**Is Power Creep Real?**

Ian Kim

**Abstract**

Is Power Creep Real?

Imagine creating a game that must remain compelling for 25 years. How would you continuously introduce new concepts to excite and engage new and returning players? This is the problem most game developers face. There are 2 ways developers solve this problem: Break the fundamental rules of the game or introduce elements that are more powerful and impactful things than the previous. While these approaches gives the audience a more compelling game, introducing overpowered characters or mechanics will only temporarily boost the excitement, and will often lead to a disruption in overall game balance. This phenomenon is called power creep.

To better understand the nature of power creep, lets look at a popular game, League of Legends. League of Legends is a MOBA (Multiplayer Online Battle Arena) where the main premise is to take down the enemy’s nexus. After about 10 years of development, Riot has released champions that have: the potential to revive teammates (Akshan), dashes and invisibility that reset on kills (Akshan, Aurora), 4 dashes in one ability (Bel’Veth), uncapped attack speed champion (Bel’Veth), dashes every time an ability is used (Ambessa), and a champion that has a dash that is an unstoppable, airborne, gives armour and magic resistance, and does damage based on the targets maximum health in one ability (K’sante).

Pokemon does not shy away from power creep either by continuing to introduce increasingly more powerful Pokemon through its stats, abilities and moves with each new generation. For instance, giving Zacian one of the best typings in the game – fairy and steel – and giving it great overall stats with an insanely broken ability, Intrepid Sword – giving it a +1 attack boost to Zacian every time it switches in. Similarly, Urshifu breaks the game mechanics by being the first Pokemon to be able to hit through Protect/Detect/Spiky Shield (without the move Feint) with the ability Unseen Fist. Furthermore, Urshifu’s signature moves – Wicked Blow and Surging Strikes – are guaranteed to critical hit, ignoring attack drops like Intimidate or defensive boosts, making typical ways of slowing down physical Pokemon ineffective.

With 25 years of development of Pokemon, this study intends to investigate has Gamefreak statistically avoided the power creep phenomenon that is present in most series-based games? To explore this, the following research questions are addressed: is there an overall increase in stats throughout the generations? Has there an increase in the number of “good to great” types per generation? How do we define a “good” or “bad” type? Is average base power of a Pokemon move affected by power creep? Do max base stats increase as the Pokemon series progresses?

**Methodology**

This study employs mainly quantitative data such as base stats of Pokemon, type chart, movebase power and accuracy, and the generation number to compare key metrics. Data in this study was mainly collected by *PokeAPI*, which is an API (Application Programming Interface) where data about any Pokemon or generation can be gathered. Additionally, the type effectiveness chart which outlines super-effectiveness, not-very-effective, and no effect was sourced through Kaggle, which is a platform for data scientists and data analysts to analyze a variety of different publicly available datasets created by users.

Given the entire dataset was forked through GitHub, there were extensive unnecessary data was required removal to ensure data relevancy. The necessary datasets were initially cleaned through the *clean.py* program, removing any empty cells in the csv. Then, data was further processed in an SQL database, and any unnecessary data was removed here. Within *PokemonCompleteStats.xlsx,* additional columns were added to using SQL joins and Python such as Generation ID, base stat total, and Pokemon’s typing.

**Findings/Results**

|  |
| --- |
| Figure 1 represents the average base stat total per generation, sorted by highest base stat total per generation to lowest |

*Figure 1* illustrates the highest average base stat total per generation, with the top 3 being Generation 9 (Scarlet and Violet), generation 7 (Sun and Moon), and Generation 4 (Diamond, Pearl, Platinum). In contrast, the bottom 3 being generation 1 (Red, Blue, Yellow), Generation 2 (Gold and Silver), and generation 3 (Ruby, Sapphire, Emerald). Generational disparities can be attributed to many factors such as amount of legendries, ultra beast, mythical, paradox, and evolution of older Pokemon. In Generaiton 9, with the introduction of paradox Pokemon, 25.83% of Pokemon in generation 9 are paradox, legendary or mythical. For example, in generation 4 high average reflects evolutions from previous generations like Togekiss and Electivire. In generation 6, there were 72 Pokemon introduced. Therefore, having Pokemon like Fletchling and Bunnelby can influence the average massively.

