

## 1. Resumen

Dentro de este análisis se aborda la problemática de la concentración de mercurio en peces que viven en lagos específicos del mundo. Esto es algo muy relevante para la sociedad ya que puede afectar directamente nuestra salud, pues altos niveles de mercurio en el organismo del cuerpo humano puede ser muy dañino para la salud. En este pequeño reporte se utilizan datos de estos lagos, con algunas de sus características más importantes. De forma general, lo que se realizó fue un análisis de normalidad y un análisis de componentes principales, para poder entender más los datos y la relación que existe entre las variables, llegando así a conclusiones importantes, como que el PH es un factor importante dentro de la concentración de mercurio que hay en los peces, así como también se pudieron identificar características importantes de las variables cómo sí existía normalidad o no.

## 2. Introducción

Por algunos años se ha estudiado el fenómeno del mercurio en los peces, pero sobre todo se ha intentado entender qué variables contribuyen o influyen en esta contaminación, para así poder crear medidas preventivas, o para detenerlo a tiempo. Se ha hablado de qué depende incluso de la edad de los peces, sin embargo no se sabe a ciencia cierta, pero esto es posible analizarlo con métodos de estadística, para poder entender mejor cuales son aquellos factores más importantes.

La importancia de abordar este problema radica en que la ingesta de mercurio a través de peces y alimentos del mar es actualmente un problema de salud pública, dada su toxicidad en el desarrollo neurológico en fetos y niños.

## 3. Análisis de los resultados

### Análisis de normalidad

Lo primero que se hizo fue un análisis de normalidad, utilizando la prueba Mardia y de Anderson Darling.

### **Prueba de Mardia:**

Test <chr>	Statistic <fctr>	p value <fctr>	Result <chr>
Mardia Skewness	410.214790601478	7.04198777815408e-23	NO
Mardia Kurtosis	4.59612555772731	4.30419392238868e-06	NO
MVN	NA	NA	NO

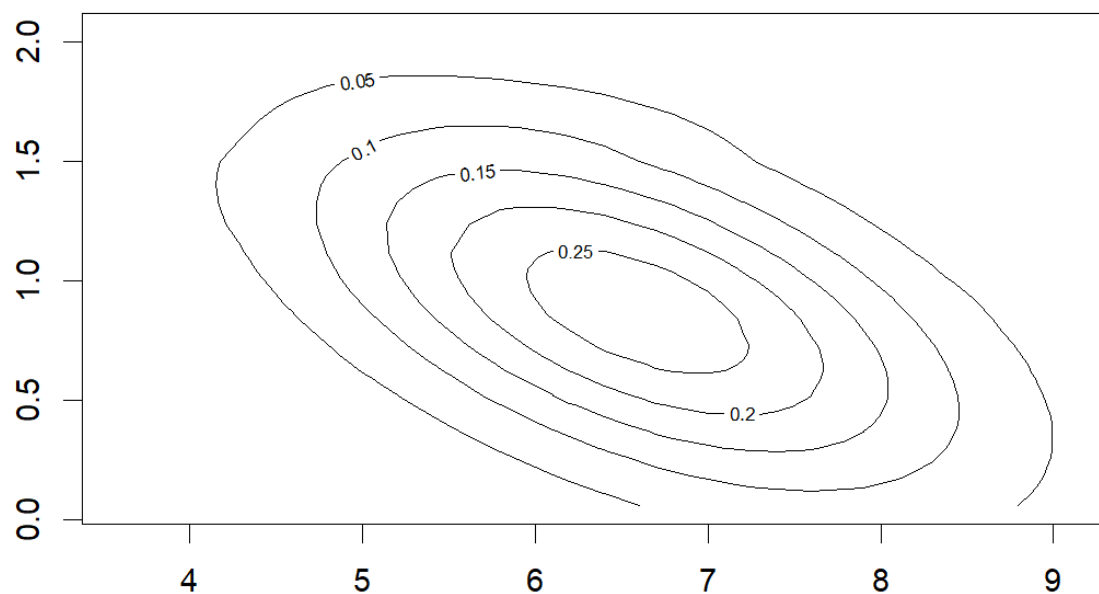
### Prueba de Anderson Darling:

