Jackson Rudoff

Modern Data Structures

November 25th, 2022

**Final Project Proposal**

Political science – as with other social sciences – have begun migrating to Python when conducting research. This tracks with the parallel developments in packages like Pandas, which has made the language more hospital to data analysis and hypothesis-based research. However, many datasets which have been adapted to R-specific data-forms or packages have not seen a similar conversion to Python. Datasets like the GSS, for example, are accessible as easily ingestible .rds files, which can be quickly and efficiently converted to tibbles with correctly formatted variables. My proposed final project would be to make an oft-used dataset – the American National Election Studies (ANES) – into a package format, which would provide users with helper functions that will streamline analysis of the data. It would include graphical and statistical testing methods that will make public opinion research more user-friendly, particularly for those just getting familiar with using Python in social science research.

The main advantage of the package is that it bypasses the step of importing the data, immediately providing the user with the ANES samples in the form of class objects. Though the package by default will provide the last 4 editions of the time series. However, users could also instantiate new ANES objects by providing new Excel sheets taken from the ANES website.