Dublin R El Dorado Competition Page

Mick Cooney

6 July 2016

SUBMISSIONS ARE OPEN. SUBMIT YOUR BIDS HERE BY 23:59:59 SEP 11, 2006 (http://www.kaybensoft.com:4114)

PLEASE REGISTER ON THIS FORM IF YOU ARE INTERESTED IN TAKING PART - WE WANT TO GET AN IDEA OF NUMBERS AND HAVE CONTACT DETAILS

(https://docs.google.com/forms/d/1YDCh4VR2znQici2Bsmcb_qA5fPXUtyACAe7Nai5XTxU/viewform)

1 Competition Summary

Starting Date: July 6, 2016

Closing Date: Sep 11, 2016

There is gold in them-there hills!

In this competition the goal is to predict the amount of gold in an area and then win an auction for the rights to extract and sell the gold.

The winner of the competition is the team that finishes with the most cash after all bids and extraction costs are paid and the mined gold has been sold.

A dataset is being provided on which teams may base their predictions, and submissions for the competition take the form of bids for all the available parcels of land.
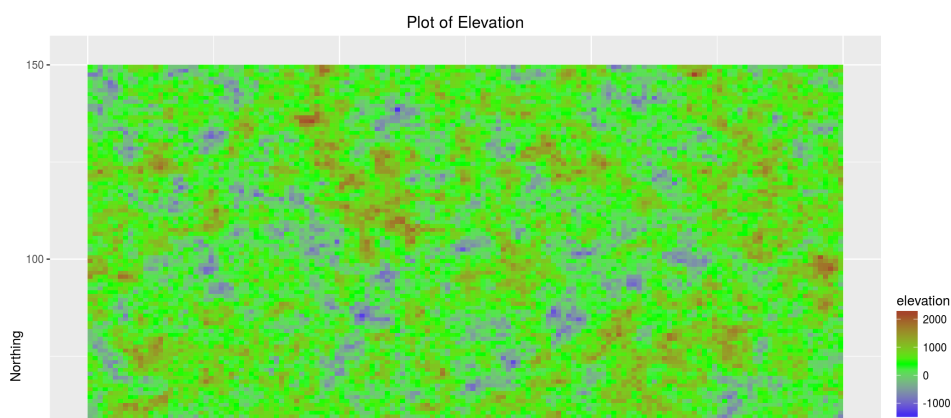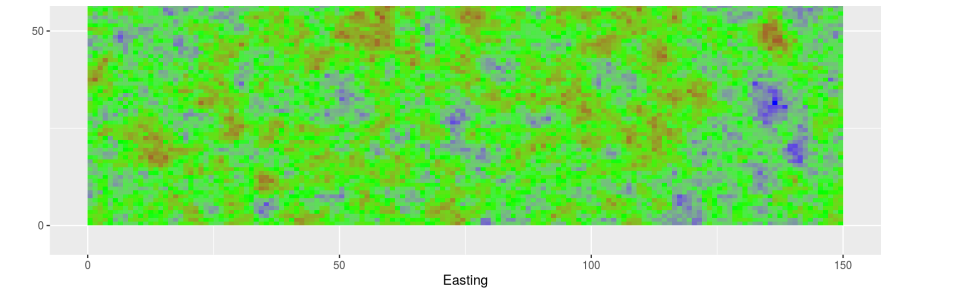
1.1 Quick Summary

- 5,000 data points with pseudo-geological information (including proven gold reserves).
- 50 (or more) new sites up for auction with limited data.
- Each team starts with $50,000,000.00 budget to bid with.
- Blind, sealed-bid auctions for rights to mine the parcel of land (min bid $100,000.00)
- Auctions happen in order by parcel_id
- Extraction costs are non-trivial.
- Winning team has most cash at the end.

2 Data Description

We have generated a random piece of land for which some basic data will be made available. This land is split up into a grid of 150x150 'parcels'. Elevation data is available for the entire grid.

5,000 squares were previously auctioned and exploited. Complete information for these parcels is available as part of the dataset.

Plot of Elevation

The 'psuedo-geological' information provided with the data is not based in any way on reality, so we are providing some domain knowledge here.

Geological compositions found in the ground are grouped into three categories: Mineral, Ore and Rock. Quantities within each category tend to co-occur in some way though the exact relationship of this is unknown.

