	
Assignme	fit Pro	Ject Exa	m H(	<del>db</del>	
LOAD  https:	/ <sup>1</sup> /tuto	rchair cS.com	56	2	
	2	Laptop	688	2	
WeC	hat: cs	statercs	1.5	2	
	4	Spoon	0.7	2	
	5	Table	59	2	

#### Item Staging Table

1. Consolidation / **Entity Matching** 

Assignme

2. Assign surrogate keys

	ItemID	CustName	Price	SourceDB	SK
	1	iPad	899	1	1
	2	Laptop	688	1	2
	3	Lemon	1.5	1	3
ianma	4 nt Dro	Pen Exe	5 <b>L</b>	1	4
ıgnme	111 110	ject Exa	m Ho	- <del>I</del> h	5
https:	/\tuto	rchair cS.com	56	2	6
	2	Laptop	688	2	2
WeC	hat: cs	statercs	1.5	2	7
	4	Spoon	0.7	2	8
	5	Table	99	2	9

#### Item Staging Table

ItemID	CustName	Price	SourceDB	SK	]	ItemSK	CustName	Price
1	iPad	899	1	1		1	iPad	899
2	Laptop	688	1	2		2	Laptop	688
3	Lemon	1.5			nt <sub>L</sub> Broject	Bxam	Hedp	1.5
4	Pen	5	1	1	- 207.13	4	Pen	5
5	Table	<b>59</b>	1	https 5	://tutores.co	m	Table	59
					14444	6	chair	56
1	chair	56			hat: cstutor	<del>CS</del> 7	Apple	1.5
2	Laptop	688	2	2		8	Spoon	0.7
3	Apple	1.5	2	7		9	Table	99
4	Spoon	0.7	2	8				
5	Table	99	2	9				

LOAD DATA INTO THE DIMENSION ITEM

## Assignment Project Exam Help

DATES:/DHAVE ENSION

# Date dimension

• It can be pre-populated (the same for time dimension).

```
Assignment Project Exam Help
CREATE TABLE d date AS
   SELECT
   n AS Date SK,
   TO DATE('31/12/2007', 'DD/MM/YYYhttps://tutorcs.com, 'day') As Full Date,
   TO_CHAR(TO_DATE('31/12/2007','DD/MM/YYYY') + NUMTODSINTERVAL(n,'day'),'DD') AS Days,
   TO_CHAR(TO_DATE('31/12/2007','DWMM(TYXXt) csNUMTOPSINTERVAL(n,'day'),'Mon') AS Month_Short,
   TO CHAR (TO DATE ('31/12/2007', 'DD/MM/YYYY') + NUMTODSINTERVAL (n, 'day'), 'MM') AS Month Num,
   TO CHAR (TO DATE ('31/12/2007', 'DD/MM/YYYY') + NUMTODSINTERVAL (n, 'day'), 'Month') AS Month Long,
   TO CHAR (TO DATE ('31/12/2007', 'DD/MM/YYYY') + NUMTODSINTERVAL (n, 'day'), 'Q') AS Quarter,
    TO CHAR(TO DATE('31/12/2007','DD/MM/YYYY') + NUMTODSINTERVAL(n,'day'),'YYYYY') AS Year
FROM
  select level n
  from dual
 connect by level <= 2000);
```

# Assignment Project Exam Help

https://tutoica.eom

#### Order Table 1

ItemID	CustID	Date	Qty
1	1	1/1/10	8
1	2	1/10/10	6
2	2	4/5/11	1
2	4	2/2/09	5
2	5	1/1/12	9
3	5	2/2/13	1
3	3	1/1/11	3 A
4	4	15/7/12	5

#### Order Table 2

ItemID	CustID	Date	Qty
1	1	5/1/10	8
1	5	4/10/10	6
2	2	6/5/11	1
2	4	7/2/12	5
2	4	3/1/12	9
3	5	21/2/13	1
3	3	13/1/14	3
4	4	15/7/12	5

#### Stage Fact Table

ItemID   CustID   Date   Qty   Price   Item   Cust   Date   Sour   SK   SK   DB	
1     2     1/10/10     6     1       2     2     4/5/11     1     1       2     4     2/2/09     5     1	е
2     2     4/5/11     1       2     4     2/2/09     5	
2 4 2/2/09 5 1	
cidnment Project Evami Helm	
1351g-mmcIII r 10 just L'Aarm 10 p	
3 5 2/2/13 1	
https://tutorcsacom 1/1/11 3	
4   4   15/7/12   5         1	
WeChaticstutores 5/1/10 8 2	
1 5 4/10/10 6 2	
2 2 6/5/11 1 2	
2 4 7/2/12 5 2	
2 4 3/1/12 9 2	
3 5 21/2/13 1 2	
3 3 13/1/14 3 2	
4 4 15/7/12 5 2	

#### Stage Fact Table

#### Assign Surrogate Key

- 1. Join the stage Fact Table with the Customer Stage Table to get the right Customer SK. <u>Join on the CustID and the SourceDB</u>
- 2. Join the stage Fact Table with the stage
- 3. Join the stage Fact Table with the Date WeCha dimension to get the right DateSK. Join on the day, month and year parsed from the field Date of the stage Fact Table.
- 4. <u>Price</u> is a derived field (Qty X Price). In the fact table it is called total

