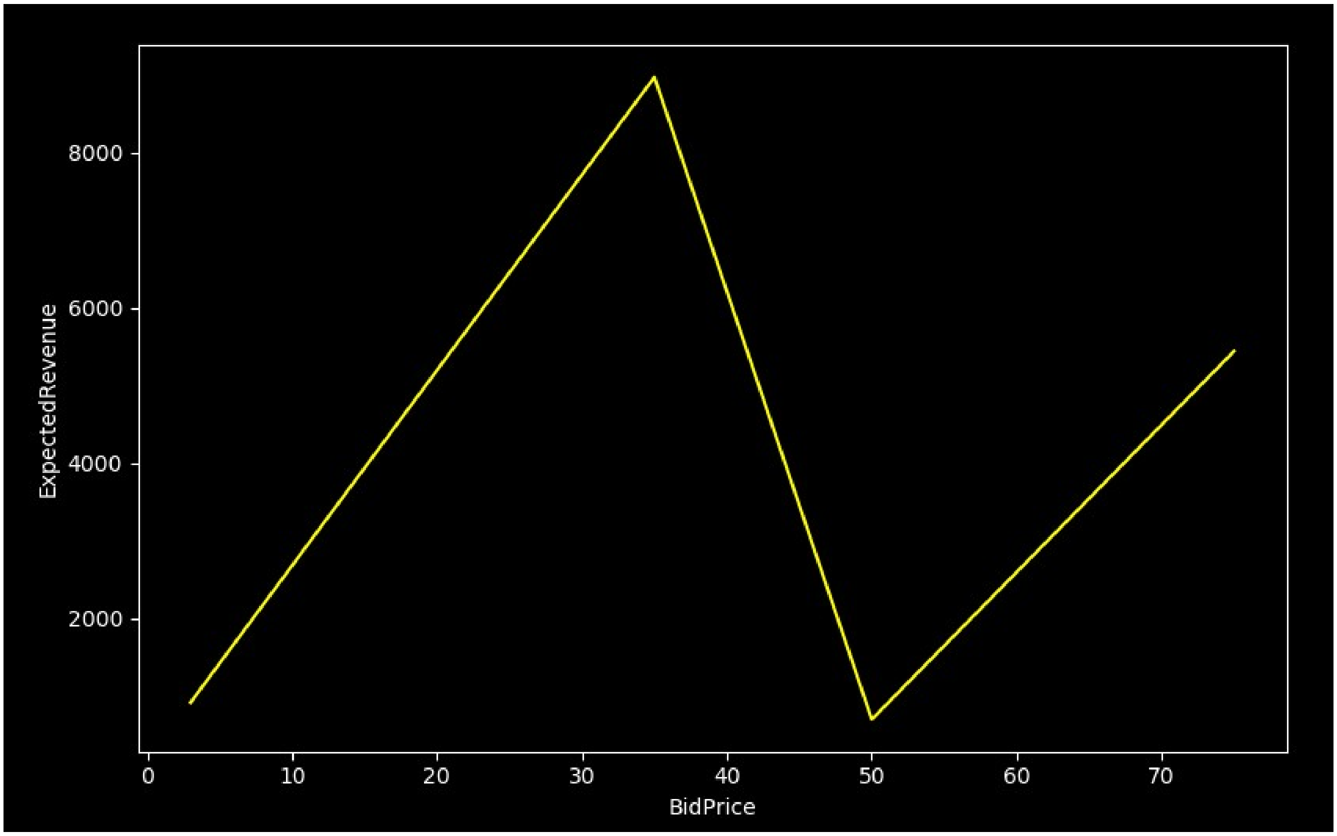
I have been tasked with analyzing and improving the decision science behind bidding on contracts consisting of $3.00, $35.00, $50.00, and $75.00,

Using the Expected Revenue, Expected (loan) Conversions, and Winning or Losing Bids… here is my analysis.

As you can see this is a general summation of revenue increase with respect to the Successful Bid Price. At first glance the $35 fixed price is the most lucrative.



Now that we are only looking at bids that were successful. We still have 16,026 total bids to analyze in one set. So let’s group them into BidPrice buckets and try and take a closer look into what’s happening.

