

Keeneland September Yearling Auction Prediction 3T's

Tapan(Taps) Mandal, Thinh Nguyen, Trung Pham

Table of contents

01

**Background
Information**

02

**Problems
Opportunities
Approach**

03

**Buyer
Segmentations
Analysis**

04

**Price Prediction
Model**

05

Key Takeaways



Background information

Equine industry



Economic Contribution

The industry's \$177 billion impact on the U.S. economy



Job Support

The industry's role in providing 2.2 million jobs



Horse Population

The 6.6 million horses in the U.S.



Horse Racing Revenue

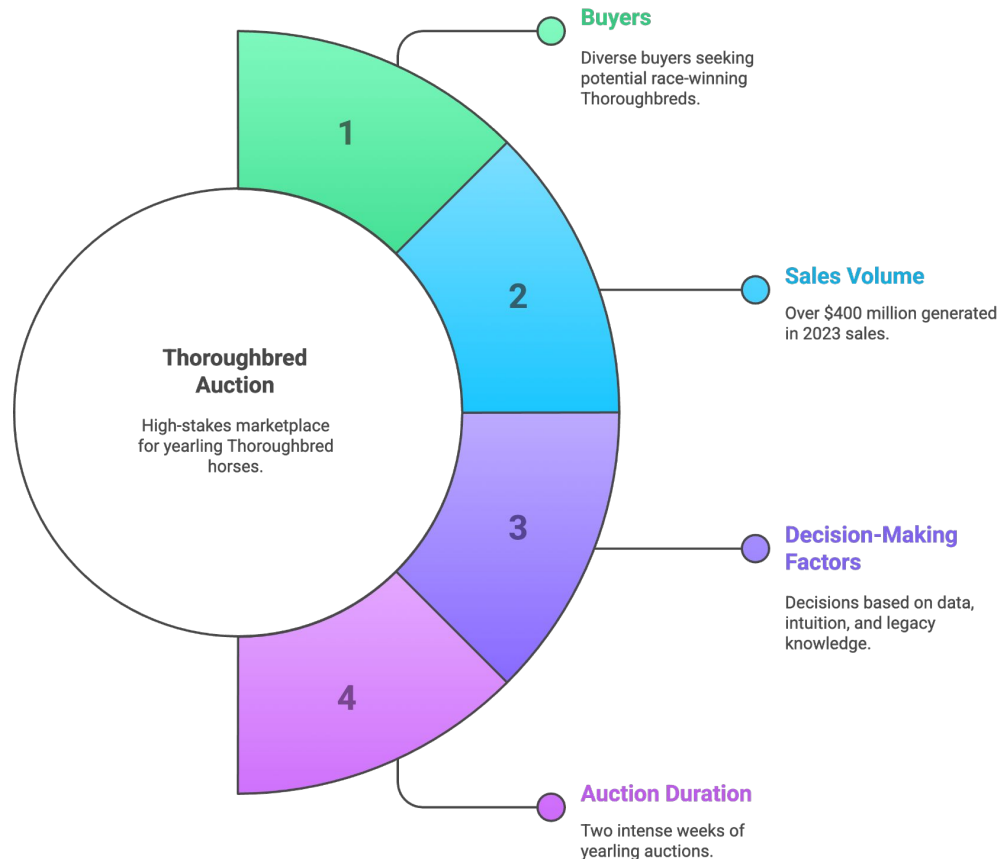
The \$38 billion generated by horse racing



