In [1]:
```
import pandas as pd
import numpy as np
from matplotlib import pyplot as plt
%matplotlib inline
import seaborn as sns
from scipy import stats

import warnings
warnings.filterwarnings('ignore')
```

Set-1_Descriptive statistics and Probability

In [2]:
```
measure =pd.DataFrame ([24.23,25.53,25.41,24.14,29.62,28.25,25.81,24.39,40.26,32.
```

In [3]:
```
measure.mean()
```

Out[3]:
```
0    33.271333
dtype: float64
```

Type *Markdown* and LaTeX: α^2

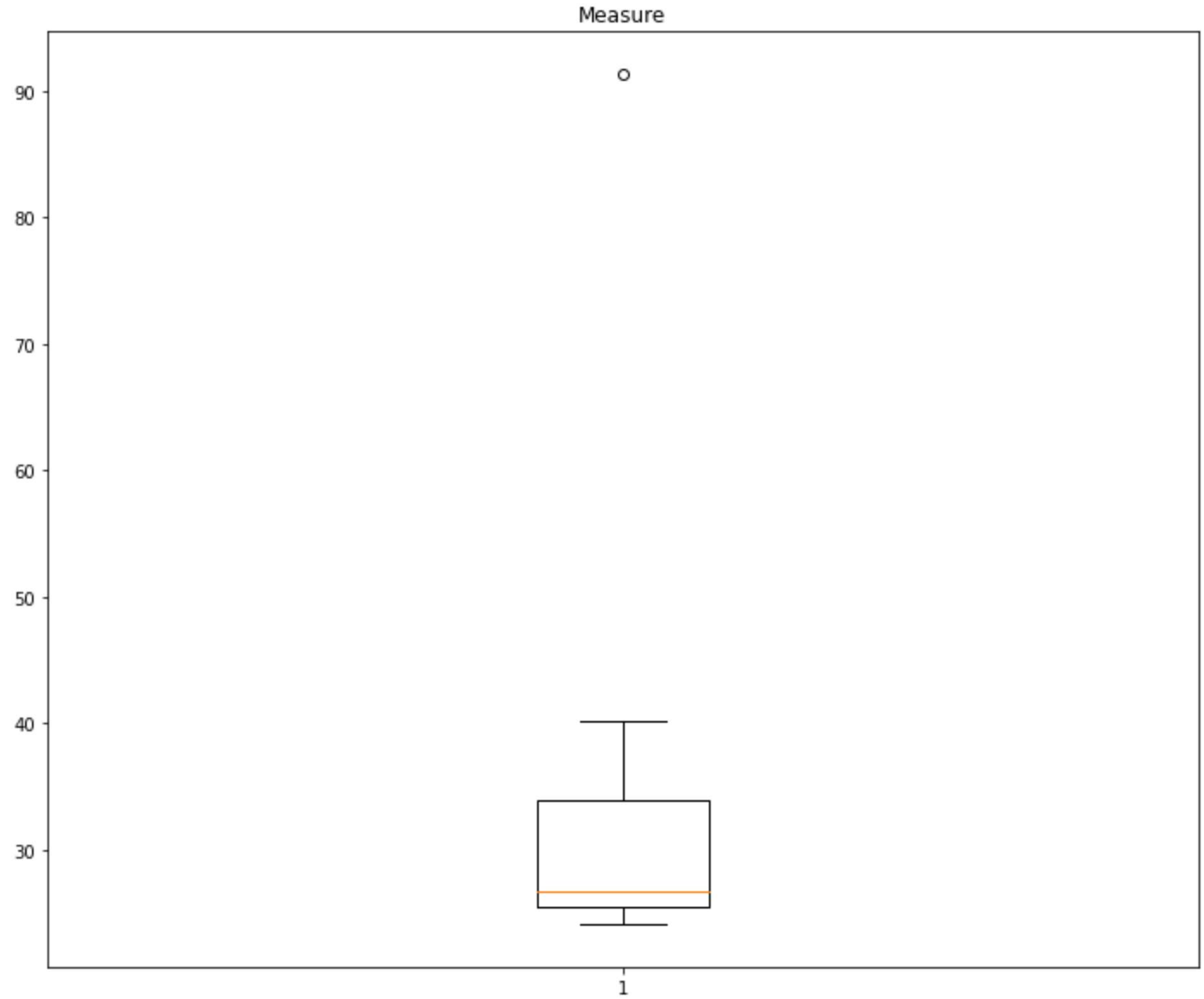
In [4]:
```
measure.var()
```

Out[4]:
```
0    287.146612
dtype: float64
```

In [5]:
```
measure.std()
```

Out[5]:
```
0    16.945401
dtype: float64
```

In [6]:
```
plt.figure(figsize=(12,10))
plt.boxplot(measure)
plt.title('Measure')
plt.show()
```



