

Cuestión 1: Los datos siguientes representan el peso en kilos de una muestra de 80 personas de la Escuela de Ingeniería Informática:

50	73	73	68	67	74	73	67	71	979
74	74	77	74	71	80	72	74	77	75
71	73	75	76	77	71	81	68	66	73
91	75	89	77	93	57	66	83	86	90
55	77	78	91	82	83	87	96	85	88
101	97	80	73	76	80	89	76	78	99
80	85	84	72	65	69	79	84	92	83
86	76	80	81	74	73	72	79	55	66

- a) Obténgase una distribución de datos en intervalos de amplitud del 5% de la distribución, construir una tabla de frecuencias absolutas y relativas y definir cada representante de la clase.

Para hallar la amplitud del 5 % tenemos que hallar el rango que restar el máximo de la tabla menos el mínimo, luego le aplicamos el 5% que se puede hacer de dos maneras una sería multiplicando por 5 y dividiendo por 100 y la otra sería dividiéndolo entre 20 que es lo mismo y ya con esto podemos aplicar la frecuencia absoluta que es las veces que se repite los intervalos según ese intervalo luego está la frecuencia relativa que es se halla dividiendo la cantidad de la frecuencia absoluta entre la frecuencia absoluta total, y la frecuencia absoluta y relativa acumulada se halla sumando los intervalos consecutivamente el primero luego el segundo con el primero y así consecutivamente tendría que darnos 1 o 0,99 que sería lo mismo que eso significa que no nos hemos equivocado en nuestros cálculos para hallar las frecuencias

```
EJERCICIO 1:
```{r}
x=c(50,73,73,68,67,74,73,67,71,79,74,74,77,74,71,80,72,74,77,75,71,73,75,76,77,71,81,68,66,73,91,75,89,7
7,93,57,66,83,86,90,55,77,78,91,82,83,87,96,85,88,101,87,80,73,76,80,89,76,78,99,80,85,84,72,65,69,79,84
,92,83,86,76,80,81,74,73,72,79,55,66)
length(x) #Cantidad de datos
rango=max(x)-min(x) #Rango de la valor
print(rango)
amplitud=rango/100*5
print(amplitud)
histograma= hist(x,breaks=seq(min(x),max(x),by=amplitud)) #Tabla de la frecuencias
final=(table.freq(histograma))
print (final)
names(final)
names(final)=c(lower="Intervalo_min",upper="Intervalo_amplitud",Main="Media_intervalos",frequency="Frecu
encia:absoluta",Percentage="Frecuencia
relativa",CF="Frecuencia_absoluta_acumulada",CPF="Frecuencia_relatva_acumulada")
print(final)
#table(x)
#str(x)
#prop.table(x) #Frecuencia relativa de cada elemento
#plot(table(x)) # visualizar la tabla de frecuencias
```
```

| Intervalo_min<br><dbl> | Intervalo_amplitud<br><dbl> | Media_intervalos<br><dbl> | Frecuencia:absoluta<br><dbl> |
|------------------------|-----------------------------|---------------------------|------------------------------|
| 50.00                  | 52.55                       | 51.275                    | 1                            |
| 52.55                  | 55.10                       | 53.825                    | 2                            |
| 55.10                  | 57.65                       | 56.375                    | 1                            |
| 57.65                  | 60.20                       | 58.925                    | 0                            |
| 60.20                  | 62.75                       | 61.475                    | 0                            |
| 62.75                  | 65.30                       | 64.025                    | 1                            |
| 65.30                  | 67.85                       | 66.575                    | 5                            |
| 67.85                  | 70.40                       | 69.125                    | 3                            |
| 70.40                  | 72.95                       | 71.675                    | 7                            |
| 72.95                  | 75.50                       | 74.225                    | 16                           |

