CO 370 Final Project

Presented by Group g 17
Ho Dac Dan (dd2ho@uwaterloo.ca), Hu Jerry (j393hu@uwaterloo.ca), Kim Ted (t88kim@uwaterloo.ca), Kim Justin (j686kim@uwaterloo.ca), Wang Anthony (c668wang@uwaterloo.ca)

Introduction

In light of the NBA's future plans for league expansion, Montreal has emerged as a prime candidate, ranking fourth in the potential cities to host a new team. This presents an exhilarating opportunity to not only introduce a fresh franchise to the basketball world but also to ambitiously position this team as the pinnacle of Canadian basketball, surpassing the Toronto Raptors. The goal is to meticulously construct a team from the ground up, one that is not just competitive but dominant, setting the standard for excellence in Canadian basketball. To achieve this, we will compare significant categories with the current Toronto Raptors roster. This analysis aims to identify key areas where Montreal's team can gain a competitive edge. This venture goes beyond mere talent acquisition; it requires a strategic blend of scounting, player development, and team chemistry, all focused on the singular objective of establishing Montreal's team as the new powerhouse in Canadian basketball. The ambition is clear: to debut not just as another expansion team, but as a formidable force that redefines the basketball landscape in Canada.

I. General Formulation

Let $\mathbf{P}$ be the set of all active NBA players,

$$\mathbf{P} = \{\text{Stephen Curry, Kevin Durant, LeBron James, Nikola Jokic, Joel Embiid, ...}\}$$

Let $\mathbf{F}$ be the set of all the skill parameters (factors) of interest of our model. For all $j \in \mathbf{F}$, j is a factor we want to include in our analysis.

For variables, we introduce x_i for player i. $x_i \in \{0, 1\}$ If $x_i = 1$, player i is selected by our team, if $x_i = 0$, player i is decided not to be included.

• For each factor $j \in \mathbf{F}$ and player $i \in \mathbf{P}$, we have:

$$\sum_{\forall i} a_{ji} \times x_i \geq b_j$$

where a_{ij} is the value of factor j in set of parameters for player i; b_j is the minimum value required of the summed abilities in our team for parameter j.

To generalize, we have

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}$$

where the numerical value of i for x_i represents the i-th element in $\mathbf{P}$ *and the numerical value of j for b_j represents j-th element in* $\mathbf{F}$*. The same applies to a_{ji}.*

• Limitations on players:

According to NBA regulations, number of players in a team cannot exceed 15, we add the constraint

$$\sum_{\forall i} x_i \leq 15$$

In professional basketball, each team is composed of five players on the court, with specific roles designated as C (Centre), PF (Power Forward), SF (Small Forward), PG (Point Guard) and SG (Shooting Guard). It is crucial for teams to maintain at least two and at most three players for each position for several reasons:

Why we need at least two players for each position?

We need Substitute Players for Fatigue and Foul Trouble: Basketball is a physically demanding sport that necessitates periodic breaks for players to manage fatigue and minimize the risk of accruing fouls since the maximum number of fouls allowed in a game is six. Having at least two players per position allows for effective rotation and rest. Injury Backup: Given the high-intensity nature of basketball, injuries are not uncommon. Maintaining a secondary player in each position ensures that a team can seamlessly continue the game in the event of an injury to a primary player, thereby upholding the team's competitiveness and consistency.

Why we can have at most three players for each position?

Roster Size limit: All professional basketball leagues such as the NBA, impose a roster size limit, typically 15 players. Adhering to a maximum of three players per position ensures a balanced distribution of players across all positions, allowing for strategic depty and versatility. Financial Constraint: Professional basketball teams operate under a salary cap, which limits the total amount that can be spent on player salaries. Allocating more than three players per position could impose financial constraints, limiting a team's ability to invest in other critical areas such as a player development, coaching, and training facilities.

