

bro

```
#!/usr/bin/env bash
```

```
# Usage: ./group_by_date_hour_symbol_source.sh path_to_logfile
```

```
LOGFILE="$1"
```

```
awk '
```

```
{
    # Split the first field (20250209-13:45:02.123456) on the dash.
    # dt[1] will hold "20250209", dt[2] will hold "13:45:02.123456".
    split($1, dt, "-")
    date_str = dt[1] # e.g. "20250209"

    # Now split dt[2] ("13:45:02.123456") on the colon to get the hour.
    split(dt[2], time_parts, ":")
    hour = time_parts[1] # e.g. "13"

    symbol = $2 # e.g. "USDCAD"
    source = $3 # e.g. "MyLiquidityProvider"

    # Increment a count in an associative array, keyed by date, hour, symbol, source.
    counts[date_str, hour, symbol, source]++
}
END {
    # Print header (optional)
    # print "DATE", "HOUR", "SYMBOL", "SOURCE", "COUNT"

    # Loop over all keys in counts.
    for (key in counts) {
        # The key format is "date_str SUBSEP hour SUBSEP symbol SUBSEP source".
        split(key, parts, SUBSEP)
        date_str = parts[1]
        hour     = parts[2]
        symbol   = parts[3]
        source   = parts[4]
        count    = counts[key]

        print date_str, hour, symbol, source, count
    }
}
' "$LOGFILE" | sort
```

Revision #3

Created 10 February 2025 16:04:52 by Guest

Updated 10 February 2025 17:27:35 by Guest