

# Pokemon Battle : Who Wins?

Team Members:

James Wan

Zexi Wang

Wenchao Zhu





# Motivation

- Pokemon, as a popular game, owns millions of loyal players all around the world. Our ultimate goal is to design a webapp to help the players make better decisions on the choice of Pokemons in a battle.



V.S.





# Data Preprocessing

- Data overview

#	Name	Type 1	Type 2	HP	Attack	Defense	Sp. Atk	Sp. Def	Speed	Generation	Legendary	
0	1	Bulbasaur	Grass	Poison	45	49	49	65	65	45	1	False
1	2	Ivysaur	Grass	Poison	60	62	63	80	80	60	1	False
2	3	Venusaur	Grass	Poison	80	82	83	100	100	80	1	False
3	4	Mega Venusaur	Grass	Poison	80	100	123	122	120	80	1	False
4	5	Charmander	Fire	NaN	39	52	43	60	50	65	1	False
5	6	Charmeleon	Fire	NaN	58	64	58	80	65	80	1	False
6	7	Charizard	Fire	Flying	78	84	78	109	85	100	1	False
7	8	Mega Charizard X	Fire	Dragon	78	130	111	130	85	100	1	False
8	9	Mega Charizard Y	Fire	Flying	78	104	78	159	115	100	1	False
9	10	Squirtle	Water	NaN	44	48	65	50	64	43	1	False

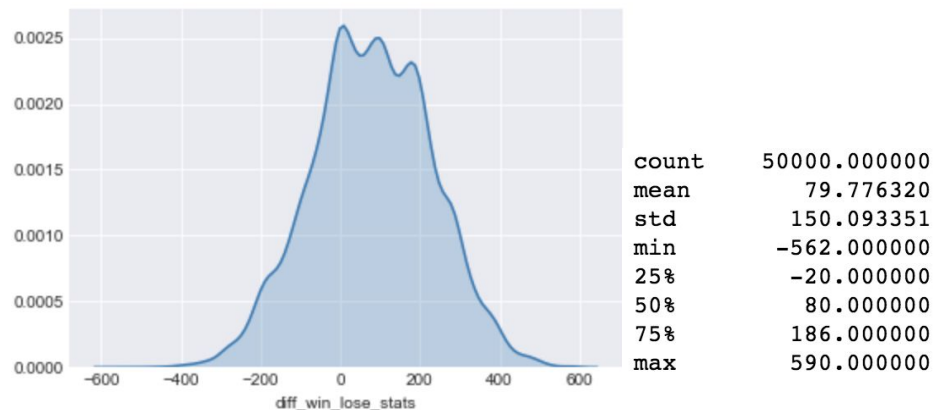
	First_pokemon	Second_pokemon	Winner
0	Larvitar	Nuzleaf	Nuzleaf
1	Virizion	Terrakion	Terrakion
2	Togetic	Beheeyem	Beheeyem
3	Slugma	Druddigon	Druddigon
4	Omastar	Shuckle	Omastar



# Data Preprocessing

- Remove the useless columns
  - Remove the 'Generation' and 'Legendary' columns
  - Add 'stats-sum' and 'winning ratio' columns

#	Name	Type 1	Type 2	HP	Attack	Defense	Sp. Atk	Sp. Def	Speed	stats_sum	Type	ratio	
0	1	Bulbasaur	Grass	Poison	45	49	49	65	65	45	318	Grass/Poison	0.278195
1	2	Ivysaur	Grass	Poison	60	62	63	80	80	60	405	Grass/Poison	0.380165
2	3	Venusaur	Grass	Poison	80	82	83	100	100	80	525	Grass/Poison	0.674242
3	4	Mega Venusaur	Grass	Poison	80	100	123	122	120	80	625	Grass/Poison	0.560000
4	5	Charmander	Fire	None	39	52	43	60	50	65	309	Fire/None	0.491071





# Data Preprocessing

- Data integrity check
  - Filling the NaN data (Type 2 and Name)
- Change the 0 to 0.25 in the type chart

