

2016-09-13

Updating the USGS Earthquake catalogue data in GMA

The existing earthquake catalogue tables in GMA are outdated (some were last updated in 2009, others in 2013). A new USGS website can be used to update the catalogue. It even goes back beyond our original catalogue start date of 1973 so we can go back to, say, 1960 to ensure that we capture the 1960 Chile and 1964 Alaska earthquakes.

USGS search page: <http://earthquake.usgs.gov/earthquakes/search/>

First, note that the search results are limited to 20,000 records and the search will return an error if greater. So, it may be necessary to download events in time-windowed chunks.

Choose magnitude (e.g. 4.5), date range, (e.g. 2014-2015), Global, CSV format, Order by Time-Newest First.
Hit Search.

Due to the large number of earthquakes, that particular searching yields lots of files, chunked by time windows - there are about 142,000 events with magnitudes between 4.5 and 5 for the inclusive period 1960-2015!

For each range of magnitudes, we can generate catalogues capturing the events from 1960-2015 inclusive.
Then, to update each catalogue, we simply add the events since 2016.

Once we lay out the procedure for generating those "base" (1960-2015 inclusive) catalogues, we can use the same steps to add the updates each year.

2016-09-13

Generate base (1960-2015 inclusive) catalogue - Magnitudes greater than 4.5

--

```
seafloor% pwd
/home/mgds/ingest/USGS_Earthquake_catalogue/mag4.5
```

Use time windows to find date ranges that are within the 20,000 record search limit.

```
seafloor% ll
total 33956
-rw-rw-r-- 1 andrewg mgds 1089420 Sep 12 2016 mag_4.5_1970_1974.csv
-rw-rw-r-- 1 andrewg mgds 2546243 Sep 12 2016 mag_4.5_1975_1979.csv
-rw-rw-r-- 1 andrewg mgds 2851504 Sep 12 2016 mag_4.5_1980_1984.csv
-rw-rw-r-- 1 andrewg mgds 2035377 Sep 12 2016 mag_4.5_1985_1987.csv
-rw-rw-r-- 1 andrewg mgds 2557184 Sep 12 2016 mag_4.5_1988_1991.csv
-rw-rw-r-- 1 andrewg mgds 2700811 Sep 12 2016 mag_4.5_1992_1995.csv
-rw-rw-r-- 1 andrewg mgds 2661103 Sep 12 2016 mag_4.5_1996_1999.csv
-rw-rw-r-- 1 andrewg mgds 3150193 Sep 12 2016 mag_4.5_2000_2003.csv
-rw-rw-r-- 1 andrewg mgds 2296676 Sep 12 2016 mag_4.5_2004_2005.csv
-rw-rw-r-- 1 andrewg mgds 2329147 Sep 12 2016 mag_4.5_2006_2007.csv
-rw-rw-r-- 1 andrewg mgds 2016415 Sep 12 2016 mag_4.5_2008_2009.csv
-rw-rw-r-- 1 andrewg mgds 2784697 Sep 12 2016 mag_4.5_2010_2011.csv
-rw-rw-r-- 1 andrewg mgds 2417590 Sep 12 2016 mag_4.5_2012_2013.csv
-rw-rw-r-- 1 andrewg mgds 2822105 Sep 12 2016 mag_4.5_2014_2015.csv
```

Each of the files extracted from USGS is in reverse chronological order and has a header as the first row.

```
seafloor% tail mag_4.5_1960_1969.csv
1960-01-
11T03:10:18.000Z,16.046,95.832,15,5.6,mw,,,,,iscgem,iscgem877983,2015-05-
13T18:52:54.000Z,"near the south coast of
Myanmar",earthquake,,,,,automatic,iscgem,iscgem
1960-01-08T14:46:02.000Z,-55.585,-
27.475,35,6,mw,,,,,iscgem,iscgem877967,2015-05-13T18:52:54.000Z,"South
Sandwich Islands region",earthquake,,,,,automatic,iscgem,iscgem
1960-01-08T11:29:21.000Z,-55.57,-
27.115,15,5.9,mw,,,,,iscgem,iscgem877965,2015-05-13T18:52:54.000Z,"South
Sandwich Islands region",earthquake,,,,,automatic,iscgem,iscgem
1960-01-
07T23:17:30.000Z,6.352,94.547,55,5.5,mw,,,,,iscgem,iscgem877960,2015-05-
13T18:52:54.000Z,"Nicobar Islands, India
region",earthquake,,,,,automatic,iscgem,iscgem
1960-01-07T13:28:20.000Z,-55.591,-
27.278,35,6.2,mw,,,,,iscgem,iscgem877956,2015-05-13T18:52:54.000Z,"South
Sandwich Islands region",earthquake,,,,,automatic,iscgem,iscgem
1960-01-
07T08:15:27.000Z,6.418,94.756,15,5.6,mw,,,,,iscgem,iscgem877954,2015-05-
```

```

13T18:52:54.000Z,"Nicobar Islands, India
region",earthquake,,,,,automatic,iscgem,iscgem
1960-01-
04T12:52:00.000Z,45.069,26.829,40,5.4,mw,,,,,iscgemsup,iscgemsup877933,20
15-07-13T17:02:23.000Z,"Romania",earthquake,,,,,automatic,iscgem,iscgem
1960-01-
04T06:16:35.000Z,11.374,42.609,15,6.1,mw,,,,,iscgemsup,iscgemsup877930,20
15-07-13T17:02:23.000Z,"Djibouti",earthquake,,,,,automatic,iscgem,iscgem
1960-01-
03T11:24:05.000Z,43.7,84.542,15,5.7,mw,,,,,iscgem,iscgem877920,2015-05-
13T18:52:54.000Z,"northern Xinjiang,
China",earthquake,,,,,automatic,iscgem,iscgem
1960-01-02T12:21:58.000Z,-55.877,-
1.89,15,6.3,mw,,,,,iscgem,iscgem877909,2015-05-13T18:52:54.000Z,"Bouvet
Island region",earthquake,,,,,automatic,iscgem,iscgem
seafloor%

```

```

seafloor% head -2 mag_4.5_1960_1969.csv
time,latitude,longitude,depth,mag,magType,nst,gap,dmin,rms,net,id,updated
,place,type,horizontalError,depthError,magError,magNst,status,locationSou
rce,magSource
1969-12-
31T19:01:56.000Z,28.532,129.075,25,6.3,mw,,,,,iscgem,iscgem801741,2015-
05-13T18:53:03.000Z,"Ryukyu Islands,
Japan",earthquake,,,,,automatic,iscgem,iscgem

```

We need to do a few things to get the files into a more suitable format for the average GMA user. The same approach is used for generating the tabels for each magnitude range:

- Combine the files into one file for everything prior to 2016.
- Open the combined file in Excel (the default for CSV files).
- Get rid of some columns that the average user will not require, move the Location column left. Move the Magnitude column left.
- On the PC, save the modified CSV file as tab-separated so that we can later use gawk to build a URL field.
- Get rid of the headers from all but the first file (we can use gawk to exclude lines containing the old header info).
- Use sed to get rid of the quotation marks around the location descriptions. (Once it's in tab-separated format, we can do that and the commas that are in some of the location descriptions will have no deleterious effect.)
- Repeat for the smaller more recent events file.
- Combine the everything-to-2016 file and the smaller recent file.
- Add to GMA menu.
- Fill in gma_menu.data_update_date field so we know when to update the file.

Combine the files (start with the most recent first):

```
seafloor% cat mag_4.5_2014_2015.csv mag_4.5_2012_2013.csv
mag_4.5_2010_2011.csv mag_4.5_2008_2009.csv mag_4.5_2006_2007.csv \
mag_4.5_2004_2005.csv mag_4.5_2000_2003.csv mag_4.5_1996_1999.csv
mag_4.5_1992_1995.csv mag_4.5_1988_1991.csv mag_4.5_1985_1987.csv \
mag_4.5_1980_1984.csv mag_4.5_1975_1979.csv mag_4.5_1970_1974.csv
mag_4.5_1960_1969.csv > mag_4.5_1960_2015_combined.csv
```

```
seafloor% ll mag_4.5_1960_2015_combined.csv
-rw-rw-r-- 1 andrewg mgds 34743789 Sep 13 13:13
mag_4.5_1960_2015_combined.csv
```

```
seafloor% wc -l mag_4.5_1960_2015_combined.csv
217308 mag_4.5_1960_2015_combined.csv
```

```
seafloor% head -3 mag_4.5_1960_2015_combined.csv
time,latitude,longitude,depth,mag,magType,nst,gap,dmin,rms,net,id,updated
,place,type,horizontalError,depthError,magError,magNst,status,locationSou
rce,magSource
2015-12-31T23:39:28.940Z,-
7.0711,129.8517,113.43,4.6,mb,,69,1.696,0.75,us,us10004anb,2016-03-
18T01:13:10.040Z,"188km WNW of Saumlaki,
Indonesia",earthquake,7.3,7,0.08,47,reviewed,us,us
2015-12-31T23:25:03.900Z,-
15.0436,167.1007,30.43,4.6,mb,,118,0.413,0.61,us,us10004bv6,2016-03-
18T01:13:10.040Z,"5km E of Port-Olry,
Vanuatu",earthquake,12.8,2.6,0.173,10,reviewed,us,us
```

Transfer the combined CSV file to the PC and open the file in Excel.
Get rid of some columns that the average user will not require, move the
"mag" and "place" columns left.
Save as a tab-separated file and transfer back to machine seafloor.

```
seafloor% ll *comb*
-rw-rw-r-- 1 andrewg mgds 34743789 Sep 13 13:13
mag_4.5_1960_2015_combined.csv
-rw-rw-r-- 1 andrewg mgds 18641979 Sep 13 14:14
mag_4.5_1960_2015_combined_tab.txt
```

```
seafloor% head -3 mag_4.5_1960_2015_combined_tab.txt
time latitude longitude mag depth place id
2015-12-31T23:39:28.940Z -7.0711 129.8517 4.6 113.43
"188km WNW of Saumlaki, Indonesia" us10004anb
2015-12-31T23:25:03.900Z -15.0436 167.1007 4.6
30.43 "5km E of Port-Olry, Vanuatu" us10004bv6
```

Get rid of all of the header lines from the combined tab-separated file
(we use lots

of redundancy to ensure that we only get rid of those header lines).
There are fifteen
(15) header lines. So, when we knock them out, the output should have 15
rows less than
the original combined file:

```
seafloor% cat mag_4.5_1960_2015_combined_tab.txt | grep -v latitude |  
grep -v longitude | grep -v depth > /tmp/a
```

```
seafloor% wc -l mag_4.5_1960_2015_combined_tab.txt /tmp/a  
217308 mag_4.5_1960_2015_combined_tab.txt  
217293 /tmp/a
```

The new temporary file has seven (7) columns:

```
seafloor% head -3 /tmp/a  
2015-12-31T23:39:28.940Z      -7.0711 129.8517      4.6      113.43  
"188km WNW of Saumlaki, Indonesia"      us10004anb  
2015-12-31T23:25:03.900Z      -15.0436      167.1007      4.6  
30.43      "5km E of Port-Olry, Vanuatu"      us10004bv6  
2015-12-31T23:15:15.690Z      -5.5171 -11.4809      4.6      10  
Ascension Island region us10004bv7
```

Use gawk to generate a USGS URL using the ID value in the last column:

```
seafloor% gawk -F"\t" '{printf  
"%s\t%s\t%s\t%s\t%s\t%s\t%thttp://earthquake.usgs.gov/earthquakes/eventpage  
/%s\n", $1, $2, $3, $4, $5, $6, $7}' /tmp/a > /tmp/b  
seafloor% head -3 /tmp/b  
2015-12-31T23:39:28.940Z      -7.0711 129.8517      4.6      113.43  
"188km WNW of Saumlaki, Indonesia"  
http://earthquake.usgs.gov/earthquakes/eventpage/us10004anb  
2015-12-31T23:25:03.900Z      -15.0436      167.1007      4.6  
30.43      "5km E of Port-Olry, Vanuatu"  
http://earthquake.usgs.gov/earthquakes/eventpage/us10004bv6  
2015-12-31T23:15:15.690Z      -5.5171 -11.4809      4.6      10  
Ascension Island region  
http://earthquake.usgs.gov/earthquakes/eventpage/us10004bv7
```

Use sed to get rid of the quotation marks around the location
descriptions:

```
seafloor% sed 's/\\"//g' /tmp/b > /tmp/c  
seafloor% ll /tmp/b /tmp/c  
-rw-rw-r-- 1 andrewg andrewg 29288667 Sep 13 16:07 /tmp/b  
-rw-rw-r-- 1 andrewg andrewg 29066455 Sep 13 16:13 /tmp/c  
  
seafloor% wc -l /tmp/b /tmp/c  
217293 /tmp/b  
217293 /tmp/c
```

```

seafloor% head -3 /tmp/c
2015-12-31T23:39:28.940Z      -7.0711 129.8517      4.6      113.43
188km WNW of Saumlaki, Indonesia
http://earthquake.usgs.gov/earthquakes/eventpage/us10004anb
2015-12-31T23:25:03.900Z      -15.0436      167.1007      4.6
30.43      5km E of Port-Olry, Vanuatu
http://earthquake.usgs.gov/earthquakes/eventpage/us10004bv6
2015-12-31T23:15:15.690Z      -5.5171 -11.4809      4.6      10
Ascension Island region
http://earthquake.usgs.gov/earthquakes/eventpage/us10004bv7

```

Finally, for this everything-to-2015 inclusive file, we need to add a header.
Generate a new header row with clearer labels (use a dummy echo command to
produce one line of input that drives the gawk command, and remember to make
the labels tab-separated):

```

seafloor% echo "hello" | gawk '{printf
"Time\tLatitude\tLongitude\tMagnitude\tDepth\tLocation\tInformation\n"}'
> header
seafloor% cat header
Time      Latitude      Longitude      Magnitude      Depth      Location
Information

