

## ساختار کلی داده‌ها:

- حدود ۹۰٪ داده‌ها نرمال و در بازه تلرانس مجاز باشند.
- ۱۰٪ داده‌ها خطادار و به‌صورت ترکیبی از الگوهای مختلف در نظر گرفته شوند.

| نوع خطا                                             | سهم از کل خطاها | توضیح شبیه‌سازی                                                                                                 |
|-----------------------------------------------------|-----------------|-----------------------------------------------------------------------------------------------------------------|
| خطاهای خوشه‌ای (Clustered Errors)                   | ۳۰٪ از خطاها    | خطاها به‌صورت متمرکز در چند بازه‌ی پیاپی ظاهر می‌شوند. مثال: ۵۰ نمونه پشت‌سر هم در یک ایستگاه دارای خطا باشند.  |
| رانش زمانی (Drift over Time)                        | ۲۰٪ از خطاها    | خطاها به‌مرور زمان بدتر می‌شوند، مثلاً افزایش تدریجی مقدار به دلیل فرسودگی ابزار.                               |
| تصادفی پراکنده (Random Errors)                      | ۲۰٪ از خطاها    | خطاهایی که به‌صورت تصادفی و پراکنده در کل دیتاست ظاهر می‌شوند.                                                  |
| خطاهای موضعی/ایستگاه‌محور (Station-Specific Errors) | ۱۵٪ از خطاها    | خطاهایی که فقط در ایستگاه خاصی (مثلاً فقط در ws۲ یا ws۳ دیده می‌شوند، مثلاً ناشی از نقص تنظیمات یک دستگاه خاص). |
| بایاس سیستمی (Systematic Bias)                      | ۱۵٪ از خطاها    | تمام مقادیر با یک شیفت مشخص از مقدار واقعی فاصله دارند (مثلاً همیشه ۰.۰۱ بیشتر)، ناشی از کالیبراسیون اشتباه.    |