

נרמול נתונים

מה זה נרמול (Normal Form)?
חוקים או כללים שהוגדרו על מנת להבטיח מבנה אופטימלי באופן כזה שהמידע יאוחסן בצורה האמינה היעילה והחסכונית ביותר. מאחר ונרמול משפר את יעילות מסד הנתונים הוא מקל על העבודה עמו באמצעות שפת SQL.


ישנם שישה חוקי נרמול. אנחנו נלמד את שלושת הראשונים שהם השכיחים יותר.
חשוב! חוקי הנרמול נבנים אחד על השני - כל רמת נרמול כוללת את הרמות שמתחתיה ותלויה בהן.

חוק הנרמול הראשון - 1NF (First Normal Form)

הגדרה - טבלה היא בצורה נורמלית ראשונה אם כל השדות של הטבלה מכילים רק ערכים אטומיים (ערכים שאינם ניתן לחלוקה).

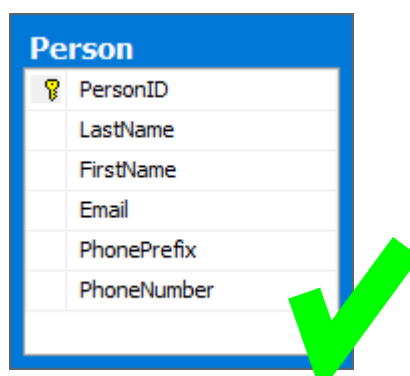
להגדרה הזו שני היבטים מעשיים:

היבט הראשון, לא נגדיר בשדה יחיד מידע אשר ניתן לחלקו לחלקים.
לדוגמה, לא נגדיר שדה שנקרא Name שמכיל גם שם פרטי וגם שם משפחה, אלא נפצל אותו לשני שדות:



PersonID
Name
Email
Phone

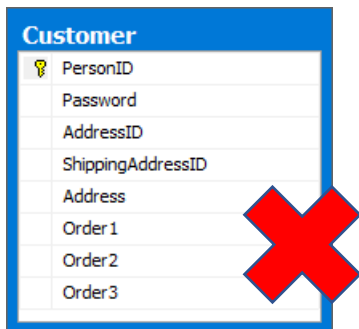
השדה Name אינו אטומי, ניתן לפצלו ל-FirstName ו-LastName.
השדה Phone אינו אטומי, ניתן לפצלו ל-PhonePrefix ו-PhoneNumber.
על מנת שהטבלה תעמוד בכלל הנרמול הראשון היא תהיה מוגדרת כך:



PersonID
LastName
FirstName
Email
PhonePrefix
PhoneNumber

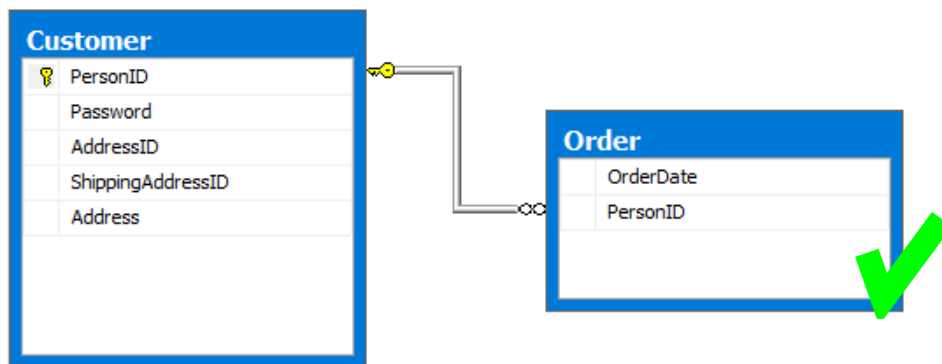
ההיבט השני, טבלה לא תכיל אוסף של שדות בעלי אותה משמעות:

לדוגמה: ע"פ הטבלה הבאה ללקוח יכולים להיות עד 3 הזמנות, מה יקרה במידה והלקוח רוצה לבצע הזמנה נוספת? לא ניתן.

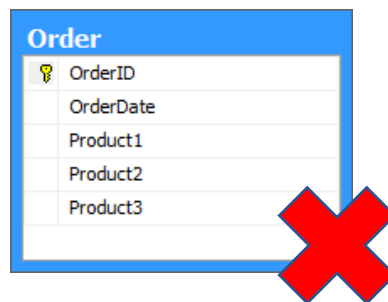


על פי חוק הנרמול הראשון אוסף ישמר בטבלה נפרדת.