	Bug	Dark	Dragon	Electric	Fairy	Fighting	Fire	Flying	Ghost	Grass	Ground	Ice	Normal	Poison	Psychic	Rock	Steel	Water
Bug	1.0	1.0	1.0	1.0	1.0	0.5	2.0	2.0	1.0	0.5	0.5	1.0	1.0	1.0	1.0	2.0	1.0	1.0
Dark	2.0	0.5	1.0	1.0	2.0	2.0	1.0	1.0	0.5	1.0	1.0	1.0	1.0	1.0	0.0	1.0	1.0	1.0
Dragon	1.0	1.0	2.0	0.5	2.0	1.0	0.5	1.0	1.0	0.5	1.0	2.0	1.0	1.0	1.0	1.0	1.0	0.5
Electric	1.0	1.0	1.0	0.5	1.0	1.0	1.0	0.5	1.0	1.0	2.0	1.0	1.0	1.0	1.0	1.0	0.5	1.0
Fairy	0.5	0.5	0.0	1.0	1.0	0.5	1.0	1.0	1.0	1.0	1.0	1.0	1.0	2.0	1.0	1.0	2.0	1.0
Fighting	0.5	0.5	1.0	1.0	2.0	1.0	1.0	2.0	1.0	1.0	1.0	1.0	1.0	1.0	2.0	0.5	1.0	1.0
Fire	0.5	1.0	1.0	1.0	0.5	1.0	0.5	1.0	1.0	0.5	2.0	0.5	1.0	1.0	1.0	2.0	0.5	2.0
Flying	0.5	1.0	1.0	2.0	1.0	0.5	1.0	1.0	1.0	0.5	0.0	2.0	1.0	1.0	1.0	2.0	1.0	1.0
Ghost	0.5	2.0	1.0	1.0	1.0	0.0	1.0	1.0	2.0	1.0	1.0	1.0	0.0	0.5	1.0	1.0	1.0	1.0
Grass	2.0	1.0	1.0	0.5	1.0	1.0	2.0	2.0	1.0	0.5	0.5	2.0	1.0	2.0	1.0	1.0	1.0	0.5
Ground	1.0	1.0	1.0	0.0	1.0	1.0	1.0	1.0	1.0	2.0	1.0	2.0	1.0	0.5	1.0	0.5	1.0	2.0
Ice	1.0	1.0	1.0	1.0	1.0	2.0	2.0	1.0	1.0	1.0	1.0	0.5	1.0	1.0	1.0	2.0	2.0	1.0
Normal	1.0	1.0	1.0	1.0	1.0	2.0	1.0	1.0	0.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Poison	0.5	1.0	1.0	1.0	0.5	0.5	1.0	1.0	1.0	0.5	2.0	1.0	1.0	0.5	2.0	1.0	1.0	1.0
Psychic	2.0	2.0	1.0	1.0	1.0	0.5	1.0	1.0	2.0	1.0	1.0	1.0	1.0	0.5	1.0	1.0	1.0	1.0
Rock	1.0	1.0	1.0	1.0	1.0	2.0	0.5	0.5	1.0	2.0	2.0	1.0	0.5	0.5	1.0	1.0	2.0	2.0
Steel	0.5	1.0	0.5	1.0	0.5	2.0	2.0	0.5	1.0	0.5	2.0	0.5	0.5	0.0	0.5	0.5	0.5	1.0
Water	1.0	1.0	1.0	2.0	1.0	1.0	0.5	1.0	1.0	2.0	1.0	0.5	1.0	1.0	1.0	1.0	0.5	0.5

