

---

# USO DE GEOMETRÍA Y TOPOLOGÍA PARA CIENCIA DE DATOS

---

## Reporte Técnico: Análisis Climatológico del Niño y Niña

Eugenio Santisteban Zolezzi - A01720932

Daniel de Zamacona Madero - A01570576

Bryan Esqueda Ruiz - A01235857

Verónica Victoria García de la Fuente - A00830383

Daniel Alfonso Garza Morales - A01232565

11 de junio de 2023

# Índice

<b>1. Introducción</b>	<b>3</b>
1.1. Acerca del niño . . . . .	3
1.2. Acerca de los datos . . . . .	4
1.2.1. Southern Oscillation Index (SOI) . . . . .	4
1.2.2. Equatorial Pacific Sea Surface Temperatures (SST) . . . . .	4
<b>2. Enfoque y objetivos</b>	<b>5</b>
<b>3. Análisis del Diagrama de Persistencia</b>	<b>7</b>
<b>4. Análisis utilizando algoritmo de Mapper</b>	<b>9</b>
<b>5. Conclusiones</b>	<b>10</b>

# 1. Introducción

## 1.1. Acerca del niño

El Niño es un fenómeno climatológico en el océano pacífico el cual consiste en una combinación entre fluctuaciones de la temperatura al nivel del mar y variaciones entre la presión del aire en la zona [for Enviromental Information, 2023]. En este evento, las temperaturas del océano suben, que usualmente provoca un cambio en la presión atmosférica en varias regiones como Tahití y Australia, específicamente en la zona de Darwin. La diferencia de presiones usualmente ocasiona vientos de Este a Oeste para equilibrar la presión a valores estables. Al tener dichos vientos, habrá un aumento de lluvias en la regiones en el poniente del océano pacífico en donde se dirigen dichos vientos y una sequía en las regiones del oriente.

Este fenómeno se divide en cuatro regiones para facilitar su análisis [?].

- **Niño 4:** Ubicado en la región occidental de los territorios donde suceden las fluctuaciones.
- **Niño 3:** Ubicado en la región oriental de los territorios donde suceden las fluctuaciones.
- **Niño 3.4:** Comprende una porción de la región 4 y la región 3.
- **Niño 1+2:** Se encuentra en las regiones costeras cercanas a la línea ecuatorial en América del Sur

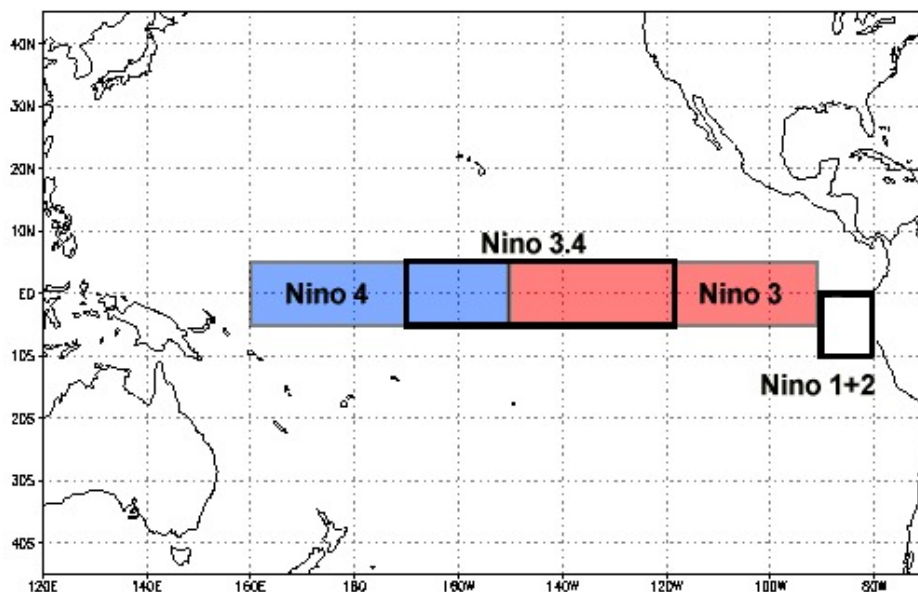


Figura 1: Regiones del niño. (National Centers for Environmental Information, s.f.) [?]