1-10 of 20 rows | 1-4 of 7 columns

Previous 1 2 Next

| Intervalo_min<br><dbl> | Intervalo_amplitud<br><dbl> | Media_intervalos<br><dbl> | Frecuencia:absoluta<br><dbl> |
|------------------------|-----------------------------|---------------------------|------------------------------|
| 75.50                  | 78.05                       | 76.775                    | 11                           |
| 78.05                  | 80.60                       | 79.325                    | 8                            |
| 80.60                  | 83.15                       | 81.875                    | 6                            |
| 83.15                  | 85.70                       | 84.425                    | 4                            |
| 85.70                  | 88.25                       | 86.975                    | 5                            |
| 88.25                  | 90.80                       | 89.525                    | 3                            |
| 90.80                  | 93.35                       | 92.075                    | 4                            |
| 93.35                  | 95.90                       | 94.625                    | 0                            |
| 95.90                  | 98.45                       | 97.175                    | 1                            |
| 98.45                  | 101.00                      | 99.725                    | 2                            |

11-20 of 20 rows | 1-4 of 7 columns

Previous 1 2 Next

| Frecuencia relativa<br><dbl> | Frecuencia_absoluta_acumulada<br><dbl> | Frecuencia_relativa_acumulada<br><dbl> |
|------------------------------|--|--|
| 1.2                          | 1                                      | 1.2                                    |
| 2.5                          | 3                                      | 3.8                                    |
| 1.2                          | 4                                      | 5.0                                    |
| 0.0                          | 4                                      | 5.0                                    |
| 0.0                          | 4                                      | 5.0                                    |
| 1.2                          | 5                                      | 6.2                                    |
| 6.2                          | 10                                     | 12.5                                   |
| 3.8                          | 13                                     | 16.2                                   |
| 8.8                          | 20                                     | 25.0                                   |
| 20.0                         | 36                                     | 45.0                                   |

1-10 of 20 rows | 5-7 of 7 columns

Previous 1 2 Next

| Frecuencia relativa<br><dbl> | Frecuencia_absoluta_acumulada<br><dbl> | Frecuencia_relativa_acumulada<br><dbl> |
|------------------------------|--|--|
| 13.8                         | 47                                     | 58.8                                   |
| 10.0                         | 55                                     | 68.8                                   |
| 7.5                          | 61                                     | 76.2                                   |
| 5.0                          | 65                                     | 81.2                                   |
| 6.2                          | 70                                     | 87.5                                   |
| 3.8                          | 73                                     | 91.2                                   |
| 5.0                          | 77                                     | 96.2                                   |
| 0.0                          | 77                                     | 96.2                                   |
| 1.2                          | 78                                     | 97.5                                   |
| 2.5                          | 80                                     | 100.0                                  |

11-20 of 20 rows | 5-7 of 7 columns

Previous 1 2 Next

b)Calcular la media muestral y la desviación estándar muestral.

Para calcular la media muestral tenemos que aplicar la media de los intervalos por la frecuencia absoluta,es decir  $f_i$  y  $x_i$  y todo dividirlo entre  $N$  q es la cantidad total, otra forma de hallarlo en R seria haciendo mean pero haría básicamente la media de todo.

```
[1] 77.31688
```

```

mean(x)
attach(final)
newdata<-subset(final, select=c(Media_intervalos,Frecuenciabsoluta))
newdata
media = data.frame(xifi =Media_intervalos*Frecuenciabsoluta)
MEDIA<-cbind(newdata,media)
print(MEDIA)
attach(MEDIA)
Mediafinal=sum(xifi)
Mediafinal=Mediafinal/length(x)
print(Mediafinal)

```

Para hallar la desviación estándar muestra básicamente lo que tenemos que hacer es aplicar la siguiente formula

$$s^2 = \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n - 1}$$

Y luego con las propiedades matemáticas aplicamos que :

$$s = \sqrt{s^2}$$

```

p=mean(x)
table(x)
desviacion=data.frame(Desvi=(p-x)^2)
desviacionFinal=sum(desviacion/(length(x)-1))
Desviacion_muestral=sqrt(desviacionFinal)
print(Desviacion_muestral)
sd(x)

```

[1] 9.487759  
[1] 9.487759

Como podemos observar tenemos dos maneras distintas de hacerlo en R. una seria aplicando la formula y la otra seria aplicando directamente el comando

- c) Encontrar la mediana, los cuartiles y el rango intercuartílico.

