



# BLACK MAMBA PREDICTION

Application of Binomial Linear Regression to predict field goal outcome from Kobe Bryant

Presented by Hannah Kosinovsky, Yat Leung, and Javier Saldaña

# KOBE "BLACK MAMBA" BRYANT

- Drafted by Charlotte Hornets in the NBA Draft of 1996
- Youngest point guard ever drafted
- Career from 1996 – 2016 (Los Angeles Lakers)
- Accolades
  - 10X All-NBA 1<sup>st</sup> Team
  - All-Defensive 2<sup>nd</sup> Team (2012)
  - 7X All-Defensive 1<sup>st</sup> Team
  - 3X All-Star MVP (2011, 2009, 2007)
  - 2X Finals MVP (2010, 2009)
  - MVP (2008)
  - All-NBA 3<sup>rd</sup> Team (2005)

Source: [http://www.espn.com/nba/player/bio/\\_/id/110/kobe-bryant](http://www.espn.com/nba/player/bio/_/id/110/kobe-bryant)





## THE DATA SET





# THE VARIABLES

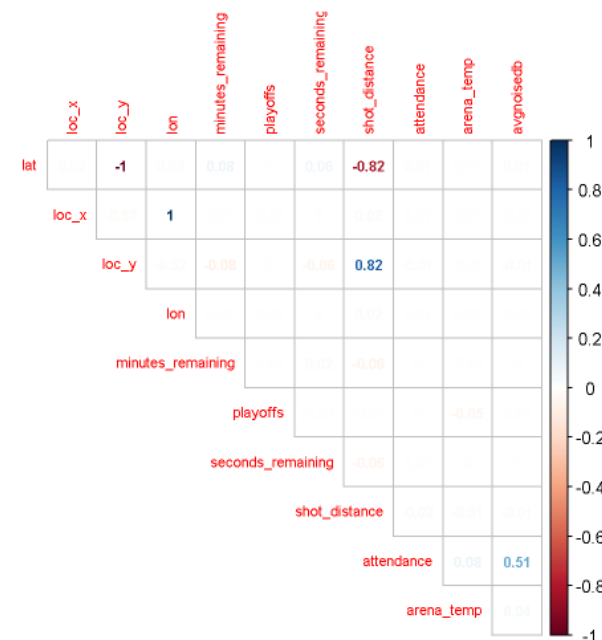
---

---

recID	Action_type	Combined_shot_type	Game_event_id	Game_id
lat	Loc_x	Loc_y	lon	Minutes_remaining
period	playoffs	season	Seconds_remaining	Shot_distance
Shot_made_flag	Shot_type	Shot_zone_area	Shot_zone_basic	Shot_zone_range
Team_id	Team_name	game_date	matchup	opponent
Shot_id	attendance	arenatemp	avgnoisedb	

# MULTICOLLINEARITY & INDEPENDENCE

- Lat  $\rightarrow$  loc\_y & shot\_distance
- Lon  $\rightarrow$  loc\_x & shot\_distance
- action\_type & combined\_shot\_type

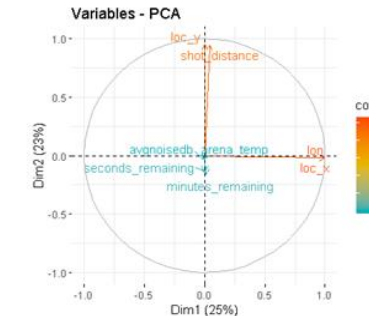
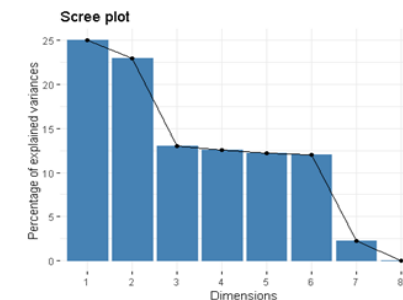
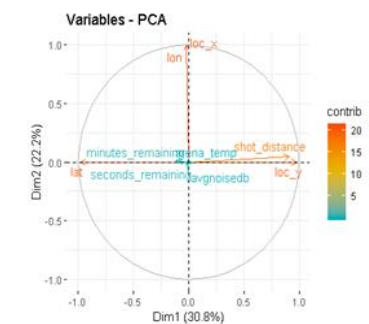
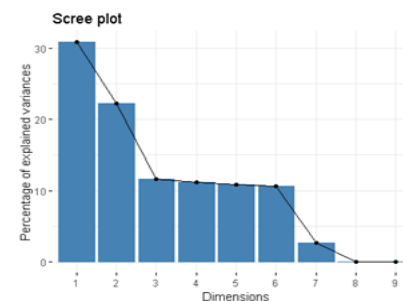


