110-2 OR Final Project
CPBL Starting Lineup Selection

經濟四 錢紫翎 B07303129

生傳四 盧家雯 B07610058

經濟四 柯逸萱 B07303045

經濟四 陳宛婷 B07303106

Table of contents

01

Problem

Problem Description

In baseball world, deciding the starting lineup is a crucial part of the game. For each game, optimizing the lineup to have a good balance between players' offensive and defensive abilities is often the important issue for many fans.

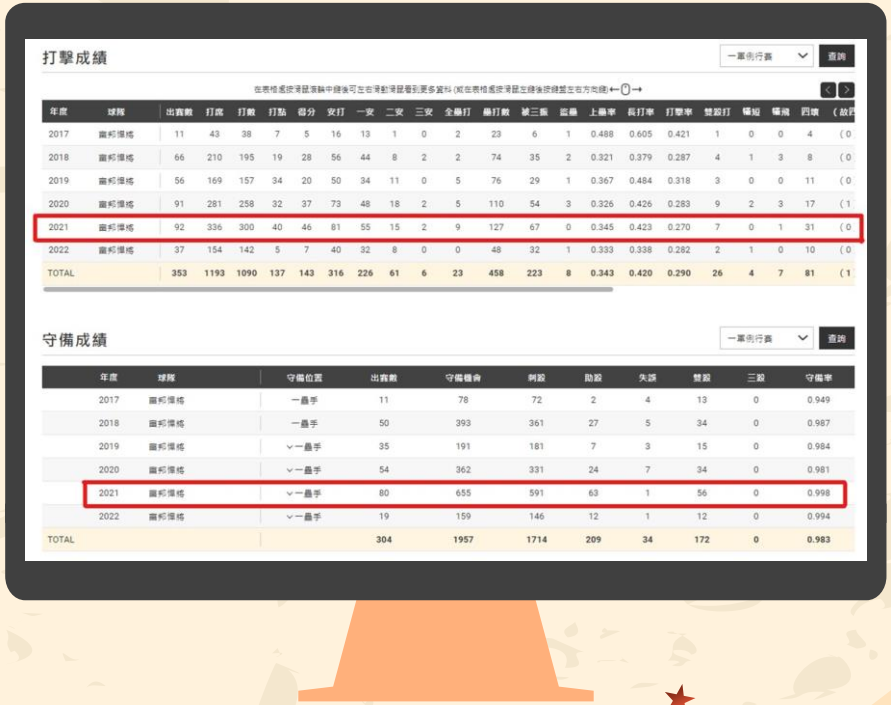
In this project, assume we are the manager of a professional team where we are given the schedule and opponent's starter pitcher of each game. We want to simulate the real-world situation to find the best lineup for our team in each game.

02

Data

Data: player stats of 2021 season

1. Use a crawler to get player stats of the 2021 season from the CPBL website.



Data: 2022 schedule

2. Get schedule of matches of the 2022 season and the opponent's starter pitcher from the CPBL website.

日期	場次	對手	場地	先發投手
4/3(日)	2	中信兄弟	洲際	象魔力
4/4(一)	4	中信兄弟	新莊	鄭凱文
4/7(四)	8	樂天桃猿	新莊	狂威
4/8(五)	9	中信兄弟	新莊	泰迪
4/9(六)	11	味全龍	新莊	布里悍
4/10(日)	13	統一7-ELEVEn獅	新莊	古林睿煬
4/14(四)	19	中信兄弟	洲際	泰迪
4/15(五)	21	中信兄弟	洲際	象魔力
4/16(六)	23	統一7-ELEVEn獅	台南	羅昂
4/21(四)	30	中信兄弟	洲際	象魔力
4/22(五)	33	味全龍	新莊	王維中
4/23(六)	35	味全龍	新莊	布里悍
4/24(日)	37	味全龍	新莊	鋼龍
4/26(二)	39	樂天桃猿	樂天桃園	黃子鵬

