Database Modeling

Brad Berezowski



Contents

[Assumptions and Adjustments 1](#_Toc499463130)

[Normalization Section 1](#_Toc499463131)

[First Normal Form (1NF) 1](#_Toc499463132)

[Definition 1](#_Toc499463133)

[1NF conversion: 1](#_Toc499463134)

[1NF table 1](#_Toc499463135)

[Second Normal Form (2NF) 2](#_Toc499463136)

[Definition: 2](#_Toc499463137)

[2NF Conversion: 2](#_Toc499463138)

[2NF Tables 3](#_Toc499463139)

[Third Normal Form (3NF) 4](#_Toc499463140)

[Definition 4](#_Toc499463141)

[3NF conversion 4](#_Toc499463142)

[3NF tables 4](#_Toc499463143)

[Database Design Decision 5](#_Toc499463144)

[MS Access Relation 5](#_Toc499463145)

[Database Relationship diagrams 6](#_Toc499463146)

[2NF Database Diagram 6](#_Toc499463147)

[3NF Database Diagram 7](#_Toc499463148)

[3NF Committee Diagrams 8](#_Toc499463149)

[Committee Table considerations 8](#_Toc499463150)

[Committee Conversion 8](#_Toc499463151)

[3NF Table as shown in Fig 1.4 8](#_Toc499463152)

[Calculated Query Field Solution 9](#_Toc499463153)

[Query Results 9](#_Toc499463154)

[5.1 Query 9](#_Toc499463155)

[5.1 Result 9](#_Toc499463156)

[5.2 Query 10](#_Toc499463157)

[5.2 Result 10](#_Toc499463158)

[5.3 Query 11](#_Toc499463159)

[5.3 Result 11](#_Toc499463160)

[5.4 Query 12](#_Toc499463161)

[5.4 Result 12](#_Toc499463162)

[5.5 Query 13](#_Toc499463163)

[5.5 Result 13](#_Toc499463164)

[5.6 Query 14](#_Toc499463165)

[5.6 Result 14](#_Toc499463166)

[5.7 Query 15](#_Toc499463167)

[5.7 Result 15](#_Toc499463168)

[References 16](#_Toc499463169)

[Appendix 17](#_Toc499463170)

# Assumptions and Adjustments

Some assumptions and adjustments had to be made due to lack of information. Below is a list of assumptions and adjustments that were made.

* Postal Code attribute’s value is unique and can not be duplicated. Each postal code is tied to a specific area within a City in a specific Province.
* Pay Week End Date Attribute’s value is tied to a unique date and can not be duplicated.
* Days Available Attribute: Is defined as how many days the store was opened for that week.
* Person Hours Worked attribute: As this is a calculated value, the attribute was removed from the table as tables within the database should not contain any calculated fields as it is at a higher risk of being incorrect and you also lose data integrity.
* Inc. Tax attribute: Is defined as an employee paying taxes or not.
* A Committee Meeting can not happen on multiple days.
* Job Code Date is when the employee started that position.
* A position can only have one supervisor
* The full supervisor name is good enough to identify the supervisor record and as such will not be broken down to first and last supervisor name.
* The senior programmer expressed the opinion that dates shouldn’t be used as a primary key when the normalization was being done.

# Normalization Section

## First Normal Form (1NF)

Definition: For a relation to be in First Normal Form(1NF), there can be no repeating groups, no duplicate primary key values and every column row value must be atomic. [1:82-83]

### 1NF conversion:

While a detailed description of this conversion was not required, the Committee Membership Table was not in 1NF as it had repeating values (Employees Names). More detail regarding this conversation will be listed in the Committee conversion section.

To remove repeating primary key values within the Employee Job and Payroll Data Table the following fields had to be removed and a new table called Work Week was created.