```

Combine the header with the body and save it as the base file that we will use for updating the earthquake catalogue:

```
seafloor% cat header /tmp/c > mag_4.5_1960_2015_combined_reformat.txt
```

The new file will have one additional row because of the header.

```

seafloor% ll
total 115324
-rw-rw-r-- 1 andrewg mgds      61 Sep 13  2016 header
-rw-rw-r-- 1 andrewg mgds 485324 Sep 12  2016 mag_4.5_1960_1969.csv
-rw-rw-r-- 1 andrewg mgds 34743789 Sep 13  2016
mag_4.5_1960_2015_combined.csv
-rw-rw-r-- 1 andrewg mgds 29066516 Sep 13  2016
mag_4.5_1960_2015_combined_reformat.txt
-rw-rw-r-- 1 andrewg mgds 18641979 Sep 13  2016
mag_4.5_1960_2015_combined_tab.txt
-rw-rw-r-- 1 andrewg mgds 1089420 Sep 12  2016 mag_4.5_1970_1974.csv
-rw-rw-r-- 1 andrewg mgds 2546243 Sep 12  2016 mag_4.5_1975_1979.csv
-rw-rw-r-- 1 andrewg mgds 2851504 Sep 12  2016 mag_4.5_1980_1984.csv
-rw-rw-r-- 1 andrewg mgds 2035377 Sep 12  2016 mag_4.5_1985_1987.csv
-rw-rw-r-- 1 andrewg mgds 2557184 Sep 12  2016 mag_4.5_1988_1991.csv
-rw-rw-r-- 1 andrewg mgds 2700811 Sep 12  2016 mag_4.5_1992_1995.csv
-rw-rw-r-- 1 andrewg mgds 2661103 Sep 12  2016 mag_4.5_1996_1999.csv

```

```

-rw-rw-r-- 1 andrewg mgds 3150193 Sep 12 2016 mag_4.5_2000_2003.csv
-rw-rw-r-- 1 andrewg mgds 2296676 Sep 12 2016 mag_4.5_2004_2005.csv
-rw-rw-r-- 1 andrewg mgds 2329147 Sep 12 2016 mag_4.5_2006_2007.csv
-rw-rw-r-- 1 andrewg mgds 2016415 Sep 12 2016 mag_4.5_2008_2009.csv
-rw-rw-r-- 1 andrewg mgds 2784697 Sep 12 2016 mag_4.5_2010_2011.csv
-rw-rw-r-- 1 andrewg mgds 2417590 Sep 12 2016 mag_4.5_2012_2013.csv
-rw-rw-r-- 1 andrewg mgds 2822105 Sep 12 2016 mag_4.5_2014_2015.csv

```

```

seafloor% head mag_4.5_1960_2015_combined_reformat.txt
Time      Latitude      Longitude      Magnitude      Depth      Location
Information
2015-12-31T23:39:28.940Z      -7.0711 129.8517      4.6      113.43
188km WNW of Saumlaki, Indonesia
http://earthquake.usgs.gov/earthquakes/eventpage/us10004anb
2015-12-31T23:25:03.900Z      -15.0436      167.1007      4.6
30.43      5km E of Port-Olry, Vanuatu
http://earthquake.usgs.gov/earthquakes/eventpage/us10004bv6
2015-12-31T23:15:15.690Z      -5.5171 -11.4809      4.6      10
Ascension Island region
http://earthquake.usgs.gov/earthquakes/eventpage/us10004bv7
2015-12-31T22:23:12.000Z      -7.25      67.75      5      10      Mid-
Indian Ridge
http://earthquake.usgs.gov/earthquakes/eventpage/gcmt20151231222312
2015-12-31T22:22:17.350Z      51.6783 -173.489      5      39.85
75km SE of Atka, Alaska
http://earthquake.usgs.gov/earthquakes/eventpage/us10004amq
2015-12-31T22:15:58.410Z      40.8895 72.5413 4.5      48.21      20km NE
of Andijon, Uzbekistan
http://earthquake.usgs.gov/earthquakes/eventpage/us10004amk
2015-12-31T21:51:11.940Z      -7.285      68.0275 4.9      10      Chagos
Archipelago region
http://earthquake.usgs.gov/earthquakes/eventpage/us10004amv
2015-12-31T21:29:38.860Z      -30.3584      -178.0074      5.4
27.28      121km S of Raoul Island, New Zealand
http://earthquake.usgs.gov/earthquakes/eventpage/us10004ama
2015-12-31T21:29:11.500Z      -50.3784      114.3863      4.9
10      Western Indian-Antarctic Ridge
http://earthquake.usgs.gov/earthquakes/eventpage/us10004amb

```

```

seafloor% wc -l mag_4.5_1960_2015_combined_reformat.txt
217294 mag_4.5_1960_2015_combined_reformat.txt

```

Tests show that the file imports fine into GMA.

When it comes to update the file by adding more recent events to the base catalogue,
remember that both the magnitude and the location columns were moved leftwards. So,
make sure that the more recent events have the same column ordering before catting
the files.

2016-09-13

Generate base (1960-2015 inclusive) catalogue - Magnitudes greater than 5.0

--

```
seafloor% pwd
/home/mgds/ingest/USGS_Earthquake_catalogue/mag5.0
```

```
seafloor% ll
total 11516
```

```
-rw-rw-r-- 1 andrewg mgds 2291450 Sep 13 16:59 mag_5.0_1960_1979.csv
-rw-rw-r-- 1 andrewg mgds 2463542 Sep 13 17:01 mag_5.0_1980_1989.csv
-rw-rw-r-- 1 andrewg mgds 2262013 Sep 13 17:01 mag_5.0_1990_1999.csv
-rw-rw-r-- 1 andrewg mgds 2784928 Sep 13 17:02 mag_5.0_2000_2009.csv
-rw-rw-r-- 1 andrewg mgds 1982078 Sep 13 17:03 mag_5.0_2010_2015.csv
```

```
seafloor% head mag_5.0_1960_1979.csv
```

```
time,latitude,longitude,depth,mag,magType,nst,gap,dmin,rms,net,id,updated
,place,type,horizontalError,depthError,magError,magNst,status,locationSource,magSource
```

```
1979-12-31T23:07:23.400Z,2.099,-79.017,33,5.5,ms,,,,,us,usp00014tb,2015-
05-13T18:53:16.000Z,"south of Panama",earthquake,,,,,reviewed,us,us
1979-12-31T11:58:39.400Z,3.616,-78.319,33,5.1,mb,,,,,us,usp00014t5,2014-
11-06T23:22:08.659Z,"south of Panama",earthquake,,,,,reviewed,us,us
1979-12-31T06:21:34.300Z,36.184,31.51,79,5.3,mb,,,,,us,usp00014sz,2015-
05-13T18:53:16.000Z,"western Turkey",earthquake,,,,,reviewed,us,us
1979-12-31T02:47:50.400Z,40.216,144.109,33,5.5,mb,,,,,us,usp00014sx,2015-
02-11T17:14:48.000Z,"off the east coast of Honshu,
Japan",earthquake,,,,,reviewed,us,us
1979-12-30T21:19:01.900Z,0.964,125.99,41,5.8,mb,,,,,us,usp00014su,2015-
05-13T18:53:16.000Z,"Molucca Sea",earthquake,,,,,reviewed,us,us
1979-12-30T17:46:50.500Z,-35.128,-
16.051,10,5.3,ms,,,,,us,usp00014st,2014-11-06T23:22:08.639Z,"southern
Mid-Atlantic Ridge",earthquake,,,,,reviewed,us,us
1979-12-30T13:53:02.200Z,-23.351,-13.373,10,5,mb,,,,,us,usp00014sn,2014-
11-06T23:22:08.601Z,"southern Mid-Atlantic
Ridge",earthquake,,,,,reviewed,us,us
1979-12-30T06:25:48.700Z,-32.602,-
70.526,78,5.2,mb,,,,,us,usp00014sg,2014-11-06T23:22:08.594Z,"Valparaiso,
Chile",earthquake,,,,,reviewed,us,us
```



```

1979-12-
30T04:18:33.800Z,52.525,152.259,555,5.4,mb,,,,us,usp00014se,2015-05-
13T18:53:16.000Z,"northwest of the Kuril
Islands",earthquake,,,,reviewed,us,us

```

Follow the above procedures to reformat the files.

Combine the files (start with the most recent first):

```

seafloor% cat mag_5.0_2010_2015.csv mag_5.0_2000_2009.csv
mag_5.0_1990_1999.csv mag_5.0_1980_1989.csv mag_5.0_1960_1979.csv >
mag_5.0_1960_2015_combined.csv

```

```

seafloor% ll
total 23024
-rw-rw-r-- 1 andrewg mgds 2291450 Sep 13 16:59 mag_5.0_1960_1979.csv
-rw-rw-r-- 1 andrewg mgds 11784011 Sep 13 17:07
mag_5.0_1960_2015_combined.csv
-rw-rw-r-- 1 andrewg mgds 2463542 Sep 13 17:01 mag_5.0_1980_1989.csv
-rw-rw-r-- 1 andrewg mgds 2262013 Sep 13 17:01 mag_5.0_1990_1999.csv
-rw-rw-r-- 1 andrewg mgds 2784928 Sep 13 17:02 mag_5.0_2000_2009.csv
-rw-rw-r-- 1 andrewg mgds 1982078 Sep 13 17:03 mag_5.0_2010_2015.csv

```

```

seafloor% wc -l mag_5.0_1960_2015_combined.csv
74488 mag_5.0_1960_2015_combined.csv

```

```

seafloor%
seafloor% head -3 mag_5.0_1960_2015_combined.csv
time,latitude,longitude,depth,mag,magType,nst,gap,dmin,rms,net,id,updated
,place,type,horizontalError,depthError,magError,magNst,status,locationSou
rce,magSource
2015-12-31T22:23:12.000Z,-
7.25,67.75,10,5,mwc,,,,gcmt,gcmt20151231222312,2016-04-
15T00:59:41.000Z,"Mid-Indian Ridge",earthquake,,,,automatic,gcmt,gcmt
2015-12-31T22:22:17.350Z,51.6783,-
173.489,39.85,5,mb,,135,0.683,0.82,us,us10004amq,2016-03-
18T01:13:09.040Z,"75km SE of Atka,
Alaska",earthquake,5.4,6.2,0.027,423,reviewed,us,us

```

Transfer the combined CSV file to the PC and open the file in Excel.
 Get rid of some columns that the average user will not require, move the
 "mag" and "place" columns left.
 Save as a tab-separated file and transfer back to machine seafloor.

```

seafloor% ll *comb*
-rw-rw-r-- 1 andrewg mgds 11784011 Sep 13 17:07
mag_5.0_1960_2015_combined.csv

```

```
-rw-rw-r-- 1 andrewg mgds 6338690 Sep 13 2016
mag_5.0_1960_2015_combined_tab.txt
```

```
seafloor% head -3 mag_5.0_1960_2015_combined_tab.txt
time      latitude      longitude      mag      depth      place      id
2015-12-31T22:23:12.000Z      -7.25      67.75      5      10      Mid-
Indian Ridge      gcmt20151231222312
2015-12-31T22:22:17.350Z      51.6783 -173.489      5      39.85
"75km SE of Atka, Alaska"      us10004amq
```

Get rid of all of the header lines from the combined tab-separated file (we use lots of redundancy to ensure that only get rid of those header lines). There are multiple header lines so that the output should have that many rows less than the original combined file:

```
seafloor% cat mag_5.0_1960_2015_combined_tab.txt | grep -v latitude |
grep -v longitude | grep -v depth > /tmp/a
```

```
seafloor% wc -l mag_5.0_1960_2015_combined_tab.txt /tmp/a
74488 mag_5.0_1960_2015_combined_tab.txt
74483 /tmp/a
```

The new temporary file has seven (7) columns:

```
seafloor% head -3 /tmp/a
2015-12-31T22:23:12.000Z      -7.25      67.75      5      10      Mid-
Indian Ridge      gcmt20151231222312
2015-12-31T22:22:17.350Z      51.6783 -173.489      5      39.85
"75km SE of Atka, Alaska"      us10004amq
2015-12-31T21:29:38.860Z      -30.3584      -178.0074      5.4
27.28      "121km S of Raoul Island, New Zealand"      us10004ama
```

Use gawk to generate a URL using the ID value in the last column:

```
seafloor% gawk -F"\t" '{printf
"%s\t%s\t%s\t%s\t%s\t%s\t%thttp://earthquake.usgs.gov/earthquakes/eventpage
/%s\n", $1, $2, $3, $4, $5, $6, $7}' /tmp/a > /tmp/b
seafloor% head -3 /tmp/b
2015-12-31T22:23:12.000Z      -7.25      67.75      5      10      Mid-
Indian Ridge
http://earthquake.usgs.gov/earthquakes/eventpage/gcmt20151231222312
2015-12-31T22:22:17.350Z      51.6783 -173.489      5      39.85
"75km SE of Atka, Alaska"
http://earthquake.usgs.gov/earthquakes/eventpage/us10004amq
2015-12-31T21:29:38.860Z      -30.3584      -178.0074      5.4
27.28      "121km S of Raoul Island, New Zealand"
http://earthquake.usgs.gov/earthquakes/eventpage/us10004ama
```

Use sed to get rid of the quotation marks around the location descriptions:

```
seafloor% sed 's/\\"//g' /tmp/b > /tmp/c
seafloor% ll /tmp/b /tmp/c
-rw-rw-r-- 1 andrewg andrewg 9988137 Sep 13 18:12 /tmp/b
-rw-rw-r-- 1 andrewg andrewg 9916595 Sep 13 18:13 /tmp/c

seafloor% wc -l /tmp/b /tmp/c
  74483 /tmp/b
  74483 /tmp/c

seafloor% head -3 /tmp/c
2015-12-31T22:23:12.000Z      -7.25    67.75    5          10          Mid-
Indian Ridge
http://earthquake.usgs.gov/earthquakes/eventpage/gcmt20151231222312
2015-12-31T22:22:17.350Z      51.6783 -173.489          5          39.85
75km SE of Atka, Alaska
http://earthquake.usgs.gov/earthquakes/eventpage/us10004amq
2015-12-31T21:29:38.860Z      -30.3584          -178.0074          5.4
27.28    121km S of Raoul Island, New Zealand
http://earthquake.usgs.gov/earthquakes/eventpage/us10004ama
```

Finally, for this everything-to-2015 inclusive file, we need to add a header.