Categorical variables		p-value
action_type	combined_shot_type	< 2.2e-16

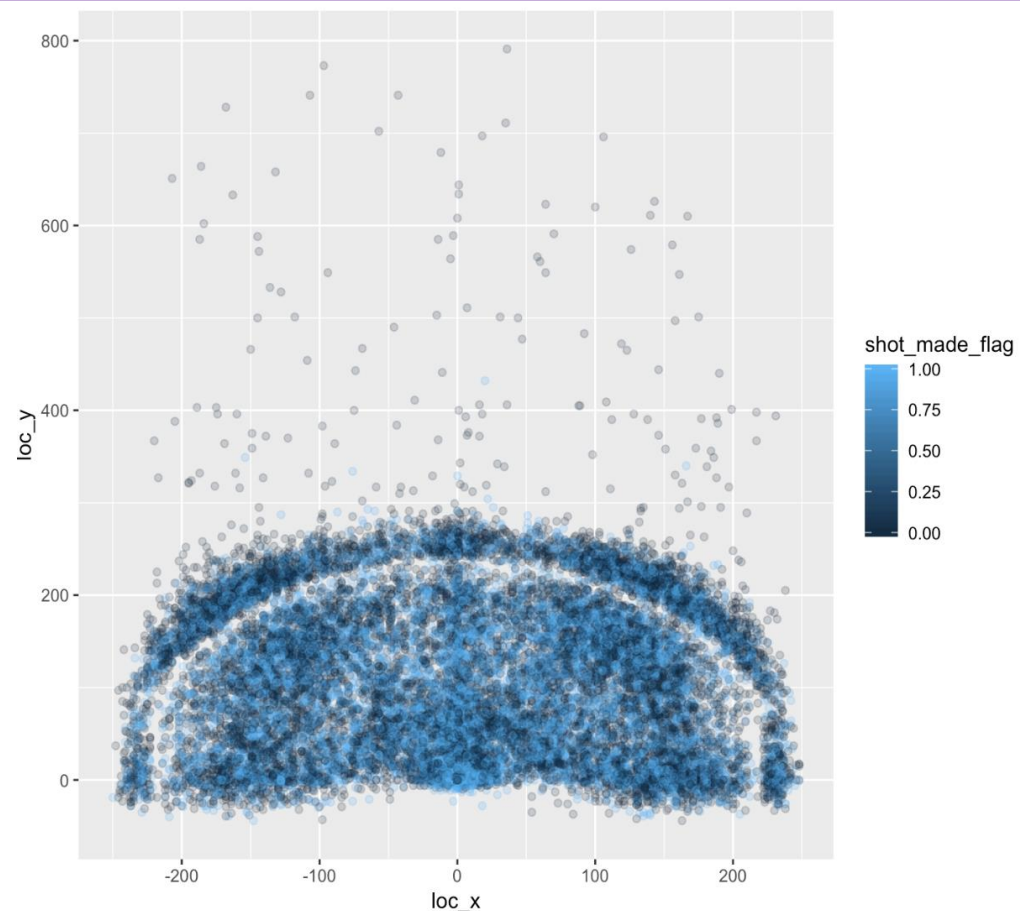
$X\text{-squared} = 128480, df = 270$

# MULTICOLLINEARITY & INDEPENDENCE

- 52% of variance explained w/ 1st
- 48% of variance explained w/o 1st



# FIELD GOALS V. DISTANCE



```
ggplot(data = Kobe)+geom_point(aes(x = loc_x, y = loc_y, color = shot_made_flag),alpha = 1 / 5)
```

