

```

1 package com.marcolussetti.opendotamatchescondenser;
2
3 import com.jsoniter.JsonIterator;
4 import com.jsoniter.any.Any;
5 import com.jsoniter.output.JsonStream;
6 import gnu.trove.map.hash.THashMap;
7 import gnu.trove.set.hash.THashSet;
8 import org.simpleflatmapper.csv.CsvParser;
9 import picocli.CommandLine;
10 import picocli.CommandLine.Option;
11 import picocli.CommandLine.Command;
12 import picocli.CommandLine.Parameters;
13
14 import java.io.*;
15 import java.time.*;
16 import java.time.format.DateTimeFormatter;
17 import java.util.*;
18 import java.util.concurrent.Callable;
19 import java.util.concurrent.TimeUnit;
20
21 @Command(description = "Process OpenDota Matches File",
22          name = "processopendota",
23          mixinStandardHelpOptions = true,
24          version = "processopendota 0.3")
25 class ProcessOpenDota implements Callable<Void> {
26     // ARGUMENTS
27     @Option(names = {"-x", "--extract-to-json"},
28            description = "Extract an existing .ser file to
29 a JSON file.")
29     private File extractToJson = null;
30
31     @Option(names = {"-c", "--condense"},
32            description = "Condense the input openDota CSV
33 file. If file is GunZipped (.gz), extract it first.")
33     private File condense = null;
34
35     @Option(names = {"-o", "--only-count"},
36            description = "Only count picks, do not record
37 wins & losses")
37     private Boolean onlyCount = false;
38
39     @Parameters(paramLabel = "OUTPUT",

```

```

40         description = "Output file for either extract or
condense")
41     private File output = null;
42
43     // CONSTANTS
44     public static final int MATCHES_NO = 1191768403;
45     public static final int DAYS_NO = 1870;
46     public static final int REPORT_THRESHOLD = 1000000; //
Report progress every million rows
47     public static final int SERIALIZE_THRESHOLD = 10000000;
// Serialize every 10 million rows
48
49     // VARIABLES
50     // Keep track of progress
51     private LocalDateTime startOfParsing;
52     private THashSet<Long> allDates = new THashSet<>();
53     private int recordCounter = 0;
54     // Store the data {date: Long, {hero#: int -> picks# int
}}
55     private THashMap<Long, THashMap<Integer, Integer[]>>
data = new THashMap<>();
56
57     private void condenseInputFile(File input, File output,
boolean onlyCount) {
58         this.startOfParsing = LocalDateTime.now();
59
60         // Main loop!
61         FileReader fileReader;
62         try {
63             fileReader = new FileReader(input);
64             Iterator<String[]> csvReader = CsvParser.
iterator(fileReader);
65             String[] headers = csvReader.next();
66             // Iterate through stuff
67             while (csvReader.hasNext()) {
68                 String[] row = csvReader.next();
69
70                 parseRow(row, onlyCount);
71
72                 if (recordCounter % REPORT_THRESHOLD == 0) {
73                     reportProgress(this.recordCounter, this.
allDates.size(), this.startOfParsing);

```

```

74
75         if (recordCounter % SERIALIZE_THRESHOLD
76             == 0) {
77             String destFolder = output.
78             getParent();
79             String[] destFile = output.getName(
80             ).split("\\.");
81             File outputFile = new File(
82             destFolder + File.separator + destFile[0] + "_" + (
83             recordCounter / SERIALIZE_THRESHOLD) + "." + destFile[1]);
84
85             serializeData(data, outputFile);
86         }
87     }
88
89     reportProgress(this.recordCounter, this.
90     allDates.size(), this.startOfParsing);
91     serializeData(data, output);
92     } catch (IOException e) {
93         e.printStackTrace();
94     }
95
96     private void extractToJson(File input, File output) {
97         THashMap<Long, THashMap<Integer, Integer[]>>
98         hashMap = deserializeData(input);
99
100         writeJSON(hashMap, output);
101     }
102
103     private void parseRow(String[] row, boolean onlyCount)
104     {
105         // Extract relevant fields
106         long startTime = Long.parseLong(row[3]);
107         String pgroup = row[26];
108         boolean radiantWin = row[2].equals("t");
109
110         // Parse date
111         Long date = extractDate(startTime).toEpochDay();

```

```

108         // Parse picks
109         ArrayList<Integer[]> heroesPicked =
            extractHeroesPicked(pgroup, radiantWin);
110
111         // Update copy of local map
112         THashMap<Integer, Integer[]> todayPicks = this.data
            .getOrDefault(date, new THashMap<Integer, Integer[]>());
113         heroesPicked.forEach(heroRecord -> {
114             int hero = heroRecord[0];
115             boolean won = heroRecord[1] == 1;
116
117             Integer[] counts = todayPicks.getOrDefault(hero
118 , new Integer[]{0, 0});
119             if (onlyCount || won)
120                 counts[0] += 1;
121             else
122                 counts[1] += 1;
123             todayPicks.put(hero, counts);
124         });
125
126         // Push to global map
127         this.data.put(date, todayPicks);
128
129         // Tracking progress
130         allDates.add(date);
131         recordCounter++;
132     }
133
134     private static LocalDate extractDate(long
epochTimeInSeconds) {
135         return LocalDateTime.ofInstant(
136             Instant.ofEpochSecond(epochTimeInSeconds),
137             ZoneId.of("UTC")
138         ).toLocalDate();
139     }
140
141     private static ArrayList<Integer[]> extractHeroesPicked
(String jsonInput, boolean radiantWon) {
142         ArrayList<Integer[]> heroes = new ArrayList<>();
143
144         JsonIterator iterator = JsonIterator.parse(
jsonInput);

