**Logotipo

Descripción generada automáticamente**

**Desarrollo y Evaluación de Sistemas de Recomendación con Análisis de Sentimientos para la Optimización de Ventas en el Mercado Online: Un Enfoque Analítico de Datos**

Lina María Beltrán Durango

Mario Earles Otero Andrade

Monografía presentada para optar al título de Especialista en Analítica y Ciencia de Datos

Asesor  
Carlos Alberto Álvarez Henao, DSc en Ing. Civil

Universidad de Antioquia  
Facultad de Ingeniería

Especialización en Analítica y Ciencia de Datos

Medellín, Antioquia, Colombia

2023

| **Cita** | (Beltrán Durango & Otero Andrade, 2023) |
| --- | --- |
| **Referencia**  **Estilo APA 7 (2020)** | Beltrán Durango, L. M., & Otero Andrade, M. E. (2023). *Desarrollo y Evaluación de Sistemas de Recomendación con Análisis de Sentimientos para la Optimización de Ventas en el Mercado Online: Un Enfoque Analítico de Datos*. Universidad de Antioquia, Medellín, Colombia. |

**** 

Especialización en Analítica y Ciencia de Datos, CohorteV.

Centro de Investigación Ambientales y de Ingeniería (CIA).

|  | Diagrama  Descripción generada automáticamente con confianza media |
| --- | --- |

Centro de Documentación Ingeniería (CENDOI)

**Repositorio Institucional:** http://bibliotecadigital.udea.edu.co

Universidad de Antioquia - www.udea.edu.co

Rector: John Jairo Arboleda Céspedes.

Decano: Julio Cesar Saldarriaga Molina

Jefe departamento: Diego José Luis Botia Valderrama

El contenido de esta obra corresponde al derecho de expresión de los autores y no compromete el pensamiento institucional de la Universidad de Antioquia ni desata su responsabilidad frente a terceros. Los autores asumen la responsabilidad por los derechos de autor y conexos.

**Dedicatoria**

Texto de dedicatoria centrado.

**Agradecimientos**

Texto de agradecimientos centrado.

**Tabla de contenido**

[Resumen 9](#_heading=h.3rdcrjn)

[Abstract 10](#_heading=h.26in1rg)

[1.](#_heading=h.35nkun2) Descripción del problema 11

[1.1.](#_heading=h.1ksv4uv) Problema de negocio 11

[1.2.](#_heading=h.44sinio) Aproximación desde la analítica de datos 11

[1.3.](#_heading=h.z337ya) Origen de los datos 11

[1.4.](#_heading=h.3j2qqm3) Métricas de desempeño 11

[2.](#_heading=h.4i7ojhp) Objetivos 12

[2.1.](#_heading=h.2xcytpi) Objetivo general 12

[2.2.](#_heading=h.3whwml4) Objetivos específicos 12

[3.](#_heading=h.2bn6wsx) Datos 13

[3.1.](#_heading=h.qsh70q) Datos originales 13

[3.2.](#_heading=h.3as4poj) Datsets 13

[3.3.](#_heading=h.1pxezwc) Analítica descriptiva 13

[4.](#_heading=h.49x2ik5) Proceso de analítica 14

[4.1.](#_heading=h.2p2csry) Pipeline principal 14

[4.2.](#_heading=h.147n2zr) Preprocesamiento 14

[4.3.](#_heading=h.3o7alnk) Modelos 14

[4.4.](#_heading=h.23ckvvd) Métricas 14

[5.](#_heading=h.ihv636) Metodología 15

[5.1.](#_heading=h.32hioqz) Baseline 15

[5.2.](#_heading=h.1hmsyys) Validación 15

[5.3.](#_heading=h.41mghml) Iteraciones y evolución 15

[5.4 Herramientas 15](#_heading=h.2grqrue)

[6.](#_heading=h.vx1227) Resultados y discusión 16

[6.1.](#_heading=h.3fwokq0) Métricas 16

[6.2.](#_heading=h.1v1yuxt) Evaluación cualitativa 16

[6.3.](#_heading=h.4f1mdlm) Consideraciones de producción 16

[7.](#_heading=h.kgcv8k) Conclusiones 20

[8.](#_heading=h.34g0dwd) Recomendaciones 21

[Referencias 22](#_heading=h.1jlao46)

[Anexos 23](#_heading=h.43ky6rz)

[Anexo 1. Autoarchivo en Repositorio y documentos de interés 24](#_heading=h.xvir7l)

[Anexo 2. Gestor de citas y referencias de Microsoft Word Microsoft Word - Wikipedia 25](#_heading=h.1baon6m)

[Anexo 3. Citas y referencias de material legal (leyes, decretos, sentencias, etc.) 27](#_heading=h.3vac5uf)

[Anexo 4. Ortografía y gramática 30](#_heading=h.2afmg28)

[Anexo 5. Buscar, reemplazar y eliminar espacios (o palabras) 32](#_heading=h.pkwqa1)

[Anexo 6. Atajos de teclado útiles en Microsoft Word 33](#_heading=h.39kk8xu)

[Anexo 7. Sinónimos y antónimos 34](#_heading=h.1opuj5n)

[Anexo 8. Copiar y pegar sin formato 35](#_heading=h.48pi1tg)

[Anexo 9. Comparar dos documentos 36](#_heading=h.2nusc19)

[Anexo 10. Control de cambios 37](#_heading=h.1302m92)

[Anexo 11. Insertar salto de página 39](#_heading=h.3mzq4wv)

[Anexo 12. Recortar y abreviar direcciones web largas 40](#_heading=h.2250f4o)

**Lista de tablas**

[**Tabla 1**  Resultados del test PBQ-SF (Personality Belief Questionnaire Short Form) 19](#_heading=h.2u6wntf)

[**Tabla 2** Características demográficas y tipo de tratamiento de hemodiálisis y diálisis peritoneal con la adherencia (SMAQ) 20](#_heading=h.19c6y18)

[**Tabla 3** Categorías de la investigación 21](#_heading=h.3tbugp1)

**Lista de figuras**

[**Figura 1** Portada Normas APA séptima edición 2020 en inglés 22](#_heading=h.2r0uhxc)

[**Figura 2** Logo Universidad de Antioquia 22](#_heading=h.1664s55)

**Siglas, acrónimos y abreviaturas**

**APA** American Psychological Association

**Cms.** Centímetros

**ERIC** Education Resources Information Center

**Esp.** Especialista

**MP** Magistrado Ponente

**MSc** Magister Scientiae

**Párr.** Párrafo

**PhD** Philosophiae Doctor

**PBQ-SF** Personality Belief Questionnaire Short Form

**PostDoc** PostDoctor

**UdeA** Universidad de Antioquia

# Resumen

Este proyecto de monografía se enfoca en el desarrollo y evaluación de sistemas de recomendación con análisis de sentimientos para la optimización de ventas en el mercado en línea, utilizando datos de reseñas de Amazon hasta el año 2018. Nuestra investigación busca responder a preguntas clave, como la utilización de datos de reseñas para mejorar las recomendaciones de productos, la aplicación del análisis de sentimientos para comprender las preferencias de los usuarios, la efectividad de diversos modelos de sistemas de recomendación y la evaluación de su rendimiento.

Hemos empleado una variedad de modelos, incluyendo ***Collaborative Filtering Systems, Content-Based Systems, Hybrid Recommendation System*** y ***Empirical Methods in Natural Language Processing***. La evaluación se basó en métricas como RMSD, MAE y k-fold cross-validation.

Durante el proceso, enfrentamos desafíos relacionados con la calidad y volumen de los datos, así como la selección adecuada de modelos. Sin embargo, superamos estos obstáculos mediante técnicas de preprocesamiento y ajuste de parámetros.

Los resultados revelaron mejoras significativas en la precisión de las recomendaciones y la comprensión de las preferencias de los usuarios. Los modelos híbridos y los enfoques basados en contenido mostraron un rendimiento particularmente prometedor. Esta investigación presenta un enfoque analítico de datos para mejorar las ventas en línea al aprovechar el análisis de sentimientos y sistemas de recomendación avanzados. Nuestros resultados respaldan la utilidad de estos enfoques y abren oportunidades para aplicaciones en el comercio electrónico y más allá.

**Repositorios GitHub:**

<https://github.com/lmbd92/DataScienceMonograph>

<https://github.com/marioeoteroa/DataScienceMonograph>

*Palabras clave*: retail, online, comercio electrónico, ML, sistemas de recomendación, EMNLP.

# Abstract

This monograph project focuses on the development and evaluation of sentiment analysis-based recommendation systems for optimizing online sales using Amazon review data. Our research aims to address key questions, such as the utilization of review data to enhance product recommendations, the application of sentiment analysis to understand user preferences, the effectiveness of various recommendation system models, and the assessment of their performance.

We have employed a variety of models, including Collaborative Filtering Systems, Content-Based Systems, Hybrid Recommendation Systems, and Empirical Methods in Natural Language Processing. Evaluation was based on metrics such as RMSD, MAE, and k-fold cross-validation.

Throughout the process, we encountered challenges related to data quality and volume, as well as the appropriate model selection. However, we overcame these obstacles through preprocessing techniques and parameter tuning.

The results unveiled significant improvements in recommendation accuracy and user preference comprehension. Hybrid models and content-based approaches exhibited particularly promising performance. This research introduces a data-driven approach to enhance online sales by leveraging advanced sentiment analysis and recommendation systems. Our findings support the utility of these approaches and open doors for applications in e-commerce and beyond.

*Keywords*: retail, online, e-commerce, machine learning, recommendation systems, EMNLP.

GitHub Repositories:

<https://github.com/lmbd92/DataScienceMonograph>

<https://github.com/marioeoteroa/DataScienceMonograph>

# Descripción del problema

La presente investigación se enfoca en el desarrollo y evaluación de sistemas de recomendación con análisis de sentimientos para la optimización de ventas en el mercado en línea. Este enfoque analítico de datos se basa en la utilización de datos de reseñas de Amazon correspondientes al periodo comprendido entre mayo de 1996 y octubre de 2018.

El problema que abordamos en esta monografía se centra en mejorar la eficiencia de las ventas en el mercado en línea, aprovechando la información contenida en las reseñas de los productos y la interacción cliente - producto. Para ello, planteamos las siguientes preguntas de investigación:

1. ¿Cómo se pueden utilizar los datos de reseñas de Amazon para mejorar las recomendaciones de productos a los usuarios del comercio electrónico?
2. ¿Cómo se puede aprovechar el análisis de sentimientos para comprender mejor las preferencias y opiniones de los usuarios?
3. ¿Cuál es la efectividad de diferentes modelos de sistemas de recomendación en la optimización de ventas en línea?
4. ¿Cuál es la mejor manera de evaluar el rendimiento de los modelos propuestos?

Para abordar estas preguntas, emplearemos una variedad de modelos de sistemas de recomendación, incluyendo *Collaborative Filtering Systems, Content-Based Systems, Hybrid Recommendation System* y *Empirical Methods in Natural Language Processing*. Además, emplearemos métricas de evaluación como RMSD, MAE y k-fold cross-validation para medir el rendimiento de estos modelos.

En síntesis, esta investigación se centra en la mejora de las ventas en línea mediante la aplicación de técnicas de análisis de sentimientos y sistemas de recomendación avanzados. A través de la exploración de estos métodos, buscamos proporcionar respuestas concretas y soluciones efectivas para el problema de negocio planteado.

## Problema de negocio

El contexto en el que se enmarca este proyecto es el del comercio electrónico, donde las empresas buscan constantemente formas de mejorar sus estrategias de ventas en línea. En particular, se aborda el desafío de optimizar las ventas en un mercado en línea, utilizando como fuente de datos reseñas de productos proporcionados por la plataforma Amazon durante un periodo aproximado de 22 años.

El problema de negocio se presenta como sigue: en el entorno altamente competitivo del comercio electrónico, las marcas y principalmente las plataformas de ventas minoristas (retail) enfrentan el desafío de ofrecer a los usuarios recomendaciones de productos precisas y personalizadas. La calidad de estas recomendaciones tiene un impacto directo en la retención de clientes y, en última instancia, en las ventas. Para abordar este desafío, es fundamental comprender a fondo las preferencias, interacciones, comportamiento de compra y opiniones de los usuarios, lo cual se puede lograr mediante los métodos de filtro colaborativo, basado en contenido e interacciones, y como es nuestro objetivo, además, robustecer estos sistemas con el análisis de sentimientos de las reseñas de productos.

La pregunta central es cómo mejorar la efectividad de los sistemas de recomendación en línea, teniendo en cuenta la información contenida en las reseñas, las interacciones usuario - ítem y las emociones expresadas por los usuarios. Este problema de negocio es crucial, ya que una recomendación precisa no solo puede aumentar las ventas, sino también mejorar la experiencia del usuario y fortalecer la fidelidad hacia la plataforma de comercio electrónico.

Por lo tanto, el proyecto aborda el desafío empresarial de mejorar la optimización de ventas en línea en un contexto de comercio electrónico altamente competitivo, aprovechando el análisis de sentimientos y los sistemas de recomendación avanzados para ofrecer a los usuarios recomendaciones más precisas y relevantes.

.

