Pearl Amin
Professor Labouseur
Database Systems
November 2, 2017

Lab 7: Normalization

Part 1:

1. What do you think of it?

I think that this dataset is not as good as it could be and has many anomalies such as inset, add and delete anomalies since it is a large table that is wide with many columns. The data would be more accurately portrayed with normalization. Then, it would be atomic, have no partial key dependencies, and no multi-key dependencies or transitive dependencies among non-key attributes, obtaining the goal of logically, clearly showing the relationships in the data set.

2. Table in 1NF:

PackageID	TagNumber	InstallDate	SoftwareCostUSD
AC01	32808	09-13-2005	754.95
DB32	32808	02-03-2005	380.00
DB32	37691	06-15-2005	380.00
DB33	57772	05-27-2005	412.77
WP08	32808	01-12-2006	185.00
WP08	37691	06-15-2005	227.50
WP08	57222	05-27-2005	170.24
WP09	59836	10-30-2005	35.00
WP09	77740	05-27-2005	35.00

3. The primary key is PackageID,TagNumber,InstallDate

Part 2:

4. New Table:

PackageID	TagNumber	InstallDate	SoftwareCostUSD	PackageName	ComputerModel
AC01	32808	09-13-2005	754.95	Zork	IBM
DB32	32808	12-03-2005	380.00	Portal	Apple
DB32	37691	06-15-2005	380.00	Portal	Apple
DB33	57772	05-27-2005	412.77	Word	Amazon
WP08	32808	01-12-2006	185.00	PowerPoint	Dell
WP08	37691	06-15-2005	227.50	PowerPoint	Dell
WP08	57222	05-27-2005	170.24	PowerPoint	Dell
WP09	59836	10-30-2005	35.00	Excel	HP
WP09	77740	05-27-2005	35.00	Excel	HP

5. **Functional Dependencies:**
   PackageID→PackageName
   TagNumber→ComputerModel
   PackageID, TagNumber, InstallDate → SoftwareCostUSD

6. This data set is not in 3NF because it is not even in 2NF. There are partial key dependencies, for example, PackageName is dependent on the PackageID, however other elements such as SoftwareCostUSD are dependent on PackageID, TagNumber, and InstallDate. It this was in 3NF, then someone should be able to determine the PackageName by the PackageID easily.

Software	
PackageID	PackageName
AC01	Zork
DB32	Portal
DB32	Portal
DB33	Word
WP08	PowerPoint
WP08	PowerPoint
WP08	PowerPoint
WP09	Excel
WP09	Excel

Computer	
TagNumber	ComputerModel
32808	IBM
32808	Apple
37691	Apple
57772	Amazon
32808	Dell
37691	Dell
57222	Dell
59836	HP
77740	HP

Installation			
PackageID	TagNumber	InstallDate	SoftwareCostUSD
AC01	32808	09-13-2005	754.95
DB32	32808	12-03-2005	380.00
DB32	37691	06-15-2005	380.00
DB33	57772	05-27-2005	412.77
WP08	32808	01-12-2006	185.00
WP08	37691	06-15-2005	227.50
WP08	57222	05-27-2005	170.24
WP09	59836	10-30-2005	35.00
WP09	77740	05-27-2005	35.00

Part 3:

7. **Primary Keys:**
   Software: PackageID
   Computer: TagNumber
   Installation: PackageID, TagNumber, InstallDate

8. Funcational Dependencies:
   PackageID→PackageName
   TagNumber→ComputerModel
   PackageID, TagNumber, InstallDate → SoftwareCostUSD

9. The new table is in third normal form because it is in 2NF and there are no longer any multikey dependencies or transitive dependencies. Both the computer and software tables have a primary key that only has one dependency. For the installation table, SoftwareCostUSD has one value that relies on each of the values of inside of composite primary key PackageID, TagNumber, and InstallDate. For the computers and software tables, the case is the same where the primary keys have one main dependency.

10. ER Diagram