# Curso de Estadística y Probabilidad

## Universidad Nacional Autónoma de México

Carlos Alberto Gómez Prado | Trabajo Final - 2 de 2

#### Tabla de contenidos

- Actividades
- Datos
- <u>Desarrollo</u>

## **Actividades**

**Objetivo:** analizar la información correspondiente a una muestra de datos del consumo de 500 clientes que pidieron un combo en un restaurante de hamburguesas, seleccionados aleatoriamente.

**El proyecto:** se divide en 2 partes, en el primer ejercicio se realizará un reporte sobre lo ocurrido en la muestra de 500 combos de hamburguesa, salsa, papas y refresco; en la segunda parte se resolverán algunas preguntas puntuales sobre la misma muestra.

## **Datos**

Los datos utilizados son de acceso libre y proporcionados por el curso de la <u>UNAM</u>.

La Tabla de datos se compone de las siguientes variables:

 $\mathbf{n}$  = 500 (n° de muestras)

Var(int)	Descripción: Cuantitativo - medidas(gramos)
Carne	Peso de la carne por hamburguesa.(grs)
Salsa	Cantidad de salsa ocupada por el cliente.(grs)

Var(cat)	Descripción: Cualitativo - medidas(Chico - Mediano- Grande)
Papas	Tamaño de las papas pedidas por el cliente.(c.m.g)
Refresco	Tamaño del refresco pedido por el cliente.(c.m.g)

# Metodología

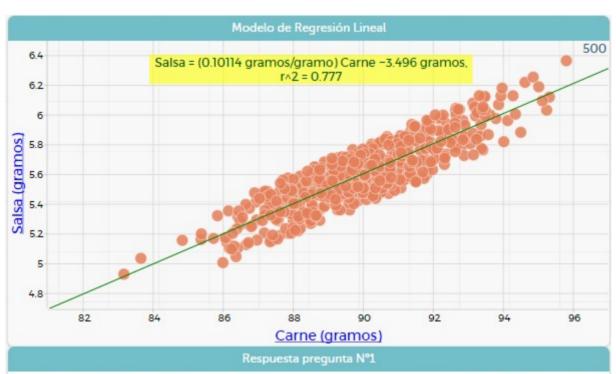
Para el desarrollo del proyecto se utilizó software y datos de uso libre siguiendo las normas y herramientas ofrecidas por el curso.

**Software**: <u>CODAP</u> es un software de ciencia de datos gratuito y de código libre creado para la educación por científicos de aprendizaje con experiencia y desarrolladores de software de código abierto, y financiado por subvenciones de la National Science Foundation.

## **Desarrollo**

#### 2ºParte | Ejercicio 2

- 1). ¿Cuánto vale el coeficiente de correlación entre las variables Carne y Salsa? Aproxima a dos decimales.
- **2).** ¿Qué cantidad de salsa, en gramos, se esperaría que un cliente le ponga a su hamburguesa si ésta tiene 89 gramos de carne? Redondea a dos decimales.



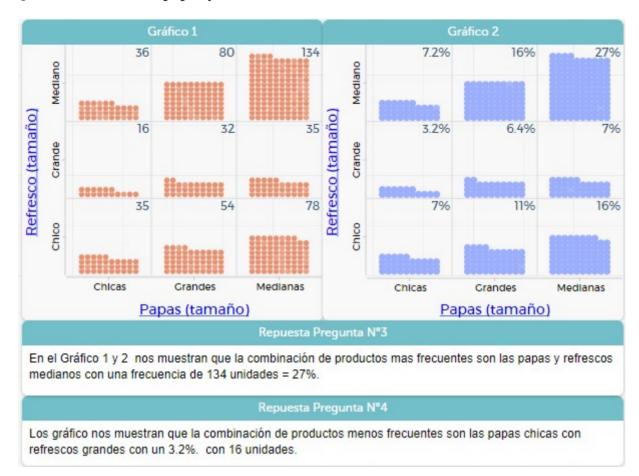
Es importante saber que el resultado del coeficiente de determinación oscila entre 0 y 1. Cuanto más cerca de 1 se sitúe su valor, mayor será el ajuste del modelo a la variable que estamos intentando explicar.

