

# GAME ON: Unveiling the future of Women's Basketball ticket sales with predictive analytics

**Team Data Jedi**

Akanksha Singh  
Aravind Teja Chikoti  
Harshraj Jadeja  
Santosh Pawan Kurada  
Sathwik Kanukuntla

"I think the beauty of women's basketball is seeing the passion, the teamwork, the camaraderie, and the pure joy of playing the game. It's not just about the sport; it's about lifting each other up and proving every day that we belong on this court just as much as anyone else."

SUE BIRD  
(WNBA PLAYER)

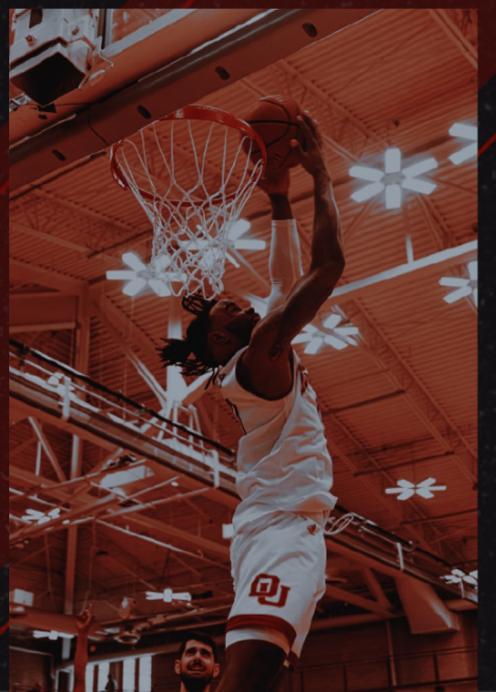




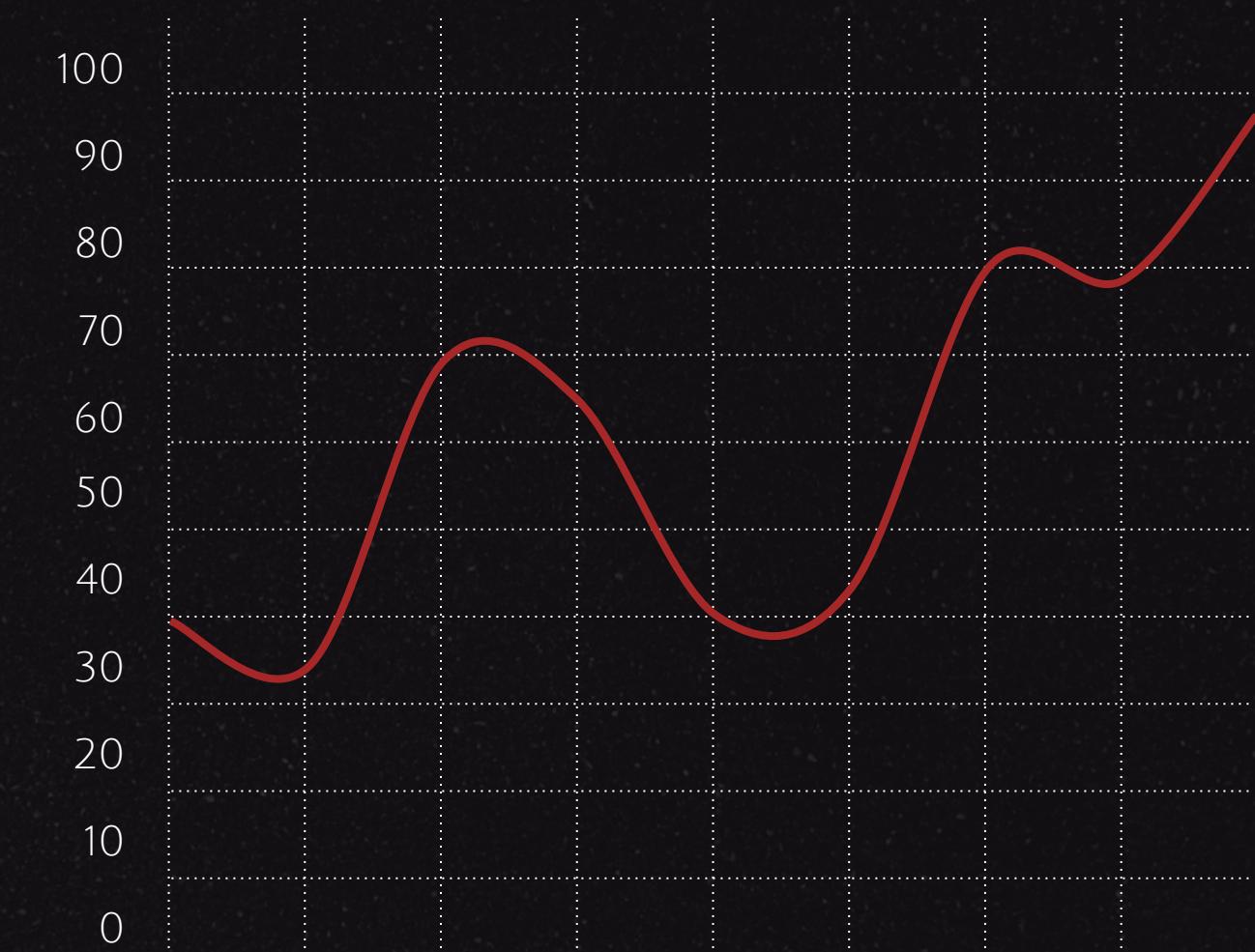
The viewership for the Women's NCAA Basketball Tournament has seen an increase, with the championship game's viewership rising by over 30% in recent years



