

A diagram of a hockey rink with blue lines and red goal creases. In the center is a white rounded rectangle containing text. On the left side of the rink, a black circle (puck) is shown with an arrow pointing towards the top goal. On the right side, a black circle is shown with an arrow pointing towards the bottom goal. There are several 'X' marks on the rink, likely representing players.

MVP - WHO HAS THE MOST HART?

PREDICTING THE HART MEMORIAL TROPHY
RECIPIENT

THE HART MEMORIAL TROPHY



- Presented to the most valuable player (MVP)
- Presented by the NHL in 1960
- Dr. David Hart donated the original trophy in 1923



WHO PICKS THE WINNER ?



- Decided through a poll of the PHWA in all NHL cities
- Poll takes place at the end of the regular season

1st Place	2nd Place	3rd Place	4th Place	5th Place
10 pts	7 pts	5 pts	3 pts	1 pts



Professional
Hockey Writers
Association

OBJECTIVES



- Correctly predict the 2018-19 Hart Memorial Trophy recipient
- Correctly predict the top 10 players voted for the Hart Memorial Trophy



TRANSFORMING THE DATA

- Training data : 2021-22 season stats
- Test data : 2018-19 season stats
- Filtered the data by keeping players

who:

- Have played ≥ 50 games
- $\text{TOI} / \text{GP} \geq 15$

THE VARIABLES

Variable	Definition
Player	The hockey player's name
Age	The player's age
Tm	The team the player is on
Pos	The players position (C, LW, RW, D)
GP	The # of games the player was on the ice
G	The # of goals the player has made
A	The # of goals the player has assisted
PTS	The sum of goals and assists the player has made
+/- (plus/minus)	Rating calculated by: + 1 point for even-strength or short-handed goal the player's team scores while they're on the ice -1 point for even-strength or short-handed goal the opposing team scores while they're on the ice



THE VARIABLES

Variable	Definition
PIM	The total # of penalty minutes the player received
PS	Point share – an estimate of the # of points contributed by a player
TOI	Time on ice (minutes)
OPS	Offensive Point Share – an estimate of the # of points contributed by a player due to their offense
DPS	Defensive Point Share – an estimate of the # of points contributed by a player due to their defense
Vote Percent	The percentage of votes a player has received in favour of winning the Hart Memorial Trophy



