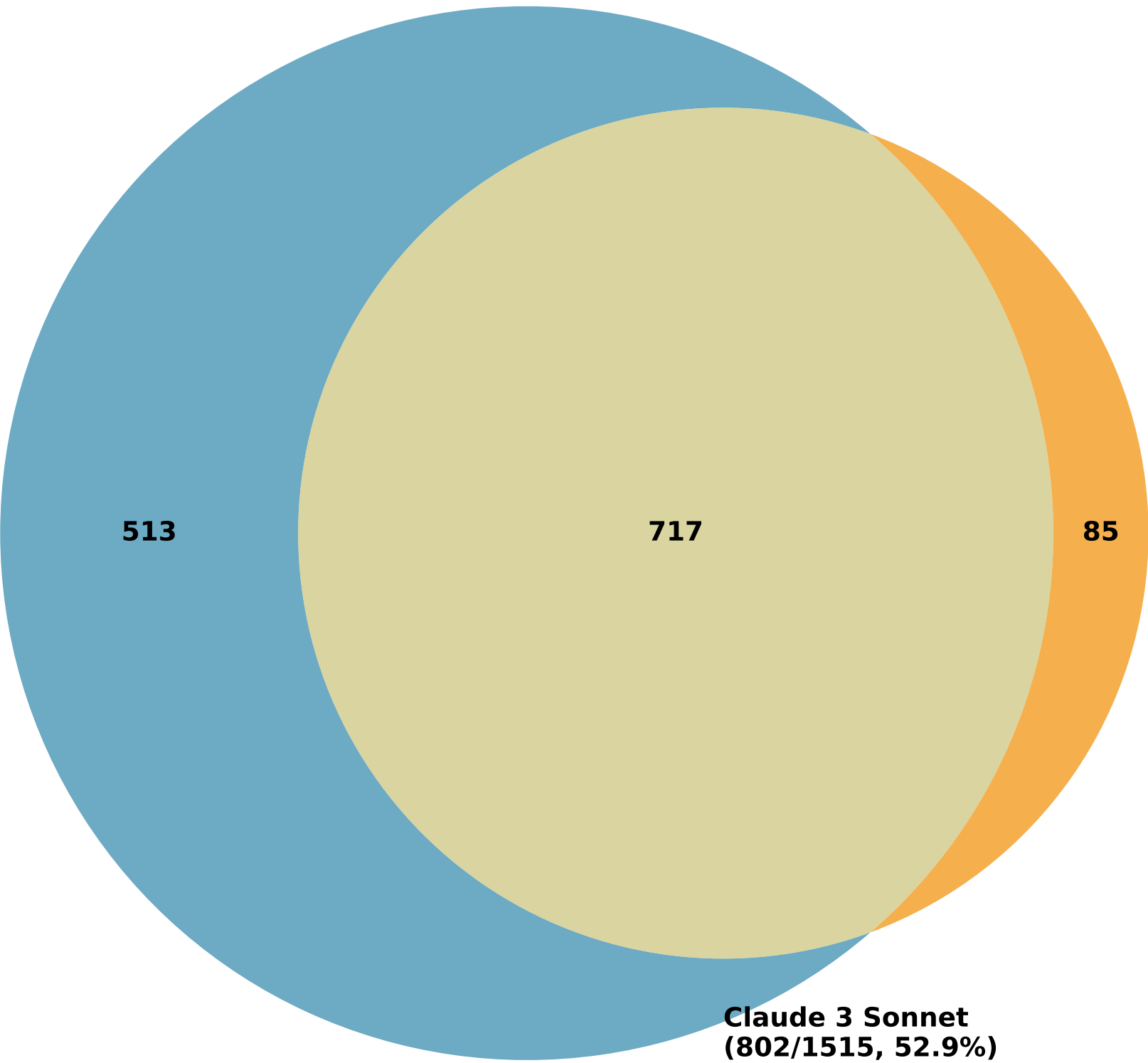


**Human vs Claude 3 Sonnet: Compiled Successes Among Baseline Files**



**Human**  
**(1230/1515, 81.2%)**

**Claude 3 Sonnet**  
**(802/1515, 52.9%)**