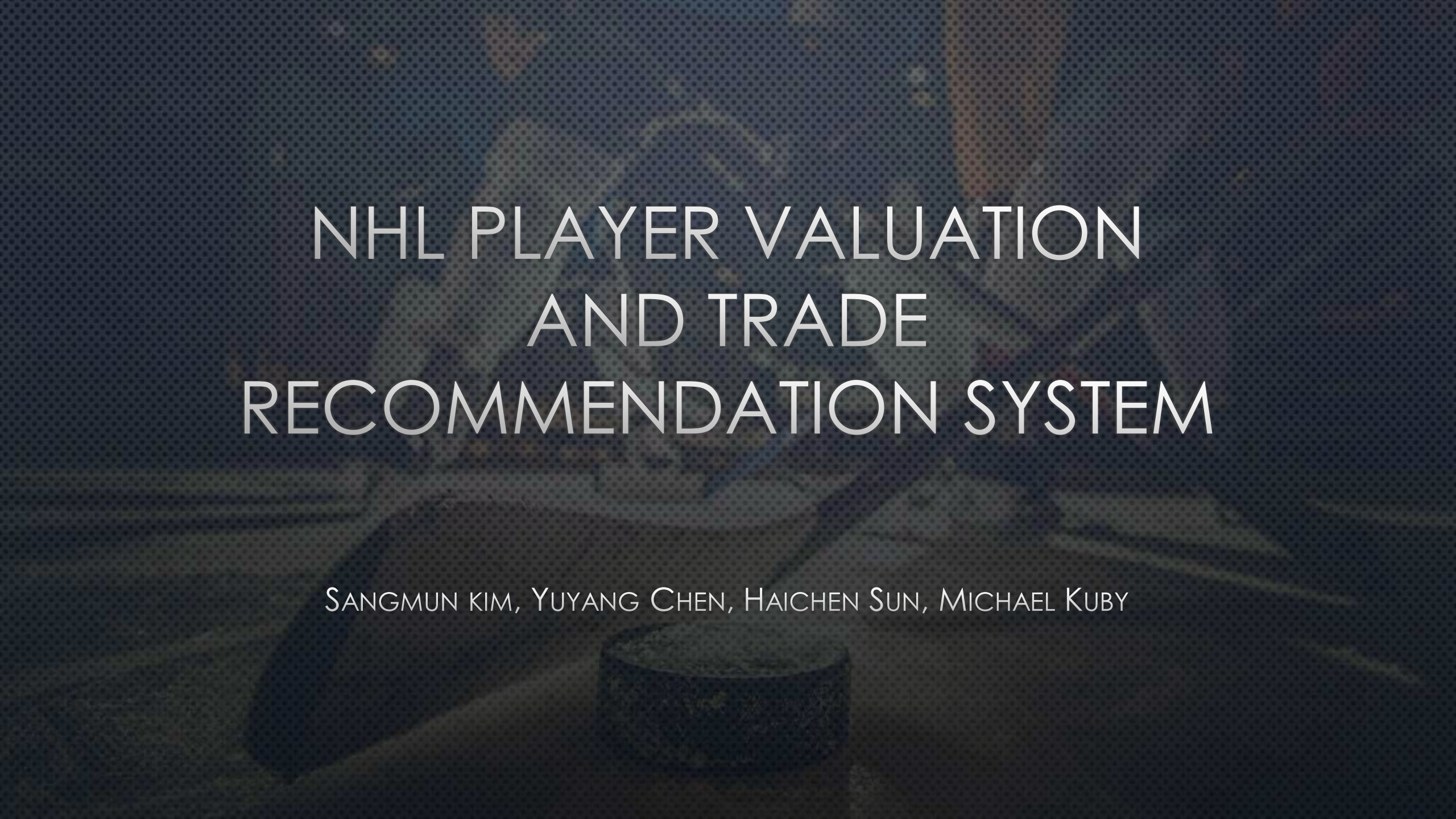


NHL PLAYER VALUATION AND TRADE RECOMMENDATION SYSTEM

SANGMUN KIM, YUYANG CHEN, HAICHEN SUN, MICHAEL KUBY

CHALLENGE: PREDICT THE SALARY OF A PLAYER BASED ON PERFORMANCE METRICS

RECALL: $Y = X\mathbf{w}$

X: PERFORMANCE METRICS FOR 2007 - PRESENT

Y: SALARY FOR 2007 – PRESENT

w: TO BE LEARNED

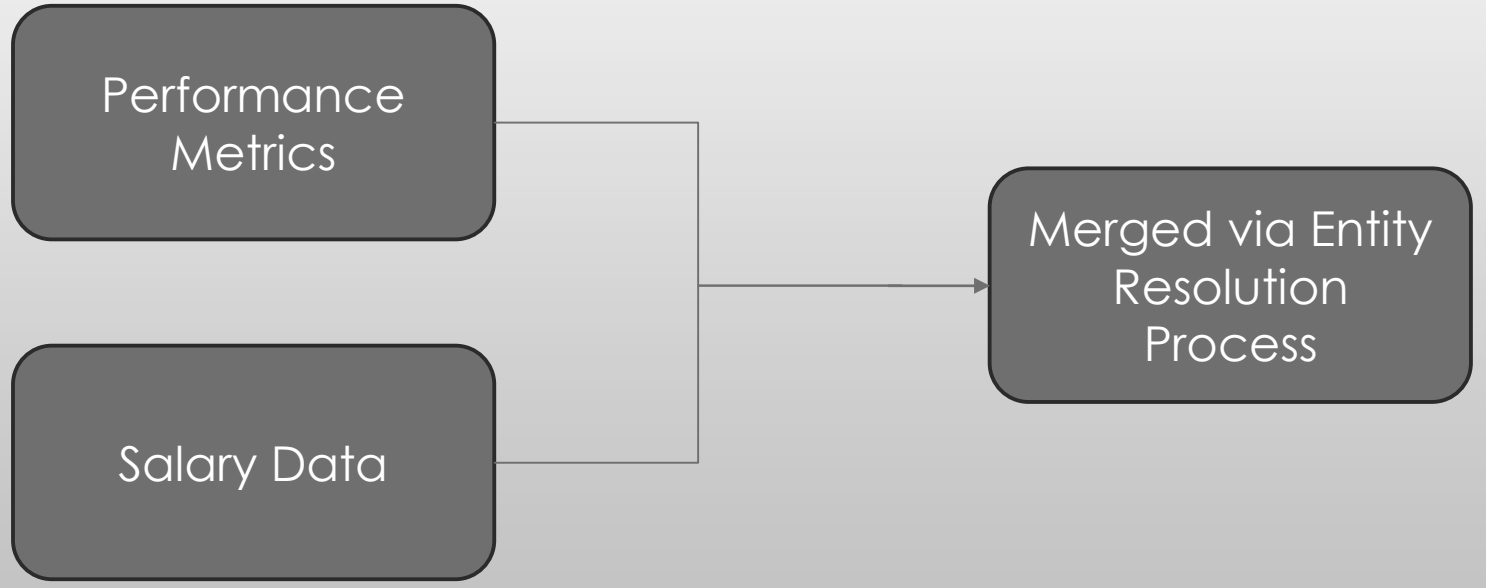
```
>>> df.head()
  Position Team         Player             TOI/GP  GP  ... Draft Team Draft Round Round Pick Overall Draft Position  Year
0        R  DET   Aaron Downey  4.5880952380952  56  ...           –           –          –                       –  2007
1        D  NYI  Aaron Johnson  13.861666666667  30  ...         CBJ           3         22                      85  2007
2        D  VAN   Aaron Miller  17.336549707602  57  ...         NYR           5          4                      88  2007
3        D  CBJ     Aaron Rome  18.178431372549  17  ...         L.A           4          8                     104  2007
4        L  MIN    Aaron Voros  9.1745454545455  55  ...         N.J           8          2                     229  2007

[5 rows x 97 columns]
```

