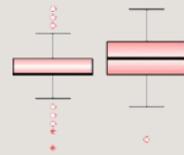


# Tema 1: Nociones Previas

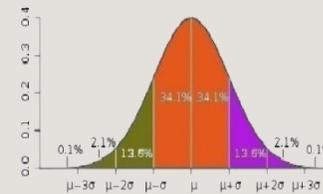
$$P(X=x) = \frac{n!}{x!(n-x)!} p^x q^{n-x}$$



$$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$



$$Z = \frac{X - \mu}{\sigma}$$



$$P(A|B) = \frac{P(A) \times P(B|A)}{\sum P(A) \times P(B|A)}$$



# Teoría de Conjuntos

La Teoría de Conjuntos es la forma en que podemos desarrollar la lógica matemática para resolver las tareas en cualquier programa de procesamiento de datos. Algunos conceptos importantes a recordar son:

- Las colecciones de objetos forman **Conjuntos**.
- Formar parte de un conjunto es **Pertenencia**.
- La subcolección de elementos forman **Subconjuntos**.
  - Todos los elementos de X forman parte de Y



- Conjunto “CG”: computadoras del gobierno del estado
- Pertenencia: las computadoras en la Dirección de Estadística (con número de inventario) pertenecen al gobierno del estado
- Subconjuntos: las computadoras de la Dirección de Estadística pertenecen al subconjunto de las computadoras del “Palacio del gobierno del estado” y no pertenecen al subconjunto de las computadoras de la Unidad de Servicios Estatales del gobierno del estado



# Teoría de Conjuntos

- Las computadoras del gobierno del estado de Sinaloa es un conjunto que pertenece a un Universo
- El Universo es la colección de los objetos disponibles, nuestros conjuntos y universos pueden ser finitos o infinitos
- Para agrupar el conjunto de las computadoras del gobierno del estado de Sinaloa, podemos definir lo que no pertenece e incluir lo que si pertenece
  - ¿Qué considero como computadora?
  - ¿Cómo identifico las computadoras no pertenecientes al Gobierno del Estado?



- La elaboración de conjuntos y subconjuntos va de la mano con la definición y reglas que establezcamos

El Universo “Trabajadores del gobierno del Estado”, puede estar conformado por los conjuntos de trabajadores según la dirección donde laboran.

Existen varios tipos de subconjuntos

- El subconjunto de los trabajadores de base y de confianza (aunque puede existir personal por honorarios o trabajadores temporales)
- El subconjunto de los trabajadores “adscritos” a cada dirección y una reclasificación de trabajadores con funciones en una dirección diferente a la que están adscritos



# Álgebra de Conjuntos

- ▶ Son operaciones básicas que nos van a permitir manipular las variables en SPSS, es decir, transformar elementos y crear conjuntos y subconjuntos:
  - Unión (Sumatoria de elementos)
  - Intersección (Elementos comunes entre 2 o más conjuntos)
  - Diferencia (Elementos de un conjunto que no pertenecen al segundo)
  - Diferencia Simétrica (Elementos no comunes pertenecientes a los conjuntos)
  - Complemento (El resto de los elementos que no pertenecen al conjunto para formar el Universo)



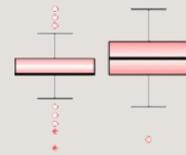
# Ejercicio 1

- En el supuesto de que todos los presentes vivimos en Sinaloa y vamos a tener días libres en Verano y todos vamos a salir a vacacionar:
  - Define el Universo (de las opciones para vacacionar)
  - Define conjuntos de opciones para agrupar los distintos destinos
  - Define un subconjunto de opciones para un destino



