Study on correlation coefficients between:

1. distances between killmails (as measured by **cosine distance** between bag-of-words model of items on a ship at the time of death), and
2. change in kill/death ratio between killmails.

I plotted their distributions and have a 6-statistic summary for long and short text representations, where **count** is the number of players involved:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Count | Min | Max | Median | Mean | Std\_dev |
| Short Text | 2372 | -0.64839 | 0.73311 | -0.032302 | -0.031026 | 0.13469 |
| Long Text | 2372 | -0.57989 | 0.68042 | -0.033684 | -0.032597 | 0.13403 |

|  |
| --- |
|  |
|  |