Ishika Prasad
ip1262@rit.edu (mailto:ip1262@rit.edu)

Import Libraries

In [31]:
```
import pandas as pd
from pandas import ExcelWriter
from pandas import ExcelFile
from matplotlib import pyplot as plt
import seaborn as sns
```

Read Excel File

In [32]:
```
df = pd.read_excel('mammal.xlsx', sheet_name='Sheet1')
```

Print the data from excel file

In [33]:
```
print(df)
```

```
    Unnamed: 0      BodyWgt     BrainWgt  NonD  Dream      Sleep
Span  \
0            1  6654.000000  5712.000000   NaN    NaN   3.300000   38.59
9998
1            2     1.000000     6.600000   6.3    2.0   8.300000    4.50
0000
2            3     3.385000    44.500000   NaN    NaN  12.500000   14.00
0000
3            4     0.920000     5.700000   NaN    NaN  16.500000
NaN
4            5  2547.000000  4603.000000   2.1    1.8   3.900000   69.00
0000
5            6    10.550000   179.500000   9.1    0.7   9.800000   27.00
0000
6            7     0.023000     0.300000  15.8    3.9  19.700001   19.00
0000
7            8   160.000000   169.000000   5.2    1.0   6.200000   30.40
0000
8            9     3.300000    25.600000  10.9    3.6  14.500000   28.00
0000
```

Descriptive statistics for Body Weight

In [34]:
```
print('Sum of Body weight =', df['BodyWgt'].sum())
print('Mean of Body weight =', df['BodyWgt'].mean())
print('Median of Body weight =', df['BodyWgt'].median())
print('Mode of Body weight =', df['BodyWgt'].mode())
print('Standard deviation of Body weight =', df['BodyWgt'].std())
```

```
Sum of Body weight = 12324.97900829371
Mean of Body weight = 198.78998400473728
Median of Body weight = 3.34249997138977
Mode of Body weight = 0    0.023
1    3.500
dtype: float64
Standard deviation of Body weight = 899.1580105092452
```

Descriptive statistics for Brain Weight

In [35]:
```
print('Sum of Brain weight =', df['BrainWgt'].sum())
print('Mean of Brain weight =', df['BrainWgt'].mean())
print('Median of Brain weight =', df['BrainWgt'].median())
print('Mode of Brain weight =', df['BrainWgt'].mode())
print('Standard deviation of Brain weight =', df['BrainWgt'].std())
```

```
Sum of Brain weight = 17554.320000723004
Mean of Brain weight = 283.13419356004846
Median of Brain weight = 17.25
Mode of Brain weight = 0      1.0
1     12.3
2    115.0
dtype: float64
Standard deviation of Brain weight = 930.2789422476623
```

Descriptive statistics for Non-Dreaming

In [36]:
```
print('Sum of Non dreaming =', df['NonD'].sum())
print('Mean of Non dreaming =', df['NonD'].mean())
print('Median of Non dreaming =', df['NonD'].median())
print('Mode of Non dreaming =', df['NonD'].mode())
print('Standard deviation of Non dreaming =', df['NonD'].std())
```

```
Sum of Non dreaming = 416.2999989986422
Mean of Non dreaming = 8.672916645805044
Median of Non dreaming = 8.349999904632565
Mode of Non dreaming = 0    11.0
dtype: float64
Standard deviation of Non dreaming = 3.666451659452768
```

Descriptive statistics for Dream

In [37]:
```
print('Sum of Dream =', df['Dream'].sum())
print('Mean of Dream =', df['Dream'].mean())
print('Median of Dream =', df['Dream'].median())
print('Mode of Dream =', df['Dream'].mode())
print('Standard deviation of Dream =', df['Dream'].std())
```

```
Sum of Dream = 98.59999936819075
Mean of Dream = 1.971999987363815
Median of Dream = 1.79999995231628
Mode of Dream = 0    0.5
1    0.9
2    1.8
3    2.0
dtype: float64
Standard deviation of Dream = 1.4426506045813576
```