Global Market Growth

The expected \$599.5 billion growth in the global horse racing market

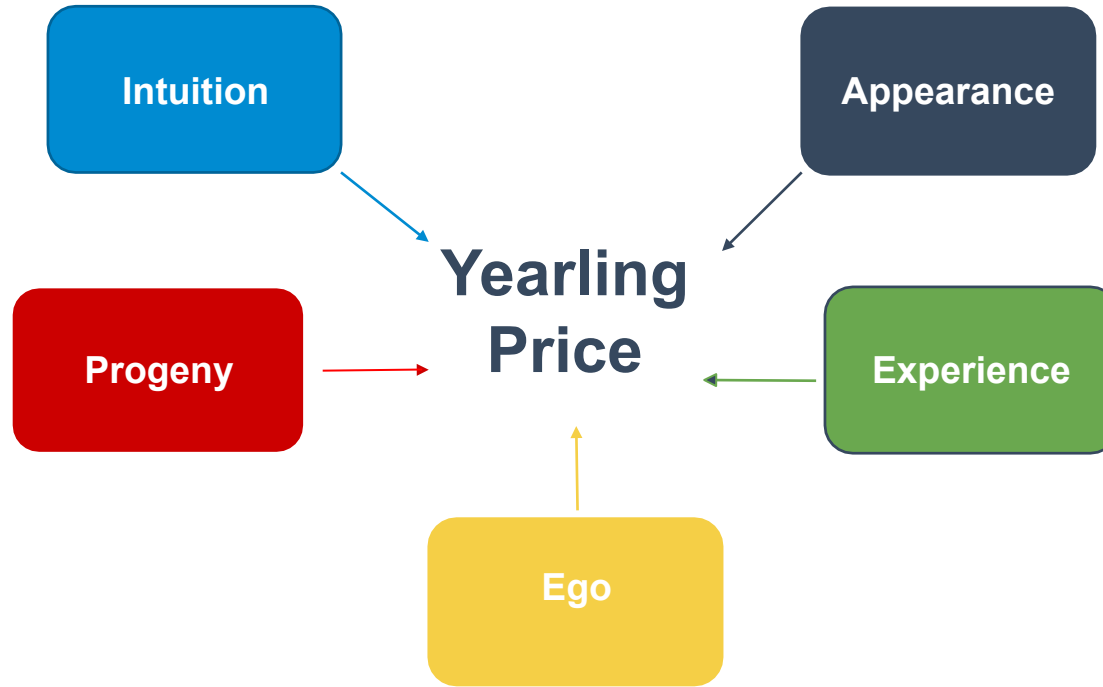
Keeneland September Yearling Auction





Problems Opportunities Approach

Current Pricing Strategy



Opportunities

- Utilizing logical and data-driven approach to setting realistic price
- Segment buyers to make better strategies for each segment.

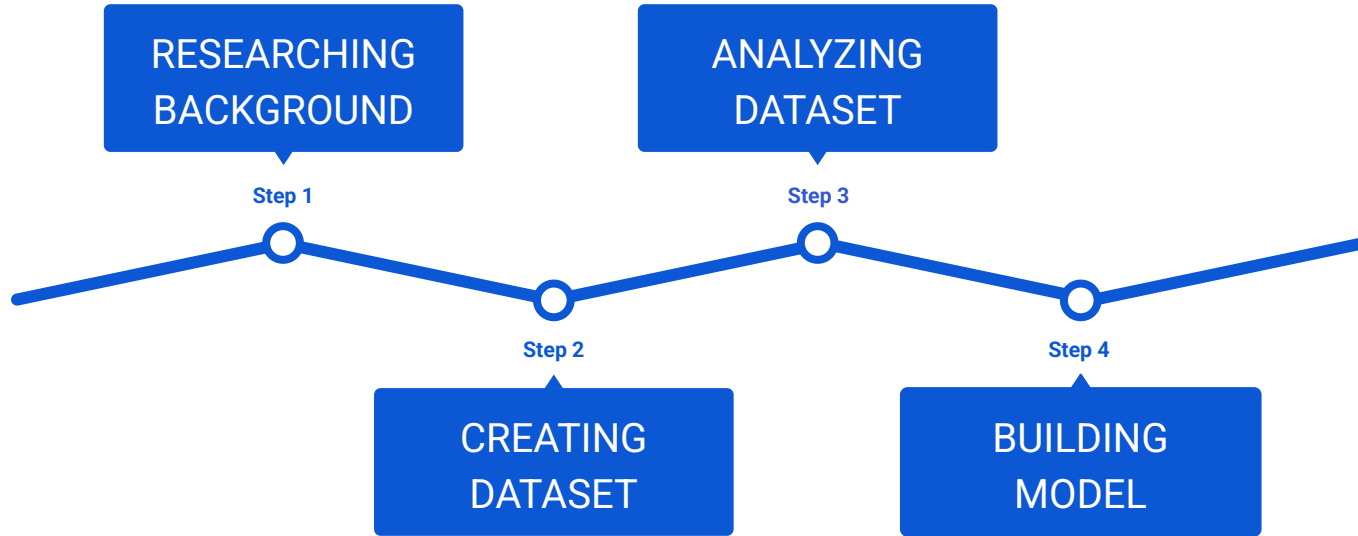
We are considering

- 5 years of Sales data (2020 - 2024)
- Sire, dam, consignors, buyers
- Pedigree information

