UKBB 67K Dataset:

DropBox: [https://www.dropbox.com/scl/fi/5q3y02g8p2img0g8au47e/ukbb\_677197\_Jan\_2024.zip?rlkey=7204q6avtrqi8abrpslnbo9q0&dl=0](https://nam02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fwww.dropbox.com%2Fscl%2Ffi%2F5q3y02g8p2img0g8au47e%2Fukbb_677197_Jan_2024.zip%3Frlkey%3D7204q6avtrqi8abrpslnbo9q0%26dl%3D0&data=05%7C02%7Cdinov%40med.umich.edu%7Cb4515d2bb8024cc7ca1808dc1e7772f8%7C1f41d613d3a14ead918d2a25b10de330%7C0%7C0%7C638418745887739847%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C3000%7C%7C%7C&sdata=wVS6sg8i3Kf3gj3rYJHjPB4EVr6hM43f%2BAE4k3bw%2Bxo%3D&reserved=0)

The tables cover all **~500k UKB participants**. There are **~25,000 columns (including repeat data points)** **and ~46,000 subjects with MRI derived phenotypes (fxxxx\_2\_0**: **first imaging visit N~46000; fxxxx\_3\_0**: **second imaging visit N~4800**). **67k is number of subjects with at least one Nifty T1 MRI**. Obviously, note all of them were included in the imaging-derived phenotypes, possibly due to bad quality or failed process.

In the zip package, I also included **ukb677197\_fields.csv** and **ukb677197.html** which you can use to *explore the available data fields*.