Generate a new header row with clearer labels (use a dummy echo command to produce one line of input that drives the gawk command, and remember to make the labels tab-separated):

```
seafloor% echo "hello" | gawk '{printf
"Time\tLatitude\tLongitude\tMagnitude\tDepth\tLocation\tInformation\n"}'
> header
seafloor% cat header
Time      Latitude      Longitude      Magnitude      Depth      Location
Information
```

Combine the header with the body and save it as the base file that we will use for updating the earthquake catalogue:

```
seafloor% cat header /tmp/c > mag_5.0_1960_2015_combined_reformat.txt
```

The new file will have one additional row because of the header.

```
seafloor% ll
total 38908
-rw-rw-r-- 1 andrewg mgds          61 Sep 13 15:30 header
-rw-rw-r-- 1 andrewg mgds 2291450 Sep 13 16:59 mag_5.0_1960_1979.csv
```

```

-rw-rw-r-- 1 andrewg mgds 11784011 Sep 13 17:07
mag_5.0_1960_2015_combined.csv
-rw-rw-r-- 1 andrewg mgds 9916656 Sep 13 18:13
mag_5.0_1960_2015_combined_reformat.txt
-rw-rw-r-- 1 andrewg mgds 6338690 Sep 13 17:18
mag_5.0_1960_2015_combined_tab.txt
-rw-rw-r-- 1 andrewg mgds 2463542 Sep 13 17:01 mag_5.0_1980_1989.csv
-rw-rw-r-- 1 andrewg mgds 2262013 Sep 13 17:01 mag_5.0_1990_1999.csv
-rw-rw-r-- 1 andrewg mgds 2784928 Sep 13 17:02 mag_5.0_2000_2009.csv
-rw-rw-r-- 1 andrewg mgds 1982078 Sep 13 17:03 mag_5.0_2010_2015.csv

```

```

seafloor% head mag_5.0_1960_2015_combined_reformat.txt
Time      Latitude      Longitude      Magnitude      Depth      Location
Information
2015-12-31T22:23:12.000Z      -7.25      67.75      5      10      Mid-
Indian Ridge
http://earthquake.usgs.gov/earthquakes/eventpage/gcmt20151231222312
2015-12-31T22:22:17.350Z      51.6783 -173.489      5      39.85
75km SE of Atka, Alaska
http://earthquake.usgs.gov/earthquakes/eventpage/us10004amq
2015-12-31T21:29:38.860Z      -30.3584      -178.0074      5.4
27.28      121km S of Raoul Island, New Zealand
http://earthquake.usgs.gov/earthquakes/eventpage/us10004ama
2015-12-31T10:56:58.120Z      11.1535 -86.6205      5.8      20
71km S of Masachapa, Nicaragua
http://earthquake.usgs.gov/earthquakes/eventpage/us10004akg
2015-12-30T11:08:36.000Z      -21.9779      -174.4701      5
46.15      86km SE of `Ohonua, Tonga
http://earthquake.usgs.gov/earthquakes/eventpage/us10004ae8
2015-12-30T03:25:34.030Z      14.137 -91.1098      5      81.05
8km NW of La Gomera, Guatemala
http://earthquake.usgs.gov/earthquakes/eventpage/us10004acp
2015-12-30T02:58:36.450Z      -34.4192      -109.3543      5
10      Southern East Pacific Rise
http://earthquake.usgs.gov/earthquakes/eventpage/us10004aci
2015-12-29T02:54:06.000Z      6.862 -76.701 5      5.7      14km SSE
of Murindo, Colombia
http://earthquake.usgs.gov/earthquakes/eventpage/us10004a6p
2015-12-29T01:51:41.550Z      -6.2452 154.7595      5.8      30
80km W of Panguna, Papua New Guinea
http://earthquake.usgs.gov/earthquakes/eventpage/us10004a6n

```

```

seafloor% wc -l mag_5.0_1960_2015_combined_reformat.txt
74484 mag_5.0_1960_2015_combined_reformat.txt

```

Tests show that the file imports fine into GMA.

When it comes to update the file by adding more recent events to the base catalogue, remember that both the magnitude and the location columns were moved leftwards. So, make sure that the more recent events have the same column ordering before catting

the files.

```
*****
*****
```

2016-09-13

Generate base (1960-2015 inclusive) catalogue - Magnitudes greater than 5.5

--

```
seafloor% pwd
/home/mgds/ingest/USGS_Earthquake_catalogue/mag5.5
```

```
seafloor% ll
total 3760
-rw-rw-r-- 1 andrewg mgds 2486357 Sep 14 09:23 mag_5.5_1960_1999.csv
-rw-rw-r-- 1 andrewg mgds 1357319 Sep 14 2016 mag_5.5_2000_2015.csv
```

```
seafloor% head mag_5.5_1960_1999.csv
time,latitude,longitude,depth,mag,magType,nst,gap,dmin,rms,net,id,updated
,place,type,horizontalError,depthError,magError,magNst,status,locationSou
rce,magSource
1999-12-
31T13:09:11.610Z,37.269,134.664,371.4,5.6,mwc,,,,,0.78,us,usp0009kj2,2015-
02-11T17:17:01.000Z,"Sea of Japan",earthquake,,,,,reviewed,us,hrv
1999-12-
30T13:21:36.710Z,47.579,154.508,33,5.6,mwc,,,,,0.99,us,usp0009kgn,2015-05-
13T18:53:41.000Z,"Kuril Islands",earthquake,,,,,reviewed,us,hrv
1999-12-29T22:53:57.190Z,-
11.165,165.33,33,6.2,mwc,,,,,1.14,us,usp0009kfj,2015-06-
30T16:59:02.582Z,"Santa Cruz Islands",earthquake,,,,,reviewed,us,hrv
```

Follow the above procedures to reformat the files.

Combine the files (start with the most recent first):

```
seafloor% cat mag_5.5_2000_2015.csv mag_5.5_1960_1999.csv >
mag_5.5_1960_2015_combined.csv
```

```
seafloor% ll
total 7516
-rw-rw-r-- 1 andrewg mgds 2486357 Sep 14 09:23 mag_5.5_1960_1999.csv
-rw-rw-r-- 1 andrewg mgds 3843676 Sep 14 09:25
mag_5.5_1960_2015_combined.csv
-rw-rw-r-- 1 andrewg mgds 1357319 Sep 14 09:24 mag_5.5_2000_2015.csv
```

```
seafloor% wc -l mag_5.5_1960_2015_combined.csv
24227 mag_5.5_1960_2015_combined.csv
```

```
seafloor% head -3 mag_5.5_1960_2015_combined.csv
time,latitude,longitude,depth,mag,magType,nst,gap,dmin,rms,net,id,updated
,place,type,horizontalError,depthError,magError,magNst,status,locationSou
rce,magSource
2015-12-31T10:56:58.120Z,11.1535,-
86.6205,20,5.8,mww,,72,1.097,1.44,us,us10004akg,2016-05-
03T16:41:21.797Z,"71km S of Masachapa,
Nicaragua",earthquake,6,1.8,,,reviewed,us,us
2015-12-29T01:51:41.550Z,-
6.2452,154.7595,30,5.8,mww,,21,3.293,0.87,us,us10004a6n,2016-04-
15T00:59:25.000Z,"80km W of Panguna, Papua New
Guinea",earthquake,7.5,1.8,,,reviewed,us,us
```

Transfer the combined CSV file to the PC and open the file in Excel.
Get rid of some columns that the average user will not require, move the
"mag" and "place" columns left.
Save as a tab-separated file and transfer back to machine seafloor.

```
seafloor% ll *comb*
-rw-rw-r-- 1 andrewg mgds 3843676 Sep 14 09:25
mag_5.5_1960_2015_combined.csv
-rw-rw-r-- 1 andrewg mgds 2068815 Sep 14 2016
mag_5.5_1960_2015_combined_tab.txt
```

```
seafloor% head -3 mag_5.0_1960_2015_combined_tab.txt
time      latitude      longitude      mag      depth      place      id
2015-12-31T10:56:58.120Z      11.1535      -86.6205      5.8      20
"71km S of Masachapa, Nicaragua"      us10004akg
2015-12-29T01:51:41.550Z      -6.2452      154.7595      5.8      30
"80km W of Panguna, Papua New Guinea"      us10004a6n
```

Get rid of all of the header lines from the combined tab-separated file
(we use lots of redundancy)

to ensure that only get rid of those header lines). There are multiple header lines so that the output should have that many rows less than the original combined file:

```
seafloor% cat mag_5.5_1960_2015_combined_tab.txt | grep -v latitude |  
grep -v longitude | grep -v depth > /tmp/a
```

```
seafloor% wc -l mag_5.5_1960_2015_combined_tab.txt /tmp/a  
24227 mag_5.5_1960_2015_combined_tab.txt  
24225 /tmp/a
```

The new temporary file has seven (7) columns:

```
seafloor% head -3 /tmp/a  
2015-12-31T10:56:58.120Z      11.1535 -86.6205      5.8      20  
"71km S of Masachapa, Nicaragua" us10004akg  
2015-12-29T01:51:41.550Z      -6.2452 154.7595      5.8      30  
"80km W of Panguna, Papua New Guinea" us10004a6n  
2015-12-28T06:55:29.880Z      14.6571 -61.3454      5.6      150  
"20km WSW of Saint-Pierre, Martinique" us10004a1v
```

Use gawk to generate a URL using the ID value in the last column:

```
seafloor% gawk -F"\t" '{printf  
"%s\t%s\t%s\t%s\t%s\t%s\t%thttp://earthquake.usgs.gov/earthquakes/eventpage  
/%s\n", $1, $2, $3, $4, $5, $6, $7}' /tmp/a > /tmp/b  
seafloor% head -3 /tmp/b  
2015-12-31T10:56:58.120Z      11.1535 -86.6205      5.8      20  
"71km S of Masachapa, Nicaragua"  
http://earthquake.usgs.gov/earthquakes/eventpage/us10004akg  
2015-12-29T01:51:41.550Z      -6.2452 154.7595      5.8      30  
"80km W of Panguna, Papua New Guinea"  
http://earthquake.usgs.gov/earthquakes/eventpage/us10004a6n  
2015-12-28T06:55:29.880Z      14.6571 -61.3454      5.6      150  
"20km WSW of Saint-Pierre, Martinique"  
http://earthquake.usgs.gov/earthquakes/eventpage/us10004a1v
```

Use sed to get rid of the quotation marks around the location descriptions:

```
seafloor% sed 's/"/g' /tmp/b > /tmp/c  
seafloor% ll /tmp/b /tmp/c  
-rw-rw-r-- 1 andrewg andrewg 3255752 Sep 14 09:32 /tmp/b  
-rw-rw-r-- 1 andrewg andrewg 3233124 Sep 14 09:33 /tmp/c  
  
seafloor% wc -l /tmp/b /tmp/c  
24225 /tmp/b  
24225 /tmp/c
```

```

seafloor% head -3 /tmp/c
2015-12-31T10:56:58.120Z      11.1535 -86.6205      5.8      20
71km S of Masachapa, Nicaragua
http://earthquake.usgs.gov/earthquakes/eventpage/us10004akg
2015-12-29T01:51:41.550Z      -6.2452 154.7595      5.8      30
80km W of Panguna, Papua New Guinea
http://earthquake.usgs.gov/earthquakes/eventpage/us10004a6n
2015-12-28T06:55:29.880Z      14.6571 -61.3454      5.6      150
20km WSW of Saint-Pierre, Martinique
http://earthquake.usgs.gov/earthquakes/eventpage/us10004a1v

```

Finally, for this everything-to-2015 inclusive file, we need to add a header.

