**Project (Due Date: July 07, 2020)**

|  |  |
| --- | --- |
| ID: 47998747 | Name: Paritosh Rai |

**Description of Data requirements:**

*A full description to the data requirements you gathered to capture all specifications that is needed for this project.*

A new tier 1 cell phone operator is trying to build a new network across USA with 100,000 sites. They wanted to build a database to keep track New Site Built (NSB) of cell site and progress made on each. This database is not only needed to track the work progress but also budget management and can be used to manage the NSB project.

Each cell site will have id (unique identifier) and site name along with site address containing street name, number, city, state, and zip, region, and market area. It also includes site type.

Turf Vendors (local contracting companies) build the sites. Turf vendor are General Contactor that will be responsible to build the sites. The database will have details of the turf vendors, like name, turf vendor id (unique identifier), address, point of contact, phone numbers, SoW No (Statement of Work), SoW signed date, and price per NSB built (New Site Built).

NSB Site trackers will contain site identification like NSB id (that will be similar to site\_id as unique identifier), site name, vendor performing the work, various milestones from lease sign, the power connection delivered, connectivity, construction complete, equipment delivery, Equipment installation, and site on air. If dates have passed, then they are actualized date i.e. milestone is achieved and if it is future date than it is forecasted date for milestones.

**ER Diagram or Model:**

*Draw an ER diagram for the above problem using this text’s ER notation, the Visio notation, or the subtypes inside super types notation, as specified by your instructor to captures all requirements.*

Entity Relational (ER) Diagram are the visual tool to capture the data requirement systematically and to help design the database without missing the critical component of database. This is a very good to communicate with users to ensure all the aspect of their requirements and needs are covered without understanding the details of each field and table. The ER model captures Entities, their attributes and relationship among them. Three prime components of ER models are:

1. **Entities:** Real world thing like place, person, or object. These are shown as Rectangles in ER diagram.

Weak Entities: Entities that do not have key attributes and need parent Entity or owner. In our ER we have not identify any Entity as weak Entity as all Entities had unique primary key.

1. **Attributes:** These are properties of Entity, like time, date, place, address etc. Represented by ellipse in ER model. We have following type of attributes used in our ER:

Simple Attributes: Can not be divided also called as atomic values. Represented by ellipse.

Composite Attributes: Possible to break down composite attributes. Site address and vendor address were used as composite attribute and was further broken down to atomic values like street number, Street name, city, state and Zip.

Derived Attribute: These are not part of physical data base. However, these values are derived from other attributes present in database. We have added on derived attribute as Time to Market. To measure how much time cell site will take to get on Air from lease sign to On Air. These attributes are shown on ER diagram by dotted ellipses.

Multivalued Attribute: Multivalued Attributes can have more than one value. We have added on multi value as telephone number of vendors. There are represented buy double eclipse.

1. **Relationship:** Relationship show association between Entities. Two critical factors of relationships are Degree of Relationship and Structural Constraints of relationship Cardinality Ratio and Participation constrained). Both are used in our ER Diagram.

Degree of Relationship: This represent number of entities participating in a relationship. All three Entities have two participating relationships, so our degree is two in all cases.

Structural Constraints:

Cardinality Ratio: Defines the number or ratio of the relationship between two entities or entity sets. Different type of cardinality ratios are 1:1, 1: n and n: m. We have only 1: n ratio in our ER diagram.

Participations (Participation Constraint): This measures all instances should participate or not all of them. Is participation total (represented by double line) or partial (shown as single line).

Recursive relation: On recursive relationship was used in ER between Site and NSB. As all the NSB will be Sites and all the sites will be in NSB.

.

We will follow following process to build ER Model:

ER Models are built in two phase Rough and Final model.

The Rough and Find ER Diagram is placed in Appendix ad Fig 1 and Fig 2 respectively.

**Rough ER Diagram** (Ref Fig 1):

In Rough ER First two steps are processed i.e. identity Entities and their attributes. In this three Entities were created Cell Site, NSB (New Site Built) and Turf Vendor.

Cell Site:

* Cell site Entity included following attributes Address included all the details of cell sites, including Site id, Site\_Name, Market, Region, and address.
* Address Entity was composite Attributes was broken down to atomic values like street number, Street name, city, state, and Zip.

