Shannon Brady
Professor Labouseur
CMPT308
30 October 2020

Lab 7: Normalization One

Part One:

1) I would explain that the design of Johnson's spreadsheet is very inefficient and leaves Tycho Manufacturing susceptible to significant errors in the future, as it is very likely that the data it holds will become inconsistent over time. In order to avoid such catastrophe, the spreadsheet will need to be organized in such a way that reduces data redundancy and ensures data integrity. This can be achieved by breaking down the initial spreadsheet into separate tables with defined relationships.

2)

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

3) The primary key is a composite key combining both the PackageID and TagNumber fields.

Part Two:

4)

PackageID	TagNumber	Package Name	Computer Model	InstallDate	SoftwareCost USD
AC01	32808	QuickBooks	Apple	09-13-2005	754.95
DB32	32808	Microsoft Office	Apple	12-03-2005	380.00
DB32	37691	Microsoft Office	IBM	06-15-2005	380.00
DB33	57772	Photoshop	HP	05-27-2005	412.77
WP08	32808	Portal	Apple	01-12-2006	185.00
WP08	37691	Portal	IBM	06-15-2005	227.50
WP08	57222	Portal	HP	05-27-2005	170.24
WP09	59836	Zork	Lenovo	10-30-2005	35.00
WP09	77740	Zork	Dell	05-27-2005	35.00

5) PackageID, TagNumber → PackageName, ComputerModel, InstallDate, SoftwareCostUSD

6) This table is not in 3NF because it does not meet the requirements for 2NF as there are partial key dependencies within the data. For example, PackageName is only determined by PackageID, which is a subset of the primary key. Similarly, ComputerModel is only determined by TagNumber, which is also a subset of the primary key.

Part Three:

Table A

PackageID	PackageName
AC01	QuickBooks
DB32	Microsoft Office
DB33	Photoshop
WP08	Portal
WP09	Zork

Table B

TagNumber	ComputerModel
32808	Apple
37691	IBM
57772	HP
59836	Lenovo
77740	Dell

Table C

PackageID	TagNumber	InstallDate
AC01	32808	09-13-2005
DB32	32808	12-03-2005
DB32	37691	06-15-2005
DB33	57772	05-27-2005
WP08	32808	01-12-2006
WP08	37691	06-15-2005
WP08	57222	05-27-2005
WP09	59836	10-30-2005
WP09	77740	05-27-2005

Table D

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

7) *Primary Keys*

Table A: PackageID
Table B: TagNumber
Table C: PackageID, TagNumber
Table D: PackageID, InstallDate

8) *Functional dependencies*

Table A: PackageID → PackageName
Table B: TagNumber → ComputerModel
Table C: PackageID, TagNumber → InstallDate
Table D: PackageID, InstallDate → SoftwareCostUSD

9) These tables are in 3NF because they now meet the requirements for 2NF, since there are no partial dependencies, and there are also no transitive dependencies. This is to say that all of the non-primary key attributes are dependent on only the primary key and no other columns in the table.

10)