

A Statistical Analysis of Pokémon Competitive Viability

STA210 Final Project - Jason Guan

Introduction

With over 480 million sales worldwide, Pokémon, short for Pocket Monsters, is one of the most successful game series in the world. Spanning 122 installations and roughly 28 years, the Pokémon games generally revolve around catching, training, and battling fictional creatures called “Pokémon” in a quest to become the Pokémon Champion, creating an entertaining gameplay experience for children and adults alike. Despite its simple premise and cartoon aesthetics, Pokémon is home to a rich competitive scene which has continued to grow and evolve since the establishment of the official Pokémon tournament circuit (Video Game Championships, or VGC) in 2009.

At the heart of this community is Smogon, an online hub which has been hosting competitions, forums, and strategy guides since the early 2000s. Above all, Smogon’s greatest contribution to the competitive scene comes in the form of its widely respected tiering system, forming the basis on which most competitive online play is based. These tiers are based on Pokémon usage rates and dictate which Pokémon are allowed in different game modes, with more powerful Pokémon in the higher tiers and lesser used Pokémon inhabiting the lower tiers. At the top lies AG (anything goes) and Ubers, which include Pokémon like Arceus and Rayquaza that are typically banned from online play. Below this lies OU (OverUsed), UU (UnderUsed), RU (Rarely Used), NU (Never Used), PU (Perfectly Useless), and their respective Banlists, which contain Pokémon banned in certain tiers but allowed in higher tiers. At the bottom we see LC (Little Cup) and NFE (Not Fully Evolved), which consist of unevolved, statistically weak pokemon that are not typically used in normal online play.

This study seeks to analyze how factors such as stats, typing, and Mega/Legendary status impact a Pokémon’s competitive rank, as well as the factors which may cause a Pokémon to be banned (AG/Uber tier) from OU competitive play. In doing so, we will hopefully be able to recognize trends in Pokémon power-scaling and identify weaknesses in game balancing which could potentially be addressed in future generations.

Our Data

Due to a lack of available recent data, the two datasets used in this report come from Generation 6 of the Pokémon games, which includes the X/Y and Omega Ruby/Alpha Sapphire games. The ‘smogon’ dataset features 499 observations and 21 columns which include all Pokémon in and above the PU (Perfectly Useless) tier. To include the Pokémon that were missing from this dataset (those below the PU tier), a subset of missing Pokémon was taken from the ‘pokemon’ dataset, which features every Pokémon up to Generation 6 but does not store their competitive tiers. This new data was cleaned by changing all variable names to match those of the ‘smogon’ dataset and removing unnecessary variables such as catch rate and egg groups. It was then merged with the existing data to create the updated ‘smogon’ data frame used throughout this project. This final data frame was then further cleaned by giving the newly imported Pokémon their respective tiers (LC or NFE) and releveling the ‘Tier’ variable to reflect the actual ordering. Important variables in this data frame include:

- **HP, Attack, Sp. Atk, Defense, Sp. Def, and Speed:** 6 continuous variables representing a Pokémon’s 6 in-game stats
- **Type.1, Type.2:** categorical variables which reflect the type(s) a Pokémon has (e.g. fire, water, grass, etc.)
- **Legendary/Mega:** boolean variables representing whether or not a Pokémon is Legendary (one of a kind in-game) or a Mega form
- **Tier:** a categorical variable storing the competitive tier (LC to AG) a Pokémon was placed in by Smogon;