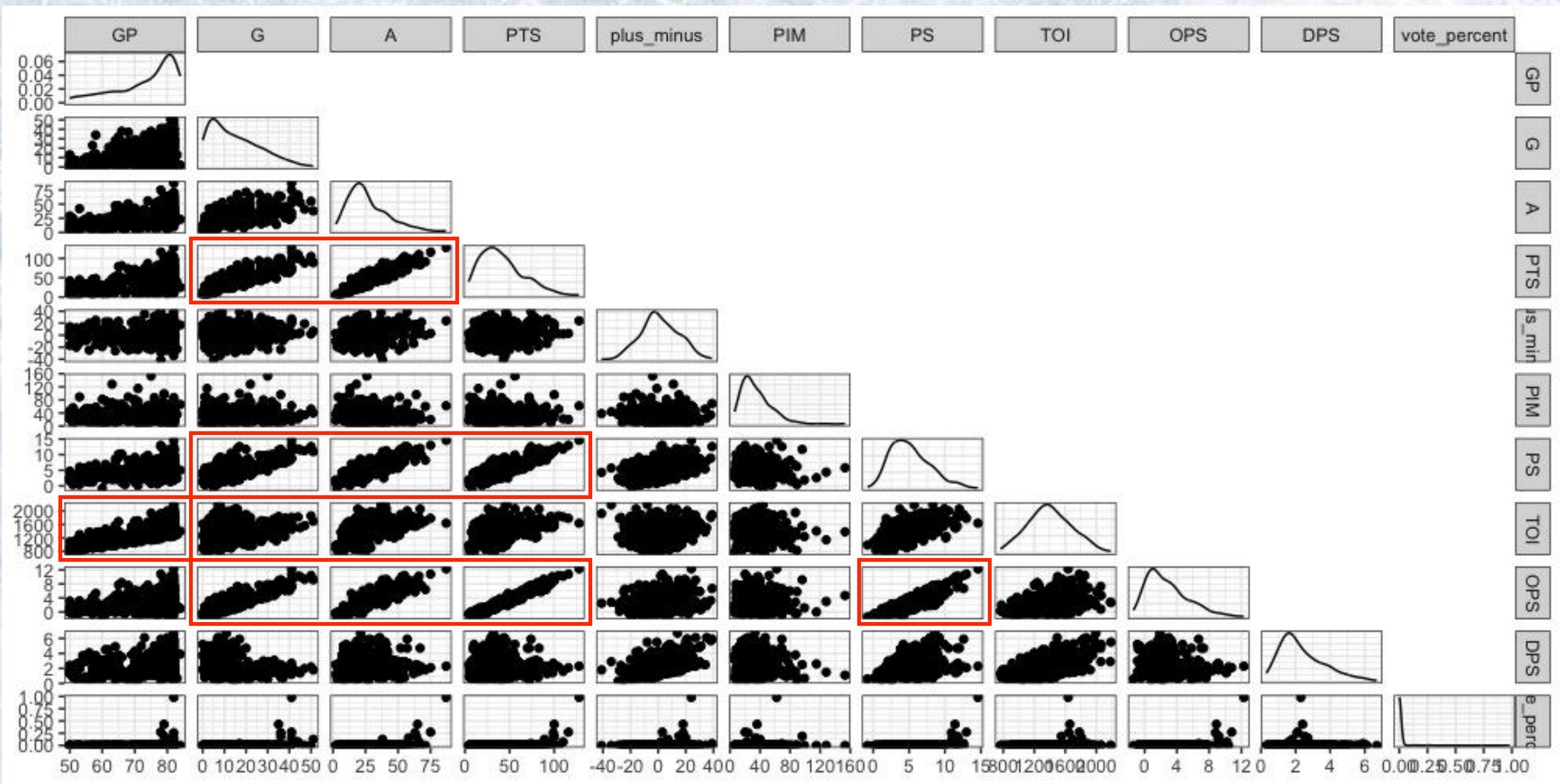


Figure 1. Pairs Plot of the Filtered 2018-19 Season Data

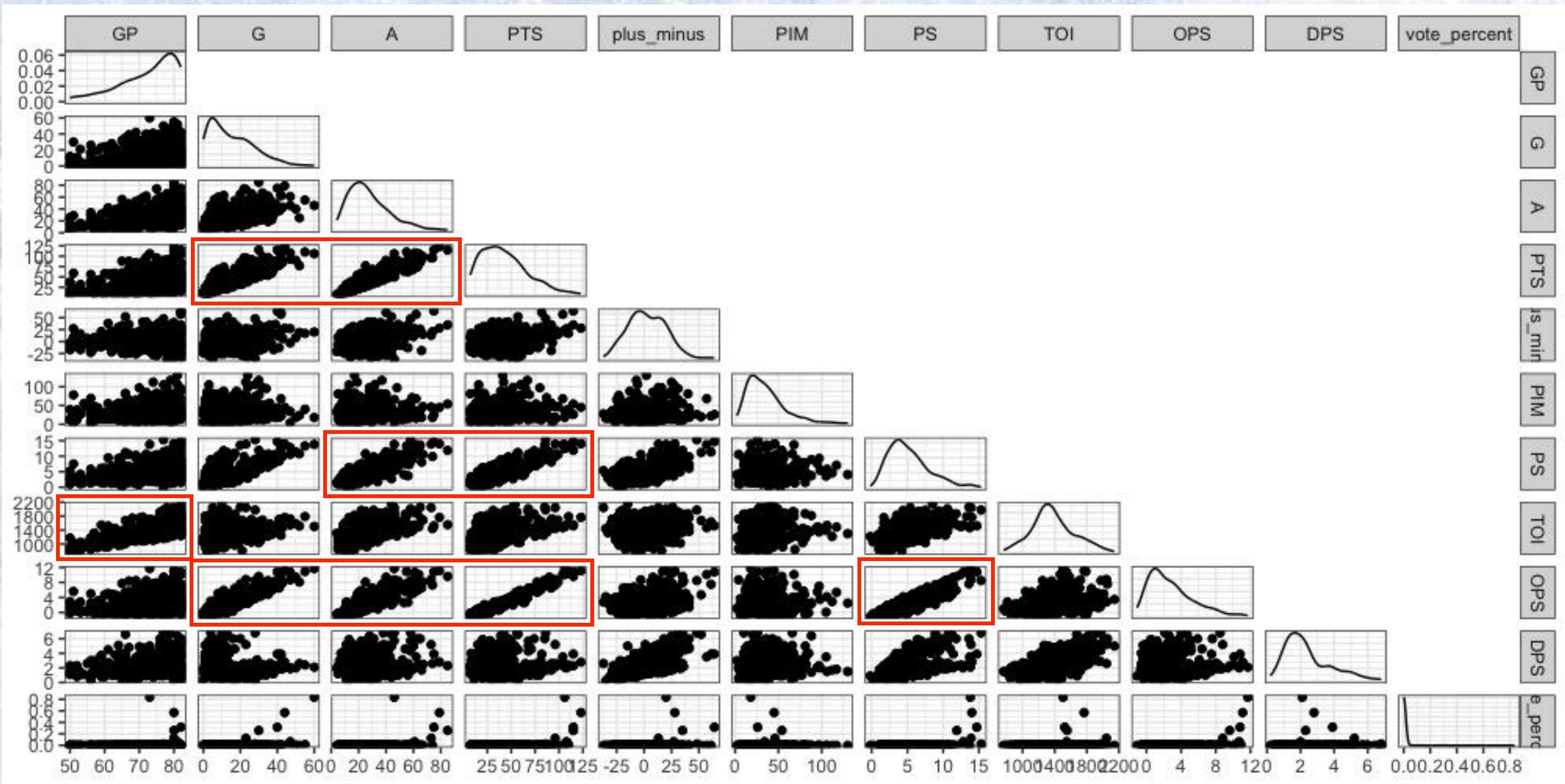


Figure 2. Pairs Plot of the Filtered 2021-22 Season Data

THE MODELS – LINEAR REGRESSION

COEFFICIENTS:

	ESTIMATE
INTERCEPT	-0.348
GP	0.00238
G	-0.179
A	-0.104
PLUS_MINUS	0.0355
PIM	-0.00487
PS	5.39
TOI	0.00195
OPS	-3.56
DPS	-5.97



Test MSE = 33.08424

LINEAR REGRESSION : PREDICTED TOP 10

PREDICTED RANK	NAME	VOTE %	ACTUAL RANK
1	Nikita Kucherov	98.07 %	1
2	Connor McDavid	27.19 %	3
3	Steven Stamkos	0.70 %	14
4	John Tavares	0.41 %	17
5	Patrick Kane	7.31 %	8
6	Leon Draisaitl	0.000 %	N/A
7	Braydon Point	0.18 %	21
8	Sidney Crosby	43.22 %	2
9	David Pastmak	0.0000 %	N/A
10	Alex Ovechkin	12.46 %	7



LINEAR REGRESSION : ACTUAL TOP 10

ACTUAL RANK	NAME	VOTE %	PREDICTED RANK
1	Nikita Kucherov	98.07 %	1
2	Sidney Crosby	43.22 %	8
3	Connor McDavid	27.19 %	2
4	Johnny Gaudreau	25.85 %	13
5	Brad Marchand	18.83 %	15
6	Nathan Mackinnon	15.50 %	11
7	Alex Ovechkin	12.46 %	10
8	Patrick Kane	7.31 %	5
9	Mark Giordano	3.86 %	20
10	Sebastian Aho Jordan Binnington	1.23 %	32 N/A



THE MODELS – RIDGE REGRESSION

COEFFICIENTS:

	ESTIMATE
INTERCEPT	-1.18
GP	-0.00440
G	0.0263
A	0.0174
PLUS_MINUS	0.00970
PIM	-0.00180
PS	0.124
TOI	0.000129
OPS	0.158
DPS	-0.0102



