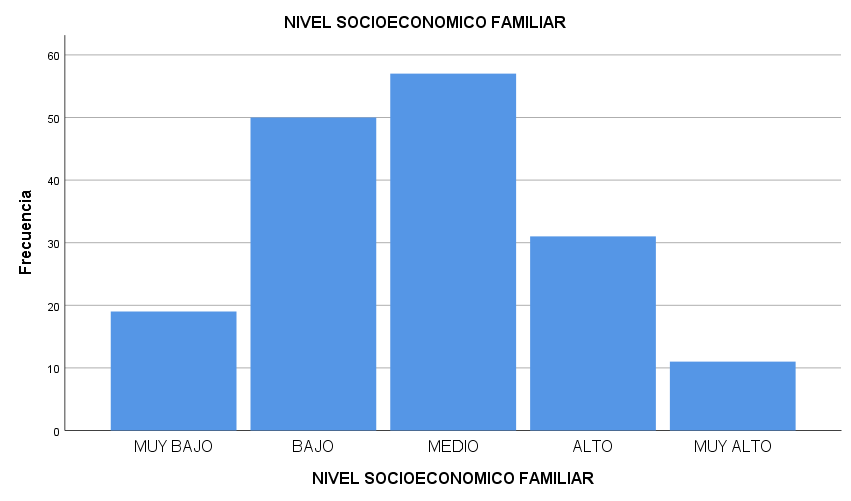
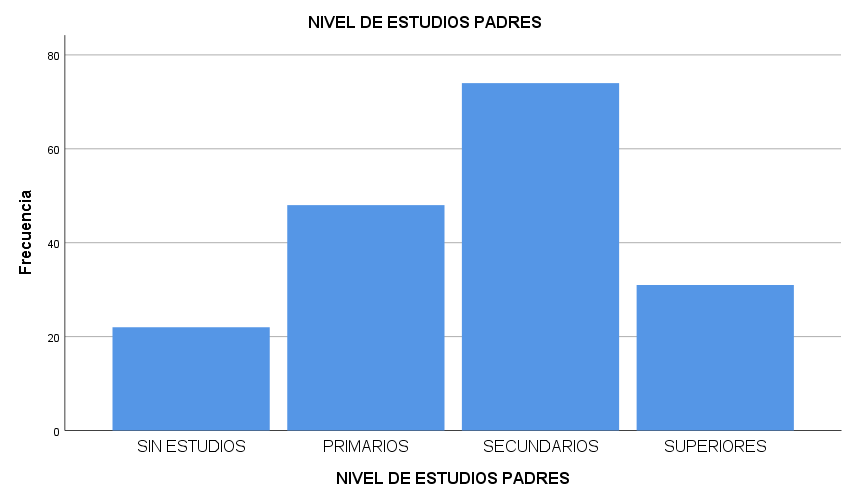
**Stefanny Arboleda 1152706891**

