**Analysis of the key determinants of birth weight using PYTHON 3.9.14**

**Course: Data science and machine learning for the biosciences**

**Chidimma Gift Omenoba-Nee Ubah (229148943), University of Bath**

**Introduction**

It is well-established that birth weight is affected by the child’s genetic endowments and on multiple factors related to the mother, particularly, maternal smoking during pregnancy (Cardenas et al. 2019; Pereira et al. 2022). The birthweight\_win dataset I worked with was provided in the SWBio DTP R-studio statistics class as an example dataset (see attached csv file). The motivation and purpose of this script is to explore further with the data and understand the specific determinants and/or key aspects of birthweight looking at maternal information available.

**Methods**

The input file contain data of up to 17 variables (including categorical and continuous). A python script was created using Anaconda, JupyterLab Notebook.

All the statistical analysis was performed using Python 3.9.14. modules used were pandas and sklearn. Seaborn and matplotlib were used for graphical visualization seaborn.

The script performed the following functions:

* Import the csv dataset from my directory
* Identified variables of interest in the .CSV file (e.g, Birthweight, Gestation, catsmoke, smoker etc)
* Variables of no interest was collapsed and those of interest were assigned to lists for each patient identity mother (patient here connotes individual mother and child)
* Loop over cells to check for strings that correlate to the variable “Birthweight”(e.g. Maternal smoking status was seen to negatively correlate to birthweight whereas, Gestation age showed a positive correlate to birthweight).

**Results**

The script performed the functions as intended. For the attached result input file (annotated codes only), output results (annotated codes and their outputs) can be seen in GITHUB LINK:

<https://github.com/CHIDIMMA-G-OMENOBA/SWBio-DTP-Machine-learning-assesment-task.git>

**Conclusion**

* There exist a clear association between gestational age and birthweight
* Babies born to mothers who are non-smokers have normal birthweight compared to others
* Maternal pre pregnancy weight also has a notable influence on Birthweight.

N/B: Though all of the codes and output are clearly explained, yet users are suggested to read the data overview attached for better understanding of the dataset.

**References**

Cardenas, A; Lutz, S.M; Everson, T.M; Perron, P; Bouchard, L; Hivert, M.F (2019). Mediation by placental DNA methylation of the association of prenatal maternal smoking and birth weight. *American journal of epidemiology*, 188: 1878-1886

Pereira, R.D; Rietveld, C.A; van Kippersluis, H (2022). The Interplay between maternal smoking and genes in offspring birth weight. *Journal of Human Resources*, 1: 1020-11266