Case Study 4
ETL Using R

1. Introduction

You are to build an ETL process for the Shippers Dimension. Business rules state that this dimension should be updated daily. Yesterday's master data and corresponding dimension table is shown as follows.

Master Data (M Table)

ShipperID	CompanyName	Phone
ERLJ-101	Speedy Express	(503) 555-9831
ERLJ-102	United Package	(503) 555-8712
ERLJ-103	Federal Shipping	(503) 555-9931
ERLJ-104	Lazaro Bulk Corporation	(503) 555-2388
ERLJ-105	Pacia Estates	(503) 555-5564
ERLJ-106	We Like to Move it	(503) 555-1234
ERLJ-107	Transpo Updates	(503) 555-6520

Dimension Data (D Table)

Shipper_Key	Shipper_ID	Shipper_Name	Current_Shipper_Phone	Previous_Shipper_Phone	Effective_Date	Current_Row_Ind
1	ERLJ-101	Speedy Express	(503) 555-9831	Previous Phone Unknown	10/31/2014	Y
2	ERLJ-102	United Package	(503) 555-8712	Previous Phone Unknown	10/31/2014	Y
3	ERLJ-103	Federal Shipping	(503) 555-9931	Previous Phone Unknown	10/31/2014	Y
4	ERLJ-104	Lazaro Bulk Corporation	(503) 555-2388	Previous Phone Unknown	10/31/2014	Y
5	ERLJ-105	Pacia Estates	(503) 555-5564	Previous Phone Unknown	10/31/2014	Y
6	ERLJ-106	We Like to Move it	(503) 555-1234	Previous Phone Unknown	10/31/2014	Y
7	ERLJ-107	Transpo Updates	(503) 555-6520	Previous Phone Unknown	10/31/2014	Y

Today's data from the source ERP system is shown as follows:

Current Source Table

ShipperID	CompanyName	Phone
ERLJ-101	Speedy International	(503) 555-9831
ERLJ-102	United Package	(503) 555-8712
ERLJ-103	Federal Shipping	(503) 555-9931
ERLJ-104	Lazaro Bulk Corporation	(503) 555-2399
ERLJ-105	Pacia Estates	(503) 555-5564
ERLJ-106	We Like to Move it	(503) 555-1234
ERLJ-107	Transpo Updates	(503) 555-6520
ERLJ-108	Log Logistics	(503) 555-2374
ERLJ-109	Fast and the Slow	(503) 555-7712
ERLJ-110	Tangera Express	(503) 555-3344
ERLJ-111		(503) 555-2334
ERLJ-112	IDLF Lofistics	
ERLJ-113	United Package	(503) 555-8712

2. Preliminaries

We first copy the source and the dw MS Access database files and place them in a folder in My Documents.

2.1. Setup of Source and DW Databases

2.1.1. Open your My Documents folder and create a New Folder named "BIModule2 ETL"

2.1.2. Copy the source.db and datawarehouse.db database files and paste in inside your "My Documents\ BIModule2 ETL" folder.

2.1.3. Open R Studio from your Programs Menu

2.1.4. In the middle right corner of the R Studio window click on the three dots "**…**"

2.1.5. Search for My Documents then Select the BIModule1 ETL Folder. Click on Ok.

2.1.6. Click on More -> Set Working Directory

3. ETL Design

3.1. Extraction of Data from Source and Insert to S

3.1.1. Type the following lines of code:

```
>  library(RSQLite)
>  #Connect to the Source System
>  db <- dbConnect(SQLite(), dbname="source.db")
>
>  #Extract Today's Shippers Data
>  shippers_extract = dbGetQuery( db,'SELECT * FROM shippers' )
>
>  #Close Source Connection
>  dbDisconnect(db)
>
>  #Connect to the DW System
>  dwcon <- dbConnect(SQLite(), dbname="datawarehouse.db")
>
>  #Insert into S
>  deletesshippers = dbGetQuery(dwcon,"DELETE FROM s_shippers" )
>  dbWriteTable(conn = dwcon, name = "s_shippers",
>               value =shippers_extract, row.names = FALSE, append = TRUE)
```

