



Escola  
Técnica  
Superior  
de Enxeñaría



## Visualización Avanzada

UNIVERSIDAD DE SANTIAGO DE COMPOSTELA

# PRÁCTICA 4.4: ANÁLISIS SOBRE LAS CARACTERÍSTICAS DE LOS POKÉMON

■ Hugo Vázquez Docampo

FECHA DE ENTREGA: 07 de Noviembre de 2022

## ÍNDICE

<b>1. Introducción</b>	<b>2</b>
<b>2. Motivación, preguntas a resolver y origen de los datos</b>	<b>2</b>
A. Fuente de datos . . . . .	2
<b>3. Diccionario de datos</b>	<b>2</b>
<b>4. Preparación y limpieza de los datos</b>	<b>3</b>
<b>5. Elaboración del Dashboard 1: General</b>	<b>4</b>
<b>6. Dashboard 2: análisis generacional</b>	<b>5</b>
<b>7. Dashboard 3: Ranking</b>	<b>6</b>
<b>8. Medidas DAX elaboradas</b>	<b>8</b>
<b>9. Análisis de los resultados observados</b>	<b>9</b>
<b>10. Conclusiones</b>	<b>10</b>

## 1. INTRODUCCIÓN

Se incluye en este informe un análisis de un conjunto de datos de libre elección sobre las **características de los pokemon** a lo largo de las generaciones. En él, se incluye un análisis de las cuestiones que se desean resolver; así como una motivación para la selección de este conjunto de datos. Además, se incorpora una breve descripción de los datos mostrados. A mayores, se introduce un apartado en el cual se relatan los pasos a seguir para realizar la limpieza de los datos. A continuación, se lleva a cabo el análisis de los mismos mediante la realización de una serie de Dashboards. Por último, se finaliza con una conclusión de los resultados observados.

## 2. MOTIVACIÓN, PREGUNTAS A RESOLVER Y ORIGEN DE LOS DATOS

La franquicia de los Pokémon es una de las más queridas en el anime por muchos niños y adultos. Así se muestra en la longevidad de esta. Cada año, nuevas criaturas son añadidas a la Pokédex, y con ellas nuevas habilidades, ataques y características asociadas a los mismos.

Por ello, se busca mostrar la gran importancia y sofisticación de estos dibujos respondiendo a las siguientes cuestiones:

1. ¿Cuáles son las características más importantes de los Pokémon?
2. ¿Cómo han ido evolucionando y creciendo este mundo a lo largo de los años (generaciones)?
3. ¿Cuántos pokémon legendarios y comunes existen en cada generación?
4. ¿Cuáles son las criaturas más poderosas?

### A. Fuente de datos

Los datos utilizados para la realización de este análisis se han obtenido de las siguientes fuentes: All Pokémon dataset, Egg Hatching Step Counts in Pokemon e Imágenes Pokémon. El primero de los conjuntos de datos ha sido el principal precursor del análisis a realizar dado que los campos que incluye otorga un amplio abanico de posibilidades para el estudio.

Por otra parte, es de interés incluir el segundo de los Datasets dado que con el se pueden analizar los pasos necesarios para que cada criatura eclosione en los videojuegos según las generaciones. Por último, el tercero de los conjuntos de datos permite asociar a los diferentes Pokémon a una imagen, para así otorgar un esquema más amigable y visual. No obstante, no ha sido posible encontrar un dataset con imágenes de todos los pokémon, lo cual afectará a algunas representaciones. Estos datos recolectados servirán como contraste para el estudio llevado a cabo.

