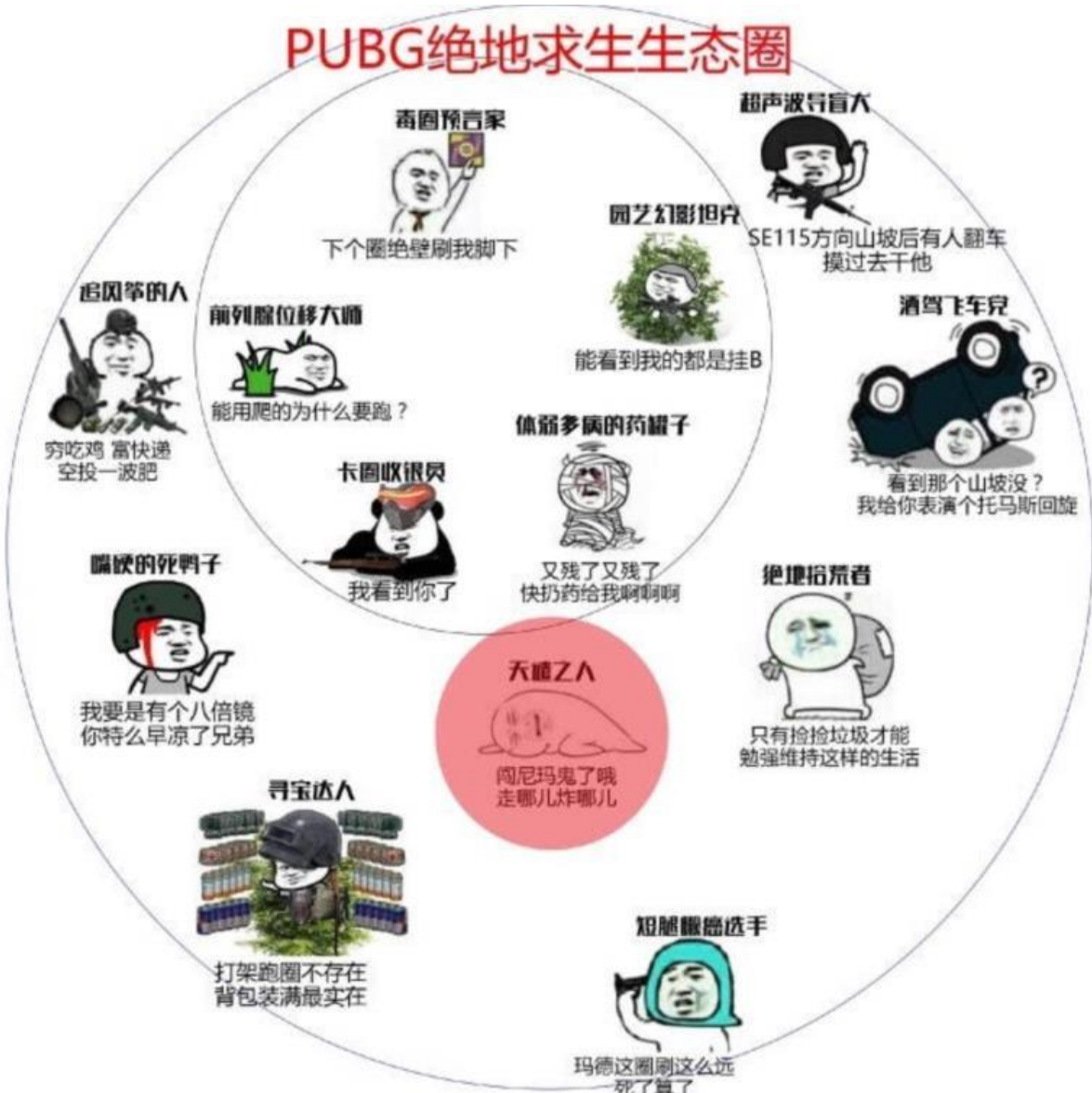


策略聚类及最终排名预测

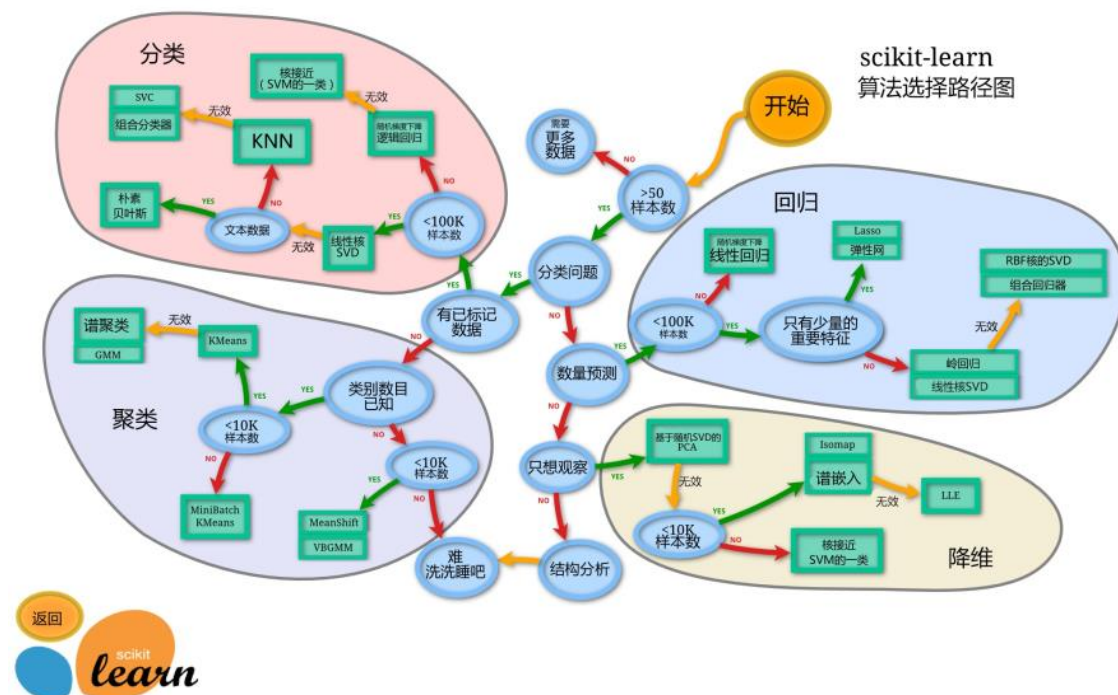
Player Unknown's Battle Grounds (PUBG)

Idea

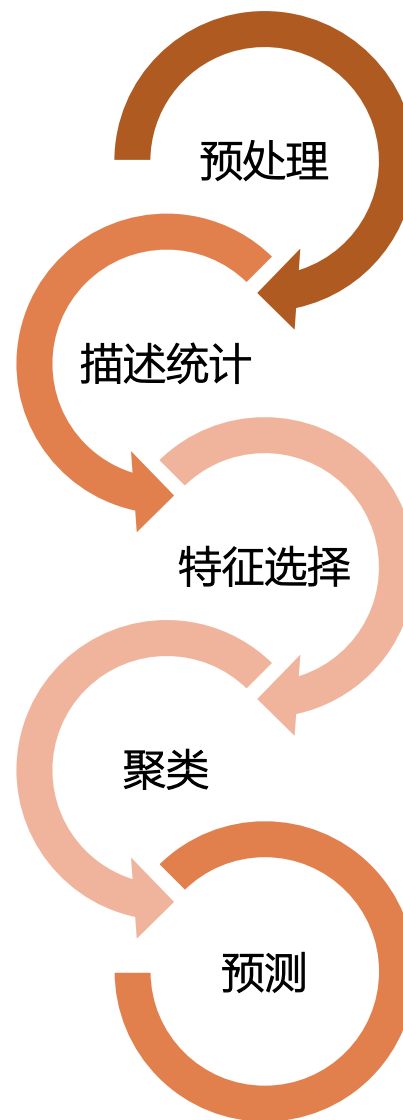


CONTENTS

- 关于在计算的过程中，如何选择合适的算法进行计算，可以参考scikit learn官方给的指导意见：



https://blog.csdn.net/cq_41424526



448万



3*5万

CONTENTS

ONE

PREPROCESS

CLICK

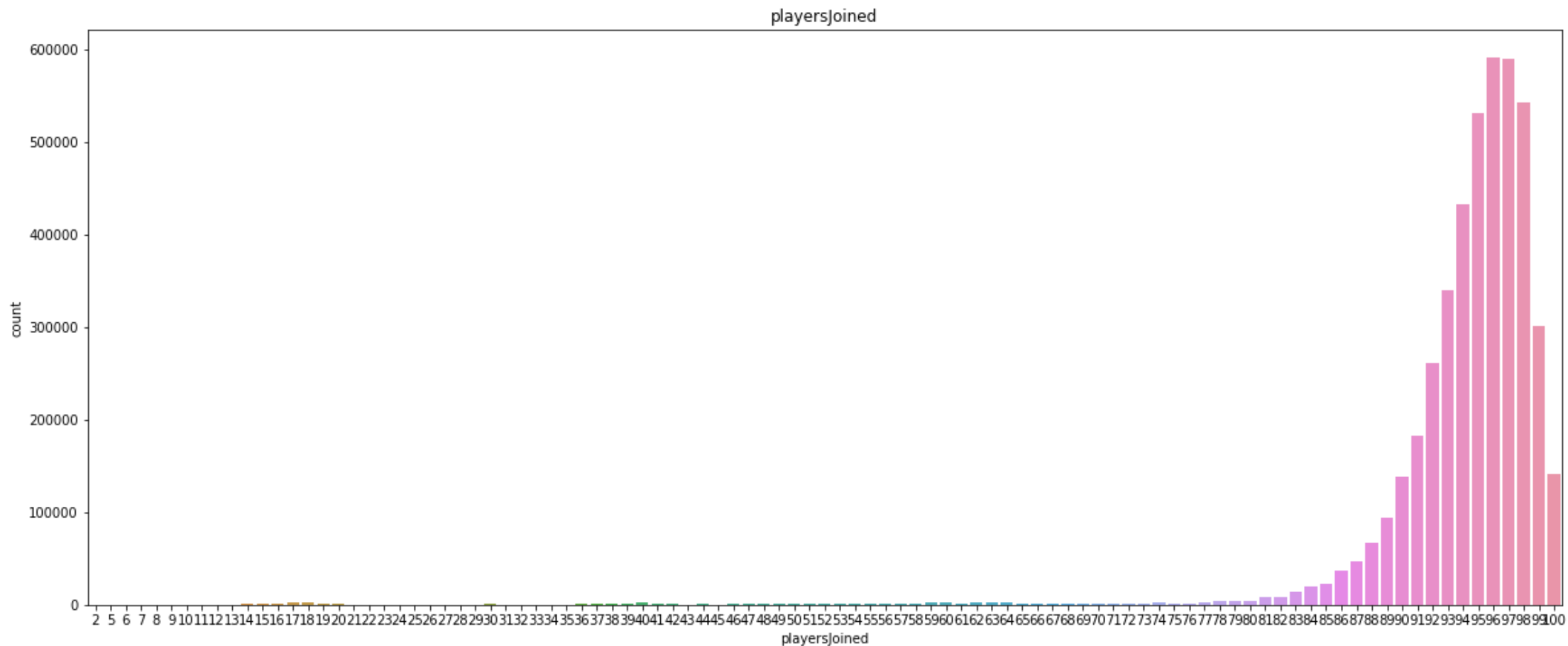
TWO

THREE

FOUR

原始数据集的变量：

- ◆ Id：用户id；
- ◆ groupId：所处小队id；
- ◆ matchId：该场比赛id；
- ◆ assists：助攻数；
- ◆ boosts：使用能量、道具数量；
- ◆ damageDealt：造成的总伤害-自己所受的伤害；
- ◆ DBNOs：击倒敌人数量；
- ◆ headshotKills：爆头数，通过爆头杀死的敌人数；
- ◆ heals：使用治疗药品数量；
- ◆ killPlace：本场比赛杀敌数量排行；
- ◆ kills：杀敌数；
- ◆ killStreaks：短时间内连续杀敌数；
- ◆ longestKill：最远杀敌距离；
- ◆ matchDuration：比赛时长；
- ◆ matchType：比赛类型(小组人数)；
- ◆ maxPlace：本局最差名次；
- ◆ numGroups：小组数量；
- ◆ revives：救活队友的次数；
- ◆ rideDistance：驾车距离；
- ◆ roadKills：驾车杀敌数；
- ◆ swimDistance：游泳距离；
- ◆ teamKills：杀死队友的次数；
- ◆ vehicleDestroys：毁坏机动车的数量；
- ◆ walkDistance：步行距离；
- ◆ weaponsAcquired：收集武器的数量；
- ◆ winPlacePerc：百分比排名，取值0-1，越接近0，排名越靠前。



