----------creating table containing data about employees---------

CREATE TABLE "Employees" (

"EmployeeID" numeric ,

"FirstName" VARCHAR(50),

"LastName" VARCHAR(50),

"Email" VARCHAR(100),

"Phone" VARCHAR(15),

"Address" VARCHAR(100),

"City" VARCHAR(50),

"State" VARCHAR(50),

"ZipCode" VARCHAR(10),

"DepartmentID" INT,

"DepartmentName" VARCHAR(50),

"ProjectID" INT,

"ProjectName" VARCHAR(100),

"RoleID" INT,

"RoleName" VARCHAR(50),

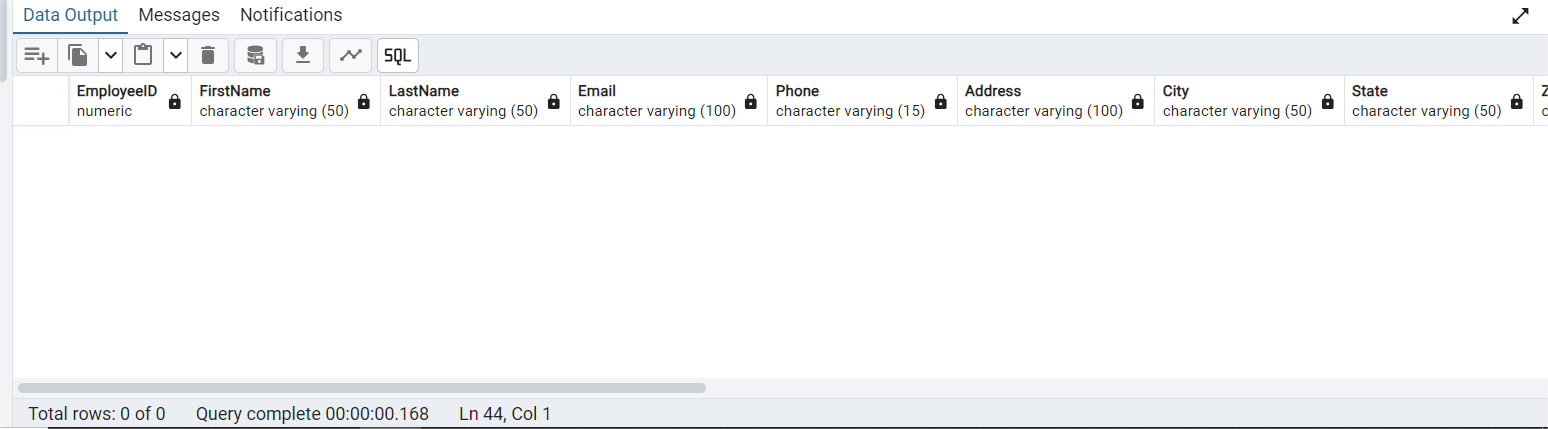
"Salary" DECIMAL(10, 2),

"HireDate" DATE,

"ManagerID" Float,

"ManagerName" VARCHAR(50),

"PerformanceRating");



-------------------------Data Extraction Using SQL--------------------------------