## 3. DICCIONARIO DE DATOS

Se recoge a continuación un diccionario de datos asociado a las columnas de los conjuntos de datos utilizados para el análisis. Tras realizar los cambios que se explicarán posteriormente en la sección de limpieza de datos, se obtienen un total de 2 tablas y una serie de medidas DAX (sobre las que se agruparán los distintos atributos):

**Tabla All\_Pokémon:**

- **Number:** número asociado a la pokédex. Tipo: numérico.
- **Name:** nombre asociado al pokémon. Tipo: texto.
- **Type 1 y Type 2:** tipos del pokémon. Tipo: texto.
- **Abilities:** habilidades competitivas de la criatura. Tipo: texto.
- **HP, Att, Def, Spa, Spd, Spe:** stats de los pokémon asociadas a los pokémon. Tipo: numérico.
- **Generation:** generación en la que ha sido introducido. Tipo: numérico.
- **Catch Rate:** dificultad para hacerse con el Pokémon. Tipo: numérico.
- **Legendary:** indica si el pokémon es legendario o común. Tipo: texto (booleano).
- **Height/Weight:** peso y altura de la criatura. Tipo: numérico.
- **Gen.N.Steps:** pasos necesarios para eclosionar un huevo de esa criatura en cada generación. Tipo: numérico.
- **Ranking:** clasificador de los pokémon en función de su poder. Tipo: numérico.

**Tabla Pokémon:**

- **Imagen:** Imagen asociada al pokémon.

**Medidas DAX:**

- **BMI:** índice de masa corporal del pokémon calculado a partir del peso y la altura. Tipo: Numérico.
- **Legendario/No Legendario:** se emplea para contar el número de pokémon de cada tipo. Tipo: Numérico.
- **Media de pasos:** calcula, a partir de las columnas de pasos, el número de pasos medios necesarios para eclosionar un huevo. Tipo: Numérico.

## 4. PREPARACIÓN Y LIMPIEZA DE LOS DATOS

En primer lugar, se lleva a cabo una modificación del fichero csv asociado a las imágenes. En la página web se introduce un csv con los campos nombre, tipo 1 y tipo 2. Las imágenes se descargan en un fichero a parte y se considera nuestro trabajo modificar el documento para que estén introducidas en el mismo antes de trabajar con Power Bi. Para ello, se cargan todas en Dropbox, se generan los enlaces de compartición de cada una de ellas y se introducen una a una en una nueva columna del fichero (tras haber importado previamente los datos del csv a un excel). Aprovechando que ya se encuentra el excel abierto, se eliminan las columnas asociadas a los tipos (dado que estos ya se incluyen en otro de los csv a utilizar).

Una vez realizado este proceso se abre Power Bi para cargar los dos csv y el excel mediante **Obtener datos**. Entonces, se abre Power Query desde el botón de **transformar datos y se comienza a realizar cambios. TABLA POKEMON (FOTOS):**

- En esta tabla es necesario sustituir, de la columna de imágenes, la subcadena "www" por "dl" para que las imágenes funcionen correctamente.
- También es importante revisar los tipos de datos para adaptarlos según corresponda.
- Por último, se cambia el formato de la columna nombre para adaptarlo a los otros ficheros, poniendo el inicio de cada palabra en mayúsculas.

**TABLA All\_POKEMON:**

- En primer lugar se eliminan las columnas que no se utilizarán para el análisis (experience type, final ev, against (tipos), el número de puntos de experiencia totales...)

- Además también se borran otras columnas que se calcularán mediante DAX como es el caso de BMI.
- La columna **gen** viene medida en decenas por lo que eliminamos el 0 asociado a cada una de las mismas (por ejemplo, 10 pasaría a 1).
- Las habilidades de los pokémon aparecen agrupadas en un único campo. Se hace uso de la herramienta de **dividir columna** para separarlas en 3 distintas. Una vez hecho, se le aplica formato borrando los corchetes y las comillas.
- La columna **Legendary** indica por medio de un booleano si es legendario o no. Se cambiarán los valores 0 o 10 por no y yes respectivamente.

#### TABLA EGG\_FINAL:

- En primer lugar se elimina la primera columna dado que carece de interés.
- Se cambia el nombre de la columna ID por **Pokedex N.**
- se cambia la forma de representar los valores nulos de NA a null para que no ocurran errores. Finalmente se guarda y se cierra el editor de power Query.