Generation 4 experiences a notable increase in base stat Pokemon because many older Pokemon received their new evolutions – Ambipom, Mismagius, Weavile, Magnezone, Lickilicky, Rhyperior, Tangrowth, Electivire, Magmortar, Togekiss, Yanmega, Gliscor, Probopass, Dusknoir, Gallade, Porygon-z, Frosslass – which’ll have naturally higher base stat totals than their pre-evolution counter parts. Pokemon Diamond, Pearl, Platinum ranks 3rd of having the most amount of legendaries and mythicals at 14. Though generation 4 has their fair share of baby Pokemon too – Mime jr. Mantyke, Happiny, Chingling, Budew, Riolu, and Munchlaw – the impact of these weaker Pokemon was offset by the abundance of legendary Pokemon, mythicals, and 2nd and 3rd stage evolutions.

Generation 2 and 3 are among the two of the lowest base stat totals due to of the amount of baby Pokemon that were introduced in this game such as Pichu (205), Cleffa (218), Igglybuff (210), Togepi (245), smoochum (305), elekid (360), magby (365) and notoriously weak Pokemon such as Sunkern (180), Unown (336), Smeargle (250), Shedinja (236), Azurill (190), Feebas (200), Wynaut (260), Surskit (269), ralts(198), etc…

Generation 3 in particular have an abundance of early-game Pokemon -- Wurmple line, Lotad line, Ralts line, Ziggzagoon line, Wingull line, and Poochyena line. This generation also introduced standalone Pokemon with low base stat totals like Spinda (60)

|  |
| --- |
| Figure 2 represents the median base stat total per generation |

*Figure 2* illustrates the minimum median base stat total per generation to compare and contrast a central tendency to a mean. The median for the top 3 are in order by generation, but the middle of the bar chart is a bit jumbled. Generation 5, with the highest number of new Pokemon introduced with 156, shows slightly skewed data because of a diversity of base stat distributions. A smaller generation like generation 6 can showcase more inconsistencies due to a limited sample size.

|  |
| --- |
| Figure 3 represents the number of abilities introduced every generation |

*Figure 3* represents the number of abilities introduced in every generation. Generation 1 and 2 have no abilities because abilities were introduced in generation 3, making it the generation with the most number of abilities. The number of abilities introduced in each generation does not matter, so there is no quantitative way of measuring abilities. However, all abilities are not equal – some abilities being stronger than others. In earlier generations, up until generation 5, great abilities were given tied to weaker Pokemon, whereas in later generations, these great abilities were paired with already strong Pokemon. For example, Huge Power and Pure Power are two really potent abilities that was introduced in generation 3,but they were balanced around being given to Pokemon with a low base stat in attack. Medicham, a Pokemon that was introduced in generation 3, has 60 base attack, so it turns into 120 base attack with the pure power ability. Azumarill, a Pokemon that was introduced in generation 2, has 50 base attack, so with huge power, it turns into 100 base stat total. However, in generation 6, 2 mega Pokemon were given huge power Medicham and Mawile. Medicham-M has a 100 base attack, turning into 200 base attack with pure power, and Mawile-M has 105 attack, and it turns into 210 base attack with huge power. To put this into perspective, the top 5 highest attack Pokemon without any abilities are: Mega Mewtwo-X (190), Heracross-M (185), Kartana (181), Deoxys (180) and Groudon-Primal (180).

In generation 3, Pokemon introduced the Speed Boost ability, which was balanced around weak Pokemon having this ability – Ninjask, Yanmega, Sharpedo, and Scolipede. How this balancing was disrupted when Speed Boost was introduced in Blaziken alongside his mega. This ability made Blaziken overwhelming strong that it instantly got sent into Ubers tier on Smogon. Ubers tier is a tier on Smogon where all the legendaries reside. To illustrate how impactful this ability is given to a mediocre Pokemon, let’s look into Espathra. Espathra’s stats are not anything to scream at. Despite its overall decent stats, this ostrich Pokemon got instantly sent to ubers tier in Smogon. In contrast, previously mentioned 4 Pokemon with Speed Boost were never classified as Ubers tier.

