PROC IMPORT

DATAFILE='D:\桌面\教学内容\大三上\数据科学与商务分析\期末报告\tap4fun竞赛数据\tap\_fun\_train.csv'

OUT=TRAIN

DBMS=CSV

REPLACE;

GETNAMES=YES;

RUN;

PROC IMPORT

DATAFILE='D:\桌面\教学内容\大三上\数据科学与商务分析\期末报告\tap4fun竞赛数据\tap\_fun\_test.csv'

OUT=TEST

DBMS=CSV

REPALCE;

GETNAMES=YES;

RUN;

PROC IMPORT OUT= WORK.OUTPOST

DATAFILE= "D:\桌面\教学内容\大三上\数据科学与商务分析\期末报告\tap4fun竞赛数据\科研等级-据点耐久.xlsx"

DBMS=EXCEL REPLACE;

RANGE="工作表1$";

GETNAMES=YES;

MIXED=NO;

SCANTEXT=YES;

USEDATE=NO;

SCANTIME=NO;

RUN;

PROC IMPORT OUT= WORK.MEAT

DATAFILE= "D:\桌面\教学内容\大三上\数据科学与商务分析\期末报告\tap4fun竞赛数据\科研等级-肉类生产.xlsx"

DBMS=EXCEL REPLACE;

RANGE="工作表1$";

GETNAMES=YES;

MIXED=NO;

SCANTEXT=YES;

USEDATE=NO;

SCANTIME=NO;

RUN;

PROC IMPORT OUT= WORK.RESOURCE

DATAFILE= "D:\桌面\教学内容\大三上\数据科学与商务分析\期末报告\tap4fun竞赛数据\原材料获取与消耗数量权重.xlsx"

DBMS=EXCEL REPLACE;

RANGE="Sheet1$";

GETNAMES=YES;

MIXED=NO;

SCANTEXT=YES;

USEDATE=NO;

SCANTIME=NO;

RUN;

PROC IMPORT OUT= WORK.VS

DATAFILE= "D:\桌面\教学内容\大三上\数据科学与商务分析\期末报告\tap4fun竞赛数据\PVP与PVE权重.xlsx"

DBMS=EXCEL REPLACE;

RANGE="Sheet1$";

GETNAMES=YES;

MIXED=NO;

SCANTEXT=YES;

USEDATE=NO;

SCANTIME=NO;

RUN;

DATA PAY;

SET TRAIN;

IF (PAY\_PRICE > 0) THEN

PAY\_7\_FLAG = 1;

ELSE

PAY\_7\_FLAG = 0;

IF (PREDICTION\_PAY\_PRICE > 0) THEN

PAY\_45\_FLAG = 1;

ELSE

PAY\_45\_FLAG = 0;

IF (PREDICTION\_PAY\_PRICE > PAY\_PRICE) THEN

CONTINUE\_PAY = 1;

ELSE

CONTINUE\_PAY = 0;

RUN;

DATA ONLINE\_45;

SET PAY;

IF (AVG\_ONLINE\_MINUTES >= 100) THEN

ONLINE\_FLAG = 100;

ELSE

ONLINE\_FLAG = INT(AVG\_ONLINE\_MINUTES);

IF (((ONLINE\_FLAG < 100) AND (PAY\_45\_FLAG = 1)) OR

((ONLINE\_FLAG >= 1) AND (ONLINE\_FLAG <= 30) AND (PAY\_45\_FLAG = 0)));

RUN;

/\*统计在线时长的频数，按照是否继续付费分类汇总\*/

PROC SORT DATA=ONLINE\_45;

BY PAY\_45\_FLAG;

RUN;

AXIS1 LABEL=("人数");

PROC GCHART DATA=ONLINE\_45;

TITLE "玩家平均在线时间情况汇总-根据是否付费区分";

BY PAY\_45\_FLAG;

VBAR ONLINE\_FLAG / MIDPOINTS = 0 TO 100 BY 2

RAXIS = AXIS1;

RUN;

QUIT;

PROC GCHART DATA=ONLINE\_45;

TITLE "玩家平均在线时间情况汇总-根据是否付费区分";

BY PAY\_45\_FLAG;

VBAR ONLINE\_FLAG / MIDPOINTS = 1 TO 30 BY 1

RAXIS = AXIS1;

RUN;

QUIT;

DATA REGISTER;

SET PAY;

REGISTER\_DATE = DATEPART(REGISTER\_TIME);

