Business Case

This business case is mainly intended to test technical skills of data processing and visualization with tools of **Python and Power BI**, and **problem-solving skills**.

Please download and install below tools for free to complete the task:

Python: [Python Release Python 3.10.5 | Python.org](https://www.python.org/downloads/release/python-3105/)

Microsoft Power BI Desktop: <https://aka.ms/pbidesktopstore>

The data set used for this business case is *Gun violence database of US from 2021 to date* (open data source). If anything in the dataset is not clear, please feel free to make assumptions as you document it.

Please use these data to:

* Extract csv files in “raw\_data” folder and transfer them into dataframe in Python.

Create one py file to complete all following operations: importing csv files, data cleansing/merging/reshaping/joining/creating new column if needed, and export the output as csv file.

* In the same py file, use your output csv file from step 1, show the trend of children/teens proportion being injured or killed in mass shooting, and the trend in gun accidental causes, please provide the summary chart by month or weeks, export the summary and save the py file.
* Use the final dataset created in step 1 as the data source of Power BI, to build a simple star schema model. Star schema is a typical dimensional model. The final schema should include at least 3 tables, save the pbix file.
* One fact table of incidents,
* One geographic dimension table (e.g. states),
* One date dimension table (e.g. year, quarter, month, week, date, weekday name, etc.).
* Further use the pbix file in step 3, create a dashboard with the data to show your Power BI skills. The topic can be
* Construct a dashboard with parameter control, YoY, YTD, Week summaries, KPIs **OR**
* Exposing some interesting insights or issues based on the data, and use visualization to illustrate the insights (hot spot, trend, root cause etc.)

Once you complete test, please submit below items, send it to concern person

* Python py file and output files from the tasks of 1 and 2
* Pbix file from tasks 3 and 4.
* Supplementary document if needed (e.g. data process assumptions, instructions of your program/visualization, etc.)