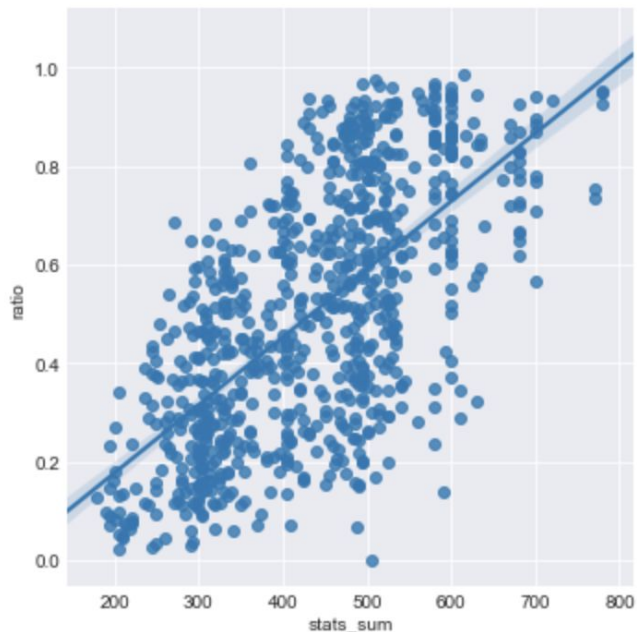


	Bug	Dark	Dragon	Electric	Fairy	Fighting	Fire	Flying	Ghost	Grass	Ground	Ice	Normal	Poison	Psychic	Rock	Steel	Water
Bug	1.0	1.0	1.00	1.00	1.0	0.50	2.0	2.0	1.00	0.5	0.50	1.0	1.00	1.00	1.00	2.0	1.0	1.0
Dark	2.0	0.5	1.00	1.00	2.0	2.00	1.0	1.0	0.50	1.0	1.00	1.0	1.00	1.00	0.25	1.0	1.0	1.0
Dragon	1.0	1.0	2.00	0.50	2.0	1.00	0.5	1.0	1.00	0.5	1.00	2.0	1.00	1.00	1.00	1.0	1.0	0.5
Electric	1.0	1.0	1.00	0.50	1.0	1.00	1.0	0.5	1.00	1.0	2.00	1.0	1.00	1.00	1.00	1.0	0.5	1.0
Fairy	0.5	0.5	0.25	1.00	1.0	0.50	1.0	1.0	1.00	1.0	1.00	1.0	1.00	2.00	1.00	1.0	2.0	1.0
Fighting	0.5	0.5	1.00	1.00	2.0	1.00	1.0	2.0	1.00	1.0	1.00	1.0	1.00	1.00	2.00	0.5	1.0	1.0
Fire	0.5	1.0	1.00	1.00	0.5	1.00	0.5	1.0	1.00	0.5	2.00	0.5	1.00	1.00	1.00	2.0	0.5	2.0
Flying	0.5	1.0	1.00	2.00	1.0	0.50	1.0	1.0	1.00	0.5	0.25	2.0	1.00	1.00	1.00	2.0	1.0	1.0
Ghost	0.5	2.0	1.00	1.00	1.0	0.25	1.0	1.0	2.00	1.0	1.00	1.0	0.25	0.50	1.00	1.0	1.0	1.0
Grass	2.0	1.0	1.00	0.50	1.0	1.00	2.0	2.0	1.00	0.5	0.50	2.0	1.00	2.00	1.00	1.0	1.0	0.5
Ground	1.0	1.0	1.00	0.25	1.0	1.00	1.0	1.0	1.00	2.0	1.00	2.0	1.00	0.50	1.00	0.5	1.0	2.0
Ice	1.0	1.0	1.00	1.00	1.0	2.00	2.0	1.0	1.00	1.0	1.00	0.5	1.00	1.00	1.00	2.0	2.0	1.0
Normal	1.0	1.0	1.00	1.00	1.0	2.00	1.0	1.0	0.25	1.0	1.00	1.0	1.00	1.00	1.00	1.0	1.0	1.0
Poison	0.5	1.0	1.00	1.00	0.5	0.50	1.0	1.0	1.00	0.5	2.00	1.0	1.00	0.50	2.00	1.0	1.0	1.0
Psychic	2.0	2.0	1.00	1.00	1.0	0.50	1.0	1.0	2.00	1.0	1.00	1.0	1.00	1.00	0.50	1.0	1.0	1.0
Rock	1.0	1.0	1.00	1.00	1.0	2.00	0.5	0.5	1.00	2.0	2.00	1.0	0.50	0.50	1.00	1.0	2.0	2.0
Steel	0.5	1.0	0.50	1.00	0.5	2.00	2.0	0.5	1.00	0.5	2.00	0.5	0.50	0.25	0.50	0.5	0.5	1.0
Water	1.0	1.0	1.00	2.00	1.0	1.00	0.5	1.0	1.00	2.0	1.00	0.5	1.00	1.00	1.00	1.0	0.5	0.5