3.0

	ExpectedConversion	ExpectedRevenue
290	0.168	196.40
816	0.109	270.61
1083	0.137	235.76
2384	0.190	340.58
3073	0.140	234.48
...	...	...
53729	0.236	126.08
53770	0.203	179.28
53812	0.346	95.36
53920	0.202	107.68
53926	0.360	88.24

[913 rows x 2 columns]

35.0

	ExpectedConversion	ExpectedRevenue
9	0.356	271.49
10	0.319	184.64
12	0.463	82.80
13	0.243	158.96
20	0.396	414.24
...	...	...
53899	0.472	92.00
53901	0.592	95.84
53914	0.383	194.64
53924	0.555	113.60
53943	0.246	221.68

[8975 rows x 2 columns]

50.0

	ExpectedConversion	ExpectedRevenue
349	0.206	474.43
396	0.238	456.97
442	0.313	385.91
549	0.279	408.15
566	0.370	365.06
...	...	...
53461	0.422	352.12
53485	0.349	374.14
53511	0.169	572.05
53835	0.276	414.24
53911	0.433	313.45

[695 rows x 2 columns]

75.0

	ExpectedConversion	ExpectedRevenue
14	0.158	614.01
15	0.337	613.61
21	0.173	676.15
38	0.598	532.30
43	0.543	323.65
...	...	...
53912	0.563	307.12
53916	0.376	411.62
53922	0.464	433.31
53936	0.267	544.99
53946	0.297	538.02

[5443 rows x 2 columns]

Note*

Now we have a nice sampling that shows we are far more successful bidding at the $35 and $75 price marker. Or at least in this sample. Let's look closer by taking the mean() of each outcome, then using that to find the profit margin with a fixed size of 20

At $3.0
ExpectedConversion 0.301
ExpectedRevenue $153.089
NetRevenue(20leads): $861
** 93.5% profit margin

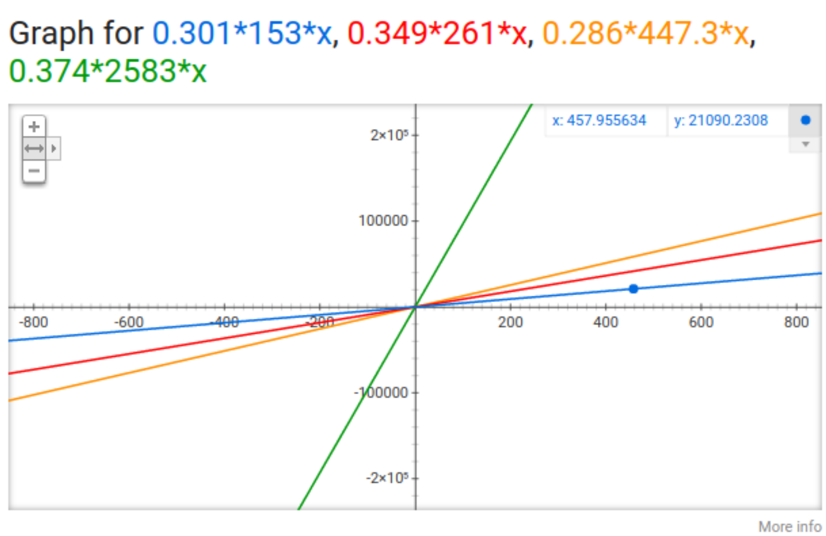
At $35.0
ExpectedConversion 0.349
ExpectedRevenue $261.681
NetRevenue(20leads): $1,126
** 61.7% profit margin

At $50.0
ExpectedConversion 0.286
ExpectedRevenue $447.308
NetRevenue(20leads): $1,236
** 60.9% profit margin

At $75.0
ExpectedConversion 0.374
ExpectedRevenue $545.888
NetRevenue(20leads): $2,583
** 63.2% profit margin

Here again we have some very telling leads into which way we should go given a finite try at bidding these contacts. One thing we can see clearly is that we are spending the most at the $35.00 bid price which actually has the lowest reward (profit) at the current ExpectedConversion rates!