Turf Vendor (Vendor):

* Vendor Entity included Vendor\_id, Vendor\_Name, PoC (Point of Contact), SoW (Statement of Work-Vendor contract), Date contract was signed and price to do NSB agreed in SoW.
* Vendor entity had a multi- value attribute as telephone number.
* Composite value attribute address was broken down to street number, Street Name, City, State, and zip.

NSB:

* NSB Entity included NSB\_ID and date for five milestones to track the built of new sites. All the payments will fall due when site goes on air.
* A derived attribute is also added to NSB to measure the time to market. Time to market will be measure time taken from lease signed to getting site on Air.

**Final ER Diagram** (Ref: Fig2):

Step 3,4 and 5 were covered in this phase. Final ER phase focuses on relationships. This will cover looking at

* Cardinality ratio
* Participation
* Recursive relationship

Cardinality: Final ER diagram is having 1:1(one to one) and 1:n (one to many or many to one) cardinality ration. No n:m ratio (Many to Many ratios were found.

Participation or Participation Constrain were added. Both Full and Part participation were leveraged to define the relationship between Entities.

Recursive Relation: One recursive relation was found between NSB and site. This was full participation relation too. As all are newly built in network so they will be found in both the entities.

**Relationships:**

Cell\_Site(Site) to NSB: This is recursive relation and full participation with 1:1 cardinality ratio. All the cells will be in NSB and all the NSB will be Cell site. Also, there is one to one relation between cell site and NSB. This is identified as recursive relationship as well as all cell sites and part of NSB as well.

Cell\_Site (Site) to Vendor: Every cell site will have vendor that has built the site and there will be only one vendor for each cell site. However, Vendors can work on multiple cell site. So this maintains multi to one relation. It maintains full participation and have 1:n cardinality ratio.

NSB to Vendor: Every NSB will a vendor that will be building the cell site. However, there might be a vendor that is not building any sites so it will partial relation on one side and full participation on other. Also, A vendor can work on multiple NSB and NSB will be assigned to only one vendor. So it maintains 1:n cardinality ration.

**Mapping Steps:**

After completing the Final ER diagram of Database, ER model is converted to Relational Model by mapping them in RDMS tables.

Step by Step Six (6) approach was followed to build full Mapping to ensure all entities, attributes and relationships s are captured.

Step1: Created table for each Regular Entity. All attributes were added except multiple values (Telephone number) and Derived Attributes (Time to market). Derived attribute will not be added to database as expectation is this will be calculated when needed from existing database and may keep changing with time. Composite Attributes are not included.

Identified the Primary Key (PK) for each Entity.



Step2: There is no Entity identified as Weak Entity, so no action was triggered.

Step3: Multi- Value entity in Final ER is Telephone, so added a table for Telephone with Primary Key as Vendor\_ID and Telephone Number. Vendor ID will also be identified as Foreign Key (FK) as well. As any primary key transferred to new table from its original table. This will enable FK is first entered as Primary Key in original location.



Step 4: There is no m:n relation so no action was required.

Step 5: This step focuses on 1:n relation. There are two 1:n relation:

* NSB ⬄ Vendor
* Cell\_Site ⬄ Vendor

Vendor\_Id Primary Key was added to NSB and Cell site mapping schema. Vendor\_Id Primary key from Vendor mapping was identified as Foreign Key (FK) in the schema of Cell\_Site and NSB.



Step 6: This step manages the 1:1 relation. There is only one (1) 1:1 relation between Cell\_Site and NSB. This is also recursive relation. So added NSB\_Id in Cell\_site table as FK.

This resulted in final schema.



