

## Water Quality Data

El presente proyecto consiste en analizar datos referentes a la calidad de agua. Para ello se utilizó una base de datos obtenida de Kaggle, facilitando la recopilación de la información que se requiere analizar. Al descargar la base de datos, se importó a Google Sheet para una mejor manipulación.

La base de datos contiene registros detallados de los parámetros de calidad del agua recopilados en el Refugio por voluntarios una vez cada dos semanas. Los parámetros incluyen turbidez, pH, oxígeno disuelto (OD), salinidad y temperatura. La toma de muestras se realiza en lugares designados dentro de diversos cuerpos de agua como la Bahía, D-Pool (estanque de pesca), C-Pool, B-Pool y A-Pool. El conjunto de datos se mantiene para garantizar la precisión y fiabilidad en el monitoreo de la calidad del agua a lo largo del tiempo. (Kaggle,2025)

El registro de los datos fue capturada una vez cada dos semanas, por lo que la información obtenida del año 1989 a 2019 requiere de un análisis amplio, por lo que solo se utilizó los datos de **Año, Ph y Temperatura del agua**, medida en grados centígrados; además se simplificó los datos por año, ya que la base de datos nos arroja la información cada dos semanas.

A	B	C	D	E	F	G	H
Date	Salinity (ppt)	DissolvedOxygen (mg/L)	pH	SecchiDepth (m)	WaterDepth (m)	WaterTemp (C)	AirTemp (C)
11/05/1989				7.5	0.3	0.9	17 -17.7777778
18/05/1989		12	7.5	0.2	0.6	17.5	20.5
25/05/1989				8	0.4	0.8	23
01/06/1989		12	8	0.4	0.9	25.5	29
11/07/1989				8.5	0.3	0.9	28.5
20/07/1989		1.6	9	0.3	0.9	30	32.5
27/07/1989		6.2	9	0.3	0.8	27	27
03/08/1989		5.8	9.5	0.2	0.7	25	28
10/08/1989		3.2	9	0.2	0.3	21.5	22.5
17/08/1989		7.3	9	0.2	0.8	28	29
31/08/1989		7.1	9	0.2	0.8	28	28
14/09/1989		6	7.5	0.3	0.9	26	26
25/10/1989		7.8	8.5	0.2	0.5	15	19
31/10/1989				7.5	0.2	0.7	21.5
20/11/1989				8	0.2	1	8.5
27/11/1989		10.3	7.5	0.2	0.6	9	4
11/12/1989		10.2	7	0.2	0.4	3	3
03/01/1990		11.9	7.5	0.2	0.6	3	5
08/01/1990		10.8	7.5	0.2	0.4	7	7.5
22/01/1990	1	9.7	7.5	0.15	0.65	10	12
30/01/1990		10.8	7.5	0.2	0.7	9	13
07/02/1990		3.4	10.8	7.5	0.2	0.6	10
12/02/1990		3.2	10.6	7	0.2	0.7	10

