

**MICROCUENCA DE LA QUEBRADA LA OSPINA/LA MUÑOZ EN EL MUNICIPIO DE LA ESTRELLA,  
ANTIOQUIA, ESCALA 1:2.500**

**CARTOGRAFIA GEOTECNICA**

**David Ricardo Ramírez Martínez**

**Profesor:  
Edier Vicente Aristizábal Giraldo**

**Universidad Nacional de Colombia  
Sede Medellín  
Facultad de Minas  
2023-II**

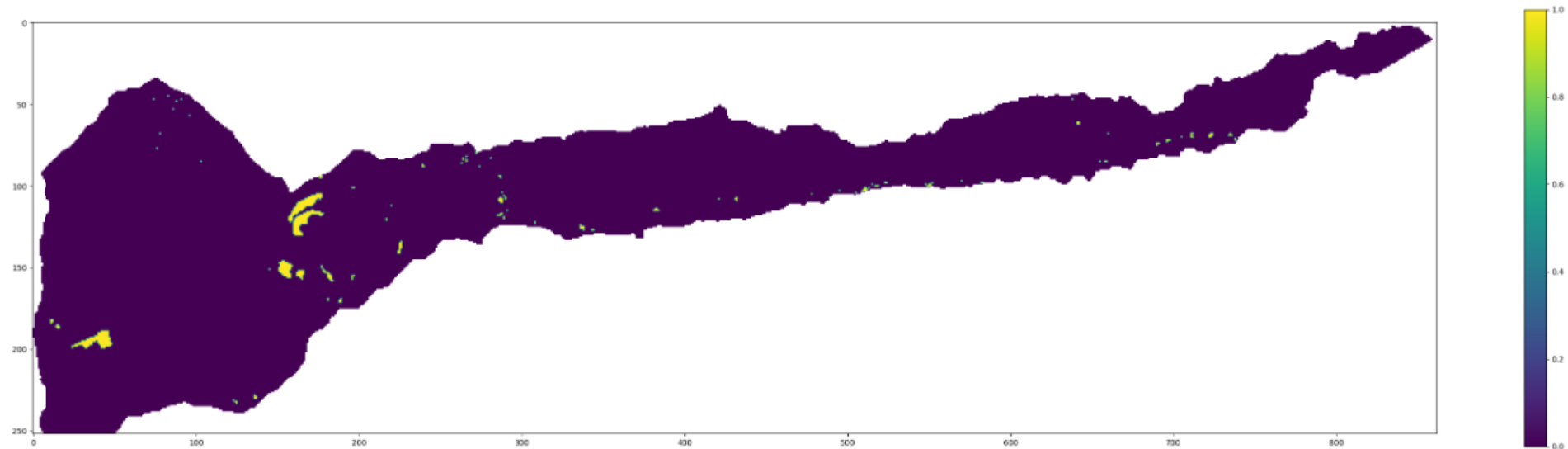
**Taller 5/Análisis exploratorio de datos**

Para elaborar un mapa de susceptibilidad y/o amenaza por movimientos en masa se debe construir un modelo lineal  $y=f(x)$  que sea predictorio, donde  $y$  es la variable dependiente y funciona con 1 y 0, es decir hay o no hay movimientos en masa. Y  $x$  las variables independientes que pueden ser continuas (Aspectos, curvatura, altitud y pendientes) o discretas (Geología y geomorfología) las cuales están encargadas de representar una relación con la ocurrencia o no de los movimientos en masa.

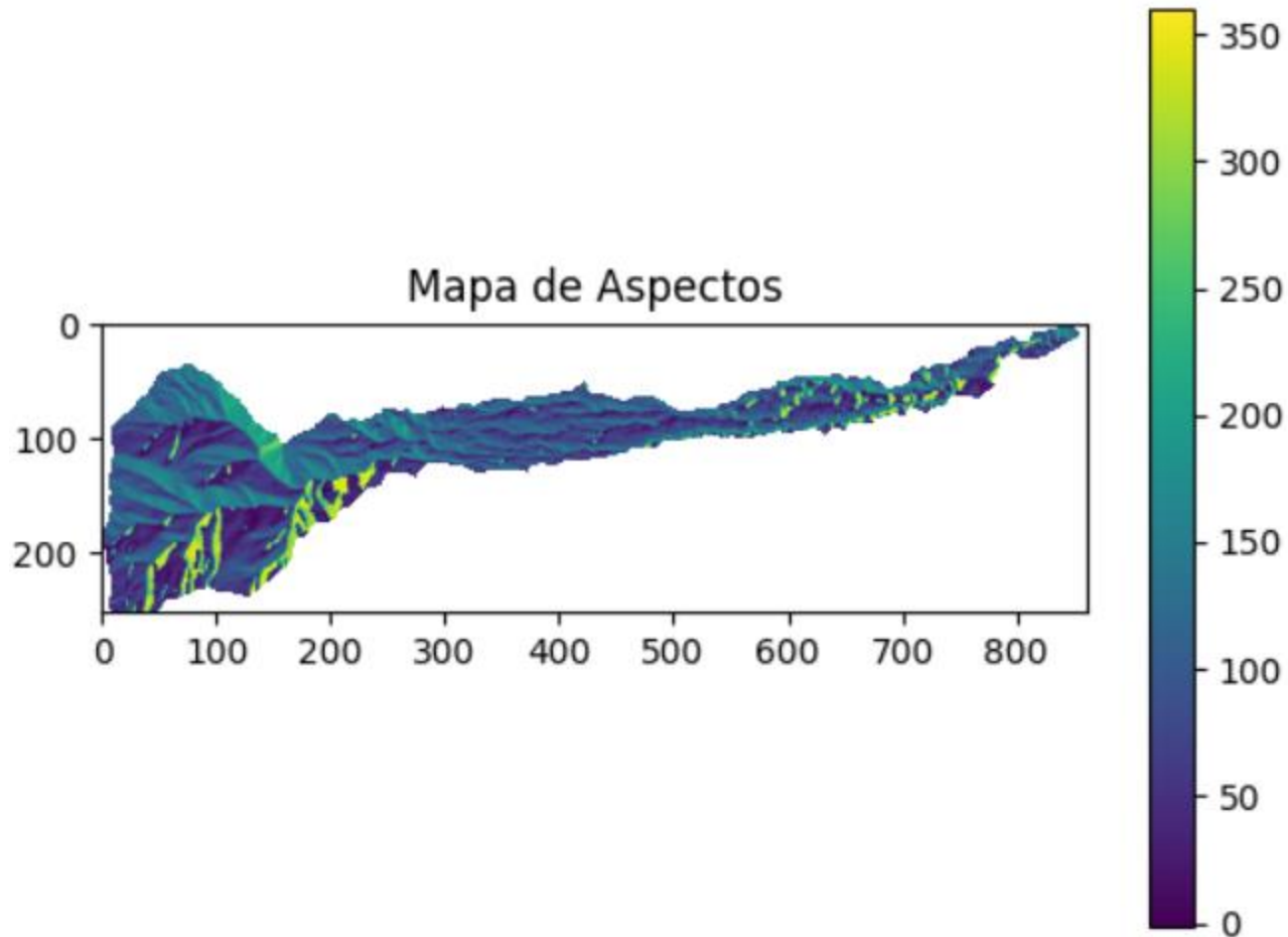
Para realizar el análisis de las variables primero debemos cargar todas las variables que están distribuidas en filas y columnas y pasarlas a tipo vector con las librerías de Python, para luego usar un dataframe que es una matriz que recopila la información de todas las variables predictorias y permite analizarlas con respecto a la variable dependiente