En México se esta observando un estatus moderado del Niño. Sin embargo. el CONAGUA ha pronósticoado que durante este trimestre de Mayo a Julio del 2023 se comenzará a observar los efectos del Niño en el territorio, con un 80 % de que sus efectos sean moderados y un 55 % de que se vuelva intenso y un 5 % de que se siga observando neutral [CONAGUA, 2023].

Si es que hay un cambio de fase en el Niño, esto implicaría una sequía para el territorio Mexicano.

## 1.2. Acerca de los datos

A través de la página oficial de los Centros Nacionales para la Información Ambiental (National Centers for Environmental Information), se puede tener acceso a una variedad de conjuntos de datos diferentes que representan los indicadores del Niño/Niña. Para el fin de este reporte, sólo se analizarán dos variables provenientes de esta base de datos: el Índice para las Oscilaciones del Sur (Southern Oscillation Index) y las Temperaturas del Nivel del Mar en el Pacífico Equatorial (Equatorial Pacific Sea Surface Temperatures). Estos datos abarcan desde Enero de 1982 hasta Abril del 2023.

### 1.2.1. Southern Oscillation Index (SOI)

Esta variable mide la presión del nivel del mar y consiste en dos secciones, una acerca de los datos anómalos y otra con los datos estandarizados del Índice de Oscilación del Sur (SOI por sus siglas en inglés). Ambas secciones se componen de 12 columnas, una para cada mes del año y 80 filas. Es importante denotar que estas variables no dependen de la región con la que se esté tratando, sino que es de forma general en todas las zonas donde se presenta la fluctuación del Niño [for Environmental Information, 2023].

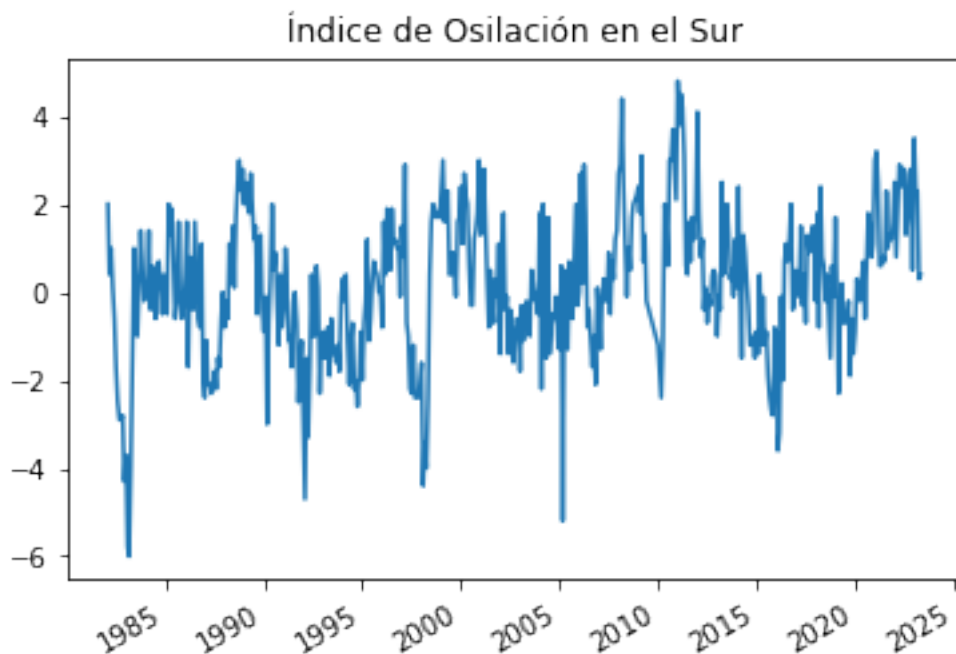


Figura 2: Serie de Tiempo para el Índice de Oscilación del Sur

### 1.2.2. Equatorial Pacific Sea Surface Temperatures (SST)

Esta variable representa la temperatura en la superficie del mar, generalmente se miden utilizando la región 3.4 y generando secciones específicas de tres meses. Se tiene un umbral

de 0.5 centígrados con respecto a la predicción de temperatura media para esos meses para considerar que se está presentando el fenómeno de El Niño al ser temperaturas por encima de lo esperado o en caso de superar el umbral por debajo de lo esperado el fenómeno sería el de La Niña. Para el SST se presentan dos conjuntos de datos [?]:

- **Oceanic Niño Index (ONI)** Datos que consisten en cuatro columnas, representando los periodos de los meses comprendidos, el año, el total (temperatura promedio en la región 3.4 durante ese trimestre) y su anomalía, esto quiere decir qué tantos grados se aleja de lo esperado para ese trimestre.
- **Niño Regions Sea Surface Temperatures (SST)** El conjunto de datos consiste en el valor de la temperatura en cada una de las regiones del niño y la cantidad de centígrados de diferencia con respecto a lo esperado para cada mes del año.

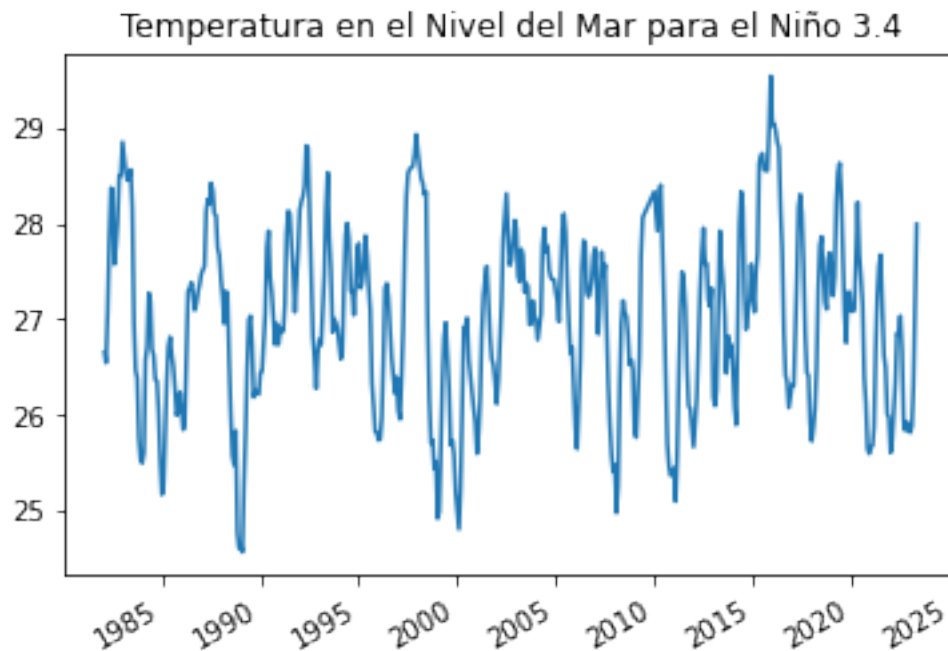


Figura 3: Serie de Tiempo para las Temperaturas en el Nivel del Mar para el Niño 3.4

Un detalle que recalcar de estas dos variables: El promedio de la Anomalía en las Temperaturas es de -0.094 con desviación estándar de 0.829. El promedio del Índice de Oscilación es de 0.155 con desviación estándar de 1.64. Con estos datos, se puede decir con un alto nivel de confianza que la media de estas dos variables es de 0.

## 2. Enfoque y objetivos

Se han realizado investigaciones estudiado las variaciones de la temperatura a lo largo del tiempo para identificar patrones relacionados con El Niño. Esto implica el análisis de registros históricos de temperatura en diferentes regiones y períodos de tiempo para detectar cambios consistentes con la presencia de El Niño, esto lo podríamos reconocer como una

