



# Universidad Autónoma de Querétaro

## Facultad de Ingeniería

### Maestría en Ingeniería de Calidad y Productividad

#### **Impacto de la colaboración internacional sobre la producción científica mexicana en las ciencias sociales, artes y humanidades**

#### TESIS

Que como parte de los requisitos para obtener el grado de  
Maestra en Ingeniería de Calidad y Productividad

Presenta:

**I.I. Adriana Ivette Zamarripa Ancona**

Dirigido por:

**Dr. Eric Leonardo Huerta Manzanilla**

#### SINODALES

Dr. Eric Leonardo Huerta Manzanilla

Presidente

Firma

M.I. Javier García Pérez

Secretario

Firma

Dr. Edwin Geovanny Vergara Ayala

Vocal

Firma

Dr. José Alejandro Rivera Gonzaga

Sinodal

Firma

Dr. Carlos Enrique Cuevas Suárez

Sinodal

Firma

Querétaro, QRO, México.

Enero 2024



Dirección General de Bibliotecas y Servicios Digitales  
de Información



Impacto de la colaboración internacional sobre la  
producción científica mexicana en las ciencias  
sociales, artes y humanidades

**por**

Adriana Ivette Zamarripa Ancona

se distribuye bajo una [Licencia Creative Commons  
Atribución-NoComercial-SinDerivadas 4.0 Internacional](#).

**Clave RI:** IGMAC-311354



© 2024 - I.I. Adriana Ivette Zamarripa Ancona

Todos los derechos reservados.



*Este trabajo va dedicado a mi familia: mi papá, mi mamá y mi hermano, por ser mi constante soporte e inspiración de todos mis locos proyectos; A Raúl, por siempre estar, por quererme tanto y recordarme continuamente el valor y capacidad que tengo; A Lua y Zuki, por hacer crecer nuestra pequeña familia y demostrar que nuestro amor no tiene límites. Les amo significativamente.*



# Agradecimientos

Primeramente, agradezco a Dios por permitir las circunstancias necesarias para que yo realizara este proyecto de vida. Estoy consciente de que soy una persona privilegiada y pretendo hacer uso de esa ventaja para promover el desarrollo de mi entorno.

Agradezco al Consejo Nacional de Humanidades, Ciencia y Tecnología (CONAHCyT) y la Universidad Autónoma de Querétaro (UAQ) por el soporte durante este periodo. De la misma manera, agradezco a la Universidad Autónoma de Hidalgo (UAEH), específicamente al Instituto de Ciencias de la Salud (ICSa), por permitir la colaboración para la obtención de los recursos necesarios para este proyecto.

Agradezco a mi comité evaluador, Dr. Eric Huerta, M.I. Javier García, Dr. Edwin Vergara, Dr. José Alejandro Rivera Gonzaga y Dr. Carlos Enrique Cuevas Suárez, por su tiempo durante las revisiones. Sus aportaciones han sido base para la mejora de esta investigación.

Por otra parte, agradezco a todos los profesores y profesoras que me fueron asignados: Dra. Aurora Femat, M.C. Guillermo Hiyane, M.I. Javier García, M.A.P. J. Agustín Martínez, M.I. René Z. Flores, Dra. Magdalena Mendoza y Dr. Antonio Sierra. La docencia es una de las profesiones más nobles que existen. Gracias a por su tiempo y orientación durante este periodo de crecimiento.

Agradezco mucho a mi padre, Dr. Juan Eliezer Zamarripa Calderón, por ser mi principal guía y soporte. Agradezco a mi madre, Dra. Adriana Leticia Ancona Meza, por ser la primera en demostrarme que no se necesita la aprobación de nadie para cumplir tus metas. A ambos, por criarme en la disciplina y perseverancia, en la diversión y en la pasión; El equilibrio de lo imperfecto hace a este universo perfecto.

Finalmente, me agradezco a mí: Por ser constante y no desistir a pesar de la gran cantidad de problemas presentados. Gracias a mi cuerpo y mente por aguantar las malas condiciones a las que lo sometí. Ha sido un periodo de gran desgaste pero con una recompensa de la misma magnitud.



# Índice general

Agradecimientos . . . . .	I
Contenidos . . . . .	III
Índice de figuras. . . . .	VII
Índice de cuadros . . . . .	IX
Siglas y abreviaturas . . . . .	XI
Resumen. . . . .	XIII
Abstract . . . . .	XV
I. Introducción . . . . .	1
I.1 Planteamiento del problema . . . . .	2
I.2 Justificación . . . . .	3
I.3 Estructura de la tesis . . . . .	4
II. Revisión de la Literatura. . . . .	7
II.1 El desarrollo científico y su evaluación . . . . .	7
II.2 Indicadores de evaluación del desarrollo científico . . . . .	8
II.3 Ciencia y tecnología en México . . . . .	10
II.4 Estrategias para mejorar el desarrollo científico en México . . . . .	11
II.5 Desafíos en las ciencias sociales, artes y humanidades . . . . .	13
II.6 Informetría como método de evaluación del desarrollo científico . . . . .	14
II.6.1 Diseño del estudio . . . . .	15
II.6.2 Recopilación de información . . . . .	15
II.6.3 Limpieza y preparación de la base de datos . . . . .	16
II.6.4 Análisis estadístico de datos . . . . .	19

II.6.5	Visualización de resultados . . . . .	23
III.	Hipótesis y Objetivos. . . . .	29
III.1	Hipótesis . . . . .	29
III.1.1	Preguntas de investigación . . . . .	30
III.2	Objetivos . . . . .	30
III.2.1	Objetivo general . . . . .	30
III.2.2	Objetivos específicos . . . . .	30
IV.	Metodología . . . . .	33
IV.1	Recopilación de información . . . . .	34
IV.1.1	Selección de la base de datos . . . . .	34
IV.1.2	Filtrado de la información . . . . .	35
IV.1.3	Exportación de la base de datos . . . . .	35
IV.1.4	Importación de información a RStudio . . . . .	35
IV.2	Limpieza y preparación de la base de datos . . . . .	36
IV.2.1	Extraer variables de interés . . . . .	36
IV.2.2	Creación de nuevas variables . . . . .	37
IV.2.3	Modificación de estructuras . . . . .	37
IV.2.4	Limpieza de atributos de estudio . . . . .	38
IV.3	Análisis estadístico . . . . .	39
IV.3.1	Análisis comparativo . . . . .	40
IV.3.2	Análisis exploratorio . . . . .	41
IV.3.3	Análisis de correlación . . . . .	44
IV.4	Visualización de la información . . . . .	45
V.	Resultados . . . . .	47
V.1	Recopilación de la información . . . . .	47
V.2	Limpieza y preparación de la base de datos . . . . .	47
V.3	Análisis estadístico y visualización de la información . . . . .	48
V.3.1	Análisis comparativo . . . . .	48
V.3.2	Análisis exploratorio . . . . .	54
V.3.3	Análisis de correlación . . . . .	66
V.4	Productos obtenidos durante el posgrado . . . . .	69
VI.	Conclusión. . . . .	71
VI.1	Recomendaciones e investigaciones futuras . . . . .	74

VI.2 Limitaciones . . . . . 74  
A. Apéndice. . . . . 77



# Índice de figuras

1	Países Estratégicos para México definidos por CONAHCyT. (CONAHCYT, 2021) . . . . .	12
2	Flujo del método informétrico. (Aria & Cuccurullo, 2017) . . . . .	15
3	Ejemplo de gráfico de burbujas (Scimago, 2022). . . . .	24
4	Ejemplo de diagrama de Pareto. (Elaboración propia) . . . . .	25
5	Ejemplo de gráfico polar. (Scimago, 2022) . . . . .	26
6	Ejemplo de gráfico treemap. (Scimago, 2022) . . . . .	26
7	Ejemplo de nube de palabras. (Scimago, 2022) . . . . .	27
8	Ejemplo de mapa de coropletas. (Scimago, 2022) . . . . .	27
9	Ejemplo de diagrama de cajas. (Elaboración propia) . . . . .	28
10	Diagrama de cajas de la metodología aplicada. . . . .	33
11	Diagrama de cajas de las actividades contenidas en la recopilación de información. . . . .	34
12	Diagrama de cajas de las actividades contenidas en la limpieza y preparación de la base de datos. . . . .	36
13	Diagrama de cajas de las actividades contenidas en el análisis estadístico. . . . .	40
14	Editores de los documentos en la población y en México. . . . .	49
15	Producción de las fuentes presentes en la población y en México. . . . .	50
16	Palabras recurrentes en la población y en México. . . . .	51
17	Rango de autores participantes por producto en la población y en México. . . . .	52
18	Productividad y visibilidad de autores en la población y en México. . . . .	53
19	Impacto científico evaluado en visibilidad en la población y en México. . . . .	55
20	Impacto científico evaluado mediante la productividad de México. . . . .	55
21	Índice de colaboración internacional de México. . . . .	57

22	Países que colaboraron con México durante el periodo de estudio. . . . .	59
23	Índice de multidisciplinariedad de México en las CSA&H. . . . .	60
24	Áreas de investigación más productivas dentro de las CSA&H de México. . .	61
25	Índice de liderazgo de México en las CSA&H. . . . .	61
26	Presupuesto asignado de México y algunos de sus países estratégicos. . . . .	63

# Índice de cuadros

1	Tipos de productos científicos (Rousseau et al., 2018b). . . . .	9
2	Coefficiente de Spearman (R) (Rousseau et al., 2018c) . . . . .	22
3	Variables disponibles en la base de datos de este proyecto. <b>Nota:</b> part. = Participantes; No. = Número . . . . .	38
3	Variables disponibles en la base de datos de este proyecto. <b>Nota:</b> part. = Participantes; No. = Número . . . . .	39
4	Rangos establecidos para el estudio de las variables. . . . .	45
5	Modelo lineal para determinar el crecimiento de la calidad de la producción científica mexicana en CSA&H. . . . .	56
6	Modelo lineal para determinar el crecimiento de la productividad de la producción científica mexicana en CSA&H. . . . .	57
7	Modelo lineal para determinar el crecimiento de la colaboración internacional en México dentro de las CSA&H. . . . .	58
8	Modelo lineal para determinar el crecimiento de la multidisciplinariedad en México dentro de las CSA&H. . . . .	60
9	Modelo lineal para determinar el crecimiento del liderazgo de México dentro de las CSA&H. . . . .	62
10	Modelos lineales para determinar el crecimiento del presupuesto asignado. . .	64
11	Matriz de correlación de Spearman entre las variables de este estudio. . . . .	67
13	Variables disponibles en Colección Principal de WoS (Web of Science Group, 2019). . . . .	77



# Siglas y abreviaturas

<b>Siglas</b>	<b>Significado</b>
CSA&H	Ciencias sociales, artes y humanidades
OECD	Organización para la Cooperación y Desarrollo Económicos
WoS	Web of Science
UAQ	Universidad Autónoma de Querétaro
UAEH	Universidad Autónoma del Estado de Hidalgo
PIB	Producto Interno Bruto
CONAHCyT	Consejo Nacional de Humanidades, Ciencia y Tecnología
SNI	Sistema Nacional de Investigadores



# Resumen

El desarrollo científico de un país implica el progreso constante en la generación de conocimiento a través de la investigación y la aplicación de teorías. En este sentido, se ven involucrados la academia, la industria y el gobierno en conjunto, promoviendo así la competitividad económica y mejorando la calidad de vida de la población.

Evaluar el progreso científico es esencial para comprender la posición de un país en estos campos. A nivel nacional, ayuda a identificar fortalezas, debilidades y guía la asignación eficiente de recursos. A nivel internacional, la comparación de indicadores científicos fomenta la colaboración. Dos indicadores clave son la productividad (cantidad de documentos generados) y la calidad de los productos científicos (medida por citas en documentos), evaluados a través de la informetría.

Este estudio se centra en el desarrollo científico en Ciencias Sociales, Artes y Humanidades (CSA&H) en México. Destaca el papel del inglés en la colaboración científica global y la diversificación de la investigación en México. La Universidad Nacional Autónoma de México (UNAM) lidera en afiliaciones de investigadores. México muestra una fuerte cultura de colaboración, pero se resalta la necesidad de promover la calidad de la investigación.

La colaboración internacional en CSA&H en México es alta y diversa, consolidando su liderazgo en esta área. La inversión en ciencia y tecnología en México está en crecimiento, lo que refleja su compromiso con el progreso científico. Se observan correlaciones interesantes entre liderazgo, colaboración internacional y multidisciplinariedad.

Este estudio proporciona información valiosa para orientar estrategias de investigación y desarrollo en México. Subraya la importancia de equilibrar cantidad y calidad en la producción científica, así como fomentar la colaboración a nivel nacional e internacional. El compromiso de México con la inversión en ciencia y tecnología es significativo, lo que refuerza su posición en un mundo impulsado por la innovación y la tecnología.

**Palabras clave:** productividad, calidad, colaboración internacional, multidisciplinariedad, liderazgo.



# Abstract

The scientific development of a country involves a continuous progression in generating knowledge through research and applying theories. It extends beyond academia, involving industry and government, thus promoting economic competitiveness, and enhancing the quality of life for the population. Achieving this requires investment, infrastructure, innovation promotion, and international collaboration. Evaluating scientific progress is crucial for understanding a country's position in research and technology. Nationally, it aids in identifying strengths and weaknesses, and guides efficient resource allocation. Internationally, comparing scientific indicators fosters collaboration. Two key indicators are productivity (the quantity of generated documents) and the quality of scientific products (measured by citations in documents), assessed through informetrics. This study focuses on scientific development in the CSA&H (Social Sciences, Arts, and Humanities) field. It highlights the role of English in global scientific collaboration and the diversification of research in Mexico.

The National Autonomous University of Mexico (UNAM) leads in researcher affiliations. Mexico exhibits a strong culture of collaboration, emphasizing promoting research quality. International cooperation in CSA&H in Mexico is highly diverse, while Mexico is consolidating its leadership in this field. Mexico's investment in science and technology is growing, reflecting its commitment to scientific progress. Interesting correlations are observed between leadership, international collaboration, and multidisciplinary. In summary, this study provides valuable information to guide research and development strategies in Mexico. It underscores the importance of balancing quantity and quality in scientific production and fostering collaboration at national and international levels. Mexico's commitment to investing in science and technology is evident, reinforcing its position in a world driven by innovation and technology.

**Key words:** productivity, quality, international collaboration, multidisciplinary, leadership.



# Introducción

La ciencia se entiende como el conjunto de actividades que, de manera sistematizada, tienen como objetivo crear y organizar conocimiento para el beneficio y desarrollo humano (Chinchilla-Rodríguez et al., 2019). El proceso científico produce salidas que regularmente se documentan en productos científicos, los cuales son clave para una evaluación integral de la ciencia. El desarrollo científico de una unidad es importante para la coordinación y administración del financiamiento disponible, así como la mejora continua de la unidad de estudio en términos de economía, seguridad, política y salud (Lancho-Barrantes & Cantú-Ortiz, 2019).

Específicamente para México, se ha identificado que las áreas de investigación con menor rendimiento en términos de calidad y productividad pertenecen a las ciencias sociales, artes y humanidades (CSA&H) (Lancho-Barrantes & Cantú-Ortiz, 2019). Siguiendo los principios de la *Teoría de Restricciones*, originalmente publicada por E. Goldratt en 1993, se plantea que para mejorar de manera eficiente los resultados de un proceso, es fundamental identificar y enfocar los esfuerzos en las áreas que presentan el rendimiento más bajo (Techt, 2016). Esta metodología se basa en la idea de que una cadena es tan fuerte como su eslabón más débil, y por lo tanto, para elevar la calidad y productividad de la investigación científica en México, resulta imperativo priorizar las áreas de investigación pertenecientes a las CSA&H, que actualmente se encuentran rezagadas en términos de indicadores de desempeño. Al hacerlo, se busca no solo elevar la media de los indicadores, sino también reducir las restricciones que limitan el potencial de desarrollo en estas áreas, lo que a su vez contribuirá a fortalecer la ciencia y la cultura académica en el país.

El presente proyecto ofrece un estudio integral de la calidad y productividad del desarrollo científico de México en las áreas de investigación de CSA&H, así como el estudio de la

colaboración internacional, la multidisciplinariedad, el liderazgo y el presupuesto designado por México durante el periodo 2010-2020. Considerando que no existen suficientes trabajos recientes que analicen el desarrollo científico de México para la mejora del rendimiento de las áreas de investigación con resultados más bajos, este proyecto pretende cubrir la brecha por medio de un proyecto informétrico exploratorio. De esta manera, se proveerá de conocimiento relevante y significativo para que autoridades nacionales fortalezcan relaciones con países y afiliaciones con las que se ha tenido los mejores resultados en calidad y productividad científica. Así mismo, se busca alcanzar directamente a investigadores con intereses de estudio enfocadas a las áreas incluidas en este proyecto, buscando ofrecer características de los productos científicos que permitirán la mejora de la visibilidad de sus productos y por ende, el rendimiento del desarrollo de la calidad científica en México.

## **I.1. Planteamiento del problema**

México es la segunda economía más grande de los países de Latino América (OECD, 2015). Sin embargo, la producción científica y la calidad de investigación en México es uno de los más bajos en comparación con los países miembros de la Organización para la Cooperación y Desarrollo Económicos (OECD). Esto se puede observar en el gasto dirigido al desarrollo de ciencia y tecnología, el cual fue el 0.30% de los ingresos totales de México durante el 2020 comparado con el más alto que fue del 5.44% de Israel (BancoMundial, 2023). La OECD emite un reporte de manera periódica, por el cual ha advertido a México de su posición en el desarrollo de ciencia y tecnología a nivel mundial. El reporte sugiere la mejora a través de la diversificación en la cultura científica, en la multidisciplinariedad, en la colaboración, así como en la inversión para obtener el crecimiento deseado (OECD, 2015).

Considerando el análisis previo, México enfrenta una limitación evidente en cuanto a la asignación de recursos para el desarrollo de ciencia y tecnología. Esta restricción afecta de manera desproporcionada a áreas de investigación con menor productividad relativa, como las ciencias sociales, las artes y las humanidades. A pesar de ocupar el segundo lugar en la economía latinoamericana, la brecha entre este logro y la inversión insuficiente en investigación revela la necesidad apremiante de reevaluar estratégicamente las estrategias para impulsar el crecimiento científico y tecnológico del país. Este desafío requiere un enfoque renovado en las políticas y estrategias, asegurando que las áreas de investigación menos visibles, pero de gran importancia, no se vean perjudicadas por la escasez de recursos, permitiendo así un avance equitativo y significativo en el panorama científico y tecnológico mexicano.

## I.2. Justificación

Para mejorar la calidad de los productos científicos nacionales, es importante cubrir las necesidades de los eslabones débiles de la unidad de estudio. En este caso, las prioridades de México deberían de ser las áreas de investigación con el rendimiento más bajo en calidad y productividad. Considerando los registros de la base de datos Clarivate - Web of Science, se han identificado las áreas de investigación con menor rendimiento dentro de los índices Social Sciences Citation Index® y Arts & Humanities Citation Index® (Lancho-Barrantes & Cantú-Ortiz, 2019).

Las ciencias sociales, humanísticas y artísticas, a menudo subestimadas en el contexto del desarrollo tecnológico, desempeñan un papel esencial y multifacético. Estas disciplinas contribuyen significativamente a la comprensión social, aportan perspectivas éticas, estimulan la creatividad e impulsan la inclusividad. Su papel va más allá de la mera producción de conocimiento; contextualizan la innovación, anticipan desafíos éticos, fomentan la creatividad y facilitan la participación ciudadana. Integrar activamente estas disciplinas en el proceso de desarrollo tecnológico no solo enriquece la calidad de la investigación, sino que también genera soluciones más contextualizadas y socialmente responsables, promoviendo un avance sostenible y equitativo en la sociedad.

Actualmente, se tienen establecidas estrategias para impulsar el desarrollo científico de México, esencialmente si se trata de un área de investigación con alto rendimiento. Sin embargo, hasta el momento, no se han definido estrategias con el objetivo de impulsar las áreas de investigación científica con un rendimiento bajo en productividad y calidad (Techt, 2016). Esta brecha presenta una oportunidad para adoptar enfoques específicos que atiendan las necesidades de estas áreas rezagadas.

La OECD ha propuesto estrategias para potenciar el desempeño científico de México, abogando por un incremento en la inversión y la diversificación de la cultura científica mediante colaboraciones. Estas medidas no solo se traducirán en la creación de ambientes innovadores, sino que también fomentarán la multidisciplinariedad, impulsando así la capacidad científica del país (OECD, 2015).

Distintos estudios respaldan la noción de que la colaboración internacional constituye una estrategia vital para el desarrollo, la mejora de la calidad y la diversificación en el ámbito científico, abarcando diversos niveles, desde instituciones educativas superiores hasta naciones o regiones (Castillo & Powell, 2020; Zhe et al., 2021). Al satisfacer la necesidad de intercambio de información, recursos y experiencias, esta estrategia no solo expande sino

que también complementa la capacidad científica de los participantes. De este modo, la colaboración internacional se erige como un medio eficaz para cumplir con la recomendación de diversificación compartida por la OECD con México (OECD, 2015).

### **I.3. Estructura de la tesis**

La presente tesis se compone de siete capítulos en donde se desarrolla a detalle el estudio del desarrollo científico mexicano en las áreas de CSA&H en México durante el periodo 2010-2020.

En el presente capítulo, que corresponde al Capítulo I, se aborda el planteamiento del problema de investigación, se expone la justificación de la presente tesis y se realiza una breve sinopsis de los contenidos que serán desarrollados en los capítulos subsiguientes del documento.

El Capítulo II aborda los conceptos necesarios para comprender la evaluación del desarrollo científico de una unidad de estudio. De la misma forma, se estudian las condiciones económicas y culturales que han permitido el desarrollo de la ciencia y tecnología como se conoce hoy en día. Por otro lado, se revisan propuestas realizadas por diversos autores e instituciones para mejorar el rendimiento de México en términos de desarrollo científico como la colaboración internacional, la multidisciplinariedad, el liderazgo y el presupuesto asignado a las actividades encaminadas a cumplir con este objetivo. Finalmente, se hace una revisión de la metodología propuesta por varios autores para el estudio de las tendencias en la ciencia: La informetría; Así como uno de los lenguajes más utilizados para este tipo de estudios: R y su interfaz RStudio (Aria & Cuccurullo, 2017; Efrain-García & García-Zorita, 2016; Rousseau et al., 2018a).

En el Capítulo III establece la hipótesis del proyecto, la cual estima que la colaboración internacional, la multidisciplinariedad y el liderazgo que México adquiere en un proyecto de investigación influye en la calidad y productividad de los productos científicos. Para confirmar esta declaración, se pretende estudiar el comportamiento de estas variables sobre las áreas de investigación CSA&H, identificar los países y las afiliaciones con las que México ha colaborado más durante el periodo 2010-2020, e identificar la correlación entre variables de estudio para determinar su influencia sobre el rendimiento científico mexicano.

En el Capítulo IV se describe el método aplicado en este estudio: La informetría. Esta metodología consiste en cinco pasos generales: “Recopilación de información”, “Limpieza y preparación de la base de datos”, “Análisis estadístico”, “Visualización de la información” y

“Resultados y discusión”. En la etapa de “Recopilación de información” se describe la base de datos seleccionada y el flujo de acciones necesarias para obtener la base de datos original para el presente estudio. En la siguiente etapa, “Limpieza y preparación de la base de datos”, se describe la lógica y los comandos utilizados para asegurar que los datos sean consistentes y libres de errores para obtener resultados fidedignos. El “Análisis estadístico” consiste en una exploración exhaustiva de los datos mediante una constante reestructuración de la información y creación de diversos indicadores, buscando identificar las tendencias de cada variable a lo largo del tiempo. La etapa de “Visualización de la información” expone la lógica utilizada para plasmar los resultados obtenidos en diversos gráficos, buscando que la visualización e interpretación de los resultados sea más amigable para el lector. Finalmente, en la etapa “Resultados y discusión” se analiza, desarrolla e interpretan los resultados obtenidos de acuerdo con el contexto y literatura revisada en el Capítulo II. Es importante mencionar que esta última etapa, “Resultados y discusión”, se revisa en un capítulo independiente, el cual es el Capítulo V.

Por último, el Capítulo VI compara la hipótesis definida en el Capítulo III con los resultados obtenidos, concluyendo si esta se cumple o no. Así mismo, da respuesta a las preguntas de investigación definidas, se resaltan hallazgos relevantes durante la investigación, recomendaciones y limitaciones de este proyecto.



# Revisión de la Literatura

La evaluación del desarrollo científico es esencial para comprender el posicionamiento de un país en la investigación y tecnología a nivel nacional e internacional. En esta sección, se analiza el desarrollo científico en México, considerando la influencia de decisiones estratégicas en la inversión en investigación científica y su impacto en la diáspora de investigadores mexicanos. Se detallan los indicadores utilizados para evaluar el desarrollo científico y se exploran los efectos de las crisis económicas en la inversión en ciencia y tecnología. Además, se abordan estrategias para mejorar el desarrollo científico en México, incluyendo la creación del Sistema Nacional de Investigadores (SNI) y la importancia de la colaboración internacional. Se destaca la relevancia de las ciencias sociales, artes y humanidades en el panorama científico y se proponen recomendaciones para fortalecer su posición. Así mismo, esta sección también profundiza en aspectos clave relacionados con la productividad en el desarrollo científico, como la producción científica, calidad, colaboración, presupuesto asignado, liderazgo y multidisciplinariedad, y ofrece una revisión exhaustiva de estos temas. La informetría es la metodología explorada en detalle para proporcionar una comprensión sólida, incluyendo herramientas y metodologías como la Ley de Pareto, medidas de tendencia central y correlación de variables. En resumidas cuentas, esta sección presenta un panorama completo del desarrollo científico en México y las herramientas para su evaluación y mejora.