## Aproximación desde la analítica de datos

La aproximación desde la analítica de datos en este proyecto se basa en la creación y evaluación de modelos predictivos avanzados para la recomendación de productos a clientes en un entorno de comercio electrónico. Estos modelos tienen como objetivo mejorar la eficiencia y la precisión de las recomendaciones, lo que a su vez puede aumentar las ventas y la satisfacción del cliente. A continuación, se describen las áreas clave de enfoque y se proporcionan referencias a artículos fundamentales de machine learning que respaldan la utilidad de los modelos predictivos desarrollados:

Modelos de Recomendación Personalizados: Se implementarán modelos de recomendación personalizados que aprovechan el comportamiento previo de los usuarios y el análisis de sentimientos de las reseñas de productos. Esto se alinea con la idea de ofrecer recomendaciones más relevantes y adaptadas a las preferencias individuales de los clientes. Un artículo clave en esta área es "Item-based collaborative filtering recommendation algorithms" (Sarwar et al., 2001), que aborda la importancia de los sistemas de filtrado colaborativo.

Análisis de Sentimientos: El análisis de sentimientos se aplicará a las reseñas de productos para comprender las opiniones y emociones de los usuarios. Esto permitirá una comprensión más profunda de las preferencias del cliente y se basará en estudios como "Mining and summarizing customer reviews" (Hu & Liu, 2004), que detalla técnicas para el análisis de reseñas.

Evaluación de Modelos: La evaluación de modelos se llevará a cabo utilizando métricas como RMSD, MAE y k-fold cross-validation para medir la precisión y el rendimiento de los sistemas de recomendación. El artículo "Evaluating recommendation systems" (Herlocker et al., 2004) proporciona una base sólida para la evaluación de sistemas de recomendación.

Enfoques Híbridos: Se explorarán enfoques híbridos que combinan sistemas de filtrado colaborativo y basados en contenido para aprovechar las fortalezas de ambos. El artículo "Hybrid Recommender Systems: Survey and Experiments" (Burke, 2002) ofrece una visión completa de estos enfoques.

## Origen de los datos

Amazon reviews es un conjunto de datos que contiene reseñas de productos y metadatos de Amazon, incluidas 142,8 millones de reseñas que abarcan desde mayo de 1996 hasta octubre de 2018. Este conjunto de datos incluye reseñas (calificaciones, texto, votos de utilidad), metadatos de productos (descripciones, información de categoría, precio, marca y características de imagen) y enlaces (también gráficos vistos/comprados).

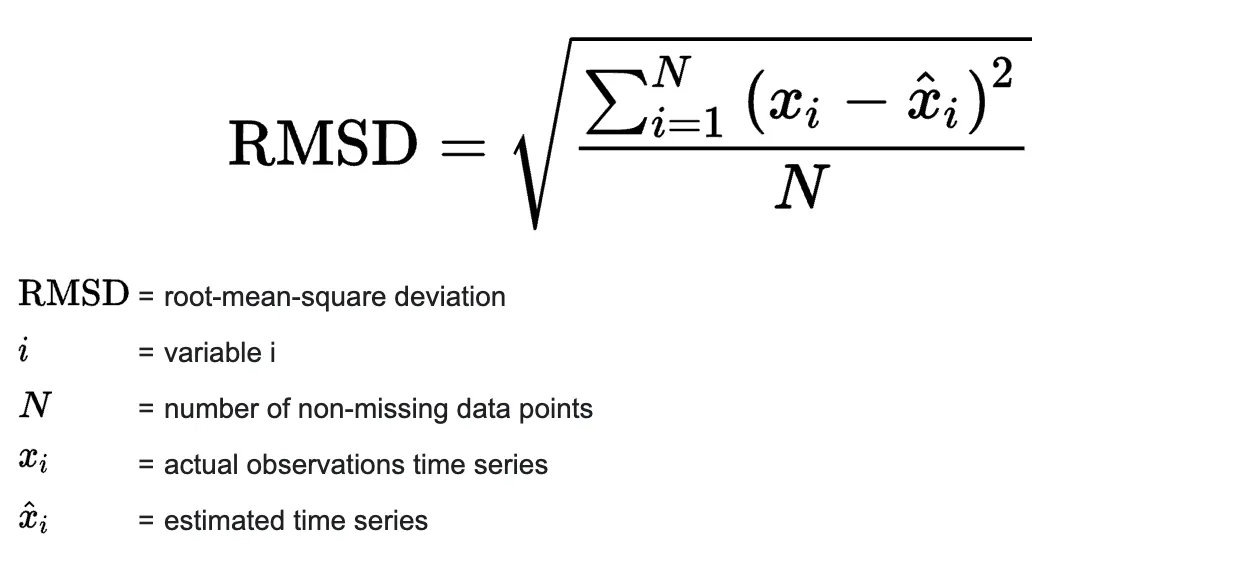
Amazon Reviews es un conjunto de datos para abordar la tarea de identificar si el sentimiento de una reseña de producto es positivo o negativo. Este conjunto de datos incluye reseñas de varias categorías de productos diferentes, tales como: libros (B) (2834 muestras), DVD (D) (1199 muestras), electrónica (E) (1883 muestras), artículos de cocina y para el hogar (K) (1755 muestras) y moda (883,636 muestras), para nuestro caso de investigación, hemos decidido tomar la categoría “moda”.

## Métricas de desempeño

En este ejercicio, utilizaremos varias métricas tanto de machine learning como de negocio para evaluar el desempeño de los modelos de recomendación desarrollados y su impacto en el proceso de negocio de ventas en línea, cabe mencionar que por el tiempo y no tener acceso a un escenario de simulación de negocio o un cliente real, las métricas de negocio se proponen a manera académica y como ideales a evaluar en un supuesto caso de implementación, sin embargo, sus resultados reales no podrán ser evaluados en este ejercicio. A continuación, se describen estas métricas y se establece un umbral mínimo en el contexto de aplicación:

**Métricas de Machine Learning (Desempeño de los Modelos):**

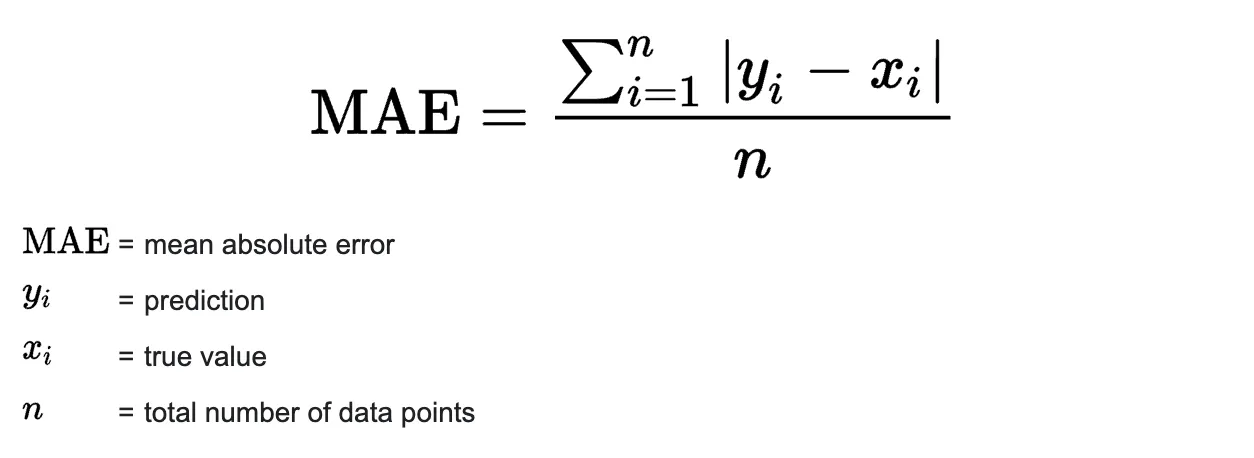
Root Mean Square Deviation (RMSD): Esta métrica se utiliza para medir la precisión de los modelos de recomendación al calcular la raíz cuadrada de la diferencia promedio entre las calificaciones reales y las calificaciones predichas. Un valor más bajo de RMSD indica una mayor precisión en las recomendaciones. En el contexto de aplicación, se considerará un umbral mínimo de RMSD para asegurar que las recomendaciones sean lo suficientemente precisas como para influir en las decisiones de compra de los usuarios.

  
Imagen tomada de [Statistics How To](https://www.statisticshowto.com/probability-and-statistics/regression-analysis/rmse-root-mean-square-error/#:~:text=If%20you%20don't%20like,square%20root%20of%20the%20result.)

RMSD es una métrica similar a MAE, pero tiene una penalización más fuerte cuando la predicción está muy lejos del valor real y una penalización más débil cuando la predicción está más cerca del valor real. Quitar los cuadrados de la diferencia entre los valores verdaderos y predichos en lugar de la suma de los valores absolutos. Esto asegura que el valor resultante sea siempre positivo y sea mayor cuando la diferencia es alta y menor cuando la diferencia es baja. Cuanto menor sea la puntuación de RMSD, mejor.

Se establecerá un umbral máximo de RMSD de 1.0. Esto significa que el RMSD debe ser menor o igual a 1.0 para considerar que los modelos de recomendación tienen un rendimiento adecuado. Un RMSD por debajo de 1.0 indicará que las recomendaciones son altamente precisas.

Mean Absolute Error (MAE): El MAE se emplea para evaluar la precisión absoluta de las predicciones de los modelos. Representa la diferencia promedio entre las calificaciones reales y las predicciones. Al igual que con el RMSD, un MAE más bajo indica una mayor precisión en las recomendaciones. Se establecerá un umbral mínimo de MAE para garantizar que las predicciones sean lo suficientemente precisas como para mejorar la experiencia del usuario.

  
Imagen tomada de [Statistics How To](https://www.statisticshowto.com/absolute-error/#:~:text=E%20%3D%20xexperimental%20%E2%80%93%20xtrue,true%20is%20the%20true%20measurement.)

Se fijará un límite máximo de MAE de 0.7. Un MAE igual o inferior a 0.7 será considerado como un nivel aceptable de precisión en las predicciones de los modelos.

**Métricas de Negocio (Desempeño del Proceso de Negocio):**

Incremento en las Ventas: Una métrica clave de negocio será el aumento en las ventas generadas por la implementación de los modelos de recomendación mejorados. Se establecerá un objetivo específico de aumento en las ventas en función de la mejora de las recomendaciones. El objetivo mínimo será lograr un aumento del 15% en las ventas después de la implementación de los modelos de recomendación mejorados. Esto significa que, en comparación con el estado previo a la implementación, las ventas deberán aumentar al menos un 15%.

Retención de Clientes: La retención de clientes es esencial en el comercio electrónico. Se medirá el impacto de los modelos en la retención de clientes y la lealtad a la plataforma. Se buscará lograr una tasa de retención de clientes del 80% o superior después de la implementación de los modelos. Esto significa que al menos el 80% de los clientes deberán realizar compras repetidas en la plataforma.

ROI Estimado: Se estimará el retorno de la inversión (ROI) en función de las mejoras en las ventas y la retención de clientes. Se establecerá un umbral mínimo de ROI que justifique la inversión en la implementación y operación de los modelos. Se considerará exitosa la implementación si se logra un ROI estimado de al menos el 20%. Esto significa que, en términos de ganancias adicionales generadas por las ventas y la retención de clientes, el retorno de la inversión será de al menos el 20%.

# Objetivos

## Objetivo general

El objetivo general de esta investigación consiste en concebir, desarrollar y evaluar sistemas de recomendación avanzados, empleando análisis de sentimientos como componente adicional al de los modelos tradicionales de este tipo de sistemas, con el propósito de optimizar las transacciones comerciales en el ámbito del comercio en línea. La delimitación temporal de esta indagación se sitúa entre el año 1996 y el año 2018 y se basa en la exploración de conjuntos de datos de reseñas de productos proporcionados por Amazon. El núcleo de este objetivo radica en la mejora substancial de la precisión y relevancia de las recomendaciones de productos, simultáneamente con una comprensión profunda de las preferencias y valoraciones emitidas por parte de los usuarios. En esencia, se persigue la formulación de un enfoque analítico de datos que desencadene un impacto positivo en el volumen de ventas y la fidelización de la clientela en el competitivo entorno del comercio electrónico contemporáneo.

En consecuencia, este estudio se posiciona como una contribución exploratoria a la convergencia de la analítica de datos y la estrategia empresarial, alentando una aproximación más precisa y personalizada para satisfacer las necesidades del consumidor y, al mismo tiempo, promoviendo una percepción más profunda de la interacción entre los usuarios y los productos ofrecidos en el mercado en línea. La consecución de este objetivo implica la implementación de modelos de recomendación innovadores, la aplicación de métricas de evaluación rigurosas y la identificación de umbrales críticos en términos de precisión y rendimiento de negocio.

## Objetivos específicos

Los objetivos específicos de esta investigación se formulan con el propósito de orientar las acciones a corto plazo que conduzcan al logro del objetivo general, y serán los siguientes:

* Realizar un análisis exhaustivo de la calidad y la estructura de los datos de reseñas proporcionados por Amazon en el periodo de tiempo descrito, identificando posibles problemas de limpieza y preprocesamiento.
* Diseñar, implementar y ajustar modelos de recomendación basados en colaboración, contenido, y enfoques híbridos, incorporando análisis de sentimientos para mejorar la precisión de las recomendaciones.
* Evaluar el Desempeño de los Modelos utilizando métricas de evaluación, como el RMSD, MAE y k-fold cross-validation, para medir el rendimiento de los modelos en términos de precisión y eficacia en la optimización de ventas.
* Determinar umbrales específicos para las métricas de precisión, como RMSD y MAE, que indiquen un rendimiento adecuado de los modelos. Estos umbrales servirán como criterios de aceptación para la implementación de los modelos en un entorno de negocio.
* (Opcional) Medir el impacto de los modelos en el negocio, evaluando el aumento en las ventas, la retención de clientes y el ROI estimado después de la implementación.
* Sintetizar los resultados de la investigación para generar conclusiones sólidas y ofrecer recomendaciones prácticas para la implementación de sistemas de recomendación basados en análisis de sentimientos en el comercio electrónico.