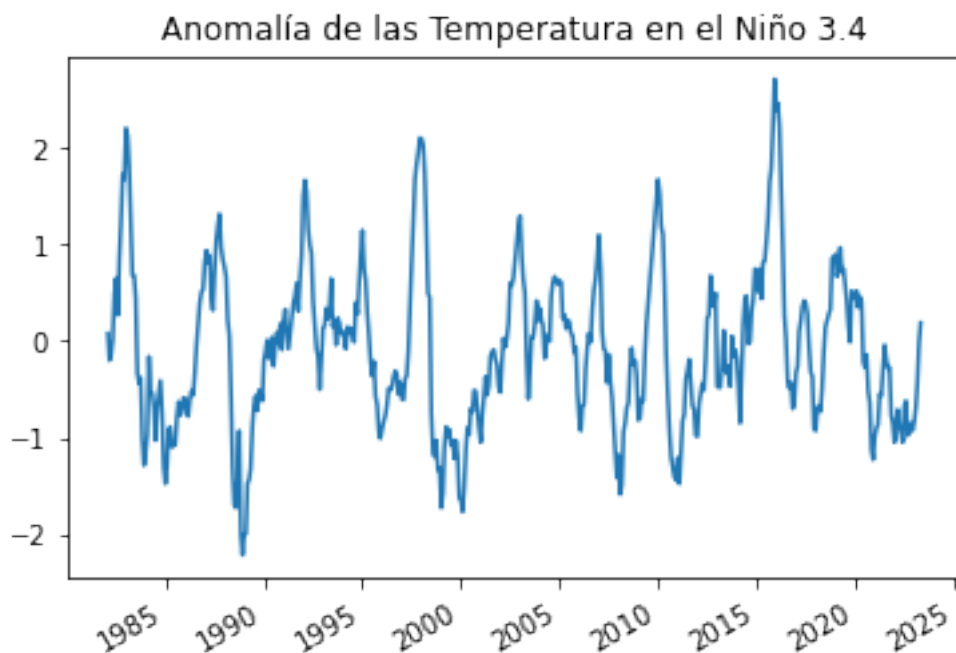


Figura 4: Serie de Tiempo para la Anomalía de las Temperaturas en el Nivel del Mar del Niño 3.4

investigación de análisis de series de tiempo. Una investigación importante que fue publicada por M. J. McPhaden llamada “El Niño and its relationship to changing background conditions in the tropical Pacific Ocean” nos habla de como el fenomeno climatológico conocido como El Niño ha sido objeto de numerosos estudios debido a su influencia en el clima global y sus impactos en diversas regiones del mundo, y como en es su investigacion se lleva a cabo un análisis de series de tiempo de las variaciones de temperatura a lo largo del tiempo con el objetivo de identificar patrones relacionados con el fenómeno de El Niño. En esta investigación se utilizan datos históricos de temperatura de diferentes estaciones meteorológicas ubicadas en áreas afectadas por El Niño para con esto poder realizar el análisis correspondiente. [Mcphaden, 2011]

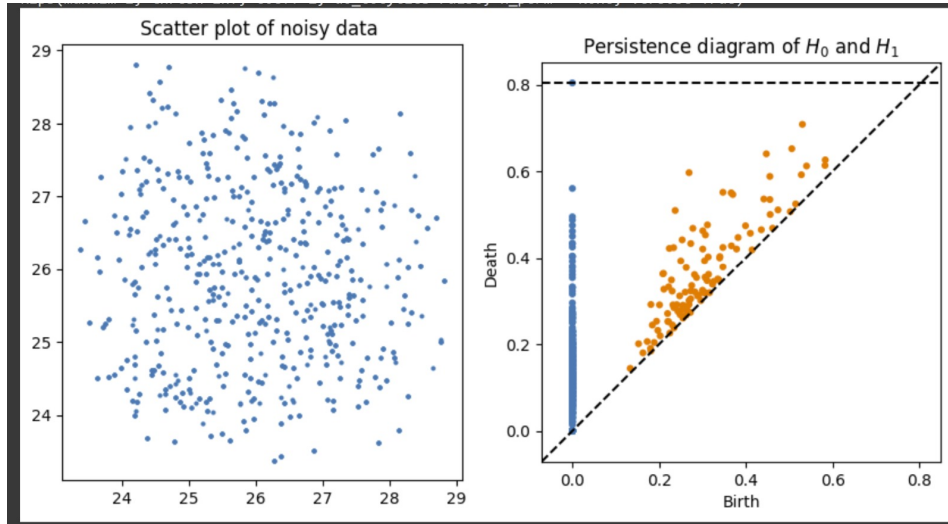
Otro tipo de investigaciones que se han realizado son aquellas en las que se implementa un análisis de correlación, en estas han realizado análisis estadísticos para evaluar la relación entre la temperatura y otras variables climáticas durante eventos de El Niño. Algunos ejemplos que podríamos examinar podrían ser cómo las anomalías de temperatura se correlacionan con los patrones de precipitación, los cambios en los vientos o las variaciones en la presión atmosférica. Una investigación en la cual utilizan este tipo de análisis fue hecha por Yatagai y Yasunari, en esta se centran en la relación entre la temperatura y la precipitación durante los eventos de El Niño en el sureste asiático, en esta investigación se recopilaron datos climáticos de estaciones meteorológicas ubicadas en varios países del sureste asiático, incluyendo Indonesia, Filipinas, entre otros. Ellos utilizaron un análisis de correlación para examinar la relación entre la temperatura y la precipitación durante los eventos de El Niño, y los resultados que obtuvieron fueron una correlación negativa significativa entre la temperatura y la precipitación, lo que indicaba que durante los eventos de El Niño, se experimentaba un aumento de la temperatura junto con una disminución de la precipitación