Una vez aplicados los cambios sobre las fuentes de datos, se analiza el modelo, observando como existe una relación 1-1 entre Egg Final y All Pokémon. En este caso, se opta por combinar dichas tablas.

Para ello, se lleva a cabo la unión por número y se quitan los campos de nombre y número. A continuación, se desactiva la carga asociada a la tabla de los huevos (dado que los datos se encuentran ubicados ya en la otra tabla. Con esto, se completa la limpieza de los datos. Es importante mencionar que las tres bases de datos no poseen los mismos pokémon. Por este motivo, existen datos de algunas criaturas que no están disponibles y pueden afectar a alguna de las visualizaciones como se explicará en las siguientes secciones.

## 5. ELABORACIÓN DEL DAHBOARD 1: GENERAL

Para este informe se incluyen dos gráficas de segmentación asociadas a las habilidades de los pokémon y a la generación para poder interactuar con las gráficas mostradas.

En lo relativo a las gráficas introducidas, se busca realizar un análisis general de la información aportada por los datos recolectados. De este modo, se incluye una gráfica circular para poder observar la cantidad de pokémon legendarios/comunes existentes.

Además, se introducen dos gráficas apiladas para comparar variables como la salud , el peso y el índice de masa corporal por cada tipo básico. Así, se observa cómo por ejemplo los pokemon de tipo acero o roca son los más pesados (lo cual encaja a la perfección con la vida real).

Para concluir, se introduce una gráfica para observar el cambio en el número de habilidades a medida que pasan las generaciones.

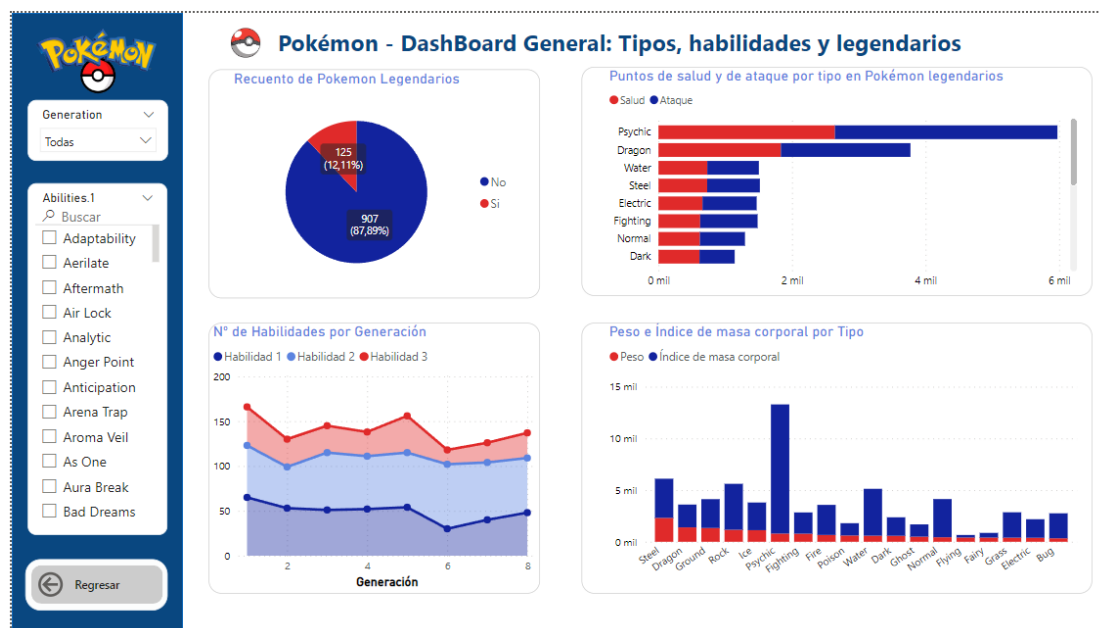


Figura 1: Dashboard 1: informe general

## 6. DASHBOARD 2: ANÁLISIS GENERACIONAL

El segundo de los esquemas realizados busca realizar un análisis generacional asociado a los pokémon. De este modo, se introduce una gráfica de segmentación para poder filtrar por generación. Además, también se incluyen segmentos de datos para los tipos y habilidades ya que pueden resultar útiles para determinados análisis.