COPY "Employees"( "EmployeeID" ,

"FirstName" ,

"LastName" ,

"Email" ,

"Phone" ,

"Address" ,

"City" ,

"State","ZipCode",

"DepartmentID" ,

"DepartmentName" ,

"ProjectID" ,

"ProjectName" ,

"RoleID" ,

"RoleName",

"Salary" ,

"HireDate" ,

"ManagerID",

"ManagerName" ,

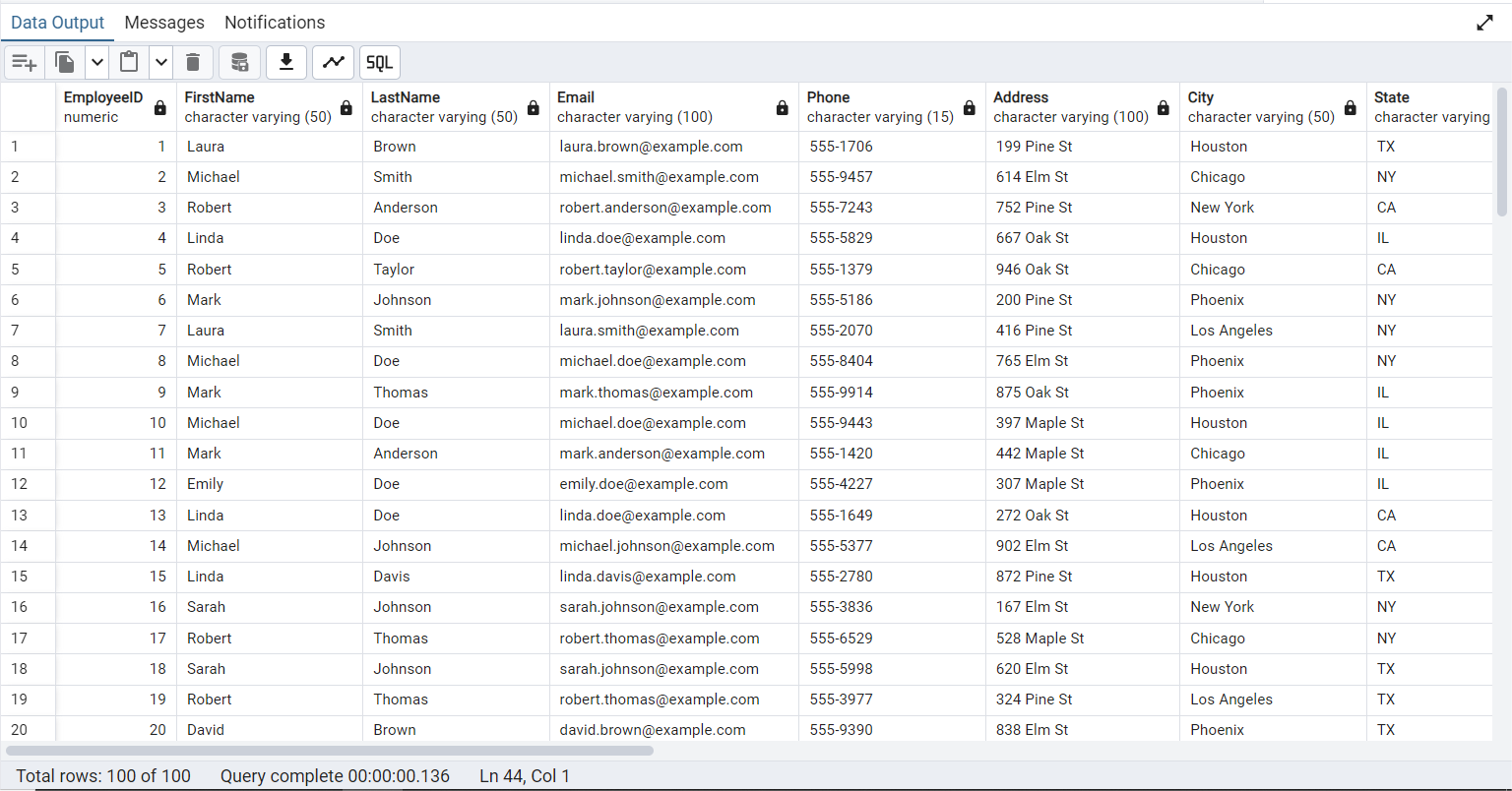
"PerformanceRating"

