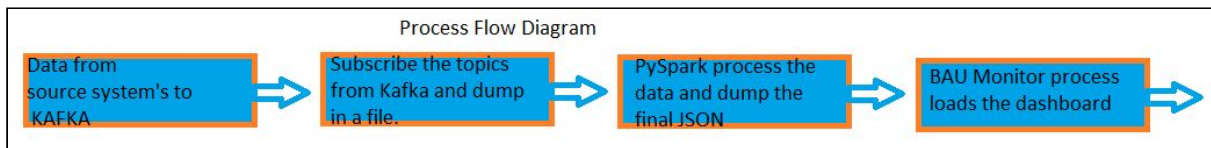


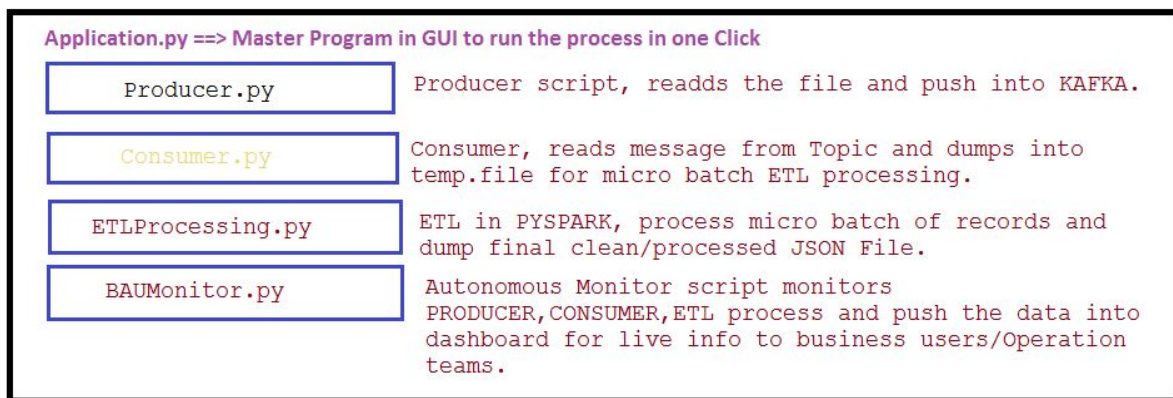
Data Streaming Application in kafka and Spark Processing

This is a standalone application for stream data processing from source, process the data in pyspark and load the data as JSON in the local directory..

The data pipeline is given below,

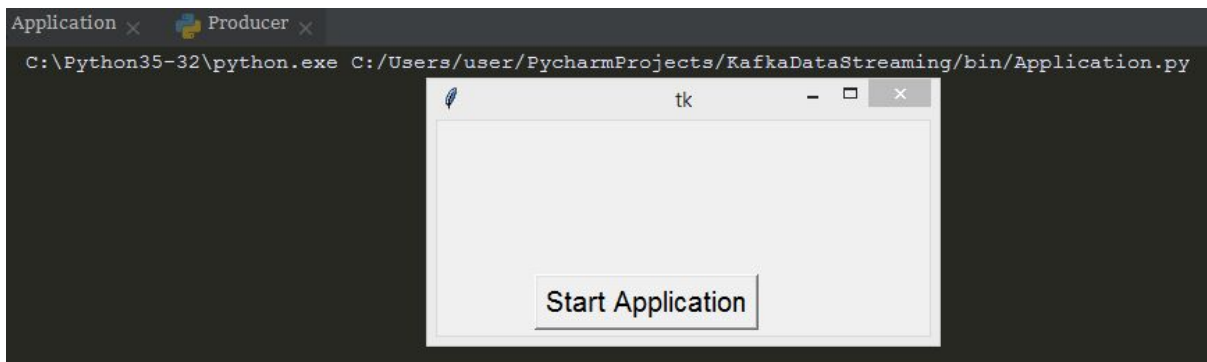


Script Flow Diagram:



Execution Steps:

User just need to execute the Applciation.py a GUI opens and should Start application as below,

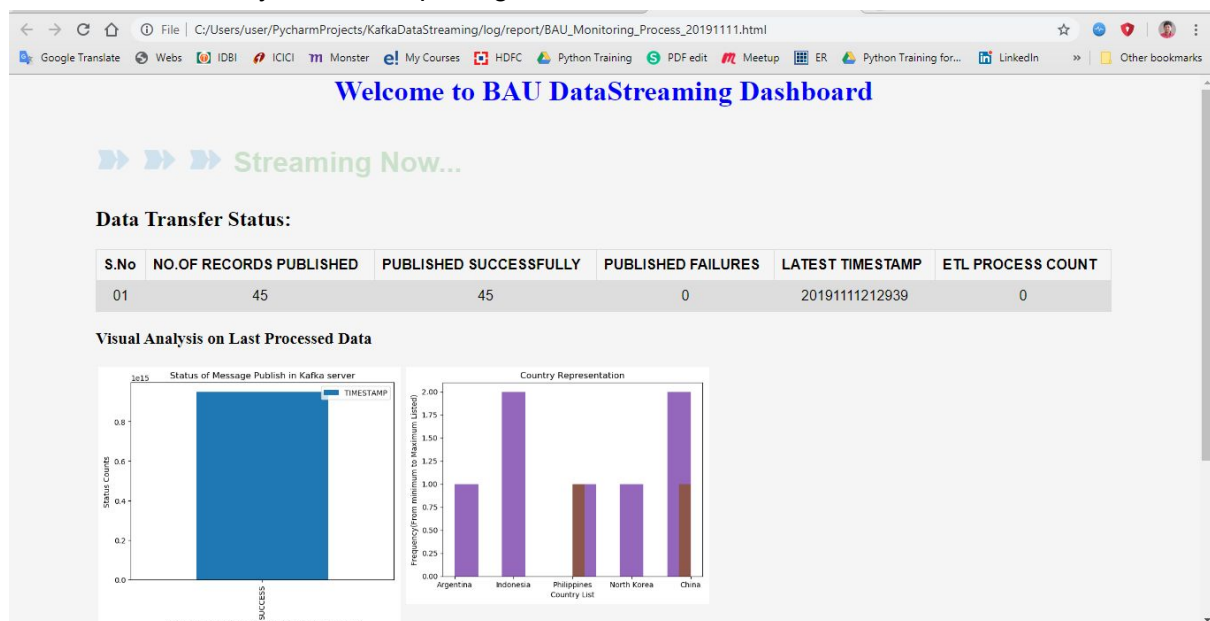


Upon clicking the Start application, process starts as,

```
Run: Application x Producer x
Data Not Streaming...
..\log\report\BAU_Monitoring_Process_20191111.html
..... ['Argentina', 'Indonesia', 'Indonesia', 'Philippines', 'North Korea', 'China', 'China']
..... ['Philippines', 'China']
15 0 20191111210545
{'first_name': 'Barthel', 'country': 'France', 'id': 1, 'last_name': 'Kittel', 'gender': 'Male', 'ip_address': '130.187.82.195', 'date': '06/
..\data\
Data Streaming...
..\log\report\BAU_Monitoring_Process_20191111.html
```

Dashboard View:

Since this is just a temporary project parameters needs to be tuned fined for excellent performance and better GUI/UI/UX performance. For now,the Dashboard(HTML) should be manually refreshed over browser due to the limitation of my personal laptop configurations, haven't linked to any realtime reporting.



SNAPSHOT 2:



The code is designed so robust and versatile in extending to other applications/further integrations with new softwares/framework.Almost the code is kind of plug-play!