

# El éxito de *Empyrean*: nostalgia estructural y evolución emocional en el romántico contemporáneo

María Yu García Muñoz

4º Inteligencia y Analítica de Negocios (BIA)

Universidad de Valencia

---

Este trabajo analiza por qué la saga *Empyrean* (2023–2025) ha alcanzado un **éxito editorial y crítico extraordinario**, con una calificación promedio de **4,50/5** en *Goodreads* y ventas combinadas que superan los **12 millones de ejemplares** en menos de dos años. Los tres volúmenes, *Fourth Wing*, *Iron Flame* y *Onyx Storm*, han dominado las **listas de ventas de *The New York Times* y Amazon**, mientras que *Onyx Storm* se convirtió en el **novel adulto más rápido en venderse en 20 años**, con **2,7 millones de copias en su primera semana**. Además, *Fourth Wing* recibió el **Premio BookTok a Libro del Año**.

A partir de un corpus de nueve novelas, se han aplicado técnicas de **procesamiento del lenguaje natural (PLN)** para evaluar dos hipótesis: (1) que *Empyrean* presenta una narrativa más colectiva y emocional; y (2) que su éxito combina **innovación** con estructuras narrativas familiares para lectores de sagas juveniles. El análisis se organiza en cuatro bloques: (1) análisis de sentimiento con los léxicos *AFINN* y *NRC*; (2) modelización temática mediante **LDA guiado**; (3) redes de coaparición de personajes; y (4) validación con las valoraciones de *Goodreads*.

Los resultados muestran que *Empyrean* posee una red relacional más densa y un tono emocional más intenso, rasgos que no aparecen con la misma fuerza en *Shatter Me* y solo parcialmente en *The Mortal Instruments*, sugiriendo que recupera fórmulas clásicas adaptadas a una sensibilidad actual. Esto indica que su éxito no depende solo de la novedad, sino de la combinación de **estructuras consolidadas** con un enfoque **colectivo**, ofreciendo bases sólidas para sistemas de recomendación narrativa más allá del género o la popularidad.

---

# Introducción

## 1.1. El surgimiento del *romantasy* como género híbrido

La literatura juvenil ha ido cambiando según las inquietudes sociales, culturales y emocionales de cada generación. Hasta bien entrada la primera década del siglo XXI, el mercado editorial para adolescentes se centraba en dos grandes tipos de historias, distintas en estructura y temática.

Por un lado, las sagas de **fantasía épica y urbana**, como *Harry Potter* (J. K. Rowling), *Juego de Tronos* (George R. R. Martin) o *El señor de los anillos* (J.R.R. Tolkien), presentaban mundos sobrenaturales con reglas propias, profecías y grandes enfrentamientos entre el bien y el mal. En estas historias, el crecimiento del protagonista dependía de su papel dentro del conflicto colectivo, mientras que el romance, aunque presente, quedaba en segundo plano.

Por otro lado, el **romance juvenil**, representado por sagas como *Twilight* (Stephenie Meyer), *Hush, Hush* (Becca Fitzpatrick) o *La selección* (Kiera Cass), se centraba en emociones, conflictos internos y relaciones amorosas. Aquí, el mundo exterior funcionaba como escenario de la historia romántica, marcada muchas veces por peligro, ambigüedad moral y dependencia emocional. La acción era más individual que colectiva.

A partir de 2020, con el auge de **BookTok**, se consolida un nuevo tipo de narrativa que combina ambos enfoques: el **romantasy** (*romantic fantasy*). Más que una mezcla de géneros, constituye un modelo narrativo con reglas propias y una comunidad lectora muy activa. Obras como *A Court of Thorns and Roses* (Sarah J. Maas) o *The Cruel Prince* (Holly Black) integran romance y conflicto bélico: las relaciones se fortalecen en la lucha compartida y la acción colectiva se apoya en los vínculos emocionales. Este enfoque también redefine a la heroína, mostrando a un personaje activo en su propio destino y en el de su comunidad.

## 1.2. Evolución narrativa: de lo individual a lo colectivo

La evolución del arquetipo femenino en la literatura juvenil refleja no solo cambios culturales, sino también **decisiones narrativas y editoriales**. Entre 2008 y 2015, las protagonistas solían estar marcadas por aislamiento emocional, trauma y dependencia de figuras masculinas que las guiaban o protegían. Por ejemplo, *Juliette Ferrars* en *Shatter Me* construye su identidad en relación con los demás, mientras que *Clary Fray* en *The Mortal Instruments* va ganando agencia lentamente, siempre ligada a personajes masculinos clave.

En contraste, las heroínas del **romantasy** actual, como *Violet Sorrengail* en *Empyrean* o *Feyre Archeron* en *A Court of Thorns and Roses*, son activas desde el inicio: participan en decisiones, lideran equipos y su desarrollo emocional ocurre dentro de comunidades narrativas que complementan la relación romántica.

El éxito de *Empyrean* no se explica solo por este cambio: también recupera elementos clásicos de sagas anteriores, como estructuras corales, equilibrio entre romance y acción y jerarquías mágicas claras. Esta combinación de novedad emocional y estructuras familiares atrae tanto a lectores nuevos como a los habituados al género. En un mercado saturado, ser original no basta: la heroína debe influir activamente en su entorno, y *Empyrean* logra transmitir esto sin romper con las convenciones que hicieron populares a sus predecesoras.

### 1.3. Objetivos e hipótesis del estudio

El objetivo de este proyecto es **analizar cuantitativamente el éxito de *Empyrean*** mediante técnicas de procesamiento del lenguaje natural y análisis de datos no estructurados. Para ello se construye un **corpus** a partir de cuatro bases de datos, que permiten comparar distintas etapas de la literatura juvenil y del **romantasy**.

Se estudian tres sagas individualmente: *Empyrean* (2023–2025), representante del romantasy contemporáneo; *The Mortal Instruments* (2007–2014) y *Shatter Me* (2011–2014), como ejemplos de sagas anteriores donde romance y fantasía conviven pero no se integran completamente. Además, se incorpora un **dataset comparativo** con sagas populares de la década de 2010 (*Vampire Academy*, *Divergent*, *Throne of Glass*, *Fallen*, *Hush, Hush*, *Percy Jackson* y *The Maze Runner*), para situar los resultados en un contexto más amplio del mercado juvenil.

Sobre este corpus se aplican **cuatro bloques metodológicos**:

1. **Análisis de sentimiento**, con léxicos AFINN, NRC y Bing y un modelo supervisado, para comparar perfiles afectivos entre etapas.
2. **Modelización temática**, mediante *seeded LDA*, para identificar los principales tópicos narrativos y contrastarlos con la estructura de *Empyrean*.
3. **Análisis de redes de personajes**, construyendo grafos de coaparición por capítulo para medir complejidad relacional y centralidad de la protagonista.
4. **Recepción externa**, usando ratings de *Goodreads* como validación de la respuesta real de los lectores.

La **hipótesis central** es doble: que el romantasy actual ha evolucionado hacia estructuras narrativas más colectivas frente al individualismo emocional de sagas anteriores, y que el éxito de *Empyrean* radica en recombinar elementos familiares con nuevas demandas emocionales.

---

## 2. Marco teórico y metodológico

Este proyecto se apoya en un enfoque interdisciplinar que combina técnicas de lingüística computacional, análisis narrativo y ciencia de datos aplicada a textos literarios. En los siguientes apartados se presentan las principales metodologías empleadas, junto con su base teórica y su utilidad para analizar la evolución del género romántico.

### 2.1. Análisis de sentimiento en textos literarios

El análisis de sentimiento permite cuantificar la valencia emocional de un texto mediante la identificación de palabras con carga afectiva [1]. En este trabajo se emplean tres enfoques complementarios. En primer lugar, los **léxicos predefinidos**:

- **AFINN** (Nielsen, 2011), que asigna puntuaciones numéricas entre -5 (muy negativo) y +5 (muy positivo) [1];
- **NRC Emotion Lexicon** (Mohammad & Turney, 2013), que categoriza palabras en ocho emociones básicas (alegría, miedo, ira, etc.) según la rueda de Plutchik [2];
- **Bing Liu's lexicon**, que clasifica términos como positivos o negativos de forma binaria [1].

Estos métodos permiten **analizar el tono emocional promedio de cada saga** y cómo este cambia a lo largo de los capítulos. Además, se entrena un **modelo supervisado de regresión logística con regularización LASSO** usando el corpus de sagas clásicas (2008–2015), con TF-IDF como representación del texto. Luego, este modelo se aplica a *Empyrean* para determinar si su perfil emocional sigue los patrones de las sagas anteriores o introduce un cambio narrativo. La combinación de **análisis léxico y modelización predictiva** asegura una interpretación sólida y contextualizada del sentimiento en la ficción [3].

### 2.2. Modelización temática con LDA guiado

La **modelización temática** permite descubrir patrones o estructuras ocultas en grandes volúmenes de texto. El **LDA estándar** [4] funciona de manera no supervisada, es decir, identifica los temas por sí solo, pero en literatura esto puede generar resultados poco claros, puesto que los temas detectados no siempre coinciden con los ejes narrativos importantes. Por ello, en contextos literarios suele ser más útil el **LDA guiado (seeded LDA)** [5], que permite **incorporar conocimiento previo** mediante diccionarios temáticos.

En este proyecto, se construyó un diccionario específico para *The Mortal Instruments*, con cinco tópicos centrales: **romance, identidad, guerra sobrenatural, política institucional y mundo subterráneo**. Cada tema se define a partir de palabras coherentes extraídas de un análisis previo de la saga, lo que ayuda a que el modelo identifique los temas de manera más precisa y fácil de interpretar.

El modelo se aplica únicamente sobre *The Mortal Instruments*, dado que su universo narrativo está bien delimitado y sus convenciones temáticas son estables. Aunque no se aplica directamente a *Empyrean*, los resultados sirven como **referencia** para analizar la evolución temática: mientras *TMI* equilibra romance y política institucional, *Empyrean* enfatiza la **acción colectiva y la resiliencia grupal**, incluso si esos temas no se modelan explícitamente.

Este enfoque muestra que la **intervención humana en la modelización temática** es fundamental para obtener resultados interpretables en literatura, permitiendo **conectar los patrones detectados con el análisis literario** y entender mejor la evolución de los géneros narrativos.

### 2.3. Redes de personajes y análisis narrativo

Las **redes de personajes** representan la narrativa como un grafo: los **nodos** son los personajes y las **aristas** muestran cuándo aparecen juntos en un mismo contexto, como un capítulo o una escena. Siguiendo la propuesta de Agarwal et al. (2012), se construyen **redes no dirigidas** basadas en la coocurrencia de nombres propios dentro de cada capítulo. La métrica principal es la **densidad de la red**, que indica el grado de interacción entre personajes al comparar las conexiones existentes con el máximo posible. También se analiza el **número de nodos** (personajes relevantes) y la **centralidad de grado** del personaje principal, que refleja su importancia como eje de la historia. [\[6\]](#)

Este enfoque permite diferenciar **dos tipos de narrativa**: la dual, centrada en una relación binaria (protagonista–pareja), y la coral, donde múltiples personajes interactúan de forma equilibrada. Se espera que *Shatter Me* y *The Mortal Instruments* muestren redes más simples, con menor densidad y la protagonista más centralizada, mientras que *Empyrean* presente una red más conectada, donde la protagonista actúa como **nexo entre distintos subgrupos** (aliados, rivales, mentores).

De este modo, esta técnica no solo analiza la estructura formal del texto, sino que también **conecta la arquitectura relacional con el contenido ideológico**, mostrando cómo las relaciones entre personajes refuerzan los temas y mensajes de la narrativa.

### 2.4. Sistemas de recomendación basados en contenido

Los **sistemas de recomendación basados en contenido** sugieren elementos similares analizando sus características propias, sin necesidad de datos de los usuarios [\[7\]](#). En este proyecto, se construye un **perfil vectorial por saga** que combina cuatro dimensiones:

- **Sentimiento promedio** (AFINN),
- **Densidad de la red de personajes** (como indicador de colectivismo),
- **Proporción de los tópicos clave** (cuando están disponibles),
- **Rating promedio en Goodreads**.

## 2.5. Datos de recepción: Goodreads y BookTok

Finalmente, se incorporan **datos de recepción externa** para validar los resultados del análisis técnico. Los **ratings promedio en Goodreads**, una plataforma con millones de reseñas, ofrecen una medida cuantitativa del éxito real de cada saga. Aunque no se analizan reseñas individuales por limitaciones de tiempo y alcance, el rating de cada libro de cada saga sirve como **referencia empírica**: si *Empyrean* muestra una estructura narrativa y un tono emocional equilibrado, y además obtiene un rating alto, esto refuerza la hipótesis de que estas características coinciden con las expectativas de la audiencia actual. [9]

Aunque no se realiza un análisis cuantitativo de **BookTok**, se menciona como fenómeno cultural que ha impulsado el *romantasy* moderno. Su lógica, centrada en la **identificación emocional**, la comunidad lectora y la viralidad de tramas colectivas, coincide con los patrones detectados en el análisis computacional.