```

```

144         Map<String, Any> jsonObject = null;
145         try {
146             jsonObject = iterator.read(Any.class).asMap();
147         } catch (IOException e) {
148             e.printStackTrace();
149         }
150
151         jsonObject.forEach((index, object) -> {
152             int heroId = object.get("hero_id").toInt();
153             boolean isRadiant = object.get("player_slot").
154             toInt() <= 127;
155             Integer[] heroRecord = {heroId, ((isRadiant &&
156             radiantWon) || (!isRadiant && !radiantWon)) ? 1 : 0 };
157             heroes.add(heroRecord);
158         });
159
160         return heroes;
161     }
162
163     private static void reportProgress(int recordCounter,
164     int days, LocalDateTime startOfParsing) {
165         Duration elapsed = Duration.between(startOfParsing,
166         LocalDateTime.now());
167         long elapsedMillis = elapsed.toMillis();
168         DateTimeFormatter dtf = DateTimeFormatter.ofPattern
169         ("yyyy/MM/dd HH:mm:ss");
170         double rowsPerSec = (double) recordCounter /
171         elapsedMillis * 1000;
172
173         System.out.printf(
174             "\n%s (%s elapsed - %s remaining) | %9.2f
175             rows/s | %,6.2f million rows (%6.2f%) | %4d days (%6.2f%)"
176             ,
177             dtf.format(LocalDateTime.now()),
178             // current
179             time
180             formatTimeDifference(elapsedMillis),
181             // elapsed time
182             formatTimeDifference((long) ((MATCHES_NO -
183             recordCounter) / rowsPerSec * 1000)), // remaining time (est
184             .)

```

```

173         (double) recordCounter / elapsedMillis *
174         1000, // rows per second
175         (double) recordCounter / 1000000,
176         // rows
177         processed (mils)
178         (double) recordCounter / MATCHES_NO * 100,
179         // % of rows
180         processed
181         days,
182         // days tracked
183         days / (float) DAYS_NO * 100
184         // % of
185         days tracked
186     );
187 }
188
189 private static String formatTimeDifference(long millis)
190 {
191     // From https://stackoverflow.com/a/44142896/
192     6238740
193     return String.format("%02d:%02d:%02d",
194         TimeUnit.MILLISECONDS.toHours(millis),
195         TimeUnit.MILLISECONDS.toMinutes(millis) -
196         TimeUnit.HOURS.toMinutes(TimeUnit.
197             MILLISECONDS.toHours(millis)),
198         TimeUnit.MILLISECONDS.toSeconds(millis) -
199         TimeUnit.MINUTES.toSeconds(TimeUnit.
200             MILLISECONDS.toMinutes(millis)));
201 }
202
203 private static void serializeData(THashMap<Long,
204     THashMap<Integer, Integer[]>> data, File output) {
205     // From https://beginnersbook.com/2013/12/how-to-
206     serialize-hashmap-in-java/
207     FileOutputStream fos = null;
208     try {
209         fos = new FileOutputStream(output);
210         ObjectOutputStream oos = new ObjectOutputStream
211             (fos);
212         oos.writeObject(data);

```

```

199         oos.close();
200         fos.close();
201     } catch (IOException e) {
202         e.printStackTrace();
203     }
204     System.out.print("\n> Saved data to " + output.
getAbsolutePath());
205 }
206
207     public static THashMap<Long, THashMap<Integer, Integer[
]>> deserializeData(File file) {
208         // From https://beginnersbook.com/2013/12/how-to-
serialize-hashmap-in-java/
209         THashMap<Long, THashMap<Integer, Integer[]>>
hashMap;
210         try {
211             FileInputStream fis = new FileInputStream(file)
;
212             ObjectInputStream ois = new ObjectInputStream(
fis);
213             hashMap = (THashMap<Long, THashMap<Integer,
Integer[]>>) ois.readObject();
214             ois.close();
215             fis.close();
216         } catch (IOException ioe) {
217             ioe.printStackTrace();
218             return null;
219         } catch (ClassNotFoundException c) {
220             System.out.println("Class not found");
221             c.printStackTrace();
222             return null;
223         }
224         return hashMap;
225     }
226
227     public static void writeJSON(THashMap<Long, THashMap<
Integer, Integer[]>> hashMap, File outputFile) {
228
229         String output = JsonStream.serialize(hashMap);
230         try {
231             outputFile.createNewFile();
232         } catch (IOException e) {

```

```

233         e.printStackTrace();
234     }
235
236     try (PrintStream out = new PrintStream(new
FileOutputStream(outputFile))) {
237         out.print(output);
238         out.flush();
239     } catch (FileNotFoundException e) {
240         e.printStackTrace();
241     }
242
243 }
244
245 public static void main(String[] args) {
246     CommandLine.call(new ProcessOpenDota(), args);
247 }
248
249 @Override
250 public Void call() throws Exception {
251     // BUSINESS LOGIC
252
253     if (extractToJson == null && condense == null) {
254         System.out.println("Well you need to select
something... try --help");
255         return null;
256     }
257     if (extractToJson != null && condense != null) {
258         System.out.println("Can't have it both ways...
try --help");
259         return null;
260     }
261     if (output == null) {
262         System.out.println("Must provide an output file
!");
263         return null;
264     }
265
266     if (extractToJson != null) {
267         System.out.println("Converting from SER to JSON
");
268         extractToJson(extractToJson, output);
269         System.out.println("Conversion complete: " +

```



```
269 output.getAbsolutePath());
270     }
271     if (condense != null) {
272         System.out.println("Condensing from CSV to SER"
273     );
274         condenseInputFile(condense, output, onlyCount);
275         System.out.println("Condensing complete: " +
276     output.getAbsolutePath());
277     }
278     return null;
279 }
280
```