- **Minerals:** Isogen, Megacyte, Mexallon, Nocxium, Pyerite, Tritanium
- **Ore:** Gneiss, Hedbergite, Plagioclase, Spudumain, Veldspar
- **Rock:** Arkonor, Bistot, Crokite, Mercoxit

3 Competition Setup

Unlike other competitions involving predictive modeling, we are not providing any kind of 'error metric' for your predictions. This competition will not reward the most accurate predictions for the amount of gold in the parcels provided. This competition also tries to emphasise effective use of the predictions, rather than the quality of the predictions themselves.

Participants will bid for parcels of land according to their perceived or predicted value and then make a profit or loss according to the amount of gold found in the parcels of land they win at auction.

The score for each team in this competition is the final cash balance after parcel bids are paid, the gold is extracted, extraction costs are paid and the extracted gold is sold.

Gold is priced in US Dollars (USD) and so this is currency of choice for this competition.

All auction bids, extraction expenses and gold amounts are priced in USD, and the final amount possessed by each team at the final accounting is the score for that team.

3.1 Gold Extraction Details

Extracting gold from the ground comes with a cost, which need to be factored into the profit calculation.

Extraction costs comprise two parts: *fixed costs* and *variable costs*, both of which depend on the elevation of the ground from which the gold is being extracted.

3.1.1 Fixed Costs

Fixed costs depend on the elevation of the land only and represent various capital costs required to mine the land such as plant installation, legal costs, etc.

Elevation Class	Elevation Range	Fixed Cost
Low Elevation	*Elevation < 0m*	$4,000,000.00
Medium Elevation	*0m < Elevation <= 500m*	$3,000,000.00
High Elevation	*Elevation > 500m*	$8,000,000.00

3.1.2 Variable Costs

Variable costs depend on the elevation of the land but also vary as the amount of gold is extracted. As more gold is mined, economies of scale mean that the cost per unit of gold extracted is reduced to a lesser or greater extent.

There are three elevation bands for variable costs: Less than 0 m, 0m – 700 m and greater than 700m.

Elevation Class	Elevation Range
Low Elevation	*Elevation < 0m*
Medium Elevation	*0m < Elevation <= 700m*
High Elevation	*Elevation > 700m*

Note that the elevation bands for fixed and variable costs differ

The variable costs are displayed here in a table and this data is available for download. Should a parcel of land contain more gold, the variable costs will scale according to trends shown in the data.

Note that the variable costs do not include the fixed costs.

```
##     gold_amount      low       med      high
## 1:           0       $0        $0        $0
## 2:        2000  $964,000  $882,000  $768,000
## 3:        4000 $1,868,000 $1,734,000 $1,359,000
## 4:        6000 $2,729,000 $2,565,000 $1,861,000
## 5:        8000 $3,559,000 $3,380,000 $2,323,000
## 6:       10000 $4,368,000 $4,184,000 $2,771,000
## 7:       12000 $5,161,000 $4,981,000 $3,218,000
## 8:       14000 $5,945,000 $5,773,000 $3,670,000
## 9:       16000 $6,723,000 $6,562,000 $4,130,000
## 10:      18000 $7,498,000 $7,349,000 $4,598,000
## 11:      20000 $8,271,000 $8,135,000 $5,073,000
```

Note that these are estimates of the underlying cost functions which maps gold amounts to the variable cost of extracting that amount of gold at the respective elevation levels.

So, if the parcel of land has 13,000 ounces of gold and is at 600m (medium elevation for variable costs), we expect the variable extraction costs for all 13,000 ounces to be somewhere between $4.98 and $5.77 million.

4 Parcel Auction

Parcels are allocated by a blind, sealed-bid auction. Your bids will be your submission to the competition.

The auction mechanism is the *Vickrey Auction* (Wikipedia Entry (https://en.wikipedia.org/wiki/Vickrey_auction)) so the highest bid for the parcel wins but pays the bid of the second highest bidder. This mechanism helps avoid gaming bids and incentivises bidders to bid the true perceived value of the parcel.

Once the competition is closed and all entries are received, the parcels of land will be sequentially auctioned off in order of their id.

A budget of $50,000,000 is allocated to each team for bidding purposes and the total amount spent by a team on their winning bids cannot exceed this total. Teams may make total bids in excess of this amount if they anticipate not winning auctions, but overspends are penalised according to the rules below.