Generate a new header row with clearer labels (use a dummy echo command to produce one line of input that drives the gawk command, and remember to make the labels tab-separated):

```

seafloor% echo "hello" | gawk '{printf
"Time\tLatitude\tLongitude\tMagnitude\tDepth\tLocation\tInformation\n"}'
> header
seafloor% cat header
Time      Latitude      Longitude      Magnitude      Depth      Location
Information

```

Combine the header with the body and save it as the base file that we will use for updating the earthquake catalogue:

```
seafloor% cat header /tmp/c > mag_5.5_1960_2015_combined_reformat.txt
```

The new file will have one additional row because of the header.

```

seafloor% ll
total 12704
-rw-rw-r-- 1 andrewg mgds      61 Sep 13 15:30 header
-rw-rw-r-- 1 andrewg mgds 2486357 Sep 14 09:23 mag_5.5_1960_1999.csv
-rw-rw-r-- 1 andrewg mgds 3843676 Sep 14 09:25
mag_5.5_1960_2015_combined.csv
-rw-rw-r-- 1 andrewg mgds 3233185 Sep 14 09:32
mag_5.5_1960_2015_combined_reformat.txt
-rw-rw-r-- 1 andrewg mgds 2068815 Sep 14 09:30
mag_5.5_1960_2015_combined_tab.txt
-rw-rw-r-- 1 andrewg mgds 1357319 Sep 14 09:24 mag_5.5_2000_2015.csv

```

```

seafloor% head mag_5.5_1960_2015_combined_reformat.txt
Time      Latitude      Longitude      Magnitude      Depth      Location
Information
2015-12-31T10:56:58.120Z      11.1535 -86.6205      5.8      20
71km S of Masachapa, Nicaragua
http://earthquake.usgs.gov/earthquakes/eventpage/us10004akg

```


2015-12-29T01:51:41.550Z	-6.2452 154.7595	5.8	30
80km W of Panguna, Papua New Guinea			
http://earthquake.usgs.gov/earthquakes/eventpage/us10004a6n			
2015-12-28T06:55:29.880Z	14.6571 -61.3454	5.6	150
20km WSW of Saint-Pierre, Martinique			
http://earthquake.usgs.gov/earthquakes/eventpage/us10004a1v			
2015-12-25T19:14:47.190Z	36.4935 71.1263 6.3	206	42km WSW
of Ashkasham, Afghanistan			
http://earthquake.usgs.gov/earthquakes/eventpage/us100049i1			
2015-12-25T17:58:04.970Z	-40.7281 -86.4356	5.8	
10 West Chile Rise			
http://earthquake.usgs.gov/earthquakes/eventpage/us100049hf			
2015-12-24T23:10:58.170Z	-7.2182 128.9775	5.8	119
270km WNW of Saumlaki, Indonesia			
http://earthquake.usgs.gov/earthquakes/eventpage/us100049dq			
2015-12-24T19:44:03.130Z	-55.755 -123.1158	6.2	12.28
Southern East Pacific Rise			
http://earthquake.usgs.gov/earthquakes/eventpage/us100049ce			
2015-12-23T16:55:09.460Z	-54.1759 -1.4943	5.8	10
Bouvet Island region			
http://earthquake.usgs.gov/earthquakes/eventpage/us1000493b			
2015-12-21T15:44:04.830Z	-4.4354 151.9782	5.7	162
32km WSW of Kokopo, Papua New Guinea			
http://earthquake.usgs.gov/earthquakes/eventpage/us100048lk			

seafloor% wc -l mag_5.5_1960_2015_combined_reformat.txt
 24226 mag_5.5_1960_2015_combined_reformat.txt

Tests show that the file imports fine into GMA.

When it comes to update the file by adding more recent events to the base catalogue,
 remember that both the magnitude and the location columns were moved leftwards. So,
 make sure that the more recent events have the same column ordering before catting
 the files.

2016-09-13

Generate base (1960-2015 inclusive) catalogue - Magnitudes greater than 6.0

```

-----
--

seafloor% pwd
/home/mgds/ingest/USGS_Earthquake_catalogue/mag6.0

seafloor% ll
total 1224
-rw-rw-r-- 1 andrewg mgds 1250036 Sep 14 2016 mag_6.0_1960_2015.csv

seafloor% head mag_6.0_1960_2015.csv
time,latitude,longitude,depth,mag,magType,nst,gap,dmin,rms,net,id,updated
,place,type,horizontalError,depthError,magError,magNst,status,locationSou
rce,magSource
2015-12-
25T19:14:47.190Z,36.4935,71.1263,206,6.3,mww,,17,0.77,0.91,us,us100049i1,
2016-04-15T00:59:02.000Z,"42km WSW of Ashkasham,
Afghanistan",earthquake,5,1.8,,,reviewed,us,us
2015-12-24T19:44:03.130Z,-55.755,-
123.1158,12.28,6.2,mww,,36,27.97,1.24,us,us100049ce,2016-04-
15T00:58:54.000Z,"Southern East Pacific
Rise",earthquake,12,3.2,,,reviewed,us,us
2015-12-
20T18:47:36.610Z,3.6455,117.6359,14,6.1,mww,,70,1.748,0.84,us,us100048hc,
2016-05-03T16:30:43.029Z,"38km N of Tarakan,
Indonesia",earthquake,5.4,1.7,,,reviewed,us,us

```

Follow the above procedures to reformat the files.

Combine the files (start with the most recent first) (For the base file, there's just one file, but we'll do that cat command so that it is here for when we later update the catalogue):

```

seafloor% cat mag_6.0_1960_2015.csv > mag_6.0_1960_2015_combined.csv

seafloor% ll
total 2448
-rw-rw-r-- 1 andrewg mgds 1250036 Sep 14 12:49
mag_6.0_1960_2015_combined.csv
-rw-rw-r-- 1 andrewg mgds 1250036 Sep 14 12:49 mag_6.0_1960_2015.csv

seafloor% wc -l mag_6.0_1960_2015_combined.csv
7870 mag_6.0_1960_2015_combined.csv

seafloor% head -3 mag_6.0_1960_2015_combined.csv

```

```

time,latitude,longitude,depth,mag,magType,nst,gap,dmin,rms,net,id,updated
,place,type,horizontalError,depthError,magError,magNst,status,locationSou
rce,magSource
2015-12-
25T19:14:47.190Z,36.4935,71.1263,206,6.3,mww,,17,0.77,0.91,us,us100049i1,
2016-04-15T00:59:02.000Z,"42km WSW of Ashkasham,
Afghanistan",earthquake,5,1.8,,,reviewed,us,us
2015-12-24T19:44:03.130Z,-55.755,-
123.1158,12.28,6.2,mww,,36,27.97,1.24,us,us100049ce,2016-04-
15T00:58:54.000Z,"Southern East Pacific
Rise",earthquake,12,3.2,,,reviewed,us,us

```

Transfer the combined CSV file to the PC and open the file in Excel.
 Get rid of some columns that the average user will not require, move the
 "mag" and "place" columns left.
 Save as a tab-separated file and transfer back to machine seafloor.

```

seafloor% ll *comb*
-rw-rw-r-- 1 andrewg mgds 1250036 Sep 14 12:49
mag_6.0_1960_2015_combined.csv
-rw-rw-r-- 1 andrewg mgds 670978 Sep 14 2016
mag_6.0_1960_2015_combined_tab.txt

seafloor% head -3 mag_6.0_1960_2015_combined_tab.txt
time      latitude      longitude      mag      depth      place      id
2015-12-25T19:14:47.190Z      36.4935 71.1263 6.3      206      "42km WSW
of Ashkasham, Afghanistan"      us100049i1
2015-12-24T19:44:03.130Z      -55.755 -123.1158      6.2      12.28
Southern East Pacific Rise      us100049ce

```

Get rid of all of the header lines from the combined tab-separated file
 (we use lots of redundancy
 to ensure that only get rid of those header lines). There are multiple
 header lines so that the
 output should have that many rows less than the original combined file:

```

seafloor% cat mag_6.0_1960_2015_combined_tab.txt | grep -v latitude |
grep -v longitude | grep -v depth > /tmp/a

```

```

seafloor% wc -l mag_6.0_1960_2015_combined_tab.txt /tmp/a
7870 mag_6.0_1960_2015_combined_tab.txt
7869 /tmp/a

```

The new temporary file has seven (7) columns:

```

seafloor% head -3 /tmp/a
2015-12-25T19:14:47.190Z      36.4935 71.1263 6.3      206      "42km WSW
of Ashkasham, Afghanistan"      us100049i1

```

2015-12-24T19:44:03.130Z	-55.755 -123.1158	6.2	12.28
Southern East Pacific Rise	us100049ce		
2015-12-20T18:47:36.610Z	3.6455 117.6359	6.1	14
"38km N of Tarakan, Indonesia"	us100048hc		

Use gawk to generate a URL using the ID value in the last column:

```
seafloor% gawk -F"\t" '{printf
"%s\t%s\t%s\t%s\t%s\t%s\t%thttp://earthquake.usgs.gov/earthquakes/eventpage
/%s\n", $1, $2, $3, $4, $5, $6, $7}' /tmp/a > /tmp/b
seafloor% head -3 /tmp/b
2015-12-25T19:14:47.190Z      36.4935 71.1263 6.3      206      "42km WSW
of Ashkasham, Afghanistan"
http://earthquake.usgs.gov/earthquakes/eventpage/us100049i1
2015-12-24T19:44:03.130Z      -55.755 -123.1158      6.2      12.28
Southern East Pacific Rise
http://earthquake.usgs.gov/earthquakes/eventpage/us100049ce
2015-12-20T18:47:36.610Z      3.6455 117.6359      6.1      14
"38km N of Tarakan, Indonesia"
http://earthquake.usgs.gov/earthquakes/eventpage/us100048hc
```

Use sed to get rid of the quotation marks around the location descriptions:

```
seafloor% sed 's/\\"//g' /tmp/b > /tmp/c
seafloor% ll /tmp/b /tmp/c
-rw-rw-r-- 1 andrewg andrewg 1056515 Sep 14 12:55 /tmp/b
-rw-rw-r-- 1 andrewg andrewg 1048925 Sep 14 12:56 /tmp/c
```

```
seafloor% wc -l /tmp/b /tmp/c
  7869 /tmp/b
  7869 /tmp/c
 15738 total
```

```
seafloor% head -3 /tmp/c
2015-12-25T19:14:47.190Z      36.4935 71.1263 6.3      206      42km WSW
of Ashkasham, Afghanistan
http://earthquake.usgs.gov/earthquakes/eventpage/us100049i1
2015-12-24T19:44:03.130Z      -55.755 -123.1158      6.2      12.28
Southern East Pacific Rise
http://earthquake.usgs.gov/earthquakes/eventpage/us100049ce
2015-12-20T18:47:36.610Z      3.6455 117.6359      6.1      14
38km N of Tarakan, Indonesia
http://earthquake.usgs.gov/earthquakes/eventpage/us100048hc
```

Finally, for this everything-to-2015 inclusive file, we need to add a header.

Generate a new header row with clearer labels (use a dummy echo command to produce one line of input that drives the gawk command, and remember to make the labels tab-separated):

```
seafloor% echo "hello" | gawk '{printf
"Time\tLatitude\tLongitude\tMagnitude\tDepth\tLocation\tInformation\n"}'
> header
seafloor% cat header
Time      Latitude      Longitude      Magnitude      Depth      Location
Information
```

Combine the header with the body and save it as the base file that we will use for updating the earthquake catalogue:

```
seafloor% cat header /tmp/c > mag_6.0_1960_2015_combined_reformat.txt
```

The new file will have one additional row because of the header.

```
seafloor% ll
total 4136
-rw-rw-r-- 1 andrewg mgds      61 Sep 13 15:30 header
-rw-rw-r-- 1 andrewg mgds 1250036 Sep 14 12:49
mag_6.0_1960_2015_combined.csv
-rw-rw-r-- 1 andrewg mgds 1048986 Sep 14 12:56
mag_6.0_1960_2015_combined_reformat.txt
-rw-rw-r-- 1 andrewg mgds  670978 Sep 14 12:54
mag_6.0_1960_2015_combined_tab.txt
-rw-rw-r-- 1 andrewg mgds 1250036 Sep 14 12:49 mag_6.0_1960_2015.csv

seafloor% head mag_6.0_1960_2015_combined_reformat.txt
Time      Latitude      Longitude      Magnitude      Depth      Location
Information
2015-12-25T19:14:47.190Z      36.4935 71.1263 6.3      206      42km WSW
of Ashkasham, Afghanistan
http://earthquake.usgs.gov/earthquakes/eventpage/us100049i1
2015-12-24T19:44:03.130Z      -55.755 -123.1158      6.2      12.28
Southern East Pacific Rise
http://earthquake.usgs.gov/earthquakes/eventpage/us100049ce
2015-12-20T18:47:36.610Z      3.6455 117.6359      6.1      14
38km N of Tarakan, Indonesia
http://earthquake.usgs.gov/earthquakes/eventpage/us100048hc
2015-12-19T02:10:53.360Z      -18.3819      169.3857      6
10      129km N of Isangel, Vanuatu
http://earthquake.usgs.gov/earthquakes/eventpage/us1000489i
2015-12-17T19:49:53.050Z      15.8015 -93.633 6.6      85      12km SW
of Tres Picos, Mexico
http://earthquake.usgs.gov/earthquakes/eventpage/us100047zq
2015-12-09T12:58:01.780Z      -16.7374      175.2475      6.1
10      255km WNW of Lautoka, Fiji
http://earthquake.usgs.gov/earthquakes/eventpage/us20004fuz
```

```

2015-12-09T10:21:48.530Z      -4.1064 129.5079      6.9      21
107km SE of Amahai, Indonesia
http://earthquake.usgs.gov/earthquakes/eventpage/us20004ft7
2015-12-07T07:50:05.950Z      38.2107 72.7797 7.2      22      104km W
of Murghob, Tajikistan
http://earthquake.usgs.gov/earthquakes/eventpage/us100044k6
2015-12-04T22:25:00.110Z      -47.6165      85.0913 7.1      35
Southeast Indian Ridge
http://earthquake.usgs.gov/earthquakes/eventpage/us100043z2

```

```

seafloor% wc -l mag_6.0_1960_2015_combined_reformat.txt
7870 mag_6.0_1960_2015_combined_reformat.txt

```

Tests show that the file imports fine into GMA.