This quick chart allows us to take into account a reshuffling and prioritize at the slightly higher margins. Furthermore, because of the negligible cost of the $3.00 price, we probably not ever pass it up if the opportunity arises. In this case net revenue is not the correct benchmark but rather the profit margin



No hard and fast rules here as there is much more variability in this universe to account for, but the priority of bidding should be $75, $3 , $50, $35

(quick google graph shows just how close the margins are for at least a few thousand winning bids.)

Of course this leaves out a large amount of variables in this universe that need exploring.

So let's see what happens we attempt regression analysis on only successful bids, first this needs organized into a sample dataframe. Here is an example of a random sampled dataframe with n=30.

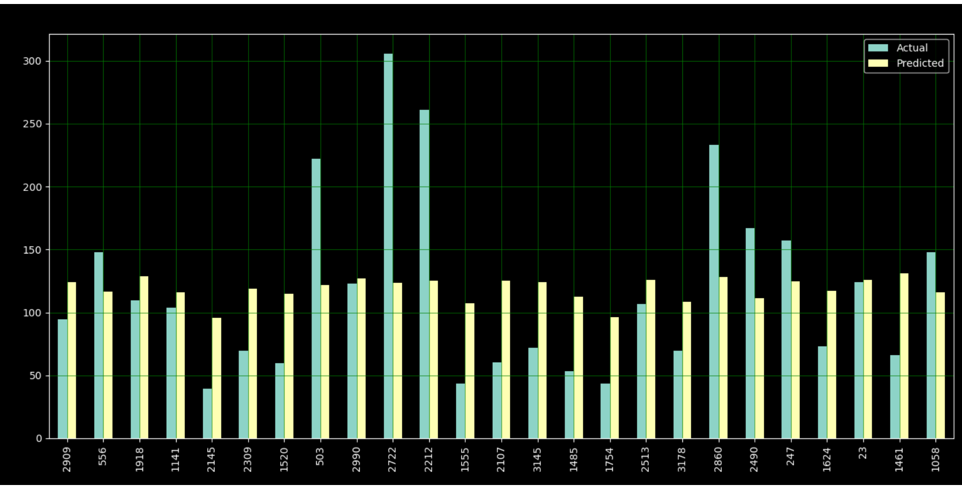
```
       BidPrice  AcceptedBid  ExpectedConversion  ExpectedRevenue
47303      35.0          1.0               0.639           160.32
42421      75.0          1.0               0.338           463.07
39813      75.0          1.0               0.234           642.31
52438      75.0          1.0               0.342           535.31
22211      35.0          1.0               0.316           148.80
30495      75.0          1.0               0.359           790.37
10362      35.0          1.0               0.569           277.01
28407      75.0          1.0               0.329           548.27
16906      35.0          1.0               0.382           189.44
14314      75.0          1.0               0.220           579.17
38834      35.0          1.0               0.341           252.96
13893      75.0          1.0               0.423           396.89
51299      35.0          1.0               0.241           373.99
19035      35.0          1.0               0.215           338.25
9335        3.0          1.0               0.212           197.12
22385      75.0          1.0               0.328           484.28
20365      75.0          1.0               0.263           561.32
27007      35.0          1.0               0.391           166.08
25926      35.0          1.0               0.105           373.70
1711       35.0          1.0               0.239           369.65
51351      75.0          1.0               0.458           542.37
25653      35.0          1.0               0.432           231.68
26705      75.0          1.0               0.133           678.61
15263      75.0          1.0               0.402           431.63
24052      35.0          1.0               0.326           196.24
17327      35.0          1.0               0.222           328.44
7843       35.0          1.0               0.276           214.00
48528      35.0          1.0               0.588           209.44
14607      35.0          1.0               0.471            81.68
37817      75.0          1.0               0.501           772.06
BidPrice                 4
AcceptedBid              1
ExpectedConversion     728
ExpectedRevenue       7619
dtype: int64
```

