



"El saber de mis hijos  
hará mi grandeza"

# UNIVERSIDAD DE SONORA

## Departamento de Matemáticas

Secuencia didáctica para la comprensión de la regresión lineal y la correlación entre universitarios, mediada con tecnología.

Predctoral que presenta  
**Felipe de Jesús Castro Lugo**

Doctorado en Ciencias  
con especialidad en Matemática Educativa

Directora de Tesis:  
Gloria Angélica Moreno Durazo



Hermosillo, Sonora

noviembre, 2025



## *Agradecimientos*

*Agradezco a la Secretaría de Ciencia, Humanidades, Tecnología e Innovación (Secihti) por el apoyo brindado para mi formación, con la beca de número 4022854.*



Figura 1. Ejemplo de tareas de tipo procedimental y muestra de aplicación incorrecta del modelo. ....	7
Figura 2. Ejemplo de uso del modelo e interpretación de diagrama de dispersión.....	7
Figura 3. Ilustración del quincux .....	12
Figura 4. Formas de representar la covariación estadística.....	19
Figura 5. La distribución de las secuencias.....	20
Figura 6. Un proceso hipotético cognitivo para trabajar con la línea de mejor ajuste .....	23
Figura 7. Dimensiones del marco para el desarrollo del pensamiento estadístico.....	31
Figura 8. Dimensión 1 ciclo investigativo (PPDAC).....	33
Figura 9. Dimensión 2: Tipos de pensamiento .....	36
Figura 10. Dimensión 3: ciclo investigativo .....	37
Figura 11. Dimensión 4: Disposiciones .....	38
Figura 12. Contexto del problema a describir.....	41
Figura 13. Captura de pantalla de CODAP con diferentes representaciones vinculadas.....	43
Figura 14. Captura de pantalla de CODAP de la recta de regresión y residuales.....	43



Tabla 1. Diferencias entre la función y la regresión lineal. ....	9
Tabla 2. Objetivo específico 1, Acciones metodológicas, y productos esperados.....	45
Tabla 3. Objetivo específico 2, Acciones metodológicas, y productos esperados.....	47
Tabla 4. Objetivo específico 3, Acciones metodológicas, y productos esperados.....	47
Tabla 5. Objetivo específico 4, Acciones metodológicas, y productos esperados.....	48
Tabla 6. Cronograma de los semestres 2014-1 al 2025-2. ....	58
Tabla 7. Cronograma de los semestres 2016-1 al 2027-2. ....	58



# Contenido

<b>1 INTRODUCCIÓN.....</b>	<b>1</b>
<b>2 ANTECEDENTES.....</b>	<b>3</b>
2.1 Planteamiento curricular de la correlación y la regresión lineal.....	3
2.1.1 La enseñanza de la regresión lineal y la correlación en la Educación Superior .....	4
2.2 La enseñanza de la regresión lineal y la correlación en libros de texto.....	6
2.3 Dificultades en la enseñanza y el aprendizaje de la regresión lineal y la correlación .....	8
2.4 Estudio epistemológico de la correlación y la regresión lineal .....	11
<b>3 ESTADO DEL ARTE.....</b>	<b>14</b>
3.1 Contribuciones a nivel local .....	14
3.2 Contribuciones a nivel nacional.....	15
3.3 Contribuciones a nivel internacional .....	18
<b>4 PROBLEMÁTICA Y OBJETIVOS.....</b>	<b>26</b>
4.1 Elementos desde el análisis de los errores y dificultades .....	26
4.2 Elementos desde la didáctica de la educación estadística.....	26
4.3 Elementos desde el ámbito profesional .....	28
4.4 Objetivos.....	28
<b>5 MARCO CONCEPTUAL.....</b>	<b>30</b>
5.1 El desarrollo del pensamiento estadístico para la investigación empírica.....	30
5.1.1 Dimensión uno: Ciclo investigativo.....	31
5.1.2 Dimensión dos: tipos de pensamiento.....	35
5.1.3 Dimensión tres: ciclo investigativo.....	37
5.1.4 Dimensión cuatro: Disposiciones.....	38
5.2 Elementos para la descripción de datos bivariados .....	39
5.3 Uso de tecnología.....	42
<b>6 ASPECTOS METODOLÓGICOS.....</b>	<b>45</b>
6.1 La ingeniería didáctica.....	45
6.1.1 Análisis preliminar.....	45
6.1.2 Concepción y análisis a priori de las situaciones didácticas de la ingeniería .....	46
6.1.3 Experimentación .....	47

6.1.4 Análisis a posteriori y evaluación .....	48
<b>7 LA PROPUESTA Y SUS CARACTERÍSTICAS .....</b>	<b>50</b>
7.1 Características de las actividades.....	50
7.1.1 Actividad uno: <i>La temperatura influye en el consumo de energía eléctrica.</i> .....	51
7.1.2 Actividad dos: <i>El tiempo de secado es factor importante para la resistencia en el concretos.</i> .....	51
7.1.3 Actividad tres: <i>La ergonomía me ayuda al diseño de productos.</i> .....	51
7.2 Ejemplo de actividad didáctica .....	52
<b>8 CRONOGRAMA.....</b>	<b>58</b>
<b>REFERENCIAS.....</b>	<b>61</b>
<b>ANEXOS .....</b>	<b>67</b>

# 1 INTRODUCCIÓN

El estudio de la regresión lineal y la correlación constituye un eje fundamental dentro de la educación estadística, tanto por su relevancia matemática como por su utilidad en la modelación y análisis de fenómenos reales. A pesar de su presencia explícita en los programas curriculares de educación superior, la literatura evidencia que su enseñanza continúa enfrentando desafíos importantes, entre ellos la falta de conexión conceptual, el predominio de enfoques procedimentales y las dificultades de los estudiantes para interpretar representaciones gráficas o establecer relaciones entre variables. En este contexto, el presente trabajo se propone diseñar, implementar y valorar una propuesta didáctica mediada por tecnología, orientada a mejorar la comprensión de estos conceptos en estudiantes de ingeniería.

Los antecedentes muestran que la regresión lineal y la correlación se han incorporado en los planes de estudio bajo el enfoque de análisis de datos, buscando que los estudiantes identifiquen patrones, modelen relaciones y tomen decisiones fundamentadas. Sin embargo, los enfoques didácticos tradicionales suelen centrarse en el cálculo algebraico o en el uso mecánico de fórmulas, lo que limita el desarrollo de un pensamiento estadístico profundo. El análisis de los libros de texto confirma esta tendencia, pues la mayoría introduce los conceptos de manera lineal, con escasa atención a la variabilidad, al razonamiento gráfico o a la interpretación contextual. Estos factores, junto con dificultades documentadas en la literatura —como la confusión entre correlación y causalidad, la identificación imprecisa de la recta de mejor ajuste o la interpretación limitada de la fuerza de la relación—, justifican la necesidad de propuestas de enseñanza que integren contextos reales, tecnología y un enfoque progresivo de formalización.

El estado del arte refuerza esta necesidad. A nivel local, los estudios destacan la importancia del uso de herramientas tecnológicas como GeoGebra, Fathom, Excel o CODAP para favorecer la visualización y la exploración de datos bivariados. Estas investigaciones muestran que los ambientes interactivos permiten a los estudiantes reconocer patrones globales, comparar modelos y comprender la variabilidad de manera más intuitiva. En el ámbito nacional, la literatura se ha enfocado en el diseño de trayectorias de aprendizaje y secuencias basadas en múltiples representaciones, evidenciando avances en la interpretación de tendencias y fuerza de relación, aunque persisten dificultades en la comprensión de la variabilidad residual y el papel del modelo en la predicción. En el plano internacional, los estudios coinciden en que el razonamiento covariacional (necesario para comprender regresión y correlación) no surge espontáneamente; más bien, requiere apoyos didácticos sistemáticos que integren intuiciones, descripciones visuales y modelos formales. Investigaciones clásicas y recientes han mostrado que los estudiantes suelen emplear criterios intuitivos para ubicar la recta de ajuste o interpretar patrones de dispersión, lo que subraya la necesidad de intervenciones que promuevan una construcción gradual de significado.

El marco conceptual del proyecto se organiza en torno al pensamiento estadístico y al análisis de datos bivariados mediado por tecnología. Se retoman los aportes de Wild y Pfannkuch (1999) para

caracterizar la estadística como un proceso dinámico basado en el ciclo investigativo PPDAC, el uso de modelos, la gestión de la variación y la argumentación con evidencia. Este enfoque es particularmente pertinente para la formación de ingenieros, donde se requiere interpretar fenómenos con datos reales, desarrollar criterio analítico y comunicar resultados de manera fundamentada. De forma complementaria, se integran los elementos propuestos por Watkins et al. (2007) para describir gráficamente conjuntos de datos bivariados —forma, tendencia y fuerza—, los cuales permiten construir un puente entre el análisis informal y el ajuste formal de un modelo lineal. La tecnología digital desempeña un papel central al facilitar la manipulación de datos, la visualización dinámica de gráficos y la experimentación con modelos, lo que promueve la consolidación de representaciones más robustas y significativas.

Para orientar el diseño y la implementación de la propuesta, el proyecto adopta la metodología de la ingeniería didáctica de Artigue (1995), que ofrece una estructura rigurosa para el análisis, la concepción, la experimentación y la evaluación de secuencias de enseñanza. En la fase de análisis preliminar se realizan estudios epistemológicos, cognitivos y didácticos que permiten identificar las necesidades formativas y las dificultades recurrentes en torno al aprendizaje de la regresión lineal y la correlación. En la fase de concepción y análisis a priori se definen las variables didácticas, el número y propósito de las actividades, el uso de la tecnología y los instrumentos de observación y recolección de datos. La fase de experimentación implica la implementación de la secuencia con estudiantes de ingeniería, bajo un esquema de interacción guiada y con un uso moderado pero significativo de herramientas tecnológicas. Finalmente, el análisis a posteriori permite contrastar los resultados obtenidos con las previsiones del análisis a priori, valorar el logro de los objetivos, validar las hipótesis de diseño y generar conclusiones y recomendaciones.

En conjunto, este proyecto busca contribuir a la mejora de la enseñanza de la regresión lineal y la correlación mediante una propuesta didáctica pertinente, contextualizada y fundamentada tanto en la investigación educativa como en las necesidades reales de los estudiantes de ingeniería. Su desarrollo pretende no solo fortalecer la comprensión conceptual, sino también promover prácticas de análisis de datos coherentes con las demandas actuales de la formación profesional y de la cultura estadística contemporánea.

En sección de las características de la propuesta, se declaran los objetivos, incluyendo las bases del marco conceptual y aspectos metodológicos que sustentan la secuencia de actividades.

Por último, en el cronograma se muestran las actividades realizadas en los semestres anteriores y las actividades que se realizarán para concretar el proyecto de tesis.

## **2 ANTECEDENTES**

### **2.1 Planteamiento curricular de la correlación y la regresión lineal**

La Asociación Americana de Estadística (ASA), define a la estadística como: “La ciencia de aprender de los datos y de medir, controlar y comunicar la incertidumbre” (Davidian, and Louis, 2012, p.12). Según autores, esta definición aborda un enfoque holístico, que incluye no solo la recolección y el análisis de datos, sino también la interpretación, la comunicación de los resultados y la generación de nuevos cuestionamientos; sobre todo, se destaca la relevancia de comprender y gestionar la presencia de la incertidumbre en los datos.

En el contexto mexicano, el estudio de nociones estadísticas relacionadas con la recolección, el análisis y la interpretación de los datos inicia desde la Educación Básica (NEM, 2022). Puntualmente, los procesos de desarrollo de aprendizaje se corresponden con la interpretación de información estadística mediante representaciones gráficas simples, como pictogramas en las etapas iniciales, y posteriormente con gráficos de barras y circulares. En fases más avanzadas, se amplían las habilidades hacia la recolección y el registro de la información en tablas, con el propósito de responder a cuestiones de interés para los estudiantes; buscando favorecer los aprendizajes significativos y contextualizados según las características de cada región. Además, se introducen nociones de probabilidad, que permiten comenzar a cuantificar la incertidumbre.

En la secundaria, se espera que los estudiantes profundicen en la interpretación de datos mediante el uso de estadísticos descriptivos, como las medidas de posición y variación, así como el manejo de frecuencias absolutas y relativas. También se incorporan gráficos más complejos, como histogramas, gráficas de líneas y polígonos de frecuencias.

Estos aprendizajes se robustecen en la Educación Media Superior (SEP, 2022), donde la estadística ocupa un lugar relevante en la formación científica de los estudiantes, y se continúan desarrollando en la mayoría de las licenciaturas e ingenierías dentro de la Educación Superior (p. ej., ITSON, 2023). Grosso modo, en el nivel superior, primeramente se estudian nociones de la estadística descriptiva, promoviendo el análisis de datos mediante recursos gráficos, tabulares y numéricos, a través de medidas de centro y dispersión, seguido del estudio de conceptos de probabilidad orientados a la cuantificación de la incertidumbre en fenómenos propios de cada disciplina, y finalmente, se aborda la inferencia estadística aplicados a la estimación de parámetros poblacionales mediante diferentes herramientas.

Como puede observarse, la formación estadística está presente a lo largo de todo el trayecto educativo, lo que resulta coherente con la naturaleza transversal de esta área del conocimiento y su aplicación en múltiples ámbitos: científicos, tecnológicos y cotidianos. En la era actual de

la información, caracterizada por la abundancia de datos, comprender las herramientas estadísticas se convierte en un componente esencial de la formación ciudadana (Wild y Phannkuch, 1999).

De lo anterior, un aspecto que resulta contradictorio en la organización curricular es que los conceptos estadísticos y probabilísticos se ubiquen al final de los programas de matemáticas. Esta ubicación tiende a limitar el tiempo y la profundidad con que los temas son abordados, reduciendo su potencial para fomentar un pensamiento estadístico sólido y aplicado.

Otro aspecto relevante es que los contenidos propuestos para el estudio de la estadística guardan relación con *las ideas estadísticas fundamentales: los datos, la variación, las distribuciones, las representaciones, la asociación y los modelos, los modelos probabilísticos y el muestreo e inferencia* (Burrill y Biehler, 2013). Sin embargo, la secuencia y el enfoque pedagógico de estos contenidos no siempre se sustentan en principios de la educación estadística, como los sugeridos por el informe GAISE (Guidelines for Assessment and Instruction in Statistics Education), lo que genera inconsistencias entre los objetivos formativos y las prácticas de enseñanza.

Esta estructura se mantiene también en el nivel superior, por ejemplo, las modificaciones recientes en el programa del curso de Probabilidad y Estadística para ingenieros (ITSON, 2023) reubicaron el tema *regresión lineal simple* a la última sección “Realiza inferencias sobre parámetros utilizando los métodos de la Estadística Inferencial para caracterizar variables en la vida cotidiana y las áreas de Ingeniería”, en contraste con el programa del curso en el 2016 que lo consideraba como una unidad independiente. De esto se infieren dificultades entre los estudiantes para reconocer la utilidad de la estadística como herramienta de análisis y toma de decisiones en contextos profesionales.

Así, considerando que el estudio de la *regresión lineal* suele ubicarse al final de los cursos de estadística, lo cual puede prestarse para que se aborde de forma superficial o incluso sea omitido por falta de tiempo, se pone de manifiesto la necesidad de replantear su papel en la trayectoria formativa promoviendo su abordaje mediante contextos significativos mediante el desarrollo de un proyecto de intervención didáctica.

### **2.1.1 La enseñanza de la regresión lineal y la correlación en la Educación Superior**

La enseñanza de la noción de *correlación y regresión lineal*, introducida en México en el bachillerato, se basa en la idea de desarrollar habilidades para analizar e interpretar datos bivariados en contexto y no solo para aplicar fórmulas y procedimientos (NEM, 2022), específicamente bajo la progresión 9 que tiene por objetivo: “analiza dos o más variables cuantitativas a través del estudio de alguna problemática o fenómenos de interés para el estudiantado, con la finalidad de identificar si existe correlación en dichas variables”. Esta

progresión involucra otros conceptos, por ejemplo, el diagrama de dispersión, la correlación, la ecuación de la recta de predicción, determinar el grado de asociación.<sup>1</sup>

En el Nivel Superior la correlación y la regresión lineal están integradas a la mayoría de los planes y programas de los programas educativos, insertados en las asignaturas de Probabilidad o Estadística, procurando el desarrollo de habilidades profesionales, por ejemplo, para el área de ingeniería fomentar la toma de decisiones a partir del análisis de la eficiencia o bien, mediante la correlación de las variables implicadas en los procesos productivos (ITSON, 2016).

Particularmente, en el modelo curricular del Instituto Tecnológico de Sonora (ITSON, 2022) se establecen las directrices para el diseño de los programas académicos y las competencias, habilidades y estrategias que se espera que el estudiante integre fortaleciendo su perfil de egreso. Algunas de las características que se declaran en el perfil de egreso en la carrera de ingeniería industrial y de sistemas está, "... orientado a incrementar la eficiencia y competitividad, tomando decisiones bajo un enfoque de sostenibilidad, con el apoyo en la analítica de datos, modelación y tecnologías de la información y la comunicación..." (ITSON, 2023). En sentido similar a lo declarado por Watkins et al. (2008) "no puedes hacer frente de forma inteligente a este mundo cuantitativo a menos que entiendas los conceptos estadísticos básicos y hayas practicado tomando decisiones informadas utilizando datos reales". Es decir, en la actualidad se demanda a los profesionistas que tomen decisiones basadas en datos.

El tema de *regresión lineal* para las carreras de ingeniería está incluido en la asignatura de Probabilidad y Estadística (ITSON, 2023) que se imparte a las siguientes carreras: Ingeniería Industrial y de Sistemas, Ingeniería Mecatrónica, Ingeniería Electromecánica, Ingeniería Civil, Ingeniería Química, Ingeniería en Electrónica e Ingeniería en Software. El curso forma parte de la academia de ciencias básicas y está dividido en unidades de competencia, la unidad de interés es la unidad de competencia IV relativa a la Regresión lineal simple; descrita de la siguiente manera: "Modelar la relación lineal existente entre dos variables", además, esta unidad de competencia incluye los siguientes elementos de competencia: concientizar sobre la presencia de aleatoriedad; diferenciar entre los modelos probabilísticos y los deterministas; ajustar un modelo de regresión a datos recolectados; diferenciar el modelo de la realidad; realizar análisis críticos de resultados obtenidos modelos de regresión; capacidad de comunicar resultados de los modelos de regresión lineal y utilizar tecnología como apoyo en el análisis e interpretación de la información. (ITSON, 2016).

---

<sup>1</sup> A raíz de las reformas realizadas en el 2025, la asignatura de Probabilidad y Estadística, anteriormente denominada Pensamiento matemático I y ubicada en el primer semestre, fue reubicada al sexto semestre del plan de estudios. Este cambio responde a una reestructuración integral del currículo, orientado a fortalecer la progresión de los aprendizajes matemáticos y a garantizar que los estudiantes cuenten con los conocimientos previos y las competencias analíticas necesarias para abordar con mayor profundidad los contenidos estadísticos (NEM, 2025). Según documentos oficiales, la nueva ubicación de la materia busca articular con otras asignaturas avanzadas del área de matemáticas aplicadas y con proyectos de investigación o intervención disciplinar, favoreciendo un enfoque más contextualizado, reflexivo y orientado al análisis de datos reales.

La organización curricular de los cursos de Probabilidad y Estadística en el ITSON tiene características similares a las reportadas en (GAISE, 2016), por ejemplo, los contenidos de los cursos están comprendidos en función del área en el que se impartirán, pero no logran grandes diferencias en la enseñanza de los conceptos en carreras distintas. Además, estos cursos están enfocados en los métodos estadísticos.

## **2.2 La enseñanza de la regresión lineal y la correlación en libros de texto**

Mayores elementos sobre la enseñanza de la regresión lineal y la correlación en la Educación Superior se obtienen al analizar los libros sugeridos en los programas de las materias. Especialmente, se analiza el programa de estudio de la materia de Probabilidad y estadística para ingeniería de ITSON, revisando los apartados que incluyen análisis de datos bivariados y regresión lineal simple y correlación. Con esta actividad se tendrá un punto de partida sobre el tipo de contextos que pueden ser incluidos en las actividades didácticas a diseñar y el tipo de cuestiones a explorar.

El programa del curso propone cuatro libros los cuales están clasificados como bibliografía básica y bibliografía de consulta, de estos apartados se seleccionaron los siguientes tres libros: Walpole (2012), Devore (2012) y Mendenhall (2015). Se dejó a un lado el libro *Probabilidad y estadística para la ingeniería: un enfoque moderno* de Spiegel y Murray (2010), dado que, en los temas de interés, no emplean las herramientas que el programa de estudio propone para la determinación del modelo lineal simple.

