DataEng S24: PubSub

A. [MUST] PubSub Tutorial

1. Get your cloud.google.com account up and running
    a. Redeem your GCP coupon
    b. Login to your GCP console
    c. Create a new, separate VM instance
2. Complete this PubSub tutorial: link Note that the tutorial instructs you to destroy your PubSub topic, but you should not destroy your topic just yet. Destroy the topic after you finish the following parts of this in-class assignment.

B. [MUST] Create Sample Data

1. Get data from https://busdata.cs.pdx.edu/api/getBreadCrumbs for two Vehicle IDs from among those that have been assigned to you for the class project.
2. Save this data in a sample file (named bcsample.json)
3. Update the publisher python program that you created in the PubSub tutorial to read and parse your bcsample.json file and send its contents, one record at a time, to the my-topic PubSub topic that you created for the tutorial.
4. Use your receiver python program (from the tutorial) to consume your records.

C. [MUST] PubSub Monitoring

1. Review the PubSub Monitoring tutorial: link and work through the steps listed there. You might need to rerun your publisher and receiver programs multiple times to trigger enough activity to monitor your my-topic effectively.

D. [MUST] PubSub Storage

1. What happens if you run your receiver multiple times while only running the publisher once?

    Answer: You will only receive one set of messages

2. Before the consumer runs, where might the data go, where might it be stored?

Answer: There must be some sort of storage service or bucket within the pub/sub service that holds the messages within the console themselves before it's received by the subscribers

3. Is there a way to determine how much data PubSub is storing for your topic? Do the PubSub monitoring tools help with this?

   Answer: Looks like you can use "Retained acked Messages" to see how many Messages are retained via the subscription

4. Create a "topic_clean.py" receiver program that reads and discards all records for a given topic. This type of program can be very useful for debugging your project code.