When it comes to update the file by adding more recent event sto the base catalogue,
remember that both the magnitude and the location columns were moved leftwards. So,
make sure that the more recent events have the same column ordering before catting
the files.

```

*****
*****

```

```

INSERT INTO gma_menu
(name,data_type,type,projection,command,info_href,url,remote,dms_project,
data_refresh_date) VALUES ('All Earthquakes, Magnitude > 6.0 (1960-
2016)','Table','2','1111','table_cmd','http://earthquake.usgs.gov/earthqu
akes/','http://app.geomapapp.org/data/datalayers/geophysics/earthquakes/U
SGS_ANSS/mag_6.0_1960_2016_combined_reformat.txt','f','MGG','2016-09-
14');
INSERT INTO gma_menu
(name,data_type,type,projection,command,info_href,url,remote,dms_project,
data_refresh_date) VALUES ('All Earthquakes, Magnitude > 5.5 (1960-
2016)','Table','2','1111','table_cmd','http://earthquake.usgs.gov/earthqu
akes/','http://app.geomapapp.org/data/datalayers/geophysics/earthquakes/U
SGS_ANSS/mag_5.5_1960_2016_combined_reformat.txt','f','MGG','2016-09-
14');
INSERT INTO gma_menu
(name,data_type,type,projection,command,info_href,url,remote,dms_project,
data_refresh_date) VALUES ('All Earthquakes, Magnitude > 5.0 (1960-
2016)','Table','2','1111','table_cmd','http://earthquake.usgs.gov/earthqu
akes/','http://app.geomapapp.org/data/datalayers/geophysics/earthquakes/U
SGS_ANSS/mag_5.0_1960_2016_combined_reformat.txt','f','MGG','2016-09-
14');
INSERT INTO gma_menu
(name,data_type,type,projection,command,info_href,url,remote,dms_project,

```

```
data_refresh_date) VALUES ('All Earthquakes, Magnitude > 4.5 (1960-
2016)', 'Table', '2', '1111', 'table_cmd', 'http://earthquake.usgs.gov/earthqu
akes/', 'http://app.geomapapp.org/data/datalayers/geophysics/earthquakes/U
SGS_ANSS/mag_4.5_1960_2016_combined_reformat.txt', 'f', 'MGG', '2016-09-
14');
```

```
INSERT INTO gma_menu_map
(node_uid,parent_node_uid,dms_project,item_order) VALUES
('20796','20512','MGG','10');
INSERT INTO gma_menu_map
(node_uid,parent_node_uid,dms_project,item_order) VALUES
('20797','20512','MGG','15');
INSERT INTO gma_menu_map
(node_uid,parent_node_uid,dms_project,item_order) VALUES
('20798','20512','MGG','20');
INSERT INTO gma_menu_map
(node_uid,parent_node_uid,dms_project,item_order) VALUES
('20799','20512','MGG','25');
```

```
*****
*****
```

2017-11-03

Update the catalogue - Magnitudes greater than 4.5

Above, we created "base catalogue" files for each magnitude step for the period 1960-2015 inclusive. To update the catalogues, we need to download the latest data. For example, for this update, we created two new files that will be catted to the base catalogue: a file for the whole of 2016 and a file for most of 2017. After the end of each calendar year, we can create a whole-year file for that year.

USGS search page: <http://earthquake.usgs.gov/earthquakes/search/>

First, note that the search results are limited to 20,000 records and the search will return an error if greater. So, it may be necessary to download events in time-windowed chunks.

Choose magnitude (e.g. 4.5), date range, (e.g. 2016-01-01 00:00:00 to 2016-12-31 23:59:59), Global, CSV format, Order by Time-Newest First. Hit Search.

The name of the downloaded CVS file is always "query.cvs". Rename it right away to avoid confusion (e.g. to "mag_4.5_2016_2016.csv" for the whole year file or to "mag_4.5_2017_20171031.csv" for the partial year file).

```
seafloor% pwd
/home/mgds/ingest/USGS_Earthquake_catalogue/mag4.5
```

Base file:

```
seafloor% ll
-rw-rw-r-- 1 andrewg mgds 29066516 Sep 13 2016
mag_4.5_1960_2015_combined_reformat.txt
```

New files since 2015 - not in final format (note that as we routinely update these catalogues, we'll generate a set of yearly files that will not change - only the partial year file will likely change):

```
seafloor% ll
-rw-rw-r-- 1 andrewg mgds 1372191 Nov 1 14:33 mag_4.5_2016_2016.csv
-rw-rw-r-- 1 andrewg mgds 918169 Nov 1 14:38 mag_4.5_2017_20171031.csv
```

Combine the **new files since 2015** starting with the most recent new file first:

```
seafloor% cat mag_4.5_2017_20171031.csv mag_4.5_2016_2016.csv >
mag_4.5_2016_20171031_combined.csv
```

```
seafloor% ll mag_4.5_2016_20171031_combined.csv
-rw-rw-r-- 1 andrewg mgds 2290360 Nov 1 15:13
mag_4.5_2016_20171031_combined.csv
```

```
seafloor% wc -l mag_4.5_2016_20171031_combined.csv
12386 mag_4.5_2016_20171031_combined.csv
```

```
seafloor% head -3 mag_4.5_2016_20171031_combined.csv
time,latitude,longitude,depth,mag,magType,nst,gap,dmin,rms,net,id,updated
,place,type,horizontalError,depthError,magError,magNst,status,locationSource,magSource
2017-10-31T23:59:47.120Z,-
21.703,168.9225,10,5,mb,,92,0.859,0.86,us,us1000azhz,2017-11-
01T00:15:47.040Z,"108km E of Tadine, New
Caledonia",earthquake,5.3,1.2,0.098,33,reviewed,us,us
```



```
2017-10-31T23:18:05.120Z,-
7.3754,155.8921,35,5,mb,,38,4.505,1.02,us,us1000azgy,2017-11-
01T03:42:22.040Z,"109km SW of Chirovanga, Solomon
Islands",earthquake,8.5,2,0.066,73,reviewed,us,us
```

Transfer the combined CSV file to the PC and open the file in Excel.
Get rid of all columns except for these 7 columns:
time,latitude,longitude,depth,mag,id,place
Move the "mag" column left.
Move the "place" column left.
Save as a tab-separated file ("xxxxx_combined_tab.txt") and transfer back
to machine seafloor.

The new column order looks like this:

time	latitude	longitude	mag	depth	place	id
------	----------	-----------	-----	-------	-------	----

Check that the columns are in this order before proceeding otherwise when
we
cat this new file to the base catalogue file, the columns will not match.
Make sure that the mag and depth columns are in the right order!

```
seafloor% ll *comb*
-rw-rw-r-- 1 andrewg mgds 34743789 Sep 13 2016
mag_4.5_1960_2015_combined.csv
-rw-rw-r-- 1 andrewg mgds 29066516 Sep 13 2016
mag_4.5_1960_2015_combined_reformat.txt
-rw-rw-r-- 1 andrewg mgds 18641979 Sep 13 2016
mag_4.5_1960_2015_combined_tab.txt
-rw-rw-r-- 1 andrewg mgds 2290360 Nov 1 15:14
mag_4.5_2016_20171031_combined.csv
-rw-rw-r-- 1 andrewg mgds 1199375 Nov 3 15:21
mag_4.5_2016_20171031_combined_tab.txt
```

```
seafloor% head -3 mag_4.5_2016_20171031_combined_tab.txt
time      latitude      longitude      mag      depth      place      id
2017-10-31T23:59:47.120Z      -21.703 168.9225      5      10
"108km E of Tadine, New Caledonia"      us1000azhz
2017-10-31T23:18:05.120Z      -7.3754 155.8921      5      35
"109km SW of Chirovanga, Solomon Islands"      us1000azgy
```

Get rid of all of the header lines from the combined tab-separated file.
We use lots
of redundancy to ensure that we only get rid of those header lines:

```
seafloor% cat mag_4.5_2016_20171031_combined_tab.txt | grep -v latitude |
grep -v longitude | grep -v depth > /tmp/a
```

```
seafloor% wc -l mag_4.5_2016_20171031_combined_tab.txt /tmp/a
12386 mag_4.5_2016_20171031_combined_tab.txt
12384 /tmp/a
```

The new temporary file has seven (7) columns:

```
seafloor% head -3 /tmp/a
2017-10-31T23:59:47.120Z      -21.703 168.9225      5      10
"108km E of Tadine, New Caledonia"      us1000azhz
2017-10-31T23:18:05.120Z      -7.3754 155.8921      5      35
"109km SW of Chirovanga, Solomon Islands"      us1000azgy
2017-10-31T23:06:04.110Z      -21.6209      168.822 5.1      12.98
"97km E of Tadine, New Caledonia"      us1000azg3
```

Use gawk to generate a USGS URL using the ID value in the last column:

```
seafloor% gawk -F"\t" '{printf
"%s\t%s\t%s\t%s\t%s\t%s\t%thttp://earthquake.usgs.gov/earthquakes/eventpage
/%s\n", $1, $2, $3, $4, $5, $6, $7}' /tmp/a > /tmp/b
seafloor% head -3 /tmp/b
2017-10-31T23:59:47.120Z      -21.703 168.9225      5      10
"108km E of Tadine, New Caledonia"
http://earthquake.usgs.gov/earthquakes/eventpage/us1000azhz
2017-10-31T23:18:05.120Z      -7.3754 155.8921      5      35
"109km SW of Chirovanga, Solomon Islands"
http://earthquake.usgs.gov/earthquakes/eventpage/us1000azgy
2017-10-31T23:06:04.110Z      -21.6209      168.822 5.1      12.98
"97km E of Tadine, New Caledonia"
http://earthquake.usgs.gov/earthquakes/eventpage/us1000azg3
```

Use sed to get rid of the quotation marks around the location descriptions:

```
seafloor% sed 's/\\"//g' /tmp/b > /tmp/c
seafloor% ll /tmp/b /tmp/c
-rw-rw-r-- 1 andrewg andrewg 1806103 Nov  3 17:21 /tmp/b
-rw-rw-r-- 1 andrewg andrewg 1784919 Nov  3 17:21 /tmp/c
```

```
seafloor% wc -l /tmp/b /tmp/c
12384 /tmp/b
12384 /tmp/c
```

```
seafloor% head -3 /tmp/c
2017-10-31T23:59:47.120Z      -21.703 168.9225      5      10
108km E of Tadine, New Caledonia
http://earthquake.usgs.gov/earthquakes/eventpage/us1000azhz
```

```

2017-10-31T23:18:05.120Z      -7.3754 155.8921      5      35
109km SW of Chirovanga, Solomon Islands
http://earthquake.usgs.gov/earthquakes/eventpage/us1000azgy
2017-10-31T23:06:04.110Z      -21.6209      168.822 5.1      12.98
97km E of Tadine, New Caledonia
http://earthquake.usgs.gov/earthquakes/eventpage/us1000azg3

```

This file is now ready to be catted (prepended) with the 1960-2015 inclusive base catalogue file.
But, before we cat them two files together, we need to strip the header off the base catalogue file.
Then, we cat the two headerless files together (remember to put the newest file first!).
Finally, add the header back at the start:

```

seafloor% tail -n +2 mag_4.5_1960_2015_combined_reformat.txt > /tmp/base
seafloor% cat /tmp/c /tmp/base > /tmp/d
seafloor% cat header /tmp/d > mag_4.5_1960_20171031_combined_reformat.txt

```

```

seafloor% ll
total 178664
-rw-rw-r-- 1 andrewg mgds      61 Sep 13  2016 header
-rw-rw-r-- 1 andrewg mgds 485324 Sep 12  2016 mag_4.5_1960_1969.csv
-rw-rw-r-- 1 andrewg mgds 34743789 Sep 13  2016
mag_4.5_1960_2015_combined.csv
-rw-rw-r-- 1 andrewg mgds 29066516 Sep 13  2016
mag_4.5_1960_2015_combined_reformat.txt
-rw-rw-r-- 1 andrewg mgds 29066516 Sep 13  2016
mag_4.5_1960_2015_combined_reformat.txt.copy_do_not_delete
-rw-rw-r-- 1 andrewg mgds 18641979 Sep 13  2016
mag_4.5_1960_2015_combined_tab.txt
-rw-rw-r-- 1 andrewg mgds 30851435 Nov  3 17:22
mag_4.5_1960_20171031_combined_reformat.txt
-rw-rw-r-- 1 andrewg mgds 1089420 Sep 12  2016 mag_4.5_1970_1974.csv
-rw-rw-r-- 1 andrewg mgds 2546243 Sep 12  2016 mag_4.5_1975_1979.csv
-rw-rw-r-- 1 andrewg mgds 2851504 Sep 12  2016 mag_4.5_1980_1984.csv
-rw-rw-r-- 1 andrewg mgds 2035377 Sep 12  2016 mag_4.5_1985_1987.csv
-rw-rw-r-- 1 andrewg mgds 2557184 Sep 12  2016 mag_4.5_1988_1991.csv
-rw-rw-r-- 1 andrewg mgds 2700811 Sep 12  2016 mag_4.5_1992_1995.csv
-rw-rw-r-- 1 andrewg mgds 2661103 Sep 12  2016 mag_4.5_1996_1999.csv
-rw-rw-r-- 1 andrewg mgds 3150193 Sep 12  2016 mag_4.5_2000_2003.csv
-rw-rw-r-- 1 andrewg mgds 2296676 Sep 12  2016 mag_4.5_2004_2005.csv
-rw-rw-r-- 1 andrewg mgds 2329147 Sep 12  2016 mag_4.5_2006_2007.csv
-rw-rw-r-- 1 andrewg mgds 2016415 Sep 12  2016 mag_4.5_2008_2009.csv
-rw-rw-r-- 1 andrewg mgds 2784697 Sep 12  2016 mag_4.5_2010_2011.csv
-rw-rw-r-- 1 andrewg mgds 2417590 Sep 12  2016 mag_4.5_2012_2013.csv
-rw-rw-r-- 1 andrewg mgds 2822105 Sep 12  2016 mag_4.5_2014_2015.csv
-rw-rw-r-- 1 andrewg mgds 1372191 Nov  1 14:33 mag_4.5_2016_2016.csv
-rw-rw-r-- 1 andrewg mgds 2290360 Nov  1 15:14
mag_4.5_2016_20171031_combined.csv

```

```
-rw-rw-r-- 1 andrewg mgds 1199375 Nov  3 15:21
mag_4.5_2016_20171031_combined_tab.txt
-rw-rw-r-- 1 andrewg mgds  918169 Nov  1 14:38 mag_4.5_2017_20171031.csv
```

```
seafloor% head mag_4.5_1960_20171031_combined_reformat.txt
Time      Latitude      Longitude      Magnitude      Depth      Location
Information
2017-10-31T23:59:47.120Z      -21.703 168.9225      5      10
108km E of Tadine, New Caledonia
http://earthquake.usgs.gov/earthquakes/eventpage/us1000azhz
2017-10-31T23:18:05.120Z      -7.3754 155.8921      5      35
109km SW of Chirovanga, Solomon Islands
http://earthquake.usgs.gov/earthquakes/eventpage/us1000azgy
2017-10-31T23:06:04.110Z      -21.6209      168.822 5.1      12.98
97km E of Tadine, New Caledonia
http://earthquake.usgs.gov/earthquakes/eventpage/us1000azg3
2017-10-31T22:59:52.370Z      -37.6252      177.3512      5
10      49km NE of Whakatane, New Zealand
http://earthquake.usgs.gov/earthquakes/eventpage/us1000azft
2017-10-31T22:50:16.450Z      -21.7537      168.821 5.3      10
99km ESE of Tadine, New Caledonia
http://earthquake.usgs.gov/earthquakes/eventpage/us1000azfp
2017-10-31T22:29:05.730Z      -21.7262      168.8365      4.7
11.89      100km ESE of Tadine, New Caledonia
http://earthquake.usgs.gov/earthquakes/eventpage/us1000azez
2017-10-31T22:26:03.150Z      -60.1029      -26.7335      5.4
101.59      119km S of Bristol Island, South Sandwich Islands
http://earthquake.usgs.gov/earthquakes/eventpage/us1000azey
2017-10-31T21:10:15.830Z      -15.6996      -173.3072      4.9
10      54km ENE of Hihifo, Tonga
http://earthquake.usgs.gov/earthquakes/eventpage/us1000azcz
2017-10-31T20:22:44.330Z      -21.7719      168.9786      5
7.31      115km ESE of Tadine, New Caledonia
http://earthquake.usgs.gov/earthquakes/eventpage/us1000azbm
```

```
seafloor% wc -l mag_4.5_1960_20171031_combined_reformat.txt
229678 mag_4.5_1960_20171031_combined_reformat.txt
```

Check that the file imports in GMA.

