



# ROCKET LEAGUE

## Championship Series CS105 Final Project

David Ryan, Kendrew Christanto, Nicholas Chang, Jonathan Thai

# What is *Rocket League*?

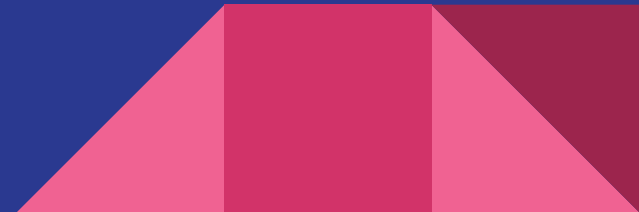
- An online multiplayer game in which up to 8 players play 5-min soccer in rocket-powered cars
- Players control cars to score goals on the opposing team's goal area.
- Players can jump to hit the ball while being mid-air
- Players can utilize speed boost to speed across the field, add momentum towards hitting the ball, or demolish opposing team's cars
- *Rocket League Championship Series* is an esport tournament based on *Rocket League*



# The Goal of this Project

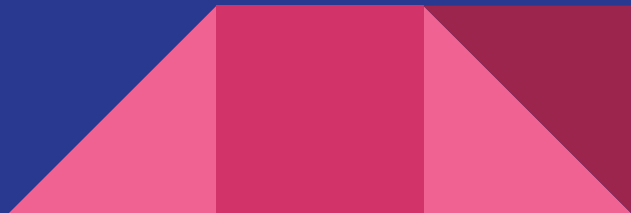


- To draw correlations between winning matches alongside various game statistics of players that perform in the *Rocket League Championship Series 2021-2022*.
- We hope to accomplish by relating winning matches to data points like goals scored, speed boost, movement, positioning, steering sensitivity, and mvp status with respect to wins.
- In order to find correlation, we plan to do:
  - Linear Regression analysis
  - K-Means Clustering
  - Chi-Square Analysis



















# The Data





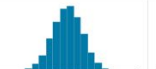






- We needed data for player and team game stats for goals, assists, boost amount, movement speed and distance, offensive and defensive positioning, and steering sensitivity.
  - Since this Esport event is annual, we took the data from 2021-2022.
- Dataset: <https://www.kaggle.com/datasets/dylanmonfret/rlds-202122>
  - Because there's so much data, we looked into specifically `games_by_players.csv` and `games_by_teams.csv`.
- The data is cleaned to only include relevant data points of interest



# Dataset (games\_by\_players.csv)

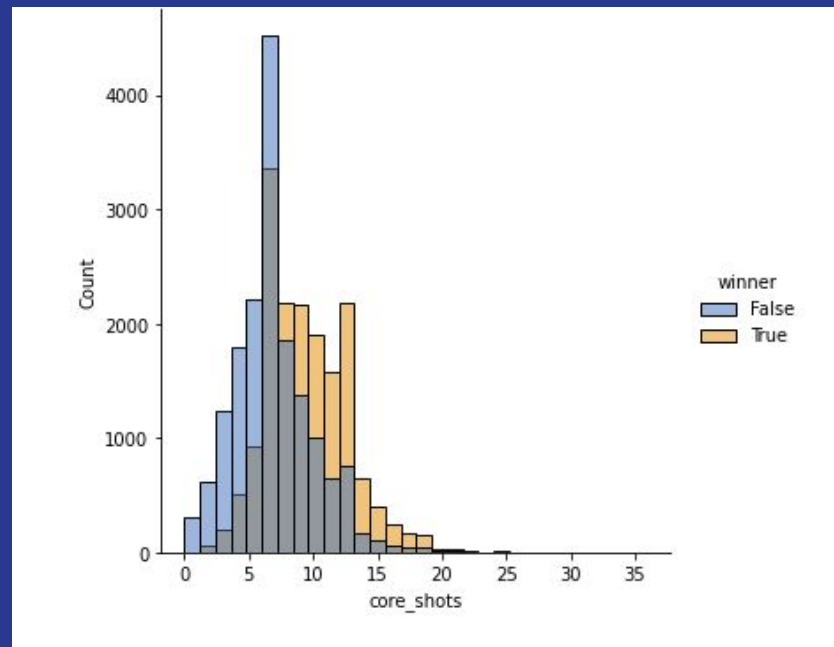
# core_goals	# core_assists	# core_score	# boost_avg...	# boost_tim...	# movement...	# movement...	# movement...	# movement...	# movement...	# positionin...	# positionin...	# positionin...	# positionin...	# positionin...	# advanced...
Goals.	Assists (last pass before a goal).	Game score.	Average amount of boost.	Time spent with full (100) boost.	Average speed.	Total distance traveled in the field.	Part of the time spent on the ground.	Part of the time spent low in the air.	Part of the time spent high in the air.	Average distance to the ball.	Time spent on defensive third.	Time spent on defensive half.	Time spent in the offensive third.	Part of the time spent in the offensive / opponent's half.	Goals participation rate.
															