# Data Preprocessing

- Predict the winning ratio of Pokemons with no battles
  - Linear regression




#		Name	Type 1	Type 2	HP	Attack	Defense	Sp. Atk	Sp. Def	Speed	stats_sum	Type	ratio
11	12	Blastoise	Water	None	79	83	100	85	105	78	530	Water/None	0.632075
32	33	Sandsrew	Ground	None	50	75	85	20	30	40	300	Ground/None	0.314821
34	35	Nidoran_	Poison	None	55	47	52	40	40	41	275	Poison/None	0.280336
45	46	Wigglytuff	Normal	Fairy	140	70	45	85	50	45	435	Normal/Fairy	0.501035
65	66	Poliwag	Water	None	40	50	40	40	40	90	300	Water/None	0.314821
77	78	Victreebel	Grass	Poison	80	105	65	100	70	70	490	Grass/Poison	0.576900
89	90	Magneton	Electric	Steel	50	60	95	120	70	70	465	Electric/Steel	0.542416
143	144	Ditto	Normal	None	48	48	48	48	48	48	288	Normal/None	0.298268
182	183	Ariados	Bug	Poison	70	90	70	60	60	40	390	Bug/Poison	0.438964
235	236	Ursaring	Normal	None	90	130	75	75	75	55	500	Normal/None	0.590694
321	322	Hariyama	Fighting	None	144	120	60	40	60	50	474	Fighting/None	0.554831
418	419	Mega Latias	Dragon	Psychic	80	100	120	140	150	110	700	Dragon/Psychic	0.866568
478	479	Honchkrow	Dark	Flying	100	125	52	105	52	71	505	Dark/Flying	0.597591
555	556	Servine	Grass	None	60	60	75	60	75	83	413	Grass/None	0.470689
617	618	Maractus	Grass	None	75	86	67	106	67	60	461	Grass/None	0.536899
654	655	Jellicent	Water	Ghost	100	60	70	85	105	60	480	Water/Ghost	0.563107
781	782	Pumpkaboo Small Size	Ghost	Grass	44	66	70	44	55	56	335	Ghost/Grass	0.363098



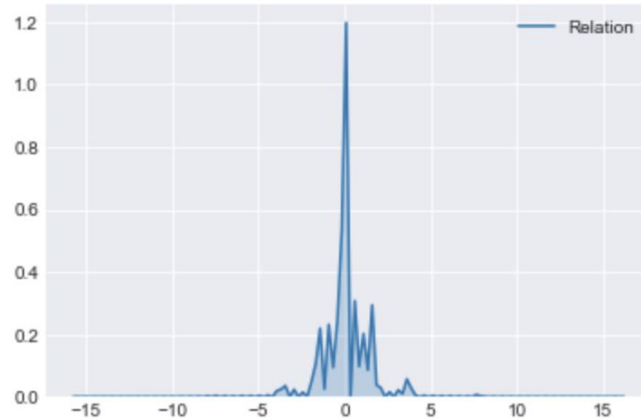
# Data Preprocessing

- Calculate the type relation between two Pokemons

	rock	ground	grass	dark
rock			2	1
ground			2	1
grass	1	0.5		
dark	1	1		

$$R = 2 \times 2 \times 1 \times 1 - 1 \times 1 \times 0.5 \times 1 = 3.5$$



count	50000.000000
mean	0.107312
std	1.296380
min	-15.500000
25%	-0.500000
50%	0.000000
75%	0.500000
max	15.937500

- Standardization
  - Z-score



# Feature Selection

- Features:

- Type relation
- Hp
- Attack
- Special attack
- Defense
- Special defense
- Speed
- Stats-sum
- Winning ratio

1	First_pokemon_hp	First_pokemon_attack	First_pokemon_defense	First_pokemon_spatack	First_pokemon_spdefense	First_pokemon_speed	First_pokemon_stats	First_pokemon_ratio
2	50	64	50	45	50	41	300	0.227272727
3	91	90	72	90	129	108	580	0.896
4	55	40	85	80	105	40	405	0.236220472
5	40	40	40	70	40	20	250	0.032520325
6	70	60	125	115	70	55	495	0.355072464
7	50	47	50	57	50	65	319	0.463768116
8	40	50	45	70	45	70	320	0.551724138
9	70	80	50	35	35	35	305	0.14
10	50	65	90	35	35	15	290	0.146788991