---

## 3. Metodología y datos

### 3.1. Selección y composición del corpus

El **corpus textual analizado** se compone de tres sagas representativas de dos momentos históricos distintos en la evolución del género juvenil:

- **Época clásica (2007–2014):** *The Mortal Instruments* (Cassandra Clare) y *Shatter Me* (Tahereh Mafi).
- **Época contemporánea (2023–2025):** *Empyrean* (Rebecca Yarros).

Esta selección responde a criterios de **representatividad cultural y comercial**. *The Mortal Instruments* se considera una de las sagas fundacionales del urban fantasy juvenil con elementos románticos, mientras que *Shatter Me* ejemplifica la tendencia post-*Twilight* hacia narrativas introspectivas centradas en el trauma individual. Por su parte, *Empyrean* representa el auge reciente del **romantasy**, un híbrido que combina fantasía épica y romance con protagonistas femeninas empoderadas, cuyo éxito masivo en plataformas como BookTok ha transformado el mercado editorial juvenil.

Además, se incluye un **corpus ampliado de 9 sagas populares publicadas entre 2008 y 2015**: *Percy Jackson & the Olympians*, *Divergent*, *Vampire Academy*, *A Court of Thorns and Roses*, *Throne of Glass*, *The Maze Runner*, *Fallen*, *Hush, Hush* y *The Mortal Instruments*. Este corpus no solo proporciona los **ratings de Goodreads** utilizados como referencia de recepción externa, sino que también sirve como **conjunto de entrenamiento para el modelo supervisado de clasificación emocional**. Al entrenar sobre este conjunto diverso de narrativas pre-romantasy, el modelo aprende patrones emocionales representativos de la era anterior a la consolidación del *romantasy* moderno.

Por ello, este enfoque permite contrastar tanto los perfiles técnicos como la percepción lectora en un contexto cultural y comercial más amplio.

### 3.2. Adquisición y estructura de los textos

Los **textos completos** de las sagas principales (*Shatter Me*, *The Mortal Instruments* y *Empyrean*) se obtuvieron en formato .txt a partir de ediciones digitales legales, respetando los derechos de autor y asegurando su uso exclusivo con fines académicos, sin distribución pública. Cada libro se procesó de manera individual, conservando **metadatos esenciales** como título, autor y número de saga [\[10\]](#)[\[11\]](#).

La **estructura jerárquica del corpus** es la siguiente:

- **Nivel 1:** Saga (3 categorías)
- **Nivel 2:** Libro (9 volúmenes en total)
- **Nivel 3:** Capítulo (unidad mínima de análisis narrativo)

Este diseño permite realizar **análisis tanto globales por saga como comparaciones detalladas por capítulo**, lo que resulta fundamental para estudiar la evolución emocional y temática a lo largo de las tramas.

### 3.3. Preprocesamiento textual

El preprocesamiento se realizó en cuatro fases [\[10\]](#)[\[11\]](#):

- Fase 1: Segmentación por capítulos
- Fase 2: Limpieza y normalización
- Fase 3: Tokenización y filtrado contextual [\[12\]](#)
- Fase 4: Validación del corpus

### 3.4. Análisis de sentimiento y emociones

El análisis de sentimiento se abordó mediante un enfoque multimétodo, combinando léxicos emocionales y modelos supervisados con el fin de capturar tanto la polaridad emocional (positiva/negativa) como la distribución de emociones específicas y su evolución a lo largo de la narrativa. El análisis se realizó a nivel de capítulo y de saga [\[12\]](#).

#### 3.4.1. Análisis de emociones con el léxico NRC

Para el análisis categórico de emociones se utilizó el léxico NRC, basado en el modelo de Plutchik, que asigna palabras a ocho emociones básicas además de polaridad positiva y negativa [\[19\]](#).

Se calcularon proporciones de cada emoción por capítulo y se agregaron por saga y por posición relativa dentro de la narrativa, lo que permitió comparar perfiles emocionales entre sagas y analizar su evolución a lo largo del progreso narrativo.

Asimismo, se identificaron las palabras más frecuentes asociadas a cada emoción para facilitar la interpretación semántica de los resultados.

### 3.4.2. Análisis de valencia emocional con AFINN

Para medir la intensidad emocional continua se empleó el léxico AFINN, que asigna puntuaciones entre  $-5$  y  $+5$  a las palabras [1].

Se calculó el sentimiento medio por capítulo y se agregaron los resultados por saga, permitiendo analizar la dinámica emocional global y realizar comparaciones entre universos narrativos.

### 3.4.4. Análisis de sentimiento con ventanas de texto y a nivel de oración

Con el objetivo de capturar variaciones emocionales más finas, se aplicaron dos enfoques complementarios:

- Léxico Bing sobre ventanas deslizantes de texto, calculando el sentimiento neto en segmentos consecutivos, lo que permitió detectar cambios abruptos de tono no necesariamente alineados con los capítulos.
- Paquete *sentimentr* a nivel de oración, incorporando efectos de negación e intensificación, con posterior agregación por capítulo para identificar picos emocionales relevantes.

### 3.4.5. Comparación entre métodos léxicos

Se compararon los resultados obtenidos con AFINN, NRC (reducido a polaridad) y Bing a nivel de capítulo, analizando la coherencia de las tendencias temporales y los valores medios por saga.

Esta comparación permitió evaluar la robustez de los patrones emocionales detectados y justificar el uso combinado de métricas en fases posteriores del análisis.

### 3.4.6. Modelo supervisado de clasificación de sentimiento

Adicionalmente, se entrenó un clasificador supervisado de sentimiento con el objetivo de evaluar la viabilidad de un modelo adaptado al dominio narrativo. Dado que no se disponía de anotaciones manuales, se utilizó etiquetado débil combinando los resultados de los léxicos AFINN y Bing para generar una variable binaria por capítulo.

Los textos se representaron mediante TF-IDF sobre matrices documento-término y se entrenó un modelo de regresión logística regularizada (LASSO) con validación cruzada. El rendimiento se evaluó mediante el AUC de la curva ROC. El modelo se aplicó posteriormente a la saga Empyrean y se compararon sus predicciones con las obtenidas por métodos léxicos, analizando coincidencias y discrepancias, así como las palabras con mayor peso predictivo.

### 3.4.7. Comparación entre sagas clásicas y contemporáneas

Finalmente, se compararon las distribuciones de sentimiento entre sagas clásicas (2008–2015) y la saga Empyrean (2023–2025), utilizando métricas derivadas de *sentimentr*.

Se analizaron diferencias en variabilidad emocional y sesgo de polaridad, permitiendo explorar posibles cambios en el tono narrativo entre períodos editoriales.

## 3.5. Modelización temática (Topic Modelling)

Para identificar los principales ejes narrativos de la saga *The Mortal Instruments* se aplicaron técnicas de modelización temática basadas en Latent Dirichlet Allocation (LDA), combinando un enfoque no supervisado estándar con un enfoque guiado mediante *seeded LDA*. El análisis se realizó a nivel de capítulos agregados, con el objetivo de capturar patrones temáticos estables a lo largo de la narrativa [13].

### 3.5.1. Selección del número de tópicos en LDA estándar

Se entrenaron múltiples modelos LDA no supervisados con distintos valores de  $K$  (entre 3 y 8 tópicos), empleando inferencia mediante muestreo de Gibbs. Para cada modelo se calculó la coherencia temática UMass a partir de los términos principales de cada tópico, y se analizó la coherencia media en función de  $K$ . Con base en este criterio y en la interpretabilidad semántica de los tópicos, se seleccionó un modelo con  $K = 5$  como compromiso entre simplicidad y capacidad descriptiva (ver [Figura 9](#)).

### 3.5.2. Interpretación de tópicos y distribución por libro

Del modelo seleccionado se analizaron:

- La distribución palabra-tópico (matriz  $\beta$ ), para identificar los términos más representativos de cada tópico.
- La distribución documento-tópico (matriz  $\theta$ ), agregada por libro, para estimar el peso relativo de cada eje temático en cada volumen de la saga.

### 3.5.3. Modelización temática guiada mediante Seeded LDA

Con el fin de evaluar la correspondencia entre los tópicos descubiertos automáticamente y categorías narrativas definidas teóricamente, se aplicó un modelo de *seeded LDA* utilizando un diccionario temático construido manualmente. Las categorías incluyeron, entre otras: relaciones románticas, conflicto sobrenatural, política de la Clave, identidad familiar y comunidades del submundo.

El modelo se entrenó también permitiendo la aparición de tópicos residuales no guiados, lo que posibilita capturar contenido no cubierto por el diccionario.

### 3.6. Redes de personajes

El análisis de redes de personajes transforma las relaciones narrativas en estructuras cuantificables. Para cada saga se construyen redes de coaparición por capítulo [15]:

- **Identificación de personajes:** Se utilizan listas controladas específicas para cada saga: 16 personajes para *Empyrean* (ej. “violet”, “xaden”, “tairn”) y 22 para *The Mortal Instruments* (ej. “clary”, “jace”, “simon”). Todos los nombres se normalizan a minúsculas.
- **Construcción de redes:** Se generan pares no dirigidos de personajes que aparecen juntos en un capítulo y se agregan todas las conexiones para cada saga, creando un grafo no dirigido mediante la librería igraph.

**Métricas calculadas:** Número de nodos y aristas, densidad de red y centralidad de grado del personaje principal, que permiten evaluar la distribución del peso narrativo y la estructura de interacción entre personajes.

### 3.7. Datos de recepción y sistema de recomendación

Para complementar los análisis textuales y evaluar la recepción externa de las sagas, se recopilaron los ratings de Goodreads. Los datos se obtuvieron a partir de un archivo Excel (goodreads\_rating.xlsx) que contiene las columnas **Libro**, **Saga** y **rating** (puntuación por libro) [17].

#### 3.7.1. Datos de ratings de Goodreads

El procesamiento se realizó en R mediante el siguiente flujo:

1. **Carga del archivo** mediante `read_excel()`.
2. **Selección de variables relevantes**, conservando únicamente la saga y la puntuación del libro.
3. **Validación de sagas clave**, comprobando que las sagas de interés (*Empyrean*, *Shadowhunters*, entre otras) estuvieran presentes en el conjunto de datos; en caso contrario, se generaría una alerta.
4. **Exportación de los datos procesados** a formato CSV para su uso posterior en los análisis comparativos y el sistema de recomendación.

Posteriormente, los ratings se agregaron por saga calculando el **rating promedio por saga**, así como el número de libros disponibles por cada una, con el fin de facilitar comparaciones globales de recepción entre sagas.

#### 3.7.2. Métricas narrativas integradas en el perfil por saga

Además de los ratings externos, el sistema de recomendación incorpora dos métricas narrativas obtenidas en fases previas del análisis:

- **Densidad de la red de personajes**, calculada a partir de grafos de coocurrencia de personajes por saga mediante la función `edge_density()` del paquete *igraph*. Esta métrica aproxima el grado de interconectividad entre personajes dentro de cada universo narrativo.
- **Sentimiento medio por saga**, calculado mediante el léxico AFINN.
  - Para *Shadowhunters*, se utilizó directamente el archivo consolidado de sentimiento por capítulo.
  - Para *Empyrean*, se realizó un procesamiento adicional: agregación de texto por capítulo, limpieza del texto, tokenización y cálculo del valor medio AFINN por capítulo, seguido del promedio global por saga.

### 3.7.3. Construcción del perfil técnico comparativo

Con las tres dimensiones, **sentimiento AFINN**, **densidad de red** y **rating promedio de Goodreads**, se construyó un **perfil técnico por saga** para las sagas principales (*Empyrean* y *Shadowhunters*). Cada saga queda representada por un vector de características:

- Sentimiento medio
- Densidad de red de personajes
- Rating promedio en Goodreads

Este perfil se utilizó tanto para análisis comparativos visuales como para contextualizar los resultados del sistema de recomendación.

### 3.7.4. Sistema de recomendación basado en similitud de ratings

Dado que el conjunto de sagas es reducido y no existen perfiles individuales de usuarios, se implementó un **sistema de recomendación simple basado en similitud entre sagas**, utilizando como criterio principal el rating promedio en Goodreads (ver [Tabla 7](#) y [Figura 17](#)).