Test MSE = 33.99761

RIDGE REGRESSION : PREDICTED TOP 10

PREDICTED RANK	NAME	VOTE %	ACTUAL RANK
1	Nikita Kucherov	98.07 %	1
2	Connor McDavid	27.19 %	3
3	Leon Draisaitl	0.000 %	N/A
4	Patrick Kane	7.31 %	8
5	Nathan Mackinnon	15.50%	6
6	Steven Stamkos	0.70 %	14
7	Brayden Point	0.18 %	21
8	Sidney Crosby	43.22 %	2
9	Brad Marchand	18.83 %	5
10	John Tavares	0.41 %	17

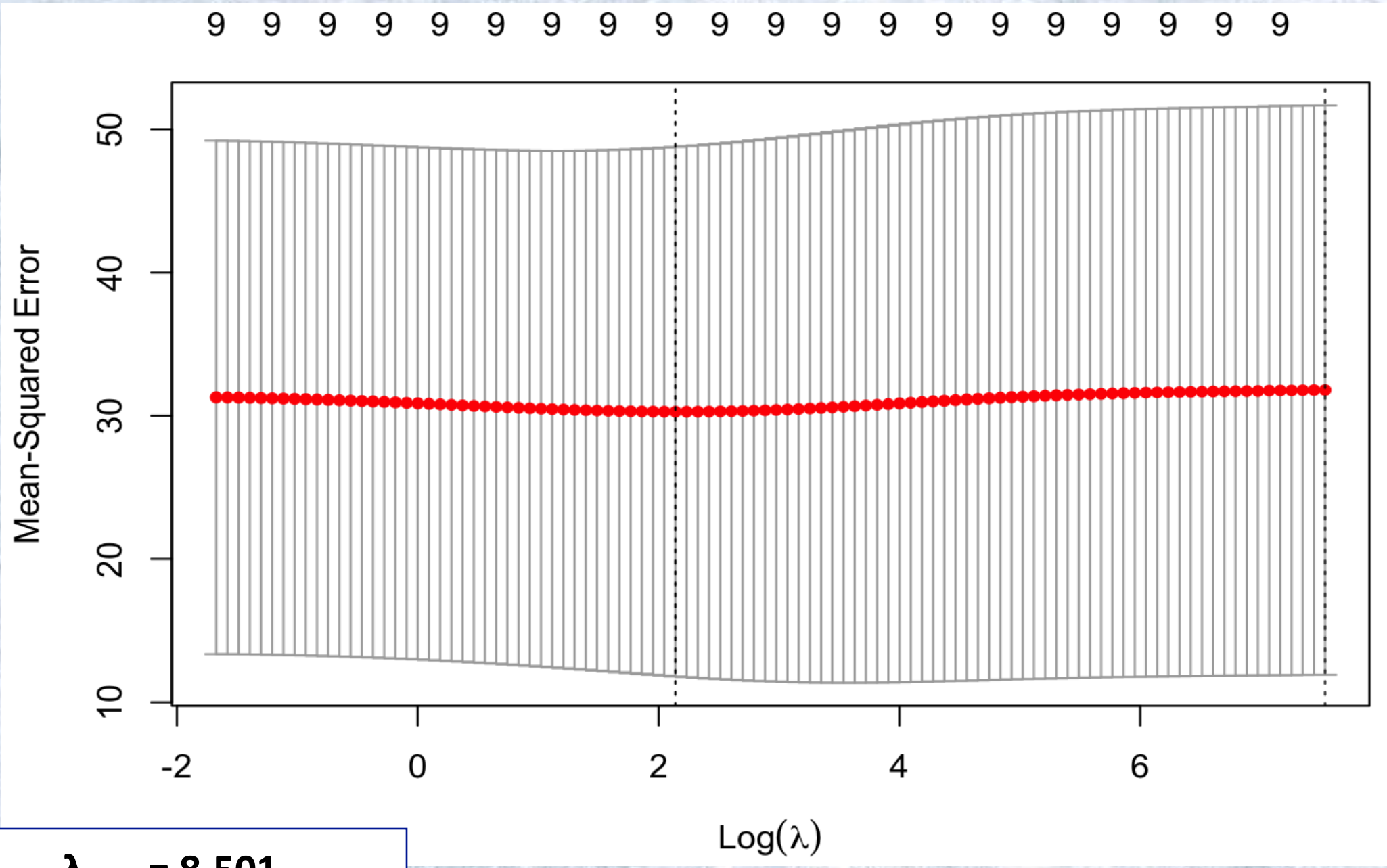


RIDGE REGRESSION : ACTUAL TOP 10

ACTUAL RANK	NAME	VOTE %	PREDICTED RANK
1	Nikita Kucherov	98.07 %	1
2	Sidney Crosby	43.22 %	8
3	Connor McDavid	27.19 %	2
4	Johnny Gaudreau	25.85 %	11
5	Brad Marchand	18.83 %	9
6	Nathan Mackinnon	15.50 %	5
7	Alex Ovechkin	12.46 %	12
8	Patrick Kane	7.31 %	4
9	Mark Giordano	3.86 %	16
10	Sebastian Aho Jordan Binnington	1.23 %	21 N/A



RIDGE REGRESSION PLOT



$\lambda_{\text{BEST}} = 8.501$



THE MODELS – LASSO

COEFFICIENTS:

	ESTIMATE
INTERCEPT	-1.38
GP	
G	
A	
PLUS_MINUS	
PIM	
PS	0.108
TOI	
OPS	0.506
DPS	



Test MSE = 33.47676

LESSO : PREDICTED TOP 10

PREDICTED RANK	NAME	VOTE %	ACTUAL RANK
1	Nikita Kucherov	98.07 %	1
2	Connor McDavid	27.19 %	3
3	Leon Draisaitl	0.000 %	N/A
4	Patrick Kane	7.31 %	8
5	Steven Stamkos	0.70 %	14
6	John Tavares	0.41 %	17
7	Brad Marchand	18.83 %	5
8	Nathan Mackinnon	15.50 %	6
9	Alex Ovechkin	12.46 %	7
10	Brayden Point	0.18 %	21

