Predicting Set Wins

Open League Darts

Will Moore

Goals:

Communicate the problem that I am attempting to solve.

Communicate my progress on addressing that problem.

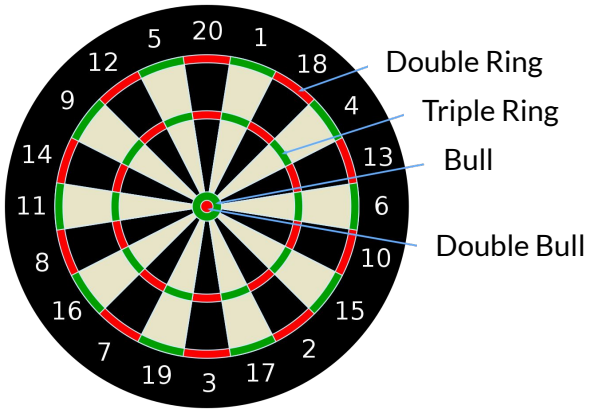
Explanation: Game

There are 3 Game variations played in an Open League darts Match:

1. 501 Single In Double Out

501 Single In Double Out

Each team counts down from 501 by scoring points on the board and wins by bringing the score to exactly zero with a double.



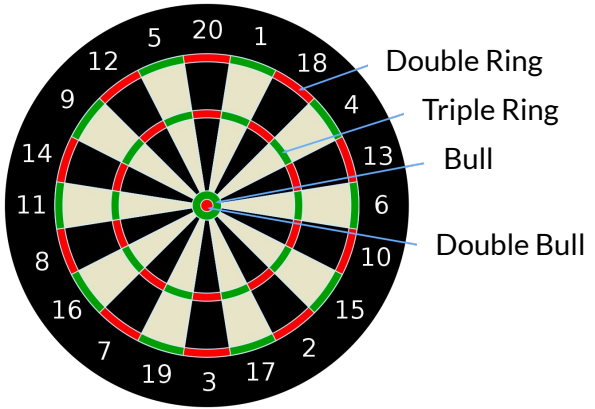
Explanation: Game

There are 3 Game variations played in an Open League darts Match:

1. 501 Single In Double Out
2. 301 Double In Double Out

301 Double In Double Out

Each team begins scoring with a double, then counts down from 301 by scoring points on the board and wins by bringing the score to exactly zero with a double.



Explanation: Game

There are 3 Game variations played in an Open League darts Match:

1. 501 Single In Double Out
2. 301 Double In Double Out
3. Cricket (Dirty)

Cricket (Dirty)

- Targets: 15-20 and bullseye
- Each team closes their targets by hitting each of those targets 3 times.
- Each team has the ability to score points on a target if they have closed a target and their opponents have not yet closed that target.
- Points scored are equal to the face value of the target and bullseye counts as 25 points.
- A team wins when they have all targets closed *and* the highest points.
- Dirty Cricket means that the order of targets does not matter.

Game Takeaway:

