Module 3: Hadoop MapReduce Framework

**Assignment** – Calculate the number of sub-patents

Code :

**package** in.edureka.mapreduce;

**import** java.io.IOException;

**import** java.util.StringTokenizer;

**import** org.apache.hadoop.io.IntWritable;

**import** org.apache.hadoop.io.LongWritable;

**import** org.apache.hadoop.io.Text;

**import** org.apache.hadoop.mapreduce.Mapper;

**import** org.apache.hadoop.mapreduce.Reducer;

**import** org.apache.hadoop.conf.Configuration;

**import** org.apache.hadoop.mapreduce.Job;

**import** org.apache.hadoop.mapreduce.lib.input.TextInputFormat;

**import** org.apache.hadoop.mapreduce.lib.output.TextOutputFormat;

**import** org.apache.hadoop.mapreduce.lib.partition.TotalOrderPartitioner;

**import** org.apache.hadoop.mapreduce.lib.input.FileInputFormat;

**import** org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;

**import** org.apache.hadoop.fs.Path;

**public** **class** PatentCount {

**public** **static** **class** Map **extends** Mapper<LongWritable,Text,Text,IntWritable>{

//Defining a local variable K of type Text

Text k= **new** Text();

//Defining a local variable v of type Text

Text v= **new** Text();

**public** **void** map(LongWritable key, Text value,

Context context)

**throws** IOException,InterruptedException {

String line = value.toString();

StringTokenizer tokenizer = **new** StringTokenizer(line," ");

**while** (tokenizer.hasMoreTokens()) {

String str1= tokenizer.nextToken();

k.set(str1);

String str2= tokenizer.nextToken();

v.set(str2);

//Sending to output collector which inturn passes the same to reducer

context.write(k,**new** IntWritable(1));

}

}

}

**public** **static** **class** Reduce **extends** Reducer<Text,IntWritable,Text,IntWritable>{

**public** **void** reduce(Text key, Iterable<IntWritable> values,

Context context)

**throws** IOException,InterruptedException {

**int** sum=0;

// **TODO** Auto-generated method stub

**for**(IntWritable x: values)

{

sum+=x.get();

}

System.***out***.println("Patent : Number of Associated Sub-patents");

System.***out***.println(key +" : "+ **new** IntWritable(sum));

context.write(key, **new** IntWritable(sum));

}

}

**public** **static** **void** main(String[] args) **throws** Exception {

// **TODO** Auto-generated method stub

//JobConf conf = new JobConf(AlphaCount.class);

Configuration conf= **new** Configuration();

//conf.setJobName("mywc");

Job job = **new** ~~Job~~(conf,"mywc");

job.setJarByClass(PatentCount.**class**);

job.setMapperClass(Map.**class**);

job.setReducerClass(Reduce.**class**);

//conf.setMapperClass(Map.class);

//conf.setReducerClass(Reduce.class);

job.setOutputKeyClass(Text.**class**);

job.setOutputValueClass(IntWritable.**class**);

job.setInputFormatClass(TextInputFormat.**class**);

job.setOutputFormatClass(TextOutputFormat.**class**);

Path outputPath = **new** Path(args[1]);

//Configuring the input/output path from the filesystem into the job

FileInputFormat.*addInputPath*(job, **new** Path(args[0]));

FileOutputFormat.*setOutputPath*(job, **new** Path(args[1]));

//deleting the output path automatically from hdfs so that we don't have delete it explicitly

outputPath.getFileSystem(conf).~~delete~~(outputPath);