We are NOT considering

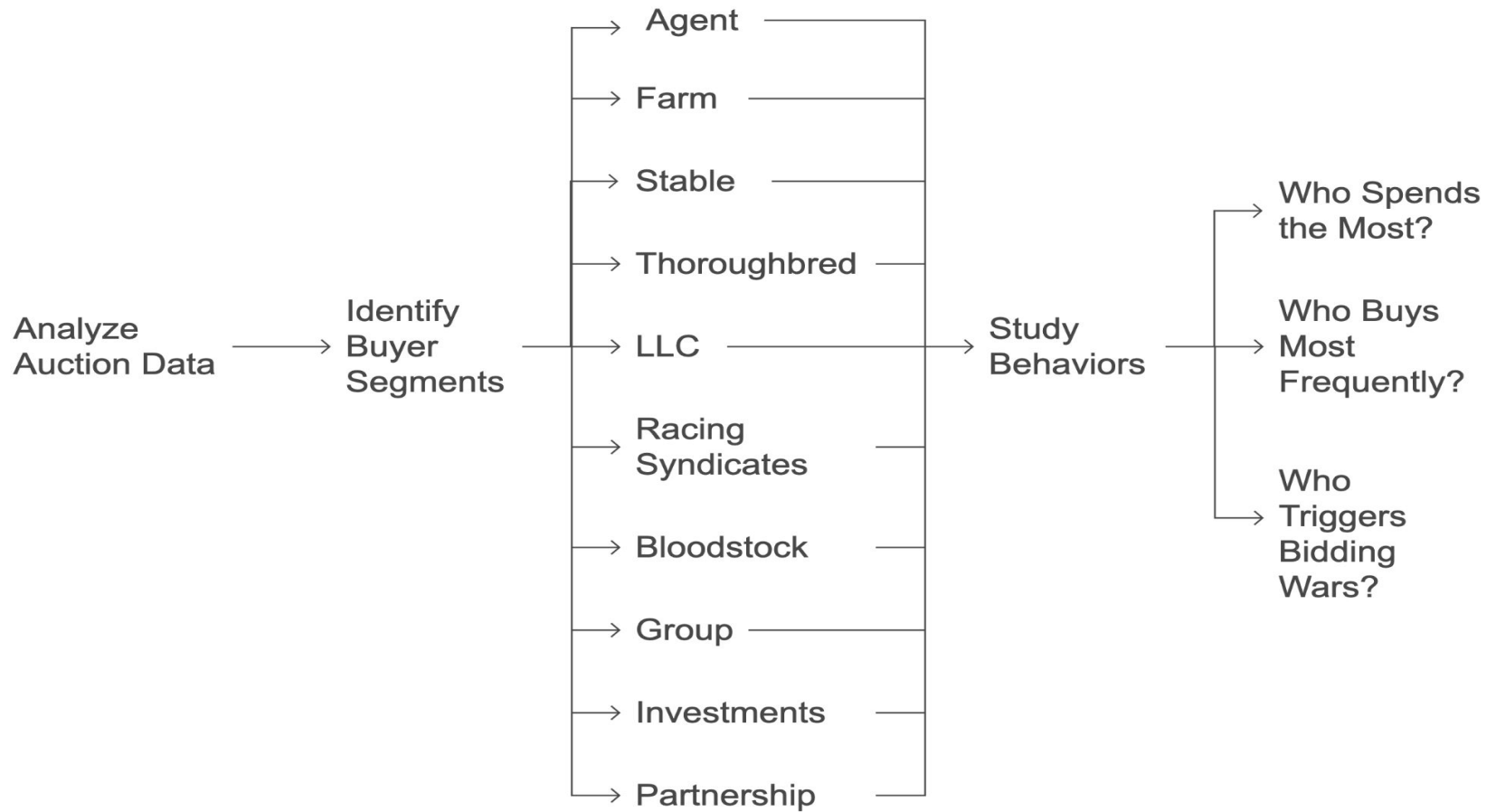
- Prices of other auction
- Private sales
- Career of the horse after the sales