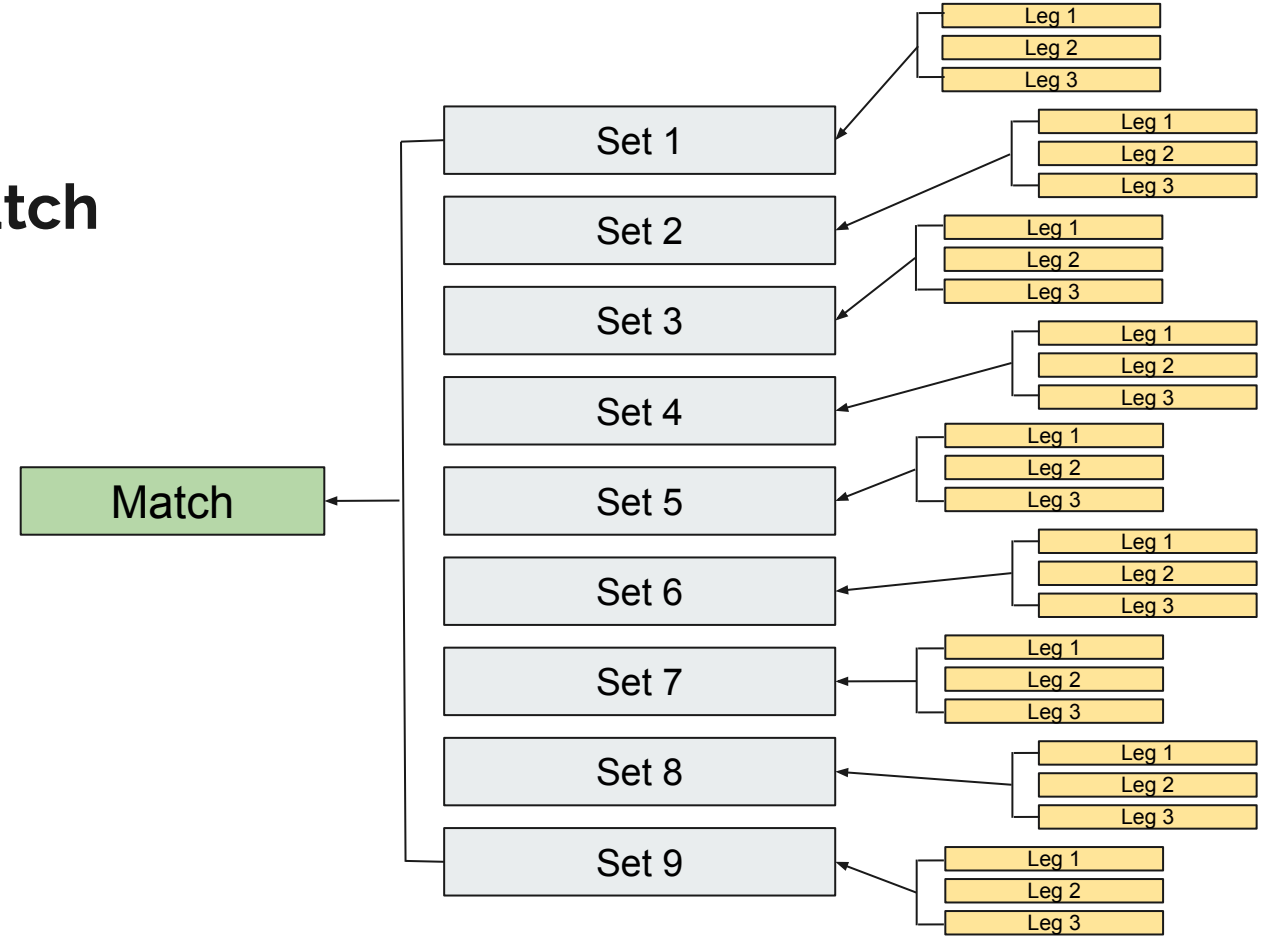
Average scores look very different between games.

Explanation: Match

A Match consists of 9 Sets.

A Set is best of 3 Legs.

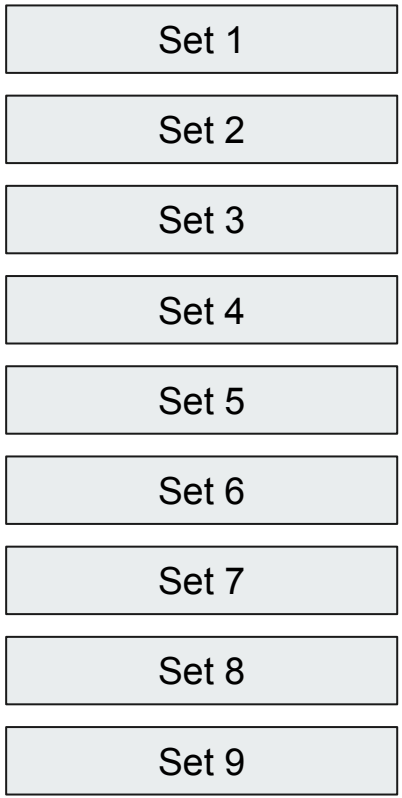
Each leg is a Game of darts.



Explanation: Set

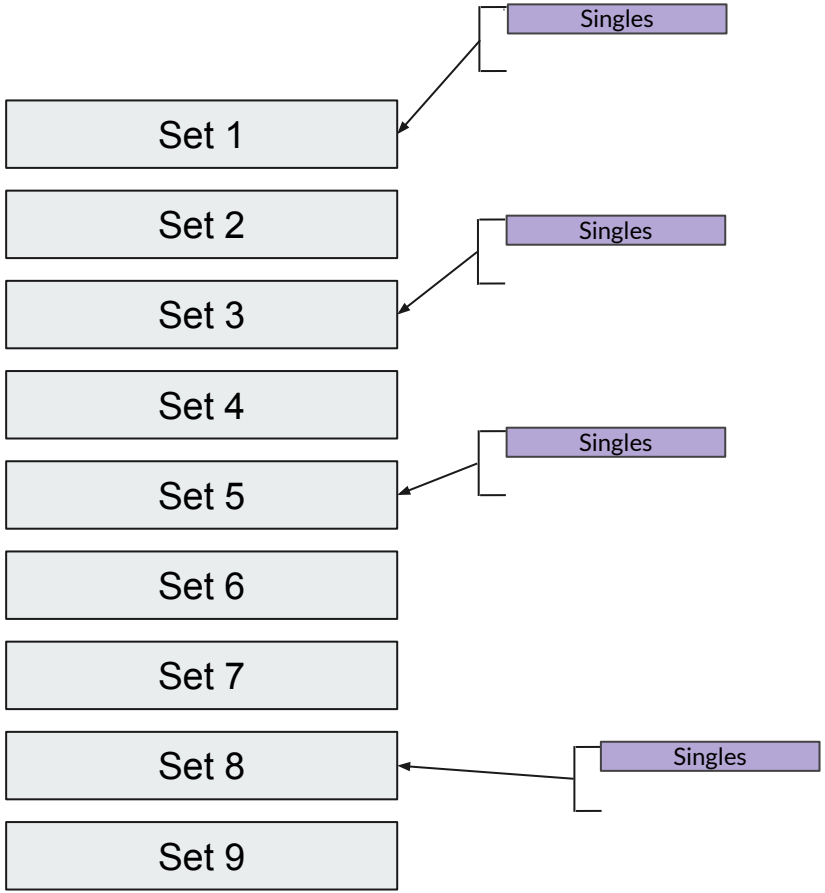
Each Set has rules that determine:

- How many players will play in the Set
- Which Games will be played in the Set



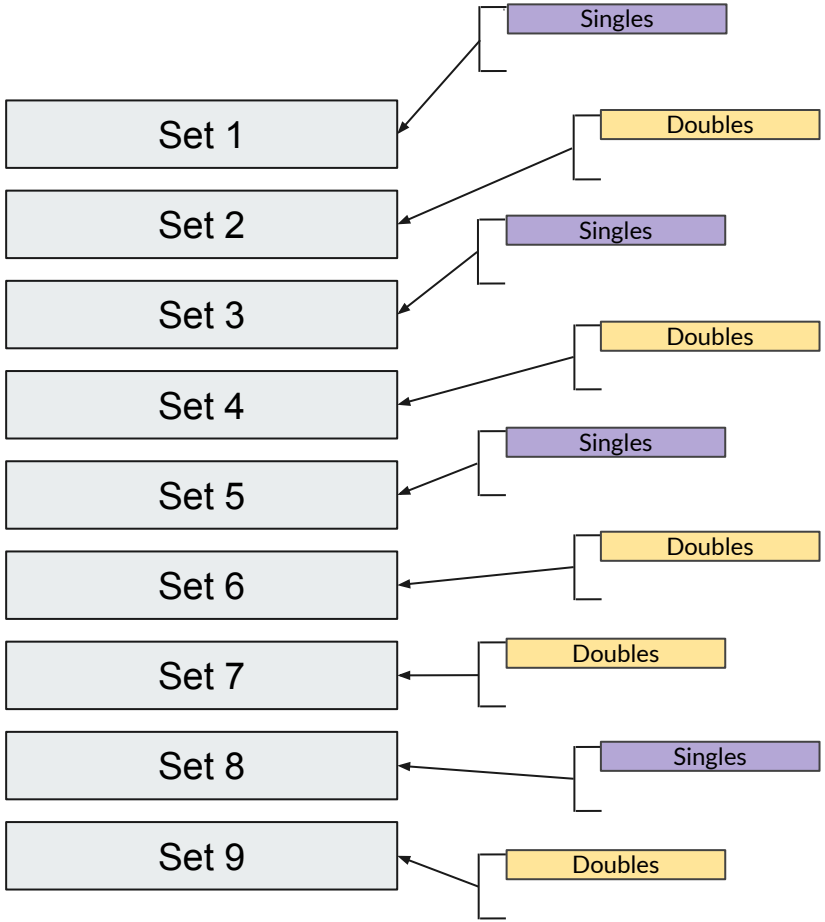
Explanation: Singles

- Singles sets have 1 player from each team.
- No player may play more than 1 singles set in a Match.



