

# Actividad N° 03 PROFVIS

## Comparativa 5 Funciones Imputación y DB\_Table\_Clientes

Autor: **Quispe Morales, Josefh Jordy** (253864)  
Curso: Inteligencia de Negocios y Analítica de Datos  
Docente: Torres Cruz, Fred

31 de agosto de 2025

### 1. Repositorio a Github

- Haga click aqui: [GitHub de la Actividad 03](#)

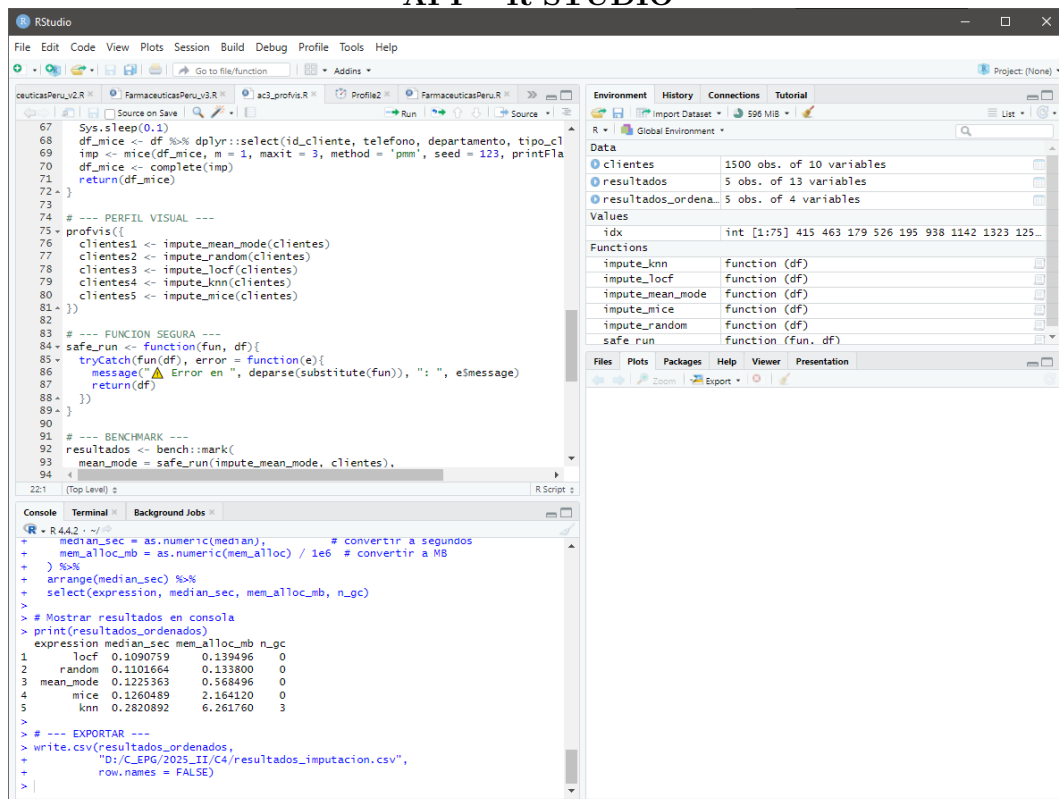
### 2. Introducción

Esta actividad recrea la comparativa usando PROFVIS mediante 5 funciones de imputación conjuntamente con la data de clientes de la base de datos Alpaca.dw.

### 3. Resultados

En cuanto a los resultados, se hace constar las evidencias del tipo captura de pantalla para el proceso de realización de la actividad.

#### APP - R STUDIO



```
67 Sys.sleep(0.1)
68 df_mice <- df %>% dplyr::select(id_cliente, telefono, departamento, tipo_cl)
69 imp <- mice(df_mice, m = 1, maxit = 3, method = "pmm", seed = 123, printF)
70 df_mice <- complete(imp)
71 return(df_mice)
72 }
73
74 # --- PERFIL VISUAL ---
75 profvis({
76   clientes1 <- impute_mean_mode(clientes)
77   clientes2 <- impute_random(clientes)
78   clientes3 <- impute_locf(clientes)
79   clientes4 <- impute_knn(clientes)
80   clientes5 <- impute_mice(clientes)
81 })
82
83 # --- FUNCION SEGURA ---
84 safe_run <- function(fun, df){
85   tryCatch(fun(df), error = function(e){
86     message("Error en ", deparse(substitute(fun)), ": ", e$message)
87   })
88 }
89
90
91 # --- BENCHMARK ---
92 resultados <- bench::mark(
93   mean_mode = safe_run(impute_mean_mode, clientes),
94 )
95
```

expression	median_sec	mem_alloc_mb	n_gc
locf	0.1090759	0.139496	0
random	0.1101664	0.133800	0
mean_mode	0.1225363	0.568496	0
mice	0.1260489	2.164120	0
knn	0.2820892	6.261760	3