- Central Limit Theorem
- Location clusters

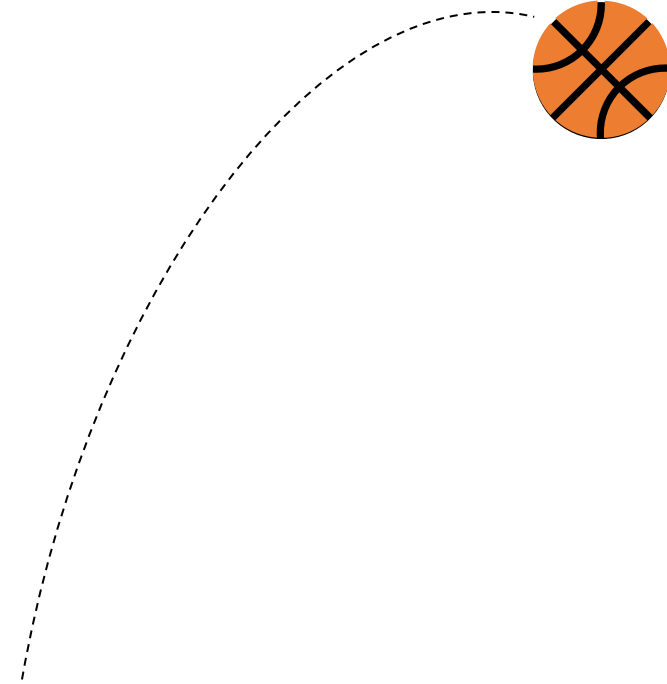


# THE QUESTIONS OF INTEREST



# QUESTIONS OF INTEREST

- Odds of making the shot as distance changes
- Probability of making the shot as distance changes
- Relationship between distance and odds of making the shot in the playoffs
- Predictive model of a successful field goal





## THE MODELS

# DISTANCE V. ODDS OF SUCCESSFUL FIELD GOALS

- Full model v. Reduced model
- Full model
  - Odds ratio = 1.010
  - Not significant
- Reduced model
  - Odds ratio = 0.969
  - Significant

Type 3 Analysis of Effects			
Effect	DF	Wald Chi-Square	Pr > ChiSq
action_type	46	1834.8480	<.0001
combined_shot_type	0	.	.
matchup	69	101.6259	0.0065
opponent	0	.	.
season	11	15.2362	0.1719
shot_type	1	0.0268	0.8700
shot_zone_area	5	34.0368	<.0001
shot_zone_basic	6	20.2411	0.0025
shot_zone_range	2	15.9026	0.0004
arena_temp	1	3.8245	0.0505
attendance	1	77.3901	<.0001
avgnoisedb	1	0.9266	0.3357
game_date	1	0.5975	0.4395
minutes_remaining	1	12.0619	0.0005
period	1	24.4270	<.0001
playoffs	1	0.1074	0.7432
seconds_remaining	1	10.2733	0.0013
shot_distance	1	1.1187	0.2902

Odds Ratio Estimates and Wald Confidence Intervals			
Odds Ratio	Estimate	95% Confidence Limits	
shot_distance	1.010	0.991	1.029