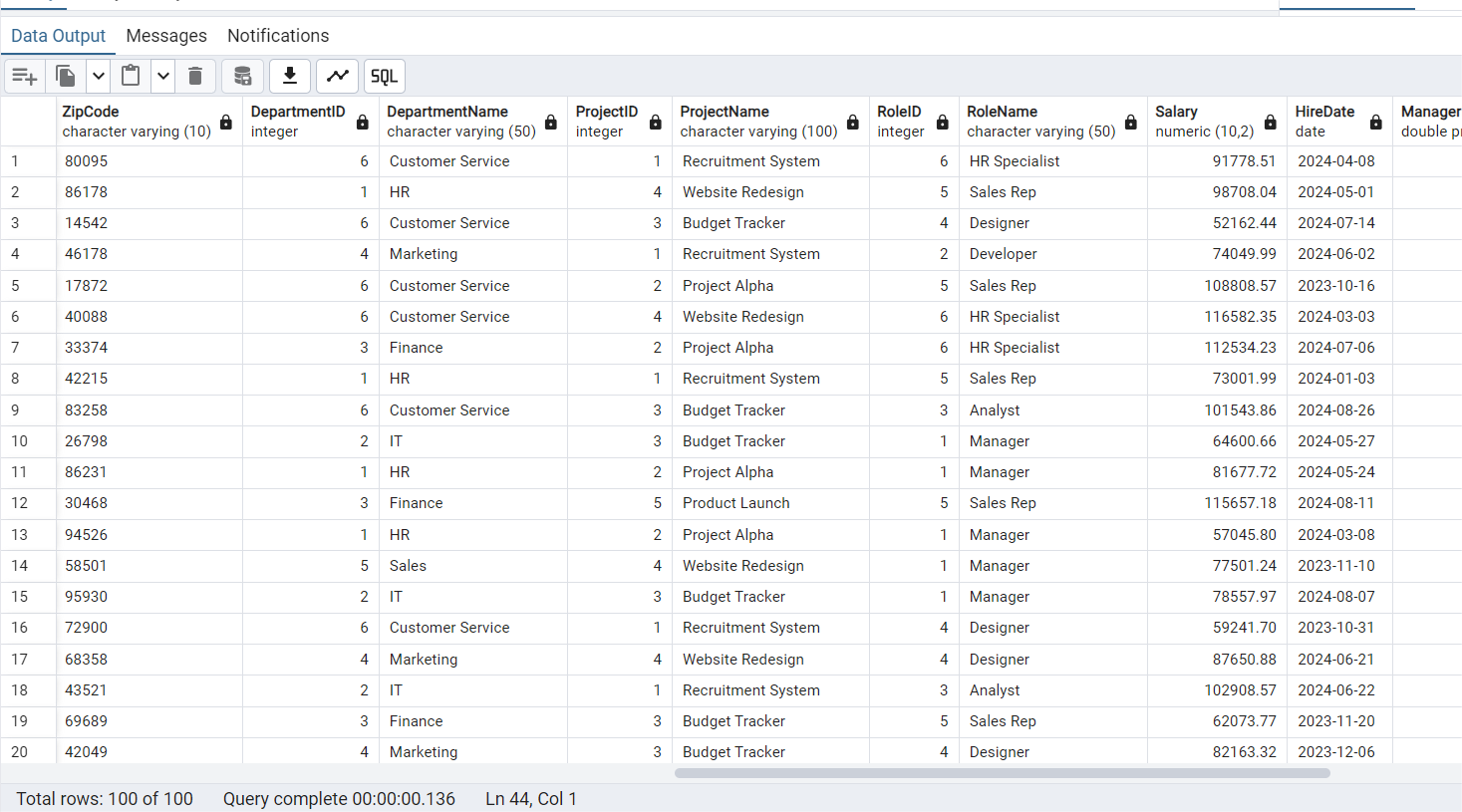
)