El análisis se basa en algunos elementos propuestos en las investigaciones en educación estadística para caracterizar las situaciones problema (Cobb et al. 1997; Batanero, C. 2001; Garfield et al. 2008 y Biehler et al 2012): 1) contexto significativo y auténtico, *que estos sean cercanos a su contexto o quehacer* 2) datos reales y manipulables, *que vivan la experiencia de la recolección* 3) exploración y análisis de datos, 4) variabilidad e incertidumbre; 5) trabajo colaborativo y discusión; 6) enfoque por indagación y resolución de problemas; 7) conexión con otra disciplinas y 8) uso de representaciones diversa.

De estos elementos se seleccionaron los que se identificaron en los libros, *el énfasis del problema/tarea* (Procedimental, conceptual o mixto); *el contexto* (intra-matemático, extra-matemático o tecnológico); *el uso de tecnología; las representaciones semióticas* que se visualizan y *la familia de problemas*. Para esta última categoría se tomaron los elementos propuestos por Gea (2012), las cuales fueron ampliadas y refinadas ya que en nuestro estudio lo realizamos con libros de nivel superior (ingeniería).

En los libros analizados, la mayoría de las tareas propuestas son de carácter procedimental (48%), en contraste con un menor porcentaje de actividades conceptuales (35%) y mixtas (17%). Asimismo, se advierte un uso limitado de tecnología, limitado a software especializado como Minitab o applets, que se plantean más como apoyo instrumental que como herramienta para la exploración y comprensión de fenómenos. Se identifica una inconsistencia en el uso de símbolos y terminología para la regresión, además, se encuentra la presencia de errores didácticos, como

el caso de un ejercicio en Walpole (2012) en el que se emplea incorrectamente la ecuación de predicción, a continuación, se presentamos ejemplos de esto, ver figura 1:

Figura 1. Ejemplo de tareas de tipo procedimental y muestra de aplicación incorrecta del modelo.

<p><b>11.11</b> La fuerza de impulso de un motor (<math>y</math>) es una función de la temperatura de escape (<math>x</math>) en °F cuando otras variables de importancia se mantienen constantes. Considere los siguientes datos.</p> <table style="margin-left: auto; margin-right: auto; border-collapse: collapse;"> <tr> <td style="border-right: 1px solid black; padding: 5px;"><math>y</math></td> <td style="padding: 5px;"><math>x</math></td> <td style="border-right: 1px solid black; padding: 5px;"><math>y</math></td> <td style="padding: 5px;"><math>x</math></td> </tr> <tr> <td style="border-right: 1px solid black; text-align: center;">4300</td> <td style="text-align: center;">1760</td> <td style="border-right: 1px solid black; text-align: center;">4010</td> <td style="text-align: center;">1665</td> </tr> <tr> <td style="border-right: 1px solid black; text-align: center;">4650</td> <td style="text-align: center;">1652</td> <td style="border-right: 1px solid black; text-align: center;">3810</td> <td style="text-align: center;">1550</td> </tr> <tr> <td style="border-right: 1px solid black; text-align: center;">3200</td> <td style="text-align: center;">1485</td> <td style="border-right: 1px solid black; text-align: center;">4500</td> <td style="text-align: center;">1700</td> </tr> <tr> <td style="border-right: 1px solid black; text-align: center;">3150</td> <td style="text-align: center;">1390</td> <td style="border-right: 1px solid black; text-align: center;">3008</td> <td style="text-align: center;">1270</td> </tr> <tr> <td style="border-right: 1px solid black; text-align: center;">4950</td> <td style="text-align: center;">1820</td> <td></td> <td></td> </tr> </table> <p>a) Grafique los datos. b) Ajuste una recta de regresión simple a los datos y grafíquela a través de ellos.</p>	$y$	$x$	$y$	$x$	4300	1760	4010	1665	4650	1652	3810	1550	3200	1485	4500	1700	3150	1390	3008	1270	4950	1820			<p><b>11.4</b> Para fines de calibración se recabaron los siguientes datos, los cuales permitirían determinar la relación entre la presión y la lectura correspondiente en la escala.</p> <table style="margin-left: auto; margin-right: auto; border-collapse: collapse;"> <thead> <tr> <th style="border-bottom: 1px solid black; padding: 5px;">Presión, <math>x</math> (lb/pulg<sup>2</sup>)</th> <th style="border-bottom: 1px solid black; padding: 5px;">Lectura en la escala, <math>y</math></th> </tr> </thead> <tbody> <tr><td style="text-align: center;">10</td><td style="text-align: center;">13</td></tr> <tr><td style="text-align: center;">10</td><td style="text-align: center;">18</td></tr> <tr><td style="text-align: center;">10</td><td style="text-align: center;">16</td></tr> <tr><td style="text-align: center;">10</td><td style="text-align: center;">15</td></tr> <tr><td style="text-align: center;">10</td><td style="text-align: center;">20</td></tr> <tr><td style="text-align: center;">50</td><td style="text-align: center;">86</td></tr> <tr><td style="text-align: center;">50</td><td style="text-align: center;">90</td></tr> <tr><td style="text-align: center;">50</td><td style="text-align: center;">88</td></tr> <tr><td style="text-align: center;">50</td><td style="text-align: center;">88</td></tr> <tr><td style="text-align: center;">50</td><td style="text-align: center;">92</td></tr> </tbody> </table> <p>a) Calcule la ecuación de la recta de regresión. b) En esta aplicación el propósito de la calibración es estimar la presión a partir de una lectura observada en la escala. Estime la presión para una lectura en la escala de 54, usando <math>\hat{x} = (54 - b_0)/b_1</math>.</p>	Presión, $x$ (lb/pulg <sup>2</sup> )	Lectura en la escala, $y$	10	13	10	18	10	16	10	15	10	20	50	86	50	90	50	88	50	88	50	92
$y$	$x$	$y$	$x$																																												
4300	1760	4010	1665																																												
4650	1652	3810	1550																																												
3200	1485	4500	1700																																												
3150	1390	3008	1270																																												
4950	1820																																														
Presión, $x$ (lb/pulg <sup>2</sup> )	Lectura en la escala, $y$																																														
10	13																																														
10	18																																														
10	16																																														
10	15																																														
10	20																																														
50	86																																														
50	90																																														
50	88																																														
50	88																																														
50	92																																														

Fuente: Walpole, (2012, pp. 399 y 400)

Se observa un énfasis marcado en tareas de construcción e interpretación de diagramas de dispersión (34%), mientras que el análisis y uso del modelo de regresión lineal simple ocupa un espacio reducido (20%). Ver siguiente figura.

Figura 2. Ejemplo de uso del modelo e interpretación de diagrama de dispersión.

<p><b>3.24 ¡Queso, otra vez!</b> La demanda de alimentos saludables que sean bajos en grasas y calorías ha resultado en un gran número de productos "bajo en grasas" y "sin grasa" en el supermercado. La tabla siguiente muestra los números de calorías y las cantidades de sodio (en miligramos) por rebanada para cinco marcas diferentes de queso estadounidense libre de grasa.</p> <table style="margin-left: auto; margin-right: auto; border-collapse: collapse;"> <thead> <tr> <th style="border-bottom: 1px solid black; padding: 5px;">Marca</th> <th style="border-bottom: 1px solid black; padding: 5px;">Sodio (mg)</th> <th style="border-bottom: 1px solid black; padding: 5px;">Calorías</th> </tr> </thead> <tbody> <tr><td style="padding: 5px;">Kraft Fat Free Singles</td><td style="text-align: center; padding: 5px;">300</td><td style="text-align: center; padding: 5px;">30</td></tr> <tr><td style="padding: 5px;">Ralphs Fat Free Singles</td><td style="text-align: center; padding: 5px;">300</td><td style="text-align: center; padding: 5px;">30</td></tr> <tr><td style="padding: 5px;">Borden Fat Free</td><td style="text-align: center; padding: 5px;">320</td><td style="text-align: center; padding: 5px;">30</td></tr> <tr><td style="padding: 5px;">Healthy Choice Fat Free</td><td style="text-align: center; padding: 5px;">290</td><td style="text-align: center; padding: 5px;">30</td></tr> <tr><td style="padding: 5px;">Smart Beat American</td><td style="text-align: center; padding: 5px;">180</td><td style="text-align: center; padding: 5px;">25</td></tr> </tbody> </table> <p>a. Trace una gráfica de dispersión para describir la relación entre la cantidad de sodio y el número de calorías. b. Describa la gráfica del inciso a). ¿Ve usted algunos resultados atípicos? ¿El resto de los puntos parece formar un modelo? c. Con base <i>sólo</i> en la relación entre sodio y calorías, ¿puede usted tomar una decisión clara acerca de cuál de las cinco marcas comprar? ¿Es razonable basar su elección en sólo estas dos variables? ¿Qué otras variables debe considerar?</p>	Marca	Sodio (mg)	Calorías	Kraft Fat Free Singles	300	30	Ralphs Fat Free Singles	300	30	Borden Fat Free	320	30	Healthy Choice Fat Free	290	30	Smart Beat American	180	25	<p><b>11.6</b> En cierto tipo de espécimen de prueba metálico se sabe que la tensión normal sobre un espécimen se relaciona funcionalmente con la resistencia al corte. El siguiente es un conjunto de datos experimentales codificados para las dos variables:</p> <table style="margin-left: auto; margin-right: auto; border-collapse: collapse;"> <thead> <tr> <th style="border-bottom: 1px solid black; padding: 5px;">Tensión normal, <math>x</math></th> <th style="border-bottom: 1px solid black; padding: 5px;">Resistencia al corte, <math>y</math></th> </tr> </thead> <tbody> <tr><td style="text-align: center;">26.8</td><td style="text-align: center;">26.5</td></tr> <tr><td style="text-align: center;">25.4</td><td style="text-align: center;">27.3</td></tr> <tr><td style="text-align: center;">28.9</td><td style="text-align: center;">24.2</td></tr> <tr><td style="text-align: center;">23.6</td><td style="text-align: center;">27.1</td></tr> <tr><td style="text-align: center;">27.7</td><td style="text-align: center;">23.6</td></tr> <tr><td style="text-align: center;">23.9</td><td style="text-align: center;">25.9</td></tr> <tr><td style="text-align: center;">24.7</td><td style="text-align: center;">26.3</td></tr> <tr><td style="text-align: center;">28.1</td><td style="text-align: center;">22.5</td></tr> <tr><td style="text-align: center;">26.9</td><td style="text-align: center;">21.7</td></tr> <tr><td style="text-align: center;">27.4</td><td style="text-align: center;">21.4</td></tr> <tr><td style="text-align: center;">22.6</td><td style="text-align: center;">25.8</td></tr> <tr><td style="text-align: center;">25.6</td><td style="text-align: center;">24.9</td></tr> </tbody> </table> <p>a) Estime la recta de regresión <math>\mu_{y x} = \beta_0 + \beta_1 x</math>. b) Estime la resistencia al corte para una tensión normal de 24.5.</p>	Tensión normal, $x$	Resistencia al corte, $y$	26.8	26.5	25.4	27.3	28.9	24.2	23.6	27.1	27.7	23.6	23.9	25.9	24.7	26.3	28.1	22.5	26.9	21.7	27.4	21.4	22.6	25.8	25.6	24.9
Marca	Sodio (mg)	Calorías																																											
Kraft Fat Free Singles	300	30																																											
Ralphs Fat Free Singles	300	30																																											
Borden Fat Free	320	30																																											
Healthy Choice Fat Free	290	30																																											
Smart Beat American	180	25																																											
Tensión normal, $x$	Resistencia al corte, $y$																																												
26.8	26.5																																												
25.4	27.3																																												
28.9	24.2																																												
23.6	27.1																																												
27.7	23.6																																												
23.9	25.9																																												
24.7	26.3																																												
28.1	22.5																																												
26.9	21.7																																												
27.4	21.4																																												
22.6	25.8																																												
25.6	24.9																																												

Fuente: Mendenhall, (2015, p. 120); Walpole, (2012, p. 399)

Las características de las actividades en libros de texto contrastan con las recomendaciones internacionales (GAISE, 2016; Wild & Pfannkuch, 1999), que sugieren privilegiar la indagación, la variabilidad y la toma de decisiones fundamentadas en datos reales. En este sentido, los libros analizados ofrecen una aproximación tradicional que, si bien cumple con los contenidos curriculares, no potencia el desarrollo de competencias estadísticas críticas y aplicadas a contextos auténticos.

Con esta actividad se tendrá un punto de partida sobre el tipo de contextos que pueden ser incluidos en las actividades didácticas a diseñar y el tipo de cuestiones a explorar

### **2.3 Dificultades en la enseñanza y el aprendizaje de la regresión lineal y la correlación**

Diversas investigaciones muestran que *la correlación y la regresión lineal* no son fáciles de comprender. Esto ha sido reportado desde diversas áreas científicas, como la psicología (interesados en las formas de pensar y razonar de los individuos) y la matemática educativa (interesados en el uso, comprensión de los conceptos y su inserción en el currículo). Desde la Psicología se comparten resultados importantes sobre la comprensión de la correlación y la regresión lineal, por un lado, Inhelder y Piaget mencionan que previo a la etapa de las operaciones formales no es posible desarrollar la idea de correlación, debido a que en las etapas tempranas del desarrollo cognitivo, el pensamiento infantil se caracteriza por una lógica más concreta y centrada en la percepción inmediata, lo que dificulta establecer relaciones abstractas entre dos variables. Por otro lado, Chapman y Chapman (1969, citado por Estepa, 1993) reportan que “La detención de la asociación no es habilidad espontáneamente adquirida y los sujetos, aún adultos, se guían preferentemente por sus teorías previas sobre las interrelaciones entre las variables más que por los datos objetivos sobre las contingencias empíricas” (p. 30).

Desde la matemática educativa, Estepa y Batanero (1996) en su estudio con estudiantes preuniversitarios, identifican tres concepciones erróneas de la correlación: 1) *La concepción determinista* se presenta cuando los estudiantes sólo admiten un valor para la variable dependiente para cualquier valor de la variable independiente, para considerar que las variables están relacionadas (consideran independientes a las variables si no se cumple esta condición), 2) *la concepción local*, cuando el estudiante realiza el juicio de asociación basándose únicamente en una parte de los datos presentados. (si detecta un patrón en la relación de las variables, por ejemplo, directa, lo generaliza para todo el conjunto de datos) y, 3) *la concepción causal*, el estudiante se basa únicamente en la existencia o no de relación causal para afirmar la existencia de correlación, sin tener en cuenta la correlación empírica observada en los datos. Otro tipo de concepción (*unidireccional*) reportada por la literatura se detectó en el estudio de Batanero et al., (1996), presentándose cuando los estudiantes perciben la existencia de asociación, solamente cuando la relación entre las variables es directa, considerando la relación inversa como independiente.

Otro tipo de dificultades reportadas en la literatura devienen de la instrucción matemática, por ejemplo, Agnelli et al., (2009) alertan que las dificultades pueden convertirse en obstáculos cuando se realizan tratamientos en situaciones escolarizadas si poner especial atención a la distinción entre conceptos, en la selección y análisis de situaciones y en la conducción de instrucción correspondiente. A continuación, se presenta un resumen con las diferencias entre la función y la regresión lineal abordadas por los autores (tabla 1), las cuales resulta importante poner atención en el proceso de enseñanza.

Tabla 1. Diferencias entre la función y la regresión lineal.

Función lineal	Regresión lineal
La variación de x e y es determinista.	La variación de una o ambas, x e y es aleatoria.
Relación funcional.	Vínculo estadístico que extiende la noción de dependencia funcional.
El cálculo de los parámetros es independiente de los puntos seleccionados sobre la recta.	La estimación de los parámetros depende de los puntos seleccionados.
Dada una tabla de valores de x e y se obtiene una recta y dada esta es posible reproducir la tabla.	Dada una tabla de valores de x e y se obtiene una recta, pero de ellas no permite reproducir la tabla.
El rol de variables independiente y dependiente es indistinto.	Los roles de las variables explicativa y respuesta no son intercambiables.
$y = a_1x + b_1 \quad y$ $x = a_2y + b_2$	$y = \alpha_1x + \beta_1 \quad y$ $x = \alpha_2y + \beta_2$
Representan el mismo conjunto de puntos.	No representan el mismo conjunto de puntos.

*Fuente:* Angnelli, et al., (2009).

Respecto a estudiantes de nivel superior, Estepa y Sánchez-Cobo (2003) se centraron en estudiar los procedimientos en conceptos relacionados con la correlación y la regresión en situaciones de resolución de problemas, particularmente el centro de gravedad del diagrama de dispersión  $(\bar{x}, \bar{y})$ , las rectas de regresión  $y = mx + n$ ;  $y - \bar{y} = \frac{S_{xy}}{S_x^2}(x - \bar{x})$ , el coeficiente de correlación  $(r = \frac{S_{xy}}{(S_x S_y)})$ , el tipo de relación entre las variables y la predicción. Sobre esto, la mayoría de los estudiantes calculan de manera correcta el coeficiente de correlación, contrario a que solo la mitad de los estudiantes llega a una expresión satisfactoria de la recta de regresión. Los estudiantes tienen dificultades al tratar de relacionar conceptos matemáticos y estadísticos en la resolución del problema al intentar interpretar los parámetros de la ecuación de la recta. Además,

la confusión de las variables dependiente e independiente genera errores que dificultan la resolución adecuada de este tipo de problemas. Otras dificultades son las confusiones entre el sentido y la intensidad de la correlación entre las variables involucradas y la creencia de que la variación conjunta es proporcional.

Por otro lado, dada la relación que tienen *la correlación y la regresión lineal* con la covariación, se retoman las investigaciones de Moritz (2004), Zieffler y Garfield (2009). En la primera investigación el objetivo es indagar sobre tres habilidades de razonamiento, la generación especulativa de datos, la interpretación verbal del gráfico y la interpretación numérica del gráfico. Los resultados son descritos en cuatro niveles de razonamiento acerca de la covariación (no estadístico, aspecto único, covariación inadecuada y covariación apropiada):

En la generación especulativa de datos se identificó el razonamiento no estadístico (cuando los estudiantes al argumentar solo utilizaban el contexto del problema y no el conjunto de datos) y la covariación apropiada (cuando los estudiantes representan ambas variables y les dan la correspondencia adecuada).

En la interpretación verbal de las gráficas de dispersión se presenta el razonamiento covariacional inadecuado (cuando los estudiantes hacen referencia a las dos variables, pero no generalizan la comparación utilizando dos o más puntos y se describen las variables, pero la correspondencia no se menciona o está en la dirección incorrecta).

En síntesis, se detectaron tres dificultades a las que se enfrentaron los estudiantes: centrarse solo en puntos bivariados aislados, centrarse solo en una variable en lugar de analizar las dos variables en conjunto y manejar una covariación negativa contraria a la creencia previa en una asociación positiva.

En el estudio de Zieffler y Garfield (2009) se estudia el desarrollo del razonamiento de estudiantes universitarios sobre los datos cuantitativos bivariados durante un semestre, donde se sugiere que los estudiantes tienen un conocimiento inicial sobre estadística muy pobre a pesar, de que el currículo en los niveles que ya habían cursado los estudiantes lo proponían. Además, mencionan que el estudio de este tema puede ser localizado en cualquier parte del curso, quedando a criterio del docente o de los autores de libros de texto; enfatizando que es necesario dedicar muchas horas en el desarrollo del razonamiento de los estudiantes sobre la distribución univariada para proporcionar bases sólidas que consecuentemente fortalecerán el razonamiento de datos bivariados.

Por su parte Zimmerman (2005), nos alerta que la asociación entre eventos –la correlación y la regresión, en particular– es un requisito esencial pero no es suficiente por si sola para establecer una relación causal, sin embargo, sirve como fundamento para realizar inferencias inductivas causales en el ámbito científico; esto es, la toma de decisiones es apoyada por el estudio de la asociación entre variables relevantes, pero es posible encontrar un coeficiente de correlación alto al comparar variables, y no existir una relación causal entre las mismas (variables espurias) (Ellet & Erickson, 1986, como se cita en Estepa, 1996).

Adicionalmente, la identificación detalla de los errores y dificultades que los estudiantes enfrentan al trabajar con regresión lineal y correlación ofrecen información valiosa para ajustar estrategias didácticas que responden a sus necesidades reales. Este análisis permite diseñar actividades más efectivas, que ayuden a superar obstáculos comunes como la interpretación determinista, la concepción local, la identificación de las variables, el enfoque local sobre los datos o problema en la extrapolación e interpolación de resultados. Abordar estos aspectos en la propuesta didáctica no solo favorece una mejor comprensión de los conceptos estadísticos, sino que también potencia el desarrollo de un razonamiento de la covariación más profundo y ampliado.

