

جامعة القدس المفتوحة



ملحق (الوحدة السابعة)

لمقرر تطبيقات حاسوبية في العلوم الإدارية 4307

تطبيع البيانات (Normalization)

إعداد: سامي زواهره

فرع بيت لحم/ مركز بيت ساحور الدراسي

آذار 2014

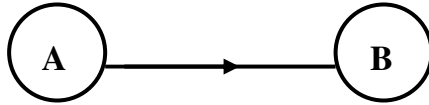
## تطبيع البيانات (Normalization)

عزيزي الطالب، أن الهدف الرئيسي لمصمم قاعدة البيانات هو تصميم البيانات بحيث تكون خالية من التكرار ويمكن استرجاعها وتعديلها والإضافة عليها دون المشاكل التي يمكن أن تحدث مع وجود التكرار فيها ويتم ذلك عن طريق إيجاد ثلاث مستويات من النماذج لقواعد البيانات تسمى نماذج التطبيع ( Normal Forms ). فالتطبيع ( Normalization ) هي عملية تجميع الخصائص بشكل علاقات مرتبة وتحويل البيانات على شكل جداول ثنائية البعد، وهذا ما يدعى التطبيع الأول (First Normal Form)، أما عمليات التطبيع الأخرى فتفحص العلاقات في التطبيع الأول ويمكن أن تقسم بعضاً منها إلى علاقات أكثر بساطة، وتتم هذه العملية في خطوتين هما: التطبيع الثاني والتطبيع الثالث، وهذه الخطوات تتعلق في التصميم المنطقي لقواعد البيانات فقط.

الاعتمادات هي عبارة عن علاقات تربط خصائص جدول أو عدة جداول ومن أهمها:

### 1. الاعتماد الوظيفي Functional Dependency

لوضع مخطط قواعد البيانات على المصمم أن يحدد أي من المحددات أو الخصائص (Attributes) تعتمد على غيرها، فمثلاً إذا كان الزبون كيان فإن رقم الزبون، اسمه، عنوانه هي خصائص. ويمكن أن نعرف المصطلح الاعتماد الوظيفي Functional Dependency بما يلي: أي محددة B في العلاقة R هي اعتماد وظيفي على المحددة A في العلاقة R، إذا كانت لكل قيمة من A لا يوجد أكثر من قيمة واحدة في B ومقرونة بها في العلاقة R في أي لحظة. ويمكن أن نكتبها كما يلي :  $A \longrightarrow B$  وتقرأ إن B تعتمد على A. وقد مثلنا ذلك باستخدام خط عليه سهم واحد من المحددة A إلى المحددة B في المخطط الفقاعي.



### 2. الاعتماد الوظيفي الكامل Full Functional Dependency

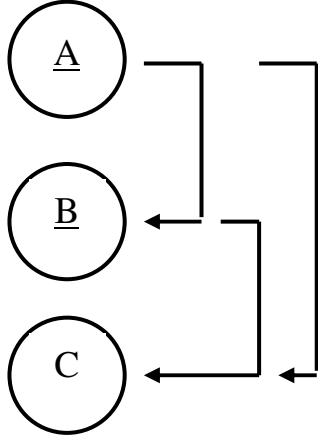
يمكن القول على أن المحددة أو مجموعة من المحددات B من العلاقة R بأنها اعتماد وظيفي كامل على محددة أو مجموعة من المحددات الأخرى A من العلاقة R إذا كانت B هي اعتماد وظيفي على جميع محددات A وليس على أي مجموعة جزئية من A. فالعلاقة R هي في التطبيع الثاني إذا كانت :

• في التطبيع الأول INF

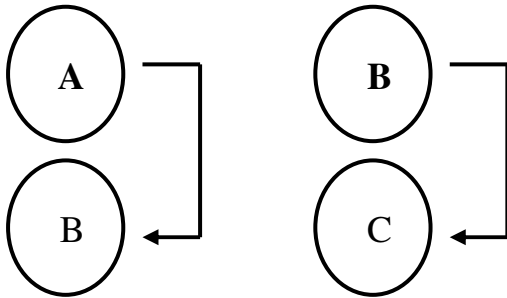
- كل محددة غير أولية Non-prime Attribute من R هي اعتماد وظيفي كامل على كل المفتاح المرشح Candidate key للعلاقة R .