0	0	348	17.9	0	546	21.8k	36.8	8.2	0	2k	12.4	19.1	0	0	0
1	0	348	45.92	55.88	1446.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
0	1	343	51.26	27.82	1619.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
1	0	238	47.91	36.14	1462.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
0	0	469	53.41	28.71	1588.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
0	0	353	48.43	51.28	1616.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
1	0	246	50.55	33.45	1632.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
2	0	499	53.65	54.26	1597.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
1	2	414	55.37	37.48	1630.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
1	1	370	50.24	28.15	1507.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
2	0	455	49.89	18.12	1552.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
0	0	294	46.19	35.09	1703.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
0	1	274	45.99	31.3	1476.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
2	1	477	49.29	40.35	1503.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
1	2	452	53.17	46.44	1651.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
0	0	272	47.53	44.39	1591.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
1	1	439	45.89	50.4	1571.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
1	1	362	48.22	37.19	1586.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
0	0	247	53.84	27.46	1609.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
5	2	823	50.49	18.96	1674.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
1	4	567	53.12	36.08	1648.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
2	2	548	53.9	39.3	1587.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
0	0	257	47.79	38.73	1607.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
0	0	247	48.07	29.37	1524.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
0	0	126	39.68	18.26	1447.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
3	3	694	49.98	24.49	1597.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
2	2	529	50.23	38.67	1581.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
1	1	344	53.88	56.96	1641.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
1	1	469	38.51	28.82	1616.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
1	0	445	39.33	28.39	1499.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
0	0	124	50.45	40.21	1578.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
2	3	875	55.41	31.98	1593.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0
4	1	802	49.06	26.09	1611.0	1.84m	91.8	58	25.2	4.71k	667	858	64.8	81	50.0

# Dataset (games\_by\_teams.csv)

# ball_possession_t...	# core_shots	# core_goals	# core_shooting_pe...	# boost_avg_amount	# movement_total...	# movement_time_L...	# movement_time...	# positioning_time...	# positioning_time...	# positioning_time...
Time spent with ball possession.	Shots on target.	Goals.	Goals / shots ratio.	Average amount of boost.	Total distance traveled in the field.	Time spent low in the air.	Time spent high in the air.	Time spent in neutral third / midfield.	Time spent in offensive third.	Time spent in defense half.
										
