

Análisis Estadístico y Ranking de los Pokémon Más Destacados por Generación

Sofia Rodriguez, Jose Aguirre y Carmen De Bolivar

2025-02-27

Planteamiento del Problema

Desde su creación, la franquicia Pokémon ha presentado una gran diversidad de especies, cada una con atributos y características únicas. Con más de 1000 Pokémon a la fecha, resulta complejo determinar objetivamente cuáles sobresalen en cada generación. Los rankings existentes a menudo incorporan sesgos subjetivos, como la popularidad o el diseño de los Pokémon. Esta investigación busca superar estas limitaciones mediante el uso de análisis estadísticos descriptivos para identificar y clasificar a los Pokémon más destacados de cada generación, basándose en sus principales estadísticas base.

Justificación de la Investigación

Este proyecto proporcionará a la comunidad de jugadores y entusiastas una evaluación objetiva y basada en datos sobre los Pokémon más sobresalientes de cada generación. El análisis estadístico permitirá identificar patrones y tendencias de los mejores Pokémon en cada generación, así como comprender mejor algunas las mecánicas internas del juego, los resultados podrán ser utilizados para optimizar equipos y estrategias de juego. Además, este proyecto servirá como una herramienta educativa para demostrar la aplicación de técnicas de las estadísticas descriptivas en un contexto más amigable y divertido para la persona promedio.

Límites y Alcances

El análisis se limitará a las estadísticas base de los Pokémon, incluyendo HP (puntos de salud), ataque, defensa, ataque especial, defensa especial y velocidad. Se utilizarán técnicas de estadística descriptiva para analizar y comparar los datos. El estudio abarcará todas las nueve generaciones de Pokémon presentes en la base de datos proporcionada. No se considerarán factores externos como habilidades, movimientos, estrategias de combate o tiers competitivos.

Marco Teórico

El proyecto se fundamentará en los conceptos de estadística descriptiva, incluyendo medidas de tendencia central, dispersión y tipificando los datos. Se utilizarán herramientas de programación como R para el análisis y visualización de los datos. Al final de cada top se tomarán en cuenta los tipos de Pokémon y sus habilidades más importantes a la hora de competir contra un líder de gimnasio.

Explicación de las Variables:

- id: Identificador único del Pokémon.
- name: Nombre del Pokémon.
- generation: Generación a la que pertenece el Pokémon.
- evolves_from: Pokémon del que evoluciona, si aplica.
- type1: Tipo primario del Pokémon.
- type2: Tipo secundario del Pokémon, si aplica.

- hp: Puntos de salud base del Pokémon.
- atk: Ataque base del Pokémon.
- def: Defensa base del Pokémon.
- spatk: Ataque especial base del Pokémon.
- spdef: Defensa especial base del Pokémon.
- speed: Velocidad base del Pokémon.
- total: Suma de todas las estadísticas base del Pokémon (HP, ataque, defensa, ataque especial, defensa especial y velocidad). Esta variable será fundamental para el ranking.
- height: Altura del Pokémon.
- weight: Peso del Pokémon.
- abilities: Habilidades del Pokémon.
- desc: Descripción del Pokémon.

Objetivos

General

Desarrollar un informe dinámico en RMarkdown que, mediante análisis estadísticos, determine el top 5 de los Pokémon más destacados por generación, utilizando una base de datos con información de las nueve generaciones y sus principales estadísticas.

Específicos:

- Dividir la base de datos en subconjuntos en base a cada generación de Pokémon.
- Realizar un análisis descriptivo de las estadísticas de manera general para todos los pokemones.
- Realizar la tipificación de los datos de las estadísticas de manera general para todos los pokemones.
- Ordenar los Pokémon de cada generación tomando en cuenta los datos tipificados de mayor a menor.
- Seleccionar los 5 mejores Pokémon de cada generación, basándonos en los datos tipificados, para formar el ranking.
- Mencionar la efectividad de los pokemons contra los lideres de gimnasio (opcional)
- Visualizar los resultados en tablas y gráficos generados en RMarkdown, que permitan la comparación entre generaciones y la identificación de los Pokémon más destacados.
- Desarrollar un dashboard interactivo que resuma gráficamente la información más relevante del estudio.

Contexto General del Tema

Desde el lanzamiento de la primera generación de Pokémon en 1996, evaluar objetivamente el desempeño de estas criaturas ha sido un desafío constante. Inicialmente, los jugadores comparaban las especies utilizando criterios cualitativos como el diseño o la rareza, o mediante métricas no estandarizadas como el Total de Estadísticas Base (BST, por sus siglas en inglés), que es la suma de las seis estadísticas fundamentales: Puntos de Salud (HP), Ataque, Defensa, Ataque Especial, Defensa Especial y Velocidad. Sin embargo, este enfoque presenta dos problemas críticos:

- **Escalada de Poder (Power Creep):** Las generaciones más recientes introducen Pokémon con BST significativamente más altos. Por ejemplo, el BST promedio de la Generación I es 420, mientras que en la Generación IX supera los 480. Esta diferencia crea una brecha artificial entre especies antiguas y nuevas, dificultando las comparaciones directas sin ajustar por variaciones generacionales (Bulbapedia, 2023).
- **Heterogeneidad de Tipos:** La distribución de tipos de Pokémon no es uniforme a lo largo de las generaciones. Algunos tipos, como Agua o Volador, son más comunes, mientras que otros, como Dragón, son menos frecuentes. Esta disparidad puede distorsionar los rankings no ajustados, ya que la abundancia o escasez de ciertos tipos influye en el desempeño relativo y en las estrategias disponibles para los jugadores (Pokémon Database, 2023).