# Tema 2: Introducción al SPSS

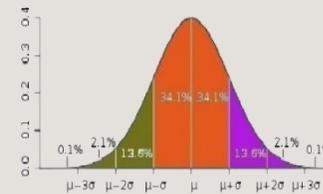
$$P(X=x) = \frac{n!}{x!(n-x)!} p^x q^{n-x}$$



$$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$



$$Z = \frac{X - \mu}{\sigma}$$



$$P(A|B) = \frac{P(A) \times P(B|A)}{\sum P(A) \times P(B|A)}$$



# Tipos de Variables (medida)

- Nominal: como su nombre lo menciona, su fin principal es nombrar (nombre de empresas o personas, puesto\*, todo pregunta abierta no numérica sin opciones de respuesta)
- Escalar: variables numéricas, se conoce exactamente que distancia hay entre las diferentes variables (edad\*, antigüedad en el trabajo, cms de cintura, número de hijos)
- Ordinal: como su nombre lo menciona, su fin principal es ordenar, son variables de texto o numéricas, no puedo decir que distancia exacta hay entre cada una pero si puedo saber cuál tiene más valor que la otra (edad\*, escolaridad, puesto\*\*, talla)



# Tipos de Variables (medida)

La edad es una variable numérica, varía con el tiempo, es comparable y se puede obtener de diferentes formas:

1.- Edad \_\_\_ años → 1 variable numérica

2.- Edad \_\_\_ años \_\_\_ meses → 2 variables numéricas

3.- 

de 0 a 5 años	}	se pierde la variable numérica y se vuelve ordinal
de 6 a 17 años		
de 18 a 29 años		
de 30 a 59 años		
60 o más años		

4.- Fecha de nacimiento → Variable fecha, se calcula en el tiempo



# Tipos de Variables (medida)

- Puesto es una variable nominal cuando se capta de forma abierta, pero se puede captar de forma ordinal cuando existe un catálogo previo
- Hay variables que se pueden obtener de diferente forma. A mayor detalle más dificultad para captarla.
- Existen rangos predefinidos confiables para un conjunto de personas: talla extra chica, chica, mediana, grande, extra grande para playeras deportivas de mujer. (las tallas cambian dependiendo el país o continente)



# Tipos de Variables (medida)

- La importancia de tener claro el tipo de variables que tenemos nos da la opción de pensar en como reclasificarlas, pasar una variable numérica a variable ordinal nos da la oportunidad de convertir la cantidad en una cualidad, ejemplos de ello son:
  - Al captar la **cantidad de ingresos** puedo dar la cualidad de pobreza vs riqueza
  - Al captar el **número de habitantes** de una vivienda por su parentesco respecto al jefe de familia, puedo clasificar una base de datos por tipo de familia
  - Con el **número de habitaciones** en una vivienda y el **número de habitantes** puedo saber si hay hacinamiento y/o medir el nivel de hacinamiento



# Variables Dicotómicas

- Variable de ausencia o presencia o de 2 opciones de respuesta excluyentes
  - Si / No
  - Hombre y Mujer



# Tipos de Variables (medida)

## Escala de Tipo Likert:

- Son variables ordinales donde los extremos son opuestos
  - Siempre vs Nunca (Ausencia y Presencia)
  - Muy bueno vs Muy malo
- Representación ordinal de variables numéricas, es decir, es darle una cualidad a lo cuantitativo.
- ¿Cuándo es mejor usar una variable escalar a una ordinal?
  - Cuando se comprometa el dato por la subjetividad
- ¿Cuándo es mejor usar una variable ordinal a una escalar?
  - Cuando lo que quiera medir sea una opinión, la satisfacción o los servicios



El SPSS es:

- Base de datos fija (excel es una hoja de cálculo dinámica) Vista de Datos
- Sistema de codificación de variables (fichas de datos) Vista de Variables
- Es un paquete estadístico que genera tablas, cruces de variables y gráficas Outputs o cuadros de salida
- Transforma o calcula nuevas variables. Calculadora o menú de opciones
- Trabaja en un ambiente amigable aunque se están generando comandos de programación y pueden ser replicados en otra computadora con la misma base, o bien, puede ser replicada con una base actualizada. Sintaxis