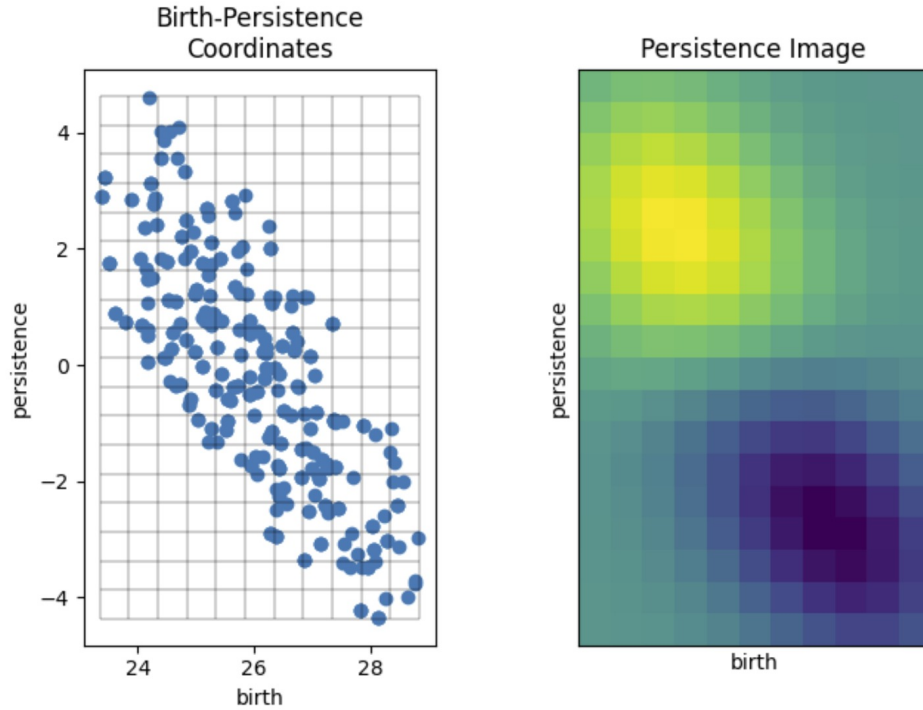
en la región. [Yatagi, 2019]

Por último, existe las investigaciones donde se utilizan modelos computacionales para simular la interacción entre el océano y la atmósfera durante eventos de El Niño. Estos modelos incorporan datos observacionales y principios físicos para predecir cómo las anomalías de temperatura en el océano Pacífico tropical pueden influir en los patrones climáticos globales y regionales. Como se puede observar en el artículo de Nan Chen [Chen, 2022], se construyó un modelo el cual combina procesos semanales, anuales y decenales para explicar las características espaciales, de intensidad máxima y evolución temporal del ENSO. Este modelo permite comprender y predecir la complejidad del ENSO la cual hasta este día no es fácil de simular y se sigue buscando nuevos modelos para encontrar uno que realice predicciones adecuadas.