Update GMA

Copy the file to the GMA data directory:

/data/mgds/web/app.geomapapp.org/htdocs/data/datalayers/geophysics/earthquakes/USGS_ANSS

In the gma_menu table, for magnitude ≥ 4.5 , node_uid = 20799.

Update these fields:

- name (to reflect the new date range)
- url (to point to the new file)
- data_refresh_date (gives the update date)

Remake the menu files, test, and ask Ed to rsynch to the commercial provider.

When it comes to update the file by adding more recent events to the base catalogue, remember that both the magnitude and the place columns were moved leftwards. So, make sure that the more recent events have the same column ordering before catting the files.

2017-11-03

Update the catalogue - Magnitudes greater than 5.0

Above, we created "base catalogue" files for each magnitude step for the period 1960-2015 inclusive. To update the catalogues, we need to download the latest data. For example, for this update, we created two new files that will be catted to the base catalogue: a file for the whole of 2016 and a file for most of 2017. After the end of each calendar year, we can create a whole-year file for that year.

USGS search page: <http://earthquake.usgs.gov/earthquakes/search/>

First, note that the search results are limited to 20,000 records and the search will return an error if greater. So, it may be necessary to download events in time-windowed chunks.

Choose magnitude (e.g. 4.5), date range, (e.g. 2016-01-01 00:00:00 to 2016-12-31 23:59:59), Global, CSV format, Order by Time-Newest First. Hit Search.

The name of the downloaded CVS file is always "query.csv". Rename it right away to avoid confusion (e.g. to "mag_4.5_2016_2016.csv" for the whole year file or to "mag_4.5_2017_20171031.csv" for the partial year file).

```
seafloor% pwd
/home/mgds/ingest/USGS_Earthquake_catalogue/mag5.0
```

Base file:

```
seafloor% ll
-rw-rw-r-- 1 andrewg mgds 9916656 Sep 13 2016
mag_5.0_1960_2015_combined_reformat.txt
```

New files since 2015 - not in final format (note that as we routinely update these catalogues, we'll generate a set of yearly files that will not change - only the partial year file will likely change):

```
seafloor% ll
-rw-rw-r-- 1 andrewg mgds 309591 Nov 3 17:32 mag_5.0_2016_2016.csv
-rw-rw-r-- 1 andrewg mgds 235584 Nov 3 17:36 mag_5.0_2017_20171031.csv
```

Combine the **new files since 2015** (start with the most recent new file first):

```
seafloor% cat mag_5.0_2017_20171031.csv mag_5.0_2016_2016.csv >
mag_5.0_2016_20171031_combined.csv
```

```
seafloor% ll mag_5.0_2016_20171031_combined.csv
-rw-rw-r-- 1 andrewg mgds 545175 Nov 3 17:38
mag_5.0_2016_20171031_combined.csv
```

```
seafloor% wc -l mag_5.0_2016_20171031_combined.csv
2974 mag_5.0_2016_20171031_combined.csv
```

```
seafloor% head -3 mag_5.0_2016_20171031_combined.csv
time,latitude,longitude,depth,mag,magType,nst,gap,dmin,rms,net,id,updated
,place,type,horizontalError,depthError,magError,magNst,status,locationSource,magSource
2017-10-31T23:59:47.120Z,-
21.703,168.9225,10,5,mb,,92,0.859,0.86,us,us1000azhz,2017-11-
01T00:15:47.040Z,"108km E of Tadine, New
Caledonia",earthquake,5.3,1.2,0.098,33,reviewed,us,us
2017-10-31T23:18:05.120Z,-
7.3754,155.8921,35,5,mb,,38,4.505,1.02,us,us1000azgy,2017-11-
01T03:42:22.040Z,"109km SW of Chirovanga, Solomon
Islands",earthquake,8.5,2,0.066,73,reviewed,us,us
```

Transfer the combined CSV file to the PC and open the file in Excel.
 Get rid of all columns except for these 7 columns:
 time,latitude,longitude,depth,mag,id,place
 Move the "mag" column left.
 Move the "place" column left.
 Save as a tab-separated file ("xxxxx_combined_tab.txt") and transfer back
 to machine seafloor.

The new column order looks like this:

```
time      latitude      longitude      mag      depth      place      id
```

Check that the columns are in this order before proceeding otherwise when
 we
 cat this new file to the base catalogue file, the columns will not match.
 Make sure that the map and depth columns are in the right order!

```
seafloor% ll *comb*
-rw-rw-r-- 1 andrewg mgds 11784011 Sep 13 2016
mag_5.0_1960_2015_combined.csv
-rw-rw-r-- 1 andrewg mgds 9916656 Sep 13 2016
mag_5.0_1960_2015_combined_reformat.txt
-rw-rw-r-- 1 andrewg mgds 6338690 Sep 13 2016
mag_5.0_1960_2015_combined_tab.txt
-rw-rw-r-- 1 andrewg mgds 545175 Nov 3 17:38
mag_5.0_2016_20171031_combined.csv
-rw-rw-r-- 1 andrewg mgds 286191 Nov 3 17:41
mag_5.0_2016_20171031_combined_tab.txt
```

```
seafloor% head -3 mag_5.0_2016_20171031_combined_tab.txt
time      latitude      longitude      mag      depth      place      id
2017-10-31T23:59:47.120Z      -21.703 168.9225      5      10
"108km E of Tadine, New Caledonia"      us1000azhz
2017-10-31T23:18:05.120Z      -7.3754 155.8921      5      35
"109km SW of Chirovanga, Solomon Islands"      us1000azgy
```

Get rid of all of the header lines from the combined tab-separated file.
We use lots
of redundancy to ensure that we only get rid of those header lines:

```
seafloor% cat mag_5.0_2016_20171031_combined_tab.txt | grep -v latitude |  
grep -v longitude | grep -v depth > /tmp/a
```

```
seafloor% wc -l mag_5.0_2016_20171031_combined_tab.txt /tmp/a  
2974 mag_5.0_2016_20171031_combined_tab.txt  
2972 /tmp/a
```

The new temporary file has seven (7) columns:

```
seafloor% head -3 /tmp/a  
2017-10-31T23:59:47.120Z      -21.703 168.9225      5      10  
"108km E of Tadine, New Caledonia"      us1000azhz  
2017-10-31T23:18:05.120Z      -7.3754 155.8921      5      35  
"109km SW of Chirovanga, Solomon Islands"      us1000azgy  
2017-10-31T23:06:04.110Z      -21.6209      168.822 5.1      12.98  
"97km E of Tadine, New Caledonia"      us1000azg3
```

Use gawk to generate a USGS URL using the ID value in the last column:

```
seafloor% gawk -F"\t" '{printf  
"%s\t%s\t%s\t%s\t%s\t%s\t%thttp://earthquake.usgs.gov/earthquakes/eventpage  
/%s\n", $1, $2, $3, $4, $5, $6, $7}' /tmp/a > /tmp/b  
seafloor% head -3 /tmp/b  
2017-10-31T23:59:47.120Z      -21.703 168.9225      5      10  
"108km E of Tadine, New Caledonia"  
http://earthquake.usgs.gov/earthquakes/eventpage/us1000azhz  
2017-10-31T23:18:05.120Z      -7.3754 155.8921      5      35  
"109km SW of Chirovanga, Solomon Islands"  
http://earthquake.usgs.gov/earthquakes/eventpage/us1000azgy  
2017-10-31T23:06:04.110Z      -21.6209      168.822 5.1      12.98  
"97km E of Tadine, New Caledonia"  
http://earthquake.usgs.gov/earthquakes/eventpage/us1000azg3
```

Use sed to get rid of the quotation marks around the location
descriptions:

```
seafloor% sed 's/"/g' /tmp/b > /tmp/c  
seafloor% ll /tmp/b /tmp/c  
-rw-rw-r-- 1 andrewg andrewg 431731 Nov  3 17:44 /tmp/b  
-rw-rw-r-- 1 andrewg andrewg 426643 Nov  3 17:44 /tmp/c  
  
seafloor% wc -l /tmp/b /tmp/c
```



```
2972 /tmp/b
2972 /tmp/c
```

```
seafloor% head -3 /tmp/c
2017-10-31T23:59:47.120Z      -21.703 168.9225      5      10
108km E of Tadine, New Caledonia
http://earthquake.usgs.gov/earthquakes/eventpage/us1000azhz
2017-10-31T23:18:05.120Z      -7.3754 155.8921      5      35
109km SW of Chirovanga, Solomon Islands
http://earthquake.usgs.gov/earthquakes/eventpage/us1000azgy
2017-10-31T23:06:04.110Z      -21.6209      168.822 5.1      12.98
97km E of Tadine, New Caledonia
http://earthquake.usgs.gov/earthquakes/eventpage/us1000azg3
```

This file is now ready to be catted (prepended) with the 1960-2015 inclusive base catalogue file.
But, before we cat them two files together, we need to strip the header off the base catalogue file.
Then, we cat the two headerless files together (remember to put the newest file first!).
Finally, add the header back at the start:

```
seafloor% tail -n +2 mag_5.0_1960_2015_combined_reformat.txt > /tmp/base
seafloor% cat /tmp/c /tmp/base > /tmp/d
seafloor% cat header /tmp/d > mag_5.0_1960_20171031_combined_reformat.txt
```

```
seafloor% ll
total 50364
-rw-rw-r-- 1 andrewg mgds      61 Sep 13  2016 header
-rw-rw-r-- 1 andrewg mgds 2291450 Sep 13  2016 mag_5.0_1960_1979.csv
-rw-rw-r-- 1 andrewg mgds 11784011 Sep 13  2016
mag_5.0_1960_2015_combined.csv
-rw-rw-r-- 1 andrewg mgds 9916656 Sep 13  2016
mag_5.0_1960_2015_combined_reformat.txt
-rw-rw-r-- 1 andrewg mgds 6338690 Sep 13  2016
mag_5.0_1960_2015_combined_tab.txt
-rw-rw-r-- 1 andrewg mgds 10343299 Nov  3 17:46
mag_5.0_1960_20171031_combined_reformat.txt
-rw-rw-r-- 1 andrewg mgds 2463542 Sep 13  2016 mag_5.0_1980_1989.csv
-rw-rw-r-- 1 andrewg mgds 2262013 Sep 13  2016 mag_5.0_1990_1999.csv
-rw-rw-r-- 1 andrewg mgds 2784928 Sep 13  2016 mag_5.0_2000_2009.csv
-rw-rw-r-- 1 andrewg mgds 1982078 Sep 13  2016 mag_5.0_2010_2015.csv
-rw-rw-r-- 1 andrewg mgds 309591 Nov  3 17:32 mag_5.0_2016_2016.csv
-rw-rw-r-- 1 andrewg mgds 545175 Nov  3 17:38
mag_5.0_2016_20171031_combined.csv
-rw-rw-r-- 1 andrewg mgds 286191 Nov  3 17:41
mag_5.0_2016_20171031_combined_tab.txt
-rw-rw-r-- 1 andrewg mgds 235584 Nov  3 17:36 mag_5.0_2017_20171031.csv
```

```

seafloor% head mag_5.0_1960_20171031_combined_reformat.txt
Time      Latitude      Longitude      Magnitude      Depth      Location
Information
2017-10-31T23:59:47.120Z      -21.703 168.9225      5      10
108km E of Tadine, New Caledonia
http://earthquake.usgs.gov/earthquakes/eventpage/us1000azhz
2017-10-31T23:18:05.120Z      -7.3754 155.8921      5      35
109km SW of Chirovanga, Solomon Islands
http://earthquake.usgs.gov/earthquakes/eventpage/us1000azgy
2017-10-31T23:06:04.110Z      -21.6209      168.822 5.1      12.98
97km E of Tadine, New Caledonia
http://earthquake.usgs.gov/earthquakes/eventpage/us1000azg3
2017-10-31T22:59:52.370Z      -37.6252      177.3512      5
10      49km NE of Whakatane, New Zealand
http://earthquake.usgs.gov/earthquakes/eventpage/us1000azft
2017-10-31T22:50:16.450Z      -21.7537      168.821 5.3      10
99km ESE of Tadine, New Caledonia
http://earthquake.usgs.gov/earthquakes/eventpage/us1000azfp
2017-10-31T22:26:03.150Z      -60.1029      -26.7335      5.4
101.59 119km S of Bristol Island, South Sandwich Islands
http://earthquake.usgs.gov/earthquakes/eventpage/us1000azey
2017-10-31T20:22:44.330Z      -21.7719      168.9786      5
7.31 115km ESE of Tadine, New Caledonia
http://earthquake.usgs.gov/earthquakes/eventpage/us1000azbm
2017-10-31T19:41:16.220Z      -21.7586      168.898 5.2      7.38
107km ESE of Tadine, New Caledonia
http://earthquake.usgs.gov/earthquakes/eventpage/us1000az9c
2017-10-31T19:22:09.830Z      -21.679 168.8173      5.4      21.1
97km E of Tadine, New Caledonia
http://earthquake.usgs.gov/earthquakes/eventpage/us1000az8y

seafloor% wc -l mag_5.0_1960_20171031_combined_reformat.txt
77456 mag_5.0_1960_20171031_combined_reformat.txt

```

Check that the file imports in GMA.