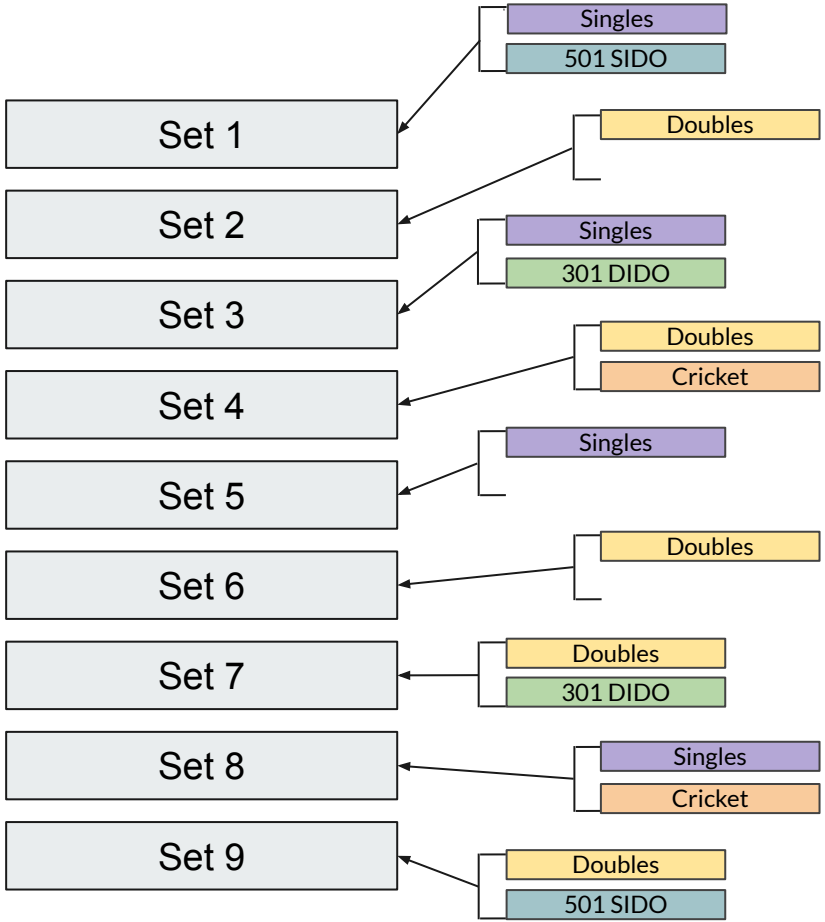
Explanation: Doubles

- Doubles sets have 2 players from each team.
- No 2 players may play more than 1 set together in a Match.



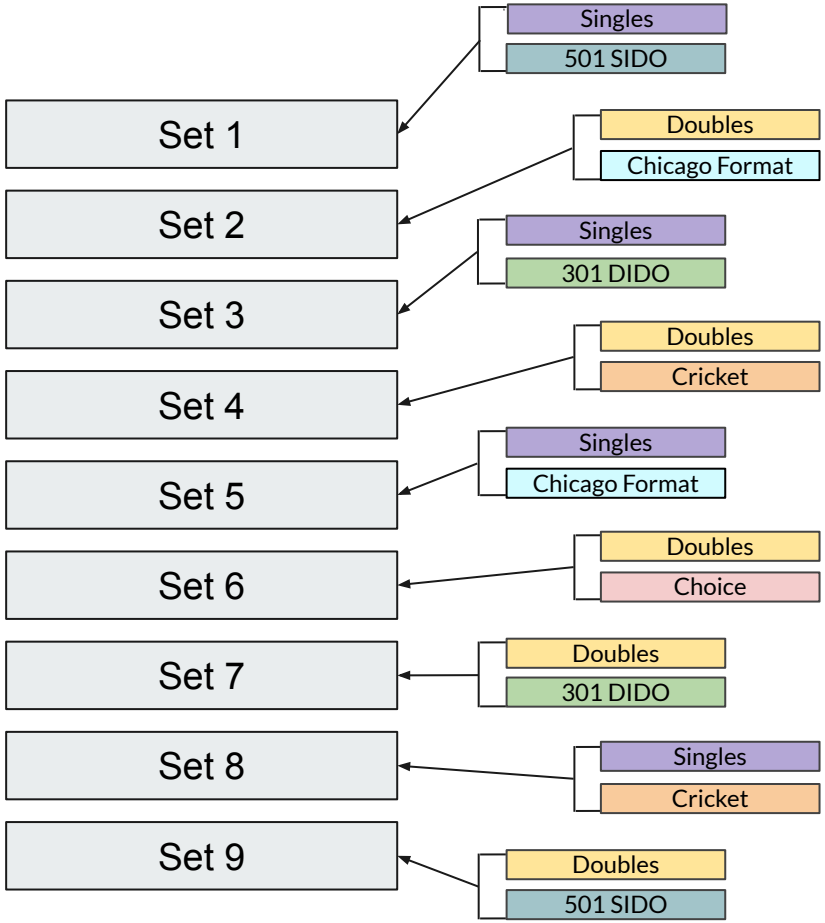
Explanation: Set Limits

- No Player may appear in more than 4 Sets
- No Player may appear in Sets 1-3 more than once.



Explanation: Chicago Format/Choice

- Chicago Format has 1 leg of each Game: 501, 301, Cricket
- Set 6 is a Doubles Set, and the team losing after Set 5 gets to choose: 501, 301, Cricket, Chicago Format



Explanation: Team

- Each team has 4 to 8 players.

Team	Home Team	Away Team
Player 1	Andrew	Alice
Player 2	Bernard	Beatrice
Player 3	Carl	Colleen
Player 4	David	Dolly
Player 5	Eric	Erin