Thus, based on the discussions, number of players of each position should be larger than 2 and smaller than 3:

$$2 \leq \sum_{\forall i,j} pos_{ij} \times x_i^j \leq 3$$

$$\sum_{\forall j} x_i^j = x_i$$

where x_i^j determines the position player plays in the season. One player can only play one position in the entire season. $pos_{ij} = 1$ if player i plays position $j \in \mathbf{Pos}$, else $pos_{ij} = 0$, and

$$\mathbf{Pos} = \{\text{PG, PF, C, SG, SF}\}$$

• To minimize sum of salaries:

For s_i for player i we obtained in Players' Salaries, we formulate the objective as:

$$\text{minimize} \quad \sum_{\forall i} s_i \times x_i$$

• Generalize:

$$\begin{aligned}
\text{minimize} \quad & \sum_{\forall i} s_i \times x_i \\
\text{subject to} \quad & \sum_{\forall i} a_{ji} \times x_i \geq b_j && \forall j \in \mathbf{F}, \\
& \sum_{\forall i} x_i \leq 15, \\
& 2 \leq \sum_{\forall i} pos_{ij} \times x_i^j \leq 3 && \forall i \in \mathbf{P}, j \in \mathbf{Pos}, \\
& \sum_{\forall j} x_i^j = x_i && \forall j \in \mathbf{P}, \\
& 0 \leq x_i \leq 1 && \forall i \in \mathbf{P}, \\
& x_i \in \mathbb{Z} && \forall i \in \mathbf{P}
\end{aligned}$$

II. Data collection and Data processing

We collect salaries of all active NBA players in 23-24 season on `https://www.espn.com/nba/salaries` as "`Salary.csv`" and all the parameters of players on `https://www.basketball-reference.com/leagues/NBA_2024_totals.html` as "`Players_with_Salary.csv`". For each player, we sum up his salaries paid by all the team he played with in the season. We filter out the players without available salary data.

We remove diacritics of players' names to avoid any confusion and mismatch following the dictionary `Names`. We also accord teams' acronyms with their "real" name defined by dictionary `Teams`.

```
Names = {"Nikola Jokić":"Nikola Jokic", "Luka Dončić":"Luka Doncic",
"Kristaps Porziņģis":"Kristaps Porzingis", "Bogdan Bogdanović":"Bogdan
Bogdanovic", "Bojan Bogdanović":"Bojan Bogdanovic", "Nikola Vučević":"Nikola
Vucevic", "Marcus Morris":"Marcus Morris", "Dāvis Bertāns":"Davis Bertans",
"Jusuf Nurkić":"Jusuf Nurkic", "Jonas Valančiūnas":"Jonas Valanciunas",
"Dennis Schröder":"Dennis Schroder", "Robert Williams":"Robert Williams
III", "Reggie Bullock":"Reggie Bullock Jr.", "Vasilije Micić":"Vasilije
Micic", "Alperen Şengün":"Alperen Sengun", "Nikola Jokić":"Nikola Jovic",
"Vlatko Čančar":"Vlatko Cancar", "Jeff Dowtin":"Jeff Dowtin Jr.", "Luka
Šamanić":"Luka Samanic", "Harry Giles":"Harry Giles III", "Boban Marjanović":"Boban
Marjanovic", "Dario Šarić":"Dario Saric", "Xavier Tillman Sr.":"Xavier
```

```
Tillman", "A.J. Green":"AJ Green", "Jeenathan Williams":"Nate Williams",
"Filip Petrušev":"Filip Petrusev"}

Teams = {"ATL":"Atlanta Hawks", "BOS":"Boston Celtics", "BRK":"Brooklyn
Nets", "CHO":"Charlotte Hornets", "CHI":"Chicago Bulls", "CLE":"Cleveland
Cavaliers", "DAL":"Dallas Mavericks", "DEN":"Denver Nuggets", "DET":"Detroit
Pistons", "GSW":"Golden State Warriors", "HOU":"Houston Rockets", "IND":"Indiana
Pacers", "LAC":"Los Angeles Clippers", "LAL":"Los Angeles Lakers",
"MEM":"Memphis Grizzlies", "MIA":"Miami Heat", "MIL":"Milwaukee Bucks",
"MIN":"Minnesota Timberwolves", "NOP":"New Orleans Pelicans", "NYK":"New
York Knicks", "OKC":"Oklahoma City Thunder", "ORL":"Orlando Magic",
"PHI":"Philadelphia 76ers", "PHO":"Phoenix Suns", "POR":"Portland Trail
Blazers", "SAC":"Sacramento Kings", "SAS":"San Antonio Spurs", "TOR":"Toronto
Raptors", "UTA":"Utah Jazz", "WAS":"Washington Wizards", "TOT":"Total"}
```

In file "`Players_with_Salary.csv`", team "TOT" is shown, it doesn't refer to a team name. Instead, it shows a player's overall performance when they've played for different teams from 2023 to 2024. There are six entries with TOT, which simply means the total of all the player's skills and abilities in one row. So, it's like a summary that captures how well the player did across various teams during that time.