## **2.4 Estudio epistemológico de la correlación y la regresión lineal**

La asociación estadística implica más que solo una relación de variables, sino una relación de cantidades medidas de características distintas, como menciona Moore (1990) “[los datos] no son simplemente números, sino números con un contexto” (p.96). En específico, las técnicas de regresión y correlación buscan determinar la interrelación entre dos o más variables para la construcción de un modelo matemático con el cual predecir en escenarios de incertidumbre, es decir, las utilidades principales de un análisis de regresión son:

Determinar la causalidad entre las variables dependiente e independiente, obtener un predictor para la variable dependiente a partir de los valores de la variable independiente, lo cual ayuda si la variable dependiente es costosa y/o difícil de medir (Estepa, 1993).

Además, la asociación en general *la correlación y la regresión, en particular* es una extensión de la dependencia funcional y es fundamental para muchos métodos que son usados al hacer predicciones y tomar decisiones en áreas como: política, economía, medicina y educación. Muchas de esas aplicaciones intentan brindar explicaciones que nos permita entender el entorno (Estepa, 1996; Koslowski, 1996; Schauble, 1996).

El surgimiento de las nociones de correlación y la regresión lineal es presentado por Estepa et al. (2012) a través de un análisis histórico – epistemológico:

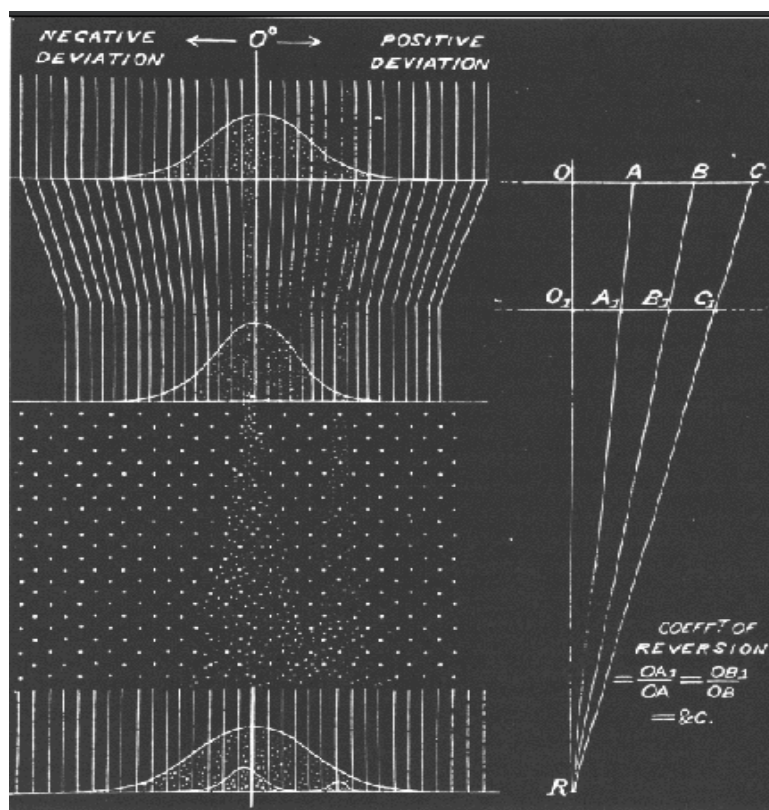
El primero en interesarse en el tema fue Lambert Adolphe-Jacques Quetelet (1796-1874). Sus aportaciones sobre la correlación y regresión lineal se originan desde sus estudios sobre “el hombre medio”, estimando empíricamente las medias y desviaciones típicas de medidas antropométricas que, suponía, dependen de varias variables independientes tales como el sexo, edad, profesión o nivel de educación.

En otro contexto, Augusto Bravais (1811-1863) desde la astronomía contribuye al desarrollo de la estadística al estudiar los errores en las medidas de las coordenadas de cuerpos celestes. Fue él quien utilizó por primera vez el término correlación en un estudio presentado en 1846 en la Academia de Ciencias en Francia. Aunque Pearson (1965) indica que Bravais no consideraba

variables aleatorias correlacionadas, sino consideraba errores independientes unos de otros; por tanto, no llegó a una verdadera idea de la correlación, tal como hoy la conocemos.

Otro avance en el desarrollo de estos conceptos de correlación y regresión lineal se origina mediante el estudio conjunto de la variación de dos medias realizado por Francis Galton (1822-1911). A lo largo del periodo 1865-1890, sus principales intereses fueron los estudios empíricos de las leyes de la herencia por medio de métodos estadísticos. Desarrolla sus propios y rudimentarios métodos para describir observaciones univariadas y bivariadas normalmente distribuidas, explicando la utilidad y el significado de la regresión y correlación, no solamente en el contexto de la herencia, sino en general. Como no dominaba con soltura la matemática de su tiempo utiliza artificios mecánicos para “probar” las propiedades de la distribución binomial como el que llamó “quincux”, hoy conocido como aparato de Galton, ver figura 3.

Figura 3. Ilustración del quincux



Fuente: Hald, (1998, p. 606)

Las ideas modernas sobre regresión se originan en los trabajos de Adrián - Marie Legendre y Carl Friedrich Gauss, sobre el método de mínimos cuadrados, para ajustar los datos sobre las órbitas de cuerpos celestes. El primer estudio documentado sobre el método de mínimos cuadrados, de donde deriva la idea de regresión, es debido a Legendre (1752-1833) en 1805.

Esta técnica de optimización intenta encontrar una la función de todas las posibles la que mejor se ajusta a los datos bivariantes, de acuerdo con el criterio de mínimo error cuadrático, y siempre que los datos cumplan con la condición de independencia entre otras.

Karl Pearson, continuando con los trabajos de Sir Francis Galton, interesado en la Biometría debido a los trabajos de éste y de Walter Frank Raphael Weldon, quien en 1892 fue el primero que publicó un artículo en el que señala el significado de un coeficiente de correlación negativo, indicando que era posible determinar una razón (coeficiente), cuyo valor se convierte en  $\pm 1$  cuando un cambio en cualesquiera de los órganos implica un cambio igual en el otro, y 0 cuando los dos órganos son bastante independientes. Pearson también había resuelto las propiedades matemáticas del coeficiente de correlación y la regresión simple utilizada para la predicción lineal entre dos variables continuas, generalizando así la idea de regresión de Galton a varias dimensiones, comprobando que muchas variables biométricas siguen la distribución normal multivariante y por tanto se podría aplicar la regresión múltiple para relacionar varias de estas medidas y defiende que el método científico es esencialmente estadístico, sus inferencias se basan en la asociación entre antecedente y consecuente.

En síntesis, el surgimiento de la correlación y la regresión lineal responde a la necesidad de resolver problemas científicos en los que la simple observación cualitativa o la dependencia funcional determinista resultaban insuficientes para explicar la variabilidad presente en fenómenos reales. Desde los estudios antropométricos de Quetelet, donde se buscaba comprender cómo múltiples características podían influir en el “hombre medio”, hasta las dificultades astronómicas enfrentadas por Bravais en la medición de cuerpos celestes, la estadística comenzó a enfrentarse al reto de describir relaciones no exactas entre cantidades sujetas a error. Galton profundizó este desafío al estudiar la herencia biológica, mostrando que, aun en procesos regidos por leyes naturales, la variación conjunta entre dos variables exhibe patrones de dependencia que no son estrictamente causales.

Posteriormente, los desarrollos matemáticos de Legendre, Gauss y Pearson permitieron formalizar métodos para cuantificar esta relación bajo criterios de optimalidad y variabilidad aleatoria, revelando que la asociación entre variables debía entenderse no como una correspondencia perfecta, sino como una tendencia sujeta a incertidumbre. Así, los conceptos de correlación y regresión lineal emergen históricamente como respuestas a problemas concretos de medición, predicción y explicación en contextos de error y variabilidad, consolidándose como herramientas fundamentales para modelar patrones en datos reales y para sustentar procesos de inferencia científica en diversos campos del conocimiento.

### 3 ESTADO DEL ARTE

Como se ha mencionado previamente, los primeros trabajos sobre la comprensión de la correlación no fueron desarrollados propiamente dentro del estudio de la educación estadística como disciplina, sino que, se desarrollaron en la psicología. El interés principal de esos estudios es determinar los modos de pensar de las personas acerca de la covariación entre dos variables o conjuntos de datos. Para la conformación del estado del arte, se centra la atención en los estudios realizados en educación estadística referidos a la comprensión de dichas nociones entre estudiantes de distintos niveles educativos, los que describiremos en los siguientes apartados.

#### 3.1 Contribuciones a nivel local

En los últimos años, la literatura en educación estadística ha destacado la complejidad cognitiva y conceptual que enfrentan los estudiantes universitarios al comprender las nociones de *covariación, correlación y regresión lineal*. Estas dificultades suelen manifestarse en la interpretación de gráficos de dispersión, en la articulación entre representaciones algebraicas y visuales, y en la comprensión del sentido de ajuste lineal más allá del cálculo mecánico (Bakker & Gravemeijer, 2004; Garfield & Ben-Zvi, 2008; Groth, Jones & Knaub, 2018). En consecuencia, se ha reconocido la importancia de promover enfoques didácticos que integren *el razonamiento covariacional, la exploración de datos y el uso de tecnologías digitales* como medios para construir significado y fortalecer el pensamiento estadístico.

En este contexto, diversos estudios han propuesto intervenciones orientadas a superar dichas dificultades, combinando la visualización, la simulación y la experimentación con datos reales. Entre ellos destacan las propuestas de Espino (2014) y Fernández (2018), quienes diseñaron estrategias de enseñanza dirigidas a estudiantes universitarios para favorecer la comprensión de *la correlación y la regresión lineal* mediante el uso de software estadístico. Estos trabajos coinciden en que el aprendizaje de la regresión lineal se potencia cuando los estudiantes participan activamente en la construcción y análisis de representaciones gráficas, estableciendo conexiones entre los datos, los modelos y los contextos de aplicación.

El estudio realizado por Espino se implementó con estudiantes de nivel superior de la Universidad de Nayarit, quienes cursaban la carrera de Psicología. Su objetivo principal fue ayudar a subsanar las dificultades didácticas y cognitivas que impiden la adquisición significativa de conceptos estadísticos abstractos, promoviendo un acercamiento intuitivo, la reflexión, la interacción entre pares y el uso de herramientas tecnológicas, como applets en GeoGebra. Por lo tanto, se desarrollaron cuatro secuencias didácticas para la correlación y regresión lineal, que permitieron al alumno realizar un análisis y una autorreflexión en su quehacer estadístico mediante el manejo de datos reales en los contextos: El camino y las localidades, los cigarrillos son causa de enfermedad, más autos implica un mayor número de muertos y la parte dos de los cigarrillos son causa de enfermedades. Dichas actividades didácticas basadas en situaciones problema de interés resultaron cruciales para conectar los

conocimientos previos de los estudiantes, fomentar la evolución de sus ideas sobre los conceptos y facilitar la adquisición de nuevos conocimientos.

En particular, el uso de datos reales o ficticios, pero con *un arreglo no tradicional*, motivó a los alumnos a realizar adaptaciones y tratamientos que no esperarían en un curso convencional, lo que consolidó la importancia de los conceptos en situaciones reales. Además, según el autor, GeoGebra facilitó la visualización dinámica de conceptos matemáticos y estadísticos al mostrar múltiples representaciones vinculadas en tiempo real (gráfica, numérica, tabular, algebraica) de un mismo objeto.

Por su parte Fernández presenta una propuesta didáctica para promover un acercamiento al estudio de la relación entre dos variables estadísticas, la cual está dirigida a estudiantes de nivel superior del área de Ciencias Sociales de la Universidad de Sonora. Uno de los objetivos de la propuesta fue incidir en el acercamiento intuitivo y razonamiento de la correlación lineal. Para lograr el objetivo se consideró importante incluir en la propuesta el manejo y uso de distintas representaciones de la correlación lineal como son: representación numérica, representación gráfica y representación verbal, ya que la investigación estadística ha demostrado que el uso de estas diferentes representaciones desarrolla las habilidades en los estudiantes potencializando su razonamiento covariacional. En esta secuencia de actividades se utilizó Microsoft Excel con el objetivo de manipular los datos y agilizar los cálculos estadísticos, como también para la construcción de gráficos estadísticos apropiados.

Se destaca que una característica común en ambos proyectos es la utilización de los elementos de la metodología ACODESA (Hitt, 2009) tanto para el diseño de las actividades didácticas como para su implementación en el aula de clases. Así, se concluye que la utilización de proyectos de intervención didácticas que integran el uso de tecnología, como GeoGebra y/o Microsoft Excel, junto con metodologías activas como ACODESA, aportan a la comprensión de conceptos estadísticos abstractos como la correlación y la regresión lineal en estudiantes universitarios; favoreciendo un aprendizaje más significativo mediante la conexión con datos reales y situaciones problemáticas, el uso de múltiples representaciones y la reflexión crítica, lo que potencia el razonamiento estadístico y facilita la visualización dinámica de las relaciones estadísticas en contextos reales.

### **3.2 Contribuciones a nivel nacional**

En el ámbito nacional, diversas investigaciones han buscado fortalecer la enseñanza y el aprendizaje de la estadística a través de propuestas didácticas innovadoras que incorporan tecnologías digitales y entornos interactivos. Estas contribuciones han permitido explorar nuevas formas de representar, analizar e interpretar datos, favoreciendo el desarrollo del razonamiento estadístico y covariacional en distintos niveles educativos. En particular, algunos estudios se han centrado en el análisis de datos bivariados y en la comprensión de las relaciones entre variables, integrando recursos tecnológicos que promueven la visualización dinámica y la experimentación.

En este contexto, destaca el trabajo de Inzunza (2016) quien reporta los resultados de una investigación sobre el análisis de datos bivariados con estudiantes de nivel superior de la carrera de Informática, quienes cursaban una asignatura introductoria de probabilidad y estadística en la Universidad de Sinaloa. La implementación de la secuencia didáctica se llevó a cabo en un entorno que integró medios digitales, como applets y software dinámico como Fathom para el análisis de datos. El objetivo principal del estudio consistió en explorar el potencial de un ambiente computacional para la enseñanza y el aprendizaje de datos bivariados, con el propósito de comprender cómo el uso de estos recursos puede mejorar dichos procesos.

Algunos de los hallazgos reportados indican que los estudiantes, al inicio, mostraban un bajo nivel de razonamiento intuitivo sobre la covariación en los diagramas de dispersión; sin embargo, mediante el uso de herramientas tecnológicas lograron identificar la dirección e intensidad de las relaciones entre variables. No obstante, la interpretación de los coeficientes de regresión y del coeficiente de determinación continuó representando un desafío. El estudio resalta el potencial de la visualización gráfica y dinámica que ofrece la tecnología, por encima de los cálculos simbólicos, en el aprendizaje de la estadística. En conclusión, aunque el análisis de datos bivariados constituye un proceso complejo, el uso de recursos computacionales puede convertirse en un elemento clave para integrarse en el diseño e implementación de propuestas didácticas efectivas.

El estudio de Medina (2021), centrado en la investigación del razonamiento covariacional de estudiantes de bachillerato al abordar temas de correlación y regresión lineal, se exploran las dificultades conceptuales y las estrategias que emplean los alumnos, tanto con métodos tradicionales donde se le da privilegio al lápiz y al papel como con la asistencia del software Fathom. Los hallazgos de las dos etapas de la investigación permitieron categorizar el razonamiento estudiantil en niveles ingenuo, embrionario y transicional. A partir de estos resultados, se propone una Trayectoria Hipotética de Aprendizaje (THA) con lecciones específicas diseñadas para fomentar una comprensión más profunda de estos conceptos estadísticos, destacando el papel crucial de la tecnología para visualizar y experimentar con los datos.

La trayectoria hipotética está determinada por cuatro lecciones, en las que se exploran distintos aspectos de la correlación y la recta de mejor ajuste:

1. Explora la relación entre la disposición de un conjunto de puntos y la correlación y tiene como objetivo: que los estudiantes asocien la disposición de una nube de puntos con su coeficiente de correlación, observando la tendencia y la fuerza de la relación;
2. La segunda lección, la relación entre el coeficiente de correlación, el centro de gravedad de la nube de puntos y la recta de mínimos cuadrados tiene como objetivo: que los estudiantes identifiquen la relación entre el coeficiente de correlación, el centro de gravedad de una nube de puntos (donde cada coordenada es la media aritmética) y la posición de la recta de mínimos cuadrados;

3. La tercera es, recta de ajuste y la recta de mínimos cuadrados, tiene como objetivo: que los estudiantes tracen una recta de ajuste y comparen características como posición y variación con respecto a la recta de mejor ajuste y:
4. Por último, el carácter predictivo/estimación del modelo de mínimos cuadrados y el objetivo fue: que los estudiantes realicen una estimación de la variable respuesta (Y) para valores dados de la variable explicativa (X)

Finalmente, algunas conclusiones importantes que se pueden considerar son: el diagrama de dispersión ayuda a dirigir la atención de los estudiantes hacia aspectos globales de los datos, permitiéndoles percibir si los puntos están alineados, su tendencia o cercanía, aunque les resulta un desafío traducir estas propiedades generales a números o ecuaciones específicas; El estudio sugiere que vincular los aspectos generales de los datos bivariados (dirección y fuerza de la correlación) con aspectos específicos (coeficiente de correlación numérico) es clave y en problemas de ajuste de una recta, la misma tendencia persiste: buscan una ley lineal que abarque todos los puntos, pero, al no encontrarla, seleccionan un subconjunto de puntos por los que pueda pasar una recta.

Por su parte, Morgado y Sánchez (2022 y 2024) investigan el razonamiento estadístico covariacional en estudiantes de bachillerato, particularmente en temas de correlación y regresión lineal, con el apoyo de tecnología digital. La investigación se basa en un experimento de diseño que explora cómo los estudiantes desarrollan la comprensión de conceptos como la noción de agregado, el modelo aleatorio lineal y la incertidumbre. Las actividades diseñadas con las plataformas CODAP (*Common Online Data Analysis Platform*) y GeoGebra, y presenta modelos de desarrollo del razonamiento. El estudio también analiza las concepciones previas y dificultades de los estudiantes, y el papel de la tecnología y el contexto en su aprendizaje.

Entre los resultados del estudio se menciona que los estudiantes de bachillerato enfrentan varios desafíos conceptuales principalmente al estudiar temas de correlación y regresión lineal, especialmente antes de recibir instrucción formal. Estos desafíos a menudo se relacionan con sus concepciones previas de las matemáticas y sus interpretaciones de los datos. El uso de la tecnología digital influye significativamente en el desarrollo del razonamiento covariacional de los estudiantes de bachillerato, al hacer que las estadísticas sean más visuales, interactivas y dinámicas, lo que permite un enfoque más centrado en los conceptos que en los algoritmos y cálculos. Esto invita a los estudiantes experimentar con los datos y relacionar múltiples representaciones (visuales, simbólicas, numéricas y verbales) de forma simultánea, facilitando así el descubrimiento y la observación de patrones y tendencias en los datos.

Los autores concluyen que la incorporación de tecnologías digitales en la construcción y análisis de diagramas de dispersión representa un avance significativo en la enseñanza de la estadística. Estas herramientas permiten a los estudiantes identificar visualmente la dirección y tendencia lineal de los datos, aunque inicialmente esta percepción pueda ser insuficiente para considerar plenamente la variabilidad inherente y el error en la relación entre variables. Este aspecto

subraya la importancia de complementar la observación gráfica con la interpretación numérica de los cálculos estadísticos, como el coeficiente de correlación, para enriquecer la comprensión de la fuerza y la dispersión de la relación, aunque se requiere reforzar la conexión entre el signo del coeficiente y la dirección de los datos.

Finalmente, las actividades diseñadas permitieron a los estudiantes reconocer la incertidumbre en situaciones específicas: la variabilidad en datos que no se ajustan perfectamente a funciones, el reconocimiento de la tendencia y fuerza de los datos desde una perspectiva estocástica, y la comprensión del criterio de mejor ajuste como una minimización del error aleatorio. Estos avances conceptuales son fundamentales para superar los razonamientos deterministas iniciales y promover una visión estadística más completa y crítica.

En su conjunto, las investigaciones revisadas coinciden en resaltar el papel fundamental de las tecnologías digitales en la enseñanza y el aprendizaje de la estadística, particularmente en la comprensión de la correlación y la regresión lineal. Los estudios anteriores ponen en evidencia que los entornos dinámicos e interactivos favorecen la visualización, la experimentación y la exploración de la relación entre variables, promoviendo un avance del razonamiento intuitivo e informal hacia formas más analíticas y conceptuales de pensamiento estadístico. No obstante, también muestran que la interpretación de los coeficientes de regresión, la comprensión de la variabilidad y el uso de modelos de predicción continua representando desafíos significativos para los estudiantes, incluso cuando se emplea tecnología.

### **3.3 Contribuciones a nivel internacional**

El trabajo de Estepa y Sánchez-Cobo (2003), con estudiantes universitarios de diferentes perfiles, tiene como objetivo identificar las concepciones de los estudiantes al finalizar un curso de estadística. Los resultados evidencian que, aunque una mayoría logra calcular correctamente el coeficiente de correlación, menos de la mitad interpreta adecuadamente la recta de regresión o distingue entre la variable dependiente e independiente. El estudio subraya además la necesidad de reforzar la enseñanza de la regresión y la correlación mediante situaciones didácticas contextualizadas, que promuevan la integración entre los conceptos matemáticos y estadísticos y favorezcan una comprensión relacional y no mecánica de estos temas.