Fields removed: Payweek end Date, Days Available, Hours and OT

### 1NF table

1NF Employee Job and Payroll Data Table

Candidate Keys: {Employee ID}

Non-Prime attributes: Sin, Last Name, First Name, Street, Hire Date, Birth Date, Inc, Tax, Postal Code, City, Prov, Postion, PayRate, Supervisor, Supervisors Cell, Hours, OT, Position ID, Job Code Date

.

1NF Work Week Table

A composite key was used as EmpID and Payweek End Dates has a many to many relation. One Employee can have many Payweek End Dates IDs and one Payweek End Dates IDs can have many Employees. Since the OT and Hours Attributes are tied to both the Employee ID and the Pay Week End Date ID these attributes were attached to this table.

Candidate Keys: {Payweek End Date, EmpID}

Non-Prime Attributes: Days Available, Hours, OT

## Second Normal Form (2NF)

Definition: For a relation to be in Second Normal form (2NF), the relation must be in the First Normal Form (1NF) relation, every non-key attribute is fully functionally dependent on the primary key and partial dependencies must be removed. [1:89]

### 2NF Conversion:

While the above tables are normalized to 1NF, Employee Job and Payroll Data table is not in 2NF because of the below non-prime attributes:

* Position, PayRate, Supervisor and Supervisor Cell# is dependent upon Job Code, they are not dependent upon EmpID. This is because these attributes provide detail to the job and not the employee.
* Job Code Date is dependent upon EmpID however it is also dependent upon Job Code as Job Code date is when the employee started this position.

These non-prime attributes violate the rule saying that every non-key attribute must be fully functionally dependent on the primary key(EmpId).

The Work Week table, while it was in 1NF was not in 2NF because of the below non-Prime Attributes:

* Days Available, while days available is dependent upon the Payweek End Date as it tells how many days that week had, it is not dependent upon the EmpID as the Days available for that period will remain the same even if there is no employee.

To get the 1NF relation into a 2NF relation the Employee Job and Payroll Data table was broken up into 3 tables named Employee Table, Job Table and Job Date Table. Work Week table was broken up into 2 tables named Work Week Table and Pay Week End Date Table. Nothing was done to the Committee tables as it was already in 2NF.

2NF Tables as shown in Fig 1.2

**2NF Employee Table**

Candidate Keys: {Emp ID}

Non-Prime Attributes: SIN, Last Name, First Name, Street, Inc. Tax, Hire Date, Birth Date, City, Prov, Postal Code, Job Code (FK).

**2NF Job Table**

Candidate Keys: {Job Code}

Non-Prime Attributes: Position, PayRate, Supervisor, Supervisor Cell#

**2NF Job Table**

Candidate Keys:{Job Code, Emp ID}

Non-Prime Attributes: Job Code Date

**2NF Pay Week End Date** **Table**

As the senior developer did not want a date type to be used as a primary Key, a new attribute was created called Pay Week Date ID and that attribute will be used as a primary key for the Payweek End Date Table and Work Week Table.

Candidate Keys: {Pay Week End Date ID}

Non-Prime Attributes: Days Available, Pay Week End Date

**2NF Work Week Table**

Pay Week End date was replaced with Pay Week End Date ID

Candidate Keys: {Pay Week End Date ID, Emp ID}

Non-Prime Attributes: Hours, OT.

## Third Normal Form (3NF)

Definition: For a relation to be in Third Normal Form (3NF), the relation must be in the 2NF relation and all non-key attributes be non-transitively dependent on the primary key [1:96].

### 3NF conversion

While the Employee table is in 2NF it is not in 3NF as the City and Prov attributes are transitively dependent on the primary key (EmpID) and the Postal Code. The Job Table also does not pass the 3NF check as Supervisor Cell # is transitively dependent on the primary key (Job Code) and the Supervisor. As such these values will have to be taking out of the tables and placed in two new tables.

3NF tables as shown in Fig 1.3

**3NF Postal Code** table

Since a Postal Code is already unique, as per my assumption, I used that as my primary key.