Base de Datos en Kaggle

Base de Datos en GoggleSheet

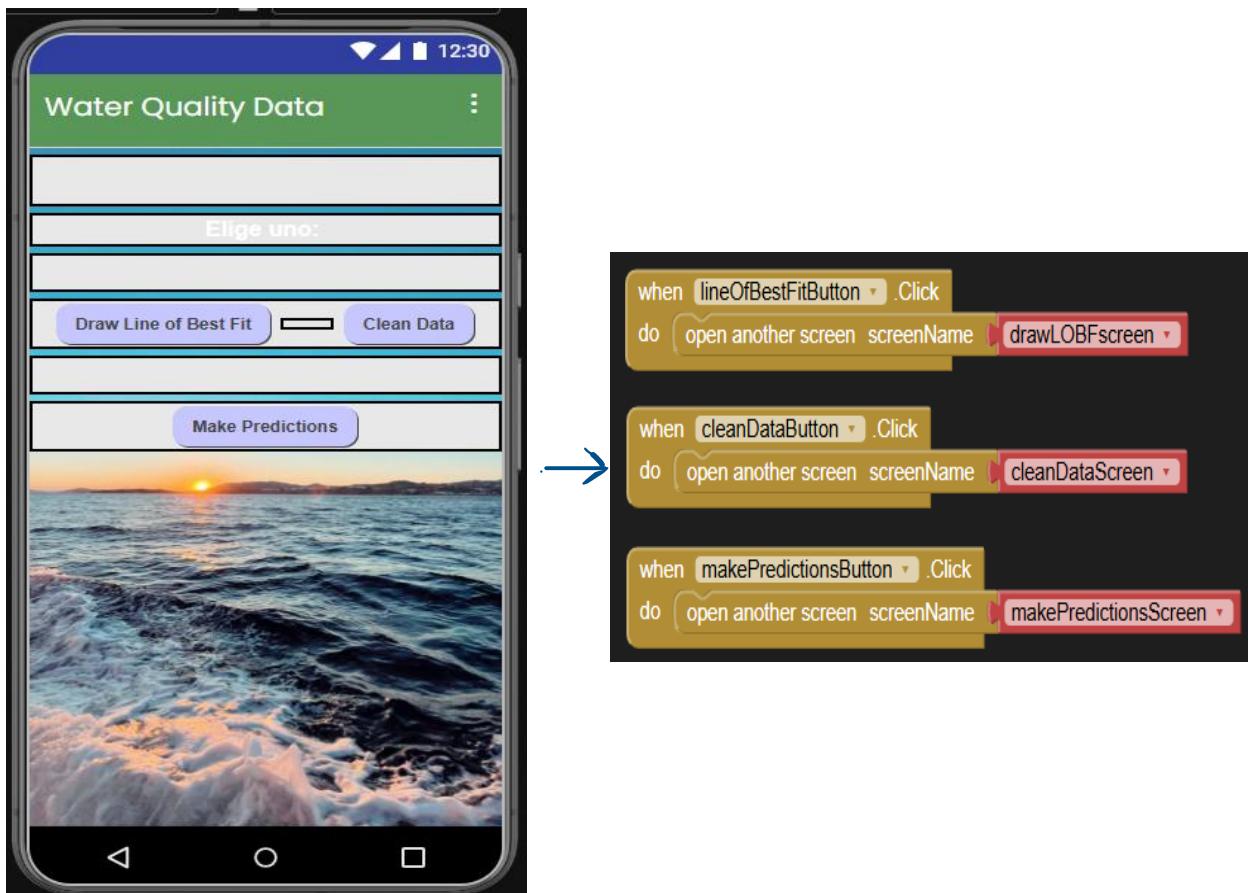
	A	B	C	D
1	Data	Year	pH	WaterTemp(C)
2	WaterQuality	1989	140	354
3	WaterQuality	1990	271.75	612
4	WaterQuality	1991	327.9	695.5
5	WaterQuality	1992	335.35	688
6	WaterQuality	1993	292.2	686.9
7	WaterQuality	1994	305	683.2
8	WaterQuality	1995	311.1	755
9	WaterQuality	1996	247.6	543.21
10	WaterQuality	1997	223.9	569.2
11	WaterQuality	1998	107.2	286.4
12	WaterQuality	1999	355.4	854.5
13	WaterQuality	2000	850	2036.4
14	WaterQuality	2001	855.7	2100.5
15	WaterQuality	2002	740.6	1248.1
16	WaterQuality	2003	641.7	1565.1
17	WaterQuality	2004	803.2	2002.7
18	WaterQuality	2005	726.5	2017.9
19	WaterQuality	2006	772.5	2116.5
20	WaterQuality	2007	751.2	2070
21	WaterQuality	2008	795.9	1984
22	WaterQuality	2009	813.8	1989.4
23	WaterQuality	2010	684	1833
24	WaterQuality	2011	650	1797
25	WaterQuality	2012	617.4	1490
26	WaterQuality	2013	612.2	1519
27	WaterQuality	2014	597	1389

Para la realización y ejecución de la aplicación se utilizó App Inventor, una plataforma de desarrollo de aplicaciones móviles, así como un dispositivo móvil para realizar las pruebas correspondientes de la aplicación.

## Diseño y Código de la Aplicación

### Screen 1

Se muestra el menú para acceder a las diferentes pantallas, así como el código que realiza su funcionamiento.



