

# **Diplomatura en Análisis de Datos, Aprendizaje Automático y sus Aplicaciones**

## **Módulo 1: Análisis y Visualización de Datos**



## **Laboratorio Nro. 2: Comunicación de Resultados.**

### **Docentes:**

**María Soledad Palacios** (soledad.palacios@gmail.com)

**Milagro Teruel** (mteruel@unc.edu.ar)

### **Autores:**

**Mariano Martín Gualpa** (mgualpa@gmail.com)

**Laercio Simoes** (laercio@hpcbrasil.com)

# Laboratorio 2

## Comunicación de resultados

Inspiración: Notas de Five thirty eight ([Ejemplo](#))

A partir de los resultados obtenidos en el laboratorio anterior, diagramar una comunicación en formato textual o interactivo describiendo un aspecto en particular de su dataset. Ejemplo:

- Variables más correlacionadas y posibles interpretaciones. Analizar la independencia de variables utilizando el test de chi-cuadrado para variables categóricas.
- Distribución de los ejemplos con respecto a una clase
- Análisis de outliers

La comunicación debe estar apuntada a un público técnico pero sin conocimiento del tema particular, como por ejemplo, sus compañeros de clase. Si quieren “simular” otra situación como una presentación a posibles clientes/inversionistas, incluyen en un documento aparte una descripción de la audiencia que quieren alcanzar.

Se evaluarán los siguientes aspectos:

- El informe debe contener un mensaje claro y presentado de forma concisa.
- Los gráficos deben aplicar los conceptos de percepción visual vistos en clase.
- Se debe describir o estimar la significancia estadística de su trabajo.

## Entrega

La entrega debe realizarse antes del **04-05**, cuando comienzan con otra materia.

En el repositorio de su grupo deben subir un archivo .html, .pdf, .jpg o .md con el informe. Deben incluir en el análisis todas las imágenes relevantes.

Pueden utilizar cualquier herramienta para generar el informe.

## Implementación

Si deciden implementar los test de chi-cuadrado, envíen la notebook que utilizan también. Tengan en cuenta que lo más recomendable es utilizar el laboratorio 1 como punto de partida.

# Características, Caracteres y “Characters”

Por: *Mariano Martín Gualpa (mgualpa@gmail.com)*  
*Laercio Simoes (laercio@hpcbrasil.com)*

Nadie desconoce hoy en día que la industria del entretenimiento mueve grandes fortunas todos los años [1]. Hoy en día, la multimillonaria industria de Hollywood se ha volcado de lleno a superproducciones cinematográficas[2][3], series televisivas y mas recientemente videojuegos basadas en personajes historietas (“comics”), cosa que no ha pasado sin críticas [4] pese a los grande réditos económicos que genera. La receta parece razonable desde el punto de vista económico: ¿por que arriesgarse con nuevas historias y personajes en un mundo en el que el público masivo parece aceptar la lógica de Gran Hermano? Este tipo de programación se convirtió en una usina de personajes que parecen no justificar su presencia en espectáculos posteriores mas allá que por el hecho de ser conocidos. Quizás sea esto mismo, historias basadas en personajes de acción conocidos sea uno de los puntos atractivos que ve la industria.

Los principales jugadores a nivel mundial son las editoriales DC Comics y Marvel. Fundadas en 1934 y 1939, se han convertido en dos usinas de creación de éxitos tan memorables como Batman, Superman y El Capitán América y tantos otros. Acumulan en la actualidad **6.896** y **16.376 personajes** respectivamente [\*] (considerando los diferentes universos que plantean e incluyendo también a los secundarios). El hecho de que sus mismos personajes sean utilizados en distintos productos de entretenimiento (cine, televisión, videojuegos y revistas) las ha convertido en grandes franquicias. Pero no es el único aspecto de interés que ha generado esta difusión de contenido multimedia, pues su forma de relato ha adoptado el tipo de **narrativa transmedia**, en el que la historia se despliega a través de múltiples medios y plataformas que van contando partes específicas y complementarias de la historia [5].

Joseph Campbell, antropólogo y mitólogo estadounidense, acuño el término de monomito o “periplo del héroe”, que resume como “El héroe se lanza a la aventura desde su mundo cotidiano a regiones de maravillas sobrenaturales; el héroe tropieza con fuerzas fabulosas y acaba obteniendo una victoria decisiva; el héroe regresa de esta misteriosa aventura con el poder de otorgar favores a sus semejantes”, agregando: “Ya sea el héroe ridículo o sublime, griego o bárbaro, gentil o judío, su aventura varía poco en cuanto al plan esencial” [6]. Desde esta perspectiva, podría considerarse que la estructura general de alguna manera se repite, teniendo necesariamente que complementar la narrativa aspectos específicos que dan marco a la historia, así como los personajes y su desarrollo. Esto da la idea de que si bien es esencial una buena historia como argumento para sostener su éxito, también es cierto que mucho de este se sustenta en la riqueza de sus personajes. Mucho se ha estudiado respecto al diseño de personajes [7], existiendo guías no solo para estudiar su diseño visual, sino también para ayudar a dotarlo de una personalidad que enriquezca la narrativa.

