

# Explorando sitios de comercio electrónico

Final AID

Leonardo Aranda

# Introducción

- ▶ Crecimiento sostenido del comercio electrónico
- ▶ ¿Qué características tiene un sitio de elevado tráfico?
- ▶ Estimaciones de tráfico: Alexa + SimilarWeb.

## Variables

Variable	Descripción
url	Dirección del sitio web
position	Pertenece al Top 50.000 mundial
bounce	Porcentaje de visitas con una sola página vista
ppv	Páginas vistas por visita
time	Tiempo promedio de la visita
paid_search	Tráfico pago
traffic_direct	Tráfico directo
traffic_display	Tráfico de redes de publicidad
traffic_search	Tráfico de buscadores
traffic_mail	Tráfico de correo electrónico
traffic_social	Tráfico de redes sociales
traffic_referrals	Tráfico referido desde otras páginas

## Vector de medias

##	[,1]	[,2]
## Group.1	"0"	"1"
## bounce	"40.48016"	"43.02896"
## paid_search	"13.92329"	"18.45615"
## ppv	"4.517698"	"5.803474"
## time	"172.9875"	"255.8759"
## traffic_direct	"24.45626"	"28.58323"
## traffic_display	"0.8779869"	"1.7365757"
## traffic_search	"54.47384"	"45.90362"
## traffic_mail	"4.214623"	"5.198238"
## traffic_social	"3.117430"	"3.837295"
## traffic_referrals	"12.86008"	"14.74079"

## Diferencias según el tiempo de visita

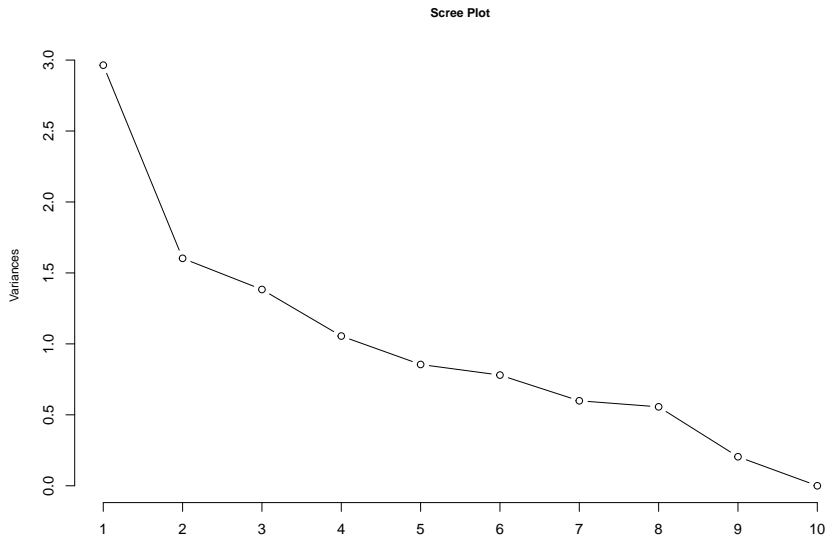
```
##  
## Shapiro-Wilk normality test  
##  
## data:  df$time  
## W = 0.96344, p-value < 2.2e-16
```

- ▶  $H_0$ : No hay diferencias significativas. Los grupos pertenecen a la misma distribución.
- ▶  $H_1$ : Existen diferencias significativas entre los grupos.

```
##  
## Wilcoxon rank sum test with continuity correction  
##  
## data:  time by position  
## W = 148840, p-value < 2.2e-16  
## alternative hypothesis: true location shift is not equal
```

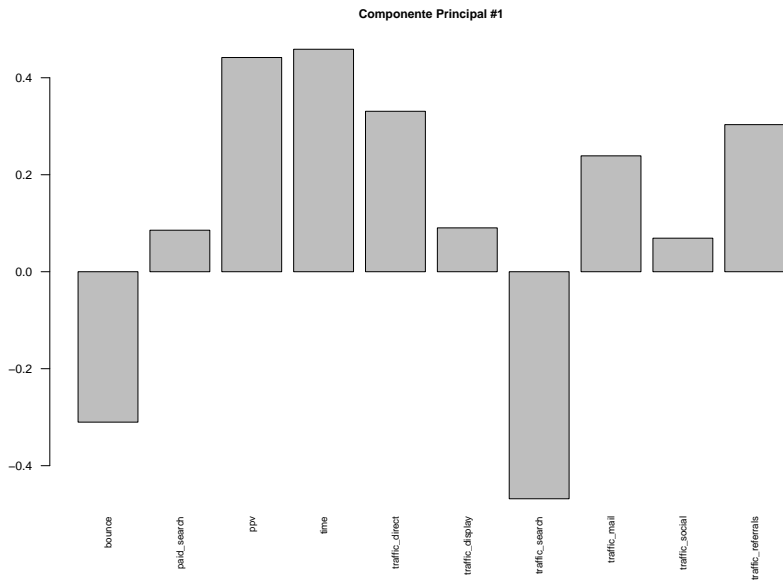
# Componentes principales - Scree Plot

- Tres componentes explican el 60% de la variabilidad.