Este estudio se enmarca en el ámbito de la estadística descriptiva estandarizada, un enfoque que permite neutralizar sesgos sistémicos al comparar Pokémon de diferentes generaciones. La relevancia de este enfoque radica en ofrecer un método replicable y transparente para establecer rankings intergeneracionales basados exclusivamente en métricas objetivas.

Revisión de Investigaciones Previas

a) Métodos Cuantitativos Sin Estandarización

- **Listas Basadas en BST Crudo:** Sitios especializados como Serebii.net (2001 - actualidad) han publicado rankings según el BST sin considerar ajustes estadísticos. Este enfoque destaca Pokémon como Slaking (Generación III, BST 670), sin tener en cuenta limitaciones específicas que afectan su rendimiento real en combate, como su habilidad Truant (Serebii.net, 2023).
- **Documentación de Estadísticas por Generación:** Bulbapedia (2005 - actualidad) ha recopilado y presentado las estadísticas base de los Pokémon por generación. Sin embargo, estas compilaciones no realizan ajustes por diferencias en las medias o desviaciones estándar entre generaciones, lo que dificulta comparaciones precisas y justas entre Pokémon de distintas épocas (Bulbapedia, 2023).

Brechas Identificadas

- **Falta de Normalización Estadística:** Comparar un BST de 500 en la Generación I (media 420) con uno en la Generación IX (media 480) es engañoso sin realizar ajustes estadísticos. Sin normalización, las diferencias naturales en las estadísticas base promedio de cada generación pueden llevar a conclusiones inexactas sobre la verdadera fortaleza de un Pokémon en su contexto generacional.
- **Variables No Consideradas:** La frecuencia y distribución de tipos entre generaciones afectan la utilidad y el desempeño real de los Pokémon. Rankings que no cuantifican y ajustan estos factores pueden ofrecer resultados sesgados, omitiendo cómo la abundancia o escasez de ciertos tipos influye en el metajuego y en las estrategias disponibles (Pokémon Database, 2023).

Aporte de Esta Investigación

Este estudio busca superar las limitaciones anteriores mediante:

- **Aplicación de Puntuaciones Z (Z-scores):** Utilizaremos la fórmula estándar de puntuación Z:

$$z = \frac{X - \mu}{\sigma}$$

Donde:

- X es el valor individual.
- μ es la media.
- σ es la desviación estándar.
- **Enfoque en Métricas Objetivas y Replicables:** Al centrarnos exclusivamente en las estadísticas base estandarizadas y evitar factores subjetivos como habilidades, movimientos o popularidad, garantizamos que el ranking resultante sea imparcial y basado en datos cuantitativos sólidos.

Importancia y Contribución del Estudio

Al ofrecer un ranking objetivo de los mejores Pokémon por generación basado en análisis estadísticos estandarizados, este estudio:

- **Proporciona una Evaluación Imparcial:** Identifica de manera precisa cuáles Pokémon sobresalen dentro de su generación y en comparación con otras, basándose en su desempeño estadístico relativo.
- **Facilita Comparaciones Intergeneracionales:** La normalización estadística elimina los sesgos introducidos por la escalada de poder, permitiendo comparaciones equitativas entre Pokémon de distintas generaciones.
- **Apoya a la Comunidad de Jugadores y Analistas:** Ofrece información valiosa para jugadores que buscan optimizar sus equipos y estrategias, así como para analistas interesados en la evolución del diseño y balance en la franquicia Pokémon.
- **Contribuye al Campo de Estudios de Videojuegos:** Demuestra la aplicación práctica de métodos estadísticos en el análisis de videojuegos, fomentando investigaciones futuras que apliquen enfoques similares en otros contextos.

Referencias

- Bulbapedia. (2023). Base Stat Total (BST). Recuperado el 15 de octubre de 2023, de <https://bulbapedia.bulbagarden.net>
- Pokémon Database. (2023). Type distribution by generation. Recuperado el 15 de octubre de 2023, de <https://pokemondb.net/type>
- Serebii.net. (2023). Base stat averages by generation. Recuperado el 15 de octubre de 2023, de <https://www.serebii.net>
- Smogon University. (2023). Smogon tier lists. Recuperado el 15 de octubre de 2023, de <https://www.smogon.com>
- The Pokémon Company. (2020). Pokémon of the Year results. Recuperado el 15 de octubre de 2023, de <https://www.pokemon.com/es/noticias/pokemon-of-the-year>

Abordamiento

Este trabajo aborda la necesidad de una evaluación objetiva y justa de los Pokémon a través de las diferentes generaciones. Al aplicar métodos estadísticos estandarizados y centrarse en métricas cuantitativas replicables, proporcionamos un ranking que refleja con mayor precisión el desempeño relativo de las especies. Este enfoque no solo beneficia a los jugadores y entusiastas de la franquicia, sino que también contribuye al ámbito académico al demostrar cómo la estadística descriptiva puede aplicarse eficazmente en el análisis de videojuegos.