# Datos

## Datos originales

Describe los datos crudos de los que se dispone, incluyendo formatos, cómo se distribuyen los datos en ficheros, descripción de las columnas, formatos y resoluciones de imágenes, etiquetado de los mismos (para aprendizaje supervisado), tamaño de los datos (número de registros, tamaño total en MB o GB), etc.

Describe también el modo de acceso a los datos y restricciones que pudiera haber sobre los mismos.

## Datsets

Describe cómo, a partir de los datos, se construyen los datasets de entrenamiento y validación

## Analítica descriptiva

Muestra algunos ejemplos de tu dataset (imágenes, registros), incluyendo una breve analítica descriptiva (distribuciones de datos, etc.)

# Proceso de analítica

## Pipeline principal

Describe con un gráfico el flujo de trabajo general de los datos en tu proyecto. Añade o modifica secciones si lo consideras necesario

## Preprocesamiento

Describe las alternativas de preprocesamiento de datos que consideraste, aumentación de datos, etc.

## Modelos

Describe los distintos modelos que consideraste, sus configuraciones, etc.

## Métricas

Describe cómo calculas las métricas de desempeño ML y de negocio (Por ejemplo: con qué funciones de sklearn, tensorflow, etc.)

# Metodología

## Baseline

Describe tu primera iteración, sus resultados y los problemas técnicos que tuviste que resolver

## Validación

Describe tu proceso de validación, particiones de train/test/validación, etc.

## Iteraciones y evolución

Describe en qué te enfocaste en las iteraciones subsiguientes. No es necesario hacer un listado exhaustivo de las iteraciones, pero sí has de describir qué perseguías y qué obtuviste con cada grupo de iteraciones. P.ej. primero hicimos un conjunto de iteraciones para definir las opciones de preprocesamiento, luego hicimos unas iteraciones probando distintos modelos de tal tipo, etc.

## 5.4 Herramientas

Menciona las herramientas que usaste para tu proyecto

# Resultados y discusión

En los resultados se comunican los hallazgos y descubrimientos del estudio. Se incluyen tablas, figuras, diagramas y demás material demostrativo. Al narrar descriptivamente una figura, tabla, etc., en un párrafo, puedes insertar una referencia cruzada, es decir, un hipervínculo al elemento mencionado dentro o fuera de paréntesis, ejemplos: estos resultados se muestran en la **Tabla 1**. Igualmente, los datos son validados con otros instrumentos (**Tabla 2**, **Tabla 3**). Lineamientos que se establecen en la nueva versión de las Normas APA séptima edición (**Figura 1**). La producción intelectual institucional se publica en el Repositorio (**Figura 2**).

La discusión es la interpretación crítica y el análisis de los resultados, que surgen de las preguntas de investigación.

## Métricas

Describe los resultados numéricos de las métricas de las iteraciones que consideres más relevantes, junto con las configuraciones. Usa tablas o gráficas siguiendo el formato mostrado a continuación.

## Evaluación cualitativa

Realiza una evaluación cualitativa de los resultados, indicando casos de overfitting, underfitting, utilidad de los resultados, relación obtenida entre la métrica de ML y la de negocio, etc.

## Consideraciones de producción

Menciona consideraciones técnicas de una posible puesta en producción. Por ejemplo: condiciones de monitoreo del desempeño de los modelos, integración con streams de datos, servicios en la nube, etc.

**Tabla 1**  *Resultados del test PBQ-SF (Personality Belief Questionnaire Short Form)*

| Trastornos | Puntaje | Media \* | Desviación \* |
| --- | --- | --- | --- |
| Esquizoide | 2.1 | 11.8 | 5 |
| Paranoide | 3.5 | 6.9 | 5.2 |
| Antisocial | 2.2 | 9.3 | 5.1 |
| Narcisista | 1.6 | 7.4 | 4.3 |
| Histriónico | 2.8 | 6.3 | 4.5 |
| Límite | 3.1 | 5.9 | 4.4 |
| Por evitación | 2.0 | 10.2 | 4.9 |
| Por dependencia | 3.1 | 7.3 | 4.6 |
| Obsesivo compulsivo | 2.9 | 11.6 | 5 |
| Pasivo agresivo | 2.7 | 9.9 | 4.6 |

*\** Las medias y las desviaciones de esta prueba fueron obtenidas en población normal (no clínica).

*Fuente.* (Ramírez H. & Guzmán, 2011).

**Tabla 2**  
*Características demográficas y tipo de tratamiento de hemodiálisis y diálisis peritoneal con la adherencia (SMAQ)*

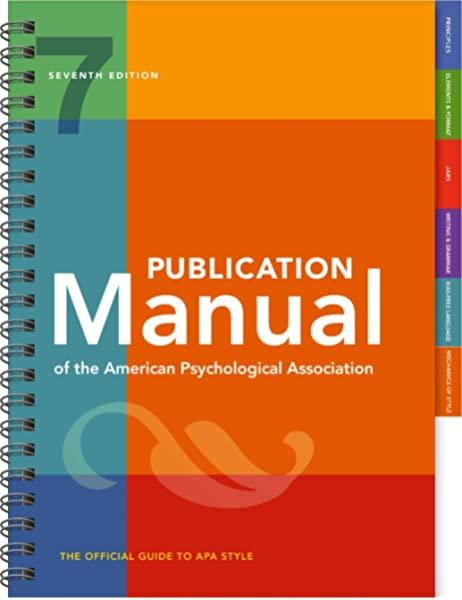
|  | Adherencia (SMAQ) | | | |  |
| --- | --- | --- | --- | --- | --- |
| No | | Sí | | Valor P |
| N | % | N | % |  |
| Sexo |  |  |  |  | 0.13 |
| Hombre | 55 | 58.5 | 45 | 70.3 |  |
| Mujer | 39 | 41.5 | 19 | 29.7 |  |
| Edad |  |  |  |  |  |
| 19 a 25 | 7 | 7.4 | 1 | 1.6 | 0.246 |
| 27 a 59 | 69 | 73.4 | 51 | 79.7 |  |
| 60 años o más | 18 | 19.1 | 12 | 18.8 |  |
| Estado civil |  |  |  |  | 0.036\* |
| Soltero | 26 | 27.7 | 11 | 17.2 |  |
| Casado / unión libre | 57 | 60.6 | 36 | 56.3 |  |
| Viudo / divorciado | 11 | 11.7 | 17 | 26.6 |  |
| Ocupación |  |  |  |  | 0.045\* |
| Hogar | 37 | 39.4 | 15 | 23.4 |  |
| Empleado | 8 | 8.5 | 3 | 4.7 |  |
| Otro | 49 | 52.1 | 46 | 71.9 |  |
| \*Valor p < 0.05 | |  |  |  |  |

**Tabla 3**  
*Categorías de la investigación*

| **Categoría** | **Subtemas** | **Definiciones** |
| --- | --- | --- |
| **Memoria** | Memoria de trabajo | Es una función ejecutiva cerebral que se encarga del almacenamiento de la información que llega del exterior, con la cual se construyen los conocimientos. |
| Bases neurológicas | Las bases neurológicas de la memoria se relacionan con el lóbulo prefrontal. |
| **Estrategias** | Estrategias lúdicas | Las estrategias lúdicas son las acciones que planean los docentes, donde intervienen el disfrute, el goce y el placer en la construcción de los conocimientos. |
| Estrategias didácticas | Las estrategias didácticas son las acciones de los maestros a partir de las cuales los estudiantes construyen los conocimientos; pueden ser estrategias de aprendizaje o de enseñanza. |
| **Proceso de aprendizaje del idioma inglés** | Aprendizaje | Es un proceso cognitivo de asimilación, donde los estudiantes unen las nuevas informaciones con saberes previos, a través de un proceso cognitivo. |
| Estilos de aprendizaje | El aprendizaje se puede dar a través de los sentidos. Es así como existen, básicamente, tres estilos o formas de adquirir los conocimientos: el visual, el auditivo y el kinestésico. |

*Nota.* Adaptado de Ruiz Rojas (2014).

**Figura 1** *Portada Normas APA séptima edición 2020 en inglés*



*Nota.* Fuente https://bit.ly/2IyrZao (American Psychological Association, 2020).

**Figura 2** *Logo Universidad de Antioquia*

****

*Nota.* Fuente http:/www.udea.edu.co

# Conclusiones

Son las interpretaciones finales que recopilan los datos de la investigación, describe lo que se obtuvo, qué se logró y cuáles son los resultados. Guardan relación directa con lo que se mencionó en el planteamiento del problema y en los objetivos. Haz un breve juicio crítico de tu proyecto.

# Recomendaciones

Las recomendaciones son las futuras y posibles líneas de investigación que llevarán a resolver problemas relacionados con la presente investigación.

# Referencias

Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2001). *Item-based collaborative filtering recommendation algorithms*. In Proceedings of the 10th International Conference on World Wide Web (WWW '01) (pp. 285-295).

Hu, M., & Liu, B. (2004). *Mining and summarizing customer reviews*. In Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '04) (pp. 168-177).

Herlocker, J. L., Konstan, J. A., Borchers, A., & Riedl, J. (2004). *Evaluating recommendation systems*. ACM Transactions on Information Systems, 22(1), 5-53.

Burke, R. (2002). *Hybrid Recommender Systems: Survey and Experiments*. User Modeling and User-Adapted Interaction, 12(4), 331-370.

# 

# Anexos

En los anexos se incluye material complementario que apoya la documentación investigativa, tales como consentimientos informados, entrevistas, material fotográfico, etc. Evite incluir material que puede estar protegido por derechos de autor, tales como pruebas psicológicas, fragmentos de libros, artículos de revistas, patentes, etc. Recuerda no incluir en tu documento datos de personas o entidades objetos de la investigación, tales como nombres, apellidos, cédulas, números telefónicos, consentimientos informados con datos personales (Resolución 8430 de 1993), nombres de empresas sin el consentimiento escrito del representante legal, fotografías en primer plano de personas (especialmente de menores de edad) y demás información que pueda contravenir los principios emitidos en la Ley Estatutaria 1581 de 2012 (Ley de protección de datos personales).

Los siguientes anexos contienen documentos de interés para el proceso de trabajo de grado, así como trucos y recomendaciones que surgen constantemente en la elaboración de un documento en Word.

## Anexo 1. Autoarchivo en Repositorio y documentos de interés

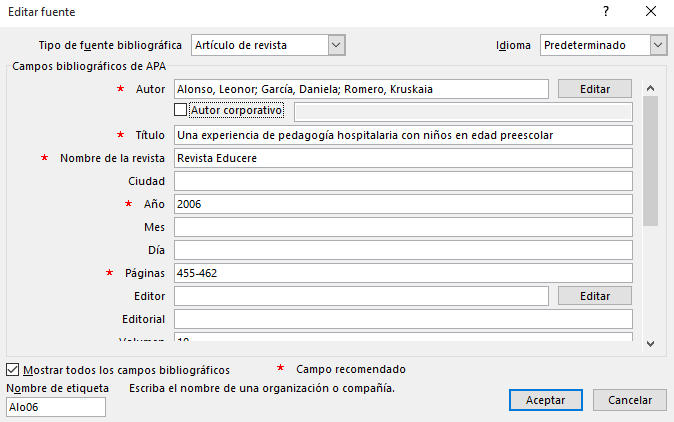
Al terminar todos los aspectos metodológicos, de redacción, de estructura y diagramación de tu tesis en Word, y con previo aval de la unidad académica, exporta el documento a versión PDF. Prepara también los anexos, si los tiene. Posteriormente, realiza la gestión de autoarchivo en el Repositorio Institucional <http://bibliotecadigital.udea.edu.co>, procedimiento que puedes consultar en video o versión PDF:

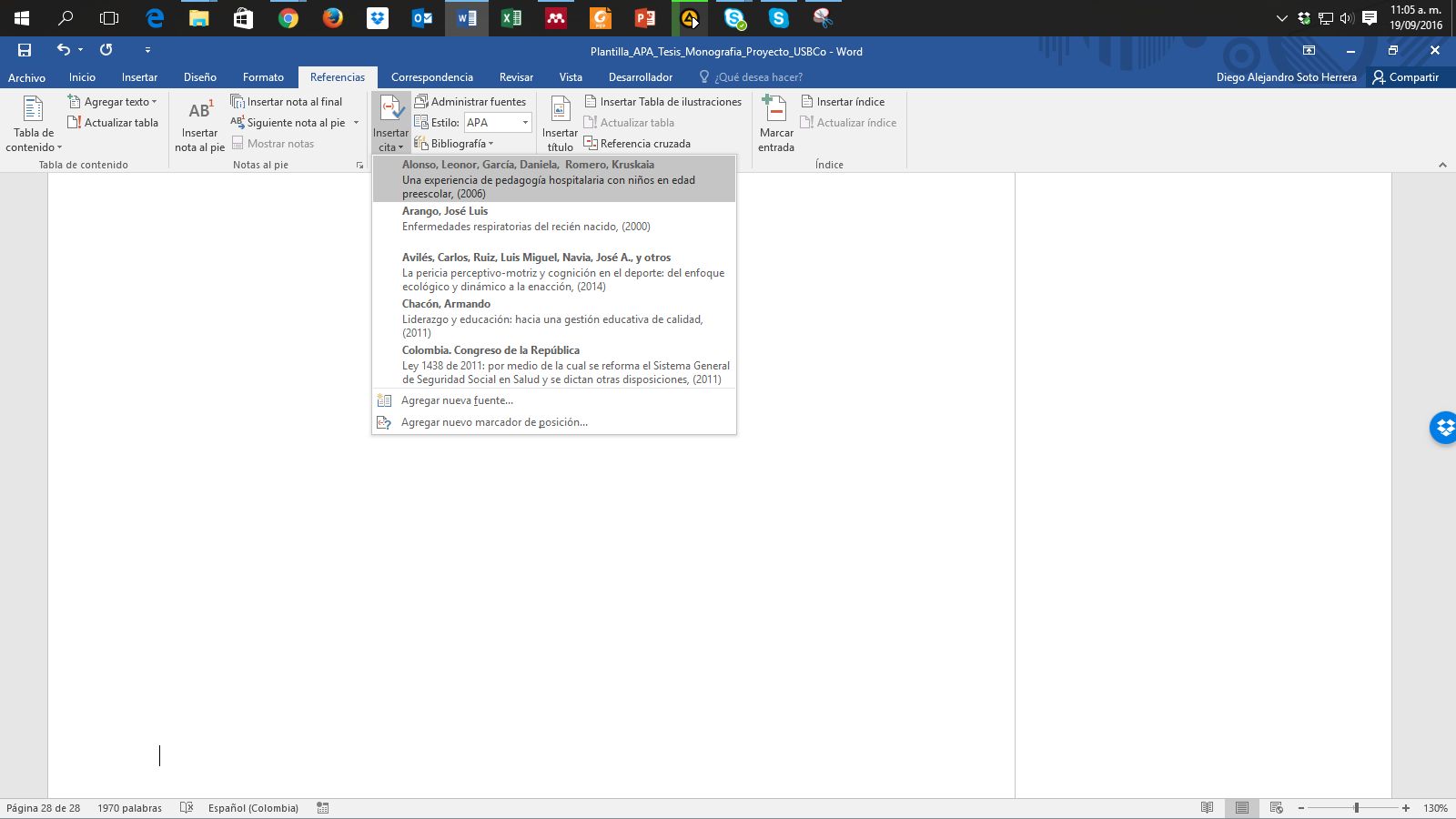
* Gestión de autoarchivo trabajos de grado (video): <https://bit.ly/3wx9U0E>
* Instructivo para el autoarchivo de trabajos de grado en el Repositorio Institucional Universidad de Antioquia (PDF): <https://bit.ly/3fOWbfB>

Recuerda que ya no se entregan trabajos de grado en CD-ROM, únicamente mediante formato digital a través del Repositorio Institucional. Otros documentos de interés para el proceso de entrega de trabajos de grado:

* Formulario institucional de entrega y autorización de trabajos de grado en la Universidad de Antioquia (diligenciar solo para 2 autores o más): <https://bit.ly/2Q0sc9P>
* Resolución Rectoral 47233 (21 de agosto de 2020): por la cual se establecen los lineamientos para la entrega de la producción académica de pregrado y posgrado en sus diferentes formatos y presentaciones al Repositorio Institucional del Departamento de Bibliotecas: <https://bit.ly/2R629hP>
* Políticas del Repositorio Institucional de la Universidad de Antioquia: <https://bit.ly/3t6dcG9>

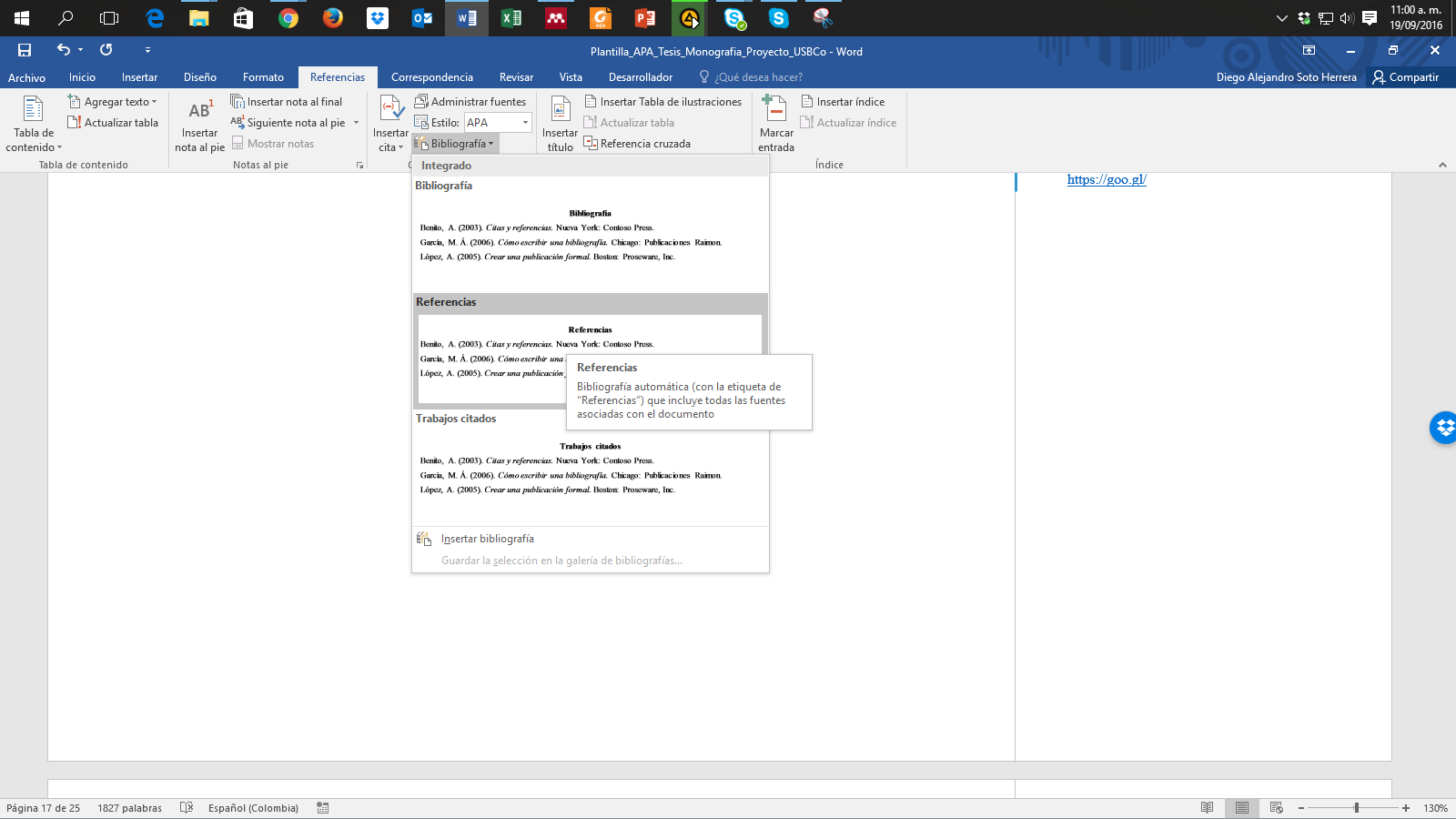
## Anexo 2. Gestor de citas y referencias de Microsoft Word Microsoft Word - Wikipedia

Ingresar las fuentes: Referencias > Administrar fuentes > Nuevo:

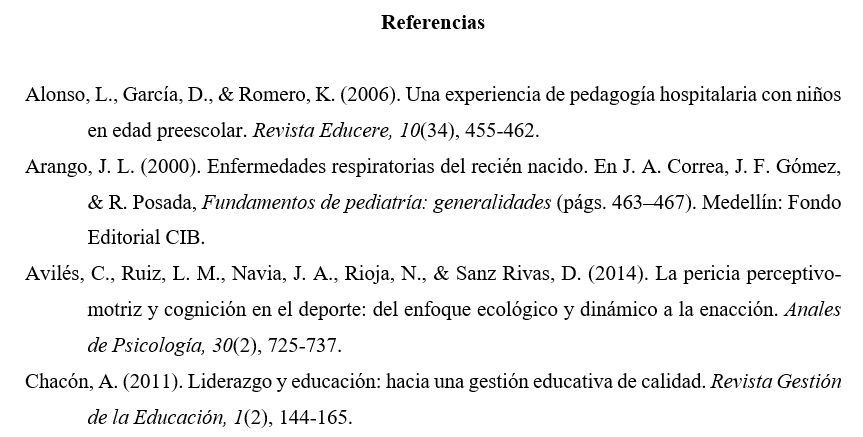
Insertar cita en el texto: Referencias > Insertar cita > Clic en fuente seleccionada:

Cita insertada dentro del texto:

Algunas experiencias significativas se han descrito mediante la pedagogía en hospitales con niños en edad preescolar (Alonso et al., 2006).

Insertar referencias (bibliografía): Referencias > Bibliografía > Referencias

Sección Referencias insertada:



## Anexo 3. Citas y referencias de material legal (leyes, decretos, sentencias, etc.)

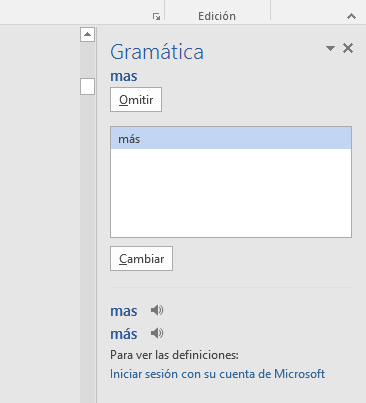
Uno de los aspectos que más puede causar confusión en Normas APA es lo referente a la citación de material legal y jurídico; de hecho, la misma American Psychological Association refiere al uso del manual internacional “Bluebook: A Uniform System of Citation” (este estilo está incluido en Zotero, Mendeley y otros gestores bibliográficos), pues estos dos estilos difieren en su formato de cita y referencia, pues las publicaciones legales citan las referencias al pie de página, en tanto que en el estilo APA se ubican todas las fuentes bibliográficas, incluyendo aquellas de materiales legales, en la lista de referencias. Si deseas conocer y adaptar los lineamientos del Bluebook, puedes consultarlos en <https://www.legalbluebook.com/>; asimismo, algunos ejemplos del manual de la APA están basados en el sistema jurídico estadounidense, lo que sin duda podría causar cierto conflicto con el entorno legal colombiano; ambos serán aceptados en los trabajos de grado y tesis de la Universidad de Antioquia. Sin embargo, para facilitar y adaptar las citas y referencias al sistema legal y jurídico colombiano, recomendamos los siguientes lineamientos basados en Normas APA como primera alternativa de citación y referenciación de los materiales más comunes en Colombia, a saber, leyes, decretos, sentencias, resoluciones, códigos, constitución política, entre otros. La primera recomendación está basada en el prefijo Colombia. como autor corporativo estatal, luego la subentidad y el año. Ejemplos de citas y referencias (se incluye un ejemplo internacional de España).

