

Nombre de alumnos: Fernanda

Patricia Hernández Díaz

Nombre del profesor: Rosario Gómez

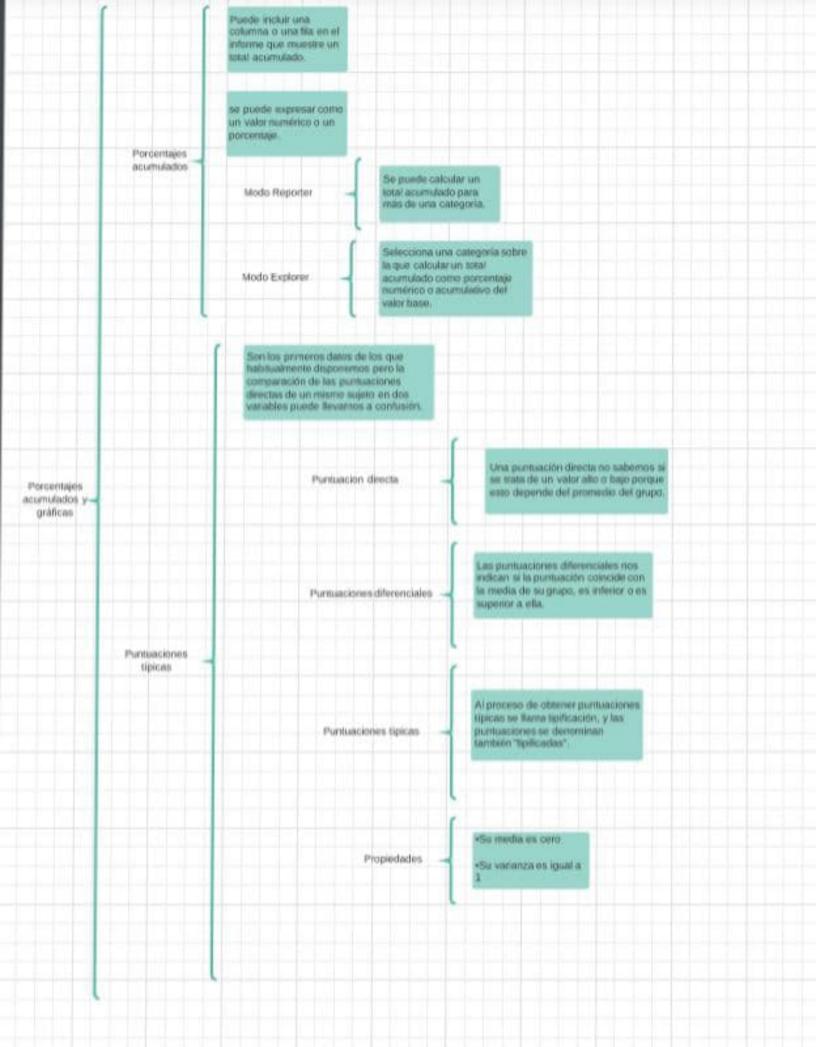
Nombre del trabajo: Porcentajes

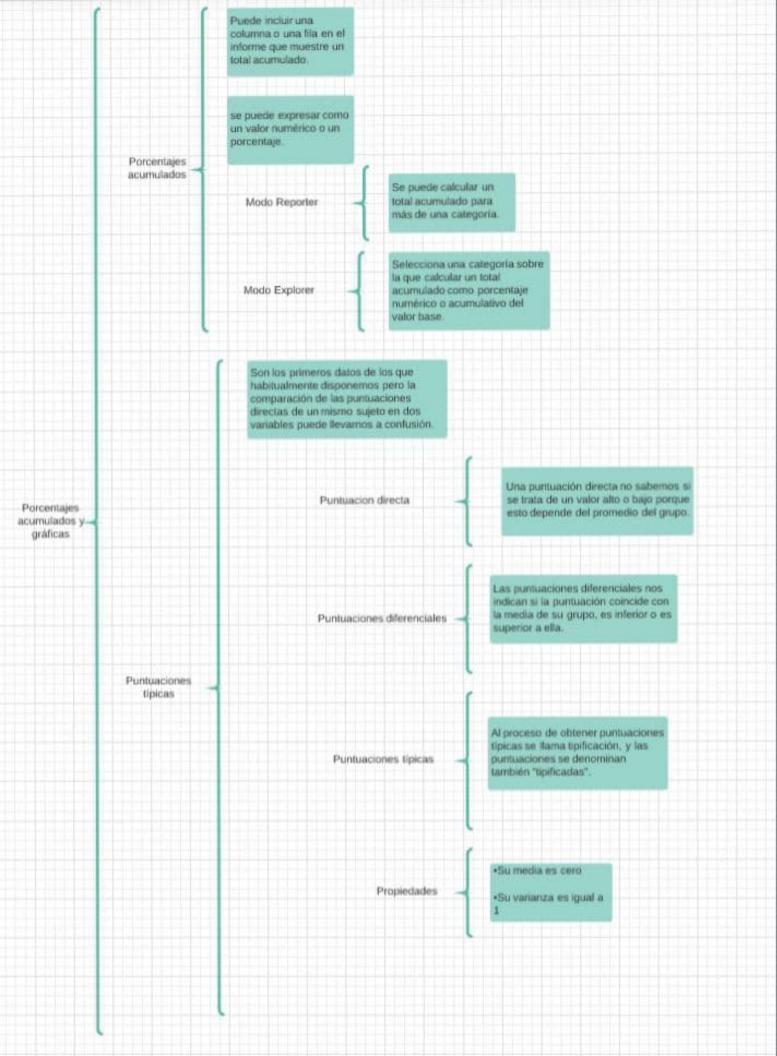
acumulados y gráficas

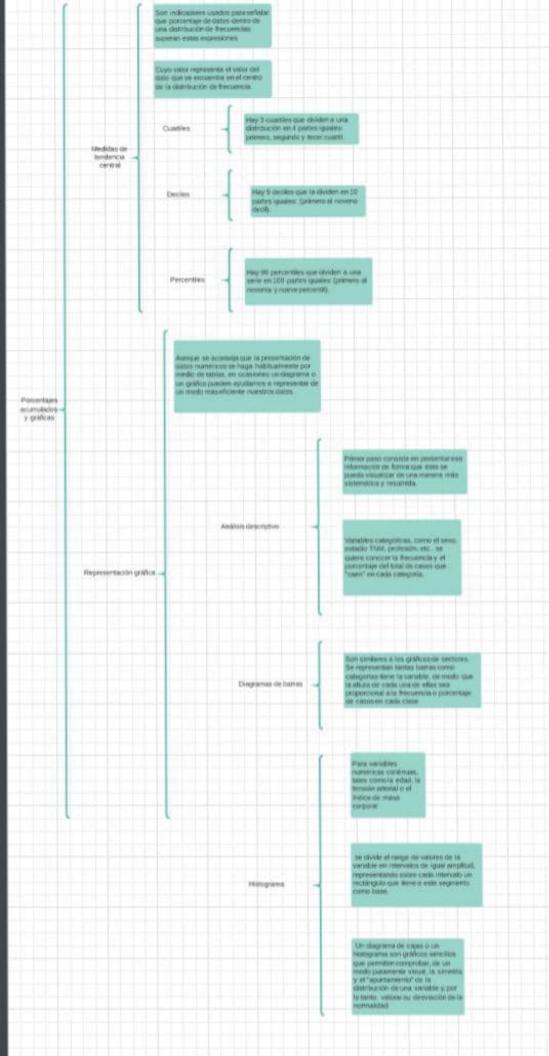
Materia: Estadística Descriptiva

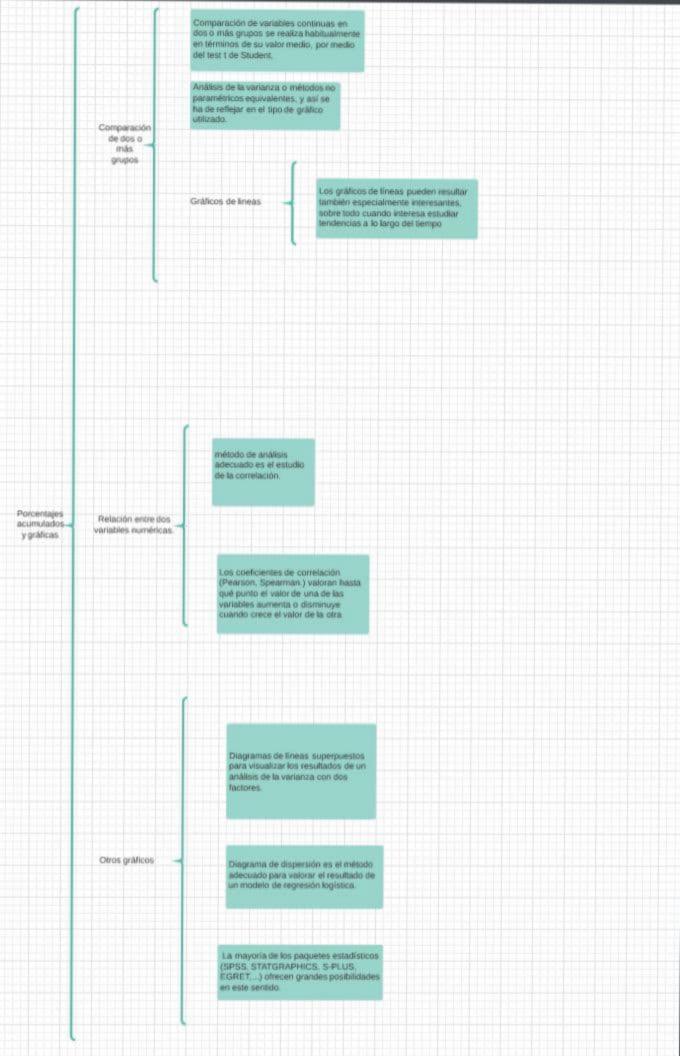
Grado: 2º Cuatrimestre

Grupo: LPS19SSCO9-A









### **EJERCICIO**

Un grupo de investigadores pertenecientes a la secretaría de seguridad pública, tomó una muestra aleatoria de las velocidades (km/h) registradas por 30 vehículos en el trayecto Pichucalco-Villahermosa, con el fin de establecer nuevos límites máximos de velocidad para una carretera. La muestra arrojo los datos siguientes: 90, 99, 104, 99, 119, 98, 95, 112, 95, 120, 100, 90, 116, 96, 114, 108, 98, 118, 100, 106, 114, 100, 112, 106, 100, 115, 111, 105, 114, 97

 Construye una distribución de frecuencia para velocidades, que tenga frecuencia absoluta acumulada, frecuencia relativa, frecuencia relativa acumulada y marca de clase así como también un histograma.