- 每局比赛的参与人数不一：选择大于75人的，并将有关变量规范化至100人局

规范化部分特征值

```
In [9]: # 规范化部分特征值, 排除赛局人数不同带来的影响
# 使得每局水平都以100人局相当
#data['assists'] = data['assists']*((100/data['playersJoined']))
data['assists'] = data['assists']*((100-data['playersJoined'])/100 + 1)
data['headshotKills'] = data['headshotKills']*((100-data['playersJoined'])/100 + 1)
data['DBNOs'] = data['DBNOs']*((100-data['playersJoined'])/100 + 1)
data['kills'] = data['kills']*((100-data['playersJoined'])/100 + 1)
data['damageDealt'] = data['damageDealt']*((100-data['playersJoined'])/100 + 1)
data['matchDuration'] = data['matchDuration']*((100-data['playersJoined'])/100 + 1)
data['killStreaks'] = data['killStreaks']*((100-data['playersJoined'])/100 + 1)
data['roadKills'] = data['roadKills']*((100-data['playersJoined'])/100 + 1)
data['killPlace'] = data['killPlace']*((100-data['playersJoined'])/100 + 1)
```

- 将会受到总人数影响的变量规范化至100人局
- 计算公式经历前后修改

异常数据清洗：

- 有载具上击杀但没有载具移动距离： $\text{data}['\text{roadKills}'] > 0 \& (\text{data}['\text{rideDistance}'] == 0)$
- 有击杀但是没有移动： $(\text{data}['\text{kills}'] > 0) \& (\text{data}['\text{totalDistance}'] == 0)$
- 有爆头击杀但没有获取武器： $(\text{data}['\text{headshotKills}'] > 0) \& (\text{data}['\text{weaponsAcquired}'] == 0)$
- 击杀人数少于爆头和载具击杀总和： $\text{data}['\text{kills}'] < (\text{data}['\text{headshotKills}'] + \text{data}['\text{roadKills}'])$
- 最远杀敌距离大于等于1.2km（枪械最远攻击距离）
- 武器获取过多（大于80个）： $\text{data}['\text{weaponsAcquired}'] \leq 80$
- 队伍人数超过当前赛局规格最高人数

特征工程：

1. 爆头率： $\text{data}['\text{headshot_rate}'] = \text{data}['\text{headshotKills}'] / \text{data}['\text{kills}']$
2. 击杀率： $\text{data}['\text{kill_rate}'] = \text{data.kills} / (\text{data.kills} + \text{data.assists} + \text{data.DBNOs})$
3. 所在队伍人数（用于后续筛选）： $\text{data.groupby}('groupId')['groupId'].transform('count')$
4. 总移动距离： $\text{data}['\text{totalDistance}'] = \text{data}['\text{rideDistance}'] + \text{data}['\text{walkDistance}'] + \text{data}['\text{swimDistance}']$
5. 移动距离与weaponsAcquired、heals、kills分别相除
6. 防御性物资的获取总数： $\text{data}['\text{healthitems}'] = \text{data}['\text{heals}'] + \text{data}['\text{boosts}']$
7. 技巧击杀数量： $\text{data}['\text{skill}'] = \text{data}['\text{headshotKills}'] + \text{data}['\text{roadKills}'] + \text{data}['\text{killStreaks}']$

空值处理:

只分母为0: inf; 只分子为0: 0; 分子与分母同为0: NaN

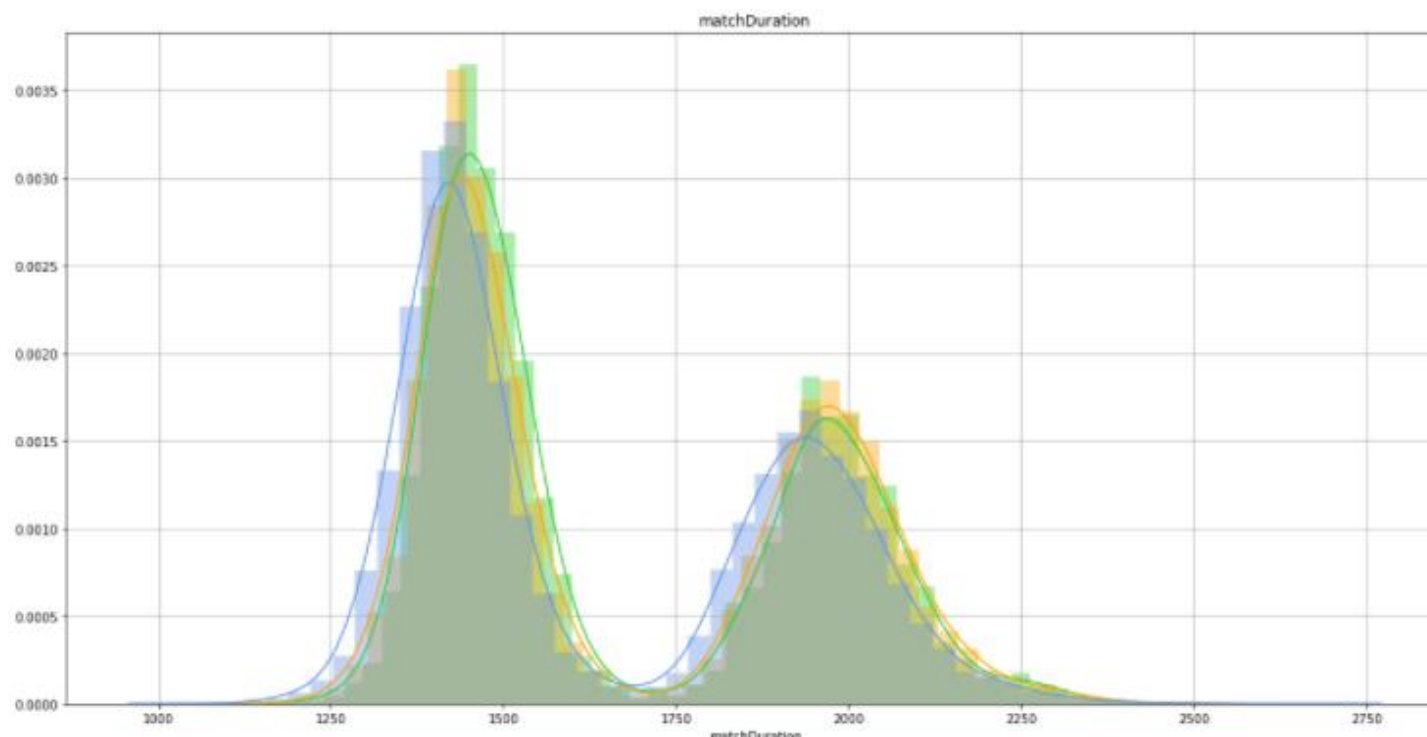
→统一填充0值

- NAN和INF值处理

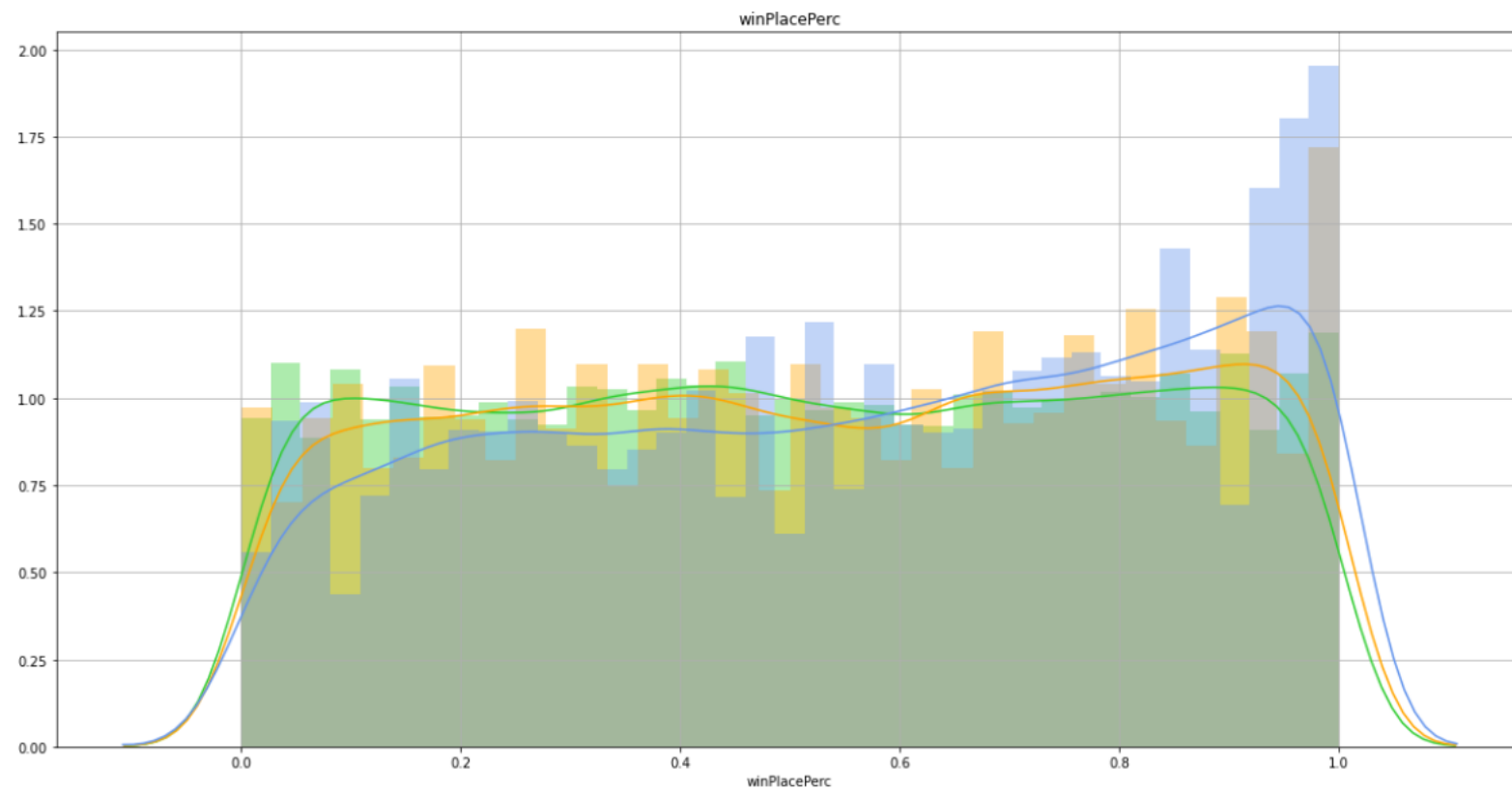
```
In [28]: import numpy as np
data[data == np.Inf] = np.NaN
data[data == np.NINF] = np.NaN
data.fillna(0, inplace=True)
```

特征剔除:

剔除matchDuration: 这一特征不属于玩家个人特征, 而与整局玩家人数、玩家水平相关, 无法反映玩家行为类型



目标列分布情况：



CONTENTS

ONE
1

TWO
2

FEATURE

CLICK

THREE
3

FOUR
4

SelectKBest:

- 用GBRT作为基础模型，选取对于分类最有作用的前K个特征

```
from sklearn.feature_selection import SelectKBest
from sklearn.metrics import mean_absolute_error
from sklearn.metrics import r2_score
from sklearn import ensemble
df = []
for i in range(27):
    SKB=SelectKBest(k=i+1)#选取对于分类最有作用的前k个特征，取PCA结果数量
    SKB.fit(x_train_ss,y_train)
    fselect_train=SKB.transform(x_train_ss)
    fselect_test=SKB.transform(x_test_ss)

    gbr_reg = ensemble.GradientBoostingRegressor()
    gbr_reg.fit(fselect_train,y_train)
    y_pr=gbr_reg.predict(fselect_test)
```

Out[58]:

	MAE	R方	column	columnNumbers
0	0.086525	0.850654	[False, False, False, False, False, False, Fal...	1
1	0.055972	0.924745	[False, False, False, False, False, True, Fals...	2
2	0.054867	0.928076	[False, True, False, False, False, True, False...	3
3	0.054858	0.927972	[False, True, False, False, False, True, False...	4
4	0.055364	0.927114	[False, True, False, False, False, True, False...	5
5	0.054854	0.928161	[False, True, False, False, False, True, False...	6
6	0.051105	0.939754	[False, True, False, False, False, True, True, ...	7
7	0.051176	0.939664	[False, True, True, False, False, True, True, ...	8
8	0.051310	0.939034	[False, True, True, False, False, True, True, ...	9
9	0.051450	0.938728	[False, True, True, False, False, True, True, ...	10
10	0.051450	0.938728	[False, True, True, False, False, True, True, ...	11
11	0.051066	0.939561	[False, True, True, False, False, True, True, ...	12
12	0.051067	0.939559	[False, True, True, False, False, True, True, ...	13
13	0.051236	0.939287	[False, True, True, False, False, True, True, ...	14
14	0.051242	0.939282	[False, True, True, False, False, True, True, ...	15
15	0.051242	0.939282	[False, True, True, True, False, True, True, T...	16
16	0.051242	0.939282	[False, True, True, True, True, True, True, Tr...	17
17	0.051375	0.939057	[False, True, True, True, True, True, True, Tr...	18
18	0.051375	0.939057	[False, True, True, True, True, True, True, Tr...	19
19	0.051382	0.938882	[False, True, True, True, True, True, True, Tr...	20
20	0.051382	0.938882	[True, True, True, True, True, True, True, Tru...	21
21	0.051382	0.938882	[True, True, True, True, True, True, True, Tru...	22
22	0.051636	0.938230	[True, True, True, True, True, True, True, Tru...	23
23	0.051425	0.938857	[True, True, True, True, True, True, True, Tru...	24
24	0.051532	0.938317	[True, True, True, True, True, True, True, Tru...	25
25	0.051491	0.938482	[True, True, True, True, True, True, True, Tru...	26
26	0.051455	0.938522	[True, True, True, True, True, True, True, Tru...	27



选择的特征:

	单排	双排	四排																																																																																																																																																																																																																																																
特征选取前	29	32	32																																																																																																																																																																																																																																																
特征选取后	13	20	25																																																																																																																																																																																																																																																
选择的特征	<table><tr><th>#</th><th>Column</th><th>Non-Null Count</th><th>Dtype</th></tr><tr><td>0</td><td>boosts</td><td>50000 non-null</td><td>int64</td></tr><tr><td>1</td><td>damageDealt</td><td>50000 non-null</td><td>float64</td></tr><tr><td>2</td><td>killPlace</td><td>50000 non-null</td><td>float64</td></tr><tr><td>3</td><td>kills</td><td>50000 non-null</td><td>float64</td></tr><tr><td>4</td><td>longestKill</td><td>50000 non-null</td><td>float64</td></tr><tr><td>5</td><td>walkDistance</td><td>50000 non-null</td><td>float64</td></tr><tr><td>6</td><td>weaponsAcquired</td><td>50000 non-null</td><td>int64</td></tr><tr><td>7</td><td>totalDistance</td><td>50000 non-null</td><td>float64</td></tr><tr><td>8</td><td>distance_over_weapons</td><td>50000 non-null</td><td>float64</td></tr><tr><td>9</td><td>distance_over_kills</td><td>50000 non-null</td><td>float64</td></tr><tr><td>10</td><td>healthitems</td><td>50000 non-null</td><td>int64</td></tr><tr><td>11</td><td>skill</td><td>50000 non-null</td><td>float64</td></tr><tr><td>12</td><td>winPlacePerc</td><td>50000 non-null</td><td>float64</td></tr></table>	#	Column	Non-Null Count	Dtype	0	boosts	50000 non-null	int64	1	damageDealt	50000 non-null	float64	2	killPlace	50000 non-null	float64	3	kills	50000 non-null	float64	4	longestKill	50000 non-null	float64	5	walkDistance	50000 non-null	float64	6	weaponsAcquired	50000 non-null	int64	7	totalDistance	50000 non-null	float64	8	distance_over_weapons	50000 non-null	float64	9	distance_over_kills	50000 non-null	float64	10	healthitems	50000 non-null	int64	11	skill	50000 non-null	float64	12	winPlacePerc	50000 non-null	float64	<table><tr><td>1</td><td>assists</td><td>50000 non-null</td><td>float64</td></tr><tr><td>2</td><td>boosts</td><td>50000 non-null</td><td>int64</td></tr><tr><td>3</td><td>damageDealt</td><td>50000 non-null</td><td>float64</td></tr><tr><td>4</td><td>DBNOs</td><td>50000 non-null</td><td>float64</td></tr><tr><td>5</td><td>heals</td><td>50000 non-null</td><td>int64</td></tr><tr><td>6</td><td>killPlace</td><td>50000 non-null</td><td>float64</td></tr><tr><td>7</td><td>kills</td><td>50000 non-null</td><td>float64</td></tr><tr><td>8</td><td>killStreaks</td><td>50000 non-null</td><td>float64</td></tr><tr><td>9</td><td>longestKill</td><td>50000 non-null</td><td>float64</td></tr><tr><td>10</td><td>rideDistance</td><td>50000 non-null</td><td>float64</td></tr><tr><td>11</td><td>walkDistance</td><td>50000 non-null</td><td>float64</td></tr><tr><td>12</td><td>weaponsAcquired</td><td>50000 non-null</td><td>int64</td></tr><tr><td>13</td><td>winPlacePerc</td><td>50000 non-null</td><td>float64</td></tr><tr><td>14</td><td>totalDistance</td><td>50000 non-null</td><td>float64</td></tr><tr><td>15</td><td>kill_rate</td><td>50000 non-null</td><td>float64</td></tr><tr><td>16</td><td>distance_over_weapons</td><td>50000 non-null</td><td>float64</td></tr><tr><td>17</td><td>distance_over_heals</td><td>50000 non-null</td><td>float64</td></tr><tr><td>18</td><td>distance_over_kills</td><td>50000 non-null</td><td>float64</td></tr><tr><td>19</td><td>healthitems</td><td>50000 non-null</td><td>int64</td></tr><tr><td>20</td><td>skill</td><td>50000 non-null</td><td>float64</td></tr></table>	1	assists	50000 non-null	float64	2	boosts	50000 non-null	int64	3	damageDealt	50000 non-null	float64	4	DBNOs	50000 non-null	float64	5	heals	50000 non-null	int64	6	killPlace	50000 non-null	float64	7	kills	50000 non-null	float64	8	killStreaks	50000 non-null	float64	9	longestKill	50000 non-null	float64	10	rideDistance	50000 non-null	float64	11	walkDistance	50000 non-null	float64	12	weaponsAcquired	50000 non-null	int64	13	winPlacePerc	50000 non-null	float64	14	totalDistance	50000 non-null	float64	15	kill_rate	50000 non-null	float64	16	distance_over_weapons	50000 non-null	float64	17	distance_over_heals	50000 non-null	float64	18	distance_over_kills	50000 non-null	float64	19	healthitems	50000 non-null	int64	20	skill	50000 non-null	float64	<table><tr><th>#</th><th>Column</th><th>Non-Null Count</th><th>Dtype</th></tr><tr><td>0</td><td>assists</td><td>50000 non-null</td><td>float64</td></tr><tr><td>1</td><td>boosts</td><td>50000 non-null</td><td>int64</td></tr><tr><td>2</td><td>damageDealt</td><td>50000 non-null</td><td>float64</td></tr><tr><td>3</td><td>DBNOs</td><td>50000 non-null</td><td>float64</td></tr><tr><td>4</td><td>headshotKills</td><td>50000 non-null</td><td>float64</td></tr><tr><td>5</td><td>heals</td><td>50000 non-null</td><td>int64</td></tr><tr><td>6</td><td>killPlace</td><td>50000 non-null</td><td>float64</td></tr><tr><td>7</td><td>kills</td><td>50000 non-null</td><td>float64</td></tr><tr><td>8</td><td>killStreaks</td><td>50000 non-null</td><td>float64</td></tr><tr><td>9</td><td>longestKill</td><td>50000 non-null</td><td>float64</td></tr><tr><td>10</td><td>revives</td><td>50000 non-null</td><td>int64</td></tr><tr><td>11</td><td>rideDistance</td><td>50000 non-null</td><td>float64</td></tr><tr><td>12</td><td>walkDistance</td><td>50000 non-null</td><td>float64</td></tr><tr><td>13</td><td>weaponsAcquired</td><td>50000 non-null</td><td>int64</td></tr><tr><td>14</td><td>totalDistance</td><td>50000 non-null</td><td>float64</td></tr><tr><td>15</td><td>headshot_rate</td><td>50000 non-null</td><td>float64</td></tr><tr><td>16</td><td>kill_rate</td><td>50000 non-null</td><td>float64</td></tr><tr><td>17</td><td>playersInGroup</td><td>50000 non-null</td><td>int64</td></tr><tr><td>18</td><td>killsPerDistance</td><td>50000 non-null</td><td>float64</td></tr><tr><td>19</td><td>distance_over_weapons</td><td>50000 non-null</td><td>float64</td></tr><tr><td>20</td><td>distance_over_heals</td><td>50000 non-null</td><td>float64</td></tr><tr><td>21</td><td>distance_over_kills</td><td>50000 non-null</td><td>float64</td></tr><tr><td>22</td><td>healthitems</td><td>50000 non-null</td><td>int64</td></tr><tr><td>23</td><td>skill</td><td>50000 non-null</td><td>float64</td></tr><tr><td>24</td><td>winPlacePerc</td><td>50000 non-null</td><td>float64</td></tr></table>	#	Column	Non-Null Count	Dtype	0	assists	50000 non-null	float64	1	boosts	50000 non-null	int64	2	damageDealt	50000 non-null	float64	3	DBNOs	50000 non-null	float64	4	headshotKills	50000 non-null	float64	5	heals	50000 non-null	int64	6	killPlace	50000 non-null	float64	7	kills	50000 non-null	float64	8	killStreaks	50000 non-null	float64	9	longestKill	50000 non-null	float64	10	revives	50000 non-null	int64	11	rideDistance	50000 non-null	float64	12	walkDistance	50000 non-null	float64	13	weaponsAcquired	50000 non-null	int64	14	totalDistance	50000 non-null	float64	15	headshot_rate	50000 non-null	float64	16	kill_rate	50000 non-null	float64	17	playersInGroup	50000 non-null	int64	18	killsPerDistance	50000 non-null	float64	19	distance_over_weapons	50000 non-null	float64	20	distance_over_heals	50000 non-null	float64	21	distance_over_kills	50000 non-null	float64	22	healthitems	50000 non-null	int64	23	skill	50000 non-null	float64	24	winPlacePerc	50000 non-null	float64
#	Column	Non-Null Count	Dtype																																																																																																																																																																																																																																																
0	boosts	50000 non-null	int64																																																																																																																																																																																																																																																
1	damageDealt	50000 non-null	float64																																																																																																																																																																																																																																																
2	killPlace	50000 non-null	float64																																																																																																																																																																																																																																																
3	kills	50000 non-null	float64																																																																																																																																																																																																																																																
4	longestKill	50000 non-null	float64																																																																																																																																																																																																																																																
5	walkDistance	50000 non-null	float64																																																																																																																																																																																																																																																
6	weaponsAcquired	50000 non-null	int64																																																																																																																																																																																																																																																
7	totalDistance	50000 non-null	float64																																																																																																																																																																																																																																																
8	distance_over_weapons	50000 non-null	float64																																																																																																																																																																																																																																																
9	distance_over_kills	50000 non-null	float64																																																																																																																																																																																																																																																
10	healthitems	50000 non-null	int64																																																																																																																																																																																																																																																
11	skill	50000 non-null	float64																																																																																																																																																																																																																																																
12	winPlacePerc	50000 non-null	float64																																																																																																																																																																																																																																																
1	assists	50000 non-null	float64																																																																																																																																																																																																																																																
2	boosts	50000 non-null	int64																																																																																																																																																																																																																																																
3	damageDealt	50000 non-null	float64																																																																																																																																																																																																																																																
4	DBNOs	50000 non-null	float64																																																																																																																																																																																																																																																
5	heals	50000 non-null	int64																																																																																																																																																																																																																																																
6	killPlace	50000 non-null	float64																																																																																																																																																																																																																																																
7	kills	50000 non-null	float64																																																																																																																																																																																																																																																
8	killStreaks	50000 non-null	float64																																																																																																																																																																																																																																																
9	longestKill	50000 non-null	float64																																																																																																																																																																																																																																																
10	rideDistance	50000 non-null	float64																																																																																																																																																																																																																																																
11	walkDistance	50000 non-null	float64																																																																																																																																																																																																																																																
12	weaponsAcquired	50000 non-null	int64																																																																																																																																																																																																																																																
13	winPlacePerc	50000 non-null	float64																																																																																																																																																																																																																																																
14	totalDistance	50000 non-null	float64																																																																																																																																																																																																																																																
15	kill_rate	50000 non-null	float64																																																																																																																																																																																																																																																
16	distance_over_weapons	50000 non-null	float64																																																																																																																																																																																																																																																
17	distance_over_heals	50000 non-null	float64																																																																																																																																																																																																																																																
18	distance_over_kills	50000 non-null	float64																																																																																																																																																																																																																																																
19	healthitems	50000 non-null	int64																																																																																																																																																																																																																																																
20	skill	50000 non-null	float64																																																																																																																																																																																																																																																
#	Column	Non-Null Count	Dtype																																																																																																																																																																																																																																																
0	assists	50000 non-null	float64																																																																																																																																																																																																																																																
1	boosts	50000 non-null	int64																																																																																																																																																																																																																																																
2	damageDealt	50000 non-null	float64																																																																																																																																																																																																																																																
3	DBNOs	50000 non-null	float64																																																																																																																																																																																																																																																
4	headshotKills	50000 non-null	float64																																																																																																																																																																																																																																																
5	heals	50000 non-null	int64																																																																																																																																																																																																																																																
6	killPlace	50000 non-null	float64																																																																																																																																																																																																																																																
7	kills	50000 non-null	float64																																																																																																																																																																																																																																																
8	killStreaks	50000 non-null	float64																																																																																																																																																																																																																																																
9	longestKill	50000 non-null	float64																																																																																																																																																																																																																																																
10	revives	50000 non-null	int64																																																																																																																																																																																																																																																
11	rideDistance	50000 non-null	float64																																																																																																																																																																																																																																																
12	walkDistance	50000 non-null	float64																																																																																																																																																																																																																																																
13	weaponsAcquired	50000 non-null	int64																																																																																																																																																																																																																																																
14	totalDistance	50000 non-null	float64																																																																																																																																																																																																																																																
15	headshot_rate	50000 non-null	float64																																																																																																																																																																																																																																																
16	kill_rate	50000 non-null	float64																																																																																																																																																																																																																																																
17	playersInGroup	50000 non-null	int64																																																																																																																																																																																																																																																
18	killsPerDistance	50000 non-null	float64																																																																																																																																																																																																																																																
19	distance_over_weapons	50000 non-null	float64																																																																																																																																																																																																																																																
20	distance_over_heals	50000 non-null	float64																																																																																																																																																																																																																																																
21	distance_over_kills	50000 non-null	float64																																																																																																																																																																																																																																																
22	healthitems	50000 non-null	int64																																																																																																																																																																																																																																																
23	skill	50000 non-null	float64																																																																																																																																																																																																																																																
24	winPlacePerc	50000 non-null	float64																																																																																																																																																																																																																																																