### 3. الاعتماد الانتقالي أو اعتماد التعدية Transitive Dependency

تحدث عند وجود اعتمادات وسيطة. والخطوة التالية يتم فيها تحويل التطبيع الثاني (2NF) إلى التطبيع الثالث (3NF)، والتي تهتم بإزالة ما يدعى اعتماد التعدية. ويمكن أن نوضح ذلك من خلال المثال التالي، افرض أن A, B, C هي ثلاثة محددات من العلاقة R. إذا كانت C اعتماد وظيفي على المحددة B والمحددة B هي اعتماد وظيفي على المحددة A، فإن C هي اعتماد وظيفي على A، أي أن C هي اعتماد تعدية على A. والشكل التالي يوضح ذلك، حيث أن C هي اعتماد تعدية على A.



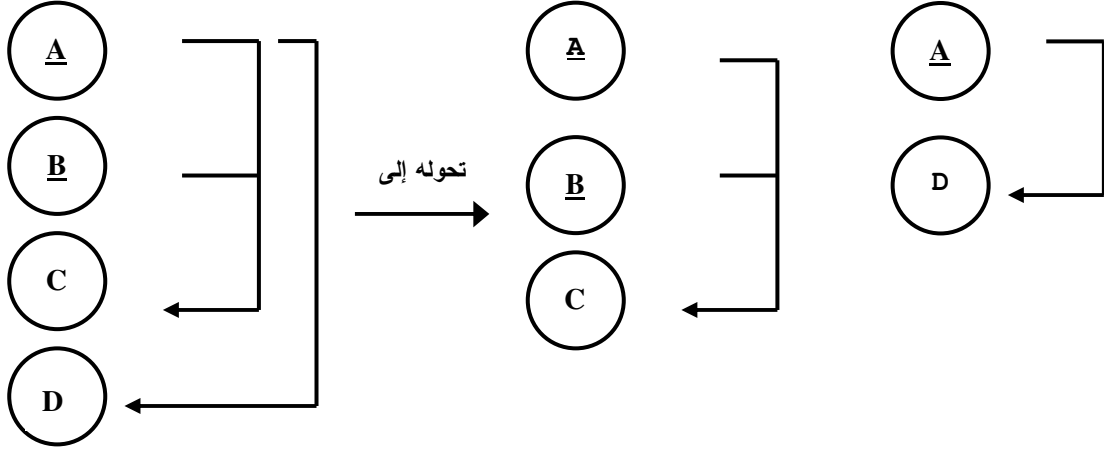
ولكي نحول ذلك إلى التطبيع الثالث (3NF)، يجب أن نتخلص من اعتماد التعدية بتقسيم العلاقة إلى علاقتين كما يلي:



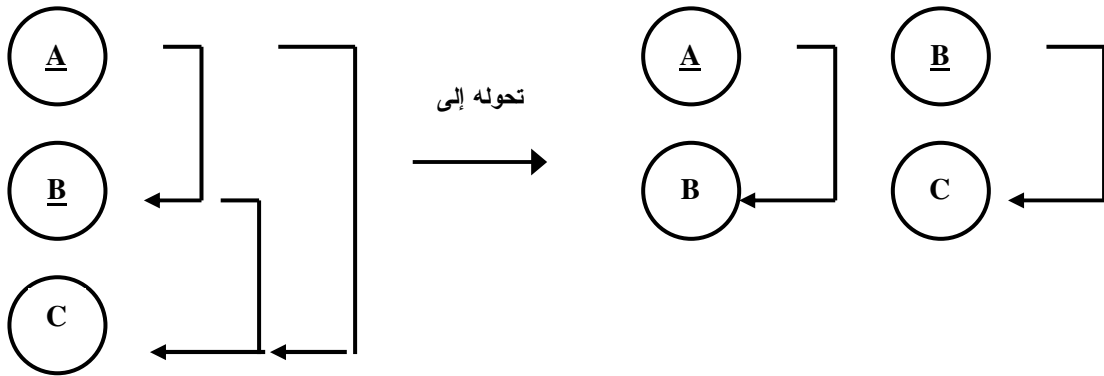
تعتبر العلاقة في التطبيع الثالث 3NF،

- إذا كانت في التطبيع الثاني.
- كل محدد غير أولية من R هي ليست اعتماد تعدية على كل المفتاح المرشح للعلاقة R.

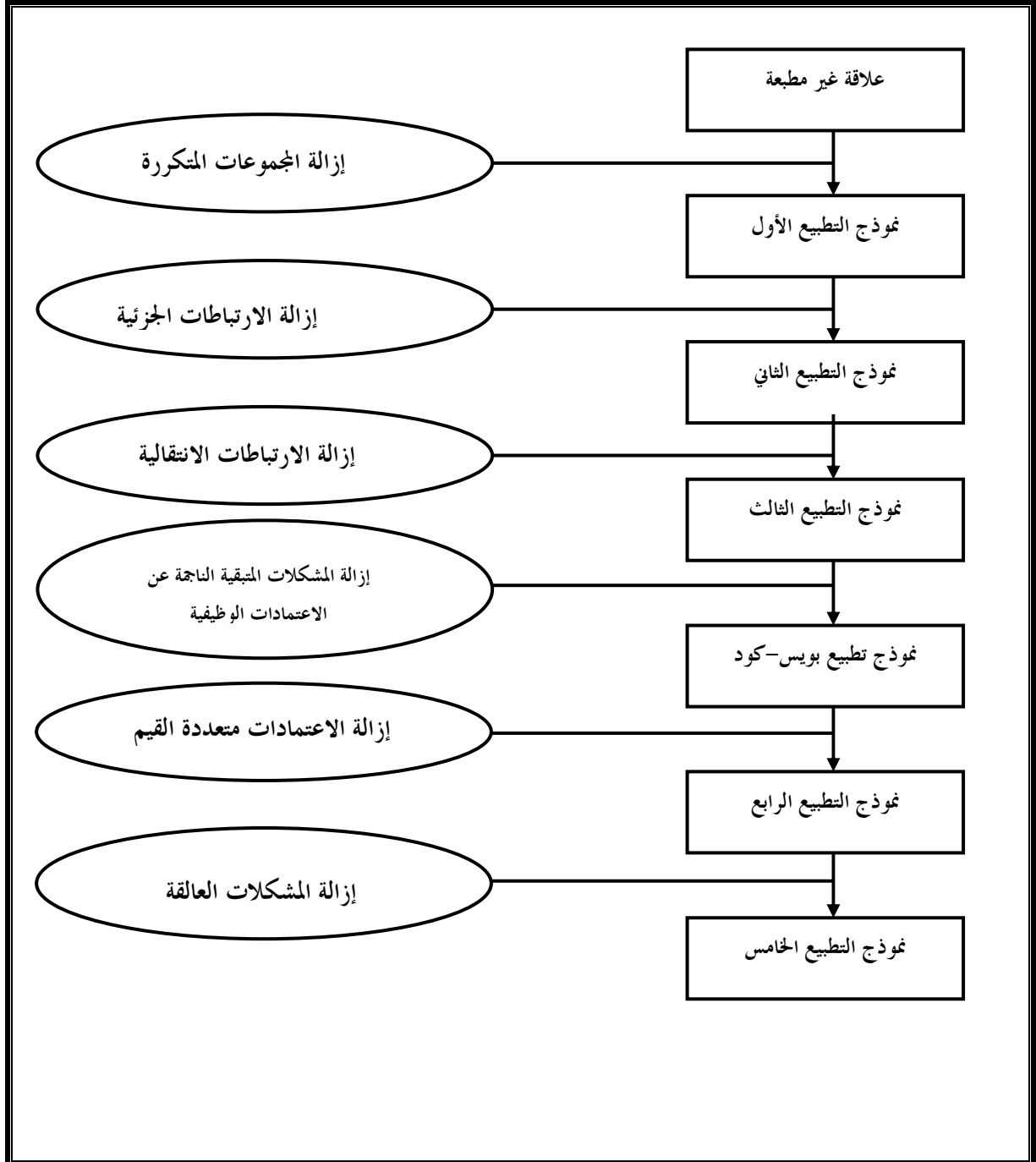
وللتحويل إلى التطبيع الثاني يكون بإزالة الارتباطات الجزئية كما هو موضح أدناه.