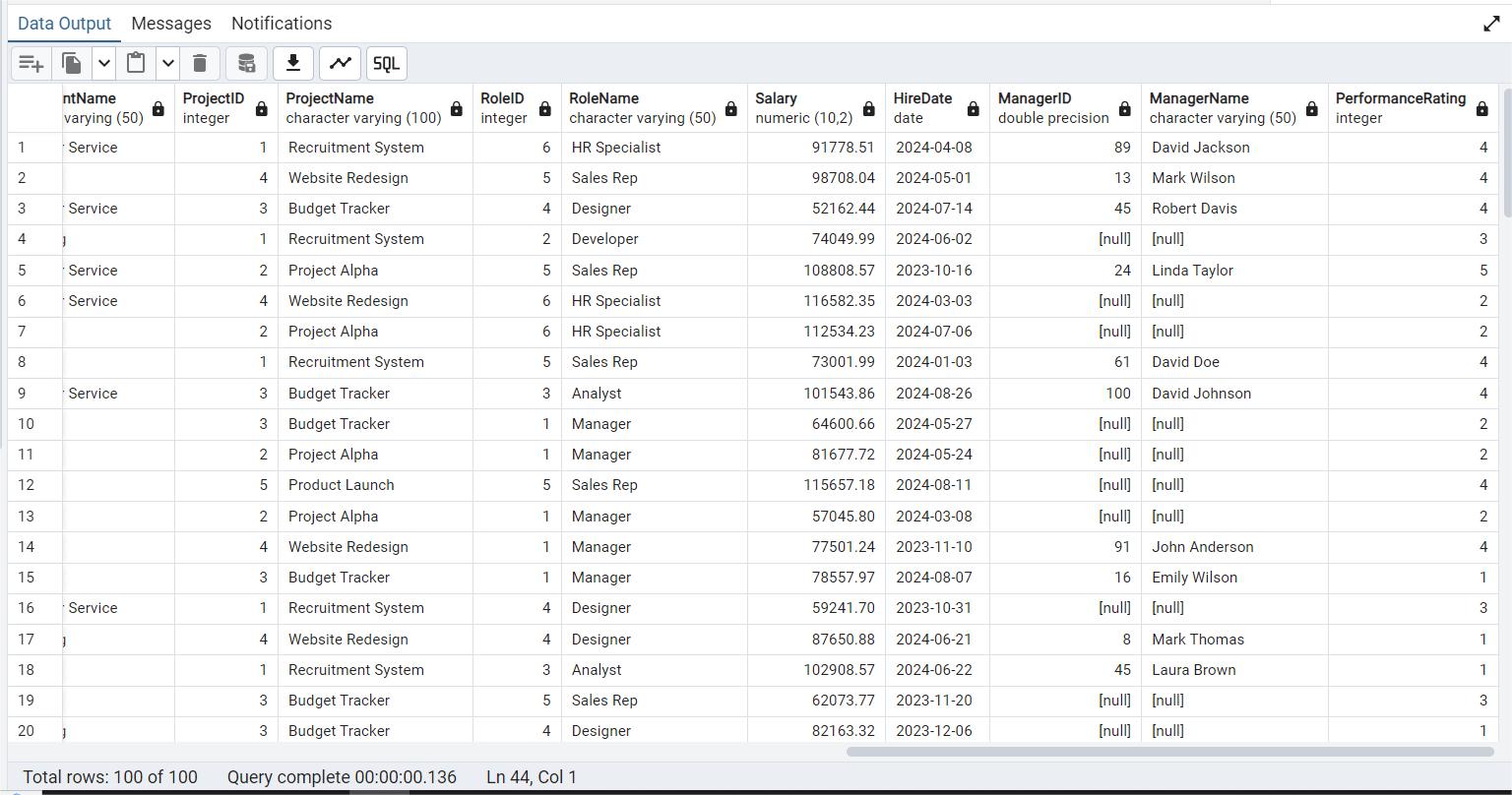
FROM 'F:\employees\_fixed.csv'

DELIMITER ','

CSV HEADER







-----------------------------------DATA NORMALIZATION TILL 3NF --------------------------------------

select \* from "Employees"

---There are no multiple values within the same column of the table.All values are atomic so we can say table saisfies the 1NF of NORMALIZATION.

--2) To meet 2NF(no partial dependency,we need to create seperate Department table)

create table "Department\_Table" as

select distinct "DepartmentID","DepartmentName" from "Employees"

where "DepartmentID" is not null

order by "DepartmentID" ---setting the order of the Deparment\_Table

select \* from "Department\_Table"

select\*from department\_table

alter table "Department\_Table" add primary key("DepartmentID")

----making DeparmentID as primary key of the table

create table "Employees\_Table" as

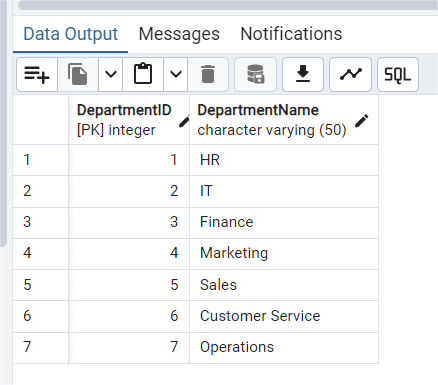
select \* from "Employees"