Dentro de la categoría, los personajes que dominan la atención por las escenas de acción suelen ser los héroes. Sin embargo, para que el héroe brille, necesita villanos que estén a la altura. Estos, no siempre tienen una estética, físico o capacidades equivalentes al héroe para las escenas de acción, pero casi siempre la carga dramática que deben tener implica una profundidad del personaje, que explica que en la mayoría de las superproducciones la estrella principal sea el villano (como botón de muestra

---

\* Considerando la totalidad de la información disponible en el dataset.

basta con pensar en la sucesión de grandes actores que han interpretado al Guasón). Resulta interesante conocer un poco mas las características que podemos encontrar en los personajes de las franquicias mas exitosas, con la expectativa de reconocer también algunos elementos que nos ayuden a conocer su carácter de héroe o villano.

## La novedad según pasan los años...

Se estudiaron los datasets sobre comics publicados en Fivethirtyeight Dataset [<sup>8</sup>], donde podemos encontrar dos archivos independientes con el registro de personajes de DC Comics y Marvel respectivamente.

Como el elemento central de análisis son los personajes, primero estudiaremos la evolución de creación de nuevos personajes por año. La gráfica (en azul) permite apreciar como la cantidad de nuevos personajes fue aumentando hasta principios de los 90, momento a partir del cual comienza a estancarse (podría estar relacionado a una baja en el interés por la lectura, en que aparece mayor acceso a televisión e internet). La suba de los últimos años podría estar relacionada a la incipiente tendencia de Hollywood en hacer películas basadas en comics.