Set-2 Normal distribution, Functions of Random Variables

Q1 Answer

In [7]:
```
z_score = (50-45)/8
z_score
```

Out[7]: 0.625

In [8]:
```
# What is the probability that the service manager cannot meet his commitment
p_value = 1 - stats.norm.cdf(0.625)
```

In [9]:
```
p_value
```

Out[9]: 0.26598552904870054

Q2 Answer

a)

In [10]:
```
##P(>44)
1-stats.norm.cdf(44,loc=38,scale=6)
```

Out[10]: 0.15865525393145707

In [11]:
```
## P(38<x<44)
stats.norm.cdf(44,loc=38,scale=6)-stats.norm.cdf(38,loc=38,scale=6)
```

Out[11]: 0.3413447460685429

True,More employees at the processing center are older than 44 than between 38 and 44

b)

In [12]:
```
##P(<30)
stats.norm.cdf(30,loc = 38,scale=6)
```

Out[12]: 0.09121121972586788

In [13]:
```
# The number of employees under the age of 30
400*0.0912
```

Out[13]:
```
36.480000000000004
```

There for the statement is also true

Q4 Answer

In [14]:
```
stats.norm.interval(.99,loc = 100, scale = 20)
```

Out[14]:
```
(48.48341392902199, 151.516586070978)
```

Q5 Answer

a)

In [15]:
```
## mean profits of two diiferent divisions of company in rupees
print('mean profit =',(7+5)*45)
```

```
mean profit = 540
```

In [16]:
```
## standered deviation
print('standered deviation =',np.sqrt(9+16)*45)
```

```
standered deviation = 225.0
```

In [17]:
```
rge = stats.norm.interval(.95,540.225)
print('Range is ',rge)
```

```
Range is  (538.2650360154599, 542.1849639845401)
```

b)

In [18]:
```
## we compute using this formula X=μ+Zσ
## in from z table, 5 percentile = -1.645
x = 540+(-1.645*225)
print('5th pecentile is',x)
```

```
5th pecentile is 169.875
```

c)

In [19]:
```
## probability of 1 making loss
round(stats.norm.cdf(0,5,3),4)
```

Out[19]: 0.0478

In [20]:
```
## probability of 2 making loss
round(stats.norm.cdf(0,7,4),4)
```

Out[20]: 0.0401

Set-3 Confidence Intervals

Q5 Answer

In [22]:
```
## Applyn One-tail z-test
z_score = (0.046-0.05)/(np.sqrt(0.05*(1-0.05)/2000))
z_score
```

Out[22]: -0.820782681668124

In [23]:
```
p_val = 1- stats.norm.cdf(abs(z_score))
```

In [25]:
```
round(p_val,4)
```

Out[25]: 0.2059

Q8 Answer

In [26]:
```
sample_size = ((1.96**2)*0.5*0.5/(0.04**2))
sample_size
```

Out[26]: 600.2499999999999

Q9

In [27]:
```
sample =  ((2.326**2)*0.5*0.5/(0.04**2))
sample
```

Out[27]: 845.355625

Set 4- Sampling Distributions and Central Limit Theorem

Q3 Answer

In [28]:
```
## probability for no investigation p(45<x<55)
## there for probability for investigation 1-p(45<x<55)
## find z-scores at x = 45 and x = 55
z_45 = (45-50)/(40/100**0.5)
z_45
```

Out[28]: -1.25

In [29]:
```
z_55 = z = (55-50)/(40/100**0.5)
z_55
```

Out[29]: 1.25

In [30]:
```
## p(x<55)-p(x<45)
stats.norm.cdf(1.25)-stats.norm.cdf(-1.25)
```

Out[30]: 0.7887004526662893

In [31]:
```
##1-p(45<x<55)
1-0.7887
```

Out[31]: 0.21130000000000004

In [ ]: