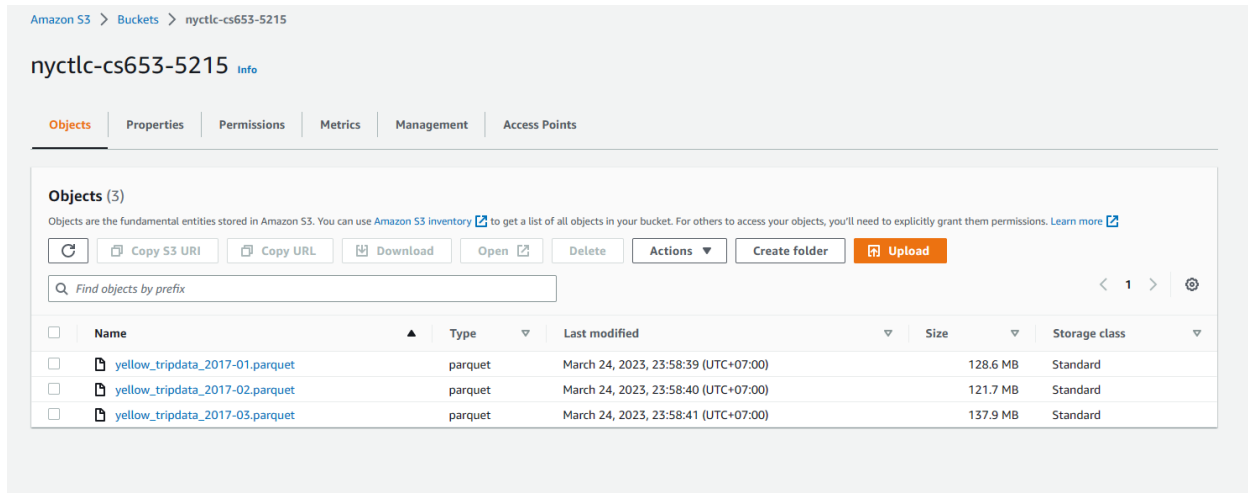


ธีรัช คุ่มไข่มุก 6509035215

1.การทำงานมีสองส่วนคือ 1. Create bucket 2.Copy จาก Bucket to another bucket



a) เนื่องจากเป็น Parquet จึง group by โดยตรงไม่ได้ จึงทำการหาPayment ซ้ำกันมากกว่า

b) หาโดยการwhere ID ใน Location ก่อนแล้วใช้คำสั่ง SQL หาค่า sumกับ avg ค่าใช้จ่ายต่อคนตามลำดับ

c) เมื่อรู้ type จากข้อ1 แล้ว ก็ทำการแปลงเป็นเก็บเป็น list แล้ววนค่าเดือน 1 ถึงเดือน 5 ตามลำดับ

เราได้ความรู้และทักษะอะไรจากการทำการบ้านครั้งนี้บ้าง และคิดว่านำไปใช้ประโยชน์อย่างไรได้บ้าง

ตอบ การดึงหรือ Copy ข้อมูลแบบอัตโนมัติ ต่อไปก็เขียน Code ให้ทำงานแบบ Automate ในการดึงข้อมูลมาได้

ธีรัช คุ่มไข่มุก 6509035215

2.สิ่งที่เราชอบและไม่ชอบในการทำการบ้านครั้งนี้

ตอบ ชอบการบ้านครั้งนี้ครับ แต่ที่ไม่ชอบคือตัวเอง เพราะ parquet นั้น เป็น สิ่งที่ไม่คุ้นเคย สมควรหาโอกาสที่ทำให้ตัวเองไปคุ้นเคยมากกว่านี้ เพราะ เป็น File ที่วงการ Big data ใช้กันอย่างหลากหลาย

3.คิดว่าตัวเองควรปรับปรุงอย่างไร หรือ มีอะไรอย่างอื่นที่ควรได้รับการปรับปรุงสำหรับการบ้านครั้งต่อไป

หาวิธีการรับมือกับ Data มากกว่านี้ ในด้านการ Coding หรือ Implement