YEAR = YEAR(REGISTER\_DATE);

MONTH = MONTH(REGISTER\_DATE);

DAY = DAY(REGISTER\_DATE);

REGISTER\_DAY = MDY(MONTH, DAY, YEAR);

FORMAT REGISTER\_DAY DATE5.;

RUN;

PROC SORT DATA=REGISTER;

BY REGISTER\_DAY;

RUN;

DATA REGISTER\_BY\_DAY;

SET REGISTER;

BY REGISTER\_DAY;

RETAIN COUNT\_REGISTER COUNT\_7 COUNT\_45 COUNT\_CONTINUE;

IF (FIRST.REGISTER\_DAY = 1) THEN

DO;

COUNT\_REGISTER = 0;

COUNT\_7 = 0;

COUNT\_45 = 0;

COUNT\_CONTINUE = 0;

END;

COUNT\_REGISTER = COUNT\_REGISTER + 1;

IF (PAY\_7\_FLAG = 1) THEN

COUNT\_7 = COUNT\_7 + 1;

IF (PAY\_45\_FLAG = 1) THEN

COUNT\_45 = COUNT\_45 + 1;

IF (CONTINUE\_PAY = 1) THEN

COUNT\_CONTINUE = COUNT\_CONTINUE + 1;

IF (LAST.REGISTER\_DAY = 1);

RATE\_7 = COUNT\_7 / COUNT\_REGISTER;

RATE\_45 = COUNT\_45 / COUNT\_REGISTER;

RATE\_CONTINUE = COUNT\_CONTINUE / COUNT\_REGISTER;

RUN;

AXIS1 ORDER=('26JAN18'd TO '06MAR18'd BY DAY)

LABEL=("注册日期");

AXIS2 LABEL=("注册玩家人数");

AXIS3 ORDER=(600 TO 5000 BY 200)

LABEL=("前7日付费人数");

AXIS4 ORDER=(600 TO 5000 BY 200)

LABEL=("前45日付费人数");

AXIS5 ORDER=(200 TO 2000 BY 100)

LABEL=("第7-45日付费人数");

PROC GPLOT DATA=REGISTER\_BY\_DAY;

TITLE "玩家注册日期统计-按日汇总";

PLOT COUNT\_REGISTER\*REGISTER\_DAY / HAXIS=AXIS1

VAXIS=AXIS2;

SYMBOL VALUE=STAR

COLOR=BLACK

INTERPOL=JOIN;

RUN;

QUIT;

PROC GPLOT DATA=REGISTER\_BY\_DAY;

TITLE "玩家注册人数与前7日付费人数汇总-按日汇总";

PLOT COUNT\_REGISTER\*REGISTER\_DAY / HAXIS=AXIS1

VAXIS=AXIS2;

PLOT2 COUNT\_7\*REGISTER\_DAY / VAXIS=AXIS3;

SYMBOL VALUE=STAR

COLOR=BLACK

INTERPOL=JOIN;

SYMBOL2 VALUE=PLUS

COLOR=RED

INTERPOL=JOIN;

RUN;

QUIT;

PROC SGPLOT DATA=REGISTER\_BY\_DAY;

TITLE "前7日付费人数占注册人数的比例-按日汇总";

VLINE REGISTER\_DAY / RESPONSE=COUNT\_REGISTER;

VBAR REGISTER\_DAY / RESPONSE=RATE\_7

Y2AXIS;

XAXIS VALUES=('26JAN18'd TO '06MAR18'd BY DAY)

LABEL="注册日期";

YAXIS VALUES=(30000 TO 120000 BY 10000)

LABEL="注册人数";

Y2AXIS VALUES=(0 TO 0.3 BY 0.05)

LABEL="前7日付费人数比例";

RUN;

PROC GPLOT DATA=REGISTER\_BY\_DAY;

TITLE "玩家注册人数与前45日付费人数汇总-按日汇总";

PLOT COUNT\_REGISTER\*REGISTER\_DAY / HAXIS=AXIS1

VAXIS=AXIS2;

PLOT2 COUNT\_45\*REGISTER\_DAY / VAXIS=AXIS4;

SYMBOL VALUE=STAR

COLOR=BLACK

INTERPOL=JOIN;

SYMBOL2 VALUE=PLUS

COLOR=RED

INTERPOL=JOIN;

RUN;

QUIT;

PROC SGPLOT DATA=REGISTER\_BY\_DAY;