| **Cita (al interior del texto)** | **Referencias** |
| --- | --- |
| (Colombia. Presidencia de la República, 1991) | Colombia. Presidencia de la República. (1991). *Constitución Política de Colombia.* Presidencia de la República. |
| (Colombia. Congreso de la República, 1994) | Colombia. Congreso de la República. (1994). *Ley 133 de 1994 (mayo 23): por la cual se desarrolla el Decreto de Libertad Religiosa y de Cultos, reconocido en el artículo 19 de la Constitución Política*. Diario Oficial. |
| (Colombia. Presidencia de la República, 1998) | Colombia. Presidencia de la República. (1998). *Decreto 1504 de 1998: por el cual se Reglamenta el Manejo del Espacio Público en los Planes de Ordenamiento Territorial*. Diario Oficial. |
| (Colombia. Congreso de la República, 2014) | Colombia. Congreso de la República. (2014). *Ley 1733 de 2014: Ley Consuelo Devis Saavedra, mediante la cual se regulan los servicios de cuidados paliativos para el manejo integral de pacientes con enfermedades terminales, crónicas, degenerativas e irreversibles en cualquier fase de la enfermedad de alto impacto en la calidad de vida*. Diario Oficial. |
| (Colombia. Corte Constitucional, 2003a) | Colombia. Corte Constitucional. (2003a). *Sentencia SU.805 de 2003: vía de hecho en proceso de lanzamiento por ocupacion de hecho / debido proceso de querellados - vulneración por actuaciones arbitrarias. M. P. Jaime Córdoba Triviño*. Corte Constitucional. |
| (Colombia. Corte Constitucional, 2003b) | Colombia. Corte Constitucional. (2003b). *Sentencia T-361 de 2003: acción de tutela instaurada por Elkis Patricia Jiménez Castro contra la Universidad Cooperativa de Colombia – Seccional Santa Marta. M. P. Manuel José Cepeda Espinosa*. Corte Constitucional. |
| (Colombia. Corte Constitucional, 2006) | Colombia. Corte Constitucional. (2006). *Sentencia T-264 de 2006: acción de tutela instaurada por Fanny Stella Lesmes Galarza, en representación de su menor hijo Paul Andrés Rodríguez Lesmes contra la Universidad de los Andes. M. P. Jaime Araújo Rentería*. Corte Constitucional. |
| (Colombia. Ministerio de Salud y Protección Social, 2012) | Colombia. Ministerio de Salud y Protección Social. (2012). *Resolución 4331 de 2012 (diciembre 19): por medio de la cual se adiciona y modifica parcialmente la Resolución 3047 de 2008 modificada por la resolución 416 de 2009*. Diario Oficial. |
| (Colombia. Ministerio de Hacienda y Crédito Público. Superintendencia Financiera, 2006) | Colombia. Ministerio de Hacienda y Crédito Público. Superintendencia Financiera. (2006). *Circular Externa 048 de 2006 (diciembre 22)*. Superfinanciera. |
| (Colombia. Ministerio de Minas y Energía, 2010) | Colombia. Ministerio de Minas y Energía. (2010). *Reglamento técnico de iluminación y alumbrado público*. Ministerio de Minas y Energía. |
| (España. Ministerio de Trabajo y Asuntos Sociales. Instituto Nacional de Seguridad e Higiene en el Trabajo, 1996) | España. Ministerio de Trabajo y Asuntos Sociales. Instituto Nacional de Seguridad e Higiene en el Trabajo. (1996). *Evaluación de riesgos laborales*. INSHT. |
| (Colombia. Ministerio de Ambiente, Vivienda y Desarrollo Territorial, 2007) | Colombia. Ministerio de Ambiente, Vivienda y Desarrollo Territorial. (2007). *Decreto 3600 de 2007: por el cual se Reglamentan las Disposiciones de las Leyes 99 de 1993 y 388 de 1997 Relativas a las Determinantes de Ordenamiento del Suelo Rural y al Desarrollo de Actuaciones Urbanísticas de Parcelación y Edificación en este tipo de suelo y se adoptan otras disposiciones.* Diario Oficial. |
| (Colombia. Ministerio de Comunicaciones, 2001) | Colombia. Ministerio de Comunicaciones. (2001). *Resolución 000797 DE 2001 (junio 8): por la cual se atribuyen unas bandas de frecuencias radioeléctricas para su libre utilización dentro del territorio nacional*. Diario Oficial. |
| (Colombia. Ministerio de Educación Nacional, 2006) | Colombia. Ministerio de Educación Nacional. (2006). *Plan Decenal de Educación 2006-2016: Pacto Social por la Educación*. Ministerio de Educación Nacional. |
| (Colombia. Congreso de la República, 2010) | Colombia. Congreso de la República. (2010). *Código penal y de procedimiento penal anotado*. Leyer. |
| (Colombia. Congreso de la República, 2006) | Colombia. Congreso de la República. (2006). *Ley 1098 de 2006 (noviembre 8): por la cual se expide el Código de la Infancia y la Adolescencia en Colombia*. Diario Oficial. |
| (Colombia. Contraloría General de la Nación, 2003) | Colombia. Contraloría General de la Nación. (2003). *La deserción escolar en la educación básica media*. Contraloría General de la Nación. |
| (Colombia. Unidad Nacional para la Gestión del Riesgo de Desastres, 2012) | Colombia. Unidad Nacional para la Gestión del Riesgo de Desastres. (2012). *Guía para la formulación del Plan Municipal de Gestión del Riesgo de Desastres*. Sistema Nacional de Gestión del Riesgo de Desastres. |
| (Colombia. Departamento Administrativo Nacional de Estadística, 2011) | Colombia. Departamento Administrativo Nacional de Estadística. (2011). *Encuesta de convivencia escolar y circunstancias que la afectan - ECECA, para estudiantes de 5o a 11o de Bogotá*. DANE. |
| (Colombia. Departamento Administrativo de Ciencia, Tecnología e Innovación, 2015) | Colombia. Departamento Administrativo de Ciencia Tecnología e Innovación. (2015). *Resultados finales de la Convocatoria Nacional para el Reconocimiento y Medición de Grupos de Investigación, Desarrollo Tecnológico o de Innovación y para el Reconocimiento de Investigadores del SNCTeI*. Colciencias. |
| (Colombia. Procuraduría General de la Nación, 2012) | Colombia. Procuraduría General de la Nación. (2012). *Financiamiento del Sistema General de Seguridad Social en Salud: seguimiento y control preventivo a las políticas públicas*. Procuraduría General de la Nación. |

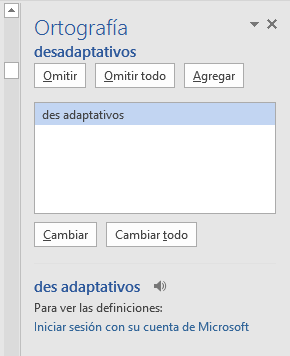
## Anexo 4. Ortografía y gramática

La ortografía y la gramática hacen parte fundamental del trabajo de grado; al finalizar la redacción de tu escrito, realiza una revisión ortográfica de todo el documento. En todo caso, siempre será recomendada y preferible la labor de un corrector de estilo que corrija redacción, ortografía, sintaxis, coherencia, citas, referencias y demás aspectos de estilo. En Microsoft Word, oprime la tecla F7. Tendrá dos tipos de sugerencias: Gramática y Ortografía, donde tendrás la opción de:

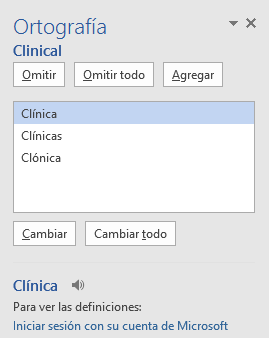
* “Cambiar”, si consideras que efectivamente había un error, ejemplo (mas, sin tilde):



* “Omitir”, si a pesar de la sugerencia consideras que está correctamente, ejemplo (desadaptativos):



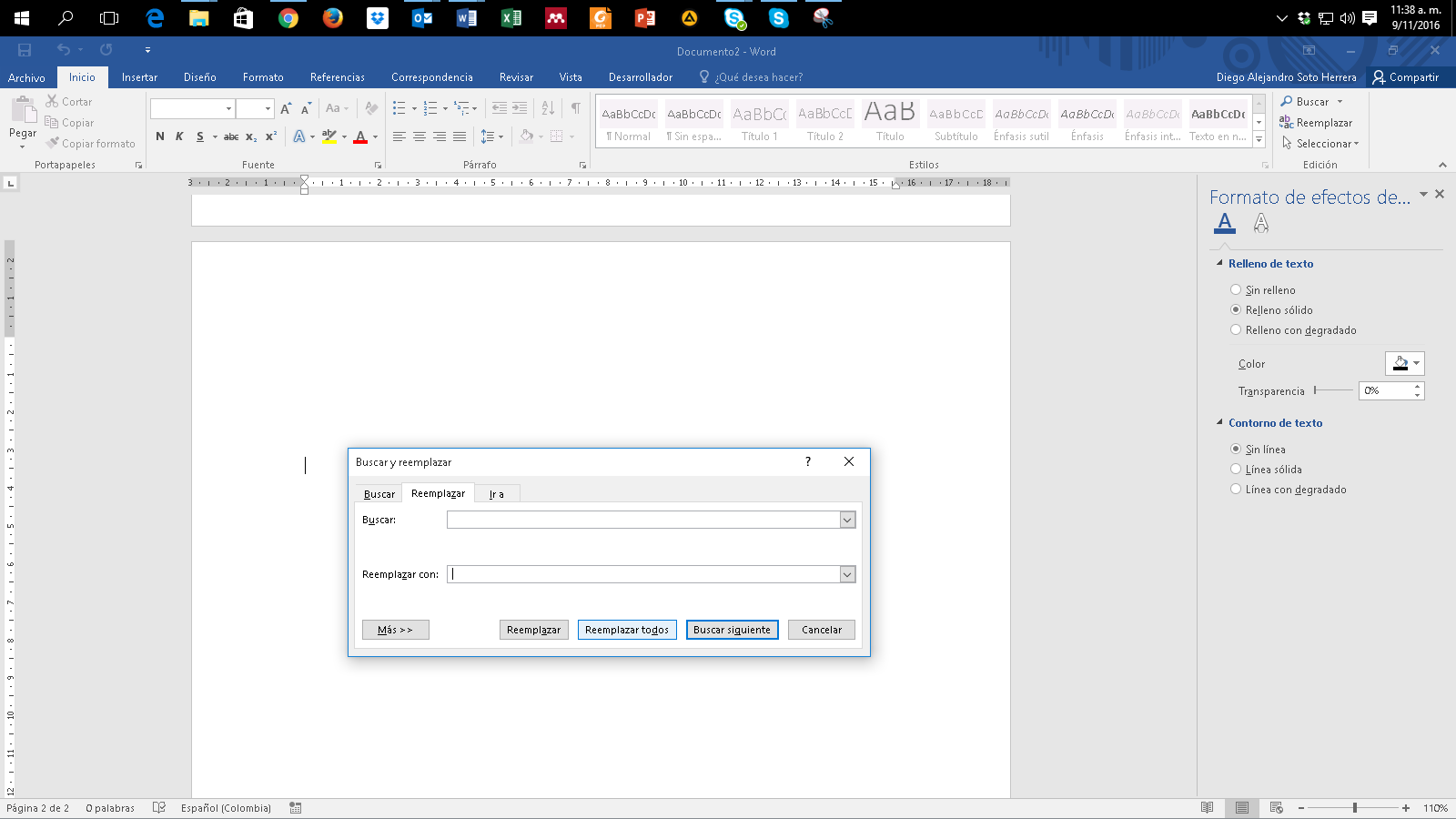
* “Omitir todo” si deseas ignorar la sugerencia de esa palabra u oración en todo el texto.
* “Agregar” si deseas incluir esa palabra en el diccionario en futuras revisiones.



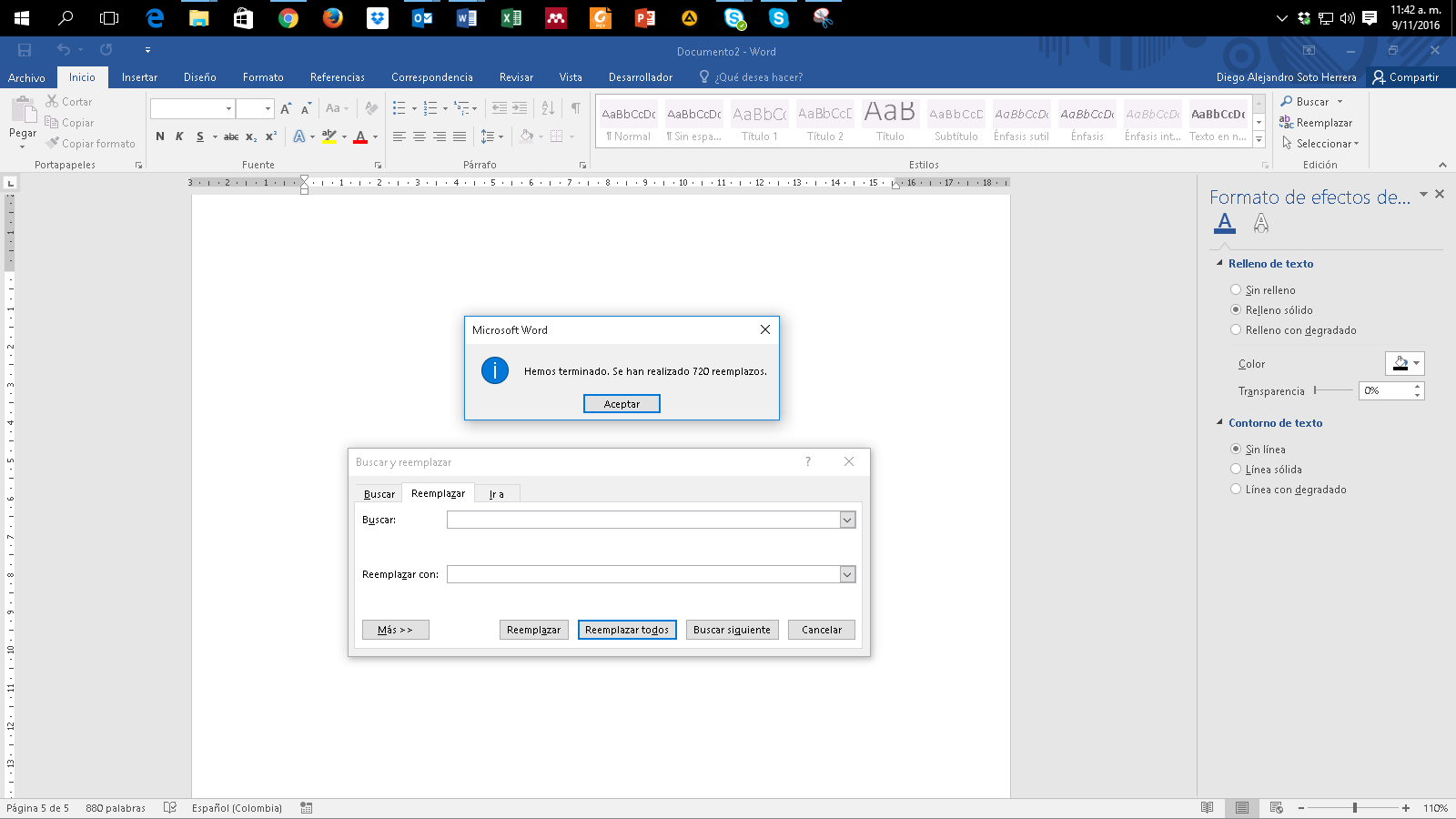
Ten precaución en aceptar cambios sugeridos, pues Microsoft Word no tiene la capacidad completa de interpretar con precisión algunos aspectos de la redacción o gramática en lengua española.