Candidate Keys: {Postal Code}

Non-Prime Attributes: City, Province

**3NF Supervisor Table**

Since, there can be another supervisor with the same first and last name, I created a primary key called Supervisor ID.

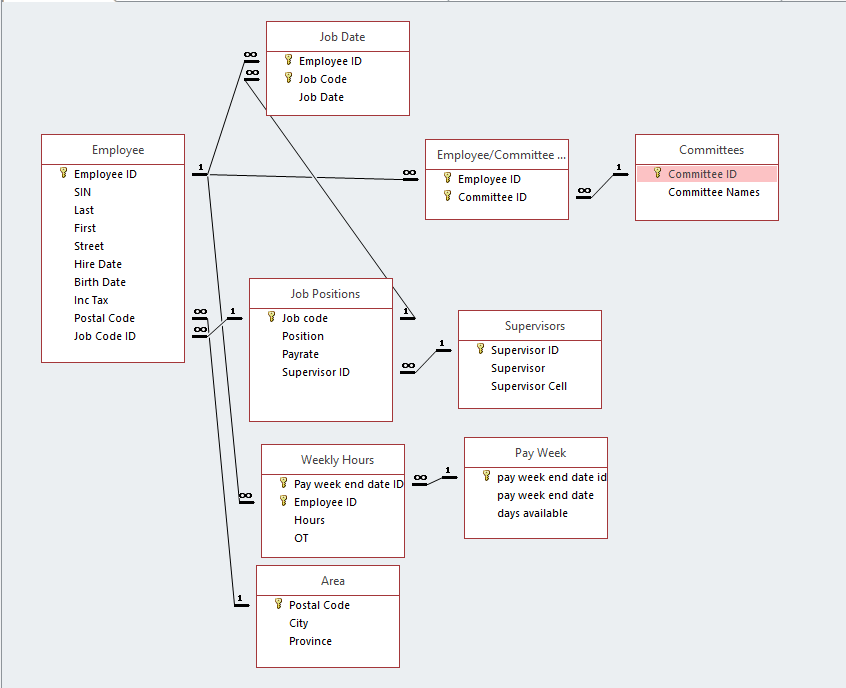
Candidate Keys: {Supervisor ID}

Non-Prime Attributes: Supervisor, Supervisor Cell #

Nothing had to be done to the Committee table as it was already in 3NF

# Database Design Decision

## MS Access Relation

 Fig 1.1

To address the point of interest about the committee being made up with many employees and an employee can serve on many committees a Junction table was created. The junction table consisted of Employee ID and Committee ID so this allowed a many to many relation to be created between the Committees table and Employees table.

To address the concern of the senior developer not liking dates as a key field, during the 2NF relation creation, a new attribute was created called Pay Week End Date ID which was made the primary key within the Pay Week End Date table and Work Week Table.

To address the concern of the accountant, the Person Hours Worked column was removed from the table. This value is to be calculated within an Access Query - Total Hours Worked per pay week: which sums the hour and OT of all employees that worked during the pay period. By doing this, it got rid of the chance of data integrity loss and bad data surrounding this previous field.

## Database Relationship diagrams

### 2NF Database Diagram



### 3NF Database Diagram



### 3NF Committee Diagrams



## Committee Table considerations

The committee membership table had multiple employees per committee and multiple committees per employee and as such I created a junction table to address this. By doing this both tables became 3NF. See Figure 1.4

# Committee Conversion

### 3NF Table as shown in Fig 1.4

**3NF Committee Membership Table**

This table is a junction table as one employee can belong to many committees and one committee can have many employees.

