Appendix 2

**Code :**

/\* Kobe Data Import \*/

%web\_drop\_table(WORK.kobeshots);

FILENAME REFFILE '/folders/myfolders/project2Data.xlsx';

/\* Kobe Predict Data Import \*/

%web\_drop\_table(WORK.kobeshotsPred);

FILENAME REFFILE1 '/folders/myfolders/project2Pred.xlsx';

PROC IMPORT DATAFILE=REFFILE

DBMS=XLSX

OUT=WORK.kobeshots;

GETNAMES=YES;

RUN;

PROC IMPORT DATAFILE=REFFILE1

DBMS=XLSX

OUT=WORK.kobeshotsPred;

GETNAMES=YES;

RUN;

PROC CONTENTS DATA=WORK.kobeshotsPred ; RUN;

%web\_open\_table(WORK.kobeshotsPred);

data kobeshots; set kobeshots;

if shot\_zone\_basic = "Above the Break" then szBasic = 4;

if shot\_zone\_basic = "Backcourt" then szBasic = 4;

if shot\_zone\_basic = "In The Paint (Non-RA)" then szBasic = 1;

if shot\_zone\_basic = "Left Corner" then szBasic = 3;

if shot\_zone\_basic = "Mid-Range" then szBasic = 2;

if shot\_zone\_basic = "Restricted Area" then szBasic = 1;

if shot\_zone\_basic = "Right Corner" then szBasic = 3;

if shot\_zone\_range = "24+ ft." then szRange = 3;

if shot\_zone\_range = "8-16 ft." then szRange = 2;

if shot\_zone\_range = "Back Court Shot" then szRange = 4;

if shot\_zone\_range = "Less Than 8ft." then szRange = 1;

if shot\_type = "2PT Field Goal" then shot\_type\_cat = 0;

if shot\_type = "3PT Field Goal" then shot\_type\_cat = 1;

if prxmatch ("/@/",matchup) > 0 then home = 1; else home = 0 ;

run;

/\*\* updated NA data in the prediction file \*/

data kobeshotsPred; set kobeshotsPred;

if shot\_made\_flag eq 'NA' then shot\_made\_flag='';

shot\_made\_flag\_new = input(shot\_made\_flag,3.0);

drop shot\_made\_flag;

rename shot\_made\_flag\_new=shot\_made\_flag;

if shot\_zone\_basic = "Above the Break" then szBasic = 4;

if shot\_zone\_basic = "Backcourt" then szBasic = 4;

if shot\_zone\_basic = "In The Paint (Non-RA)" then szBasic = 1;

if shot\_zone\_basic = "Left Corner" then szBasic = 3;

if shot\_zone\_basic = "Mid-Range" then szBasic = 2;

if shot\_zone\_basic = "Restricted Area" then szBasic = 1;

if shot\_zone\_basic = "Right Corner" then szBasic = 3;

if shot\_zone\_range = "24+ ft." then szRange = 3;

if shot\_zone\_range = "8-16 ft." then szRange = 2;

if shot\_zone\_range = "Back Court Shot" then szRange = 4;

if shot\_zone\_range = "Less Than 8ft." then szRange = 1;

if shot\_type = "2PT Field Goal" then shot\_type\_cat = 0;

if shot\_type = "3PT Field Goal" then shot\_type\_cat = 1;

if prxmatch ("/@/",matchup) > 0 then home = 1; else home = 0 ;

run;

/\*\* Transforming the time variables \*/

data kobeshots;

set kobeshots;

time\_remaining = 60\*minutes\_remaining+seconds\_remaining;

run;

data kobeshotsPred;

set kobeshotsPred;

time\_remaining = 60\*minutes\_remaining+seconds\_remaining;

run;

PROC SURVEYSELECT DATA=kobeshots outall OUT=kobe METHOD=srs SAMPRATE=0.1;

RUN;

PROC SURVEYSELECT DATA=kobeshotsPred outall OUT=kobePred METHOD=srs SAMPRATE=0.1;

RUN;