## Anexo 5. Buscar, reemplazar y eliminar espacios (o palabras)

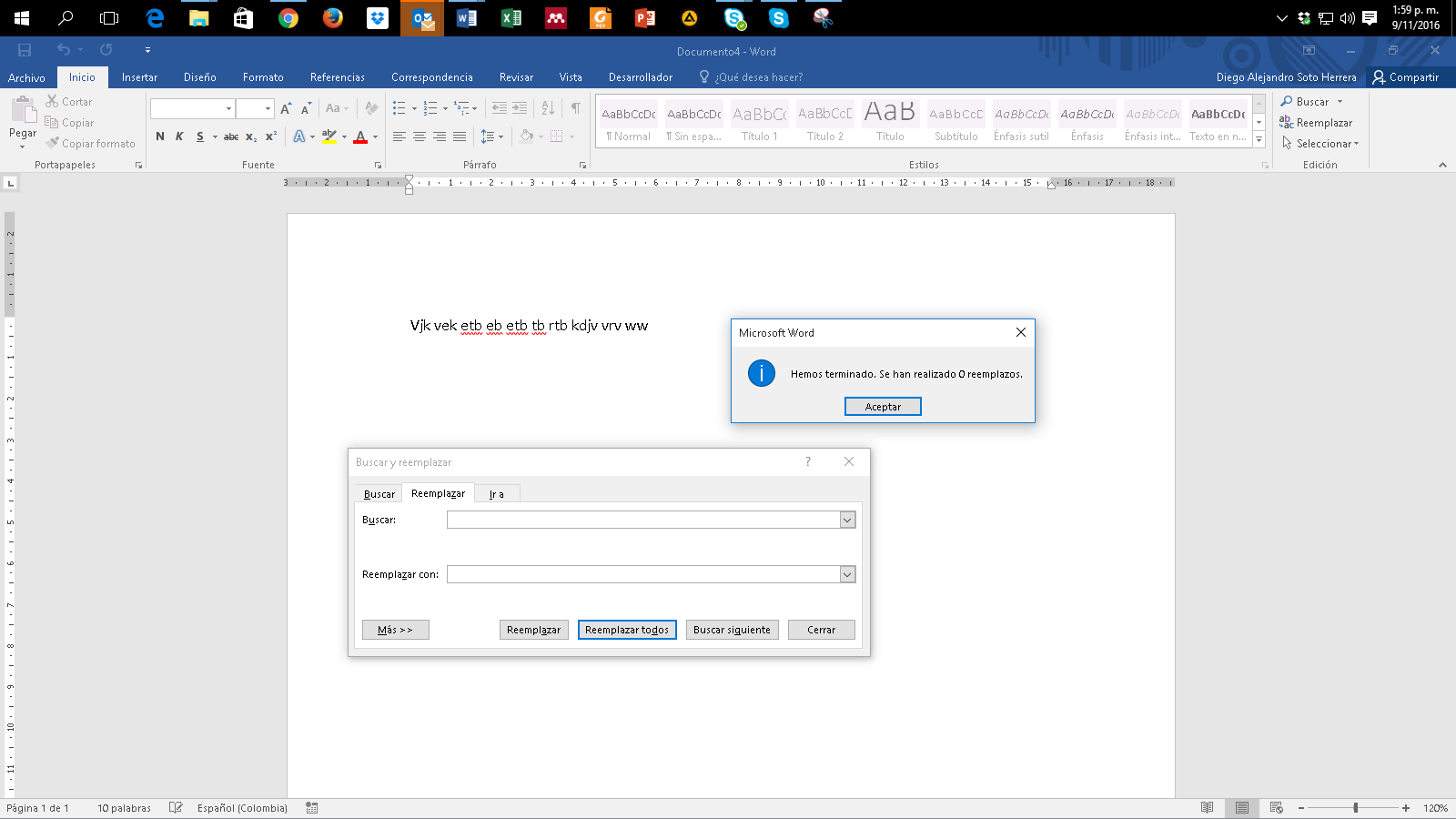
Uno de los errores más comunes al redactar un texto es incluir dobles, triples o cuádruples espacios, que en esencia son casi imperceptibles pero que afectan la distribución del documento. Para eliminar dobles espacios, realiza estos pasos: Ctrl + L (Reemplazar) > Digita 2 espacios en “Buscar” > Digita 1 espacio en “Reemplazar” > Clic en “Reemplazar todos”.



Word notificará cuántos espacios dobles se eliminaron y se han reemplazado por un espacio.



Haz clic en “Reemplazar todos” para eliminar dobles espacios que quedan, por ejemplo, de triples o cuádruples espacios anteriores, hasta que aparezca este mensaje:



## Anexo 6. Atajos de teclado útiles en Microsoft Word

**Ctrl +**

A= Abrir

B= Buscar

C= Copiar

D= Alinear derecha

E= Seleccionar todo

G= Guardar

H= Sangría

I= Ir a página

J= Justificar

K= Cursiva

L= Reemplazar

M= Formato

N= Negrilla

O= Disminuir tamaño

P= Imprimir

Q= Alinear izquierda

R= Cerrar documento

S= Subrayado

T= Centrar

U= Nuevo documento

V= Pegar

X=Cortar

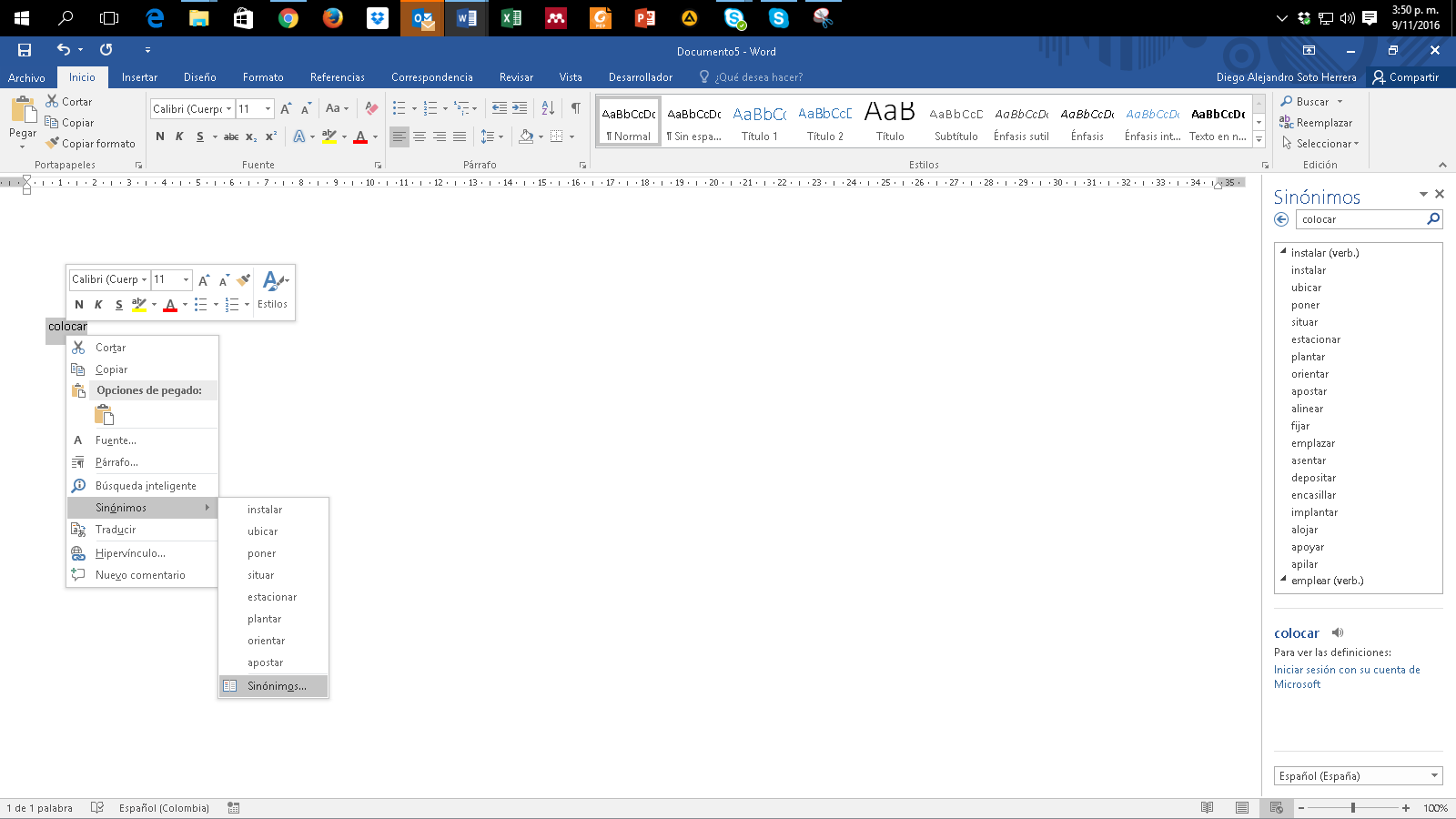
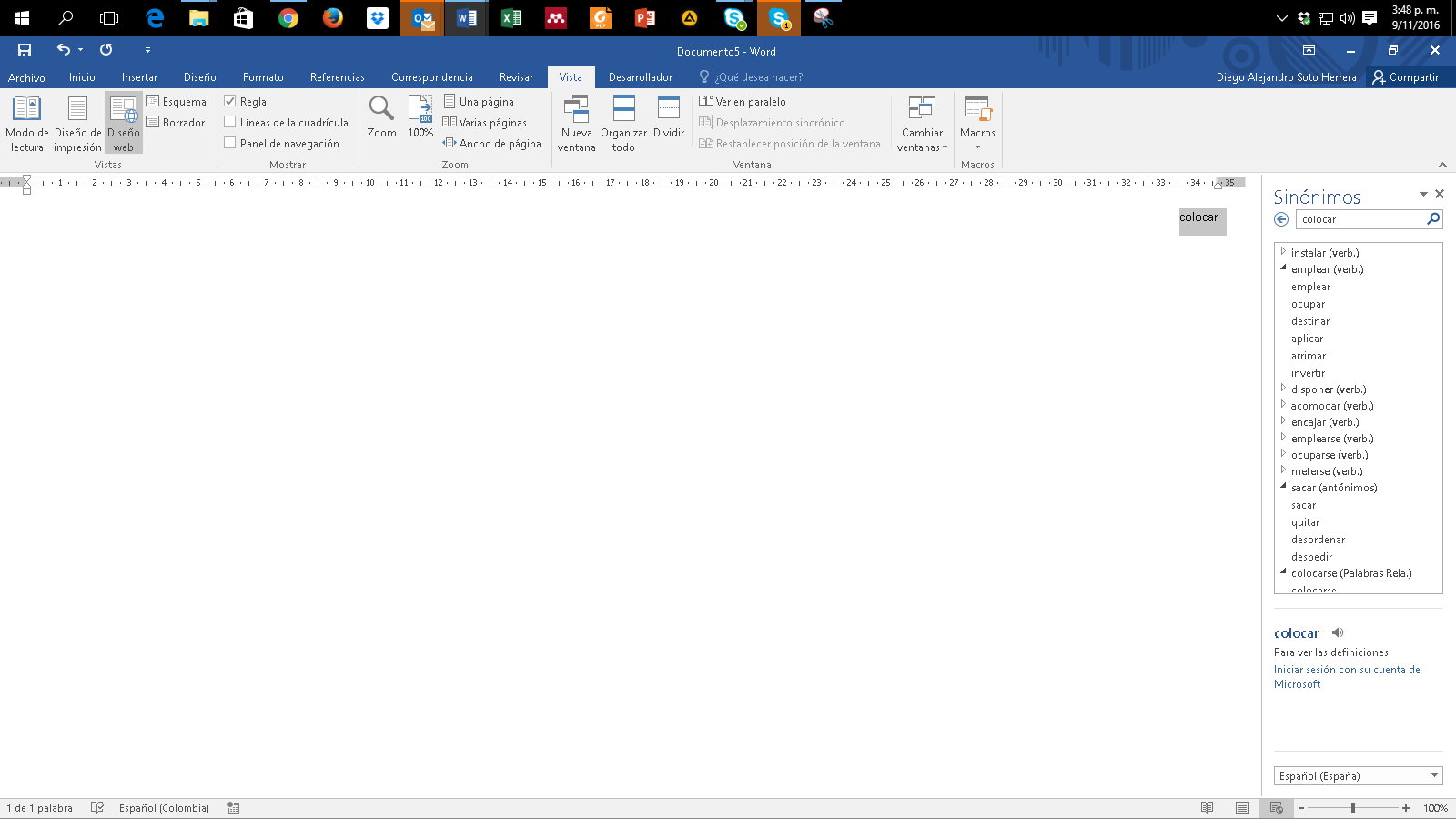
Y= Rehacer