El procedimiento fue el siguiente:

1. Cálculo del **rating medio por saga**.
2. Generación de todas las combinaciones posibles de pares de sagas.
3. Cálculo de la **diferencia absoluta de ratings** entre cada par.
4. Transformación de la diferencia en una medida de **similitud normalizada**, definida como:

$$\text{similitud} = 1 - \frac{\text{diferencia de rating}}{\max(\text{diferencia de rating})}$$

5. Ordenación de las sagas candidatas en función de la mayor similitud respecto a la saga de referencia.

Este enfoque corresponde a un **sistema de recomendación basado en contenido agregado por ítem (saga)**, no a nivel de usuario, y se utiliza como apoyo interpretativo para identificar sagas con patrones de recepción similares.

---

## 4. Resultados

### 4.1. Análisis de sentimiento emocional

El **análisis de sentimiento** permite cuantificar las dimensiones afectivas del texto literario, traduciendo emociones subjetivas en métricas comparables.

#### 4.1.1. El miedo como emoción narrativa clave (ver [Tabla 1-2](#))

En *Shatter Me*, el miedo aparece estrechamente ligado a la identidad de la protagonista y a su percepción de sí misma como amenaza para los demás. Se trata, por tanto, de un miedo interiorizado, que se vive como culpa, confusión y ruptura emocional.

En *Empyrean*, aunque también se observan picos elevados de miedo, estos se concentran en escenas de combate y peligro físico colectivo. A diferencia de la saga anterior, el miedo no conduce a la parálisis, sino que se integra en la acción, la estrategia y la protección del grupo. El resultado es una experiencia emocional intensa, pero orientada hacia la superación del obstáculo inmediato, no hacia la autodestrucción psicológica.

#### 4.1.2. Evolución del tono emocional a lo largo de la saga (ver [Figura 3-5](#))

El análisis del sentimiento con el léxico AFINN permite observar la evolución global del tono emocional capítulo a capítulo. En *Shatter Me*, la curva muestra una caída pronunciada desde valores positivos hacia territorios claramente negativos, con una recuperación lenta y frágil. Los picos positivos son breves y suelen verse anulados rápidamente por nuevos episodios de sufrimiento.

Este patrón es coherente con una narrativa centrada en el trauma, donde la mejora emocional no es acumulativa, sino interrumpida de forma constante por nuevas crisis. El conflicto no se resuelve, sino que se reconfigura una y otra vez, manteniendo a la protagonista en un estado prolongado de vulnerabilidad.

En *Empyrean*, en cambio, la evolución es mucho más estable. Aunque existen descensos marcados en momentos de batalla o pérdida, el tono general se mantiene cercano a valores positivos, y los picos de recuperación son más frecuentes y sostenidos. Esto sugiere una estructura narrativa donde el conflicto se supera de forma progresiva, y donde cada victoria contribuye al crecimiento del grupo y de la protagonista.

#### 4.1.3. Distribución de emociones básicas (ver [Figura 2 y 4](#))

La comparación de la proporción media de emociones por saga confirma estas diferencias estructurales. En *Shatter Me*, las emociones dominantes son el miedo y la tristeza, seguidas por la ira y el desprecio. Las emociones positivas, como la alegría y la confianza, aparecen en niveles significativamente más bajos.

En *Empyrean*, el perfil emocional es mucho más equilibrado. La confianza y la anticipación ocupan un lugar central, reflejando una narrativa basada en la preparación, las alianzas y la expectativa de confrontación. La alegría y la sorpresa también presentan valores elevados, asociados a logros colectivos, descubrimientos y evolución de los vínculos entre personajes. Las emociones negativas están presentes, pero no monopolizan el tono general.

#### 4.1.4. Núcleos semánticos de las emociones (ver [Figura 6-7](#))

El análisis de las palabras más frecuentes asociadas a cada emoción revela diferencias cualitativas relevantes. En *Shatter Me*, el miedo y la tristeza se articulan en torno a términos como kill, pain, broken o hurt, que remiten directamente al daño físico y psicológico. La ira y el desprecio se expresan mediante lenguaje intensamente emocional y reactivo, lo que refuerza la sensación de descontrol y opresión.

En *Empyrean*, las emociones positivas se vinculan a palabras relacionadas con el grupo y la fortaleza compartida, como together, squad, strong o safe. Incluso las emociones negativas aparecen contextualizadas dentro de la acción y el sacrificio colectivo, más que en el aislamiento personal.

#### 4.1.5 Comparación de métodos de medición emocional (ver [Figura 8](#))

La comparación entre los resultados obtenidos con AFINN y con la polaridad binaria de NRC muestra trayectorias coherentes entre ambos métodos, aunque con diferencias en la sensibilidad a la intensidad emocional. AFINN permite captar mejor los extremos emocionales, especialmente en *Shatter Me*, donde las palabras con carga negativa fuerte generan descensos más pronunciados.

En *Empyrean*, la combinación de términos negativos intensos con un alto volumen de vocabulario positivo produce un equilibrio más estable. Esto refuerza la idea de que la saga no reduce la intensidad emocional, sino que la redistribuye entre peligro, vínculo y superación.

#### 4.1.6. Capítulos extremos

La identificación de **capítulos extremos** permite vincular los hallazgos cuantitativos con la trama [\[12\]](#).

Como se observa en las [Tablas 1 y 2](#), los capítulos más negativos de ambas sagas giran en torno a la **traición** o la **pérdida**, pero con una diferencia crucial en el papel de la **protagonista**:

- En *Shatter Me* (**Capítulo 81 de *Unravel Me***, sentimiento = -0.162), constituye el final del libro donde se describe a Juliette como víctima pasiva de una traición, con un lenguaje centrado en la **parálisis** (“can’t”, “frozen”, “alone”).
- En *Empyrean* (**Capítulo 87 de *Fourth Wing***, sentimiento = -0.108), el sentimiento también es negativo, pero el discurso se centra en la **toma de decisiones bajo presión** (“command”, “choose”, “fight”), evidenciando un **rol activo** incluso en la adversidad.

#### 4.1.7. Evaluación del modelo personalizado

El modelo personalizado, entrenado sobre capítulos de sagas clásicas, alcanzó un **AUC de 0.668** en el conjunto de prueba, lo que indica una capacidad moderada de discriminar entre capítulos positivos y negativos. Este resultado confirma que el modelo aprende patrones emocionales contextuales propios del género, más allá de la simple polaridad de palabras aisladas.

#### 4.1.8. Comparación de tono emocional con el léxico Bing (ver [Tabla 3](#))

Los valores promedio son similares sin embargo, la desviación de las sagas clásicas es mucho mayor, reflejando una mayor volatilidad emocional en esos textos.

#### 4.1.9. Evolución del sentimiento por método

Al comparar la evolución del sentimiento capítulo a capítulo, se observa un patrón distinto entre los métodos (ver [Figura 11](#)):

- **Léxico Bing:** La línea roja muestra fluctuaciones bruscas con picos de negatividad profunda, interpretando literalmente palabras como *dragon* o *battle* como negativas, aunque en contexto no lo sean.
- **Modelo personalizado:** La línea azul presenta una evolución más estable, con valores mayoritariamente positivos y una ligera tendencia hacia la neutralidad en los capítulos finales. Esto indica que el modelo reconoce la carga contextual de palabras que, aunque asociadas a conflicto o acción, no implican negatividad en la narrativa de *Empyrean*.

#### 4.1.10. Palabras más predictivas

El modelo identifica las palabras con mayor peso para predecir el tono emocional (ver [Figura 10](#)):

**Palabras positivas:** *love, brilliant, romantic, beauty, polite, smiling, protective, include, courage*

- Reflejan conexiones humanas, superación personal y pertenencia a grupos.

**Palabras negativas:** *scream, hell*

- Indican conflicto físico y emocional, vulnerabilidad o presión social, más que fracaso o tristeza general.

#### 4.1.11. Distribución del tono emocional

La densidad de los valores de sentimiento por capítulo muestra diferencias claras entre sagas (ver [Figura 12](#)):

- **Sagas Clásicas:** Curva con un pico centrado cerca de 0.0 y colas largas hacia valores negativos, indicando predominio de tonos neutros o ligeramente negativos y presencia frecuente de capítulos con emociones intensamente negativas.
- **Empyrean:** Curva más ancha y desplazada ligeramente hacia valores positivos (0.0–0.1), reflejando mayor estabilidad emocional y menor enfoque en el sufrimiento individual.

### 4.2. Modelización temática con LDA guiado

La **modelización temática** permite identificar patrones y estructuras subyacentes en textos narrativos. En este estudio, se utiliza el **LDA guiado (seeded LDA)**, una versión del modelo de asignación latente de Dirichlet que incorpora conocimiento previo mediante **diccionarios temáticos**, lo que mejora la interpretación de los resultados en contextos literarios. [\[14\]](#)

#### 4.2.1. Tópicos identificados

El **seeded LDA** se aplicó sobre el corpus de *The Mortal Instruments*, segmentado en macro-capítulos (agrupaciones de tres capítulos consecutivos) [\[13\]](#). Los resultados muestran una distribución temática clara (ver [Tabla 4](#)):

- **Guerra sobrenatural** es el tópico más frecuente, con palabras como “demon”, “blade” y “angel”.
- **Política institucional** sigue en frecuencia, destacando términos como “council” e “inquisitor”.
- **Romance** aparece con palabras íntimas (“love”, “heart”, “together”), pero con menor peso que en sagas puramente románticas.
- **Identidad y Mundo subterráneo** funcionan como ejes secundarios que enriquecen la trama y aportan profundidad.

Estos resultados sugieren que *The Mortal Instruments* combina **acción colectiva** e **introspección individual**, operando con una arquitectura temática dual que equilibra conflicto, política y desarrollo personal.

### 4.3. Redes de personajes y estructura narrativa

El análisis de redes de personajes aplicado a la narrativa literaria permite transformar las relaciones entre personajes en estructuras cuantificables [\[16\]](#). En este estudio, se construyen redes de coaparición por capítulo, donde los **nodos** representan personajes relevantes y las **aristas** reflejan su presencia conjunta en el mismo contexto narrativo [\[15\]](#).

Las redes de coaparición muestran diferencias significativas entre sagas (ver [Tabla 5](#)):

- *The Mortal Instruments*: 21 nodos y 438 aristas, densidad 2.09; la centralidad máxima (42) indica que los protagonistas (Clary, Jace, Simon, Isabelle y Alec) están conectados con muchos personajes secundarios, reflejando un modelo narrativo centrado en varios personajes principales (ver [Figura 13](#)).
- *Empyrean*: 14 nodos y 194 aristas, densidad 2.13; la centralidad máxima (28) refleja que la protagonista (Violet), actúa como un vínculo dentro de un grupo más equilibrado (Xaden, Tairn, Andarna, Rhiannon, etc.), mostrando un enfoque colectivo de la narrativa (ver [Figura 14](#)).

Como podemos observar en la [Figura 15](#), estos resultados permiten interpretar cómo se distribuye el peso narrativo: mientras *The Mortal Instruments* se centra en un protagonista altamente conectado, *Empyrean* organiza la interacción de manera más equitativa entre los miembros del grupo, reforzando la dimensión colectiva de la historia.

#### 4.4. Sistema de recomendación basado en perfil emocional-estructural

Finalmente, este proyecto propone un sistema de recomendación basado en contenido que combina tres dimensiones cuantificables: el tono emocional, la interacción entre personajes y la recepción externa [\[18\]](#).

##### 4.4.1. Perfil de cada saga

Para cada saga se construye un perfil que integra tres métricas principales (ver [Tabla 6](#) y [Figura 16](#)):

- **Sentimiento promedio (AFINN)**: Empyrean muestra un valor de -0.254, mientras que Shadowhunters tiene un tono más negativo, de -0.571.
- **Densidad de red**: Empyrean presenta una densidad ligeramente mayor (2.13 frente a 2.09), lo que refleja que sus personajes están más interconectados.
- **Rating promedio en Goodreads**: Empyrean alcanza un 4.50, en comparación con 4.18 de Shadowhunters.

Estas tres métricas se combinan para formar un vector que representa cada saga, integrando tanto su estructura interna como la respuesta de los lectores.

##### 4.4.2. Comparación entre sagas

El **sistema de recomendación** propuesto no se basa en *preferencias individuales de usuarios*, sino en **similitud entre sagas** a partir de sus **perfíles narrativos agregados**. Por tanto, las recomendaciones deben interpretarse como **afinidad estructurales y emocionales** entre universos narrativos, y no como predicciones de comportamiento lector a nivel individual.

Los resultados en la [Tabla 8](#) muestran que *Throne of Glass* es la saga más similar a *Empyrean*, con una **similitud cercana a 0.95** y una **diferencia mínima en ratings**. Esta cercanía puede explicarse por la combinación de **tramas épicas**, fuerte presencia de **vínculos interpersonales y protagonistas jóvenes en procesos de formación**, características compartidas por ambas sagas.

