

# CCT College Dublin

## Assessment Cover Page

---

Module Title:	Data Visualisation Techniques
Assessment Title:	CA 2 - Individual
Lecturer Name:	David McQuaid
Student Full Name:	Victor Ferreira Silva
Student Number:	2021324
Assessment Due Date:	28th May 2023
Date of Submission:	28th May 2023

### Declaration

By submitting this assessment, I confirm that I have read the CCT policy on Academic Misconduct and understand the implications of submitting work that is not my own or does not appropriately reference material taken from a third party or other source. I declare it to be my own work and that all material from third parties has been appropriately referenced. I further confirm that this work has not previously been submitted for assessment by myself or someone else in CCT College Dublin or any other higher education institution.

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Imports &amp; Configurations</b>	<b>2</b>
<b>3</b>	<b>Data Understanding</b>	<b>2</b>
<b>4</b>	<b>Insights</b>	<b>2</b>
4.0.1	Distribution of Bids Per Bidder . . . . .	2
4.0.2	Distribution of Bids Per Auction . . . . .	2
4.0.3	Distribution of Bidding Ratio By Class . . . . .	2
<b>5</b>	<b>Clustering Visualisation</b>	<b>2</b>
<b>6</b>	<b>Interactive Plots</b>	<b>2</b>
6.0.1	Interactive 2D Histogram Contour Plot . . . . .	2
6.0.2	Interactive 3D Scatter Plot . . . . .	2
<b>7</b>	<b>Dashboard</b>	<b>6</b>
<b>8</b>	<b>Conclusion</b>	<b>6</b>
<b>9</b>	<b>References</b>	<b>6</b>

# Assessment Task

Victor Ferreira Silva

May 28, 2023

## 1 Introduction

The ability to predict normal and abnormal bidding behaviour of eBay users can help companies identify scams and other undesirable users on the platform. This work has as object of study the Shill Bidding Data set (SBD) and aims to find insights and detect patterns that can help subsequent phases of analyses that can mitigate such behaviours. In turn, the SBD aims to facilitate academic research and support the advancement of machine learning techniques by providing a realistic and preprocessed auction dataset (Alzahrani and Sadaoui 2018) and basically it consists of eBay auctions that have various features, including auction duration, bidder tendency and class.

## 2 Imports & Configurations

## 3 Data Understanding

## 4 Insights

### 4.0.1 Distribution of Bids Per Bidder

Figure 1 shows Distribution of Bids Per Bidder.

### 4.0.2 Distribution of Bids Per Auction

Figure 2 shows Distribution of Bids Per Auction.

### 4.0.3 Distribution of Bidding Ratio By Class

Figure 3 shows Distribution of Bidding Ratio By Class

## 5 Clustering Visualisation

Figure 4 shows a k-means cluster Analysis of Bidding Behaviour using in a 3D Space

## 6 Interactive Plots

Two interactive plots were created: a 2D Histogram Contour Plot and a 3D Scatter Plot.

### 6.0.1 Interactive 2D Histogram Contour Plot

Figure 5 shows an Interactive Contour Plot.

### 6.0.2 Interactive 3D Scatter Plot

Figure 6 shows a Class Distribution in a Interactive 3D Scatter Plot

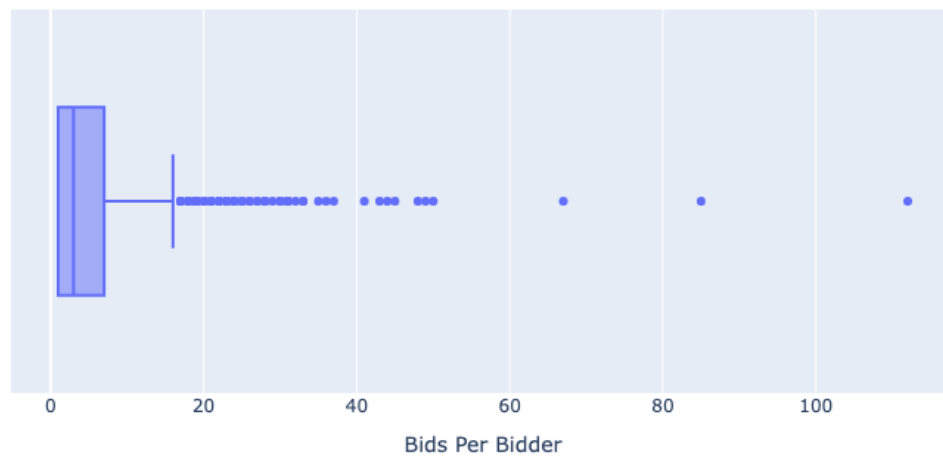


Figure 1: Distribution of Bids Per Bidder, illustrating the variability in the number of bids made by individual bidders