En lo relativo a la sección de gráficas, se incluye una circular para mostrar la cantidad de criaturas introducidas en cada generación. Además, en relación con la anterior representación, se realiza una gráfica de líneas para comprobar el número de criaturas legendarias asociadas a cada fase.

Se introducen también dos gráficas de tarjetas para mostrar de forma más visual el número de criaturas legendarias y normales existentes (es de interés combinar esta visualización con los segmentos de datos).

Por último, se añade una tabla con información asociada al nombre, generación tipo(s) y masa corporal de las criaturas.

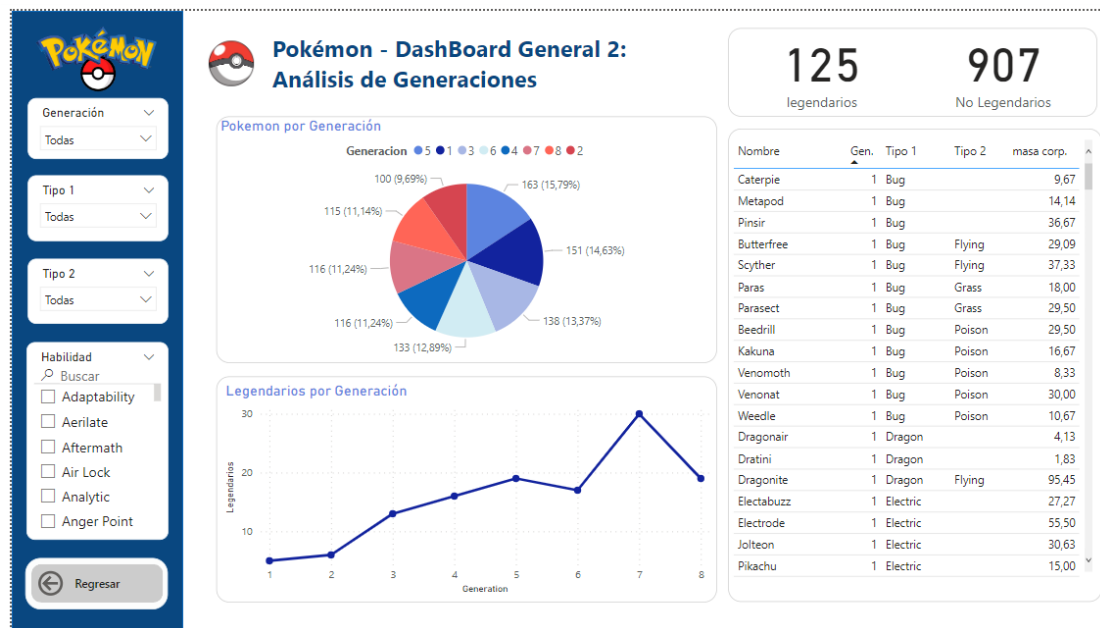


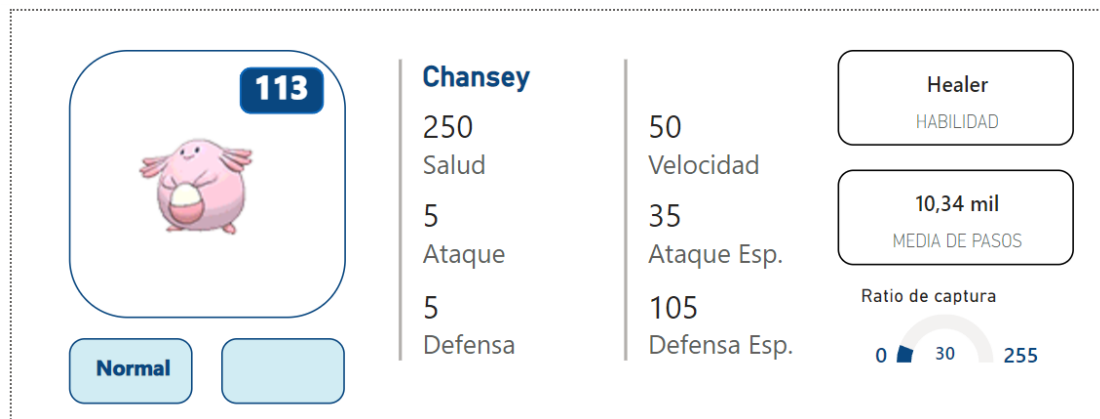
Figura 2: Dashboard 2: análisis generacional