En posiciones intermedias aparecen *Percy Jackson* y *Vampire Academy*, que, aunque difieren en **tono y ambientación**, mantienen **estructuras narrativas centradas en grupos de personajes recurrentes y progresión emocional clara**.

*Shadowhunters* aparece en una posición inferior, lo que concuerda con los análisis previos que mostraban diferencias tanto en el **tono emocional medio** como en la **distribución temática**, con mayor peso de **estructuras institucionales y conflictos políticos**.

---

## 5. Discusión y conclusiones

### 5.1. Interpretación de los resultados

Los resultados respaldan claramente la hipótesis central del estudio: el éxito de *Empyrean* no se explica únicamente por tendencias comerciales, sino por una **reconfiguración narrativa** que combina **innovación emocional** con **estructuras conocidas**. A diferencia de *Shatter Me*, donde el miedo se experimenta como aislamiento y parálisis, *Empyrean* transforma emociones intensas, miedo, anticipación, confianza, en **acción colectiva**. Esta diferencia no es solo temática: mientras *Shatter Me* se centra en una protagonista aislada, *Empyrean* construye una **red de relaciones densa** (2.13) en la que la **agencia se distribuye entre varios personajes**, haciendo que la historia sea más coral y participativa.

Asimismo, la **similitud narrativa** entre *Empyrean* y *Shadowhunters* (0.96) indica que el romanticismo moderno no rompe con el pasado, sino que **recicla fórmulas exitosas**, como la integración de romance y conflicto institucional, adaptándolas a un enfoque más **colectivo y emocionalmente complejo**. Esto refuerza la idea de que el éxito editorial no depende solo de la originalidad: es clave ofrecer **novedad suficiente para sorprender**, pero también **elementos familiares que generen confort** en los lectores [\[20\]](#).

### 5.2. Implicaciones prácticas

Estos hallazgos tienen **aplicaciones prácticas** tanto para editoriales como para plataformas de recomendación:

- **Para editoriales:** el perfil propuesto, que combina **sentimiento, densidad de red y rating**, puede servir como **herramienta de evaluación previa** de manuscritos. Un texto con una **red relacional sólida**, vocabulario que fomente la colectividad

(“squad”, “together”) y un **equilibrio emocional** tiene más probabilidades de conectar con la audiencia actual.

- **Para plataformas como Goodreads o BookTok:** los sistemas de recomendación basados únicamente en género o autor resultan limitados. Integrar **señales textuales más profundas**, como la **estructura de relaciones** o el **perfil emocional**, permite generar sugerencias más **precisas y relevantes** para los lectores.

### 5.3. Limitaciones

El estudio presenta tres limitaciones clave:

1. **Alcance del corpus:** el análisis detallado se centra únicamente en tres sagas. Esto permite un estudio profundo, pero no refleja toda la diversidad del **romántico contemporáneo**.
2. **Simplificación del sentimiento:** los léxicos **AFINN** y **NRC** no capturan ironía, contexto o ambigüedad emocional, lo que puede generar cierta **distorsión en la interpretación de algunos capítulos**.
3. **Dependencia del diccionario en sagas específicas:** en el caso de *Shadowhunters*, se observó que el rendimiento de los modelos de análisis de sentimiento y de seeded LDA dependía fuertemente de la especificidad del léxico. Si los diccionarios no estaban suficientemente adaptados al universo narrativo, los resultados podían ser inconsistentes o erróneos. Esta limitación resalta la importancia de personalizar los diccionarios según el contexto de cada saga.

### 5.4. Líneas futuras

Futuras investigaciones podrían:

- **Ampliar el corpus** a otras sagas contemporáneas (**Powerless** o **Blood and Ash**) para validar la generalización del **patrón colectivo**.
- **Desarrollar un diccionario temático específico** para **Empyrean** y aplicar **LDA guiado** de forma comparativa.
- **Integrar datos de BookTok** (*comentarios, hashtags*) para vincular el **análisis textual** con la **viralidad real**.
- **Implementar un prototipo funcional de recomendación narrativa** en una plataforma tipo **Shiny**, permitiendo a los lectores explorar **afinidad basadas en arquitectura emocional**, no solo en género.
- **Explorar la división de los capítulos en segmentos más pequeños (*micro-documentos*)** para capturar **variaciones emocionales o temáticas** dentro de un mismo capítulo. Sin embargo, esta estrategia presentó **dificultades durante la implementación**: aumentaba de forma exponencial el **número de documentos y tópicos**, encareciendo la interpretación y degradando la **coherencia general del**

**análisis.** Aunque no se adoptó como enfoque definitivo, esta idea podría retomarse en **estudios futuros** para conseguir una mayor **granularidad** sin comprometer la *interpretabilidad*.

- En conjunto, este trabajo demuestra que la literatura juvenil no solo refleja su contexto cultural, sino que **lo codifica en su estructura narrativa**, y esa codificación, hoy, puede leerse tanto con los ojos como con algoritmos.
- 

## 6. Referencias

1. Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). *Lexicon-based methods for sentiment analysis*. Computational Linguistics, 37, 267–307. [https://doi.org/10.1162/coli\\_a\\_00049](https://doi.org/10.1162/coli_a_00049)
2. Zad, S., Jimenez, J., & Finlayson, M. (2021). *Hell hath no fury? Correcting bias in the NRC emotion lexicon*. Proceedings of the 5th Workshop on Online Abuse and Harms (WOAH 2021). <https://doi.org/10.18653/v1/2021.woah-1.11>
3. Chiny, M., Chihab, M., Bencharef, O., & Chihab, Y. (2021). *LSTM, VADER and TF-IDF based hybrid sentiment analysis model*. <https://doi.org/10.14569/ijacsa.2021.0120730>
4. Jacobs, T., & Tschötschel, R. (2019). *Topic models meet discourse analysis: A quantitative tool for a qualitative approach*. International Journal of Social Research Methodology, 22, 469–485. <https://doi.org/10.1080/13645579.2019.1576317>
5. Harandizadeh, B., & Morstatter, F. (2021). *Keyword assisted embedded topic model*. Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining. <https://doi.org/10.1145/3488560.3498518>
6. Min, S., & Park, J. (2016). *Modeling narrative structure and dynamics with networks, sentiment analysis, and topic modeling*. PLoS ONE, 14. <https://doi.org/10.1371/journal.pone.0226025>
7. Sarma, D., Mittra, T., & Shahadat, M. (2021). *Personalized book recommendation system using machine learning algorithm*. International Journal of Advanced Computer Science and Applications. <https://doi.org/10.14569/ijacsa.2021.0120126>
8. Watanabe, M., & Fukazawa, Y. (2025). *Quantitative analysis of emotional changes in narrative progression of popular web novels*. Transactions of the Japanese Society for Artificial Intelligence. [https://doi.org/10.1527/tjsai.40-5\\_mo25-c](https://doi.org/10.1527/tjsai.40-5_mo25-c)

9. Rani, N. (2025). *The rise of BookTok: How TikTok is transforming literary trends and reader communities*. International Journal of Scientific Research in Engineering and Management. <https://doi.org/10.55041/ijserem49758>
10. García, M. (2026). *01a\_preprocesado\_corpus\_focales.R*. GitHub repository. [https://github.com/mariayugarcia/proyecto\\_romantasy\\_final/blob/main/scripts/01a\\_preprocesado\\_corpus\\_focales.R](https://github.com/mariayugarcia/proyecto_romantasy_final/blob/main/scripts/01a_preprocesado_corpus_focales.R)
11. García, M. (2026). *01b\_preprocesado\_shadowhunters.R*. GitHub repository. [https://github.com/mariayugarcia/proyecto\\_romantasy\\_final/blob/main/scripts/01b\\_preprocesado\\_shadowhunters.R](https://github.com/mariayugarcia/proyecto_romantasy_final/blob/main/scripts/01b_preprocesado_shadowhunters.R)
12. García, M. (2026). *02\_analisis\_sentimiento.R*. GitHub repository. [https://github.com/mariayugarcia/proyecto\\_romantasy\\_final/blob/main/scripts/02\\_analisis\\_sentimiento.R](https://github.com/mariayugarcia/proyecto_romantasy_final/blob/main/scripts/02_analisis_sentimiento.R)
13. García, M. (2026). *03\_topic\_modeling.R*. GitHub repository. [https://github.com/mariayugarcia/proyecto\\_romantasy\\_final/blob/main/scripts/03\\_topic\\_modeling.R](https://github.com/mariayugarcia/proyecto_romantasy_final/blob/main/scripts/03_topic_modeling.R)
14. Joyeux-Prunel, B. (2024). *Digital humanities in the era of digital reproducibility: Towards a fairest and post-computational framework*. International Journal of Digital Humanities, 6, 23–43. <https://doi.org/10.1007/s42803-023-00079-6>
15. García, M. (2026). *04\_redes\_personajes.R*. GitHub repository. [https://github.com/mariayugarcia/proyecto\\_romantasy\\_final/blob/main/scripts/04\\_redes\\_personajes.R](https://github.com/mariayugarcia/proyecto_romantasy_final/blob/main/scripts/04_redes_personajes.R)
16. Rodda, I., & Bhavani, D. (2023). *Visualization of the dynamics in character networks using social network analysis*. Proceedings of the 6th Joint International Conference on Data Science & Management of Data (10th ACM IKDD CODS and 28th COMAD). <https://doi.org/10.1145/3570991.3571019>
17. García, M. (2026). *05\_datos\_repcion.R*. GitHub repository. [https://github.com/mariayugarcia/proyecto\\_romantasy\\_final/blob/main/scripts/05\\_datos\\_repcion.R](https://github.com/mariayugarcia/proyecto_romantasy_final/blob/main/scripts/05_datos_repcion.R)
18. García, M. (2026). *06\_sistema\_recomendacion.R*. GitHub repository. [https://github.com/mariayugarcia/proyecto\\_romantasy\\_final/blob/main/scripts/06\\_sistema\\_recomendacion.R](https://github.com/mariayugarcia/proyecto_romantasy_final/blob/main/scripts/06_sistema_recomendacion.R)
19. Semeraro, A., Vilella, S., & Ruffo, G. (2021). *PyPlutchik: Visualising and comparing emotion-annotated corpora*. PLoS ONE, 16. <https://doi.org/10.1371/journal.pone.0256503>
20. Form, S. (2019). *Reaching Wuthering Heights with Brave New Words: The influence of originality of words on the success of outstanding best-sellers*. Journal of Creative Behavior, 53, 508–518. <https://doi.org/10.1002/jocb.230>

21. García, M. (2026). *proyecto\_romantasy\_final* [Repositorio de GitHub].  
[https://github.com/mariayugarcia/proyecto\\_romantasy\\_final/tree/main](https://github.com/mariayugarcia/proyecto_romantasy_final/tree/main)
22. Montalva Candel, I. (s. f.). *Tema 5*. En **moncanig.quarto.pub/dne2526**.  
 Recuperado 23 de enero de 2026, de  
<https://moncanig.quarto.pub/dne2526/Tema5.html>
23. Montalva Candel, I. (s. f.). *Tema 6*. En **moncanig.quarto.pub/dne2526**.  
 Recuperado 23 de enero de 2026, de  
<https://moncanig.quarto.pub/dne2526/Tema6.html>
24. Montalva Candel, I. (s. f.). *Tema 7*. En **moncanig.quarto.pub/dne2526**.  
 Recuperado 23 de enero de 2026, de  
<https://moncanig.quarto.pub/dne2526/Tema7.html>
25. Montalva Candel, I. (s. f.). *Tema 8*. En **moncanig.quarto.pub/dne2526**.  
 Recuperado 23 de enero de 2026, de  
<https://moncanig.quarto.pub/dne2526/Tema8.html>
- 

## 7. Anexos

### 7.1. Tablas

Libro	Capítulo	Sentimiento	Frecuencia	Total de palabras	Saga	Proporción
Shatter Me	14	fear	96	332	Shatter Me	0.289
Unravel Me	8	fear	7	41	Shatter Me	0.171
Iron Flame	120	fear	91	552	Empyrean	0.165
Fourth Wing	85	fear	345	2119	Empyrean	0.163
Fourth Wing	87	fear	61	382	Empyrean	0.160

Tabla A1. Distribución del sentimiento “fear” por capítulo y saga

Saga	Tipo	Libro	Capítulo	Sentimiento promedio	Frases
Empyrean	MÁS POSITIVO	Onyx Storm (The Empyrean Book 3)	104	0.047	308
Empyrean	MÁS NEGATIVO	Fourth Wing	87	-0.108	126
Shatter Me	MÁS POSITIVO	Shatter Me	14	0.216	126
Shatter Me	MÁS NEGATIVO	Unravel Me	81	-0.162	12