Explanation: Roster

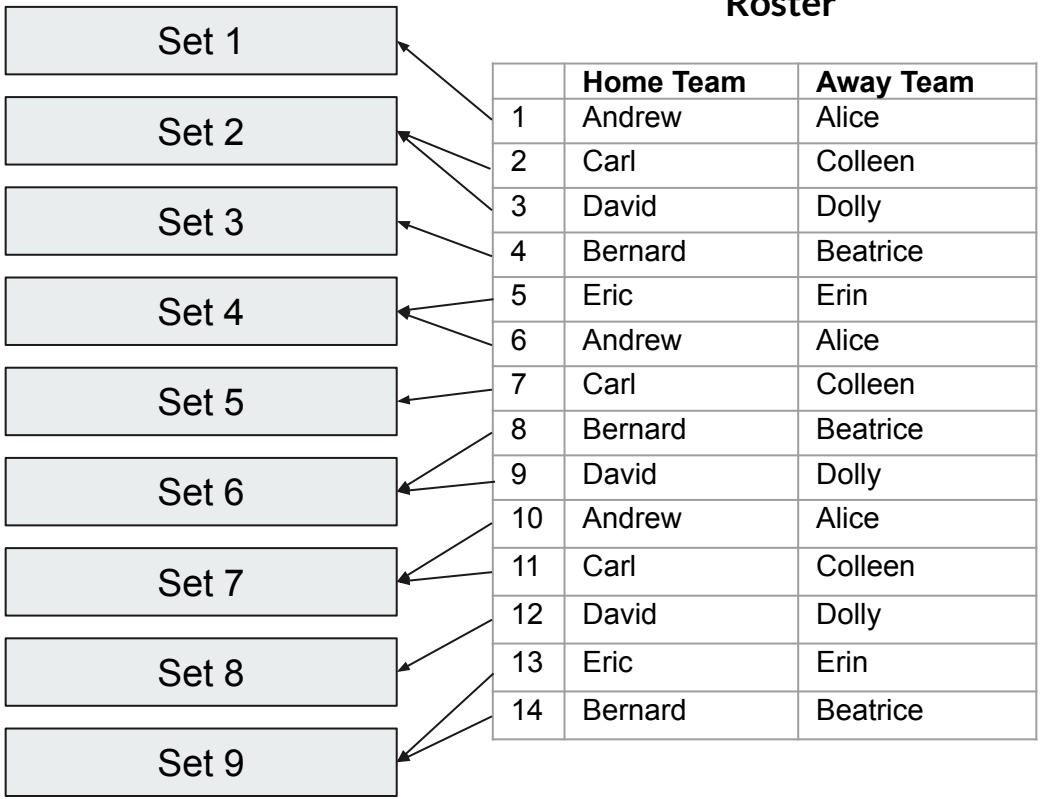
For the match, each team will fill 14 slots with players from their team.

Roster

	Home Team	Away Team
1	Andrew	Alice
2	Carl	Colleen
3	David	Dolly
4	Bernard	Beatrice
5	Eric	Erin
6	Andrew	Alice
7	Carl	Colleen
8	Bernard	Beatrice
9	David	Dolly
10	Andrew	Alice
11	Carl	Colleen
12	David	Dolly
13	Eric	Erin
14	Bernard	Beatrice

Explanation: Roster

The roster determines who will play in each set.