Data: 2022 roster swaps

3. Get roster swaps for each game of the 2022 season from the CPBL website.

異動日期	球員	球隊	異動原因
2022/05/31	林子昱	味全龍	升一軍
	林孝程	味全龍	升一軍
	林岱安	統一7-ELEVEn獅	升一軍
	高國麟	富邦悍將	升一軍
	林益全	富邦悍將	升一軍
	林哲瑄	富邦悍將	升一軍
	鍾玉成	樂天桃猿	新註冊
2022/05/30	孔念恩	富邦悍將	降二軍
	蔣智賢	富邦悍將	降二軍
	張進德	富邦悍將	降二軍

03

Model Introduction

Model Introduction

Our model will be split into 2 stages:

- The 1st stage is to choose the best player for each position in each game and maximize the sum of all starting players' score in games of a week.
    1. We calculate the sum of defensive score and offensive score of each player on each position according to the opponent's starter pitcher, game field, month of the game and their fielding stats.
    2. Both formulas for the offensive and defensive score are defined by ourselves.

Model Introduction

Our model will be split into 2 stages:

- The 2nd stage is to assign the batter order for the starting lineup from results of 1st stage in each game.
    1. Lastly, we'll assign the batting order for the chosen players from the 1st stage by their overall offensive contributions per plate appearance in each batting order.

With implementation of our model, we'll be able to get the optimal player lineup for each game for the whole week

04

IP Model

Sets

1. Set $G = \{1, \dots, N\}$: all games in a week
2. Set S_1: star players who must be a starting player in every game during the week
3. Set S_2: rookie players who should be a starting player at least once during the week
4. Set S_3: old or injured players who must be bench player at least once during the week
5. Set I: all selectable fielders in a week
6. Set $J = \{2, \dots, 9\}$: numbers corresponding to each position, excluding the pitcher
7. Set $J' = \{1, \dots, 9\}$: numbers corresponding to each position, including the pitcher

Parameters

1. $A_{ij} = 1$ if player $i \in I$ can be in position $j \in J$ or 0 otherwise
2. F_{ij} is the defensive score of player $i \in I$ in position $j \in J$
3. B_i is the offensive score of player $i \in I$
4. w_F is the coefficient of the defensive score
5. w_B is the coefficient of the offensive score
6. V_{ij} is the sum of the offensive and defensive score of player $i \in I$ in position $j \in J$

$$\Longrightarrow V_{ij} = A_{ij}(w_B B_i + w_F F_{ij})$$

Parameters

7. P_{jk} is the PA (Plate Appearance) of the k^{th} hitter ($k \in K$) of position $j \in J$
8. O_{jk} wOBA (Weighted On Base Average) of the k^{th} hitter ($k \in K$) of position $j \in J'$
9. Z_g is the weight of game $g \in G$ depending on the opponent team

Self-defined Formulas

- Offensive score B

$$B = \frac{(OPS{+}_{Season} \times PA_{Season} + OPS{+}_{month} \times PA_{month} + OPS{+}_{field} \times PA_{field} + OPS{+}_{vsp} \times PA_{vsp})}{PA_{season} + PA_{month} + PA_{field} + PA_{vsp}}$$

- Offensive Score F

$$F = \frac{FPCT_{season}}{FPCT_{\mu pos}}$$

IP Model – Decision Variable

1. $x_{ijg} = 1$ if player $i \in I$ is in position $j \in J$ in game $g \in G$ or 0 otherwise
2. $y_{jk} = 1$ if player of position $j \in J'$ is the k^{th} hitter ($k \in K$) in the batting order or 0 otherwise

