

TABLAS DE CONTINGENCIAS

Seguimos avanzando con los conceptos de estadística descriptiva.

Ahora vas a practicar con las tablas de contingencias.

Para entenderlas es mejor ir paso a paso practicando.

ABRE LA TABLA DE DATOS “esoph”

Si has hecho la primera hoja de trabajo, ahora sólo tendrás que copiar la siguiente tabla:

Más información: <https://stat.ethz.ch/R-manual/R-devel/library/datasets/html/esoph.html>

Nombre Variable	¿Qué característica mide?	¿Qué unidades tiene?	¿Qué tipo de variable es?
agegp	Mide la pertenencia de un individuo a un grupo, es decir en qué rango de edad se encuentra	No tiene, unidades, representa grupos de edades cómo una etiqueta	Categórica ordinal los valores tienen un orden natural pero no se pueden medir o cuantificar en términos de valores específicos.
alcgp	Consumo de alcohol	Gramos / día	Categorica ordinal, ya que los valores se dividen en categorías ordinales que indican diferentes niveles de consumo de alcohol, pero no representan valores numéricos continuos.
tobgp	Consumo de tabaco por día de cada individuo	Gramos / día	Categorica ordinal, ya que los valores se dividen en categorías ordinales que indican diferentes niveles de consumo de tabaco
ncases	Número de casos de personas con cáncer de esófago	No tiene unidades específicas	Cuantitativa discreta
ncontrols	¿Número de personas del grupo de control?	No tiene unidades específicas	Cuantitativa discreta

TABLAS DE CONTINGENCIAS

"Oct-19" parece una convención para indicar que no consume tabaco. Es cierto?			
---	--	--	--

TABLAS DE FRECUENCIAS y PROPORCIONES

Calcula la tabla de contingencias para la variable cualitativa: "agegp"

Grupos en años	Frecuencia	Frec. Relativa	Proporción (porcentaje)
25-34	15	0.1705	17.05
35-44	15	0.1818	18.18
45-54	16	0.1818	18.18
55-64	16	0.1818	18.18
65-74	15	0.1705	17.05
75+	11	0.1250	12.50
Total	88		

Esta tabla representa por categorías el número diferente de combinaciones de consumo de tabaco y de consumo de alcohol, no el número de consumidores que contiene cada categoría!

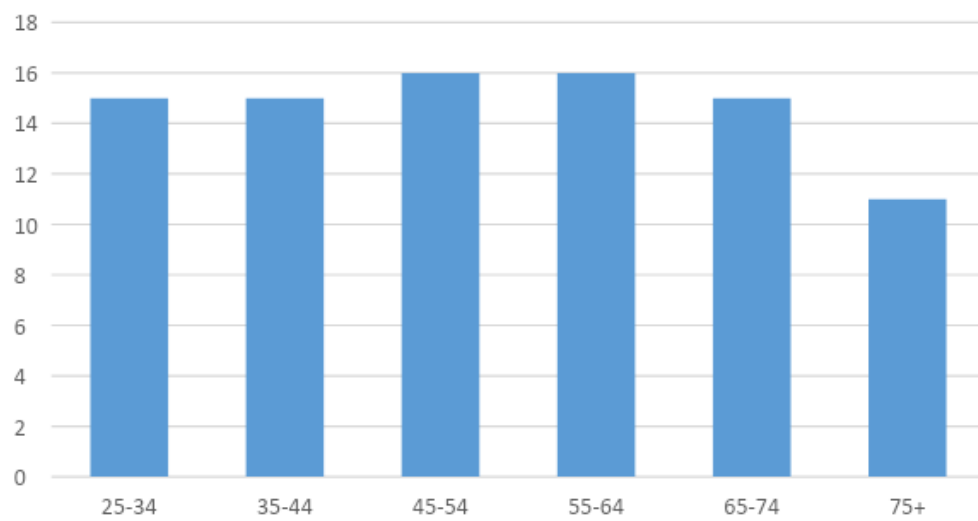
DIAGRAMA DE BARRAS y SECTORES

Describe esta tabla de datos con un diagrama de sectores y de barras.