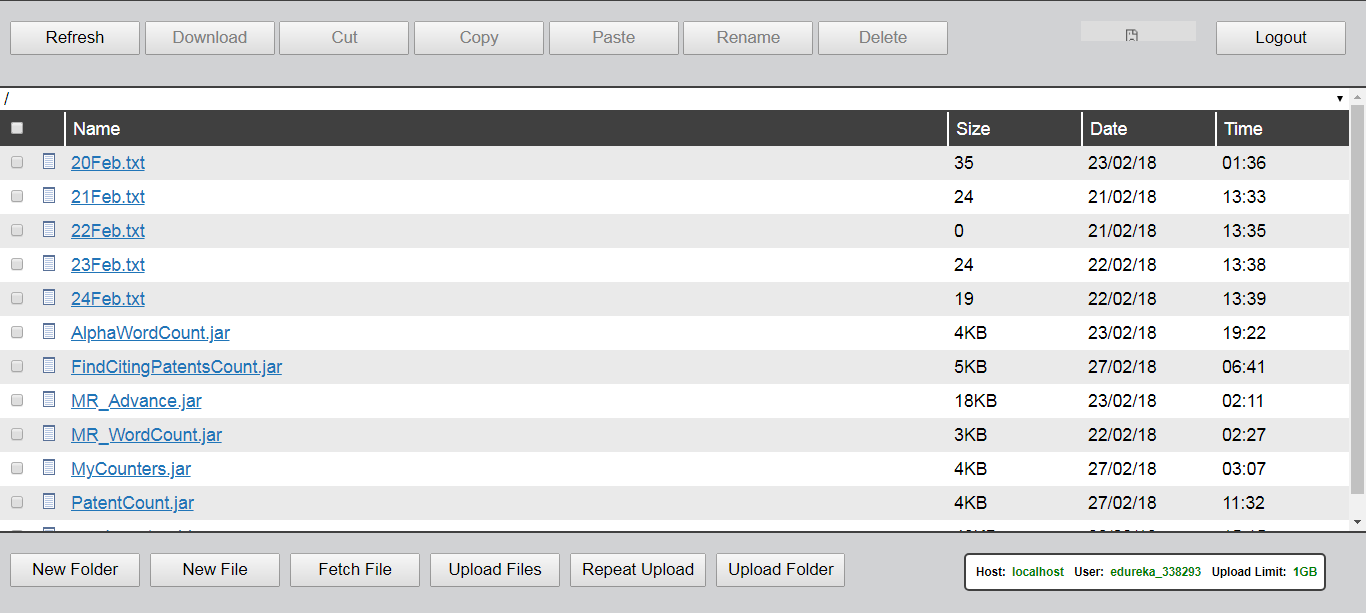
//exiting the job only if the flag value becomes false

System.*exit*(job.waitForCompletion(**true**) ? 0 : 1);