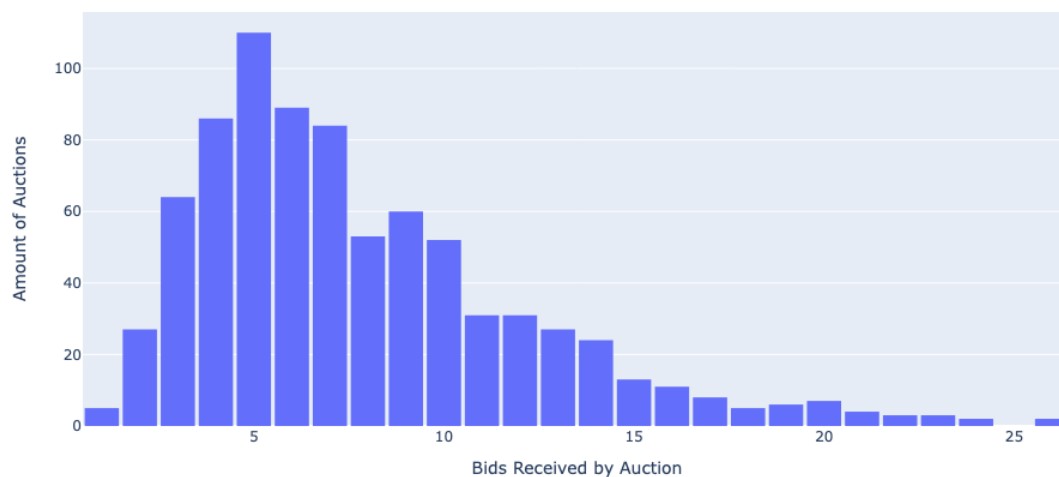


Figure 2: Distribution of Bids per Auction, demonstrating the frequency of bids received by each auction, with bins representing the count of bids

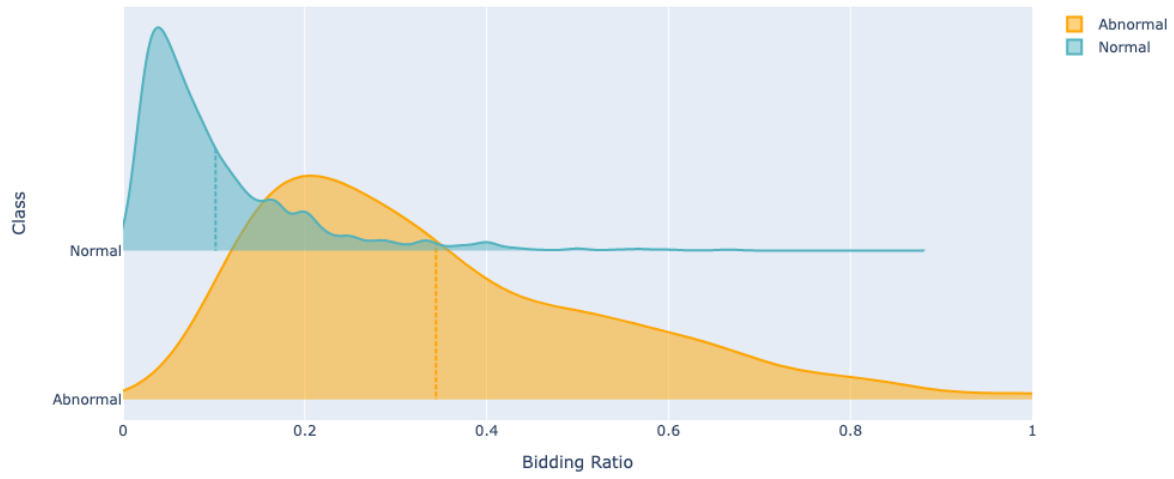


Figure 3: Distribution of Bidding Ratio by Class, illustrating the comparison between abnormal and normal bidding behaviour