Generation 8 introduced an unprecedented ability: a Pokemon’s ability that can always hit through Protect – Urshifu’s Unseen Fist ability. This ability breaks the fundamental concept of the game because Protect is a move that should always guarantee the Pokemon that utilizes does not get damaged (barring the move Feint, but that move is so niche, it is never used). Another example of broken design in Generation 8 is Zacian-Crowned, which boasts the ability Intrepid Sword and a base Attack stat of 170. Intrepid Sword increases Zacian’s Attack by 50% upon switching in, and combined with its exceptional Fairy/Steel typing, Zacian-Crowned was so overpowered that its ability was nerfed in Generation 9 to activate only on its first switch-in.

In generation 9, new abilities pushed the game in terms of balance, the runination ability – Vessels of Ruin, Tablets of Ruin, Swords of Ruin and Beads of Ruin -- reduces specific stat of all the Pokemon on the field (except the user) by 25%. Another game changing ability, Zero to Hero was introduced, which is an ability that increase a Pokemon’s stat by 193 by just switching out once. These abilities only scrape the surface of broken abilities that were introduced in newer generations, and how supposedly broken abilities were balanced in earlier generations, and how these same abilities were given to great Pokemon, making them broken.

|  |
| --- |
| Figure 4 represents the average base power between damage classes (physical/special) per generation |

*Figure 4* represents the average base power between damage classes (physical and special), and examine if there is a correlation between an increase in base power with physical or special moves per generations. Certain moves are excluded from this analysis (see Appendix 1). Based on the data, there doesn’t seem to be any correlation between moves and their base power over generations. Moves introduced in earlier generations are ‘staple moves’ or ‘common moves’ such as Tackle or Pound for early game or Fire Blast, Hydro Pump, Draco Meteor. The prevalence of general-use moves in generation 1 and generation 2 contributes to a lower average base stat for both physical and special. In contrast, later generations introduced more signature moves or niche/utility move. For example, generation 9, 64 moves were introduced in generation 9, and 87.5% of moves are signature.

|  |
| --- |
| Figure 5 represents a column heatmap of the max base stats per generation |

*Figure 5* illustrates the maximum base stats (attack, defence, hit points, special attack, special defence, and speed) per generation. This heatmap highlights a key trend: while earlier generations feature Pokémon with exceptionally high individual stats, however, they are often offset by significant weaknesses in their other stats. For example, Blissey, a Pokemon introduced in generation 2, has the max HP stat of 255, however Blissey is balanced by having low attack, defence , speed , and special attack, and any physical attack will do a great deal of damage. In comparison, Regidrago, a Pokemon introduced in generation 8 with 200 HP has more balanced stats despite the 55 less stats in HP. The same could be said for Alomamola, the Pokemon with a 165 HP, its stats are more well-rounded even though Alomamola have 70 less base stats. Neither of these Pokemon have 10 defence and 10 attack, which provides a huge liability for the Pokemon. *Appendix* 2 showcases the base stat totals between Blissey, Regidrago, and Alomamola. A Pokemon with 10 attack is Magikarp.

The Pokémon with the highest Defense and Special Defense stats in the series is Shuckle, with an impressive 230 in both categories. However, Shuckle’s other stats are abysmally low: 20 HP, 10 Attack, 10 Special Attack, and 5 Speed, which limits its usability in most battle scenarios. By contrast, Pecharunt, a Generation 9 Pokémon with 160 base Defense, has a much more balanced stat distribution, with 88 base stats in all its other categories. This makes Pecharunt far more viable in battle compared to Shuckle, whose extreme stat specialization comes at a steep cost.

|  |
| --- |
| Figure 6 represent the ranking of typing’s based on offensive and defensive capabilities |

*Figure 6* illustrates the typing based on offensive and defensive capabilities using *RankAlgorithm.py*. This algorithm calculates an offensive and defensive score based on the type effectiveness, which is determined by how each type interacts with other types. The scoring system is as follows:

* Offensively: No effect giving a -2, super effective +2, not very effective -1, and neutral attack +1
* Defensively: neutral +1, not very effective +1, super effective giving -2, no effect +2.