# Tipos de Variables (caracteres)

- Numérico
- Fecha
- Dólar (moneda)
- Cadena (alfanumérico)
- Número restringido (entero con ceros iniciales)
  - En algunos casos se puede definir el ancho de las variables y el número de decimales
  - Elegir el tipo de variable indicado facilita la validación de los datos
  - Elegir el tipo correcto permite crear fórmulas u operaciones por ejemplo: saber cuanto transcurrió de una fecha a otra



# Vista de Datos

- Función de captura o análisis: en la primera se definen las variables en vista de código, para el análisis se activa la vista de etiquetas
- Se ordenan los datos por columna.
- Un renglón es un caso

9 : acc\_alim15

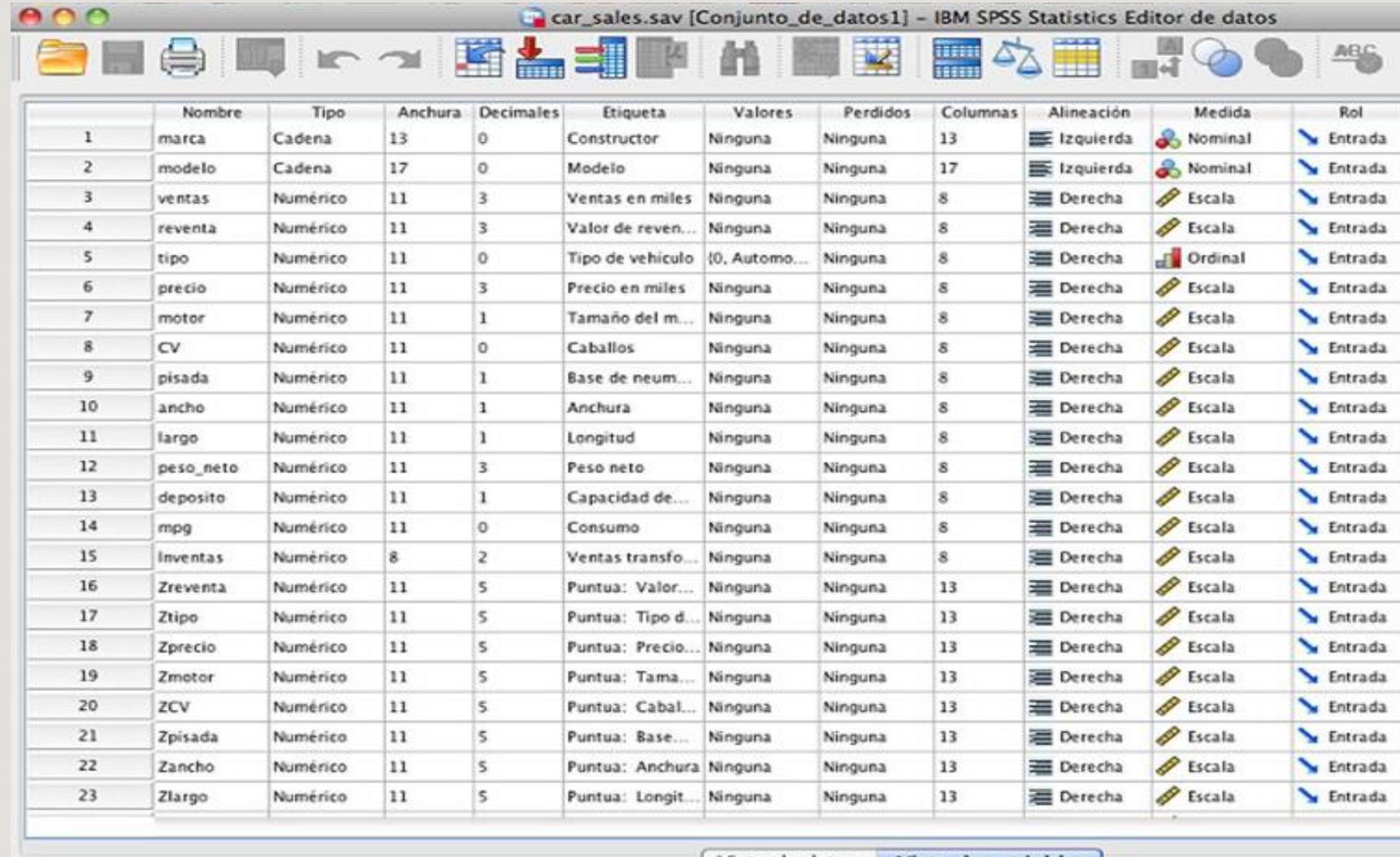
	trab_com_e	acc_alim_1	acc_alim_2	acc_alim_3	acc_alim_4	acc_alim_5	acc_alim_6	acc_alim_7	acc_alim_8	acc_alim_9	acc_alim_10	acc_alim_11	acc_alim_12
1	. Sí	No	No	No									
2	. No	No	No	No	No	No	No	No	No	No	No	No	No
3	. No	No	No	No	No	No	No	No	No	No	No	No	No
4	. No	No	No	No	No	No	No	No	No	No	No	No	No
5	. Sí	No	Sí	No	No	No							
6	. No	No	No	No	No	No	No	No	No	No	No	No	No