The number of women participating in NCAA basketball has grown by approximately 25% since the early 2000s



Approximately 6% of NCAA Division I women's basketball players go on to compete professionally, either in the WNBA or overseas



The Women's NCAA Tournament has seen a surge in its economic impact, with host cities reporting over \$20 million in revenue from the event, underscoring the significant financial contribution of women's basketball to local economies.

# THE 'WHY' OF PREDICTION

Develop predictive models using Division I Women's Basketball customer data, supplemented with external datasets, to accurately forecast whether a customer will purchase a ticket and through which market (primary or secondary).

## Objective

## Stakeholders

### NCAA

- Increased Revenue
- Customer Insights
- Operational Efficiency

### For Customers:

- Improved Experience
- Accessibility to tickets

### For Sponsors & Partners

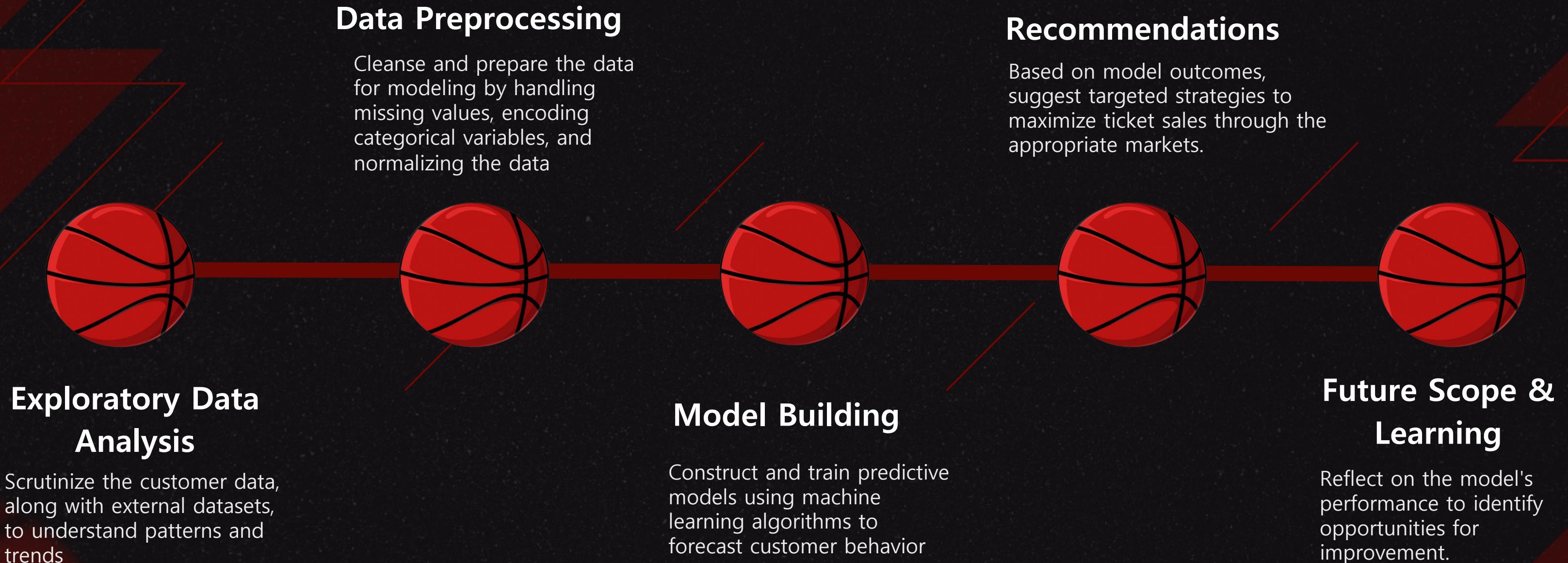
- Brand Exposure
- Data-Driven Partnerships