/\*\* Checking correlation between the selected variables \*/

proc sgscatter data=kobe;

matrix lat lon time\_remaining playoffs shot\_distance attendance arena\_temp avgnoisedb / diagonal=(kernel histogram);

run;

/\*\* Checking outliers \*/

proc reg data=kobe ;

model shot\_made\_flag = lat lon time\_remaining period playoffs shot\_distance attendance arena\_temp avgnoisedb / r;

output out=kbCook cookd=cooks student=students rstudent=studresid;

run;

proc sort data=kbCook out=outSortKB;

by descending cooks;

run;

proc print data=outSortKB (obs=10);

var recId cooks;

run;

/\*\* Running principal components for dataset \*/

proc princomp cov prefix=k data=kobe out=kobe;

var time\_remaining attendance arena\_temp avgnoisedb;

run;

proc princomp cov prefix=k data=kobeshotsPred out=kobeshotsPred;

var time\_remaining attendance arena\_temp avgnoisedb;

run;

PROC SQL;

CREATE TABLE WORK.kobetrain

AS

SELECT

DISTINCT \* FROM WORK.kobe kobe

where kobe.selected = 0;

QUIT;

PROC SQL;

CREATE TABLE WORK.kobetest

AS

SELECT

DISTINCT \* FROM WORK.kobe kobe

where kobe.selected = 1;

QUIT;

proc sort data = kobetrain;

by shot\_type;

run;

/\* shot\_type Analysis - • The odds of Kobe making a shot decrease with respect to the distance he is from the hoop \*/

proc logistic data = kobe plots = all;

class shot\_type home combined\_shot\_type(ref='Jump Shot') /param=ref;

model shot\_made\_flag(event='1') = shot\_distance shot\_type k1 combined\_shot\_type

/ ctable lackfit clparm=wald cl pcorr pprob=.4 .5 .6;

contrast 'shot\_distance vs. shot\_type' shot\_distance 1 -1 0 0 0 0 ;

output out=logisticOut predprobs=I p=predprob resdev=resdev reschi=pearres;

run;

title "Examine Senisitivity and Speicificity for Logistic";

proc freq data=logisticOut; tables shot\_made\_flag\*\_into\_/nocol nopercent; run;

/\* playoffs Analysis - • The relationship between the distance Kobe is from the basket and the odds of him making the shot is different if they are in the playoffs \*/

proc logistic data = kobetrain plots = all;

class shot\_type combined\_shot\_type playoffs /param=ref;

model shot\_made\_flag(event='1') = shot\_distance k1 playoffs combined\_shot\_type

/ ctable lackfit clparm=wald cl pcorr;

contrast 'shot\_distance vs playoffs' shot\_distance 2 0 -1 -1;

output out=logisticOut predprobs=I p=predprob resdev=resdev reschi=pearres;

run;

/\* LDA with Shot\_type principle componenet analysis and shot\_distance \*/

/\* Predicting for shot\_made\_flag for the prediction file data using LDA \*/

proc discrim data=kobe pool=test testdata=kobeshotsPred testout=discrimPredictKB out=discrimKB crossvalidate ;

class shot\_made\_flag ;

var shot\_distance k1 shot\_type\_cat;

priors '0' = 0.5538 '1' = 0.4462;

run;

title "Examine Senisitivity and Speicificity for Discrim";

proc freq data =discrimKB; tables shot\_made\_flag\*\_into\_/nocol nopercent; run;

/\*\* Predicting for shot\_made\_flag for the prediction file data \*/

proc logistic data = kobe plots = all;

class shot\_type home combined\_shot\_type(ref='Jump Shot') /param=ref;

model shot\_made\_flag(event='1') = shot\_distance shot\_type k1 combined\_shot\_type

/ ctable lackfit clparm=wald cl pcorr pprob=.4 .5 .6;

contrast 'shot\_distance vs. shot\_type' shot\_distance 1 -1 0 0 0 0 ;

output out=logisticOut predprobs=I p=predprob resdev=resdev reschi=pearres;

Score data=kobeshotsPred out = logisticPred;

run;