## 7. DASHBOARD 3: RANKING

Además, se lleva a cabo un ranking para evaluar el nivel de poder de los pokémon. En este caso, se observa como la falta de imágenes para todas las criaturas afecta negativamente a las visualizaciones. Así, en la tabla incluida en el lateral derecho, se evaden las entradas con pokémon sin fotografía y en la representación de los primeros (Dado que son formas especiales de las criaturas) se muestran las características de las diferentes especies pero no las imágenes.

Las gráficas seleccionadas para esta representación buscan realizar un ranking similar al creado en la última práctica de la asignatura. Así, se crea una composición (que posteriormente servirá como plantilla) con las siguientes características:

1. Imagen del pokémon
2. Tipo(s): en caso de que solo disponga de uno, uno de los cuadrados aparecerá vacío.
3. Nombre y stats en dos gráficas de tarjetas con varias entradas.
4. Habilidad principal.
5. Tarjeta con la media de pasos necesarios para eclosionar los respectivos huevos.
6. Gráfica asociada al ratio de captura del pokémon.



**Figura 3:** Tarjeta con información general

De esta forma, se crean dos hojas en power BI: una para los primeros y otra para los últimos. En ambas se realizan las mismas gráficas pero se modifican los filtros empleados para transmitir información diferente.

Se incluye una tabla con el puesto que ocupa en el ranking, la imagen del pokémon y su nombre. Esta se configura para que, al pasar por encima con el ratón se muestre una tarjeta como la incluida en la figura anterior con la información específica del pokémon asociado a la entrada de la celda. Dicha tabla se configura ordenándola en función del valor del ranking (de mayor a menor o viceversa en función de si se trata de la hoja de primeros o últimos). Además, se define un filtro para excluir entradas repetidas (dado que los 3 valores extremos, mayores y menores, se incluirán en tarjetas independientes en grande).

Como acaba de mencionarse, se introducen 3 grupos de gráficas usando la platilla, cada una con el número que ocupa en el ranking) para representar las características asociadas a las mejores y peores criaturas. Así, es necesario definir un filtro para cada una de las gráficas presentes en la representación.

Por último, se incluyen una serie de gráficas de segmentación (las cuales se definen para que solo afecten a la tabla y no al resto de grupos de gráficas. Además se definen marcadores para cambiar entre las páginas de primeros y últimos del ranking.