7	. Sí	No	No	No									
8	. No	No	No	No	No	No	No	No	No	No	No	No	No
9	. No	No	No	No	No	No	No	No	No	No	No	No	No
10	. No	No	No	Sí	No	No	No						
11	. No	No	No	No	No	No	No	No	No	No	No	No	No
12	. Sí	Sí	Sí	No	No	No							
13	. No	No	No	No	No	No	No	No	No	No	No	No	No
14	. No	No	No	No	No	No	No	No	No	No	No	No	No
15	. No	No	Sí	Sí	No	No	No						
16	. Sí	No	Sí	Sí	No	No	No						
17	. No	No	No	No	No	No	No	No	No	No	No	No	No
18	. No	No	No	No	No	No	No	No	No	No	No	No	No
19	. No	No	No	No	No	No	No	No	No	No	No	No	No
20	. No	No	No	No	No	No	No	No	No	No	No	No	No
21	. No	No	No	No	No	No	No	No	No	No	No	No	No
22	. No	No	No	No	No	No	No	No	No	No	No	No	No
23	. No	No	No	No	No	No	No	No	No	No	No	No	No
24	. No	No	No	No	No	No	No	No	No	No	No	No	No
25	. No	No	No	No	No	No	No	No	No	No	No	No	No
26	. No	No	No	No	No	No	No	No	No	No	No	No	No
27	. No	No	No	No	No	No	No	No	No	No	No	No	No
28	. No	No	No	No	No	No	No	No	No	No	No	No	No
29	. No	No	No	No	No	No	No	No	No	No	No	No	No
30	. No	No	No	No	No	No	No	No	No	No	No	No	No
31	. Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí
32	. No	No	No	No	No	No	No	No	No	No	No	No	No