---removing department name (from Employees table)as now we have a seperate table for department

alter table "Employees\_Table"

drop column "DepartmentName"

select \* from "Employees\_Table"



------creating seperate project table for meet 2NF requirements----

create table "Project\_Table" as

select distinct "ProjectID","ProjectName" from "Employees\_Table"

order by "ProjectID"

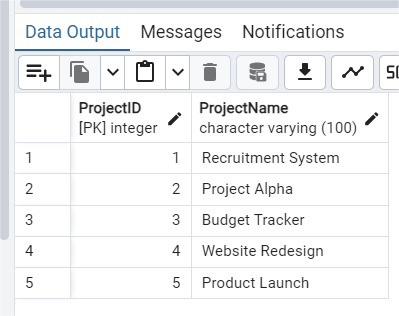
select \* from "Project\_Table"

---removing Project name (from Employees table)as now we have a seperate table Project\_Table

alter table "Employees\_Table"

drop column "ProjectName"

alter table "Project\_Table" add primary key("ProjectID") ----making ProjectID as primary key of the table



------creating seperate Role table for meet 2NF requirements----

create table "Role\_Table" as

select distinct "RoleID" ,"RoleName" from "Employees\_Table"

order by "RoleID"

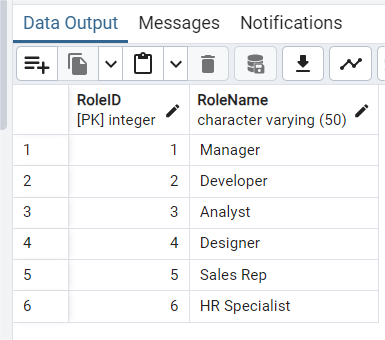
select \* from "Role\_Table"

---removing Role name (from Employees table)as now we have a seperate table Role\_Table

alter table "Employees\_Table"

drop column "RoleName"

alter table "Role\_Table" add primary key("RoleID") ----making RoleID as primary key of the table



-----creating seperate Manager table for meet 2NF requirements----

create table "Manager\_Table" as

select distinct "ManagerID" ,"ManagerName" from "Employees\_Table"

where "ManagerID" is not null

and "ManagerName" is not null

order by "ManagerID"

select \* from "Manager\_Table"

-------creating a primary key for Manager-Table using cte----

alter table "Manager\_Table"

add column "ManagerID\_PK" int;

with cte as(

select ctid,row\_number() over(order by ctid) as rn

from "Manager\_Table"

)

update "Manager\_Table"

set "ManagerID\_PK" = cte.rn

from cte

where "Manager\_Table".ctid = cte.ctid

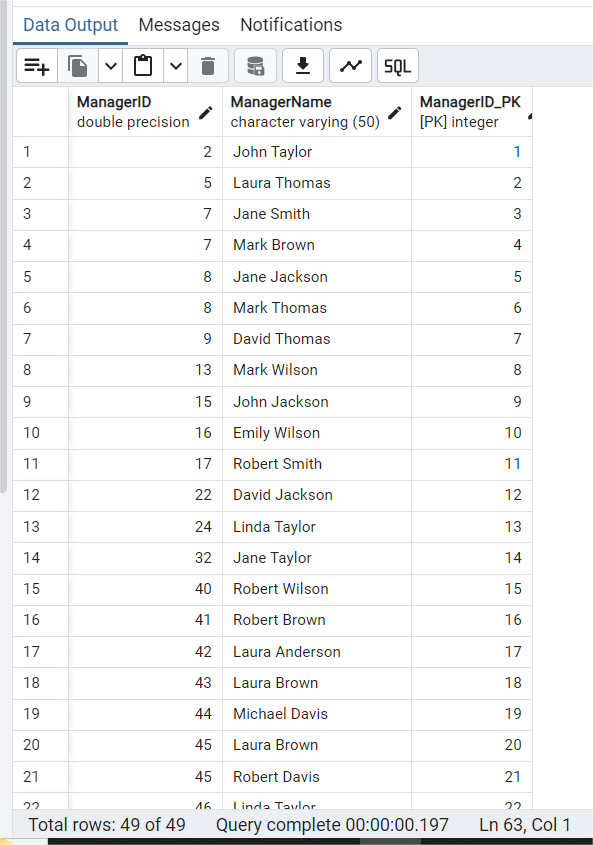
alter table "Manager\_Table" add primary key("ManagerID\_PK") ----making ManagerID\_PK as primary key of the table