1	Second_pokemon_hp	Second_pokemon_attack	Second_pokemon_defense	Second_pokemon_spatack	Second_pokemon_spdefense	Second_pokemon_speed	Second_pokemon_stats	Second_pokemon_ratio	Relation
2	70	70	40	60	40	60	340	0.424778761	-3.5
3	91	129	90	72	90	108	580	0.853658537	3.5
4	75	75	75	125	95	40	485	0.156716418	0
5	77	120	90	60	90	48	485	0.278571429	-0.5
6	20	10	230	10	230	5	505	0	3
7	60	50	150	50	150	60	520	0.520661157	-0.25
8	65	50	35	115	95	95	455	0.733333333	-0.5
9	150	100	120	100	120	90	680	0.666666667	-0.75
10	50	53	62	58	63	44	330	0.137096774	0





# Feature Reduction

- Calculate Attack / Special attack to get the 'ra' values

- $ra \geq 1.5$

No use of special attack, correspondingly,  
no use of the special defense for its opponent.

- $ra \leq 0.667$

No use of attack, correspondingly,  
no use of the defense for its opponent.

- $0.667 < ra < 1.5$

No features need to remove.

First_pokemon_ra	Second_pokemon_ra
1.422222	1.166667
1.000000	1.791667
0.500000	0.600000
0.571429	2.000000
0.521739	1.000000
0.824561	1.000000
0.714286	0.434783
2.285714	1.000000
1.857143	0.913793
0.545455	1.100000

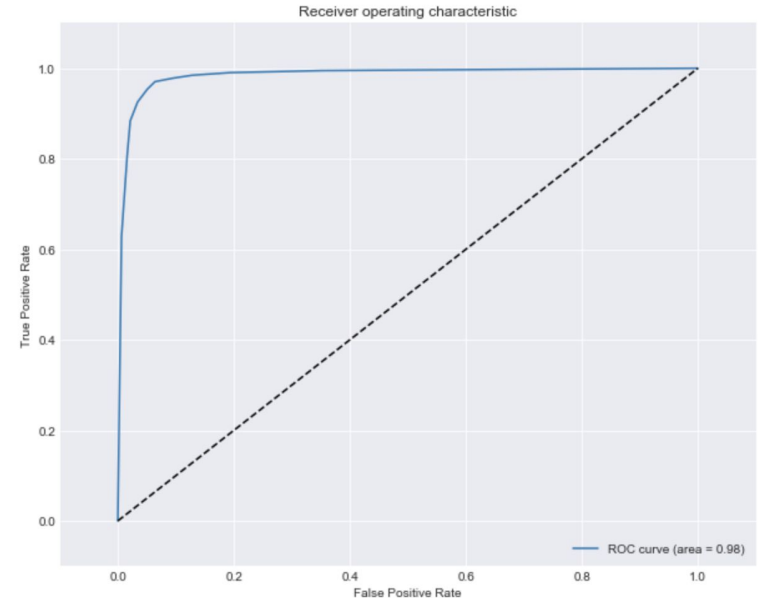


# Model Selection and Cross Validation

- Cross validation score

	Model	Score
4	Random Forest	95.777617
3	Decision Tree	94.804111
0	Logistic Regression	93.459483
1	KNN	87.600408
2	Naive Bayes	79.374606

- ROC





# Model Selection and Cross Validation

- Compare scores with/without reduction

```
score before reduction:95.77761712672537  
score after reduction:95.77761712672537
```

- Accuracy estimation for 9 categories of battles

<https://github.com/Columbia-Intro-Data-Science/python-introduction-wan3231998/blob/master/Project.ipynb>

	Model	Score
0	attack attack	90.42
1	attack balance	93.01
2	attack spattack	87.26
3	balance attack	93.69
4	balance balance	95.77
5	balance spattack	92.00
6	spattack attack	88.18
7	spattack balance	93.14
8	spattack spattack	94.05



# Prediction and Web app

- Django