Figura 1: Evidencia de ejecución del Código

## APP - R STUDIO

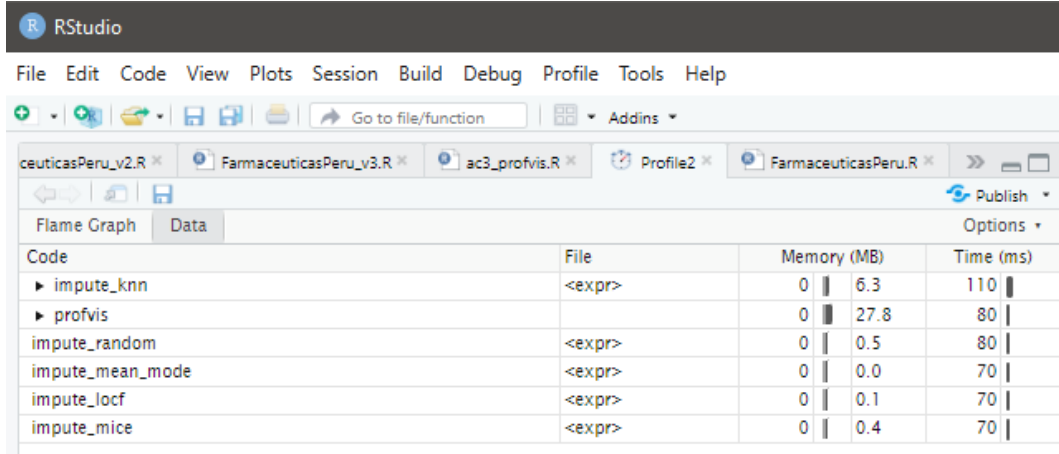


Figura 2: Resultados entregados

## APP - R STUDIO

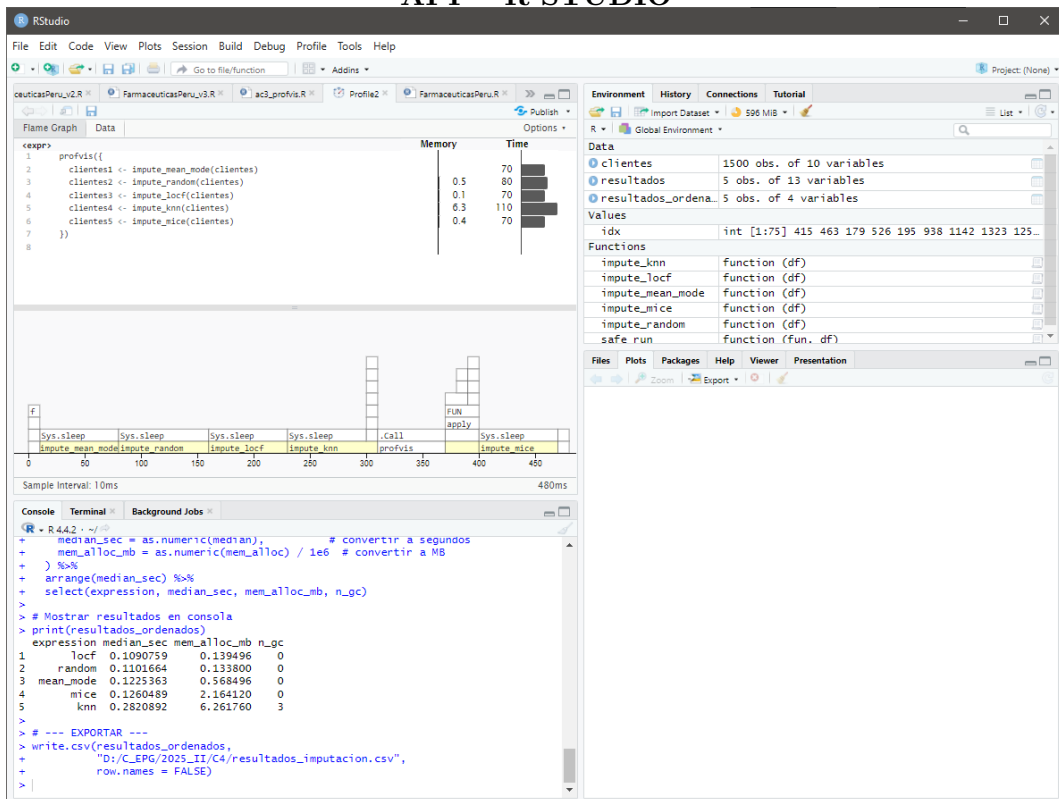


Figura 3: Evidencia Grafica de la Comparativa usando PROFVIS

## 4. Interpretación de Resultados

### Conclusión práctica:

- Si buscas **rapidez y eficiencia**: locf o random.
- Si quieres un **balance**: mice o mean\_mode.
- Si tienes más recursos y buscas **robustez**: knn.

Método	Tiempo (seg)	Memoria (MB)	Observaciones
<b>locf</b>	0.109	0.14	Muy rápido y liviano.
<b>random</b>	0.110	0.13	Muy rápido y liviano.
<b>mean_mode</b>	0.123	0.56	Intermedio en tiempo y memoria.
<b>mice</b>	0.126	2.16	Ligeramente más lento, memoria moderada.
<b>knn</b>	0.282	6.26	El más lento y pesado; forzó 3 GC.

Cuadro 1: Comparación de métodos de imputación: tiempo, memoria y estabilidad.