Cuenca quebrada la Ospina en color morado con movimientos en masa tipo polígono y multitemporales en color amarillo

Cuenca B1  
Forma: (252, 862)  
Valores únicos: [ 0. 1. nan]

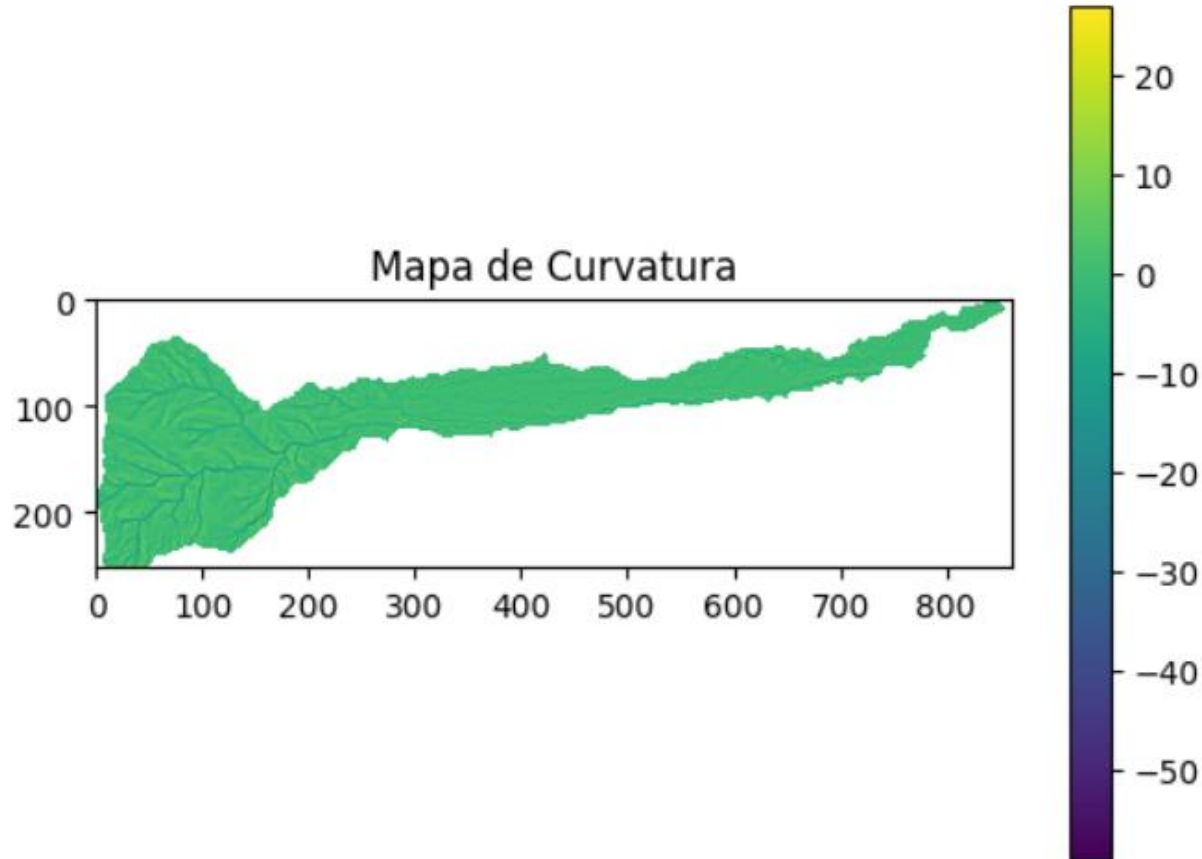


# Aspectos de la cuenca de la Ospina



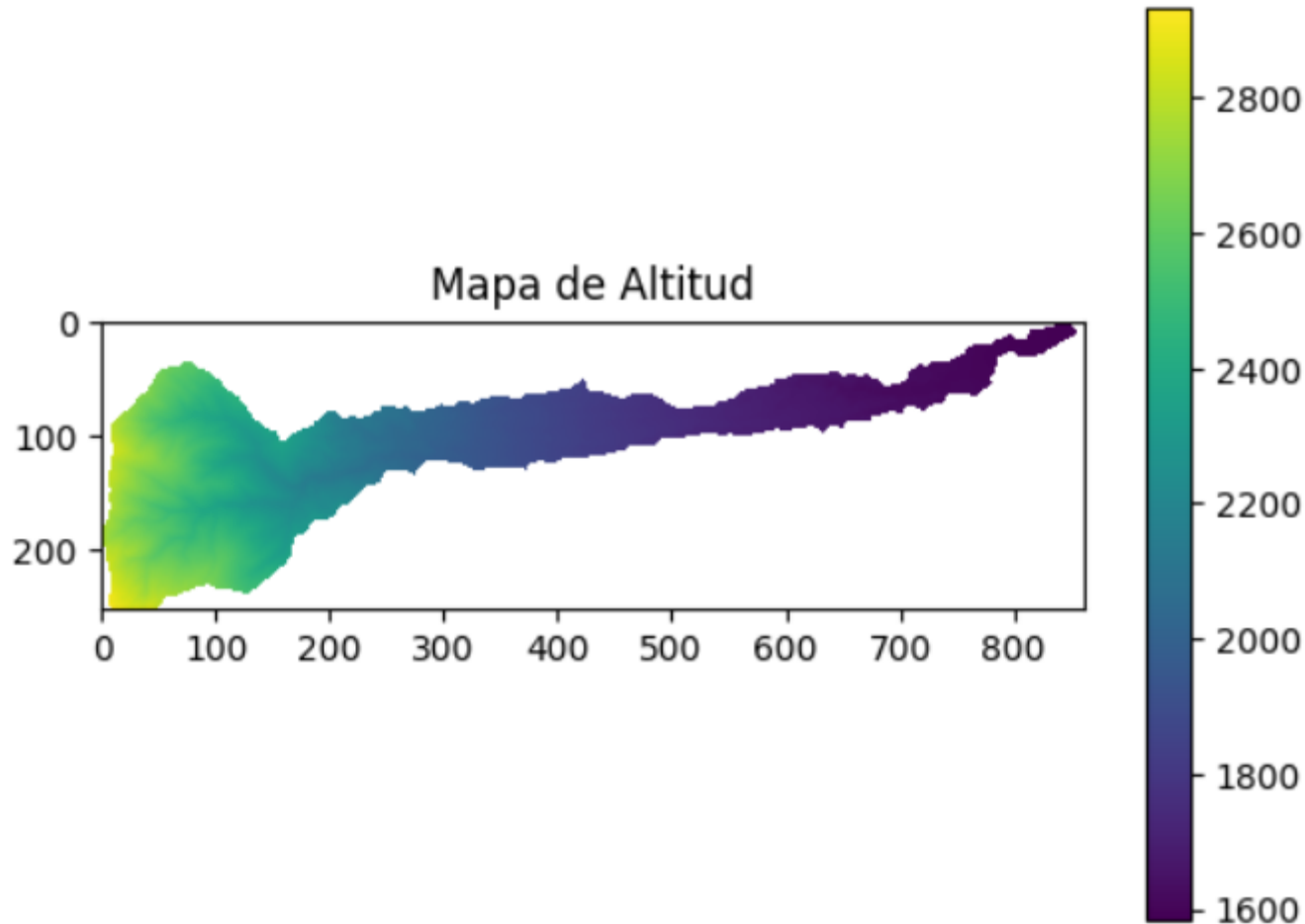
## Curvatura de la cuenca de la Ospina

Se observa como los drenajes y líneas de escorrentía presentan valores negativos que