TITLE "前45日付费人数占注册人数的比例-按日汇总";

VLINE REGISTER\_DAY / RESPONSE=COUNT\_REGISTER;

VBAR REGISTER\_DAY / RESPONSE=RATE\_45

Y2AXIS;

XAXIS VALUES=('26JAN18'd TO '06MAR18'd BY DAY)

LABEL="注册日期";

YAXIS VALUES=(30000 TO 120000 BY 10000)

LABEL="注册人数";

Y2AXIS VALUES=(0 TO 0.3 BY 0.05)

LABEL="前45日付费人数比例";

RUN;

PROC GPLOT DATA=REGISTER\_BY\_DAY;

TITLE "玩家注册人数与第7-45日付费人数汇总-按日汇总";

PLOT COUNT\_REGISTER\*REGISTER\_DAY / HAXIS=AXIS1

VAXIS=AXIS2;

PLOT2 COUNT\_CONTINUE\*REGISTER\_DAY / VAXIS=AXIS5;

SYMBOL VALUE=STAR

COLOR=BLACK

INTERPOL=JOIN;

SYMBOL2 VALUE=PLUS

COLOR=RED

INTERPOL=JOIN;

RUN;

QUIT;

PROC SGPLOT DATA=REGISTER\_BY\_DAY;

TITLE "第7-45日付费人数占注册人数的比例-按日汇总";

VLINE REGISTER\_DAY / RESPONSE=COUNT\_REGISTER;

VBAR REGISTER\_DAY / RESPONSE=RATE\_CONTINUE

Y2AXIS;

XAXIS VALUES=('26JAN18'd TO '06MAR18'd BY DAY)

LABEL="注册日期";

YAXIS VALUES=(30000 TO 120000 BY 10000)

LABEL="注册人数";

Y2AXIS VALUES=(0 TO 0.1 BY 0.01)

LABEL="第7-45日付费人数比例";

RUN;

DATA REGISTER\_ACTIVE;

SET REGISTER;

IF (AVG\_ONLINE\_MINUTES >= 10) THEN

ACTIVE = 1;

IF (ACTIVE = 1);

RUN;

PROC SORT DATA=REGISTER\_ACTIVE;

BY REGISTER\_DAY;

RUN;

DATA REGISTER\_ACTIVE;

SET REGISTER\_ACTIVE;

BY REGISTER\_DAY;

RETAIN COUNT\_ACTIVE COUNT\_PAY;

IF (FIRST.REGISTER\_DAY = 1) THEN

DO;

COUNT\_ACTIVE = 0;

COUNT\_PAY = 0;

END;

COUNT\_ACTIVE = COUNT\_ACTIVE + 1;

IF (PAY\_45\_FLAG = 1) THEN

COUNT\_PAY = COUNT\_PAY + 1;

IF (LAST.REGISTER\_DAY = 1);

RATE\_PAY = COUNT\_PAY / COUNT\_ACTIVE;

RUN;

AXIS1 ORDER=('26JAN18'd TO '06MAR18'd BY DAY)

LABEL=("注册日期");

AXIS2 LABEL=("活跃玩家人数");

PROC GPLOT DATA=REGISTER\_ACTIVE;

TITLE "活跃玩家注册日期统计-按日汇总";

PLOT COUNT\_ACTIVE\*REGISTER\_DAY / HAXIS=AXIS1

VAXIS=AXIS2;

SYMBOL VALUE=STAR

COLOR=BLACK

INTERPOL=JOIN;

RUN;

QUIT;

PROC SGPLOT DATA=REGISTER\_ACTIVE;

TITLE "前45日付费人数占活跃人数的比例-按日汇总";

VLINE REGISTER\_DAY / RESPONSE=RATE\_PAY;

VBAR REGISTER\_DAY / RESPONSE=COUNT\_ACTIVE

Y2AXIS;

XAXIS VALUES=('26JAN18'd TO '06MAR18'd BY DAY)

LABEL="注册日期";

YAXIS VALUES=(0 TO 0.2 BY 0.01)

LABEL="付费人数比例";

Y2AXIS VALUES=(5000 TO 20000 BY 1000)

LABEL="活跃人数";

RUN;

PROC SORT DATA=REGISTER;

BY REGISTER\_DAY;

RUN;

DATA REGISTER\_PAYSUM;

SET REGISTER;

BY REGISTER\_DAY;

RETAIN PAYSUM\_7 PAYSUM\_CONTINUE;

