Data and replication package README file for "Identifying Partisan Gerrymandering and its Consequences: Evidence from the 1990 US Census Redistricting" (Navid Sabet and Noam Yuchtman)

There are five (5) datasets used in the analysis which are posted for replication. This file briefly describes them. Fuller descriptions of the variables, their sources and construction can be found under Section 3 (Data) of the paper.

1. irca_gerrymander.dta
This is the main dataset of the analysis. It is a County × Congress dataset that is used to generate:
- Figures 2, 3, 4, 5, 6, 8 and 9
- Tables 1, C.1, C.2, C.3 and D.1.

2. state_irca_congress.dta
This is a State × Congress dataset that is used to generate Table 3.

3. state_irca_districts.dta
This is a State × Congressional District × Congress that is used to generate Figure 7 and Table 2.

4. anes.dta
This is data directly obtained from the ANES and is used to generate Table B.1 in the Appendix.

5. hispanic_members_of_congress_list.dta
This is a dataset of Hispanic members in congress as well as the representatives they replaced. It is used to generate Table E.1 in the Appendix.

There are 16 .do files and one R-script, each of which generates a single exhibit (table or figure) included in the paper as well as the online appendix. Each .do file or R-script is titled according to the figure or table that it produces (and each file is self-contained).