ItemID	CustID	Date	Qty	Price	Item SK	Cust SK	Date SK	Source DB
1	1	1/1/10	8	7192	1	1	34	1
1	2	1/10/10	6	5394	1	2	45	1
2	2	4/5/11	1	688	2	2	163	1
2	4	2/2/09	5	3440	2	4	54	1
t <sup>2</sup> Proi	et F	1/1/1 <sup>2</sup> H	ip	6192	2	5	34	1
3	5	2/2/13	1	1.5	3	5	65	1
titoro	s con	1/1/11	3	4.5	3	3	34	1
4	4	15/7/12	5	25	4	4	54	1
at est	utorc	5/1/10	8	448	6	1	39	2
1	5	4/10/10	6	336	6	5	59	2
2	2	6/5/11	1	688	2	6	133	2
2	4	7/2/12	5	3440	2	8	178	2
2	4	3/1/12	9	6192	2	8	33	2
3	5	21/2/13	1	1.5	7	5	61	2
3	3	13/1/14	3	4.5	7	7	342	2
4	4	15/7/12	5	3.5	8	8	128	2

#### Stage Fact Table

ItemID	Cust ID	Date	Qty	Price	Item SK	Cust SK	Date SK	Source DB		
1	1	1/1/10	8	7192	1	1	34	1		
1	2	1/10/10	6	5394	1	2	45	1		
2	2	4/5/11	1	688	2	2	163	1		
2	4	2/2/09	5	3440	2	4	54	1		
2	5	1/1/12	9	6192	7 00	5	34	Droise	t <b>Exa</b> m	,
3	5	2/2/13	1	1.5	ASS	gnn	es 65			
3	3	1/1/11	3	4.5	3	http	34 / /4:	1 utores	com	
4	4	15/7/12	5	25	4	4	54		com	
1	1	5/1/10	8	448	6	We	39	cstut	orce	
1	5	4/10/10	6	336	6	5	59 59	. CStut	ores	
2	2	6/5/11	1	688	2	6	133	2		
2	4	7/2/12	5	3440	2	8	178	2		
2	4	3/1/12	9	6192	2	8	33	2		
3	5	21/2/13	1	1.5	7	5	61	2		
3	3	13/1/14	3	4.5	7	7	342	2		
4	4	15/7/12	5	3.5	8	8	128	2		

#### Fact Table

Qty	TOTAL	Item SK	Cust SK	Date SK
8	7192	1	1	34
6	5394	1	2	45
1	688	2	2	163
5	3440	2	4	54
9	6192	2	5	34
цегр	1.5	3	5	65
3	4.5	3	3	34
5	25	4	4	54
8	448	6	1	39
6	336	6	5	59
1	688	2	6	133
5	3440	2	8	178
9	6192	2	8	33
1	1.5	7	5	61
3	4.5	7	7	342
5	3.5	8	8	128

#### LOAD INTO THE FACT TABLE

# Assignment Project Exam Help WHAT HAPPENS IF THERE ARE NEW DATA AND I NEED TO LOAD THEM IN THE STAR SCHEMA? WeChat: cstutorcs

#### NEW DATA FROM DB2

#### **Customer Staging Table**

#### DB2

CustID	CustName	Age	Gender
1	Mark	34	1
2	Joe	30	1
3	Martha	34	0
4	Pauline	19	0
5	Philip	25	1
6	Ryan	28	1
7	Patricia	19	0

	CustID	CustName	Age	Gender	SourceDB	CustSK
	1	Mark	34	Male	1	1
	2	John	33	Male	1	2
	3	Mila	51	Female	1	3
Aggianma	4 Dro	Ann	19 II	Female	1	4
Assignme LOAD	<u>nt F10</u>	Philip EXA	25	Male	1	5
https:	/tuto	Mark CS.COM	34	Male	2	1
	2	Joe	30	Male	2	6
WeC	hat: cs	stutores	34	Female	2	7
	4	Paula	19	Female	2	8
	5	Philip	25	Male	2	5
+						

	CustID	CustName	Gender	SourceDB	CustSK
	1	Mark	Male	1	1
	2	John	Male	1	2
	3	Mila	Female	Female 1	
Aggianma	4 nt Dro	Ann	Female	1	4
Assignme	11t F10	Philip EXA	Male	1	5
https:	/\tuto	res.com	Male	2	1
перы	2	Joe	Male	2	6
WeC	hat: cs	stutores	Female	2	7
	4	Pauline	Female	2	8
	5	Philip	Male	2	5
	6	Ryan	Male	2	9
	7	Patricia	Female	2	10

ASSIGN SURROGATE KEYS AND STANDARDIZE DATA

#### **Updated Customer Dimension**

CustID	CustName	Gender	SourceDB	CustSK		CustSK	CustName	Gender
1	Mark	Male	1	1		1	Mark	Male
2	John	Male	1	2		2	John	Male
3	Mila	Female	1	3		3	Mila	Female
4	Ann	Female	1	4	, D , , E	4	Ann	Female
5	Philip	Male	1 ASSI	gnme	ent Project E	<del>xam</del> l	Philip	Male
1	Mark	Male	2	httns	://tutores.com	6	Joe	Male
2	Joe	Male	2	6	.//tatores.eor	7	Martha	Female
3	Martha	Female	2	WeC	hat: cstutorc	<b>S</b> 8	Pauline	Female
4	Pauline	Female	2	8		9	Ryan	Male
5	Philip	Male	2	5		10	Patricia	Female
6	Ryan	Male	2	9				
7	Patricia	Female	2	10				

INCREMENTAL LOAD DATA INTO THE DIMENSION CUSTOMER

# Changes in the Fact Table and Dimensions

- New facts are 99% appended in the fact table (very few exceptions)
- Dimensions change slower than the fact table. However, when a dimension is modified by an update, multiple strategies are possible (see next presentation)https://tutorcs.com