

**Instituto Tecnológico y de Estudios Superiores de Monterrey**  
**Series de tiempo**  
**Abril – Julio de 2022**

**Tarea 1**  
Abril 18, 2022

Durante el curso utilizaremos R con nuestras series de tiempo. Las tareas tienen la finalidad de que los estudiantes se familiaricen con el *software*, si es que no lo conocen, y de que aprendan a utilizarlo para realizar pronósticos con datos reales. Por lo tanto, antes de comenzar esta tarea, lo primero es:

1. [Descarga e instala R.](#)
2. [Descarga e instala RStudio.](#)
3. Abre RStudio.
4. Instala los paquetes de uso común en nuestra clase

```
install.packages("tidyverse")  
install.packages("fpp3")  
install.packages("forecast")  
install.packages("ggplot2")  
install.packages("astsa")  
install.packages("GGally")  
install.packages("fma")  
install.packages("expsmooth")  
install.packages("tsibble")
```

Todos los ejemplos, en clase y en las tareas, asumen que ya has instalado y cargado el paquete fpp3 (library(fpp3)).

1. Una serie de tiempo es una secuencia de observaciones (o números) sobre intervalos de tiempo separados de manera regular (el índice). En R, esta información se puede almacenar en un formato llamado “tsibble”, el cual nos permite graficar la información.
  - a. Convierte la serie de tiempo de la siguiente tabla en “tsibble” en R.

Año	Aprobación presidencial (% promedio en México, anual)
2015	48
2016	36
2017	25
2018	30
2019	68

Puedes usar este código:

```
y <- tsibble(  
  Year = 2015:2019,  
  Observation = c(48, 36, 25, 30, 68),  
  index = Year  
)
```

**b. Gráfica la serie. Puedes usar el siguiente código:**

```
autoplot(y) +  
  ggtitle("Aprobación presidencial (% promedio en México, anual)") +  
  ylab("%") +  
  xlab("Año")
```

**c. ¿La serie sigue alguna tendencia, estacionalidad o ciclo?**

2. Carga la base de datos "tute1.csv", que puedes ubicar en el sitio del curso en Canvas. Tiene 4 columnas: Quarter, Sales (ventas trimestrales de una tienda de teléfonos), AdBudget (es el presupuesto para publicidad de la tienda (en dólares australianos) y GDP (Producto interno bruto de Australia, en dólares australianos). La serie ya fue ajustada por inflación.

**a. Carga la base de datos en R (recuerda asignar un directorio de trabajo y ahí guardar la base de datos** (ejemplo: `setwd("E:\\Tareas\\Curso\\Curso TS")`).

```
tute1 <- readr::read_csv("tute1.csv")  
View(tute1)
```

**b. Transforma los datos en una serie de tiempo para R.**

```
mytimeseries <- tute1 %>%  
  mutate(Quarter = yearmonth(Quarter)) %>%  
  as_tsibble(index = Quarter)
```

**c. Gráfica las series de tiempo de las variables en la base de datos.**

```
mytimeseries %>%  
  pivot_longer(-Quarter) %>%  
  ggplot(aes(x = Quarter, y = value, colour = name)) +  
  geom_line() +  
  facet_grid(name ~ ., scales = "free_y")
```

**d. Revisa qué pasa cuando no incluyes `facet_grid()`.**

3. Usa la base de datos “hh\_budget” de la librería de R. Esta base de datos incluye indicadores anuales de los presupuestos de los hogares de Australia, Japón, Canadá y los Estados Unidos, de 1995 a 2016.

Variable	Descripción
Debt	Deuda como porcentaje del ingreso disponible del hogar.
DI	Tasa anual de crecimiento del ingreso disponible del hogar.
Expenditure	Tasa anual de crecimiento del gasto del hogar.
Savings	Ahorro como porcentaje del ingreso disponible del hogar.
Wealth	Riqueza como porcentaje del ingreso disponible del hogar.
Unemployment	Porcentaje de desempleo en la fuerza de trabajo.
Country	País.

- a. Crea un *training set* que excluya los últimos cuatro años y los convierta en un *test set*.

```
library(tsimbledata)

hh_budget_train <-
  hh_budget %>%
  filter_index(. ~ '2012')

hh_budget_test <-
  hh_budget %>%
  filter_index('2013' ~ .)
```

- b. Ejecuta los métodos simples de pronóstico con el *training set* y pronostica para los correspondientes al *test set*.

```
hh_budget_models <-
  hh_budget_train %>%
  model(
    mean = MEAN(Wealth),
    naive = NAIVE(Wealth),
    snaive = SNAIVE(Wealth),
  )

hh_budget_forecast <-
  hh_budget_models %>%
  forecast(hh_budget_test)
```

**c. Calcula la exactitud de los métodos. ¿Cuál es el mejor?**

```
hh_budget_forecast %>% accuracy(hh_budget_test) %>%  
  group_by(Country) %>%  
  arrange(RMSE) %>%  
  slice(1)
```

**4. Establece si las siguientes frases son verdaderas o falsas.**

- a. El mejor indicador para evaluar la exactitud de un pronóstico es el MAPE.**
- b. Los modelos explicativos predicen mejor las series de tiempo.**
- c. Siempre debe escogerse el modelo con el mejor desempeño en el *test set*.**
- d. Los modelos simples de pronóstico son los mejores, por su sencillez.**