Intervalo de clase	f	fa	fr	fra	MC	%	% acumulados
[90-95)	2	2	0.07	0.07	92.5	7	7
[95-100)	8	10	0.27	0.34	97.5	27	34
[100-105)	5	15	0.17	0.51	102.5	17	51
[105-110)	4	19	0.13	0.64	107.5	13	64
[110-115)	6	25	0.2	0.84	112.5	20	84
[115-120]	5	30	0.16	1	117.5	16	100
Total	30		1			100	

# Operación

$$K = 1 + 3.3 \text{Log} (30) = 1 + 3.3 (1.4771212547) = 1 + 4.87 = 5.87 \approx 6$$

Fr=f/N

Fr = 2/30 = 0.07

Fr = 8/30 = 0.27

Fr = 5/30 = 0.17

Fr = 4/30 = 0.13

Fr = 6/30 = 0.2

Fr = 5/30 = 0.16

## ESTADÍSTICOS DE ASOCIACIÓN ENTRE VARIABLES.

#### Introducción

Para empezar, para determinar la relación entre dos variables cuantitativas en un grupo; debe seguir los objetivos de dicho análisis:

- a. Determinar si las dos variables están correlacionadas, es decir si los valores de una variable tienden a ser más altos o más bajos para valores más altos o más bajos de la otra variable.
- b. Poder predecir el valor de una variable dado un valor determinado de la otra variable.
- valorar el nivel de concordancia entre los valores de las dos variables.

También está el método de correlación, ya que dicho cálculo es el primer paso para determinar la relación entre las variables.

- El coeficiente de correlación posee las siguientes características:
- a. El valor del coeficiente de correlación es independiente de cualquier unidad usada para medir las variables.
- b. El valor del coeficiente de correlación se altera de forma importante ante la presencia de un valor extremo, como

sucede con la desviación típica. Ante estas situaciones conviene realizar una transformación de datos que cambia la escala de medición y modera el efecto de valores extremos (como la transformación logarítmica).

- c. El coeficiente de correlación mide solo la relación con una línea recta. Dos variables pueden tener una relación curvilínea fuerte, a pesar de que su correlación sea pequeña. Por tanto cuando analicemos las relaciones entre dos variables debemos representarlas gráficamente y posteriormente calcular el coeficiente de correlación.
- d. El coeficiente de correlación no se debe extrapolar más allá del rango de valores observado de las variables a estudio ya que la relación existente entre X e Y puede cambiar fuera de dicho rango.
- e. La correlación no implica causalidad.
  La causalidad es un juicio de valor que requiere más información que un simple valor cuantitativo de un coeficiente de correlación.

Además, el coeficiente de correlación de Pearson (r) puede calcularse en cualquier grupo de datos, sin embargo la validez del test de hipótesis sobre la correlación entre las variables requiere en sentido estricto lo siguiente:

- a) que las dos variables procedan de una muestra aleatoria de individuos.
- b) que al menos una de las variables tenga una distribución normal en la población de la cual la muestra procede.

#### REGRESION LINEAL

Analizamos que el término regresión se utilizó por primera vez en el estudio de variables antropométricas: al comparar la estatura de padres e hijos, donde resultó que los hijos cuyos padres tenían una estatura muy superior al valor medio, tendían a igualarse a éste, mientras que aquellos cuyos padres eran muy bajos tendían a reducir su diferencia respecto a la estatura media; es decir, "regresaban" al promedio.

Y también el término lineal se emplea para distinguirlo del resto de técnicas de regresión, que emplean modelos basados en cualquier clase de función matemática.

El modelo lineal relaciona la variable dependiente Y con K variables explícitas (k = 1,...K), o cualquier transformación de éstas que generen un hiperplano de parámetros desconocidos: donde es la perturbación aleatoria que recoge todos aquellos factores de la realidad no controlables u observables y que por tanto se asocian con el azar, y es la que

## confiere al modelo su carácter estocástico.

El problema de la regresión consiste en elegir unos valores determinados para los parámetros desconocidos, de modo que la ecuación quede completamente especificada. Para ello se necesita un conjunto de observaciones.

Los valores escogidos como estimadores de los parámetros, son los coeficientes de regresión sin que se pueda garantizar que coincidan con parámetros reales del proceso generador.

Los valores son por su parte estimaciones o errores de la perturbación aleatoria.