3.1.2. All of these Rows then click on Run. ⟶ Run

3.1.3. To Verify the Extraction Type and Run:

```
>  dbGetQuery(dwcon, "SELECT * FROM s_shippers")
```

3.1.4. The Console should look like this:

```
> sqlQuery(dwcon, "SELECT * FROM S_Shippers")
    ShipperID              CompanyName             Phone
1   ERLJ-101     Speedy International  (503) 555-9831
2   ERLJ-102           United Package  (503) 555-8712
3   ERLJ-103          Federal Shipping (503) 555-9931
4   ERLJ-104 Lazaro Bulk Corporation  (503) 555-2399
5   ERLJ-105             Pacia Estates (503) 555-5564
6   ERLJ-106       We Like to Move it  (503) 555-1234
7   ERLJ-107          Transpo Updates  (503) 555-6520
8   ERLJ-108            Log Logistics  (503) 555-2374
9   ERLJ-109        Fast and the Slow  (503) 555-7712
10  ERLJ-110          Tangera Express  (503) 555-3344
11  ERLJ-111                     <NA>  (503) 555-2334
12  ERLJ-112           IDLF Lofistics             <NA>
13  ERLJ-113           United Package  (503) 555-8712
```

3.2. Get New and Changed Data and Insert into X

3.2.1. Go back to R Studio and type the following lines of code:

```
> #Get New and Changed Data
> s_table_new_data <- dbGetQuery(dwcon, "SELECT * FROM s_shippers WHERE
>              shipperid NOT IN (SELECT shipperid FROM m_shippers)")
> s_table_changed_company_name <-dbGetQuery(dwcon, "SELECT s.ShipperID,
  s.CompanyName,
>              s.Phone FROM S_Shippers s INNER JOIN
>              M_Shippers m ON s.ShipperID = m.ShipperID
>              WHERE NOT s.CompanyName = m.CompanyName")
> s_table_changed_phone_number <-dbGetQuery(dwcon, "SELECT s.ShipperID,
  s.CompanyName,
>              s.Phone FROM S_Shippers s INNER JOIN
>              M_Shippers m ON s.ShipperID = m.ShipperID
>              WHERE NOT s.Phone = m.Phone")
> s_table_changed_data <-
  rbind(s_table_changed_company_name,s_table_changed_phone_number)
> s_table_extract <- rbind(s_table_new_data,s_table_changed_data)
>
> #Insert into X
> deletequery=dbGetQuery(dwcon, "DELETE FROM X_Shippers")
> dbWriteTable(conn = dwcon, name = "X_Shippers", value =s_table_extract,
  row.names = FALSE, append = TRUE)
```

3.2.2. Select all these lines and click on Run. **Run**

3.2.3. To Verify the code, Type and Run:

```
> dbGetQuery(dwcon, "SELECT * FROM X_shippers")
```

3.2.4. The Console should look like this:

```
> sqlQuery(dwcon, "SELECT * FROM X_Shippers")
    ShipperID              CompanyName             Phone
1   ERLJ-108            Log Logistics  (503) 555-2374
2   ERLJ-109        Fast and the Slow  (503) 555-7712
3   ERLJ-110          Tangera Express  (503) 555-3344
4   ERLJ-111                     <NA>  (503) 555-2334
5   ERLJ-112           IDLF Lofistics             <NA>
6   ERLJ-113           United Package  (503) 555-8712
7   ERLJ-101     Speedy International  (503) 555-9831
8   ERLJ-104 Lazaro Bulk Corporation  (503) 555-2399
```