126.54	13	2	15.384615384615385	145.09	1433486.0	481.42999999999995	63.91	291.06	189.48000000000002	660.06
153.6	4	1	25.0	152.39	1539398.0	382.19	64.11	329.25	214.03000000000003	633.39
155.79	18	4	40.0	159.26	1581436.0	413.58	77.14	388.51	264.91999999999996	659.74
126.81	6	2	33.33333333333333	142.07	1561261.0	441.26000000000005	53.86	368.96000000000004	147.48000000000002	762.21
165.76	11	3	27.27272727272727	149.99	1679856.0	459.13	69.78	373.65999999999997	279.05	648.19
141.17	5	2	40.0	147.95	1674109.0	436.89	58.36	378.5	165.27	813.96
140.93	14	8	57.14285714285714	157.51	1686848.0	433.08999999999999	89.81	406.49	338.77	600.6899999999999
126.65	3	0	0.0	135.54	1567035.0	417.30000000000007	58.44	286.2	98.08	916.8199999999999
142.42	14	6	42.857142857142854	154.09	1742004.0	490.64	70.88	455.41999999999996	326.1	627.0
151.1	5	2	40.0	128.29000000000002	1679917.0	435.36	59.050000000000004	346.16999999999996	140.91	883.18000000000006
130.82	18	7	38.88888888888889	157.97	1669160.0	462.01	77.59	385.25	235.89000000000001	744.96
139.44	8	2	25.0	135.26	1646360.0	457.86	56.98	369.59000000000003	195.3	800.17000000000006
148.71	14	2	14.285714285714285	167.74	1461339.0	337.09	69.34	367.22	284.62	505.14
132.09	3	1	33.33333333333333	155.01	1397388.0	363.92	67.14	284.25	82.2	782.5799999999999
142.62	11	7	63.63636363636363	155.42000000000002	1618589.0	454.74	64.94	375.28	241.7	765.94
129.13	6	2	33.33333333333333	149.62	1631433.0	439.64	50.489999999999995	370.63	152.53	857.78
230.29	10	0	0.0	166.95999999999998	2113964.0	525.19	88.44999999999999	493.09000000000003	269.67	953.75
223.13	9	1	11.111111111111111	160.95	2210247.0	526.88	70.83	527.59	297.77000000000004	880.95
132.03	11	3	27.27272727272727	156.72	1486993.0	375.41999999999996	68.12	351.31	202.95	675.83
159.91	6	2	33.33333333333333	160.77	1484975.0	368.86	58.730000000000004	319.93	137.10999999999999	787.0

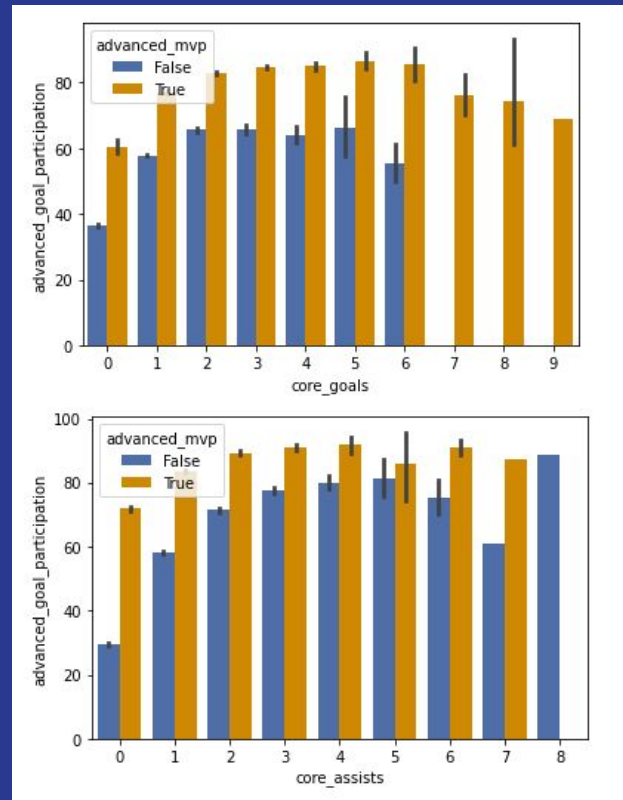
# EDA: Shots based on winning teams

- Winning teams tend to shoot towards the goal more often than losing teams because it leads to higher chances of scoring.