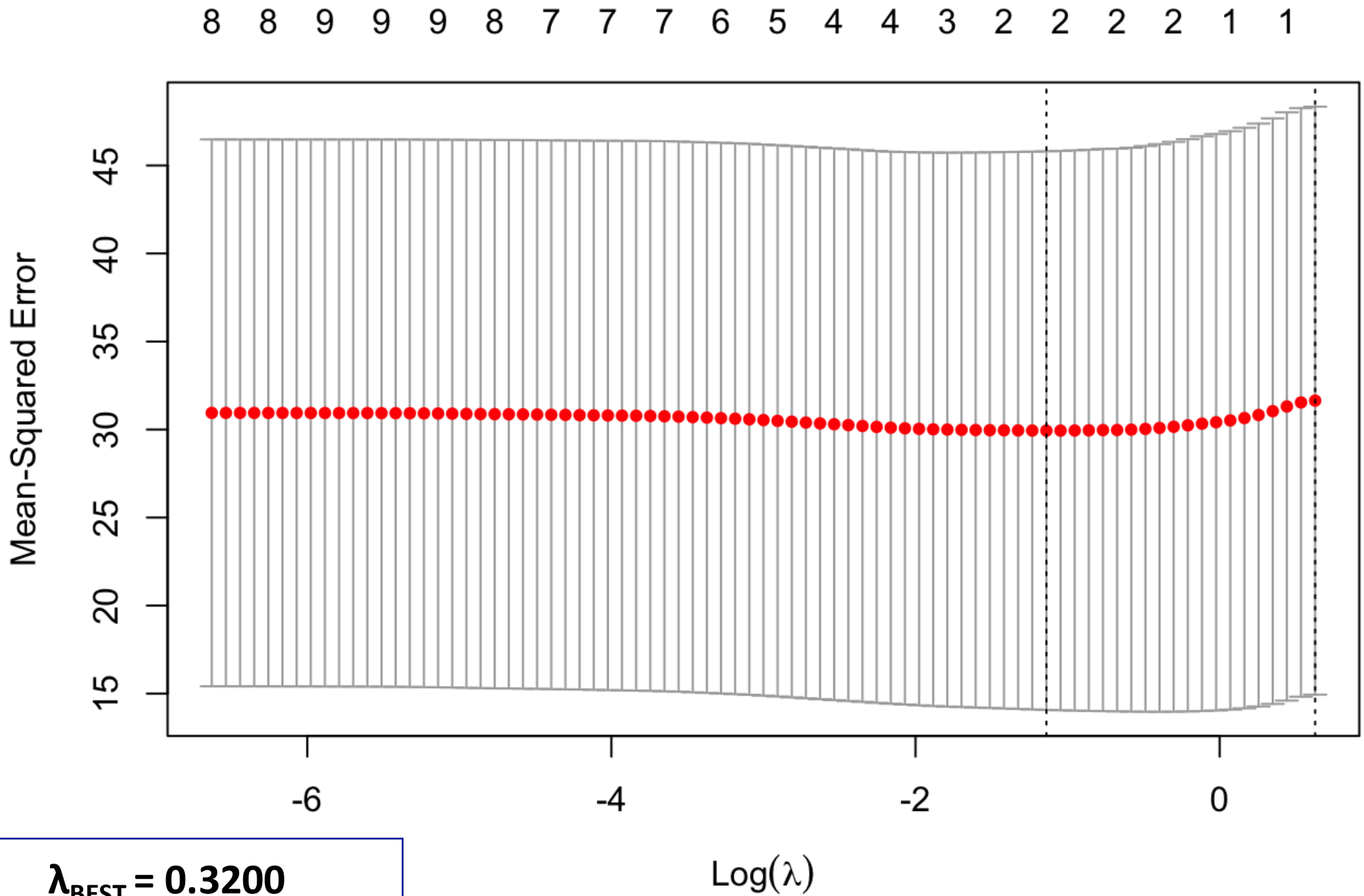


LESSO : ACTUAL TOP 10

ACTUAL RANK	NAME	VOTE %	PREDICTED RANK
1	Nikita Kucherov	98.07 %	1
2	Sidney Crosby	43.22 %	11
3	Connor McDavid	27.19 %	2
4	Johnny Gaudreau	25.85 %	12
5	Brad Marchand	18.83 %	7
6	Nathan Mackinnon	15.50 %	8
7	Alex Ovechkin	12.46 %	9
8	Patrick Kane	7.31 %	4
9	Mark Giordano	3.86 %	23
10	Sebastian Aho Jordan Binnington	1.23 %	28 N/A

