# Distribution of Ratings by Model

Absolute counts and percentages of Correct, Inaccurate, and Wrong ratings for each model.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Model Category** | **Correct** | **Inaccurate** | **Wrong** | **Total** | **Correct (%)** | **Inaccurate (%)** | **Wrong (%)** |
| Sonar-Reasoning-Pro + WL | 463 | 39 | 18 | 520 | 89.0 | 7.5 | 3.5 |
| Sonar-Pro + WL | 456 | 45 | 19 | 520 | 87.7 | 8.7 | 3.7 |
| OpenEvidence (Assumed WL) | 429 | 60 | 31 | 520 | 82.5 | 11.5 | 6.0 |
| Sonar-Reasoning-Pro | 422 | 80 | 18 | 520 | 81.2 | 15.4 | 3.5 |
| Sonar-Pro | 415 | 85 | 20 | 520 | 79.8 | 16.3 | 3.8 |
| Sonar + WL | 406 | 79 | 35 | 520 | 78.1 | 15.2 | 6.7 |
| Sonar | 314 | 145 | 61 | 520 | 60.4 | 27.9 | 11.7 |