IF (FIRST.REGISTER\_DAY = 1) THEN

DO;

PAYSUM\_7 = 0;

PAYSUM\_CONTINUE = 0;

END;

PAYSUM\_7 = PAYSUM\_7 + PAY\_PRICE;

PAYSUM\_CONTINUE = PAYSUM\_CONTINUE + PREDICTION\_PAY\_PRICE - PAY\_PRICE;

IF (LAST.REGISTER\_DAY = 1);

RUN;

PROC GCHART DATA=REGISTER\_PAYSUM;

TITLE "玩家前7日付费金额汇总-按日汇总";

VBAR REGISTER\_DAY / SUMVAR=PAYSUM\_7

TYPE=SUM

DISCRETE;

RUN;

QUIT;

PROC GCHART DATA=REGISTER\_PAYSUM;

TITLE "玩家第7-45日付费金额汇总-按日汇总";

VBAR REGISTER\_DAY / SUMVAR=PAYSUM\_CONTINUE

TYPE=SUM

DISCRETE;

RUN;

QUIT;

AXIS1 ORDER=('26JAN18'd TO '06MAR18'd BY DAY)

LABEL=("注册日期");

AXIS2 ORDER=(0 TO 120000 BY 10000)

LABEL=("前7日付费金额");

AXIS3 ORDER=(0 TO 120000 BY 10000)

LABEL=("第7-45日付费金额");

PROC GPLOT DATA=REGISTER\_PAYSUM;

TITLE "玩家前7日与第7-45日付费金额汇总-按日汇总";

PLOT PAYSUM\_7\*REGISTER\_DAY / HAXIS=AXIS1

VAXIS=AXIS2;

PLOT2 PAYSUM\_CONTINUE\*REGISTER\_DAY / VAXIS=AXIS3;

SYMBOL VALUE=STAR

COLOR=BLACK

INTERPOL=JOIN;

SYMBOL2 VALUE=PLUS

COLOR=RED

INTERPOL=JOIN;

RUN;

QUIT;

PROC SGPLOT DATA=REGISTER\_PAYSUM;

TITLE "玩家前7日与第7-45日付费金额汇总-按日汇总";

VBAR REGISTER\_DAY / RESPONSE=PAYSUM\_7;

VBAR REGISTER\_DAY / RESPONSE=PAYSUM\_CONTINUE

TRANSPARENCY=0.5

Y2AXIS;

RUN;

AXIS1 LABEL=("科研：据点耐久等级");

AXIS2 LABEL=("比例");

AXIS3 LABEL=("付费能力水平");

PROC GCHART DATA=OUTPOST;

TITLE "各付费能力水平玩家比例汇总-按“科研：据点耐久等级”分类";

VBAR CLASS\_LEVEL / SUMVAR=PROPORTION

TYPE=SUM

GROUP=sr\_outpost\_durability\_level

DISCRETE

GAXIS=AXIS1

RAXIS=AXIS2

MAXIS=AXIS3;

RUN;

QUIT;

AXIS1 LABEL=("科研：肉类生产等级");

AXIS2 LABEL=("玩家人数");

AXIS3 LABEL=("付费能力水平");

PROC GCHART DATA=MEAT;

TITLE "各付费能力水平玩家人数汇总-按“科研：肉类生产等级”分类";

VBAR CLASS\_LEVEL / SUMVAR=COUNT

TYPE=SUM

GROUP=sr\_rss\_d\_prod\_level

DISCRETE

GAXIS=AXIS1

RAXIS=AXIS2

MAXIS=AXIS3;

RUN;

QUIT;

AXIS1 LABEL=("权重")

ORDER=(0 TO 0.3 BY 0.05);

AXIS2 LABEL=("字段解释");

PROC GCHART DATA=RESOURCE;

TITLE "原材料获取与消耗数量权重";

VBAR VARIABLE / SUMVAR=PROPORTION

TYPE=SUM

DISCRETE

RAXIS=AXIS1

MAXIS=AXIS2;

RUN;

QUIT;

AXIS1 LABEL=("权重")

ORDER=(0 TO 0.5 BY 0.05);

AXIS2 LABEL=("字段解释");

PROC GCHART DATA=VS;

TITLE "PVP与PVE权重";

VBAR VARIABLE / SUMVAR=PROPORTION

TYPE=SUM

DISCRETE

RAXIS=AXIS1

MAXIS=AXIS2;

RUN;

QUIT;