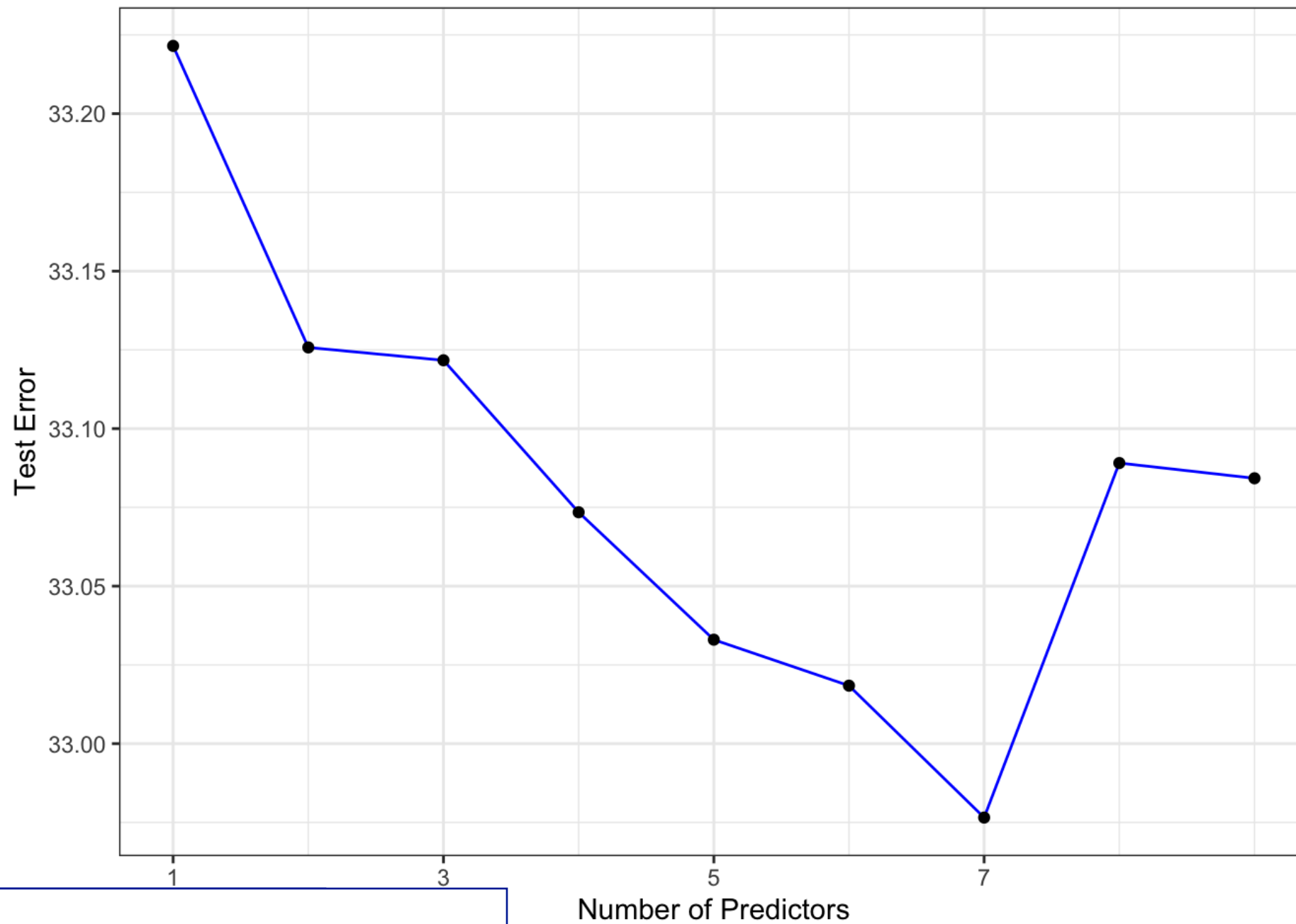


LASSO REGRESSION PLOT



$\lambda_{\text{BEST}} = 0.3200$

SUBSET SELECTION PLOT



OPTIMAL # OF PREDICTORS = 7

THE MODELS – SUBSET SELECTION

COEFFICIENTS:

	ESTIMATE
INTERCEPT	1.058
GP	
G	-0.181
A	0.0555
PLUS_MINUS	0.0399
PIM	
PS	4.435
TOI	-0.00140
OPS	-3.149
DPS	-4.778



Test MSE = 31.78458

SUBSET SELECTION : PREDICTED TOP 10

PREDICTED RANK	NAME	VOTE %	ACTUAL RANK
1	Nikita Kucherov	98.07 %	1
2	Connor McDavid	27.19 %	3
3	Sydney Crosby	43.22 %	2
4	Mitch Marner	0.000 %	N/A
5	Brad Marchand	18.83 %	5
6	Brent Burns	0.29 %	19
7	Johnny Gaudreau	25.85 %	4
8	Patrick Kane	7.31 %	8
9	Patrice Bergeron	0.47 %	15
10	Mark Giordano	3.86 %	9



SUBSET SELECTION : ACTUAL TOP 10

ACTUAL RANK	NAME	VOTE %	PREDICTED RANK
1	Nikita Kucherov	98.07 %	1
2	Sidney Crosby	43.22 %	3
3	Connor McDavid	27.19 %	2
4	Johnny Gaudreau	25.85 %	7
5	Brad Marchand	18.83 %	5
6	Nathan Mackinnon	15.50 %	18
7	Alex Ovechkin	12.46 %	56
8	Patrick Kane	7.31 %	8
9	Mark Giordano	3.86 %	10
10	Sebastian Aho Jordan Binnington	1.23 %	25 N/A



SUMMARY & COMPARING MODEL RESULTS

	LINEAR	RIDGE	LASSO	SUBSET
MSE_{TEST}	33.08424	33.99761	33.47676	31.78458
PREDICTED WINNER	Nikita Kucherov	Nikita Kucherov	Nikita Kucherov	Nikita Kucherov
ACTUAL WINNER	Nikita Kucherov	Nikita Kucherov	Nikita Kucherov	Nikita Kucherov
PREDICTED TOP 10	5 correct players	6 correct players	6 correct players	7 correct players

BEST MODEL : SUBSET SELECTION