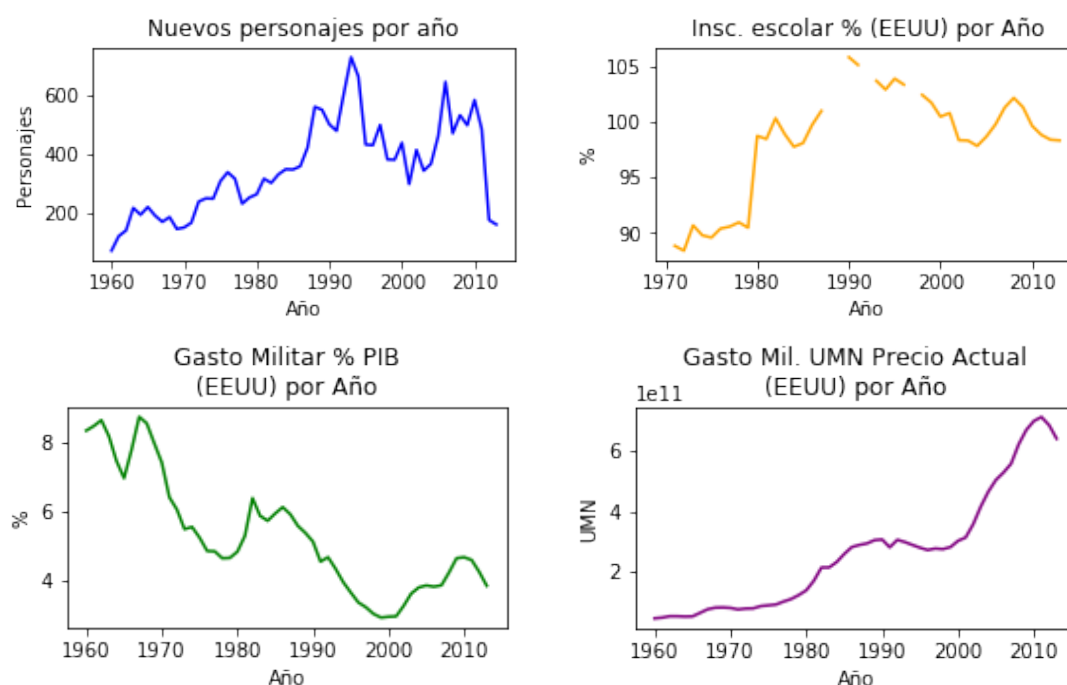


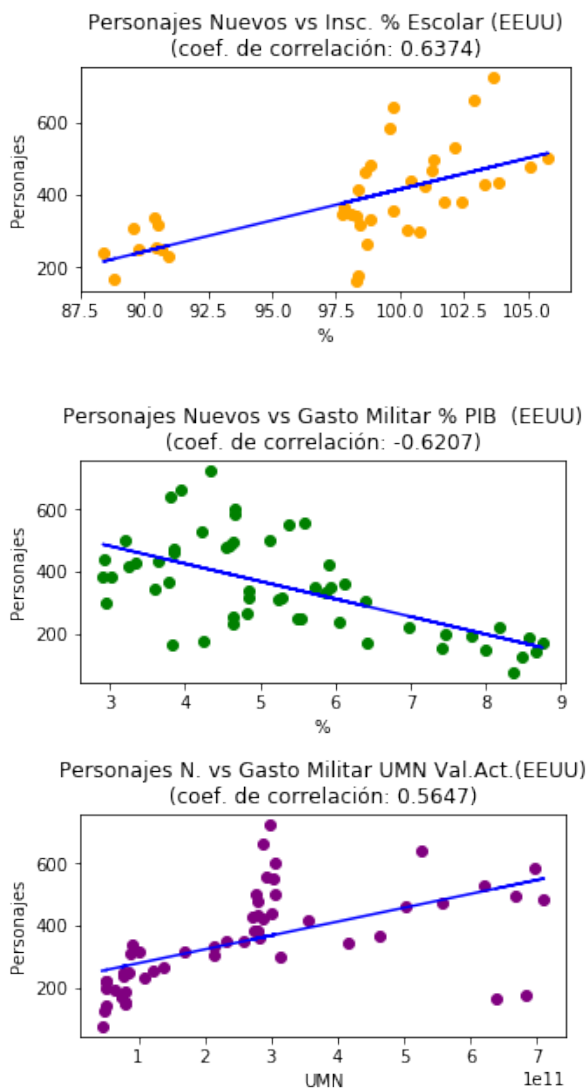
Figura 1. Evolución de las variables a lo largo del tiempo.

Resulta interesante cruzar esta información con la provista por el Banco Mundial respecto a indicadores de Estados Unidos. Los mas interesantes (exceptuando algunos descartados por tener pocas observaciones) fueron: Inscripción Escolar Porcentual (en naranja), Gasto Militar Porcentual Respecto al PIB (en verde), Gasto Militar UMN a Precio Actual (en violeta).

Resulta interesante encontrarse con que efectivamente existe una correlación entre estas variables y la principal (nuevos personajes por año).

En la siguiente imagen se presenta la cantidad de personajes correlacionada con cada una de estas variables correlacionadas. Es interesante encontrar que existe algún nivel

de correlación, como puede intuirse por la dispersión de las observaciones y verificarse por el coeficiente de correlación (indicado en el título de cada gráfica). Además, se ha agregado una recta con la regresión lineal utilizada como referencia.



Figuras 2, 3 y 4. Correlaciones de cantidad de nuevos personajes respecto a variables económicas de EEUU

Se puede observar que existe una correlación media (0,6374) entre el porcentaje de inscripción escolar en y la cantidad de nuevos personajes (Figura 2). Para establecer una hipótesis, primero es interesante observar el vacío existente entre 91% y 97,5%, lo cual se explica por el salto que da la variable inscripción escolar a inicios de los años 80. Es decir, seguramente existieron cambios importantes de política educativa que produjeron un incremento sustancial en muy poco tiempo. La consecuencia de esto es que la nube de puntos tiene dos agrupaciones. Como se observa, a medida que aumenta la escolaridad existe un incremento en la cantidad de personajes. Tiene sentido, pues mientras mas personas tienen la posibilidad de leer, crece la audiencia potencial de los comics. Por otro lado, existe una correlación negativa (-0,6207) entre personajes nuevos y gasto militar medido como porcentaje del PIB (Figura 3). Observando las series temporales, se observa que el mismo baja a lo largo del tiempo (en contraposición de la variable personajes nuevos). Debe considerarse que el PIB de Estados Unidos es una variable que ha crecido sostenidamente en el período