Candidate Keys: {Committee ID, Emp ID}

**3NF Committee Table**

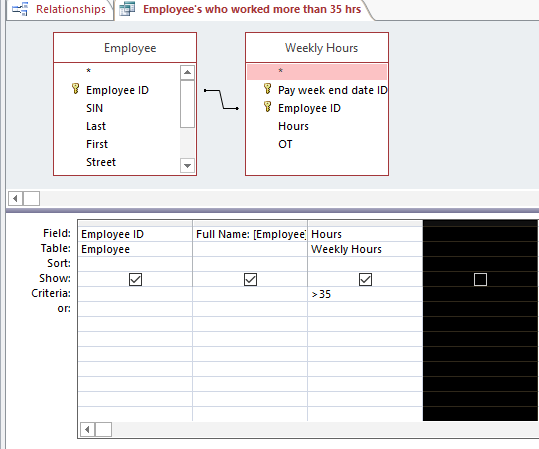
Candidate Keys: {Committee ID}

Non-Primary Attributes: Committee Name, Meeting Night

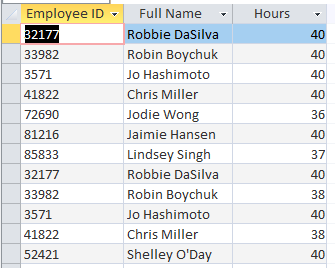
## Calculated Query Field Solution

# Query Results

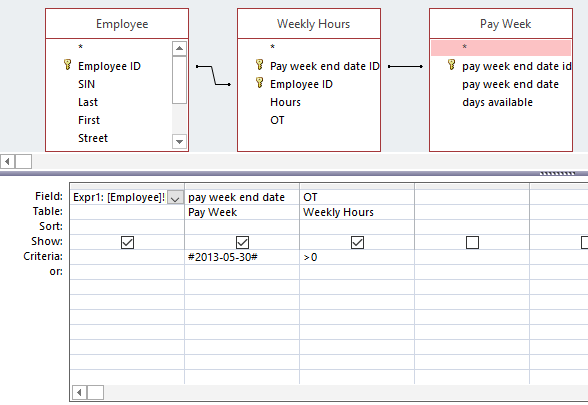
## 5.1 Query



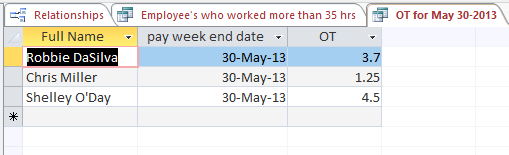
## 5.1 Result



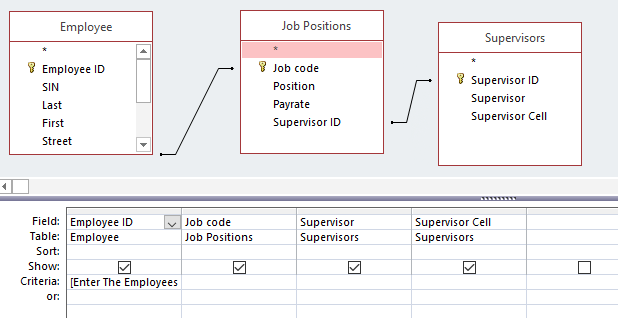
## 5.2 Query



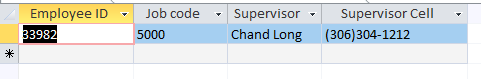
## 5.2 Result



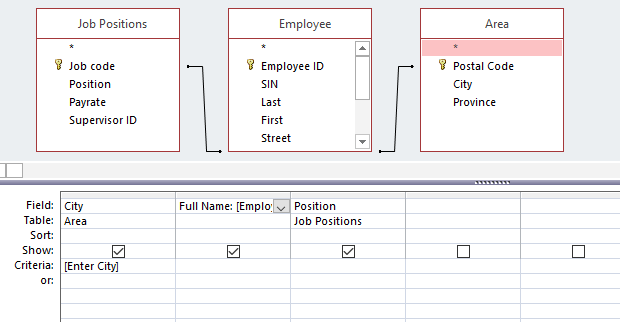
## 5.3 Query



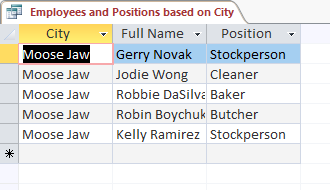
## 5.3 Result



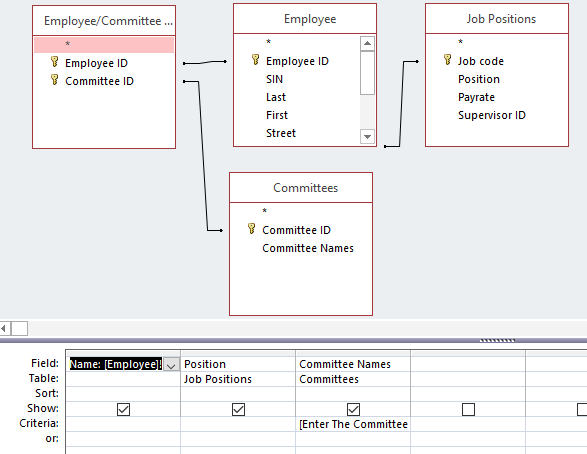
## 5.4 Query



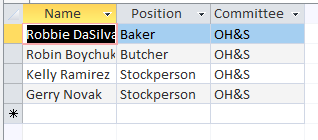
## 5.4 Result



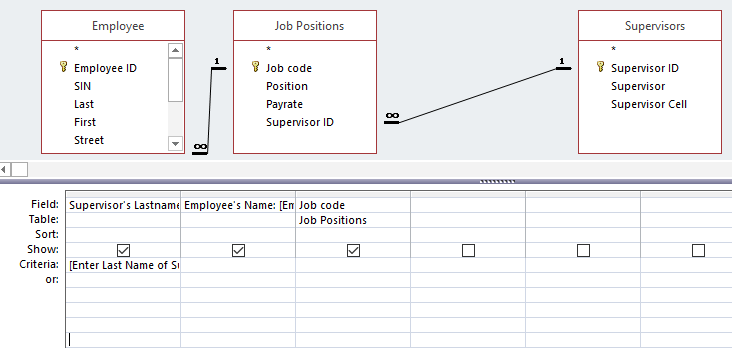