3.3. Clean X and Insert Errors into E

3.3.1. Go back to R Studio and type the following lines of code:

```
>   #Clean X
>   #Select Companies with Null Names
>   x_table_no_companyname = dbGetQuery(dwcon, "SELECT * FROM x_shippers
>                                        WHERE companyname =''")
>   x_table_no_companyname$ErrorType = "No Company Name"
>
>   #Select Duplicate Companies
>   x_table_duplicate_companies= dbGetQuery(dwcon, "SELECT * FROM x_shippers
>                                   WHERE companyname IN
>                                   (SELECT companyname FROM s_shippers
>                                   GROUP BY companyname HAVING
>                                   COUNT(companyname) > 1)")
>   x_table_duplicate_companies$ErrorType = "Duplicate Company Name"
>   x_table_errors = rbind(x_table_duplicate_companies,x_table_no_companyname)
>
>   #Set Unknown to Missing Phone Number
>   updatequery =dbGetQuery(dwcon, "UPDATE c_shippers SET Phone='Unknown Phone
>                                   Number' WHERE Phone =''")
>
>   #Insert into E
>   deletequery=dbGetQuery(dwcon, "DELETE FROM e_shippers")
>   dbWriteTable(conn = dwcon, name = "e_shippers", value =x_table_errors,
    row.names = FALSE, append = TRUE)
```

3.3.2. Select these lines and click on Run. [Run]

3.3.3. To Verify the code, Type and Run:

```
>   dbGetQuery(dwcon, "SELECT * FROM e_shippers")
```

3.3.4. The Console should look like this:

```
> sqlQuery(dwcon, "SELECT * FROM E_Shippers")
  ShipperID     CompanyName          Phone              ErrorType
1 ERLJ-113 United Package (503) 555-8712 Duplicate Company Name
2 ERLJ-111          <NA> (503) 555-2334        No Company Name
```

3.4. Select Clean Data and Insert into C

3.4.1. Go back to R Studio and type the following lines of code:

```
>   #Select Clean Data
>   x_table_clean_data = dbGetQuery(dwcon, "SELECT * FROM X_Shippers
>                                   WHERE ShipperID NOT IN (SELECT ShipperID FROM
>                                   E_Shippers)")
>
>   #Insert into C
>   query=dbGetQuery(dwcon, "DELETE FROM C_Shippers")
>   dbWriteTable(conn = dwcon, name = "C_Shippers", value =x_table_clean_data,
>                                   row.names = FALSE, append = TRUE)
```

3.4.2. Select these lines and click on Run. [Run]

3.4.3. To Verify the code, Type and Run:

```
>   dbGetQuery(dwcon, "SELECT * FROM C_Shippers")
```

3.4.4. The Console should look like this:
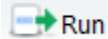
```
> sqlQuery(dwcon, "SELECT * FROM C_Shippers")
  ShipperID        CompanyName               Phone
1  ERLJ-108       Log Logistics      (503) 555-2374
2  ERLJ-109    Fast and the Slow     (503) 555-7712
3  ERLJ-110     Tangera Express      (503) 555-3344
4  ERLJ-112      IDLF Lofistics Unknown Phone Number
5  ERLJ-101   Speedy International   (503) 555-9831
6  ERLJ-104 Lazaro Bulk Corporation (503) 555-2399
```

3.5.   Update M Table

3.5.1. Go back to R Studio and type the following lines of code:

```
>   #Update M Table
>   #Select All New From C
>   c_table_new_data = dbGetQuery(dwcon, "SELECT * from C_Shippers c
>                           WHERE c.ShipperID NOT IN (SELECT m.ShipperID FROM
>                           M_shippers m )")
>   dbWriteTable(conn = dwcon, name = "M_Shippers_Test", value =c_table_new_data,
>                           row.names = FALSE, append = TRUE)
>   #Select All Changed From C
>   c_table_changed_data = dbGetQuery(dwcon, "SELECT c.* FROM C_Shippers c,
>                             M_Shippers m
>                             WHERE c.ShipperID = m.ShipperID and
>                             (c.CompanyName <> m.CompanyName or c.Phone <>
>                              m.Phone)")
>   deletequery = dbGetQuery(dwcon, "DELETE FROM M_Shippers_Test
>                             WHERE ShipperID IN (SELECT m.ShipperID FROM
>                             C_Shippers c, M_Shippers m
>                             WHERE c.ShipperID = m.ShipperID and
>                             (c.CompanyName <> m.CompanyName or c.Phone <>
>                              m.Phone))")
>   dbWriteTable(conn = dwcon, name = "M_Shippers_Test", value
>          =c_table_changed_data, row.names = FALSE, append = TRUE)
```

