



Ice Around the World

Brianna Pagan

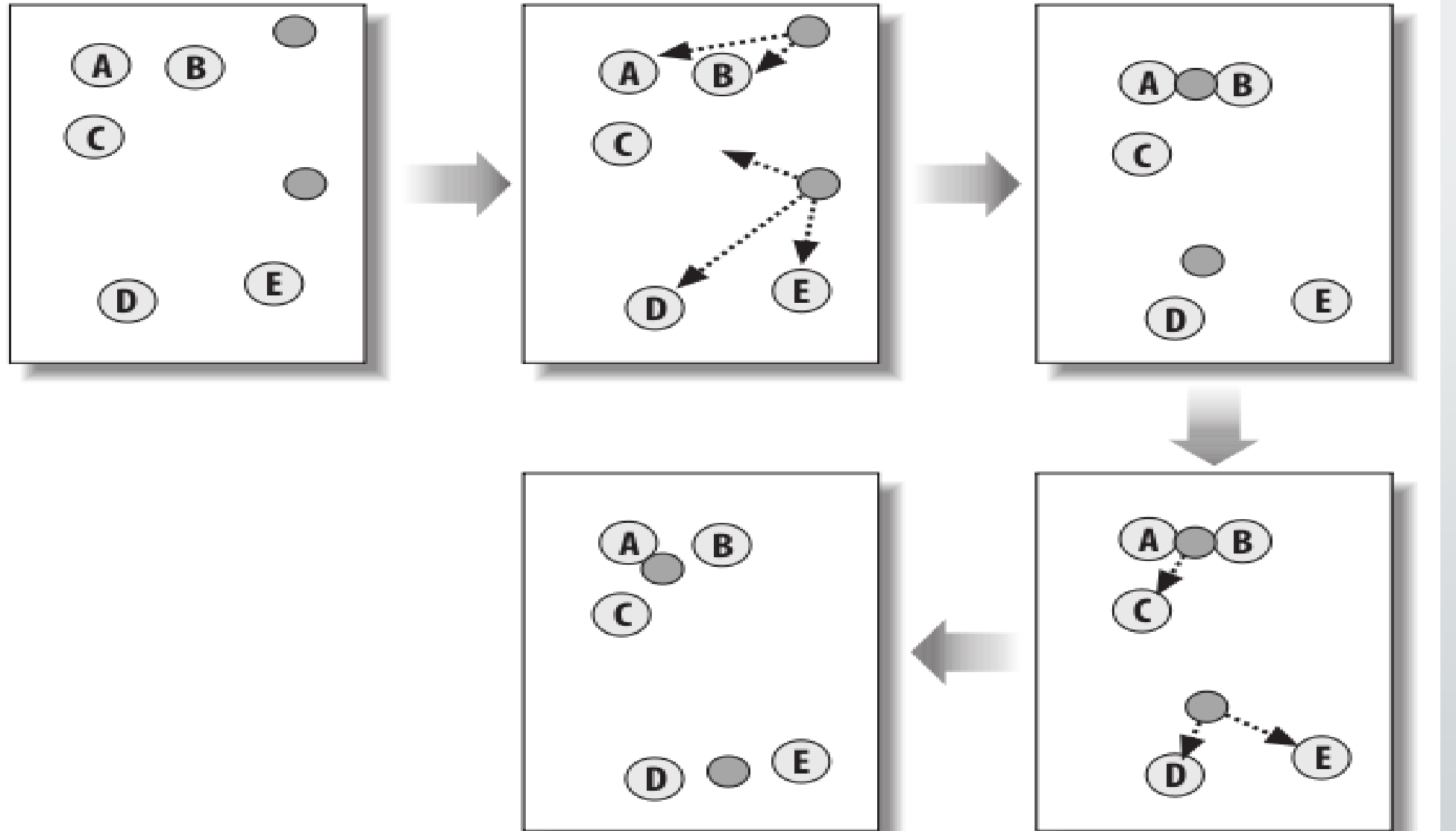
Roadmap

- Hockey dataset
- K-means Clustering Algorithm
- Experiment
- Discovered Information
- Results and Issues
- Summary