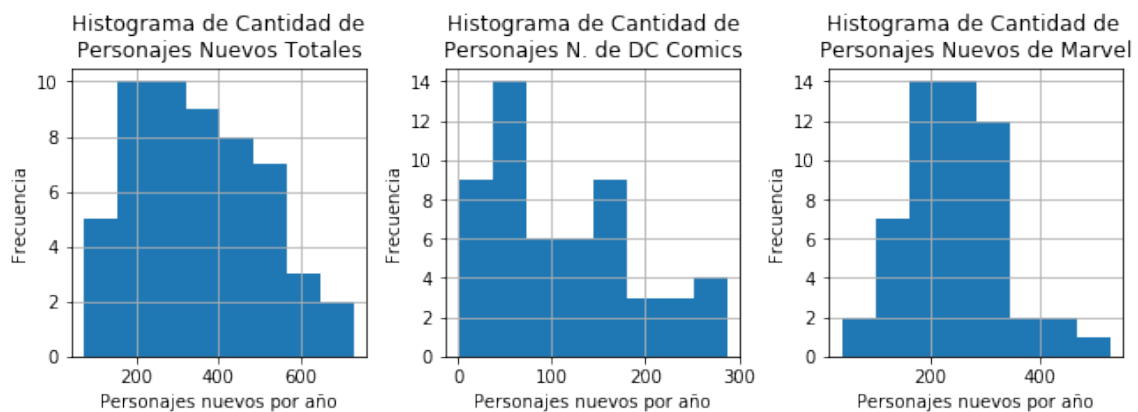
de tiempo bajo análisis. Esto explica por que comparando las dos series de gasto militar, aunque el porcentaje del PIB destinado decrece, el monto en UMN (Unidades Monetarias) crece. De esta manera, y en relación al estudio de los personajes nuevos, se observa que la correlación negativa de la figura 2 podría estar relacionada con que mayor PIB general puede indicar mayor bienestar de la economía y con ello mayor consumo de comics y productos relacionados. La correlación entre gasto militar en UMN y personajes nuevos es menor (0,5647), probablemente influido por los últimos años en que por conflictos de ese país en el mundo el gasto militar creció significativamente, cambio que no fue acompañado por la cantidad de personajes que además decreció en eso años.

Como los datos analizados se están presentando agregados, sería interesante observar si las diferentes editoriales han tenido distribuciones similares de personajes nuevos a lo largo del tiempo.



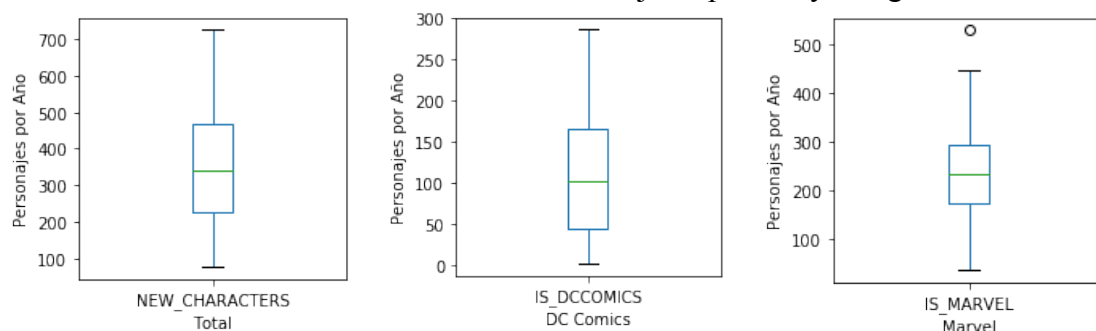
Figura 5. Cantidad de personajes nuevos desglosados por editorial.

La cantidad de personajes de ambas editoriales parecen tener una evolución consistente entre ellas. Por otro lado, es interesante analizar este aspecto desde la distribución de frecuencias de cada una de estas series. Para ello, se presentan los histogramas de las figuras 6, 7 y 8.



Figuras 6, 7 y 8. Histograma de personajes nuevos por año.

Puede observarse que en todos los casos podríamos asumir una distribución de probabilidad del tipo Poisson. Esto es consistente pues estamos midiendo cantidad de eventos por unidad de tiempo (personajes nuevos por año). Debería no obstante realizarse un test de Chi-cuadrado de bondad de ajuste para mayor seguridad.



Figuras 9, 10 y 11. Gráficos de cajas para cantidad de nuevos personajes por año (totales y por editorial). Puede observarse en la figura 11 (Marvel) outliers.

Estas mismas distribuciones pueden observarse desde la perspectiva que nos provee el gráfico de cajas. Aquí, podemos ver la distribución de los cuartiles, completamente consistente con los histogramas. Además, se puede observar en la figura 11, correspondiente a Marvel, como aparecen outliers, también completamente consistentes con la forma de la distribución de frecuencias (la probabilidad de valores altos cae abruptamente).

### Características y Carácter

Como se comentó anteriormente, resulta crucial para la historia las características no solo del héroe sino también del antagonista. En este sentido, resulto interesante también observar como es la relación de estos personajes con respecto al resto de su sociedad. Para ello, se analizaron dos variables:

- Alineación (“Align”) que se refiere a si el personajes es bueno, malo, neutral, agente doble, etc.
  - Identidad (“Id”) que se refiere a la identidad pública, secreta, desconocida, etc.
- Como se observa en la figura 12 (donde se presenta la probabilidad conjunta), existe un “área caliente” de interés en la sección buenos y malos en relación a identidad

