Q20

In [8]:
```
import pandas as pd
import matplotlib.pyplot as plt
import scipy.stats as stats
```

In [7]:
```
df1=pd.read_csv("Cars.csv")
df1.head()
```

Out[7]:

	HP	MPG	VOL	SP	WT
0	49	53.700681	89	104.185353	28.762059
1	55	50.013401	92	105.461264	30.466833
2	55	50.013401	92	105.461264	30.193597
3	70	45.696322	92	113.461264	30.632114
4	53	50.504232	92	104.461264	29.889149

In [25]:
```
df1.shape
```

Out[25]:
```
(81, 5)
```

In [14]:
```
df1.describe()
```

Out[14]:

	HP	MPG	VOL	SP	WT
count	81.000000	81.000000	81.000000	81.000000	81.000000
mean	117.469136	34.422076	98.765432	121.540272	32.412577
std	57.113502	9.131445	22.301497	14.181432	7.492813
min	49.000000	12.101263	50.000000	99.564907	15.712859
25%	84.000000	27.856252	89.000000	113.829145	29.591768
50%	100.000000	35.152727	101.000000	118.208698	32.734518
75%	140.000000	39.531633	113.000000	126.404312	37.392524
max	322.000000	53.700681	160.000000	169.598513	52.997752

In [15]:
```
df1.dtypes
```

Out[15]:
```
HP       int64
MPG    float64
VOL      int64
SP     float64
WT     float64
dtype: object
```

In [16]:
```
df1.isnull().sum()
```

Out[16]:
```
HP     0
MPG    0
VOL    0
SP     0
WT     0
dtype: int64
```

- a. P(MPG>38)

In [16]:
```
round(1-stats.norm.cdf(x = 38, loc = 34.42, scale = 9.13),2)#34.42 mean, 9.13 sta
```

Out[16]:
```
0.35
```

- b. P(MPG<40)

In [17]:
```
round(stats.norm.cdf(x = 40, loc = 34.42, scale = 9.13),2)
```

Out[17]:
```
0.73
```

- c. P (20<MPG<50)

In [22]:
```
round(stats.norm.cdf(x = 50, loc = 34.42, scale = 9.13)-stats.norm.cdf(x = 20, lo
```

Out[22]:
```
0.9
```

Q21

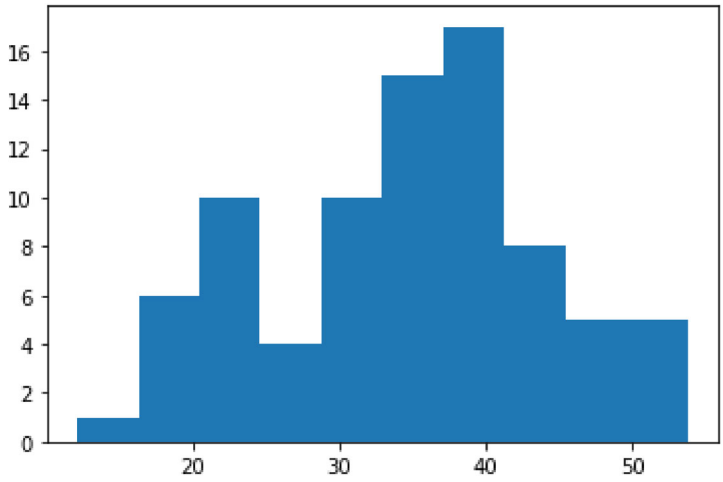
In [8]:
```
df.HP.median()
```

Out[8]:
```
100.0
```

*a) Check whether the MPG of Cars follows Normal Distribution

In [40]:
```
plt.hist(df1.MPG)
#Yes it is a normally distributed graph
```

Out[40]:
```
(array([ 1.,  6., 10.,  4., 10., 15., 17.,  8.,  5.,  5.]),
 array([12.10126289, 16.26120474, 20.42114659, 24.58108844, 28.74103029,
        32.90097213, 37.06091398, 41.22085583, 45.38079768, 49.54073953,
        53.70068138]),
 <BarContainer object of 10 artists>)
```



*b) Check Whether the Adipose Tissue (AT) and Waist Circumference(Waist) from wc-at data set follows Normal Distribution

In [17]:
```
df2=pd.read_csv("wc-at.csv")
df2.head()
```

Out[17]:

	Waist	AT
0	74.75	25.72
1	72.60	25.89
2	81.80	42.60
3	83.95	42.80
4	74.65	29.84

In [24]:
```
df2.shape
```

Out[24]:
```
(109, 2)
```

In [43]:
```
df2.describe()
```

Out[43]:

	Waist	AT
count	109.000000	109.000000
mean	91.901835	101.894037
std	13.559116	57.294763
min	63.500000	11.440000
25%	80.000000	50.880000
50%	90.800000	96.540000
75%	104.000000	137.000000
max	121.000000	253.000000

In [44]:
```
df2.isnull().sum()
```

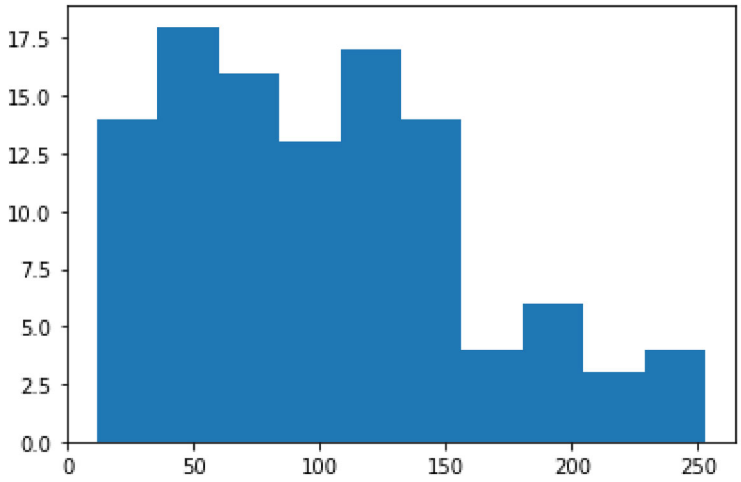
Out[44]:
```
Waist    0
AT       0
dtype: int64
```

In [45]:
```
df2.dtypes
```

Out[45]:
```
Waist    float64
AT       float64
dtype: object
```

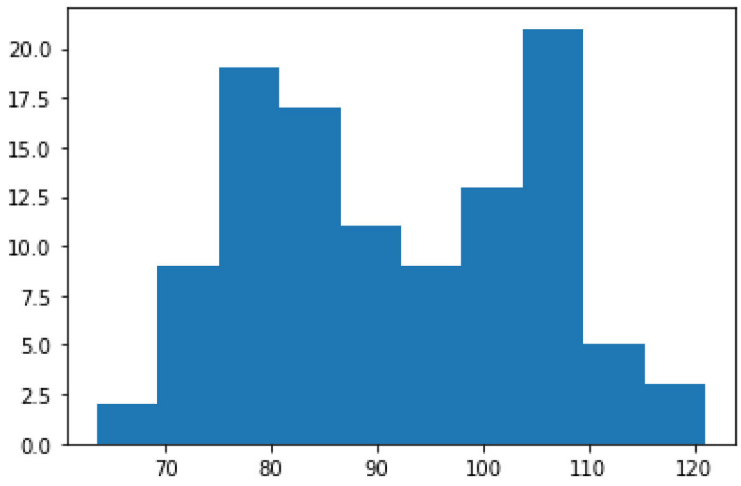
In [42]:
```
plt.hist(df2["AT"])
#No it is not a normally distributed graph as its right skewed
```

Out[42]:
```
(array([14., 18., 16., 13., 17., 14.,  4.,  6.,  3.,  4.]),
 array([ 11.44 ,  35.596,  59.752,  83.908, 108.064, 132.22 , 156.376,
        180.532, 204.688, 228.844, 253.   ]),
 <BarContainer object of 10 artists>)
```



In [50]:
```
plt.hist(df2["Waist"])
#No it is not a normally distributed graph as its decreasing from center
```

Out[50]:
```
(array([ 2.,  9., 19., 17., 11.,  9., 13., 21.,  5.,  3.]),
 array([ 63.5 ,  69.25,  75.  ,  80.75,  86.5 ,  92.25,  98.  , 103.75,
        109.5 , 115.25, 121.  ]),
 <BarContainer object of 10 artists>)
```



Q22

In [63]:
```
round(stats.norm.ppf(0.95),3) # for 90%
```

Out[63]:
```
1.645
```

In [60]:
```
round(stats.norm.ppf(0.97),3) # for 94%
```

Out[60]: 1.881

In [5]:
```
round(stats.norm.ppf(0.80),3) # for 60%
```

Out[5]: 0.842

Q23

In [28]:
```
# df= n-1, 25-1=24
round(stats.t.ppf(0.975,24),3) # for 95%
```

Out[28]: 2.064

In [29]:
```
round(stats.t.ppf(0.98,24),3) # for 96%
```

Out[29]: 2.172

In [30]:
```
round(stats.t.ppf(0.995,24),3) # for 99%
```

Out[30]: 2.797