**\*Appendix**

**.** Data cleaning and transformation:

Observations on Data Cleaning and Transformation

Column Reduction: The transformed dataset have 10 columns even the old the raw dataset had some 21 columns. I have removed unnecessary columns such as DGUID, UOM, UOM\_ID, SCALAR\_FACTOR, SCALAR\_ID, VECTOR, COORDINATE, STATUS, SYMBOL, TERMINATED, and DECIMALS from the data because this columns were containing the information which I thought would not be useful for data analysis.

Renaming: For better clarity the original column named value was split into Median employment income and number of graduates. Because the column contained two different numerical values in just one column so I separated it by using filter in excel

Missing Data Handling: missing values in the columns such as Median employment income and number of graduates got cleaned up from the dataset.

Certain field of study labels were simplified to improve the identification:

From the dataset, I renamed Science and science technology into just **"**Science Tech**"**

And to make it more simplify the Engineering and engineering technology into just "Engineering"

Columns Used for Charts:

Median Employment Income:

This column basically talks about basic economics that help users analyze the economic data as a result of different fields of study.

Number of Graduates:

This column helped in understanding how graduates are diversified across the fields based on the interest need and their passions.

Field of Study:

To get the idea of the results between the different academic fields such as Science Tech , Engineering against non stem degrees such from BHASE discipline such as huminity,business,arts,social science

Educational Qualification:

It compares the graduate performance across the levels of educational qualifications such as certificate and diploma level with Ph.D. and Master degrees.

REF\_DATE:

This column indicates how the year's progression was in each of 2018,2019,2020.

Reasons for Chart Selections:

1.Bar Chart (Maximum Median Employment Income):

A graph of a graph of a number of squares

Description automatically generated with medium confidence

income levels of the biggest one to the years spread across the different fields of study.

Purpose: It recognizes knowledge categories that help to understand the top most income after graduation reward for students.

2.Bar Chart (Sum of Number of Graduates):

A graph of graduation

Description automatically generated with medium confidence

Grouping of numbers of grad by their degree programs to get the idea of educational qualification regarding graduate statistics.

Purpose: It shows the popular degree programs by time trends for the year for 2018, 2019, and 2020 respectively.

3.Pie Chart (Sum of Median Employment Income):

A pie chart with different colored sections

Description automatically generated

all incomes within a particular year from the point of view of different educational qualification.

Purpose: This data points gives the overview about the various sum of income based on different education program.

4.line chart of average median income by different education qualifications:

A graph of a person with a green orange and blue line

Description automatically generated with medium confidence

purpose: This line chart explains and helps to get trend Analys on employment income changes for diverse academic subjects over three years so different fields can be easily assessed for growth or decline.