son los que indican forma cóncava y como las crestas de los lomos tienen valores positivos amarillos que indican forma convexa y el resto color verde que no presenta forma convexa muy marcada y se interpreta como zonas con poca curvatura.



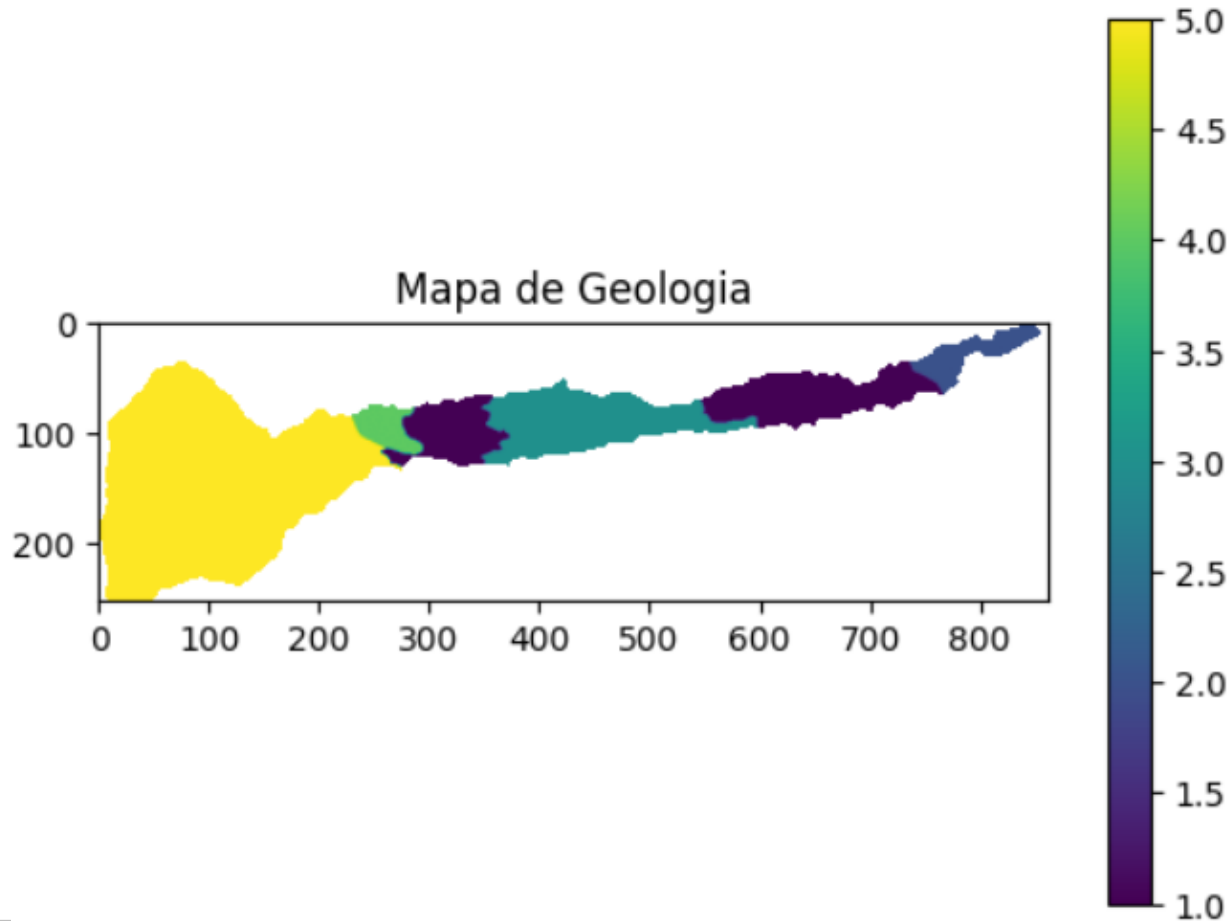
## Altitud de la cuenca de la Ospina

Este mapa indica que la zona mas alta queda al oeste del área de estudio y que disminuye en dirección este.



