

# ¿Para qué son útil las variables suplementarios?

Diego Figueroa

2025-05-15

## Variables Suplementarias

**Definición:** Son variables que no participan en la construcción de los componentes principales. Las proyectamos a posteriori en el espacio definido por los componentes principales creados con las variables activas. Esto es útil por varias razones:

- Evitar Influencia No Deseada: A veces, tienes variables que podrían dominar el PCA debido a su escala o alta varianza, enmascarando patrones importantes en otras variables. Al tratarlas como suplementarias, te aseguras de que la estructura principal del PCA se base en las variables que consideras primarias para definir la similitud entre tus individuos.
- Interpretar los Componentes Principales: Las variables suplementarias pueden ayudarte a entender qué “significan” los componentes principales que has encontrado. Por ejemplo, si un componente principal separa los autos por tamaño (basado en LARGO y ANCHO como variables activas), observar dónde se proyecta la variable suplementaria “PRECIO” puede indicarte si los autos más grandes tienden a ser más caros.
- Visualizar Información Adicional (Cualitativa): Las variables cualitativas suplementarias (como “LUJO”) no tienen una “dirección” en el espacio de las variables. En el gráfico de individuos, sus categorías se representan mediante los centroides (el punto promedio) de los individuos que pertenecen a cada categoría. Esto te permite ver si los grupos definidos por la variable cualitativa se corresponden con la estructura revelada por el PCA basado en las variables activas.
- Validación Externa: Puedes tener variables que no estaban disponibles para la fase inicial de la recolección de datos, o variables que representan una clasificación externa conocida. Proyectarlas como suplementarias te permite ver si la estructura de tus datos, basada en las variables activas, se alinea con esta información externa.