En el trabajo de Watkins et al. (2004), se describen seis características de una relación de datos bivariados que puede ser utilizadas para promover la examinación, argumentación, comprensión y razonamiento sobre los datos, estas giran en torno al análisis de: a) las variables individuales y su variabilidad; b) la forma de los datos mediante la construcción de un gráfico de dispersión en función de su linealidad, los grupos y los valores atípicos, c) la tendencia, en caso de tenerla (positiva o negativa), d) la fuerza de la tendencia (fuerte o débil, variable o constante), e) la generalización del patrón, de ser posible y f) las explicaciones para el patrón.

Garfield y Ben-Zvi (2008) presentan una secuencia de actividades para desarrollar la comprensión de la correlación y la regresión lineal partiendo de las ideas intuitivas e informales. Además, en las actividades propuestas no solo se razona acerca de las ideas antes mencionadas

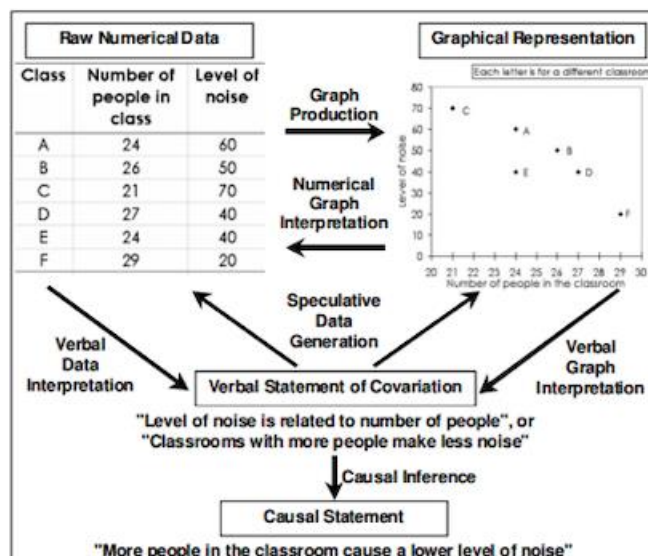
sino también, en todos los conceptos implicados en los procesos de enseñanza aprendizaje de la regresión lineal y la correlación, por ejemplo: los estudiantes razonan acerca de los datos, los procesos llevados a cabo al recabar información, las gráficas de dispersión, de la misma manera, interpretando y comprendiendo las características de los datos bivariados.

Además, consideran que, el tema de la relación entre variables puede abordarse de manera informal cuando se introduce la unidad relacionada con la recopilación y producción de datos, y que los gráficos de datos bivariados pueden presentarse de forma preliminar durante el análisis de datos. Sin embargo, el estudio formal de los datos bivariados es más apropiado realizarlo después de abordar la inferencia estadística. Esta recomendación se basa en que la covariación estadística integra y profundiza múltiples conceptos estudiados previamente, lo que permite retomar y afianzar ideas esenciales hacia el final del curso introductorio.

Otras investigaciones importantes, referencias de las investigaciones antes mencionadas, son las que realizaron Moritz (2004) y Zieffler y Garfield (2009). En la primera investigación se indaga sobre las habilidades de razonamiento que desarrollan estudiantes entre los 8 y los 15 años sobre la covariación estadística, entendida como “la correspondencia de variación de dos variables estadísticas que varían sobre escalas numéricas” (Moritz, 2004, p. 228). Estas habilidades se presentan en la figura 2 como conexiones entre los registros de representación:

- La generación especulativa de datos (transformar una afirmación verbal en un gráfico),
- La interpretación verbal del gráfico y de la tabla numérica (traduciendo un diagrama de dispersión en una declaración verbal).
- La interpretación numérica del gráfico (lectura de valores o interpolación),
- La construcción del gráfico.

Figura 4. Formas de representar la covariación estadística



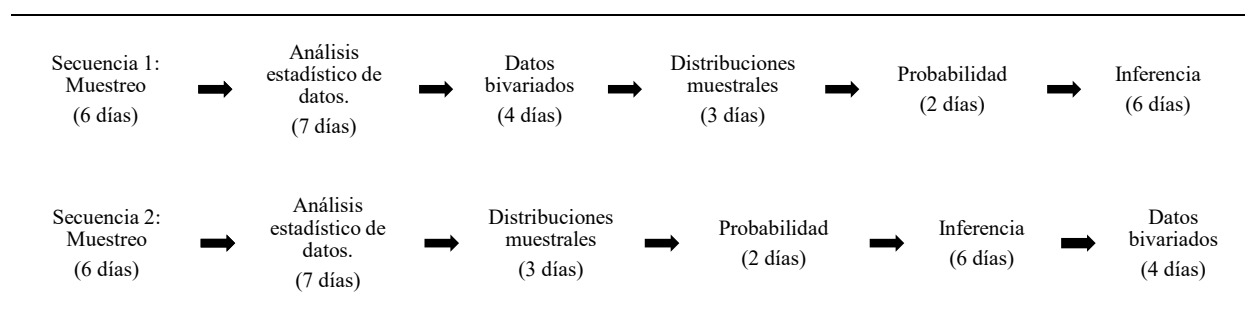
Fuente: Moritz, (2004, p. 230)

Los resultados son descritos en cuatro niveles de razonamiento, *no estadístico* (los estudiantes no apoyan la evidencia con el conjunto de datos), *aspecto único* (los estudiantes solo muestran valores aislados de datos bivariados o una sola variable), *covariación inadecuada* (en la interpretación verbal de las gráficas se presenta cuando los estudiantes hacen referencia a las dos variables, pero no generalizan la comparación utilizando dos o más puntos y se describen las variables, pero la correspondencia no se menciona o está en la dirección incorrecta) y *covariación apropiada* (cuando los estudiantes representan ambas variables y les dan la correspondencia adecuada, son clasificados en el nivel que utilizan razonamiento de covariación apropiada).

En el estudio de Zieffler y Garfield (2009), se estudia el desarrollo del razonamiento de los estudiantes sobre los datos cuantitativos bivariados, durante un semestre en un curso de estadística con estudiantes de nivel universitario. La investigación se propone indagar sobre el patrón de cambio del razonamiento de los estudiantes a lo largo del curso, la influencia de la secuencia de datos bivariados en los patrones de cambio en el razonamiento de los estudiantes y el análisis de los cambios de razonamiento en los conceptos fundamentales de distribución (si están asociados en los cambios en los patrones de razonamiento). Para ello, se diseñan actividades con las siguientes secuencias (Figura 3).

La característica que diferencia las dos secuencias propuestas en el estudio fue la ubicación del tema referente a los *datos bivariados*. En la primera secuencia se ubicó después del tema de análisis estadístico de datos, antes de los temas de distribuciones muestrales, probabilidad e inferencia. En la secuencia dos el tema de datos bivariados se ubicó al final del curso cuando los estudiantes tenían nociones de distribuciones muestrales, probabilidad e inferencia.

Figura 5. La distribución de las secuencias



Fuente: Zieffler & Garfield, (2009, p. 13).

Los autores concluyen que, al no encontrar diferencias o influencias sobre el orden de los temas en las dos secuencias, este tópico puede ser localizado en cualquier parte del curso y queda a criterio del docente o de los autores de libros de texto. Además, a pesar de las limitaciones encontradas en el estudio, los resultados sugieren algunas implicaciones prácticas para los profesores de un curso introductorio de estadística. por ejemplo, los resultados sugieren que es necesario dedicar muchas horas en el desarrollo del razonamiento de los estudiantes sobre la

distribución univariada para proporcionar bases sólidas que consecuentemente fortalecerán el razonamiento de datos bivariados.

La investigación de Sorto et al. (2011) se realizó con estudiantes universitarios y fue uno de los primeros estudios sobre las concepciones de los estudiantes acerca de la interpretación de la línea de mejor ajuste cuando se les pide dibujar a lápiz y papel en diferentes diagramas de dispersión, esto antes de que se les enseñe el método de mínimos cuadrados. Los autores encontraron que las ideas iniciales de los estudiantes no se centran en minimizar la distancia entre los puntos y la línea, sino más bien en conceptos de centro o tendencia como dividir los puntos a la mitad o usar promedios.

Los investigadores categorizaron las concepciones iniciales de los estudiantes y establecieron las siguientes categorías:

- Generalización del concepto de Mediana: Trazar una línea que divide los puntos de datos a la mitad en el sentido de que la mitad de los puntos están sobre la línea y la mitad están debajo de la línea.
- Generalización del concepto de media: Trazar la línea a través de puntos que son puntos medios de diferentes grupos de puntos.
- Ajustar la tendencia de los datos: Trazar la línea que une el primer y último punto y trazar "la línea está más estrechamente relacionada con todos los puntos y la tendencia de los datos".

La conclusión más relevante del estudio es que el método de mínimos cuadrados para ajustar una línea no emerge de manera espontánea en el razonamiento inicial de los estudiantes. Además, el criterio de minimizar distancias para determinar la mejor recta de ajuste no forma parte de sus concepciones intuitivas, lo que evidencia la complejidad cognitiva que implica este concepto. En su lugar, los estudiantes tienden a apoyarse en estrategias más intuitivas, como ubicar la línea en el centro del conjunto de datos, dividir visualmente los puntos, considerar el promedio de los datos bivariados o estimar una tendencia general.

En un trabajo reportado por Casey (2014) se describen los resultados de tres investigaciones acerca de la enseñanza y el aprendizaje de la regresión lineal. El autor afirma que en muchos casos la regresión lineal representa el primer acercamiento que los estudiantes tienen con conceptos fundamentales de la asociación estadística y la dependencia. En el reporte se presenta una trayectoria hipotética de aprendizaje (THA) para la regresión lineal, creada a través de un proceso iterativo en el que interactuaron un grupo de practicantes y un equipo de investigadores. La primera fase de la THA se enfoca en el aprendizaje de la recta informal de mejor ajuste, la segunda fase en la recta de mínimos cuadrados, el concepto de correlación es la base de la tercera fase, las ecuaciones de regresión para muestras se estudian en la fase cuatro y, la última fase se enfoca en la distribución muestral para la pendiente de la recta de regresión de una muestra.

La conclusión principal del trabajo es la descripción del conocimiento que deben tener los profesores acerca de las concepciones y la evolución que tienen los estudiantes sobre regresión lineal; algunas de estas conclusiones son:

1. Dado que la determinación informal de la recta de mejor ajuste es un proceso visual, una necesidad de los estudiantes es convertir datos recopilados en tablas en una representación gráfica (diagrama de dispersión).
2. Los estudiantes tienen dificultad en identificar cuál es la variable dependiente y en asignar las escalas adecuadas para los ejes en las gráficas. Además, presentan confusión al analizar el comportamiento de los datos, por ejemplo, las funciones lineales crecen o decrecen de manera monótona, lo que no ocurre necesariamente con datos que tienen una asociación lineal.
3. Muchos estudiantes al trazar la recta de ajuste de manera informal enfocan su atención en algunos puntos de la gráfica como: el primero, el último, el más alto o bajo, o en un subconjunto de puntos, sin darle importancia al resto de los puntos al analizar su tendencia.
4. Los estudiantes igualmente tienen dificultad al observar la tendencia global de un conjunto de datos, cuando leen una gráfica de dispersión, partiendo del hecho de que enfocan su atención en puntos aislados y perciben los datos como una serie de casos individuales, en lugar de considerarlos holísticamente (con características invisibles para puntos aislados).

Otros estudios con este mismo enfoque son los presentados por García y Yáñez (2018) y Groth (2018). En el primer estudio se exploró el razonamiento covariacional informal (RCI) de estudiantes de octavo grado en relación con la recta de mejor ajuste en diagramas de dispersión. En este estudio se aplicaron dos pruebas a grupos de estudiantes sin instrucción previa en análisis de datos bivariados para evaluar su capacidad de establecer relaciones entre variables, predecir valores y determinar los criterios que usan para trazar la recta de mejor ajuste. Los hallazgos confirman estrategias y concepciones previas, como la tendencia a buscar un patrón de asociación incluso cuando no existe, y la utilización de criterios como la proximidad o alineamiento de puntos para el ajuste de la recta.

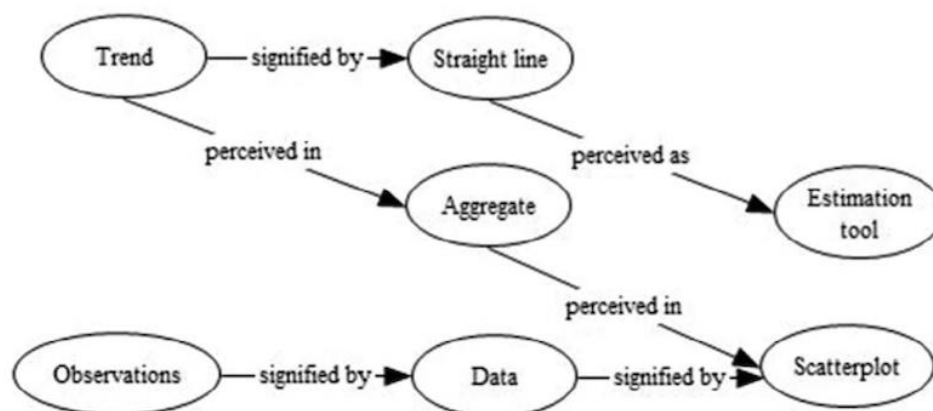
El estudio contribuye a la comprensión del RCI actual de los estudiantes y sugiere implicaciones para el diseño de actividades de aprendizaje informal en estadística. *Una de ellas se relaciona con el manejo del contexto*, los resultados muestran que el tipo de contexto influye directamente en el tipo de razonamiento que los estudiantes ponen en juego; por ello, se recomienda evitar situaciones altamente familiares que puedan inducir interpretaciones causales espontáneas. Otra implicación se refiere a la comprensión del criterio de “*cercanía*”, fundamental para interpretar la recta de mejor ajuste. Se observó que los estudiantes no siempre comprenden que la mejor recta es aquella que minimiza la distancia entre los puntos y la línea.

En síntesis, reconocer los errores, intuiciones y formas de razonamiento que los estudiantes movilizan al analizar datos bivariados constituye el punto de partida para diseñar trayectorias de aprendizaje más efectivas. Estas consideraciones no solo orientan la secuencia de actividades, sino que permiten construir puentes entre el razonamiento inicial y el pensamiento covariacional que se espera desarrollar. Diseñar actividades de aprendizaje informal en estadística es, en este sentido, similar a construir un puente apoyado en los escalones irregulares que los estudiantes ya han colocado. En lugar de ignorar esos peldaños o intentar sustituirlos de inmediato, se aprovechan sus intuiciones y se añaden nuevos escalones estratégicos para guiar su tránsito hacia una comprensión estadística más sólida y formal.

Por su parte Groth, et al. (2018), propone un marco para caracterizar los procesos cognitivos de los estudiantes al trabajar con la noción de recta de mejor ajuste informal en el análisis de datos bivariados. El trabajo emplea la metodología de diseño basada en la investigación, examinando cómo los estudiantes procesan información al significar observaciones como datos, percibir agregados en diagramas de dispersión y finalmente, usar líneas rectas como herramientas de estimación.

El marco representado en la figura... detalla un conjunto de procesos interconectados que se construyen a partir del análisis semiótico.

Figura 6. Un proceso hipotético cognitivo para trabajar con la línea de mejor ajuste



Fuente: Groth et al., (2018, p. 253)

El marco incluye seis procesos esenciales del pensamiento del estudiante, que funcionan como los nodos o elementos principales:

1. Significación de observaciones como datos; este proceso es la capacidad del estudiante para codificar observaciones empíricas (como las de un experimento) en datos, típicamente registrándolos en tablas. También incluye la capacidad de inferir las circunstancias contextuales que produjeron los datos, yendo de los datos al contexto, lo cual es un componente esencial del pensamiento estadístico.

2. Significación de datos con diagrama de dispersión; este elemento implica que los estudiantes sean capaces de leer los valores de los puntos en los diagramas de dispersión y de construir los gráficos. Sin embargo, la dificultad a menudo reside en manejar escalas inusuales (que requieren razonamiento proporcional) y en el sentido gráfico, es decir, decidir cuando un diagrama de dispersión es la representación más útil para un conjunto de datos bivariados.

3. Percepción de agregados en diagramas de dispersión; superar la tendencia a ver los puntos de datos solo como casos individuales es fundamental para poder realizar inferencias estadísticas. La percepción de agregados puede ocurrir de dos maneras: percepción de dentro hacia afuera, reconocer características dentro del agregado, como cúmulos y brechas y percepción de fuera hacia adentro, ocurre cuando una característica fuera del cúmulo principal (como en valor atípico o una estimación irrazonable trazada) llama la atención hacia el agregado central de puntos.

4. Percepción de tendencias en agregados, este proceso implica discernir las características de la colección de datos. Los estudiantes utilizan verbalizaciones y gestos para indicar sus percepciones de tendencias, a veces de manera holística (por ejemplo, cuesta arriba o cuesta abajo), cuando la variabilidad es baja, la percepción del agregado y la percepción de la tendencia lineal pueden ocurrir casi simultáneamente. A medida que aumenta la variabilidad, estas percepciones se vuelven procesos más separados en el pensamiento del estudiante.

5. Significación de tendencias con líneas rectas, esta es la etapa en la que la tendencia percibida se representa formalmente. La línea recta debe servir como una abreviatura de la percepción de la tendencia lineal por parte del estudiante. Es crucial que la línea no sea un signo impuesto prematuramente que se dibuja solo para satisfacer un requisito de la tarea, sino que debe reflejar la percepción de la covariación lineal subyacente.

6. Percepción de líneas rectas como herramientas de estimación estadística, este proceso requiere una capa adicional de trabajo cognitivo donde la línea ajustada a los datos se percibe como una herramienta para hacer interpolaciones y extrapolaciones. Para ser una herramienta de estimación efectiva, la línea debe significar una tendencia global sobre todo el cuerpo de datos, no solo una porción local.

El marco también incluye procesos de desvío (presentados por flechas discontinuas en el diagrama) que puede desviar a los estudiantes del uso de la línea como herramienta de estimación global. Uso de tendencias locales, hacer predicciones basándose únicamente en los puntos de datos cercanos a la estimación solicitada, ignorando la tendencia global y Razonamiento incremental, analizar los cambios de una observación a la siguiente, a menudo visto al examinar tablas o al dibujar patrones en los diagramas de dispersión, lo cual puede entrar en conflicto con las estimaciones producidas por la línea de tendencia. Los autores argumentan que, aunque a menudo se etiqueta como informal o simple, la inferencia estadística visual implica procesos cognitivos complejos que merecen ser comprendidos y apoyados pedagógicamente.

En conjunto, este marco proporciona una visión articulada y profunda de los procesos cognitivos que intervienen en la construcción de la noción de la recta de mejor ajuste informal, mostrando que dicha comprensión no surge de manera inmediata, sino que requiere transitar por distintas formas de significar, percibir y representar la información. Al detallar cómo los estudiantes avanzan desde la interpretación de observaciones hasta el uso de la línea como herramienta de estimación global, el aporte de estos autores destaca la complejidad inherente al razonamiento estadístico visual y subraya la necesidad de diseñar experiencias de aprendizaje que apoyen explícitamente estos procesos. Así, este marco no solo enriquece la perspectiva teórica sobre el pensamiento sobre la covariación estadística y la modelación informal, sino que también ofrece orientaciones valiosas para la práctica docente en el análisis de datos bivariados.

Recientemente encontramos el estudio realizado por Alvarado et al., (2025), El cual examina la eficacia de metodologías activas, específicamente la creación de problemas, para fortalecer el razonamiento estadístico en 450 estudiantes de diversos programas de ingeniería civil. La investigación se basa en el modelo de pensamiento estadístico de Wild y Phannkuch (1999), evaluando las competencias de los estudiantes a través de un ciclo de investigación que incluye la formulación de preguntas e hipótesis, la recolección de datos y el análisis estadístico. Aunque se observó una actitud positiva y un alto rendimiento en la formulación de problemas y conclusiones, los estudiantes mostraron dificultades notables en el análisis estadístico univariante y bivariante, incluyendo la interpretación de medidas de dispersión y regresión lineal. Los resultados sugieren que este enfoque fomenta la autonomía y el pensamiento crítico, pero también subraya la necesidad de mayor apoyo en la aplicación del conocimiento técnico de estadística descriptiva.