Tabla A2. Capítulos con mayor y menor sentimiento promedio por saga

Saga	Capítulos	Tono promedio	Desviación
Empyrean (2023–2025)	108	-0.370	0.130
Sagas Clásicas (2008–2015)	1058	-0.361	0.265

Tabla A3. Comparación de tono emocional promedio por saga usando el léxico Bing

Ran go	Romance	Identidad Padres	Shadowhunting /Guerra	Políti ca clave	Downworld ers	Otros
1	love	mother	demon	clave	vampire	portal
2	heart	father	blade	inquisitor	pack	magic
3	bed	past	angel	council	warlock	mundane
4	together	family	rune	power	water	lilith
5	touch	mom	mark	gard	faerie	spell
6	brother	cup	stele	city	werewolf	new
7	closer	morgenstern	circle	accords	praetor	dress
8	remembered	truth	fire	law	metal	parabatai

Tabla A4. Palabras más frecuentes por categoría temática

Saga	Personajes	Conexiones	Centralidad máxima
Empyrean	14	194	28
Shadowhunters	21	438	42

Tabla A5. Red de personajes por saga

Saga	Sentimiento Positivo	Densidad Red	Rating Goodreads
Empyrean	-0.254	2.13	4.50
Shadowhunters	-0.571	2.09	4.18

Tabla A6. Sentimiento promedio, densidad de red y rating en Goodreads

Saga	Nº de libros	Rating promedio
Throne of Glass	2	4.540
Empyrean	3	4.503
Percy Jackson & the Olympians	5	4.378
Vampire Academy	6	4.277
Shadowhunters	6	4.183
Shatter Me	3	4.130
Divergent	2	4.045
Fallen	3	3.847
The Maze Runner	3	3.840

Tabla A7. Número de libros y rating promedio por saga

Saga actual	Saga candidata	Similitud	Diferencia rating
Empyrean	Throne of Glass	0.948	0.037
Empyrean	Percy Jackson & the Olympians	0.821	0.125
Empyrean	Vampire Academy	0.676	0.227
Empyrean	Shadowhunters	0.543	0.320
Empyrean	Shatter Me	0.467	0.373

Tabla A8. Similitud entre sagas candidatas y Empyrean

## 7.2. Gráficos

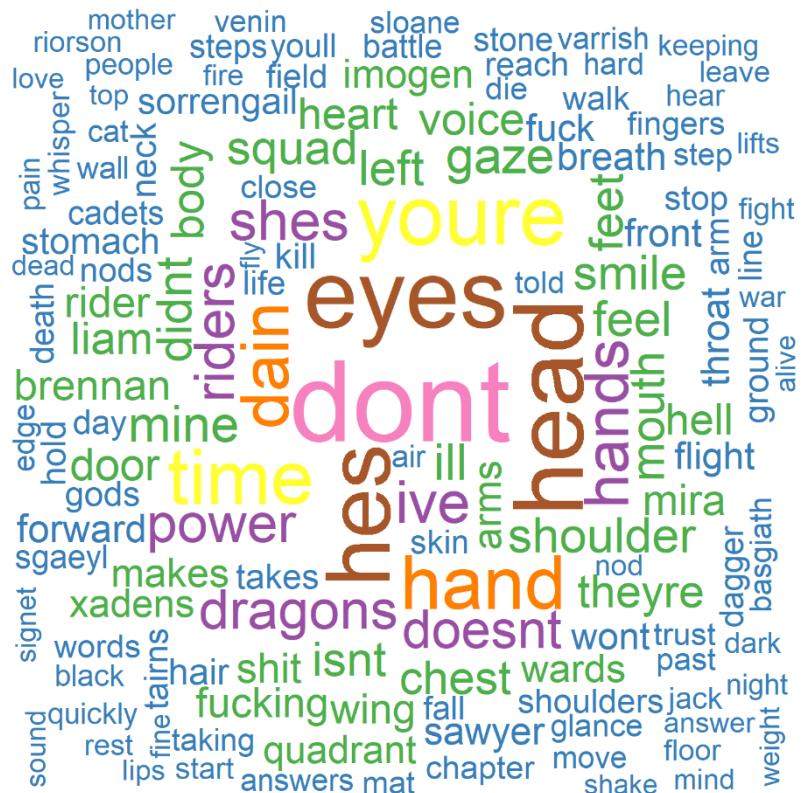


Figura A1. Nube de palabras Empyrean

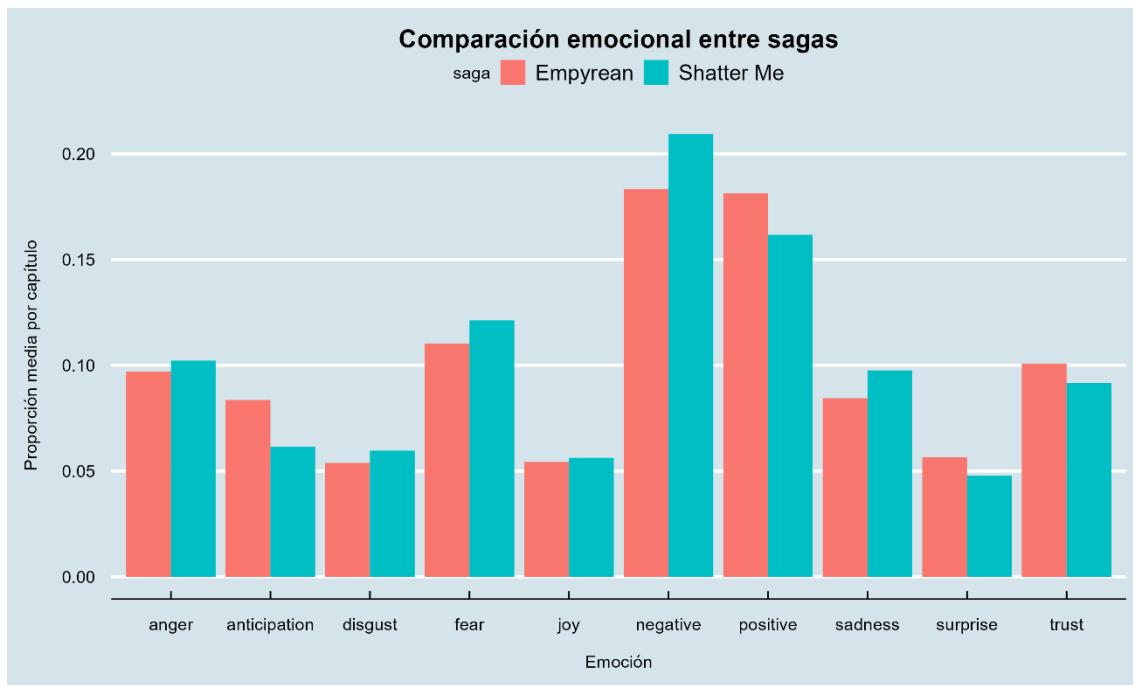


Figura A2. Comparación emocional entre sagas.

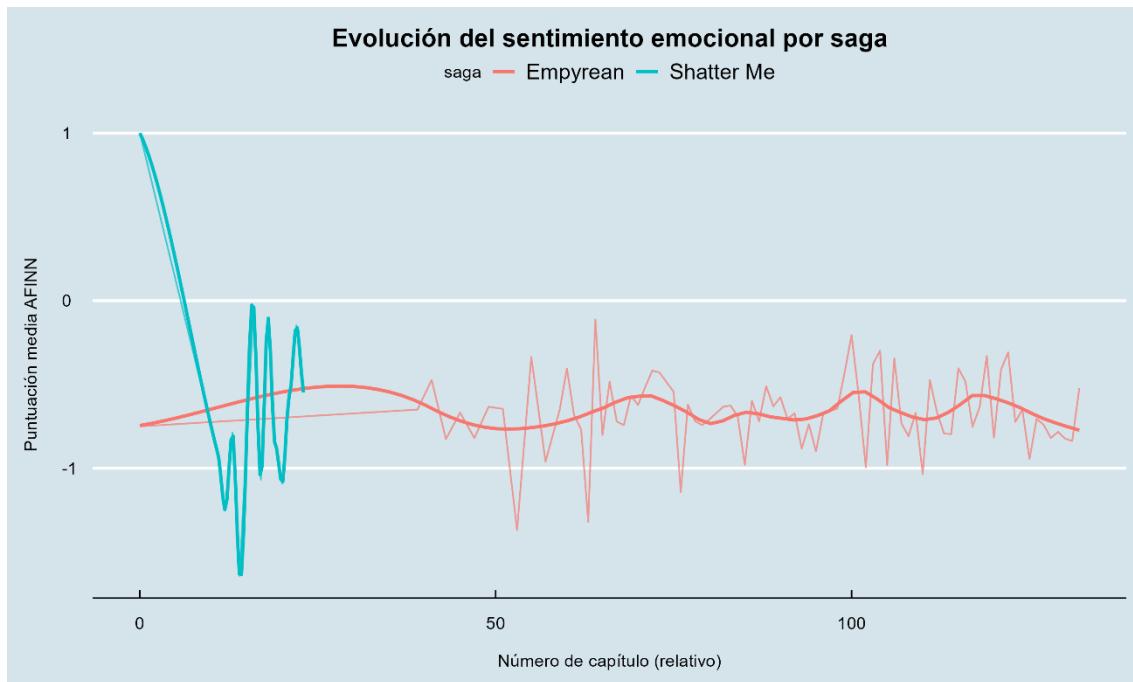


Figura A3. Evolución del sentimiento emocional por saga.