באמצעות קשר **יחיד לרבים** בין הטבלה Customer לטבלה Order נאפשר ללקוח להזמין מספר לא מוגדר של הזמנות:

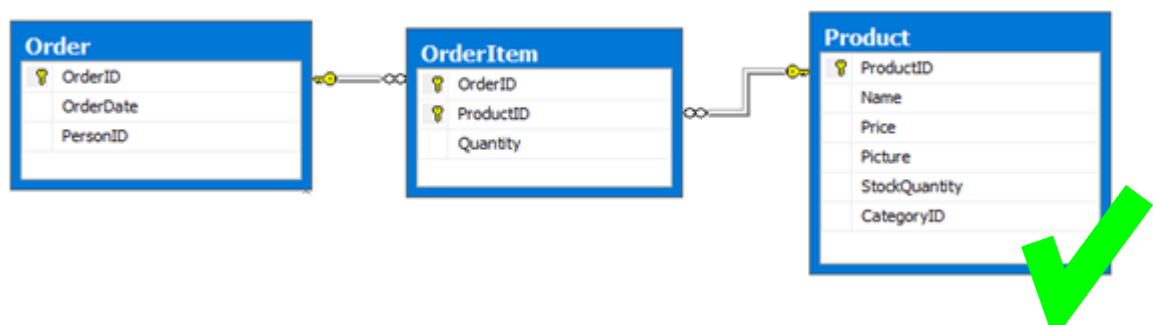


דוגמה נוספת, בהזמנה אחת יכולים להיות מספר מוצרים. אסור על פי חוק הנרמול הראשון להגדיר את הטבלה בצורה הבאה:



משום שיתכן ובהזמנה יהיה רק מוצר אחד ואז סתם בזבזנו משאבים, אבל חמור מכך, לא ניתן לטפל בהזמנות שיש בהן יותר משלושה מוצרים וזה כבר מאוד מגביל את המערכת.

נפצל אותה למספר טבלאות. הפעם נעדיף קשר **רבים לרבים** (כי בכל הזמנה יכולים להיות הרבה מוצרים, וכל מוצר יכול להיות שייך להרבה הזמנות):




חוק הנרמול השני - 2NF (Second Normal Form)

הגדרה - טבלה היא בצורה נורמלית שנייה רק אם מתקיימת רמת נרמול 1NF ואם כל השדות שאינם חלק מהמפתח תלויים באופן מלא בכל שדה המפתח.

לדוגמה:

OrderItem	
OrderID	
ProductID	
Quantity	
ProductName	
OrderDate	




השדה Quantity תלוי במלוא המפתח (תזכורת - מלוא המפתח הינו מושג הבא לציין את כל השדות המרכיבים את המפתח, במקרה הנ"ל מדובר בשני שדות) מכיוון שהוא מתייחס לכמות שהלקוח הזמין ממוצר מסוים בהזמנה מסוימת, לדוגמה: 3 יחידות ממוצר שהמזהה שלו הוא 234.

השדה ProductName אינו קשור למלוא המפתח אלא רק ל-ProductID ולכן מקומו לא בטבלה OrderItem (אלא בטבלה Product), השדה OrderDate אינו תלוי במלא המפתח אלא רק ל-OrderID ולכן מקומו בטבלה Order בלבד.

ולכן ע"פ 2NF הטבלה OrderItem צריכה להיראות כך:

OrderItem	
OrderID	
ProductID	
Quantity	




חוק הנרמול השלישי - 3NF (Third Normal Form)

הגדרה - טבלה היא בצורה נורמלית שלישית רק אם מתקיימת רמת נרמול 2NF ואם כל שדה שאיננו שדה מפתח (או חלק מהמפתח) אינו תלוי בשדה אחר שאיננו שדה מפתח.

לדוגמה:

Order	
🔑	OrderID
	OrderDate
	PersonID
	CustomerLastName
	CustomerFirstName



השדה PersonID תלוי במפתח משום שאנו חייבים לדעת לאיזה לקוח שייכת ההזמנה.

השדות CustomerFirstName ו-CustomerLastName אינם תלויים בשדה המפתח אלא קשורים ל-PersonID ולכן אינם שייכים לטבלה Order אלא לטבלה Customer שם הם כן תלויים בשדה המפתח.

Order	
🔑	OrderID
	OrderDate
	PersonID