والتحويل إلى التطبيع الثالث يكون بإزالة الارتباطات الانتقالية كما هو موضح أدناه.



فالتطبيع (Normalization) إذاً هو عبارة عن عملية تجميع الخصائص بشكل علاقات مرتبة وتنفيذ عملية تطبيع البيانات في عدة مراحل كما يبين ذلك الشكل (1).



الشكل (1)

إن كل مرحلة تمثل شكل من أشكال التطبيع وأشكال التطبيع Normal Forms عبارة عن شكل من أشكال العلاقة تقابل نوع من الاعتمادات التي تحكم العلاقة. وفيما يلي وصف لأشكال التطبيع المختلفة:

(1) نموذج التطبيع الأول (1NF) **First Normal Form (1NF)**: عبارة عن نموذج يتحدد فيه قيمة واحدة للعلاقة التي يتقاطع فيها صف مع عمود. أي يجب إزالة الجداول المركبة (جدول داخل جدول).

(2) نموذج التطبيع الثاني (2NF) **Second Normal Form (2NF)**: عبارة عن علاقة لنموذج التطبيع الأول ويرتبط فيها كل مفتاح غير رئيس non key attribute وظيفياً مع المفتاح الرئيس Primary Key أي أن كل قيمة تعتمد على المفتاح الرئيس.

(3) نموذج التطبيع الثالث (3NF) **Third Normal Form (3NF)**: عبارة عن علاقة لنموذج التطبيع الثاني والتي لا يرتبط فيها أي مفتاح غير رئيس non key attribute وظيفياً مع أي مفتاح غير رئيس آخر.

مثال (1): يبين الجدول (1.1) علاقة غير مطبوعة، ولنعمل الآن على وضع هذه العلاقة في التطبيع الأول والثاني والثالث مع تحديد المفاتيح الرئيسية.

### علاقة غير مطبوعة ( Unnormalized Relation )

#### جدول (1.1): Grade \_ Report

Student Number	Student Name	Major	Course Number	Course Title	Instructor Name	Instructor Department	Grade
8100	Ahmad	Computer	1383	Dbase	Sami	Comp.sc.	B
			6101	Math 1	Maher	Math	A
91100	Rania	Math	1181	C++	Sami	Comp.sc.	C
			6101	Math 1	Maher	Math	B
			1280	Algebra	Reem	Math	C

لتغيير اسم المقرر Math 1 مثلاً يتطلب الأمر بحث جميع التكرارات ، والجدول أعلاه ليس في التطبيع الأول لأنه جدول غير ثنائي البعد .

وعمليات التطبيع تكون كما يلي :

**جدول (2.1): Student**

<u>Student Number</u>	<u>Student Name</u>	<u>Major</u>
8100	Ahmad	Computer
91100	Rania	Math

ومن الملاحظ أن الجدول (2.1) في التطبيع الثالث (3NF) لأنه لا يرتبط فيه أي مفتاح غير رئيس وظيفياً مع أي مفتاح غير رئيس آخر .