Para hallar la mediana los cuartiles y el rango intercuartilico

```

median(x)
quantile(x)
IQR(x)

```

```

> median(x)
[1] 76.5
> quantile(x)
      0%      25%      50%      75%     100%
50.00  72.75  76.50  83.00 101.00
> IQR(x)
[1] 10.25
>

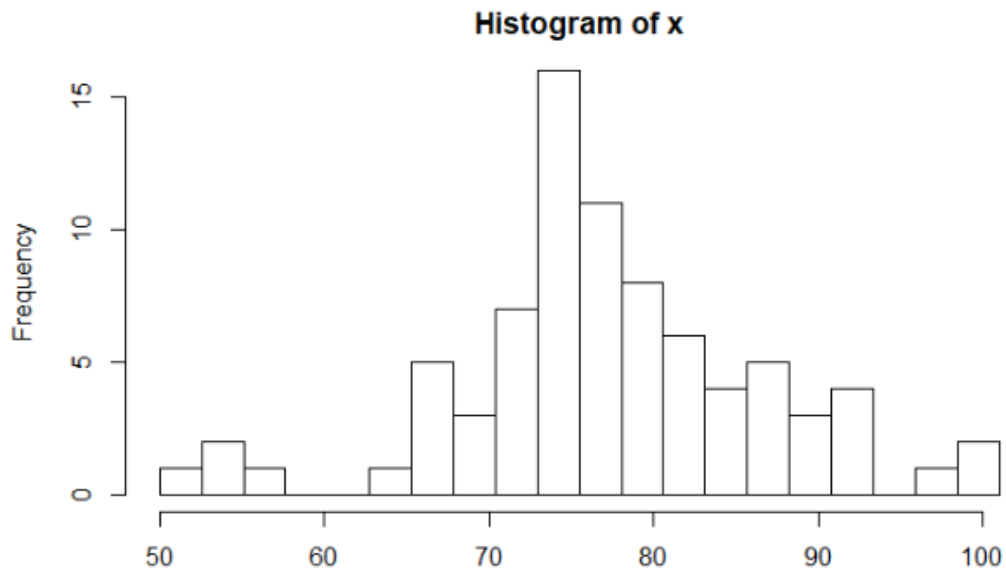
```

Para hallar los cuartiles básicamente lo que tenemos que hacer es dividirlo en 4 cajones. todas las cantidades y hay las multiplicamos por sus respectivos porcentajes Que son de 25 en 25 por que se dividen en 4 partes  $4/100=25$ . Y el rango intercuartil forma parte del final del primer intercuartil menos el inicio del tercer intercuartil , lo consecuente con restar  $83-72,75=10.25$

- d) Elaborar un histograma con los datos.

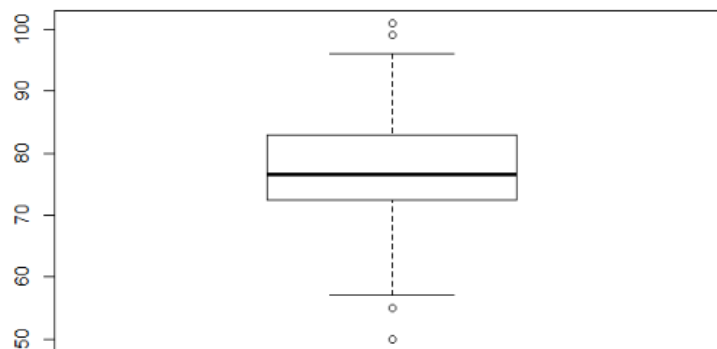
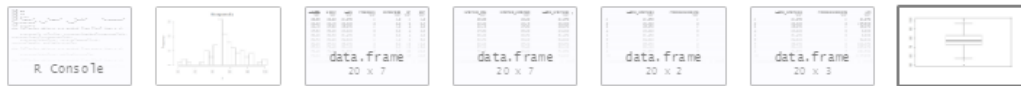
El histograma se realiza extrayendo los datos con sus respectivas repeticiones y representarlo gráficamente en R seria:

```
histograma= hist(x,breaks=seq(min(x),max(x),by=amplitud)) #Tabla de la frecuencias
```