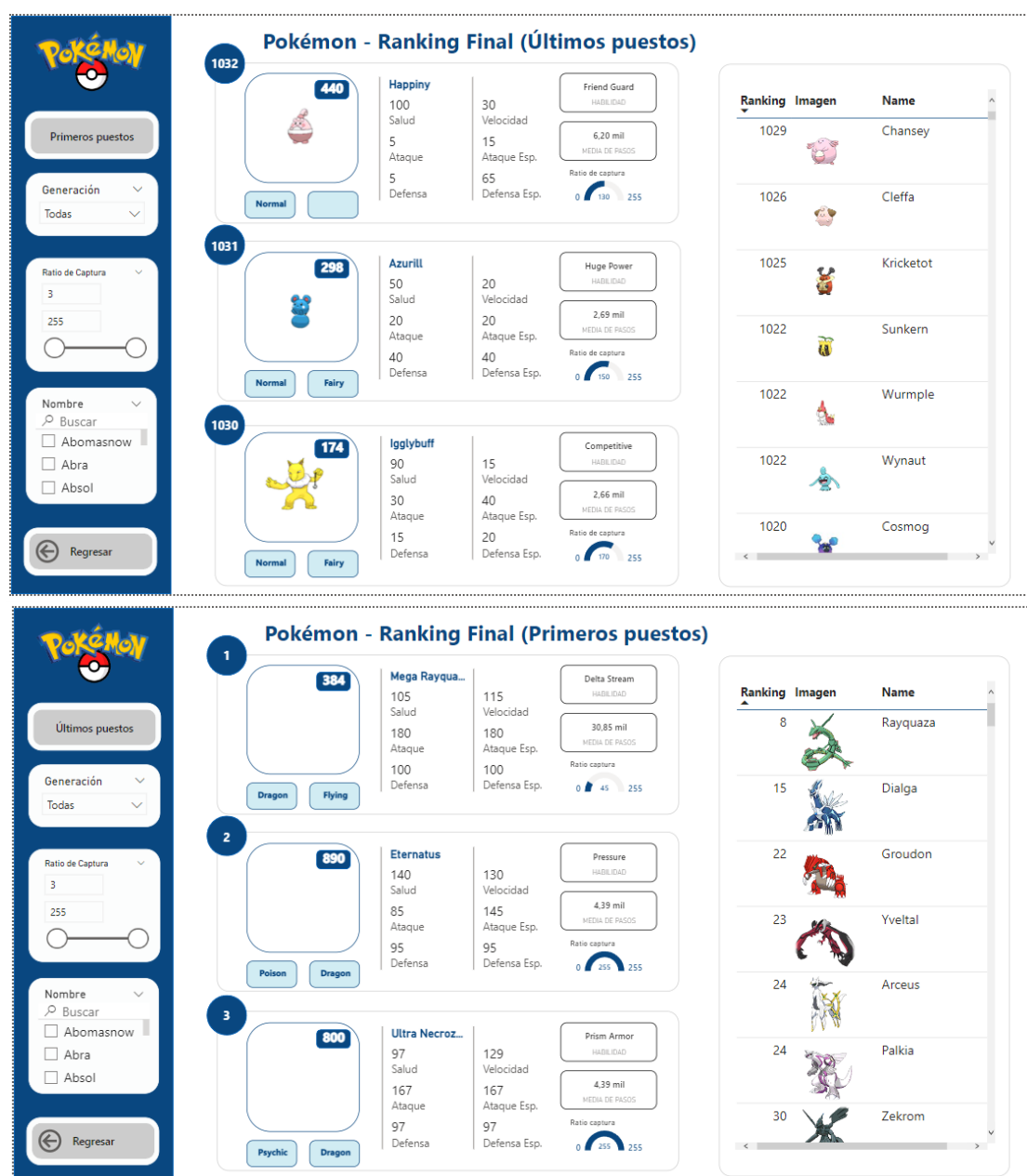


Figura 4: Dashboard: ranking pokémon

## 8. MEDIDAS DAX ELABORADAS

Para poder realizar algunas de las representaciones anteriormente incluidas, ha sido necesario definir una serie de medidas DAX. Estas han sido:

- **BMI:**  $BMI = \text{SUMX}('All\_Pokemon', 'All\_Pokemon'[Weight] / 'All\_Pokemon'[Height])$ . Esta medida busca calcular el índice de masa corporal de los pokémon en función de las columnas altura y peso.
- **Legendario:** `legendarios = CALCULATE ( COUNT(All_Pokemon[Legendary]), FILTER (All_Pokemon, All_Pokemon[Legendary] = "Si"))`. Esta medida busca contar el número de

pokémon legendarios haciendo uso de la función filter y contando los datos de la columna legendary cuyo valor sea si.