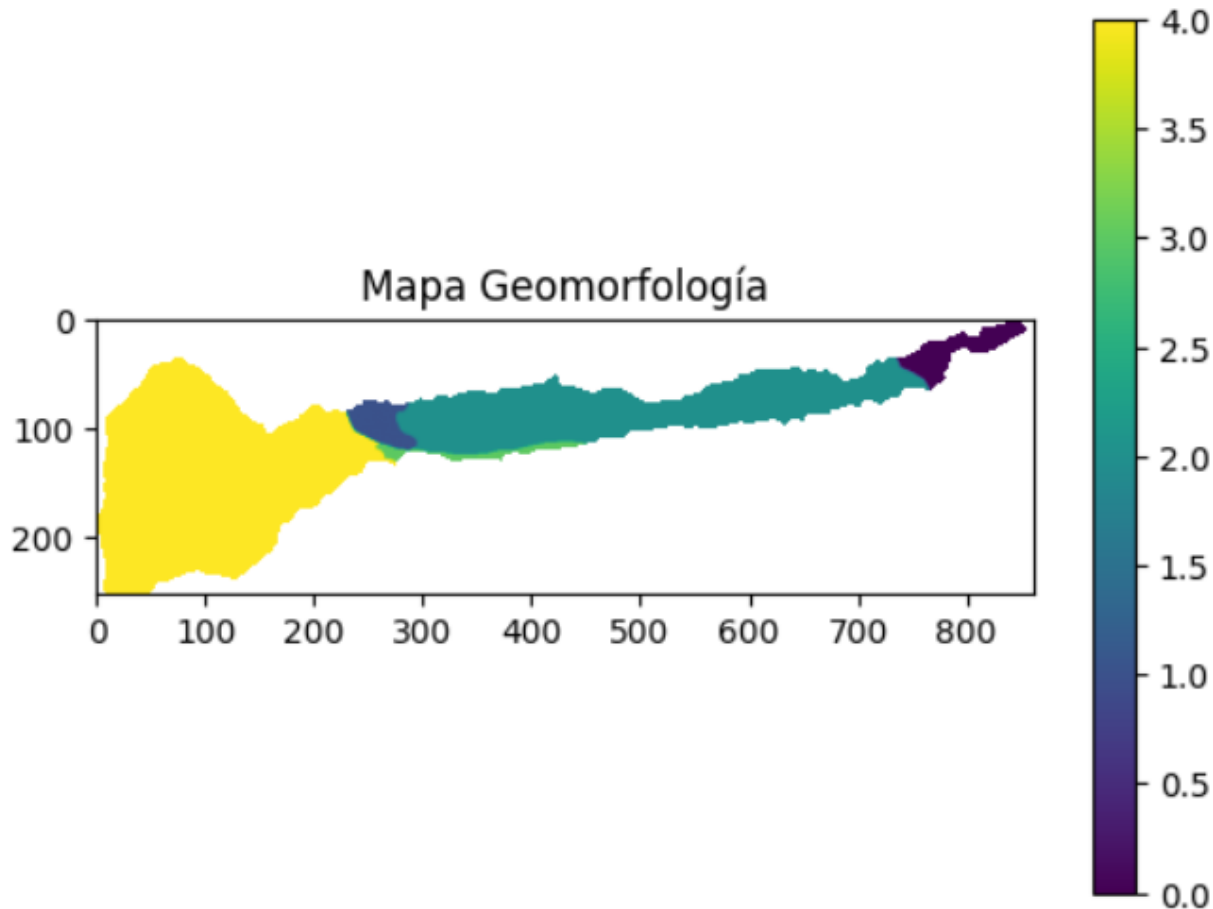
## Geología de la cuenca de la Ospina

Nos muestra que el área de estudio tiene 5 unidades geológicas, compuestas principalmente por depósitos de flujo de lodos y escombros en el Este y una Secuencia volcanosedimentaria y volcánica hacia el Oeste.



## Geomorfología de la cuenca de la Ospina

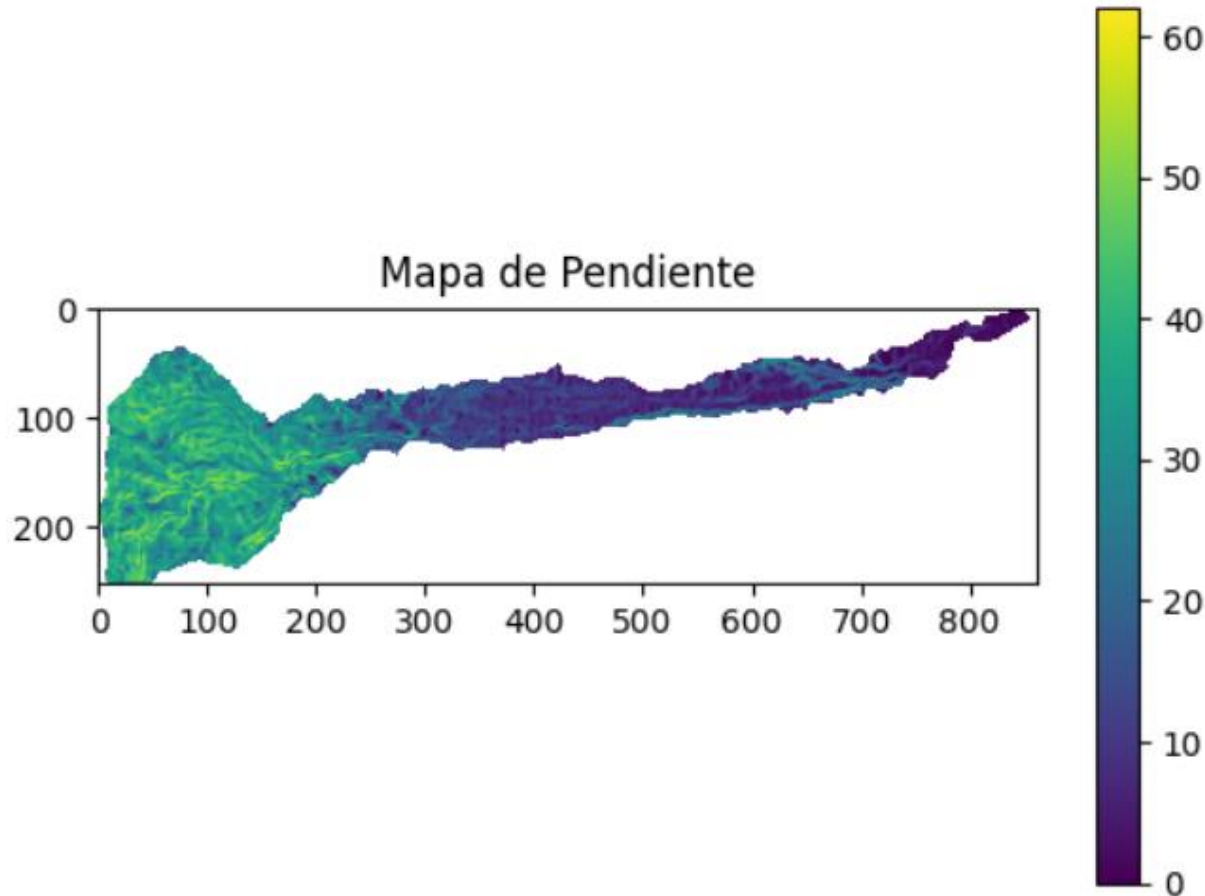
Nos muestra que el área de estudio se puede agrupar en 5 formas del paisaje con una estrecha correlación con respecto a la geología, compuestas principalmente por superficies suaves en el Este y una Secuencia de lomos hacia el Oeste.





## Geomorfología de la cuenca de la Ospina

Las pendientes de la zona tienen correlación con la geología y la geomorfología, siendo las pendientes mas suaves pertenecientes a las superficies suaves y los depósitos y las pendientes altas se asocian a los lomos en rocas ígneas y sedimentarias.