	Test <S3: Asls>	Variable <S3: Asls>	Statistic <S3: Asls>	p value <S3: Asls>	Normality <S3: Asls>
1	Anderson-Darling	X3	3.6725	<0.001	NO
2	Anderson-Darling	X4	0.3496	0.4611	YES
3	Anderson-Darling	X5	4.0510	<0.001	NO
4	Anderson-Darling	X6	5.4286	<0.001	NO
5	Anderson-Darling	X7	0.9253	0.0174	NO
6	Anderson-Darling	X8	8.6943	<0.001	NO
7	Anderson-Darling	X9	1.9770	<0.001	NO
8	Anderson-Darling	X10	0.6585	0.081	YES
9	Anderson-Darling	X11	1.0469	0.0086	NO

Como se puede observar, se obtuvo que solamente 2 variables son normales. Después se volvió a realizar esta prueba, pero ahora solo tomando en cuenta estas dos variables.

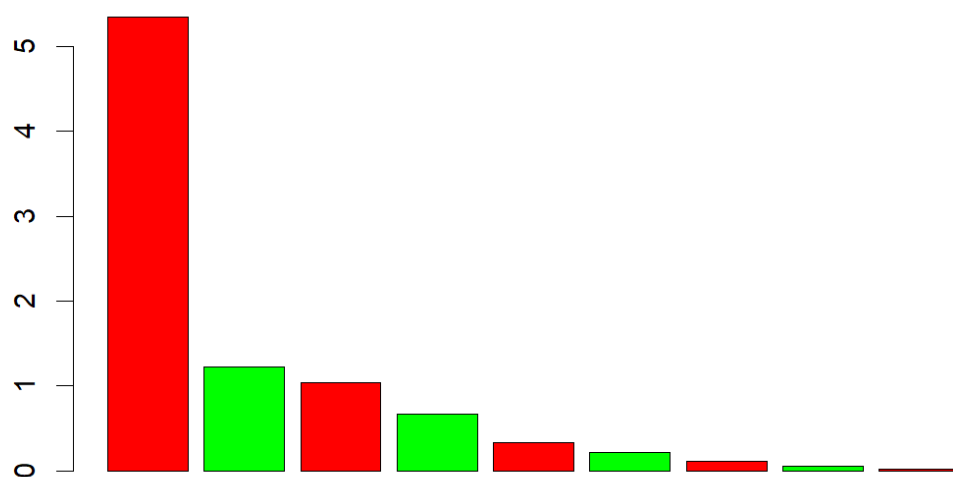
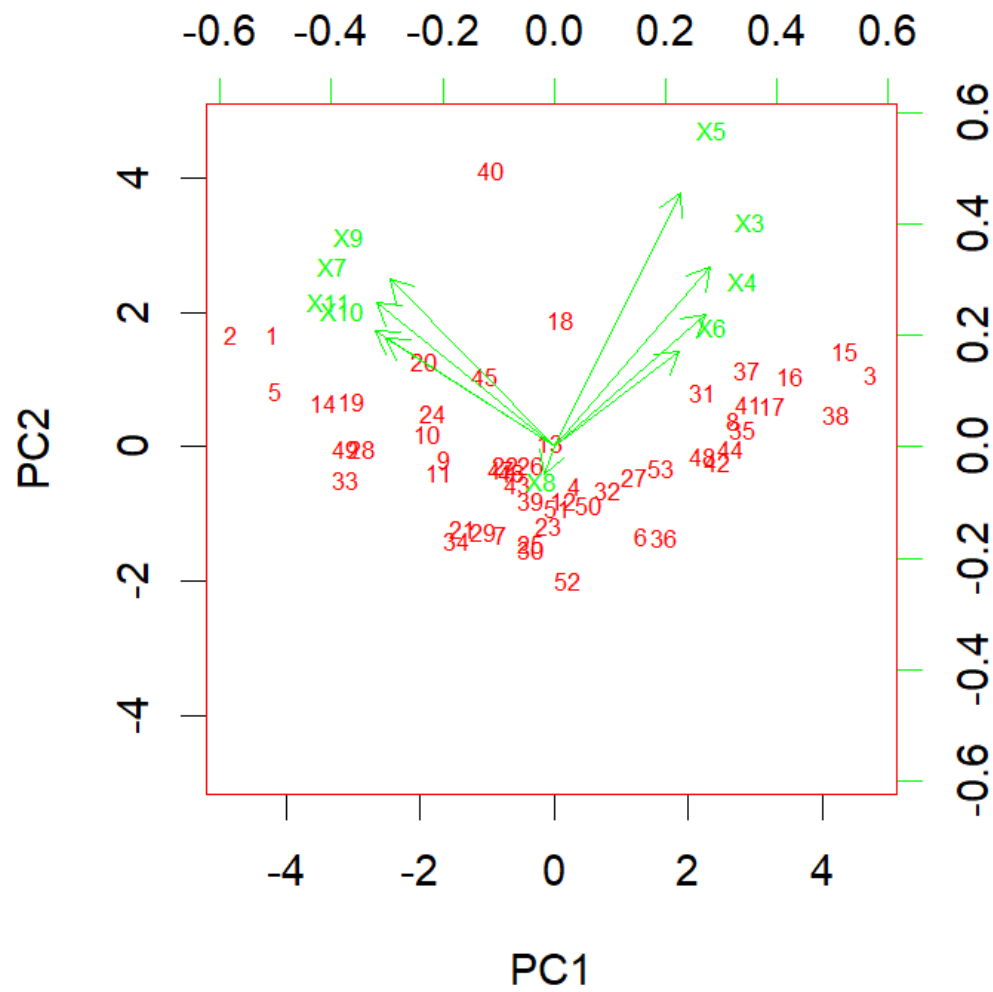
Test <chr>	Statistic <fctr>	p value <fctr>	Result <chr>
Mardia Skewness	6.17538668676458	0.186427564928852	YES
Mardia Kurtosis	-1.12820795824432	0.25923210375991	YES
MVN	NA	NA	YES