Vista de datos Vista de variables



# Vista de Variables

- Se definen los 2 nombres: “observado y descriptivo”
- Se definen validaciones como el tipo de variable, si es numérico no acepta símbolos ni letras, la anchura solo aceptará la indicada
- Se definen los significados de las codificaciones. (variables con medidas ordinales)



	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos	Columnas	Alineación	Medida	Rol
1	marca	Cadena	13	0	Constructor	Ninguna	Ninguna	13	Izquierda	Nominal	Entrada
2	modelo	Cadena	17	0	Modelo	Ninguna	Ninguna	17	Izquierda	Nominal	Entrada
3	ventas	Numérico	11	3	Ventas en miles	Ninguna	Ninguna	8	Derecha	Escala	Entrada
4	reventa	Numérico	11	3	Valor de reven...	Ninguna	Ninguna	8	Derecha	Escala	Entrada
5	tipo	Numérico	11	0	Tipo de vehiculo	{0, Automo...	Ninguna	8	Derecha	Ordinal	Entrada
6	precio	Numérico	11	3	Precio en miles	Ninguna	Ninguna	8	Derecha	Escala	Entrada
7	motor	Numérico	11	1	Tamaño del m...	Ninguna	Ninguna	8	Derecha	Escala	Entrada
8	CV	Numérico	11	0	Caballos	Ninguna	Ninguna	8	Derecha	Escala	Entrada
9	pisada	Numérico	11	1	Base de neum...	Ninguna	Ninguna	8	Derecha	Escala	Entrada
10	ancho	Numérico	11	1	Anchura	Ninguna	Ninguna	8	Derecha	Escala	Entrada
11	largo	Numérico	11	1	Longitud	Ninguna	Ninguna	8	Derecha	Escala	Entrada
12	peso_net	Numérico	11	3	Peso neto	Ninguna	Ninguna	8	Derecha	Escala	Entrada
13	deposito	Numérico	11	1	Capacidad de...	Ninguna	Ninguna	8	Derecha	Escala	Entrada
14	mpg	Numérico	11	0	Consumo	Ninguna	Ninguna	8	Derecha	Escala	Entrada
15	inventas	Numérico	8	2	Ventas transfo...	Ninguna	Ninguna	8	Derecha	Escala	Entrada
16	Zreventa	Numérico	11	5	Puntua: Valor...	Ninguna	Ninguna	13	Derecha	Escala	Entrada
17	Ztipo	Numérico	11	5	Puntua: Tipo d...	Ninguna	Ninguna	13	Derecha	Escala	Entrada
18	Zprecio	Numérico	11	5	Puntua: Precio...	Ninguna	Ninguna	13	Derecha	Escala	Entrada
19	Zmotor	Numérico	11	5	Puntua: Tama...	Ninguna	Ninguna	13	Derecha	Escala	Entrada
20	ZCV	Numérico	11	5	Puntua: Cabal...	Ninguna	Ninguna	13	Derecha	Escala	Entrada
21	Zpisada	Numérico	11	5	Puntua: Base...	Ninguna	Ninguna	13	Derecha	Escala	Entrada
22	Zancho	Numérico	11	5	Puntua: Anchura	Ninguna	Ninguna	13	Derecha	Escala	Entrada
23	Zlargo	Numérico	11	5	Puntua: Longit...	Ninguna	Ninguna	13	Derecha	Escala	Entrada



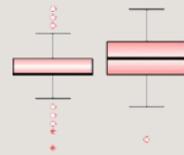
- Bitácora para guardar recodificaciones, transformaciones, cálculos etc. Se debe dar clic en pegar en lugar de aceptar a cada operación
- Facilita el proceso para replicar tareas
- No se necesita aprender el lenguaje, se realiza una tarea manual y se pega

```
1 DATASET ACTIVATE ConjuntoDatos1.
2
3 DESCRIPTIVES VARIABLES=num_trab_d
4 /STATISTICS=MEAN STDDEV MIN MAX.
5
```



# Tema 3: Operaciones básicas

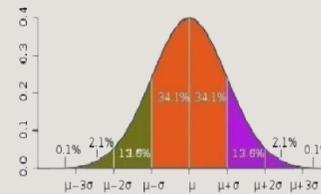
$$P(X=x) = \frac{n!}{x!(n-x)!} p^x q^{n-x}$$



$$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$



$$Z = \frac{X - \mu}{\sigma}$$



$$P(A|B) = \frac{P(A) \times P(B|A)}{\sum P(A) \times P(B|A)}$$



# Tipos de bases de datos a usar en SPSS

Título Variable 1	Título Variable 2	Título Variable 3	Título Variable 4	Título Variable 5
Caso 1				
Caso 2				
Caso 3				
Caso 4				
Caso 5				