Las investigaciones revisadas evidencian la complejidad cognitiva y didáctica que implica el aprendizaje de la regresión lineal y la correlación, presentando mecanismos que intervienen en el desarrollo del razonamiento covariacional estadístico. Dichos estudios muestran que este tipo de pensamiento no surge de manera espontánea, sino que se construye progresivamente a partir de experiencias de aprendizaje estructuradas, reflexivas y contextualizadas. Asimismo, diversas propuestas coinciden en la importancia de promover la exploración de datos reales y el uso de representaciones gráficas como medios para favorecer una comprensión significativa de la correlación y la regresión lineal.

En síntesis, la enseñanza de la regresión lineal y del razonamiento covariacional requiere de la fundamentación en experiencias de aprendizaje que integren la exploración, la reflexión y la formalización progresiva.

## 4 PROBLEMÁTICA Y OBJETIVOS

En este apartado se expone la problemática en la que se enmarca el presente proyecto de intervención didáctica respecto a la enseñanza y el aprendizaje de la regresión lineal y correlación. Se integra la información presentada en los antecedentes y el estado del arte con el propósito de contextualizar la situación educativa que motiva la propuesta. Con base en esto, se plantea el objetivo general del proyecto, junto con los objetivos específicos que estructuran y guían el desarrollo del trabajo.

### 4.1 Elementos desde el análisis de los errores y dificultades

La correlación y la regresión son herramientas de gran utilidad ya que permiten determinar si existe una relación entre dos o más variables aleatorias, además, posibilitan conocer la fuerza y la dirección de esa relación. Adicionalmente, pueden emplearse para estimar el valor de una variable a partir de la otra siempre y cuando esto sea posible (Gea, 2014). Estos conceptos constituyen la base de numerosos métodos estadísticos, ya que amplían y formalizan el concepto de dependencia funcional entre variables (Batanero, 2001).

A pesar de la utilidad de la regresión lineal y la correlación, que se ha argumentado en apartados anterior, la comprensión de estos conceptos genera múltiples dificultades. Las dificultades para comprender estos conceptos no son solo de carácter epistemológico o conceptual; por ejemplo, investigaciones en psicología han mostrado que la noción de la correlación no se desarrolla de forma espontánea, pues requiere un tipo de razonamiento abstracto ajeno al pensamiento intuitivo infantil e incluso adulto, que tiene a operar desde creencias previas, causalidades aparentes o patrones percibidos de manera subjetiva. A ello se suman las concepciones erróneas identificadas desde la educación estadística, *determinista, local, causal o unidireccional* que reflejan tensiones profundas entre la dependencia funcional aprendida en matemáticas y la dependencia estadística propia del análisis de datos reales.

Estas dificultades se amplifican en la enseñanza formal, donde la ambigüedad entre los roles de las variables, la interpretación de la recta de regresión, la relación entre sentido e intensidad de la correlación, o la tendencia a extrapolar sin fundamento muestran que los estudiantes enfrentan retos persistentes al articular herramientas estadísticas con razonamiento estadístico.

La mayoría de las investigaciones que han estudiado las concepciones de los estudiantes, concuerdan en la necesidad de un diseño adecuado de situaciones de enseñanza que ofrezca un camino para utilizar las concepciones previas con el objetivo de promover el razonamiento covariacional estadístico y que a su vez ayuden a superar las concepciones erróneas que se pueden presentar.

### 4.2 Elementos desde la didáctica de la educación estadística

A pesar de que, en la última década, ha aumentado el énfasis en que la enseñanza de la estadística debe centrarse más en la alfabetización, el razonamiento y el pensamiento estadístico. Uno de los principales argumentos es que los métodos tradicionales de enseñanza se enfocan en habilidades, procedimientos y cálculos, lo que no promueve que los estudiantes desarrollen un razonamiento o una forma de pensar estadística propia. Esta perspectiva impulsa un cambio hacia enfoques educativos que fomenten la comprensión profunda y crítica de los conceptos estadísticos en contextos relevantes para los estudiantes (Ben-Zvi y Garfield, 2004).

No obstante, el dominio de los procedimientos formales en la educación superior, especialmente en estadística, se ve a menudo obstaculizado por el fuerte énfasis en el conocimiento procedimental, lo que impide a los estudiantes alcanzar una comprensión conceptual sólida de métodos como la regresión y la correlación. Para contrarrestar esta laguna, que puede llevar a concepciones erróneas comunes como la interpretación de las relaciones como deterministas o causales, existe una tendencia en la investigación a introducir la asociación estadística de forma informal como un punto de partida para el desarrollo conceptual. El enfoque de la línea de mejor ajuste informal, definido como el ajuste visual de una línea a un diagrama de dispersión sin recurrir a cálculos o tecnología, es crucial para que los educadores y estudiantes construyan un puente entre la visión determinista de una función y una perspectiva estadística que aprecia el rol de la variación en el análisis de datos bivariados.

Por otro lado, se considera necesario el diseño de actividades orientadas al desarrollo del pensamiento estadístico requiere situar a los estudiantes frente a datos reales y contextualizados, ya que esto favorece la comprensión del papel de la estadística en la vida cotidiana y en la toma de decisiones. Trabajar con datos auténticos obliga a los aprendices a enfrentarse a la variabilidad, la incertidumbre y la complejidad propia de los fenómenos estudiados, lo cual fortalece la capacidad de interpretar evidencias de manera crítica. Diversos autores coinciden en que la autenticidad de los datos es clave para que los estudiantes comprendan la naturaleza de los problemas estadísticos y reconozcan la importancia del análisis en contextos reales (Franklin et al., 2007; Wild & Pfannkuch, 1999).

Asimismo, es imprescindible que las actividades promuevan procesos de indagación en torno al ciclo investigativo estadístico, de manera que los estudiantes puedan formular preguntas, recolectar datos, analizarlos e interpretar sus resultados para responder a problemas relevantes. La participación en todo el ciclo potencia la comprensión de la estadística como una herramienta para explorar fenómenos y no solo como un conjunto de procedimientos aislados. En este sentido, la literatura en educación estadística destaca que involucrar a los estudiantes en procesos completos de investigación aumenta su comprensión conceptual y promueve un pensamiento más integrado y flexible (Garfield & Ben-Zvi, 2008; Wild & Pfannkuch, 1999).

Además, las actividades deben fomentar la interpretación y argumentación basadas en evidencia, de modo que los estudiantes aprendan a justificar conclusiones a partir de datos y no únicamente de intuiciones. Este tipo de experiencias desarrolla habilidades críticas para evaluar afirmaciones y entender las limitaciones de los análisis. Como señalan Moore (1997) y Wild &

Pfannkuch (1999), la estadística tiene como función central apoyar el razonamiento basado en datos, por ello, es necesario promover entornos donde los estudiantes desarrollen la capacidad de explicar, defender y comunicar sus resultados.

Finalmente, es crucial que las actividades integren oportunidades para la comunicación estadística, tanto oral como escrita. Comunicar hallazgos obliga a organizar el pensamiento, clarificar ideas y reflexionar sobre la interpretación de los resultados. Además, contribuye al desarrollo de una ciudadanía capaz de comprender y producir argumentos estadísticos en contextos sociales, académico y profesionales. La comunicación es reconocida como una dimensión esencial del pensamiento estadístico y debe ser incorporada intencionalmente en el diseño de tareas (Garfield & Ben-Zvi, 2008; Wild & Pfannkuch, 1999).

### **4.3 Elementos desde el ámbito profesional**

Según Alvarado, et al. (2025), “La educación superior en ingeniería todavía enfrenta el desafío de promover metodologías activas que fomenten el desarrollo de ciclos de investigación en los que los estudiantes puedan formular preguntas, crear problemas contextualizados y asimilar conocimiento científico” (p. 1).

Al respecto, los docentes no cuentan con referencias apropiadas ya que en la literatura son pocas las propuestas didácticas mediadas con tecnología para el nivel superior con estudiantes de ingeniería. Por consiguiente, los docentes en su práctica utilizan los medios a su alcance, un medio de estos son libros de texto que el programa analítico propone. Que como se mencionó anteriormente, presentan un enfoque tradicionalista (se enfatiza en las definiciones, conceptos y los ejemplos promueven aspectos operativos, proponen escasa reflexión). Otra característica que las situaciones problema contienen pocos datos y las temáticas no son afines a la ingeniería. Estos rasgos del libro de texto parecen insuficientes y no estar alineados al enfoque por competencias que se propone en ITSON.

Todo lo anterior, justifica la pertinencia de este trabajo de tesis que pretende desarrollar una secuencia didáctica para ayudar la comprensión de la regresión y la correlación lineal (vinculados a la asociación y sus modelos) entre estudiantes universitarios. Ya que dicho tema es introductorio para herramientas más avanzadas que se utiliza en la ingeniería. Por lo que consideramos que nuestro trabajo abona a tener materiales didácticos acordes a las propuestas tanto nacionales e internacionales y que por ende ayudará a fortalecer el perfil de egreso de los futuros profesionistas que se formen en la institución.

### **4.4 Objetivos**

El objetivo general consiste *en diseñar una secuencia didáctica para favorecer la comprensión de la regresión lineal y la correlación con el apoyo de tecnología, dirigida a estudiantes de las carreras de ingeniería.*

Objetivos específicos:

- Realizar un análisis preliminar para identificar las necesidades en la enseñanza y el aprendizaje de la regresión y correlación.
- Diseñar actividades didácticas que promuevan un acercamiento informal sobre la regresión lineal y la correlación, con el apoyo de tecnología, como un puente hacia el acercamiento formal.
- Implementar las actividades didácticas con estudiantes de las carreras de ingeniería en el Instituto tecnológico de Sonora, utilizando un marco para la solución de problemas empíricos.
- Valorar los resultados de la implementación.

## 5 MARCO CONCEPTUAL

En este apartado se describen los elementos teóricos que fundamentan la propuesta intervención didáctica, cuyos elementos son relevantes para el diseño, implementación y valoración de la propuesta. En primer lugar, como marco global de la intervención, se toman elementos del marco para el desarrollo del pensamiento estadístico propuesto por Wild y Pfannkuch (1999), el cual incluye un ciclo investigativo para la solución de problemas, Problema, Plan, Datos, Análisis y Conclusión (PPDAC). Se considera este marco por su congruencia con la organización del programa del curso de la materia de Probabilidad y Estadística de ITSON (ITSON, 2023), donde se solicita el desarrollo de proyectos incluyendo planteamiento del problema, toma de datos, análisis y conclusiones.

En segundo lugar, se consideran aspectos sobre el análisis de datos bivariados de Watkins et al. (2007) en el registro gráfico, que se compaginan con las potencialidades ofrecidas por el uso de tecnología digital para ayudar a la visualización, agilizar el cómputo de estadísticos y la construcción de múltiples representaciones (Biehler, et al. 2013). El uso de la tecnología también se acompaña con los requisitos del curso de Probabilidad y Estadística, donde se solicita el uso de software (como Excel, R Studio) para recolección y análisis de datos.

### 5.1 El desarrollo del pensamiento estadístico para la investigación empírica

El Marco para el desarrollo del pensamiento estadístico de Wild & Phannkuch (1999) se enfoca en la resolución de problemas reales en estadística y consta de 4 dimensiones: la primera dimensión *el ciclo investigativo PPDAC*, la segunda dimensión *los tipos de pensamientos*, la tercera dimensión *el ciclo interrogativo* y la dimensión cuatro *las cualidades* (figura 7). Estas dimensiones interactúan entre sí para desarrollar el *pensamiento estadístico* implica un nivel de pensamiento superior al razonamiento estadístico.

El *pensamiento estadístico* es la forma en que piensan los estadísticos profesionales (Wild & Pfannkuch, 1999). Además, el *pensamiento estadístico* implica comprender cómo se generan y estructuran los datos, reconocer la variabilidad como un componente central de los fenómenos estudiados, y utilizar modelos estadísticos para describir, explicar y predecir comportamientos en contextos reales. Para Chance (2002), pensar estadísticamente requiere entender cómo y por qué se aplican determinados métodos y procedimientos, considerar las limitaciones inherentes a los procesos estadísticos, y emplear el contexto del problema para planear investigaciones, interpretar resultados y sustentar conclusiones fundamentadas. En esta perspectiva, el pensamiento estadístico supone una visión integral que combina el uso adecuado de herramientas estadísticas, el razonamiento probabilístico y la toma de decisiones informada dentro de situaciones de incertidumbre.

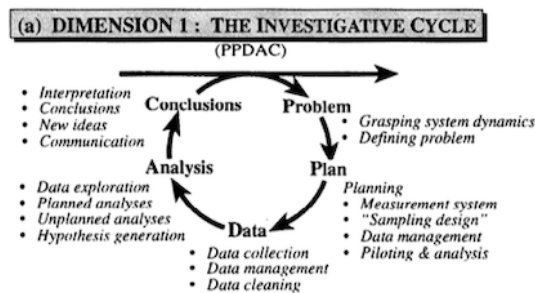
El *ciclo investigativo* PPDAC es la estructura operativa del pensamiento estadístico, es la guía de la actuación del estudiante, profesor o investigador. Wild y Pfannkuch muestran que el pensamiento estadístico implica tomar decisiones y gestionar razonamientos en todas las fases del ciclo. Por ejemplo, definir adecuadamente el problema puede determinar la pertinencia de los datos y la validez de las conclusiones, por lo que, el ciclo permite ver la estadística como un proceso dinámico, no como un conjunto de reglas o técnicas aisladas.

### 5.1.1 Dimensión uno: Ciclo investigativo

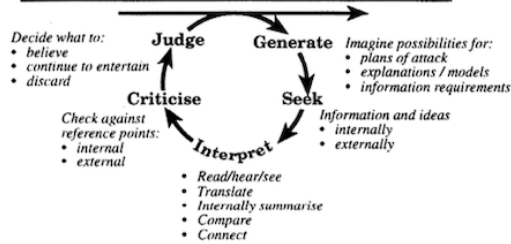
La etapa del *Problema* (P) tiene los siguientes elementos: *definición del problema* y *entender la dinámica del problema*. Para definir el problema estadístico es fundamental comprender el contexto en el que surge el problema, incluye conocer el fenómeno, las variables potenciales y las relaciones posibles. Si el estudiante no tiene suficiente información contextual, su definición del problema será vaga o incompleta. Además, identificar el propósito guía la formulación del problema y determina el tipo de análisis posterior, lo que resulta esencial para el planteamiento de preguntas específicas. Para formular un buen problema, el estudiante debe preguntarse *¿Qué información se necesita?*, *¿Qué datos ayudarían a responder la pregunta?* O *¿Qué características deben tener esos datos?*, abonando directamente en la precisión y relevancia del problema planteado; así también reconocer las restricciones y limitaciones del problema desde el inicio permite plantear problemas realistas y viables, contribuyendo a reducir el tiempo, garantizar el acceso a los datos, a los recursos o reconocer la variabilidad natural del fenómeno de estudio.

Al transitar por la etapa de problema, los estudiantes deben articular contexto, propósito, necesidad de la información, dinámica del fenómeno y limitaciones reales para poder definir un problema estadístico investigable.

Figura 7. Dimensiones del marco para el desarrollo del pensamiento estadístico



**(c) DIMENSION 3 : THE INTERROGATIVE CYCLE**



**(b) DIMENSION 2 : TYPES OF THINKING**

- GENERAL TYPES**
- **Strategic**
    - planning, anticipating problems
    - awareness of practical constraints
  - **Seeking Explanations**
  - **Modelling**
    - construction followed by use
  - **Applying Techniques**
    - following precedents
    - recognition and use of archetypes
    - use of problem solving tools
- TYPES FUNDAMENTAL TO STATISTICAL THINKING**  
(Foundations)
- **Recognition of need for data**
  - **Transnumeration**  
(Changing representations to engender understanding)
    - capturing "measures" from real system
    - changing data representations
    - communicating messages in data
  - **Consideration of variation**
    - noticing and acknowledging
    - measuring and modelling for the purposes of prediction, explanation, or control
    - explaining and dealing with
    - investigative strategies
  - **Reasoning with statistical models**
  - **Integrating the statistical and contextual**
    - information, knowledge, conceptions

**(d) DIMENSION 4 : DISPOSITIONS**

- **Scepticism**
- **Imagination**
- **Curiosity and awareness**
  - observant, noticing
- **Openness**
  - to ideas that challenge preconceptions
- **A propensity to seek deeper meaning**
- **Being Logical**
- **Engagement**
- **Perseverance**

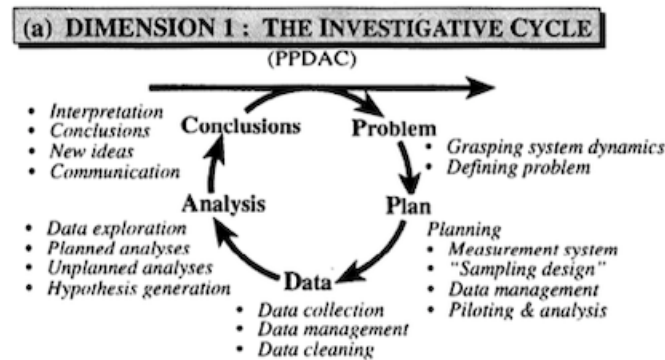
Fuente: Wild & Pfannkuch, (1999, p. 226)

La segunda etapa del ciclo investigativo es el *plan (P)*, el modelo propone en esta etapa al menos los siguientes elementos, *sistema de medición, el diseño del muestreo, gestión de los datos y el pilotaje y análisis previos*. El sistema de medición determina qué se va a medir (definir operacionalmente las variables), cómo se va a medir y con qué instrumentos (seleccionar herramientas válidas, confiables y adecuadas), además, es necesario prever los errores de medición y estandarizar del proceso. El diseño del muestro establece de qué o de quién se obtendrán los datos (fijar la población objetivo, seleccionar el tipo de muestro más adecuado en función de los objetivos, determinar el tamaño de la muestra con el fin de garantizar la representatividad y evitar sesgos en las conclusiones). En la gestión de datos es necesario registrar y organizar los datos (determinar los formatos de registro, definir los procedimientos de almacenamiento, controlar la calidad de los datos, respetar las consideraciones éticas con el manejo de la información confidencial y si es necesario codificar las variables). Los pilotajes son importantes ya que permiten evaluar la redacción de las preguntas, si se producen datos

útiles, se detectan errores y si es viable recabar la información y estos cumplen con las condiciones necesarias para su posterior análisis y de considerarse necesario ajustar el plan.

El desarrollo de la etapa de problema y de planificación resultan importantes, ya que en el modelo tradicional de enseñanza estas etapas son omitidas y simplemente en los problemas que se resuelven en clase los datos son proporcionados a priori y no es necesario discriminar la información a analizar ni la cantidad de datos.

Figura 8. Dimensión 1 ciclo investigativo (PPDAC).



Fuente: Wild & Pfannkuch, (1999, p. 226)

En la etapa de los *Datos* (D) se proponen los siguientes elementos, *la recolección de la información, la gestión de la información, la organización y limpieza para su posterior análisis*. La etapa de *datos* dentro del ciclo PPDAC es un momento crucial en el proceso investigativo, pues constituye el puente entre la planificación y el análisis, determinando en gran medida la calidad y la validez de las conclusiones posteriores. Transitar por esta etapa implica considerar diversos elementos que influyen tanto en la eficiencia del proceso como en la integridad de la información obtenida.

- A) La gestión de la información es fundamental, ya que permite organizar las fuentes, definir los formatos de registro y establecer criterios consistentes para el almacenamiento de los datos. Una gestión deficiente puede derivar en pérdida de información o en inconsistencias que comprometan el análisis.
- B) La recolección de la información constituye un proceso central, donde se garantiza que los datos obtenidos sean pertinentes, suficientes y coherentes con los objetivos definidos en la etapa de problema y con las decisiones tomadas en la fase del plan. La precisión en la recolección depende tanto de los instrumentos utilizados como del control sobre las condiciones del muestreo o del experimento.
- C) La necesidad de realizar una organización y limpieza de los datos, actividad indispensable para identificar valores atípicos, errores de digitación, omisiones o inconsistencias que pudieran distorsionar la interpretación, este proceso de depuración

contribuye a fortalecer la validez interna del estudio y prepara la base para un análisis más riguroso.

Finalmente, la etapa de datos también requiere reflexionar sobre la optimización de los datos entendida como el conjunto de acciones encaminadas a mejorar la estructura, formato y accesibilidad de la información. Esto puede incluir la transformación de las variables, la estandarización de escalas, la codificación adecuada o la integración de datos provenientes de diversas fuentes. Optimizar los datos nos solo facilita su análisis, sino que además genera conclusiones sólidas.

En la etapa de *Análisis de los datos (A)* se incluyen, *las exploraciones previas, los análisis planeados y los no planeados y la generación de hipótesis iniciales*. En esta etapa, los datos adquieren significado y se transforman en evidencia que sustenta o cuestiona las hipótesis plantadas, transitar por esta etapa implica considerar diversos elementos que influyen directamente en la calidad interpretativa del proceso. En primer lugar, las exploraciones previas cumplen un papel fundamental, pues permiten al investigador familiarizarse con el comportamiento general de los datos, identificar posibles patrones iniciales y reconocer anomalías que ameriten atención. Estas exploraciones tempranas constituyen una aproximación flexible y abierta que facilita la comprensión preliminar del conjunto de datos antes de aplicar métodos más formales.