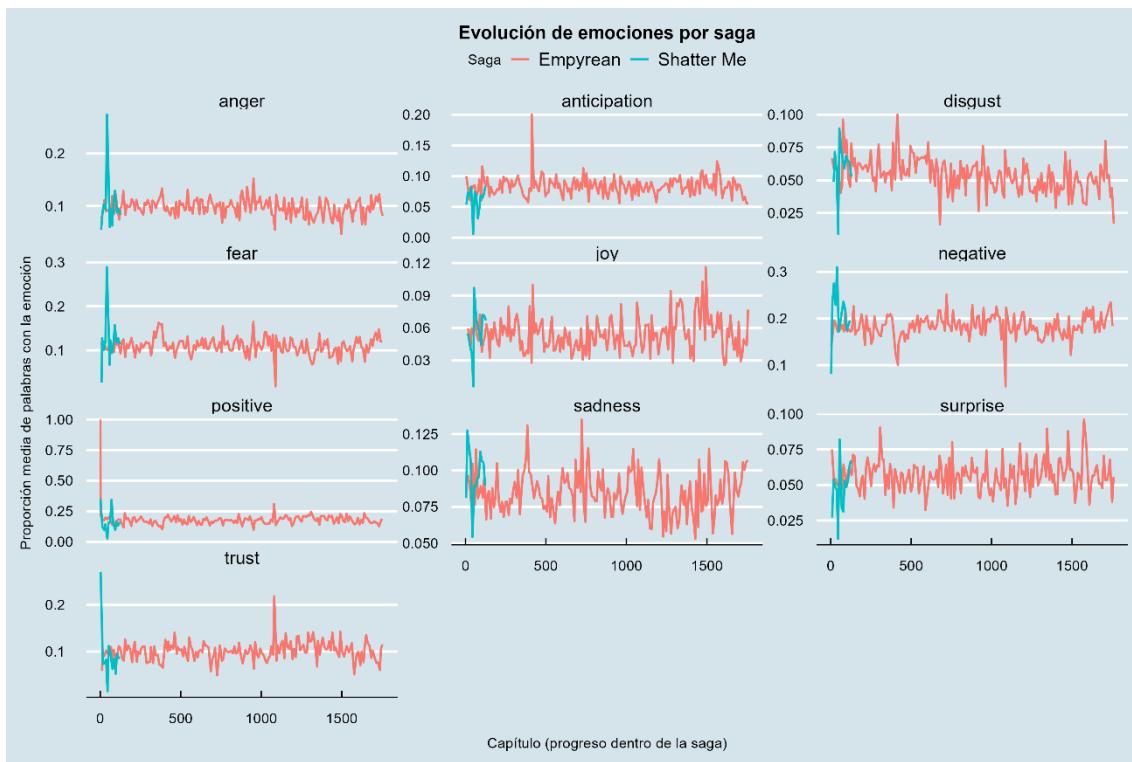


Figura A4. Evolución de emociones por saga

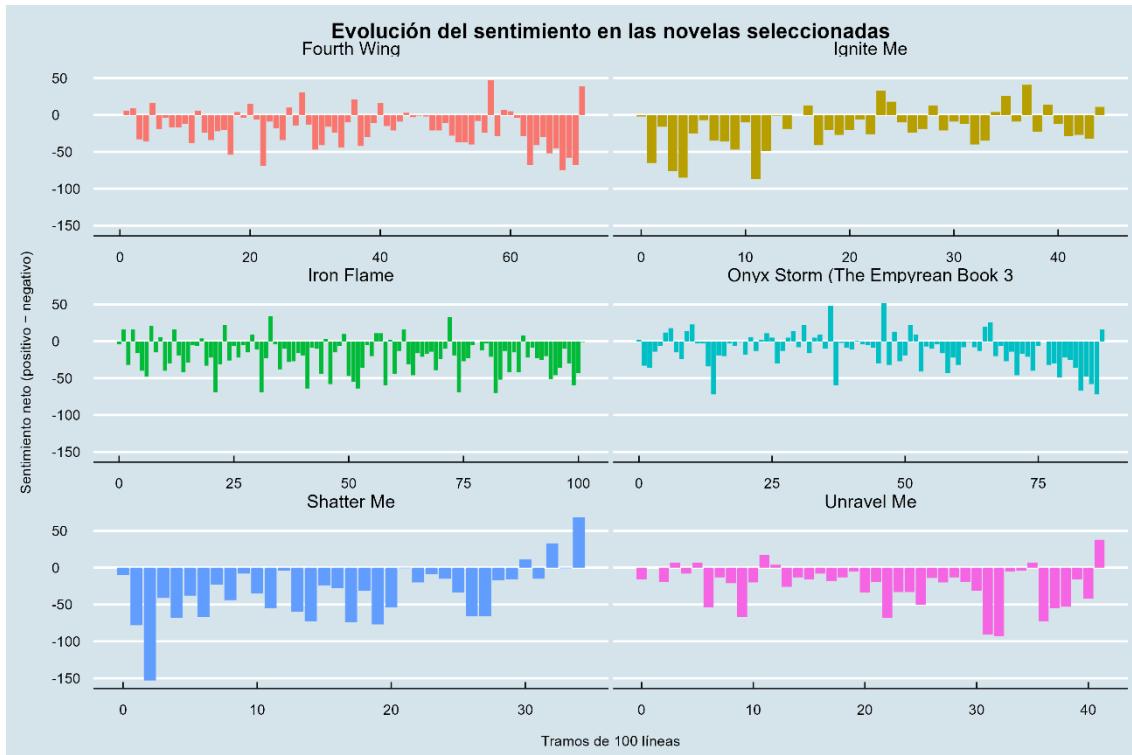


Figura A5. Evolución del sentimiento en las sagas Shatter Me y Empyrean

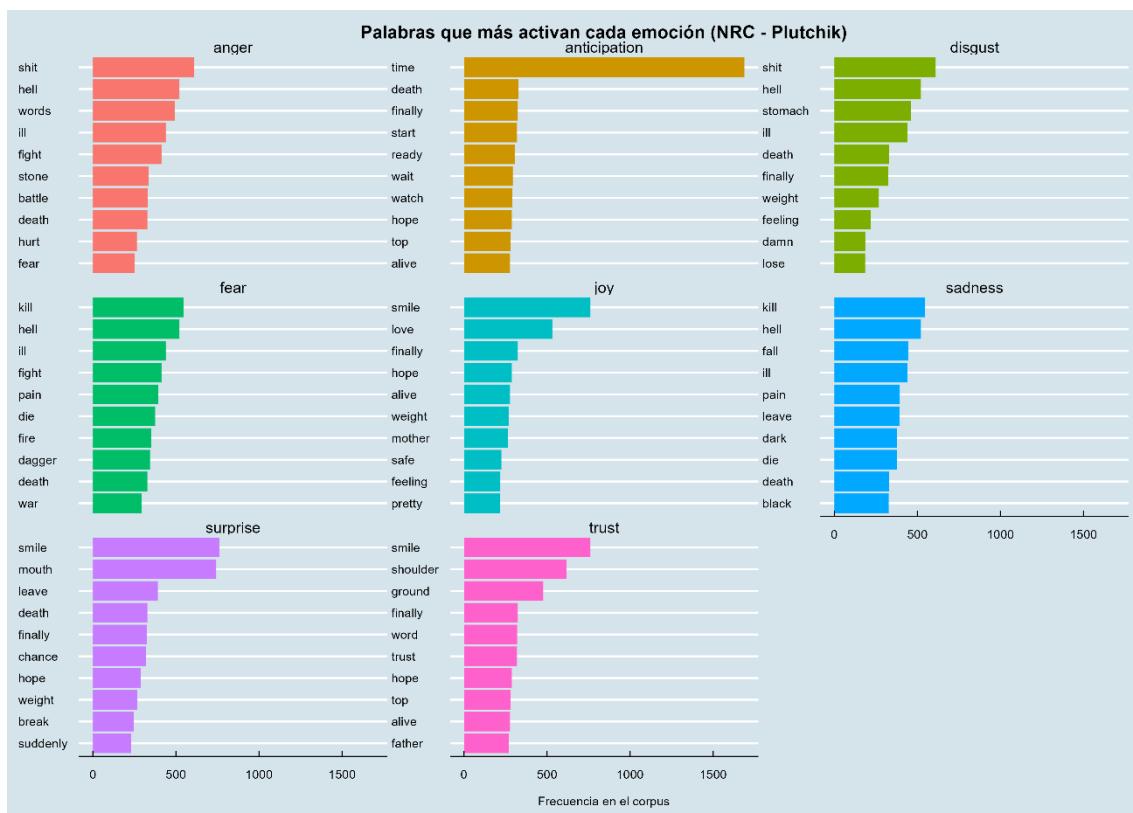


Figura A6. Palabras que más activan cada emoción (NRC – Plutchik)