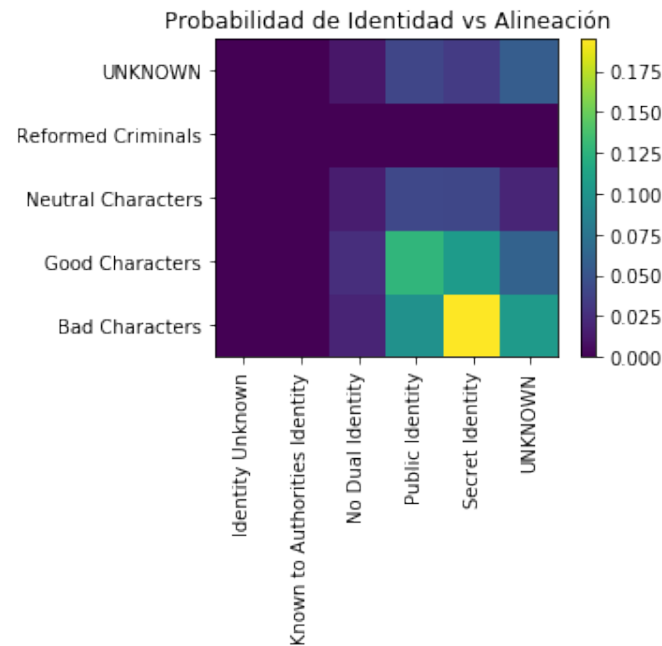


Figura 12. Los colores mas claros indican mayor probabilidad. En ellos se concentraron la mayoría de los casos.

pública y privada. La mayoría de los casos se concentran en este sector (los casos asociados a la clase UNKNOWN se refieren a casos donde el dato no estaba indicado). Existen secciones muy frías, asociados a la baja cantidad de casos, por ejemplo al no haber casi personajes de alineación “criminales reformados”, la probabilidad de cualquier situación contingente con la identidad es baja. Lo mismo puede decirse de “Identidad Desconocida” (con confundir con valores faltantes UNKNOWN).

Nos detenemos en este punto pues parece existir una relación entre la alineación y la identidad que asumen. Por ello, y para simplificar el análisis, se agrupan varias clases de cada variable a fin de despejar los casos de interés.

ID_AGREGATE	Other	Public Identity	Secret Identity	All
ALIGN_AGREGATE				
Bad Characters	0.125366	0.097371	0.194694	0.417431
Good Characters	0.086872	0.127636	0.106261	0.320770
Other	0.104559	0.082616	0.074624	0.261799
All	0.316798	0.307623	0.375579	1.000000

Tabla 1. Probabilidades conjuntas y marginales para variables estudiadas (con clases agregadas).



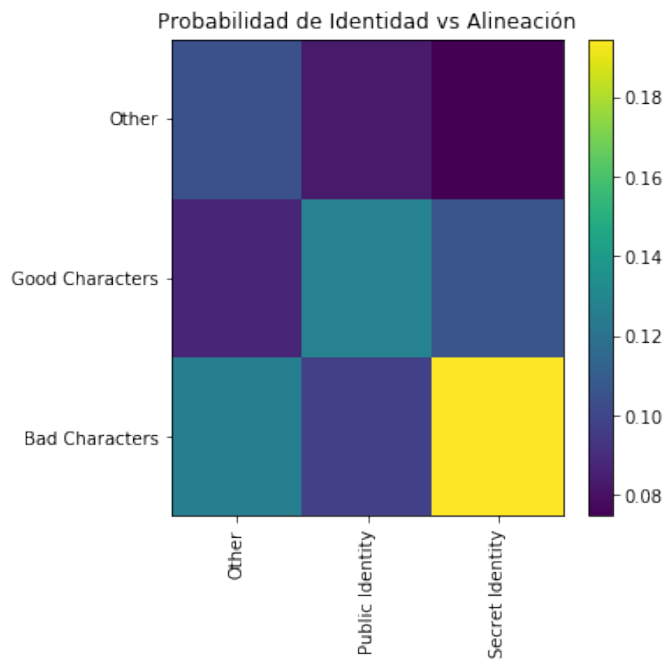


Figura 13. Probabilidad conjunta con algunas clases agrupadas.

Puede observarse una probabilidad notablemente mayor de que un personaje sea malvado y al mismo tiempo tenga identidad secreta, contra una probabilidad mucho menor de que esta sea pública. Por otro lado, los personajes buenos, tienen mayor probabilidad a tener identidad pública, aunque no tan marcada en relación a la posibilidad de que sea secreta. Para comprobar que la alineación influye sobre la tendencia a tener un tipo de identidad dato, simplemente podemos estudiar el problema desde las herramientas del cálculo de probabilidad.

