

# laboratorio – Regresión y Series de Tiempo

Para cada uno de los ejercicios:

- Suba el archivo .R indicado
- Suba el archivo .Rdata con la data de su environment
- Adjunte el PLOT como imagen en esta hoja de trabajo
- Responda las preguntas indicadas

## 1. Regresión en Repuestos Web

Usando el siguiente query en la base de datos de “RepuestosWeb”:

```
SELECT v.Anio,  
       v.Marca,  
       v.Modelo,  
       p.Nombre AS NombreParte,  
       d.NombreDescuento,  
       do.Cantidad  
FROM   dbo.Detalle_orden do  
JOIN   dbo.Partes p  
       ON p.ID_Parte = do.ID_Parte  
JOIN   dbo.Descuento d  
       ON d.ID_Descuento = do.ID_Descuento  
JOIN   dbo.Vehiculo v  
       ON v.VehiculoID = do.VehiculoID;
```

Cree una vista llamada V\_CantidadPorVehiculoDescuento

- a. Cree un nuevo script .R llamado “Regresión Lineal Repuestos Web”
- b. Cree un data frame que lea la data de la vista anteriormente creada
- c. Borre registros que tengan datos incompletos
- d. Separe la data en dos data frames a partir del data frame principal, el primero tendrá el primer 80% del total de los registros y el segundo tendrá el restante 20%
- e. Cree un modelo de Regresión Lineal que permita predecir la cantidad de partes que se compraran en función de las características de un vehículo, el nombre de la parte y el descuento aplicado
- f. Adjunte un PLOT de cada modelo de regresión
- g. Enumere y grafique cuáles son las variables de entrada que más influyen en el pronóstico de la cantidad de partes que un cliente puede comprar
  - i. ¿Cuál es su interpretación de los atributos que más y que menos influyen?

## 2. Series de Tiempo en Repuestos Web

Usando el siguiente query en la base de datos de “RepuestosWeb”:

```

SELECT CAST(DATEPART(YEAR,c.FechaCreacion) AS VARCHAR(4)) + '-' +
        CAST(DATEPART(QUARTER,c.FechaCreacion) AS VARCHAR(1)) AS
AnioTrimestreCotizacion,
        SUM(cd.Cantidad * p.Precio) AS TotalCotizacion,
        SUM(cd.Cantidad) AS TotalPartesVendidas,
        AVG(cd.Cantidad * p.Precio) AS PromedioCotizacion
FROM dbo.Cotizacion c
        INNER JOIN dbo.CotizacionDetalle cd
        ON cd.IDCotizacion = c.IDCotizacion
        INNER JOIN dbo.Partes p
        ON p.ID_Parte = cd.ID_Parte
        JOIN dbo.Vehiculo v
        ON v.VehiculoID = cd.VehiculoID
GROUP BY CAST(DATEPART(YEAR,c.FechaCreacion) AS VARCHAR(4)) + '-' +
        CAST(DATEPART(QUARTER,c.FechaCreacion) AS VARCHAR(1))
ORDER BY 1 ASC;
  
```

- Cree un nuevo script .R llamado “Serie de Tiempo RepuestosWeb”
- Cree un tres objetos de tipo “ts” (time series) que lea la data de la vista anteriormente creada, uno para cada métrica (TotalCotizacion, TotalPartesCotizadas, PromedioCotizacion), el query devuelve la columna “AnioTrimestreCotizacion” ya que separa los datos en trimestres, por lo cual puede usar los parámetros de TS de “frequency” para colocar sus datos en el formato correcto (trimestral)
- Dibuje un plot con el objeto anterior
- Evalúe los comportamientos de la data de cada objeto “ts” y encuentre que componentes tiene (autocorrelación, estacionario, etc)
  - ¿Qué puede interpretar de los resultados?
  - Adjunte los plots
- Evalúe la propiedad “residuals” del análisis de series de tiempo e interprete dicha propiedad con las funciones “mean” y “var”
  - ¿Qué puede interpretar de los resultados?
- Para cada objeto “ts” de cada métrica (TotalCotizacion, TotalPartesCotizadas, PromedioCotizacion), Separe la data en dos objetos “ts” a partir del objeto principal usando la función “window”, el primero tendrá el primer 80% del total de los registros y el segundo tendrá el restante 20%
- Aplique las tres técnicas de pronósticos simples vistas en clase y evalúe su accuracy
  - ¿Cuál de las tres tiene un mejor desempeño?
  - Basado en el comportamiento de la data ¿cree que era necesario aplicar una técnica simple? ¿Por qué?
- Aplique las dos técnicas de pronósticos avanzados vistas en clase y evalúe su accuracy
  - ¿Cuál de las tres tiene un mejor desempeño?
  - Realice un pronostico de por lo menos 30 días en un plot
  - Basado en el comportamiento de la data ¿cree que era necesario aplicar una técnica avanzada? ¿Por qué?