III. Implementation

We write a python program called "`final_project_coding.py`" to formulate our problem. To perform optimization, we need to define a dictionary `Constraints`.

Example:

```
Constraints = {"PTS":[1, 200], "3PA":[1, 30], "ORB":[1,600]}
```

Each key represents a factor $j \in \mathbf{F}$, the list corresponds to the key is `[d, b]`. The first element of the list `d` is the indicator of the direction of comparison of the constraint, the second element `b` is the right hand side of $a^T x(\leq, \geq)b$. If `d` $= 1$, $a^T x \geq$ `b` is used, if `d` $= -1$, $a^T x \leq$ `b` is implement for constraint on factor j. In our study, `b` is the lowest amount of skill of $j \in \mathbf{F}$ required in our team.

Our goal is to beat Toronto Raptors. We set our constraints as:

```
Constraints = {"FG":[1, 980], "3P":[1, 260], "2P":[1,730], "FT":[1,
380], "ORB":[1, 300], "DRB":[1, 760], "TRB":[1,1160], "AST":[1, 660],
"STL":[1, 190], "BLK":[1, 110], "TOV":[1,325], "PF":[1, 420], "PTS":[1,
2580]}
```

Or in terms of vector b,

$$b = \begin{pmatrix} 980 \\ 260 \\ 730 \\ 380 \\ 300 \\ 760 \\ 1160 \\ 660 \\ 190 \\ 110 \\ 325 \\ 420 \\ 2580 \end{pmatrix} \begin{matrix} \text{FG} \\ \text{3P} \\ \text{2P} \\ \text{FT} \\ \text{ORB} \\ \text{DRB} \\ \text{TRB} \\ \text{AST} \\ \text{STL} \\ \text{BLK} \\ \text{TOV} \\ \text{PF} \\ \text{PTS} \end{matrix}$$

Solving this integer program, we get

```
    Variable X
-------------------------
Santi Aldama 1
Desmond Bane 1
Malik Beasley 1
Goga Bitadze 1
Toumani Camara 1
Jalen Johnson 1
Derrick Jones Jr.  1
Tyrese Maxey 1
Andrew Nembhard 1
Dario Saric 1
Alperen Sengun 1
Jabari Walker 1
Russell Westbrook 1
```

Tyrese Maxey (Philadelphia 76ers; \$4.344M), Russell Westbrook (LA Clippers; \$3.836M), Andrew Nembhard (Indiana Pacers; \$2.132M) are chosen for Point Guard; Desmond Bane (Memphis Grizzlies; \$3.845M) and Malik Beasley (Milwaukee Bucks; \$2.710M) are for Shooting Guard; Derrick Jones Jr (Dallas Mavericks; \$2.710M), Jalen Johnson (Atlanta Hawks; \$2.925M), and Jabari Walker (Portland Trail Blazers; \$1.720M) are for Small Forward; Santi Aldama (Memphis Grizzlies; \$2.194M), Toumani Camara (Portland Trail Blazers; \$1.120M) are for Power Forward; Alperen Sengun (Houston Rockets; \$3.536M), Goga Bitadze (Orlando Magic; \$2.067M), Dario Saric (Golden State Warriors; \$2.710M) are for Centre. Visit **Appendix** to examine abilities and salaries of our selected players compared with their counterparts in Toronto Raptors.

• Concerns about Salary cap:

If our summed salary exceeds assigned salary cap, our solution is not feasible.

IV. Salaries Explained

There are significant differences between NBA rookie contracts and veteran contracts that influence team finances and roster construction. Rookie Contracts: Rookies sign four year contracts, which are typically in line with a rookie scale set by the NBA's Collective Bargaining Agreement. The scale changes each year based on the salary cap and the player's draft position. The collective salary of our players remains modest, largely because they are still on their rookie contracts, which are substantially lower than the lucrative deals typically signed post-rookie period. Veteran Contracts: Minimum Salary Scale: Veteran salaries vary based on their years of experience in the NBA. For instance, a player with 10+ years of experience signing a new minimum contract would be around $3 million. Players such as Russell Westbrrok and Dario Saric from our team belong here. In contrast, the salary total for the Raptors' roster is comparatively higher, a reflection of their players' extensive NBA experience. Specifically, players like Siakam and Barnes, having earned accolades such as All-Star and Rookie Team selections, command higher salaries due to their proven track records and accolades.