Type 3 Analysis of Effects			
Effect	DF	Wald Chi-Square	Pr > ChiSq
action	1	276.5885	<.0001
period	1	4.7552	0.0292
playoffs	1	0.0937	0.7595
season	11	13.9284	0.2370
time_remaining	1	10.3643	0.0013
shot_distance	1	12.0334	0.0005
shot_type	1	0.0294	0.8638
shot_zone_area	5	14.7072	0.0117
shot_zone_basic	6	13.1413	0.0408
shot_zone_range	2	16.4571	0.0003
game_date	1	0.0005	0.9827
opponent	31	27.9176	0.6254
attendance	1	75.5634	<.0001
arena_temp	1	9.2452	0.0024
avgnoisedb	1	1.8231	0.1769
home_away	1	0.0861	0.7693

Odds Ratio Estimates and Wald Confidence Intervals			
Odds Ratio	Estimate	95% Confidence Limits	
shot_distance	0.969	0.952	0.986

# DISTANCE V. PROBABILITY OF SUCCESSFUL FIELD GOALS

- Negative slope for shot\_distance
- Significant shot\_distance

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	0.3681	0.0224	270.3278	<.0001
shot_distance	1	-0.0441	0.00141	983.1629	<.0001



# FIELD GOALS IN PLAYOFFS & PREDICTIONS

- Odds Ratio in Playoffs
  - Odds increased in “8 – 16 ft” and “8 ft or less”
- Stepwise
  - Playoffs removed
  - Not significant
  - Confusion Matrix

	actual_values	
predicted_values	0	1
0	11764	7630
1	2468	3835

```

**Base logistic model**
glm(shot_made_flag~loc_x+loc_y+lon+playoffs+shot_distance+shot_zone_area
+shot_zone_basic+shot_zone_range,family=binomial(link="logit"),data=kobe)

**Stepwise logistic model**
glm(formula = shot_made_flag ~ shot_distance + shot_zone_area +
shot_zone_basic + shot_zone_range, family = binomial(link = "logit"), data =
kobe)

```

	Odds ratio	2.5 %	97.5 %
(Intercept)	1.962484e-05	3.713612e-85	1.037088e+75
shot_distance	9.594545e-01	9.447178e-01	9.744210e-01
shot_zone_areaCenter(C)	7.100505e+04	1.344892e-75	3.748791e+84
shot_zone_areaLeft Side Center(LC)	6.974884e+04	1.321102e-75	3.682456e+84
shot_zone_areaLeft Side(L)	5.799215e+04	1.098375e-75	3.061877e+84
shot_zone_areaRight Side Center(RC)	7.491551e+04	1.418964e-75	3.955233e+84
shot_zone_areaRight Side(R)	5.987611e+04	1.134059e-75	3.161341e+84
shot_zone_basicBackcourt	8.958763e+03	1.678990e-76	4.780221e+83
shot_zone_basicIn The Paint (Non-RA)	1.138011e+00	9.275034e-01	1.396295e+00
shot_zone_basicleft Corner 3	1.343206e+00	1.005611e+00	1.794135e+00
shot_zone_basicsMid-Range	1.091381e+00	9.565173e-01	1.245259e+00
shot_zone_basicsRestricted Area	1.918734e+00	1.436307e+00	2.563198e+00
shot_zone_basicsRight Corner 3	1.127703e+00	8.730093e-01	1.456700e+00
shot_zone_range24+ ft.	NA	NA	NA
shot_zone_range8-16 ft.	9.383795e-01	8.265318e-01	1.065363e+00
shot_zone_rangeBack Court Shot	NA	NA	NA
shot_zone_rangeLess Than 8 ft.	6.155917e-01	4.917847e-01	7.705673e-01



## THE CONCLUSIONS & INFERENCES

# RESULTS



---

---

## Conclusion

- Odds ratio increases slightly w/ distance
- Probability decreases w/ distance
- Playoffs don't have a significant effect on odds ratio of field goal

## Inferences & Future Studies

- Results only applicable to Kobe
- Kobe played within his skill level
- Broaden scope to all NBA players



**THE END**

Hannah Kosinovsky

Yat Leung

Javier Saldaña