## Lesson 2 Assignment: Roger André

## R-Script:

library(sqldf)

library(tidyverse)

require(data.table)

# Load the 3 CSV files in R using the fread function from data.table package.

address <- fread("Address.csv")

business\_entity\_address <- fread("BusinessEntityAddress.csv")

employee <- fread("Employee.csv")

# Assume the 3 files are database tables and write a SQL query to join all

# the tables together and filter for Job Titles that start with the word

# 'Research'

sql\_query <- ("SELECT b.BusinessEntityID, e.LoginID, e.JobTitle, a.City

FROM business\_entity\_address b

INNER JOIN employee e

ON b.BusinessEntityID = e.BusinessEntityID

INNER JOIN address a

ON b.AddressID = a.AddressID

WHERE e.JobTitle like 'Research%';")

# Use the sqldf package to run the sql from the step above.

results <- sqldf(sql\_query)

## Results:

**> results**

**BusinessEntityID LoginID JobTitle City**

**1 7 adventure-works\\dylan0 Research and Development Manager Kenmore**

**2 8 adventure-works\\diane1 Research and Development Engineer Everett**

**3 9 adventure-works\\gigi0 Research and Development Engineer Bellevue**

**4 10 adventure-works\\michael6 Research and Development Manager San Francisco**

## Question: Suggest a location that is close to most of the employees.

## Answer: Given that 3 out of 4 employees live near Seattle, it seems like a good place for the new office would be in North Seattle, to make it easy for East Side commuters.