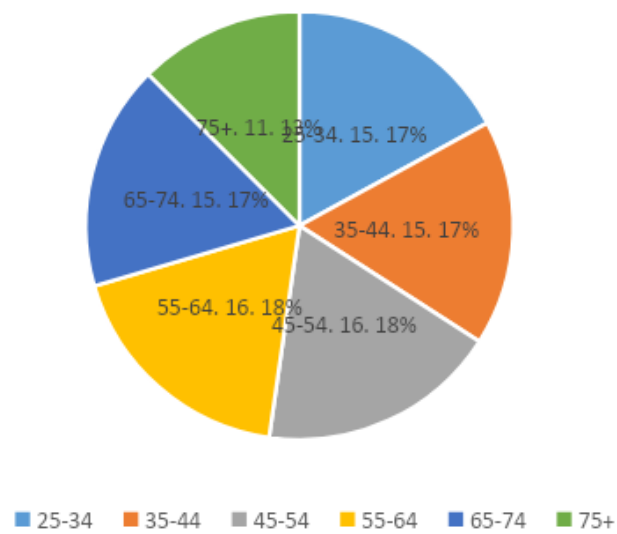
Copia el diagrama de barras y el de sectores para representar la tabla anterior. Puedes hacerlo en frecuencia o en proporciones.

TABLAS DE CONTINGENCIAS

Distribución etaria población experimento



Distribución etaria población experimento



TABLAS DE CONTINGENCIAS

TABLA DE CONTINGENCIAS

Ahora vamos a observar las parejas de grupos. Entre los grupos de edad ("agegp") y los grupos de casos ("ncases").

De esta forma sabremos la distribución de casos de Cáncer por grupos de edad:

Grupos de años	Nº casos grupos								
	1	2	3	4	5	6	8	9	17
25-34	1								
35-44	2	2	1						
45-54	2	2	2	3	2	2			
55-64		2	4	3	2	2	1	2	
65-74	4	2	2	2	2	1			1
75+	7	3							

Observa la tabla y calcula con Excel el número de pacientes con cáncer por cada grupo de edad. (pista: si tenemos 7 casos de 1 y 10 de 6, el número de pacientes es $7*1+10*6=76$)

Grupos en años	Pacientes totales
25-34	1
35-44	9
45-54	46
55-64	76
65-74	55
75+	13
Total	200

ABRE LA TABLA DE DATOS "WorldPhones"

Esta tabla de datos es en sí una tabla de contingencias.

Tenemos dos variables:

- Los años: ordinal (en filas)
- Los continentes: nominal (en columnas)

CALCULA LAS FRECUENCIAS CONDICIONADAS

Ahora se trata de trabajar el concepto de frecuencias condicionadas e interpretar su significado.

Recuerda que las frecuencias condicionadas son frecuencias relativas condicionando a los grupos (dividir la frecuencia absoluta por el total de observaciones del grupo)

Frecuencias Condicionadas por columnas (continentes)

Calcula las frecuencias condicionadas por columnas.

En este caso a las columnas tenemos la variable “continentes”.

Para cada continente calcularemos el porcentaje condicionado para los diferentes años.

Year	N.Amer	Europe	Asia	S.Amer	Oceania	Africa	Mid.Amer
1951	9.83	8.97	6.60	9.35	8.96	0.86	9.42
1956	12.93	12.47	10.80	13.23	12.88	13.58	12.44
1957	13.85	13.52	11.99	13.89	13.75	14.88	13.12
1958	14.66	14.65	15.28	14.66	14.64	16.01	14.19
1959	15.37	15.64	15.72	15.46	15.61	17.03	15.46
1960	16.27	16.78	18.85	16.21	16.62	18.34	17.11
1961	17.09	17.96	20.76	17.20	17.55	19.30	18.26
TOTAL	100	100	100	100	100	100	100

