Chris Ognibene (*help from Clayton Enright for Part III)

Professor Labouseur

CMPT 308 – 111

26 March 2014

Lab 7 – Normalization 1

Part One:

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

PackID	TagNum	InstallDate	SoftwareCost
AC01	32808	09-13-1995	754.95
DB32	32808	12-03-1995	380.00
	37691	06-15-1995	380.00
DB33	57772	05-27-1995	412.77
WP08	32808	01-12-1996	185.00
	37691	06-15-1995	227.50
	57222	05-27-1995	170.24
WP09	59836	10-30-1995	35.00
	77740	05-27-1995	35.00

I reply that for manual, record-keeping purposes, the table does the job. However, transforming the table and inputting it into a database, the table needs work done. In terms of the field (column) names, they were chosen appropriately and descriptively. The main work that needs to be done, however, is that the multiple values in most of the cells need to be changed to one single value per cell (atomicity violation, 1NF, not that I would assume s/he knew this term).

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

PackID	TagNum	InstallMon	InstallDay	InstallYr	SoftwareCost
AC01	32808	9	13	1995	754.95
DB32	32808	12	3	1995	380
DB32	37691	6	15	1995	380
DB33	57772	5	27	1995	412.77
WP08	32808	1	12	1996	185
WP08	37691	6	15	1995	227.5
WP08	57222	5	27	1995	170.24
WP09	59836	10	30	1995	35
WP09	77740	5	27	1995	35

What is the primary key?

The primary key is a composite key, (PackID, TagNum).

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.

Hint for Packages: http://en.wikipedia.org/wiki/List_of_open-source_software_for_mathematics

Hint for Computer Brands: http://www.thetoptens.com/best-computer-brands/

Display the new table.

PackID	TagNum	InstallMon	InstallDay	InstallYr	SoftwareCost	SoftPackage	CompModel
AC01	32808	9	13	1995	754.95	Mesa3D	Dell
DB32	32808	12	3	1995	380	ROS	Dell
DB32	37691	6	15	1995	380	ROS	Asus
DB33	57772	5	27	1995	412.77	Blender	Toshiba
WP08	32808	1	12	1996	185	Bitcoin	Dell
WP08	37691	6	15	1995	227.5	Bitcoin	Asus
WP08	57222	5	27	1995	170.24	Bitcoin	Lenovo
WP09	59836	10	30	1995	35	Sage	Alienware
WP09	77740	5	27	1995	35	Sage	Samsung

Identify and document all the functional dependencies.

(PackID, TagNum) → (InstallMon, InstallDay, InstallYr, SoftPackName, CompModel)

PackID → SoftPackage

TagNum → CompModel

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

Third Normal Form means that the table is in second normal form (atomic values and no partial key dependencies) with the added requirement that there are no multiple primary key dependencies. The reason is because the table is not in second normal form because there are partial key dependencies, or PackID functionally determines a subset of the non-key attributes, specifically Software_Package, and TagNum functionally determines CompModel.

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.

Software_Package	
PackID	**SoftPackage**
AC01	Mesa3D
DB32	ROS
DB33	Blender
WP08	Bitcoin
WP09	Sage

Comp_Model	
TagNum	**CompModel**
59836	Alienware
37691	Asus
32808	Dell
57222	Lenovo
77740	Samsung
57772	Toshiba

Install_Time					
PackID	**TagNum**	**InstallMon**	**InstallDay**	**InstallYr**	**SoftwareCost**
AC01	32808	9	13	1995	754.95
DB32	32808	12	3	1995	380
DB32	37691	6	15	1995	380
DB33	57772	5	27	1995	412.77
WP08	32808	1	12	1996	185
WP08	37691	6	15	1995	227.5
WP08	57222	5	27	1995	170.24
WP09	59836	10	30	1995	35
WP09	77740	5	27	1995	35

Identify all primary keys (determinants) for all tables.

Software_Package Table: PackID, a primary key

Comp_Model Table: TagNum, a primary key

Install_Time Table: (PackID, TagNum), a composite key (still a PK)

- Taken individually, PackID and TagNum are foreign keys to their respective tables

Identify all functional dependencies for all tables.

Software_Package Table: PackID → SoftPackage

Comp_Model Table: TagNum → CompModel

Install_Time Table: (PackID, TagNum) → (InstallMon, InstallDay, InstallYr, SoftwareCost)

Explain why the new tables are in third normal form.

The tables are in third normal form because they are in second normal form (1NF and no partial key dependencies), with the added condition that there are no multiple primary key dependencies.

Draw a beautiful E/R diagram.