V. Future applications

Our model can determine who to hire to gives the most cost-worthy performances and who is over or under paid. Generally, we can conclude based on our optimization that all the players in Toronto Raptors are overpaid.

According to (Sampaio, Jaime, et al.), between-season variations of performances of players can only be affected by their playing time, and the time players spend on-field is determined by their coach. Thus, on-field performances of players cannot always be the same as is shown in the previous seasons. Hence, we need to perform robust optimization to optimize our model with the uncertain data.

VI. Robust optimization

For constraint

$$\sum_{\forall i} a_{ji} \times x_i \geq b$$

rewrite, we get

$$-\sum_{\forall i} a_{ji} \times x_i \leq -b$$

Applying robust optimization on uncertainty set $\mathcal{E} = \{u \in \mathbb{R}^n : ||x|| \leq r\}$ where $||x|| = \sqrt{x_1^2 + x_2^2 + ... + x_m^2}$, we get

$$-\sum_{\forall i} (a_{ji} \times x_i) + r \times ||x|| \leq -b$$

Since x_i are binary variables, $x_i^2 = \{0, 1\}$,

$$||x|| = \sqrt{\sum_{\forall i} x_i^2} = \sqrt{\sum_{\forall i} x_i}$$

We hence introduce a new variable `Norm`, where `Norm` $= ||x||$.

$$\texttt{Norm} \times \texttt{Norm} = \sum_{\forall i} x_i$$

and

$$-\sum_{\forall i}(a_{ji} \times x_i) + r \times \texttt{Norm} \leq -b$$

Note: this is not linear, but is solvable by gurobi under the setting `m.setParam("NonConvex", 2)`*!*

Since there is a lack of research on this, it is difficult to predict a statistically significant uncertainty set. Therefore, we write a code "`robust_optimization.py`", which can be used to perform robust optimization in future analysis.

VII. Why not Sensitivity Analysis?

Sensitivity analysis was not incorporated into the project for several reasons

• Fixed Salaries of NBA Players: The analysis assumes that the salaries of NBA players are stable and not open to negotiation. Since these salaries are considered constant, there's no need for sensitivity analysis, which typically examines how changes in input variables affect the output.

• Static Parameters: In the model, "b" represents a set of numbers specifically chosen to achieve success against Toronto Raptors. These numbers should remain static for the analysis, eliminating the need to analyze the effects of their variation.

• Integer Constraints and Non-Continuous Solution Space: It is noteworthy that Sensitivity Analysis, typically utilized in Linear Programming (LP), facilitates understanding by allowing variables to assume any real value within a specified range, leading to a smooth and continuous solution space. However, Integer Programming involves variables restricted to integer values, resulting in a non-continuous solution space. In this context, the impact of minor changes in coefficients or constraints may not be as straightforward to analyze, and sensitivity analysis methods developed for LP may not be directly applicable.

Conclusion:

In conclusion, the project utilized data optimization strategies and data analysis to design a method for building a basketball team in Montreal that aims to surpass the Toronto Raptors. This method ensures high performance while being financially efficient. Balancing performance and making strategic financial decisions are crucial for the success of a sports team.

General formulation introduced a general mathematical optimization model that focuses on skill parameters (metrics), and constraints; Mainly, the general

formula formulation involves limitations on players and focuses on minimizing total salaries. For data analysis, we collect salaries of all active NBA players in the 23-24 season. The "Data Collection and Processing" section describes how we process salary and performance data for analytical purposes. Afterward, we will implement the optimization model using python, and solving the program generates the optimal results for our objective.

When implementing our optimization model, it is important to understand the importance of robust optimization. It is utilized to deal with data uncertainty, which accounts for fluctuations of market and player performance. Due to the lack of research, predicting a statistically significant uncertainty set can be difficult; therefore, our group used Python code to perform robust optimization.