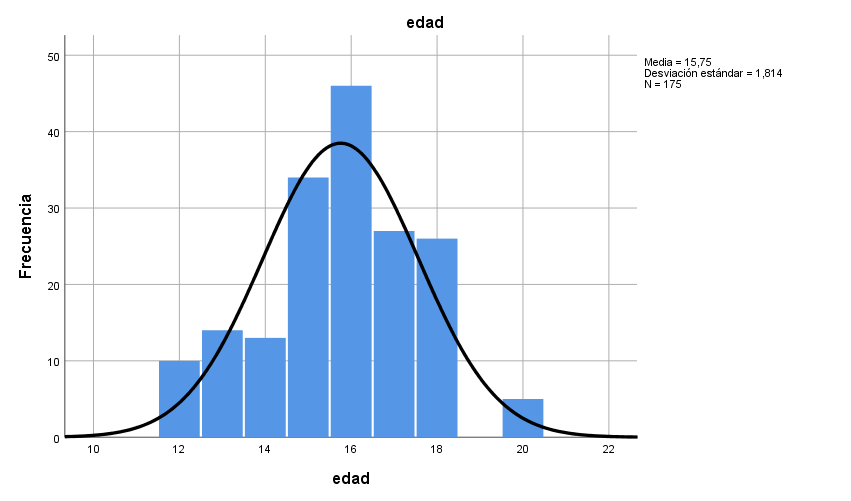
**PRÁCTICA SPSS**

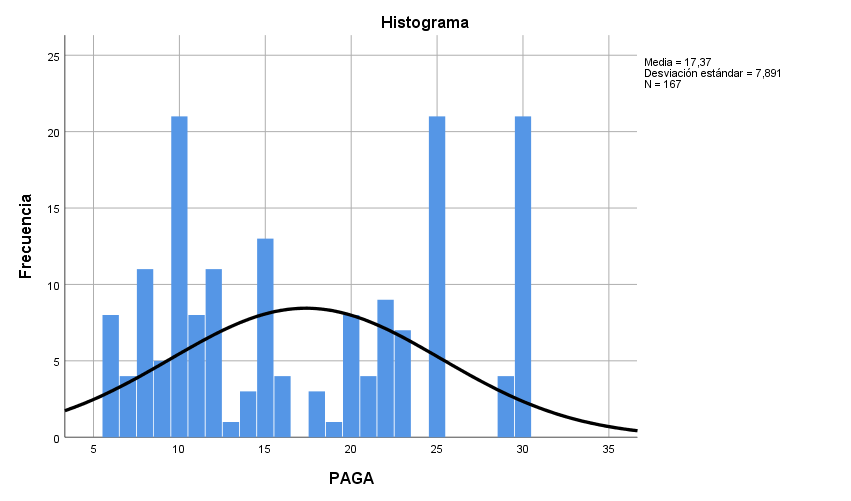
De acuerdo con el análisis individual de cada variable cualitativa, hemos encontrado que más del 50% de la población son mujeres con una frecuencia de 98; el 63% del total encuestado estudian en un centro público y el 37% restante en un centro privado. Por otro lado, el 62.3% están en un hábitat urbano; con relación al estudio de los padres, el 82,3% se encuentra acumulado hasta estudios secundarios siendo éste mismo el que presenta mayor frecuencia con 74 casos de los 175 posibles, tan sólo el 17,7% afirmaron tener estudios superiores. En general, el nivel socioeconómico es medio con un 33.9% y en este caso, encontramos 9 casos perdidos o valores faltantes en la base de datos.





Ahora bien, al realizar el análisis de cada variable cuantitativa, se encuentra que la media para la variable riqueza verbal es 30.78, para la variable asistencia anual al cine es 28.39, para la variable paga 17.37, para la variable horas semanales tv 18.63 y para edad es de 15.75. Las desviaciones estándar respectivas encontradas fueron 7.156, 16.851, 7.891, 4.77, 1.814. En estos valores no se tuvieron en cuenta los valores perdidos marcados como “99”, en el caso de la variable “PAGA” se encontró un total de 8 de estos elementos, lo que claramente al realizar este procedimiento alteraba el resultado real de la media, moda, varianza, desviación estándar, entre otras.





Dado que, la curtosis para la variable “PAGA” es negativa, podemos suponer que su curva será achatada o platicúrtica. Por el contrario, la curva de la edad es menos achatada aún cuando su curtosis también es negativa. Por otro lado, el análisis de asimetría respectivo indica que la variable “Edad” es asimétrica negativa y la variable “PAGA” asimétrica positiva.

Con relación al rendimiento escolar de acuerdo con el hábitat de la persona, encontramos que si el hábitat es rural la media de rendimiento se encuentra en 6,70 con un intervalo de confianza del 95% con valores extremos 6.34 y 7.05. La mediana para este caso es 7 y la varianza es 2.061. Por otro lado, en el hábitat urbano, la media es 6.18, el intervalo de confianza también del 95% es 5.99 a 6.37, la mediana es 6 y la varianza es 1.003.

Los percentiles de la variable “PAGA” son los siguientes

|  |  |  |
| --- | --- | --- |
| **Estadísticos** | | |
| PAGA | | |
| N | Válido | 175 |
| Perdidos | 0 |
| Percentiles | 10 | 8,00 |
| 20 | 10,00 |
| 30 | 11,00 |
| 40 | 14,00 |
| 50 | 16,00 |
| 60 | 21,60 |
| 70 | 25,00 |
| 80 | 25,00 |
| 90 | 30,00 |

De la tabla, podemos concluir que hay mayor acumulación de observaciones en la cola derecha de la curva que en la cola izquierda, esto es, se presenta una asimetría a derecha, algo que se vería respaldado con el gráfico adjunto y la afirmación que la variable “PAGA” es asimétrica positiva.