Del gráfico de regresión Lineal podemos extraer la información de su cociente de determinación  $R^2$ , Sacando su raíz obtenemos un cociente de correlación = CC = 0.88 = 88%, lo que nos da un valor muy cercano al  $n^{\circ}1$ , haciendo de este modelo muy bueno para predecir.

#### Respuesta pregunta Nº2

Observando el gráfico de regresión lineal podemos observar la recta que se ajusta a nuestros datos es: Y = 0.1011 X-3.5; Reemplazando la X por el valor a predecir. nuestro resultado nos dice que para 89 gr de carne se espera que el cliente utilice y = 5.4979, aprox = 5.50 gr de salsa. en su hamburguesa.

- 3). ¿Qué combinación de papas y refresco es la más frecuente?
- 4). ¿Qué combinación de papas y refresco es la menos frecuente?

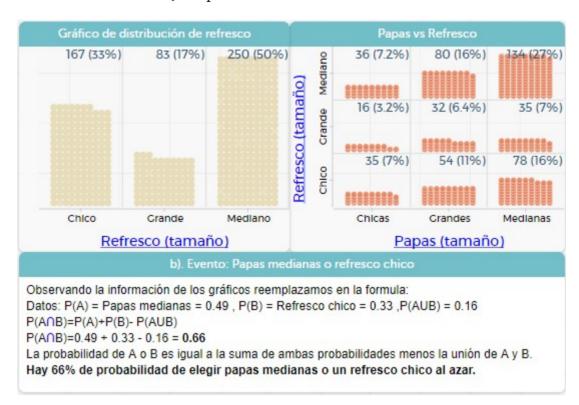


**5).** Calcula la probabilidad que hay de que un cliente seleccionado al azar haya pedido...

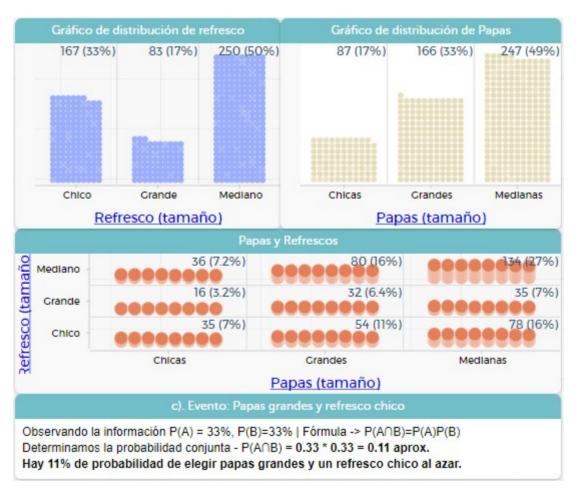
a)- Papas medianas.



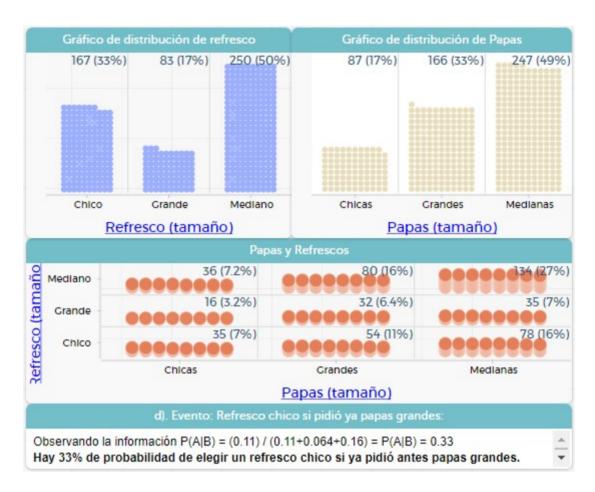
#### **b)-** Papas medianas o refresco chico.



### c)- Papas grandes y refresco chico.



#### **d)**- Refresco chico si pidio ya papas grandes.



6). ¿Los eventos papas grandes y refresco grande son independientes? Sí, No y Por qué.