**DDL Statements:**

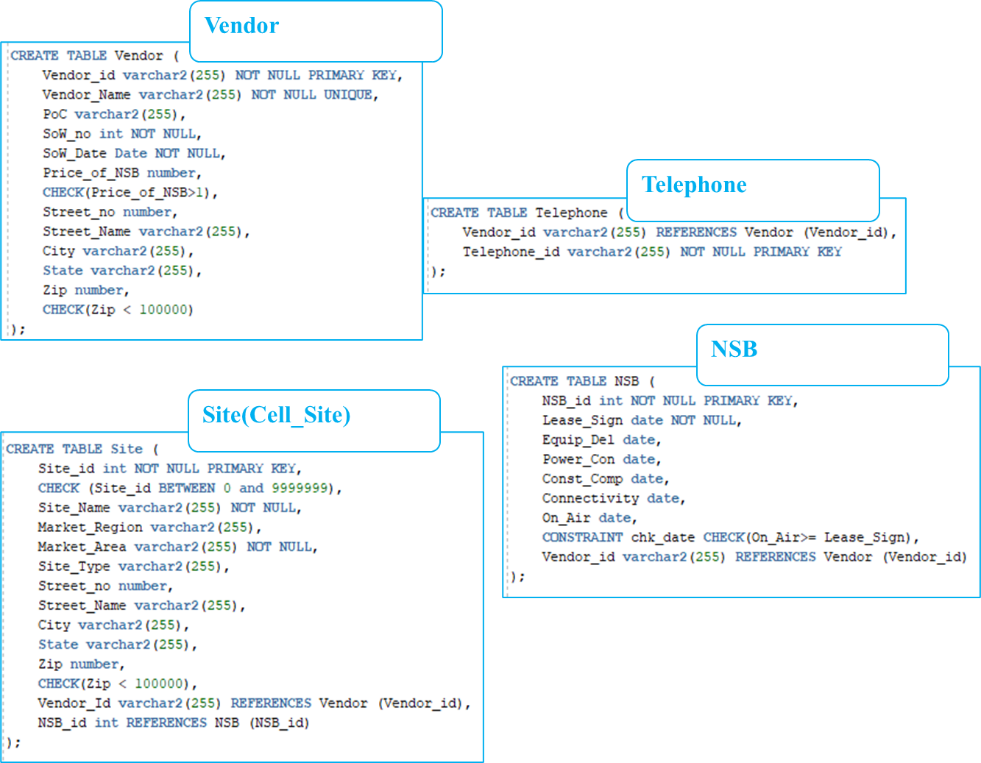
*A full DDL statements that includes all primary keys, unique keys, foreign keys, and, check constraints.*

Data Defination Language explains how data should resides in the data base. Four (4) tables were created based on final schema shared in Fig 3 of Appendix. and primary keys were defined along with foreign keys in following order:

1. Vendor
2. Telephone
3. NSB
4. Site (Cell\_Site)

Refer Fig 4 in Appendix

Ref: Create Table are in Attachment 2 in Appendix



Built following constrain to ensure data is appropriately entered:

|  |  |
| --- | --- |
| **Constrain** | **Description** |
| Not NULL | This constrain ensures column cannot store NULL Value |
| UNIQUE | Ensure each row of the column must have different values |
| Primary Key | This constrain is combination of NOT NULL and UNIQUE. This constrain ensures unique identifier. |
| Foreign Key | This ensure referential integrity of the data |
| Between | Ensure value entered stay between specified values |
| Date Check | This constrain ensure activities are happening in logical order in time domain |
| Numeric Value | Ensue value is less than, greater than or equal to certain specified value. |

**Filling of Data in Tables:**

*You should fill your tables with many rows (main transaction tables at least 100,000 records). It is allowed to use import data utilities(like excel ….etc), or writing a transcript insert statements to create random data.*

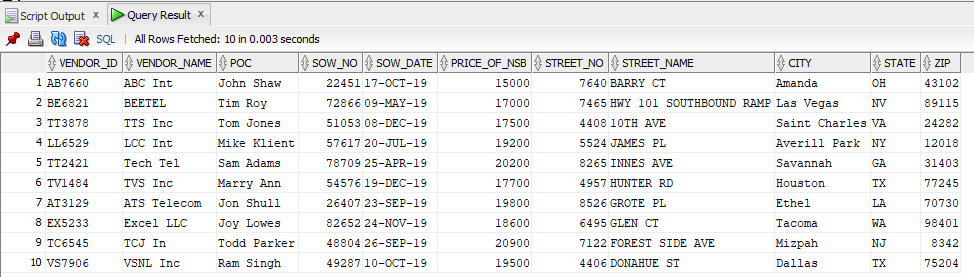
Ref: Insert Data into Table are in Attachment 3 in Appendix.

Created four tables with multiple columns and created two tables with 100,000 rows. Following is description of tables structure:

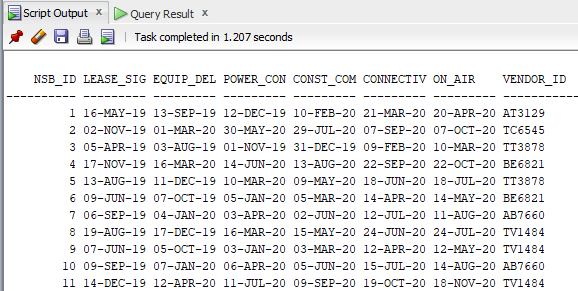
|  |  |  |  |
| --- | --- | --- | --- |
| **Table Name** | **# of Columns** | **# of Rows** | **Description** |
| **Site (Cell\_Site)** | 12 | 100,000 | Site level information like name and address |
| **NSB** | 8 | 100,000 | Track NSB built Milestone |
| **Vendor** | 11 | 10 | Turf vendor details |
| **Telephone** | 2 | 34 | Vendor Telephone number |

Detail excel sheet is enclosed in Appendix as Attachment 1. Excel was leveraged to create INSERT INTO VALUE Script to inter data into table. Excel sheet has the raw data and formula used to create the script.

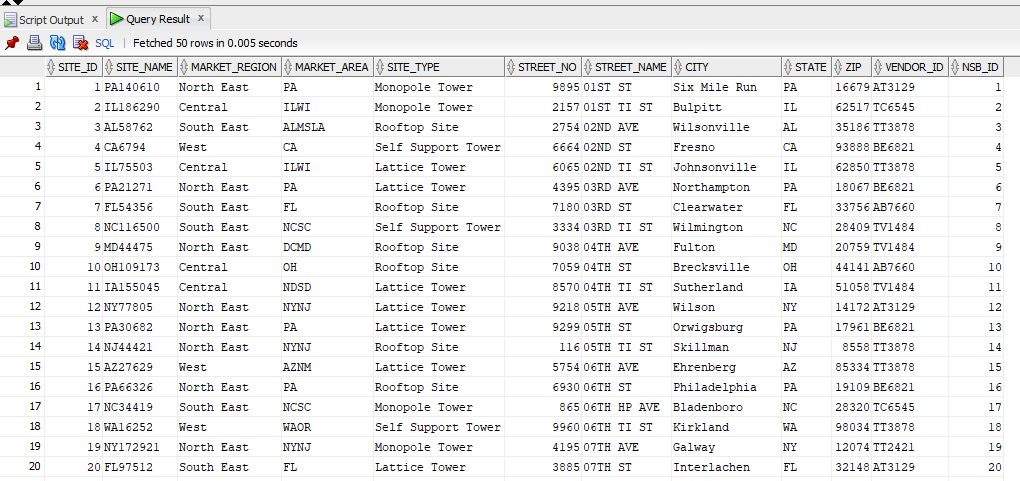
Vendor Table View: SELECT \* FROM Vendor;



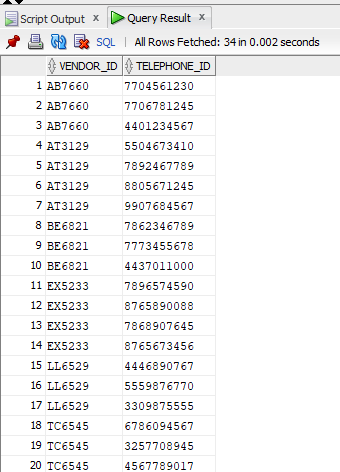
NSB Table Sample View: SELECT \* FROM NSB;



Site Table Sample View: SELECT \* FROM Site;



Telephone Table Sample View: SELECT \* FROM Telephone;



PercCa

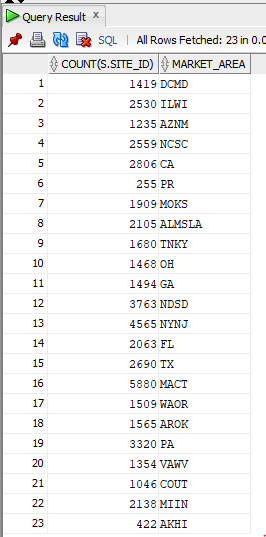
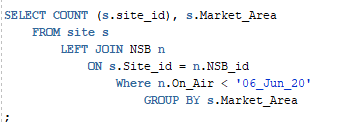
*Writing 5 complex quires to retrieve main reports you find they important in the system.(The complex queries means it should include at least one of or more of the following: aggregate function, inline queries, correlated inner query, different types of joins, set operations etc).*

Ref: SQL Statement for five queries are in Attachment 4 of Appendix

Query 1: Determine Number of sites on Air as of Jun 06th in a given Market Area.