## II.1. El desarrollo científico y su evaluación

El desarrollo científico de un país se define como el continuo progreso en la generación de conocimiento, materializado a través de la investigación y la aplicación de teorías (Leite & Pinho, 2017). Este concepto va más allá de los confines académicos, ya que impulsa

la competitividad económica mediante la colaboración entre la academia, la industria y el gobierno. Además, tiene un impacto directo en la mejora de la calidad de vida de la población y en el progreso de la sociedad, gracias a la implementación de soluciones científicas en áreas de salud, energía, medio ambiente, entre otros (Ebadi & Schiffauerova, 2013). El desarrollo científico está vinculado a factores esenciales como la inversión, la infraestructura y recursos disponibles, la promoción de la innovación y la colaboración internacional (Chinchilla-Rodríguez et al., 2019; CONAHCYT, 2021; Sierra-González & Ramos-Pérez, 2021).

Evaluar el avance científico es fundamental para comprender cómo un país se posiciona en términos de investigación y tecnología en el mundo. A nivel nacional, esta evaluación permite identificar fortalezas y áreas de mejora, orientar la toma de decisiones en políticas de investigación y desarrollo, y asignar recursos de manera eficiente (Rousseau et al., 2018b; Zhe et al., 2021). A nivel internacional, la comparación de indicadores científicos y tecnológicos entre países fomenta la colaboración y el intercambio de conocimientos, además de identificar oportunidades para la cooperación en investigación y desarrollo (Francisco, 2015; Sierra-González & Ramos-Pérez, 2021). Esta evaluación también es crucial para identificar tendencias emergentes y desafíos compartidos que requieren soluciones científicas y tecnológicas conjuntas (Melin & Persson, 1996; Sandoval-Romero & Larivière, 2020).

## **II.2. Indicadores de evaluación del desarrollo científico**

La ciencia es el conjunto de actividades que crean y organizan conocimiento en beneficio y desarrollo humano, y es el punto central para estudiar el progreso científico (Chinchilla-Rodríguez et al., 2019; Rousseau et al., 2018b). Este proceso sistemático cuenta con dos resultados: Una salida intelectual y una salida física conocida como “producto científico”. Este último es esencial para comprender adecuadamente un proyecto, ya que contiene información de gran relevancia que engloba antecedentes, métodos y resultados del trabajo (Lancho-Barrantes & Cantú-Ortiz, 2019). Algunos ejemplos de estos productos científicos se detallan en el Cuadro 1.

Para medir el progreso científico, se destacan dos indicadores clave: La productividad y la calidad de los productos científicos (Prathap, 2013). La productividad se refiere a la cantidad de documentos generados por un individuo, reflejando su habilidad para impulsar el avance científico a través de la creación de conocimiento (Ebadi & Schiffauerova, 2013; Lancho-Barrantes & Cantú-Ortiz, 2019). En contraste, la calidad de los productos científicos se evalúa

Cuadro 1: Tipos de productos científicos (Rousseau et al., 2018b).

<b>Tipos de productos científicos</b>	
<b>Publicaciones</b>	Documento publicado, no necesariamente con revisión de pares. Puede ser un libro, una revista, o una impresión electrónica.
<b>Libros editados</b>	Conjunto de capítulos escritos por diferentes autores, conjuntados y armonizados por uno o más editores.
<b>Libro de texto</b>	Una monografía escrita con propósitos educativos.
<b>Artículo científico</b>	Artículo que contiene consideraciones teóricas y una revisión de la literatura. Está escrito para una audiencia escolar.
<b>Literatura gris</b>	Documentos no controlados por publicistas comerciales, como los reportes técnicos.

de varias maneras, siendo la cantidad de citas por producto científico el indicador más usado (Aldieri et al., 2018; Katz & Martin, 1997). El análisis de las citas en documentos mide el nivel de confianza por parte de la comunidad científica, en los resultados y métodos del proyecto citado (Sooryamoorthy, 2019; Uddin et al., 2015). Al incluir un producto científico en las referencias de un proyecto propio, el investigador demuestra reconocimiento ante el contenido del documento, lo que amplía la influencia y el alcance sobre otros trabajos (Ebadi & Schiffauerova, 2016; Guo et al., 2016).

A pesar de la inversión limitada en desarrollo científico (menos del 1% del Producto Interno Bruto - PIB), México ha demostrado que tiene capacidad para aumentar su productividad científica (Castaños-Lomnitz, 2006). En el periodo de 1989 a 2014, México contribuyó al 0.45% de la productividad científica global, posicionándose en el lugar 35 entre los productores científicos a nivel mundial (Uddin et al., 2015). En 2020, México ocupó el puesto 17 con una contribución de 0.69% a la producción científica global (CONAHCYT, 2021). En cuanto a la calidad de las disciplinas de investigación, la bioquímica, biología molecular, neurociencias, medicina general e interna, inmunología y microbiología representan la mayoría de las citas generadas en México (Barahona et al., 2014). Las áreas de ciencias sociales (4.65 citas por producto) y artes y humanidades (4.38 citas por producto) son las de menor calidad investigativa en el país (Lancho-Barrantes & Cantú-Ortiz, 2019).

## II.3. Ciencia y tecnología en México

México, reconocido como la segunda economía más grande de América Latina, ha enfrentado a lo largo de su historia diversas crisis financieras, incluyendo las ocurridas en 1982, 1994-95, 2008-09 (Benneworth & Jongbloed, 2010; OECD, 2015) y, más recientemente, durante el período 2020-23 debido a la pandemia global. Ante la necesidad de respuestas inmediatas en muchas de estas situaciones, el gobierno federal recurrió al endeudamiento con el Fondo Monetario Internacional y el Banco Mundial, desencadenando incluso una inflación de hasta el 140% (Sandoval-Romero & Larivière, 2020). Este aumento inflacionario sumergió al país en una fragilidad financiera que planteó la posibilidad de inestabilidad e incluso colapso económico (Sierra-González & Ramos-Pérez, 2021).

En este contexto desafiante, México se vio forzado a reajustar las prioridades de inversión en investigación científica, enfocándose en áreas que abordan necesidades fundamentales del país. Lancho-Barrantes y Cantú-Ortiz (2019) identificaron que las disciplinas de bioquímica, biología molecular, neurociencias, medicina general e interna, inmunología y microbiología recibieron la mayor inversión durante el período 2007-2016. Sin embargo, como resultado de esta priorización, muchos miembros de la comunidad científica mexicana, en campos no destacados, optaron por emigrar a naciones con recursos más propicios para el desarrollo de sus carreras científicas. Este fenómeno dio origen a una red de conexiones dispersas, conocida como diáspora, en la cual investigadores ubicados en diferentes partes del mundo colaboran y confían entre sí para impulsar su crecimiento profesional (Sandoval-Romero & Larivière, 2020). En un reporte del Consejo Nacional de Humanidades, Ciencia y Tecnología (CONAHCyT) del 2020, se identificó que estos investigadores tuvieron una residencia principalmente en Estados Unidos, España e Inglaterra, mientras que sus países de origen abarcan principalmente España y Colombia (CONAHCYT, 2021).

A pesar de las dificultades económicas que el país ha afrontado, México destaca por su valiosa contribución científica desde instituciones públicas y universidades. No obstante, el sector empresarial aporta apenas un 47% de financiamiento, siendo inferior al promedio (50%) de los países miembros de la Organización para la Cooperación y el Desarrollo Económicos (OECD) (OECD, 2015). Dado que las instituciones respaldadas por financiamiento gubernamental constituyen la base del conocimiento científico en México, es esencial abordar la necesidad de implementar estrategias para mitigar el impacto de las crisis económicas. Estas crisis han limitado la inversión en ciencia y tecnología a un modesto promedio del 0.50% del PIB en la ciencia y tecnología (Lancho-Barrantes & Cantú-Ortiz, 2019).

## **II.4. Estrategias para mejorar el desarrollo científico en México**

En 1984, el gobierno federal y el CONAHCyT crearon la iniciativa del Sistema Nacional de Investigadores (SNI) que tiene como propósito elevar la producción, calidad, trascendencia e impacto científico tanto a nivel nacional como internacional de México mediante el reconocimiento y respaldo de los investigadores nacionales (CONAHCYT, 2021; Sandoval-Romero & Larivière, 2020). Desde su creación, el SNI ha fomentado un período de crecimiento científico, en gran parte atribuible a la colaboración internacional facilitando el acceso a recursos que de otro modo no estarían disponibles (Uddin et al., 2015).

La colaboración internacional implica la cooperación entre entidades geográficamente dispersas que comparten el objetivo fundamental de generar conocimiento. Esta estrategia recibe reconocimiento y promoción gracias a la mejora en las comunicaciones y al compromiso de elevar tanto la calidad como el alcance de la investigación en un contexto global (Kyvik & Reymert, 2017; Sooryamoorthy, 2019; Zhe et al., 2021). Un informe de la OECD señala que la colaboración en México se acerca al promedio objetivo, lo cual sugiere que la diáspora de investigadores mexicanos y la promoción del SNI en proyectos colaborativos han contribuido a establecer una sólida red global de colaboración (OECD, 2015). Para enfocar y optimizar las acciones colaborativas de México, CONAHCyT ha clasificado ciertas naciones como prioritarias en su estrategia de cooperación (Fig. 1). Entre los países estratégicos para México se incluyen Alemania, Argentina, Brasil, Canadá, Chile, China, Colombia, Corea del Sur, España, Francia, India, Israel, Japón, Rusia, Sudáfrica, Turquía, Estados Unidos y el Reino Unido (CONAHCYT, 2021).

La estrategia de colaboración internacional para potenciar la producción científica nacional puede identificarse a través de su impacto en diversos factores clave en el ámbito científico de un país (Zhe et al., 2021). Entre las formas en que esta colaboración puede incidir positivamente en la investigación científica se incluyen el acceso a recursos complementarios, la diversidad de perspectivas y enfoques, la expansión de redes profesionales, el incremento de la productividad y calidad científica, el establecimiento de colaboraciones a largo plazo, la obtención de financiamiento internacional y la contribución a problemas globales (Ebadi & Schiffauerova, 2013; Katz & Martin, 1997; Lancho-Barrantes & Cantú-Ortiz, 2019; Sooryamoorthy, 2019). Estos beneficios suelen evaluarse a través de factores generales relacionados con las condiciones de investigación en el país de estudio, como la coautoría, el liderazgo, la multidisciplinariedad y el presupuesto destinado a la investigación (Dumitru, 2019).

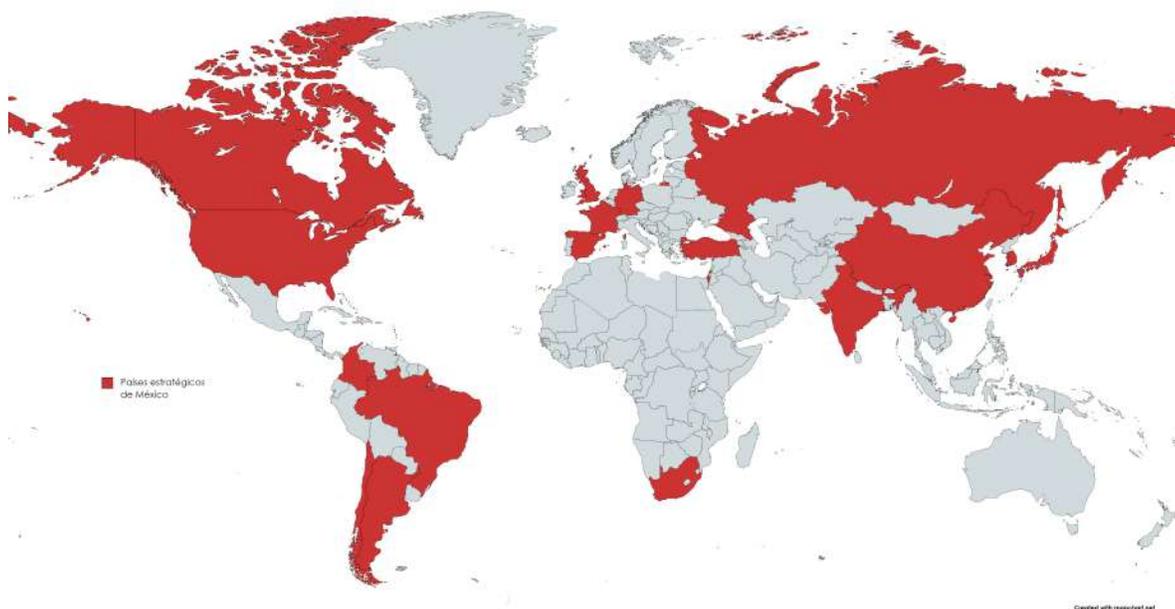


Figura 1: Países Estratégicos para México definidos por CONAHCyT. (CONAHCYT, 2021)

Los países líderes en las investigaciones científicas usualmente son aquellos con una mayor capacidad científica, inversión y recursos disponibles (Chinchilla-Rodríguez et al., 2019). Diversos estudios confirman que la colaboración, especialmente con entidades científicas más avanzadas, robustece la capacidad investigadora, mejorando la calidad y la producción de los países con menores capacidades científicas (Aldieri et al., 2018; Katz & Martin, 1997). Esto explica por qué los países latinoamericanos tienen mucha presencia en la colaboración internacional (Monge-Nájera & Ho, 2018). Castillo y Powell (2020) resaltan que Brasil y México, como líderes en producción científica en América Latina, exhiben tasas de colaboración internacional del 27% y 40.5%, respectivamente, superando el promedio mundial. Esto indica que ambos países tienen una marcada presencia en la colaboración científica a nivel internacional, sin necesariamente implicar que la colaboración sea exclusiva entre ellos

La clave de la colaboración reside en crear vínculos en los que todos los involucrados se benefician, combinando fortalezas y distribuyendo costos para lograr resultados superiores (Villaseñor et al., 2017; Zhe et al., 2021). La conformación de equipos diversos en habilidades, valores, culturas y especialidades propicia la innovación y la resolución de problemas complejos (Ebadi & Schiffauerova, 2013; Katz & Martin, 1997). Esta exposición a la diversidad conduce a productos de investigación más atractivos, con metodologías estandarizadas, soluciones innovadoras y un alcance global, resultando en alta calidad y demanda (Sooryamoorthy, 2019).

En cuanto a inversión en ciencia y tecnología, esta desempeña un papel crucial en la productividad científica de un país ya que define la capacidad de adquisición de recursos para proyectos de investigación científica (Chinchilla-Rodríguez et al., 2019; Lancho-Barrantes & Cantú-Ortiz, 2019). En México, el gasto asignado a la ciencia y tecnología en 2016 fue de 0.39 % y en 2020 de 0.28 % del PIB, respectivamente, siendo ambos inferiores al promedio de la OECD y América Latina (0.58 %) (BancoMundial, 2023; OECD, 2015). En un informe del 2020 (CONAHCYT, 2021), el CONAHCyT evidenció que el porcentaje del PIB destinado al desarrollo científico y tecnológico en ciencias sociales, artes y humanidades no superó el 0.05 % del total del 0.28 % invertido ese mismo año. Este porcentaje refleja una asignación considerablemente menor en comparación con las áreas de ciencias de la salud, naturales e ingeniería, que recibieron la mayor inversión durante dicho periodo.

Con el propósito de elevar tanto la productividad como la calidad de la investigación científica en México, la OCDE recomienda estrategias que abarcan la diversificación de proyectos, el aumento de la inversión, el fomento de la cultura científica y la atracción de talento internacional (Ebadi & Schiffauerova, 2013; OECD, 2015). Estas estrategias pueden ser fortalecidas a través de la colaboración internacional. La multidisciplinariedad de los equipos colaborativos no solo promueve la innovación, sino que también concede acceso a recursos que originalmente no estarían disponibles, además de atraer a expertos de renombre internacional. Esto, a su vez, culmina en una mejora sustancial en la producción, calidad e innovación de la investigación científica en México (Ebadi & Schiffauerova, 2016; Kyvik & Reymert, 2017; Lancho-Barrantes & Cantú-Ortiz, 2019).

## **II.5. Desafíos en las ciencias sociales, artes y humanidades**

Las Ciencias Sociales, Artes y Humanidades (CSA&H) desempeñan un papel fundamental en el desarrollo integral de una nación, proporcionando ambientes propicios para el pensamiento crítico y la interpretación de significados (Dumitru, 2019). Sin embargo, en México, estas áreas han sido relegadas en términos de inversión y recursos, con disciplinas como la inmunología y microbiología, la bioquímica, genética y biología molecular, y las neurociencias acaparando la prioridad y los fondos (Sandoval-Romero & Larivière, 2020). Este enfoque desequilibrado no solo limita el potencial de las CSA&H, sino que también afecta negativamente la diversidad y amplitud de la investigación en el país.

Dentro de las CSA&H se encuentran disciplinas como sociología, ciencias económicas, ciencias jurídicas y derecho en el ámbito de las ciencias sociales. Las artes y humanidades

presentan características distintivas, como un enfoque en temas locales y una inclinación por el trabajo individual (Lámbarry-Vilchis & Moreno-Jiménez, 2020; Taşkın & Doğan, 2020). A pesar de su importancia, estas áreas enfrentan barreras significativas para su mejora, careciendo de estrategias específicas de apoyo y enfrentándose a la falta de visibilidad y recursos (Ball, 2017; Ebadi & Schiffauerova, 2016).

La "Teoría de Restricciones" de E. Goldratt resalta la necesidad de enfocarse en las etapas con bajo rendimiento para mejorar los resultados del proceso (Techt, 2016). A pesar de esta perspectiva, los programas de inversión actuales en México continúan concentrándose en disciplinas más productivas, desatendiendo las áreas de menor rendimiento. Esta falta de estrategias específicas para apoyar a las CSA&H ha llevado a un escaso estudio de su desarrollo científico en comparación con campos como la ciencia, tecnología y medicina (Ball, 2017; Ebadi & Schiffauerova, 2016).

La diversidad y singularidad de enfoques en CSA&H han limitado su visibilidad y, por ende, su acceso a recursos y oportunidades. Adoptar una visión que respalde a las áreas de investigación más débiles podría generar mejoras sustanciales en la calidad de la investigación nacional en CSA&H, contribuyendo así a un desarrollo más equitativo y enriquecedor para la sociedad mexicana. Este enfoque no solo enriquecería la base de conocimientos del país, sino que también fomentaría la innovación, la creatividad y la inclusividad en la investigación y el desarrollo en México.

Reconocer y priorizar las CSA&H en la agenda de desarrollo científico y tecnológico de México es esencial para construir una base de conocimientos sólida y equitativa. La diversidad de disciplinas dentro de CSA&H aporta perspectivas únicas y fundamentales para comprender la complejidad de la sociedad y enfrentar los desafíos contemporáneos. El apoyo adecuado a estas áreas no solo promovería la excelencia científica, sino que también enriquecería la vida cultural y social del país, contribuyendo a un desarrollo integral y sostenible.

## **II.6. Informetría como método de evaluación del desarrollo científico**

Diversas metodologías han sido identificadas para la evaluación del desarrollo de la ciencia como el análisis estadístico, la revisión de literatura, la bibliometría y la cienciometría, siendo las últimas dos las más recurrentes.

Los términos "bibliometría" y "cienciometría" a menudo generan confusión, dado que ambos son enfoques estadísticos para cuantificar características de productos documentados.

La distinción radica en las particularidades que analizan; la cienciometría incluye atributos políticos y administrativos, a diferencia de la bibliometría (Guo et al., 2016; Lancho-Barrantes & Cantú-Ortiz, 2019). Ante esta confusión, los expertos sugieren usar el término “informetría” para englobar ambas metodologías. La informetría se enfoca en analizar y respaldar el proceso de registro y transferencia de información (Aria & Cuccurullo, 2017; Rousseau et al., 2018a). Este enfoque brinda una visión integral de la producción científica de una entidad, mapear comportamientos, rastrear la internacionalización de la ciencia, inferir tendencias y respaldar la formulación de políticas (Sandoval-Romero & Larivière, 2020).

La informetría consta de cinco etapas generales y con frecuencia se requieren diversas herramientas de software, las cuales no siempre son gratuitas. El flujo propuesto por Aria y Cuccurullo (2017) muestra las etapas de este proceso (Fig. 2). A continuación, se describirán las etapas del flujo de la informetría.



Figura 2: Flujo del método informétrico. (Aria & Cuccurullo, 2017)

### **II.6.1. Diseño del estudio**

El diseño del estudio es la primera etapa de la informetría. Así como en cualquier otro proyecto, es importante desarrollar un plan de las actividades que se deben realizar para cumplir con el objetivo del proyecto. Dentro de esta etapa se encuentra la definición de las preguntas de investigación y la selección de los métodos apropiados para dar respuesta a las preguntas definidas (Aria & Cuccurullo, 2017).

### **II.6.2. Recopilación de información**

En esta etapa se realizan tres actividades para la recolección de la información a analizar: La selección de la base de datos, el filtrado de la información relevante para el estudio y finalmente, la exportación de la base de datos en un formato amigable con alguna herramienta de tipo software para el análisis estadístico y de visualización (Aria & Cuccurullo, 2017). El principal objetivo de esta etapa es la preparación de un conjunto de observaciones que cumplan las características necesarias para su manipulación y análisis buscando cumplir con los objetivos definidos para el proyecto (Efrain-García & García-Zorita, 2016).

Existen varias bases de datos bibliográficas disponibles que permiten obtener metadatos sobre trabajos científicos ya estructurados. Algunos ejemplos de estas bases de datos son Clarivate -WoS, Scopus, Google Scholar y Science Direct (Aria & Cuccurullo, 2017). Una vez seleccionada la fuente de información, es importante filtrar los registros que se contemplarán para el estudio. Afortunadamente, la mayoría de las bases de datos cuentan con herramientas de filtrado como las listas de verificación y los queries. Un query es una consulta realizada a una base de datos por medio de un código en un lenguaje específico (Efrain-García & García-Zorita, 2016). Finalmente, la exportación de datos es la descarga de la información seleccionada para su tratamiento y análisis (Fernandez-Casal et al., 2021).

### **II.6.3. Limpieza y preparación de la base de datos**

La limpieza de la base de datos se enfoca en mejorar la estructura de la información a analizar, con el objetivo de aumentar la confiabilidad de los resultados obtenidos. Esta tarea involucra el procesamiento de los datos para identificar y corregir elementos duplicados o incorrectamente escritos (Aria & Cuccurullo, 2017; Efrain-García & García-Zorita, 2016). Por su parte, la preparación de los datos implica recopilar las variables necesarias para abordar las preguntas de investigación planteadas (Efrain-García & García-Zorita, 2016). Estas variables, que se convertirán en indicadores informétricos, desempeñarán un papel fundamental en la evaluación del rendimiento del sistema científico de la entidad analizada (Lancho-Barrantes & Cantú-Ortiz, 2019). Aunque es posible utilizar las variables directamente desde la base de datos, es más común generar variables derivadas de las originales, lo que permite un análisis más profundo y orientado a los objetivos del proyecto. En el Anexo 13 se presentan las variables disponibles en la base de datos de Clarivate - WoS.

Es común que, al analizar el impacto de la colaboración internacional, se consideren diversas variables. Entre ellas se encuentran la productividad o número de productos científicos emitidos por la unidad de estudio (Aldieri et al., 2018; Bornmann, 2017; Coccia & Wang, 2015; Ebadi & Schiffauerova, 2013); la calidad de un producto científico o bien, el número de citas por producto (Chinchilla-Rodríguez et al., 2019; Lancho-Barrantes & Cantú-Ortiz, 2019; Leydesdorff et al., 2019; Monge-Nájera & Ho, 2018; Villaseñor et al., 2017); la colaboración entre autores o países, o número de participantes en el proyecto (Chung et al., 2009; Ebadi & Schiffauerova, 2013; Sooryamoorthy, 2019; Villaseñor et al., 2017; Zhe et al., 2021); el presupuesto asignado al desarrollo de la ciencia y tecnología (Ebadi & Schiffauerova, 2016); la posición de liderazgo que tiene la unidad de estudio (Chinchilla-Rodríguez et al., 2019; Zhu & Liu, 2020); así como la cantidad de disciplinas involucradas del produc-

to, también entendido como multidisciplinariedad (Lancho-Barrantes & Cantú-Ortiz, 2019; Zhe et al., 2021; Zhu & Liu, 2020). Estos indicadores se asocian estrechamente con los beneficios derivados de la implementación de la estrategia de colaboración internacional, y a continuación, se ofrecerá una descripción concisa de cada uno de ellos para comprender su contribución al éxito de esta estrategia en el ámbito científico nacional.

### **II.6.3.1. Productividad: Producción científica**

La productividad en el ámbito del desarrollo científico se refiere a la medición y evaluación de la producción de las actividades de investigación y desarrollo (Castillo & Powell, 2020). Este indicador suele medirse mediante el conteo de publicaciones, patentes, y otros documentos que evidencien la generación de resultados de investigación innovadores e impactantes que contribuyan al avance del conocimiento y la economía (Aldieri et al., 2018; Bornmann & Williams, 2020). La productividad es una demanda de los formuladores de políticas para asegurar la utilización efectiva de los recursos asignados al desarrollo de ciencia y tecnología en un país o instituto (Chinchilla-Rodríguez et al., 2018). Sin embargo, el énfasis actual en maximizar el número de publicaciones y citas por producto científico podría comprometer la integridad de la literatura científica (Chung et al., 2009).

