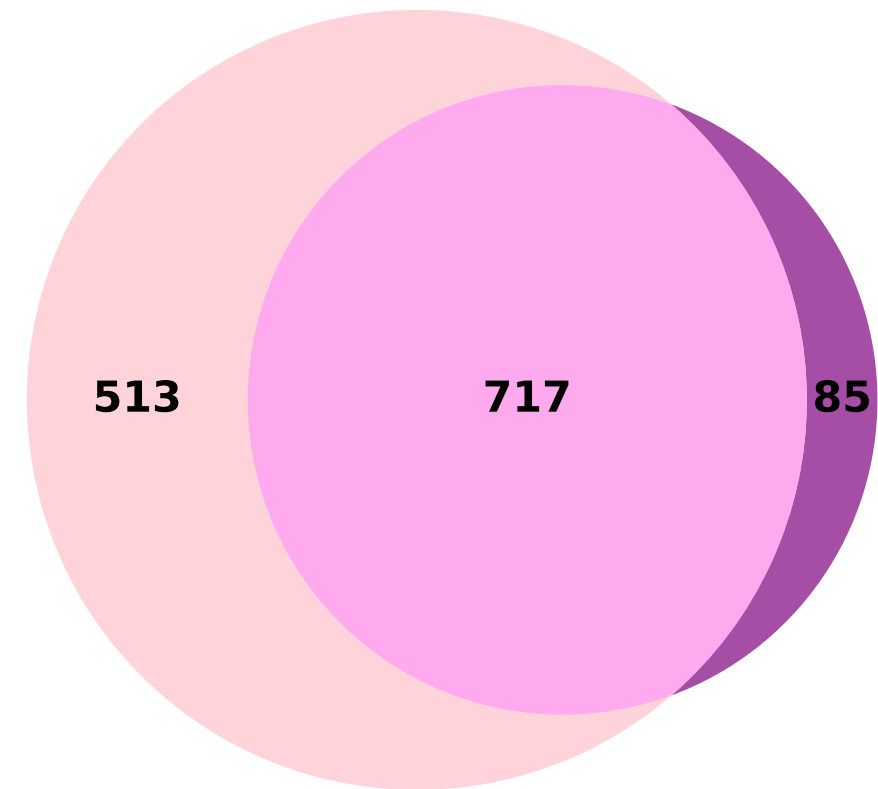


Human
(1230/1515, 81.2%)
o3-mini
(916/1515, 60.5%)



Human
(1230/1515, 81.2%)
claude3 sonnet
(802/1515, 52.9%)