### Screen 2 – Draw Line Of Best Fit

En la pantalla 2 se puede visualizar la creación del modelo, dónde se muestran los datos y la línea de mejor ajuste.

```

when [drawLineOfBestFitButton] .Click
do
  set [topTrendline] .ChartData to [topChartData2D]
  set [bottomTrendline] .ChartData to [bottomChartData2D]
  set [topSlopeValueLabel] .Text to [topTrendline] .LinearCoefficient
  set [topY_intValueLabel] .Text to [topTrendline] .YIntercept
  set [topCor_coefValueLabel] .Text to [topTrendline] .CorrelationCoefficient
  set [bottomSlopeValueLabel] .Text to [bottomTrendline] .LinearCoefficient
  set [bottomY_intValueLabel] .Text to [bottomTrendline] .YIntercept
  set [bottomCor_coefValueLabel] .Text to [bottomTrendline] .CorrelationCoefficient
end

```

```

when [showDataButton] .Click
do
  call [topChartData2D] .Clear
  call [bottomChartData2D] .Clear
  call [spreadsheet1] .ReadSheet
  sheetName [WaterQuality]
end

```

```

when [homeButton] .Click
do
  open another screen screenName [Screen1]
end

```

### Screen 3 – Clean Data

En la pantalla 3, se muestran los datos anteriores, pero se agrega un botón *Detect Anomalies*, que permite detectar anomalías o datos atípicos que se obtienen en los registros de la base datos. En este caso, se estableció para detectar 1 anomalía.

The diagram illustrates the mobile application "Cleaning the Data" and its corresponding Scratch script.

**Mobile Application Interface:**

- Top Bar:** Cleaning the Data
- Buttons:** Home, Show Data, Detect Anomalies, Draw Line of Best Fit
- Data Cleaning Window:** A graph showing four data points (1,0), (2,1), (3,2), (4,3) connected by a line. Below the graph are three input fields: M =, B =, and R =.
- Bottom Buttons:** Back, Home, Stop

**Scratch Script (Top Panel):**

```

when homeButton [ ] .Click
do [open another screen screenName Screen1]

```

**Scratch Script (Middle Panel):**

```

when showDataButton [ ] .Click
do [call cleanedChartData2D [Clear v] 
call dataCleaningChartData2D [Clear v] 
call spreadsheet1 [ReadSheet v] 
sheetName WaterQuality]

```

**Scratch Script (Bottom Panel):**

```

when spreadsheet1 [GotSheetData]
sheetData
do [set topChartLabel [Text v] to [0] 
call cleanedChartData2D [ImportFromSpreadsheet v] 
spreadsheet spreadsheet1 
xColumn Year 
yColumn pH 
useHeaders true] 
call dataCleaningChartData2D [ImportFromSpreadsheet v] 
spreadsheet spreadsheet1 
xColumn Year 
yColumn WaterTemp(C) 
useHeaders true]

when drawLineOfBestFitButton [ ] .Click
do [set Trendline1 [ChartData v] to [cleanedChartData2D]]

```

**Scratch Script (Bottom Panel):**

```

when detectAnomaliesButton [ ] .Click
do [call dataCleaningChartData2D [HighlightDataPoints v] 
dataPoints color red]

```

**Scratch Script (Middle Panel):**

```

when dataCleaningChartData2D [EntryClick]
x y
do [if is in list? thing [make a list [get x] [get y]] 
list call AnomalyDetection1 [DetectAnomaliesInChartData v] 
chartData dataCleaningChartData2D 
threshold 1]
then [call dataCleaningChartData2D [RemoveEntry v] 
x get x 
y get y] 
call cleanedChartData2D [Clear v] 
call cleanedChartData2D [ImportFromList v] 
list call dataCleaningChartData2D [GetAllEntries v]

```

**Scratch Script (Bottom Panel):**

```

when Trendline1 [Updated]
results
do [set topSlopeValueLabel [Text v] to [Trendline1 [LinearCoefficient v]] 
set topY_intValueLabel [Text v] to [Trendline1 [YIntercept v]] 
set topCor_coefValueLabel [Text v] to [Trendline1 [CorrelationCoefficient v]]]