La sintaxis para la recodificación y la creación de variables es la siguiente:

-Rendimiento escolar:

STRING rende\_re (A8).

RECODE rendesco (3 thru 5='Bajo') (6 thru 7='Medio') (8 thru 9='Alto') INTO rende\_re.

VARIABLE LABELS rende\_re 'Rendimiento escolar codificada'.

EXECUTE.

-Ciudad:

RECODE CIUDAD (1=08) (2=76) (3=05) (4=08).

EXECUTE.

-OCIO:

COMPUTE OCIO=cine + teatro + conciert.

VARIABLE LABELS OCIO 'Ocio (Asistencia evento)'.

EXECUTE.

El promedio para esta variable, completa, es 41.31

|  |  |  |
| --- | --- | --- |
| **Estadísticos descriptivos** | | |
|  | N | Media |
| Ocio (Asistencia evento) | 175 | 41,3086 |
| N válido (por lista) | 175 |  |

Sin embargo, para cada variable por separado encontramos la siguiente tabla de medias. Nótese que, al sumar estos tres valores, encontramos el valor anterior.

|  |  |  |
| --- | --- | --- |
| **Estadísticos descriptivos** | | |
|  | N | Media |
| ASISTENCIA ANUAL CINE | 175 | 28,39 |
| ASISTENCIA ANUAL TEATRO | 175 | 4,37 |
| ASISTENCIA ANUAL CONCIERTOS | 175 | 8,55 |
| N válido (por lista) | 175 |  |

-Sexo y hábitat:

STRING sex\_hab (A8).

COMPUTE sex\_hab=concat(sex,habita).

VARIABLE LABELS sex\_hab 'Concatenación de sexo y hábitat'.

EXECUTE.

RECODE sex\_hab ('HU'='1') ('HR'='2') ('MU'='3') ('MR'='4').

EXECUTE.

-Diez:

Sintaxis para (Y)

IF (cine>28.39 & teatro>4.37 & conciert>8.55) diez=10.

VARIABLE LABELS diez 'Asigna un 10 a aquellos alumnos que superen las medias de asistencia al '+

'cine, teatro y conciertos.'.

EXECUTE.

En la tabla, encontramos que la cantidad de personas que cumplen con la condición (28) corresponden al 16% del total, para esto hemos usado simplemente un análisis de moda, donde claramente esta vale 10 al ser todos los datos iguales a éste.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Asigna un 10 a aquellos alumnos que superen las medias de asistencia al cine, teatro y conciertos.** | | | | | |
|  | | Frecuencia | Porcentaje | Porcentaje válido | Porcentaje acumulado |
| Válido | 10,00 | 28 | 16,0 | 100,0 | 100,0 |
| Perdidos | Sistema | 147 | 84,0 |  |  |
| Total | | 175 | 100,0 |  |  |

Sintaxis para (O)

IF (cine>28.39 | teatro>4.37 | conciert>8.55) diez2=10.

VARIABLE LABELS diez2 'Asigna un 10 a aquellos alumnos que superen en algún caso el promedio.'.

EXECUTE.

Nuevamente, en la tabla se encuentra el recuento de personas que cumplen la condición (123), en este caso, corresponde al 70.3% del total.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Asigna un 10 a aquellos alumnos que superen en algún caso el promedio.** | | | | | |
|  | | Frecuencia | Porcentaje | Porcentaje válido | Porcentaje acumulado |
| Válido | 10,00 | 123 | 70,3 | 100,0 | 100,0 |
| Perdidos | Sistema | 52 | 29,7 |  |  |
| Total | | 175 | 100,0 |  |  |

Por consiguiente, luego de ordenar la tabla de datos por medio de sexo y nivel socioeconómico hemos encontrado que el primer caso que aparece ahora al inicio de nuestra base de datos es el número 4, viéndolo desde la columna “NUMSUJ”, corresponde a un hombre de centro público, cuyos estudios los realiza en EGB ubicado en zona urbana, cuyos padres no tienen estudios y su nivel socioeconómico es muy bajo.