3.5.2. Select these lines and click on Run. **➡Run**

3.5.3. To Verify the code, Type and Run:

```
>   dbGetQuery(dwcon, "SELECT * FROM M_Shippers_Test")
```

3.5.4. The Console should look like this:

```
> sqlQuery(dwcon, "SELECT * FROM M_Shippers_Test")
   ShipperID          CompanyName               Phone
1   ERLJ-101    Speedy International    (503) 555-9831
2   ERLJ-102         United Package     (503) 555-8712
3   ERLJ-103        Federal Shipping    (503) 555-9931
4   ERLJ-104 Lazaro Bulk Corporation   (503) 555-2399
5   ERLJ-105           Pacia Estates    (503) 555-5564
6   ERLJ-106       We Like to Move it   (503) 555-1234
7   ERLJ-107        Transpo Updates     (503) 555-6520
8   ERLJ-108          Log Logistics     (503) 555-2374
9   ERLJ-109       Fast and the Slow    (503) 555-7712
10  ERLJ-110         Tangera Express    (503) 555-3344
11  ERLJ-112          IDLF Lofistics Unknown Phone Number
```

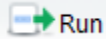**3.6.   Select From C, Transform and Insert into T**

3.6.1. Go back to R Studio and type the following lines of code:

```
>   #Select From C and Transform to DW Format
```

```
>    c_table_data = dbGetQuery(dwcon, "SELECT ShipperID as [Shipper_ID],
>                              CompanyName as [Shipper_Name], Phone as
>                              [Current_Shipper_Phone],
>                              DATE() as [Effective_Date] FROM C_Shippers" )
>    c_table_data$Previous_Shipper_Phone = "Previous_Shipper_Phone"
>    c_table_data = c_table_data[,c("Shipper_ID", "Shipper_Name",
>                              "Current_Shipper_Phone","Previous_Shipper_Phone",
>                              "Effective_Date"  ) ]
>    #Insert into T
>    query=dbGetQuery(dwcon, "DELETE FROM T_Shipper")
>    dbWriteTable(conn = dwcon, name = "T_Shipper", value =c_table_data, row.names
>                              = FALSE, append = TRUE)
```

3.6.2. Select these lines and click on Run. **Run**

3.6.3. To Verify the code, Type and Run:

```
>    dbGetQuery(dwcon, "SELECT * FROM T_Shipper")
```

3.6.4. The Console should look like this:
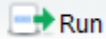
```
> sqlQuery(dwcon, "SELECT * FROM T_Shipper")
  Shipper_ID           Shipper_Name Current_Shipper_Phone Effective_Date Previous_Shipper_Phone
1   ERLJ-108           Log Logistics        (503) 555-2374     2016-03-18 Previous_Shipper_Phone
2   ERLJ-109       Fast and the Slow        (503) 555-7712     2016-03-18 Previous_Shipper_Phone
3   ERLJ-110         Tangera Express        (503) 555-3344     2016-03-18 Previous_Shipper_Phone
4   ERLJ-112         IDLF Lofistics  Unknown Phone Number     2016-03-18 Previous_Shipper_Phone
5   ERLJ-101     Speedy International        (503) 555-9831     2016-03-18 Previous_Shipper_Phone
6   ERLJ-104 Lazaro Bulk Corporation        (503) 555-2399     2016-03-18 Previous_Shipper_Phone
```

3.7. Select New From T Insert Into I

3.7.1. Go back to R Studio and type the following lines of code:

```
>    #Select New from T
>    t_table_new_data = dbGetQuery(dwcon, "SELECT t.*
>                       FROM t_shipper t LEFT JOIN d_shipper d ON t.Shipper_ID =
>                       d.Shipper_ID
>                       WHERE d.Shipper_ID IS NULL")
>    t_table_new_data$Current_Row_Ind = 'Y'
>
>    #Insert New into I
>    query=dbGetQuery(dwcon, "DELETE FROM I_Shipper")
>    dbWriteTable(conn = dwcon, name = "I_Shipper", value =t_table_new_data,
>                          row.names = FALSE, append = TRUE)
```