alter table "Employees\_Table"

drop column "ManagerID"

alter table "Employees\_Table"

drop column "ManagerName"



----joinning ManagerID\_PK with Employee-Table-----

create table "Employees\_Table2" as

select "EmployeeID","FirstName","LastName","Email","Phone","Address","City", "State","ZipCode",

"DepartmentID","ProjectID","RoleID","Salary","PerformanceRating",m."ManagerID\_PK" from "Employees\_Table" as e

left join "Manager\_Table" as m

on e."ManagerID"= m."ManagerID"

order by "ManagerID\_PK"

alter table "Employees\_Table2" rename to "Employees\_Table"

select \* from "Employees\_Table"

-------creating a primary key for Employees\_Table using cte---------

alter table "Employees\_Table"

add column "EmployeeID-PK" int

with cte as(

select ctid,row\_number() over(order by ctid) as rn

from "Employees\_Table"

)

update "Employees\_Table"

set "EmployeeID-PK" = cte.rn

from cte

where "Employees\_Table".ctid = cte.ctid

alter table "Employees\_Table2" add primary key("EmployeeID-PK") ----making EmployeeID-PK as primary key of the the table

create table "Employees\_Table2" as

select \* from "Employees\_Table"

order by "EmployeeID-PK"

select \* from "Employees\_Table"

drop table "Employees\_Table"

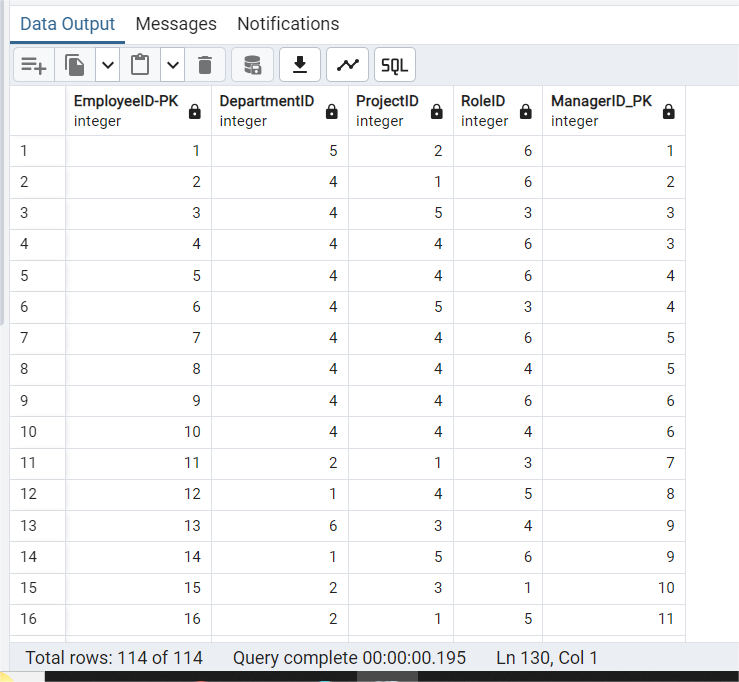
alter table "Employees\_Table2" rename to "Employees\_Table"

----------creating Relationships\_Table-------------------------------------------

create table "Relationships\_Table" as

select "EmployeeID-PK","DepartmentID","ProjectID","RoleID","ManagerID\_PK" from "Employees\_Table"

select \* from "Relationships\_Table"



