MEMORANDUM

To:	Carole Voulgaris, Course Instructor, SES 5215
From:	Chunfeng Yang, Student, SES 5215
Date:	November 14, 2022
Subject:	Assignment 4, Regression with one predictor

The purpose of this memo is to present the results of a set of linear regression models that are relevant to the question:

What is the effect of the distance to an Airbnb's nearest gallery on its listing price for each night, after accounting for the effects of the distance to the nearest subway station and room type in Manhattan, New York?

I will be addressing this question using Airbnb listing data from a mission-driven project "Inside Airbnb" (http://insideairbnb.com/explore/) and public amenities data from NYC OpenData (https://data.cityofnewyork.us/Transportation/Subway-Stations/arq3-7z49). My dataset includes 21,598 Airbnb listings in Manhattan, New York.
The variables in my data set are:

1. The listing price of an Airbnb listing (outcome): the expense of one night stay in an Airbnb listing, in USD
2. Vicinity of the gallery (predictor): The distance from one Airbnb listing to its nearest gallery.
3. Vicinity of the subway Station (predictor): The distance from one Airbnb listing to its nearest subway station.
4. Room Type (predictor): A categorical variable representing the types of Airbnb listing. *Inside my categorical variable Room Type, I choose Entire Home/Apartment as my reference category.

Relationship between an Airbnb listing room type and listing price

Table 1 shows the results of the model that predicts an Airbnb listing price based on room type.The R-squared value for this model was 0.1333, suggesting that about 13 percent of the variation on listing price can be explained by differences in rome type. The coefficients for <u>private room</u> and <u>shared room</u> both had significant coefficients, indicating that when booking different room types of Airbnb had listing prices that were significantly different from the listing price for the <u>entire home or apartment</u>. The coefficients for <u>private room</u> and <u>shared room</u> are both negative, indicating that booking these two room types of airbnb would cost less money than booking an <u>entire home or apartment</u>.

Table 1: Results of regression model predicting listing price based on room type

Coefficient	Estimated value	p-value
Intercept	233.413	<2e-16
Private room	-121.004	<2e-16
Shared room	-144.436	<2e-16

Relationship between an Airbnb listing price and continuous variables (distance to the nearest gallery and distance to the nearest subway station)

Table 2 shows the results of two different regression models: one predicting Airbnb <u>listing price</u> based on the <u>distance to its nearest gallery</u>, and the other predicting <u>listing price</u> based on the <u>distance to its nearest subway station</u>.

The first model with <u>the distance to its nearest gallery</u> predicts about 3.76% of the total variation in <u>listing price</u>, with R-squared values of 0.0376; The second model with <u>the distance to its nearest subway station</u> predicts about 0.04% of the total variation in <u>listing price</u>, with R-squared values of 0.0004.

The coefficient in the first model is significant and negative, indicating that, without controlling for other factors, a shorter <u>distances to an Airbnb's nearest gallery</u> is associated with higher <u>listing price</u>; The coefficient in the second model, however, is negative but not significant, indicating that, without controlling for other factors, a shorter <u>distance to an Airbnb's nearest subway station</u> is less associated with higher <u>listing price</u> than a shorter <u>distance to its nearest gallery</u>.

Table 2: Results of regression models predicting listing price based on distance to its nearest gallery and distance to its nearest subway station

Predictor	Estimated coefficient (in regression with a single predictor)	Estimated intercept value	p-value	Model R2
distance to its nearest gallery	-0.11	219.74	<2e-16	0.0376
distance to its nearest subway station	-0.014	191.26	0.0032	0.0004