3.7.2. Select these lines and click on Run. **Run**

3.7.3. To Verify the code, Type and Run:

```
>    dbGetQuery(dwcon, "SELECT * FROM I_Shipper")
```

3.7.4. The Console should look like this:

```
> sqlQuery(dwcon, "SELECT * FROM I_Shipper")
  Shipper_ID        Shipper_Name Current_Shipper_Phone Previous_Shipper_Phone Effective_Date Current_Row_Ind
1   ERLJ-108       Log Logistics        (503) 555-2374 Previous_Shipper_Phone     2016-03-18               Y
2   ERLJ-109   Fast and the Slow        (503) 555-7712 Previous_Shipper_Phone     2016-03-18               Y
3   ERLJ-110     Tangera Express        (503) 555-3344 Previous_Shipper_Phone     2016-03-18               Y
4   ERLJ-112      IDLF Lofistics  Unknown Phone Number Previous_Shipper_Phone     2016-03-18               Y
```
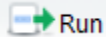
3.8. Select Updated From T Insert Into U

3.8.1. Go back to R Studio and type the following lines of code:

```
>    #Select Changed from T
```

NATIONAL ENGINEERING CENTER
University of the Philippines
Diliman, Quezon City

```
>   t_table_changed_data=dbGetQuery(dwcon, "SELECT t.*
>                     FROM t_shipper t inner join d_shipper d
>                     ON t.Shipper_ID = d.Shipper_ID
>                     WHERE (NOT t.Shipper_Name = d.Shipper_Name or NOT
>                     t.Current_Shipper_Phone = d.Current_Shipper_Phone)
>                     AND d.Current_Row_Ind = 'Y'")
>   t_table_changed_data$Current_Row_Ind = 'Y'
>
>   #Insert into U
>   deletequery=dbGetQuery(dwcon, "DELETE FROM U_Shipper")
>   dbWriteTable(conn = dwcon, name = "U_Shipper", value =t_table_changed_data,
>                     row.names = FALSE, append = TRUE)
```

3.8.2. Select these lines and click on Run. ⬛➡Run

3.8.3. To Verify the code, Type and Run:

```
>   dbGetQuery(dwcon, "SELECT * FROM U_Shipper")
```

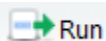3.8.4. The Console should look like this:

```
> sqlQuery(dwcon, "SELECT * FROM U_Shipper")
  Shipper_ID           Shipper_Name Current_Shipper_Phone Previous_Shipper_Phone Effective_Date
1   ERLJ-101    Speedy International       (503) 555-9831 Previous_Shipper_Phone     2016-03-18
2   ERLJ-104 Lazaro Bulk Corporation       (503) 555-2399 Previous_Shipper_Phone     2016-03-18
  Current_Row_Ind
1               Y
2               Y
```

3.9.   Select from I Insert into D

3.9.1. Go back to R Studio and type the following lines of code:

```
>   maxkey <- dbGetQuery(dwcon, "SELECT MAX(Shipper_Key) as MAX FROM D_Shipper")
>   i_table_data <- dbGetQuery(dwcon, "SELECT * FROM I_Shipper")
>   i_table_data$Shipper_Key = (maxkey[1,1]+1):(maxkey[1,1]+nrow(i_table_data))
>
>   #Insert New into D
>   i_table_data = i_table_data[,c("Shipper_Key","Shipper_ID","Shipper_Name",
>     "Current_Shipper_Phone", "Previous_Shipper_Phone", "Effective_Date",
>     "Current_Row_Ind" )]
>   dbWriteTable(conn = dwcon, name = "D_Shipper", value =i_table_data, row.names
>     = FALSE, append = TRUE)
```