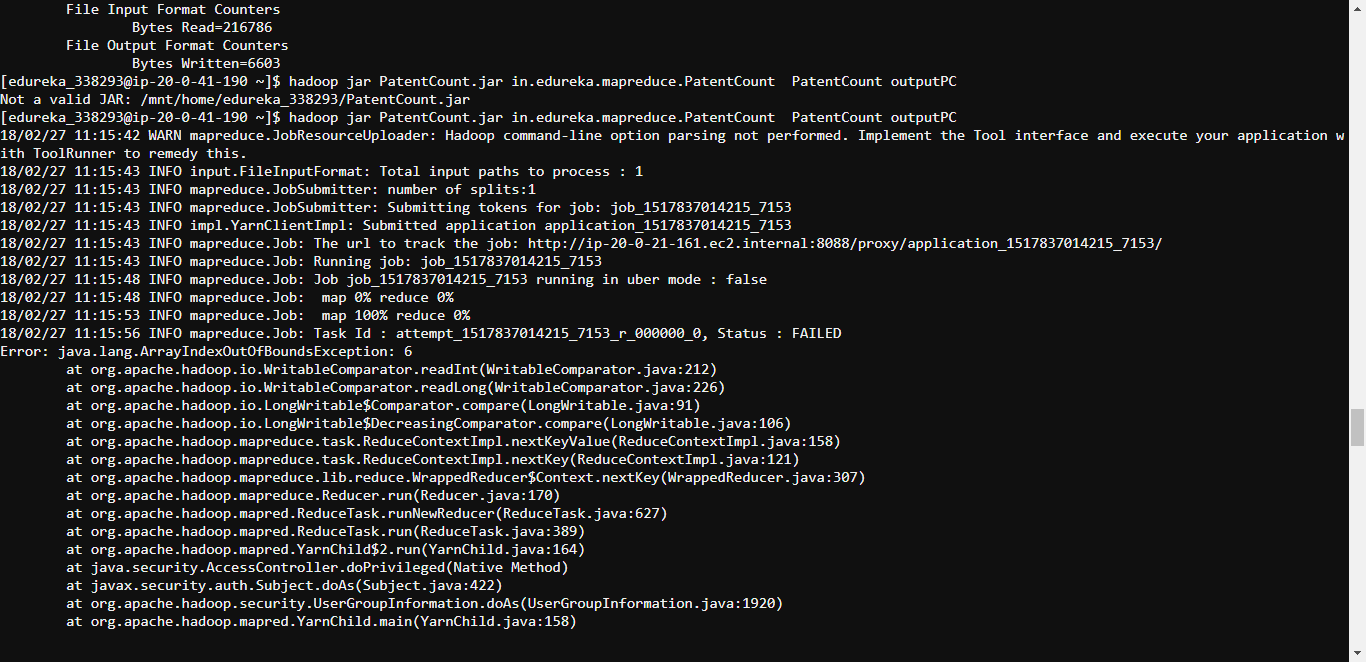
}

}

**FTP:**

****

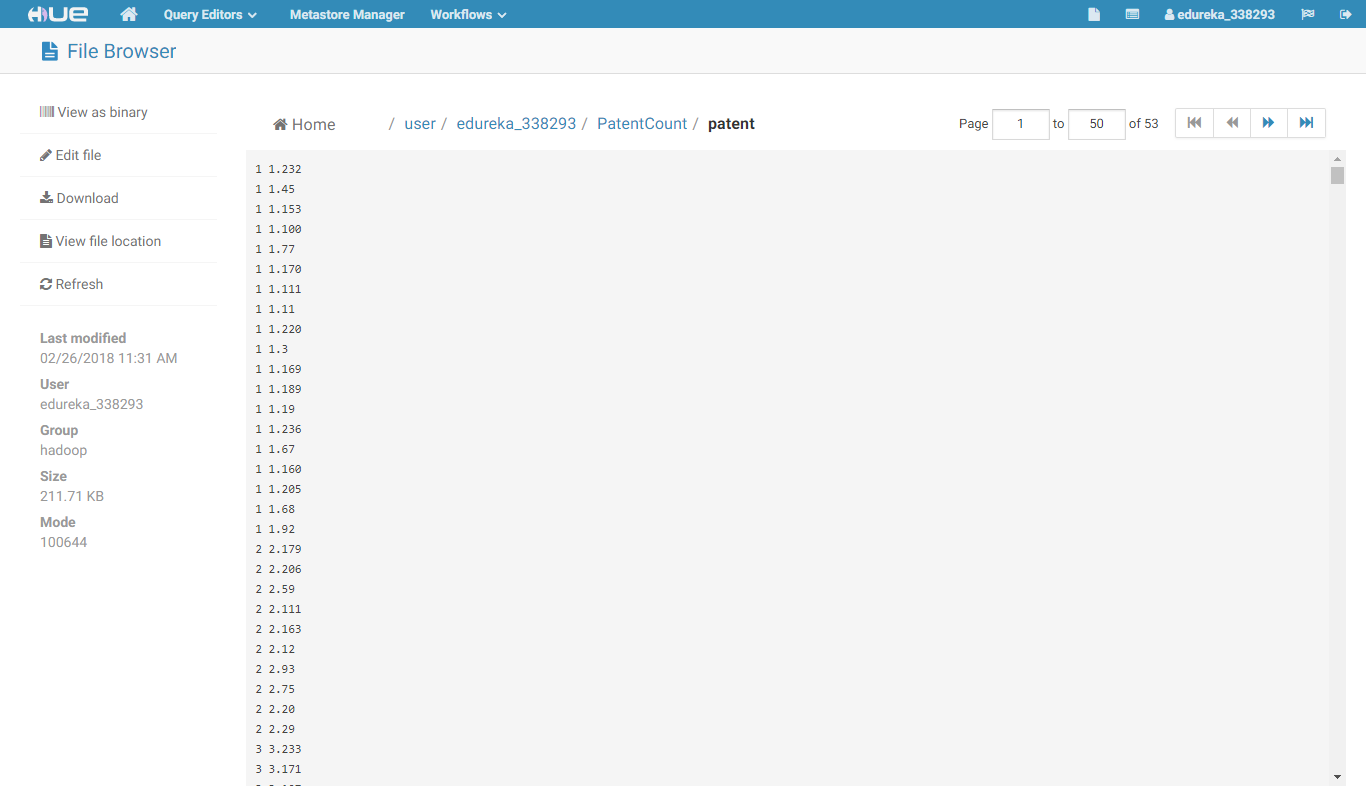
**WebConsole**:

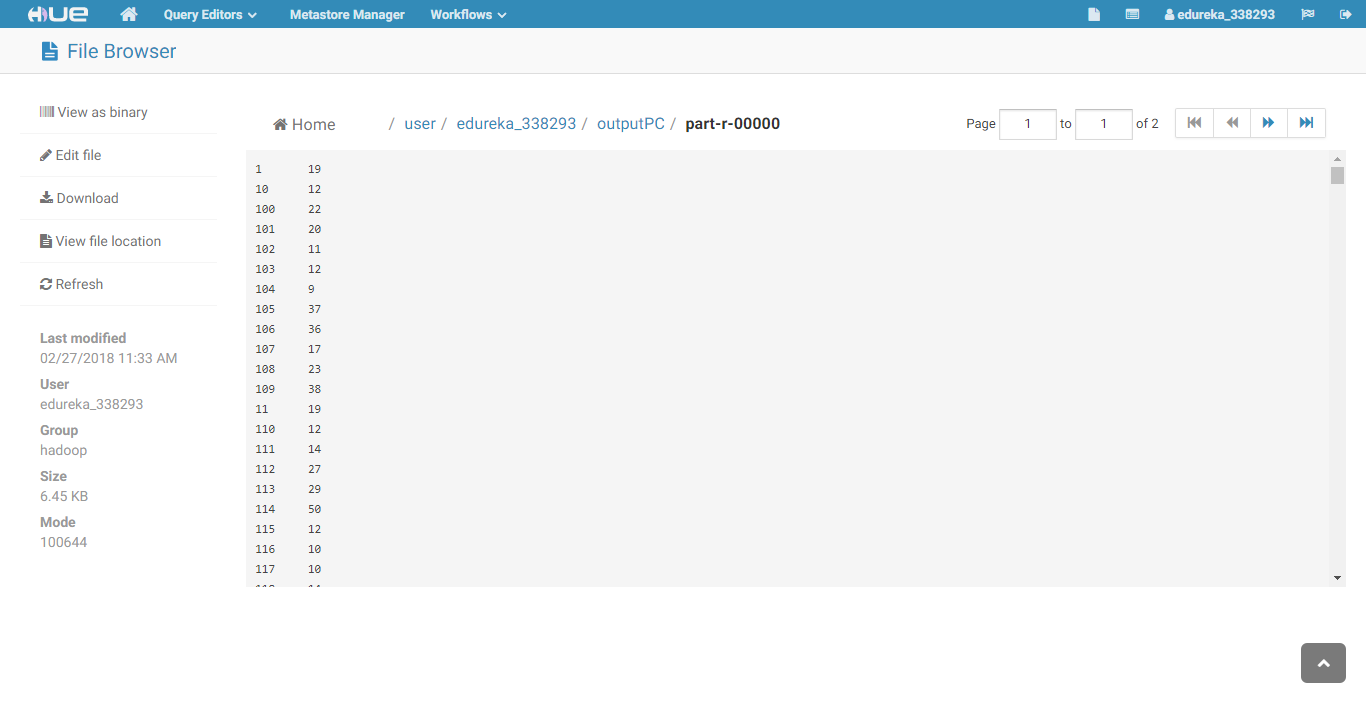


**Command**:

hadoop jar PatentCount.jar in.edureka.mapreduce.PatentCount PatentCount outputPC

**HUE:**

****

****

**Output:**

1 19

10 12

100 22

101 20

102 11

103 12

104 9

105 37

106 36

107 17

108 23

109 38

11 19

110 12

111 14

112 27

113 29

114 50

115 12

116 10

117 10

118 14

119 8

12 28

120 18

121 21

122 6

123 12

124 17

125 5

126 38

127 29

128 16

129 18

13 6

130 26

131 24

132 10

133 25

134 22

135 23

136 24

137 6

138 27

139 4

14 22

140 19

141 23

142 15

143 11

144 19

145 26

146 24

147 28

148 7

149 28

15 18

150 15

151 14

152 20

153 16

154 12

155 9

156 12

157 13

158 10

159 53

16 12

160 13

161 21

162 23

163 7

164 9

165 16

166 7

167 19

168 8

169 32

17 22

170 3

171 13

172 14

173 7

174 14

175 28

176 13

177 5

178 20

179 21

18 49

180 32

181 21

182 24

183 11

184 15

185 2

186 29

187 17

188 20

189 37

19 22

190 27

191 20

192 6

193 30

194 19

195 7

196 20

197 15

198 21

199 9

2 10

20 14

200 16

201 29

202 35

203 14

204 13

205 6

206 13

207 26

208 14

209 2

21 30

210 14

211 33

212 20

213 9

214 17

215 12

216 23

217 33

218 42

219 20

22 28

220 23

221 17

222 13

223 21

224 40

225 27

226 33

227 21

228 24

229 19

23 22

230 8

231 25

232 19

233 8

234 7

235 21

236 11

237 26

238 18

239 15

24 4

240 22

241 20

242 9

243 23

244 7

245 14

246 30

247 15

248 9

249 34

25 24

250 15

251 24

252 14

253 13

254 20

255 18

256 16

257 20

258 40

259 14

26 9

260 21

261 21

262 8

263 1

264 12

265 14

266 9

267 33

268 17

269 1

27 6

270 26

271 12

272 21

273 29

274 7

275 9

276 46

277 11

278 25

279 8

28 18

280 26

281 11

282 37

283 17

284 26

285 11

286 4

287 17

288 33

289 20

29 16

290 19

291 11

292 4

293 17

294 10

295 10

296 14

297 27

298 7

299 18

3 15

30 17

300 12

301 24

302 31

303 16

304 18

305 9

306 22

307 15

308 20

309 33

31 9

310 21

311 18

312 31

313 47

314 10

315 20

316 24

317 22

318 26

319 4

32 29

320 11

321 10

322 18

323 19

324 26

325 11

326 10

327 10

328 14

329 22

33 20

330 33

331 6

333 26

334 5

335 10

336 28

337 5

338 19

339 11

34 12

340 24

341 13

342 25

343 29

344 11

345 4

346 10

347 15

348 3

349 2

35 28

350 23

351 20

352 24

353 18

354 32

355 16

356 12

357 14

358 17

359 39

36 19

360 12

361 32

362 21

363 20

364 6

365 12

366 17

367 9

368 29

369 10

37 30

370 18

371 12

372 38

373 28

374 36

375 2

376 9

377 2

378 12

379 14

38 2

380 10

381 29