Posteriormente, los análisis planeados representan el núcleo metodológico de la etapa, al corresponder con las técnicas seleccionadas durante la fase del plan, estos análisis se alinean con los objetivos del estudio y permiten evaluar las relaciones entre variables, contrastar hipótesis y generar descripciones o inferencias sustentadas en métodos estadísticos adecuados. Sin embargo, el proceso investigativo también da lugar a análisis no planeados, los cuales emergen de hallazgos inesperados o de preguntas nuevas que surgen a partir de la exploración inicial. Estos análisis inesperados o de preguntas nuevas que surgen a partir de la exploración inicial, aunque no previstos originalmente, pueden enriquecer la investigación al abrir nuevas líneas de interpretación o al revelar aspectos que no habían sido considerados en la planificación inicial.

Por último, la etapa de análisis incluye la generación de hipótesis iniciales, que surge como un proceso interactivo entre lo observado en los datos y los conocimientos previos del investigador. Este tipo de hipótesis, muchas veces de carácter exploratorio, guían la profundización en los análisis subsiguientes y orienta la búsqueda de explicaciones más robustas. En conjunto, estos elementos muestran que el análisis de datos no es un procedimiento lineal ni estrictamente técnico; por lo contrario, requiere flexibilidad, reflexión y una constante interacción entre los datos, los métodos y el razonamiento estadístico.

En la etapa final de Conclusiones (C) se encuentra *la interpretación de resultados, la generación de conclusiones en relación con el contexto, aparecen nuevas ideas y se comunican los hallazgos*. Representa el cierre reflexivo y articulado del proceso de indagación estadística,

transita adecuadamente por esta fase implica considerar diversos elementos que influyen de manera significativa en la calidad y profundidad de razonamiento estadístico que el estudiante o investigadores puede alcanzar. En primer lugar, la interpretación de resultados constituye un aspecto esencial, ya que demanda ir más allá de la información descriptiva numérica o gráfica para establecer el significado de los hallazgos en función del problema planteado. Esta interpretación requiere integrar los datos analizados con el conocimiento previo, identificando patrones relevantes y conocer las limitaciones inherentes del estudio.

Otro elemento fundamental es la generación de conclusiones en relación con el contexto, lo cual implica que los resultados no pueden entenderse de manera aislada, sino que deben situarse en el marco conceptual, social disciplinar o profesional desde el cual surge el problema. Es esta fase, el investigador debe valorar la pertinencia y alcance de las conclusiones, así como su coherencia con la evidencia obtenida, esta vinculación con el contexto permite que las conclusiones tengan utilidad real y que respondan de manera efectiva a la pregunta inicial que motivó la investigación.

Asimismo, durante esta etapa suelen emerger nuevas ideas, ya sea en forma de preguntas adicionales, hipótesis alternativas o líneas potenciales de investigación. Este surgimiento de nuevas perspectivas evidencia la naturaleza dinámica del pensamiento estadístico, que no se limita a cerrar un proceso, sino que propicia la apertura hacia nuevas reflexiones y posibilidades de estudio.

Finalmente, un componente central de esta fase es la comunicación de los hallazgos ya sea en informes escritos, presentaciones orales o visualizaciones. Comunicar implica seleccionar, organizar y expresar la información de manera clara y rigurosa, de modo que otros puedan comprender, evaluar y eventualmente replicar el estudio. Esta habilidad comunicativa es parte integral del quehacer estadístico, pues la utilidad del análisis depende en gran medida de la capacidad del investigador para transmitir adecuadamente lo que los datos revelan.

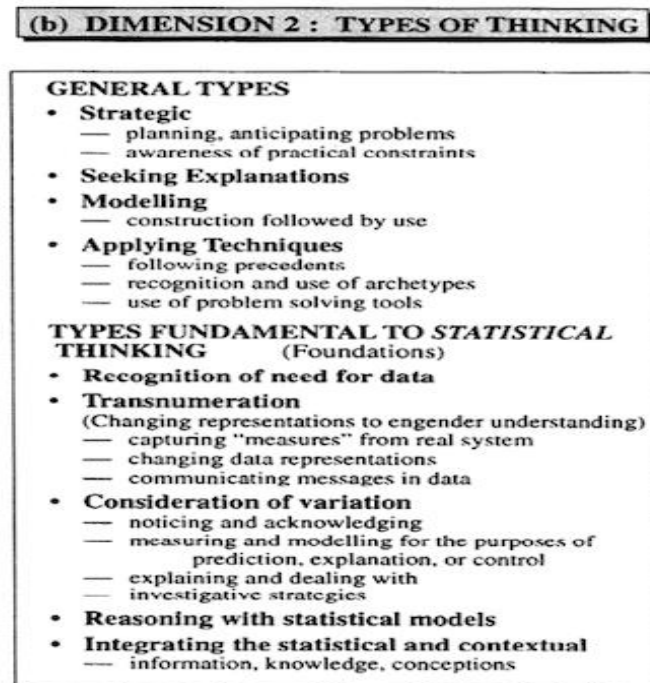
### **5.1.2 Dimensión dos: tipos de pensamiento**

La dimensión dos, los tipos de pensamiento incluye los tipos de pensamientos generales y los tipos fundamentales para el pensamiento estadístico. En los primeros tipos de pensamientos generales están el pensamiento estratégico que se refiere a como se describió en el Plan (Dimensión Uno), implica la planificación y la conciencia de las restricciones prácticas. Además, la búsqueda de explicaciones es otro tipo de pensamiento ya que se trata de generar explicaciones posibles, mecanismos o explicaciones confusas para fenómenos y características de los datos. También según los autores se relaciona con la búsqueda de patrones de variación. El modelado, implica la construcción y el uso de modelos, esto incluye la formación de modelos mentales de la realidad del contexto, respaldados por información y conocimiento experto. La estadística proporciona modelos distintivos y marcos para pensar genéricamente sobre ciertos aspectos de la investigación, todos los modelos son simplificaciones de la realidad. Por último,

la aplicación de técnicas implica seguir precedentes, reconocer y utilizar arquetipos y usar herramientas de resolución de problemas.

Lo referente a los tipos fundamentales de pensamiento estadístico.

Figura 9. Dimensión 2: Tipos de pensamiento



Fuente: Wild & Pfannkuch, (1999, p. 226)

El primer elemento es el reconocimiento de la necesidad de los datos, los estadísticos deben de reconocer las insuficiencias de las experiencias personales y la evidencia anecdótica, llevando al deseo de basar las decisiones en datos recopilados deliberadamente.

*La transnumeración*, se define como "transformaciones de numeración realizadas para facilitar la comprensión", la idea fundamental es formar y cambiar representaciones de datos para llegar a una mejor comprensión de un sistema, como una herramienta importante permea todo el análisis de datos (explorar gráficos, re-expresar datos, probar modelos).

Consideración de la variación, el pensamiento estadístico moderno se preocupa por el aprendizaje y la toma de decisiones bajo incertidumbre, gran parte de la cual proviene de la variación omnipresente, la variación tiene importantes impactos prácticos, haciendo difícil resolver preguntas de causa y efecto, los estadísticos modelan la variación para predecir, explicar o controlar.

Razonamiento con modelos estadísticos, implica el uso de modelos abstractos, a menudo probabilísticos, para lidiar con la complejidad y separar los fenómenos "duraderos" de los puramente locales. Integración de lo estadístico y lo contextual, es la síntesis de la información de los datos, el conocimiento estadístico y el conocimiento del contexto, no se puede pensar

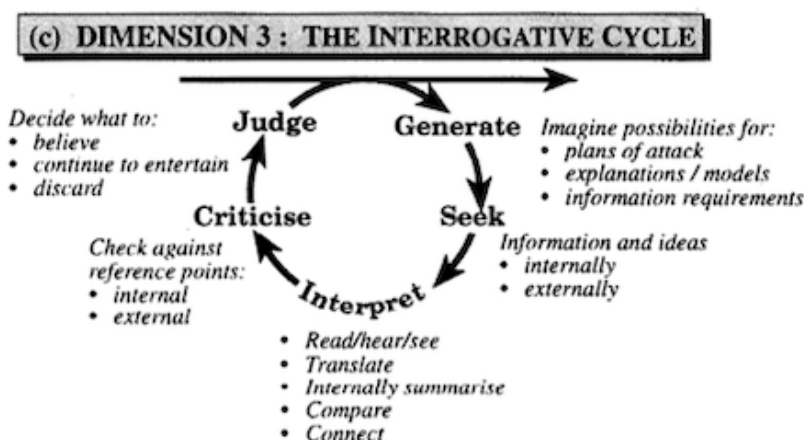
estadísticamente sin conocimiento del contexto. Implica aplicar todo el conocimiento relevante y hacer conexiones entre el conocimiento del contexto existente y los resultados del análisis para llegar al significado. Hay un continuo ir y venir entre pensar en la esfera del contexto y la esfera estadística a lo largo de todo el ciclo PPDAC.

### 5.1.3 Dimensión tres: ciclo investigativo

En la tercera dimensión del marco se presenta el ciclo interrogativo, Este es un proceso de pensamiento genérico que se utiliza constantemente en la resolución de problemas estadísticos. Es recursivo y se aplica tanto a niveles macro como muy detallados.

Generar, significa imaginar y generar posibilidades. Se aplica a la búsqueda de posibles causas, explicaciones y mecanismos, a las formas en que las partes de un sistema pueden interrelacionarse (bloques de modelos mentales/estadísticos). La generación puede provenir del contexto, los datos o el conocimiento estadístico, y puede aplicar al problema actual o registrarse para investigación futura (generación de hipótesis).

Figura 10. Dimensión 3: ciclo investigativo



Fuente: Wild & Pfannkuch, (1999, p. 226)

Buscar, generalmente sigue a la generación y es la búsqueda o recuperación de información. Puede ser interna (memoria) o externa (hablar con otros, leer literatura, recolectar datos estadísticos, consultar datos disponibles).

Interpretar, significa tomar y procesar los resultados de la búsqueda. El proceso es: Leer/oír/ver → Traducir → Resumir internamente → Comparar → Conectar. Se aplica a todas las formas de información. "Conectar" se refiere a interconectar las nuevas ideas e información con nuestros modelos mentales existentes y ampliar esos modelos.

Criticar, la fase de crítica se aplica a la información e ideas entrantes e implica verificar la consistencia interna y contrastar con puntos de referencia. Se pregunta: "¿Es esto correcto?", "¿Tiene sentido?", "¿Concuerda con lo que yo u otros sabemos?". Se verifica contra puntos de

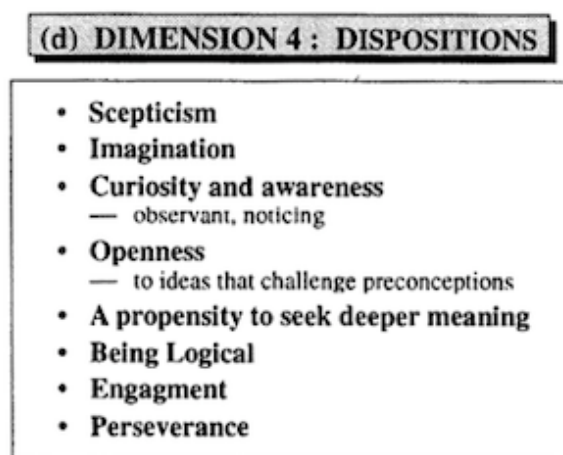
referencia internos (conocimiento del contexto/estadístico, restricciones, anticipación de problemas) y externos (otras personas, literatura, datos). Incluye la metacognición, monitorear el propio pensamiento, verificando contra el propósito, los sistemas de creencias (preconceptos) y las respuestas emocionales.

Juzgar, es el punto final de decisión de la crítica, se decide qué conservar, qué descartar o ignorar, qué seguir considerando tentativamente, qué creer. Se aplica a la fiabilidad de la información, la utilidad de las ideas, la practicidad de los planes, la adecuación de la encapsulación, la conformidad con la comprensión del contexto y estadística, la plausibilidad de explicaciones, la necesidad de más investigación, etc.

#### 5.1.4 Dimensión cuatro: Disposiciones

Como última Dimensión del marco se describen las *Disposiciones*. Esta dimensión abarca las cualidades personales que influyen o inician un modo de pensamiento. Emergieron de las entrevistas con estadísticos y se reconocieron en los estudiantes. Se consideran genéricas, pero se discuten en el contexto de la resolución de problemas estadísticos. Las disposiciones pueden ser dependientes del problema, cambiando según el grado de compromiso de la persona con el problema.

Figura 11. Dimensión 4: Disposiciones



Fuente: Wild & Pfannkuch, (1999, p. 226)

Escepticismo, una tendencia a estar constantemente buscando fallos lógicos y fácticos al recibir nuevas ideas e información. Es adoptar una "actitud crítica", preguntando si las conclusiones están justificadas. Implica preocupación por la motivación, predisposiciones y objetividad. Evoca "preguntas de preocupación" técnicas (adecuación de mediciones, diseño, calidad de datos, método de análisis, apoyo de conclusiones). También implica verificar la credibilidad ("¿Esto es siquiera creíble?") a menudo usando el sentido numérico.

Imaginación, fundamental para el pensamiento estadístico. Crucial para la formación de modelos mentales que capturen la dinámica esencial de un problema, para ver una situación desde diferentes perspectivas y para generar posibles explicaciones (o explicaciones confusas) para fenómenos y características de los datos. Curiosidad y Conciencia, son los manantiales del proceso de generación de preguntas. Se desencadenan al notar algo y reaccionar con preguntas internas como "¿Por qué?", "¿Cómo sucedió eso?". Ser observador (consciente) y curioso es fundamental para el aprendizaje innovador. "Notar la variación y preguntarse por qué" es crítico para generar ideas y para la exploración/análisis de datos.

Apertura, estar abierto a ideas que desafían las preconcepciones. Ayuda a registrar y considerar nuevas ideas e información que entran en conflicto con las propias suposiciones. Una propensión a buscar un significado más profundo, no tomar las cosas al pie de la letra y estar dispuesto a indagar un poco más.

Ser Lógico, la capacidad de detectar cuándo una idea se deriva de otra y cuándo no, y de construir un argumento lógico. La síntesis de nueva información con conocimiento existente implica ver implicaciones. El razonamiento lógico es el camino seguro para llegar a conclusiones válidas.

Compromiso, un interés intenso en un problema o área que lleva a una sensibilidad y conciencia aumentadas hacia la información periférica. Intensifica cada una de las otras "disposiciones": curiosidad, conciencia, imaginación y perseverancia.

El interés espontáneo es innato, pero puede ser ayudado por el conocimiento previo, el pago o la importancia del problema. La falta de compromiso puede ser una dificultad para que los estudiantes piensen estadísticamente.

Y por último la perseverancia, una cualidad autoevidente que es importante en la resolución de problemas. Se intensifica con el compromiso.

En resumen:

El ciclo de investigación proporciona la estructura general para abordar el problema real. Los tipos de pensamiento son los procesos cognitivos utilizados dentro de esa estructura. El Ciclo Interrogativo es el proceso dinámico constante de interactuar y evaluar información e ideas a lo largo del camino, a menudo estimulado por preguntas disparadoras. Y las Disposiciones son las actitudes personales que influyen y energizan todo el proceso, determinando la profundidad, la amplitud y la rigurosidad del pensamiento aplicado a la solución del problema real.

El pensamiento estadístico se describe como integración de estos elementos, junto con el conocimiento estadístico y el conocimiento del contexto, para producir percepciones e inferencias sobre el problema real. Ocurre en el continuo ir y venir entre la esfera del contexto y la esfera estadística.

## **5.2 Elementos para la descripción de datos bivariados**

Para los análisis iniciales y la descripción de los datos a partir de las representaciones gráfica y tabular se utilizarán los elementos propuestos por Watkins y colaboradores, 2007. Para describir y cuantificar la información de los datos bivariados se utilizan las siguientes rasgos: *la forma*, este rasgo permite a los estudiantes visualizar si el conjunto de datos sigue un patrón lineal o no (recta de regresión lineal), *la tendencia*, proporciona la dirección y por consiguiente el signo al identificar si las variables crecen proporcionalmente o si no lo hacen (signo de la pendiente de la recta de regresión) y *la fuerza*, los estudiantes perciban el ajuste de bondad a partir de una línea imaginaria de mejor ajuste (la fuerza de la correlación), por lo tanto los estudiantes pueden argumentar si es fuerte o débil. Se recomiendan una serie de pasos a seguir al momento de realizar la descripción de los datos bivariados.

Primero, *identifica las variables y los casos*. En un diagrama de dispersión cada punto representa un caso, con la coordenada “x” igual al valor de la variable independiente y la coordenada “y” igual a la variable dependiente. También se debería describir la escala (unidades de medida) y el rango de cada variable.

Segundo, *describir la forma general de la relación*, poniendo atención en:

- La linealidad: ¿El patrón es lineal (disperso sobre una línea) o curvo?
- Los grupos: ¿Hay solo un grupo, o hay más de uno?
- Valores atípicos: ¿Hay alguna excepción sorprendente al patrón general?

Tercero, *describir la tendencia*; si a medida que “x” se hace más grande, “y” tiende a hacerse más grande, hay tendencia positiva, (la nube de puntos tiende a inclinarse hacia arriba a medida que vas de izquierda a derecha). Si a medida que “x” se hace más grande “y” tiende a hacerse más pequeño, hay una tendencia negativa, (la nube de puntos tiende a inclinarse hacia abajo a medida que vas de izquierda a derecha).

Cuarto, *describir la fuerza de la relación*. Si los puntos se agrupan estrechamente alrededor de una línea de mejor ajuste o curva imaginaria, la asociación es fuerte. Si los puntos son dispersos más lejos de la línea, la asociación es débil. Si los puntos tienden a dispersarse en un extremo (la tendencia es llamada *heteroscedasticidad*), la relación varía en fuerza. Si no, tiene fuerza constante.

Cinco, realizarse la siguiente pregunta, ¿El patrón se generaliza a otros casos?

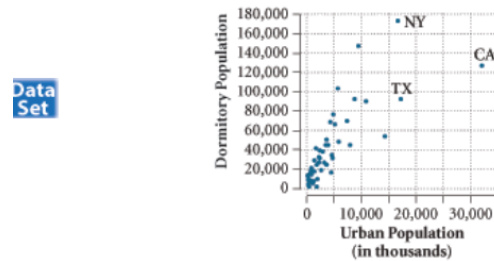
Seis, Por último, ¿Hay explicaciones plausibles para el patrón? ¿Qué variable causa la otra?

A continuación, se presenta un ejemplo explícito extraído de Watkins, et al. 2007.

Figura 12. Contexto del problema a describir

**Example: Dormitory Populations**

The plot in Display 3.3 shows, for the 50 states in the United States, the number of people living in college dormitories versus the number of people living in cities, in thousands. Describe the pattern in the plot.



**Display 3.3** Number of people living in college dormitories versus number of people living in cities for the 50 states in the United States. [Source: U.S. Census Bureau, 2000 Census of Population and Housing.]

Fuente: Watkins, et al., (2007, p. 108)

**Solución:**

*Las variables y los casos:* El diagrama de dispersión representa la población de dormitorios universitarios frente a la población urbana, en miles, para los 50 estados de EE.UU. La población en dormitorios varía desde cerca de 0 hasta más de 174 000 en Nueva York. La población urbana varía desde cerca de 0 hasta aproximadamente 17 millones en Texas y Nueva York y 32 millones en California.

*Forma:* Aunque la mayoría de los estados siguen una tendencia lineal, los tres estados con las poblaciones urbanas más grandes sugieren curvatura en el gráfico, porque en esos estados el número de personas que viven en dormitorios es proporcionalmente menor que en los estados pequeños. California puede considerarse un valor atípico respecto a su población urbana, que es mucho mayor que la de otros estados. También es un valor atípico respecto al patrón general, porque se encuentra muy por debajo de la tendencia lineal general.

*Tendencia:* La tendencia es positiva: los estados con poblaciones urbanas más grandes tienden a tener poblaciones mayores en dormitorios universitarios, y los estados con poblaciones urbanas menores tienden a tener poblaciones más pequeñas en dormitorios.

*Fuerza:* La relación varía en su intensidad. Para los estados con las poblaciones urbanas más pequeñas, los puntos se agrupan bastante cerca de una línea. Para los estados con las poblaciones urbanas más grandes, los puntos están más dispersos respecto a la línea. En general, la fuerza de la relación es moderada.

*Generalización:* Los 50 estados no son una muestra de una población más amplia de casos; por lo tanto, la relación aquí no necesariamente se generaliza a otros casos. Debido a que ambas variables tienden a cambiar lentamente a lo largo del tiempo, podemos esperar que la relación en el diagrama de dispersión sea similar a la de otros años.