3.9.2. Select these lines and click on Run. ⬛➡Run

3.9.3. To Verify the code, Type and Run:

```
>   dbGetQuery(dwcon, "SELECT * FROM D_Shipper")
```

3.9.4. The Console should look like this:

```
> sqlQuery(dwcon, "SELECT * FROM D_Shipper")
   Shipper_Key Shipper_ID           Shipper_Name Current_Shipper_Phone Previous_Shipper_Phone Effective_Date Current_Row_Ind
1            1   ERLJ-101          Speedy Express       (503) 555-9831 Previous Phone Unknown     2014-10-31               Y
2            2   ERLJ-102          United Package       (503) 555-8712 Previous Phone Unknown     2014-10-31               Y
3            3   ERLJ-103         Federal Shipping       (503) 555-9931 Previous Phone Unknown     2014-10-31               Y
4            4   ERLJ-104 Lazaro Bulk Corporation       (503) 555-2388 Previous Phone Unknown     2016-03-18               Y
5            5   ERLJ-105            Pacia Estates       (503) 555-5564 Previous Phone Unknown     2014-10-31               Y
6            6   ERLJ-106        We Like to Move it       (503) 555-1234 Previous Phone Unknown     2014-10-31               Y
7            7   ERLJ-107          Transpo Updates       (503) 555-6520 Previous Phone Unknown     2014-10-31               Y
8            8   ERLJ-108            Log Logistics       (503) 555-2374 Previous_Shipper_Phone     2016-03-19               Y
9            9   ERLJ-109        Fast and the Slow       (503) 555-7712 Previous_Shipper_Phone     2016-03-19               Y
10          10   ERLJ-110          Tangera Express       (503) 555-3344 Previous_Shipper_Phone     2016-03-19               Y
11          11   ERLJ-112           IDLF Lofistics  Unknown Phone Number Previous_Shipper_Phone     2016-03-19               Y
```

3.10. Select Type 3 Updates from U, Update D

3.10.1. Go back to R Studio and type the following lines of code:

```
>   #Select Type 3 from U and D
>   u_table_type3_data <- dbGetQuery(dwcon, "SELECT u.*
>                        from u_shipper u inner join d_shipper d
>                        on u.Shipper_ID = d.Shipper_ID
>                        where NOT (u.Current_Shipper_Phone =
>                        d.Current_Shipper_Phone)
>                        and d.Current_Row_Ind = 'Y'")
>
>   d_table_type3_data <- dbGetQuery(dwcon, "SELECT d.*
>                        from u_shipper u inner join d_shipper d
>                        on u.Shipper_ID = d.Shipper_ID
>                        where NOT (u.Current_Shipper_Phone =
>                        d.Current_Shipper_Phone)
>                        and d.Current_Row_Ind = 'Y'")
>
>   u_table_type3_data$Previous_Shipper_Phone =
    d_table_type3_data$Current_Shipper_Phone
>   u_table_type3_data$Shipper_Key = d_table_type3_data$Shipper_Key
>   u_table_type3_data =
>     u_table_type3_data[,c("Shipper_Key","Shipper_ID","Shipper_Name",
>     "Current_Shipper_Phone", "Previous_Shipper_Phone", "Effective_Date",
>     "Current_Row_Ind" )]
>
>   deletequery = dbGetQuery(dwcon, "DELETE FROM d_shipper
>                          WHERE Shipper_ID IN (SELECT d.Shipper_ID
>                        from u_shipper u inner join d_shipper d
>                        on u.Shipper_ID = d.Shipper_ID
>                        where NOT (u.Current_Shipper_Phone =
>                        d.Current_Shipper_Phone)
>                        and d.Current_Row_Ind = 'Y')")
>   dbWriteTable(conn = dwcon, name = "d_shipper", value =u_table_type3_data,
>           row.names = FALSE, append = TRUE)
```