### 3. Análisis del Diagrama de Persistencia

Al realizar un diagrama de persistencia para cada una de las variables, nos permite visualizar la evolución de las características topológicas de los datos a lo largo del tiempo. Ahora, debemos analizar si las variables seleccionadas presentan persistencia o no. Para realizar este análisis, nos enfocamos en las gráficas obtenidas para las variables, donde se nos muestran los datos en los niveles  $H_0$ ,  $H_1$  y  $H_2$ . Nos centramos especialmente en los niveles  $H_1$  y  $H_2$ , ya que nos brindan información sobre los agujeros y las esferas presentes en los datos. Al observar estas gráficas, podemos notar que en algunas variables el tiempo de nacimiento y muerte de las características topológicas es relativamente alto. Esto significa que ciertas estructuras topológicas, como los agujeros y las esferas, existen durante períodos de tiempo prolongados. Esta persistencia en las características topológicas indica que los datos presentan cierta estabilidad en su forma y estructura a lo largo del tiempo. Además, al observar las gráficas, es posible identificar patrones repetitivos en las características topológicas. Estos patrones sugieren la presencia de regularidades o ciclos en los datos. La existencia de patrones repetitivos junto con la persistencia de características topológicas nos indica que los datos exhiben periodicidad.





Cada componentes conexos representa un grupo de puntos en el conjunto de datos que están conectados entre sí. Estos componentes conexos pueden ser considerados como regiones o áreas en el espacio de datos donde los puntos están agrupados de manera continua. Las componentes disconexas en un diagrama de persistencia representan los agujeros. Una componente disconexa indica la formación de un agujero en el conjunto de datos. Estos agujeros pueden ser considerados como regiones vacías o espacios rodeados de puntos.

La persistencia la podemos determinar observando la distancia horizontal entre los componentes convexos y disconexos correspondientes. Si la distancia entre un punto 0 y un punto 8 es grande, indica una persistencia prolongada del componente conexo o del agujero. Por otro lado, si la distancia entre ellos es pequeña, indica que el componente conexo o el agujero es menos persistente o se forma y desaparece rápidamente.

Este análisis de persistencia nos ayuda a entender mejor los conjuntos de datos y las formas en que se conectan entre sí, y específicamente en este caso nos es útil en el reconocimiento de patrones. Ya que este análisis nos permite identificar características importantes en los datos, como agujeros o estructuras conectadas, y nos ayuda a estudiar cómo cambian estas características a lo largo del tiempo para clasificar y reconocer patrones en los datos. Por lo tanto este análisis nos es importante ya que nos permite identificar patrones y ciclos que nos ayudaran a predecir y pronosticar como se comportaran los eventos de EL NIÑO, y además nos permite una mejor comprensión de la duración y la frecuencia de los eventos de El Niño, y cómo se relacionan con otros factores climáticos



## 4. Análisis utilizando algoritmo de Mapper

De forma general el algoritmo Mapper es utilizado para tener una representación visual de conjuntos específicos de datos, así como observar patrones y tener una comprensión más amplia de la relación dentro de los mismos datos. En el caso de este proyecto, además de utilizar el método de persistencia, se utilizó el algoritmo Mapper para generar agrupaciones entre clusters compuestos de años en los cuales la presión atmosférica en el nivel del mar y la temperatura del mismo sean similares entre los años de cada conjunto.

Entre los diferentes clusters obtenidos se observó una tendencia que muestra que efectivamente existe una correlación negativa entre las variables contempladas.

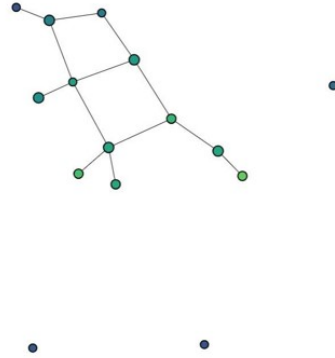


Figura 5: Representación en grafos de agrupaciones de los datos utilizando el algoritmo Mapper.

