Week 2 Azure Case Study Assessment

**Graded Assignment**

**Source:**

NYC Open Data is a vast trove of City government datasets that have been made available to the public. One such dataset, 311 Service Requests from 2010 to Present Is handled and kept in the following website https://opendata.cityofnewyork.us/

**Data Set Information:**

These 311 data contain information about more than 24 million service requests made since 2010. For those who aren’t familiar, 311 is a phone number used in the U.S. that allows callers to access non-emergency municipal services, report problems to government agencies, and request information.

This dataset explores trends in a recent subset of the data (June-Nov. 2020] The government agency that responds to a given call. The NYPD is responsible for responding to most 311 calls, and some agencies respond to very few.

NYC 311's mission is to provide the public with quick and easy access to all New York City government services and information while offering the best customer service. Each day, NYC311 receives thousands of requests related to several hundred types of non-emergency services, including noise complaints, plumbing issues, and illegally parked cars. These requests are received by NYC311 and forwarded to the relevant agencies such as the police, buildings, or transportation. The agency responds to the request, addresses it, and then closes it.

**Attribute Information:**

|  |
| --- |
| 1. Unique Key |
| 1. Created Date |
| 1. Closed Date |
| 1. Agency |
| 1. Agency Name |
| 1. Complaint Type |
| 1. Descriptor |
| 1. Location Type |
| 1. Incident Zip |
| 1. Incident Address |
| 1. Street Name |
| 1. Cross Street 1 |
| 1. Cross Street 2 |
| 1. Intersection Street 1 |
| 1. Intersection Street 2 |
| 1. Address Type |
| 1. City |
| 1. Landmark |
| 1. Facility Type |
| 1. Status |
| 1. Due Date |
| 1. Resolution description |
| 1. Resolution Action Updated Date |
| 1. Community Board |
| 1. Borough |
| 1. X Coordinate (State Plane) |
| 1. Y Coordinate (State Plane) |
| 1. Park Facility Name |
| 1. Park Borough |
| 1. School Name |
| 1. School Number |
| 1. School Region |
| 1. School Code |
| 1. School Phone Number |
| 1. School Address |
| 1. School City |
| 1. School State |
| 1. School Zip |
| 1. School Not Found |
| 1. Latitude |
| 1. Longitude |
| 1. Location |

**Task 1(Azure) – (50 marks)**

**Azure Task:**

**Task 1: 20 marks**

1. **Visualization on Azure Databricks** 
   1. Move the dataset onto Azure Blob. **(05-marks)**
   2. Transfer the data to Azure Data bricks and Perform PySpark operations to drive core reports out of the data **(15 marks)**
2. **Task 2: 30 marks**

Move the dataset to Azure Machine Learning and train the model -

* 1. Create notebook instance and perform analysis on the data, identifying correct algorithm **(10 marks)**
  2. Predicting model output – **20 marks**

1. **Submission guidelines:**
2. Take a screenshot of every step clearly in a word document, Mention your name and login ID at the top of the document. Capture step by Step Screenshot for each of the above task, convert Word doc as PDF and upload it on LMS. Note Azure instances should have your name defined.