Reference:

• 2023-24 NBA player stats: Totals. Basketball. (n.d.-a). https://www.basketball-reference.com/leagues/NBA_2024_totals.html

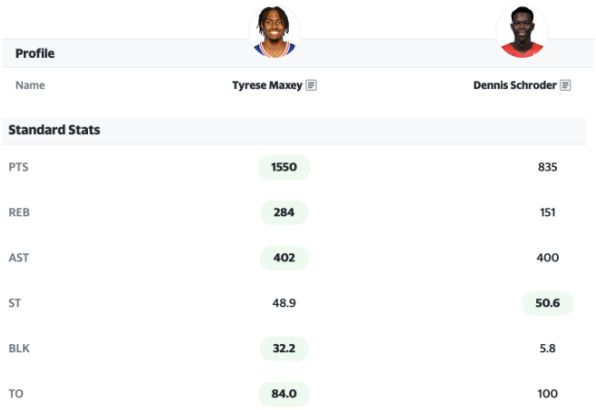
• 2023-24 Toronto Raptors roster and stats. Basketball. (n.d.-b). https://www.basketball-reference.com/teams/TOR/2024.html

• Sampaio, J., Drinkwater, E. J., & Leite, N. M. (2010). Effects of season period, team quality, and playing time on basketball players' game-related statistics. European Journal of Sport Science, 10(2), 141–149. https://doi.org/10.1080/17461390903311935

• Vancouver, Vegas or ... Dakar? Ranking the candidates for NBA expansion cities

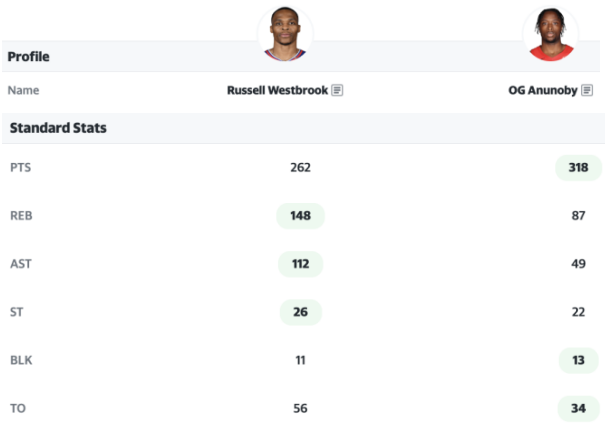
Appendix

- Point Guard
    - Tyrese Maxey (Philadelphia 76ers; $4.344M) vs Dennis Schroeder (Toronto Raptors; $12.405M)



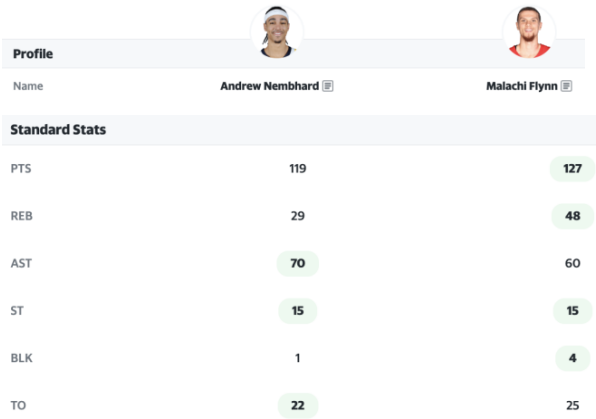
Profile		
Name	Tyrese Maxey	Dennis Schroder
Standard Stats		
PTS	1550	835
REB	284	151
AST	402	400
ST	48.9	50.6
BLK	32.2	5.8
TO	84.0	100

    - Russell Westbrook (LA Clippers; $3.836M) vs OG Anunoby (Toronto Raptors; $18.643M)



Profile		
Name	Russell Westbrook	OG Anunoby
Standard Stats		
PTS	262	318
REB	148	87
AST	112	49
ST	26	22
BLK	11	13
TO	56	34