### **II.6.3.2. Calidad: Número de citas por producto científico**

Las citas en un documento científico representan un reconocimiento de su importancia, impacto y utilidad por parte de otros trabajos académicos. Expertos sostienen que una publicación con un alto número de citas está relacionada con contenido de calidad (Chinchilla-Rodríguez et al., 2019; Zhe et al., 2021). Por esta razón, el análisis de citas por producto es constantemente considerado para evaluar la calidad de un producto científico (Melin & Persson, 1996). Sin embargo, es crucial normalizar esta variable para una comparación justa. Uno de los principales desafíos es la variación en las citas a lo largo del tiempo: Un artículo publicado en 2015 tendrá más oportunidades de recibir citas que uno publicado en 2019 debido al tiempo que ha estado disponible para la comunidad científica (Bornmann et al., 2019; Lancho-Barrantes & Cantú-Ortiz, 2019; Leydesdorff et al., 2019; Sooryamoorthy, 2019). Para abordar este desafío, se han desarrollado diversas técnicas, como la ponderación percentil o el uso de indicadores compuestos. Sin embargo, también es posible recurrir a enfoques más simples, como la división del número de citas entre la cantidad de años transcurridos desde su publicación (Bornmann & Williams, 2020; Rousseau et al., 2018d). La elección de la técnica dependerá de las particularidades de cada estudio.

### **II.6.3.3. Colaboración**

La colaboración en la investigación se identifica a menudo por la participación de múltiples individuos en un proyecto, siendo la coautoría un indicador para evaluar esta variable (Katz & Martin, 1997; Rousseau et al., 2018b). En un estudio informétrico, la coautoría puede proporcionar resultados en tres niveles distintos: geográfico, institucional y de redes. A nivel geográfico, se exploran las colaboraciones entre países, ciudades o regiones. A nivel institucional, se analiza la coautoría entre universidades, industrias y entidades gubernamentales. En el ámbito de las redes, se examina la coautoría entre investigadores individuales (Lancho-Barrantes & Cantú-Ortiz, 2019). Un documento científico se considera colaborativo a nivel institucional si los autores tienen afiliaciones en múltiples instituciones, y se considera internacional si los autores pertenecen a diferentes países (Melin & Persson, 1996).

### **II.6.3.4. Presupuesto asignado**

El presupuesto para investigación científica proviene de diversas fuentes como agencias gubernamentales, fundaciones privadas e industria. Una asignación efectiva de fondos es esencial para fomentar investigaciones innovadoras y originales, manteniéndose a la vanguardia en campos científicos (Ebadi & Schiffauerova, 2013). Se suele asociar un mayor presupuesto con un impacto positivo en el avance científico: Países con mayor inversión tienden a ser más especializados y visibles (Lancho-Barrantes & Cantú-Ortiz, 2019), mientras que naciones con recursos limitados dependen de colaboraciones externas para competir globalmente (Ebadi & Schiffauerova, 2016). Asimismo, a mayor inversión también aumenta la probabilidad de liderazgo y la capacidad de generar innovación (Chinchilla-Rodríguez et al., 2019).

En el caso de México, la inversión en investigación ha permanecido estancada, promediando el 0.50% del PIB (Lancho-Barrantes & Cantú-Ortiz, 2019). En 2020, el financiamiento se distribuyó de la siguiente manera: El gobierno federal aportó el 66.36%, las instituciones de educación superior el 2.18%, las instituciones privadas no lucrativas el 0.70%, el sector empresarial (incluyendo empresas nacionales y transnacionales) el 22.83%, las familias el 7.33%, y el sector externo el 0.59% del gasto total en ciencia, tecnología e innovación (CONAHCYT, 2021). Estos datos subrayan que México continúa siendo uno de los países donde el sector empresarial nacional y transnacional contribuye menos en comparación con otros, destacando la necesidad urgente de implementar estrategias que fomenten la inversión privada para fortalecer la investigación y mejorar la competitividad a nivel global.

### **II.6.3.5. Liderazgo**

El liderazgo en un proyecto de investigación se refiere a la posición que ocupa la unidad de estudio. La contribución de los autores al proyecto determina su lugar en la lista de autores del documento, siendo los primeros y últimos autores generalmente los más influyentes (Zhe et al., 2021). Las unidades con mayor inversión en investigación, capacidad científica y *expertise* tienden a ser líderes en productos científicos. Sin embargo, las publicaciones que no tienen liderazgo pueden tener un impacto y una especialización mayores (Chinchilla-Rodríguez et al., 2018). Esto sugiere que la colaboración internacional con socios líderes puede aprovechar las fortalezas de los países. Este patrón es especialmente notorio en países con producción y liderazgo científico por debajo del promedio global, donde las naciones con menos recursos a menudo logran un mayor impacto científico cuando no lideran proyectos (Chinchilla-Rodríguez et al., 2019).

### **II.6.3.6. Multidisciplinariedad**

La ciencia experimenta un notable avance gracias a la multidisciplinariedad, un enfoque que une diversas perspectivas y enfoques provenientes de distintas disciplinas. Esta integración permite una comprensión más profunda de fenómenos complejos y conduce a resultados más precisos (Ebadi & Schiffauerova, 2016). Este concepto implica la colaboración e integración de especialistas de distintos campos para colaborar, compartir saberes y aportar sus perspectivas y habilidades únicas. En última instancia, esto se refleja en una mejora sustancial de la calidad del contenido del proyecto y por ende, del producto científico (Fanelli et al., 2017). En el análisis informétrico, la multidisciplinariedad suele medirse a través de la cantidad de áreas de investigación o disciplinas involucradas en el proyecto de investigación (Aria & Cuccurullo, 2017; Lancho-Barrantes & Cantú-Ortiz, 2019).

## **II.6.4. Análisis estadístico de datos**

El análisis estadístico constituye una fase fundamental del proceso informétrico, ya que en esta serie de actividades se proporcionarán respuestas a las preguntas de investigación planteadas al inicio del proyecto informétrico (Efrain-García & García-Zorita, 2016). Es esencial destacar que no todos los proyectos informétricos incorporarán idénticos tipos de análisis estadístico. Esto dependerá de los objetivos establecidos, los métodos utilizados y las variables disponibles para el estudio (Aria & Cuccurullo, 2017). Los análisis comúnmente empleados pueden clasificarse en descriptivos e inferenciales, los cuales se describirán a continuación.

### II.6.4.1. Estadística descriptiva

El propósito del análisis estadístico descriptivo, también conocido como análisis estadístico exploratorio, radica en resumir y visualizar los datos de una muestra o población (Rousseau et al., 2018c). El término “población” se refiere al conjunto total de registros sobre los cuales se busca obtener información. Sin embargo, en ocasiones, la población puede tener una cantidad de registros demasiado grande para analizar en su totalidad. En estos casos, los investigadores optan por seleccionar registros que cumplan ciertas características, lo que se denomina una “muestra representativa”. Siempre que sea factible, en investigaciones informáticas, es preferible incluir la totalidad de la población (Rousseau et al., 2018e). A continuación, se presentan algunos conceptos de estadística descriptiva que respaldan esta fase del análisis.

#### Ley de Pareto

La Ley de Pareto, también conocida como el principio del 80/20, establece que aproximadamente el 80 % de los efectos provienen del 20 % de las causas. Este concepto se aplica para describir escenarios en los que una gran proporción de los efectos es atribuible a un reducido número de factores. En el ámbito científico, esta ley puede ilustrarse con ejemplos como un equipo de investigación, donde el 80 % de las publicaciones pueden originarse a partir del 20 % de los miembros más activos. Para visualizar este principio, se utiliza el gráfico de Pareto, un recurso que se analiza con mayor detalle en la sección “Visualización de resultados” II.6.5.

#### Medidas de tendencia central

Las medidas de tendencia central son valores que intentan resumir un conjunto completo de observaciones mediante un único número que representa la centralidad de la muestra o población. Estas medidas buscan encontrar un valor representativo en el centro del conjunto de datos, facilitando la formulación de una conclusión general (Rousseau et al., 2018c). A continuación, se presentan las tres principales medidas de tendencia central (Diez et al., 2019).

- **Media aritmética ( $\bar{X}$ ):** Es el promedio de todos los datos en la muestra, obtenido al sumar todos los valores  $(x_1, x_2, x_3, \dots, x_n)$  y dividir entre el tamaño total de la muestra ( $N$ ). La Fórmula II.1 muestra su expresión matemática.

$$\bar{X} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{N} \quad (\text{II.1})$$

Donde  $\bar{X}$  es la media aritmética,  $x_i$  es un valor individual en la muestra y  $N$  es el total de observaciones.

- Mediana ( $Me$ ): La mediana es el valor que se encuentra en el centro de los datos ordenados de menor a mayor. Se calcula sumando el producto de cada valor ( $x_i$ ) y su frecuencia ( $f_i$ ), y dividiendo entre el tamaño de la muestra ( $n$ ). La Fórmula II.2 muestra su expresión matemática.

$$Me = \frac{\sum_{i=1}^n x_i \cdot f_i}{n} \quad (\text{II.2})$$

Donde  $Me$  es la mediana,  $x_i$  es un valor individual,  $f_i$  es la frecuencia correspondiente a  $x_i$  y  $n$  es el tamaño de la muestra.

- Moda ( $M_0$ ): La moda es el valor con mayor frecuencia en el conjunto de datos. Su cálculo involucra la clase modal ( $L_{M_0}$ ), la diferencia entre las frecuencias adyacentes ( $d_1$  y  $d_2$ ) y el ancho de la clase ( $\omega$ ). La Fórmula II.3 muestra su expresión matemática.

$$M_0 = L_{M_0} + \left(\frac{d_1}{d_1 + d_2}\right)\omega \quad (\text{II.3})$$

Donde  $M_0$  es la moda,  $L_{M_0}$  es la clase modal,  $d_1$  es la diferencia entre la frecuencia modal y la frecuencia de la clase anterior,  $d_2$  es la diferencia entre la frecuencia modal y la frecuencia de la clase siguiente, y  $\omega$  es el ancho de la clase.

### Correlación de variables

La correlación se refiere a la fuerza con la que se establece una relación entre dos variables. En este contexto, el coeficiente de correlación de Spearman emerge como una medida estadística no paramétrica que evalúa la existencia de una relación de monotonía entre dos variables categorizadas (Devore, 2016). A diferencia de la correlación de Pearson, la correlación de Spearman no supone distribución normal ni relación lineal entre las variables. Este coeficiente adquiere valores en el rango de -1 a 1 (Diez et al., 2019). Un valor positivo de  $R$  indica una relación directa entre las variables, con una pendiente positiva. Un valor negativo de  $R$  señala una relación inversa y una pendiente negativa en el modelo. Cuando  $R = 0$ , se

sugiere la ausencia de relación entre las variables. En el Cuadro 2, se proporcionan detalles sobre la interpretación de los valores de correlación en contextos de análisis informétrico. Para calcular el coeficiente de Spearman entre dos variables que pertenecen a una población, se emplea la Fórmula II.4 (Diez et al., 2019).

$$\rho_{XY} = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \quad (\text{II.4})$$

En la Fórmula II.4,  $n$  representa el número de observaciones (pares de valores clasificados), y  $d_i$  corresponde a la diferencia entre los rangos de los valores en la variable X y la variable Y para el  $i$ -ésimo par. Este coeficiente proporciona una medida de la relación de monotonía entre las variables evaluadas en el análisis de correlación de Spearman.

Cuadro 2: Coeficiente de Spearman (R) (Rousseau et al., 2018c)

Coeficiente de Spearman (R)						
-1	-0.7	-0.5	0	0.5	0.7	1
Negativa ideal	Negativa fuerte	Negativa media	Nulo	Positiva media	Positiva fuerte	Positiva ideal

### II.6.4.2. Estadística inferencial

El objetivo de la inferencia estadística es identificar relaciones entre las variables seleccionadas. Se requieren tanto variables de entrada como de salida para realizar evaluaciones inferenciales. Las entradas corresponden a metadatos obtenidos de bases de datos, mientras que las salidas representan el rendimiento de las publicaciones, ya sea en términos de productividad o calidad del país evaluado (Moya-Anegón & Herrero-Solana, 2006).

Un enfoque común en el estudio del desarrollo científico es la aplicación de modelos lineales (Lancho-Barrantes & Cantú-Ortiz, 2019). La regresión lineal, una técnica estadística ampliamente utilizada, modela la relación entre una variable dependiente continua y una o más variables independientes. Su objetivo es encontrar la “línea de regresión”, que mejor se ajuste a los datos y permita realizar predicciones (Bates et al., 2015). La ecuación de la regresión lineal se expresa con la Fórmula II.5.

$$y = \beta_0 + \beta_1 x + \varepsilon \quad (\text{II.5})$$

Donde  $x$  es la variable independiente,  $y$  es la variable dependiente relacionada con  $x$ ,  $\varepsilon$

denota el error o desviación aleatoria en el modelo, y  $\beta_0$  y  $\beta_1$  son coeficientes de regresión estimados a través del método de mínimos cuadrados (Kutner et al., 2004).

Los resultados clave de un análisis de regresión lineal incluyen los siguientes aspectos (Montgomery et al., 2012):

1. **Coeficientes de regresión ( $\beta$ ):** Estos coeficientes ofrecen información sobre la dirección y magnitud de la relación entre las variables en estudio. Un coeficiente positivo señala una relación directa, mientras que uno negativo indica una relación inversa. En particular, el coeficiente  $\beta_1$  muestra la tasa de cambio en la variable dependiente en relación con la variable independiente, en caso de que la correlación sea lineal.
2. **Coeficiente de determinación ( $R^2$ ):** Este valor indica la calidad del ajuste del modelo, representando la proporción de variabilidad explicada por el modelo. Un valor más cercano al valor 1, indica un ajuste más adecuado del modelo a los datos observados.
3.  **$p$ -value asociados a los coeficientes:** Los  $p$ -values indican la relevancia estadística de cada variable independiente en el modelo. Un  $p$ -value bajo, sugiere que la variable tiene un impacto significativo en la variable dependiente.

## II.6.5. Visualización de resultados

La visualización de datos es la etapa en la que se representa la información obtenida de manera gráfica y visualmente atractiva para facilitar la comprensión de los resultados. En un estudio informétrico, la visualización de datos desempeña un papel fundamental al permitir la exploración y presentación de los atributos y relaciones de los productos científicos de manera intuitiva (Ebadi & Schiffauerova, 2013). A continuación, se describen brevemente y se ejemplifican algunos tipos de visualizaciones.

### Tablas de contingencia

La tabla de contingencia es una herramienta estadística que organiza la distribución conjunta de variables categóricas en una estructura rectangular, donde cada celda representa el recuento de observaciones en una combinación específica de categorías. Su utilidad radica en analizar relaciones y dependencias entre estas variables, identificando patrones, asociaciones y diferencias significativas (Rousseau et al., 2018c).

## Gráfico de burbujas

Un gráfico de burbujas es una representación tridimensional de datos que utiliza burbujas o nodos para mostrar observaciones. Cada burbuja en el gráfico se posiciona en un plano bidimensional que representa dos variables, mientras que su tamaño refleja una tercera variable. Esto posibilita la visualización simultánea de relaciones y patrones entre tres variables. Las burbujas se sitúan en el plano según sus coordenadas y su tamaño se ajusta en proporción a la magnitud de la tercera variable (Scimago, 2022). La Figura 3 proporciona un ejemplo ilustrativo de este tipo de gráfico.

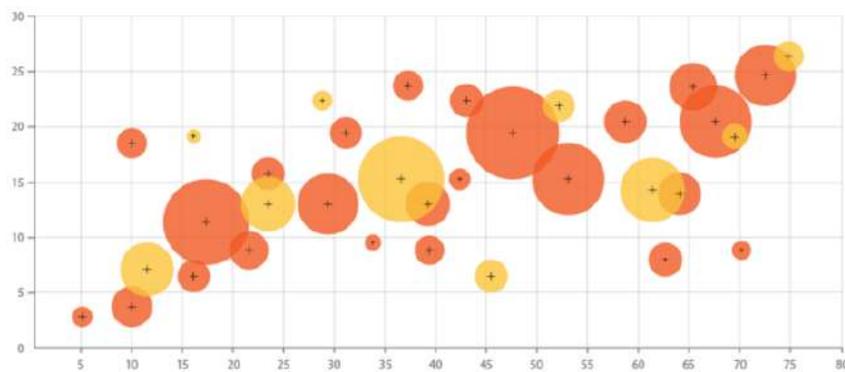


Figura 3: Ejemplo de gráfico de burbujas (Scimago, 2022).

## Gráfico de línea

Un gráfico de líneas conecta puntos de datos mediante líneas rectas para representar la tendencia o patrón a lo largo de la variable independiente. Este tipo de gráfico facilita la identificación de cambios, fluctuaciones y tendencias en los datos a lo largo del tiempo, permitiendo la detección de comportamientos, ciclos o evoluciones (Scimago, 2022).

## Gráfico de barras

Un gráfico de barras emplea barras verticales u horizontales para representar y comparar la magnitud de distintas variables. Cada barra corresponde a una categoría y su longitud refleja la frecuencia. Existen versiones simples, que muestran una única variable, y versiones agrupadas, que permiten comparar múltiples variables dentro de cada categoría (Scimago, 2022).

## Diagrama de Pareto

El diagrama de Pareto es una herramienta gráfica que combina un gráfico de barras con una línea acumulativa para visualizar la distribución de frecuencias de diferentes categorías en función de su importancia relativa. Su uso radica en identificar y priorizar los factores más significativos en un conjunto de datos o problema. Las barras representan las frecuencias de cada categoría en orden descendente, mientras que la línea acumulativa muestra la suma acumulada de estas frecuencias. Esto permite identificar de manera clara las categorías más relevantes que contribuyen en mayor medida al total acumulado (Rousseau et al., 2018c). Un ejemplo de este diagrama se presenta en la Figura 4.

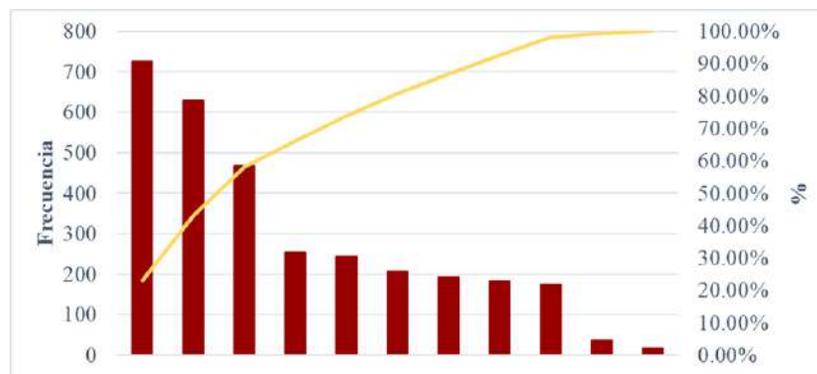


Figura 4: Ejemplo de diagrama de Pareto. (Elaboración propia)

## Gráfico de pastel

Un gráfico de pastel muestra la proporción o porcentaje de diferentes categorías en un conjunto de datos. Se utiliza comúnmente para ilustrar la distribución relativa de las partes en relación con el todo (Rousseau et al., 2018c).

## Diagrama de dispersión

Un diagrama de dispersión es una representación gráfica que exhibe la relación entre dos variables continuas. En este tipo de gráfico, se colocan puntos en un plano cartesiano, donde cada punto representa un par de valores correspondientes a las dos variables. La disposición de los puntos en el diagrama ilustra cómo se distribuyen los valores en relación con ambas variables (Rousseau et al., 2018c).

## Gráfico polar

El gráfico polar es una representación visual que emplea un sistema de coordenadas polar en lugar del sistema cartesiano convencional. En este tipo de gráfico, los datos se representan mediante puntos que se posicionan de acuerdo con su distancia radial y su ángulo. La distancia radial refleja una variable numérica, mientras que el ángulo representa una variable categórica (Scimago, 2022). La Figura 5 muestra un ejemplo de este tipo de gráfico.

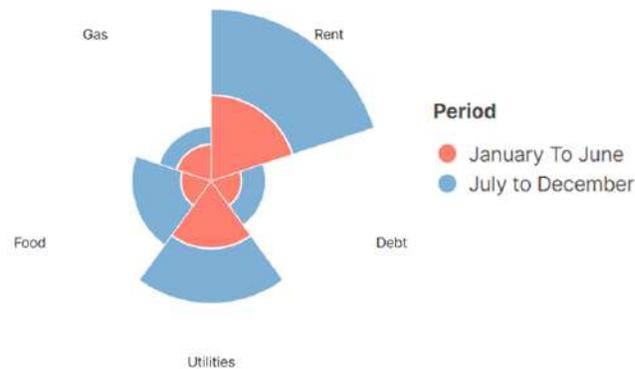


Figura 5: Ejemplo de gráfico polar. (Scimago, 2022)

## Gráfico de treemap

El gráfico de treemap utiliza rectángulos de distintos tamaños para representar la jerarquía y proporción de los datos analizados. Siendo el área de cada rectángulo la frecuencia de la categoría, mientras que la jerarquía se muestra mediante la ubicación y el anidamiento de los rectángulos. Los gráficos de treemap permiten una identificación visual rápida de las partes más significativas y menos significativas de un conjunto de datos (Scimago, 2022). La Figura 6 presenta un ejemplo de este tipo de gráfico.



Figura 6: Ejemplo de gráfico treemap. (Scimago, 2022)



## Diagrama de cajas

El diagrama de caja es una herramienta gráfica que permite visualizar la distribución y dispersión de datos. Cada caja refleja el rango intercuartil (desde el primer cuartil hasta el tercer cuartil), con la línea interna indicando la mediana. Los “bigotes” que se extienden desde la caja muestran la extensión de los datos más allá del rango intercuartil, y los valores atípicos se representan como puntos. Los diagramas de caja son ideales para evaluar simetría, identificar valores atípicos y comparar la variabilidad entre grupos de datos (Rousseau et al., 2018c). La Figura 9 presenta un ejemplo de esta visualización.

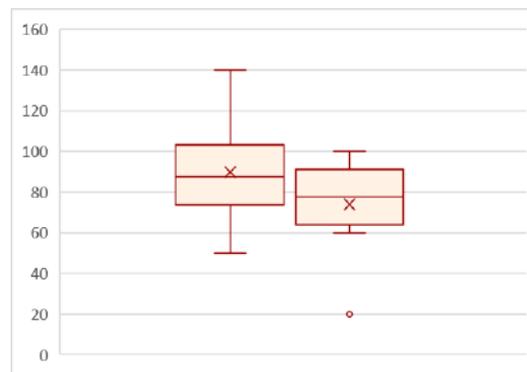


Figura 9: Ejemplo de diagrama de cajas. (Elaboración propia)

# Hipótesis y Objetivos

En este capítulo, se aborda la hipótesis central de la tesis, que se enfoca en la influencia de la colaboración internacional, la multidisciplinariedad y el liderazgo de México en proyectos de investigación en las ciencias sociales, artes y humanidades durante el período de 2010 a 2020. Además, se plantean preguntas de investigación para analizar estos factores y se describen los objetivos del estudio, que incluyen la comparación de tendencias, la exploración de patrones de colaboración y la identificación de correlaciones clave.

### III.1. Hipótesis

Evaluando en el periodo 2010-2020 y tomando en cuenta la limitación de recursos asignados a las ciencias sociales, artes y humanidades en México, se postula que la colaboración internacional, la multidisciplinariedad y el liderazgo en proyectos de investigación impactan significativamente en la calidad de los productos científicos generados en estas áreas. Este impacto se medirá a través de la cantidad de citas por producto científico.

Esta hipótesis surge como respuesta a la disparidad evidente entre el estatus económico de México y la asignación insuficiente de recursos para el desarrollo científico, especialmente en áreas con menor productividad. Busca explorar cómo la colaboración internacional y otros factores han influido en la calidad de la investigación en ciencias sociales, artes y humanidades, destacando la importancia de abordar estas áreas para un desarrollo científico más equitativo y significativo.

### **III.1.1. Preguntas de investigación**

Para el desarrollo de este proyecto de tesis, se establecen las siguientes preguntas de investigación:

- ¿Cómo ha evolucionado la colaboración internacional, la multidisciplinariedad y el liderazgo en proyectos de investigación en México en las ciencias sociales, artes y humanidades durante el periodo 2010-2020?
- ¿Cuáles son los países y afiliaciones con los que México ha colaborado de manera más significativa en las ciencias sociales, artes y humanidades durante el mismo periodo?
- ¿Existe una relación significativa entre la colaboración internacional, la multidisciplinariedad, el gasto en investigación y desarrollo, el liderazgo y la calidad de los productos científicos en las áreas de ciencias sociales, artes y humanidades en México durante 2010-2020?

Estas preguntas buscan profundizar en la situación específica de México en las ciencias sociales, artes y humanidades, abordando la necesidad de estrategias que impulsen estas áreas en el contexto de la limitación de recursos. Exploran las colaboraciones específicas y buscan determinar la influencia de factores clave en la calidad y producción científica, alineándose con la recomendación de la OECD de diversificar la cultura científica y fomentar la colaboración para mejorar el rendimiento científico y tecnológico del país.

## **III.2. Objetivos**

### **III.2.1. Objetivo general**

Analizar y describir las tendencias del 2010 al 2020 en colaboración internacional, multidisciplinariedad y liderazgo de la producción y calidad científica de México en las ciencias sociales, artes y humanidades, utilizando información obtenida de los índices Social Sciences Citation Index® y Arts & Humanities Citation Index® de la base de datos Clarivate - Web of Science.

### **III.2.2. Objetivos específicos**

1. Comparar las tendencias de la producción científica global contra las preferencias de México dentro de las áreas de investigación y el periodo de estudio.

2. Explorar las tendencias de la colaboración internacional, multidisciplinariedad y liderazgo de México dentro de las áreas de investigación y el periodo de estudio.
3. Identificar la correlación entre el presupuesto destinado al desarrollo científico nacional, la colaboración internacional, la multidisciplinariedad, el liderazgo, la producción y la calidad científica de México dentro de las áreas de investigación y el periodo de estudio.



## Metodología

La metodología en este trabajo se divide en cuatro etapas generales: “Recopilación de información”, “Limpieza y preparación de la base de datos”, “Análisis estadístico” y “Visualización de la información”. La Figura 10 permite observar este proceso mediante un diagrama de cajas.