# DATAFRAME

5 primeros datos de la matriz de vectores que contienen las variables predictoras y el inventario (variable dependiente)

	inventario	aspectos	curvatura	altitud	geologia	geomorfologia	pendiente
<b>63277</b>	0.0	67.359482	-2.209582	2822.357666	5.0	4.0	33.464508
<b>63278</b>	0.0	31.936472	13.475179	2820.012939	5.0	4.0	28.379774
<b>63279</b>	0.0	33.755505	11.376484	2816.632080	5.0	4.0	27.254066
<b>63280</b>	0.0	37.180698	6.638794	2812.989258	5.0	4.0	21.701637
<b>63281</b>	0.0	28.403448	1.510018	2810.272217	5.0	4.0	15.197203

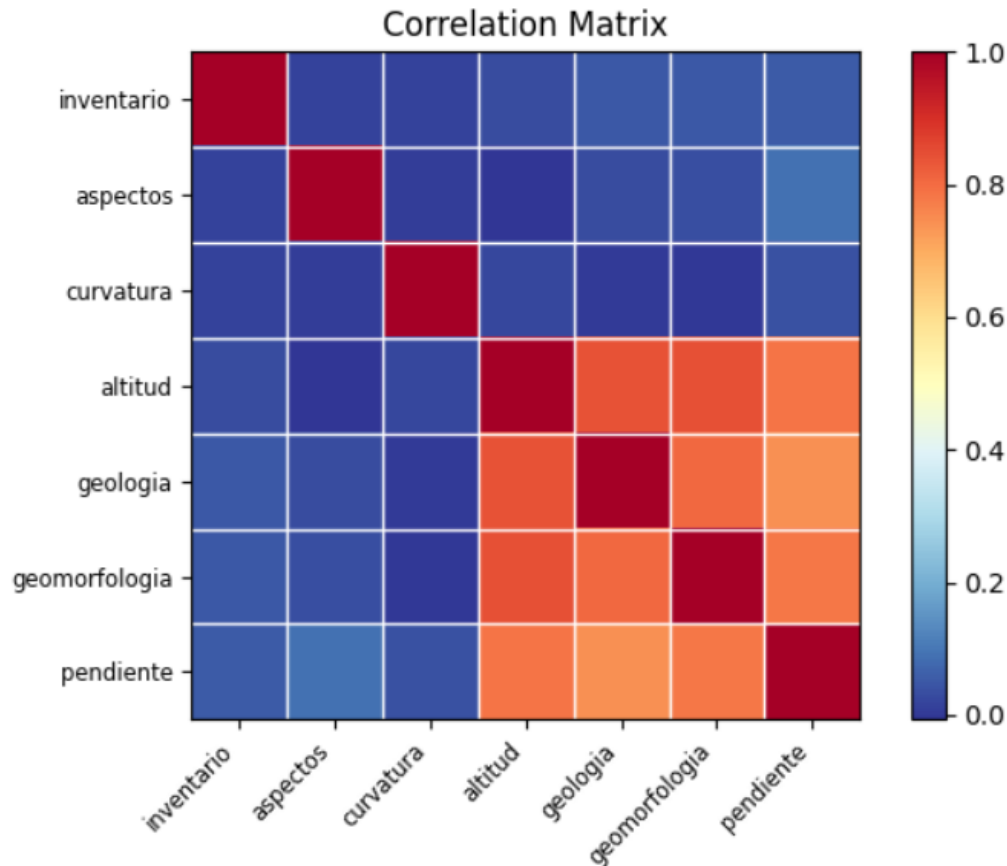
# DATAFRAME

Resumen estadístico de los datos de la matriz de vectores que contienen las variables predictoras y el inventario (variable dependiente)

	count	mean	std	min	25%	\
inventario	63282.0	1.019247e-02	0.100443	0.000000	0.000000	
aspectos	63282.0	1.108404e+02	84.190811	-1.000000	52.710374	
curvatura	63282.0	2.893484e-09	4.624277	-59.041771	-1.008415	
altitud	63282.0	2.166668e+03	382.658539	1582.284058	1812.693939	
geologia	63282.0	3.696343e+00	1.621941	1.000000	3.000000	
geomorfologia	63282.0	3.013622e+00	1.204152	0.000000	2.000000	
pendiente	63282.0	2.399402e+01	14.347461	0.000000	10.019297	
	50%	75%	max			
inventario	0.000000	0.000000	1.000000			
aspectos	93.580257	141.737938	359.992706			
curvatura	0.341596	1.956921	26.930264			
altitud	2189.159668	2508.885803	2934.191895			
geologia	5.000000	5.000000	5.000000			
geomorfologia	4.000000	4.000000	4.000000			
pendiente	25.882936	36.152896	62.054195			

## Matrix de correlación

Relación entre las variables, valores de 1 significan que tan parecidas son entre si y por tanto que tanta información aporta al modelo utilizar las dos variables correlacionadas al tiempo, para este caso la altitud, geología, geomorfología y pendientes aportan la misma información al modelo.