Descriptive statistics for Sleep

In [38]:
```
print('Sum of Sleep =', df['Sleep'].sum())
print('Mean of Sleep =', df['Sleep'].mean())
print('Median of Sleep =', df['Sleep'].median())
print('Mode of Sleep =', df['Sleep'].mode())
print('Standard devation of Sleep =', df['Sleep'].std())
```

```
Sum of Sleep = 610.8999993801118
Mean of Sleep = 10.532758610001926
Median of Sleep = 10.4500002861023
Mode of Sleep = 0     8.4
1    10.3
2    12.5
dtype: float64
Standard devation of Sleep = 4.606760394482433
```

Descriptive statistics for Span

In [39]:
```
print('Sum of Span =', df['Span'].sum())
print('Mean of Span =', df['Span'].mean())
print('Median of Span =', df['Span'].median())
print('Mode of Span =', df['Span'].mode())
print('Standard deviation of Span =', df['Span'].std())
```

```
Sum of Span = 1152.8999972343445
Mean of Span = 19.877586159212836
Median of Span = 15.10000038146975
Mode of Span = 0    7.0
dtype: float64
Standard deviation of Span = 18.20625532827711
```

Descriptive statistics for Gestation

In [40]:
```
print('Sum of Gestation =', df['Gest'].sum())
print('Mean of Gestation =', df['Gest'].mean())
print('Median of Gestation =', df['Gest'].median())
print('Mode of Gestation =', df['Gest'].mode())
print('Standard deviation of Gestation =', df['Gest'].std())
```

```
Sum of Gestation = 8256.5
Mean of Gestation = 142.35344827586206
Median of Gestation = 79.0
Mode of Gestation = 0    42.0
1    63.0
dtype: float64
Standard deviation of Gestation = 146.8050386528531
```

Descriptive statistics for Predation

In [41]:
```
print('Sum of Predation =', df['Pred'].sum())
print('Mean of Predation =', df['Pred'].mean())
print('Median of Predation =', df['Pred'].median())
print('Mode of Predation =', df['Pred'].mode())
print('Standard deviation of Predation =', df['Pred'].std())
```

```
Sum of Predation = 178
Mean of Predation = 2.870967741935484
Median of Predation = 3.0
Mode of Predation = 0    2
dtype: int64
Standard deviation of Predation = 1.4764142535625584
```

Descriptive statistics for Exposure

In [42]:
```
print('Sum of Exposure =', df['Exp'].sum())
print('Mean of Exposure =', df['Exp'].mean())
print('Median of Exposure =', df['Exp'].median())
print('Mode of Exposure =', df['Exp'].mode())
print('Standard deviation of Exposure =', df['Exp'].std())
```

```
Sum of Exposure = 150
Mean of Exposure = 2.4193548387096775
Median of Exposure = 2.0
Mode of Exposure = 0    1
dtype: int64
Standard deviation of Exposure = 1.6047918724845893
```

Descriptive statistics for Danger

In [43]:
```
print('Sum of Danger =', df['Danger'].sum())
print('Mean of Danger =', df['Danger'].mean())
print('Median of Danger =', df['Danger'].median())
print('Mode of Danger =', df['Danger'].mode())
print('Stndard deviation of Danger =', df['Danger'].std())
```

```
Sum of Danger = 162
Mean of Danger = 2.6129032258064515
Median of Danger = 2.0
Mode of Danger = 0    1
dtype: int64
Stndard deviation of Danger = 1.4412521731282144
```

Effect of Missing Data

It can have an adverse effect in analyzing standard deviation based on the attributes present in the data set which means that how the parametrs of each attribute varies can be unknown.

Handling the Missing Data

Missing data can be handled by Average imputation. The missing value can be filled using the result of average value from the other responses of the same attribute value.