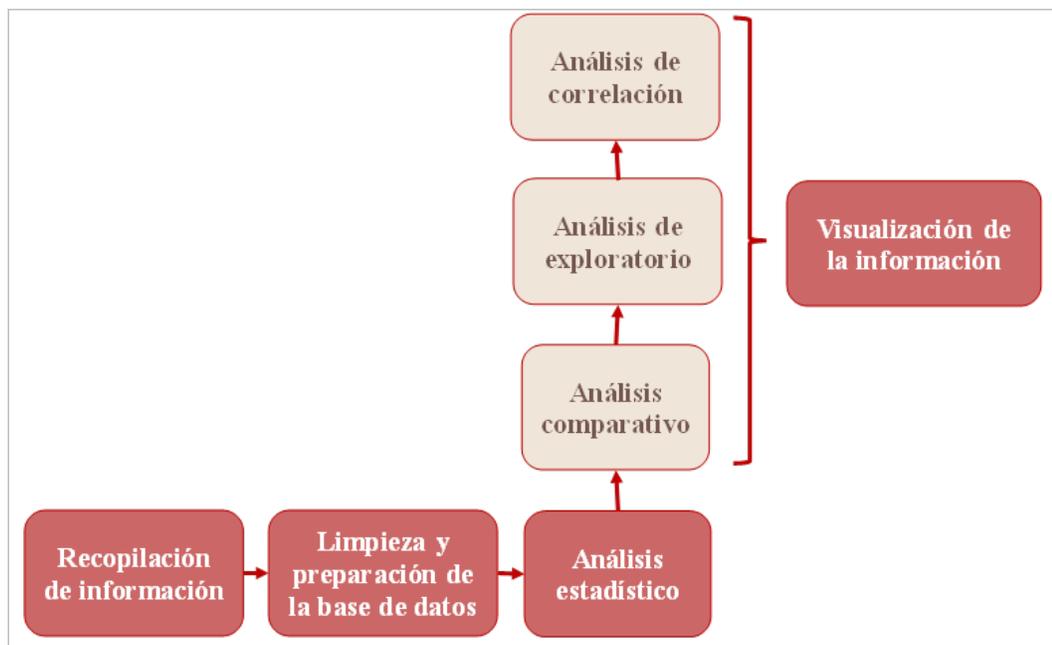


Figura 10: Diagrama de cajas de la metodología aplicada.

Este proyecto dio inicio con la recopilación de la información, considerando la calidad y disponibilidad de los datos de interés. Es importante asegurar que los datos de estudio, aún si se obtuvieron de una fuente confiable, sean estandarizados y estructurados de forma conve-

niente para la exploración de cada variable. Una vez asegurando la calidad de los datos por mano propia, el análisis estadístico permitió dar respuesta a las preguntas de investigación e hipótesis definidas. En este proyecto, esto se realiza por medio de un análisis comparativo entre México y la población, un análisis exploratorio de variables de estudio seleccionadas, así como un análisis de correlación entre las mismas variables (productividad, calidad, colaboración internacional, multidisciplinariedad, liderazgo y presupuesto). Finalmente, la visualización de los resultados es clave para poder emitir y comprender de una manera gráfica los resultados obtenidos.

Para el desarrollo de este proyecto, fueron de utilidad recursos de tipo informático, principalmente, en donde las herramientas y datos tuvieron forma digital. El único recurso físico necesario fue una computadora personal con acceso a Internet para llevar todas las etapas de este proyecto. Cabe destacar que no fue necesario ningún otro tipo de material o equipo especializado.

## IV.1. Recopilación de información

La recopilación de la información se compone de cuatro actividades generales, en donde se encuentran la selección de la base de datos, el filtrado de la información, la exportación de la base de datos y la importación al ambiente de RStudio. En la Figura 11 se muestra el flujo de estas actividades.



Figura 11: Diagrama de cajas de las actividades contenidas en la recopilación de información.

### IV.1.1. Selección de la base de datos

La base de datos seleccionada para este proyecto de tesis fue Clarivate - WoS debido a la alta calidad de la información que ofrece, la basta colección de documentos con la que cuenta y el nivel de especificación que tienen los datos ofrecidos. Clarivate - WoS es una base de datos que permite acceso a su repositorio por medio de Internet en su página (<https://access.clarivate.com/>) por medio de una cuenta registrada. La Universidad Autónoma de Querétaro (UAQ) cuenta con una biblioteca digital que ofrece repositorios de documentos

trabajados en dicha universidad, así como en colaboración con otros institutos. Sin embargo, la UAQ no cuenta con bases de datos como la seleccionada dentro de los recursos ofrecidos a la comunidad. Debido a ello, fue necesario contactar a investigadores con acceso a esta fuente de datos y trabajar en colaboración para la obtención de la información que se utilizaría en este proyecto. Se agradece al equipo perteneciente a la Universidad Autónoma del Estado de Hidalgo (UAEH), específicamente al Instituto de Ciencias de la Salud (ICSa), por tal facilidad.

#### **IV.1.2. Filtrado de la información**

Actualmente, el repositorio de Clarivate - WoS cuenta con millones de registros de literatura científica, lo que resulta una cantidad de registros complicada para descargar y analizar en su totalidad (6,097,338 resultados dentro del periodo de estudio). La interfaz de Clarivate - WoS facilita el filtrado de la información en diferentes formas, incluyendo listas de verificación y queries. El filtrado de la base de datos utilizada para el presente proyecto se realizó mediante el query que se puede consultar en el Código A.1. Este algoritmo considera cuatro variables: El año de publicación dentro del periodo 2010-2020, los índices Social Sciences Citation Index ® y Arts & Humanities Citation Index ® , las categorías WoS pertenecientes a los índices seleccionados y el país registrado de acuerdo con la lista de países estratégicos de México, incluyendo al mismo.

#### **IV.1.3. Exportación de la base de datos**

Una vez obtenida la información acotada de acuerdo con el interés del estudio, fue necesario exportarla en una herramienta para su análisis estadístico. Clarivate - WoS permite la descarga de la base de datos deseada con ciertas barreras. Algunas que se presentaron durante la exportación de la base de datos de este proyecto fueron la cantidad de registros permitida total de descarga y el rango de registros por descarga. Por el tipo de contrato que se tuvo con Clarivate - WoS, la interfaz permitió únicamente la descarga de los primeros 100,000 registros de la consulta en grupos de 1,000 con un formato de exportación de tipo *Tab delimited file* por su compatibilidad con R y Bibliometrix.

#### **IV.1.4. Importación de información a RStudio**

El presente trabajo utilizó el entorno RStudio, basado en el lenguaje de programación R, para realizar análisis estadísticos y la generación de gráficos. R está integrado por un conjunto

de librerías que permite aplicar técnicas y métodos estadísticos a un conjunto de datos. Esta herramienta es preferida por su gran capacidad de análisis matemático y la facilidad de crear visualizaciones de alta calidad (Efrain-García & García-Zorita, 2016). Es posible descargar esta herramienta de forma gratuita desde su página web (<http://www.r-project.org>).

Para importar la base de datos original al ambiente de RStudio, se aplicó la lógica del Código A.2. En este código se importan todos los archivos ubicados en la ruta definida que tengan un formato *.txt*. Esta importación no es la final, ya que es necesario para un mejor manejo de los datos, así como asegurar que el tipo de datos sea el correcto en cada variable para evitar errores. Para ello, se iteró por todas las columnas en búsqueda de los nombres definidos en la siguiente condición: Si el nombre de la columna coincide con alguno de los nombres definidos, se asegura que el tipo de dato sea numérico, en caso contrario, se realizará un cambio de tipo de dato a carácter.

## IV.2. Limpieza y preparación de la base de datos

La limpieza y preparación de la base de datos se compone de cuatro actividades generales, en donde se encuentran la extracción de variables de interés, creación de nuevas variables, la modificación de estructuras y la limpieza de los atributos de estudio. En la Figura 12 se muestra el flujo de estas actividades.

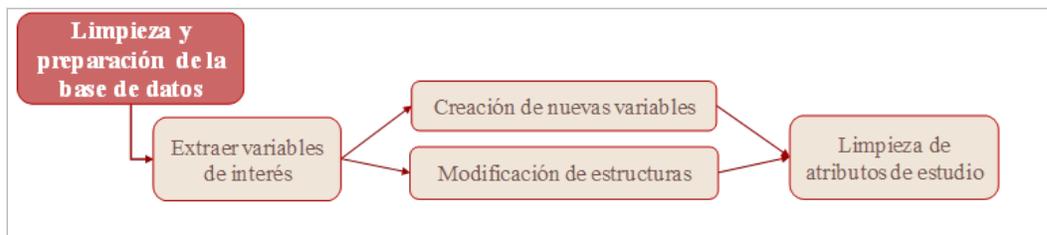


Figura 12: Diagrama de cajas de las actividades contenidas en la limpieza y preparación de la base de datos.

### IV.2.1. Extraer variables de interés

Las variables que interesan al estudio para dar respuesta a las preguntas de investigación y comprobar la hipótesis definida son el tipo de publicación, los nombres de los autores, el título del documento, el idioma en el que se publicó, las palabras clave definidas, las direcciones registradas de los autores, las afiliaciones a las que pertenecen los contribuyentes del proyecto, la cantidad de citas generadas hasta el momento de descarga, el editor, el nombre

de la fuente, el año de publicación, las áreas de investigación involucradas, las categorías e índices WoS a las que pertenecen los documentos. El Código A.3 muestra el algoritmo aplicado para la extracción de variables y en él se describen las variables utilizadas durante el estudio.

#### **IV.2.2. Creación de nuevas variables**

La creación de variables se basa en una ya existente de tipo cadena que contiene aquellos campos que sean de interés, contabilizando y categorizando los atributos. Se emplearon cinco métodos de creación de variables, los cuales fueron aplicados en función de la naturaleza de la variable original y los requerimientos del proyecto. El Código A.5 presenta ejemplos de cada tipo de creación.

El primer caso de creación de variables se realiza para el conteo de las cadenas delimitadas por el carácter “;” como el número de autores participantes en el proyecto, número de palabras clave, número de países colaboradores, número de categorías e índices WoS, número de afiliaciones registradas y número de áreas de investigación.

El segundo tipo de creación realiza una conversión de variables continuas a variables categóricas, buscando establecer rangos para su análisis. Esta lógica fue aplicada para las siguientes variables: Rangos de número de autores, rangos de número de palabras y rangos de números de países participantes.

El algoritmo de creación tipo tres se utilizó específicamente para la variable de países en colaboración ya que la estructura original de la variable de origen difería de las demás variables, conteniendo tanto el nombre de los autores como su país de registro.

El cuarto caso de creación se utilizó específicamente para la generación del atributo de liderazgo. Este consiste en extraer el primer país registrado en el documento interpretado como el origen geográfico del primer autor del proyecto.

Finalmente, el quinto tipo de creación se aplicó para obtener la posición de autoría de México en el documento, el cual se basa en la búsqueda de la posición de la primera aparición de México en la autoría del documento.

#### **IV.2.3. Modificación de estructuras**

La modificación de estructuras de las variables resulta en la modificación del tipo de dato de una variable original. A partir de una cadena, se separaron los valores y se almacenó cada uno de ellos en un vector con el objetivo de determinar frecuencias y repetibilidad de

las separaciones obtenidas. Esta reestructuración fue aplicada para las variables de autores participantes, título de los documentos, palabras clave, categorías e índices WoS, número de afiliaciones y áreas de investigación. El Código A.6 muestra un ejemplo de la lógica aplicada para modificar una cadena a vector.

#### IV.2.4. Limpieza de atributos de estudio

Respetando la naturaleza de cada tipo de variable, se utilizaron cuatro métodos diferentes para la limpieza de las variables. Estos algoritmos se pueden encontrar en el Código A.4.

El primer tipo de limpieza consiste en el aseguramiento de que exista información registrada en todas las observaciones y eliminar espacios dentro de la cadena que puedan provocar duplicados (Bornmann, 2017). Este tipo de limpieza fue aplicada para todos los atributos de estudio adaptándolo al algoritmo ya establecido para cada uno. El tipo de limpieza dos se aplicó especialmente para la variable de países en colaboración. Este tipo de limpieza consiste en evitar el duplicado de países participantes, extrayendo las cadenas de texto únicas que contienen el nombre de los países que colaboran. El caso tres de limpieza es un filtrado de la información para asegurar que los registros que se contemplan en el estudio hayan sido verdaderamente publicados dentro del periodo 2010 - 2020. Finalmente, el tipo de limpieza cuatro fue aplicado a toda la base de datos con el objetivo de filtrar los registros que únicamente cuentan con la presencia de México. Este último tipo de limpieza fue aplicada para la obtención de una base de datos acotada a México para realizar el estudio comparativo y posteriormente, el análisis exploratorio y correlacional.

Al terminar la “Limpieza y preparación de la base de datos”, se obtuvo una base de datos con los atributos descritos en el Cuadro 3, con un nivel de calidad confiable para poder analizar estadísticamente el comportamiento y tendencias de los valores seleccionados.

Cuadro 3: Variables disponibles en la base de datos de este proyecto.

**Nota:** part. = Participantes; No. = Número

Código	Variable	Código	Variable
PT	Tipo de publicación	AU	Autores
nAU	No. de autores part.	rnAU	Rango de autores part.
Continuación en la siguiente página.			

Cuadro 3: Variables disponibles en la base de datos de este proyecto.

**Nota:** part. = Participantes; No. = Número

Código	Variable	Código	Variable
TI	Título del artículo	LA	Idioma
DE	Palabras clave del autor	nDE	No. de palabras clave del autor
Z9	Cuenta de citas	PU	Editor
JI	Nombre de la fuente	PY	Año de publicación
WC	Categorías de WoS	nWC	No. de categorías de WoS
WE	Índice de WoS	nWE	No. de índices de WoS
SC	Áreas de investigación	nSC	No. de áreas de investigación
P_Colab	Países part.	P_Lider	País del primer autor
PosMex	Posición de México en la colaboración	nP_Colab	No. de países part.
rnP_Colab	Rango de países part.	Afiliac	Afiliaciones part.
nAfiliac	No. de afiliaciones part.	-	-

### IV.3. Análisis estadístico

El análisis estadístico de este proyecto consta de tres actividades generales mostradas en la Figura 13: “Análisis comparativo”, “Análisis exploratorio” y “Análisis de correlación”. Esta etapa del estudio es una de las más relevantes ya que aquí se obtienen los principales resultados del estudio, respondiendo a las preguntas de investigación y corroboración de la

hipótesis del estudio.

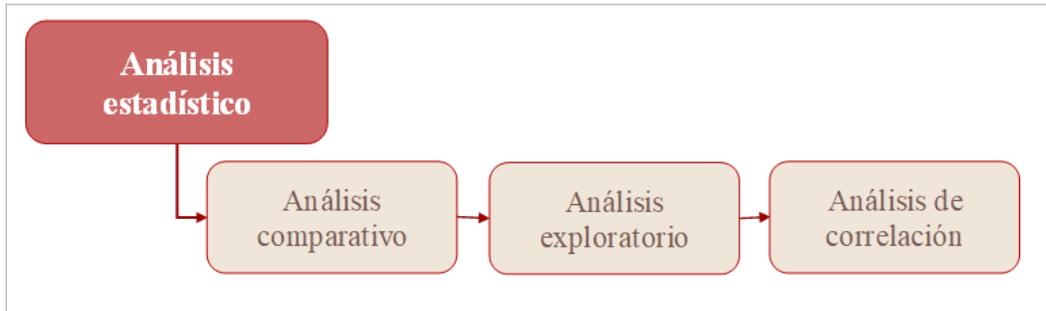


Figura 13: Diagrama de cajas de las actividades contenidas en el análisis estadístico.

### IV.3.1. Análisis comparativo

Para realizar el estudio comparativo fue necesario recurrir a tres herramientas complementarias: Clarivate - WoS, Bibliometrix y el diseño de algoritmos para el manejo de datos. A continuación, se describen detalladamente estas tres actividades. Es importante recalcar que estas fueron aplicadas en dos bases de datos, en donde la primera contiene los 100,000 registros originales descargados, representando la población de estudio, y la segunda conteniendo únicamente los registros en los que México participó.

#### IV.3.1.1. Clarivate - WoS

La interfaz de Clarivate - WoS permite extraer tablas de información específica de la base de datos filtrada. Para este proyecto, se extrajo información específica acerca del idioma de los documentos, las categorías e índices WoS, el editor y las áreas de investigación, en formato *.txt*. Se importó la información obtenida a RStudio para su posterior interpretación y visualización.

#### IV.3.1.2. Bibliometrix

Para hacer uso de las ventajas que Bibliometrix ofrece, fue necesario iniciar con la carga de la base de datos del proyecto a su interfaz por medio del menú *Data*. Una vez importada la base de datos del proyecto, se aseguró que la herramienta realizara el análisis únicamente a los registros que cumplieran con el objetivo del estudio. Para ello, se filtró la información al año de publicación dentro del periodo 2010-2020. Posteriormente, se hizo uso de sus múltiples menús disponibles para obtener datos relevantes para este proyecto de estudio: Por medio del menú *Dataset* se obtuvo la información principal del total de los documentos, la producción

científica anual y la citación promedio por año; del menú *Sources* se obtuvieron las veinte fuentes más relevantes y el impacto por fuente evaluado por citas totales; por medio del menú *Authors* se obtuvieron los datos de los veinte autores más relevantes, su producción a través del tiempo, las afiliaciones más relevantes, los países de los autores de correspondencia, producción científica por país y los países más citados; para finalizar, del menú *Documents* se obtuvieron las palabras más frecuentes presentes en la base de datos considerando los títulos de los documentos estudiados.. Los datos obtenidos de esta fase fueron descargados en formato *.xlsx* para su posterior importación a RStudio.

#### **IV.3.1.3. Algoritmos en R**

Complementando la información obtenida de Bibliometrix y Clarivate - WoS, se utilizaron comandos que R ofrece para la manipulación y análisis estadístico de la información. Las librerías utilizadas para esta etapa del estudio fueron *tidyverse*, *readxl* y *stats*.

La manipulación de los datos se refiere al cambio de la estructura de los datos, buscando obtener información relevante para el estudio. Para ello, se utilizaron comandos y operadores que permitieron el análisis estadístico descriptivo de las variables de estudio. La lógica aplicada para la manipulación y análisis descriptivo de los datos se encuentra en el Código A.7. Este algoritmo es aplicado para el estudio del tipo de producto, los títulos, el editor, el liderazgo de México, las áreas de investigación y las categorías de WoS.

El análisis estadístico de las variables categóricas (rangos de número de autores participantes, rangos de palabras clave, rangos de áreas de investigación, rangos de países colaboradores, rangos de categoría e índices WoS) consistió en ordenar la información de acuerdo con la frecuencia de los rangos especificados. Posteriormente, se obtuvo el porcentaje representativo de la población para su posterior visualización. El algoritmo utilizado para esta actividad, asumiendo la adaptación particular a cada variable, se muestra en el Código A.8.

#### **IV.3.2. Análisis exploratorio**

El análisis exploratorio consistió en estudiar las variables de productividad, visibilidad, colaboración internacional, multidisciplinarietàad, liderazgo y presupuesto de México dentro de las áreas de investigación y periodo definidos para este estudio. Para ello, fue de gran utilidad la información obtenida en el análisis comparativo.

### IV.3.2.1. Productividad

Para estudiar la productividad de México, se hizo el análisis del comportamiento de la acumulación de documentos por año, así como el cálculo del porcentaje que México aportó con base en la productividad global. Se deseaba saber si existía crecimiento en la productividad científica mexicana en este periodo, por lo que se elaboró un modelo de regresión lineal que permitió ver este cambio a través de los años. Se repitió el análisis anterior pero con la división de los datos para cada área de investigación y cada afiliación presente en los registros, considerando las áreas de CSA&H durante el periodo 2010-2020. El Código A.9 muestra la lógica aplicada para el estudio de esta variable. Para la elaboración del diagrama de Pareto fue necesario calcular las columnas de agregación “Porcentaje” y “Porcentaje Acumulado” mediante las Fórmulas IV.1 y IV.2.

$$\text{Porcentaje} = \frac{\text{Productos\_Año}}{\text{Productos\_Total}} \times 100 \quad (\text{IV.1})$$

$$\text{Porcentaje\_Acumulado} = \text{Porcentaje}_1 + \text{Porcentaje}_2 + \dots + \text{Porcentaje}_n \quad (\text{IV.2})$$

En donde *Productos\_Año* es la cantidad de productos científicos generados por año, *Productos\_Total* es la cantidad de registros que se consideran en todo el estudio y *Porcentaje<sub>n</sub>* es el porcentaje en el punto n que se desea incluir en el cálculo acumulativo.

### IV.3.2.2. Visibilidad

El análisis de la visibilidad de la producción científica mexicana consideró la cantidad de citas generadas por documento científico para formar el indicador de evaluación. Se hizo una normalización de las citas por medio de la Fórmula IV.3.

$$\text{Citas\_Normalizadas} = \frac{\text{Citas\_Promedio\_Anual}}{2020 - \text{Año\_Publicación} + 1} \quad (\text{IV.3})$$

En donde *Citas\_Promedio\_Anual* representa el promedio anual de las citas generadas y el *Año\_Publicación* es el año en que se publicó el documento científico. Esta sencilla operación permitió demostrar un comportamiento sin influencias de la maduración de los proyectos científicos evaluados. A partir de este punto, el estudio de esta variable fue muy similar al estudio de la productividad. Se analizó la acumulación de citas anualmente de México en el periodo de estudio 2010-2020, así como el análisis enfocado a cada área de investigación perteneciente a las CSA&H y por afiliación. Posterior a ello, se comparó la pendiente de crecimiento de México con la global mediante la creación de dos modelos lineales, uno para

México y otro para el comportamiento global. Así mismo, para la elaboración del diagrama de Pareto se calcularon las columnas de agregación “Porcentaje” y “Porcentaje acumulado” como en el estudio de la productividad por medio de las Fórmulas IV.1 y IV.2. El Código A.10 muestra la lógica aplicada para el estudio de la visibilidad de la ciencia en México.

#### IV.3.2.3. Colaboración Internacional

Para este proyecto, la colaboración internacional es entendida como la existencia de más de un país contribuyente a un proyecto científico. La lógica aplicada para el estudio de esta variable se muestra en el Código A.11. Buscando entender la tendencia de esta característica particular, se calculó el porcentaje de los registros que representan la colaboraron internacional y aquellos que no se trabajaron mediante esta estrategia. Posteriormente, se calculó el índice de colaboración internacional anual por medio de la Fórmula IV.4.

$$\text{Índice\_Colaboración\_Internacional} = \frac{n\text{Paises}_1 + n\text{Paises}_2 + \dots + n\text{Paises}_n}{n\text{Registros\_Año}} \quad (\text{IV.4})$$

En donde,  $n\text{Paises}_n$  representa el número de países que colaboraron en el registro  $n$  y  $n\text{Registros\_Año}$  es el número de registros que se publicaron en el año de estudio durante el periodo 2010-2020. Se identificó la pendiente de cambio del índice durante la década de estudio por medio de un modelo lineal. Este mismo análisis se llevó a cabo considerando una división entre áreas de investigación pertenecientes a las CSA&H.

#### IV.3.2.4. Multidisciplinariedad

En este estudio, la multidisciplinariedad es entendida como la existencia de más de un área de investigación participante en un proyecto científico. Para evaluar su comportamiento, se llevaron a cabo dos tipos de estudios mostrados en el Código A.12. El primero evalúa las tendencias que hubo durante el periodo de estudio mediante el análisis de frecuencia de las palabras en los títulos de los documentos científicos. Por otro lado, se evalúa el comportamiento del índice de multidisciplinariedad anual por medio de la Fórmula IV.5.

$$\text{Índice\_Multidisciplinariedad} = \frac{n\text{Areas}_1 + n\text{Areas}_2 + \dots + n\text{Areas}_n}{n\text{Registros\_Año}} \quad (\text{IV.5})$$

En donde  $n\text{Areas}_n$  representa el número de áreas de investigación participantes en el registro  $n$  y  $n\text{Registros\_Año}$  el número de registros que se publicaron en el año de estudio dentro del periodo 2010-2020. Así como en el estudio de las variables anteriores, se evaluó

el crecimiento que se tuvo a lo largo de los diez años de estudio mediante un modelo lineal.

#### **IV.3.2.5. Liderazgo**

Para este proyecto, el liderazgo se refiere a la asunción de una posición importante en el proyecto, ya sea en calidad de primer autor o autor de correspondencia, siendo la primera o última posición del autor respectivamente. El Código A.13 muestra la lógica aplicada para estudiar la variable de liderazgo. Se calcularon los porcentajes representativos de los registros en donde México asume una posición de liderazgo, es decir, en donde la posición de México fue la primera o la última. Es posible identificar el liderazgo de México considerando estos dos porcentajes ya que ambas posiciones suelen ser ocupadas por los autores con mayor participación en el proyecto científico. Por otro lado, se hizo el estudio del índice de liderazgo calculado mediante la Fórmula IV.6.

$$\text{Índice Liderazgo} = \frac{\text{Posición}_1 + \text{Posición}_2 + \dots + \text{Posición}_n}{n\text{Registros\_Año}} \quad (\text{IV.6})$$

En donde  $\text{Posición}_n$  representa la posición que México asumió en el registro n y la variable  $n\text{Registros\_Año}$  es el número de registros que se publicaron en el año de estudio. Para definir la pendiente de crecimiento de este periodo, se recurrió a la creación de una regresión lineal.

#### **IV.3.2.6. Presupuesto**

El presupuesto asignado al desarrollo de la ciencia y tecnología de México se estudia por medio del porcentaje de PIB asignado a las actividades que conlleven al cumplimiento de sus objetivos. La información para el estudio de esta variable fue obtenida de la base de datos del Banco Mundial (BancoMundial, 2023). La lógica aplicada para el estudio de esta variable puede encontrarse en el Código A.14. Se determinó el crecimiento mediante una regresión lineal para México y algunos de sus países estratégicos.