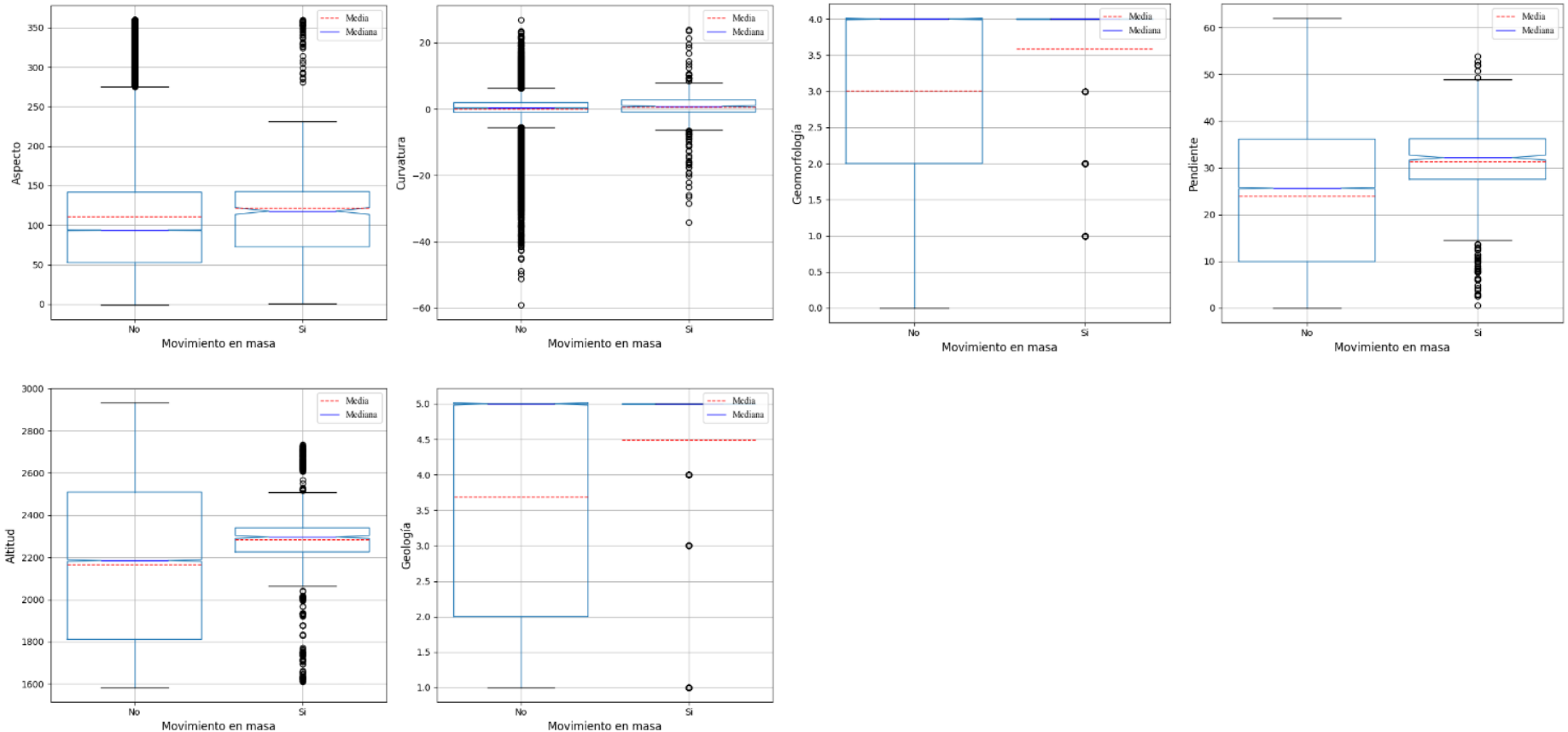


# Análisis univariado

Este análisis estudia la información estadística de una sola variable.

El diagrama de caja representa el 50% de la información de los datos dentro de la caja y la media, así como el 25% y 75% de los datos. Es útil para categorizar las variables continuas.

Boxplot grouped by inventario

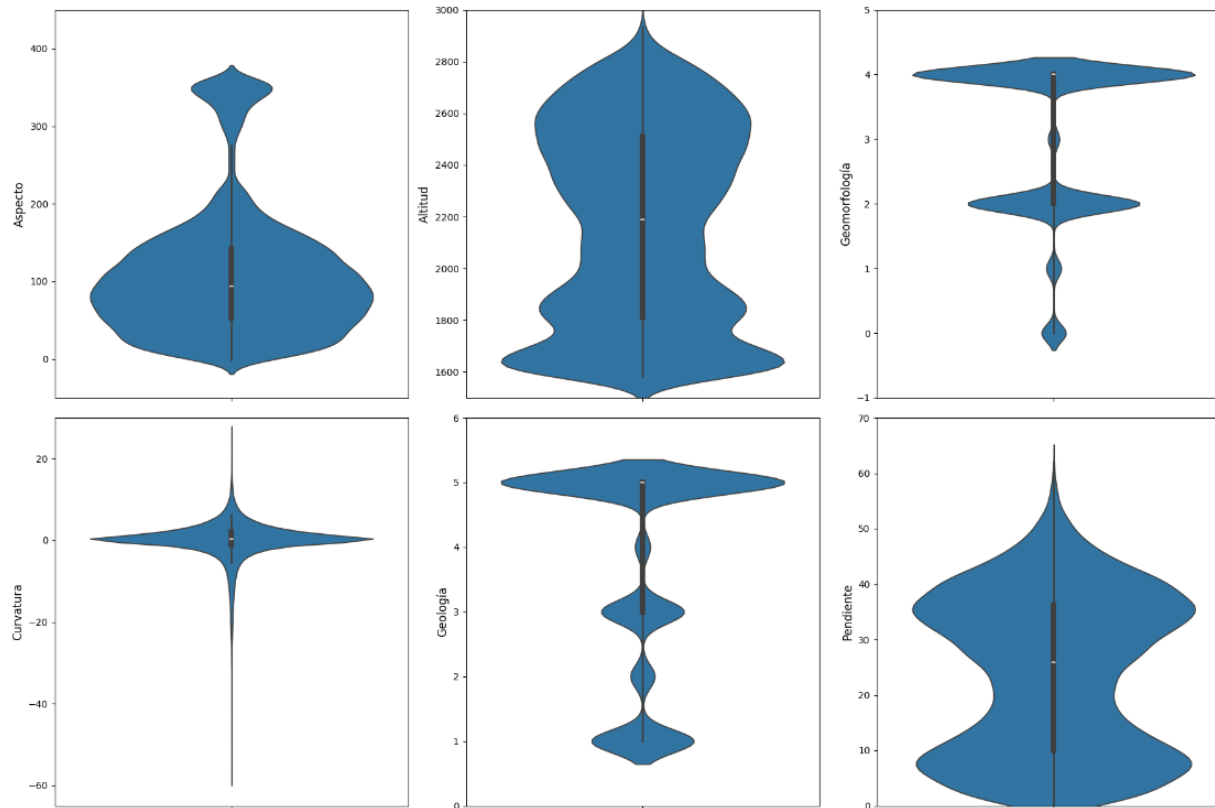


## Análisis univariado

Este análisis estudia la información estadística de una sola variable.

El diagrama de violín representa en que valor se agrupa la mayor cantidad de pixeles.