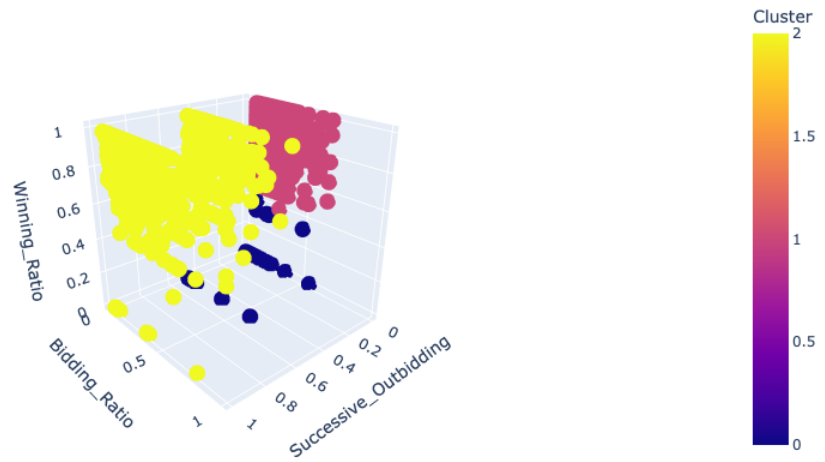


Figure 4: Visualisation of K-means Cluster Analysis in a 3D Space, illustrating the grouping patterns and relationships between successive outbidding, bidding ratio, winning ratio, and cluster assignments

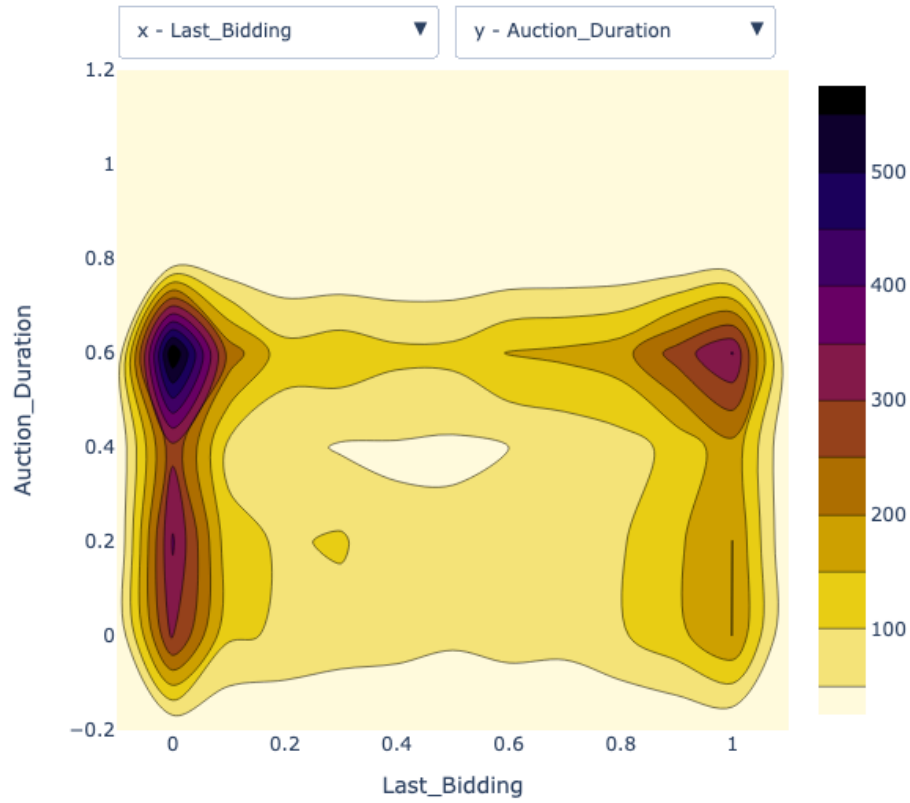


Figure 5: Interactive Contour Plot showcasing the relationship between Last Bidding and Auction Duration

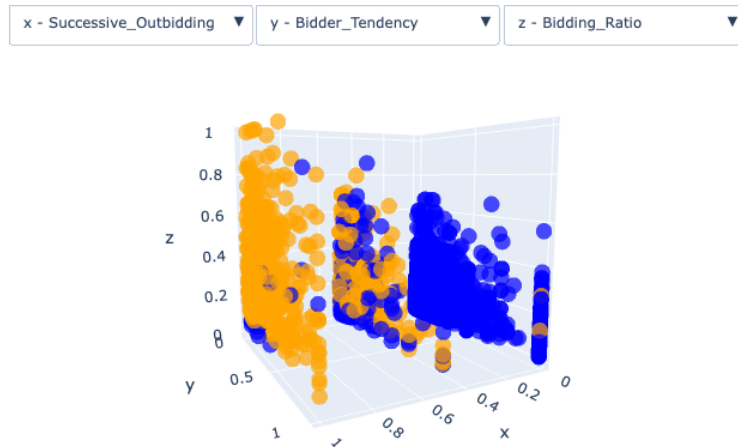


Figure 6: 3D Scatter Plot displaying bid-related features with adjustable dimensions and class markers. Blue markers represent normal bids, while orange markers indicate abnormal bids.