### **IV.3.3. Análisis de correlación**

El análisis de correlación de este estudio abarca las variables de productividad, visibilidad, colaboración internacional, multidisciplinariedad, liderazgo y presupuesto. Antes de ello, es importante identificar los valores atípicos en los datos de estudio para su consideración en la interpretación de los resultados de la correlación. Para ello, se utilizaron diagramas de cajas y se identificaron gran cantidad de estos valores. Considerando esto, fue necesario

normalizar las variables de forma ponderada. La Tabla 4 muestra los rangos para cada variable de estudio normalizada y el Código A.15 muestra la lógica aplicada durante esta etapa del análisis estadístico.

Cuadro 4: Rangos establecidos para el estudio de las variables.

Variable	Código	Rangos	Unidad de media
Colaboración Internacional	rnPaises	0, 1, 2-3, 4-5, 6-7, 8 o más	Países involucrados
Liderazgo	rPosMex	0, 1, 2-3, 4-5, 6 o más	Posición de México
Multidisci.	rnSC	1-2, 3-4, 5 o más	Areas de investigación involucradas
Colaboración Afiliaciones	rnAfiliac	0, 1-3, 4-6, 7-9, 10-19, 20 o más	Afiliaciones involucradas
Citas por documento	rcitas	1*, 2**	* Sin citas. ** Con citas

Posteriormente, se generaron gráficos de dispersión para cada par de variables buscando determinar si la correlación entre variables era monótonica o no. Para finalizar, se obtuvieron los coeficientes de Spearman para cada par de variables identificando la existencia de correlación entre ellas y si era positiva o negativa.

#### IV.4. Visualización de la información

Todas las visualizaciones presentadas en este documento fueron elaboradas mediante la librería de R *ggplot2*. A continuación, se enlistan los tipos de gráficos utilizados y su aplicación en este proyecto:

- **Gráfico de línea:** Utilizados para demostrar las tendencias a través del tiempo en un análisis de la productividad, citas por documento, multidisciplinariedad, colaboración internacional, liderazgo, presupuesto y colaboración entre autores (Código A.16);
- **Gráfico de pastel:** Fue aplicado para determinar la presencia de colaboración entre autores y entre países (Código A.17);
- **Gráfico de barras horizontales:** Utilizado para el estudio de las áreas de investigación con mayor colaboración internacional, así como el impacto en citas de las fuentes presentes (Código A.18);

- **Gráfico de treemap:** Se usó para representar la clasificación de la información en más de cinco categorías, específicamente para el tipo de producto, idioma, editor, fuente e índice de Wos (Código A.19);
- **Gráfico polar:** Fue aplicado a la comparación de frecuencias de los rangos definidos para las variables categóricas (A.21);
- **Gráfico de burbujas:** Permitió la comparación de autores en términos de calidad y productividad (Código A.20);
- **Nube de palabras:** Fue utilizada para el análisis exploratorio de las palabras con mayor frecuencia en los títulos de los documentos registrados (Código A.22);
- **Mapa de coropletas:** Usado para identificar visualmente los países que colaboraron con México durante el periodo de estudio (Código A.23);
- **Diagrama de cajas:** Permitió detectar valores atípicos dentro de los datos de estudio en el análisis de correlación (Código A.24);
- **Diagrama de Pareto:** Utilizado para identificar las áreas de investigación y afiliaciones que más aportaron a la productividad y calidad de la ciencia mexicana dentro de las CSA&H (Código A.25); y
- **Diagrama de dispersión:** Utilizado para confirmar el tipo de correlación que hay entre las variables del análisis de correlación (Código A.26).

# Resultados

A lo largo de este capítulo se describirán y discutirán los resultados obtenidos durante la aplicación de la metodología descrita en el capítulo “Metodología” IV. Las etapas “Análisis estadístico” y “Visualización de la información” se presentarán en una misma sección de este capítulo, ya que las visualizaciones fueron generadas a partir de los resultados del análisis estadístico.

### V.1. Recopilación de la información

A partir de la selección de la base de datos Clarivate - WoS y el uso del query proporcionado, se obtuvo un total de 104,166 resultados en la interfaz de la página. Se descargó un conjunto de 100,000 observaciones, lo que representa un 96.00% del total de registros disponibles en la página. Esta base de datos original consta de 71 atributos, los cuales se encuentran detallados en el Anexo 13. Para facilitar su uso en diferentes herramientas, se realizó la exportación de la información en formato de archivo de texto plano.

### V.2. Limpieza y preparación de la base de datos

Se utilizaron dos bases de datos para comparar el desarrollo científico entre México y las tendencias globales. La primera contiene 99,992 registros y la segunda, 594. La diferencia principal radica en que la primera abarca todos los registros que cumplen los criterios, mientras que la segunda se limita a observaciones relacionadas con México. La calidad de la información en ambas bases es sólida, con más del 99% de los atributos registrados, excepto las palabras clave de los autores, presentes en alrededor del 24.78% en la base de

datos completa y el 20.53 % en la centrada en México. Esto se debe principalmente a la prevalencia de “resúmenes de reuniones,” que generalmente no incluyen palabras clave de los autores (Rousseau et al., 2018b). Esto contrasta con las expectativas previas basadas en la producción de libros y artículos científicos en las ciencias sociales, artes y humanidades (Sandoval-Romero & Larivière, 2020).

### **V.3. Análisis estadístico y visualización de la información**

El análisis estadístico de este proyecto fue constituido por tres tipos de análisis: “Análisis comparativo”, “Análisis exploratorio” y “Análisis de correlación”. A continuación, se detallan los resultados y discusiones en casa una de las etapas mencionadas.

#### **V.3.1. Análisis comparativo**

El análisis comparativo tuvo como objetivo contrastar las tendencias y patrones de México en las CSA&H contra lo global, buscando características de los productos científicos que tengan mayor recurrencia o bien, que coincidan entre ambas bases de datos.

##### **V.3.1.1. Idiomas**

En ambas bases de datos, el inglés prevalece como el idioma principal, representando el 99.72 % en la población general y el 98.65 % en la base de datos centrada en México. Esto se alinea con investigaciones anteriores que señalan un sesgo hacia el inglés en la plataforma WoS Sandoval-Romero y Larivière, 2020. Además del inglés, se encuentran otros idiomas en la base de datos global, aunque en proporciones reducidas, como el francés (0.130 %), español (0.060 %), portugués (0.046 %), alemán (0.030 %), entre otros. En contraste, la base de datos de México registra principalmente documentos en inglés y español (1.346 %). Aunque se esperaba una mayor presencia de otros idiomas en la base de datos de México, el predominio del inglés se debe a su estatus como idioma universal en la colaboración científica internacional.

##### **V.3.1.2. Editores y fuentes**

Se identificaron 254 editores en la base de datos global y 36 en la base de datos de México (Fig. 14). En la base global, Elsevier lidera con el 20.39 % de los productos científicos, seguido por Wiley (19.66 %) y Springer Nature (17.11 %). En la base de datos de México, Wiley encabeza con el 22.72 %, seguido por Wiley-Blackwell (13.80 %) y la American Society

of Tropical Medicine and Hygiene (12.62%). Wiley es el editor principal en ambas bases, abarcando diversas áreas de investigación. La inclusión de la American Society of Tropical Medicine and Hygiene es relevante debido a su asociación con la salud global y no solo con las CSA&H.

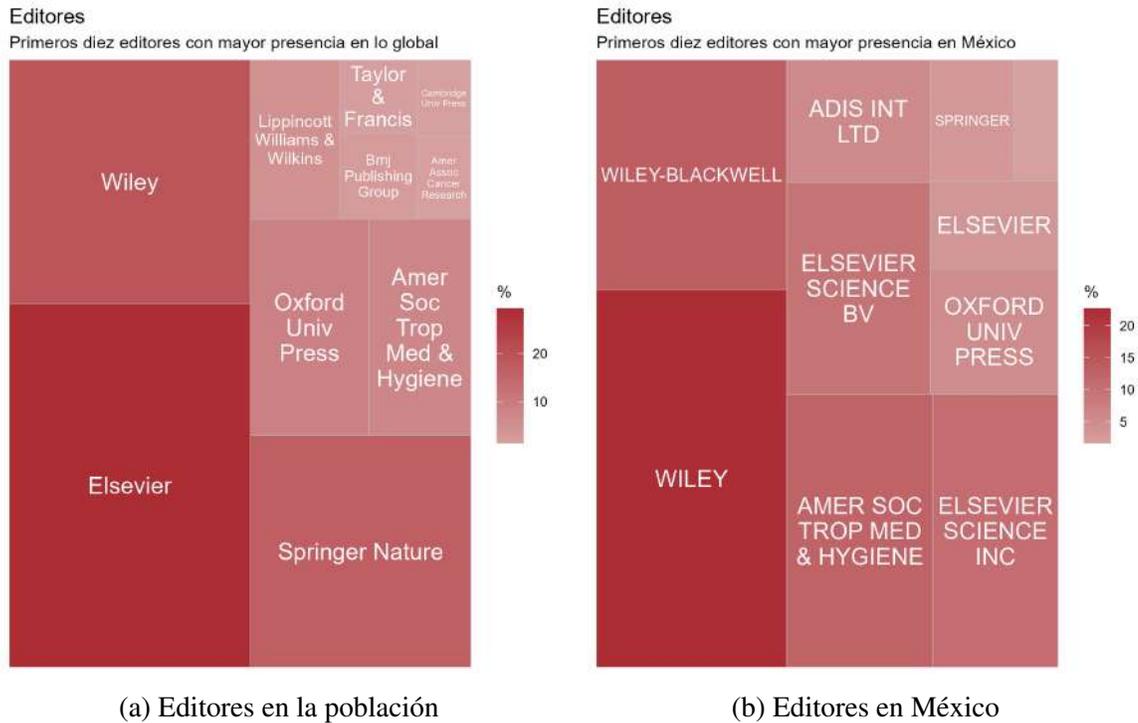


Figura 14: Editores de los documentos en la población y en México.

Las revistas más productivas en la base de datos global son “Journal of General Internal Medicine” (11.12%), “American Journal of Physical Anthropology” (10.49%) y “Biological Psychiatry” (9.09%) (Fig. 15). Estas cubren diversas áreas, desde medicina interna hasta la evolución humana y trastornos mentales. Entre las veinte revistas más influyentes, solo “American Journal of Physical Anthropology” aborda temas de ciencias sociales y desarrollo, y ninguna se centra en artes y humanidades. Juntas, “Journal of General Internal Medicine” y “American Journal of Physical Anthropology” representan el 21.61% de la producción científica en la base de datos global.

En México, las revistas más productivas son “American Journal of Physical Anthropology” (15.99%), “Bipolar Disorders” (15.82%) y “American Journal of Tropical Medicine and Hygiene” (12.63%) (Fig. 15). Sin embargo, en la base de datos mexicana hay una diversificación mayor, incluyendo revistas como “Value in Health” (3.87%), “Scientometrics” (1.52%), “Ancient Mesoamerica” (1.01%), “American Journal of Human Biology” (0.84%)

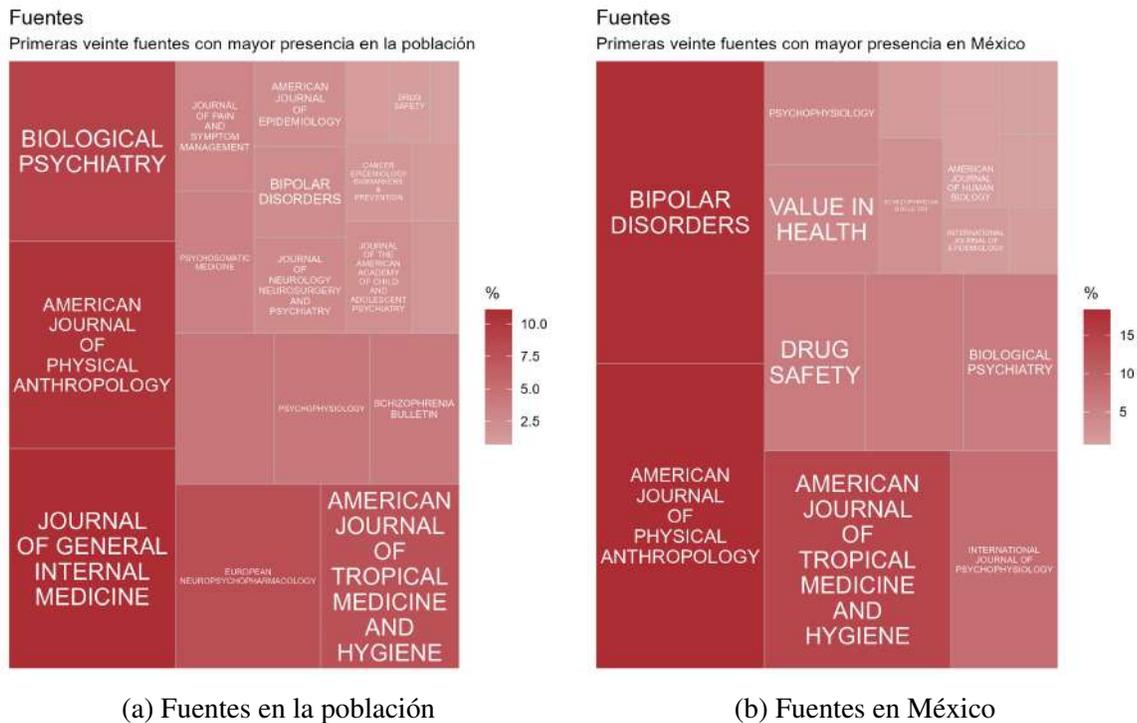


Figura 15: Producción de las fuentes presentes en la población y en México.

y “Design Journal” (0.67%). En conjunto, “American Journal of Physical Anthropology” y “Bipolar Disorders” representan el 31.82% de los registros en esta base de datos.

En cuanto a calidad, las fuentes más citadas en la base de datos global son “Radiation Protection Dosimetry (5,211 citas desde 2010), “Biological Psychiatry” (2,522 citas desde 2010) y “Schizophrenia Bulletin” (2,458 citas desde 2011) (Fig. ??). En la base de datos centrada en México, destacan “Scientometrics” (46 citas desde 2017), “American Journal of Human Biology” (31 citas desde 2015) y “Ancient Mesoamerica” (11 citas desde 2015).

Es importante destacar que en la base de datos global no se encuentran fuentes relacionadas con áreas de investigación en CSA&H, tanto en términos de registros de citas como de producción científica. En contraste, en la base de datos centrada en México, sí se identifican fuentes en estas áreas. Sin embargo, es notable la presencia significativa de revistas médicas en ambas bases de datos, particularmente relacionadas con la salud mental.

### V.3.1.3. Afiliaciones

Las principales afiliaciones en la base de datos global incluyen la Universidad de Pittsburgh (2.96%), la Universidad de California Los Ángeles (2.34%), la Universidad de Harvard

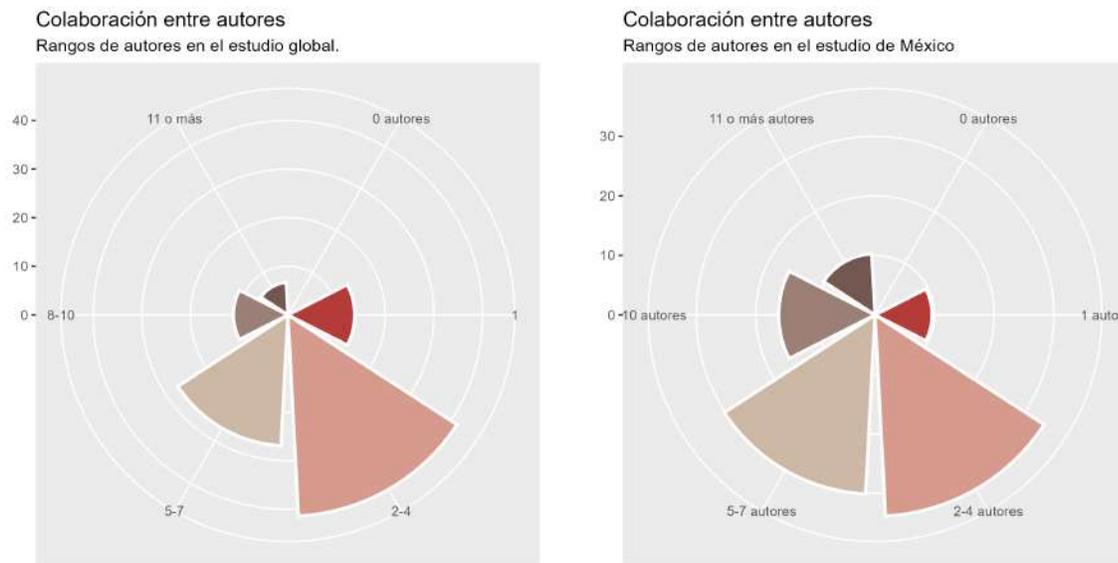


ciencias sociales, artes o humanidades entre las más frecuentes, pero algunas menos usuales como “maya,” “latin america,” y “financiamiento” indican documentos relacionados con la cultura mexicana, latinoamericana y desarrollo socioeconómico.

En cuanto a la cantidad de palabras clave por documento, en ambas bases de datos, la mayoría de los registros tiene de 2 a 4 palabras (12.76 % y 11.11 %), seguido por 5 a 7 palabras (11.38 % y 8.58 %). Esto refleja que la mayoría de los documentos no emplea muchas palabras clave, permitiendo una descripción concisa de su contenido.

### V.3.1.5. Rendimiento de autores

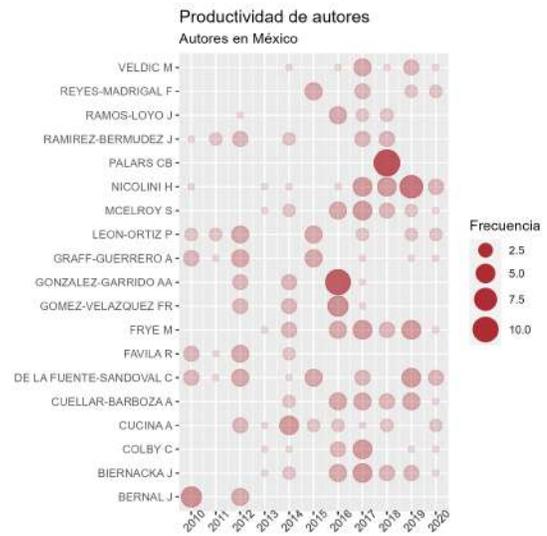
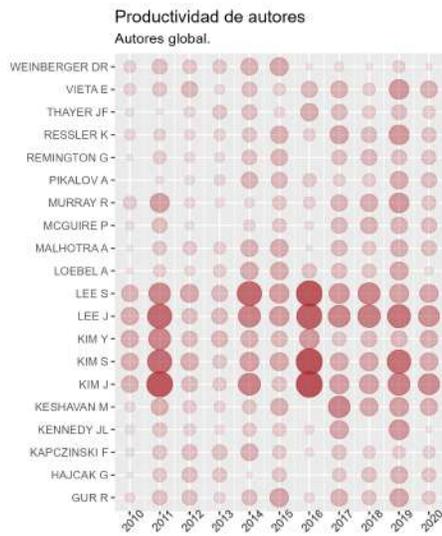
Las Figuras 18 muestran el rendimiento de autores individuales por año. En cuanto a productividad, se resaltan nombres como “LEE J” con 478 productos científicos, “LEE S” con 435, y “KIM J” con 429. Por otro lado, evaluando su desempeño en términos de citas normalizadas, “KIM S” lidera con un promedio de 104.04 citas al año, seguido por “LEE J” con 88.75, y “LEE S” con 80.05. Estos autores han tenido un papel esencial en la producción y visibilidad de la investigación en la última década.



(a) Rango de autores en la población

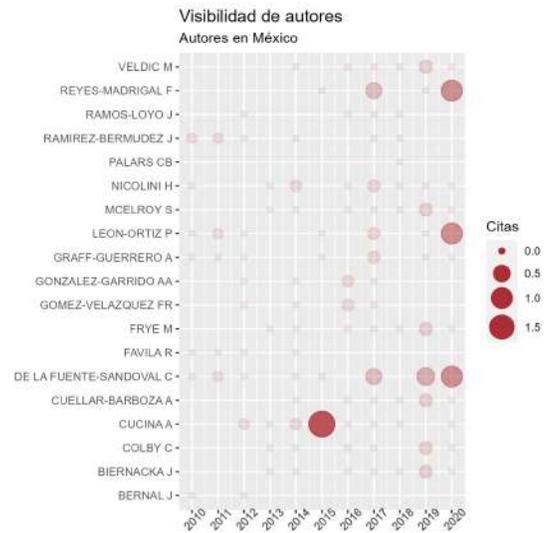
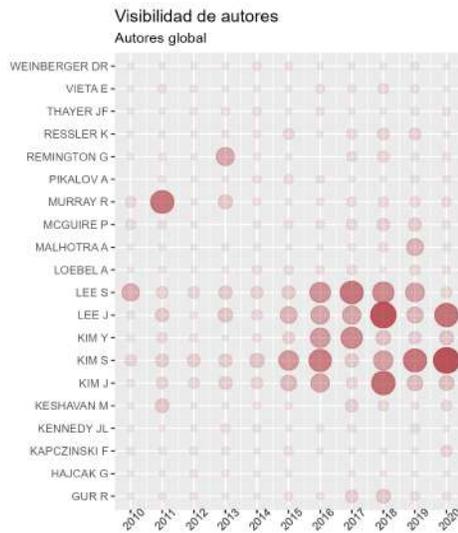
(b) Rango de autores en México

Figura 17: Rango de autores participantes por producto en la población y en México.



(a) Productividad de autores en la población

(b) Productividad de autores en México



(c) Visibilidad de autores en la población

(d) Visibilidad de autores en México

Figura 18: Productividad y visibilidad de autores en la población y en México.

En cuanto a la colaboración entre autores, el rango más común implica de 2 a 4 autores, representando el 41.44% en la base de datos global y el 33.83% en la centrada en México (Fig. 17). Sin embargo, en la base de datos mexicana, se observa una mayor frecuencia de colaboración con 5 a 7 autores (30.13%), 8 a 10 autores (16.16%), y 11 o más autores (10.26%). Esto indica una tendencia más marcada a trabajar en equipos con múltiples autores en la investigación científica mexicana en comparación con la base de datos global. En México, el 90.40% de los registros en la base de datos son colaborativos, mientras que en la base de datos global, este porcentaje es del 86.26%.

La colaboración entre autores en documentos en los que México participa supera significativamente a la base de datos global. El índice de colaboración en estos documentos es de 4.04, el doble que en la base de datos global (2.07). Además, el número promedio de autores por documento en la base de datos centrada en México (4.04 autores) es un 213.75% más alto que en la base de datos global (1.89 autores). Estos indicadores reflejan tanto la capacidad científica de México como una sólida cultura de colaboración en la investigación científica en el país (OECD, 2015).

### **V.3.2. Análisis exploratorio**

El análisis exploratorio se centró en examinar el comportamiento de las variables clave de este estudio: colaboración internacional, multidisciplinaria, liderazgo, presupuesto, calidad y productividad en el contexto de las CSA&H en México durante el periodo de estudio 2010-2020. Es esencial subrayar que la selección de las variables para esta sección del análisis se basó en los factores previamente demostrados como impactados por la implementación de la estrategia de colaboración internacional.

#### **V.3.2.1. Calidad y productividad**

En términos de calidad y citación de documentos científicos, se observa que, en promedio, hay 2.289 citas por documento a nivel global y 1.988 en México (Fig. 19). Sin embargo, es importante destacar que es poco común que un producto científico tenga citas registradas, con solo el 19.88% de los registros en la base de datos global y el 19.86% en la base de datos de México que cuentan con ellas. Esto puede sugerir una presión por la producción científica en lugar de la calidad del trabajo, lo que podría distorsionar los procesos de investigación y desincentivar los proyectos a largo plazo.

La Figura 20 muestran la producción científica en México. En promedio, se publicaron 7.24 documentos al año a nivel global y 7.03 en México. El año más productivo en México



Figura 19: Impacto científico evaluado en visibilidad en la población y en México.

fue 2018 con 125 artículos, mientras que el menos productivo fue 2013 con 17 artículos. Globalmente, se observa una tendencia al crecimiento con un índice anual del 6.2%. En el gráfico de México, se nota un crecimiento con un índice anual del 3%.

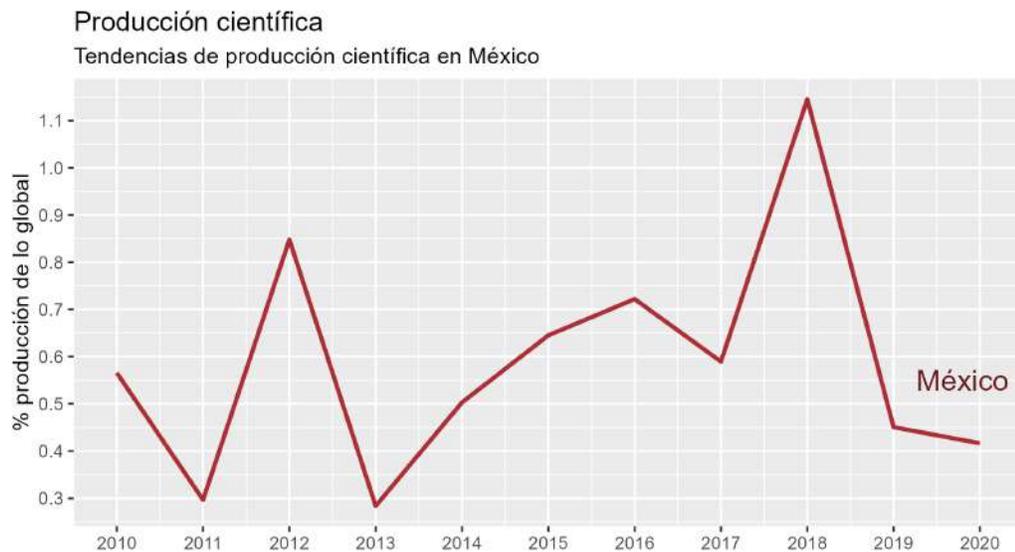


Figura 20: Impacto científico evaluado mediante la productividad de México.

