

Lab 7: Normalization 1

Tycho Manufacturing:Software Packages

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

Part One:

1. Looking at the state of this spreadsheet, there are a few issues that can arise if it were to be transferred over into a table in a database. One of the quirks that are in this spreadsheet are the names of the columns. If one were to look at the spreadsheet without context, it would be somewhat difficult to decipher the meaning behind this data. For example the column 'packageID' does not give much information about the software being installed. This column essentially tells us only that something was installed. The names of each software are missing, therefore one would not know what software is on what computer without memorizing the packageID's. We can only tell software is being installed because there is a column 'SoftwareCostUSD' and 'InstallDate', implying the packages are software. The column 'TagNumber' does not explicitly state if asset tag number belongs to a computer; for all we know, one of those numbers could be an espresso machine.

2. The spreadsheet into First Normal Form:

SoftwarePkgID	CompTagNumber	InstallDate	SoftwareCostUSD
AC01	32808	09-13-2005	754.96
DB32	32808	12-03-2005	380.00
DB32	37691	06-15-2005	380.00
DB33	57772	05-27-2005	412.77
WP08	32808	01-12-2005	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 for this new table would be a composite key consisting of
 'SoftwarePkgID' and 'CompTagNumber.'

Part Two:

4.

SoftwarePkgID	CompTagNumber	CompModel	SoftwareName	InstallDate	SoftwareCostUSD
AC01	32808	Lenovo	Unreal Engine	09-13-2005	754.96
DB32	32808	Lenovo	Adobe Acrobat	12-03-2005	380.00
DB32	37691	Lenovo	Adobe Acrobat	06-15-2005	380.00
DB38	57772	Apple	Parallels	05-27-2005	412.77
WP08	32808	Lenovo	Adobe Premier	01-12-2005	185.00
WP08	37691	Lenovo	Adobe Premier	06-15-2005	227.50
WP08	57222	Apple	Adobe Premier	05-27-2005	170.24
WP09	59836	Apple	Mail Pilot 2	10-30-2005	35.00

WP09	77740	Apple	Mail Pilot 2	05-27-2005	35.00
------	-------	-------	--------------	------------	-------

5. Functional Dependencies:

SoftwarePkgID \rightarrow SoftwareName, InstallDate, SoftwareCostUSD

CompTagNumber \rightarrow CompModel

6. 'SoftwareCosUSD' violates 3NF due to the fact that the cost of the software is dependant on the install time as well as 'SoftwarePkgID'. The price of each software changes at various install dates. There is also a violation of 2NF as well. There are a few partial key dependencies such as 'SoftwareName' only depending on 'SoftwarePkgID' and not 'CompTagNumber.'

Part Three:

7. There are 4 tables: Computers, Software, Prices, and Installs

Primary Keys:

Computers: TagNumber

Software: PackageID

Prices: PackageID, InstallDate

Installs: PackageID, TagNumber

8. Functional Dependencies:

TagNumber \rightarrow Model

PackageID \rightarrow Name

PackageID, InstallDate \rightarrow CostUSD

PackageID, TagNumber → InstallDate

9. The new tables are in Third Normal Form because we have eliminated the issues of partial key dependencies, making the tables 2NF. The new tables also do not have any multi-key dependencies such as 'SoftwareCostUSD' depending on both packageID and InstallDate.

10.

