

# Diseño y Evaluación de RCTs

Isidoro Garcia

2021

```
library(tidyverse)
library(data.table)
library(broom)
library(knitr)
library(lubridate)
library(RCT)
```

## Contexto

Rappi te contrata para hacer una intervención que active a sus usuarios en la plataforma. La empresa quiere evaluar si es necesario darles cash los usuarios (y cuánto) para reactivar a los usuarios o si es suficiente con una campaña informativa.

Para ello, te decides a realizar un experimento factorial donde evalúas:

- El impacto de mandar un mensaje informativo donde muestres las nuevas tiendas afiliadas a Rappi, y
- El impacto de dar 100 ó 200 pesos en cupones

Finalmente, les gustaría entender cómo interactuar el mostrar las nuevas tiendas aunado con dar dinero en cupones.

A la empresa le gustaría entender el impacto de la intervención sobre:

- Las compras
- La tasa de usuarios activos en la app (transaccionar: abrir app o hacer compras)

## Datos

Los datos para asignar los pueden encontrar en `universo.Rdata`.

Cargemos los datos

```
load('Bases input/universo.Rdata')
```

**1. Cuántos grupos de tratamiento debe de haber? Elabora sobre que intervención va a recibir cada uno**

- Control: Ninguna intervención
- Trat 1: Sólo mensaje
- Treat 2: Sólo cupón \$100
- Treat 3: Sólo cupón de \$200
- Treat 4: Sólo cupón de \$300
- Treat 5: Mensaje + \$100
- Treat 6: Mensaje + \$200
- Treat 7: Mensaje + \$300

**2 (2pts).** Como pueden notar, tenemos 2 poblaciones: Usuarios inactivos y usuarios que nunca estuvieron activos. Para ellos, las tasas de transaccionalidad son hasta ahora 7.94% y 0%. Utiliza esta información para hacer pruebas de poder: Dada esta tasa y población, cuál es el efecto mínimo detectable sobre la tasa de transaccionalidad como función de cuantas observaciones asignamos al grupo control? Interpreta. (Tip: asegurate de dejar claros los grupos comparados en esta prueba)

**3.** Repite el mismo ejercicio pero ahora para usando las compras totales como variable objetivo. Elige un share de control con base en tu respuesta de esta y la anterior pregunta

**4 (2ptos)** Qué variables crees que puedan estar más correlacionadas con el impacto? Justifica tu respuesta y elige un set

**5 (2ptos)** Realiza la asignación aleatoria. Muestra la distribución de los grupos por estrato, los misfits. Sin mirar el balance, lograron una asignación aleatoria exitosa? Justifica tu respuesta

**6.** Qué elección tomaron sobre como manejar los misfits? Elaboren sus razones

**7.** Realiza las pruebas de balance t sobre todas las variables (Tip: transforma las categóricas en dummies). Parece haber balance?

**8.** Repite el ejercicio pero ahora con pruebas de balance conjuntas. Muestra los resultados (incluyendo el estadístico de prueba, grados de libertad y p values) Interpreta

**9.** Elabora porqué parecen cumplirse los 3 supuestos de la asignación

**10.** Elabora un pitch de negocio sobre los beneficios que este experimento podría dejar a Rappi.

## Evaluación

Pasemos a la evaluación de tu intervención. En este ejercicio, Rappi diseñó un nuevo experimento con tus enseñanzas algo distinto al tuyo.

Este consistió en 6 grupos de tratamiento y un control:

- T1: Dar 100 pesos en cupones (con mensaje)

- T2: Dar 200 pesos en cupones (con mensaje)
- T3: Dar un descuento de 20% en la siguiente compra
- T4: Dar un descuento de 25% en la siguiente compra
- T5: Ofrecer 2% de descuento en la siguiente compra por cada usuario que refieran
- T6: Ofrecer 4% de descuento en la siguiente compra por cada usuario que refieran

Te piden ahora medir este experimento (estratificado por `phone_verified`, `population` y `device_value`) en la base `base_evaluacion.Rdata`. Las variables endógenas son `total_purchases_after` que refleja el gasto total post-tratamiento y `transacted` que refleja abrir la app o hacer compras.

Carguemos la base

```
rm(list = ls())

load('Bases input/base_evaluacion.Rdata')
```

10 (2ptos). Muestra el estimador ITT para la tasa de transaccionalidad. Recuerda que tu cliente es un grupo empresarial. Por ende, muestra una gráfica donde se aprecie la diferencia entre los grupos de tratamiento y las significancias de manera sencilla. Interpreta tus resultados

11. Repite el ejercicio sobre compras totales. Que resultados se aprecian? Que indica esto sobre la rentabilidad del sistema de incentivos?

12. Interpreta el impacto del grupo de referidos 4%. Porque el estimador es tan diferente y a la vez es no significativo? Por que esto no paso en la tasa de transaccionalidad?

13 (2ptos). Repite la medición en 11 pero ahora con  $\log(\text{total\_purchases\_after}+1)$ . Que encuentras ahora? Interpreta las diferencias

14 (4ptos). Describe que variables necesitas para hacer un análisis costo beneficio completo. Les doy algunas: Ticket promedio \$100, Customer Lifetime value: \$1,100. Con esto, que sistema de incentivos recomendarías? Porqué? Muestra el razonamiento detrás de tu recomendación

15 (2ptos). Realiza la estimación de efectos heterogeneos para ambas variables usando `population`. Que encuentras? existe alguna subpoblación para la que los efectos difieran del promedio? Para cada efecto, muestra gráficas como lo hiciste en los ITTs

16 (2ptos). Repite el ejercicio para `'phone_verified'`.

17. Presenta una propuesta de focalización con base en tus resultados generales y heterogeneos.