# EDA: MVP players v. Regular players on winning teams

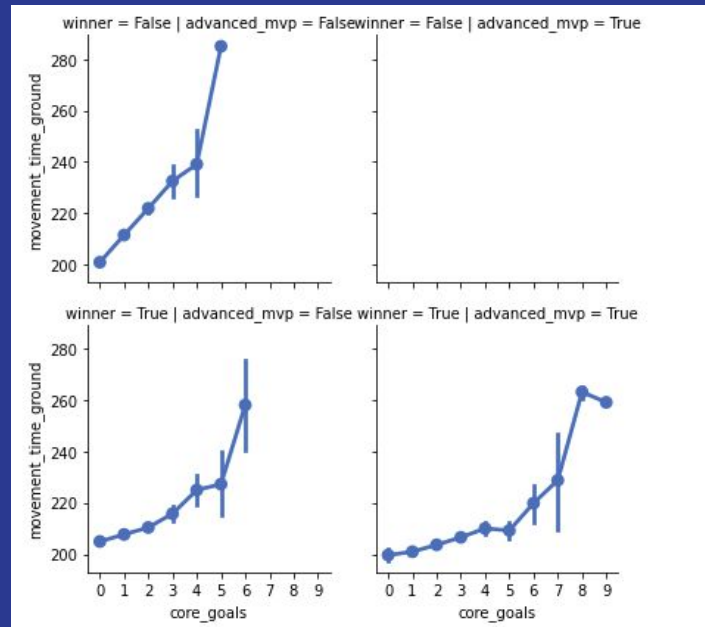
- When comparing MVP players to regular players on winning teams, we find that MVP players tend to score goals more often than regular players, leading to wins
- They also happen to assist more often than regular players
- This means you can win matches in RLCS if you have an MVP player who can shoot goals and assist





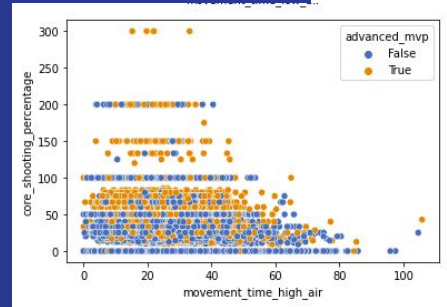
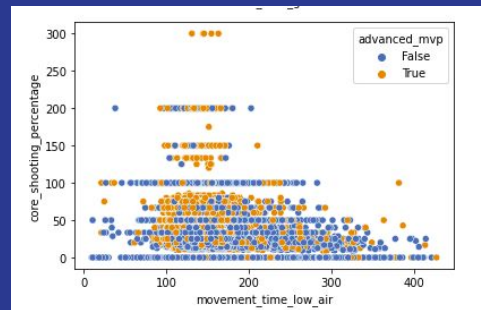
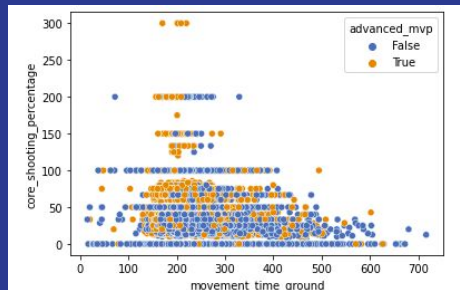
# EDA: Time spent in the ground or air

- Those who are winning players, regardless of MVP status tend to stay on the ground less when scoring early goals.
- MVP players, in fact, score early goals on the ground much less that regular winning players
- This leads to the fact that MVP players score more often in the air