382 16

383 14

384 9

385 3

386 33

387 12

388 24

389 10

39 6

390 14

391 1

392 18

393 2

394 36

395 3

396 10

397 14

398 8

399 17

4 9

40 17

400 1

401 19

402 33

403 11

404 41

405 40

406 8

407 15

408 10

409 21

41 8

410 38

411 22

412 6

413 5

414 21

415 25

416 11

417 18

418 13

419 11

42 14

420 12

421 10

422 26

423 17

424 27

425 9

426 14

427 3

428 37

429 33

43 35

430 7

431 8

432 11

433 6

434 24

435 32

436 18

437 16

438 18

439 17

44 16

440 32

441 32

442 9

443 2

444 2

445 31

446 22

447 31

448 12

449 9

45 1

450 17

451 26

452 19

453 31

454 14

455 18

456 19

457 33

458 15

459 29

46 7

460 26

461 33

462 25

463 24

465 5

466 24

467 23

468 9

469 17

47 11

470 20

471 24

472 11

473 35

474 7

475 29

476 8

477 19

478 14

479 10

48 14

480 10

481 11

482 5

483 13

484 17

485 3

486 6

487 5

488 18

489 11

49 16

490 16

491 24

492 38

493 12

494 8

495 37

496 12

497 19

498 20

499 23

5 11

50 22

500 29

501 11

502 36

503 2

504 26

505 53

506 12

507 2

508 6

509 41

51 19

510 10

511 27

512 24

513 26

514 9

515 13

516 5

517 11

518 2

519 33

52 14

520 36

521 17

522 12

523 26

524 8

525 26

526 3

527 20

528 8

529 40

53 26

530 15

531 1

532 23

533 28

534 17

535 29

536 4

537 14

538 16

539 15

54 8

540 5

541 7

542 29

543 21

544 17

545 7

546 19

547 25

548 10

549 30

55 6

550 22

551 13

552 30

553 19

554 6

555 46

556 2

557 3

558 35

559 13

56 13

560 4

561 15

562 22

563 19

564 7

565 35

566 6

567 4

568 26

569 10

57 7

571 11

572 9

573 9

574 17

575 18

576 17

577 9

578 37

579 27

58 4

581 40

582 5

583 37

584 11

585 23

586 7

587 12

588 8

589 25

59 17

590 11

591 10

592 16

593 9

594 8

595 19

596 14

597 21

598 10

599 39

6 21

60 40

600 28

601 13

602 6

603 19

604 22

605 13

606 6

607 4

608 10

609 30

61 19

610 17

611 15

612 35

613 25

614 28

615 33

616 12

618 29

619 6

62 20

620 4

621 3

622 7

623 25

624 32

625 7

626 28

627 1

628 20

629 13

63 13

630 26

631 13

632 12

633 12

634 61

636 3

637 3

638 35

639 34

64 29

640 8

641 13

642 12

643 14

644 12

645 18

646 2

647 24

648 25

649 15

65 9

650 30

651 29

652 14

653 12

654 19

655 11

656 4

657 22

659 14

66 3

660 16

661 5

662 18

663 13

664 20

665 26

666 3

667 28

668 24

669 22

67 26

670 24

671 25

672 31

673 27

674 30

675 12

676 9

677 22

678 8

679 7

68 2

680 32

681 8

682 18

683 8

684 5

685 12

686 10

687 7

688 26

689 16

69 47

690 19

691 22

692 17

693 1