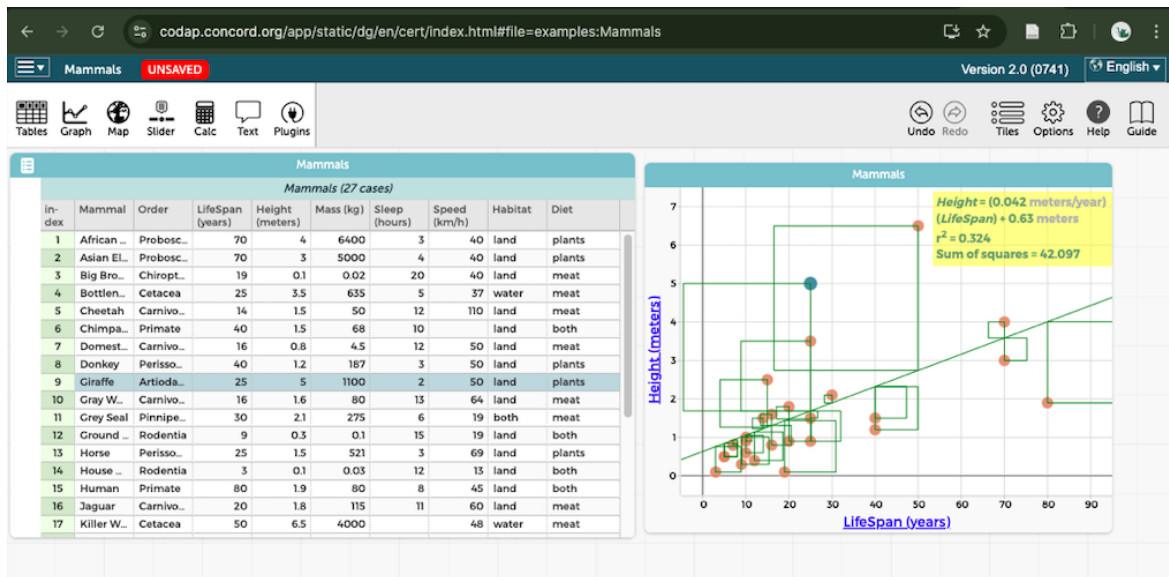
*Explicación:* Es tentador atribuir la relación positiva a la idea de que las ciudades atraen universidades. (solo piensa en una ciudad grande cercana e identifica cuántas universidades están ubicadas allí). Sin embargo, la razón principal de la relación positiva no es tan interesante: ambas variables están relacionadas con la población del estado. Cuantas más personas hay en un estado, más personas viven en dormitorios universitarios y más personas viven en ciudades. (Aquí hay una moraleja: interpretar asociaciones puede ser más complicado, en parte porque las dos variables que ves en un gráfico a menudo están relacionadas con alguna variable oculta que no estás observando).

### **5.3 Uso de tecnología**

Los avances tecnológicos han facilitado y mejorado la precisión de los cálculos en los cursos de estadística. Esto ha disminuido el tiempo que los profesores dedican a los procedimientos, permitiendo así, que se enfoquen más en los conceptos estadísticos (Chance, et al., 2008). Además, el uso adecuado de visualizaciones puede transformar la manera en que los estudiantes comprenden conceptos estadísticos, especialmente si se emplean programas que ofrecen opciones dinámicas, flexibles y múltiples representaciones vinculadas (Batanero, et al. 1996; Cobb, et al., 2003). Estas herramientas permiten que los estudiantes desplieguen, manipulen y analicen datos, lo que facilita su razonamiento sobre *la regresión lineal y la correlación* (Biehler, et al., 2013).

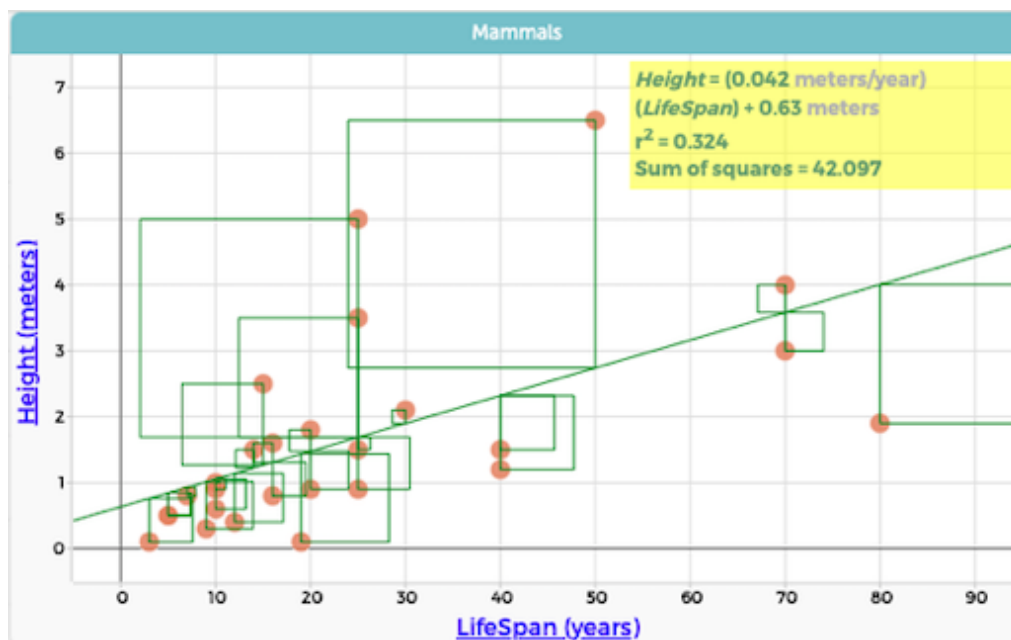
Tecnologías como CODAP contribuyen a mejorar la comprensión y aplicación de ideas relacionadas con *la regresión lineal y la correlación*, ya que permite a los alumnos trabajar con diferentes tipos de gráficos y realizar ajustes visuales, como la colocación de líneas de regresión. Por ejemplo, este software facilita el ajuste visual de una recta de regresión y muestra de manera gráfica cómo varían las desviaciones cuadráticas al ir modificando la línea sobre los datos. Esta y otras capacidades del software promueven la exploración activa y el desarrollo de inferencias estadísticas en el aula. (figura...)

Figura 13. Captura de pantalla de CODAP con diferentes representaciones vinculadas.



Fuente: Elaboración propia.

Figura 14. Captura de pantalla de CODAP de la recta de regresión y residuales.



Fuente: Elaboración propia.

En conclusión, como lo señala Fuentes (2016), el uso de la estadística en problemas reales es esencial para consolidar el aprendizaje de los estudiantes. Coincidimos en que esta estrategia resulta clave para desarrollar el pensamiento estadístico, una competencia fundamental en la formación del futuro ingeniero. Trabajar con situaciones auténticas favorece no sola la

comprensión al permitir que los estudiantes creen y exploren sus propios problemas, sino también su capacidad para identificar situaciones relevantes, investigar y ampliar su perspectiva sobre la estadística, fortaleciendo así su formación disciplinar. Asimismo, posibilita que vivan un proceso de iniciación científica que estimula la formulación de preguntas, la identificación de problemas, la construcción de modelos estadísticos y el desarrollo de la creatividad, aspecto que incrementan significativamente su motivación y compromiso

## 6 ASPECTOS METODOLÓGICOS

En el apartado siguiente se describen de manera detallada las acciones necesarias para alcanzar los objetivos planteados en este proyecto. Dichas acciones se estructuran y orientas a partir de los principios de la ingeniería didáctica propuestos por Artigue (1995), lo que permite organizar el proceso de diseño, implementación y análisis con un enfoque sistemático y fundamentado. Además, se presenta un esquema que articula cada uno de los objetivos específicos con las acciones necesarias para su cumplimiento y con los productos esperados en cada fase.

### 6.1 La ingeniería didáctica

La ingeniería didáctica se caracteriza por un diseño experimental basado en realizaciones didácticas en el aula y una validación esencialmente interna, confrontando análisis a priori y a posteriori. La cual consta de cuatro fases o etapas las cuales son: primera fase análisis preliminar; además como segunda fase esta las concepciones y el análisis a priori; como tercera fase se propone una experimentación para finalizar con un análisis a posteriori y la validación.

#### 6.1.1 Análisis preliminar

Fase 1: Análisis preliminar: Esta fase es el fundamento para la fase de concepción, se basa en un marco teórico didáctico general, conocimientos didácticos previos en el campo de estudio y una serie de análisis preliminares, los análisis más frecuentes incluyen: Análisis epistemológico de los contenidos a enseñar; Análisis de la enseñanza tradicional y sus efectos; Análisis de las concepciones de los estudiantes, sus dificultades y los obstáculos que determinan su evolución; Análisis del campo de restricciones en el que se situará la realización didáctica y la consideración de los objetivos específicos de la investigación.

Estos análisis preliminares no siempre son completamente explícitos en las publicaciones, y a menudo se retoman y profundizan a lo largo de las fases de la ingeniería. Las exigencias de esta fase varían según el objetivo de la investigación, un ejemplo ilustra cómo esta fase puede centrarse en analizar el funcionamiento de un sistema, identificando restricciones desde una perspectiva sistémica que distingue dimensiones epistemológica, cognitiva y didáctica. Aunque a veces la dimensión cognitiva recibe menos énfasis, un análisis preliminar detallado de las concepciones de los estudiantes es a menudo un apoyo esencial para la concepción, buscando provocar una evolución controlada de las mismas.

Tabla 2. Objetivo específico 1, Acciones metodológicas, y productos esperados.

<i>Objetivo</i>	<i>Acciones metodológicas</i>	<i>Productos esperados</i>
-----------------	-------------------------------	----------------------------

<p>OB1. Realizar un análisis preliminar para identificar las necesidades en la enseñanza y aprendizaje de la regresión lineal y la correlación.</p>	<p>A1. Elaborar el análisis epistemológico.</p> <p>A2. Desarrollar el análisis didáctico.</p> <p>A3. Elaborar el análisis cognitivo.</p> <p>A4. Determinar los elementos a priori de las actividades.</p> <p>A5. Determinar la trayectoria de enseñanza a priori.</p> <p>A6. Determinar los contextos que se incluirán en la propuesta de intervención.</p> <p>A7. Definir las etapas del modelo de pensamiento estadístico en la investigación empírica, que se utilizarán en cada actividad.</p>	<p>P1. Elementos para el diseño de las actividades.</p> <p>P2. Instrumentos a priori para el análisis el diseño de las actividades.</p> <p>P3. Instrumentos a priori para el análisis de la implementación.</p>
---	--	---

Fuente: Elaboración propia.

### 6.1.2 Concepción y análisis a priori de las situaciones didácticas de la ingeniería

En esta fase 2, el investigador toma decisiones sobre las "variables de comando" del sistema que no están fijadas por las restricciones. Se distinguen dos tipos de variables de comando: macro-didácticas o globales, relativas a la organización global de la ingeniería, y micro-didácticas o locales, relativas a la organización de una secuencia o fase. Estas pueden ser generales o dependientes del contenido didáctico, las selecciones globales se hacen primero y preceden a las selecciones locales.

El análisis a priori comienza en esta fase y es parte del proceso de validación interna, su objetivo es determinar cómo las elecciones realizadas permiten controlar los comportamientos de los estudiantes y su significado. Tradicionalmente, el análisis a priori tiene una parte descriptiva y una predictiva; describe las selecciones locales, analiza lo que está en juego para el estudiante en la situación (sus posibilidades de acción, decisión, control, validación) y previene los posibles comportamientos. El objetivo es asegurar que los comportamientos esperados, si ocurren, sean el resultado de la aplicación del conocimiento previsto.

Tabla 3. Objetivo específico 2, Acciones metodológicas, y productos esperados.

Objetivo	Acciones	Productos
OB2. Diseñar actividades didácticas que promuevan un acercamiento informal sobre la regresión lineal y la correlación, con el apoyo de tecnología como un puente hacia el acercamiento formal.	<p>A8. Determinar el número de actividades y los objetivos de cada actividad.</p> <p>A9. Determinar la estructura del proyecto de intervención.</p> <p>A10. Determinar los objetivos de las actividades del proyecto de intervención.</p> <p>A11. Elaborar los instrumentos de recogida de la información.</p> <p>A12. Elaborar los instrumentos de observación.</p>	<p>P4. Secuencia de actividades didácticas.</p> <p>P5. Instrumentos para la observación.</p> <p>P6. Instrumentos para el análisis.</p>

Fuente: Elaboración propia.

### 6.1.3 Experimentación

Esta fase 3, implica la realización efectiva de las secuencias de enseñanza concebidas en las fases anteriores. La secuencia de actividades se implementará en el Instituto Tecnológico de sonora con estudiantes de las carreras de ingeniería y que estén cursando la materia de Probabilidad y Estadística. Generalmente estos estudiantes cursan dicha materia en el tercer semestre, el tamaño de la muestra dependerá de las condiciones administrativas en dicho curso. Es necesario generar instrumentos para recabar la información, videograbar de permitirse y contar con un observador, el cual dará recomendaciones derivadas de la implementación.

Tabla 4. Objetivo específico 3, Acciones metodológicas, y productos esperados.

<i>Objetivo</i>	<i>Acciones</i>	<i>Productos</i>
OB3. Implementar las actividades didácticas con estudiantes de las carreras de	A13. Definir el papel del profesor que implemente el proyecto.	P7. Hojas de respuestas de los estudiantes.

ingeniería en el Instituto Tecnológico de Sonora, utilizando el marco	<p>A14. Definir los momentos de interacción entre los estudiantes y el profesor.</p> <p>A15. Moderar el uso de las herramientas tecnológicas.</p> <p>A16. Definir el papel del observador.</p>	P8. Grabaciones de la interacción grupal en la solución de las situaciones problema.
---	--	--

Fuente: Elaboración propia.

#### 6.1.4 Análisis a posteriori y evaluación

La fase 4 se basa en los datos recogidos durante la experimentación, como observaciones de clase y producciones de los estudiantes. Estos datos se complementan frecuentemente con información de metodologías externas como cuestionarios y entrevistas, la validación de las hipótesis de investigación se fundamenta esencialmente en la confrontación entre el análisis a priori y el análisis a posteriori.

Este proceso de validación interna es distinto de los esquemas de validación estadística usuales en experimentaciones en clase. Sin embargo, la validación presenta dificultades, como la dificultad de comunicar completamente el análisis a priori en las publicaciones y la tendencia a proponer modificaciones a la ingeniería en lugar de analizar qué hipótesis son invalidadas por las distorsiones observadas, las hipótesis a menudo son globales y de largo plazo, lo que dificulta un verdadero proceso de validación dentro de la ingeniería.

Por lo que la ingeniería didáctica no solo es un medio para investigar los procesos de enseñanza y aprendizaje, sino también un motor para el progreso teórico y metodológico de la didáctica. Se caracteriza por un esquema experimental basado en la concepción, realización, observación y análisis de secuencias de enseñanza, denominadas "realizaciones didácticas" en clase. A diferencia de otras investigaciones basadas en la experimentación en el aula que recurren a comparaciones estadísticas y validación externa, la ingeniería didáctica se sitúa en el registro de los estudios de caso con una validación esencialmente interna, fundamentada en la confrontación entre el análisis a priori y el análisis a posteriori.

Tabla 5. Objetivo específico 4, Acciones metodológicas, y productos esperados.

<i>Objetivo</i>	<i>Acciones</i>	<i>Productos</i>
OB4. Valorar los resultados de la implementación.	A17. Contrastar los análisis a priori con los análisis a posteriori.	P9. Resultados del alcance logrado de los objetivos del trabajo.

<p>A18. Generar las conclusiones derivadas de los análisis de la implementación.</p>	<p>P10. Resultados del logros en la implementación.</p> <p>P11. Conclusiones parciales de los logros.</p>
<p>A19. Generar recomendaciones sobre la implementación del proyecto de intervención.</p>	<p>P12. Recomendaciones parciales de los logros.</p>
<p>A20. Generar las conclusiones sobre el uso de la tecnología.</p>	

Fuente: Elaboración propia.

## 7 LA PROPUESTA Y SUS CARACTERÍSTICAS

Una vez delimitados los objetivos y los elementos fundamentales del marco conceptual, así como los aspectos metodológicos que sustentan este trabajo desde la perspectiva de la ingeniería didáctica de Artigue (1995), en este apartado se presentan las características generales de la secuencia de actividades que aportarán elementos para el desarrollo del pensamiento estadístico, a través del marco de Wild y Pfannkuch (1999).

La construcción de la secuencia de actividades considera los resultados derivados de los estudios preliminares, las demandas específicas del marco conceptual adoptado y las características del contexto educativo donde se implementará, particularmente su orientación hacia carreras de ingeniería. Por ello, las tareas propuestas se desarrollan en torno a contextos reales y problemas auténticos que permiten a los estudiantes aplicar conceptos estadísticos (como la regresión lineal y la correlación) a situaciones propias de su futura práctica profesional. De esta forma, la secuencia de actividades no solo busca promover la comprensión conceptual, sino también fortalecer habilidades de análisis, interpretación y toma de decisiones fundamentadas en datos alienándose con los requisitos formativos de la educación superior en ingeniería.

### 7.1 Características de las actividades

La propuesta está dirigida a estudiantes de las carreras de ingeniería (Industrial, Civil, Electrónica, Electromecánica, Mecatrónica, Manufactura y Química) del Instituto Tecnológico de Sonora, quienes al momento de la implementación cursarán el tercer semestre. A través de la secuencia didáctica se busca mejorar la comprensión de la regresión lineal y la correlación, promoviendo el uso de herramientas tecnológicas como CODAP y otros recursos para la elaboración de instrumentos de recolección de información. Asimismo, se fomentará el trabajo colaborativo, de manera que los estudiantes argumenten, contrasten y compartan información que les permita avanzar en cada etapa; además, se prevén momentos de trabajo individual y en parejas para la discusión y consensos conceptuales. El diseño de la secuencia contempla tres actividades orientadas al desarrollo del pensamiento estadístico. A continuación, se describen los elementos específicos que conforman cada una de ellas.

El objetivo general del proyecto de intervención es: *Desarrollar el pensamiento estadístico en los estudiantes mediante una secuencia de actividades orientadas a la comprensión e interpretación de la relación entre variables a través del estudio de la regresión lineal simple y la correlación, promoviendo la exploración de datos reales, el razonamiento covariacional y la toma de decisiones fundamentadas en el análisis estadístico.*

### **7.1.1 Actividad uno: *La temperatura influye en el consumo de energía eléctrica.***

Objetivo: Analizar la relación entre la temperatura ambiente y el consumo de energía eléctrica para identificar patrones de covariación y modelar la dependencia entre ambas variables, promoviendo la interpretación crítica de los resultados y la toma de decisiones orientadas a la eficiencia energética.

En esta actividad se promoverá el análisis de las representaciones tabular y gráfica de manera informal en una primera instancia utilizando los elementos propuestos por Watkins y col., posteriormente realizar interpretaciones formales apoyados en el uso de CODAP. Además, se contempla que los estudiantes discutan sus ideas al trabajar de manera grupal, individual y en parejas.

Con respecto a las etapas del ciclo investigativo. En esta actividad se definirá el problema, se discutirán elementos generales del plan y los datos, los cuales por cuestión de tiempo serán proporcionados por el profesor, los análisis se realizan con los elementos que propone Watkins y col. Y con los resultados de estos análisis se contará con elementos para generar conclusiones.

El tiempo estimado para realizar esta actividad es de 5 hrs clase.

### **7.1.2 Actividad dos: *El tiempo de secado es factor importante para la resistencia en el concretos.***

Objetivo: Explorar la influencia del tiempo de secado y la humedad en la resistencia del concreto mediante el análisis de datos experimentales, con el fin de comprender cómo las variables ambientales afectan el comportamiento del material y fortalecer la capacidad de modelar fenómenos mediante regresión lineal y correlación.

En esta actividad se contempla la integración de todas las etapas del ciclo investigativo, en interacción con las dimensiones propuestas por el marco para el desarrollo del pensamiento estadístico (los tipos de pensamiento, el ciclo interrogativo y las disposiciones). Asimismo, se incorpora el uso de recursos tecnológicos, particularmente la plataforma CODAP, que facilitará el análisis de la información y la generación de representaciones dinámicas y vinculadas. El tiempo estimado para el desarrollo de esta actividad es de 8 horas clase.

### **7.1.3 Actividad tres: *La ergonomía me ayuda al diseño de productos.***

Objetivo: Investigar la relación entre las dimensiones antropométricas y las características de diseño de implementos ergonómicos, empleando análisis de correlación y regresión para fundamentar decisiones de diseño que mejoren la funcionalidad y la adaptación al usuario.