Por otro lado, al crear el filtro para la variable “Edad”, haciendo ésta igual a 15, obtenemos la siguiente tabla con la media para la variable “PAGA” para las personas que tienen 15 años. Vemos que, el número de personas con esta edad es 32 y reciben en promedio 15.94 de paga semanal.

|  |  |  |
| --- | --- | --- |
| **Estadísticos descriptivos** | | |
|  | N | Media |
| PAGA | 32 | 15,94 |
| N válido (por lista) | 32 |  |

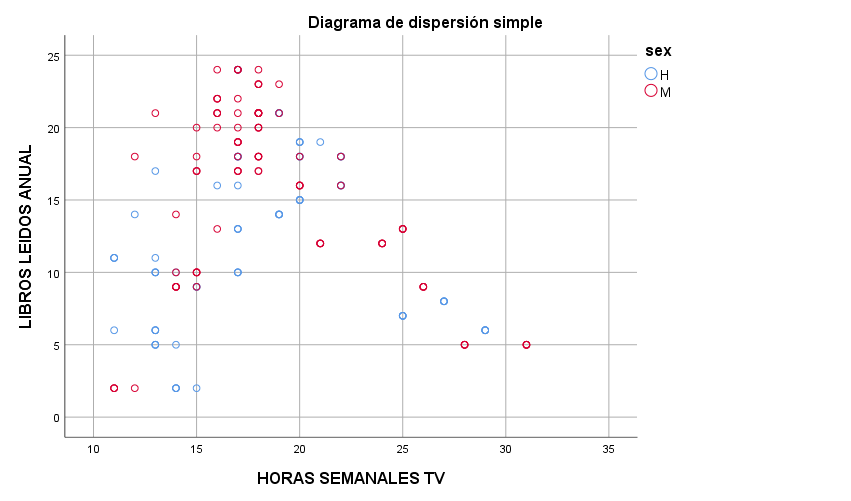
Ahora, se adjuntan las tres tablas de contingencia pedidas. De ellas, podemos concluir que al sector privado pertenecen 48 de las 98 mujeres que hay, de las cuales 23 asisten a escuelas rurales y 25 a escuelas urbanas. En el caso de los hombres, 10 asisten a escuelas rurales y 7 a escuelas urbanas. Del total general, hay igual asistencia a centro públicos como privados en la zona rural y se ve una gran desproporción en zonas urbanas, donde la asistencia a centros públicos es más del doble de la asistencia a centros privados sin importar el sexo en cuestión. Podemos notar también, que el colegio que más asistencia tiene siempre en el hábitat rural es el “BUP”, mientras que el “EGB” es el de menor. En el hábitat urbano ocurre todo lo contrario.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Hombre** | | | | |
|  | | PUBLICO | PRIVADO | Total |
| R | EGB | 1 | 1 | 2 |
| BUP | 9 | 9 | 18 |
| FP | 3 |  | 3 |
| Total | 13 | 10 | 23 |
| U | EGB | 21 | 2 | 23 |
| BUP | 11 | 4 | 15 |
| FP | 15 | 1 | 16 |
| Total | 47 | 7 | 54 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Mujer** | | | | |
|  | | PUBLICO | PRIVADO | Total |
| R | EGB | 2 | 7 | 9 |
| BUP | 11 | 9 | 20 |
| FP | 7 | 7 | 14 |
| Total | 20 | 23 | 43 |
| U | EGB | 16 | 3 | 19 |
| BUP | 8 | 10 | 18 |
| FP | 6 | 12 | 18 |
| Total | 30 | 25 | 55 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Total** | | | | |
|  | | PUBLICO | PRIVADO | Total |
| R | EGB | 3 | 8 | 11 |
| BUP | 20 | 18 | 38 |
| FP | 10 | 7 | 17 |
| Total | 33 | 33 | 66 |
| U | EGB | 37 | 5 | 42 |
| BUP | 19 | 14 | 33 |
| FP | 21 | 13 | 34 |
| Total | 77 | 32 | 109 |

De acuerdo con la gráfica, podemos concluir que las mujeres le dedican mayor tiempo a la lectura de libros anualmente, notándose en la región (20-25) mayor conglomeración de mujeres que de hombres. Así mismo, son ellas mismas quienes dedican más horas semanales a ver televisión (15-20) y tan sólo una mujer le dedica más de 30 horas a esta actividad. Podemos notar que los hombres se acumulan sobre todo en las regiones de (10-20) en ambas actividades.