**جدول (3.1): Course- Student**

<u>Student Number</u>	<u>Course Number</u>	<u>Course Title</u>	<u>Instructor Name</u>	<u>Instructor Department</u>	<u>Grade</u>
8100	1383	Dbase	Sami	Comp.sc.	B
8100	6101	Math 1	Maher	Math	A
91100	1181	C++	Sami	Comp.sc.	C
91100	6101	Math 1	Maher	Math	B
91100	1280	Algebra	Reem	Math	C

الجدول أعلاه في التطبيع الأول لأنه جدول ثنائي البعد. لاحظ أن (Grade) تعتمد على مفتاح مركب مكون من Course Number + Student Number .

حيث لا يمكن إضافة أي مقرر جديد إلا إذا سجل له طالب واحد على الأقل لان رقم الطالب جزء من المفتاح المركب، لذلك يجب إزالة أي محددة غير أولية ليست اعتماداً كاملاً على المفتاح الرئيس وعندها نحصل على جدول (4.1) أدناه في التطبيع الثاني (2NF) .

**جدول (4.1): Course - Instructor**

<u>Course Number</u>	<u>Course</u>	<u>Instructor</u>	<u>Instructor</u>
----------------------	---------------	-------------------	-------------------

	<b>Title</b>	<b>Name</b>	<b>Department</b>
1383	Dbase	Sami	Comp.sc.
6101	Math 1	Maher	Math
1181	C++	Sami	Comp.sc.
6101	Math 1	Maher	Math
1280	Algebra	Reem	Math

أما الجدول (5.1) (Registration) فإنه في التطبيع الثالث (3NF).

### جدول (5.1): Registration

<u>Student Number</u>	<u>Course Number</u>	<u>Grade</u>
8100	1383	B
8100	6101	A
91100	1181	C
91100	6101	B
91100	1280	C

وبدراسة الجدول (4.1) نجد أن العلاقة بين (Course Number) و (Instructor Department) علاقة متعدية ، حيث أن العلاقة الأساسية هي بين (Instructor Name) و (Instructor Department). فبعد إزالة اعتماد تعدية للمحددات غير الأولية عن المفتاح الرئيس نحصل على التطبيع الثالث (3NF) كما في الجدولين (6.1)، (7.1) أدناه.

### جدول (6.1)



<u>Course Number</u>	<u>Course Title</u>	<u>Instructor Name</u>
1383	Dbase	Sami
6101	Math 1	Maher
1181	C++	Sami
1280	Algebra	Reem

ومن الجدير ملاحظة أن Instructor Name هو المفتاح الأجنبي هنا حيث انه مفتاح غير رئيس في علاقة ما ويصبح مفتاحا رئيسا في علاقة أخرى .

### جدول (7.1): Instructor

<u>Instructor Name</u>	<u>Instructor Department</u>
Sami	Comp.sc.
Maher	Math
Reem	Math

مثال(2): أوجد التطبيق الأول والثاني والثالث للجدول (8.1).

### الجدول (8.1): علاقة تمثل قائمة طلبيات

رقم الطلبة	تاريخ الاستلام	رقم المورد	أسم المورد	رقم القطعة	اسم القطعة	الكمية المطلوبة	الكمية الموجودة
1000	2005/05/1	200003	شركة مرسيدس	B101 C102	.....	.....	....

والتي تمثل قائمة بالطلبات لمؤسسة تقوم باستيراد قطع سيارات؛ حيث تقوم هذه المؤسسة بطلب عدد من قطع السيارات ضمن طلبية واحدة من مورد واحد وبكمية معينة لكل قطعة.