Approach

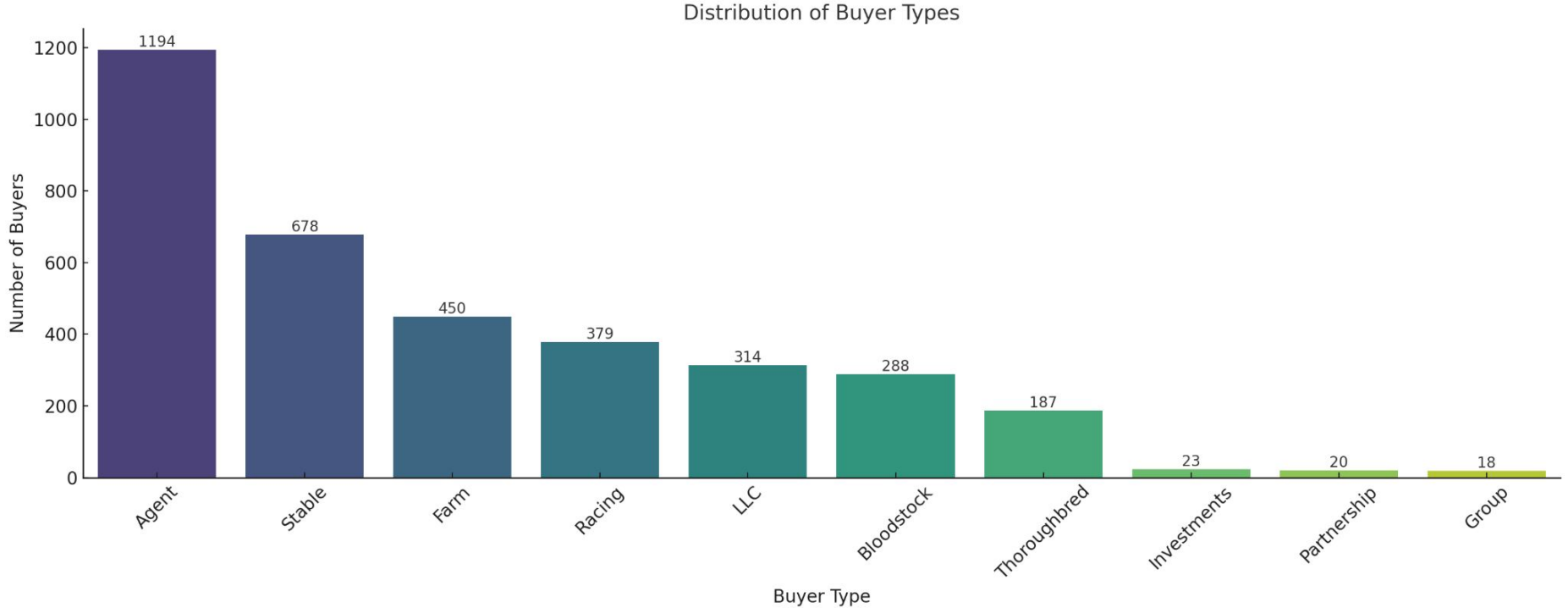




Buyer Segmentations Analysis

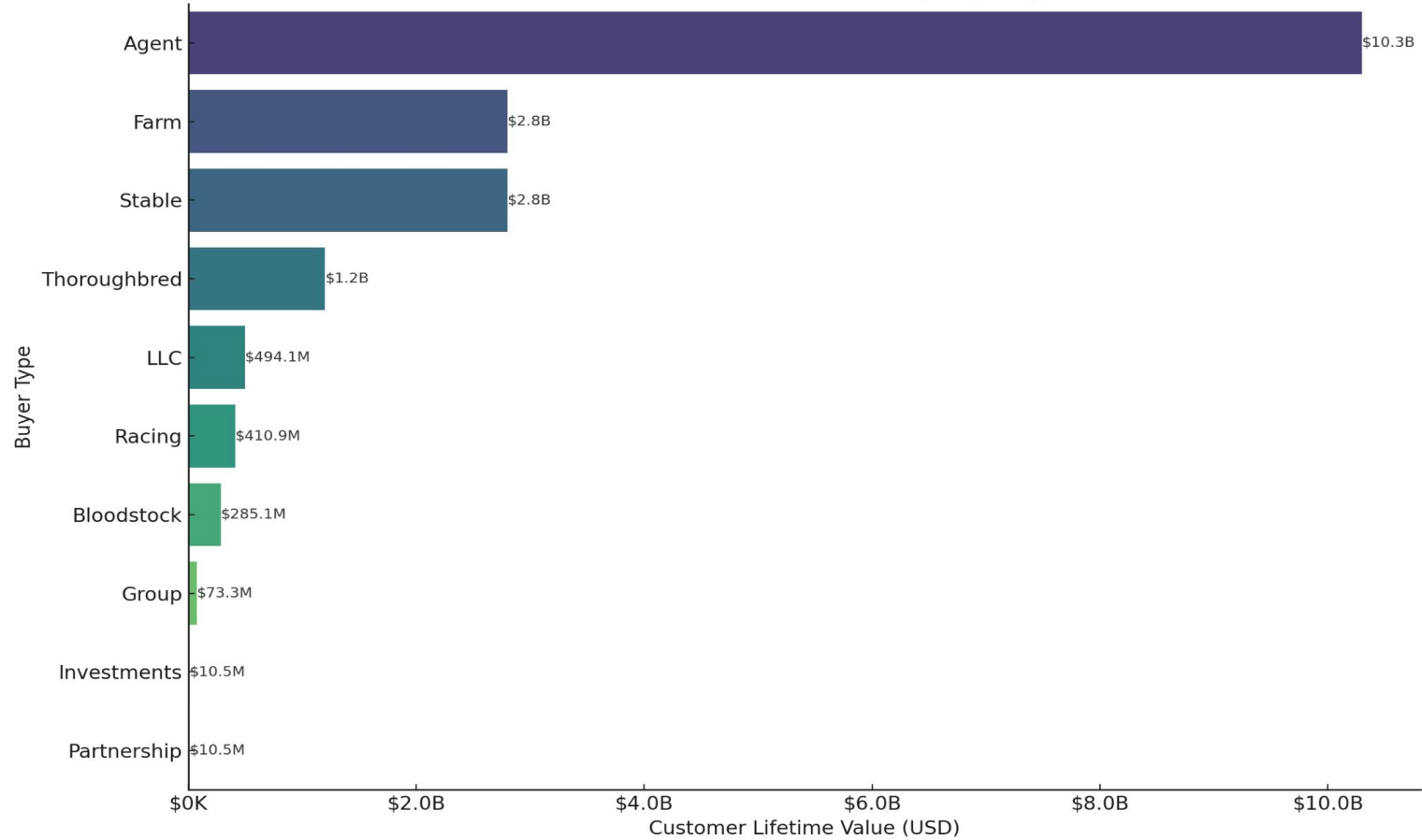


Type of Buyers

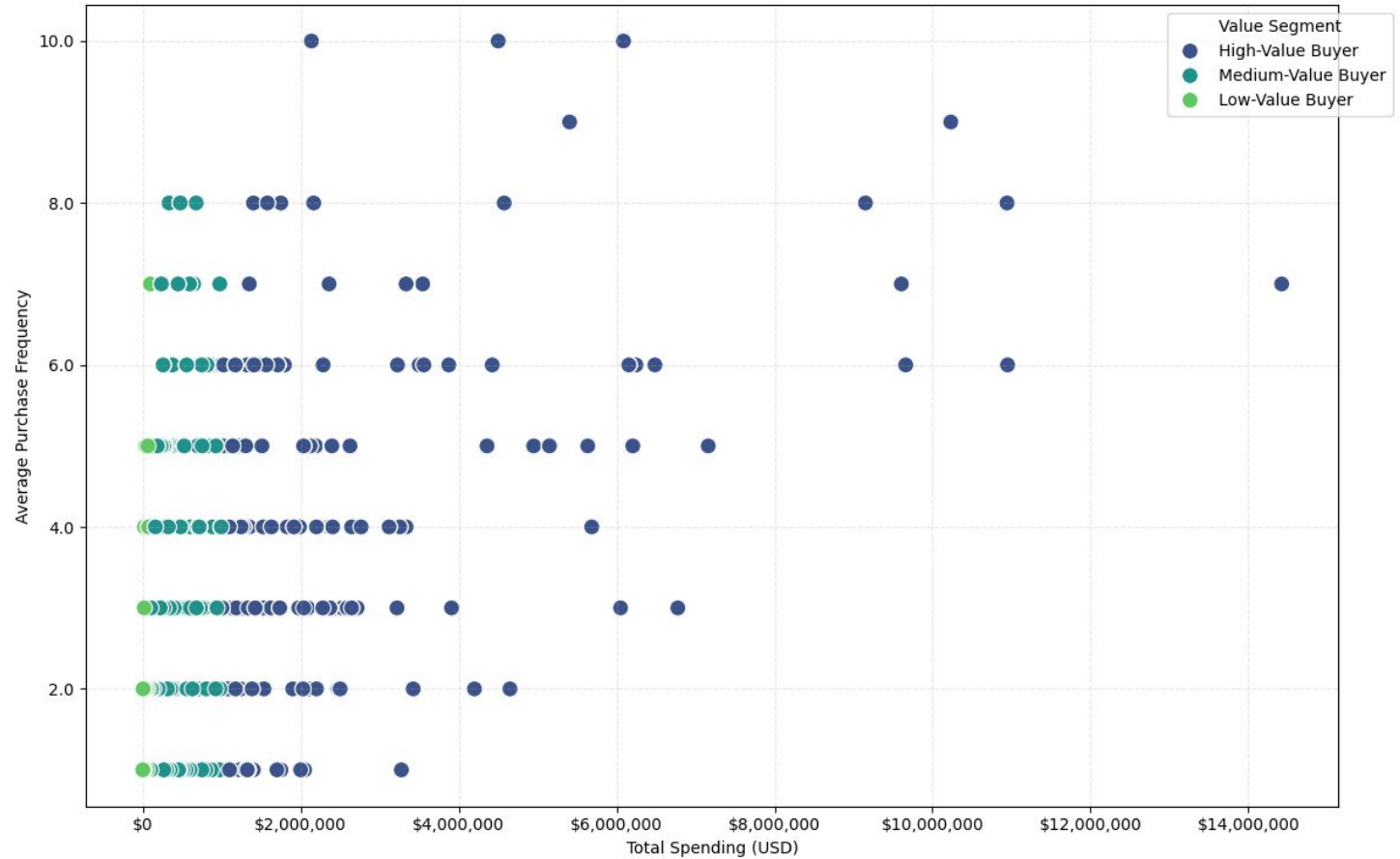