En relación con el punto 1, se ha adjuntado el archivo de Excel el cual consta de tres hojas, la primera es la base de datos adquirida al realizar todo el procedimiento de la práctica, la segunda es la sintaxis realizada para etiquetar las variables y la tercera, es la sintaxis usada para darle valores a cuyas variables lo necesitaran. También, se adjunta el archivo de sintaxis con extensión de SPSS, para facilitar la revisión de la práctica completa, además del siguiente de texto en este documento conteniendo dicha sintaxis de la manera en que estaría dispuesta en el archivo SPSS.

PRESERVE.

SET DECIMAL COMMA.

GET DATA /TYPE=TXT

/FILE="C:\Users\StefannyAC\Documents\Datos\Parcial2\inicio1.dat"

/FIXCASE=1

/ARRANGEMENT=FIXED

/FIRSTCASE=1

/VARIABLES=

/1 NUMSUJ 0-2 AUTO

SEX 3-3 AUTO

CENTRO 4-4 AUTO

ESTUDIOS 5-5 AUTO

HABITAT 6-6 AUTO

ESTUPADR 7-7 AUTO

NIVSOC 8-8 AUTO

PAGA 9-10 AUTO

NUMHER 11-11 AUTO

EDAD 12-13 AUTO

RENDESCO 14-14 AUTO

VERBAL 15-16 AUTO

LIBRO 17-18 AUTO

CINE 19-20 AUTO

TEATRO 21-21 AUTO

CONCIERT 22-23 AUTO

TV 24-25 AUTO

CIUDAD 26-26 AUTO.

RESTORE.

CACHE.

EXECUTE.

DATASET NAME ConjuntoDatos1 WINDOW=FRONT.

\*Renombrar la variable V6

VARIABLE LABELS NUMSUJ 'Identificador'.

VARIABLE LABELS SEX 'Sexo'.

VARIABLE LABELS CENTRO 'Centro de estudios del alumno (público o privado)'.

VARIABLE LABELS ESTUDIOS 'Estudios del alumno'.

VARIABLE LABELS HABITAT 'Habitat rural o urbano'.

VARIABLE LABELS ESTUPADR 'Nivel de estudio de los padres'.

VARIABLE LABELS NIVSOC 'Nivel socioeconómico'.

VARIABLE LABELS PAGA 'Paga semanal'.

VARIABLE LABELS NUMHER 'Número de hermanos incluido el sujeto'.

VARIABLE LABELS EDAD 'Edad del alumno'.

VARIABLE LABELS RENDESCO 'Rendimiento escolar'.

VARIABLE LABELS VERBAL 'Riqueza verbal'.

VARIABLE LABELS LIBRO 'Libros leídos anual'.

VARIABLE LABELS CINE 'Asistencia anual al cine'.

VARIABLE LABELS TEATRO 'Asistencia anual teatro'.

VARIABLE LABELS CONCIERT 'Asistencia anual conciertos'.

VARIABLE LABELS TV 'Horas semanales TV'.

VARIABLE LABELS CIUDAD 'Ciudad de residencia'.

VALUE LABELS SEX

1 "H"

2 "M"

EXECUTE.

VALUE LABELS CENTRO

1 "Público"

2 "Privado"

EXECUTE.

VALUE LABELS ESTUDIOS

1 "EGB"

2 "BUP"

3 "FP"

EXECUTE.

VALUE LABELS HABITAT

1 "R"

2 "U"

EXECUTE.

VALUE LABELS ESTUPADR

1 "Sin estudios"

2 "Primarios"

3 "Secundarios"

4 "Superiores"

EXECUTE.

VALUE LABELS NIVSOC

1 "Muy bajo"

2 "Bajo"

3 "Medio"

4 "Alto"

5 "Muy alto"

EXECUTE.

VALUE LABELS CIUDAD

1 "Bogotá"

2 "Cali"

3 "Medellín"

4 "Barranquilla"

EXECUTE.