## Activity Types:

- No Activity
- Primary Purchase
- Secondary Purchase
- Other Secondary Activity
- Multiple Activities
- Transfer Recipient

## Target

# SOLUTION APPROACH



# DATASET INFORMATION



**No. of Rows & Columns**  
200k Rows & 25 Features



**10+**  
Additional Features



**6**  
Unique Activity Types



**Level of data**  
Record ID



**23+**  
Teams Accounted For



# DATA PREPROCESSING

## Layer 1 | Null Imputation & Feature Engineering



### Feature Engineering

- Duration for all events, purchases and actions
- College Basketball Rankings
- Engagement Score (Based on the email engagements of customers)
- Temporal attributes
- Consumer information by zipcode



### Null Imputation

- Unknown for Categorical Columns (Nulls, 'Unk', 'NA')
- Median imputation for Numerical columns

All the features have been scaled using Standard Scaler separately for test and train dataset

# DATA PREPROCESSING

## Layer 2 | Dropping columns with high multicollinearity

Before

feature	VIF
const	2748.429405
RecordID	1.310240
CustomerInstitutionAffinity	1.002686
IsCustomerInNCAAMembership	1.002162
HasCustomerClickedOrOpenedEmailsSixMonthsPrior	2.906244
CustomerFirstWBBActionDate	7.715084
CustomerFirstWBBPurchaseDate	6.809485
EventRoundName	21.954628
EventSession	73.708241
EventBeginDate	40.335491
HostingInstitution	4.774687
ActivityType	6.226794
DurationForAction	2.497854
DurationForPurchase	2.407557
DurationForEvent	4.814581
ZipCodeDistance	1.306582
Is_event_first_weekend	16.974883
Is_event_last_weekend	5.243461
DaysToNextEvent	1.331038
population	4.503272
density	6.630137
Ranking	3.237788
EngagementScore	2.770337