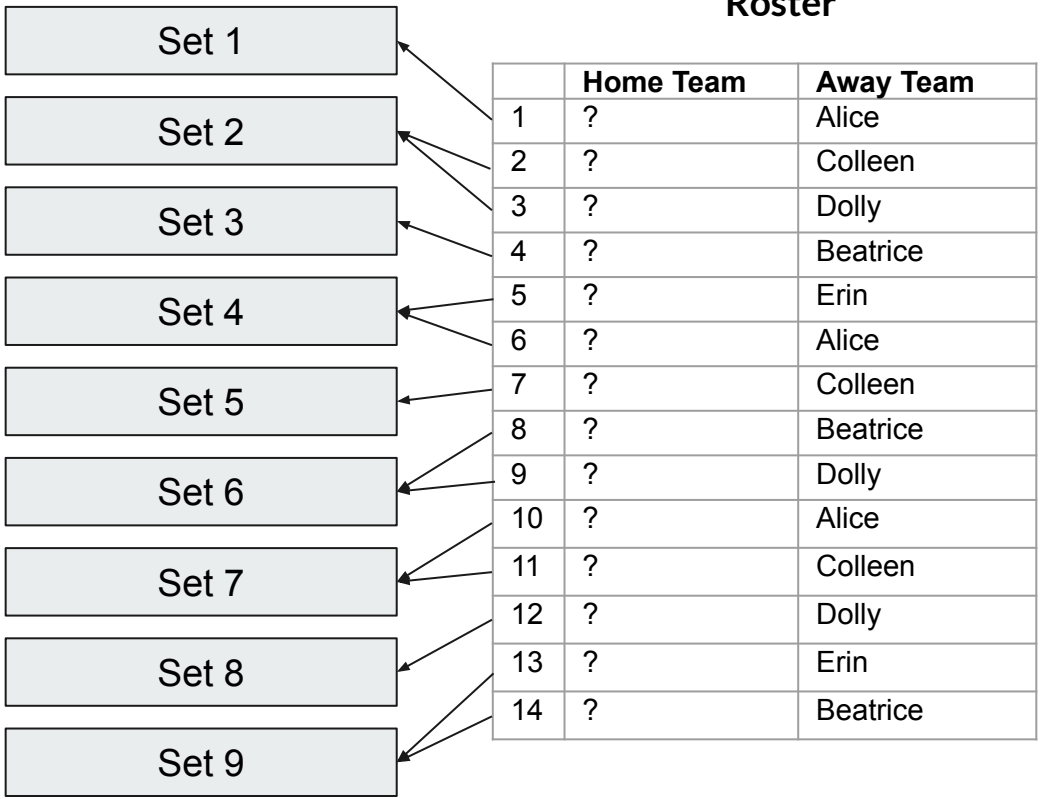


Roster

	Home Team	Away Team
1	Andrew	Alice
2	Carl	Colleen
3	David	Dolly
4	Bernard	Beatrice
5	Eric	Erin
6	Andrew	Alice
7	Carl	Colleen
8	Bernard	Beatrice
9	David	Dolly
10	Andrew	Alice
11	Carl	Colleen
12	David	Dolly
13	Eric	Erin
14	Bernard	Beatrice

The Problem

How do we determine the Roster configuration that will result in the largest number of Set wins for the Home Team?

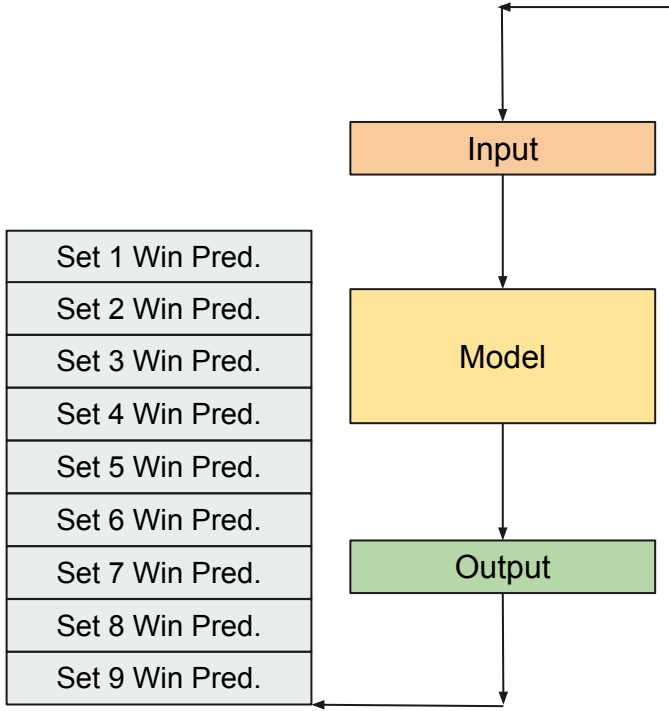


Roster

	Home Team	Away Team
1	?	Alice
2	?	Colleen
3	?	Dolly
4	?	Beatrice
5	?	Erin
6	?	Alice
7	?	Colleen
8	?	Beatrice
9	?	Dolly
10	?	Alice
11	?	Colleen
12	?	Dolly
13	?	Erin
14	?	Beatrice

The Model

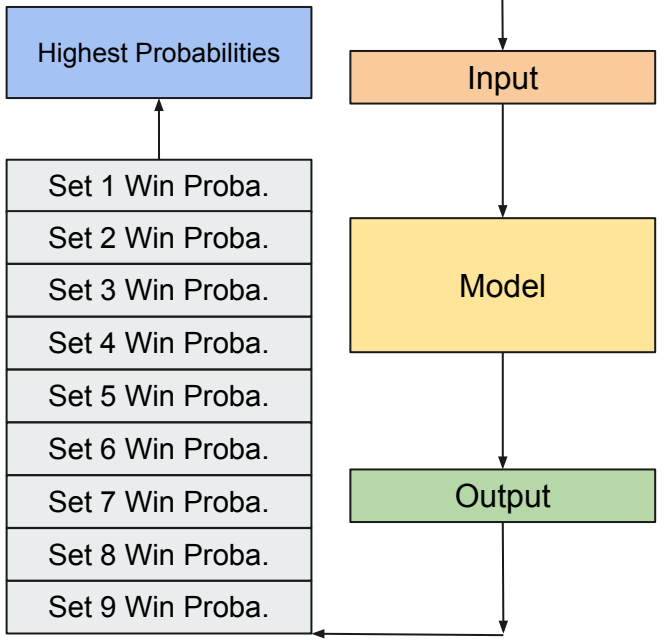
We'll need a model that will predict whether the Home Team will win a given Set when we input generalized historical performance for each team.



Home Team	Away Team
Average Set 1 Player	Average Set 1 Player
Average Set 2 Player	Average Set 2 Player
Average Set 3 Player	Average Set 3 Player
Average Set 4 Player	Average Set 4 Player
Average Set 5 Player	Average Set 5 Player
Average Set 6 Player	Average Set 6 Player
Average Set 7 Player	Average Set 7 Player
Average Set 8 Player	Average Set 8 Player
Average Set 9 Player	Average Set 9 Player

The Model

Then we can input various rosters for the Home Team and we can choose the Roster with the highest Set Win probabilities.



