

Taller de práctica 7

Interpretación de parámetros estimados
Métricas y pruebas de hipótesis asociadas a una regresión lineal múltiple
(secciones 3.3.2, 3.3.3 y 3.3.4)

La siguiente base de datos ([link](#)) contiene los datos del mercado de valores de Airbnb desde su creación. Las variables que contiene son las siguientes:

- **Date**: Fecha de inclusión del inmueble al sistema de Airbnb
- **Open**: Precio cuando se abre el mercado.
- **High**: Precio más alto registrado para el día.
- **Low**: Precio más bajo registrado para el día
- **Close**: Precio cuando el mercado cierra.
- **Adj Close**: Precio de cierre modificado basado en acciones corporativas.
- **Volume**: Cantidad de existencias vendidas en un día (estadias).

Nota: todos los precios están en dólares.

- 1) Modifique la columna **Date**, de tal forma, que refleje únicamente el año (número) de la fecha de inclusión del inmueble. Para ello, utilice el siguiente código:

```
df$Date = as.numeric(format(as.Date(df$Date), "%Y"))
```

- 2) Para estudiar el Cantidad de existencias vendidas, realice los siguientes pasos:

- a) Aplique logaritmo sobre la variable **Volume** reescribiéndola en la base de datos.
- b) Filtre La base de datos, conservando solo las observaciones correspondientes al año 2021.
- c) Construya un modelo con nombre **modelo1** que estudie el logaritmo de la cantidad de evidencias vendidas a base de el precio más alto, más bajo registrado en el día y el precio de apertura.
- d) Interprete los parámetros estimados y las pruebas de no nulidad asociadas al **modelo1**.
- e) Construya un nuevo modelo **modelo1.1** igual al **modelo1**, excluyendo la covariable **Open**.
- f) Interprete los parámetros estimados y las pruebas de no nulidad asociadas al **modelo1.1**.
- g) Construya un modelo con nombre **modelo2** que estudie el logaritmo de la cantidad de evidencias vendidas a base de el precio de cierre y el precio más bajo registrado en el día.
- h) Interprete los beta estimados y las pruebas de no nulidad asociadas al **modelo2**.
- i) Compare los tres modelos construidos, para ello utilice el R^2 ajustado y el error estándar residual. ¿Qué modelo utilizaría en un caso de negocio real?