Last Name:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ First Name:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Last Name:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ First Name:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Last Name:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ First Name:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Last Name:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ First Name:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Intro to Stats Homework for Week 1

Module 5

Due Nov 25th at 17:00

This part of the assignment will cover Sessions 1-4 and be worth 100 points. Each question is worth 5 points unless otherwise marked.

You may submit your answers electronically with this sheet via MyCourses, or on paper in class to me on Friday. If you submit answers on paper, please include all associated student names, write legibly, and preserve the order of the questions as listed below.

I would like to see your work, but regardless, for any questions that require calculations, please make it clear to me what your final answer is by either making it the final line (electronic) or by circling it (handwritten).

**Session 1:**

(1) Pg 51: 1.7

(2) Pg 55: 1.23

(3) Pg 59: 1.27

(4) Pg 65: 1.53

**Session 2:**

(1) Pg77: 2.9

(2) Pg 85: 2.15

(3) Pg 86: 2.23

(4) Pg 91: 2.25

**Session 3:**

(1) Pg 166: 3.3

(2) Pg 175: 3.33 (Note that page 173, second paragraph, discusses two different forms of box plots. Feel free to you whichever makes more sense to you.)

(3) Pg 1854: 3.47

(4) Pg 188: 3.69, a, b, c, and e (you can skip part d)

**Session 4:**

(1) Pg 209: 4.13 (7 pts)

(2) Pg 216: 4.21 (7 pts)

(3) Pg 224: 4.42 (6 pts)

**Data Exercise** (20 points total, 5 points each part)

In MyCourses is a data file, SoldHomes1.xls which contains data on a sample of sold homes in Savannah, GA, from 2018-2021. Variable definitions are located on the “VarDefs” tab.

These data are taken from the database of homes that were sold through real estate agents 2018-2021 and represent 82% of transactions in that time. This represents about 2% of total homes in Savannah.

As you’ll see, the data aren’t perfect and contain some missing or erroneous values.

A home insurance firm is interested in moving into the Savannah market and they hire you analyze data on the homes in the area to get a sense as to what the housing stock looks like.

i) First, summarize the typical home value by using the homes’ selling prices. Report mean, median, and standard deviation. Also, consider that some houses in the data aren’t “real” transactions in that they are transfers from one person to another as gifts, such as when a parent passes a home onto his/her children. Those transactions are usually recorded with a sales price of $0 or something really low ($1, $100, or something like that). Should those transactions be included in your analysis?

ii) Create a histogram of home prices.

iii) Your client is interested in the sizes of the houses in Savannah. Create a graph that summarizes this information. Check to see if there are any observations that should be excluded from the analysis.

iv) Your client is interested in the rate of home price appreciation in this region over the last four years. Create a chart or graph that illustrates the rate of home price change over this time. Again, review the data to see if there are any observations that should be excluded.