Podemos observar que probabilidad conjunta (calculada a partir de las frecuencias en la tabla de contingencia):

$$P('Bad Character', 'Secret Identity') = 0.194694$$

Además, vemos que:

$$P('Bad Character') = 0.194694 \text{ y } P('Secret Identity') = 0.156778$$

Sabemos que si:

$$P('Bad Character', 'Secret Identity') = P('Bad Character') \cdot P('Secret Identity')$$

entonces podemos considerarlas independientes. Pero observamos que al calcular obtenemos:

$$P('Bad Character') \cdot P('Secret Identity') = 0.156778$$

Que es distinto al  $P('Bad Character', 'Secret Identity') = 0.194694$  **por lo que podemos inferir que no existe independencia entre ambas variables.**

**Sería interesante observar si esta distribución, es consistente entre los personajes de ambas editoriales.** Para ello, simplificaremos aún mas las tablas quedándonos solamente con los valores que queremos analizar. En este caso, las alineaciones bueno y malvado y las identidades secreta y pública. Nos centraremos únicamente en la forma en que se distribuyen las probabilidades para estas clases.

Tomando a Marvel como referencia, intentaremos establecer si los personajes de DC Comics tiene la misma distribución. Es decir, la hipótesis nula  $H_0$  es que DC Comics tiene la misma distribución que Marvel respecto a identidades (secreta y pública) de los personajes buenos y malos. La hipótesis alternativa  $H_1$  es que tienen diferentes



distribuciones.

Para ello, deberemos calcular la frecuencia esperada en DC Comics. Se calcula primero la probabilidad esperada a partir de las observaciones de Marvel. Con ellas, se calculan las frecuencias esperadas de los personajes de DC Comics, que será comparada con la observada. Esta información se resume en las tablas 2 y 3.

	Public Identity	Secret Identity	_All
Bad Characters	696	1187	1883
Good Characters	1224	809	2033
_All	1920	1996	3916

Tabla 2. Frecuencia observada de personajes DC Comic.

	Public Identity	Secret Identity	_All
Bad Characters	740.703303	1592.267555	2332.970858
Good Characters	801.568138	781.461005	1583.029142
_All	1542.271440	2373.728560	3916.000000

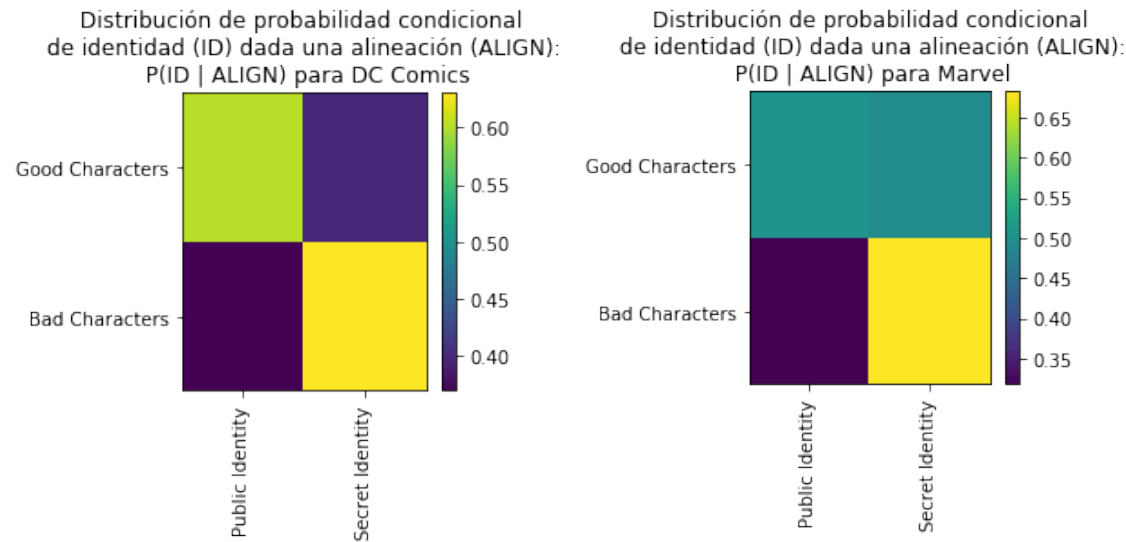
Tabla 3. Frecuencia esperada de personajes DC Comic.

Calculamos el estimados de la prueba de Chi-cuadrado, que da un resultado de 329,442525

Aplicando un nivel de confianza de 0,95 y 1 grado de libertad, obtenemos un valor crítico de 3,841458.

**Podemos observar que el estimador de la prueba de Chi-Cuadrado da un valor muy por arriba del valor crítico. Podemos rechazar la hipótesis nula que dice que ambas distribuciones son similares y aceptar la hipótesis alternativa que dice que**

**ambas distribuciones pueden ser consideradas diferentes.** A continuación, calcularemos las distribuciones de **probabilidad condicional** de la identidad ("ID"), dada una alineación ("ALIGN"). Con ella, podremos observar en mejor detalle estas diferencias.



