<u>Name:</u> **Payal Sharma**
<u>Date:</u> **20th March' 2014**
<u>Assignment 7</u>: **Normalization 1**

Part One: Kramerica CEO Miles Meservy has put together a spreadsheet of all the data he has so far, which he personally collected.

1. As he shows you the spreadsheet, having just signed your consulting agreement, he asks what you think of it. How do you reply?

   <u>Solution:</u> I would reply "It is a mere collection of unorganized data which needs to get normalized in order to get some meaning."

2. Put his data in 1NF and display it. (Show me the table; no SQL.)

   <u>Solution:</u> Since there is **many to many relationship** between Software package and a computer i.e. a software package can be installed in many computers, similarly a computer can have many software packages installed. So, the first normal form of above data is given below:

Package_info	
PID	**PackID**
1	AC01
2	DB32
3	DB33
4	WP08
5	WP09

Package_Installation_Info			
PID	**TAGID**	**InstallDate**	**SoftwareCost_USD**
1	32808	9/13/1995	754.95
2	32808	12/3/1995	380
2	37691	6/15/1995	380
3	57772	5/27/1995	412.77
4	32808	1/12/1996	185
4	37691	6/15/1995	227.5
4	57222	5/27/1995	170.24
5	59836	10/30/1995	35
5	77740	5/27/1995	35

3. What is the primary key?

> **Solution:** **PID** is the **primary key** in table Package_info and **PID, TagNum** is the **composite key** in Package_Installation_Info.

Part Two: Add two columns of new data: one for software package name (e.g., Zork, Portal, etc.) and one for computer model (e.g., HP, Apple, etc.). Be sure that your new data is consistent with the original data. Do not add any additional columns.

4. Display the new table.

> **Solution:** Since the above data was in first normal form, so after adding data in above tables below are the new tables:

Package_info		
PID	**PackID**	**Package_Name**
1	AC01	DOS
2	DB32	Firefox
3	DB33	Microsoft Office
4	WP08	Telnet
5	WP09	Scandisk

Package_Installation_Info				
PID	**TAGID**	**InstallDate**	**SoftwareCost_USD**	**Computer_Model**
1	32808	9/13/1995	754.95	Apple
2	32808	12/3/1995	380	Apple
2	37691	6/15/1995	380	HP
3	57772	5/27/1995	412.77	Lenovo
4	32808	1/12/1996	185	Apple
4	37691	6/15/1995	227.5	HP
4	57222	5/27/1995	170.24	Lenovo
5	59836	10/30/1995	35	Acer
5	77740	5/27/1995	35	Dell

5. Identify and document all the functional dependencies.

> **Solution:** In table Package_info, PackID and Package_Name are functionally determined by PID.

i.e. PID →PackID, Package_Name

In table Package_Installation_Info, InstallDate and SoaftwareCost_USD are functionally determined by composite key <PID,TAGID>

i.e. <PID, TAGID>→InstallDate,SoftwareCost_USD

6. Explain why this new table is not in third normal form.

   Solution: These new tables are not in 3rd normal form because it is not in 2nd normal form i.e. the non key attributes does not depend on all primary key attributes. In table Package_Installation_Info, Computer_Model is not dependent on composite key. It is determined by only TAGID.

Part Three: Decompose your 1NF table into a set of tables that are in at least third normal form. (BCNF would be better.) Remember that it's wrong to add artificial keys to associative entities.

New tables are given below:

Package_info		
PID	PackID	Package_Name
1	AC01	DOS
2	DB32	Firefox
3	DB33	Microsoft Office
4	WP08	Telnet
5	WP09	Scandisk

Computer_info	
TagNum	Computer_Model
32808	Apple
37691	HP
57772	Lenovo
59836	Acer
77740	Dell

Package_Installation_Info			
PID	TAGID	InstallDate	SoftwareCost_USD
1	32808	9/13/1995	754.95
2	32808	12/3/1995	380
2	37691	6/15/1995	380
3	57772	5/27/1995	412.77
4	32808	1/12/1996	185
4	37691	6/15/1995	227.5
4	57222	5/27/1995	170.24
5	59836	10/30/1995	35
5	77740	5/27/1995	35

7. Identify all primary keys (determinants) for all tables.

Solution: **PID** is the **primary key** in table Package_info.

TagNum is the **primary key** of Computer_info table.

PID, TagNum is the **composite key** in Package_Installation_Info.

8. Identify all functional dependencies for all tables.

Solution: In table Package_info, PackID and Package_Name are functionally determined by PID.

i.e. PID →PackID, Package_Name

In table Computer_info, Computer_Model is functionally dependent on TagNum.

i.e. TagNum→ Computer_Model

In table Package_Installation_Info, InstallDate and SoftwareCost_USD are functionally determined by composite key <PID, TAGID>

i.e. <PID, TAGID>→InstallDate, SoftwareCost_USD

9. Explain why the new tables are in third normal form.

Solution: The tables are in third normal form as they meet all four conditions mentioned below:

- It is in second normal form.
- All non-key attributes are not dependent on any other non-key attributes.
- Each field has unique name.
- It has a primary key.

10. Draw a beautiful E/R diagram.

Solution: The ER diagram of above tables is drawn below:

Package_info
- 🔑 PID
- 🔷 PackID
- 🔶 PackageName

Indexes ▶

Computer_info
- 🔑 TagNum
- 🔷 Computer_Model

Indexes ▶

Package_Installation_Info
- 🔷 InstallDate
- 🔷 SoftwareCost_USD
- 🔑 PID
- 🔑 TagNum

Indexes ▶