IP Model – Stage 1

Fielding position assignment for a week

Objective function:

$$\max \sum_{g \in G} \sum_{j \in J} Z_g V_{ij} x_{ijg}$$

IP Model – Stage 1

Fielding position assignment for a week

Constraints:

$$\text{s.t.} \quad \sum_{i \in I} x_{ijg} = 1 \qquad \forall j \in J, g \in G$$

$$\sum_{j \in J} x_{ijg} \leq 1 \qquad \forall i \in I, g \in G$$

$$x_{ijg} \leq A_{ij} \qquad \forall i \in I, j \in J, g \in G$$

$$\sum_{g-n}^{n+k} x_{i2g} \leq 3 \qquad \forall i \in I, n \in G, n + k \leq \mathrm{N}.$$

IP Model – Stage 1

Fielding position assignment for a week

Constraints:

$$\text{s.t.} \quad \sum_{g \in \mathrm{G}} x_{ijg} = N \qquad \forall i \in I, j \in J, \text{ if player } i \in S_1$$

$$\sum_{g \in G} x_{ijg} \geq 1 \qquad \forall i \in I, j \in J, \text{ if player } i \in S_2$$

$$\sum_{g \in G} x_{ijg} < N \qquad \forall i \in I, j \in J, \text{ if player } i \in S_3$$

$$x_{ijg} \in \{0,1\} \qquad \forall i \in I, j \in J, g \in G.$$

IP Model – Stage 2
Batting order assignment of each game

1. After Stage 1, let the player with the highest offensive score B_i among the remaining players be DH (Designated Player)
2. Generate the batting order for each game by conducting the Stage 2 model N times

IP Model – Stage 2

Batting order assignment of each game

Objective Function:

$$\max \sum_{k \in K} \sum_{j \in J'} O_{jk} y_{jk}$$

IP Model – Stage 2

Batting order assignment of each game

Constraints:

$$\text{s.t.} \quad \sum_{j \in J'} y_{jk} = 1 \qquad \forall k \in K$$

$$\sum_{k \in K} y_{jk} = 1 \qquad \forall j \in J'$$

$$y_{jk} \in \{0,1\} \qquad \forall j \in J',\ k \in K$$

05

Performance Analysis

Performance Analysis

- Ex: Comparison between stimulated results and the actual lineup of **富邦悍將**

	5/17		5/18		5/20	
	Real Version	IP Model	Real Version	IP Model	Real Version	IP Model
1	王正棠 2B	葉竹軒 2B	王正棠 2B	申皓瑋 CF	莊韋恩 2B	葉竹軒 2B
2	董子恩 3B	范國宸 1B	董子恩 3B	林琨笙 C	董子恩 3B	蔣智賢 1B
3	申皓瑋 CF	申皓瑋 CF	申皓瑋 CF	蔣智賢 1B	申皓瑋 CF	申皓瑋 CF
4	范國宸 1B	李宗賢 SS	范國宸 1B	葉竹軒 2B	范國宸 1B	李宗賢 SS
5	蔣智賢 DH	陳真 LF	蔣智賢 DH	董子恩 3B	蔣智賢 DH	陳真 LF
6	張進德 C	張進德 C	陳真 LF	張進德 DH	張進德 C	張進德 C
7	陳真 LF	王正棠 RF	張進德 C	李宗賢 SS	陳真 LF	范國宸 RF
8	孔念恩 RF	蔣智賢 DH	孔念恩 RF	范國宸 RF	葉竹軒 RF	張冠廷 DH
9	王勝偉 SS	董子恩 3B	王勝偉 SS	陳真 LF	王勝偉 SS	董子恩 3B
Stage 2 obj	2.455	3.136	2.455	3.045	2.625	3.103

Performance Analysis

- Average of our stimulated results and the actual results of 富邦悍將's 2 months games

	Stage 1	Stage 1 wAVG	Stage 2
Real Version	2241.33	580.90	2.57
IP Model	2611.57	666.45	3.14

06

Conclusion

Conclusion

From the results of our IP model, we can see that our algorithm can determine the lineup more efficiently than lineups that are chosen manually without any consideration of player performance.

Any team with complete data can use our model to generate a starting lineup for better reference.

Thanks