### حل مثال (2):

بالنظر إلى الجدول (8.1) نجد أن كل سطر من الجدول يمثل طلبية وبذلك فإنه عند تقاطع هذا السطر مع العمود المعنون رقم القطعة نجد وجود أكثر من قطعة (B101,C102). وتهدف عملية التطبيع إلى تبسيط هذه العلاقة بحيث يمكن الحصول على علاقة أخرى لها مفتاح رئيس يؤدي استخدامه إلى جعل التقاطع المذكور يحتوي قطعة واحدة. وباختيار المفتاح الرئيس مكوناً من رقم الطلبية + رقم القطعة فإننا نحقق هذا الهدف. وبذلك فإننا نحصل على التطبيع الأول (First Normal Form) وباستخدام هذا المبدأ نحصل على الجدول (9.1) والمكافئ للجدول الأصلي ولكنه في التطبيع الأول.

### الجدول (9.1): التطبيع الأول

رقم الطلبية + رقم القطعة	رقم المورد	اسم المورد	تاريخ الاستلام
B101+1000	.....	.....	.....
C102+1000	.....	.....	.....

وبحصولنا على هذا الجدول نجد أن بعض الخصائص تعرف بجزء من المفتاح الرئيس. وهنا نضع هدفاً آخر هو أن تكون الخصائص معتمدة على المفتاح الرئيس كله وليس على جزء منه. وهذا يؤدي إلى تقسيم هذا الجدول إلى الجداول التالية ونكون بذلك قد حصلنا على التطبيع الثاني (الجدول 10.1 ، 11.1 ، 12.1).

### الجدول (10.1)

الكمية الموجودة	رقم القطعة
.....	B101
.....	C102

الجدول (11.1)

الكمية المطلوبة	رقم الطلبية + رقم القطعة
.....	B101+1000
.....	C102+1000

الجدول (12.1)

اسم المورد	رقم المورد	تاريخ الاستلام	رقم الطلبية
شركة مرسيدس	200003	2005/05/1	1000

وبدراسة الجدول (12.1) نجد أن العلاقة بين رقم الطلبية واسم المورد علاقة متعدية، حيث أن العلاقة الأساسية هي بين رقم المورد وأسم المورد. وهنا نضع هدفاً جديداً هو إزالة هذه العلاقة المتعدية وبذلك نحصل على جدولين من الجدول (12.1) كما هو مبين في الجدول (13.1) ، (14.1) وهي علاقات في التطبيع الثالث.

الجدول (14.1)

الجدول (13.1)

رقم المورد	تاريخ الاستلام	رقم الطلبية	اسم المورد	رقم المورد
200003	2005/05/1	1000	شركة مرسيديس	200003

ويمكن تمثيل الجداول (10.1، 11.1، 13.1، 14.1) بالعلاقات التالية:  
 علاقة القطعة (رقم القطعة، الكمية الموجودة)  
 علاقة الطلبية (رقم الطلبية، تاريخ الاستلام، رقم المورد)  
 علاقة محتويات الطلبية (رقم الطلبية + رقم القطعة ، الكمية المطلوبة)  
 علاقة المورد (رقم المورد، اسم المورد)  
 وهذه العلاقات جميعها في التطبيع الثالث.

### من هنا حتى نهاية الملحق - اختياري للطالب للاستفادة

(1) نموذج تطبيع بويس - كود (Boyce -Codd Normal Form (BCNF): عبارة عن نموذج لوصف التطبيع العلائقي الذي يرشح كل محدد فيه لأن يكون مفتاحاً رئيساً.

مثال (3): على تطبيع بويس كود (Boyce-Codd Normal Form)

جدول (15.4) في التطبيع الثالث (3NF) ويمكن تحويله إلى (BCNF) بعد إزالة المشكلات المتبقية الناجمة عن الاعتمادات الوظيفية .