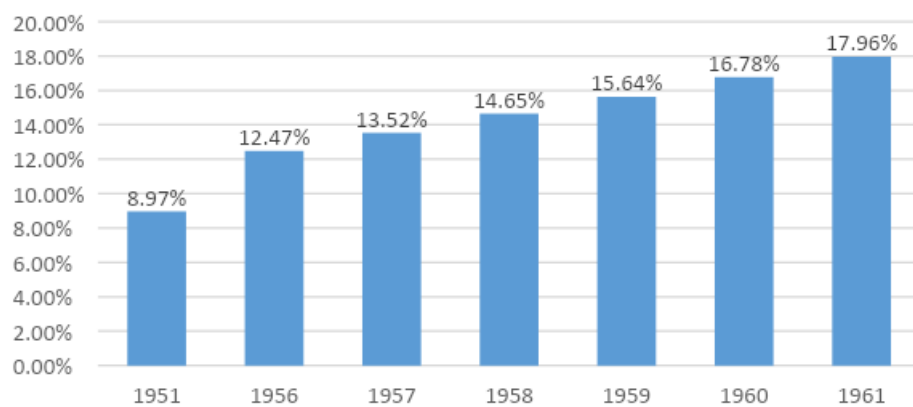
Dibuja el diagrama de barras o sectores para el continente europeo

Dibuja el diagrama de barras para el continente europeo e interpreta los resultados. ¿Tiene algún sentido?

Si, muestra la distribución por año del porcentaje del total de teléfonos adquiridos en Europa en la década 1951-1961.

TABLAS DE CONTINGENCIAS

Distribución por año del porcentaje del total de
teléfonos adquiridos en Europa en la década
1951-1961



TABLAS DE CONTINGENCIAS

Frecuencias Condicionadas por filas (años)

Calcula las frecuencias condicionadas por filas.

En este caso en las filas tenemos la variable “años”.

Para cada año calcularemos el porcentaje de cada continente. (Yo creo que tiene más sentido este cálculo en filas que en columnas, ¿verdad?)

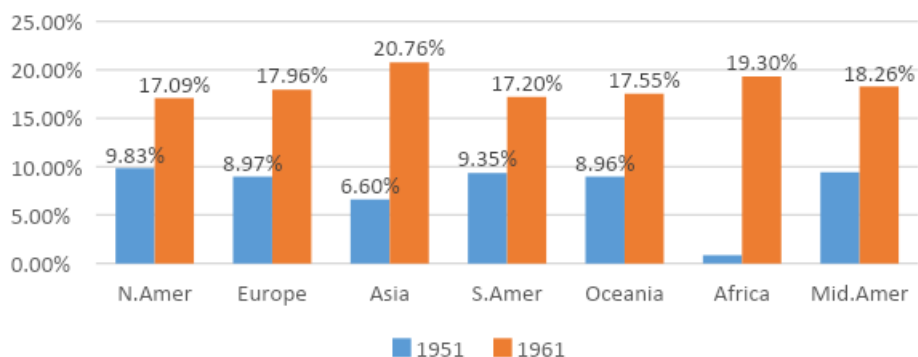
Year	N.Amer	Europe	Asia	S. Amer	Oceania	Africa	Mid. Amer	TOTAL
1951	61.67	28.96	3.86	2.44	2.21	0.12	0.75	100
1956	59.12	29.34	4.61	2.51	2.32	1.38	0.72	100
1957	58.84	29.55	4.75	2.45	2.30	1.41	0.70	100
1958	57.84	29.75	5.63	2.40	2.27	1.40	0.71	100
1959	56.87	30.17	6.15	2.35	2.28	1.42	0.75	100
1960	56.87	30.17	6.15	2.35	2.28	1.42	0.75	100
1961	56.34	30.47	6.39	2.36	2.28	1.41	0.76	100

Dibuja el diagrama de barras o sectores para el año 1951 y el 1961. Visualiza el cambio de proporciones.

Dibuja los dos diagramas de barras o sectores e interpreta los resultados.

El gráfico muestra la participación porcentual de los años 1951 y 1961 del numero total de teléfonos llegados a cada Continente en dicha década.

Participación porcentual de los años 1951 y 1961
del numero total de teléfonos llegados a cada
Continente en dicha década



TABLAS DE CONTINGENCIAS

¡Estás haciendo un gran trabajo para entender los conceptos a base de ejercicios prácticos!

¡Bien!

¡Vamos a seguir avanzando!