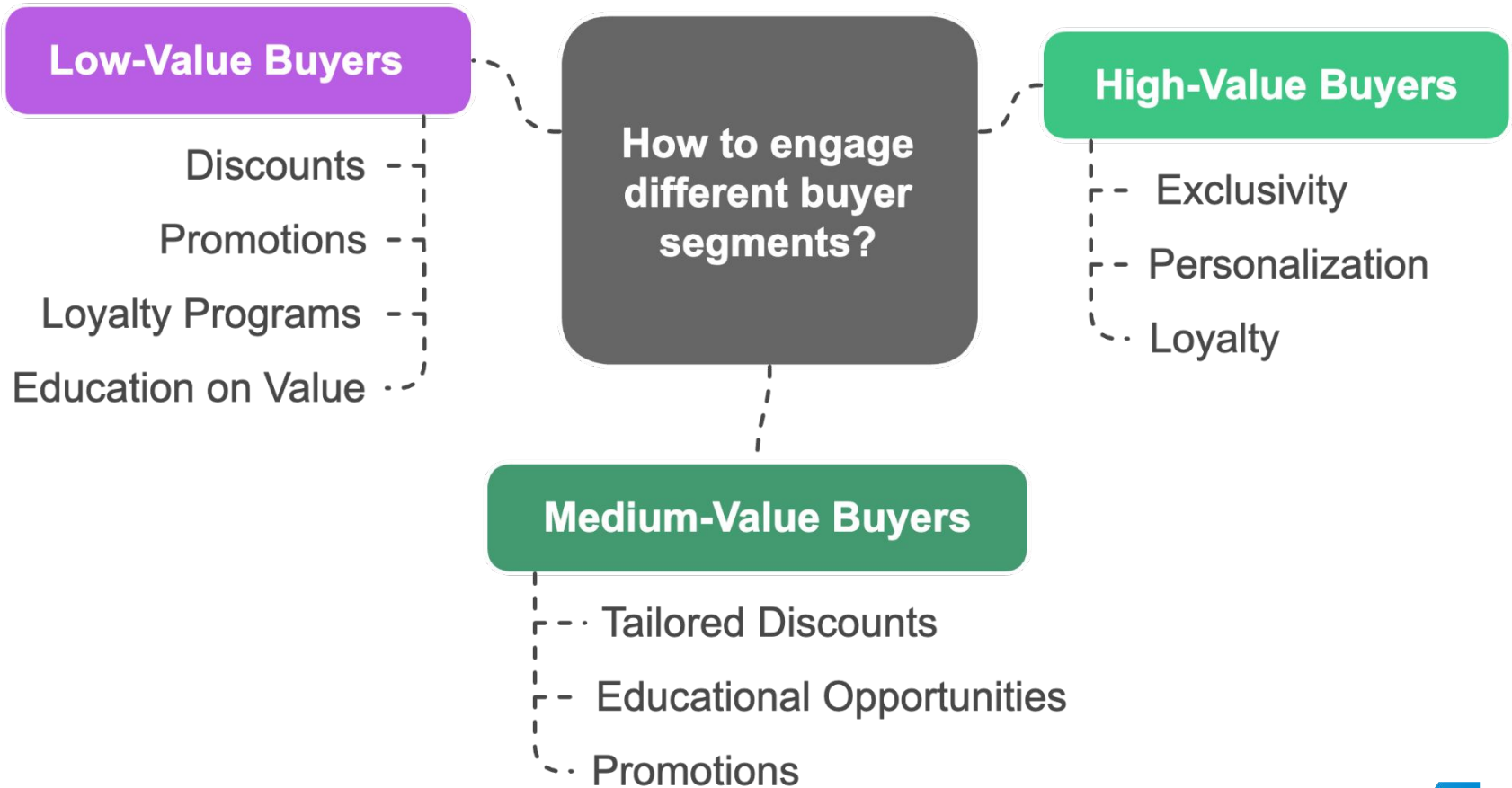
Total Buyers: 3550


Customer Lifetime Value (CLV) by Buyer Type



Buyer Segmentation by Spending & Frequency







Price Prediction Model

Demo

Quick guide

Step 1: Enter Horse Basic information

Input sire/dam names, seller, session number, gender, and color from your Keeneland documentation.

Step 2: Enter pedigree information

Visit [FREE 5-Cross Pedigree](#) and search using either the horse's name (if known) or the dam's name to retrieve: sire statistics, broodmare sire statistics and dosage profile.