Figuras 14 y 15. Probabilidades **condicionales** de una identidad secreta o pública dada la alineación bueno o malvado, para personajes de DC Comics y Marvel. Se observa como para alineación "Bad Character" las distribuciones se parecen pero hay mayor diferencia entre las distribuciones asociadas a "Good Characters".

El resultado es muy interesante. Analizando primero el caso de los personajes malvados: Puede observarse que se distribuyen de manera parecida (en ambos casos

aproximadamente un tercio de los casos tienen identidad pública y 2/3 identidad secreta), pero con una tendencia en Marvel a mayor probabilidad de tener identidad secreta.

Por otro lado, en el caso de los buenos: Puede observarse que mientras que los buenos en Marvel tienden a tener igual probabilidad de tener identidad secreta o pública, en DC Comics hay mayor probabilidad (0,20 mas) de tener identidad pública que identidad secreta.

	Public Identity	Secret Identity	_All		Public Identity	Secret Identity	_All
Bad Characters	0.369623	0.630377	1.0	Bad Characters	0.317494	0.682506	1.0
Good Characters	0.602066	0.397934	1.0	Good Characters	0.506351	0.493649	1.0
_All	0.490296	0.509704	1.0	_All	0.393838	0.606162	1.0

Tablas 4 y 5. Probabilidades **condicionales** de una identidad secreta o pública dada la alineación bueno o malvado, para personajes de DC Comics y Marvel respectivamente.

En conclusión: **Se observa que la alineación de los personajes (bueno o malo) afecta la distribución de probabilidad de asumir una identidad secreta o pública. En general, los malvados tienen el doble de probabilidades de asumir una identidad secreta. Mientras que en los buenos, se distribuye de manera similar aunque es un poco mayor la probabilidad de que adopten identidad pública.** Esto tiene sentido pues es lógico que los personajes malvados (tanto los líderes como sus esbirros) al actuar en los mundos de fantasía desde la ilegalidad, requieren proteger su identidad o evitar que se les sospeche la responsabilidad de los crímenes. Por otro lado, si bien existen muchos superhéroes con identidad secreta, suelen tener personajes aliados de identidad pública.

**Si analizamos estas probabilidades desagregando por publicadoras, los personajes malvados mantienen la tendencia de sostener una identidad secreta (un poco mas en Marvel que en DC Comics). Sin embargo, los personajes buenos de Marvel pueden tener identidad secreta o pública con prácticamente igual probabilidad, mientras que en DC Comics los personajes buenos tienen mayor tendencia a tener identidad pública.**

## Carácter y “Character”: ¿Constituyen los pelados un peligro para el Universo Marvel?

A partir de esta hipótesis, se analizaron los datos disponibles. Lamentablemente, la información del dataset correspondiente a DC Comics no contiene la categoría "Bald" o "No Hair" y en su lugar dejan a estos personajes con un valor NaN (pero que impide diferenciar los casos en que simplemente no existe dicha información). Por ello, este análisis solo se hará en este caso con la información disponible de personajes del Universo Marvel. Dada la gran cantidad de categorías, se agruparon todos los “no pelados” en una sola.

	ALIGN	Bad Characters	Good Characters	Neutral Characters	UNKNOWN	All
HAIR_AGGREGATE						
Bald		1137	329	331	217	2014
Other		5583	4307	1877	2595	14362
All		6720	4636	2208	2812	16376

Tablas 6. Tabla de contingencia con frecuencia de personajes de Marvel.

Ya se puede observar sobre la tabla 6 una distribución de frecuencias interesante para los personajes pelados. Para seguir analizando, se agruparon todas las categorías de alineación (excepto “Bad Character”).

ALIGN_AGGREGATE	Bad Characters	Other	All
HAIR_AGGREGATE			
Bald	1137	877	2014
Other	5583	8779	14362
All	6720	9656	16376

Tablas 7. Tabla de contingencia con frecuencia de personajes de Marvel discriminando únicamente las clases Bald (pelado) y Bad Character (malvado).

La observación sobre la tabla 7 da la idea de que no se distribuyen uniformemente y que existiría una relación entre las variables. Nuestra hipótesis es que la característica de tener alineación con los malos ("Bad Character") está mas presente entre los pelados que entre el resto de los personajes. Es decir: La hipótesis nula diría que personajes de la categoría Bald se distribuyen de la misma manera que Others. La hipótesis alternativa dice que los personajes de la categoría Bald no se distribuyen de la misma manera que Others. Es importante aclarar que al agregar de esta manera las clases, un personaje que no esté en la categoría “malvado” (“Bad Character”) no es necesariamente “bueno”, sino que podría ser en su lugar neutral, agente doble, etc.