-------------keeping only non\_key columns in Employees\_Table-----

alter table "Employees\_Table"

drop column "DepartmentID"

alter table "Employees\_Table"

drop column "ProjectID"

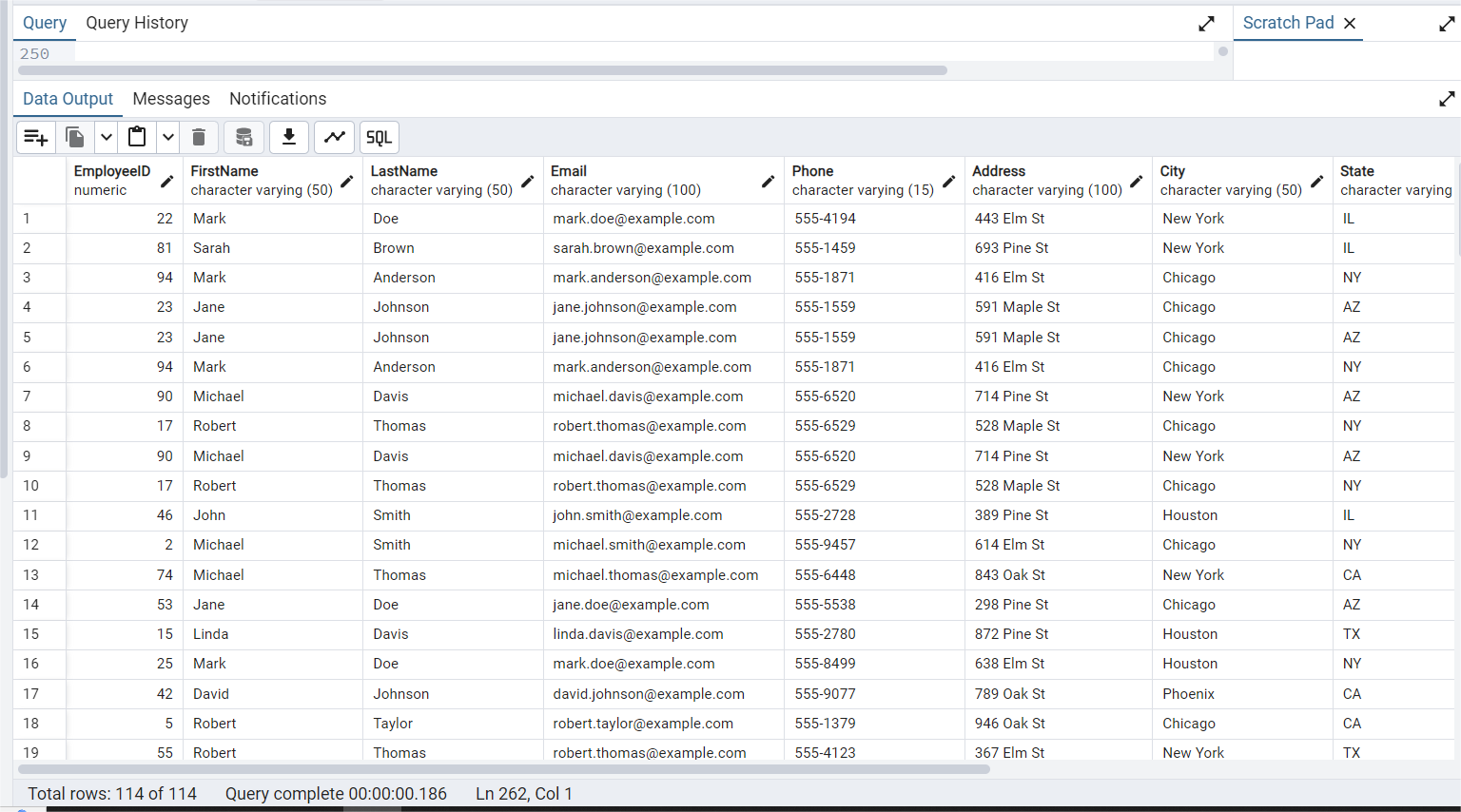
alter table "Employees\_Table"