694 11

695 20

696 6

697 12

698 24

699 33

7 25

70 25

700 24

701 26

702 28

703 22

704 4

705 14

706 21

707 9

708 31

709 12

71 31

710 32

711 9

712 32

713 17

714 20

715 9

716 12

717 23

718 21

719 5

72 10

720 20

721 23

722 18

723 7

724 30

725 20

726 39

727 26

728 8

729 7

73 44

730 33

731 9

732 29

733 7

734 24

735 25

736 22

737 32

738 17

739 18

74 11

740 16

741 13

742 14

743 8

744 16

745 3

746 16

747 19

748 20

749 12

75 29

750 14

751 4

752 15

753 40

754 21

755 6

756 18

757 16

758 11

759 28

76 9

760 14

761 10

762 25

763 6

764 32

765 7

766 20

767 4

768 18

769 4

77 39

770 18

771 31

772 18

773 17

774 9

775 21

776 18

777 9

778 4

779 23

78 9

780 9

781 11

782 26

783 15

784 20

785 22

786 21

787 8

788 2

789 9

79 51

790 35

791 16

792 8

793 17

794 25

795 16

796 30

797 34

798 17

799 16

8 7

80 16

800 3

801 21

802 19

803 19

804 7

805 12

806 25

807 13

808 29

809 20

81 10

810 8

811 17

812 7

813 10

814 48

815 7

816 14

817 19

818 18

819 20

82 9

820 12

821 11

822 38

823 38

824 5

825 11

826 12

827 13

828 26

829 25

83 28

830 2

831 21

832 22

833 25

834 27

835 27

836 11

837 18

838 23

839 13

84 15

840 15

841 17

842 34

843 19

844 26

845 12

846 34

847 17

848 13

849 18

85 23

850 3

851 15

852 26

853 32

854 16

855 34

856 15

857 15

858 3

859 13

86 6

860 4

861 22

862 24

863 18

864 16

865 17

866 9

867 3

868 17

869 8

87 16

870 16

871 19

872 19

873 10

874 18

875 6

876 20

877 26

878 14

879 23

88 13

880 9

881 32

882 22

883 5

884 19

885 14

886 31

887 3

888 17

889 33

89 7

890 7

891 9

892 7

893 17

894 13

895 5

896 25

897 10

898 13

899 5

9 15

90 17

900 8

901 15

902 15

903 18

904 17

905 14

906 10

907 17

908 14

909 15

91 33

910 15

911 31

912 17

913 9

914 16

915 21

917 24

918 17

919 13

92 37

920 22

921 15

922 15

923 18

924 32

925 15

926 22

927 7

928 31

929 16

93 19

930 17

931 33

932 16

933 7

934 33

935 15

936 10

937 34

938 8

939 31

94 27

940 6

941 46

942 21

943 27

944 13

945 26

946 9

947 3

948 15

949 10

95 31

950 7

951 37

952 15

953 13

954 4

955 3

956 28

957 19

958 7

959 13

96 13

960 17

961 38

962 3

963 17

964 8

965 32

966 5

967 5

968 57

969 10

97 21

970 28

971 19

972 16

973 30

974 36

975 21

976 23

977 26

978 1

979 16

98 17

980 3

981 16

982 32

983 6

984 22

985 15

986 36

987 13

988 13

989 14

99 28

990 8

991 28

992 15

993 3

994 17

995 29

996 14

997 18

998 22

999 14