## 5.5 Query



## 5.5 Result



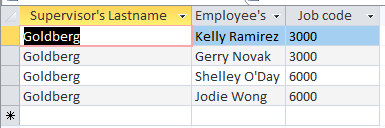
## 5.6 Query



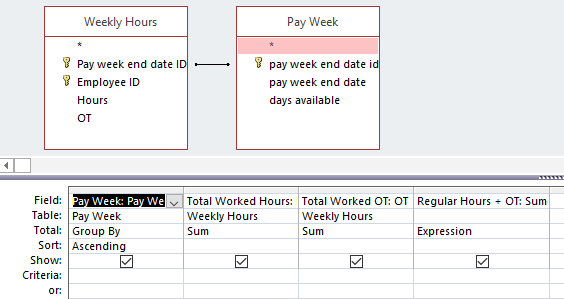
Query that is contained in the field “Supervisor’s LastName”

Supervisor's Lastname: Right([Supervisors]![Supervisor],Len([Supervisors]![Supervisor])-InStr([Supervisors]![Supervisor]," "))

## 5.6 Result



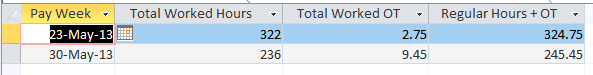
## 5.7 Query



Query that is contained in the filed Regular hours + OT

Regular Hours + OT: Sum([Weekly Hours]![Hours]+[Weekly Hours]![OT])

## 5.7 Result



# References

[1]G. Powel. “Understanding Normalization,” in *Beginning Database Design*. Indianapolis, IN: Wiley Publishing, 2006. pp. 73-104.

# Appendix

Report based on Query 5.1

Report based on Query 5.4



Report based on Query 5.5



Report based on Query 5.6