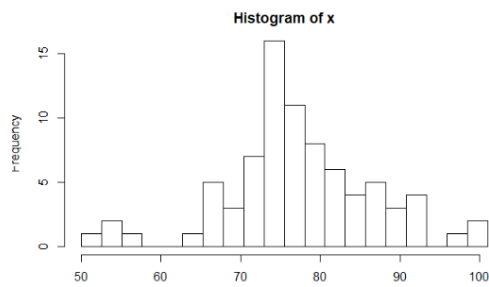
- e) Construir un diagrama de caja y mostrar los casos atípicos y otros elementos relevantes del mismo.

```
boxplot(x)
```



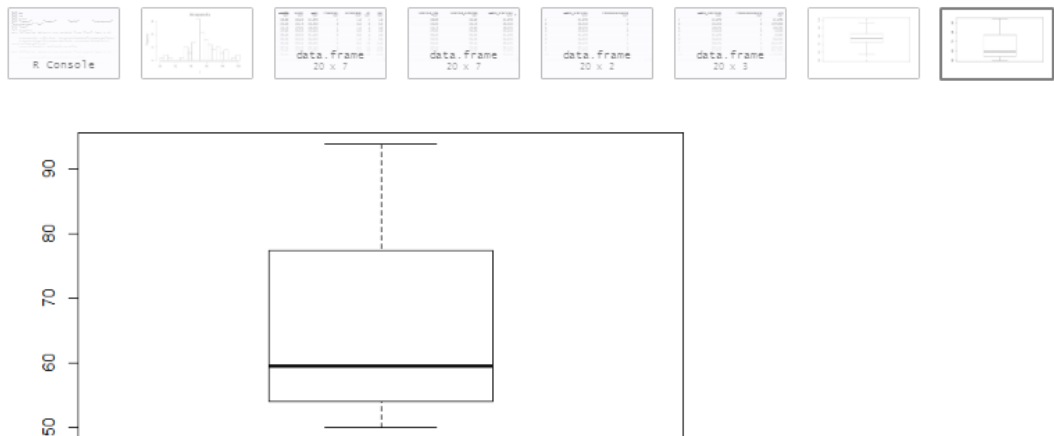
Como podemos ver nos encontramos con 4 puntos atípicos que son 99 ,101,50,55

- f) Analizar los valores fuera de rango y la posible corrección de estos “outliers”  
 Estos valores son están fuera de rango por que no tienen un valor que los represente  
 .Aqui en el histograma podemos observar como en 60 y 95 no tenemos valores .



Rellenando los outliers faltantes podemos resolver estos casos

```
library(outliers)
x_outliers=data.frame(b=c(58,61,94))
x2=rbind(x,x_outliers)
boxplot(x2)
```