En esta actividad se contempla la integración de todas las etapas del ciclo investigativo, en interacción con las dimensiones propuestas por el marco para el desarrollo del pensamiento

estadístico (los tipos de pensamiento, el ciclo interrogativo y las disposiciones). Asimismo, se incorpora el uso de recursos tecnológicos, particularmente la plataforma CODAP, que facilitará el análisis de la información y la generación de representaciones dinámicas y vinculadas. El tiempo estimado para el desarrollo de esta actividad es de 8 horas clase.

## 7.2 Ejemplo de actividad didáctica

Como ejemplo de una actividad inicial donde se promueve la descripción de datos bivariados a partir de la representación tabular y gráfica, se presenta a continuación:

**¿La  
temperatura  
afecta a la  
demanda de  
energía  
eléctrica?**



En las noticias se pueden encontrar que debido al cambio climático se han rotó récords de temperatura a nivel mundial. Aseguran que en los últimos años hemos tenido incrementos anuales constantes de temperatura máximas que estas van desde los 1.34 y 1.41 grados Celsius. Localmente en la región en junio del 2024 se alcanzó una temperatura máxima de 49.5°. Estos incrementos de temperatura tienen un impacto diverso, aumentando el costo de energía eléctrica y los riesgos de salud para la humanidad. Derivado de estos cambios de temperatura en Sonora se han modificado períodos para los horarios para las tarifas de verano, estos determinados por la comisión federal de electricidad. Podemos preguntarnos: ¿Fue una medida necesaria? ¿Crees que en el corto plazo debido a los cambios constantes de temperatura se deba de modificar este periodo? ¿Se puede anticipar como cambiará la cantidad de energía eléctrica que se tiene que producir en la zona noroeste para satisfacer la demanda?

A continuación, se muestran un conjunto de 24 datos:

Mes: meses del año etiquetados con el número correspondiente del mes.

Año: el año con cuatro dígitos (ej. 2024).

Temperatura\_Prom: temperatura promedio en el mes correspondiente en grados centígrados.

Temperatura\_Max: temperatura máxima alcanzada en el mes correspondiente en grados centígrados.

Demanda\_Prom (MW): demanda promedio en Mega Watts de la zona noroeste.

Demanda\_Max (MW): demanda máxima en Mega Watts de la zona noroeste.

Los datos fueron recabados de los siguientes sitios web:  
(<https://www.cenace.gob.mx/graficademanda.aspx>);  
(<https://smn.conagua.gob.mx/es/climatologia/temperaturas-y-lluvias/resumenes-mensuales-de-temperaturas-y-lluvias>).

Mes	Año	Temperatura_Prom	Temperatura_Max	Demanda_Prom (MW)	Demanda_Max (MW)
1	2023	13.3	20.9	1267.23	1846.79
2	2023	13.3	20.8	1199.42	1807.21
3	2023	16	23.7	1226.52	1933.37
4	2023	18.7	27.5	1316.64	2307.55
5	2023	21.6	29.8	1647.53	2521.93
6	2023	25	33.1	1697.71	2660.78
7	2023	30	37.1	2214.36	2909.65
8	2023	29.3	36	2179.93	3032.43
9	2023	27.3	34.4	2055.08	2869.44
10	2023	24.2	31.7	1714.65	2586.33
11	2023	19.2	27	1329.45	2192.5
12	2023	15.6	23.9	1047.42	1623.8
1	2024	13.6	21.3	1131.34	1659.89
2	2024	15	22.2	1115.35	1855.94
3	2024	16.1	23.85	1049.84	1644.37
4	2024	18.7	27.05	1223.22	1920.59
5	2024	22.1	30.95	1508.46	2255.32
6	2024	27.3	35.25	1924.47	2718.52
7	2024	29.9	36.7	2056.83	2773.91
8	2024	29.3	36.3	2069.73	2871.86
9	2024	27.7	35.65	2014.7	2982.04
10	2024	24.7	33.45	1934.94	2945.45
11	2024	16.5	25.3	1132.34	1615.57
12	2024	15.4	24.65	1108.52	1598.85

Fuente: Elaboración propia.

Trabajo grupal:

Con las preguntas iniciales se promoverá una discusión grupal, para generar ideas generales sobre la pertinencia de la información y de lo relevante para especular sobre los cambios en el futuro debidos al fenómeno.

Se espera que los estudiantes logren determinar el problema, identificar sus limitaciones y reconocer que las variables involucradas no son de carácter determinista, pues están influenciadas por diversos factores que introducen incertidumbre. Dado que se trata de una actividad inicial, se proporcionarán datos previamente recopilados de distintas fuentes, como el CENACE y la CONAGUA. Es importante señalar que dichos datos fueron depurados y organizados por el profesor, ya que el propósito central de la actividad es que los estudiantes se enfoquen en la descripción y el análisis de distintas representaciones.

Trabajo individual:

Analizando a simple vista las variables, Temperatura\_Prom (°C) y la Demanda\_Prom (MW) de la tabla de datos anterior.

¿Cuál es la variable dependiente? explicar

¿Cuál es el rango de Demanda\_Prom (MW)?

¿Qué cantidad de Demanda\_Prom (MW) le corresponde el valor de la Temperatura\_Prom de 13.3 °C? ¿Es razonable esta relación?, explica.

¿Qué cantidad de Demanda\_Prom (MW) le corresponde el valor de la Temperatura\_Prom de 28 °C?

¿Parece haber una relación entre la Temperatura\_Prom (°C) y la Demanda\_Prom (MW) de energía eléctrica?, Describe esa relación.

¿Los datos de manera general tienen una forma (lineal, curva, etc.)?

¿Existen datos que parecen ser no razonables (no siguen la tendencia)? Explique.

¿Los datos se agrupan torno a algún rango de valores?

¿Los datos guardan una relación entre ellos?

¿Parece que los datos se agrupan o se dispersan o mantienen la tendencia?

¿Se pueden generalizar la tendencia o se debe solo a este conjunto de datos en particular?

¿Crees que exista otra variable que este influyendo en la tendencia de los datos que deba ser considerada?

Con las preguntas anteriores se orienta a los estudiantes a emplear los elementos propuestos por Watkins y colaboradores, iniciando con el análisis de la representación tabular. Aunque esta tarea puede resultar desafiante, su propósito es que los estudiantes realicen transnumeraciones

a partir de los datos, de modo que estén en condiciones de generar argumentos pertinentes frente a las preguntas planteadas.

Indicaciones para el estudiante:

De considerarlo necesario, ingresa al siguiente hipervínculo:

<https://codap.concord.org/app/static/dg/es/cert/index.html#shared=https%3A%2F%2Fcfm-shared.concord.org%2F7yneRc6hio1OPyxzBLzH%2Ffile.json>

Encontrarás los datos cargados en CODAP, puedes manipular la tabla para contestar de mejor manera las preguntas anteriores. (en caso de necesitar ayuda solicitarla a tu profesor)

Trabajo en parejas.

Discute tus respuestas con las de tus compañeros de equipo y lleguen a un consenso sobre las mejores respuestas y la razón de la decisión.

Continuando el trabajo en parejas. Sigue las instrucciones de tu profesor para generar un gráfico de dispersión.

Analizando el gráfico generado por CODAP, contesta de nuevo las preguntas anteriores.

Con las respuestas a las preguntas anteriores generen un párrafo para comunicar a tus compañeros sobre la forma, la tendencia y la fuerza que tienen los datos, agregando información adicional que consideren pertinente.

El objetivo de estas tareas es que los estudiantes reconozcan que la representación elegida influye directamente en la profundidad y calidad de los argumentos que pueden formular al interpretar la información.

De manera grupal se discutirá.

¿Qué tan fuerte percibes la relación entre la Temperatura\_Prom (°C) y la Demanda\_Prom (MW) de energía?

¿Qué factores externos influyen para que la correlación entre las variables no sea perfecta?

¿Creen que, con los aumentos anuales constantes de la temperatura, se mantendrá también el aumento constante de demanda de energía eléctrica?

¿Qué otros factores (variables) se deberían analizar antes de tomar decisiones?

## 8 CRONOGRAMA

Con el propósito de sistematizar y dar seguimiento a las actividades desarrolladas durante los primeros cuatro semestres del programa de posgrado, así como de planificar de manera estructurada las acciones correspondientes a los siguientes cuatro semestres, se presenta a continuación un cronograma detallado. En este se incluye un listado organizado de tareas y metas que orientan la conclusión del trabajo de tesis. Los rectángulos sombreados indican los periodos semestres en los que cada acción fue realizada o se proyecta realizar, permitiendo visualizar la secuencia temporal, la duración y la superposición de las distintas etapas del proyecto.

En la tabla 6 se presenta el trabajo desarrollado en los semestres 2024-1, 2024-2, 2025-1 y 2025-2. La tabla 7 contiene la información de las tareas por realizar en los semestres; 2026-1, 2026-2, 2027-1 y 2027-2.

Tabla 6. Cronograma de los semestres 2014-1 al 2025-2.

Acciones	Semestres				Comentarios
	2024-1	2024-2	2025-1	2025-2	
Antecedentes					
Determinar los elementos del estado de arte.					
Identificar la problemática y los objetivos.					
Definir los elementos teóricos.					
Determinar las características de la propuesta de intervención.					
Escribir documento predoctoral.					

Fuente: Elaboración propia.

Tabla 7. Cronograma de los semestres 2016-1 al 2027-2.

Acciones	Semestres				Comentarios
	2026-1	2026-2	2027-1	2027-2	
Determinar objetivos de la propuesta de intervención.					
Fundamentar teórica y metodológicamente.					
Diseñar actividades de la propuesta de intervención.					
Elaborar instrumentos de recolección de datos.					
Identificar y gestionar condiciones del contexto real.					
Realizar la intervención didáctica.					
Registrar sistemáticamente las acciones producto de la intervención.					
Presentar avances parciales de la implementación.					
Organizar la información obtenida de la implementación.					
Analizar la información organizada.					
Identificar áreas de mejora en la propuesta e implementación.					
Establecer conclusiones locales.					

Integrar los resultados de los análisis a la tesis.					
Determinar el logro de los objetivos específicos de la tesis.					
Estructurar las conclusiones generales de la tesis.					
Elaborar la versión final de la tesis.					

Fuente: Elaboración propia.

Al concluir con las acciones previstas en este cronograma se tendrán elementos suficientes para realizar los seminarios y presentar los resultados de la tesis en el mes de febrero del 2028.

## REFERENCIAS

- Agnelli, H., Konic, P., Peperelli, S., Zón, N., & Flores, P. (2009). La función lineal obstáculo didáctico para la enseñanza de la regresión lineal. *UNIÓN-Revista Iberoamericana de Educación Matemática*, 5(17).
- Burrill, G., & Biehler, R. (2011). Fundamental statistical ideas in the school curriculum and in training teachers. *Teaching statistics in school mathematics-challenges for teaching and teacher education: A Joint ICMI/IASE study: The 18th ICMI study*, 57-69.
- Batanero, C., Estepa, A., Godino, J. D., & Green, D. R. (1996). Intuitive strategies and preconceptions about association in contingency tables. *Journal for Research in Mathematics Education*, 27(2), 151-169.
- Batanero, C., Estepa, A., & Godino, J. D. (1997). Evolution of students' understanding of statistical association in a computer-based teaching environment. *Research on the role of technology in teaching and learning statistics*, 191-205.
- Batanero, C., Gea, M. M., del Mar López-Martín, M., & Arteaga, P. (2017). Análisis de los conceptos asociados a la correlación y regresión en los textos de bachillerato. *Didacticae: Revista de Investigación en Didácticas Específicas*, (1), 60-76.
- Batanero, C., Godino, J. D., & Estepa, A. (1998). Building the meaning of statistical association through data analysis activities. In *PME CONFERENCE* (Vol. 1, pp. 1-221).
- Bakker, A., & Gravemeijer, K. P. (2004). Learning to reason about distribution. In *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 147-168). Dordrecht: Springer Netherlands.
- Ben-Zvi, D., & Garfield, J. (2004). Statistical literacy, reasoning, and thinking: Goals, definitions, and challenges. *The challenge of developing statistical literacy, reasoning, and thinking*, 66, 3-15.
- Biehler, R., Ben-Zvi, D., Bakker, A., & Makar, K. (2013). Technology for enhancing statistical reasoning at the school level. In M. A. Clements, A. Bishop, C. Keitel, J. Kilpatrick, & F. Leung (Eds.), *Third international handbook of mathematics education* (pp. 643-690). New York: Springer.
- Biehler, R., Frischemeier, D., Reading, C. y Shaughnessy, J. M. (2018). Reasoning about data. En Ben-Zvi, K. Makar, J. Garfield (Eds.). *International Handbook of Research in Statistics Education*. Springer.
- Burrill, G. y Biehler, R. (2011). Fundamental statistical ideas in the school curriculum and in training teachers. En C. Batanero, G. Burrill y C. Reading (Eds.), *Teaching statistics in school mathematics. Challenges for teaching and teacher education. A joint ICMI/IASE study* (pp. 57-69). Dordrecht, The Netherlands: Springer.

- Carver, R., Everson, M., Gabrosek, J., Horton, N., Lock, R., Mocko, M., ... & Wood, B. (2016). Guidelines for assessment and instruction in statistics education (GAISE) college report 2016.
- Casey, S. (2014). *Teachers' knowledge of students' conceptions and their development when learning linear regression*, en K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics. ICOTS9 Invited Paper-Refereed*. Voorburg, The Netherlands: International Statistical Institute.
- Casey, S. A. (2015). *Examining Student Conceptions of Covariation: A Focus on the Line of Best Fit*. *Journal of Statistics Education*, 23(1). <https://doi.org/10.1080/10691898.2015.11889722>.
- Chance, B. L. (2002, November). *Components of statistical thinking and implications for instruction and assessment*. *Journal of Statistics Education*, 10(3). Retrieved July 15, 2007, from <http://www.amstat.org/publications/jse/v10n3/chance.html>
- Chance, B., Ben-Zvi, D., Garfield, J., & Medina, E. (2008). *The role of technology in improving student learning of statistics*. *Technology Innovations in Statistics Education*, 1(1). Retrieved from <http://escholarship.org/uc/item/8sd2t4rr>.
- CODAP (Common Online Data Analysis Platform) [Computer software]. (2022). Retrieved from <https://codap.concord.org>
- Davidian, M. & Louis, T. *Why Statistics?*. Science, 06 Apr 2012, Vol. 336, Issue 6077, p.12 DOI:10.1126/science.1218685. <http://science.sciencemag.org/content/336/6077/12>.
- Devore, J. L. (2008). *Probabilidad y estadística para ingeniería y ciencias*. (7.a ed). México: Cengage.
- Engel, J., & Sedlmeier, P. (2011). *Correlation and regression in the training of teachers*. *Teaching Statistics in School Mathematics-Challenges for Teaching and Teacher Education: A Joint ICMI/IASE Study: The 18th ICMI Study*, 247-258.
- Estepa, A., & Batanero, C. (1996). Judgments of correlation in scatter plots: Students' intuitive strategies and preconceptions. *Hiroshima Journal of Mathematics Education*, 4(25), 25-41.
- Estepa, A., Gea, M. M., Cañadas, G., & Contreras, J. M. (2012). Algunas notas históricas sobre la correlación y regresión y su uso en el aula. *Números. Revista de didáctica de las Matemáticas*, 81, 5-14.
- Estepa, A., & Sánchez-Cobo, F. (2003). Evaluación de la comprensión de la correlación y regresión a partir de la resolución de problemas. *Statistics Education Research Journal*, 2(1), 54-68.

- Espino, G. (2014) La significación de la correlación en estudiantes de la carrera de psicología. (tesis de maestría, Universidad de Sonora)
- Fernández, G. (2018) Propuesta didáctica para un acercamiento intuitivo de la correlación lineal en estudiantes del área de Ciencias Sociales de la Universidad de Sonora (tesis de maestría, Universidad de Sonora)
- Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report: A pre-K–12 curriculum framework*. American Statistical Association.
- Fuentes Leyva, J. M. (2016). El desarrollo de habilidades para la resolución de problemas prácticos en la asignatura de Estadística. *Revista Cubana de Educación Superior*, 35(3), 30-46.
- García, A., & Yáñez, G. (2018). Predicción informal en un diagrama de puntos. *Acta Latinoamericana de Matemática Educativa*, 31(1), 32-39.
- Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International statistical review*, 70(1), 1-25.
- Garfield, J., & Ben-Zvi, D. (2008). *Developing students' statistical reasoning: Connecting research and teaching practice*. Springer Science & Business Media.
- Garfield, J., Ben-Zvi, D., Chance, B. & Zieffler, A. (2008). Learning to reason about covariation. *Developing Students' Statistical Reasoning: Connecting Research and Teaching Practice*, 289 – 308.
- Garfield, J., Chance, B., & Snell, J. L. (2001). Technology in college statistics courses. In *The teaching and learning of mathematics at university level: An ICMI study* (pp. 357-370). Dordrecht: Springer Netherlands.
- Groth, R. E., Jones, M., & Knaub, M. (2018). A framework for characterizing students' cognitive processes related to informal best fit lines. *Mathematical Thinking and Learning*, 20(4), 251–276. <https://doi.org/10.1080/10986065.2018.1509418>
- Inzunza Cazares, S. (2007). Recursos de Internet para apoyo de la investigación y educación estadística. *Revista Iberoamericana De Educación*, 41(4), 1–12. <https://doi.org/10.35362/rie4142448>
- Inzunza Cazares, S. (2016). Análisis de datos bivariados en un ambiente basado en applets y software dinámico. *Educación matemática*, 28(3), 61-90.
- ITSON (2016). Programa del curso Probabilidad y Estadística. Cd. Obregón, México: Instituto Tecnológico de Sonora.
- Hatil, J. G., & Akar, G. K. (2023). Investigating Mathematics Teacher Educators' Conceptions and Criteria for an Informal Line of Best Fit. In *Research on Reasoning with Data and*

- Statistical Thinking: International Perspectives* (pp. 119-142). Cham: Springer International Publishing.
- Medina, M., & Silvestre, E. (2020). Covariational Reasoning Patterns of High School Students in Problems of Correlation and Linear Regression. *Patrones de razonamiento covariacional de estudiantes de bachillerato en problemas de correlación y regresión lineal*. North American Chapter of the International Group for the Psychology of Mathematics Education.
- Medina delgado, M. N. (2021). Niveles de razonamiento covariacional derivados de las respuestas de estudiantes de bachillerato a problemas de correlación y regresión lineal. Cinvestav. Tesis de doctorado.
- Mendenhall, W., Beaver, R. J. & Beaver, B. M. (2010). *Introducción a la probabilidad y Estadística*. (10.a ed). México: Cengage.
- Morgado, C., & Sánchez, E. (2022). Razonamientos informales de estudiantes de bachillerato al enfrentar una tarea para introducir la recta de mejor ajuste. *Números: Revista de didáctica de las matemáticas*, (112), 65-82.
- Morgado, C., & Sánchez, E. (2024). Conceptualización de la recta de mejor ajuste desde un enfoque informal por estudiantes de bachillerato. *Educación matemática*, 36(2), 125-155.
- Moritz, J. (2004). Reasoning about covariation. In *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 227-255). Dordrecht: Springer Netherlands.
- Moore, D. S. (1997). New pedagogy and new content: The case of statistics. *International Statistical Review*, 65(2), 123–137. <https://doi.org/10.2307/1403773>
- Sorto, M. A., White, A., & Lesser, L. M. (2011). Understanding student attempts to find a line of fit. *Teaching Statistics*, 33(2), 49-52.
- Osorio Amaya, E. A., Inzunza Cazares, S., & Ward Bringas, S. E. (2023). Modelización estadística para el aprendizaje de la correlación y regresión lineal. PNA.
- Pfannkuch, M., & Wild, C. J. (2000). Statistical thinking an statistical practice: Themes gleaned from professional statisticians. *Statistical science*, 15(2), 132-152.
- Pfannkuch, M., & Wild, C. (2002). Statistical thinking models. *The University of Auckland. NewZealand. ICOTS6*.
- Walpole, R. E., Myers, R. H. & Myers, S. L. (2012). Probabilidad y estadística para ingeniería y ciencias. (9.a ed). México: Pearson Educación.
- Watkins, A. E., Scheaffer, R. L. & Cobb, G. W. (2008). *Statistics in Action: Understanding a World of Data* (2.a ed). Key Curriculum Press

- Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67(3), 223–265. <https://doi.org/10.1111/j.1751-5823.1999.tb00442.x>
- Wild, C. J., Utts, J. M., & Horton, N. J. (2018). What is statistics? *International handbook of research in statistics education*, 5-36.
- Zieffler, A. S., & Garfield, J. B. (2009). Modeling The Growth of Students' Covariational Reasoning During an Introductory Statistics Course. *Statistics Education Research Journal*, 8(1), 7-31.
- Zieffler, A., Garfield, J., & Fry, E. (2018). What is statistics education? *International handbook of research in statistics education*, 37-70.



## **ANEXOS**



## **Anexo 1**