CONTENTS



ONE
1



TWO
2



THREE
3

CLUSTER

CLICK



FOUR
4



聚类方法与模型选择：

	单排		双排		四排	
	轮廓系数	簇类数	轮廓系数	簇类数	轮廓系数	簇类数
Kmeans	0.33	6	0.26	6	0.18	6
AGG	0.29	5	0.26	7	0.16	6
DBSCAN	0.50	6	0.55	5	0.48	5
特征选择后 DBSCAN	0.57	5	0.41	4	0.30	5

- 单排数据的轮廓系数表现更优，相比双排和四排更适合运用聚类

DBSCAN调参:

- 遍历调参
- 计算速度慢
- 会尽量选择簇类数大于等于4，轮廓系数大于0.5或接近0.5的参数组合。

```
from sklearn.cluster import DBSCAN
from sklearn.metrics import silhouette_score
res = []
# 迭代不同的eps值
for eps in np.arange(2.1, 2.5, 0.1):
    # 迭代不同的min_samples值
    for min_samples in np.arange(4, 7, 1):
        dbscan = DBSCAN(eps = eps, min_samples = min_samples)
        # 模型拟合
        db_pre = dbscan.fit_predict(data_input)
        # 统计轮廓系数
        k = silhouette_score(data_input, db_pre)
        clu_num = len(np.unique(db_pre))
        res.append({'eps':eps, 'min_samples':min_samples, 'score':k, 'clu_num':clu_num})
# 将迭代后的结果存储到数据框中
df = pd.DataFrame(res)

df
```

	eps	min_samples	score	clu_num
0	2.1	4	0.508548	8
1	2.1	5	0.592410	4
2	2.1	6	0.602268	3
3	2.2	4	0.578130	5
4	2.2	5	0.632466	3
5	2.2	6	0.594182	3
6	2.3	4	0.592160	6
7	2.3	5	0.645134	3
8	2.3	6	0.648510	2
9	2.4	4	0.583985	8
10	2.4	5	0.654038	3
11	2.4	6	0.657835	2

聚类效果:

单排

双排

四排

特征选择前

```
In [12]: data['dbscan'].value_counts()
```

```
Out[12]: 0    48961  
         1      530  
         2      218  
         3      154  
        -1       82  
         4       55  
         Name: dbscan, dtype: int64
```

```
In [11]: data['dbscan'].value_counts()
```

```
Out[11]: 0    49477  
         2     373  
         1      89  
        -1     50  
         3      11  
         Name: dbscan, dtype: int64
```

```
In [13]: data['dbscan'].value_counts()
```

```
Out[13]: 0    49335  
         1     489  
         2     116  
        -1      42  
         3       18  
         Name: dbscan, dtype: int64
```

特征选择后

```
In [42]: data['dbscan'].value_counts()
```

```
Out[42]: 0    48961  
         1      530  
         2      218  
         3      154  
        -1       82  
         4       55  
         Name: dbscan, dtype: int64
```

```
In [23]: data['dbscan'].value_counts()
```

```
Out[23]: 0    48797  
        -1    1196  
         1         4  
         2         3  
         Name: dbscan, dtype: int64
```

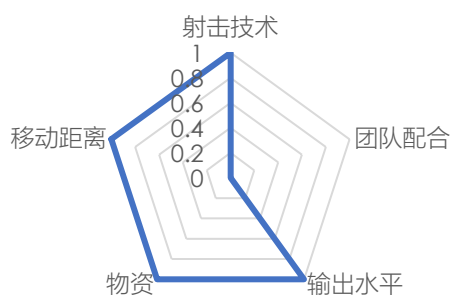
```
In [26]: data['dbscan'].value_counts()
```

```
Out[26]: 0    47707  
        -1    2279  
         2         8  
         3         4  
         1         2  
         Name: dbscan, dtype: int64
```

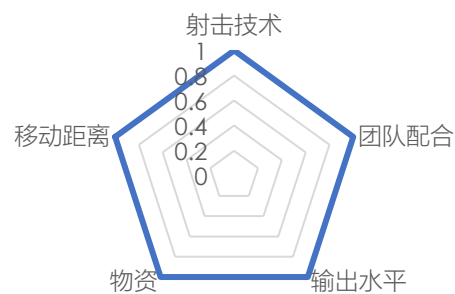
- -1为outliner, 不进行聚类中心的分析
- 各类数量分布严重不均, 难以聚成明显的几个策略类
- 特征选择后的聚类并没有优化

单排\双排\四排的分类0玩家:

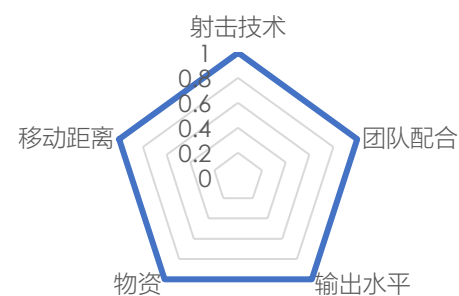
类别0 排名0.5



类别0 排名0.52



类别0 排名0.55

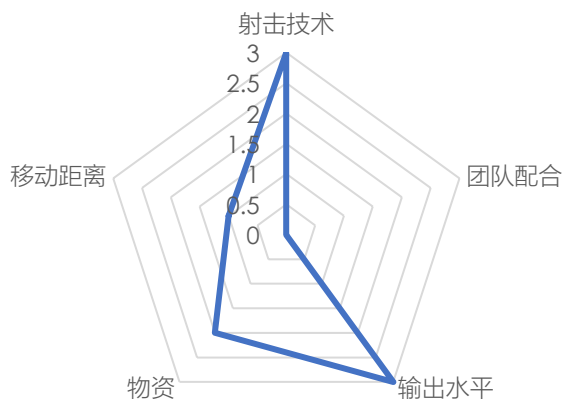


类别0玩家单位距离内击杀敌人水平最高，单位距离内获得武器、增益道具数量最多，总体移动距离最少、物资获取水平最少，这类玩家很可能是选择了大家熟知的资源密集区降落，但在激烈的厮杀争夺中落败，中早期就成盒了。这类玩家数量最多，说明相当大部分玩家的策略会选择这种策略。

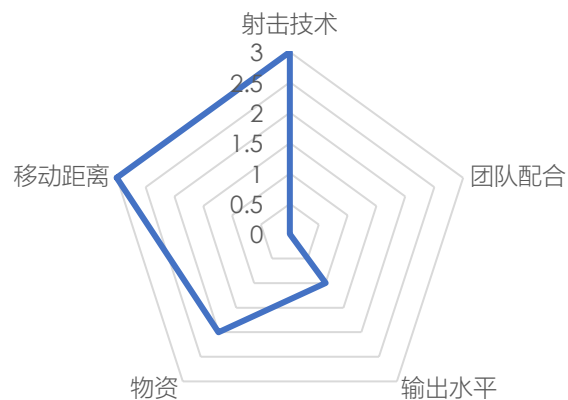
P.S.雷达图为相对指标，类别为5时有5个等级，类别为4时有4个等级，1为最低级

单排:

类别1 排名0.55



类别4 排名0.61



类别1玩家和类别4玩家都存在击杀队友行为，但单排没有队友，即自杀了，但存在区别，类别1玩家各项指标、移动距离和最终排名均劣于类别4玩家，因此类别1玩家是在游戏前期就自杀了，可能属于不熟悉操作的玩家，类别4玩家属于游戏中后期自杀的玩家，可能是误操作导致的。

单排:



类别2玩家

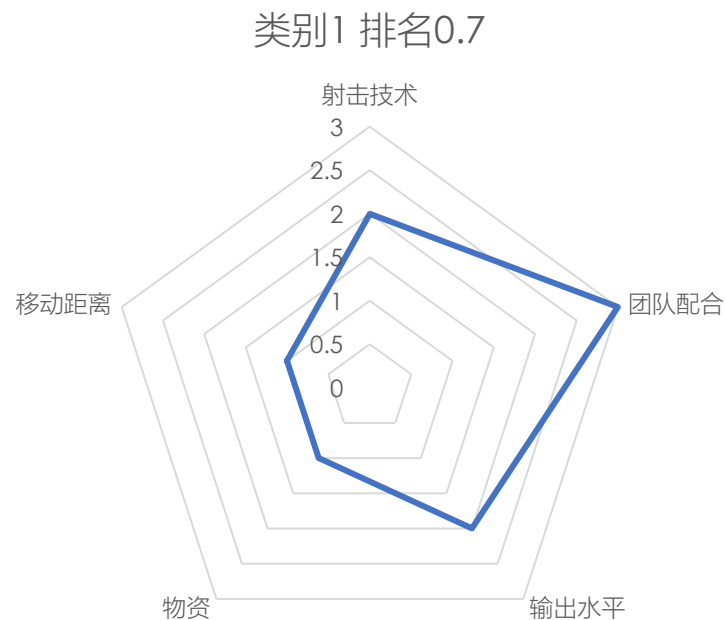
类别2玩家的连续击杀能力最强，击杀排名最靠前，借助载具移动距离最多，还会利用载具击杀等技巧性击杀方式，其总的移动距离也最多，属于在地图上移动范围较广的玩家，但其物资获取能力并不突出，因此最终的排名虽然属于较高水平0.74，但并不是最可能吃鸡的那类人。

单排：



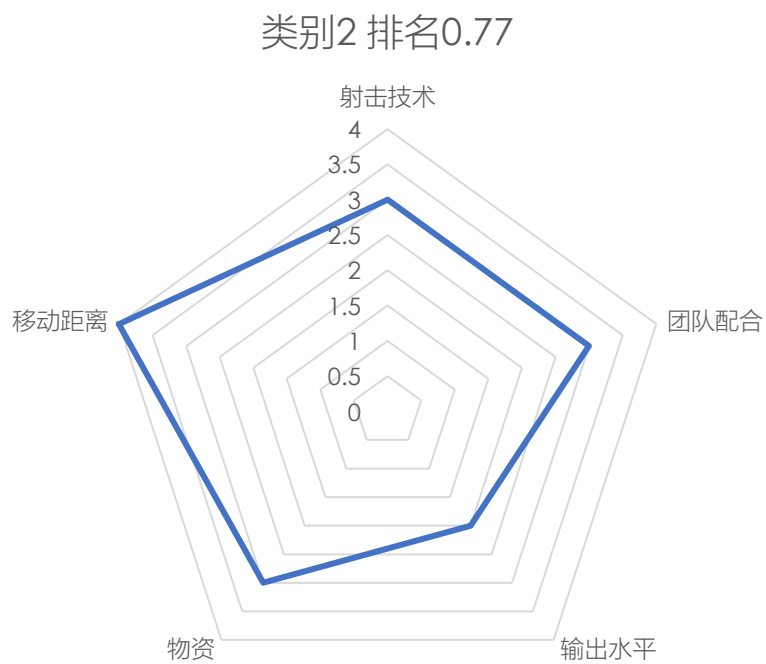
类别3玩家在爆头击杀能力、击杀量、伤害输出较多，同时在武器和增益道具的获取总量最多，同时每单位距离内获得heals的能力最高，说明综合搜集物资能力较强。这类玩家虽然总移动距离不是最多，但其各种移动方式的距离数都较高，说明这类玩家的移动方式较灵活，同时借助载具移动的距离较高，是有利用载具的玩家。总的来说，各方面综合素质较强的玩家才最接近吃鸡水平。

双排：



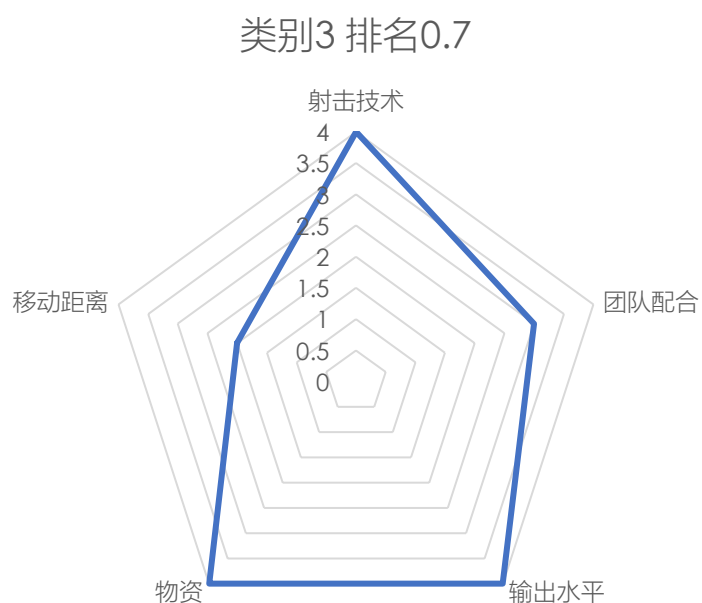
类别1玩家击杀敌人最远距离最大，同时击倒敌人但没杀死敌人最多，也表明这类玩家多采用远距离狙击作战方式，其载具移动距离远高于步行移动距离，同时救援队友次数最多，队友数接近2，说明多数是和队友一起乘载具移动，能够及时援救队友，这类玩家最终排名水平也较高，0.70。

双排:



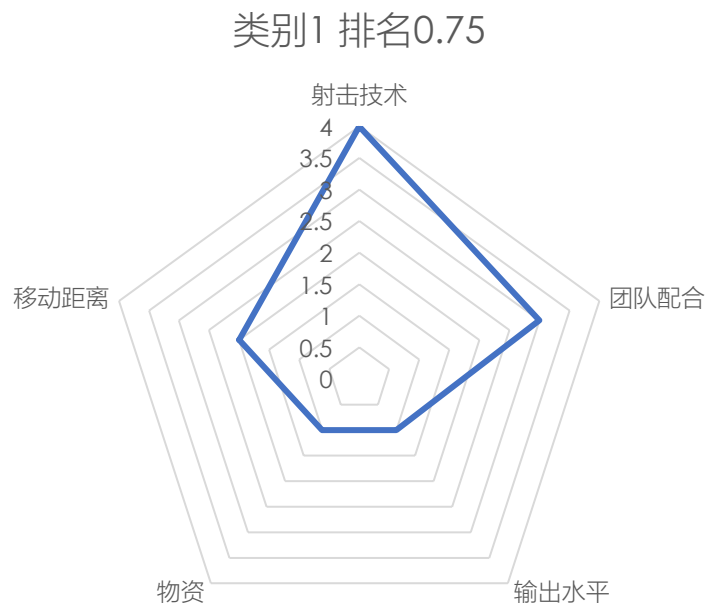
类别2玩家使用boosts和武器获得数最多，使其爆头率最高。类别2玩家步行距离最远，载具移动距离也较多，同时有破坏载具的行为，有可能是玩家使用完载具就摧毁不让别人有机会使用或是破坏他人载具的策略。这类玩家最终排名也较高，0.77。

双排:



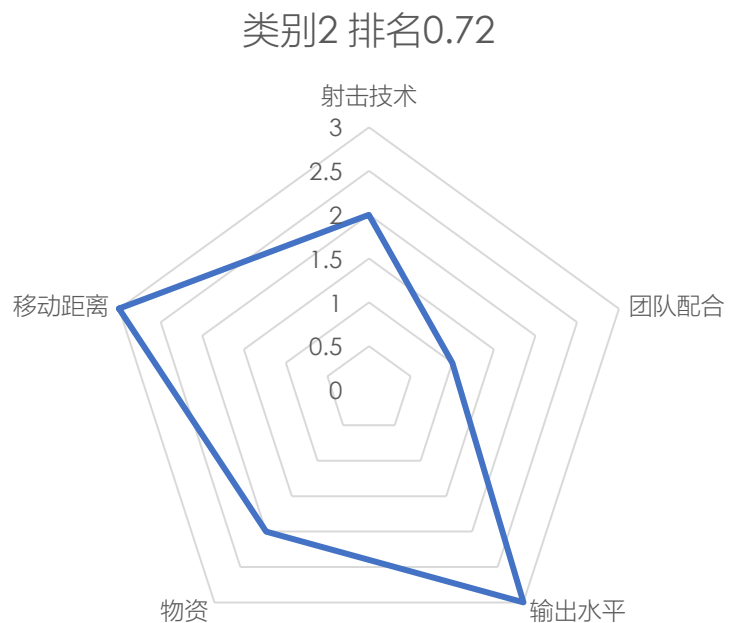
类别3玩家击杀水平最高，输出伤害也最多，增益道具使用最多，载具移动距离远大于步行距离，同时载具击杀最多，是善于利用道具的类型。这类玩家都有队友，团队合作使其综合表现更佳。这类玩家最终排名也较高，0.71。不高于类别2玩家，很可能因为样本问题，聚出的这类玩家数量只有11个，远低于类别2玩家数量。

四排：



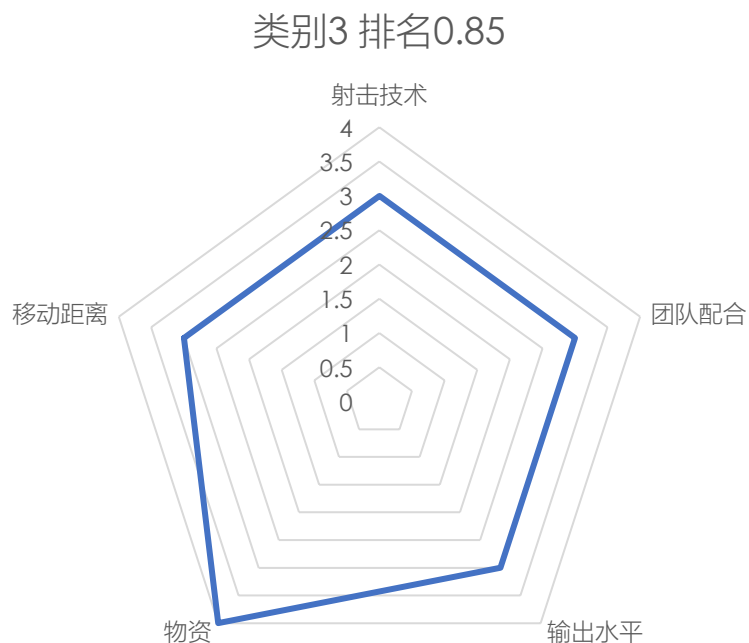
类别1玩家爆头击杀水平最高，但使用boosts数量属于中等，是本身射击技术较高的玩家。这类玩家团队配合水平很高，伤害输出不是很多，但击杀效率高，这类玩家最终排名位于0.76左右。

四排：



类别2玩家击杀率最高，助攻数相对较少，即这类玩家是队内的主要负责击杀的角色，被抢“人头”的概率较低。同时，这类玩家使用载具移动最多，也是最会利用载具辅助击杀的玩家，这类玩家最终排名位于0.73左右。

四排:




类别3玩家队友人数最多，绝大多数在四人队内，击杀排名最高，伤害输出最多，远距离击杀和连杀能力最强，物资获取数量也最多，其载具移动和步行移动距离都很多，游泳距离几乎没有，属于团队作战配合水平相当高的类型，团队配合使综合表现更佳。这类也是最接近吃鸡的玩家类别，0.86。侧面反映团队很少采用游泳的移动方式。这类玩家同时有最多的载具损毁水平，高于双排模式玩家水平，考虑因为四排玩家集体行动往往借助载具，攻击载具更可能达到让载具上敌队团灭效果，或削弱敌队使用载具可能性。

CONTENTS



ONE



TWO



THREE



FOUR

PREDICT

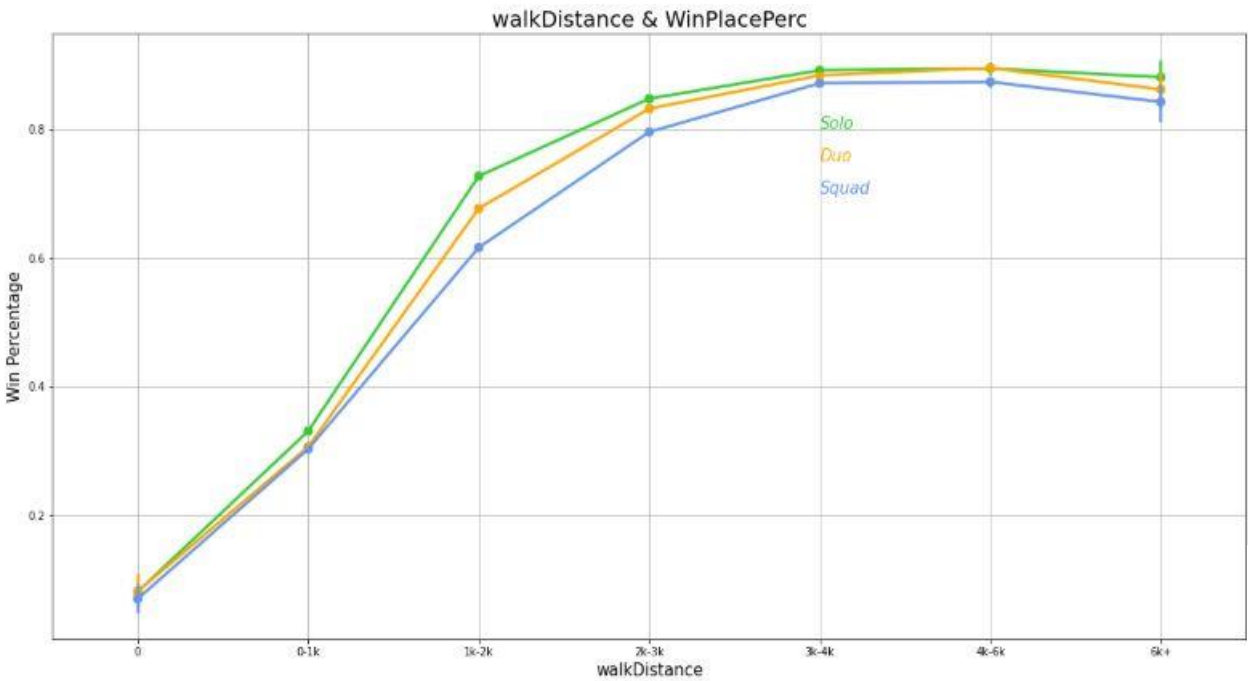
CLICK

全特征预测

模型		单排 MAE	单排 R ²	双排 MAE	双排 R ²	四排 MAE	四排 R ²
神经网络	聚类前	0.0491	—	0.0573	—	0.0754	—
线性回归	聚类前	0.0640	0.9057	0.0745	0.8790	0.0980	0.8040
决策树	聚类前	0.0682	0.8906	0.0785	0.8627	0.1038	0.7731
	调参后	0.0534	0.9312	0.0612	0.9160	0.8057	0.8653
KNN	聚类前	0.0610	0.9134	0.0717	0.8884	0.0965	0.8091
GBRT	聚类前	0.0515	0.9385	0.0574	0.9269	0.0768	0.8792
Lasso	聚类前	0.0654	0.9024	0.0761	0.8756	0.0987	0.8017

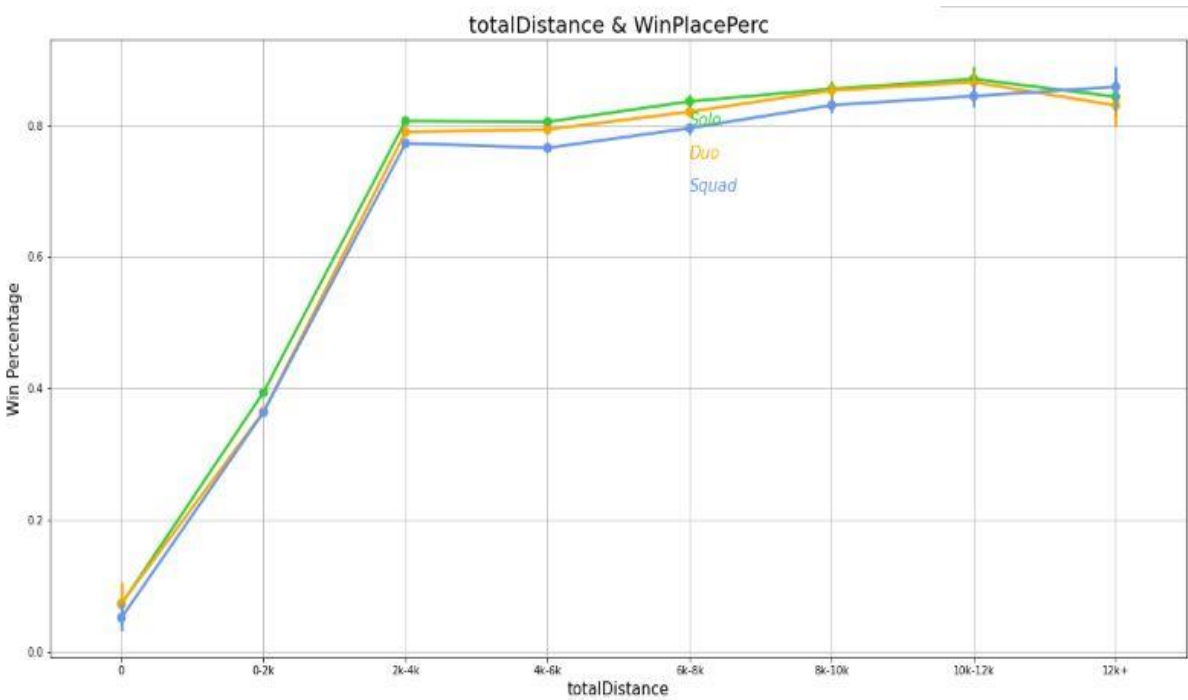


单排	双排	四排
walkDistance, 0. 7548197852803508	walkDistance, 0. 7206809587704923	walkDistance, 0. 6300887999218612
killPlace, 0. 16546545054626557	killPlace, 0. 17269912234573978	killPlace, 0. 23854085451219054
totalDistance, 0. 03938223960936025	totalDistance, 0. 05165553899265136	totalDistance, 0. 04516326959702657
boosts, 0. 017936822929918352	killsPerDistance, 0. 02531516079917443	killsPerDistance, 0. 0383961548525234
killsPerDistance, 0. 01059216986064035	boosts, 0. 0133572178842733	distance_over_kills, 0. 011202708700145777
healthitems, 0. 0025056016702223517	healthitems, 0. 004054505071639147	boosts, 0. 010071662360096519
kill_rate, 0. 0022667918622579315	kill_rate, 0. 003852820253108385	kill_rate, 0. 008672847570570607
kills, 0. 00187034864603156	kills, 0. 0024926977261432543	kills, 0. 008586370009037196
distance_over_kills, 0. 00141052175104514	longestKill, 0. 001902109941769581	assists, 0. 0029498182256406016
skill, 0. 0010034062477565036	assists, 0. 0008894443422498914	healthitems, 0. 0016532970527482383