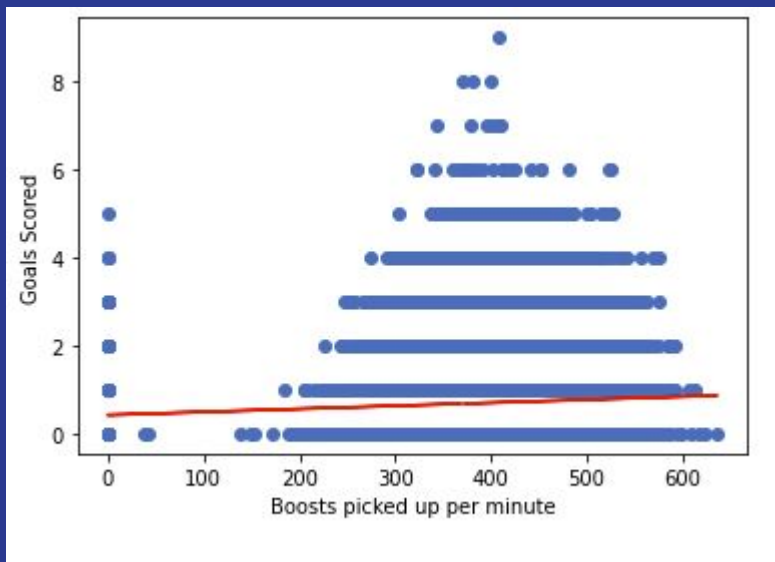
# EDA: Time spent in the ground or air

- This is backed up by these 3 scatter plots here.
- This shows that MVPs with a greater shooting percentage tend to spend more time high in the air, which is most likely where they take their shots from.



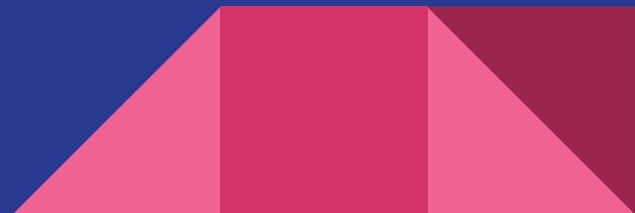
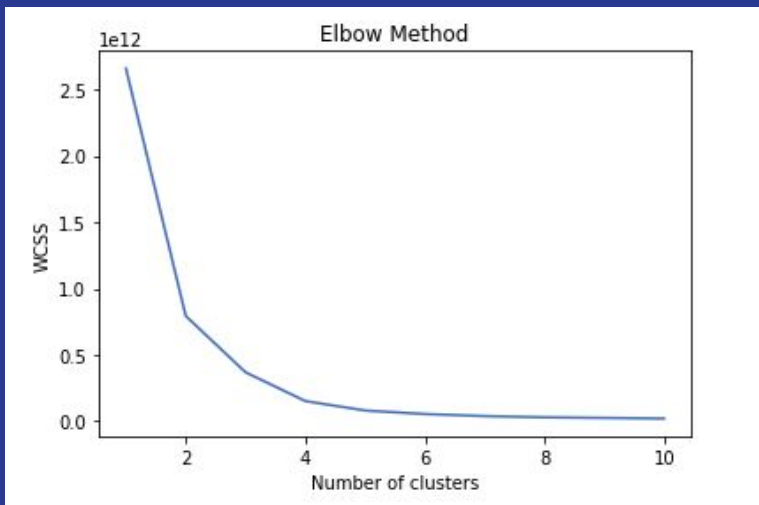
# Linear Regression

- For predicting goals scored based on boost pickups per minute.
- Not a good predictor of goals.



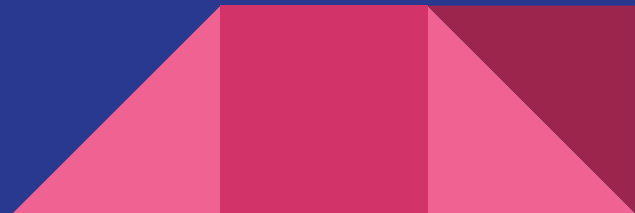
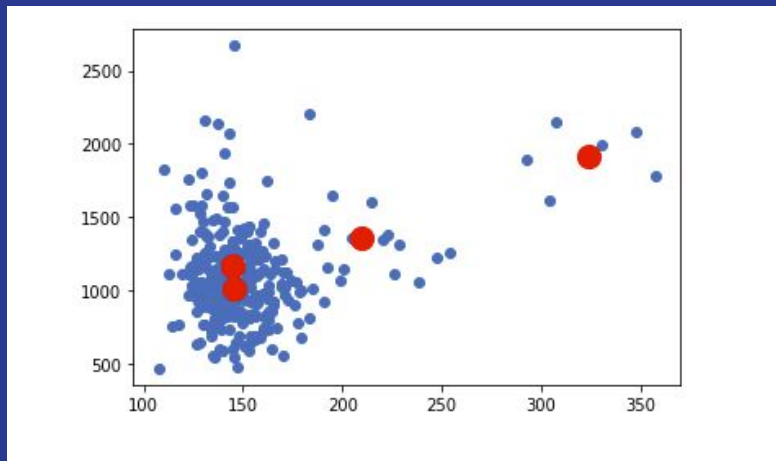
# K-Mean Clustering