4.1 Example Auction

To illustrate the mechanics of the auction, we illustrate a simple one. Suppose we have five bids for a parcel of land with values:

- A - $1,500,000.00
- B - $1,200,000.00
- C - $1,000,000.00
- D - $1,000,000.00
- E - $800,000.00

In this case, Team A wins the auction with their bid of $1,500,000.00. As the second bid was $1,200,000, Team A pays this amount to secure the rights for that parcel of land.
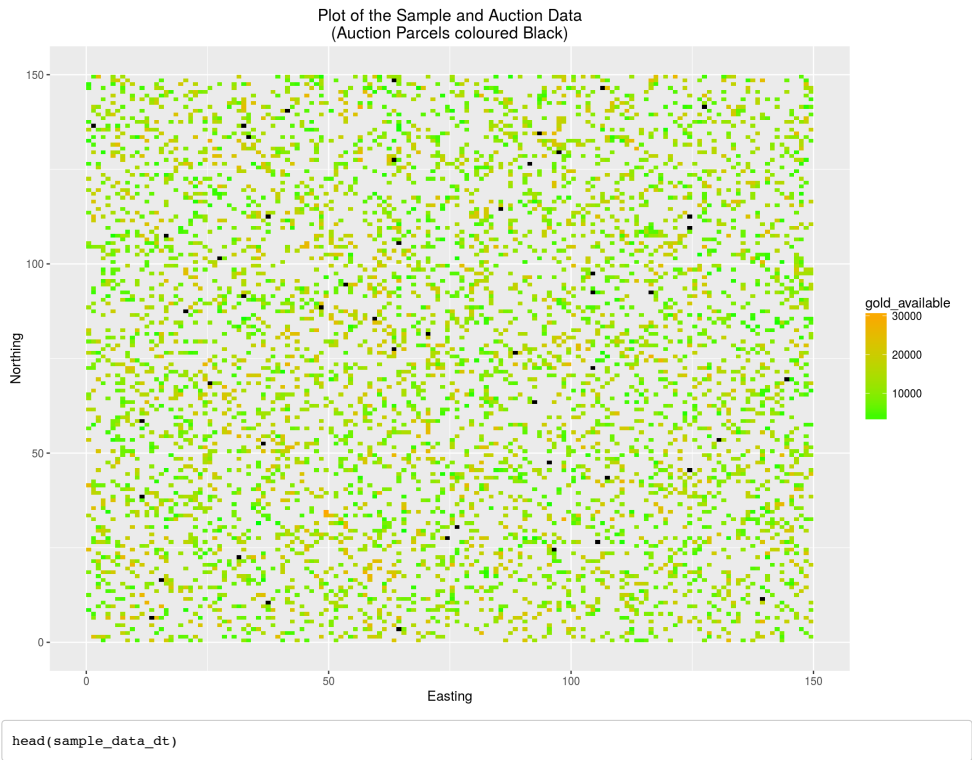
Suppose this auction is one of the later ones and Team A only has $1,000,000 left to spend. Without the cash to pay for the parcel, Team A is removed from the auction and it is redone. In this case, Team B wins the auction and pays the new-second-highest bid: $1,000,000.00.

When the final scores are tallied, a fine of $100,000 is assessed on Team A's score at the end.

5 Competition Rules

1. A competition signup sheet will be published a few days after the competition goes live. Signups will have a team name and an email address for contact details. Any and all correspondance for this competition will be done via that email address.
2. This competition is just for fun so there are no prizes beyond some kudos from me and the rest of the community. That said, if groups of you decide to form small pools amongst yourselves, let me know. It may be possible to accommodate such arrangements in an informal way.
3. It is expected that Dublin R will run a 'Data Dojo'-style event after the competition concludes where participants will be invited to speak about the methods they employed. Participation in any such events is entirely optional but is always interesting and a lot of fun!
4. All quantities of gold in the competition are in ounces.
5. It is assumed that all teams have hedged the future price of gold at $1,500 per ounce. This is the price at which profits will be calculated.
6. If we have a larger than anticipated interest, I reserve the right to release additional parcels of land for auction, so that all entrants feel they have a shot at winning at least one parcel of land. Should this happen, all signed up participants will be informed via email and the datasets on the website will be updated.
7. A website through which participants submit their bids will be made available in due course.
8. Entries consist of a single CSV file with two columns: `parcel_id` (integer) and `bid_amount` (floating point with 2 decimal places).
9. The minimum bid amount for each parcel of land is $100,000 USD. Bids below that amount will be considered 0.
10. Teams do not have to make a bid for every available parcel of land.
11. Not all parcels of land are guaranteed to make a profit - especially once bids are paid.
12. In the unlikely event of two winning bids being equal, the winner for that auction will be the team with the earlier timestamp on their submission. Please note that this means that they will pay the amount they bid, as the second-highest bid amount is the same amount!
13. Auctions are resolved in order of 'parcel_id' and are performed sequentially right at the start of the resolution process.
14. If, for whatever reason, a team is unable to pay an otherwise-winning bid, that team will be removed from the current auction, the auction will be redone without that team's participation and a penalty of $100,000 will be assessed to the final score for the team (but not while the auctions are being determined). The team will still participate in subsequent auctions. Note that this means should that team fail to fund another winning bid, further fines will be assessed to their final score.
15. Fixed and variable costs for gold extraction are not paid out of the budget. It is assumed that money is paid out of the expected gold sales and are accounted at the end.
16. The top 5 scores will be published after the close of the competition so that those competitors may be lauded for their predictive prowess, the auction acumen and their overall level of general awesome. Other participants may also get their scores upon request.
17. Code for the auction will be made publicly available after completion of the competition. The data-generation code may be made available, though I may decide to keep it private in order to run future competitions with a similar structure.
18. If any of the rules appear unclear, first re-read them just in case. If you are still confused, drop me a mail at mickcooney AT gmail
19. I, Mick Cooney, hereby declare myself "Dublin R El Dorado Competition Dictator-for-Life". While not a cruel man, I do consider myself more a member of the Augustus Caesar / Sulla academy, and not so much the Julius Caesar one.
20. Anyone misbehaving or generally trying to cheat in some way will fall afoul of the above Dictator-for-Life clause. This is for fun, so whatever you are thinking of doing - just don't. :)