Workflow Step 1: Procure the Data

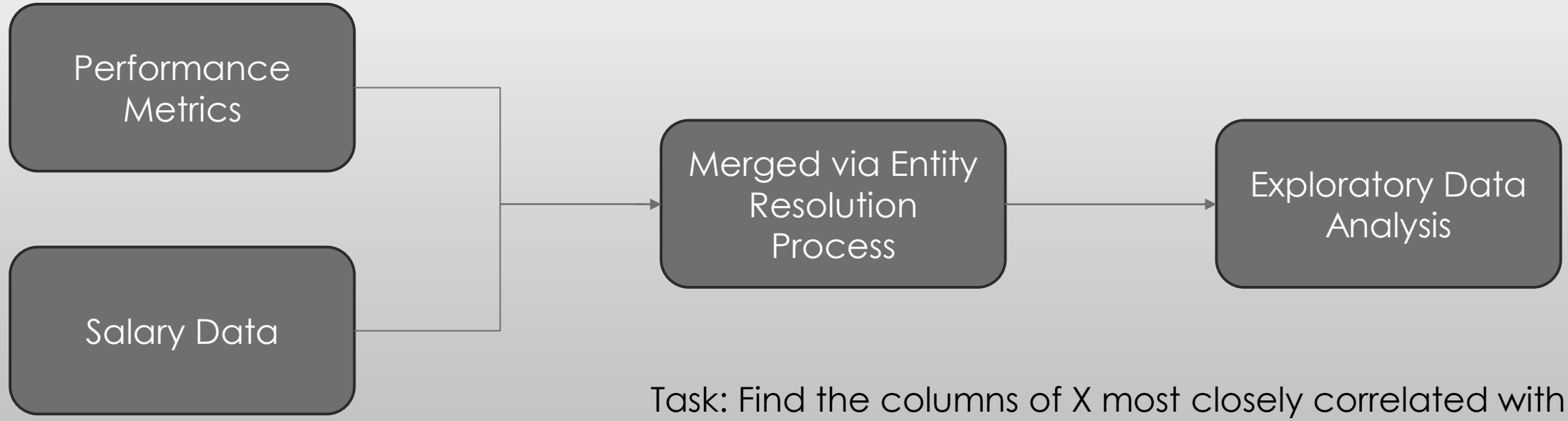
X (Performance Metrics): Scraped from www.NaturalStatTrick.com

Y (Salary Data): Scraped from www.CapFriendly.com

WORKFLOW STEP 2: ENTITY RESOLUTION



WORKFLOW STEP 3: EXPLORATORY DATA ANALYSIS



Task: Find the columns of X most closely correlated with Y.
Tools: Correlation Analysis, Matplotlib/Seaborn
Other: Principal Component Analysis

WORKFLOW STEP 4: TRAIN A REGRESSION MODEL

Decisions and Challenges:

- Multiple models (one for each position)?
- Format for our Y's (how to adjust for inflation)?
- Tools (AutoML)?

WORKFLOW STEP 5: PREDICT

Understanding the predictions:

"Predictions are what the model thinks the player is *worth*, given their *performance*."

A measure of **value**.

WORKFLOW STEP 6: PLAYER EVALUATION

Consider: Every player **already has a salary.**

Recall: Prediction = value

Player Evaluation:

If Value – Salary > 0:
Player is **overperforming**

If Value – Salary < 0:
Player is **underperforming**

WORKFLOW STEP 7A: RECOMMEND

Recommend: Given a player that we would like to trade, what players should we **target**?

Player to trade: **underperformer**
Player to target: **overperformer**
Constraint: Should have a comparable **salary.**

WORKFLOW STEP 7B: RECOMMEND

Time Permitting: Include sentiment analysis by analysis /r/{hockey, teamA, teamB, …}

Players to target: **overperformer + sentiment analysis**

WORKFLOW STEP 8: PRODUCT

Product: A web-based interface

Features:

1. Select Team → Select Player → View valuation and relevant metrics.
2. Produce a list of players identified as possible trade targets.


THANK YOU!