**MY REPORT ON ORIGINAL DATASET. (ID 2483538 Micheal Okon)**

***Extracted my dataset from***: [Search Results - data.gov.uk](https://www.data.gov.uk/search)

Publisher: Aberdeen City Council

Topic: Business & Economy

Format: CSV

Selection and importing of a suitable dataset 5% Effective and appropriate using tidyverse. There are four fundamental processes that help me turn my data into understanding, knowledge, and insight. The four are: data manipulation, data visualization, statistical analysis/modeling, and deployment of results. The one that is focused on this assignment is data manipulation. The two purposes to data manipulation are it makes data suitable for processing and reveals insight to readers. My main objectives is to packages utilized for cleaning, processing, and manipulating data are tidyr and dplyr. The function of gather reshapes a wide format of data and puts into a long format. The separate function splits a single variable into two. The function of unite is the opposite of separate, in which it merges two variables into one. Lastly, the spread function reshapes long format to wide format. To remember this function, it is the opposite of gather.

Within the dplyr operations, there are seven fundamental functions of data transformation. The first function, select, reduces the data frame size to only the variable you want for the current situation. To find the data you are looking for with the function “select”, you can do special functions like starts with, ends with, contains, and matches to find the data you want to use. The filter functions reduces the numbers of rows and observations to your liking. Within this function you can use special functions, The function that I find the most useful is the summaries function that can provide analysis of the data.

The function of computers is to store, transform, visualize, and model their data. The term data tidying is what the author uses to describe the best way to organize your data for R.

Dplyr can work with data sets and data tables without leaving R.

Within Dplyr, when you use the function of tbl\_df it prints nicely and formats the data in a display that is appealing to the eye.

Dplyr is a way of cleaning up your data in R Studio. The chaining command is imported from the magrittr package and is used to replace nesting commands.