Three key steps:

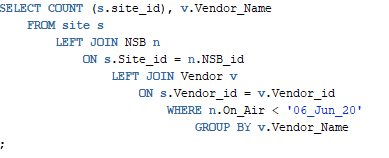
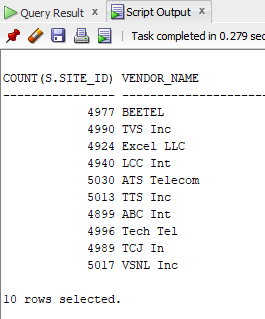
* Join table NSB and Site
* Determine sites that are on Air by Jun 06th
* Group it by Market Area



Query 2: Determine Number of sites on Air as of Jun 06th by vendor.

Three key steps:

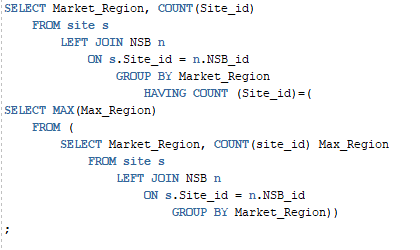
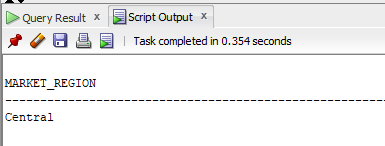
* Join table NSB, Site and Vendor Table
* Determine sites that are on Air by Jun 06th
* Group it by Vendor

Query 3: Find the Market region that has planned maximum numbers of sites.

Three key steps:

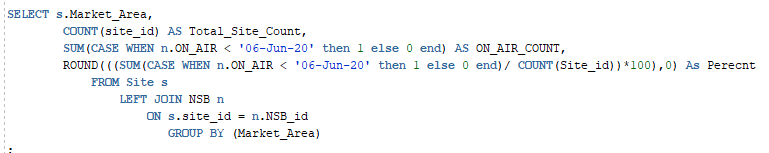
* Join table NSB and Site
* Group it by Region
* Find Max Count in having statement and interleaving sub query

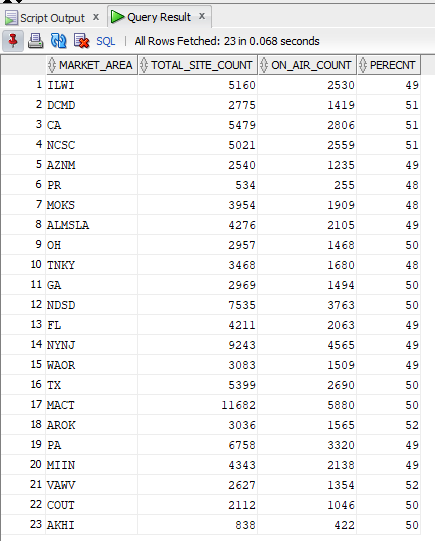
 

Query 4: Calculate percentage of sites On Air per Market Area by Jun 06

Three key steps:

* Join table NSB and Site
* Calculate total number of sites in Market area, Site on Air and Percentage
* Group it by Market Area

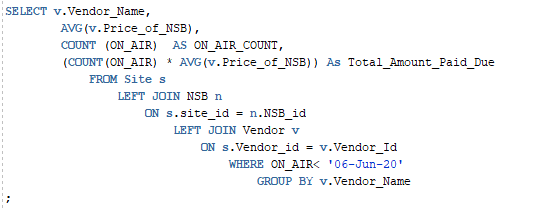


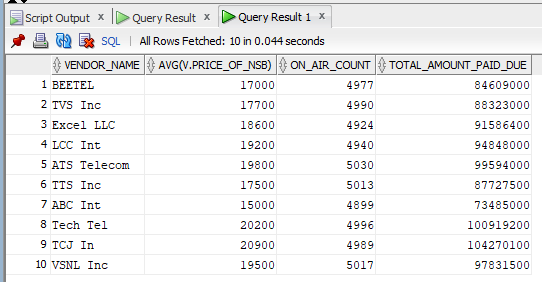


Query 5: Calculate Amount falling due to vendor as Jun 06th. Amount will fall after site is On Air.