# Tipos de bases de datos a usar en SPSS

Título Variable 1	Título Variable 2	Título Variable 3	Título Variable 4	Título Variable 5
Grupo 1	Caso 1	Caso 1	Caso 1	Caso 1
Grupo 1	Caso 2	Caso 2	Caso 2	Caso 2
Grupo 1	Caso 3	Caso 3	Caso 3	Caso 3
Grupo 2	Caso 1	Caso 1	Caso 1	Caso 1
Grupo 2	Caso 2	Caso 2	Caso 2	Caso 2



# Tipos de bases de datos a usar en SPSS

Título Variable 1	Título Variable 2	Título Variable 3	Título Variable 4	Título Variable 5
Grupo 1	Caso 1	Caso 1	Variable común del grupo 1	Caso 1
Grupo 1	Caso 2	Caso 2	Variable común del grupo 1	Caso 2
Grupo 1	Caso 3	Caso 3	Variable común del grupo 1	Caso 3
Grupo 2	Caso 1	Caso 1	Variable común del grupo 2	Caso 1
Grupo 2	Caso 2	Caso 2	Variable común del grupo 2	Caso 2



# Reglas para el llenado de la vista de Variables

Nombre:	Inicio con letra, van sin espacios, se usa guión bajo (corto)
Tipo:	Numérico, Fecha, Dólar, Cadena, Número Restringido
Decimales:	Numérico, Dólar
Etiqueta:	Nombre largo, pregunta completa, descripción
Valores	Significado de codificación para las variables numéricas (ordinales)
Perdidos	Valores codificados que no se deben tomar en cuenta para cálculos
Medida	Escalar (acepta números*), Nominal y Ordinal aceptan todos los valores**



# Operaciones Básicas

- Estadísticas
  - Frecuencias
  - Descriptivos
  - Tablas Cruzadas
  - Comparar Medias



# Operaciones Básicas: Frecuencias

- ANALIZAR/
  - ESTADÍSTICOS DESCRIPTIVOS/
    - FRECUENCIAS/

Frecuencia: Cantidad de veces que se presenta cada caso de las diferentes variables y su porcentaje respecto del total. \*Los valores perdidos se toman en cuenta solo en variables ordinales y numéricas, en las variables nominales un caso vacío es un caso válido.



# Operaciones Básicas: Frecuencias

- ANALIZAR/
  - ESTADÍSTICOS DESCRIPTIVOS/
    - FRECUENCIAS/

Frecuencia: Para variables numéricas como edad, número de hijos.

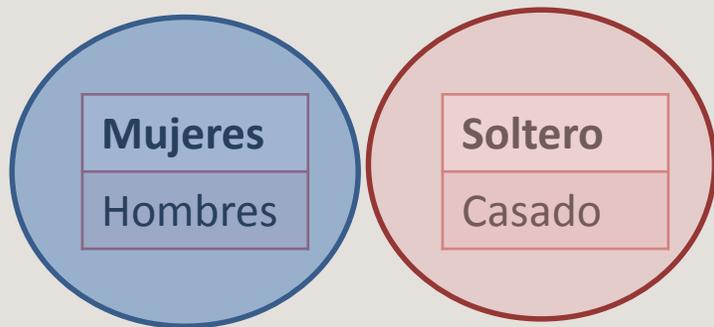
- » Moda
- » Mediana
- » Media
- » Mínimo
- » Máximo



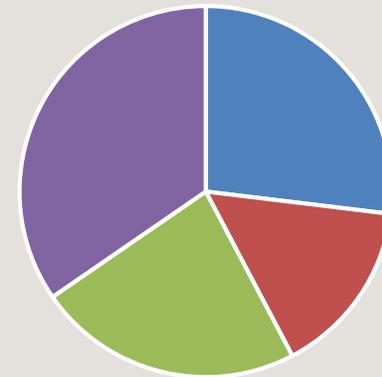
# Operaciones Básicas: Tablas Cruzadas

- ANALIZAR/
  - ESTADÍSTICOS DESCRIPTIVOS/
    - TABLAS CRUZADAS/

Cruce de Tablas: es hacer una división de 2 conjuntos entre todos sus elementos y posteriormente unir cada elementos del subconjunto 1 con el conjunto 2