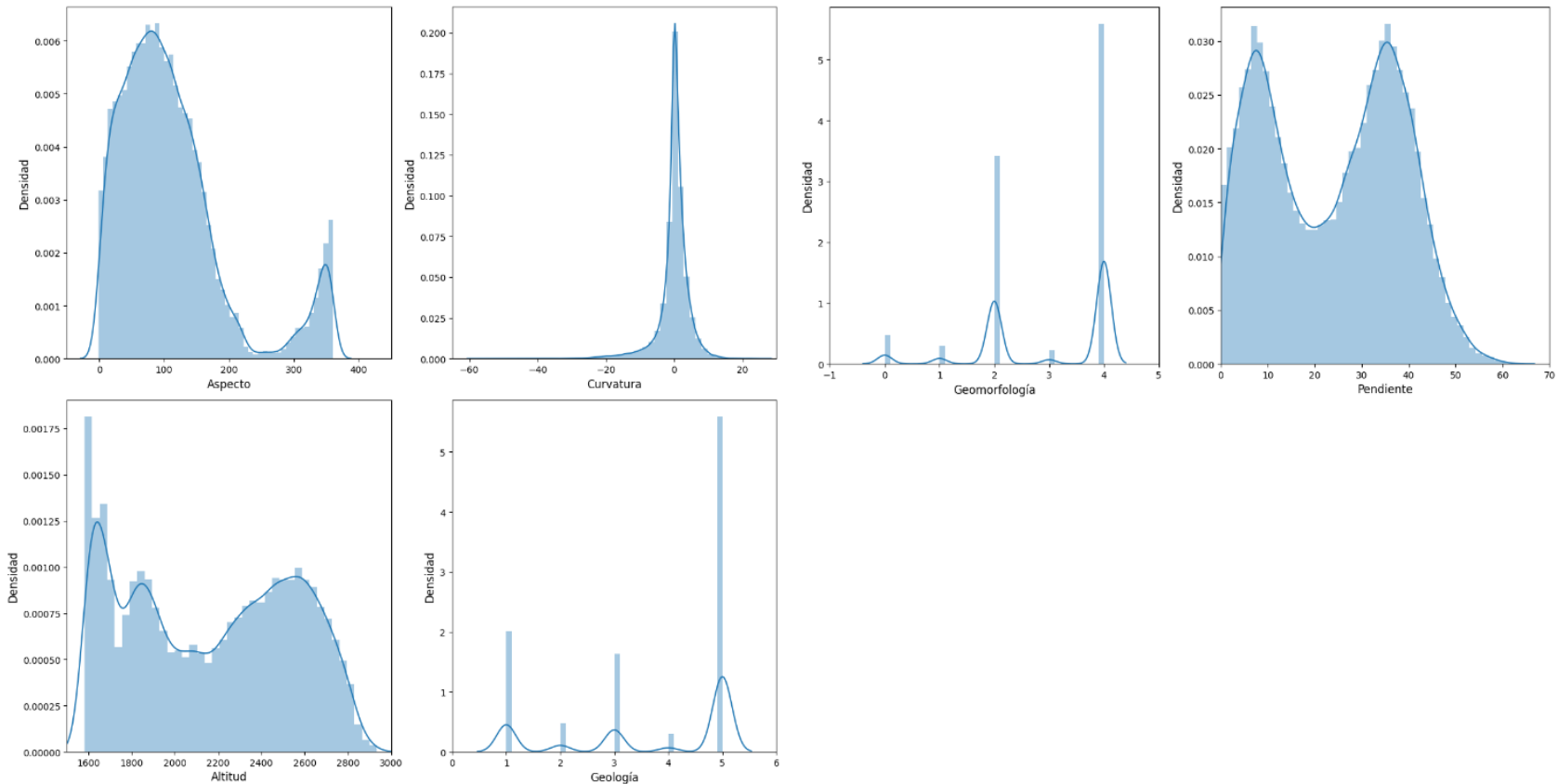
Es útil para categorizar las variables continuas.



# Análisis univariado

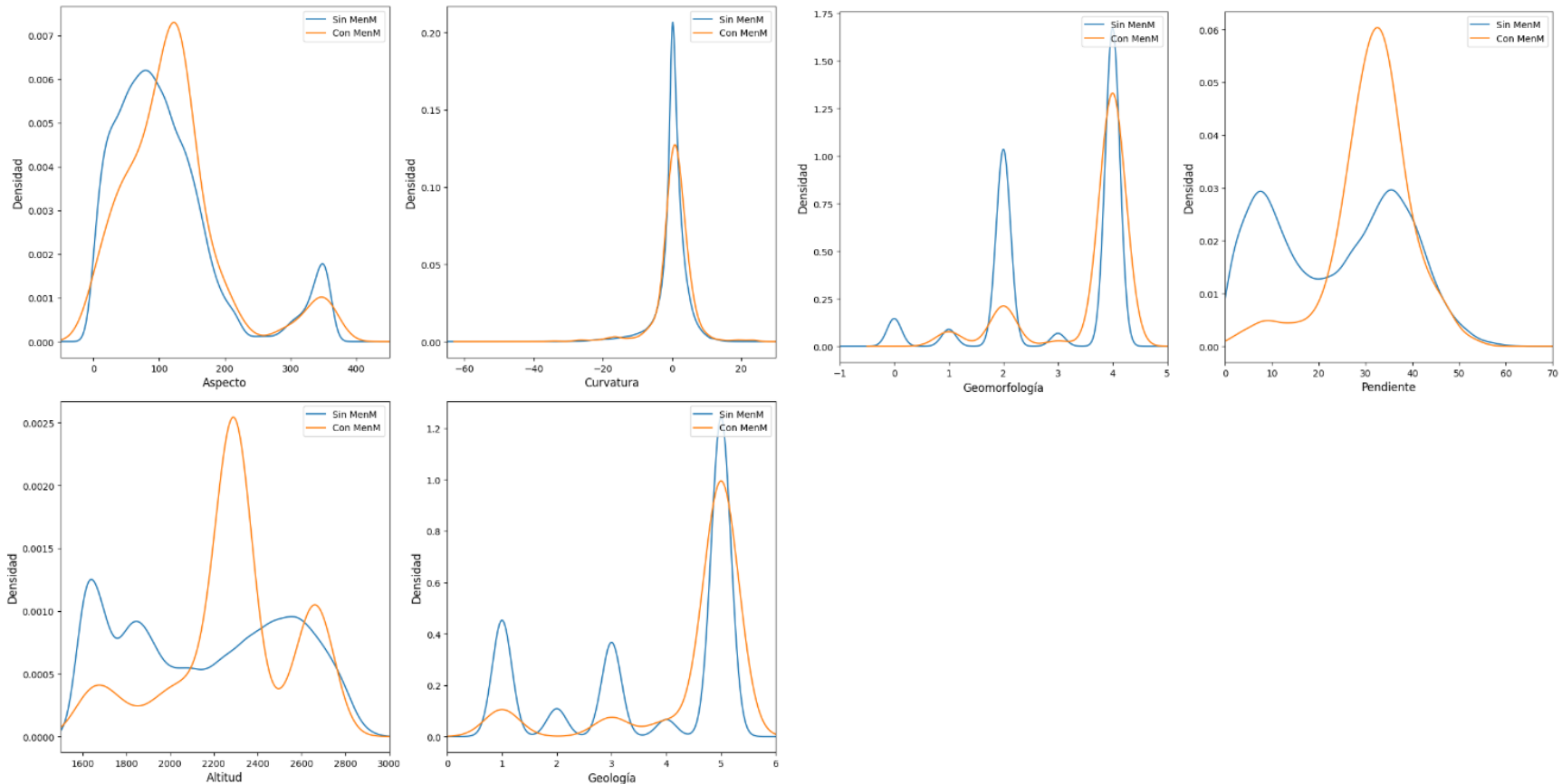
Este análisis estudia la información estadística de una sola variable.

El diagrama de densidad representa la cantidad de píxeles por cada valor, las variables categóricas solo mostraran la cantidad de píxeles por clase entonces no es útil usar este análisis en ella; Pero es útil para categorizar las variables continuas.