The average score is derived from adding the offensive and defensive score then dividing by 2. As shown on the graph, grass and bug type Pokemon are the weakest type, reflected by their weaknesses and limited resistances. Interestingly, **Ghost** types have the highest average score, despite being known more for their offensive capabilities rather than their defensive strength. Additionally, it is surprising that Steel and Dragon types, both regarded as strong and popular typings with a broad range of Pokémon, fall into the middle range of the ranking.

|  |
| --- |
| Figure 7 represents the weighted average of typing based on how many Pokemon of a certain type were introduced in that generation |

*Figure 7* represents the weighted average of typing effectiveness based on the number of Pokémon introduced per generation. These scores were derived by multiplying average score of types by the number of types of Pokemon in that generation divided by the number of Pokemon in that generation to get the weighted average per generation. Generations 9 and 6 rank the highest, likely due to the introduction of powerful typings like Fairy and the prevalence of paradox and legendary Pokémon being mostly dark type. In contrast, Generation 8 ranks the lowest, possibly due to the smaller number of new Pokémon introduced and the lack of impactful new typings.

**Discussion**

Statistically speaking, there is little to no correlation between generation introduced and power creep. However, there is a reallocation of stat spreads which makes Pokemon seem stronger. For example, there are more balanced Pokemon rather than specialized Pokemon like Blissey or Shuckle. With a bit less base stat total in some stats, some of these Pokemon are stronger and have more usage than Shuckle or Blissey. However, looking at abilities from the pure eye test, and the different Pokemon that gets abilities in future generations showcases that there has been power creep.

Within this research, there are a few limitations, such as determining type effectiveness. Counting type effectiveness does not account for intangibles such as the Pokémon's movepool, ability, and synergy with other Pokémon in a team. Further research opportunities could include exploring additional aspects of type effectiveness, such as how certain abilities or moves impact the effectiveness of types in battle. Another avenue for research could be to investigate whether the inclusion of Legendary, Mythical, Ultra Beast, and Paradox Pokémon is a driving force behind power creep. Additionally, it would be valuable to analyze how non-Legendary Pokémon compare to each other in terms of base stat totals, as well as how Legendary Pokémon compare to one another in terms of base stat totals.

Another potential research question could be whether the average evolution stage affects the average base stat total per generation. In this context, it would be insightful to explore the typical evolution stages of Pokémon in each generation and examine how the average base stat total varies across these stages.

Conclusion

* + Summary of findings
  + Significance
  + Final remarks

**Appendices**

|  |
| --- |
| Appendix 1:   1. These moves are very strong "one time moves" or "you can use for 3 turns" (Z-Moves & Dynamax + Gigantamax moves) and will skew the results 2. Moves that are calculated based off "something", because the base power (or fixed damage) is variable 3. Some moves are not useable ingame (like eternabeam, light of ruin) 4. Fixed damage does not equal base power   Moves that were excluded:   * z-moves, dynamax moves, eternabeam,return,frustration,gyro ball,wring out, counter crush grip, electro ball, OHKO (One Hit KO) [fissure, horn drill, sheer cold,guillotine],sonic boom, low kick,seismic toss,dragon rage, fissure, night shade, bide, psywave,super-fang, flail,reversal,present,magnitude,mirror-coat, beat-up,spit-up,endeavor,natural-gift,metal burst, fling, trump card,punishment, grass knot,heavy slam, final gambit, heat crash, pika papow, veevee-volly |

|  |
| --- |
| Appendix 2:  Blissey stats: |
| Alolmamola stats: |
| Regidrago stats: |

**Works Cited**

Bailey, J. (2020, November 3). *Pokemon Type Chart*. Kaggle. https://www.kaggle.com/datasets/jadenbailey/pokemon-type-chart

Pokéapi. (n.d.). https://pokeapi.co/