Update GMA

Copy the file to the GMA data directory:

```

/data/mgds/web/app.geomapapp.org/htdocs/data/datalayers/geophysics/earthq
uakes/USGS_ANSS

```

In the gma_menu table, for magnitude >= 5.0, node_uid = 20798.

Update these fields:

- name (to reflect the new date range)
- url (to point to the new file)

- data_refresh_date (gives the update date)

Remake the menu files, test, and ask Ed to rsynch to the commercial provider.

When it comes to update the file by adding more recent events to the base catalogue, remember that both the magnitude and the place columns were moved leftwards. So, make sure that the more recent events have the same column ordering before catting the files.

2017-11-03

Update the catalogue - Magnitudes greater than 5.5

Above, we created "base catalogue" files for each magnitude step for the period 1960-2015 inclusive. To update the catalogues, we need to download the latest data. For example, for this update, we created two new files that will be catted to the base catalogue: a file for the whole of 2016 and a file for most of 2017. After the end of each calendar year, we can create a whole-year file for that year.

USGS search page: <http://earthquake.usgs.gov/earthquakes/search/>

First, note that the search results are limited to 20,000 records and the search will return an error if greater. So, it may be necessary to download events in time-windowed chunks.

Choose magnitude (e.g. 4.5), date range, (e.g. 2016-01-01 00:00:00 to 2016-12-31 23:59:59), Global, CSV format, Order by Time-Newest First.

Hit Search.

The name of the downloaded CVS file is always "query.cvs". Rename it right away to avoid confusion (e.g. to "mag_4.5_2016_2016.csv" for the whole year file or to "mag_4.5_2017_20171031.csv" for the partial year file).

```
seafloor% pwd
/home/mgds/ingest/USGS_Earthquake_catalogue/mag5.5
```

Base file:

```
seafloor% ll
-rw-rw-r-- 1 andrewg mgds 3233185 Sep 14 2016
mag_5.5_1960_2015_combined_reformat.txt
```

New files since 2015 - not in final format (note that as we routinely update these catalogues, we'll generate a set of yearly files that will not change - only the partial year file will likely change):

```
seafloor% ll
-rw-rw-r-- 1 andrewg mgds 84253 Nov 3 17:57 mag_5.5_2016_2016.csv
-rw-rw-r-- 1 andrewg mgds 57229 Nov 3 17:56 mag_5.5_2017_20171031.csv
```

Combine the **new files since 2015** (start with the most recent new file first):

```
seafloor% cat mag_5.5_2017_20171031.csv mag_5.5_2016_2016.csv >
mag_5.5_2016_20171031_combined.csv
```

```
seafloor% ll mag_5.5_2016_20171031_combined.csv
-rw-rw-r-- 1 andrewg mgds 141482 Nov 3 18:00
mag_5.5_2016_20171031_combined.csv
```

```
seafloor% wc -l mag_5.5_2016_20171031_combined.csv
784 mag_5.5_2016_20171031_combined.csv
```

```
seafloor% head -3 mag_5.5_2016_20171031_combined.csv
time,latitude,longitude,depth,mag,magType,nst,gap,dmin,rms,net,id,updated
,place,type,horizontalError,depthError,magError,magNst,status,locationSou
rce,magSource
2017-10-31T12:37:50.690Z,-
3.8124,127.8091,38.73,5.7,mww,,46,4.52,0.95,us,us1000az05,2017-11-
```

```

02T21:27:35.040Z,"38km WSW of Amahusu,
Indonesia",earthquake,6.5,5.7,0.073,18,reviewed,us,us
2017-10-31T11:50:51.980Z,-
3.7013,127.8157,32.38,6.3,mww,,47,4.465,1.1,us,us1000ayz8,2017-11-
01T13:05:14.548Z,"32km WSW of Hila,
Indonesia",earthquake,6.7,4.4,0.06,27,reviewed,us,us

```

Transfer the combined CSV file to the PC and open the file in Excel.
 Get rid of all columns except for these 7 columns:
 time,latitude,longitude,depth,mag,id,place
 Move the "mag" column left.
 Move the "place" column left.
 Save as a tab-separated file ("xxxxx_combined_tab.txt") and transfer back
 to machine seafloor.

The new column order looks like this:

```

time      latitude      longitude      mag      depth      place      id

```

Check that the columns are in this order before proceeding otherwise when
 we
 cat this new file to the base catalogue file, the columns will not match.
 Make sure that the map and depth columns are in the right order!

```

seafloor% ll *comb*
-rw-rw-r-- 1 andrewg mgds 3843676 Sep 14 2016
mag_5.5_1960_2015_combined.csv
-rw-rw-r-- 1 andrewg mgds 3233185 Sep 14 2016
mag_5.5_1960_2015_combined_reformat.txt
-rw-rw-r-- 1 andrewg mgds 2068815 Sep 14 2016
mag_5.5_1960_2015_combined_tab.txt
-rw-rw-r-- 1 andrewg mgds 141482 Nov 3 18:00
mag_5.5_2016_20171031_combined.csv
-rw-rw-r-- 1 andrewg mgds 74691 Nov 3 18:02
mag_5.5_2016_20171031_combined_tab.txt

```

```

seafloor% head -3 mag_5.5_2016_20171031_combined_tab.txt
time      latitude      longitude      mag      depth      place      id
2017-10-31T12:37:50.690Z      -3.8124 127.8091      5.7      38.73
"38km WSW of Amahusu, Indonesia"      us1000az05
2017-10-31T11:50:51.980Z      -3.7013 127.8157      6.3      32.38
"32km WSW of Hila, Indonesia"      us1000ayz8

```

Get rid of all of the header lines from the combined tab-separated file.
 We use lots
 of redundancy to ensure that we only get rid of those header lines:

```
seafloor% cat mag_5.5_2016_20171031_combined_tab.txt | grep -v latitude |  
grep -v longitude | grep -v depth > /tmp/a
```

```
seafloor% wc -l mag_5.5_2016_20171031_combined_tab.txt /tmp/a  
784 mag_5.5_2016_20171031_combined_tab.txt  
782 /tmp/a
```

The new temporary file has seven (7) columns:

```
seafloor% head -3 /tmp/a  
2017-10-31T12:37:50.690Z      -3.8124 127.8091      5.7      38.73  
"38km WSW of Amahusu, Indonesia"      us1000az05  
2017-10-31T11:50:51.980Z      -3.7013 127.8157      6.3      32.38  
"32km WSW of Hila, Indonesia"      us1000ayz8  
2017-10-31T11:34:44.460Z      -3.8206 127.6642      5.7      47.52  
"53km WSW of Hila, Indonesia"      us1000ayz4
```

Use gawk to generate a USGS URL using the ID value in the last column:

```
seafloor% gawk -F"\t" '{printf  
"%s\t%s\t%s\t%s\t%s\t%s\t%thttp://earthquake.usgs.gov/earthquakes/eventpage  
/%s\n", $1, $2, $3, $4, $5, $6, $7}' /tmp/a > /tmp/b  
seafloor% head -3 /tmp/b  
2017-10-31T12:37:50.690Z      -3.8124 127.8091      5.7      38.73  
"38km WSW of Amahusu, Indonesia"  
http://earthquake.usgs.gov/earthquakes/eventpage/us1000az05  
2017-10-31T11:50:51.980Z      -3.7013 127.8157      6.3      32.38  
"32km WSW of Hila, Indonesia"  
http://earthquake.usgs.gov/earthquakes/eventpage/us1000ayz8  
2017-10-31T11:34:44.460Z      -3.8206 127.6642      5.7      47.52  
"53km WSW of Hila, Indonesia"  
http://earthquake.usgs.gov/earthquakes/eventpage/us1000ayz4
```

Use sed to get rid of the quotation marks around the location descriptions:

```
seafloor% sed 's/\\"//g' /tmp/b > /tmp/c  
seafloor% ll /tmp/b /tmp/c  
-rw-rw-r-- 1 andrewg andrewg 112921 Nov  3 18:04 /tmp/b  
-rw-rw-r-- 1 andrewg andrewg 111585 Nov  3 18:04 /tmp/c  
  
seafloor% wc -l /tmp/b /tmp/c  
782 /tmp/b  
782 /tmp/c  
  
seafloor% head -3 /tmp/c
```

```

2017-10-31T12:37:50.690Z      -3.8124 127.8091      5.7      38.73
38km WSW of Amahusu, Indonesia
http://earthquake.usgs.gov/earthquakes/eventpage/us1000az05
2017-10-31T11:50:51.980Z      -3.7013 127.8157      6.3      32.38
32km WSW of Hila, Indonesia
http://earthquake.usgs.gov/earthquakes/eventpage/us1000ayz8
2017-10-31T11:34:44.460Z      -3.8206 127.6642      5.7      47.52
53km WSW of Hila, Indonesia
http://earthquake.usgs.gov/earthquakes/eventpage/us1000ayz4

```

This file is now ready to be catted (prepended) with the 1960-2015 inclusive base catalogue file.
But, before we cat them two files together, we need to strip the header off the base catalogue file.
Then, we cat the two headerless files together (remember to put the newest file first!).
Finally, add the header back at the start:

```

seafloor% tail -n +2 mag_5.5_1960_2015_combined_reformat.txt > /tmp/base
seafloor% cat /tmp/c /tmp/base > /tmp/d
seafloor% cat header /tmp/d > mag_5.5_1960_20171031_combined_reformat.txt

```

```

seafloor% ll
total 16328
-rw-rw-r-- 1 andrewg mgds      61 Sep 13  2016 header
-rw-rw-r-- 1 andrewg mgds 2486357 Sep 14  2016 mag_5.5_1960_1999.csv
-rw-rw-r-- 1 andrewg mgds 3843676 Sep 14  2016
mag_5.5_1960_2015_combined.csv
-rw-rw-r-- 1 andrewg mgds 3233185 Sep 14  2016
mag_5.5_1960_2015_combined_reformat.txt
-rw-rw-r-- 1 andrewg mgds 2068815 Sep 14  2016
mag_5.5_1960_2015_combined_tab.txt
-rw-rw-r-- 1 andrewg mgds 3344770 Nov  3 18:05
mag_5.5_1960_20171031_combined_reformat.txt
-rw-rw-r-- 1 andrewg mgds 1357319 Sep 14  2016 mag_5.5_2000_2015.csv
-rw-rw-r-- 1 andrewg mgds  84253 Nov  3 17:57 mag_5.5_2016_2016.csv
-rw-rw-r-- 1 andrewg mgds 141482 Nov  3 18:00
mag_5.5_2016_20171031_combined.csv
-rw-rw-r-- 1 andrewg mgds  74691 Nov  3 18:02
mag_5.5_2016_20171031_combined_tab.txt
-rw-rw-r-- 1 andrewg mgds  57229 Nov  3 17:56 mag_5.5_2017_20171031.csv

```

```

seafloor% head mag_5.5_1960_20171031_combined_reformat.txt
Time      Latitude      Longitude      Magnitude      Depth      Location
Information
2017-10-31T12:37:50.690Z      -3.8124 127.8091      5.7      38.73
38km WSW of Amahusu, Indonesia
http://earthquake.usgs.gov/earthquakes/eventpage/us1000az05

```

2017-10-31T11:50:51.980Z	-3.7013 127.8157	6.3	32.38
32km WSW of Hila, Indonesia			
http://earthquake.usgs.gov/earthquakes/eventpage/us1000ayz8			
2017-10-31T11:34:44.460Z	-3.8206 127.6642	5.7	47.52
53km WSW of Hila, Indonesia			
http://earthquake.usgs.gov/earthquakes/eventpage/us1000ayz4			
2017-10-31T11:31:42.810Z	-3.7973 127.7212	5.9	27.3
47km WSW of Hila, Indonesia			
http://earthquake.usgs.gov/earthquakes/eventpage/us1000ayz3			
2017-10-31T04:58:58.120Z	-15.5309 -173.3206		5.8
35	65km NE of Hihifo, Tonga		
http://earthquake.usgs.gov/earthquakes/eventpage/us1000aywi			
2017-10-31T02:20:46.420Z	-21.4682 168.8953		5.9
10	105km E of Tadine, New Caledonia		
http://earthquake.usgs.gov/earthquakes/eventpage/us1000ayv5			
2017-10-31T00:42:06.880Z	-21.6598 169.2034		6.8
11.06	137km E of Tadine, New Caledonia		
http://earthquake.usgs.gov/earthquakes/eventpage/us1000aytk			
2017-10-30T17:12:06.720Z	-10.7359 -74.693	5.7	100.89
57km N of Satipo, Peru			
http://earthquake.usgs.gov/earthquakes/eventpage/us1000ayky			
2017-10-28T19:11:01.690Z	86.8932 55.1348	6	10
North of Franz Josef Land			
http://earthquake.usgs.gov/earthquakes/eventpage/us1000ay3r			

```
seafloor% wc -l mag_5.5_1960_20171031_combined_reformat.txt
25008 mag_5.5_1960_20171031_combined_reformat.txt
```

Check that the file imports in GMA.