El modelo de regresión lineal en CSA&H en México no mostró una relación significativa entre el año y la calidad científica (Cuadro 5), con un coeficiente estimado de 0.0003453

citas promedio por año y un R cuadrado ajustado de -0.1085, lo que sugiere que el tiempo no influye significativamente en la calidad de la investigación en esta área.

Cuadro 5: Modelo lineal para determinar el crecimiento de la calidad de la producción científica mexicana en CSA&H.

Calidad				
Variable	Coefficiente	Error Estándar	Valor t	Valor p
(Intercept)	-0.6680	4.8148	-0.139	0.893
Año	0.0003	0.0023	0.144	0.888
Residual SE		0.0250		
R-cuadrado ( $R^2$ )			0.0023	
$R^2$ ajustado			-0.1085	

En cuanto a la productividad científica, se observó un crecimiento gradual en la producción promedio anual de 0.01447 productos científicos por año (Cuadro 6). Sin embargo, este aumento no alcanzó significación estadística, con un valor p de 0.574, lo que indica que no hay evidencia suficiente para afirmar que la tendencia es diferente de cero. A pesar de la falta de significación estadística, estos resultados ofrecen una visión general de la evolución de la productividad científica en CSA&H en México, que podría ser útil para futuros análisis y estrategias de investigación en el país.

### V.3.2.2. Colaboración internacional

La Figura 21 presenta un análisis del índice de colaboración internacional en México dentro del campo de las CSA&H. La colaboración internacional en México en el ámbito de CSA&H es más frecuente que en la población global, representando el 40.48% en México y el 21.74% en la población general. Un modelo de regresión lineal (Cuadro 21), mostró una tendencia significativa de crecimiento en el índice de colaboración internacional en CSA&H en México, con un coeficiente anual estimado de 0.06360, lo que sugiere que ha aumentado en promedio 0.06360 países por año durante el período estudiado. Además, alrededor del 48.61% de la variabilidad en el índice de colaboración puede ser explicada por el factor temporal.

Cuadro 6: Modelo lineal para determinar el crecimiento de la productividad de la producción científica mexicana en CSA&H.

Productividad				
Variable	Coficiente	Error Estándar	Valor t	Valor p
(Intercept)	-28.5634	50.0063	-0.571	0.582
Año	0.0144	0.0248	0.583	0.574
Residual SE		0.2603		
R-cuadrado ( $R^2$ )			0.0363	
$R^2$ ajustado			-0.0706	

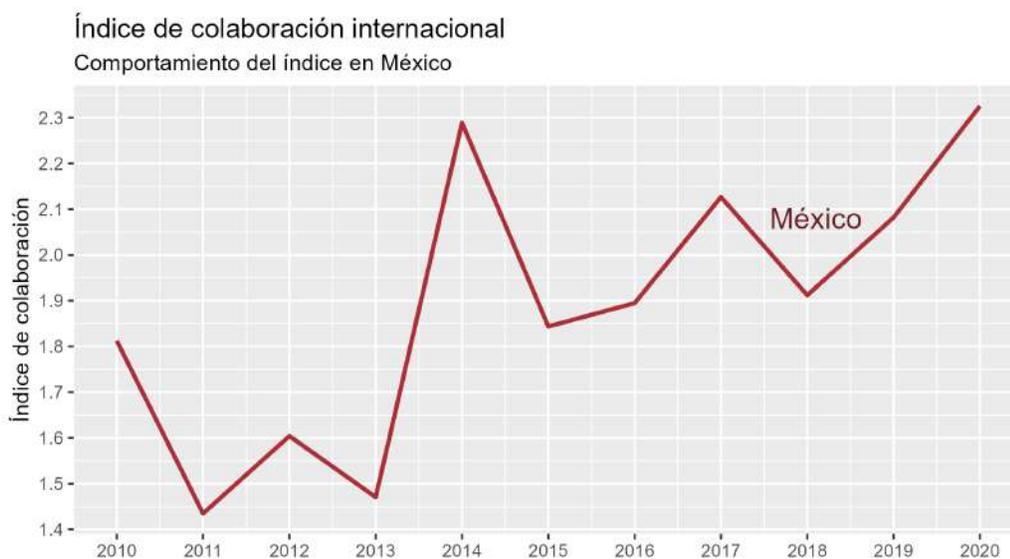


Figura 21: Índice de colaboración internacional de México.

México se destaca como el segundo país latinoamericano con mayor presencia en la producción científica mundial, después de Brasil, superando a Colombia y Argentina. La colaboración internacional de México se da principalmente con países europeos (72.41%), seguidos de América (62.70%), Asia (12.63%), África (3.96%) y Oceanía (1.93%). Los principales países con los que México colabora incluyen Estados Unidos, Canadá, Colombia, España, Francia, Reino Unido, Argentina, Australia, Alemania e Italia, representando el 21.04% de los registros en la base de datos. La Figura 22 muestra un mapa de coropletas que identifica los países presentes y su producción durante la década de estudio.

Cuadro 7: Modelo lineal para determinar el crecimiento de la colaboración internacional en México dentro de las CSA&H.

Índice de colaboración internacional				
Variable	Coefficiente	Error Estándar	Valor t	Valor p
(Intercept)	-126.254	43.9166	-2.875	0.0183 *
Año	0.0636	0.0217	2.918	0.0171 *
Residual SE		0.2286		
R-cuadrado ( $R^2$ )			0.4861	
$R^2$ ajustado			0.429	

Melin y Persson, 1996 mencionan que la relación de México con los países mencionados puede ser derivado a la influencia de factores como la distancia geográfica, barreras culturales, lingüísticas y políticas en las interacciones de colaboración internacional, especialmente con países de América. Estos resultados coinciden con investigaciones anteriores que destacan la colaboración de México con Estados Unidos, España y Francia (Zhe et al., 2021). Además, se confirma una colaboración significativa con Italia, como se sugiere en otro estudio (Lancho-Barrantes & Cantú-Ortiz, 2019). De los 25 países con los que México colabora, no se encontraron colaboraciones estratégicas con China, Corea del Sur, India, Israel, Rusia, Sudáfrica y Turquía, lo que respalda los hallazgos previos.

El área de investigación “psiquiatría” es la que principalmente genera productos en colaboración internacional con el 43.68 % de sus productos. En segundo lugar, se encuentra el área de “antropología”, con el 38.94 % de sus productos. Las áreas de investigación “historia,” “literatura,” “religión,” “sociología,” “filosofía” y “lingüística” obtuvieron nula colaboración con otros países. Es relevante mencionar que estas áreas fueron igualmente las áreas con menor producción científica durante la década, pero no necesariamente las de menor visibilidad.

En cuanto a la cantidad de países que tienden a colaborar en la base de datos global, es predominante el trabajo individual a una escala país (78.26%) y es relativamente nula la colaboración con más de tres países. Para México, también es más recurrente trabajar de manera individual (51.51%), pero la tendencia a trabajar con 2 o 3 países incrementa a comparación de la población (39.56%).

## Países en colaboración con México

Identificación de países que colaboraron con México durante el periodo 2010 y 2020

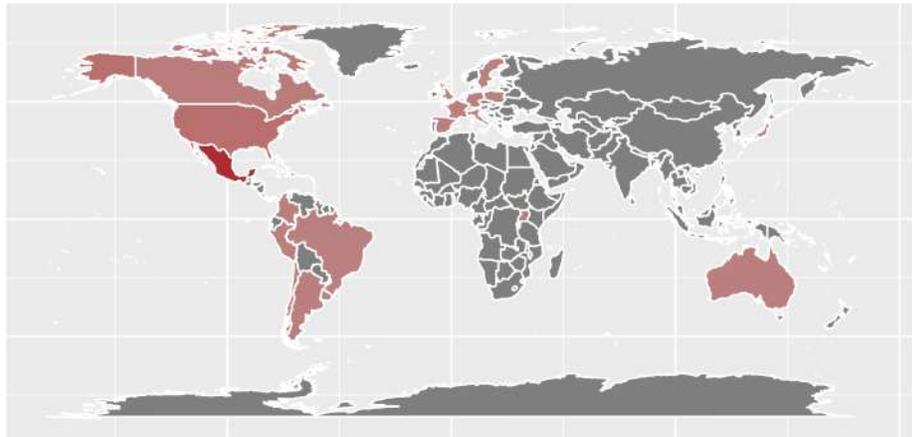


Figura 22: Países que colaboraron con México durante el periodo de estudio.

### V.3.2.3. Multidisciplinariedad

La Figura 23 presenta el estudio del índice de multidisciplinariedad en el contexto de las CSA&H en México. Este análisis no mostró una relación estadísticamente significativa entre el tiempo y la multidisciplinariedad (Cuadro 8). El coeficiente de año estimado fue muy bajo ( $-0.001332$  áreas por año), y el modelo no logró explicar eficazmente la variabilidad en los datos, con un valor de R-cuadrado múltiple extremadamente bajo ( $0.0003252$ ). Esto indica que no hay una tendencia clara en la multidisciplinariedad de la ciencia en CSA&H en México a lo largo de los años estudiados.

En la base de datos centrada en México, se identificaron 49 áreas de investigación, pero las 10 más comunes representan la mayoría de los registros (Fig. 24). Destacan áreas como “psiquiatría” (14.91 %) y “antropología” (7.80 %), que en conjunto constituyen el 22.72 % de los registros. Además, ocho de las diez áreas más comunes no están relacionadas con la salud, lo que sugiere una diversidad de áreas en la investigación científica de México.

Se observa que las áreas médicas siguen siendo prominentes en ambas bases de datos, lo que refleja la creciente importancia de la salud mental en la última década (Kara, 2015; Sandoval-Romero & Larivière, 2020). México tiende a abordar dos (62.62 %) o tres (25.42 %) áreas de investigación en sus productos científicos, lo que indica la multidisciplinariedad en los documentos. La investigación científica en México muestra una diversidad de áreas y una multidisciplinariedad que reflejan las necesidades y tendencias actuales en la sociedad.

Cuadro 8: Modelo lineal para determinar el crecimiento de la multidisciplinariedad en México dentro de las CSA&H.

Índice de multidisciplinariedad				
Variable	Coficiente	Error Estándar	Valor t	Valor p
(Intercept)	4.8653	49.6163	0.098	0.924
Año	-0.0013	0.0246	-0.054	0.958
Residual SE		0.2583		
R-cuadrado ( $R^2$ )			0.0003	
$R^2$ ajustado			-0.1107	

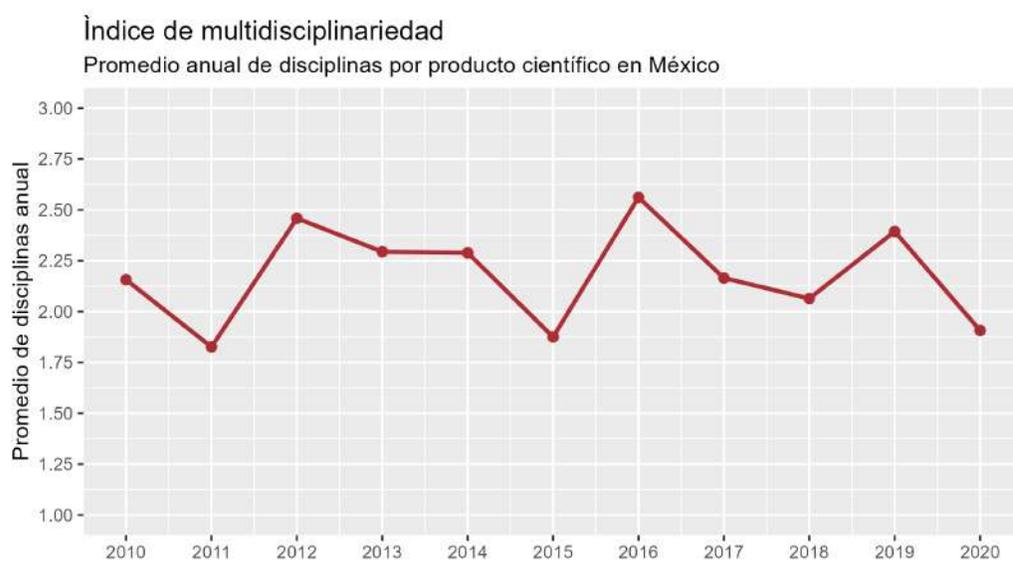


Figura 23: Índice de multidisciplinariedad de México en las CSA&H.

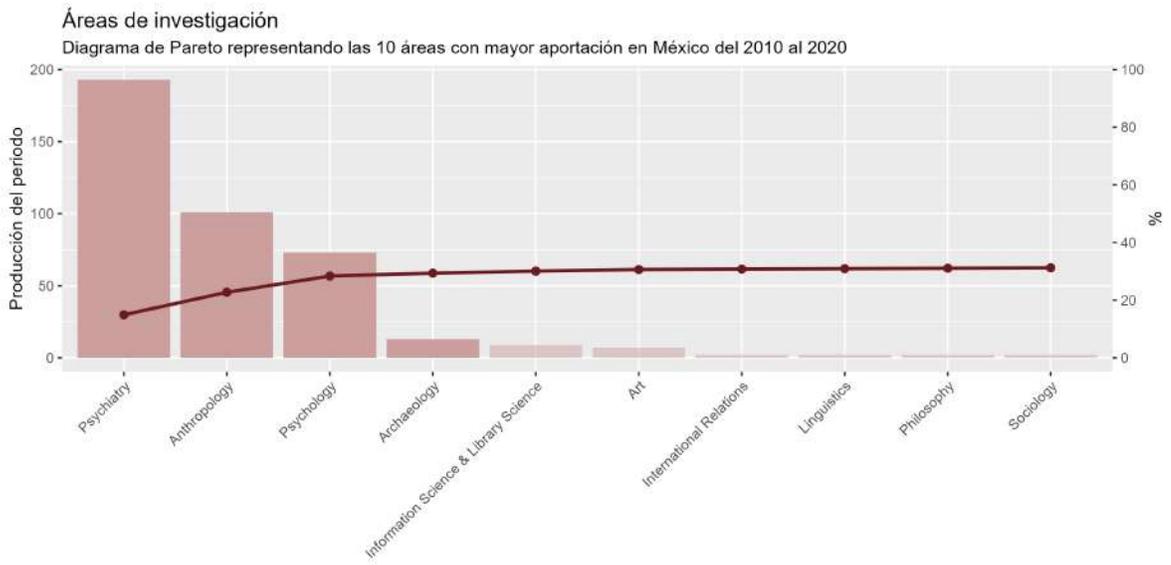


Figura 24: Áreas de investigación más productivas dentro de las CSA&H de México.

#### V.3.2.4. Liderazgo

La Figura 25 presenta la evolución del liderazgo de México en el ámbito de las CSA&H en el país, mostrando una tendencia positiva a lo largo de los años estudiados (Cuadro 9).

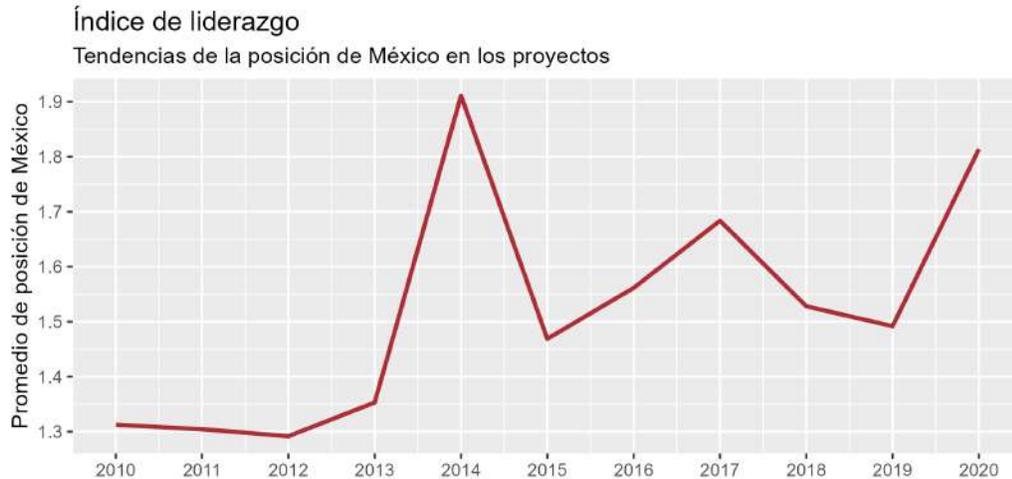


Figura 25: Índice de liderazgo de México en las CSA&H.

El coeficiente de año estimado fue de 0.03889, lo que indica un crecimiento promedio de 0.03889 posiciones por año en el liderazgo de México en este campo. Además, el valor de R-

cuadrado múltiple sugiere que alrededor del 37.73 % de la variabilidad en el liderazgo puede explicarse por el factor tiempo. Estos resultados respaldan la creciente capacidad investigativa de México en CSA&H y su liderazgo en este campo.

Cuadro 9: Modelo lineal para determinar el crecimiento del liderazgo de México dentro de las CSA&H.

Índice de liderazgo				
Variable	Coefficiente	Error Estándar	Valor t	Valor p
(Intercept)	-76.8376	33.5581	-2.290	0.0478 *
Año	0.0388	0.0166	2.335	0.0444 *
Residual SE		0.1747		
R-cuadrado ( $R^2$ )			0.3773	
$R^2$ ajustado			0.3081	

En la base de datos centrada en México, se identificaron 28 países líderes en la investigación en CSA&H. Los diez países más frecuentes en términos de liderazgo en estos registros son México (64.98 %), Estados Unidos (19.19 %), Canadá (2.02 %), Italia (1.85 %), Inglaterra (1.34 %), Brasil (1.17 %), España (1.17 %), Chile (1.17 %), Suiza (0.84 %) y Cuba (0.67 %). Estos países provienen de diversos continentes, con una destacada representación de Europa (33.33 %) y América (33.33 %). México se destaca aún más al ocupar una posición de liderazgo en el 24.75 % de los registros como autor de correspondencia, lo que refleja su influencia destacada en CSA&H en la región.

### V.3.2.5. Presupuesto

El presupuesto de México asignado a su desarrollo científico y tecnológico se estudió con el objetivo de detectar un crecimiento significativo durante el periodo de estudio. Así mismo, se obtuvieron resultados para los países estratégicos de México con mayor y peor presupuesto asignado. En el Cuadro 10 y en la Figura 26 se exponen los resultados de cada país.

En el caso de México, se observa que el coeficiente estimado para el año es de -0.0002064, lo que indica una disminución promedio de 0.0002064 unidades por año en el porcentaje del

Gasto en investigación y desarrollo  
 Porcentaje del PIB de México y algunos de sus países estratégicos

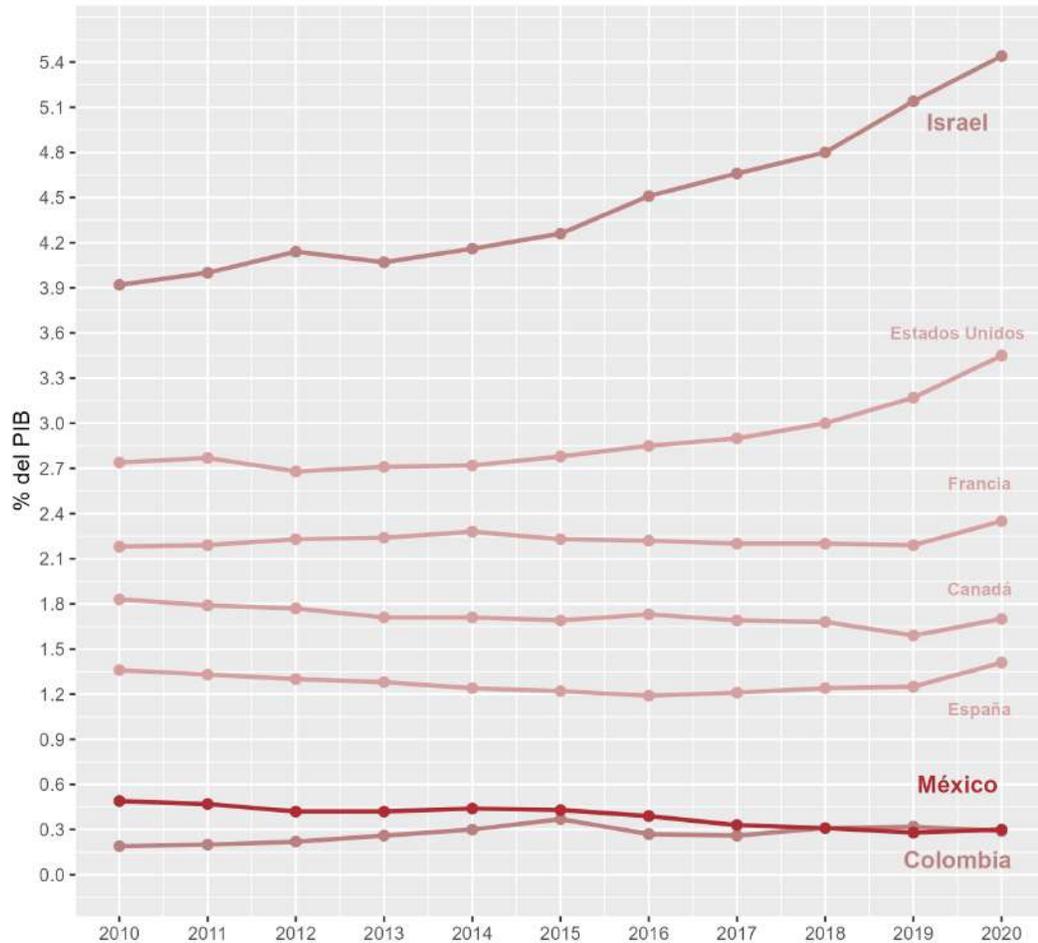


Figura 26: Presupuesto asignado de México y algunos de sus países estratégicos.

PIB asignado a la ciencia y tecnología durante el período analizado. Estos resultados sugieren que México ha experimentado un crecimiento significativo en su presupuesto asignado a la ciencia y tecnología en el tiempo estudiado. Vale la pena destacar que otros países como Israel, Estados Unidos, Canadá y Colombia también han mostrado un crecimiento significativo en esta área a lo largo del tiempo. En particular, Israel se destaca por exhibir el mayor crecimiento en el porcentaje del PIB asignado a la ciencia y tecnología, con un coeficiente de año estimado de 0.0014245, lo que implica un aumento promedio de 0.0014245 unidades por año en la inversión en investigación científica y tecnológica en Israel. Por otro lado, España presenta el menor crecimiento en este indicador, con un coeficiente de año estimado de -0.0000400, lo que indica una disminución promedio de 0.0000400 unidades por año en

el porcentaje del PIB destinado a la investigación científica y tecnológica en España. Estos hallazgos proporcionan una visión general de cómo los presupuestos asignados a la ciencia y tecnología han evolucionado en diferentes naciones estratégicas a lo largo del tiempo.

Cuadro 10: Modelos lineales para determinar el crecimiento del presupuesto asignado.

<b>Presupuesto asignado</b>				
<b>Variable</b>	<b>Coficiente</b>	<b>Error Estándar</b>	<b>Valor t</b>	<b>Valor p</b>
<b>Israel</b>				
(Intercept)	-2.8258	0.3112	-9.079	7.95e-06 ***
Año	0.0014	0.0001	9.223	6.99e-06 ***
Residual SE		0.0016		
R-cuadrado ( $R^2$ )			0.9043	
$R^2$ ajustado			0.8937	
<b>Estados Unidos</b>				
(Intercept)	-1.1837	0.2575	-4.596	0.0013 **
Año	0.0006	0.0001	4.708	0.0011 **
Residual SE		0.0013		
R-cuadrado ( $R^2$ )			0.7112	
$R^2$ ajustado			0.6791	
<b>Francia</b>				
(Intercept)	-9.1e-02	9.2e-02	-0.982	0.352
Año	5.6e-05	4.6e-05	1.222	0.253
Residual SE		0.0004		

Continúa en la siguiente página

Cuadro 10 – Continuación desde la página anterior

<b>Presupuesto asignado</b>				
<b>Variable</b>	<b>Coficiente</b>	<b>Error Estándar</b>	<b>Valor t</b>	<b>Valor p</b>
R-cuadrado ( $R^2$ )			0.1423	
$R^2$ ajustado			0.0469	
<b>Canadá</b>				
(Intercept)	3.3e-01	7.2e-02	4.619	0.0012 **
Año	-1.5e-04	3.6e-05	-4.383	0.0017 **
Residual SE		0.0003		
R-cuadrado ( $R^2$ )			0.681	
$R^2$ ajustado			0.6455	
<b>España</b>				
(Intercept)	9.3e-02	1.3e-01	0.688	0.509
Año	-4.0e-05	6.7e-05	-0.594	0.567
Residual SE		0.0007		
R-cuadrado ( $R^2$ )			0.0377	
$R^2$ ajustado			-0.0692	
<b>México</b>				
(Intercept)	4.1e-01	4.8e-02	8.722	1.1e-05 ***
Año	-2.0e-04	2.3e-05	-8.641	1.1e-05 ***
Residual SE		0.0002		

Continúa en la siguiente página

Cuadro 10 – Continuación desde la página anterior

Presupuesto asignado				
Variable	Coficiente	Error Estándar	Valor t	Valor p
R-cuadrado ( $R^2$ )			0.8924	
$R^2$ ajustado			0.8805	
Colombia				
(Intercept)	-2.2e-01	8.0e-02	-2.740	0.0229 *
Año	1.1e-04	3.9e-05	2.773	0.0216 *
Residual SE		0.0004		
R-cuadrado ( $R^2$ )			0.4608	
$R^2$ ajustado			0.4009	

### V.3.3. Análisis de correlación

En el Cuadro 11, se evidencian relaciones estadísticas entre los indicadores previamente identificados como afectados por la implementación de la estrategia de colaboración internacional, específicamente en el contexto del desarrollo científico en México. Estos indicadores proporcionan información valiosa sobre cómo se interrelacionan diferentes aspectos del entorno científico y pueden ayudar a orientar futuras estrategias en investigación y desarrollo. Estas correlaciones resaltan la complejidad de las dinámicas en el entorno científico mexicano y subrayan la importancia de un enfoque equilibrado para impulsar el desarrollo científico y tecnológico del país.

