Multiple Linear Regression

In this application, brain weight needs to be predicted for the given input features. In the dataset, Brain weight (grams) for 237 adults is classified by gender and age group and head size. The features are Gender, Age Range, Head size, Brain weight. For Gender feature, value 1 is assigned for male and value 2 is assigned for Female. Similarly for Age feature, value 1 is assigned for a person having an age in the range of 20-46 and value 2 is assigned for a person whose age happens to be greater than 46. Head size is in cm3 . Brain weight is in grams. Create the MultipleLinearRegression Model to predict the Brain weight using gender, age and head size as input features

Importing Libraries

```
In [2]:  import matplotlib.pyplot as plt
         import pandas as pd
         import pylab as pl
         import numpy as np
         from sklearn import linear_model
         import seaborn as sns
         from sklearn.model_selection import train_test_split
         from sklearn.metrics import mean_squared_error
```

Load the Data- brainhead.csv load into brainhead dataframe

Type *Markdown* and LaTeX: α^2

Print Head, describe and info of data

```
In [3]:  # head the data
         df = pd.read_csv("brainhead.csv")
         df.head()
```

Out[3]:

	Gender	Age_range	Head_size	Brain_weight
0	1	1	4512	1530
1	1	1	3738	1297
2	1	1	4261	1335
3	1	1	3777	1282
4	1	1	4177	1590

```
In [2]:  # info of data
```

```
In [1]:  # describe
```

Relationship between Features and Response

```
In [4]:   # Generate pairplot between 'Gender','Age_range','Head_size' as X and Brain_weight as Y
          g = sns.pairplot(df,x_vars=["Gender",'Age_range','Head_size'],y_vars='Brain_weight')
          g.fig.set_size_inches(25,10)
          plt.show()
```
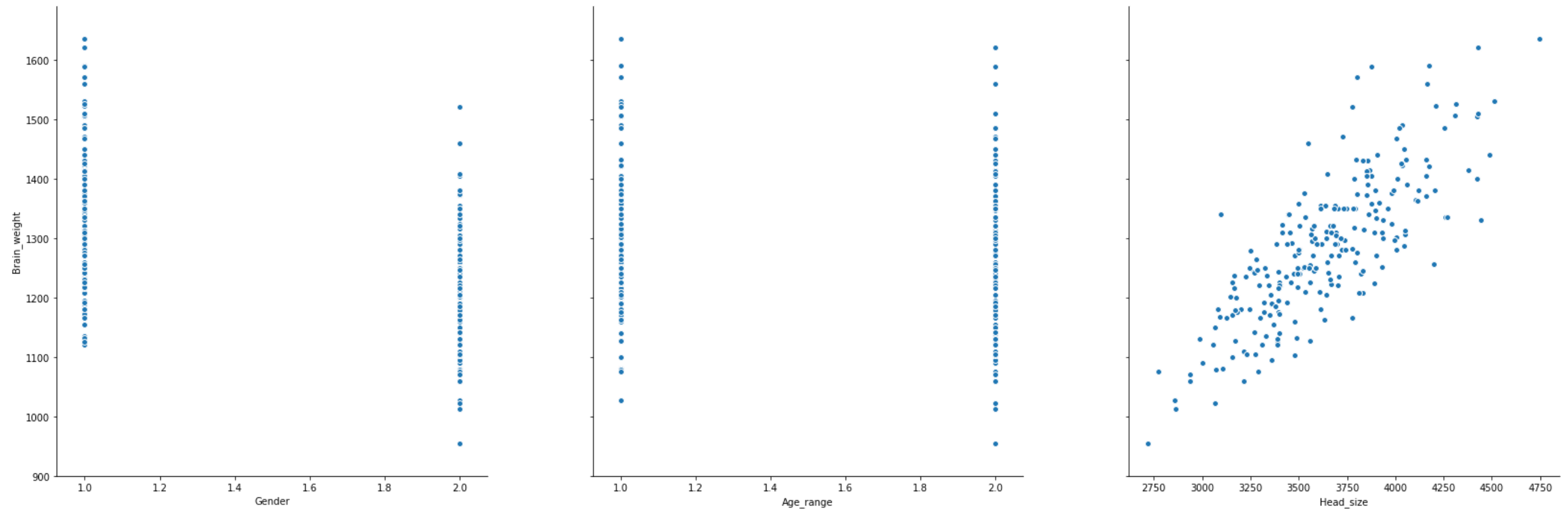

Define X as 'Gender','Age_range','Head_size' and Brain_weight as y.

Drop column of Brain_weight into X

Assign Brain_weight column to y

```
In [5]:   X= df.drop(["Brain_weight"], axis = 1)
          y=df["Brain_weight"]
```

Multiple Linear Regression - Estimating Coefficients

In [7]:
```python
#Create model and fit by defining X as input features and Y as output feature


from sklearn.linear_model import LinearRegression
X_train,X_test,y_train,y_test=train_test_split(X,y,random_state=1)
lm=LinearRegression().fit(X_train,y_train)
print(lm.intercept_)
print(lm.coef_)
# print coefficient and intercepts
```

```
402.0936515629046
[-22.38120018 -11.36913222   0.25614559]
```

Predict the brain_weight from test dataset

In [8]:
```python
y_pred = lm.predict(X_test)
```

Calculate RMSE

In [10]:
```python
print(np.sqrt(mean_squared_error(y_test,y_pred)))
```

```
74.34684798264809
```