## 7 Dashboard

Figure 7 shows a Interactive Dashboard featuring an Interactive 2D Histogram Contour Plot and Interactive 3D Scatter Plot.

## 8 Conclusion

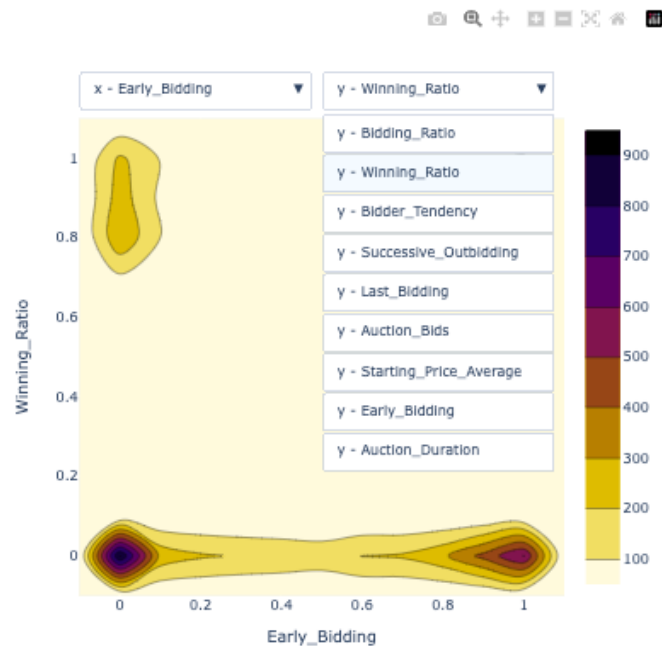
## 9 References

### References

- Alzahrani, Ahmad and Samira Sadaoui (June 2018). *Scraping and preprocessing commercial auction data for fraud classification*. URL: <https://arxiv.org/abs/1806.00656>.
- UCI (2018). *UCI Machine Learning Repository: Shill bidding dataset data set*. URL: <https://archive.ics.uci.edu/ml/datasets/Shill+Bidding+Dataset>.

## CA2 - Dashboard

### Interactive 2D Histogram Contour Plot



### Interactive 3D Scatter Plot

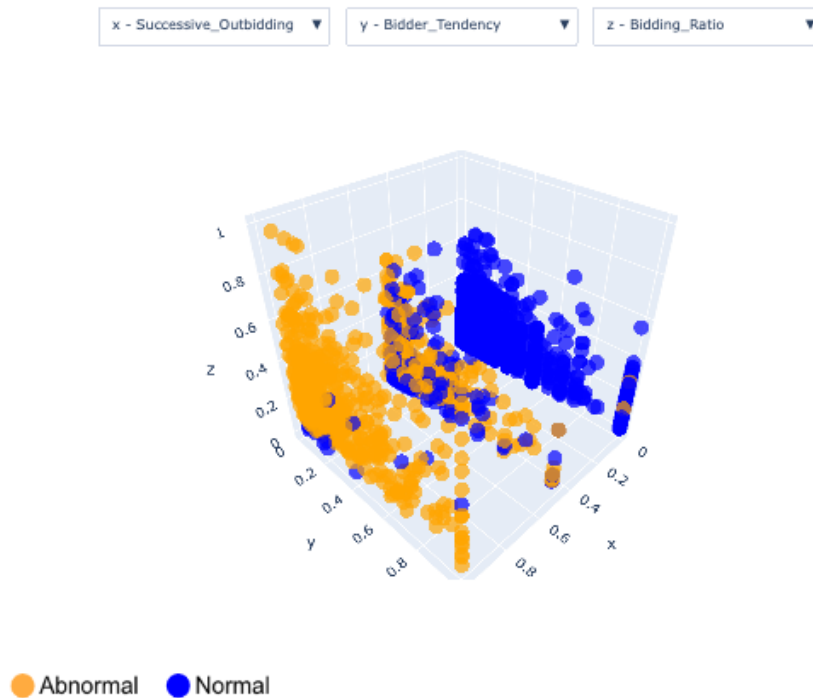


Figure 7: Screenshot of a dashboard powered by Dash library, featuring an Interactive 2D Histogram Contour Plot and Interactive 3D Scatter Plot.