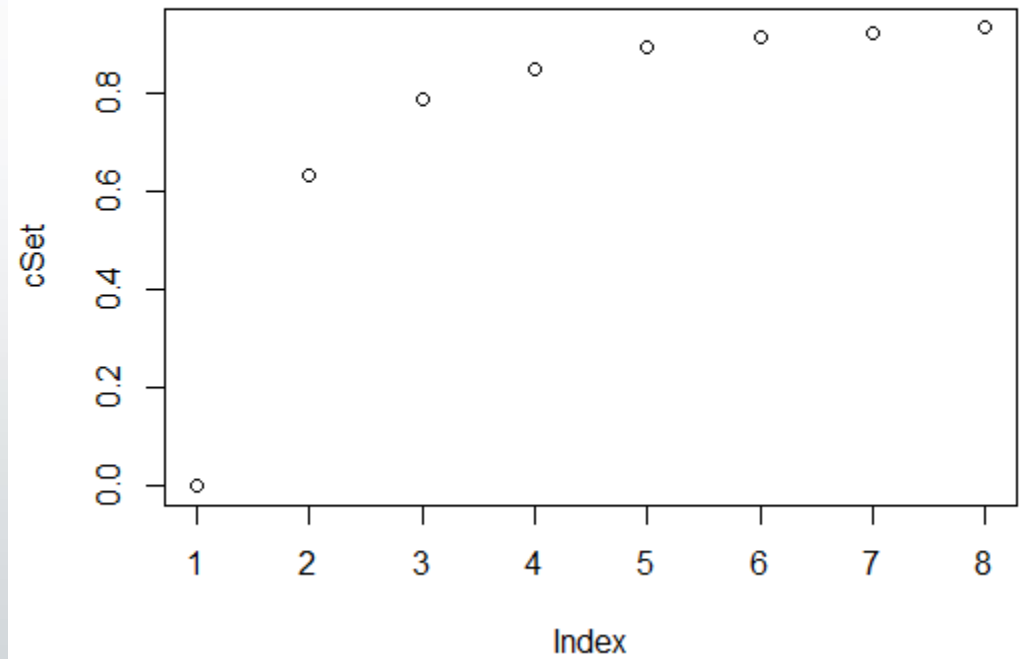
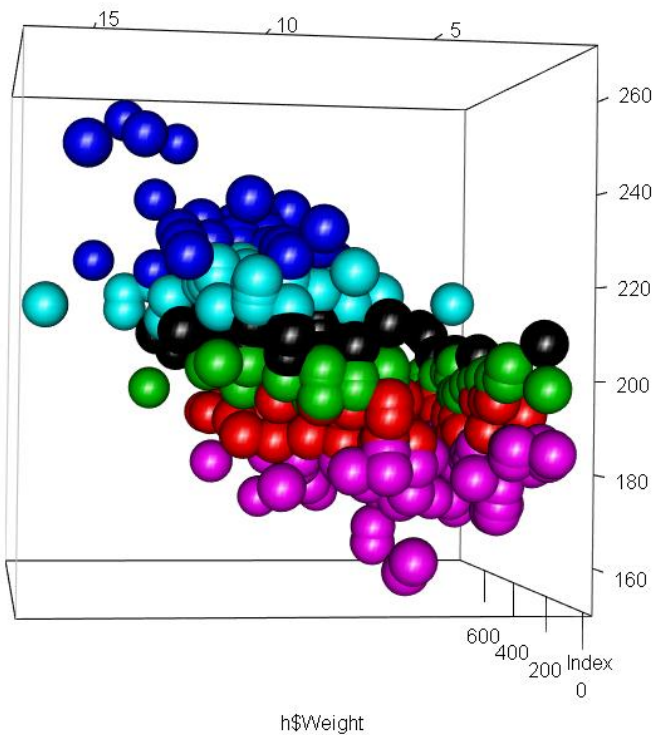
Hockey Dataset

- Contains 741 instances
- Contains 6 attributes for each instance
 - Player name
 - Position
 - Shooting Hand
 - Height
 - Weight
 - Country
- Collected from the official website of the National Hockey League (NHL)

K-means



Choosing Clusters



Clusters

Cluster	Range of Heights	Average Weight
1	5'10-6'6	207.8494
2	5'9-6'4	191.4331
3	5'10-6'6	199.3939
4	6'1-6'9	233.0222
5	5'11-6'7	217.8462
6	5'7-6'4	179.4054

Body Type by Country

- 30.72% of Americans are from Cluster 1

	1	2	3	4	5	6
AUT	1	0	1	0	1	2
CAN	80	79	61	19	68	48
CHE	1	2	2	0	0	1
CZE	6	7	3	3	3	4
DEU	0	0	2	1	2	2
DNK	2	2	0	1	0	2
EST	1	0	0	0	0	0
FIN	5	8	4	0	7	8
FRA	1	1	2	0	0	0
ITA	1	0	0	0	0	0
KAZ	0	0	1	0	0	0
LVA	0	0	1	0	0	0
NOR	0	0	0	0	1	1
RUS	8	8	5	3	4	5
SVK	4	1	2	1	0	2
SVN	0	0	0	0	1	0
SWE	9	19	18	3	10	7
USA	47	30	30	14	33	29

Body Type by Position

- 46.67% of players in Cluster 4 are defensemen
- 34.39% of players in Cluster 2 are centers
- 40.96% of players in Cluster 1 are defensemen
- 27.84% of centers are in Cluster 2
- 28.81% of defensemen are in Cluster 1