Z= Deshacer

Lista completa de atajos https://bit.ly/3oHliCj

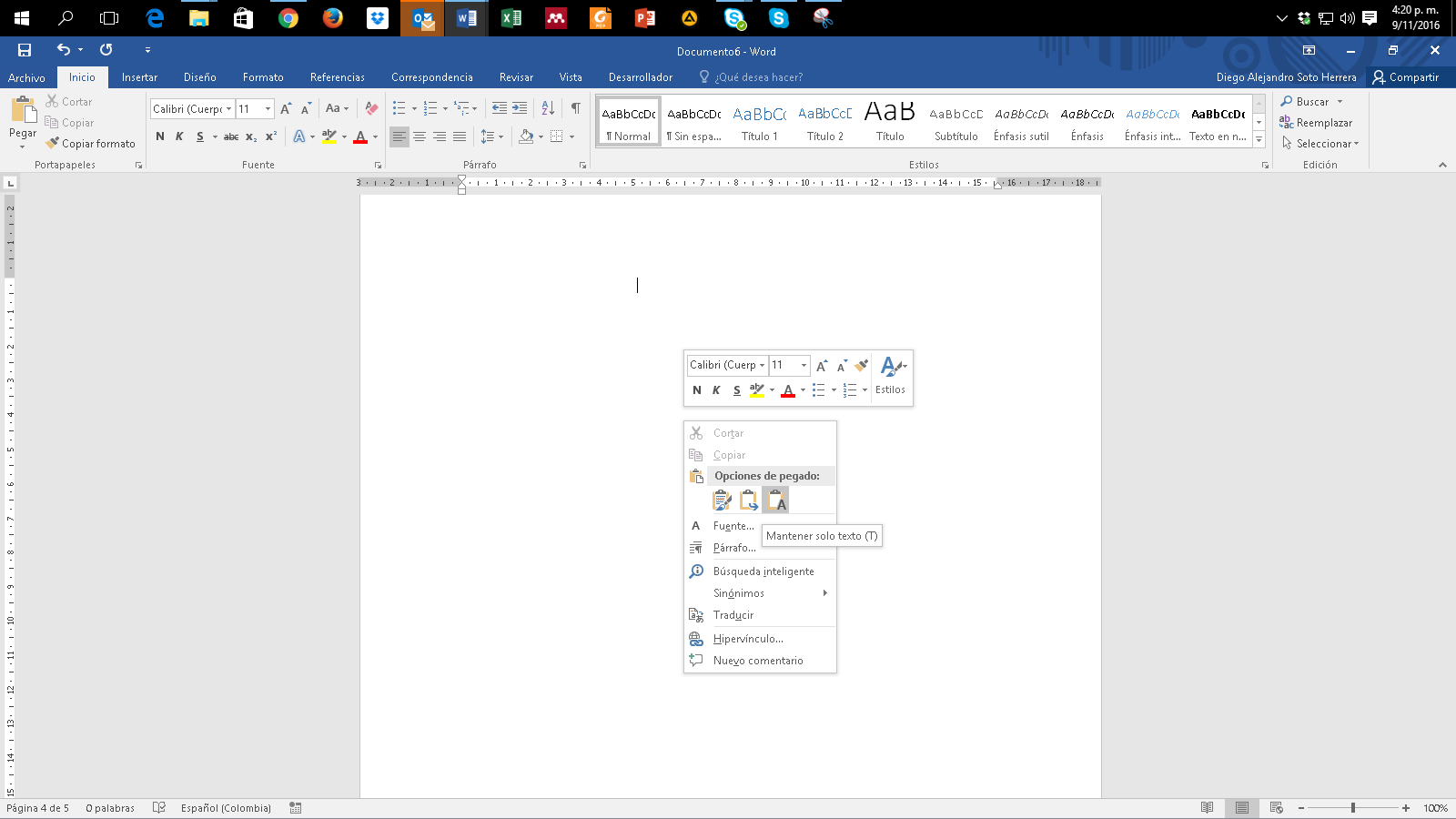
## Anexo 7. Sinónimos y antónimos

Constantemente surgen inconvenientes al redactar una oración, al no tener la palabra adecuada, un sinónimo o un antónimo. Microsoft Word apoya estas inquietudes, así:



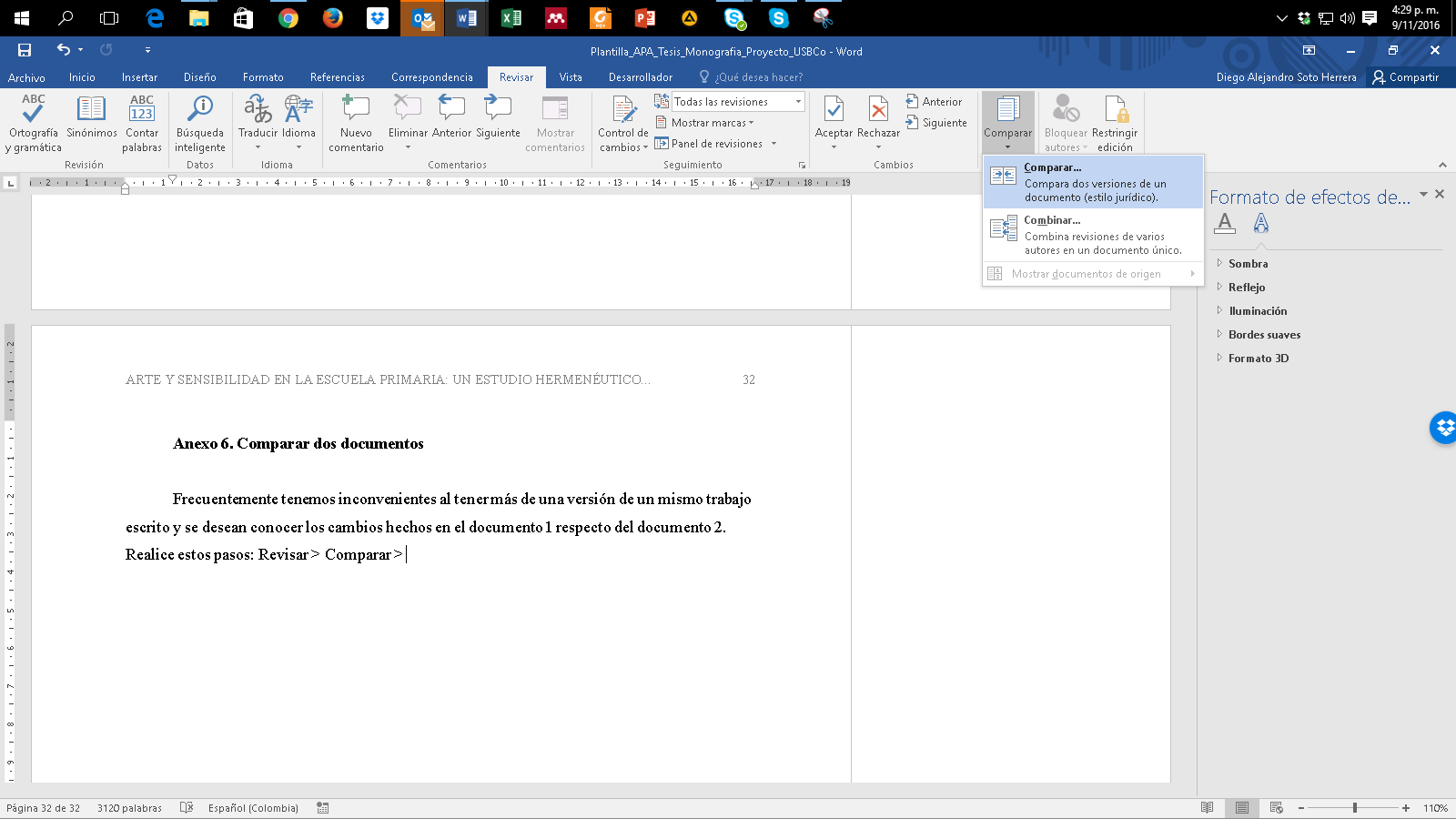
## Anexo 8. Copiar y pegar sin formato

En ocasiones copiamos y pegamos objetos o texto desde páginas web u otras fuentes hacia Word con el conocido Ctrl + C y Ctrl + V; sin embargo, se conservan colores, tipos de letras, tablas, y otros formatos indeseados. Para pegar solo el texto y sin formato alguno, clic derecho > “Mantener solo texto (T)”

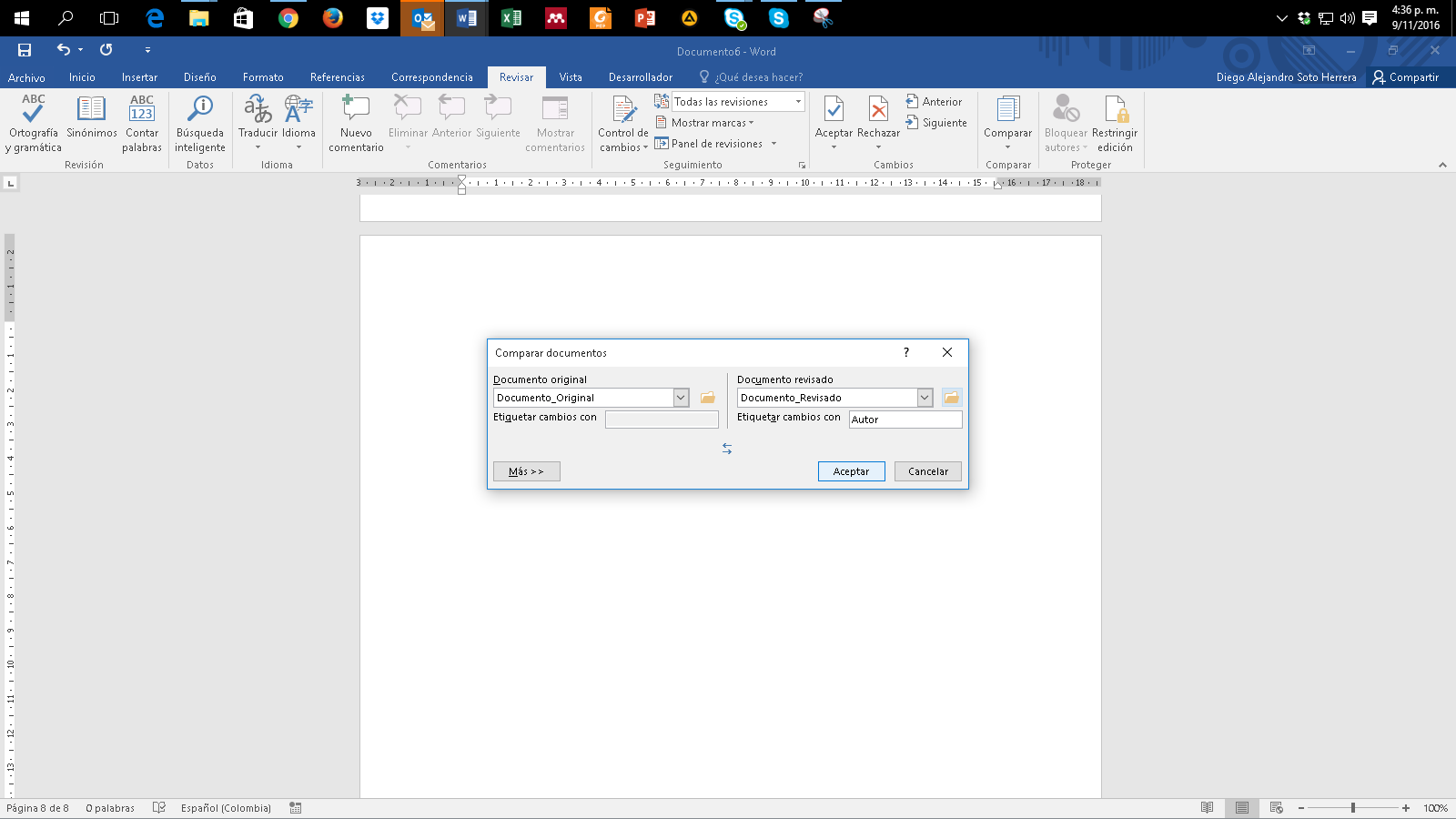


## Anexo 9. Comparar dos documentos

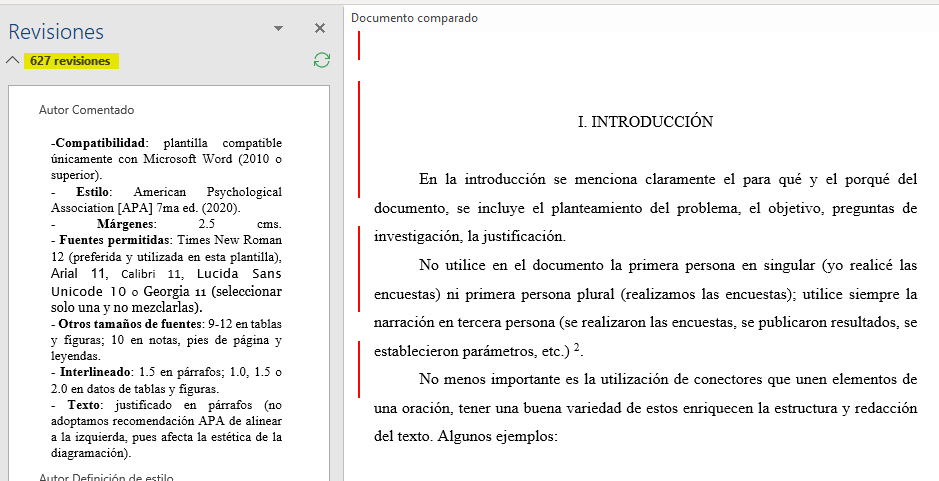
Frecuentemente tenemos inconvenientes al tener más de una versión de un mismo trabajo escrito y se desean conocer los cambios hechos en el documento 1 respecto del documento 2. Realiza estos pasos: Revisar > Comparar > Comparar... Compara dos versiones de un documento (estilo jurídico).



Busca la ruta en tu dispositivo donde se encuentra el documento original (izquierda) y luego el mismo procedimiento con el documento revisado (derecha). > clic en Aceptar.