After

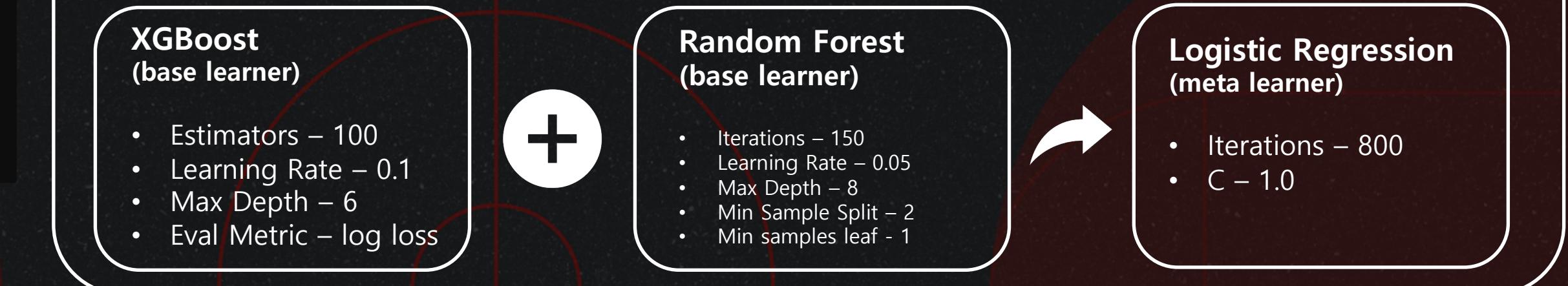
feature	VIF
const	269.498964
RecordID	1.600517
IsCustomerInNCAAMembership	1.004329
HasCustomerClickedOrOpenedEmailsSixMonthsPrior	1.041300
CustomerFirstWBBActionDate	1.651902
CustomerFirstWBBPurchaseDate	1.569928
EventRoundName	11.065122
EventBeginDate	3.689953
ActivityType	6.363331
DurationForAction	2.912807
DurationForPurchase	2.847593
ZipCodeDistance	1.255875
DaysToNextEvent	1.557782
CustEventActionPhase	1.211238

The VIF significantly reduced after columns like "CustomerInstitutionAffinity" were dropped

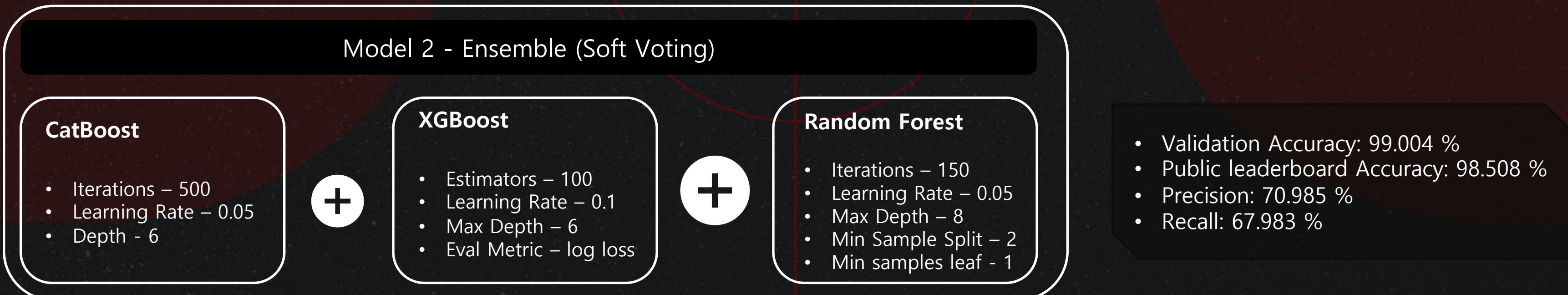
# PREDICTIVE MODELING

- Validation Accuracy: 99.100 %
- Public leaderboard Accuracy: 98.575 %
- Precision: 71.031 %
- Recall: 69.673 %

Model 1 – Stacking Classifier



Model 2 - Ensemble (Soft Voting)



# RECOMMENDATIONS | ENHANCING TICKET SALES



## Loyalty & Rewards program

Introduce loyalty programs that reward customers for purchasing through the primary market



## Data Driven Event Management

Use predictive insights to better plan event logistics, including venue selection & staffing levels



## Use External Data

Include economic trends, social media sentiment analysis, etc. to improve model performance



## Feedback Loops

Implement mechanisms to capture customer feedback post event to refine the model



## Enhance Customer Service

Understanding customer's likelihood to purchase can help representatives provide tailored assistance

# FUTURE SCOPE | APP-BASED TRACKER



## Benefits to NCAA

- Real-time sales tracking
  - Fan-Insights
  - Goal-Setting & Tracking
  - Community Building
  - Optimize Pricing & Promotions
  - Real-Time Feedback
  - Enhanced User Experience

TO A FUTURE OF WOMAN BASKETBALL DOMINANCE



THANK YOU!