- First K-Means used to find clustering for ball possession time and the core score.
- We first find the optimal number of clusters using the elbow method
- We pick the number at the elbow which is 4 in this case.



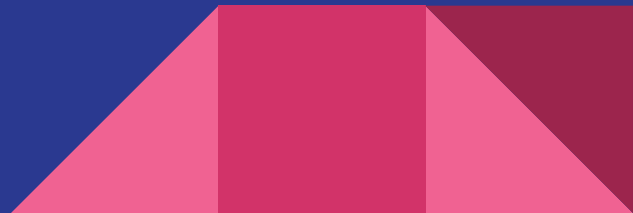
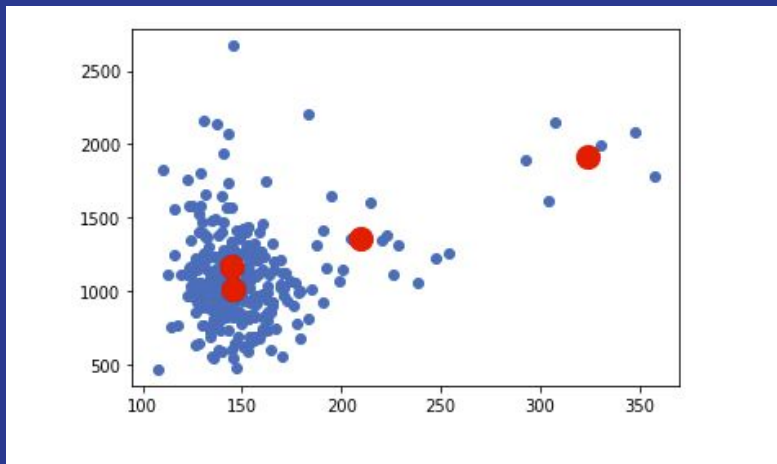
# K-Mean Clustering

- Using the optimal number of clusters we found we plug them
- The points put into the graph which are called centroids
- We calculate the distance of each point to the centroid and find the average
- Those steps are repeated until clusters are found.



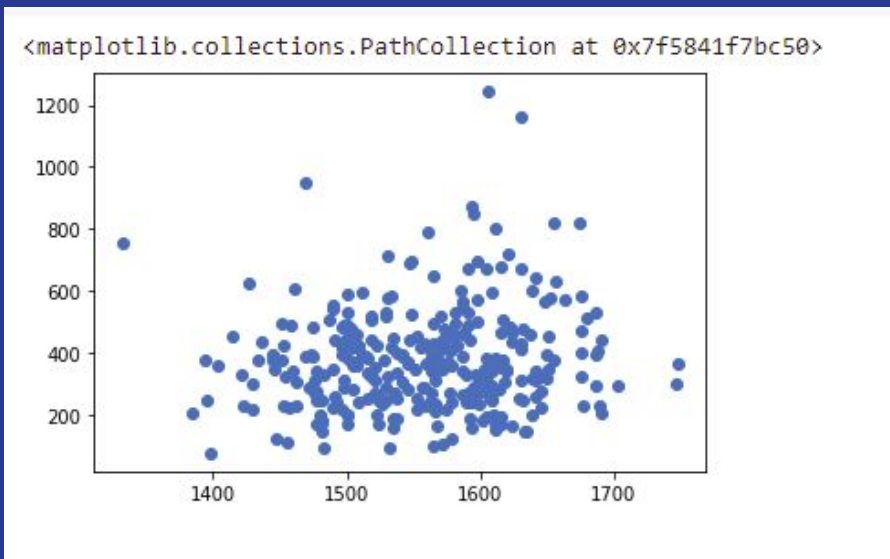
# K-Mean Clustering

- The groups are formed where the clusters are.
- This means that there are clusters of data when looking at a ball possession time and core score graph