Home Team	Away Team
Ideal Set 1 Player	Average Set 1 Player
Ideal Set 2 Players	Average Set 2 Player
Ideal Set 3 Player	Average Set 3 Player
Ideal Set 4 Players	Average Set 4 Player
Ideal Set 5 Player	Average Set 5 Player
Ideal Set 6 Players	Average Set 6 Player
Ideal Set 7 Players	Average Set 7 Player
Ideal Set 8 Player	Average Set 8 Player
Ideal Set 9 Players	Average Set 9 Player

Model Design - Features

Recall that Chicago Format and Choice (Set 6) may involve 501, 301, and Cricket.

Features:

Five Match Moving Averages for each Set, for each Game of the following:

- Three Dart Average = (Darts Thrown/Points or Marks Scored) * 3
- Leg Win %
- Assist % (Partner of winner in doubles)
- Start/Cork % (Going first is a huge advantage in all games)

Additionally, we will one-hot encode the Set number as a feature.

Data

150K Open League Leg performances scraped from DartConnect

	Game	Player	C	P-3DA	3DA	PPD	T	WA	RD	CO	SS	PR	DT	PTS	MKS	180	HT	9M
1.1	501 SIDO	Chris Liszkowski			44.2	14.7	2	W	12	6	501		34	501				
		Dustin Scholtes	C		41.1	13.7	1				501	8	36	493				
1.2	501 SIDO	Chris Liszkowski			33.0	11.0	2				501	303	18	198				
		Dustin Scholtes	S		75.2	25.1	1	W	7	42	501		20	501				
1.3	501 SIDO	Chris Liszkowski	C		48.3	16.1	1				501	66	27	435				
		Dustin Scholtes			55.7	18.6	2	W	9	98	501		27	501				
2.1	301 DIDO	Daniel K Siemaszko		37.7	34.0	11.3	2				301	75	9	102				
		Samuel Casiano			41.3	13.8	4						9	124				
		Chuck Haisler	C	47.5	55.8	18.6	1	W	4	4	301		10	186				
		Kevin Lally			38.3	12.8	3	A					9	115				
2.2	501 SIDO	Daniel K Siemaszko	S	47.0	59.3	19.8	1	W	6	10	501		17	336				
		Samuel Casiano			33.0	11.0	3	A					15	165				
		Chuck Haisler		46.1	46.4	15.5	2				501	40	15	232				
		Kevin Lally			45.8	15.3	4						15	229				
2.3	Cricket	Daniel K Siemaszko		1.7	1.7	0.6	2				0M		21		12M			
		Samuel Casiano			1.7	0.6	4						18		10M			
		Kevin Lally	C	2.2	1.6	0.5	1	A			0M		21		11M			
		Chuck Haisler			2.8	0.9	3	W	7				19		18M			
3.1	301 DIDO	Joe Pena			26.8	8.9	2				301	6	33	295				
		Sean Ryan	C		25.1	8.4	1	W	12	12	301		36	301				
3.2	301 DIDO	Joe Pena	S		50.2	16.7	1	W	6	35	301		18	301				
		Sean Ryan			27.8	9.3	2				301	162	15	139				
3.3	301 DIDO	Joe Pena			32.1	10.7	2				301	12	27	289				
		Sean Ryan	C		32.3	10.8	1	W	10	6	301		28	301				
4.1	Cricket	Samuel Casiano		1.3	1.4	0.5	2				0M		27		13M			
		Joe Pena			1.1	0.4	4						27		10M			
		Chuck Haisler	C	1.7	1.7	0.6	1	W	10		0M		30		17M			
		Patrick O'Krongley Jr			1.8	0.6	3	A					27		16M			
4.2	Cricket	Samuel Casiano	S	1.6	1.6	0.5	1				0M		33		18M			
		Joe Pena			1.6	0.5	3						33		17M			
		Patrick O'Krongley Jr		1.6	1.6	0.5	2	A			0M		33		17M			
		Chuck Haisler			1.7	0.6	4	W	11				31		18M			
5.1	301 DIDO	Daniel K Siemaszko			43.5	14.5	2				301	40	18	261				
		Patrick O'Krongley Jr	C		47.5	15.8	1	W	7	20	301		19	301				