# Análisis multivariado

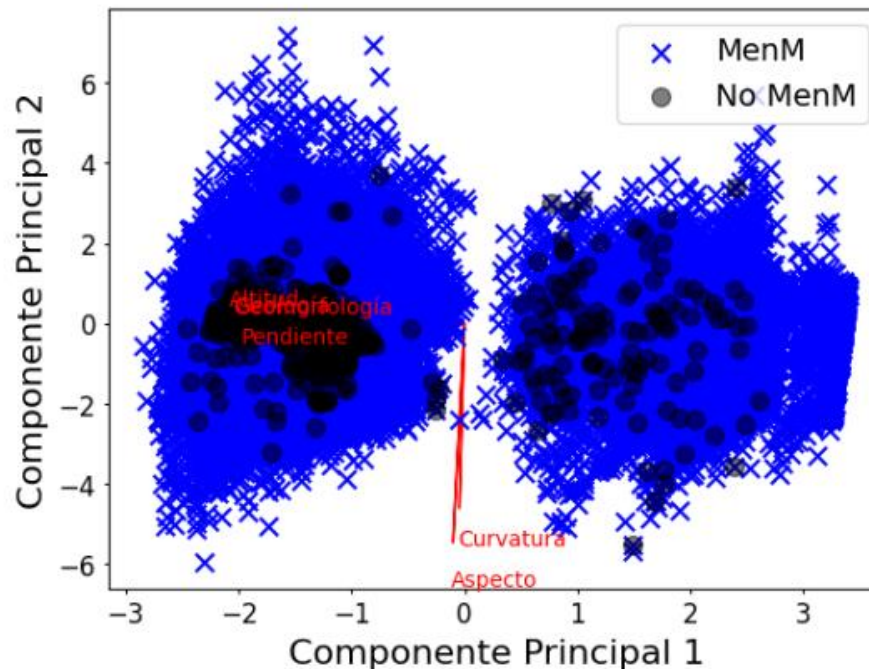
Este análisis estudia la correlación de las variables predictoras con la variable dependiente, en esta análisis se observa la densidad de pixeles por categoría cuando hay o no hay movimientos en masa, se puede ver que la pendiente da una buena separación de los datos por tanto es bueno incluirla en el modelo.





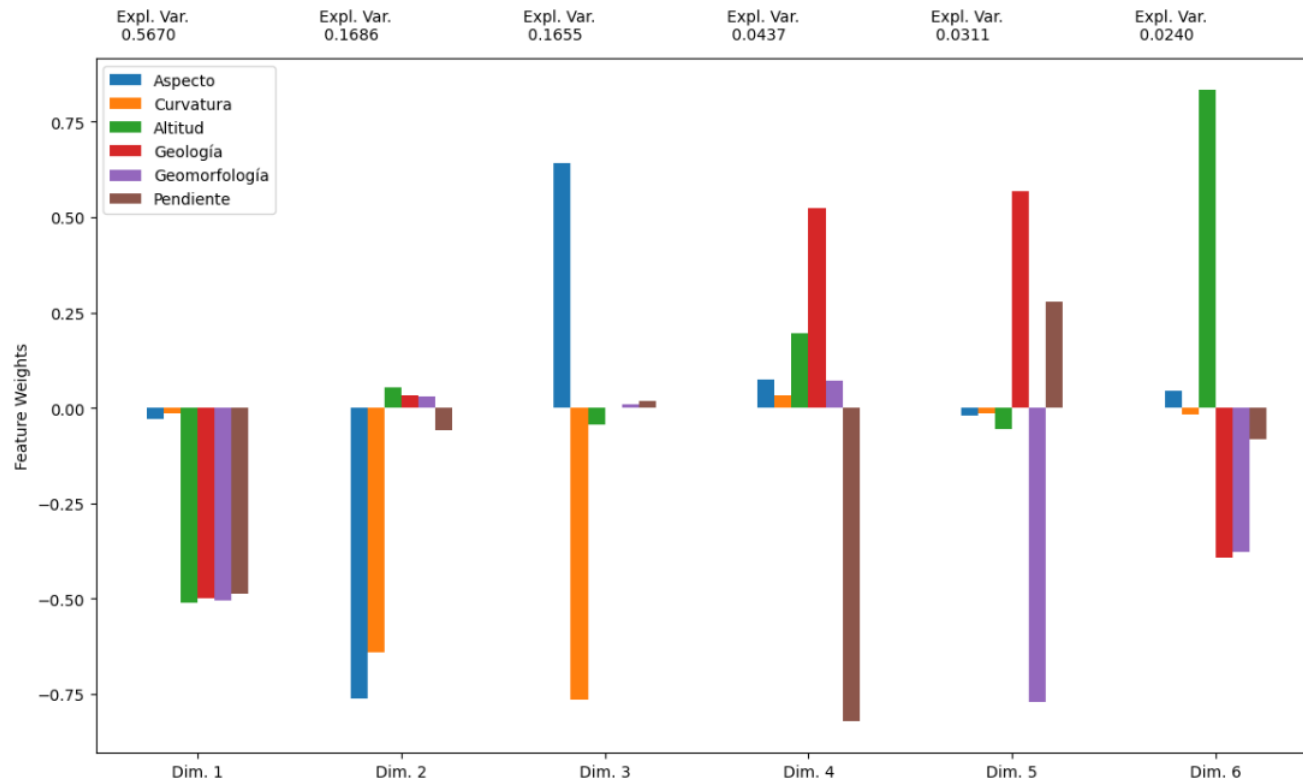
## Análisis de componentes principales (PCA)

Este análisis busca encontrar las variables que mas varianza le aportan al modelo y así reducir el numero de variables que puede generar error (ruido). En otras palabras, se buscan las variables que menos se parecen entre si pero que son buenas predictoras de movimientos en masa. Para la cuenca de la Ospina se observa que dos variables que le aportarían bien al modelo seria la geomorfología (componente principal 1) y la curvatura (componente principal 2). Esto quiere decir que el modelo se puede reducir a estas dos variables.



## Análisis de componentes principales (PCA)

Este análisis busca encontrar las variables que mas varianza le aporta al modelo y así reducir el numero de variables que puede generar error (ruido). En otras palabras, se puede escoger entre altitud, geología, geomorfología y pendiente para el componente principal 1 y entre aspectos y curvatura para el componente principal 2. Esto quiere decir que el modelo se puede reducir a dos variables.



## Análisis de componentes principales (PCA)

Conclusión, se puede escoger geomorfología en el componente principal 1, curvatura como componente principal 2, aspectos como componente principal 3 y geología como componente principal 4. Esto quiere decir que el modelo se puede reducir a cuatro variables y tener un buen porcentaje de predicción.