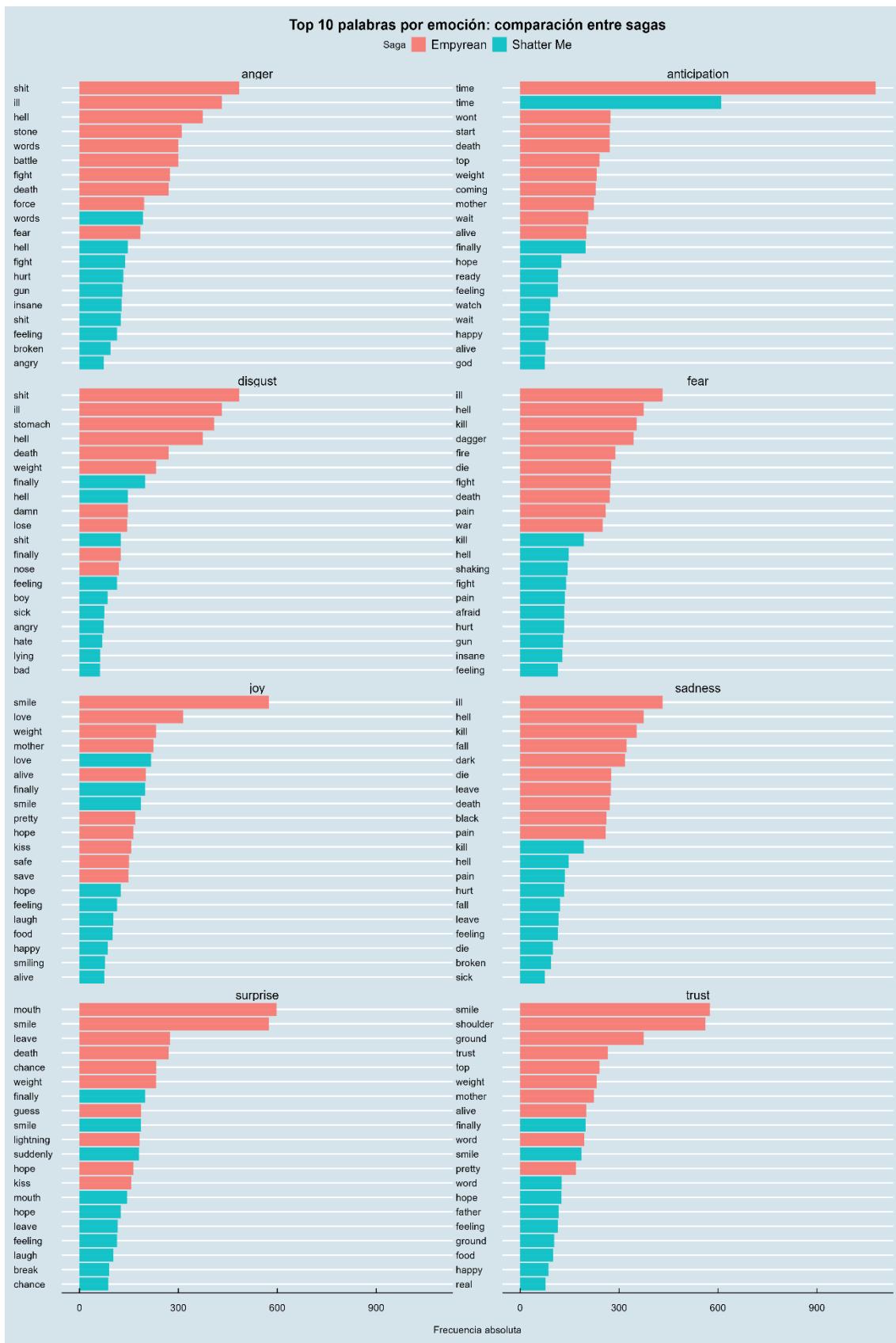


Figura A7. Top 10 palabras por emoción: comparación entre sagas

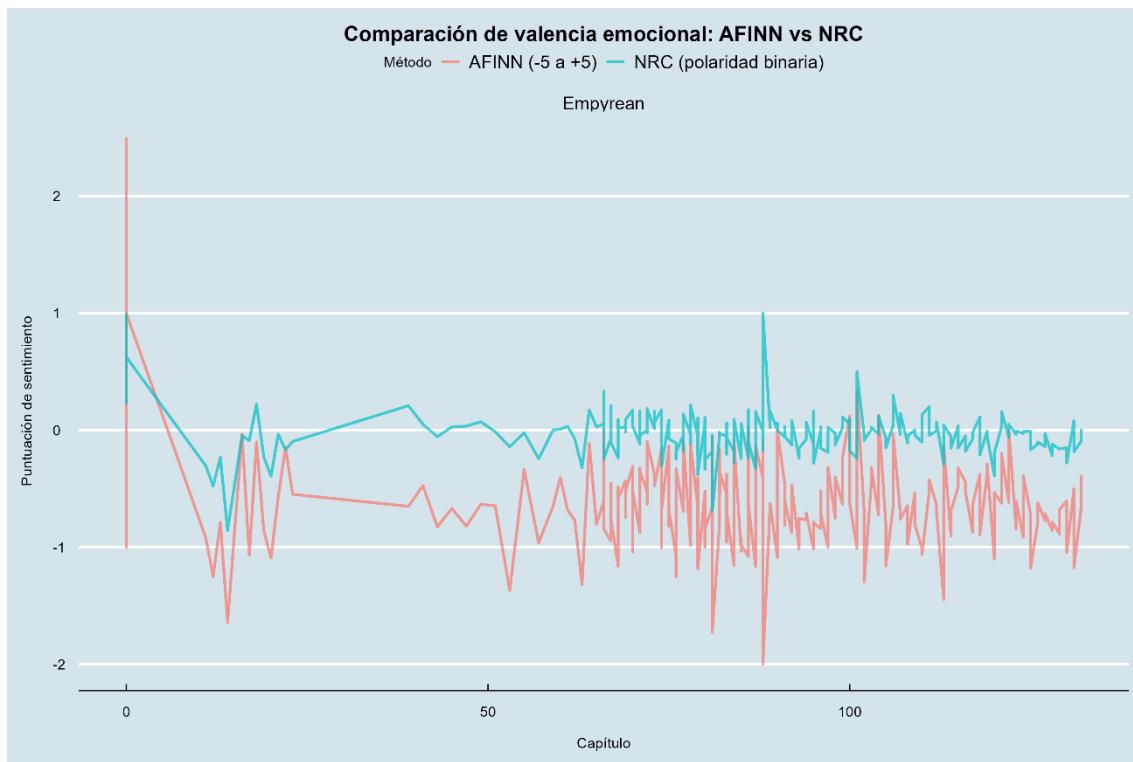


Figura A8. Comparación valencia emocional: AFINN vs NRC

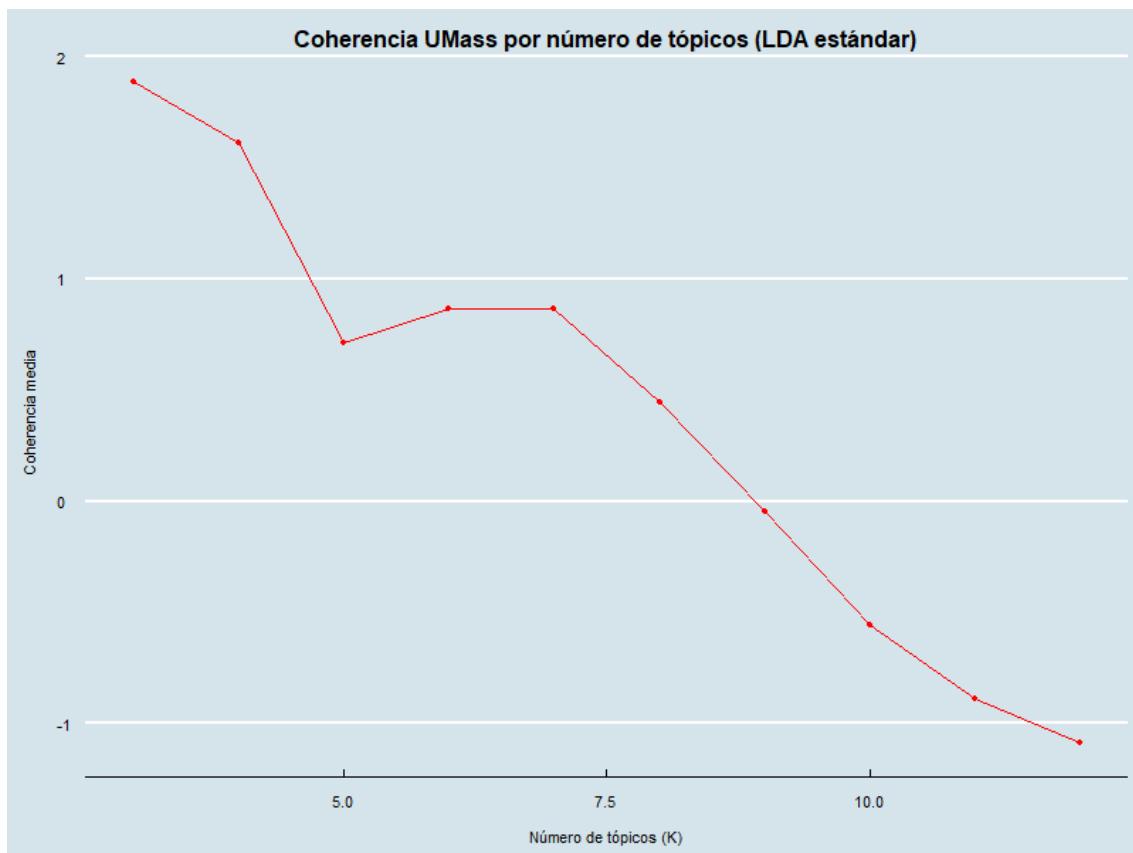
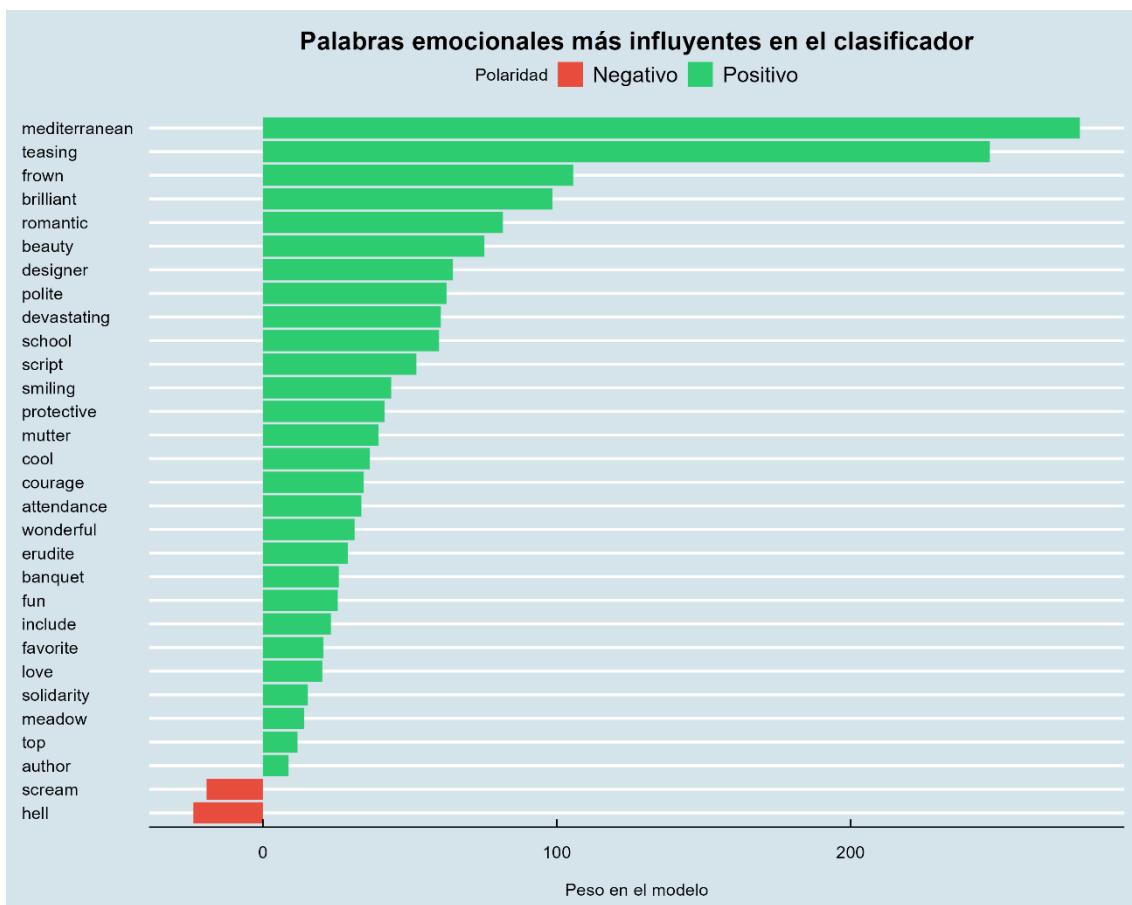


Figura A9. Coherencia UMass por número de tópicos (LDA estándar)