Three key steps:

* Join table NSB, Site and Vendor
* Calculate site on Air by Jun 06th and multiple by Avg NSB price
* Group it by Vendor Name





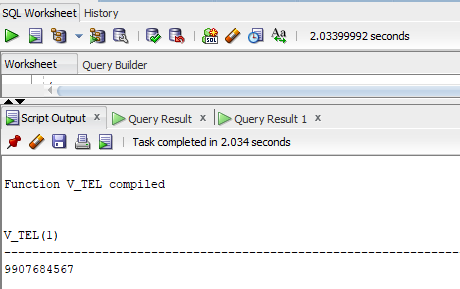
**Stored Procedures and Function:**

*Add business functionalities using stored procedures and function.*

Ref: SQL statements for creating function are in Attachment 5 in Appendix.

This is one of the daily tasks for project management team that they keep calling vendors when there is issue or problem with specific site. This Function will help us pull out Phone number when we enter the site id .



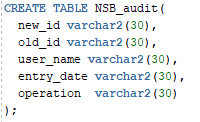


**Audit Trigger**

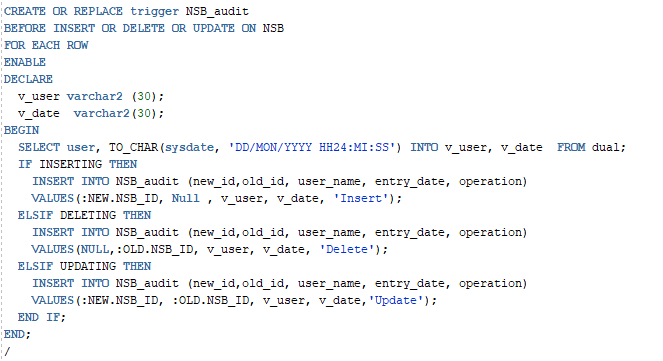
*Use DB triggers to audit the DML operations on critical tables of your DB .*

Ref: SQL statement to create and test Audit Trigger are in Attachment 6 in Appendix.

This Audit Trigger will allow us to keep an eye on the user who are inserting, updating or deleting data from the ‘NSB’ table.

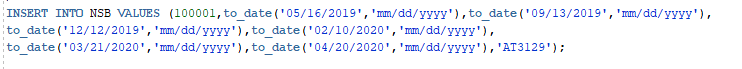
We created an NSB\_Audit table to ensure these entries are saved. This table NBS\_Audit has 5 columns which will store the auditing information such as the new data inserted or updated, old data which is updated or deleted from the table, name of the user who did it along with the date and time, also the type of DML operation performed.

 Create a trigger on the source table NSB and will store the data into the auditing table NSB\_Audit.

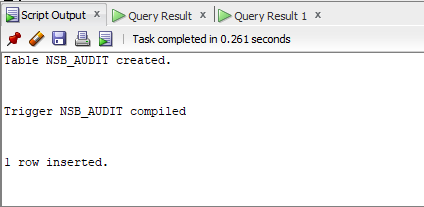


Testing the trigger: Inserted a row to test the trigger. Following are the outcome:

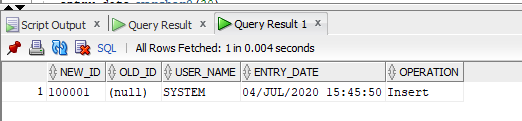
Insert Statement



Script Output:



Audit Table output: After running Insert command



Appendix

Fig 1: Rough ER Diagram



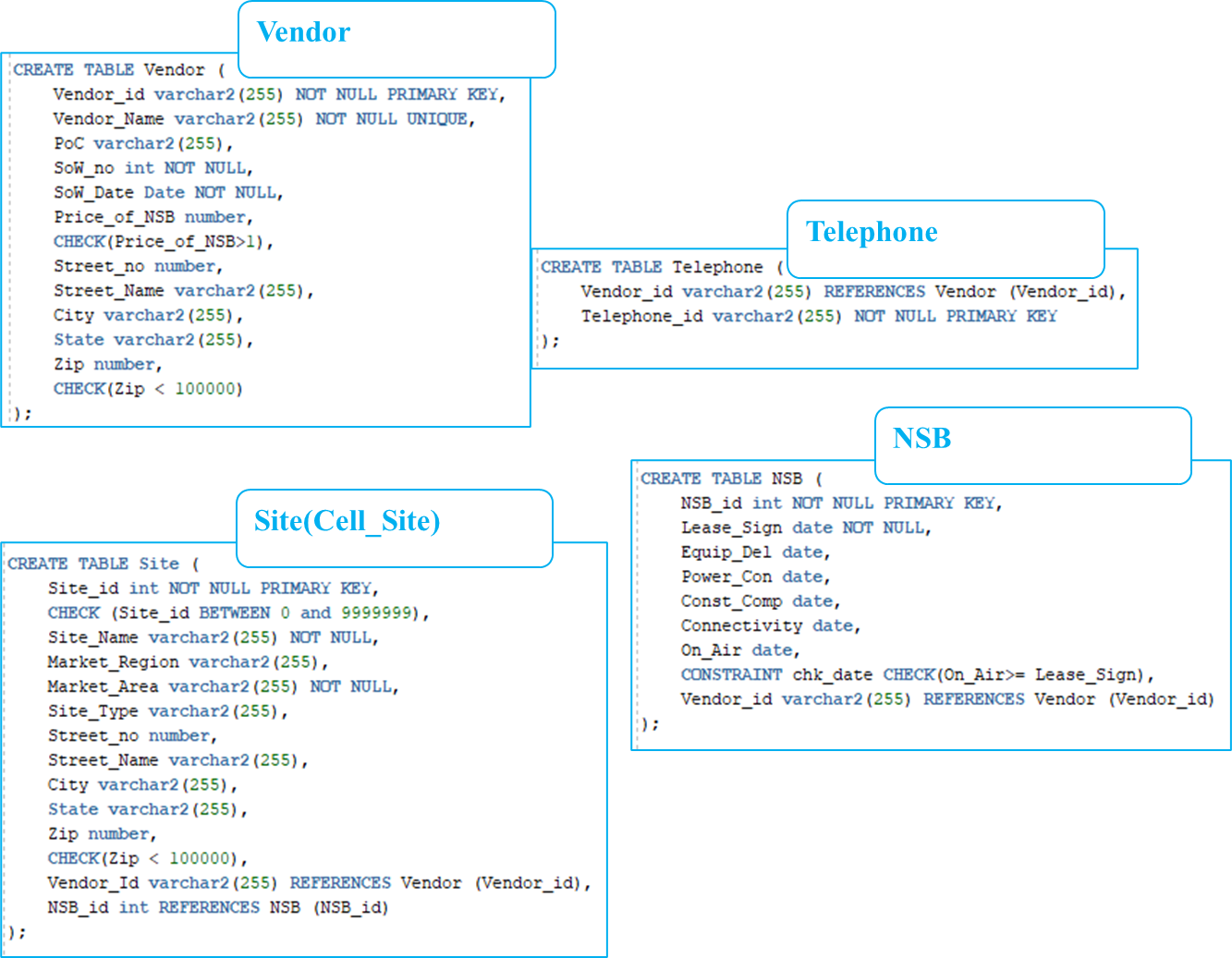
Fig 2: Final ER Diagram



Fig 3: Final Mapping Schema



Fig 4 Creating four tables:



Attachment 1: Fill data into table



Please use Note++ to open the SQL files.

Attachment 2: Create a Table in Note++ files:



Attachment 3: Insert Data into Table in Note++ files:



Attachment 4: 5 Complex Query in Note++ files



Attachment 5: Create Function



Attachment 6: Create Audit Trigger



References:

<https://www.guru99.com/er-diagram-tutorial-dbms.html>

Class Notes from DS 7330

<https://www.w3resource.com/sql/creating-and-maintaining-tables/creating-table-constraint.php>

<http://www.rebellionrider.com/table-auditing-using-dml-triggers-in-oracle-pl-sql/>