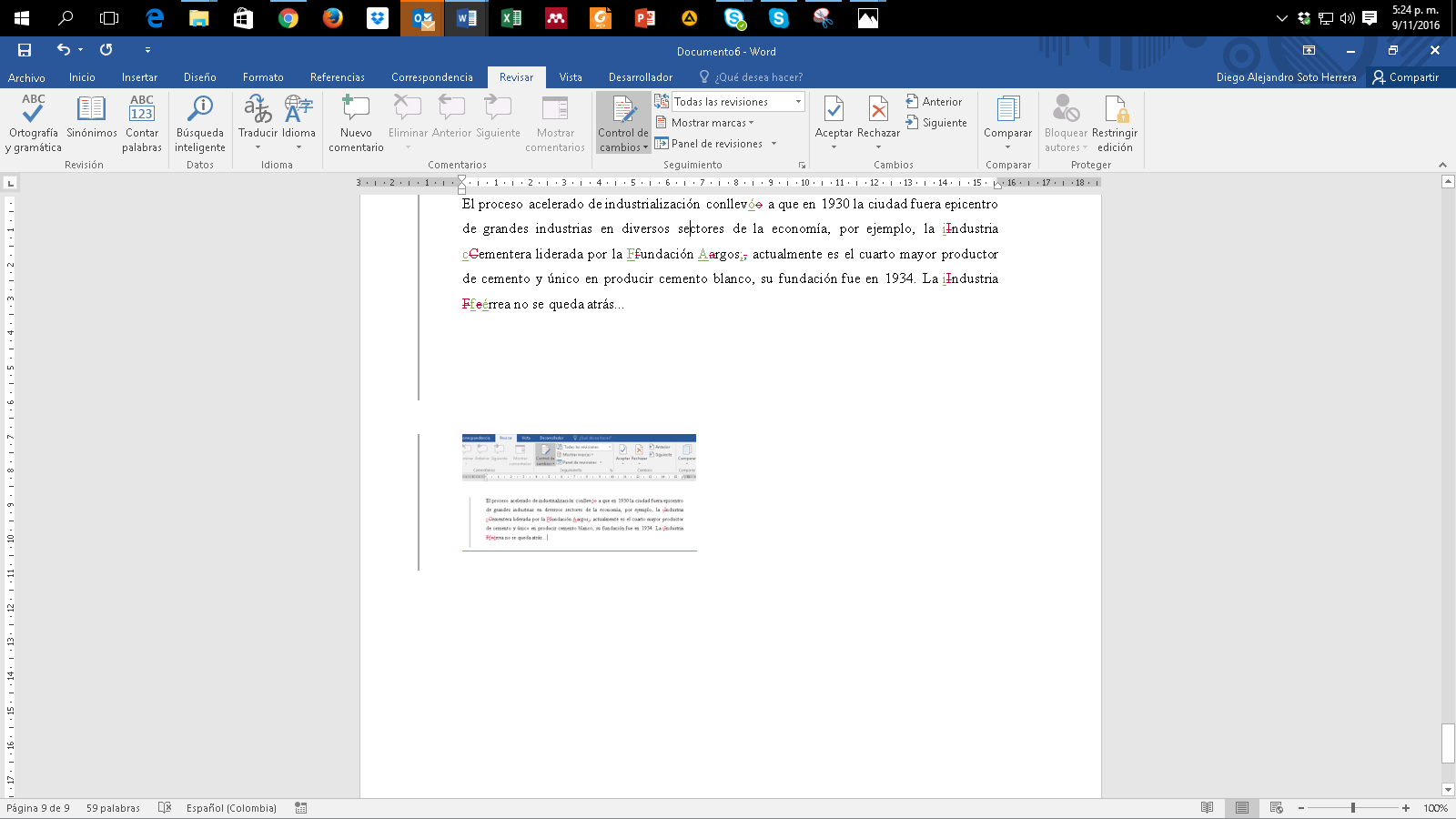
Posteriormente aparece el informe con la cantidad de revisiones hechas en el documento:



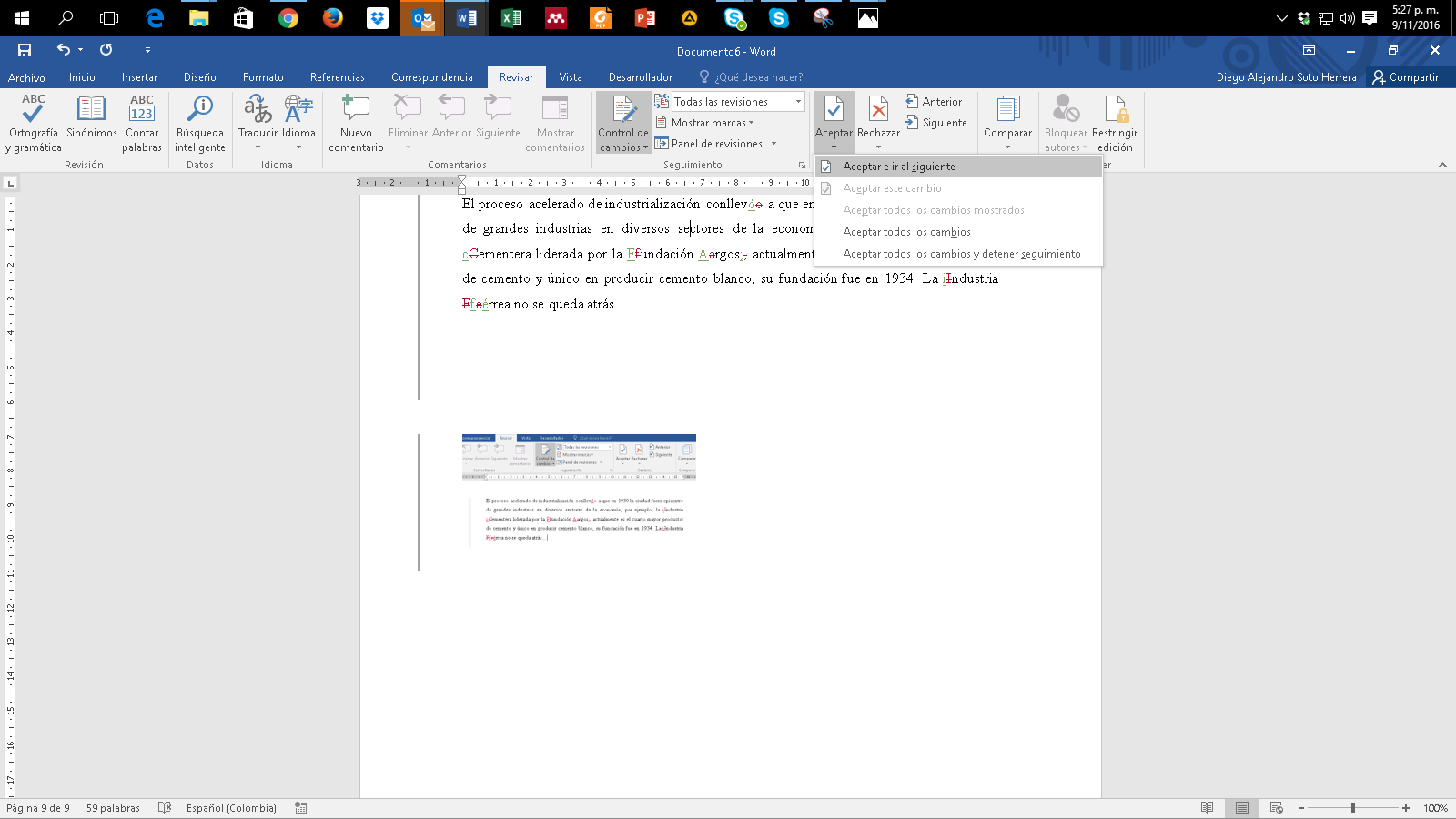
## Anexo 10. Control de cambios

Es una de las funciones más útiles, especialmente cuando se desea vigilar, revisar y aceptar cualquier cambio en un documento. Supongamos la interacción entre un estudiante que elabora la tesis y su asesor. El asesor considera que hay que hacer cambios, pero no desea modificar sin que el estudiante se entere y que, por consiguiente, acepte o rechace los cambios y aprenda de las sugerencias. Activa esta opción, así: Revisar > Control de Cambios.

**Modo asesor:** el asesor corrige los errores; es visible lo que se pretende eliminar con tachado guion medio (color rojo) y la sugerencia con guion bajo (color verde):

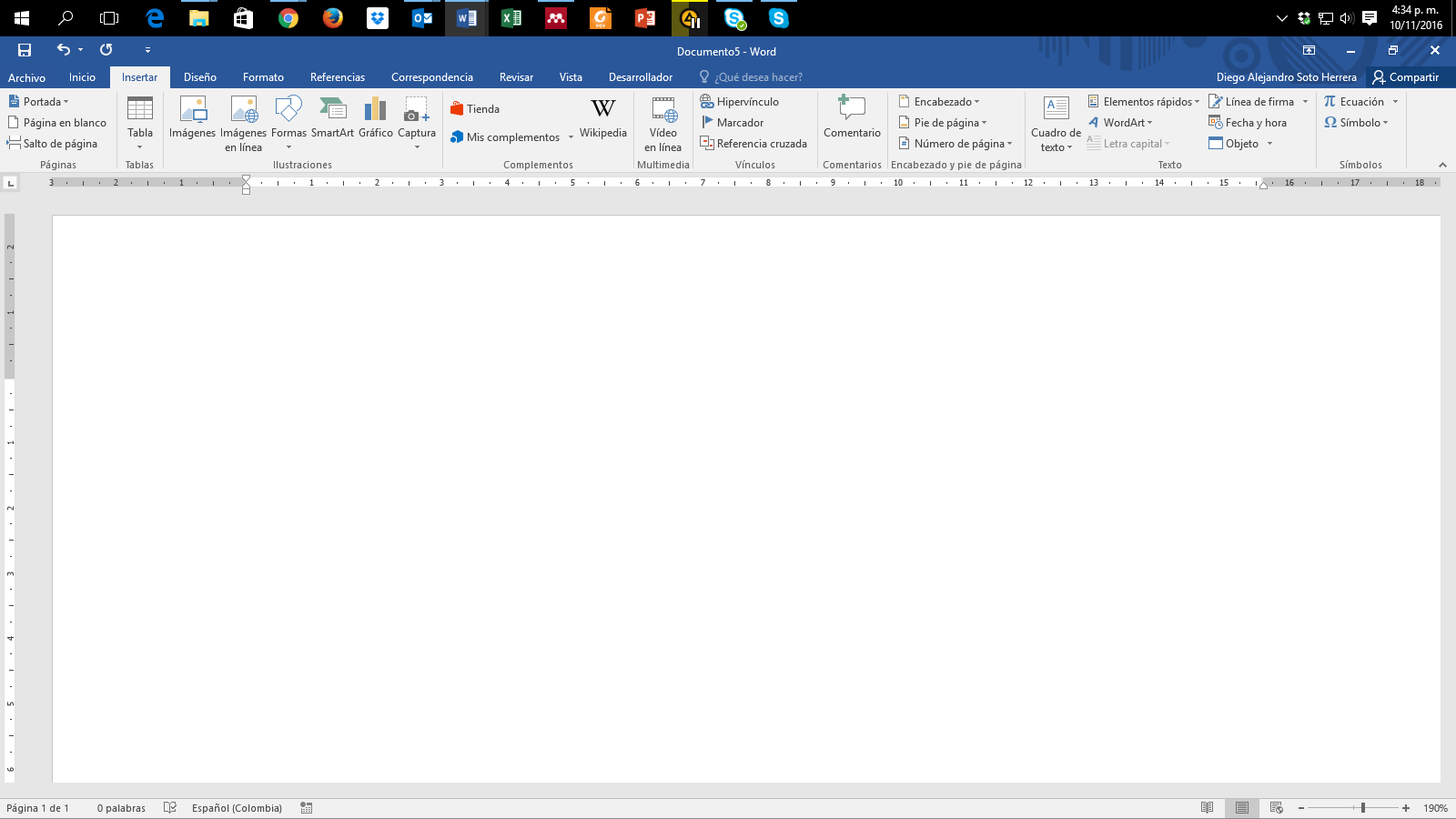


**Modo estudiante:** estudiante recibe archivo con sugerencias (el botón “Control de cambios” debe estar activo), clic en “Siguiente” y tiene la opción de “Aceptar” o “Rechazar” una a una las sugerencias visibles del asesor.



## Anexo 11. Insertar salto de página

Existe una sencilla función llamada “Salto de página” que ahorra tiempo en la estructura del texto, cuando se requiere iniciar en una nueva página en blanco, sin necesidad de insertar “Enter” una y otra vez en cada línea: Insertar > Salto de página. Su método abreviado con el teclado es: Ctrl + Enter.



## Anexo 12. Recortar y abreviar direcciones web largas

Eventualmente utilizamos páginas web, imágenes, documentos en línea, entre otros, y es necesario citarlas o mencionarlas en el texto; sin embargo, esos enlaces son supremamente largos, lo que le resta estética a la presentación del documento, ejemplo:

**Largo**: https://www.youtube.com/watch?reload=9&v=tRH59E1aybE&feature=youtu.be

**Corto**: https://bit.ly/3abhsgE

Utiliza una herramienta en línea para hacer de este enlace mucho más corto. Existe gran variedad de ellos, recomendamos algunos.

https://cutt.ly/ https://bitly.com/ https://tiny.cc/ https://tinyurl.com/

Ejemplo realizado con Bitly https://bitly.com/

Copiar y pega la URL larga en la casilla Shorten your link > Clic en Shorten > Posteriormente aparece la nueva URL corta > Clic en Copy > Pégala en el lugar del texto que la necesites.