Se ha identificado que, en el contexto de estudio específico dentro de México, en las áreas de ciencias sociales, artes y humanidades, la colaboración internacional ha demostrado impactar positivamente dos factores clave: el liderazgo de México en la producción científica y la colaboración entre afiliaciones. Por otro lado, se encontró que existe una correlación significativa negativa entre las variables de productividad científica y presupuesto asignado al desarrollo científico de México en estas áreas.

Cuadro 11: Matriz de correlación de Spearman entre las variables de este estudio.

Matriz de correlación de Spearman							
	Liderazgo	Multi	ColabAfilia	ColabInter	Productividad	Presupuesto	Calidad
Liderazgo	1.0000	-0.1455	0.6073	0.7886	0.0279	-0.0694	0.0176
Multi	-0.1455	1.0000	-0.1554	-0.1157	-0.1753	0.0631	-0.0545
ColabAfilia	0.6073	-0.1554	1.0000	0.6586	0.0150	-0.0770	-0.0615
ColabInter	0.7886	-0.1157	0.6586	1.0000	0.0308	-0.0865	0.0462
Productividad	0.0279	-0.1753	0.0150	0.0308	1.0000	-0.6102	-0.0477
Presupuesto	-0.0694	0.0631	-0.0770	-0.0865	-0.6102	1.0000	0.0934
Calidad	0.0176	-0.0545	-0.0615	0.0462	-0.0477	0.0934	1.0000

**Nota:** *Multi* = Multidisciplinaria, *ColabAfilia* = Colaboración entre afiliaciones, *ColabInter* = Colaboración internacional.

La asociación positiva entre el liderazgo y la colaboración, ya sea internacional (0.7886) o entre afiliaciones (0.6073), se explica de manera detallada a través de un proceso de selección meticulosa de miembros, donde la proximidad, definida por trayectorias profesionales y colaboraciones previas, facilita la construcción de la red (Barrera, 2018). La selección y la posterior propuesta de colaboración a menudo provienen de participantes que obtienen beneficios mayores al trabajar colectivamente en comparación con el trabajo individual (Chung et al., 2009). En este contexto, considerando que los líderes del proyecto generalmente son quienes proponen la colaboración, resulta comprensible que México asuma un papel destacado como líder en colaboraciones internacionales, siendo esta una estrategia beneficiosa para el intercambio de conocimientos y el fortalecimiento de las redes científicas (Lancho-Barrantes & Cantú-Ortiz, 2019). Dada la limitación de recursos para la investigación científica en las CSA&H en México, surge la necesidad de involucrar a contribuyentes externos en los proyectos nacionales, posicionando a los investigadores e investigadoras mexicanas como líderes en la generación de conocimiento en áreas específicas de CSA&H. La relación positiva entre colaboración internacional y colaboración entre afiliaciones (0.6586) subraya la convergencia de estos dos factores bajo la influencia de la selección de miembros, fortaleciendo la colaboración y la responsabilidad en la red, incluso cuando las afiliaciones no están necesariamente ubicadas dentro de México.

Por otro lado, se destaca una correlación negativa significativa entre la productividad científica de México y el porcentaje del PIB asignado al desarrollo científico. Esta relación plantea interrogantes cruciales sobre la eficiencia y la inversión en la ciencia en México. En el estudio de Ortiz, 2015 (2015), se identificaron cambios significativos en la percepción de los programas de estímulos en el país. Anteriormente, iniciativas como el Sistema Nacional

de Investigadores (SNI) y el Programa para el Desarrollo Profesional Docente (PRODEP) podían ser simplemente considerados como reconocimientos al rendimiento eficiente de los investigadores. Sin embargo, en lugar de ser solo premios por eficiencia, estos programas ahora se perciben también como vías para obtener prestigio, beneficios económicos e incluso como indicadores de calidad y productividad institucional. Uno de los requisitos más destacados para ingresar al SNI es la experiencia y producción científica (Sandoval-Romero & Larivière, 2020). En consideración a esto, los investigadores pueden sentir una presión adicional para participar en estos programas, lo que tiene implicaciones significativas en la producción científica en las CSA&H a pesar del presupuesto limitado, con el objetivo de formar parte del SNI y aprovechar todos los beneficios asociados.

## V.4. Productos obtenidos durante el posgrado

Productos obtenidos durante el posgrado	
Primer semestre	
Producto	Resultado
Registro de protocolo ante la División de Investigación y Posgrado de la FI.	13364
Examen de Manejo de Lengua Inglesa	8-
Segundo semestre	
Producto	Resultado
Artículo “ <i>Estudio de desigualdad entre antecedentes migratorios: Un análisis de las expectativas educativas y el rendimiento académico en un contexto argentino</i> ”	Publicado en Transdigital el 22 de noviembre del 2023
Asistencia y gestión del XVI Coloquio de Posgrado de la Facultad de Ingeniería de la UAQ.	Constancia
Asistencia al I Congreso Nacional de Ciencias de la Educación por parte de la UPP.	Constancia
Asistencia al 1er Congreso Nacional de Productividad y Mejora Continua por parte del conaii.	Constancia
Artículo “ <i>Países aliados en la ciencia: ¿Qué es la colaboración internacional?</i> ”	Aceptado en Boletín Institucional de Prepa 1 - UAEH el 20 de octubre del 2023.
Miembro del Comité de Ética Aplicada a la Investigación de la Facultad de Ingeniería - UAQ	Constancia
Continuación en la siguiente página.	

<b>Productos obtenidos durante el posgrado</b>	
<b>Tercer semestre</b>	
<b>Producto</b>	<b>Resultado</b>
Miembro del Comité de Ética Aplicada a la Investigación de la Facultad de Ingeniería - UAQ	Constancia
Charla en Preparatoria Norte: “¿Ciencia? Una opción a profesión.”	Constancia
Recurrencia de hules caducos en el proceso de molero: Diseño de experimentos y análisis de datos	Reporte técnico y carta de uso
<b>Cuarto semestre</b>	
<b>Producto</b>	<b>Resultado</b>
Artículo “ <i>Colaboración internacional, multidisciplinariedad y liderazgo en las ciencias sociales, artes y humanidades en México (2010-2020)</i> ”	Aceptado en el XVII Coloquio de Posgrado de la FI - UAQ el 17 de noviembre del 2023
Examen de Comprensión de Textos en inglés	Aprobado
Participación como <i>Jefa de logística virtual</i> en el XVII Coloquio de la Facultad de Ingeniería de la UAQ.	Constancia

# Conclusión

En el marco de esta investigación, se ha realizado un análisis exhaustivo del panorama científico de México en Ciencias Sociales, Artes y Humanidades (CSA&H), siguiendo un enfoque metodológico en varias etapas.

Primero, se ha llevado a cabo un análisis comparativo que destaca el predominio del idioma inglés en la colaboración científica global, tanto en México como a nivel mundial, consolidándose como un vehículo predominante en la comunicación e investigación científica a nivel internacional. Además, se ha resaltado la diversidad y amplitud de la investigación en México, evidenciada por la variedad de fuentes, editores y temas de investigación. Este enfoque diversificado subraya la riqueza y la interconexión de la investigación en ambas perspectivas.

En lo que respecta a las afiliaciones de los investigadores, se ha observado que, aunque las instituciones estadounidenses lideran en ambas poblaciones, en México destaca la presencia de la Universidad Nacional Autónoma de México (UNAM) como una entidad líder en este ámbito. Estos hallazgos sugieren una oportunidad para fomentar la diversificación de las instituciones que contribuyen a las áreas de investigación estudiadas.

En términos de colaboración entre autores, se ha identificado que la colaboración en la investigación científica mexicana supera significativamente a la población global. Este hallazgo refleja una sólida cultura de colaboración arraigada en el país, como se evidencia en el mayor número promedio de autores por documento y un índice de colaboración más alto en la base de datos centrada en México en comparación con la global. Estos resultados corroboran la fortaleza de la capacidad científica mexicana y su activa participación en la investigación.

Posteriormente, el análisis exploratorio ha permitido profundizar en la calidad y la citación de los documentos científicos. Si bien se ha encontrado un alto promedio de citas por

documento a nivel global y en México, es relevante destacar que una fracción significativa de los registros carece de citas registradas. Esto genera preocupaciones legítimas acerca de si la presión por aumentar la producción científica podría comprometer la calidad. En consecuencia, se enfatiza la necesidad de promover la investigación de alta calidad y de evaluar el impacto real de la investigación más allá de las métricas de cantidad de publicaciones.

En cuanto a la colaboración internacional de CSA&H en México, los resultados indican una frecuencia significativamente mayor en comparación con lo global. Este hecho subraya la importancia de la cooperación internacional en estas áreas en el contexto mexicano y sugiere una mayor apertura a la colaboración en investigaciones multidisciplinarias y transnacionales. Además, la diversidad de socios de colaboración y la disposición a colaborar con dos o tres países distintos ponen de manifiesto una cultura de colaboración arraigada en la investigación científica mexicana en estas áreas. La identificación de los países con los que México colabora con mayor frecuencia (Estados Unidos, Canadá, Colombia y España) y la ausencia de colaboración con ciertos países (China, Corea del Sur, India, Israel, Rusia, Sudáfrica y Turquía), proporcionan información valiosa sobre las dinámicas de colaboración internacional con los países estratégicos definidos para México por CONAHCyT.

El crecimiento sostenido del liderazgo de México en el ámbito de las CSA&H es un descubrimiento de gran relevancia respaldado por una relación estadísticamente significativa entre el factor tiempo y la consolidación de su posición líder en este campo. Estos hallazgos apuntan a que México está desempeñando un rol cada vez más influyente en la esfera global de la investigación en CSA&H, con repercusiones significativas tanto a nivel nacional como internacional. La identificación de México como líder predominante entre los 28 países líderes en investigación en CSA&H, enfatiza la posición destacada de la nación en este dominio y su potencial impacto en la comunidad científica mundial. Estos logros deberían servir de catalizador para continuar invirtiendo en investigación científica y fomentar la colaboración tanto a nivel nacional como internacional. Es evidente que México ha implementado una estrategia de colaboración internacional alineada con las recomendaciones de la OECD, buscando diversificar su conocimiento y, en consecuencia, mejorar la calidad y el desarrollo científico del país.

El análisis correlacional revela relaciones significativas entre las variables estudiadas. La conexión positiva entre liderazgo y colaboración, tanto internacional como entre afiliaciones, se basa en una selección cuidadosa de miembros, favoreciendo la construcción de redes. México destaca como líder en colaboraciones internacionales, aprovechando esta estrategia para el intercambio de conocimientos. La limitación de recursos impulsa la inclusión

de contribuyentes externos, consolidando a los investigadores mexicanos como líderes. La relación positiva entre colaboración internacional y afiliaciones subraya la convergencia de estos factores, fortaleciendo la colaboración y responsabilidad en la red. En conjunto, destaca la importancia de estrategias colaborativas y liderazgo en la investigación científica, permitiendo a México destacar internacionalmente. Por otro lado, la correlación negativa entre la productividad científica de México y el porcentaje del PIB asignado al desarrollo científico plantea cuestionamientos sobre la eficiencia de la inversión en ciencia. En estudios anteriores (Barrera, 2018), se destacan cambios en la percepción de programas de estímulos, ahora vistos como medios para obtener prestigio e indicadores de calidad. Requisitos como experiencia y producción científica para el SNI generan presiones adicionales, con implicaciones para la producción científica en CSA&H y un presupuesto limitado. Estas dinámicas subrayan la complejidad de la relación entre inversión, reconocimiento institucional y producción científica en el entorno científico mexicano.

Este estudio ha arrojado luz sobre diversas facetas del panorama científico en México en el ámbito de las CSA&H, proporcionando una base sólida para orientar futuras estrategias de investigación y desarrollo en el país. Estos hallazgos respaldan la importancia de mantener un equilibrio entre la cantidad y la calidad en la producción científica, así como promover la colaboración tanto a nivel nacional como internacional. Sin embargo, es fundamental considerar la estrecha interconexión entre los hallazgos presentados y las políticas educativas del país. La colaboración internacional y la cooperación en investigaciones multidisciplinarias y transnacionales son aspectos que pueden ser directamente relevantes para la formulación de políticas educativas que fomenten la movilidad académica y la adquisición de conocimientos diversificados. En conjunto, este análisis pone de manifiesto la estrecha relación entre la investigación científica, la inversión en ciencia y tecnología, y las políticas educativas del país, destacando la necesidad de una estrategia integral que promueva la excelencia científica, la colaboración internacional y el progreso educativo en México. No obstante, se deben considerar las limitaciones, como la descarga incompleta de datos y el poder explicativo limitado de los modelos de regresión, lo que requiere precaución en la interpretación de los resultados y resalta la necesidad de investigaciones adicionales para comprender mejor las dinámicas de investigación en México en este ámbito.

## **VI.1. Recomendaciones e investigaciones futuras**

Durante el desarrollo de este proyecto, se ha observado una marcada tendencia hacia la producción de contenido relacionado con el tratamiento y la comprensión de enfermedades mentales en el período comprendido entre 2010 y 2020, en contraste con la atención prestada a las disciplinas de CSA&H. Esta tendencia plantea cuestionamientos significativos y justifica la necesidad de continuar investigando y analizando sus posibles causas.

Es importante destacar que se ha identificado un aumento sustancial en la visibilidad de los documentos científicos mexicanos durante el año 2014. Sería provechoso llevar a cabo un análisis detallado de los metadatos exclusivamente para este período con el fin de identificar las condiciones que llevaron a una mayor cantidad de citas por documento científico. Asimismo, es necesario llevar a cabo un estudio que analice el aumento en la producción científica de México durante el año 2018, ya que este fenómeno podría proporcionar información valiosa para comprender mejor las dinámicas de investigación en el país.

Además de estos aspectos, resulta fundamental investigar y comprender las variables que influyen de manera significativa en la producción y calidad de la ciencia en el ámbito de las CSA&H a través de la colaboración internacional. Esta investigación es esencial porque permite identificar los factores clave que impulsan la excelencia científica y la innovación en un contexto globalizado. Discutir estas variables ayudará a comprender la importancia de este análisis y cómo puede contribuir al fortalecimiento de la investigación y la colaboración científica de las CSA&H en México a nivel internacional.

## **VI.2. Limitaciones**

Solo fue posible obtener el 96.00% de los registros de la base de datos filtrada, equivalentes a 100,000 registros de los 104,166 disponibles. Esto se debió a las restricciones de membresía que limitaron la cantidad de registros descargables. Asimismo, la descarga de información se vio restringida por el tipo de membresía, lo que obligó a realizar descargas en intervalos de 1,000 registros.

Es relevante mencionar que el poder explicativo de los modelos de regresión es bajo, una característica común en ciencias sociales. Las colaboraciones individuales en investigación se ven influenciadas por diversas condiciones, como normas disciplinarias, estrategias de grupos de investigación, políticas institucionales y financiamiento. Por lo tanto, existen numerosos factores que afectan la intensidad de la colaboración en investigación en grupos

y redes. Además, las diferencias en la productividad y calidad de la investigación están influenciadas por una multitud de factores que van más allá de los considerados en este estudio. Estas dinámicas han sido discutidas previamente por Kyvik (2017), requiriendo futuras investigaciones.



## Apéndice

Cuadro 13: Variables disponibles en Colección Principal de WoS (Web of Science Group, 2019).

<b>Variables disponibles en Colección Principal de WoS</b>			
<b>Código</b>	<b>Variable</b>	<b>Código</b>	<b>Variable</b>
PT	Tipo de publicación	AU	Authors
BA	Autores de libro	BE	Editores de libro
GP	Grupos de autores de libro	AF	Nombres completos de autores
BF	Nombres completos de autores de libro	CA	Grupos de autores
TI	Título del artículo	SO	Título de la fuente
SE	Título de la serie del libro	BS	Subtítulo de la serie del libro
LA	Idioma	DT	Tipo de documento
Continuación en la siguiente página.			

<b>Código</b>	<b>Variable</b>	<b>Código</b>	<b>Variable</b>
CT	Título de la conferencia	CY	Fecha de la conferencia
CL	Ubicación de la conferencia	SP	Patrocinador de la conferencia
HO	Anfitrión de la conferencia	DE	Palabras clave del autor
ID	Palabras clave plus	AB	Resumen
C1	Direcciones	C3	Afiliaciones
RP	Dirección de reimpresión	EM	Correo electrónico
RI	ID's de investigadores	OI	ORCID's
FU	Organizaciones de financiación	FP	Nombre preferido de la organización de financiación
FX	Texto de financiación	CR	Referencias citadas
NR	Cuenta de referencias citadas	TC	Cuenta de citas en WoS Core
Z9	Cuenta de citas en todas las bases de datos	U1	Cuenta de uso en 180 días previos
U2	Cuenta de uso desde 2013	PU	Editor
PI	Ciudad de editor	PA	Dirección de editor
SN	ISSN	EI	eISSN
Continuación en la siguiente página.			

<b>Código</b>	<b>Variable</b>	<b>Código</b>	<b>Variable</b>
BN	ISBN	J9	Abreviación de la revista
JI	Abreviación ISO de la revista	PD	Fecha de publicación
PY	Año de publicación	VL	Volumen
IS	Issue	PN	Número de parte
SU	Suplemento	SI	Issue especial
MA	Resumen de reunión	BP	Página de inicio
EP	Página de fin	AR	Número de artículo
DI	DOI	DL	DOI link
D2	DOI de libro	EA	Fecha de acceso temprano
PG	Número de páginas	WC	Categorías de WoS
WE	Índice de WoS	SC	Áreas de investigación
GA	Número de IDS	PM	Pubmed ID
OA	Designaciones Open Access	HC	Estado de alta citación
HP	Estatus de tendencia	DA	Fecha de exportación
UT	ID único de WoS	-	-

Código A.1: Query utilizado para filtrar la base de datos *Clarivate - Web of Science*.

```
PY=(2010-2020))
AND (TASCA==( "ANTHROPOLOGY" OR "COMMUNICATION"
OR "ECONOMICS" OR "EDUCATION SCIENTIFIC
DISCIPLINES" OR "EDUCATION EDUCATIONAL
RESEARCH" OR "ENVIRONMENTAL STUDIES" OR
"GEOGRAPHY" OR "DEVELOPMENT STUDIES" OR
"LAW" OR "INFORMATION SCIENCE LIBRARY
SCIENCE" OR "MANAGEMENT" OR "POLITICAL
SCIENCE" OR "PUBLIC ADMINISTRATION" OR
"PSYCHIATRY" OR "PSYCHOLOGY" OR "PUBLIC
ENVIRONMENTAL OCCUPATIONAL HEALTH" OR
"HEALTH CARE SCIENCES SERVICES" OR
"REHABILITATION" OR "SOCIAL SCIENCES
INTERDISCIPLINARY" OR "SOCIOLOGY" OR
"URBAN STUDIES" OR "ARCHAEOLOGY" OR
"ART" OR "ARCHITECTURE" OR "HISTORY"
OR "LINGUISTICS" OR "LANGUAGE
LINGUISTICS" OR "RELIGION"))
AND CU==( "ARGENTINA" OR "BRAZIL" OR "CANADA" OR
"CHILE" OR "COLOMBIA" OR "PEOPLES R CHINA"
OR "FRANCE" OR "GERMANY" OR "INDIA" OR
"ISRAEL" OR "JAPAN" OR "MEXICO" OR "RUSSIA"
OR "SOUTH AFRICA" OR "SOUTH KOREA" OR "USA"
"SPAIN" OR "TURKEY" OR "ENGLAND" OR
"SCOTLAND" OR "WALES")
AND EDN==( "WOS.SSCI" OR "WOS.AHCI" OR "WOS.SCI")
```

Código A.2: Importación de base de datos a R y conversión de los tipos de datos.

```
#Conversion de tipo de datos
for (i in 1:100) {
  for (col_nombre in nombres_columnas) {
    if (col_nombre %in% c("NR", "TC", "Z9", "U1",
      "U2", "PY", "VL", "SU", "BP", "EP", "PG")) {
      #Conversion a integer
      lista_archivos[[i]][[col_nombre]] <-
      as.integer(lista_archivos[[i]][[col_nombre]])
    } else { #Si no, se hace conversion a character
      lista_archivos[[i]][[col_nombre]] <-
      as.character(lista_archivos[[i]][col_nombre])
    }
  }
}
```

Código A.3: Selección de variables de interés.

```
base_datos_filtrada <- base_datos %>%
  select = ("PT", #Tipo de publicacion
    "AU", #Autores
    "TI", #Titulo
    "LA", #Idioma
    "DE", #Palabras clave
    "C1", #Direcciones
    "C3", #Afiliaiones
    "Z9", #Veces citado
    "PU", #Editor
    "JI", #Fuente
    "PY", #Anio de publicacion
    "WC", #Categorias WoS
    "WE", #Indice WoS
    "SC" #Areas de investigacion))
```

Código A.4: Limpieza de variables en la base de datos.

```
# Limpieza 1: PT
for (i in 1:nrow(base_datos)) {
  if (length(base_datos$PT[[i]]) == 0){
    base_datos$PT[[i]] <- "Sin_registro" }}

# Limpieza 2: P_Colab
P_Colab <- vector("list", nrow(base_datos))
for (i in 1:nrow(base_datos)) {
  P_Colab[[i]] <- unique(base_datos$P_Colab[[i]])}

# Limpieza 3: PY
base_datos <- base_datos %>%
  filter(base_datos$PY >= 2010 &
         base_datos$PY <= 2020)

# Limpieza 4: Presencia de Mexico
MxPresente <- logical(length(base_datos$P_Colab))
for (i in seq_along(base_datos$P_Colab)) {
  if ("Mexico" %in% base_datos$P_Colab[[i]]) {
    MxPresente[i] <- TRUE
  } else {
    MxPresente[i] <- FALSE }}
```

Código A.5: Creación de variables en la base de datos.

```
# Creacion 1: nAU
nAU <- sapply(base_datos$AU, function(x)
  length(unlist(strsplit(x, ";"))))

# Creacion 2: rnAU
rangos <- c(0, 1, 2, 5, 8, 11, Inf)
rnAU <- cut(base_datos$nAU, breaks = rangos,
            labels = c("0", "1", "2-4", "5-7",
                      "8-10", "11-o-mas"))
```

```

# Creacion 3: P_Colab
P_Colab <- vector("list", nrow(base_datos))
patron <- "\\b[A-Za-z]+(?:=;|_\\|)\\b"
for (i in 1:nrow(base_datos)) {
  if (base_datos$C1[[i]] == "") {
    P_Colab[[i]] <- "Sin_registro"
  } else {
    paises <- str_extract_all(base_datos$C1[i],
    patron)
    ult_p <- tail(strsplit(base_datos$C1[i],
    "\\s+|_" )[[1]], 1)
    if (length(paises[[1]]) > 0) {
      P_Colab[[i]] <- unlist(paises)
    }
    P_Colab[i] <- paste0(c(P_Colab[i], ult_p)) }}

# Creacion 4: P_Lider
P_Lider <- vector("list", nrow(base_datos))
for (i in 1:nrow(base_datos)) {
  P_Lider[i] <- base_datos$P_Colab[[i]][1] }

# Creacion 5: PosMex
PosM <- vector("numeric", nrow(base_datos))
for (i in 1:nrow(base_datos)) {
  if ("Mexico" %in% base_datos$P_Colab[[i]]) {
    PosM[i] <- min(which(base_datos$P_Colab[i] ==
    "Mexico"))
  } else {PosM[i] <- 0 # Asignar 0 }}

```

Código A.6: Modificación a vector de variables.

```

# Modificacion a vector
AU_vect <- vector("list", nrow(base_datos))
patron <- "[^;]+"
for (i in 1:nrow(base_datos)) {
  autores <- str_extract_all(base_datos$AU[i])

```

```

ult_aut <- tail(strsplit(base_datos$AU[i],
                        "\\s*;" )[[1]], 1)
if (length(autores[[1]]) > 0) {
  AU_vect[[i]] <- unlist(autores)
}else {
  AU_vect[[i]] <- "Sin_registro" }
AU_vect[[i]] <- unique(c(AU_vect[[i]], ult_aut)) }

```

Código A.7: Contar y filtrar.

```

# Ejemplo P_Lider.
# Contar la presencia de una categoria.
PaisLider <- base_datos %>%
  count(P_Lider, name = "Count") %>%
  mutate(Percent = Count / sum(Count) * 100)

# Ordenar y filtrar informacion.
PaisLider <- PaisLider %>%
  arrange(desc(Percent)) %>%
  top_n(25)

```

Código A.8: Orden de variables categóricas.

```

# Ejemplo rnAU.
# Orden descendente: Frecuencia.
sort(table(base_datos$rnAU), decreasing = TRUE)
# Orden descendente: Porcentaje.
sort(prop.table(table(base_datos$rnAU)),
      decreasing = TRUE)

