

# Lab07: Regression Model Building


---

**Handout date:** Wednesday, October 16, 2019

**Due date:** Friday, October 25, 2019 by noon as hardcopy

*This lab counts 4 % toward your total grade:*

**Objectives:** You engage an empirical dataset to build an **informed** regression model to explain the variation of a dependent variable.

**Format of answer:** Your answers (graphs and verbal description) should be handed in as **hard-copy** in **one** document. Add a running title into the header of the document with the following information: **your name, Lab07 and page numbers**. Format any code and computer output properly before inserting it into the document with your answer. -code and text output need to be in a **monospaced** font (i.e., fixed-pitch font) such as Courier New so proper spacing and alignments are preserved. Excessive, but irrelevant, output will lead to a deduction of your accumulated points.

This lab is based on the **tractShp** spatial data frame in the package **DallasTracts**. You can use the script **Lab07StarterCode.r** to set up the data.

## Task 1: Model Specification (1point)

Select a set of no more than 6 independent variables which explain the variation in the percentage of the civilian population in a census tract who does not have health insurance. I.e., the dependent variable is **PCTNOHINS**.

Note: Independent variables should be on relative scale rather than absolute values because the dependent variable is on a relative scale. (see the lecture notes **Chapt04Relationships.pdf** on page 6)


Justify your choice of independent variables and formulate explicit hypotheses how and in which direction these independent variables are expected to influence the dependent variables.

Explore the univariate and pairwise relationships of the variables. For variables with a strong positive skewness and all observations larger than zero consider a log-transformation.

## Task 2: Model Building (2 points)

Build a proper regression model and perform model diagnostics. Based on the information gained by the model diagnostics revise, if necessary, your model specification. Interpret and document each relevant step of your analysis (skip broad trial analysis).

## Task 3: Analysis Documentation (1 point)

Write a **professional report** with an objective statement in an introduction and a conclusions section based on tasks 1 and 2 that also includes relevant analysis results, figures and maps as well as their interpretation. Include important intermediate analysis steps into your report. In an appendix add the -code that you used for your report.