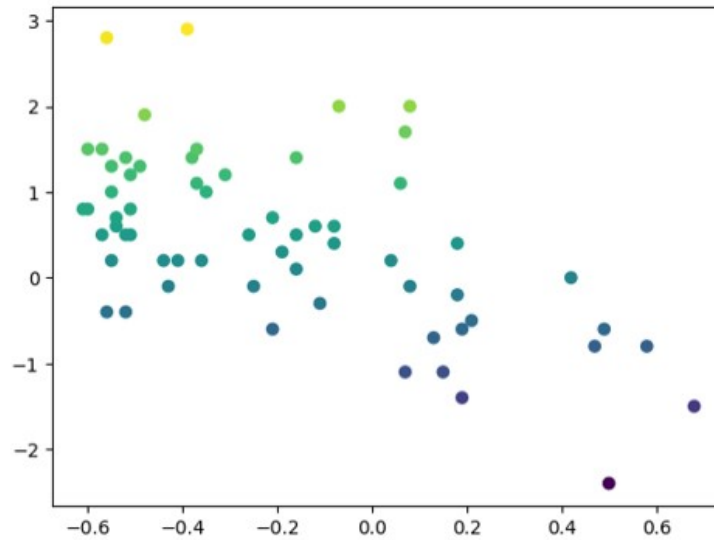


Figura 6: Ejemplo de Cluster que nos permite observar la relación entre la presión y la temperatura, donde se tiene como eje X el índice de la presión y eje Y las anomalías en la temperatura.

## 5. Conclusiones

A través de la implementación de los dos métodos se logró determinar que la hipótesis inicial, la cual denotaba que debía haber una correlación negativa entre las variables de la presión atmosférica en el nivel del mar y la temperatura en el mismo se cumple. Esto también es fundamentado a través de nuestra documentación donde se investigó que la temperatura encontraba afectaba a su vez el movimiento de las corrientes del aire, por ejemplo una presión baja en un espacio y una presión alta en el opuesto fomentaría la circulación de aire en cierto sentido, donde el sentido indicaría el lugar donde está el aumento en la temperatura. En nuestra conclusión, proponemos desarrollar un modelo de documentación mejorado para comprender las influencias globales de los fenómenos investigados, apoyándonos en la efectividad de nuestra agrupación. Tomando en cuenta que los modelos interdecadales de la circulación océano-atmósfera en el Pacífico ejercen una influencia significativa en la temperatura global. Por otra parte, al realizar la implementación del algoritmo de Mapper observamos que también permitía generar una clasificación confiable entre los fenómenos del Niño y la Niña agrupando por año aquellos que forman parte del mismo fenómeno. Por lo tanto, se propone para futuras investigaciones analizar este comportamiento y comprobar su eficacia.

## Referencias

- [Chen, 2022] Chen, N. (2022). A multiscale model for el niño complexity. <https://www.nature.com/articles/s41612-022-00241-x>.
- [CONAGUA, 2023] CONAGUA (2023). El niño osilación del sur (enos). <https://smn.conagua.gob.mx/es/climatologia/diagnostico-climatico/estatus-del-nino>.
- [for Enviromental Information, 2023] for Enviromental Information, N. C. (2023). El niño/southern oscillation (enso). <https://www.ncei.noaa.gov/access/monitoring/enso/olr>.
- [for Environmental Information, 2023] for Environmental Information, N. C. (2023). Southern oscillation index (soi). <https://www.ncei.noaa.gov/access/monitoring/enso/soi>.
- [Mcphaden, 2011] Mcphaden, M. J. (2011). El niño and its relationship to changing background conditions in the tropical pacific ocean. <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2011GL048275>.
- [Yatagi, 2019] Yatagi, A. e. a. (2019). El ninio-related summer precipitation anomalies in southeast asia modulated by the atlantic multidecadal oscillation. [https://www.researchgate.net/publication/335412508\\_El\\_NinoRelated\\_Summer\\_Precipitation\\_Anomalies\\_in\\_Southeast\\_Asia\\_Modulated\\_by\\_the\\_Atlantic\\_Multidecadal\\_Oscillation](https://www.researchgate.net/publication/335412508_El_NinoRelated_Summer_Precipitation_Anomalies_in_Southeast_Asia_Modulated_by_the_Atlantic_Multidecadal_Oscillation).