```

Código A.9: Análisis de la productividad de México en las CSA&H.

```

# Analisis anual de Mexico
produc_anual$Porcentaje <- (produc_anual$Art_Mex /
  produc_anual$Art_Global) * 100
mean(produc_anual$Porcentaje)
modelo <- lm(Porcentaje ~ Anio, data = produc_anual)

```

```

pendiente <- coef(modelo)[2]

# Identificacion de area mas productiva
tabla_areas_inv <- tabla_areas_inv %>%
  filter(area_invest %in% areas_interes)
tabla_areas_inv$Porcentaje <-
  (tabla_areas_inv$frecuencia /
   sum(tabla_areas_inv$frecuencia)) * 100

# Analisis por area de investigacion
base_datos_area_desag <- base_datos_filtrada %>%
  unnest(SC) %>%
  mutate(SC_limpia = trimws(SC)) %>%
  count(SC_limpia, PY, name = "Frecuencia",
        .drop = FALSE) %>%
  filter(SC_limpia %in% areas_interes) %>%
  arrange(PY, SC_limpia)

# Identificacion de afiliacion mas productiva
tabla_afiliaciones <- data.frame(afiliacion =
  names(frecuencias), frecuencia =
  as.numeric(frecuencias)) %>%
  arrange(-frecuencia)

```

Código A.10: Análisis de la Visibilidad de la ciencia en México en las CSA&H.

```

# Analisis anual de Mexico
citacion_anual <- citacion_anual %>%
  mutate(citas_promedio_por_año =
         PromedioCitasxAño / AñosCitables)

modelo1 <- lm(Crecimiento ~ Año, data =
  subset(citacion_anual, Region == "MEXICO"))
pendienteMx <- coef(modelo1)[2]

```

```

# Identificacion de area con mayor calidad
area_investigacion_Pareto <-
  area_investigacion_extraido %>%
  select(SC, citas_normalizada) %>%
  mutate(SC = trimws(SC)) %>%
  group_by(SC) %>%
  summarize(total_citas =
    sum(citas_normalizada))

# Analisis por area de investigacion
area_investigacion_extraido <-
  area_investigacion_extraido %>%
  group_by(PY, SC) %>%
  summarise(citas_normalizada =
    sum(citas_normalizada)) %>%
  filter(SC %in% areas_interes)

# Analisis por afiliacion
base_datos_afiliacion_desagregada <-
  base_datos_filtrada %>%
  select(Z9, Afiliac, nAfiliac, PY, SC) %>%
  filter(SC %in% areas_interes) %>%
  filter(Afiliac != "Sin_registro") %>%
  unnest(cols = c(Afiliac))

```

Código A.11: Análisis de la colaboración internacional de la ciencia mexicana en CSA&H.

```

# Porcentaje de colaboracion internacional
colab_internacional <- base_datos_filtrada %>%
  summarise(Conteo = sum(nPaises > 1),
    Porcentaje = (sum(nPaises > 1) / n()) * 100)

# Indice de colaboracion internacional
indice_colab_inter <- base_datos_filtrada %>%
  group_by(PY) %>%

```

```

summarise(Indice_colaboracion =
  mean(nPaises)) %>%
ungroup()

modelo <- lm(Indice_colaboracion ~ PY,
  data = indice_colab_inter)
pendiente <- coef(modelo)[2]

# Analisis por area de investigacion
areas_en_colaboracion <- base_datos_filtrada %>%
  select(SC, nPaises, PY) %>%
  unnest(cols = SC) %>%
  filter(SC %in% areas_interes) %>%
  filter(nPaises>1) %>%
  count(SC, name = "Frec")

```

Código A.12: Análisis de la multidisciplinariedad de la ciencia mexicana en CSA&H.

```

# Estudio de tendencias en el periodo
palabras_analisis <- base_datos_filtrada %>%
  select(PY, TI, SC) %>%
  filter(SC %in% areas_interes) %>%
  unnest(cols = c(TI)) %>%
  group_by(PY, TI)

# Indice de multidisciplinariedad
multidisci <- base_datos_filtrada %>%
  select(PY, nSC) %>%
  arrange(PY) %>%
  group_by(PY) %>%
  summarize(Promedio = mean(nSC))

modelo <- lm(Promedio ~ PY, data = multidisci)
pendiente <- coef(modelo)[2]

```

Código A.13: Análisis del liderazgo de México en CSA&H.

```
# Calculo de proporcion primer autor
lider <- base_datos_filtrada %>%
  summarise(Conteo = sum(PosMex == 1),
    Porcentaje = (sum(PosMex == 1) / n()) * 100)

# Calculo de proporcion autor de correspondencia
ultimo <- base_datos_filtrada %>%
  summarise(Conteo = sum(PosMex == nPaises),
    Porcentaje = (sum(PosMex == nPaises &
      PosMex != 1) / n()) * 100)

# Indice de liderazgo
indice_liderazgo <- base_datos_filtrada %>%
  select(PY, PosMex) %>%
  group_by(PY) %>%
  summarise(liderazgo = mean(PosMex, na.rm=TRUE))

modelo <- lm(liderazgo ~ PY,
  data = indice_liderazgo)
pendiente <- coef(modelo)[2]
```

Código A.14: Análisis del presupuesto asignado al desarrollo de la ciencia y tecnología. México, Colombia, España, Canadá, Francia, Estados Unidos e Israel.

```
# Analisis del porcentaje de PIB de Mexico y algunos
# de sus paises estrategicos
modelo <- lm(Mexico ~ Anio, data = presupuesto)
pendiente <- coef(modelo)[2]
```

Código A.15: Análisis correlacional de la calidad, productividad, multidisciplinariedad, liderazgo y presupuesto en México dentro de las CSA&H.

```
# Matriz de correlacion
correlation_matrix <- cor(base_box)
correlation_matrix
```

Código A.16: Gráfico de Línea.

```
#Visualizacion grafico de linea. Ejemplo productiv.
ggplot(produccion_anual)+
  geom_line(mapping = aes(x=Anio, y=Porcentaje),
    color = colores[1], size = 1)+
  scale_x_continuous(breaks = c(2010:2020))+
  scale_y_continuous(breaks = seq(0,1.20, 0.1))
```

Código A.17: Gráfico de Pastel.

```
#Visualizacion grafico de pastel. Ejemplo rnAU.
ggplot(datos, aes(x = "", y = Proporciones*100,
  fill = Rangos)) +
  geom_bar(stat = "identity", color = "white") +
  coord_polar(theta = "y")
```

Código A.18: Gráfico de Barras Horizontales.

```
# Visualizacion grafico de barras horizontales
ggplot(areas_en_colaboracion)+
  geom_bar(mapping = aes(x = Frecuencia, y = SC),
  stat = "identity", fill = colores[4])
```

Código A.19: Gráfico de TreeMap.

```
#Visualizacion treemap. Ejemplo PT.
ggplot(data = DocType, aes(area = Porcentaje,
  label = TipoDoc, fill = Porcentaje)) +
  geom_treemap() +
  geom_treemap_text(colour = "white", size = "15",
  place = "centre", reflow = T)
```

Código A.20: Gráfico de Burbujas.

```
#Visualizacion grafico de burbuja. Ejemplo AU.
ggplot(datos, aes(x=year, y = Author,
  size = Frecuencia)) +
  geom_point(alpha=transparencia(datos$Frecuencia)) +
```

```
scale_size(range = c(2,9)) +  
scale_alpha(guide='none')
```

#### Código A.21: Gráfico Polar.

```
#Visualizacion grafico polar. Ejemplo rnAU.  
ggplot(datos, aes(x = Proporciones*100, y = Rangos,  
  fill = Rangos)) +  
  geom_bar(stat = "identity", color = "white",  
    lwd = 1, show.legend = FALSE) +  
  coord_polar(theta = "y")
```

#### Código A.22: Gráfico de Nube de Palabras.

```
# Visualizacion nube de palabras.  
ggplot(palabras_2010_2020)+  
  geom_text_wordcloud(mapping = aes(label = TI,  
    size = Frecuencia_2, color = Frecuencia)) +  
  scale_size_continuous(range = c(3.9, 20),  
    breaks = c(3, 7, 10))+  
  scale_color_gradient(low = colores[3],  
    high = colores[7])
```

#### Código A.23: Gráfico de Mapa de Coropletas.

```
# Visualizacion mapa de coropletas  
ggplot() +  
  geom_map(data = merged_data, map = map_data,  
    aes(x = long, y = lat, map_id = region,  
      fill = Ranke), color = "white") +  
  theme(plot.title = element_text(size = 20,  
    color = "#4D4545"),  
    plot.subtitle = element_text(size = 12,  
    color = "#4D4545"),  
    axis.title = element_blank(),  
    axis.text = element_blank(),  
    axis.ticks = element_blank())
```

Código A.24: Diagrama de Cajas.

```
# Visualizacion diagrama de cajas
ggplot(base_box, aes(x = Variable, y = Valor)) +
  geom_boxplot() +
  theme_minimal()+
  facet_wrap(vars(Variable), scales = "free_y",
    ncol = 2, strip.position = "bottom")
```

Código A.25: Diagrama de Pareto.

```
# Visualizacion diagrama de Pareto
ggplot(tabla_afiliaciones) +
  geom_col(mapping = aes(x = afiliacion,
    y = frecuencia, alpha = transparencia),
    fill = colores[5]) +
  geom_point(mapping = aes(x = afiliacion,
    y = percent_transform(PorcenAcumu), group = 1),
    color = colores[3], size = 2)+
  geom_line(mapping = aes(x = afiliacion,
    y = percent_transform(PorcenAcumu), group = 1),
    color = colores[3], size = 1) +
  scale_alpha(range = c(0.5, 1))
```

Código A.26: Diagrama de Dispersion.

```
# Visualizacion diagrama de dispersion
grafico_combinado <- (
ggplot(base_datos_filtrada) +
  geom_point(aes(x = Presupuesto, y = nAfiliac)) +
  labs(title = "Presupuesto_□-□Colab_□Afiliacion") +
  theme_minimal()
+
...)
```



---

## Referencias

- Aldieri, L., Kotsemir, M., & Vinci, C. P. (2018). The impact of research collaboration on academic performance: An empirical analysis for some European countries. *Socio-Economic Planning Sciences*, 62, 13-30. <https://doi.org/10.1016/j.seps.2017.05.003>
- Aria, M., & Cuccurullo, C. (2017). Bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, 11(4), 959-975. <https://doi.org/10.1016/j.joi.2017.08.007>
- Ball, R. (2017, 18 de septiembre). *An Introduction to Bibliometrics: New Development and Trends*. Chandos Publishing.
- BancoMundial. (2023, 1 de enero). *World bank open data*. World Bank Open Data. Consultado el 10 de agosto de 2023, desde <https://data.worldbank.org>
- Barahona, A., Chamizo, J. A., Garritz, A., & Slisko, J. (2014). The History and Philosophy of Science and Their Relationship to the Teaching of Sciences in Mexico. En M. R. Matthews (Ed.), *International Handbook of Research in History, Philosophy and Science Teaching* (pp. 2247-2269). Springer Netherlands. [https://doi.org/10.1007/978-94-007-7654-8\\_69](https://doi.org/10.1007/978-94-007-7654-8_69)
- Barrera, H. A. (2018). TEORÍA FUNDAMENTADA SOBRE REDES DE INVESTIGACIÓN EN EL CENTRO DE MÉXICO. Consultado el 13 de enero de 2024, desde <https://ri-ng.uaq.mx/handle/123456789/1026>  
Accepted: 2019-01-29T01:18:47Z.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using **lme4**. *Journal of Statistical Software*, 67(1). <https://doi.org/10.18637/jss.v067.i01>
- Benneworth, P., & Jongbloed, B. W. (2010). Who matters to universities? A stakeholder perspective on humanities, arts and social sciences valorisation. *Higher Education*, 59(5), 567-588. <https://doi.org/10.1007/s10734-009-9265-2>

- Bornmann, L. (2017). Is collaboration among scientists related to the citation impact of papers because their quality increases with collaboration? An analysis based on data from F1000Prime and normalized citation scores. *Journal of the Association for Information Science and Technology*, 68(4), 1036-1047. <https://doi.org/10.1002/asi.23728>
- Bornmann, L., Tekles, A., & Leydesdorff, L. (2019). How well does I3 perform for impact measurement compared to other bibliometric indicators? The convergent validity of several (field-normalized) indicators. *Scientometrics*, 119(2), 1187-1205. <https://doi.org/10.1007/s11192-019-03071-6>
- Bornmann, L., & Williams, R. (2020). An evaluation of percentile measures of citation impact, and a proposal for making them better. *Scientometrics*, 124(2), 1457-1478. <https://doi.org/10.1007/s11192-020-03512-7>
- Castañón-Lomnitz, H. (2006). Social Sciences and Science Policies in Mexico. *Science and Public Policy*, 33(2), 115-123. <https://doi.org/10.3152/147154306781779127>
- Castillo, J. A., & Powell, M. A. (2020). Research Productivity and International Collaboration: A Study of Ecuadorian Science. *Journal of Hispanic Higher Education*, 19(4), 369-387. <https://doi.org/10.1177/1538192718792151>
- Chinchilla-Rodríguez, Z., Larivière, V., Costas Comesaña, R., Robinson-García, N., & Sugimoto, C. R. (2018). Building ties across countries: International collaboration, field specialization, and global leadership. <https://doi.org/10.13039/501100010198>  
Accepted: 2018-09-23T20:22:36Z.
- Chinchilla-Rodríguez, Z., Sugimoto, C. R., & Larivière, V. (2019). Follow the leader: On the relationship between leadership and scholarly impact in international collaborations. *PLOS ONE*, 14(6), e0218309. <https://doi.org/10.1371/journal.pone.0218309>
- Chung, K. H., Cox, R. A. K., & Kim, K. A. (2009). On the relation between intellectual collaboration and intellectual output: Evidence from the finance academe. *The Quarterly Review of Economics and Finance*, 49(3), 893-916. <https://doi.org/10.1016/j.qref.2008.08.001>
- Coccia, M., & Wang, L. (2015, 16 de marzo). *Evolutionary Convergence of the Patterns of International Research Collaborations Across Scientific Fields*. <https://doi.org/10.2139/ssrn.2579452>
- CONAHCYT. (2021, noviembre). *Informe General Del Estado de La Ciencia, La Tecnología y La Innovación: México 2020*. Consejo Nacional de Humanidades, Ciencia y Tecnología, Conahcyt. <https://www.siicyt.gob.mx/index.php/transparencia/informes->

conacyt/informe-general-del-estado-de-la-ciencia-tecnologia-e-innovacion/informe-general-2020-1/4987-informe-general-2020/file

- Devore, J. L. (2016). *Probability and statistics for engineering and the sciences* (7ma ed). Thomson-Brooks/Cole. <https://latam.cengage.com/libros/probabilidad-y-estadistica-para-ingenieria-y-ciencias/@article{devore20121, title={1 Probabilidad 1 y estadística}, author={Devore, Jay L}, year={2012} }>.
- Diez, D. M., Barr, C. D., & Cetinkaya-Rundel, M. (2019). *OpenIntro Statistics* (4.ª ed.). Consultado el 8 de marzo de 2022, desde <https://www.openintro.org/book/os/>
- Dumitru, D. (2019). Creating Meaning. The Importance of Arts, Humanities and Culture for Critical Thinking Development. *Studies in Higher Education*, 44(5), 870-879. <https://doi.org/10.1080/03075079.2019.1586345>
- Ebadi, A., & Schiffauerova, A. (2013). Impact of Funding on Scientific Output and Collaboration: A Survey of Literature. *Journal of Information & Knowledge Management*, 12(04), 1350037. <https://doi.org/10.1142/S0219649213500378>
- Ebadi, A., & Schiffauerova, A. (2016). How to boost scientific production? A statistical analysis of research funding and other influencing factors. *Scientometrics*, 106(3), 1093-1116. <https://doi.org/10.1007/s11192-015-1825-x>
- Efrain-García, P., & García-Zorita, C. (2016). “Desktop Scientometrics”: una metodología para el uso de datos procedentes de WoS mediante el programa estadístico R. *Revista Española de Documentación Científica*, 39(1), e122-e122. <https://doi.org/10.3989/redc.2016.1.1223>
- Fanelli, R., Lombardi, M., Michelassi, S., Zizzo, C., & Duro, G. (2017). The importance of multidisciplinary. *Giornale di Tecniche Nefrologiche e Dialitiche*, 29, S32-S34. <https://doi.org/10.5301/GTND.2017.17482>
- Fernandez-Casal, R., Lafuente-Rego, B., Lombardia, M.-J., Tarrío-Saavedra, J., & Costa-Bouzas, J. (2021). *Análisis bibliométrico con R*. Consultado el 19 de abril de 2023, desde [https://rubenfcasal.github.io/scimetr/articles/docs/R\\_packages.html](https://rubenfcasal.github.io/scimetr/articles/docs/R_packages.html)
- Francisco, J. S. (2015). International Scientific Collaborations: A Key to Scientific Success. *Angewandte Chemie International Edition*, 54(50), 14984-14985. <https://doi.org/10.1002/anie.201505267>
- Guo, S., Zhang, G., & Guo, Y. (2016). Social Network Analysis of 50 Years of International Collaboration in the Research of Educational Technology. *Journal of Educational Computing Research*, 53(4), 499-518. <https://doi.org/10.1177/0735633115611114>

- Kara, H. (2015, 10 de abril). *Creative Research Methods in the Social Sciences: A Practical Guide*. Policy Press.
- Katz, J. S., & Martin, B. R. (1997). What is research collaboration? *Research Policy*, 26(1), 1-18. [https://doi.org/10.1016/S0048-7333\(96\)00917-1](https://doi.org/10.1016/S0048-7333(96)00917-1)
- Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2004). *Applied Linear Statistical Models*. Consultado el 16 de mayo de 2023, desde <http://83.136.219.140:8080/handle/123456789/554>  
Accepted: 2023-04-18T12:21:34Z.
- Kyvik, S., & Reymert, I. (2017). Research collaboration in groups and networks: Differences across academic fields. *Scientometrics*, 113(2), 951-967. <https://doi.org/10.1007/s11192-017-2497-5>
- Lámbarry-Vilchis, F., & Moreno-Jiménez, J. C. (2020). Co-authorship network of national researchers of Social Sciences in Mexico //Red de coautoría de investigadores nacionales en Ciencias Sociales en México - Dialnet. 11(31), 8-25. Consultado el 22 de agosto de 2023, desde <https://dialnet.unirioja.es/servlet/articulo?codigo=8809930>
- Lancho-Barrantes, B. S., & Cantú-Ortiz, F. J. (2019). Science in Mexico: A bibliometric analysis. *Scientometrics*, 118(2), 499-517. <https://doi.org/10.1007/s11192-018-2985-2>
- Leite, D., & Pinho, I. (2017). Science Geography and International Research Collaboration. En D. Leite & I. Pinho (Eds.), *Evaluating Collaboration Networks in Higher Education Research: Drivers of Excellence* (pp. 1-9). Springer International Publishing. [https://doi.org/10.1007/978-3-319-45225-8\\_1](https://doi.org/10.1007/978-3-319-45225-8_1)
- Leydesdorff, L., Bornmann, L., & Adams, J. (2019). The integrated impact indicator revisited (I3\*): A non-parametric alternative to the journal impact factor. *Scientometrics*, 119(3), 1669-1694. <https://doi.org/10.1007/s11192-019-03099-8>
- Melin, G., & Persson, O. (1996). Studying research collaboration using co-authorships. *Scientometrics*, 36(3), 363-377. <https://doi.org/10.1007/BF02129600>
- Monge-Nájera, J., & Ho, Y.-S. (2018). Guatemala articles in the Science Citation Index Expanded: bibliometry of subjects, collaboration, institutions and authors. *Revista de Biología Tropical*, 66(1), 312-320. <https://doi.org/10.15517/rbt.v66i1.29875>
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to linear regression analysis 5th ed.* John Wiley.

- Moya-Anegón, F. D., & Herrero-Solana, V. (2006). Science in america latina: A comparison of bibliometric and scientific-technical indicators. *Scientometrics*, 46(2), 299-320. <https://doi.org/10.1007/bf02464780>
- OECD. (2015). *OECD Science, Technology and Industry Outlook 2014* (Text). Consultado el 7 de febrero de 2023, desde [https://www.oecd-ilibrary.org/science-and-technology/oecd-science-technology-and-industry-outlook\\_19991428](https://www.oecd-ilibrary.org/science-and-technology/oecd-science-technology-and-industry-outlook_19991428) [https://read.oecd-ilibrary.org/science-and-technology/oecd-science-technology-and-industry-outlook-2014/mexico\\_sti\\_outlook-2014-65-en#page1](https://read.oecd-ilibrary.org/science-and-technology/oecd-science-technology-and-industry-outlook-2014/mexico_sti_outlook-2014-65-en#page1).
- Ortiz, C. U. (2015). El ethos científico en tres facultades de la UAQ. Un estudio a partir de los programas de estímulos. Consultado el 13 de enero de 2024, desde <https://ring.uaq.mx/handle/123456789/742>  
Accepted: 2018-12-14T15:32:27Z.
- Prathap, G. (2013). Second order indicators for evaluating international scientific collaboration. *Scientometrics*, 95(2), 563-570. <https://doi.org/10.1007/s11192-012-0804-8>
- Rousseau, R., Egghe, L., & Guns, R. (2018a, 1 de enero). Chapter 1 - Introduction. En R. Rousseau, L. Egghe & R. Guns (Eds.), *Becoming Metric-Wise* (pp. 1-10). Chandos Publishing. <https://doi.org/10.1016/B978-0-08-102474-4.00001-7>
- Rousseau, R., Egghe, L., & Guns, R. (2018b, 1 de enero). Chapter 2 - Scientific Research and Communication. En R. Rousseau, L. Egghe & R. Guns (Eds.), *Becoming Metric-Wise* (pp. 11-35). Chandos Publishing. <https://doi.org/10.1016/B978-0-08-102474-4.00002-9>
- Rousseau, R., Egghe, L., & Guns, R. (2018c, 1 de enero). Chapter 4 - Statistics. En R. Rousseau, L. Egghe & R. Guns (Eds.), *Becoming Metric-Wise* (pp. 67-97). Chandos Publishing. <https://doi.org/10.1016/B978-0-08-102474-4.00004-2>
- Rousseau, R., Egghe, L., & Guns, R. (2018d, 1 de enero). Chapter 5 - Publication and Citation Analysis. En R. Rousseau, L. Egghe & R. Guns (Eds.), *Becoming Metric-Wise* (pp. 99-153). Chandos Publishing. <https://doi.org/10.1016/B978-0-08-102474-4.00005-4>
- Rousseau, R., Egghe, L., & Guns, R. (2018e, 1 de enero). Chapter 5 - Publication and Citation Analysis. En R. Rousseau, L. Egghe & R. Guns (Eds.), *Becoming Metric-Wise* (pp. 99-153). Chandos Publishing. <https://doi.org/10.1016/B978-0-08-102474-4.00005-4>

- Sandoval-Romero, V., & Larivière, V. (2020). The national system of researchers in Mexico: Implications of publication incentives for researchers in social sciences. *Scientometrics*, 122(1), 99-126. <https://doi.org/10.1007/s11192-019-03285-8>
- Scimago. (2022). *Data Viz Catalogue*. Consultado el 16 de marzo de 2023, desde <https://graphica.app/catalogue>
- Sierra-González, J. H., & Ramos-Pérez, C. E. (2021). Science, Technology, Innovation, and Inclusive Development: A Country Comparison Between Colombia and Mexico. En L. A. Orozco, G. Ordóñez-Matamoros, J. H. Sierra-González, J. García-Estévez & I. Bortagaray (Eds.), *Science, Technology, and Higher Education: Governance Approaches on Social Inclusion and Sustainability in Latin America* (pp. 287-343). Springer International Publishing. [https://doi.org/10.1007/978-3-030-80720-7\\_11](https://doi.org/10.1007/978-3-030-80720-7_11)
- Sooryamoorthy, R. (2019). Scientific knowledge in South Africa: Information trends, patterns and collaboration. *Scientometrics*, 119(3), 1365-1386. <https://doi.org/10.1007/s11192-019-03096-x>
- Taşkın, Z., & Doğan, G. (2020). Arts and Humanities and the others: Why can't we measure arts and humanities. *ITM Web of Conferences*, 33, 02003. <https://doi.org/10.1051/itmconf/20203302003>
- Techt, U. (2016, 1 de marzo). *Goldratt y la Teoría de Restricciones: El Salto Cuántico en Gerencia*. ibidem-Verlag / ibidem Press.
- Uddin, A., Singh, V. K., Pinto, D., & Olmos, I. (2015). Scientometric mapping of computer science research in Mexico. *Scientometrics*, 105(1), 97-114. <https://doi.org/10.1007/s11192-015-1654-y>
- Villaseñor, E. A., Arencibia-Jorge, R., & Carrillo-Calvet, H. (2017). Multiparametric characterization of scientometric performance profiles assisted by neural networks: A study of Mexican higher education institutions. *Scientometrics*, 110(1), 77-104. <https://doi.org/10.1007/s11192-016-2166-0>
- WebofScienceGroup. (2019). *WEB OF SCIENCE: Manual de Uso*. [https://www.recursoscientificos.fecyt.es/sites/default/files/spanish\\_manual\\_09\\_10\\_2019.pdf](https://www.recursoscientificos.fecyt.es/sites/default/files/spanish_manual_09_10_2019.pdf)
- Zhe, C., Lu, X., & Xiong, X. (2021). Analysis of Influence Factors on the Quality of International Collaboration Research in the Field of Social Sciences and Humanities: The Case of Chinese World Class Universities (2015–2019). *SAGE Open*, 11(4), 21582440211050381. <https://doi.org/10.1177/21582440211050381>

Zhu, J., & Liu, W. (2020). A tale of two databases: The use of Web of Science and Scopus in academic papers. *Scientometrics*, *123*(1), 321-335. <https://doi.org/10.1007/s11192-020-03387-8>