```

#### Screen 4 – Make Predictions

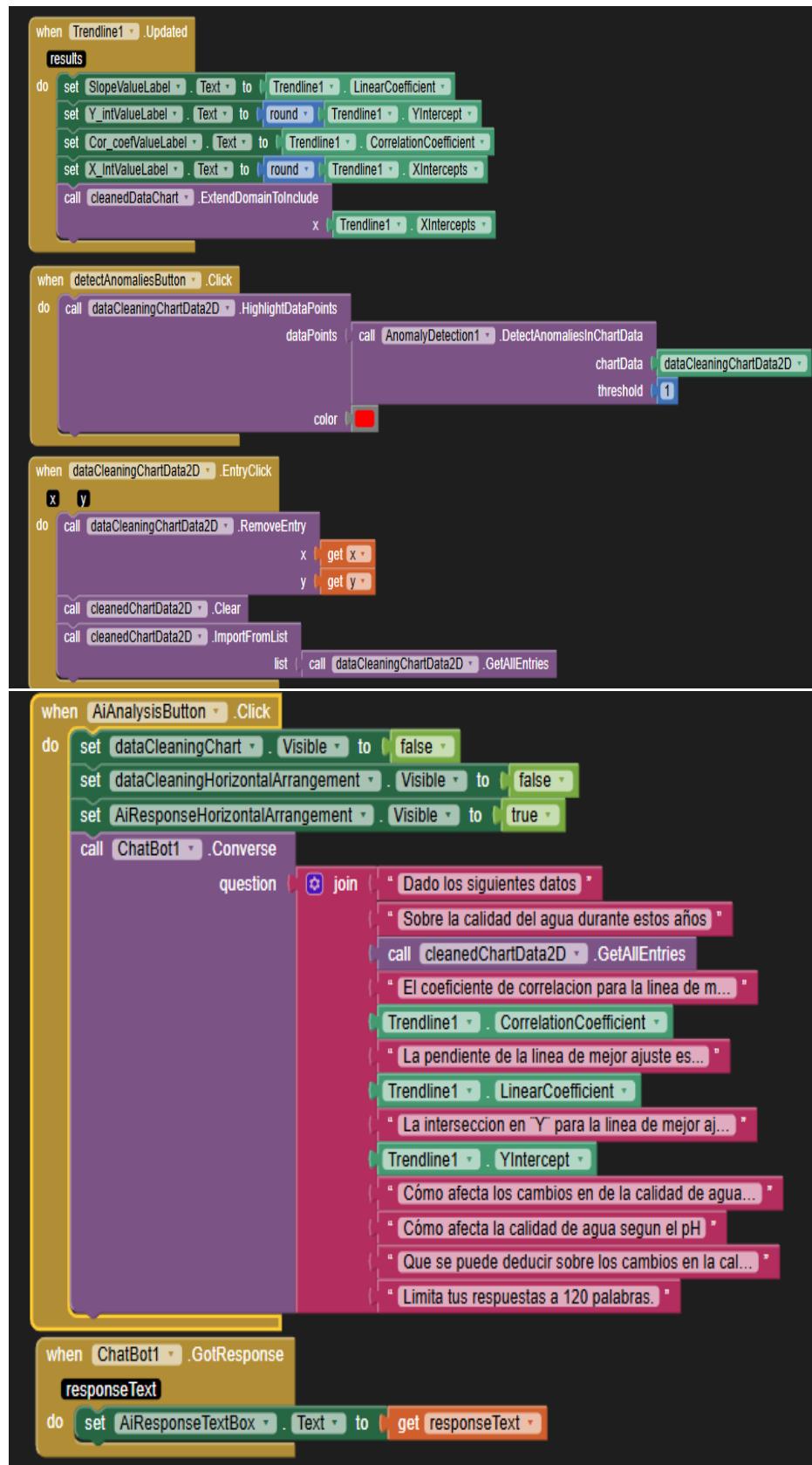
En la pantalla 4 se agrega el ultimo botón *IA Analysis*, donde con la ayuda de la inteligencia artificial, podrá responder a preguntas ya establecidas dentro del código.



Preguntas establecidas en el código:

1. ¿Cómo afecta los cambios en la calidad de agua en la población?
2. ¿Cómo afecta la calidad de agua según el pH?
3. ¿Qué se puede deducir sobre los cambios en la calidad del agua?