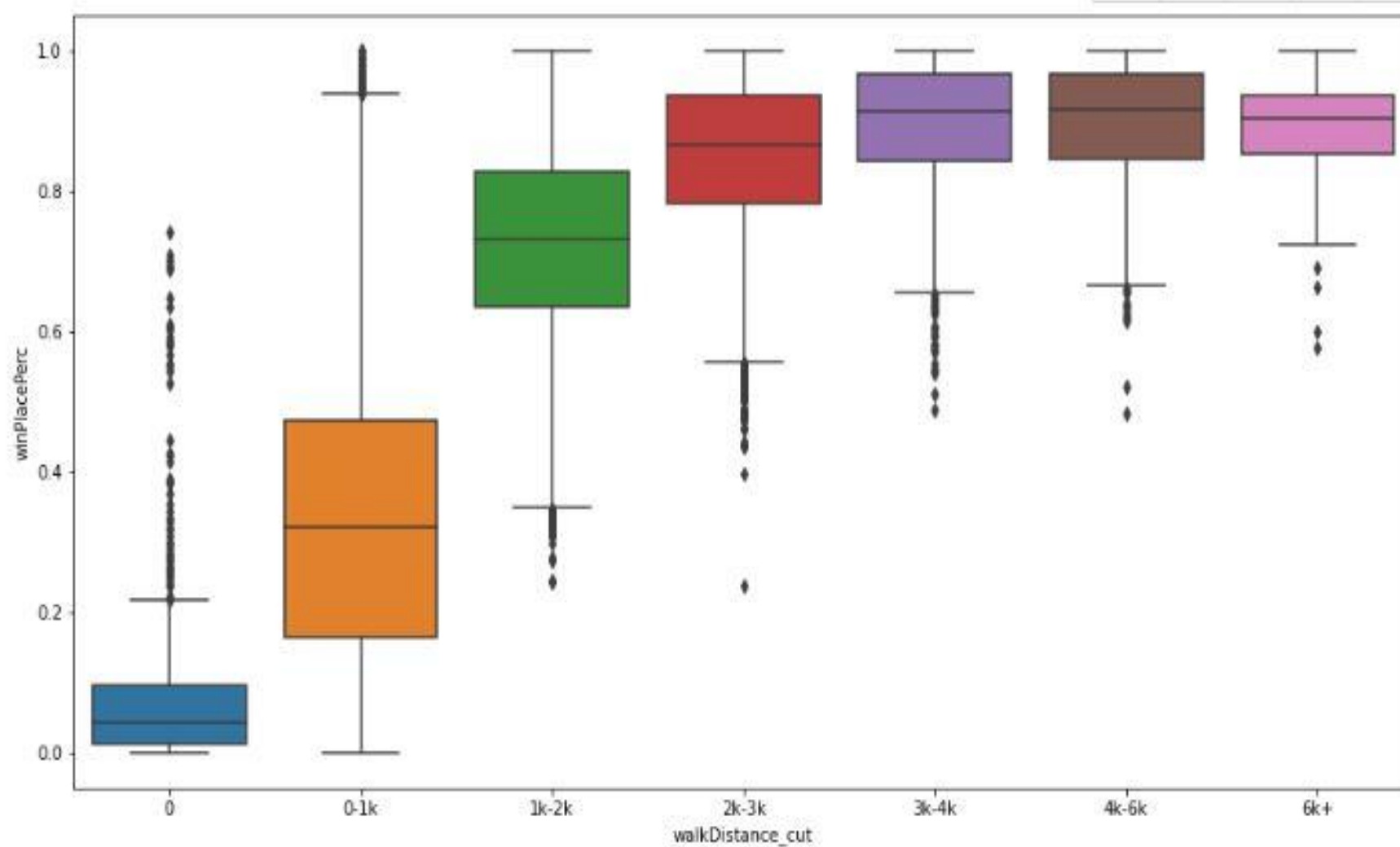
WalkDistance

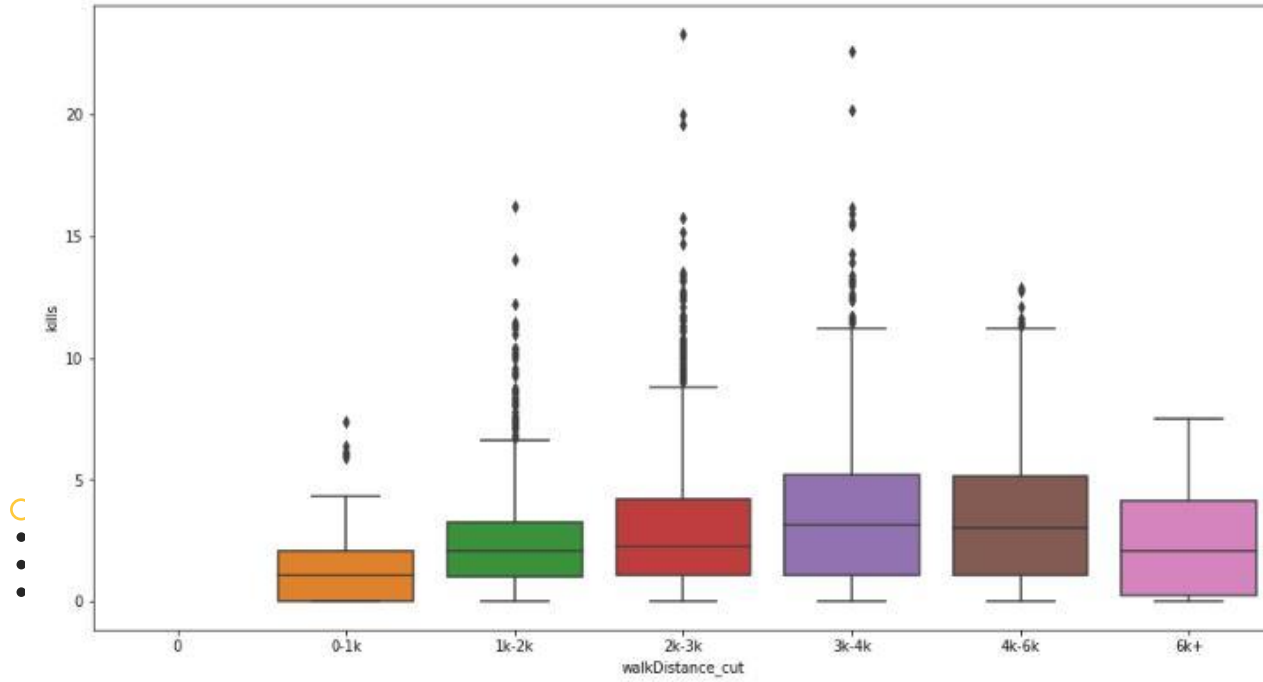
totalDistance





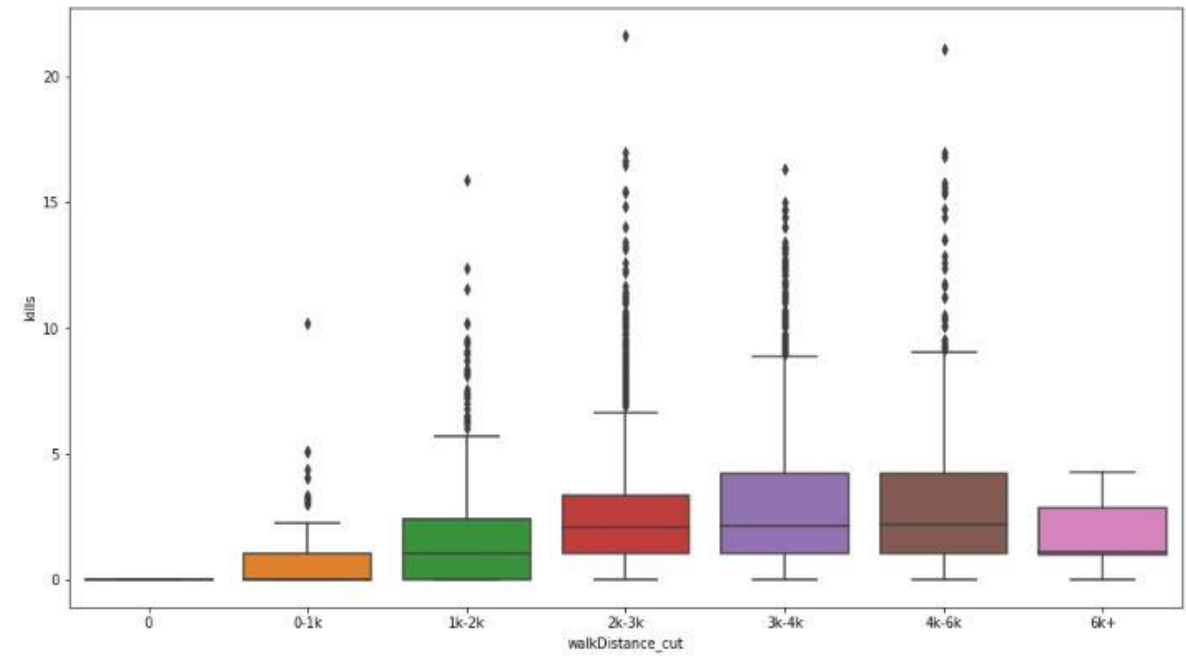
walkDistance—Solo



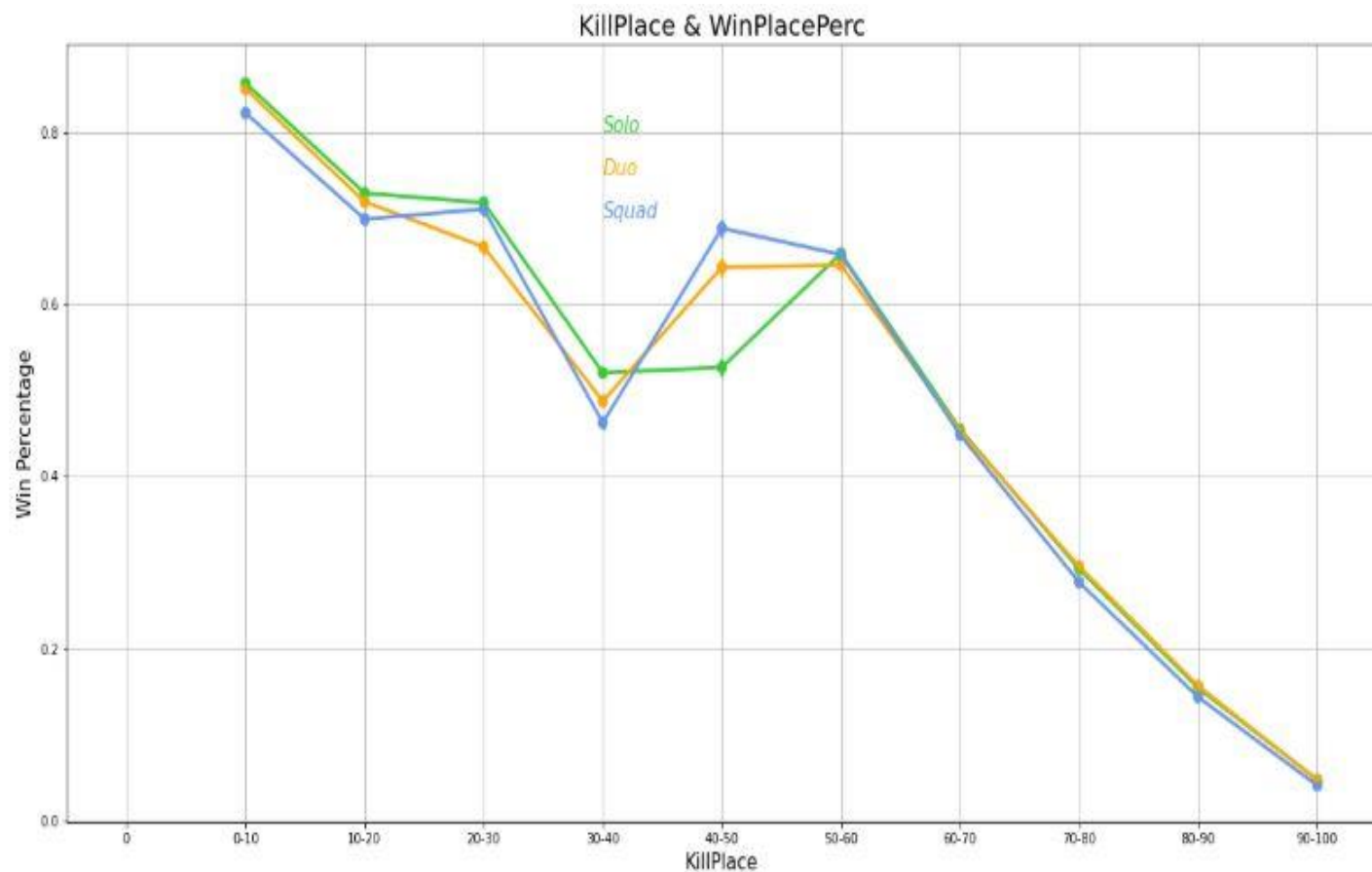


Solo

Squad



KillPlace、kills

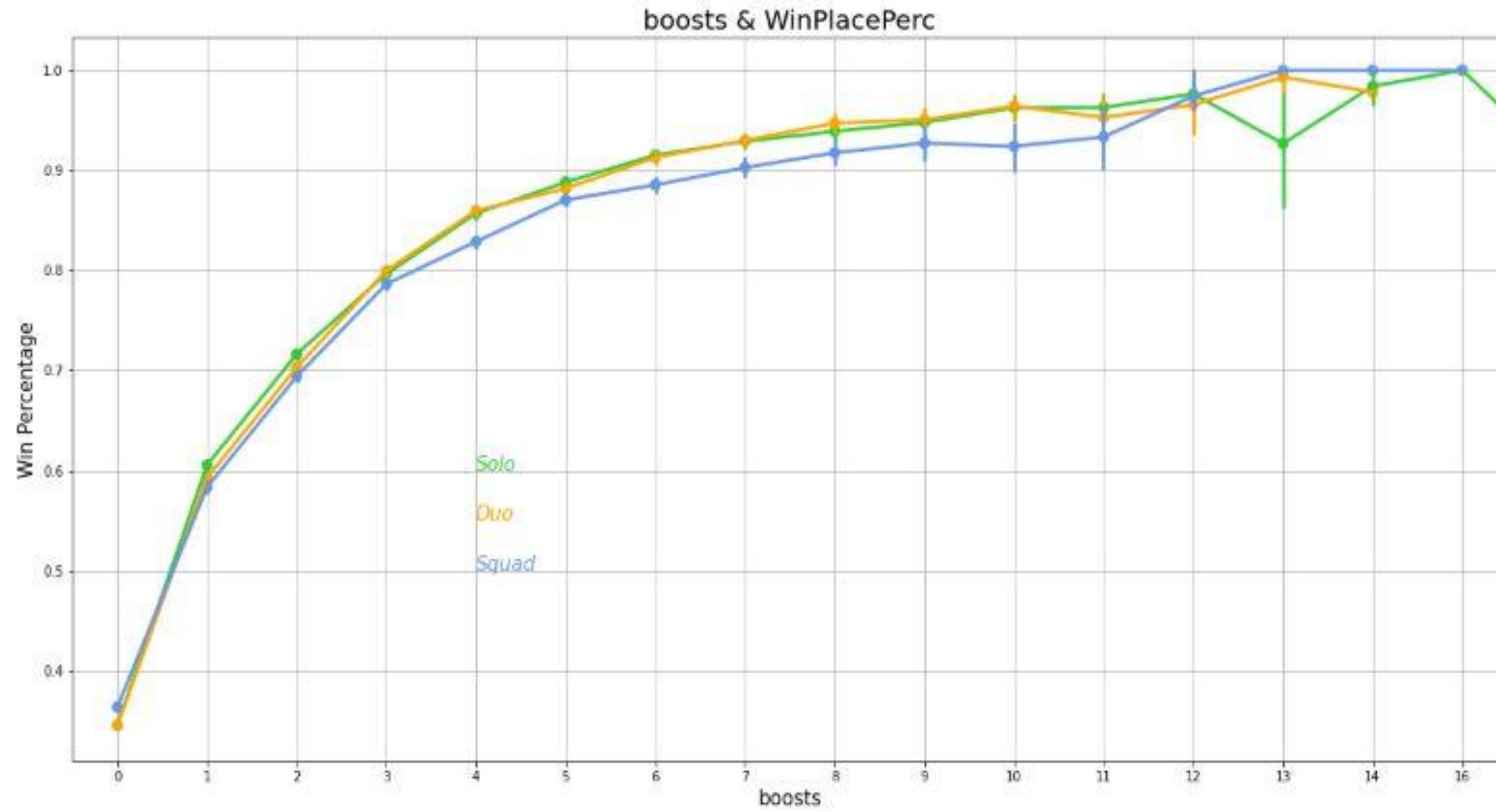


单排中1 个玩家 (0.00%) 不用杀一个人就可以赢得比赛

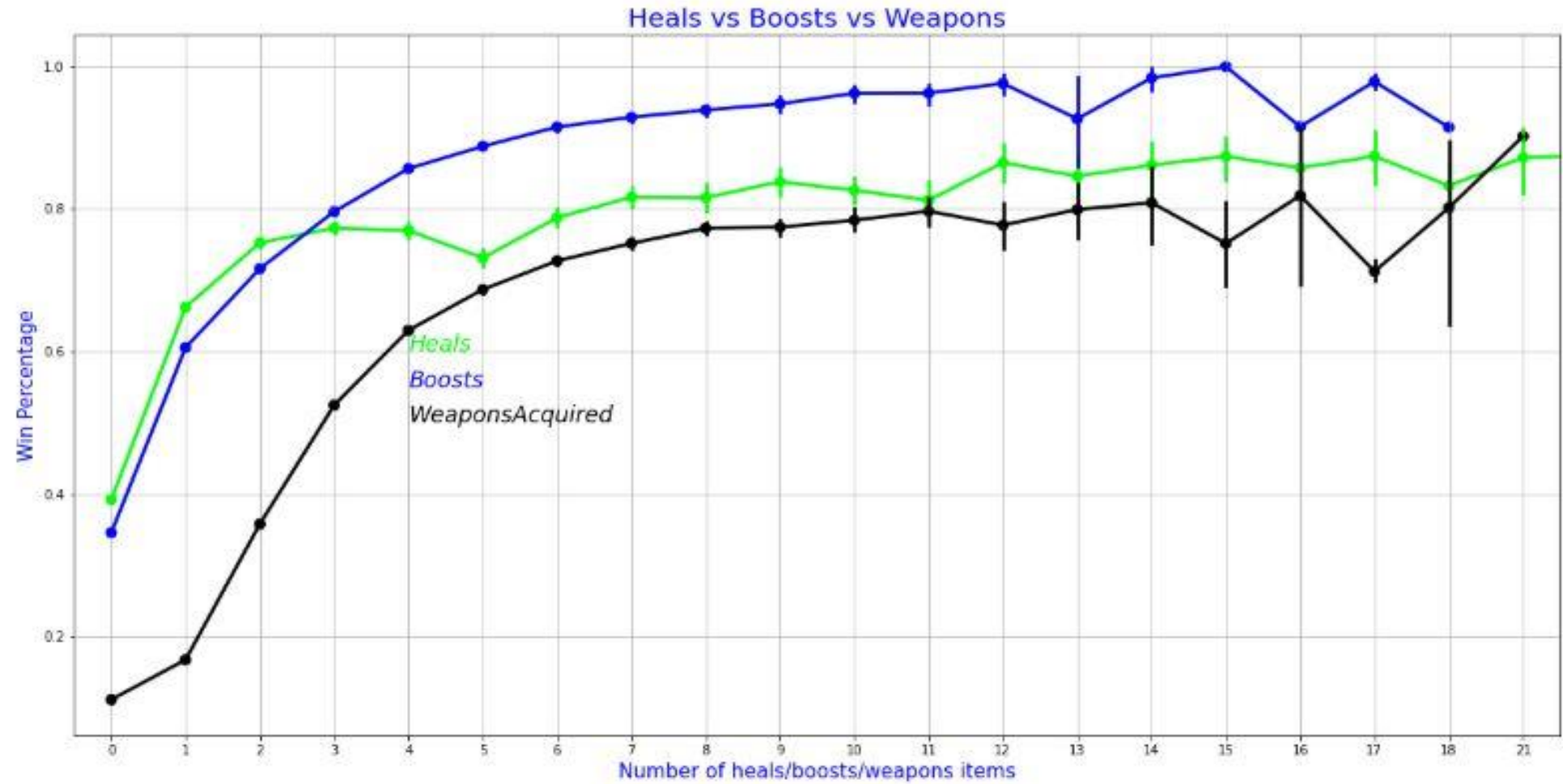
双排中84 个玩家 (0.17%) 不用杀一个人就可以赢得比赛

四排中403 个玩家 (0.81%) 不用杀一个人就可以赢得比赛

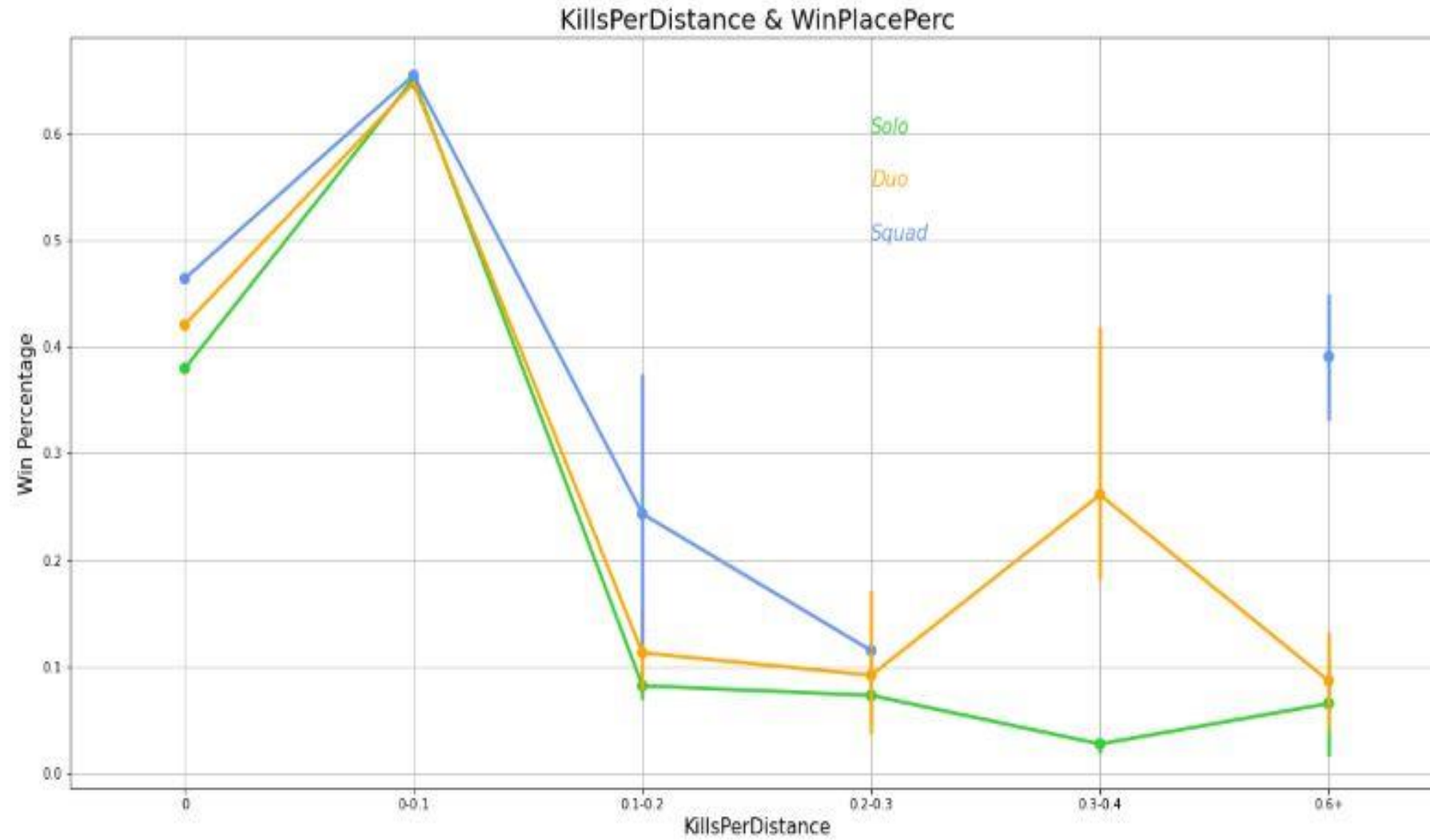
boosts



boosts, heals, weapons



KillsPerDistance





全特征聚类后预测



模型		单排 MAE	单排 R ²
GBRT 模型	聚类前	0.0515	0.9385
	类别 0	0.1385	0.4988
	类别 1	0.0512	0.9400
	类别 2	0.0690	0.8120
	类别 3	0.0788	0.6939
	类别 4	0.0781	0.6866
	类别 5	0.0883	0.3371

模型		双排 MAE	双排 R ²
GBRT 模型	聚类前	0.0574	0.9269
	类别 0	0.05606	0.9297
	类别 1	0.0889	0.5414
	类别 2	0.0753	0.7584
	类别 3	0.1893	-0.6813

模型		四排 MAE	四排 R ²
GBRT 模型	聚类前	0.0768	0.8792
	类别 0	0.0747	0.8876
	类别 1	0.0916	0.6902
	类别 2	0.1632	0.3077
	类别 3	0.1672	-2.4395

DISCUSS AND CONCLUSION

CLICK



项目亮点



研究的内容是时下流行的大逃杀类游戏。