Update GMA

Copy the file to the GMA data directory:

```
/data/mgds/web/app.geomapapp.org/htdocs/data/datalayers/geophysics/earthquakes/USGS_ANSS
```

In the gma_menu table, for magnitude >= 5.5, node_uid = 20797.

Update these fields:

- name (to reflect the new date range)
- url (to point to the new file)
- data_refresh_date (gives the update date)

Remake the menu files, test, and ask Ed to rsynch to the commercial provider.

When it comes to update the file by adding more recent events to the base catalogue, remember that both the magnitude and the place columns were moved leftwards. So, make sure that the more recent events have the same column ordering before catting the files.

2017-11-03

Update the catalogue - Magnitudes greater than 6.0

Above, we created "base catalogue" files for each magnitude step for the period 1960-2015 inclusive. To update the catalogues, we need to download the latest data. For example, for this update, we created two new files that will be catted to the base catalogue: a file for the whole of 2016 and a file for most of 2017. After the end of each calendar year, we can create a whole-year file for that year.

USGS search page: <http://earthquake.usgs.gov/earthquakes/search/>

First, note that the search results are limited to 20,000 records and the search will return an error if greater. So, it may be necessary to download events in time-windowed chunks.

Choose magnitude (e.g. 4.5), date range, (e.g. 2016-01-01 00:00:00 to 2016-12-31 23:59:59), Global, CSV format, Order by Time-Newest First. Hit Search.

The name of the downloaded CVS file is always "query.cvs". Rename it right away to avoid confusion (e.g. to "mag_4.5_2016_2016.csv" for the whole year file or to "mag_4.5_2017_20171031.csv" for the partial year file).

```
seafloor% pwd
/home/mgds/ingest/USGS_Earthquake_catalogue/mag6.0
```

Base file:

```
seafloor% ll
-rw-rw-r-- 1 andrewg mgds 1048986 Sep 14 2016
mag_6.0_1960_2015_combined_reformat.txt
```

New files since 2015 - not in final format (note that as we routinely update these catalogues, we'll generate a set of yearly files that will not change - only the partial year file will likely change):

```
seafloor% ll
-rw-rw-r-- 1 andrewg mgds 25999 Nov 3 18:11 mag_6.0_2016_2016.csv
-rw-rw-r-- 1 andrewg mgds 15885 Nov 3 18:11 mag_6.0_2017_20171031.csv
```

Combine the **new files since 2015** (start with the most recent new file first):

```
seafloor% cat mag_6.0_2017_20171031.csv mag_6.0_2016_2016.csv >
mag_6.0_2016_20171031_combined.csv
```

```
seafloor% ll mag_6.0_2016_20171031_combined.csv
-rw-rw-r-- 1 andrewg mgds 41884 Nov 3 18:13
mag_6.0_2016_20171031_combined.csv
```

```
seafloor% wc -l mag_6.0_2016_20171031_combined.csv
234 mag_6.0_2016_20171031_combined.csv
```

```
seafloor% head -3 mag_6.0_2016_20171031_combined.csv
time,latitude,longitude,depth,mag,magType,nst,gap,dmin,rms,net,id,updated
,place,type,horizontalError,depthError,magError,magNst,status,locationSou
rce,magSource
2017-10-31T11:50:51.980Z,-
3.7013,127.8157,32.38,6.3,mww,,47,4.465,1.1,us,us1000ayz8,2017-11-
01T13:05:14.548Z,"32km WSW of Hila,
Indonesia",earthquake,6.7,4.4,0.06,27,reviewed,us,us
2017-10-31T00:42:06.880Z,-
21.6598,169.2034,11.06,6.8,mww,,55,1.106,1.39,us,us1000aytk,2017-10-
31T22:10:12.151Z,"137km E of Tadine, New
Caledonia",earthquake,6.9,3,0.071,19,reviewed,us,us
```

Transfer the combined CSV file to the PC and open the file in Excel.

Get rid of all columns except for these 7 columns:

time,latitude,longitude,depth,mag,id,place

Move the "mag" column left.

Move the "place" column left.

Save as a tab-separated file ("xxxxx_combined_tab.txt") and transfer back to machine seafloor.

The new column order looks like this:

time	latitude	longitude	mag	depth	place	id
------	----------	-----------	-----	-------	-------	----

Check that the columns are in this order before proceeding otherwise when we

cat this new file to the base catalogue file, the columns will not match.

Make sure that the map and depth columns are in the right order!

```
seafloor% ll *comb*
```

```
-rw-rw-r-- 1 andrewg mgds 1250036 Sep 14 2016
```

```
mag_6.0_1960_2015_combined.csv
```

```
-rw-rw-r-- 1 andrewg mgds 1048986 Sep 14 2016
```

```
mag_6.0_1960_2015_combined_reformat.txt
```

```
-rw-rw-r-- 1 andrewg mgds 670978 Sep 14 2016
```

```
mag_6.0_1960_2015_combined_tab.txt
```

```
-rw-rw-r-- 1 andrewg mgds 41884 Nov 3 18:13
```

```
mag_6.0_2016_20171031_combined.csv
```

```
-rw-rw-r-- 1 andrewg mgds 22187 Nov 3 18:16
```

```
mag_6.0_2016_20171031_combined_tab.txt
```

```
seafloor% head -3 mag_6.0_2016_20171031_combined_tab.txt
```

time	latitude	longitude	mag	depth	place	id
2017-10-31T11:50:51.980Z		-3.7013	127.8157		6.3	32.38

```
"32km WSW of Hila, Indonesia" us1000ayz8
```

2017-10-31T00:42:06.880Z		-21.6598		169.2034		6.8
--------------------------	--	----------	--	----------	--	-----

11.06	"137km E of Tadine, New Caledonia"				us1000aytk	
-------	------------------------------------	--	--	--	------------	--

Get rid of all of the header lines from the combined tab-separated file.

We use lots

of redundancy to ensure that we only get rid of those header lines:

```
seafloor% cat mag_6.0_2016_20171031_combined_tab.txt | grep -v latitude |  
grep -v longitude | grep -v depth > /tmp/a
```

```
seafloor% wc -l mag_6.0_2016_20171031_combined_tab.txt /tmp/a
```

```
234 mag_6.0_2016_20171031_combined_tab.txt
```

```
232 /tmp/a
```

The new temporary file has seven (7) columns:

```
seafloor% head -3 /tmp/a
2017-10-31T11:50:51.980Z      -3.7013 127.8157      6.3      32.38
"32km WSW of Hila, Indonesia" us1000ayz8
2017-10-31T00:42:06.880Z      -21.6598      169.2034      6.8
11.06 "137km E of Tadine, New Caledonia" us1000aytk
2017-10-28T19:11:01.690Z      86.8932 55.1348 6      10      North of
Franz Josef Land      us1000ay3r
```

Use gawk to generate a USGS URL using the ID value in the last column:

```
seafloor% gawk -F"\t" '{printf
"%s\t%s\t%s\t%s\t%s\t%s\t%thttp://earthquake.usgs.gov/earthquakes/eventpage
/%s\n", $1, $2, $3, $4, $5, $6, $7}' /tmp/a > /tmp/b
seafloor% head -3 /tmp/b
2017-10-31T11:50:51.980Z      -3.7013 127.8157      6.3      32.38
"32km WSW of Hila, Indonesia"
http://earthquake.usgs.gov/earthquakes/eventpage/us1000ayz8
2017-10-31T00:42:06.880Z      -21.6598      169.2034      6.8
11.06 "137km E of Tadine, New Caledonia"
http://earthquake.usgs.gov/earthquakes/eventpage/us1000aytk
2017-10-28T19:11:01.690Z      86.8932 55.1348 6      10      North of
Franz Josef Land
http://earthquake.usgs.gov/earthquakes/eventpage/us1000ay3r
```

Use sed to get rid of the quotation marks around the location descriptions:

```
seafloor% sed 's/\\"//g' /tmp/b > /tmp/c
seafloor% ll /tmp/b /tmp/c
-rw-rw-r-- 1 andrewg andrewg 33467 Nov  3 18:20 /tmp/b
-rw-rw-r-- 1 andrewg andrewg 33045 Nov  3 18:20 /tmp/c

seafloor% wc -l /tmp/b /tmp/c
  232 /tmp/b
  232 /tmp/c

seafloor% head -3 /tmp/c
2017-10-31T11:50:51.980Z      -3.7013 127.8157      6.3      32.38
32km WSW of Hila, Indonesia
http://earthquake.usgs.gov/earthquakes/eventpage/us1000ayz8
2017-10-31T00:42:06.880Z      -21.6598      169.2034      6.8
11.06 137km E of Tadine, New Caledonia
http://earthquake.usgs.gov/earthquakes/eventpage/us1000aytk
2017-10-28T19:11:01.690Z      86.8932 55.1348 6      10      North of
Franz Josef Land
http://earthquake.usgs.gov/earthquakes/eventpage/us1000ay3r
```

This file is now ready to be catted (prepended) with the 1960-2015 inclusive base catalogue file.
 But, before we cat them two files together, we need to strip the header off the base catalogue file.
 Then, we cat the two headerless files together (remember to put the newest file first!).
 Finally, add the header back at the start:

```
seafloor% tail -n +2 mag_6.0_1960_2015_combined_reformat.txt > /tmp/base
seafloor% cat /tmp/c /tmp/base > /tmp/d
seafloor% cat header /tmp/d > mag_6.0_1960_20171031_combined_reformat.txt
```

```
seafloor% ll
total 5308
-rw-rw-r-- 1 andrewg mgds      61 Sep 13  2016 header
-rw-rw-r-- 1 andrewg mgds 1250036 Sep 14  2016
mag_6.0_1960_2015_combined.csv
-rw-rw-r-- 1 andrewg mgds 1048986 Sep 14  2016
mag_6.0_1960_2015_combined_reformat.txt
-rw-rw-r-- 1 andrewg mgds  670978 Sep 14  2016
mag_6.0_1960_2015_combined_tab.txt
-rw-rw-r-- 1 andrewg mgds 1250036 Sep 14  2016 mag_6.0_1960_2015.csv
-rw-rw-r-- 1 andrewg mgds 1082031 Nov  3 18:23
mag_6.0_1960_20171031_combined_reformat.txt
-rw-rw-r-- 1 andrewg mgds  25999 Nov  3 18:11 mag_6.0_2016_2016.csv
-rw-rw-r-- 1 andrewg mgds  41884 Nov  3 18:13
mag_6.0_2016_20171031_combined.csv
-rw-rw-r-- 1 andrewg mgds  22187 Nov  3 18:16
mag_6.0_2016_20171031_combined_tab.txt
-rw-rw-r-- 1 andrewg mgds  15885 Nov  3 18:11 mag_6.0_2017_20171031.csv
```

```
seafloor% head mag_6.0_1960_20171031_combined_reformat.txt
Time      Latitude      Longitude      Magnitude      Depth      Location
Information
2017-10-31T11:50:51.980Z      -3.7013 127.8157      6.3      32.38
32km WSW of Hila, Indonesia
http://earthquake.usgs.gov/earthquakes/eventpage/us1000ayz8
2017-10-31T00:42:06.880Z      -21.6598      169.2034      6.8
11.06 137km E of Tadine, New Caledonia
http://earthquake.usgs.gov/earthquakes/eventpage/us1000aytk
2017-10-28T19:11:01.690Z      86.8932 55.1348 6      10      North of
Franz Josef Land
http://earthquake.usgs.gov/earthquakes/eventpage/us1000ay3r
2017-10-24T10:47:47.540Z      -7.2364 123.0401      6.7      549.19
141km NNE of Palua, Indonesia
http://earthquake.usgs.gov/earthquakes/eventpage/us1000aw8q
```

2017-10-18T12:00:59.810Z	-20.5981	-173.8949	6
10	100km SSE of Pangai, Tonga		
http://earthquake.usgs.gov/earthquakes/eventpage/us1000atnb			
2017-10-10T18:53:27.770Z	-54.2632	8.6363 6.7	9
Bouvet Island region			
http://earthquake.usgs.gov/earthquakes/eventpage/us1000apq1			
2017-10-10T06:32:20.800Z	-18.5212	-69.6411	6.3
82.43	36km SSW of Putre, Chile		
http://earthquake.usgs.gov/earthquakes/eventpage/us2000b3dm			
2017-10-08T22:34:33.960Z	52.4496 176.8078	6.6	111.84
60km E of Buldir Island, Alaska			
http://earthquake.usgs.gov/earthquakes/eventpage/us2000b2vq			
2017-10-08T20:48:58.720Z	-61.7431	154.5526	6.3
10	Balleny Islands region		
http://earthquake.usgs.gov/earthquakes/eventpage/us2000b2v1			

seafloor% wc -l mag_6.0_1960_20171031_combined_reformat.txt
8102 mag_6.0_1960_20171031_combined_reformat.txt

Check that the file imports in GMA.

Update GMA

Copy the file to the GMA data directory:

/data/mgds/web/app.geomapapp.org/htdocs/data/datalayers/geophysics/earthquakes/USGS_ANSS

In the gma_menu table, for magnitude >= 6.0, node_uid = 20796.

Update these fields:

- name (to reflect the new date range)
- url (to point to the new file)
- data_refresh_date (gives the update date)

Remake the menu files, test, and ask Ed to rsynch to the commercial provider.

When it comes to update the file by adding more recent events to the base catalogue, remember that both the magnitude and the place columns were moved leftwards. So, make sure that the more recent events have the same column ordering before catting the files.