3.10.2. Select these lines and click on Run. ⮕ Run

3.10.3. To Verify the code, Type and Run:

```
>   dbGetQuery(dwcon, "SELECT * FROM D_Shipper")
```

3.10.4. The Console should look like this:

```
> sqlQuery(dwcon, "SELECT * FROM D_Shipper")
    Shipper_Key Shipper_ID          Shipper_Name Current_Shipper_Phone Previous_Shipper_Phone Effective_Date Current_Row_Ind
1             1   ERLJ-101         Speedy Express      (503) 555-9831   Previous Phone Unknown     2014-10-31               Y
2             2   ERLJ-102         United Package      (503) 555-8712   Previous Phone Unknown     2014-10-31               Y
3             3   ERLJ-103        Federal Shipping      (503) 555-9931   Previous Phone Unknown     2014-10-31               Y
4             4   ERLJ-104 Lazaro Bulk Corporation     (503) 555-2399          (503) 555-2388     2016-03-19               Y
5             5   ERLJ-105          Pacia Estates      (503) 555-5564   Previous Phone Unknown     2014-10-31               Y
6             6   ERLJ-106       We Like to Move it     (503) 555-1234   Previous Phone Unknown     2014-10-31               Y
7             7   ERLJ-107        Transpo Updates      (503) 555-6520   Previous Phone Unknown     2014-10-31               Y
8             8   ERLJ-108          Log Logistics      (503) 555-2374   Previous_Shipper_Phone     2016-03-19               Y
9             9   ERLJ-109        Fast and the Slow     (503) 555-7712   Previous_Shipper_Phone     2016-03-19               Y
10           10   ERLJ-110        Tangera Express      (503) 555-3344   Previous_Shipper_Phone     2016-03-19               Y
11           11   ERLJ-112         IDLF Lofistics   Unknown Phone Number Previous_Shipper_Phone     2016-03-19               Y
```

3.11. Select Type 2 Updates from U, Insert into D

    3.11.1.   Go back to R Studio and type the following lines of code:

```
> #Select Type 2
> maxkey <- dbGetQuery(dwcon, "SELECT MAX(Shipper_Key) as MAX FROM D_Shipper")
> u_table_type2_data <- dbGetQuery(dwcon, "SELECT u.*
>                   FROM u_shipper u INNER JOIN d_shipper d
>                   ON u.Shipper_ID = d.Shipper_ID
>                   WHERE NOT (u.Shipper_Name = d.Shipper_Name)
>                         AND d.Current_Row_Ind = 'Y'")
> u_table_type2_data$Shipper_Key =
>  (maxkey[1,1]+1):(maxkey[1,1]+nrow(u_table_type2_data))
>
> #Insert New into D
> u_table_type2_data
>  =u_table_type2_data[,c("Shipper_Key","Shipper_ID","Shipper_Name",
>  "Current_Shipper_Phone", "Previous_Shipper_Phone", "Effective_Date",
>  "Current_Row_Ind" )]
>
> dbWriteTable(conn = dwcon, name = "d_shipper", value =u_table_type2_data,
>  row.names = FALSE, append = TRUE)
```

    3.11.2.   Select these lines and click on Run. `Run`

    3.11.3.   To Verify the code, Type and Run:

```
> dbGetQuery(dwcon, "SELECT * FROM D_Shipper")
```

    3.11.4.   The Console should look like this:

```
> sqlQuery(dwcon, "SELECT * FROM D_Shipper")
   Shipper_Key Shipper_ID          Shipper_Name Current_Shipper_Phone Previous_Shipper_Phone Effective_Date Current_Row_Ind
1            1   ERLJ-101         Speedy Express        (503) 555-9831  Previous Phone Unknown     2014-10-31               Y
2            2   ERLJ-102         United Package        (503) 555-8712  Previous Phone Unknown     2014-10-31               Y
3            3   ERLJ-103        Federal Shipping        (503) 555-9931  Previous Phone Unknown     2014-10-31               Y
4            4   ERLJ-104 Lazaro Bulk Corporation        (503) 555-2399         (503) 555-2388     2016-03-19               Y
5            5   ERLJ-105           Pacia Estates        (503) 555-5564  Previous Phone Unknown     2014-10-31               Y
6            6   ERLJ-106       We Like to Move it        (503) 555-1234  Previous Phone Unknown     2014-10-31               Y
7            7   ERLJ-107         Transpo Updates        (503) 555-6520  Previous Phone Unknown     2014-10-31               Y
8            8   ERLJ-108           Log Logistics        (503) 555-2374  Previous_Shipper_Phone     2016-03-19               Y
9            9   ERLJ-109        Fast and the Slow        (503) 555-7712  Previous_Shipper_Phone     2016-03-19               Y
10          10   ERLJ-110         Tangera Express        (503) 555-3344  Previous_Shipper_Phone     2016-03-19               Y
11          11   ERLJ-112          IDLF Lofistics  Unknown Phone Number  Previous_Shipper_Phone     2016-03-19               Y
12          12   ERLJ-101     Speedy International        (503) 555-9831  Previous_Shipper_Phone     2016-03-19               Y
```