6 Data Downloads

Plot of the Sample and Auction Data
(Auction Parcels coloured Black)

```
head(sample_data_dt)
```

```
##    parcel_id Easting Northing   Pyerite Mexallon Tritanium Megacyte Nocxium
## 1:         4     3.5      0.5  133.2717 108.2557   95.8811 129.6954 111.146
## 2:         9     8.5      0.5  144.4469 110.1732   97.8410 104.6556 104.430
## 3:        11    10.5      0.5  145.6990 110.6321  116.2038 158.5974 118.838
## 4:        14    13.5      0.5  126.2608 116.7307  120.1670 116.6437 127.243
## 5:        17    16.5      0.5  107.3763  98.9368  120.7785  98.7749 119.904
## 6:        18    17.5      0.5   91.9581 110.6819   98.0557 111.3876 105.708
##     Isogen Veldspar Plagioclase Hedbergite Spudumain   Gneiss Arkonor Mercoxit
## 1: 138.742 109.8161     148.298    130.353   145.597 114.181 105.516  107.604
## 2: 123.081 104.5781     117.522    128.958   126.252 119.666 116.793  109.627
## 3: 109.344  98.5723     148.772    129.747   149.727 119.894 103.669  114.768
## 4: 103.685 117.2867     101.197    191.688   178.005 138.917 101.045  107.515
## 5: 121.311 121.4102     167.341    158.549   178.469 131.626 109.365  101.883
## 6: 129.003 101.6654     175.911    114.814   184.340 104.648 103.643  104.924
##      Bistot Crokite gold_available
## 1:  94.1169 108.024        3901.52
## 2: 130.8569 113.133       17523.53
## 3: 119.5031 106.386        4537.82
## 4: 104.5877 105.245       21601.11
## 5: 111.5630 104.275       21186.86
## 6: 123.2936 102.417        8731.34
```

```
head(auction_data_dt)
```

```
##    parcel_id Easting Northing elevation Arkonor Mercoxit   Bistot  Crokite
## 1:       515    64.5      3.5   23.1895 107.207  101.978 105.7088 102.4967
## 2:       914    13.5      6.5  784.9871 102.436  104.951 108.8340 109.3543
## 3:      1538    37.5     10.5 1258.7271 103.497  111.473 105.5494  98.0687
## 4:      1790   139.5     11.5  474.1757 102.066  109.974 105.2102 104.4218
## 5:      2416    15.5     16.5 1584.8595 114.476  116.988 113.1447 108.8117
## 6:      3332    31.5     22.5  209.1741 105.475  109.398  84.8997 110.8472
```

All the files available for this competition are available in a single zip file.

The datafile is downloadable here (data/drcomp_files.zip)

The zip file contains the following files:

- **auction_parcels.csv**: This file contains the basic data available for the parcels of land available for auction.
- **costs_data.csv**: This file gives estimations for the variable costs for extracting the gold.
- **elevation_data.csv**: This file contains the elevation data for the full 150x150 grid of parcels.
- **sample_data.csv**: This is the training data supplied to all the participating teams.