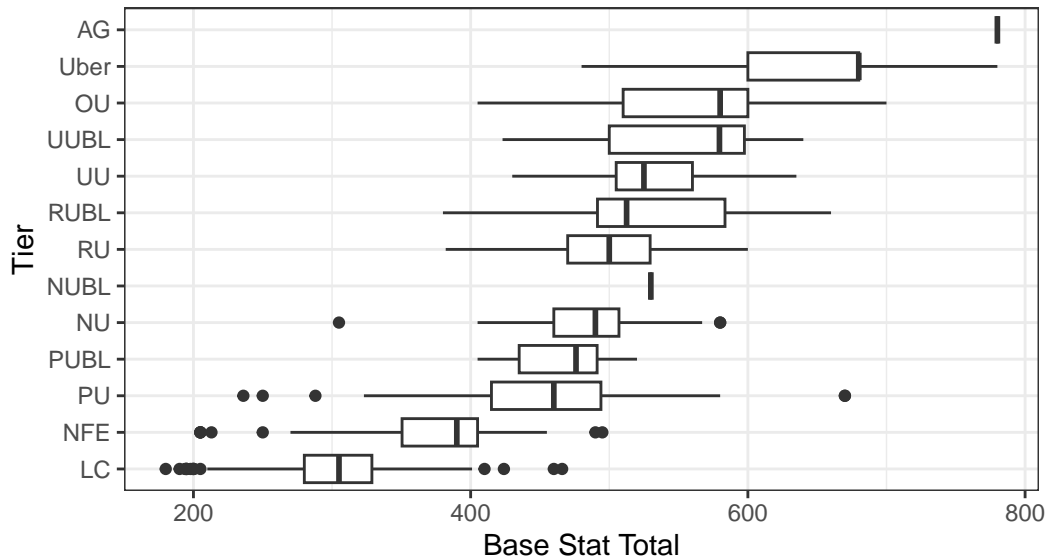
These tiers are ordered from lowest to highest as follows: LC, NFE, PU, PUBL, NU, NUBL, RU, RUBL, UU, UUBL, OU, Uber, AG

Exploratory Data Analysis

Given that the goal of this project is to analyze the factors that impact a Pokémon’s competitive rank, our exploratory data analysis will focus primarily on examining the distribution of various attributes among Pokémon across different tiers, such as stat totals, typings, and legendary or mega status. To start, the boxplot on the following page illustrates the distribution of Pokémon base stat totals across different competitive tiers. We can see in this plot that, with the exception of the NUBL tier, which only contains one Pokémon, the median base stat totals (BSTs) seem to **increase** as we move up the tiers, meaning that Pokémon in higher competitive tiers tend to have higher base stat totals. In the lower tiers, there are three outliers: Articuno (PU), Slaking (PU), and Regirock (NU). These Pokémon have high base stat totals, though all carry some competitive hindrance which keeps them in low competitive tiers. Regirock and Articuno boast weak/vulnerable typings, while Slaking’s immense 670 BST is hindered severely by the “Truant” ability, which renders it useless every second turn.