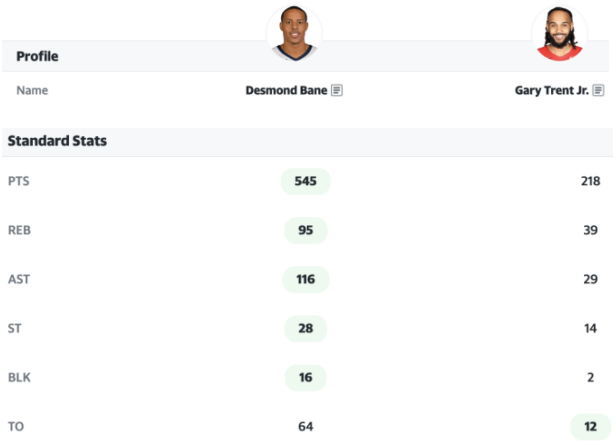
    - Andrew Nembhard (Indiana Pacers; $2.132M) vs Malachi Flynn (Toronto Raptors; $3.873M)



Profile		
Name	Andrew Nembhard	Malachi Flynn
Standard Stats		
PTS	119	127
REB	29	48
AST	70	60
ST	15	15
BLK	1	4
TO	22	25

- Shooting Guard

- Desmond Bane (Memphis Grizzlies; $3.845M) vs Gary Trent Jr (Toronto Raptors; $18.560M)



Profile		
Name	Desmond Bane	Gary Trent Jr.
Standard Stats		
PTS	545	218
REB	95	39
AST	116	29
ST	28	14
BLK	16	2
TO	64	12

- Malik Beasley (Milwaukee Bucks; $2.710M) vs Gradey Dick (Toronto Raptors; $4.536M)

Profile		
Name	Malik Beasley	Gradey Dick
Standard Stats		
PTS	287	60
REB	106	22
AST	35	15
ST	23	4
BLK	4	0
TO	16	12

- Small Forward
  - Derrick Jones Jr (Dallas Mavericks; $2.710M) vs Scottie Barnes (Toronto Raptors; $8.008M)

Profile		
Name	Derrick Jones Jr.	Scottie Barnes
Standard Stats		
PTS	218	482
REB	83	218
AST	21	135
ST	14	40
BLK	20	29
TO	20	58

- Jalen Johnson (Atlanta Hawks; $2.925M) vs Precious Achiuwa (Toronto Raptors; $4.379M)


Profile		
Name	Jalen Johnson	Precious Achiuwa
Standard Stats		
PTS	**212**	144
REB	109	**111**
AST	**36**	35
ST	**17**	11
BLK	**15**	8
TO	30	**26**

- Jabari Walker (Portland Trail Blazers; $1.720M) vs Jalen McDaniels (Toronto Raptors; $4.516M)


Profile		
Name	Jabari Walker	Jalen McDaniels
Standard Stats		
PTS	**164**	32
REB	**108**	18
AST	**16**	9
ST	**17**	5
BLK	**9**	3
TO	21	**6**

- Power Forward
  - Santi Aldama (Memphis Grizzlies; $2.194M) vs Pascal Siakam (Toronto Raptors; $37.893M)


Profile		
Name	Santi Aldama	Pascal Siakam
Standard Stats		
PTS	210	**508**
REB	102	**163**
AST	38	**120**
ST	13	**20**
BLK	**14**	5
TO	**19**	59

- Toumani Camara (Portland Trail Blazers; $1.120M) vs Otto Porter Jr (Toronto Raptors; $6.300M)


Profile		
Name	Toumani Camara	Otto Porter Jr.
Standard Stats		
PTS	147	28
REB	103	24
AST	26	3
ST	20	4
BLK	13	4
TO	24	2

- Centre
  - Alperen Sengun (Houston Rockets; $3.536M) vs Jakob Poeltl (Toronto Raptors; $19.500M)


Profile		
Name	Alperen Sengun	Jakob Poeltl
Standard Stats		
PTS	419	263
REB	191	205
AST	114	56
ST	19	17
BLK	17	28
TO	56	36

  - Goga Bitadze (Orlando Magic; $2.067M) vs Chris Boucher (Toronto Raptors; $11.750M)


Profile		
Name	Goga Bitadze	Chris Boucher
Standard Stats		
PTS	154	141
REB	126	81
AST	39	8
ST	14	7
BLK	35	12
TO	14	10

- Dario Saric (Golden State Warriors; $2.710M) vs Thaddeus Young (Toronto Raptors; $8.000M)


Profile		
Name	Dario Saric	Thaddeus Young
Standard Stats		
PTS	251	4
REB	139	3
AST	53	3
ST	15	0
BLK	6	0
TO	28	1