D-FLARE ■■■■■■ O(1)/■■■■■■■■■■■■

1. log_cleaning.py — ■■■■■■■■■■■■

- ■■■■■■■Raw Log Files■■■■■■■■■■■■■■■
- ■■■■■■■■■■■■■■■■■■raw_log ■■■
- ■■■■■■■■■■■■
- ■■■■■ CSV■processed_logs.csv■■■■■■■■■■log_unique_values.json■■
- ■■■■■■■■■■■■■■ OOM■CLI ■■■■■■■■■■■■■■■■■■■

2. log_mapping.py — ■■■■■■■■■■■

- ■■■■■■■■■■■■■Mapping to Integer IDs■■
- ■■ idseq ■■■■■■■■■■■■■■■■
- ■■ raw_log ■■■■■■■■■■
- ■■■■■■■■■■■ log_unique_values.json■■
- ■■■■■■■■■■■■■■■■■■■■

3. feature_engineering.py — ■■■■■■

- ■■■■■ 5 ■■■■■■
- 1) ■■■■■■
- 2) ■■■■■■■
- 3) ■■■■■■
- 4) ■■■■■
- 5) ■■■■■■
- ■■ TB ■■■■■ chunk ■■■■■■■■■■■■■■■ OOM ■■■

O(1)/■■■■■■■■

■■
O(W)■■■■■■■■■■■■■
■■■■■■■■■■■■■Sliding Counters with Minute Buckets■■
- ■■ deque ■■■■■■■■■■ (minute, Counter_src, Counter_dst, Counter_pair)■
- ■■■■■■ Counter ■■■■■■■■■■■srcip■dstip■(srcip,dstip)■■
- ■■■■■
- 1) ■■■■■■■■■■■■■■■■ Counter ■■■■■■■
- 2) ■■■■ Counter■■■■■■■ key ■■■■■■■O(1)■■
- 3) ■■■■■■■■■■■■■■■ Counter■
- ■■■■ O(N×W) ■■ O(N)■■■■■■■■■■■■■■■■■■■■■■■■■■■■

■■■■

```
■■■■deque window_buckets, Counter agg_src, agg_dst, agg_pair for ■■■■ (ts, src,
dst): minute_key = floor_to_minute(ts) while window_buckets ■■ ■ ■■ minute <
current_minute - WINDOW_SIZE: ■■■■■■■ agg_src/agg_dst/agg_pair ■■■■ count_src =
agg_src[src] count_dst = agg_dst[dst] count_pair = agg_pair[(src,dst)] ■■■■■■■■
agg_src/agg_dst/agg_pair
```