#### جدول (15.1)

<u>Student #</u>	<u>Major</u>	<u>Advisor</u>
150	Physics	Sami
250	Math	Maher
400	Biology	Reem
500	Physics	Salim
700	Physics	Sami

لاحظ أن العلاقات في جدول (16.1) و جدول (17.1) في (BCNF)

جدول (16.1)

<u>Student #</u>	<u>Advisor</u>
150	Sami
250	Maher
400	Reem
500	Salim
700	Sami

جدول (17.1)

<u>Advisor</u>	<u>Major</u>
Sami	Physics
Maher	Math
Reem	Biology
Salim	Physics

(2) نموذج التطبيع الرابع (4NF) Fourth Normal Form: عبارة عن علاقة لنموذج تطبيع بويس - كود والتي لا تحتوي على ارتباطات متعددة القيم.

مثال (4): على التطبيع الرابع (4NF)

عزيزي الطالب انتبه إلى الملاحظات التالية أثناء اطلاعك على الجدول (18.1).

جدول (18.1)

<u>Computer</u>	<u>Package</u>	<u>Outlet</u>
Apple	Visicalc	Computerland
Apple	Applestar	Computerland
Apple	Visicalc	Byteshop
Zenith	Wordstar	Computershop
Zenith	Supercalc	Computershop
Zenith	Wordstar	Byteshop

1. لكل Computer يوجد عدة حزم packages و عدة محلات تجارية Outlets.

2. Packages و Outlets تعتمدان على آل Computer ولكنهما مستقلان عن بعضهما البعض.

فالعلاقة أعلاه تحتوي اعتماد متعدد القيم "multi valued dependency".  
يوجد مفتاح مركب Computer + package، كما ويوجد مفتاح مركب آخر هو Computer + Outlet.  
ولكي نحصل على (4NF) نتخلص من multi valued dependency وعندها نحصل على جدولين في التطبيق الرابع هما: جدول (19.1) وجدول (20.1).

جدول (19.1)

<u>Computer</u>	<u>Package</u>
Apple	Visicalc
Apple	Applestar
Zenith	Wordstar
Zenith	Supercalc

جدول (20.1)

<u>Computer</u>	<u>Outlet</u>
Apple	Computerland
Apple	Byteshop
Zenith	Computershop
Zenith	Byteshop

(3) نموذج التطبيق الخامس (5NF) Fifth Normal Form: عبارة عن علاقة لا تحتوي على اعتمادات Join Dependencies .

مثال (5): على التطبيق الخامس (5NF)

عزيزي الطالب انتبه إلى الملاحظتين التاليتين عند اطلاعك على الجدول (21.1) وهو في التطبيق الرابع (4NF) ويوجد به عدة مشاكل.

جدول (21.1):

Persons

<u>Person id</u>	<u>Skill</u>	<u>Project-id</u>
Rami	Computing	Proj 1
Rami	Computing	Proj3
Rami	Math	Proj3
Samir	Math	-----
Samir	Economics	Proj1
Ahmad	Computing	Proj1
Ahmad	Computing	Proj2

1. بعض الحقائق مخزنة مرتين ( "Ahmad" و "Rami" وتمتلك المهارة Skill من نوع "Computing").

2. يوجد حقول فارغة ( "Samir" يمتلك مهارة "Math" ولكن حالياً لا يوجد لها أي مشروع).

ولحل المشاكل أعلاه فإننا يجب أن نقسم الجدول (21.1) Persons إلى ثلاثة جداول وعندها نحصل على التطبيع الخامس (5NF).

جدول (22.1):

**Knowledge**

<u>Person-id</u>	<u>Skill</u>
Rami	Computing
Rami	Math
Samir	Math
Samir	Economics
Ahmad	Computing

جدول (23.1):

**Assignments**

<u>Person-id</u>	<u>Project-id</u>
Rami	Proj1
Rami	Proj3
Samir	Proj1
Ahmad	Proj1
Ahmad	Proj2

جدول (24.1):

**Needs-Skill**

<u>Project-id</u>	<u>Skill</u>
Proj1	Computing
Proj1	Economics
Proj2	Computing
Proj3	Computing
Proj3	Math

" انتھی "