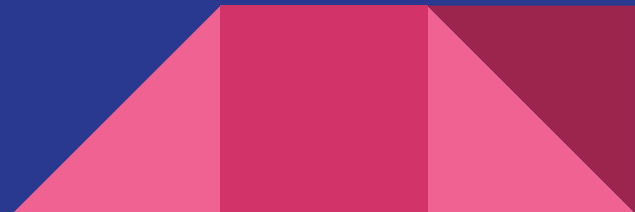
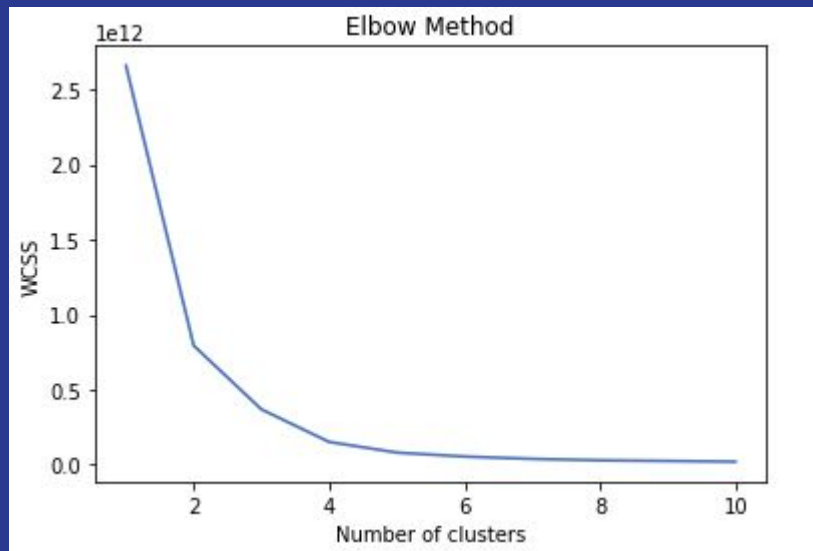
# K-Mean Clustering

- The same is done for the other one for average movement speed and core score



# K-Mean Clustering

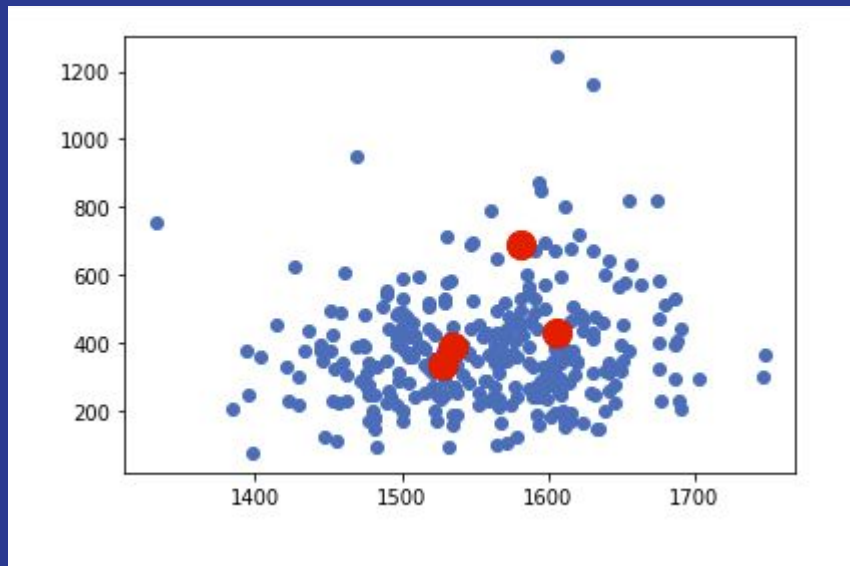
- We find the optimal number of clusters using the elbow method which is 4





# K-Mean Clustering

- And we find the groups using the clusters.