Now we can split the dataframe into a test set and a training set. Adding in some columns for purview and potential training features.

```
[16026 rows x 3 columns]
       BidPrice  ExpectedConversion  ExpectedRevenue  ExpectedNetRevenue  PotentialMargin(%)
290         3.0               0.168           196.40              32.995               1.527
816         3.0               0.109           270.61              29.496               1.109
1083        3.0               0.137           235.76              32.299               1.272
2384        3.0               0.190           340.58              64.710               0.881
3073        3.0               0.140           234.48              32.827               1.279
...         ...                 ...              ...                 ...                 ...
53912      75.0               0.563           307.12             172.909              24.420
53916      75.0               0.376           411.62             154.769              18.221
53922      75.0               0.464           433.31             201.056              17.309
53936      75.0               0.267           544.99             145.512              13.762
53946      75.0               0.297           538.02             159.792              13.940
```

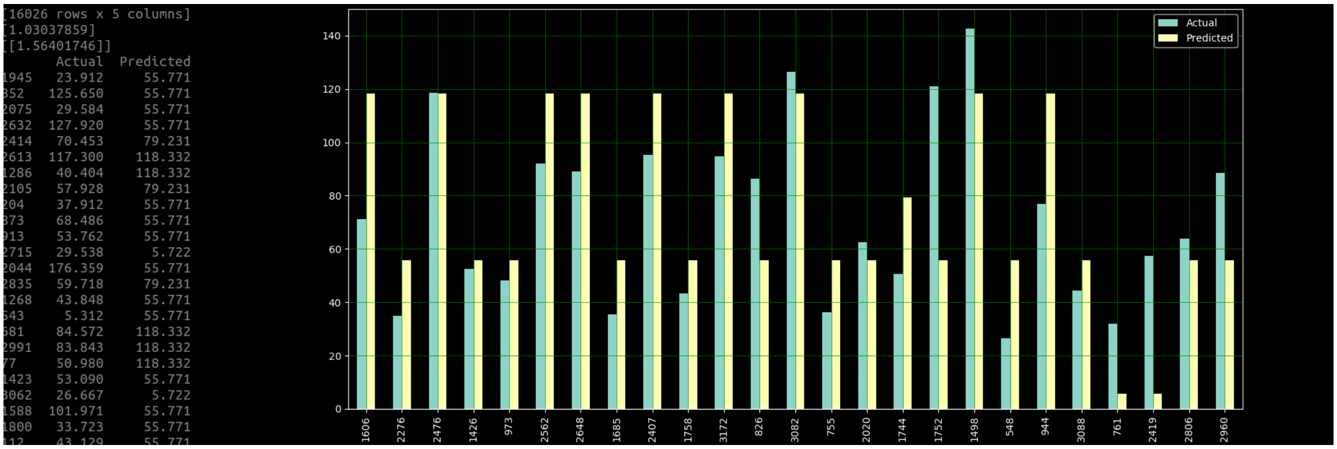
Now with our nicely organized dataset that assumes all bids were successful we can train a linear regression model and make a prediction. Here’s one that is training with one of the main coefficients being PotentialMargin.
Hypothesis: Potential Margin percentage affects the Total Expected Revenue
Result:

```
[16026 rows x 5 columns]
[139.27257126]
[[-1.17302227]]
       Actual  Predicted
2434   66.430    130.373
1596   48.992    126.702
2572  182.365    118.094
1698   65.724    123.094
2038   71.319    129.256
979    68.749    117.416
850    63.445    125.166
2938   55.786    114.324
1719  100.958    122.599
2325   67.374    118.554
2671   57.731    119.218
2748  132.183    122.900
2225   34.008    136.220
2518   98.777    130.710
351    46.160    121.484
2602  107.202    125.212
462   182.067    124.841
134   131.377    126.210
3167  108.610    123.801
3059   96.105    127.695
33     28.579    136.896
2169   55.919    124.001
2887   91.482    126.527
2833  193.794    125.880
1362   39.479    136.162
```



Not so great…. It appears that we are regularly missing by 50% much of the time.

Hypothesis: Bid Price is greater factor in Total Expected Revenue
Result:



This seems like it’s moving quicker into the right direction even with the overfitting problem (notice the 55.771 repeats etc.)

One thing is very clear, more work is needed because of the abundance of missing features, not even considering the failed bids as missed potential revenue. However, this leads me to believe that BidPrice is actually a very important indicator of long-term growth in bottom line.

Thanks for the opportunity to work on this and I look forward to continuing it!

Respectfully,

Ted Morrison
theodorejmorrison@gmail.com
708-490-9553