Data

150K Open League Leg performances scraped from DartConnect

Magic/Suffering

Data

150K Open League Leg performances scraped from DartConnect

Magic/Suffering

18,591 Sets ready for the model

```
Unnamed: 0                                3060
Match                595304336438e23c5dc1495c
Team                                   N.W.O.
Set                                         2
FiGMA_TDAC                            2.86165
FiGMA_TDA3                            52.8429
FiGMA_TDA5                             59.445
FiGMA_Win                            0.233333
FiGMA_Assist                         0.233333
FiGMA_Start                          0.366667
Home                                        1
Set_Win                                     1
Opp_Team                           140 Darts
Opp_FiGMA_TDAC                        2.83543
Opp_FiGMA_TDA3                        48.0088
Opp_FiGMA_TDA5                        64.7867
Opp_FiGMA_Win                             0.2
Opp_FiGMA_Assist                          0.2
Opp_FiGMA_Start                      0.133333
CumFigTDA                              115.15
OppCumFigTDA                           115.631
CumFigTDADiff                       -0.481378
FigStartDiff                         0.233333
FigWinDiff                           0.0333333
FigAsstDiff                          0.0333333
CumFigTDADiff%                     -0.00418046
Set 1                                   False
Set 2                                    True
Set 3                                   False
Set 4                                   False
Set 5                                   False
Set 6                                   False
Set 7                                   False
Set 8                                   False
LogProba                             0.594238
GNBProba                             0.792879
AVGVoteProba                         0.693558
Name: 3060, dtype: object
```

Model Benchmarks

Precision above .87 - “Safe Win” Classification

If 5 Set wins are predicted, chances are that you will win those 5 Sets.

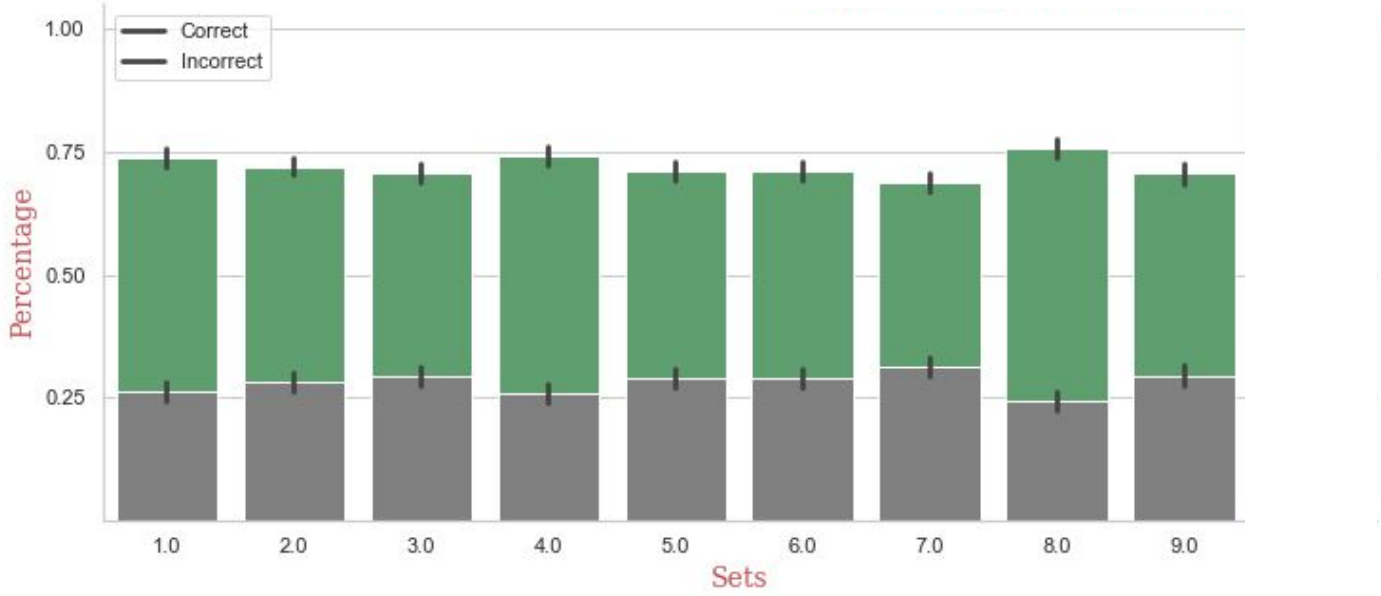
This model is an Average Vote of Logistic Regression and Gaussian Naive Bayes

```
avg_proba = (((clf.predict_proba(set_test[X])[:,1]) + (logit.predict_proba(set_test[X])[:,1])) /2) > .8
```

```
print("Average Vote:")
print("Precision: {:6.4f},   Recall: {:6.4f}".format(precision_score(set_test[y], avg_proba),
                                                    recall_score(set_test[y], avg_proba)))
```

```
Average Vote:
Precision: 0.8775,   Recall: 0.5404
```

Safe Win Chart - Accuracy




Open League Home Set Wins
Win Percentage
0.55
0.50
0.45
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
Set Number

Insights

- Cricket and 501 are easier for this model to predict correctly.
- The Home Team usually wins Cricket and 501 and usually loses Set 6 and 7.

Future Work

- Lay out possible Roster configurations for all team sizes. Each of these Rosters can then be evaluated by the model.
- Create web app.

End

Thank you for your time