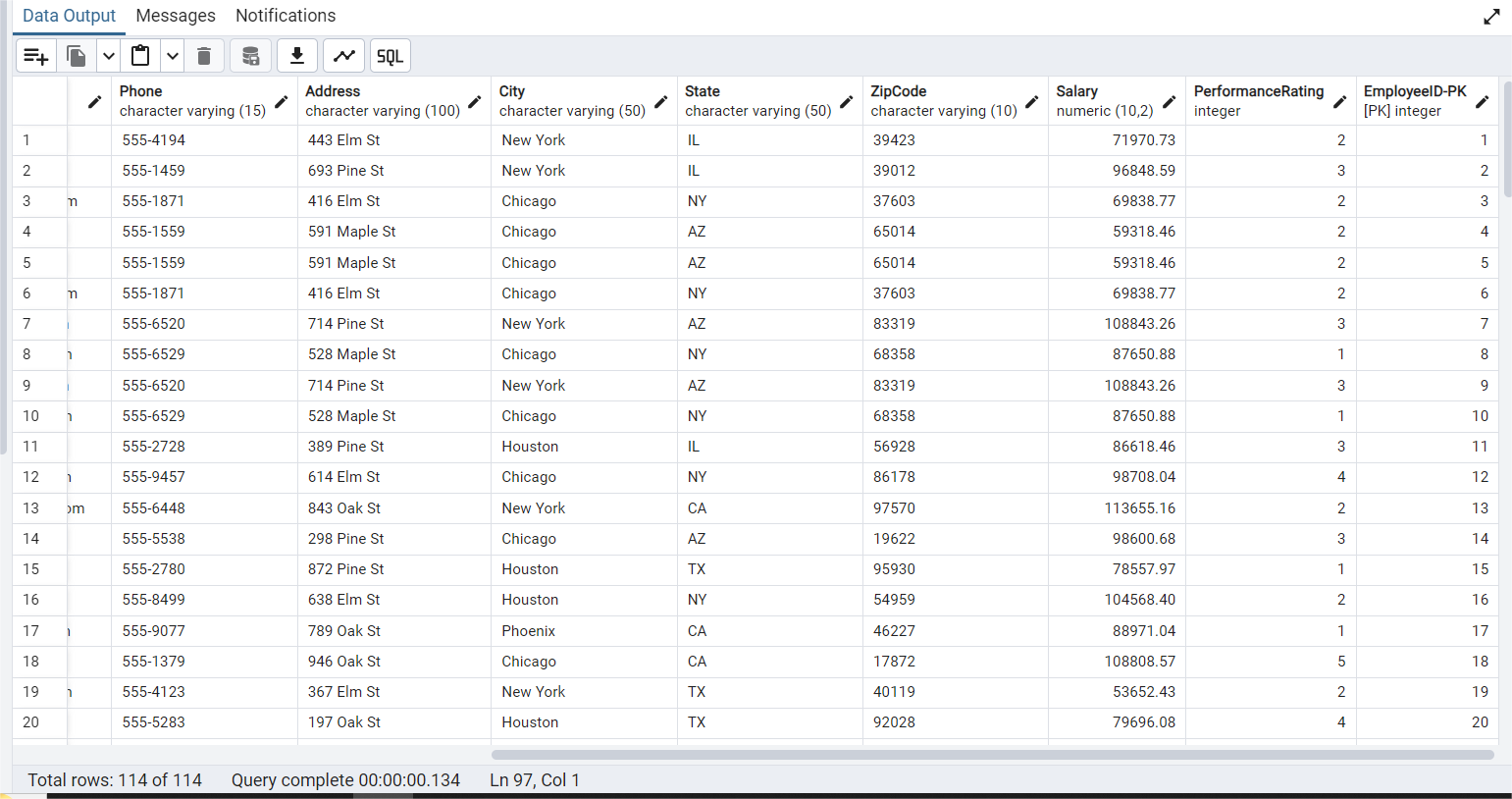
drop column "RoleID"

alter table "Employees\_Table"

drop column "ManagerID\_PK"

select \* from "Employees\_Table"





------------------------Hence the data is now normalized till 3NF-------------------------------------