3.12. Update Current Indicators in D

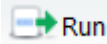   3.12.1.   Go back to R Studio and type the following lines of code:

```
> #Update Shipper Current Value
> forupdatedata <- dbGetQuery(dwcon, "SELECT d.*
>                         FROM u_shipper u INNER JOIN d_shipper d ON
>                         u.Shipper_Id = d.Shipper_Id
>                         WHERE d.Current_Row_Ind = 'Y' AND Not
>  u.Shipper_Name = d.Shipper_Name ")
> forupdatedata$Current_Row_Ind = 'N'
>
>
> deletequery = dbGetQuery(dwcon, "DELETE FROM d_shipper
>                         WHERE Shipper_Key IN (SELECT d.Shipper_Key
>                         FROM u_shipper u INNER JOIN d_shipper d ON
```

```
>                              u.Shipper_Id = d.Shipper_Id
>                              WHERE d.Current_Row_Ind = 'Y' AND Not
>    u.Shipper_Name = d.Shipper_Name)")
>
>    dbWriteTable(conn = dwcon, name = "d_shipper", value =forupdatedata,
>     row.names = FALSE, append = TRUE)
>    dbGetQuery(dwcon, "SELECT * FROM D_Shipper")
>     dbDisconnect(dwcon)
```

3.12.2. Select these lines and click on Run. ⬛→ Run

3.12.3. The Console should look like this:

```
> sqlQuery(dwcon, "SELECT * FROM D_Shipper")
   Shipper_Key Shipper_ID           Shipper_Name Current_Shipper_Phone Previous_Shipper_Phone Effective_Date Current_Row_Ind
1            1    ERLJ-101         Speedy Express        (503) 555-9831 Previous Phone Unknown     2014-10-31               N
2            2    ERLJ-102         United Package        (503) 555-8712 Previous Phone Unknown     2014-10-31               Y
3            3    ERLJ-103        Federal Shipping        (503) 555-9931 Previous Phone Unknown     2014-10-31               Y
4            4    ERLJ-104 Lazaro Bulk Corporation        (503) 555-2399         (503) 555-2388     2016-03-19               Y
5            5    ERLJ-105           Pacia Estates        (503) 555-5564 Previous Phone Unknown     2014-10-31               Y
6            6    ERLJ-106        We Like to Move it        (503) 555-1234 Previous Phone Unknown     2014-10-31               Y
7            7    ERLJ-107         Transpo Updates        (503) 555-6520 Previous Phone Unknown     2014-10-31               Y
8            8    ERLJ-108           Log Logistics        (503) 555-2374 Previous_Shipper_Phone     2016-03-19               Y
9            9    ERLJ-109         Fast and the Slow        (503) 555-7712 Previous_Shipper_Phone     2016-03-19               Y
10          10    ERLJ-110         Tangera Express        (503) 555-3344 Previous_Shipper_Phone     2016-03-19               Y
11          11    ERLJ-112           IDLF Lofistics Unknown Phone Number Previous_Shipper_Phone     2016-03-19               Y
12          12    ERLJ-101     Speedy International        (503) 555-9831 Previous_Shipper_Phone     2016-03-19               Y
```