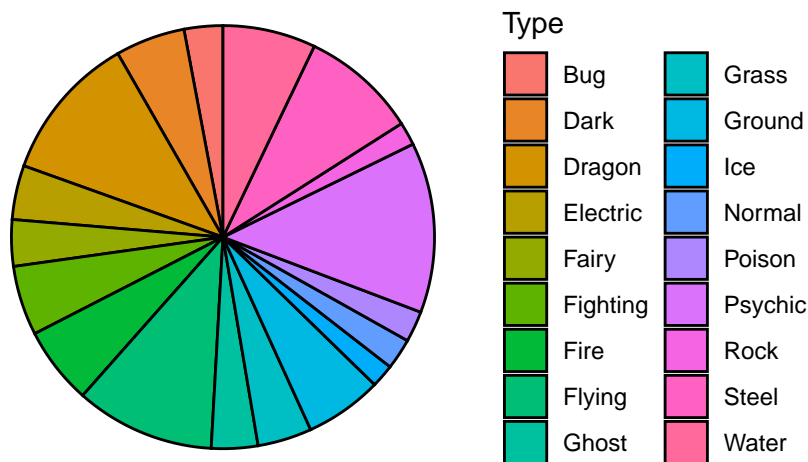
Pokemon Base Stat Totals by Smogon Tier

Median BSTs Increase Across Tiers



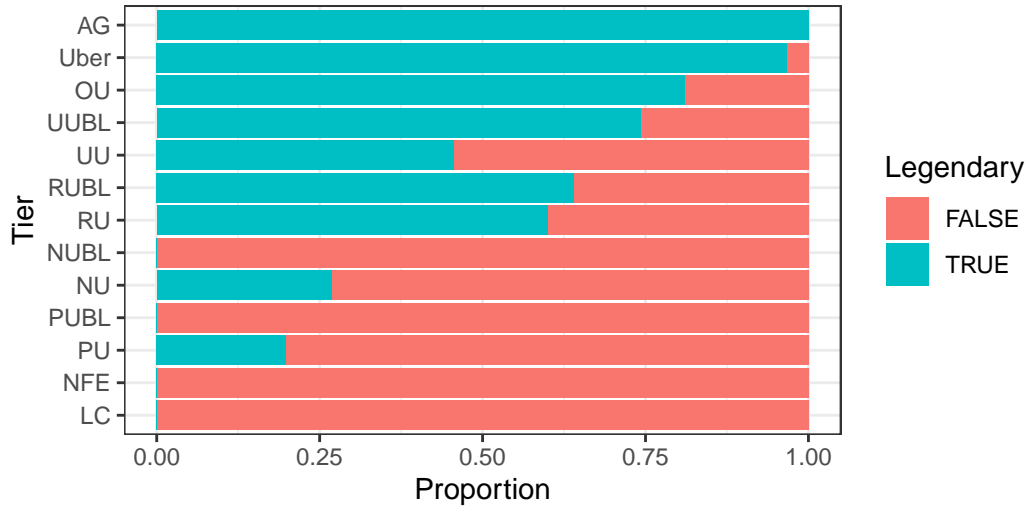
The figure below illustrates the distribution of typings across the OU, Uber, and AG tiers, the tiers where Pokémon are typically considered competitively viable. From this, we can see that Psychic, Flying, Dragon, and Steel type Pokémon are the most common in these tiers. However, note that Fairy, generally considered as one of the best typings in the game, was introduced in this generation, meaning there were significantly less Fairy Pokémon in the dataset compared to other typings.

