

Predicting Award Prices of First Price Sealed Bid Procurement Auctions

Fabian Blasch

Supervisor: Dr. Katharina Fenz

07/14/2022

Contents

1	Introduction	2
2	Data	2
2.1	Scraping	3
2.1.1	Text Based Information	4
2.2	Descriptive Statistics	6
3	Economic Operationalization	6
3.1	Auctioneer	6
3.2	Firms	6
4	Methods	6
4.1	Elastic Nets	6
4.2	Ensemble Methods	6
4.2.1	Random Forests	6
4.2.2	eXtreme Gradient Boosting	6
4.3	Nested Cross Validation	6
4.3.1	Logistic PCA	6
4.3.2	Recursive Feature Elimination	6
5	Results	6
5.1	Prediction	6
5.2	Unsupervised Colusion Detection	6
6	Conclusion	6

1 Introduction

2 Data

All the information about the procurement contracts, is obtainable through the bid tab archive on the official website of the Colorado Department of Transportation. The information is provided in PDF documents. In each of those documents the following information of the respective auction is covered.

- A table listing all submitted bids, including a unique identifier for each of the participating bidders
- An Engineer's Estimate
- The contract ID
- The letting date
- Either the amount of time in business days given to complete all the contracted obligations, or a completion due date
- The County in which the contract is to be completed in

For illustrative, Figure 1 displays an example of a bid tab, in particular the second page, which contains the vendor ranking as well as the contract description and the remaining information listed above.

Colorado Department Of Transportation

Printed On: 11/17/2015

Vendor Ranking

Page 1 of 1

Letting No: 20151112

Contract ID: C19868

Project(s): STU1211-084

Letting Date: November 12, 2015

Region: 1

Letting Time: 10:00 AM

Contract Time: 260 WORKING DAYS

Counties: JEFFERSON, REGION 1

Contract Description:

SH121(WADSWORTH)-HIGHLAND DR-10TH AVE-JEFFERSON CO

THIS PROJECT IS LOCATED ON WADSWORTH BETWEEN HIGHLAND AND 10TH.

CONSTRUCTION WILL INCLUDE A FULL CONSTRUCTION WITH WIDENING OF ONE LANE IN BOTH DIRECTIONS, AND A MULTI MODAL TRAIL ON BOTH SIDES. THE MAINLINE PAVING WILL BE CONCRETE. THE WORK ALSO INCLUDES A CONCRETE BOX CULVERT NEAR HIGHLAND TO CARRY LAKEWOOD GULCH UNDER WADSWORTH.

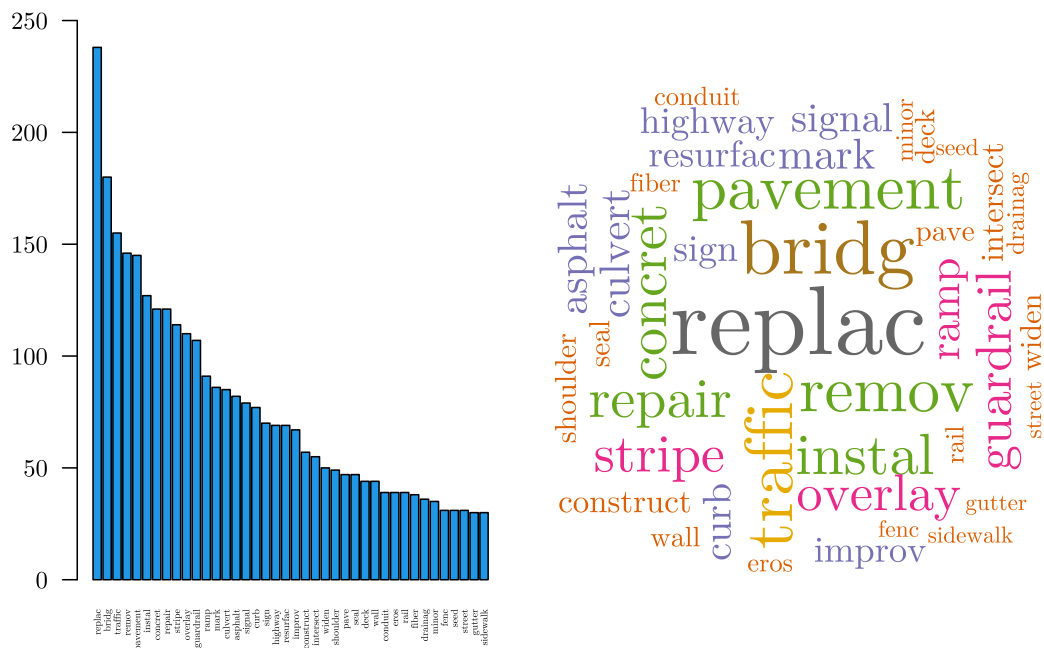
CDOT WILL ONLY BE ACCEPTING ELECTRONIC BIDS FOR THIS PROJECT. PLEASE CONTACT BID EXPRESS CUSTOMER SERVICE AT 1-888-352-2439 TO OBTAIN AN ACCOUNT IF NECESSARY.

Rank	Vendor ID	Vendor Name	Total Bid	Percent Of Low Bid	Percent Of Estimate
0	-EST-	Engineer's Estimate	\$9,821,027.20	91.58%	100.00%
1	870A	SEMA CONSTRUCTION, INC.	\$10,723,550.00	100.00%	109.19%
2	884A	HAMON INFRASTRUCTURE, INC.	\$10,817,000.00	100.87%	110.14%
3	1275A	CASTLE ROCK CONSTRUCTION COMPANY OF COLORADO, LLC	\$10,817,845.03	100.88%	110.15%
4	065A	CONCRETE WORKS OF COLORADO INCORPORATED	\$11,614,565.78	108.31%	118.26%
5	232A	AMERICAN CIVIL CONSTRUCTORS, INC. dba ACC Mountain West	\$12,338,888.00	115.06%	125.64%

Figure 1: Bid Tab Example

2.1 Scraping

In order to obtain all the archived bid tabs, the html code of the website was first examined using a google chrome extension called SelectorGadget. This tool allows one to identify html nodes, that website contents are associated with. In the case of the bid tab archive, the html node carrying the links to the individual bid tabs is “<td a>”. Once this html node is discovered and the consistency across different years in the archive is ensured, the download is easily achieved by looping over the links and downloading the respective PDF. The hyperlink extraction was performed utilizing *rvest*, by Wickham (2022). For the remaining steps in the data extraction, a distinction will be made for text based information and tabular data.



2.2 Descriptive Statistics

3 Economic Operationalization

3.1 Auctioneer

3.2 Firms

4 Methods

4.1 Elastic Nets

4.2 Ensemble Methods

4.2.1 Random Forests

4.2.2 eXtreme Gradient Boosting

4.3 Nested Cross Validation

4.3.1 Logistic PCA

4.3.2 Recursive Feature Elimination

5 Results

5.1 Prediction

5.2 Unsupervised Colusion Detection

6 Conclusion

References

Wickham, H. (2022). *Rvest: Easily harvest (scrape) web pages* [<https://rvest.tidyverse.org/>,
<https://github.com/tidyverse/rvest>].