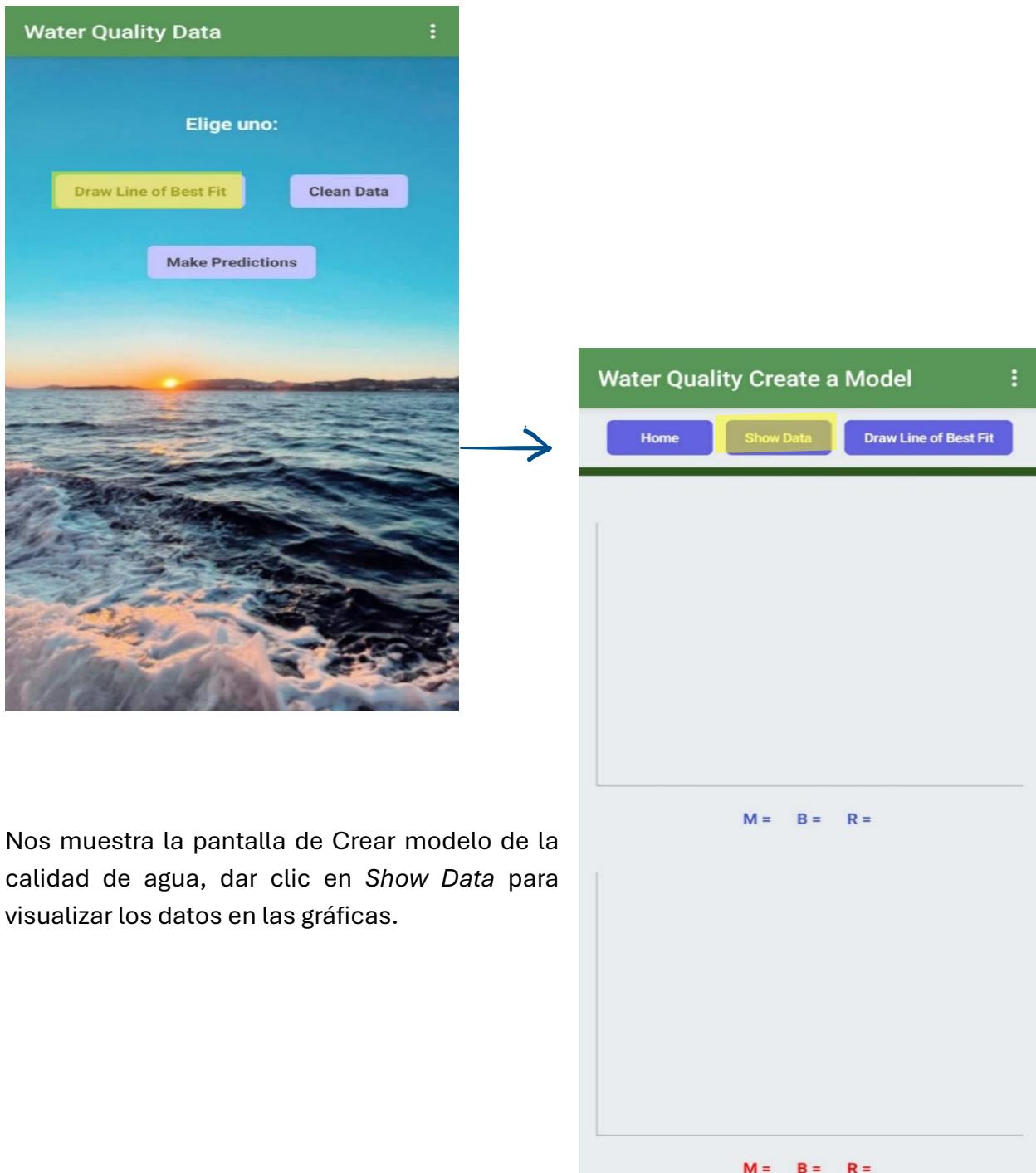
# KAREN CRUZ SANCHEZ



## Fucionamiento

Para el funcionamiento de la aplicación se utilizó un dispositivo móvil que su sistema operativo sea Android, para una mejor compatibilidad.