# Chi Squared Test [Correlation Analysis]

- Mathematical test to determine whether there is a statistically significant difference between observed and expected frequencies at the intersection of multiple categorical variables with respect to one variable.
- Can be used to compute a p-value

Formula

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

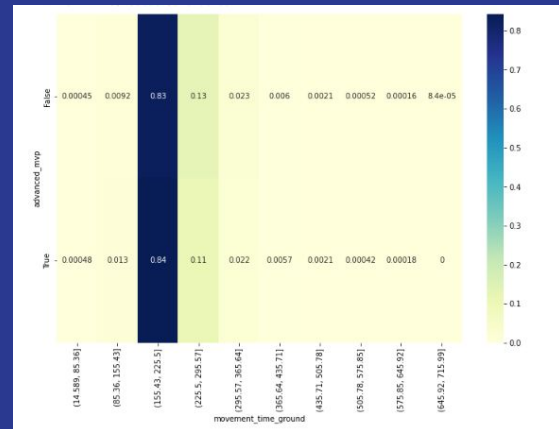
$\chi^2$  = chi squared

$O_i$  = observed value

$E_i$  = expected value

# Findings [MVP Status and Ground Movement Time]

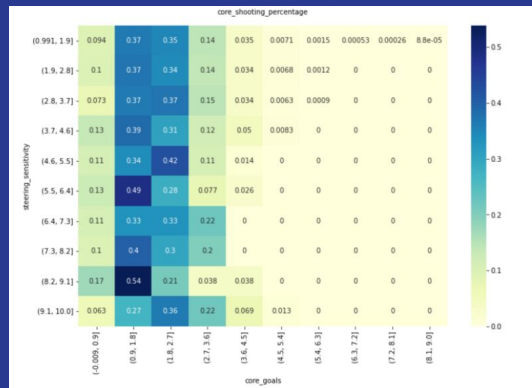
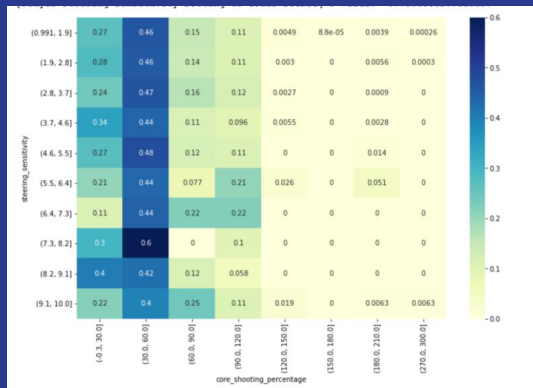
- Produced a contingency table and ran chi squared test. The resulting p-value was 2.4394338384981347e-08



movement_time_ground	(14.589, 85.36]	(85.36, 155.43]	(155.43, 225.5]	(225.5, 295.57]	(295.57, 365.64]	(365.64, 435.71]	(435.71, 505.78]	(505.78, 575.85]	(575.85, 645.92]	(645.92, 715.99]
advanced_mvp										
False	37	763	68944	10632	1951	500	175	43	13	7
True	8	221	13998	1884	364	94	35	7	3	0

# MVP Findings [Steering Sensitivity and Performance]

- [Player Steering Sensitivity Setting to Shooting Percentage] P-Value: 0.000178140588134638
- [Player Steering Sensitivity Setting to Goals Scored] P-Value: 0.7884866377610897



# Question 1:

What statistics are most important for determining an MVP?

# Question 2:

What irregularity can be found when comparing offensive vs defensive positioning statistics?

# Question 3:

Are there any correlations between diasporic metrics and MVP status?