<b>Mujeres Solteras</b>
Mujeres Casadas
Hombres Solteros
Hombres Casados



# Operaciones Básicas: Tablas Cruzadas

## COLUMNAS

## FILAS

	Solter@s	Casad@s	
Mujeres	X	X	Total
Hombres	X	X	Total
	Total	Total	Total



# Operaciones Básicas: Tablas Cruzadas

- ANALIZAR/
  - ESTADÍSTICOS DESCRIPTIVOS/
    - TABLAS CRUZADAS/
      - ELEGIR FILAS Y COLUMNAS
        - CASILLAS (PORCENTAJES EN FILAS)

Las opciones de elección se deben reducir a nuestro objeto de estudio en las filas así como su porcentaje.

En el presente ejemplo nos interesa saber que porcentaje de mujeres son solteras y que porcentaje son casadas, al igual que los hombres solteros y casados los cuales nos sumarán el 100%.

## COLUMNAS

## FILAS

	Solter@s	Casad@s	
Mujeres	X	X	Total
Hombres	X	X	Total
	Total	Total	Total



# Operaciones Básicas: Tablas Cruzadas

- ANALIZAR/
  - ESTADÍSTICOS DESCRIPTIVOS/
    - FRECUENCIAS/
      - ELEGIR FILAS Y COLUMNAS
        - CASILLAS (PORCENTAJES EN FILAS)

En el presente ejemplo nos interesa saber que porcentaje de mujeres si usan internet y que porcentaje no lo usa, al mismo tiempo puedo saber los mismos porcentajes respecto a los hombres y comparar el porcentaje de mujeres que si usan VS. el porcentaje de los hombres que si lo hacen.

## COLUMNAS

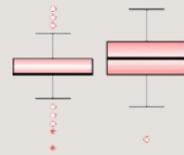
## FILAS

	Si usan Internet	No usan Internet	
Mujeres	X	X	Total
Hombres	X	X	Total
	Total	Total	Total



# Tema 4: Manejo de Bases

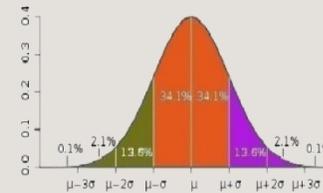
$$P(X=x) = \frac{n!}{x!(n-x)!} p^x q^{n-x}$$



$$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$



$$Z = \frac{X - \mu}{\sigma}$$



$$P(A|B) = \frac{P(A) \times P(B|A)}{\sum P(A) \times P(B|A)}$$

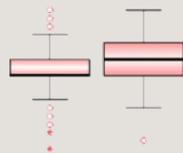


- Operaciones en bases y manejo de los datos
  - Transformar (calcular variables)
  - Recodificación de variables
  - Reemplazar valores perdidos
  - Segmentar archivo
  - Dividir en archivos
  - Seleccionar casos
  - Ponderar casos
  - Ordenar casos
  - Fusionar archivos
    - Añadir casos
    - Añadir variables
  - Reestructurar bases de datos



# Tema 5: Operaciones básicas con manejo avanzado de bases

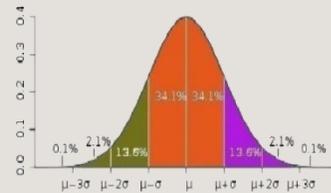
$$P(X=x) = \frac{n!}{x!(n-x)!} p^x q^{n-x}$$



$$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$



$$Z = \frac{X - \mu}{\sigma}$$



$$P(A|B) = \frac{P(A) \times P(B|A)}{\sum P(A) \times P(B|A)}$$



- Operaciones en bases y manejo de los datos
  - Transformar (calcular variables) con condiciones de selección de casos
  - Frecuencias para segmentar archivos en cuartiles, deciles o grupos iguales
  - Recodificación con selección de casos
  - Medias con capas
  - Cruces de tablas con capas