	1	2	3	4	5	6
C	35	54	43	4	27	31
D	68	40	41	21	51	15
G	12	12	10	7	13	13
LW	30	27	23	9	15	25
RW	21	24	15	4	24	27

Body Types by Country and Position

- 33.66% of Canadian centers are from Cluster 2
- 52.06% of centers come from Canada

	Centers					
	1	2	3	4	5	6
AUT	0	0	0	0	0	1
CAN	19	34	19	2	13	14
CHE	0	0	0	0	0	1
CZE	2	1	3	1	1	0
DEU	0	0	0	0	1	0
DNK	1	1	0	0	0	0
EST	1	0	0	0	0	0
FIN	1	1	1	0	2	2
LVA	0	0	1	0	0	0
RUS	1	2	2	0	0	2
SVN	0	0	0	0	1	0
SWE	2	6	6	0	3	2
USA	8	9	11	1	6	9

Body Types by Country and Position

	Defensemen					
	1	2	3	4	5	6
CAN	32	20	17	11	25	6
CHE	0	1	2	0	0	0
CZE	3	1	0	1	1	0
DEU	0	0	1	0	1	0
DNK	0	0	0	0	0	1
FIN	1	2	1	0	2	1
FRA	1	1	0	0	0	0
ITA	1	0	0	0	0	0
RUS	3	2	2	2	2	0
SVK	1	0	1	1	0	0
SWE	4	6	5	0	6	2
USA	22	7	12	6	14	5

- 28.83% of Canadian defensemen are from Cluster 1
- 33.33% of American defensemen are from Cluster 1
- 47.03% of defensemen come from Canada

Body Types by Country and Position

- 41.79% of goalies come from Canada

	Goalies					
	1	2	3	4	5	6
AUT	0	0	0	0	0	0
CAN	4	5	5	1	8	5
CZE	1	0	0	0	0	1
DEU	0	0	0	1	0	1
DNK	0	0	0	1	0	0
FIN	0	2	0	0	2	2
KAZ	0	0	1	0	0	0
RUS	2	0	0	0	0	1
SVK	0	0	1	0	0	1
SVN	0	0	0	0	0	0
SWE	0	2	2	2	0	1
USA	5	3	1	2	3	1

Body Types by Country and Position

Right Wing Forwards

	1	2	3	4	5	6
AUT	0	0	0	0	0	1
CAN	10	7	7	2	14	9
CHE	1	0	0	0	0	0
CZE	0	3	0	1	1	3
DEU	0	0	1	0	0	1
DNK	0	1	0	0	0	1
FIN	3	1	1	0	1	1
NOR	0	0	0	0	0	1
RUS	1	3	1	0	1	1
SVK	2	1	0	0	0	0
SWE	0	2	2	1	0	1
USA	4	6	3	0	7	8

- 51.16% of left wing forwards come from Canada
- 42.61% of right wing forwards come from Canada

Left wing Forwards

	1	2	3	4	5	6
AUT	1	0	1	0	1	0
CAN	15	13	13	3	8	14
CHE	0	1	0	0	0	0
CZE	0	2	0	0	0	0
DNK	1	0	0	0	0	0
FIN	0	2	1	0	0	2
FRA	0	0	2	0	0	0
NOR	0	0	0	0	1	0
RUS	1	1	0	1	1	1
SVK	1	0	0	0	0	1
SWE	3	3	3	0	1	1
USA	8	5	3	5	3	6

Results & Issues

- The results are not conclusive
 - No specific body type
 - Only a few results actually dominate a cluster
- The spread of body types for the NHL is not diverse enough to make distinct clusters that do not overlap
- K-means clustering divided the clusters based on mainly a player's weight

Summary

- Hockey dataset
- K-means Clustering Algorithm
- Experiment
- Discovered Information
- Results and Issues
- Summary

References

- Cluster Analysis: Basic Concepts and Algorithms. (n.d.). Retrieved from <https://www.bing.com/cr?IG=50515960EDD742D48D7B9F02F77A927A&CID=35969EA9A9256AC0052F9741A8146BD7&rd=1&h=X75Mhp4mTdZCWPWooywIriWKxZtfgbssdyqBc5BjRao&v=1&r=https://www-users.cs.umn.edu/~kumar/dmbook/ch8.pdf&p=DevEx,5082.1>
- Gove, R. (2015, December 3). *Using the Elbow Method to Determine the Number of Clusters for K-Means Clustering*. Retrieved from Robert Gove's Block: <https://bl.ocks.org/rpgove/0060ff3b656618e9136b>
- Teams. (n.d). Retrieved from <https://www.nhl.com/info/teams>



QUESTIONS?