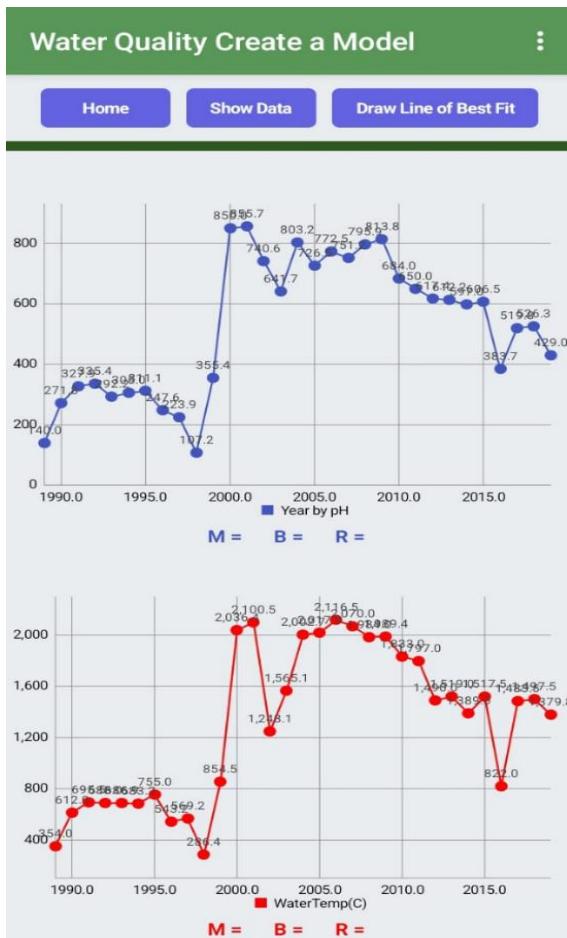
Al ejecutar la aplicación muestra la pantalla 1, dónde se visualiza el menú. Dar clic sobre el botón *Draw Line of Best Fit*.



Nos muestra la pantalla de Crear modelo de la calidad de agua, dar clic en *Show Data* para visualizar los datos en las gráficas.

## KAREN CRUZ SANCHEZ

En los datos observamos el pH según el año, así como la temperatura que presenta la calidad de agua.



Para visualizar la línea de mejor ajuste, solo es cuestión de dar clic en el botón *Draw Line of Best Fit*. Al final se observa el trazo en cada grafica.

## KAREN CRUZ SANCHEZ

Al conocer los datos anteriores, damos clic en *Home* para regresar a la pantalla principal. Luego seleccionar el botón *Clic Data*.

Dentro de la pantalla *Cleaning the Data*, se observarán los botones anteriores, que al seleccionar cada uno mostrará la información anterior, después de seguir el procedimiento antes mostrado, finalmente llegamos al botón *Detect Anomalies*, dar clic para visualizar las anomalías o datos atípicos de la información,



En este caso, se observan anomalías en la gráfica de temperatura en distintos años. Nuevamente, dar clic en el botón *Home*, para regresar a la pantalla principal.

KAREN CRUZ SANCHEZ

En la pantalla principal, seleccionar el último botón restante *Make Predictions*.

Seleccionar el botón *AI Analysis* para hacer uso de la Inteligencia Artificial, donde responderá las preguntas ya establecidas.

**Water Quality: Making Predictions**

Home Show Data

Detect Anomalies Draw LOBF AI Analysis

Data Cleaning: Click on anomalies in this window to remove.

**Water Quality: Making Predictions**

Home Show Data

Detect Anomalies Draw LOBF AI Analysis

El coeficiente de correlación de 0.50 indica una relación moderada entre el tiempo y la calidad del agua. La pendiente positiva (12.65) sugiere que, en general, la calidad del agua ha empeorado con el tiempo.

Los cambios en la calidad del agua pueden afectar a la población al:

- \*\*Dañar la salud:\*\* Agua contaminada puede causar enfermedades.
- \*\*Restringir acceso:\*\* La escasez de agua potable afecta la vida diaria.
- \*\*Impacto económico:\*\* Sectores como la agricultura y el turismo sufren.

El pH del agua es crucial para su calidad: niveles muy ácidos o básicos pueden ser dañinos. Se puede deducir que la

**Water Quality: Making Predictions**

Home Show Data

Detect Anomalies Draw LOBF AI Analysis

Data Cleaning: Click on anomalies in this window to remove.

**Water Quality: Making Predictions**

Home Show Data

Detect Anomalies Draw LOBF AI Analysis

Data Cleaning: Click on anomalies in this window to remove.

*Fuentes de información - URL*

**Base de Datos en Kaggle**

- <https://www.kaggle.com/datasets/supriyoain/water-quality-data>