通过官方API获得上百万玩家数据集，丰富的无差异数据集使得研究具有普遍意义。



通过探索挖掘发现玩家行为特征之间以及与最后排名水平之间的潜在联系，并在原数据基础上通过特征工程创建出新的特征，更全面反映玩家的行为表现。



使用了丰富的预测模型和聚类模型，探究过程中一步步优化玩家排名水平预测精度，从而对玩家给出建议以帮助他们选择更适合自己的作战策略。



- 需要一段时间的研究了解才能确定是否合适

项目定题坎坷

试错成本高

- 整个流程前后为串联关系
- 特征工程、特征选择、聚类方法、根据赛局时长分割

数据获取难

- 尝试API接口爬取
- 数据丰富度较低
- 数据获取时间成本较大

遇到的问题

算法复杂度高

- 数据维度较高
- 存在无意义维度

特征信息量不足

- 原数据集的特征是直观数据，信息待挖掘

维度灾难

- 数据维度较高
- 存在无意义维度

聚类效果差

- 尝试多种聚类方法
- 轮廓系数普遍较低

聚类后获得信息与预期不一致

- 聚类结果与预期不一致
- 簇类数据量分布不均，预测效果较差

PUBG

策略聚类 排名预测

CONCLUSION

This part shows the importance of GBRT model and clustering center, and gives strategy explanation according to different clustering center.

1. 相较于单排赛局，双排和四排中需要多考虑团队因素，一方面是己方的战力水平、生存水平、物资获取等的相互配合，另一方面是防止敌方团战、埋伏。
2. 对于作战水平较弱的玩家，自边缘苟活发育的战术可以有效提升最终排名水平，注意躲避人员密集区。但是若此时处于团队作战赛局，应在发育成功后及时与队友汇合，在物资上支援队友的同时寻求队友保护。
3. 作战水平较强的玩家在前期可选择装备集中区降落，从而迅速发育自己，在快速确定使用的枪支后不要轻易更换，通过前面分析看出频繁更换枪支对于后期获胜没有帮助。
4. 前期发育过程率先选择武器和提升装备，治疗物资在前期作用不大。
5. 团队局人数不均衡时需要更多的击杀、路程、装备等才可能获得较好的排名水平，故而需要根据自身水平慎重考虑

THANKS YOUR SUPPORT



汪瑶 | 余佳熹 | 赵娅妮 | 陈秋燕 | 林艺婧