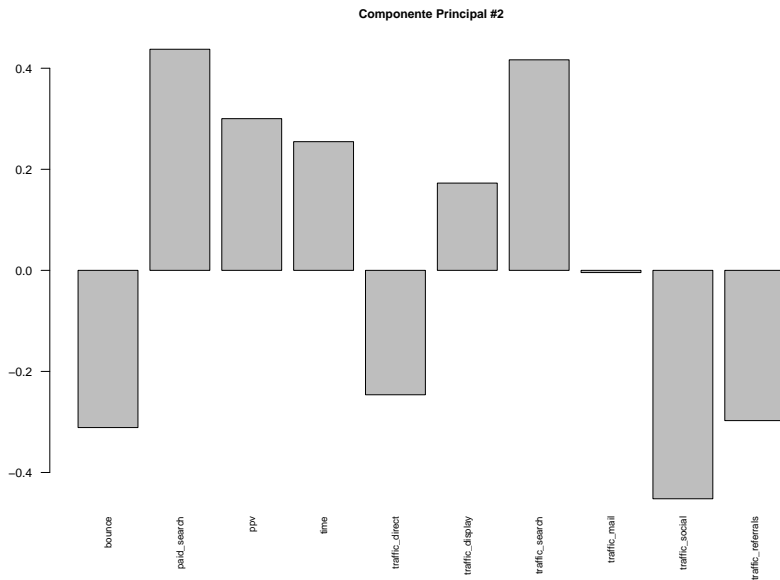
# Componentes principales - Componente 1

## ► Nivel de interacción



# Componentes principales - Componente 2

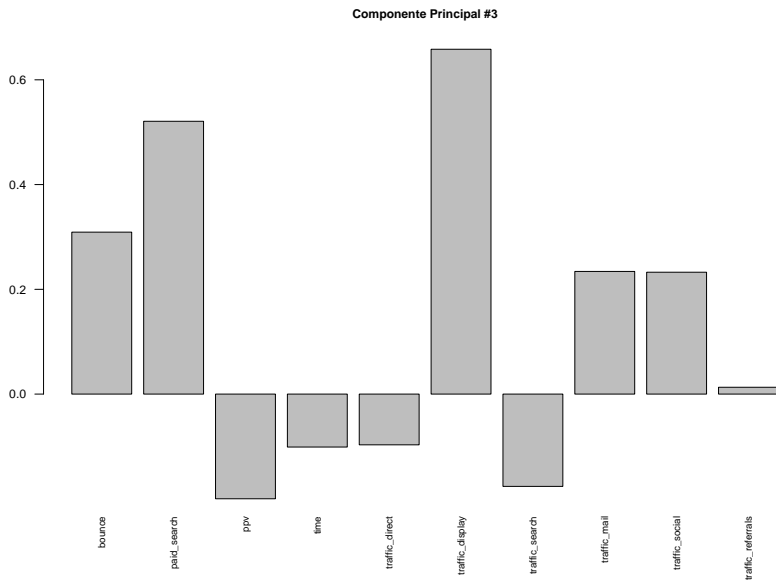
## ► Conocimiento de marca





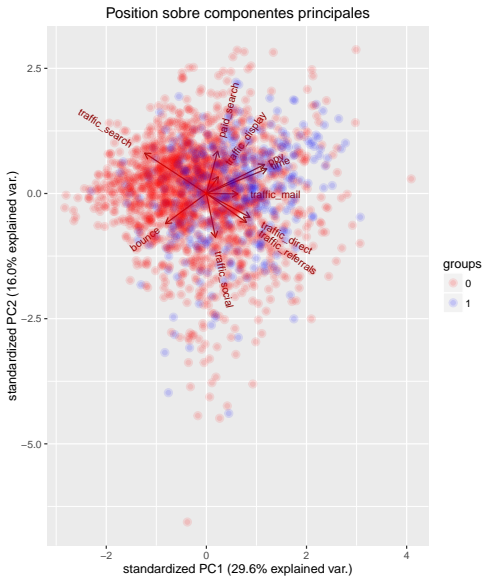
# Componentes principales - Componente 3

## ► Tráfico pago



# Componentes principales - Biplot

- Agrupamiento basado en el nivel de tráfico



# Análisis discriminante

# Software

- ▶ Python + Jupyter: Preparación de datos.
- ▶ Rstudio: Análisis y documentación.
- ▶ GitHub: Almacenamiento y versionado del código.

# Conclusiones

- ▶ Diferencias significativas según el nivel de tráfico