Programming Challenge

Feel free to use any languages or tools. Please create a solution based upon what you know best. We want to see you do your best demonstration of your talent and ability. We are not testing whether you can figure out the latest trendy framework/language/technique. The solution should reflect code you would release to a production environment and maintain for multiple years.

You've been given a series of log files from a malfunctioning cloud server, and have been tasked to run some diagnostics and post processing against those files. The logs are a fictitious representation of our cloud disposition server based upon client connector lookups. The problem represents typical debugging problems we are presented with. All data (especially file paths :wink:) are fictitious.

The data is structured as JSON with a line of the file representing a single entry. Below is an example line:

```
{
    "ts": "1455575401",
    "pt": "100",
    "si": "1fadedbc-e6a5-4f49-84cf-d077717ea500",
    "uu": "52fde82d-e4c9-4018-9de9-6d92ea447a66",
    "bg": "27f0c467-a74e-4f6c-a432-50f44bb58fe2",
    "sha":
"2BAF1F40105D9501FE319A8EC463FDF4325A2A5DF445ADF3F572F626253678C9",
    "nm": "program.exe",
    "ph": "/Users/tracey/Desktop/program.exe",
    "dp": "2"
}
```

The translation key for the fields is as follows:

```
ts:  timestamp
pt:  processing time
si:  session ID
uu:  connector GUID
bg:  business GUID
sha: sha 256 of the file
nm:  file name
ph:  path
dp:  disposition (valid values: MALICIOUS (1), CLEAN (2), UNKNOWN
(3))
```

Tasks to Complete:

1. Read & parse all the log data provided

2. For each entry marked with an UNKNOWN disposition, perform a lookup in the included sqlite database to determine:

- Whether the file has been seen before
    - if so increment the database count (`cnt`) by 1
- Whether the file is new
    - if so insert a new entry with `count == 1` and disposition (`dp`) == UNKNOWN

3. We know the cloud server is malfunctioning, what types of patterns or anomalies did you observe?

- If possible, what types of problem hypothesis can you form from your observations?