# **Learning Journal - Unit 3**

Computer Science, University of the People

CS 2203-01 Databases 1 - AY2024-T3

Instructor, Irfan Rashid Thoker

February 22, 2024

For this task, I was required to continue my journey into database design with this assignment focusing on redundancy of data and normalization. I am required to take an existing data structure and use normalization methods learned in this week's learning guide.

| **spID** | **spName** | **spBirthDate** | **spCitySelling** | **bonusPercent** |
| --- | --- | --- | --- | --- |
| **1** | Luke Skywalker | 1979-05-21 | Mumbai | 10% |
| **2** | Leia Organa | 1995-03-20 | Mumbai | 10% |
| **3** | Han Solo | 1963-01-06 | Calcutta | 8% |
| **4** | Rey Skywalker | 1991-10-18 | Calcutta | 8% |
| **5** | James T. Kirk | 1993-11-16 | Delhi | 9% |

First, I need to identify the source of our redundancy. In this case, we can see that the source is the connection between the city and the bonus amount. This issue is creating a state where every person who sells in a given city has the same value for the bonus set too. This redundancy will cause us to store a duplicate value along the bonus column that is based entirely on the city column. This will also see this issue causing us trouble when we want to update or modify the bonus values since we would need to verify that all cities are updated to the correct value, which could lead to mistakes and inconsistencies. We can see this if we were to imagine all salespeople working in the same city, this extreme would have us having the bonus record also being the same throughout the table.

**Primary keys & dependencies**

* The primary key: in our data is the ‘spID’ where each salesperson has his own unique identifier in the data.
* Dependencies:
  + spID 🡪 spName, spBirthDate, spCitySelling, bonusPercent
  + spCitySelling 🡪 bonusPercent
* The dependencies above clearly show us that the spID is unique to all the other attributes. We can also see that the spCitySelling determins the value in bonusPercent.

**Steps to resolve and reduce the found redundancies:**

1. First we identify the source and causes, as we saw above, the source of our redundancy is that the bonusPercent is dependant on the cpCitySelling instead of the salesperson.
2. Next we will decompose our relationship into two new relationships
   1. Salesperson (based on the old)
      1. New attributes:
         1. spID (Primary Key)
         2. spName
         3. spBirthDate
         4. spCitySelling (Foreign Key)
      2. Dependencies
         1. spID 🡪 spName, spBirthDate, spCitySelling
   2. CityBonus (new)
      1. Attributes:
         1. spCitySelling (Primary Key)
         2. bonusPercent
      2. Dependencies
         1. spCitySelling 🡪 bonusPercent

A diagram of a computer

Description automatically generated

The above diagram represents the new structure and relationship and the following tables represent the demo data from above in the new structure:

**SalesPerson Table:**

| **spID** | **spName** | **spBirthDate** | **spCitySelling** |
| --- | --- | --- | --- |
| **1** | Luke Skywalker | 1979-05-21 | Mumbai |
| **2** | Leia Organa | 1995-03-20 | Mumbai |
| **3** | Han Solo | 1963-01-06 | Calcutta |
| **4** | Rey Skywalker | 1991-10-18 | Calcutta |
| **5** | James T. Kirk | 1993-11-16 | Delhi |

**CityBonus Table:**

| **spCitySelling** | **bonusPercent** |
| --- | --- |
| **Mumbai** | 10% |
| **Calcutta** | 8% |
| **Delhi** | 9% |

**Justification and results:**

* In our new design, all the attributes in the salesperson table are fully dependent on the primary key (spID). With no extra dependencies, we are now in full compliance with 2NF and 3NF.
* In the new CityBonus table we also now have our bonusPercent dependant only on our primary key (spCitySelling) also making us compliant with 3NF.

These changes have removed the redundancy and duplicate data from our structure. And on top of that, we have made maintaining and modifying the bonusPercent easier and much more maintainable. And by changing the value in a single place we can effectively apply the change to all salespeople attached to the city the new value is applied to.

## References

* Learning Guide Unit 1-3  
  <https://my.uopeople.edu/course/view.php?id=7455>
* Sharma, N., Perniu, L., Chong, R. F., Iyer, A., Nandan, C., Mitea, A. C., Nonvinkere, M. & Danubianu, M. (2010). Database fundamentals. IBM Canada.  
  <https://my.uopeople.edu/pluginfile.php/1827130/mod_book/chapter/484065/Database_Fundamentals.pdf>
* Watt, A., & Eng, N. (2014). Database design,  2nd ed. BCcampus, BC Open Textbook Project.   
  <https://opentextbc.ca/dbdesign01/>  
  <https://my.uopeople.edu/pluginfile.php/1827130/mod_book/chapter/484065/Database-Design-2nd-Edition-1560272109.pdf>