Figuras A10. Palabras emocionales más influyentes en el clasificador

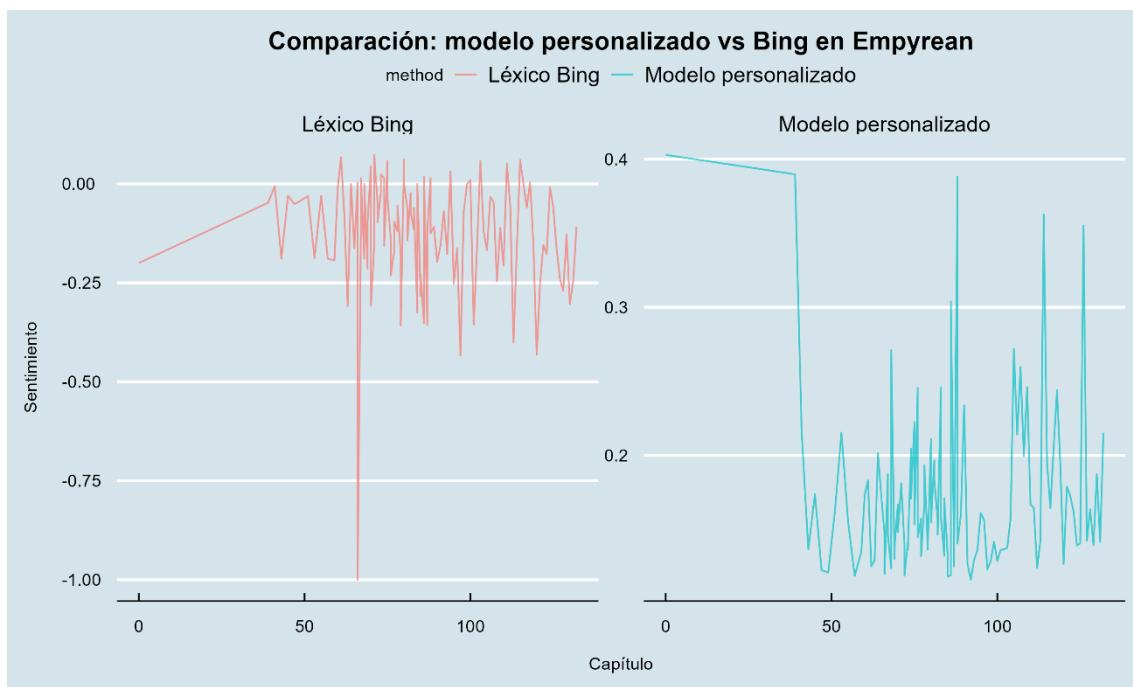


Figura A11. Comparación: modelo personalizado vs Bing saga Empyrean

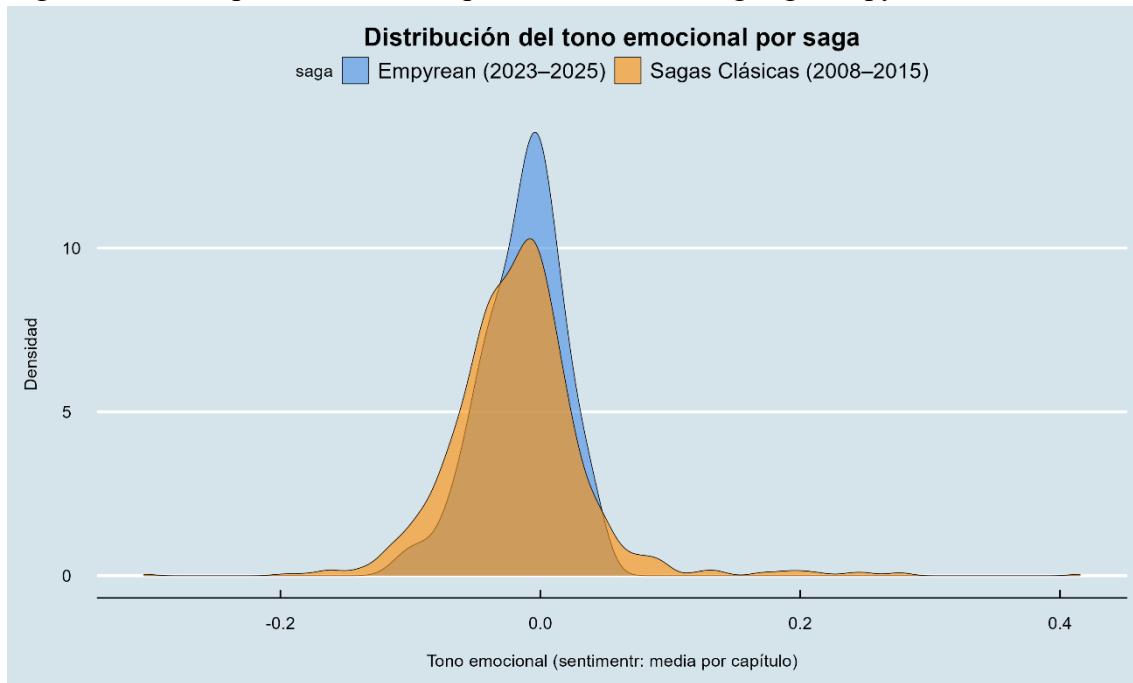


Figura A12. Distribución del tono emocional por saga

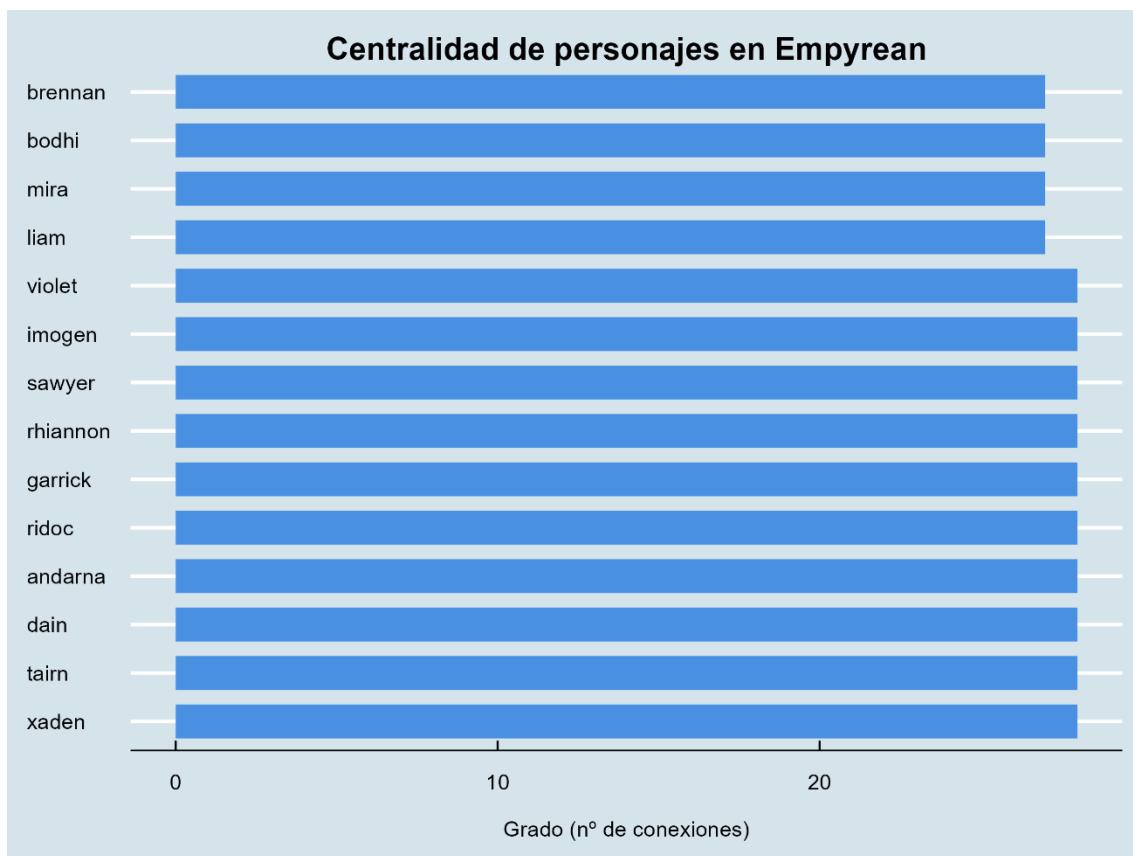


Figura A13. Centralidad de personajes en Empyrean

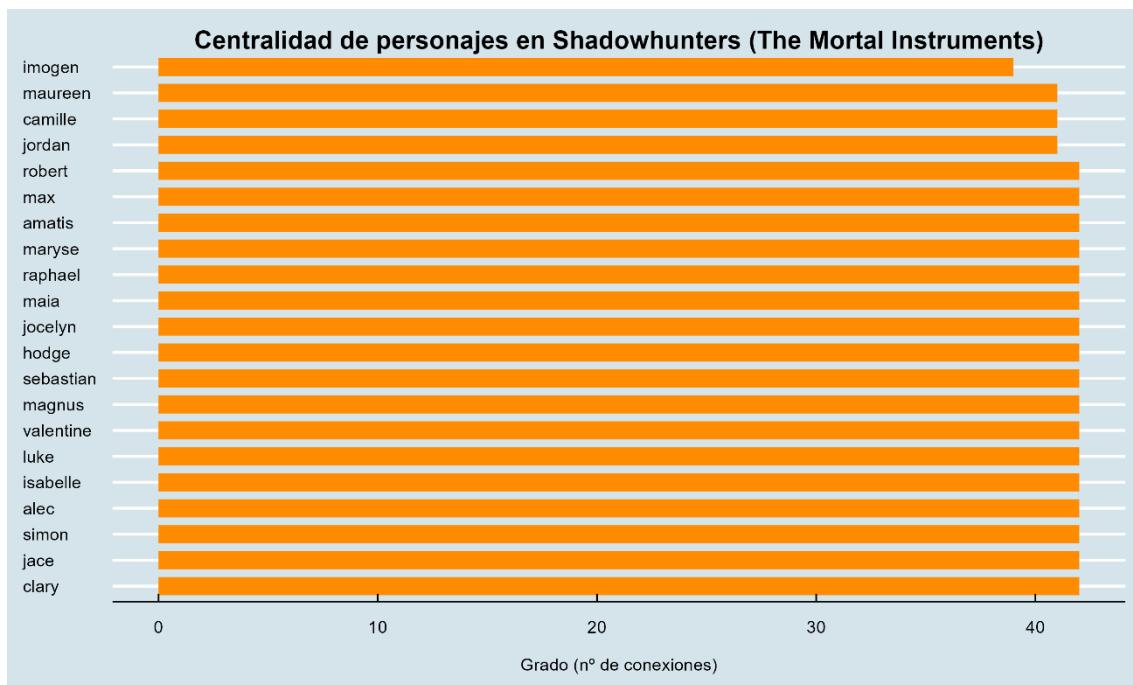


Figura A14. Centralidad de personajes en Shadowhunters

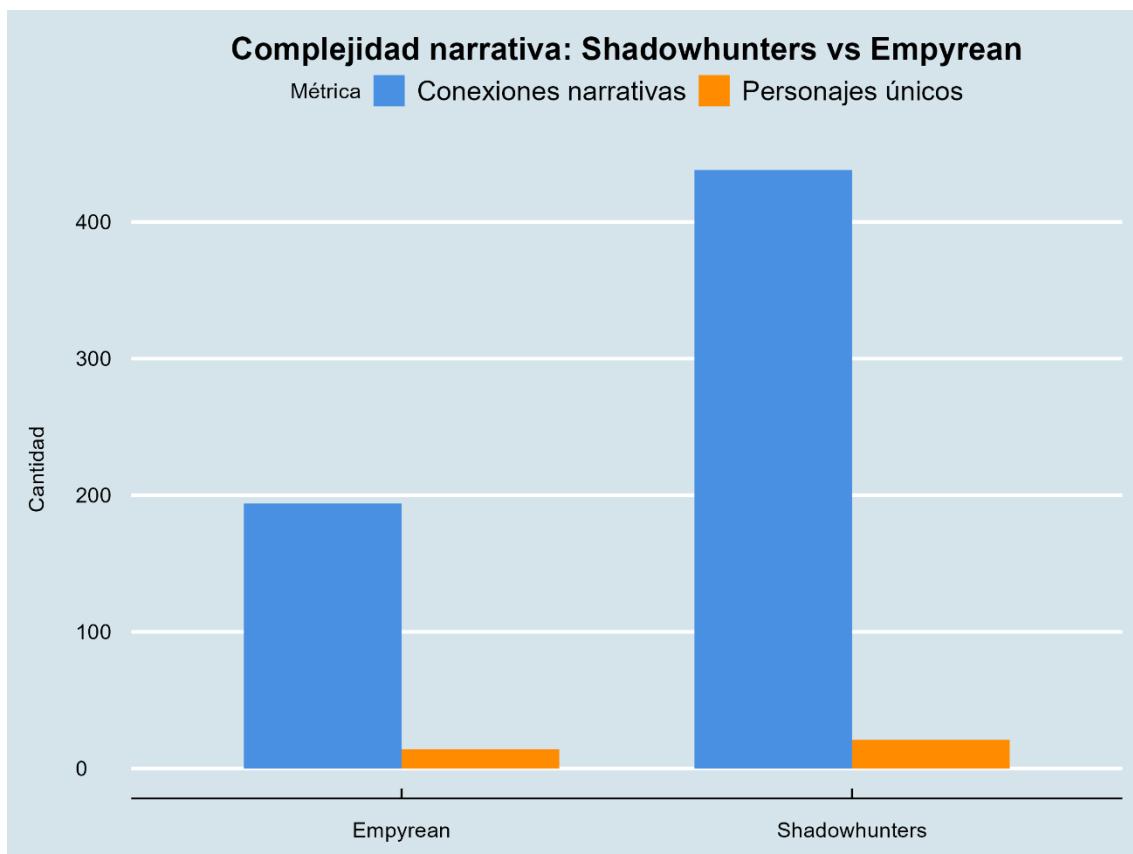


Figura A15. Complejidad narrativa: Shadowhunters vs Empyrean

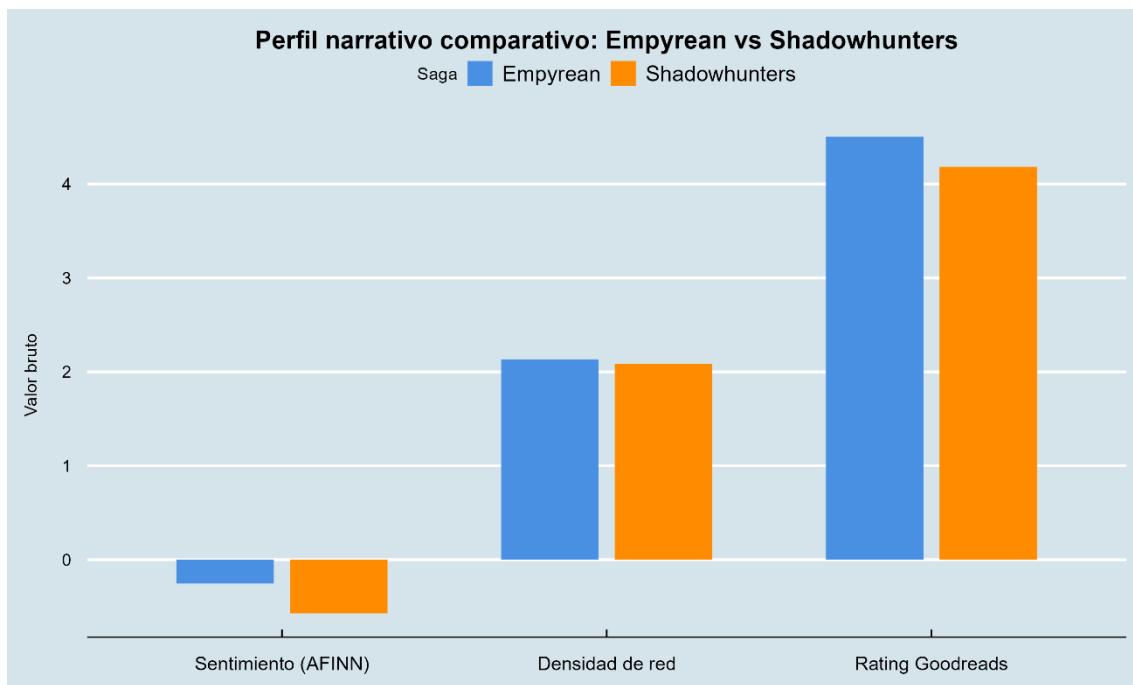


Figura A16. Perfil narrative comparativo: Empyrean vs Shadowhunters

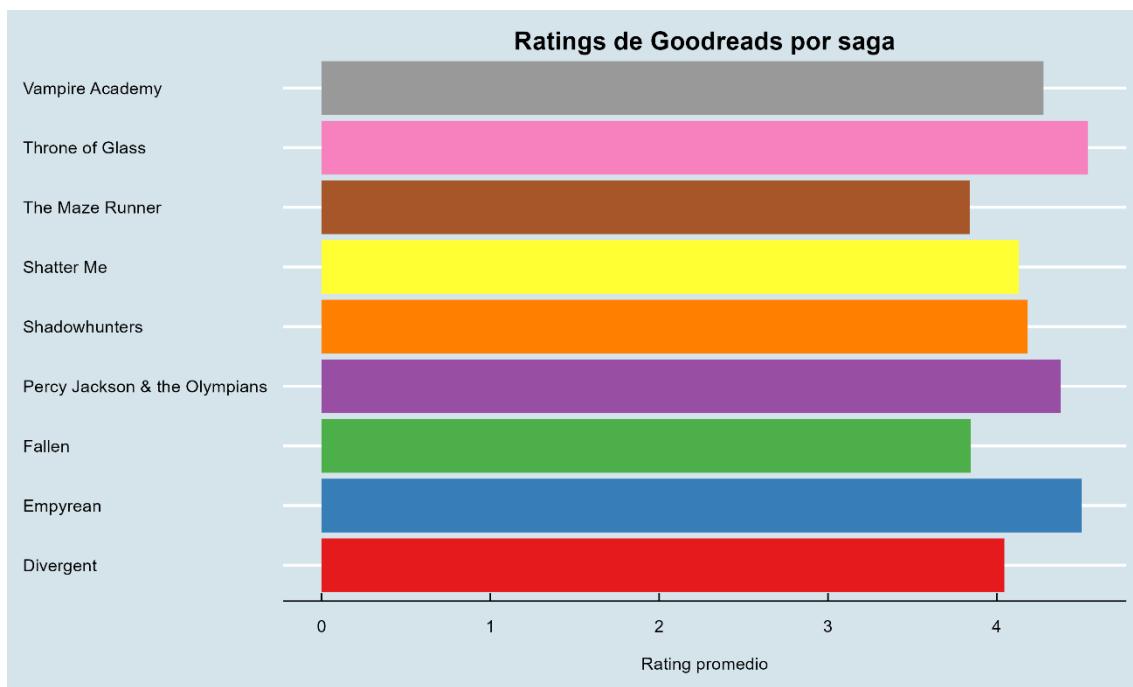


Figura A17. Rating de Goodreads por saga