Una forma de comprobar la hipótesis nula es aplicando el test de Chi-Cuadrado. Para ello debemos construir una tabla de frecuencias esperadas, basada en la probabilidad teórica de la hipótesis nula. El resultado del test da un valor de 225,643389 que comparado con el valor crítico de 3,841458 (correspondiente a una confianza del 0,95 y 1 grado de libertad), por lo que podemos rechazar la hipótesis nula (que en universo marvel los pelados se distribuyen entre las categorías Bad Character (malvado) y Others siguiendo la misma distribución que los otros tipos de personajes). Entonces, puede aceptarse la hipótesis alternativa, que dice que tiene una distribución diferente. Otra forma de confirmar esta hipótesis es calculando las probabilidades conjuntas:

$$P(\text{Bald, Bad Character}) = 0.069430$$

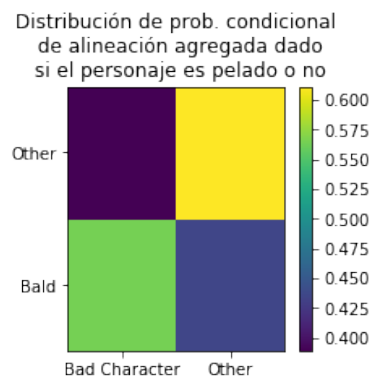
$$P(\text{Bald}) \cdot P(\text{Bad Character}) = 0.122984 \cdot 0.410356 = 0.050467$$

De ser variables independientes, ambas deberían dar el mismo valor. Esta diferencia también muestra que no existe independencia entre las variables, pues una está condicionada por la otra.

Podemos observar esta situación calculando las probabilidades condicionales dado el atributo HAIR\_AGGREGATE:

ALIGN_AGGREGATE	Bad Characters	Other	All
HAIR_AGGREGATE			
Bald	0.564548	0.435452	1.0
Other	0.388734	0.611266	1.0
All	0.410357	0.589643	1.0

Tablas 8. Tabla de probabilidades condicionales de Alineación Agregada (que sea malvado o no) para personajes de Marvel, dado si el personaje es Bald (pelado) o no.



Figuras 16. Probabilidad conjunta para el universo Marvel, de que un personaje que sea malvado o no, según si es pelado o no.

Podemos observar que dentro de la población seleccionada, dado un personaje que sea pelado, se tiene una menor probabilidad de que pertenezca a la alineación Others (formado por buenos, neutrales y desconocidos) y existe mayor probabilidad de que sea un malvado que quiera dominar el mundo, o al menos el Universo Marvel.

Como conclusiones generales, la mas interesante a destacar a partir de las series analizadas es que mas allá de la creatividad y singularidad de cada caso, la creación de **personajes que ayuden a enriquecer las distintas historias en las dos editoriales mas importantes del mundo, parecen cumplir con determinadas reglas y proporciones**. Sería interesante poder estudiar estos aspectos con mayor profundidad, y disponiendo de mas información (un atributo que hubiese sido interesante de tener es la serie a la que pertenece cada personaje, mas allá de los crossover de los últimos tiempos, y el nivel de éxito de la misma). **Esta información podría ayudar a detectar estrategias o “recetas” exitosas compartidas para el diseño de las características de personajes de ficción.**

## Referencias

<sup>1</sup> Las industrias que mas dinero mueven en el Mundo. El Economista. Fuente:

<http://listas.eleconomista.es/economia/240-las-industrias-que-ms-dinero-mueven-en-el-mundo>

<sup>2</sup> Hollywood, obsesionado con los cómics . La Nación. Fuente: <https://www.lanacion.com.ar/1746879-hollywood-obsesionado-con-los-comics>

<sup>3</sup> List of films based on English-language comics . Wikipedia. Fuente:

[https://en.wikipedia.org/wiki/List\\_of\\_films\\_based\\_on\\_English-language\\_comics](https://en.wikipedia.org/wiki/List_of_films_based_on_English-language_comics)

<sup>4</sup> Los superhéroes ¿son buenos o malos? Hollywood, dividido por el rumbo de la industry. El Día.

Fuente: <https://www.eldia.com/nota/2018-1-8-4-48-33-los-superheroes-son-buenos-o-malos-hollywood-dividido-por-el-rumbo-de-la-industria-espectaculos>

<sup>5</sup> Qué es transmedia y storytelling?. Agencia Best.

Fuente: <https://agencia.best/blog/transmedia-y-storytelling>

<sup>6</sup> Monomito. Wikipedia.

Fuente: [https://es.wikipedia.org/wiki/Monomito#cite\\_note-monomyth-website-2](https://es.wikipedia.org/wiki/Monomito#cite_note-monomyth-website-2)

<sup>7</sup> Great books for Character Design. Goodreads.

Fuente: [https://www.goodreads.com/list/show/23070.Great\\_books\\_for\\_Character\\_Design](https://www.goodreads.com/list/show/23070.Great_books_for_Character_Design)

<sup>8</sup> Fivethirtyeight Dataset. Fuente: <https://github.com/fivethirtyeight/data>