Member Task

**Task-2:**

Section in the report describing the enrichment data and datatype - variable dictionary:

**Employment Dataset:**

The employment dataset provides the level of employment and the earning potential by Geographical impact.

**Datatype-Variable dictionary:**

|  |  |  |
| --- | --- | --- |
| **Name** | **DataType** | **Description** |
| Area\nCode | object | Five-digit unique code assigned for each area |
| St Name | object | Name of the state |
| Year | int64 | Specific year |
| Area | object | Name of the area |
| Employment wages Details | int64 | Wages for the three months |
| Total Quarterly Wages | int64 | Total quarterly wages |
| Average Weekly Wage | int64 | Average of the weekly wages |

**How can you merge the data with the primary COVID-19 dataset. Identify the individual variable which map between the datasets:**

#### To merge data with the primary COVID-19 dataset, we need to identify the variables that exist in both datasets and use them as the key for the merge. Here are the steps we can follow:

1. Identify common variables: Look for variables that exist in both datasets and have the same or similar names. For example, if both datasets have a column such as "Area" or "State" or "St Name" or "County Name\_x", you can use that as the key to merge the data.
2. Check the data type and format: Make sure that the data type and format of the common variables match in both datasets. For example, if one dataset has the "Area" variable and the other has "County Name\_x", we need to convert one of them to match the other. Stripping covid data County Name\_x column to avoid white spaces at the end. Replacing Area Column of Employee Data To have only County Name. And convert St Name into State codes according to the abbreviation.
3. Perform the merge: Once we have identified the common variables, checked the data type and format, and decided on the type of merge, we can perform the merge using a software tool or programming language of our choice. For example, in Pandas, we can use the merge function to perform a merge.

Note: The Reason we have chosen employee data set to show the super Covid-19 data set to show the impact of covid\_19 with respect to the employee enrichment data within each "Area" and "State".

**Describe how your enrichment data can help in the analysis of COVID-19 spread. Pose initial hypothesis questions:**

An employment dataset can help in the analysis of COVID-19 spread by providing information on the industries and professions that have been most affected by the pandemic, as well as the economic impact of COVID-19 on different regions and demographic groups.

Some initial hypothesis questions that could be posed include:

1. Geographical impact: How has the economic impact of COVID-19 varied by Area & State , and what factors may be contributing to these differences?