Distribution of Pokémon Types in OU and Above

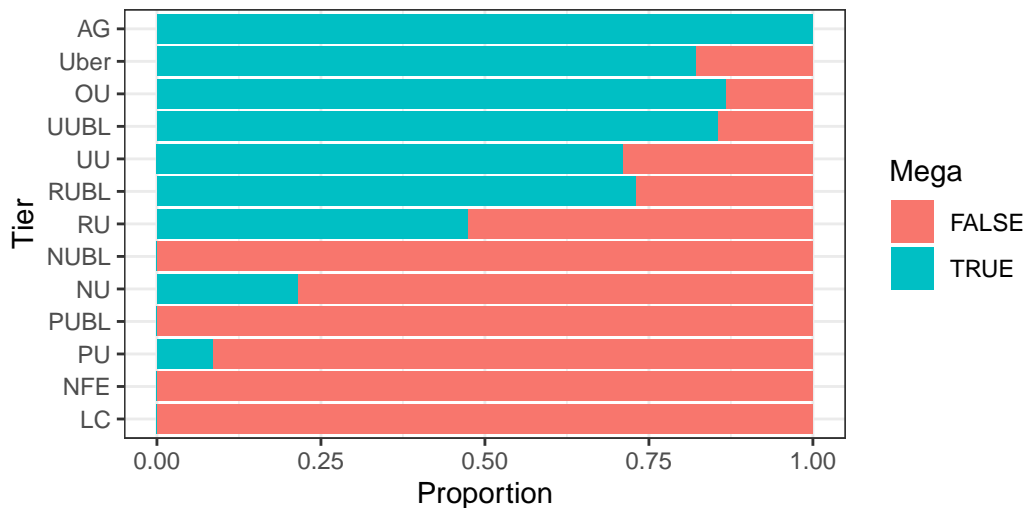


The next two visualizations show the proportion of Pokémon in each tier that are legendary or mega-forms. Legendary Pokémon are characterized by their rarity, high BSTs, and prominent roles in the stories of each Pokémon game, while mega evolutions are new Pokémon forms introduced in Generation 6 that give existing Pokémon a new “mega” evolution, often with very high BSTs and new abilities. We see in both plots that, as we move up to more competitive tiers, the proportion of Pokémon in each tier that are legendary or mega tends to increase. We can also notice that proportion for both visualizations peaks in the AG tier. This is because it only contains one Pokémon, Mega Rayquaza, which is both legendary and a mega.

Proportion of Legendary Pokémon in Each Tier



Proportion of Mega Pokémon in Each Tier



Methodology

[insert methodology]

These variables were chosen after it was discovered through exploratory data analysis that they may...

Call:

```
polr(formula = Tier ~ HP + Attack + Defense + Sp..Atk + Sp..Def +  
      Speed + Normal + Fire + Water + Electric + Grass + Ice +  
      Fighting + Poison + Ground + Flying + Psychic + Bug + Rock +  
      Ghost + Dragon + Dark + Steel + Fairy + Mega + Legendary,  
      data = smogon)
```

Coefficients:

	Value	Std. Error	t value
HP	0.04342	0.004003	10.8493
Attack	0.03060	0.003657	8.3687
Defense	0.03768	0.003781	9.9658
Sp..Atk	0.03953	0.003702	10.6784
Sp..Def	0.02471	0.003892	6.3482
Speed	0.04676	0.003710	12.6055
Normal	0.72170	0.314850	2.2922
Fire	-0.20352	0.321038	-0.6339
Water	0.51173	0.265680	1.9261
Electric	0.50271	0.350101	1.4359
Grass	-0.13755	0.276496	-0.4975
Ice	-0.32963	0.397887	-0.8285
Fighting	1.93017	0.338572	5.7009
Poison	1.22038	0.302561	4.0335
Ground	1.04172	0.291442	3.5744
Flying	0.67564	0.239745	2.8182
Psychic	0.55969	0.296561	1.8873
Bug	1.67412	0.306831	5.4562
Rock	-0.80080	0.321982	-2.4871
Ghost	1.21553	0.361285	3.3645
Dragon	0.12281	0.335892	0.3656
Dark	1.47346	0.316887	4.6498
Steel	1.42463	0.354364	4.0202
Fairy	2.22264	0.389162	5.7113
MegaTRUE	-1.07882	0.350984	-3.0737
LegendaryTRUE	-0.24876	0.331984	-0.7493

Intercepts:

	Value	Std. Error	t value
LC NFE	14.1872	0.6999	20.2704
NFE PU	15.6910	0.7593	20.6646
PU PUBL	18.5775	0.8681	21.4005
PUBL NU	18.6948	0.8707	21.4717
NU NUBL	19.5623	0.8904	21.9704
NUBL RU	19.5781	0.8907	21.9798
RU RUBL	20.4791	0.9118	22.4611
RUBL UU	20.8363	0.9201	22.6465
UU UUBL	22.1192	0.9478	23.3373
UUBL OU	22.5374	0.9557	23.5824
OU Uber	24.5850	1.0123	24.2871
Uber AG	30.6729	1.5913	19.2748

Residual Deviance: 1958.38

AIC: 2034.38

	HP	Attack	Defense	Sp..Atk	Sp..Def
	1.0443814	1.0310772	1.0384013	1.0403195	1.0250162
Speed		Normal	Fire	Water	Electric
	1.0478742	2.0579361	0.8158548	1.6681801	1.6532029
Grass		Ice	Fighting	Poison	Ground
	0.8714908	0.7191876	6.8906596	3.3884788	2.8340858
Flying		Psychic	Bug	Rock	Ghost
	1.9652854	1.7501212	5.3341063	0.4489678	3.3720892
Dragon		Dark	Steel	Fairy	MegaTRUE
	1.1306751	4.3643223	4.1563215	9.2316649	0.3399962
LegendaryTRUE					
	0.7797634				

