

Python for Regression Analysis

February 19, 2021

Brooke Luetgert, PhD



THE UNIVERSITY OF
CHICAGO

Office of Research and
National Laboratories
Research Computing Center

Contact Information

Brooke Luetgert, PhD. Senior Computational Scientist at Research and Computing Center (RCC)

Email: luetgert@uchicago.edu

Office: TAAC 2, 5607 South Drexel

Telephone: (773)-834-5313 (currently out of office)

RCC Help Desk: Reg. 216, Mon-Fri 9 AM-5 PM

Materials on GitHub- use search bar, enter **user:luetgert** Our folder is **luetgert/IntroRegression**

Plan for Today

We will use Python within a Jupyter Notebook and several libraries for regression analysis and diagnostic plots.

1. Open a Jupyter Notebook, install patsy and statsmodels
2. Import data from .txt
3. Learn about our data, clean and label
4. Begin regression analysis, explore diagnostics
5. Consider backwards feature selection for multiple linear regression

Survey Request

We use Survey Monkey to record your feedback and improve our workshops. Please let us know what you would like to see in the future.

<https://www.surveymonkey.com/r/QK3MM9M>

Conceptual Review of OLS