- **No legendario:** No Legendarios = CALCULATE ( COUNT(All\_Pokemon[Legendary]), FILTER (All\_Pokemon,All\_Pokemon[Legendary] = "No")). Esta medida se calcula siguiendo la misma lógica que para el caso anterior.
- **Media de pasos:** Este cálculo se basa en definir la media entre el número de pasos necesarios para eclosionar un huevo en cada generación. Como no todos los pokemon existen en todas las generaciones, es necesario realizar la cuenta de los huecos en blanco (es decir, sin datos) para tenerlo en cuenta a la hora de llevar a cabo la división asociada a la media.

## 9. ANÁLISIS DE LOS RESULTADOS OBSERVADOS

Para justificar los resultados observados, se utiliza como punto de partida las preguntas planteadas al inicio del informe. Así, empleando como material de apoyo los dashboard elaborados, se recogen los siguientes resultados:

### 1. ¿Cuáles son las características más importantes de los Pokémon?

Para dar respuesta a esta pregunta se hace uso del primer Dashboard en el cual se muestran gráficas asociadas a las diferentes características de los Pokémon. Se observa que cada Pokémon es único y que las características más importantes de cada uno son distintas en función de las características de las criaturas. Se observa como, para las diferentes generaciones, el número de habilidades ha ido oscilando de forma irregular, cambiando de esta forma la importancia que se le da a esta característica.

Por otra parte, es importante tener en cuenta las características fisionómicas de los Pokémon, pues algunos tipos dependen considerablemente de estas. Por ejemplo, se observa como el índice de masa corporal es notablemente alto en los Pokémon de tipo psíquico.

Además, se aprecia la relevancia de otras características como los puntos de salud o el ataque asociado a los Pokémon.

### 2. ¿Cómo han ido evolucionando y creciendo este mundo a lo largo de los años (generaciones)?

La respuesta a esta pregunta es fácilmente contestable con ayuda del segundo de los Dashboards. En él, es posible observar el crecimiento en el número de Pokémon, las criaturas asociadas a cada una de las generaciones (tanto los datos como las cantidades), la cantidad de criaturas legendarias incluidas o el porcentaje de individuos introducidos en cada entrega.

Así, se observa como la tendencia asociada a esta franquicia se encuentra en auge y cada vez son más los sprites creados, pudiendo llegar a alcanzar los 1000 en poco tiempo.

### 3. ¿Cuántos Pokémon legendarios y comunes existen en cada generación?

Para dar respuesta a esta pregunta se hace uso del mismo dashboard que para la pregunta anterior. En él es posible filtrar por generaciones y, gracias a los gráficos de tarjetas incluidos, es fácil observar el número de Pokémon asociado a cada generación de cada tipo. Por ejemplo, en la primera generación se incluyen 146 Pokémon no legendarios y 5 legendarios mientras que, por ejemplo en la segunda, son 6 los legendarios y 94 los no legendarios.

**4. ¿Cuáles son las criaturas más poderosas?**

La respuesta a esta pregunta se obtiene a raíz del último dashboard; pudiendo observar en el ranking fácilmente los pokémon con mejores características y aquellos con peores propiedades. Además, es posible apreciar un informe detallado para cada una de las mismas pasando con el ratón por encima. Se concluye así que los pokémon legendarios son aquellas criaturas con niveles de poder mayores.

## 10. CONCLUSIONES

Atendiendo a los resultados de los análisis realizados, se observa cómo la industria de los pokémon se encuentra en auge; llegando a formar su propio mundo e introduciendo numerosas características que poco a poco conforman nuevas maneras de jugar para poder crear numerosas estrategias. Además, con la realización de este informe se pudo ver cómo todas las propiedades están interconectadas entre sí, y es fácil elaborar estudios y gráficas para analizar diversas características (tal y como se ha ido narrando a lo largo del informe).

Así, es posible observar por ejemplo la relación entre el poder de los pokémon y su característica de legendario o no, la relación entre los tipos y las características fisionómicas de las criaturas...