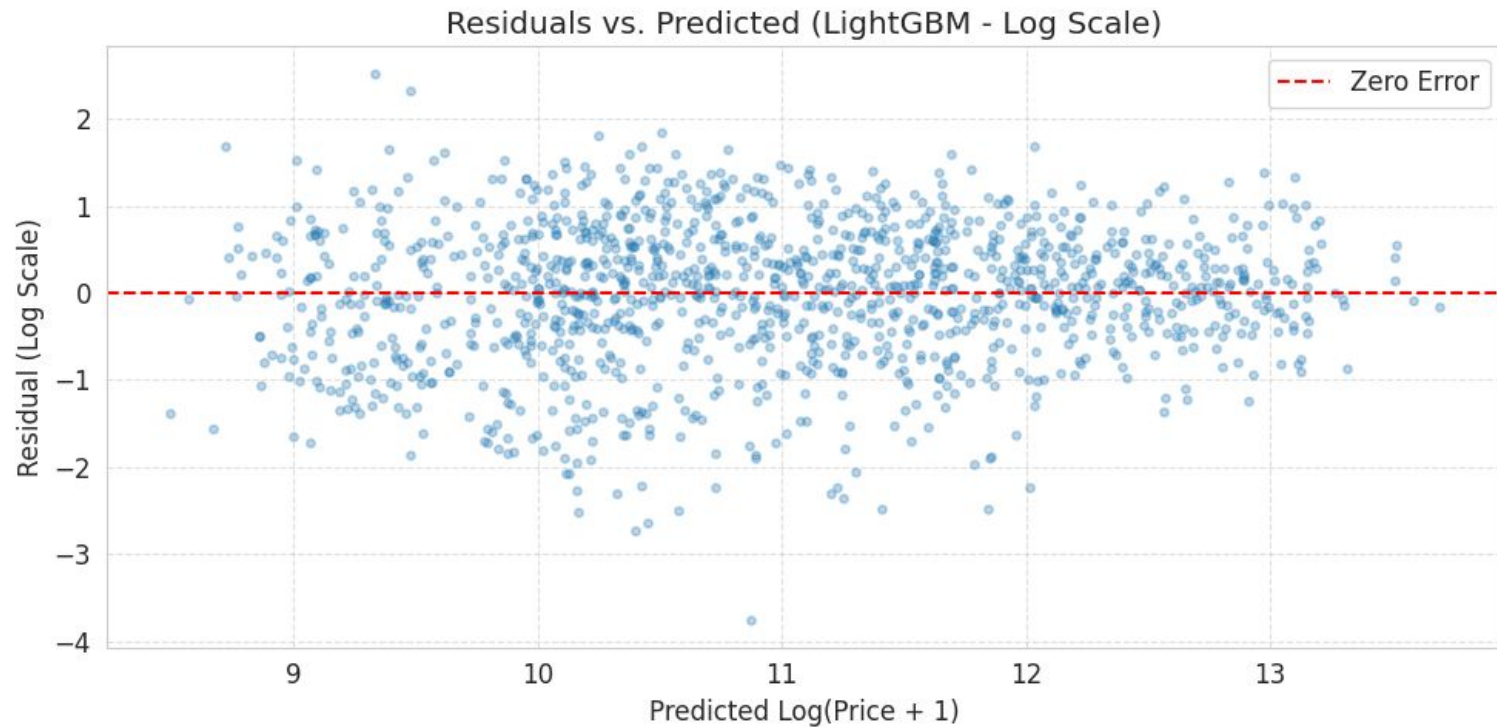
Step 3: Generate Price Prediction

Click the "Predict Price" button to receive the estimated auction price and plausible range (90% Conformal Prediction Interval).

Major Price Drivers

1	Session
2	Sire and dam reputation
3	Sire's foals average earning index

Model Assessment



Model Assessment

	Log MAE	Log R ²	MAE	R ²
Train	0.54	0.74	\$59,421	0.63
Validation	0.62	0.70	\$72,042	0.54
Test	0.62	0.69	\$64,363	0.57

Current Limitation

- Buyer Segmentation Analysis:
 - Relying solely based on the **given buyer names**
 - Some types of buyers **could overlap** with each other.
- Price Prediction Model:
 - Overall accuracy needs further refinement
 - Model did **not cover** some important factors such as stud fees, physical traits, injury records
 - Model needs a better way to normalize the price over the years

What Would We Do To Improve?

- Gather more data:
 - Longer time range
 - Important data that we could not get access to (stud fees, physical traits, injury record,...)
 - Deeper look into horses' family
 - More advanced data gathering methods (computer vision,...)
- Research more about economic and auction dynamic factors
- More advanced modelling approaches
- Gain experience from people in the field



Key Takeaways

Equine industry



Provided background information about the equine industry and Keeneland September Yearling Auction

Buyer segmentation analysis



Segmented buyers into distinct groups based on their organization name and analyzed their purchasing behavior

Price prediction model



Developed a predictive model based on data of Keeneland September Yearling to set realistic expectations for the price and identify valued horses

KEENELAND 

HIP 977

\$1,100,000

Questions?





Thank you

Sources

- <https://www.equineline.com/>
- <https://www.equibase.com/>
- https://uknowledge.uky.edu/agecon_facpub/14/
- <https://ageconsearch.umn.edu/record/260085?v=pdf>