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

Call:

```
glm(formula = Banned ~ Total + Mega + Legendary + Normal + Fire +
  Water + Electric + Grass + Ice + Fighting + Poison + Ground +
  Flying + Psychic + Bug + Rock + Ghost + Dragon + Dark + Steel +
  Fairy, family = "binomial", data = smogon)
```

Coefficients:

Estimate	Std. Error	z value	Pr(> z)
----------	------------	---------	----------

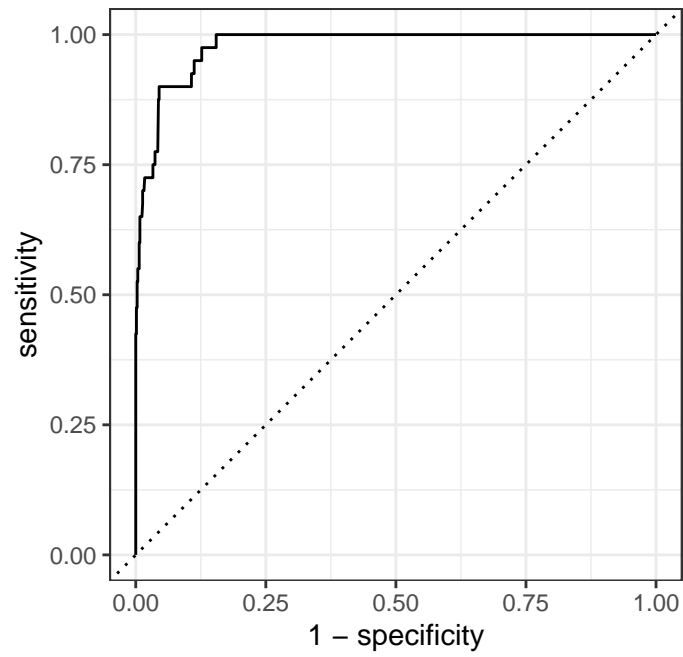
(Intercept)	-21.35722	4.02072	-5.312	1.09e-07	***
Total	0.03205	0.00714	4.489	7.16e-06	***
MegaTRUE	-0.36351	0.87023	-0.418	0.6762	
LegendaryTRUE	1.90393	0.86216	2.208	0.0272	*
Normal	-1.46804	1.54391	-0.951	0.3417	
Fire	-0.40522	0.90167	-0.449	0.6531	
Water	-0.76659	1.08974	-0.703	0.4818	
Electric	-1.34954	1.37849	-0.979	0.3276	
Grass	-1.29288	1.29672	-0.997	0.3187	
Ice	-2.41179	1.48788	-1.621	0.1050	
Fighting	1.34427	0.95675	1.405	0.1600	
Poison	0.34403	1.53578	0.224	0.8227	
Ground	-0.76224	1.14756	-0.664	0.5065	
Flying	0.22855	0.82716	0.276	0.7823	
Psychic	-1.35615	0.94031	-1.442	0.1492	
Bug	-0.13896	1.39605	-0.100	0.9207	
Rock	-20.75568	1647.66062	-0.013	0.9899	
Ghost	2.80746	1.09959	2.553	0.0107	*
Dragon	-1.06995	0.88424	-1.210	0.2263	
Dark	1.89754	1.00858	1.881	0.0599	.
Steel	0.42386	0.89559	0.473	0.6360	
Fairy	2.13994	1.37297	1.559	0.1191	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 317.32 on 796 degrees of freedom
 Residual deviance: 112.87 on 775 degrees of freedom
 AIC: 156.87

Number of Fisher Scoring iterations: 19



```
# A tibble: 1 x 3
  .metric .estimator .estimate
  <chr>   <chr>       <dbl>
1 roc_auc binary      0.978
```

Results