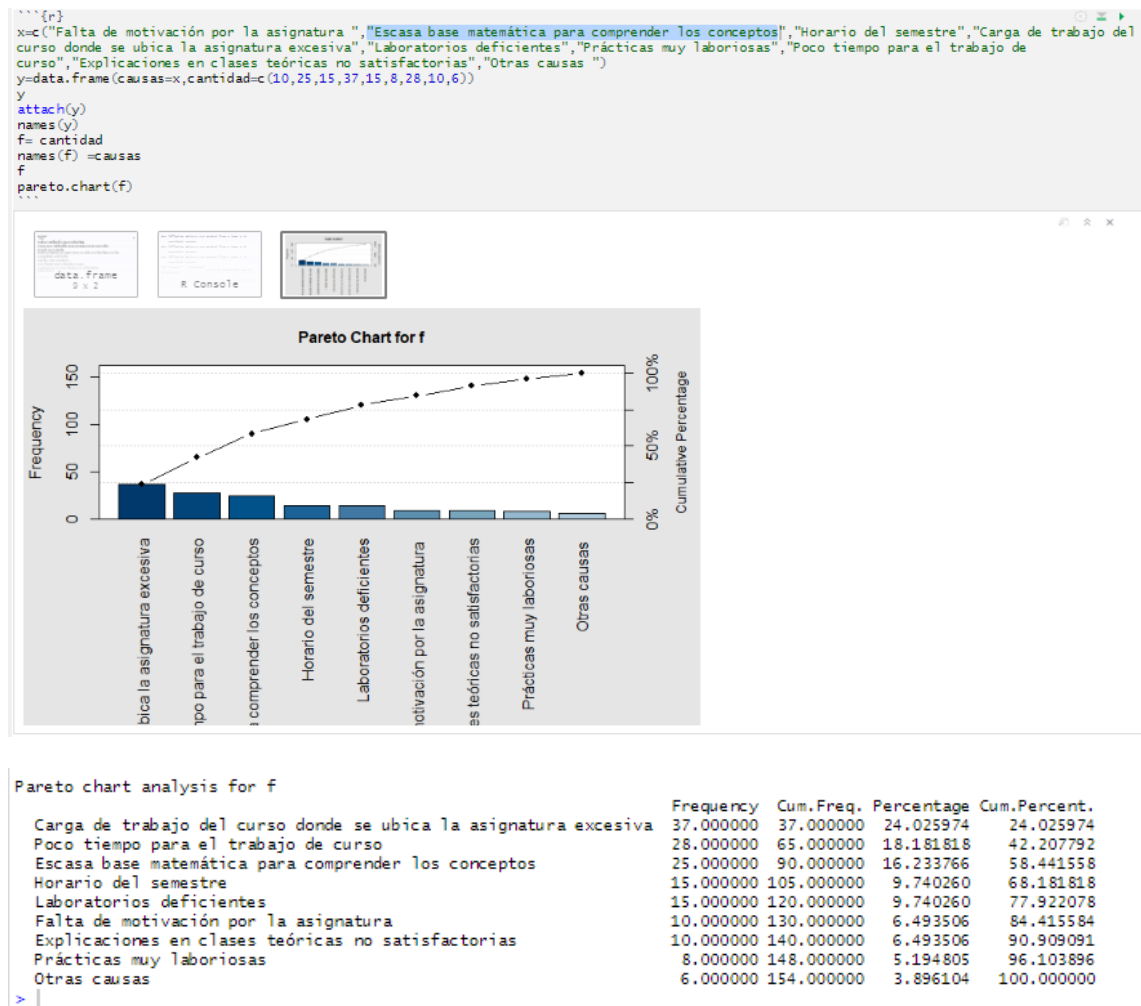
- g) Construir un programa R para verificar los distintos apartados anteriores.

Hecho

Cuestión 3. Los datos siguientes se corresponden con las causas más frecuentes de suspenso o abandono de la asignatura de Métodos Estadísticos

- Construir un diagrama de Pareto y evaluar los porcentajes de causas que se pueden explicar por categorías.
- Analizar gráficamente el problema y establecer conclusiones y recomendaciones para el/la profesor/a, el departamento y el centro.
- Construir un programa R que permita realizar los distintos apartados anteriores.

a)



Como podemos ver las mayores causas de suspenso según los porcentajes son la carga de trabajo que ocupan un 24 % luego el poco tiempo para el trabajo que ocupa un 18% y por ultima la escasa base matematica que es un 16 % .Nos enfocamos en estos tres causas por que son las que principalmente están generando el suspenso en la asignatura y en lo que nos debemos priorizar e intentar resolverlo. Se ordenan de izquierda a derecha de la frecuencia mas baja a la mas alta

b)

Como podemos observar el principal problema es el excesivo carga de trabajo, para dominar o entender una asignatura no hay que mandar muchísimos ejercicios para practicarlos o mandar muchos deberes eso solo motiva al desapego del alumno por la asignatura y la frustración de no verse capaz de llevarla al día ,además la causa de que hay mas asignaturas además de ME, para poder mejorar esto lo que realizaría yo sería que en los laboratorios los alumnos se vieran más implicados con los temas y fueran temas mas dinámicos con un buen guion estructurado de las cosas que se van a dar no un power point de unas 100 o 200 o 300 páginas , eso no hay por donde cogerlo. Como profesor a la hora de hacer un power point los power points deben ser cortos y no tener demasiadas hojas ya que al alumno ya le estas provocando un rechazo a la asignatura, Imagínese la situación de un alumno que tenga que cada vez que quiera ver un tema de la asignatura se encuentre un power con esta densidad de páginas a la mínima el alumno se hartara y lo dejara para otro día si es que vuelve claro.

