CADS Project 2: Regression Instructions

Regression Project on Real Estate Dataset

The data for this assignment is a collection of variables for a series of real estate transactions. The most important feature of the data is the Price Column. We will determine the correlation between the Price and several other variables. Please see a full detailed description of the (columns/variables below.

Use R-Studio to compute the following statistics for the data in Date columns and Age column, accurate to 3 decimal places
intsall.packages()
library(dplyr)
library(ggplot2)

Make sure that the file RealEstateDataSet.csv is in the appropriate directory on your computer. In R-Studio, type getwd() to know exactly where to save the dataset.

(a) Compute the correlation between Price and Date.
(b) Compute the correlation between Price and Age.
(c) Compute the correlation between Price and Distance.
(d) Compute the correlation between Price and Stores.
(e) Select the title of the column with the strongest positive correlation.
(f) Select the title of the column with the strongest negative correlation.
(g) Compute the slope of the least squares regression line between Price (response variable) and Distance (explanatory variable). Put your formula in …….
(h) Compute least squares regression line between Price (response variable) and Distance (explanatory variable). Put your formula in ….
(i) Compute the value of r2, using the RSQ function, for the least squares regression line between Price (response variable) and Distance (explanatory variable). Put your formula in …….
(j) Use the slope and intercept you've calculated above to predict y values-prices per unit area for the x values-distance to nearest metro station, given as 500, 700, 725 and 1000. Put your answers in ……...
(k) Create a scatterplot of the data in Distance and Price. Be sure to put the correct variable on the horizontal axis. Include a meaningful title. Place the scatterplot below the values you've calculated.

Column (Variable Names) in Dataset:

A. Date of Transaction
B. Age of House
C. Distance from the nearest metro station.
D. Number of Convenience Stores

E. Latitude
F. Longitude
G. Home Price per square foot