Como siguiente paso, se realizó la gráfica de contorno de la normal multivariada que se obtuvo en el paso anterior.



### Análisis de componentes principales

Este método se utilizó para poder identificar aquellas variables más influyentes en nuestro estudio, y así generar nuevas variables más descriptivas de los datos.



Importance of components:

	PC1	PC2	PC3	PC4	PC5
standard deviation	2.312	1.1049	1.0210	0.81723	0.5794
Proportion of Variance	0.594	0.1357	0.1158	0.07421	0.0373
Cumulative Proportion	0.594	0.7297	0.8455	0.91969	0.9570
	PC6	PC7	PC8	PC9	
Standard deviation	0.45710	0.32750	0.22810	0.13731	
Proportion of Variance	0.02322	0.01192	0.00578	0.00209	
Cumulative Proportion	0.98021	0.99212	0.99791	1.00000	

Como se puede observar en el segundo gráfico, la primera componente ya cubre el 50% de la variabilidad de los datos, mientras que la segunda cubre otro 13% de esta misma. Si tomamos valores absolutos, la variable 2 es la que más influye en el primer componente, esta variable es el PH.

#### 4. Conclusiones

La utilización de métodos estadísticos para analizar problemáticas de la vida real, ayuda a poder entender mejor los factores que influyen, así como la relación que existe entre cada una de las variables, dentro de este pequeño estudio fue posible simplificar las variables a partir del método de componentes principales, facilitando así la interpretación y además preparando los datos para poder realizar métodos más sofisticados que ayuden a la toma de decisiones.

#### 5. Anexos:

<https://drive.google.com/drive/folders/1F7MecFCo3uoXWbRQ7sOHkOthdycRI22w?usp=sharing>