Por otra parte también ,Los alumnos se ven superado por la cantidad de ejercicios que se mandan, es más recomendable que tengan un par de ejercicios que hacer pero luego tengan ejercicios opcionales ya solucionados pero que no tengan que hacerlos que los utilicen como referencia para tener una ayuda a la hora de hacer los ejercicios a entregar.

En cuanto al poco tiempo de trabajo para el curso el problema no es el tiempo ni mucho menos el problema es que quieren enseñar mucho o dar muchas cosas en un tiempo reducido lo que no se puede hacer es exigir dar mucho del alumno y que el tiempo y las clases no estén acordes con ello. Por mucho que el alumno tenga un tiempo en casa para poder repasar e ir practicando hay que llevar a cabo mas asignaturas no se puede llevar todo y eso produce que el alumno cada vez se vea más alejado de la asignatura o que cada vez la vea mas difícil, porque se intenta dar mucho temario y que al final cuando se lleva a la práctica no es útil hay muchas cosas que se dan que no son nada útiles a la hora de imaginarse el caso de trabajar en una empresa y que tengas que usar el R te encontraras en la situación de que casi siempre tienes unas cosas esenciales para manejar el programa e intentar dar lo mejor de él, pero una gran cantidad de información inútil o formulas o métodos inútiles que con el tiempo la persona se olvida del temario que se dio y de cómo se aplicaba. Es muchísimo mejor enseñar lo básico y lo más útil a la horade trabajar que el alumno tenga muchísimos conocimiento de formulas de algebra

"Escasa base matemática para comprender los conceptos" este tema no solo está provocado por esta asignatura sino por todas las asignaturas anteriores y cursos anteriores y sistema de educación actual, en la universidad por mi experiencia se premia mas a la persona que es capaz de almacenar una cierta capacidad de información aunque luego la olvide o no sepa llevarlo a la práctica o explicárselo a otra persona, una frase que yo suelo usar es "Nos premian por ser Monos" en el sentido de que no se califica al alumno que tiene un conocimiento de la asignatura o comprende y luego a la hora de manejarlo por falta de práctica se ve incapaz pero tiene unos conocimientos que si deberían premiar, no se debería premiar a las personas que son capaces de aprenderse un temario en un cierto tiempo pero luego se olvidan completamente de todo o se aprende cosas de memoria pero luego no entienden nada de lo que han hecho y aprueban la asignatura . Siguiendo con el tema el sistema de educación actual no está bien hay muchas deficiencias

matemáticas o alumnos que llegan a la universidad si en verdad tener un buen nivel de conocimientos matemáticos que en mi opinión ya va dependiendo de la carrera que se vaya a cursos deberían haber asignaturas en el propio curso que sean muy practicas que no tengan calificación que te lleven a tener ese nivel mínimo para poder seguir en la universidad. Por culpa de quien ocurre esto, principalmente de que los profesores de años anteriores no dan lo que va a ser útil para los años siguientes o cursos siguientes o solo dan una pincelada de los primero pero luego la mayoría de la asignatura es desconocida por el alumno, también influye es que se requieren unos conocimientos que el alumno no tiene o esta adquiriendo pero no tiene la capacidad para a la hora de llevar a cabo las clases tener un buen entendimiento de ellas .

Por otra parte luego algunos profesores de laboratorios por lo que me cuentan mis compañeros no los están enseñando bien o les están enseñando por así decirlo “el lado malo de la asignatura “ enseñándola como si fuera un verdadero aburrimiento que nunca acaba, en cuanto a las lecciones teóricas ocurre el caso de que el profesor explica rápido la asignatura y explica mucho y de nada sirve tener ejemplos en los powers hay que practicarlos pero de manera moderada hacer 30 mil veces lo mismo o de manera muy seguida acaba sofocando o aburriendo a más de uno, también por otro lado muchos alumnos que ya se han desapegado de la asignatura pos vienen a clases pero se pasan el rato hablando de otras asignatura u otros temas y molestan al demás alumnado al comprendimiento de la asignatura .Es preferible dar poco y darlo bien y hacer varios ejercicios en la pizarra que dar mucho temario y que salgan a la pizarra , debido a que el alumno a los 20 minutos de pasar la asignatura ya se empieza a cansar de estar todo el rato igual .