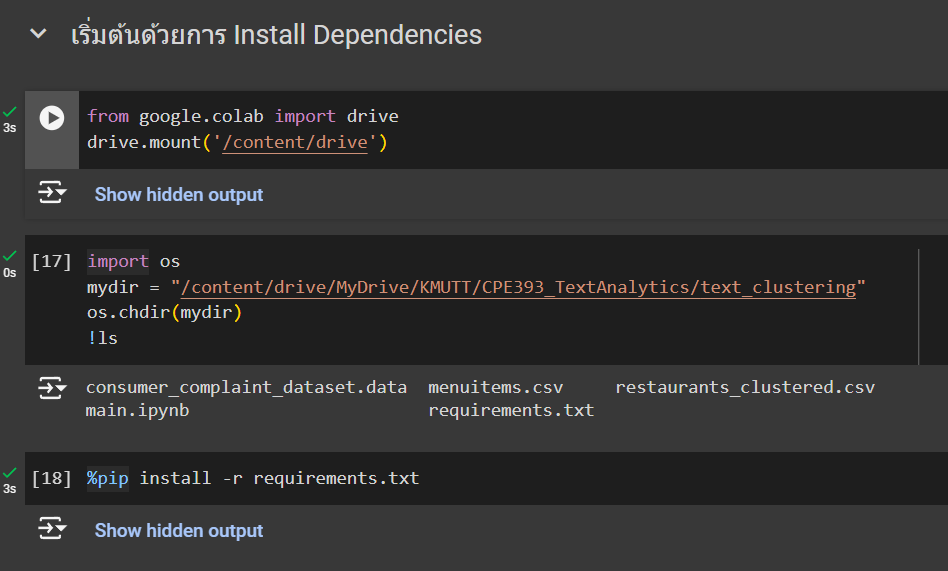
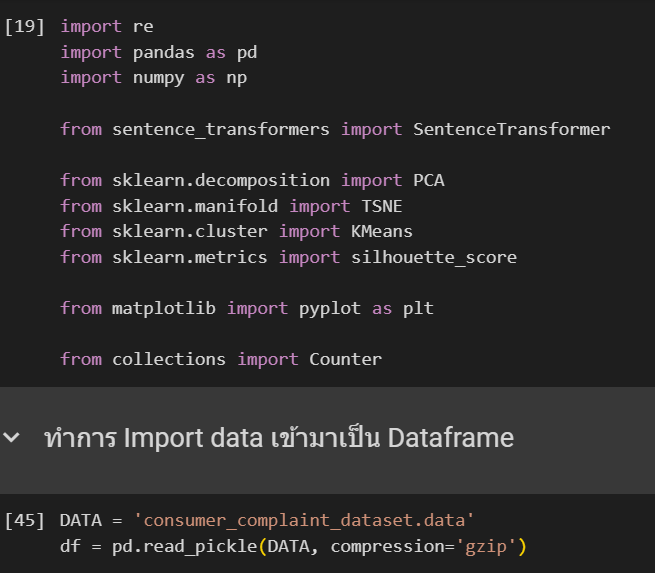
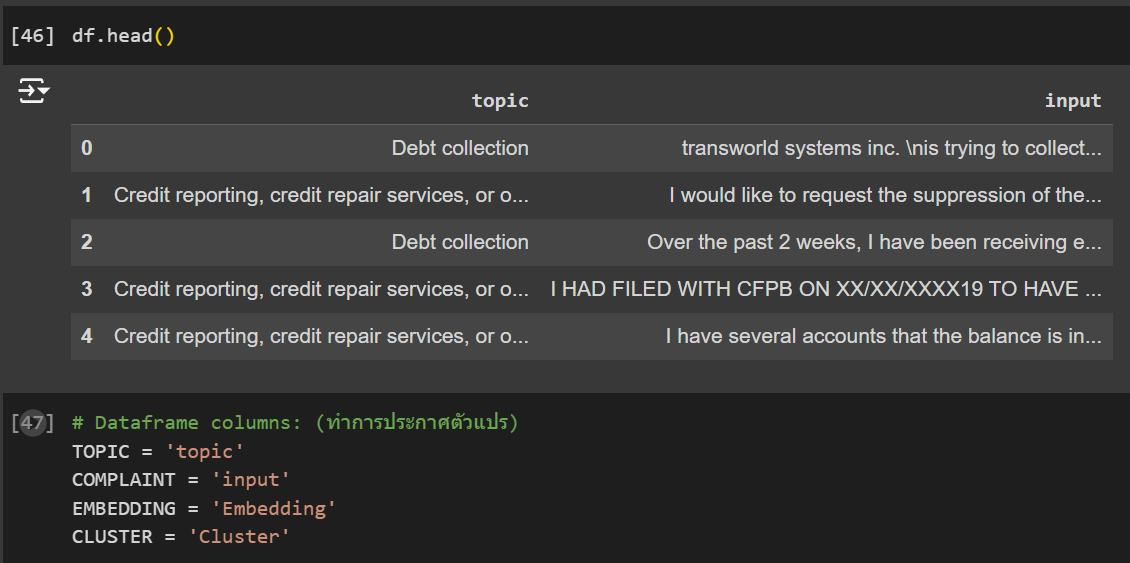
**Lab 07 – Text Clustering**

* เริ่มต้นด้วยการ Install Dependencies

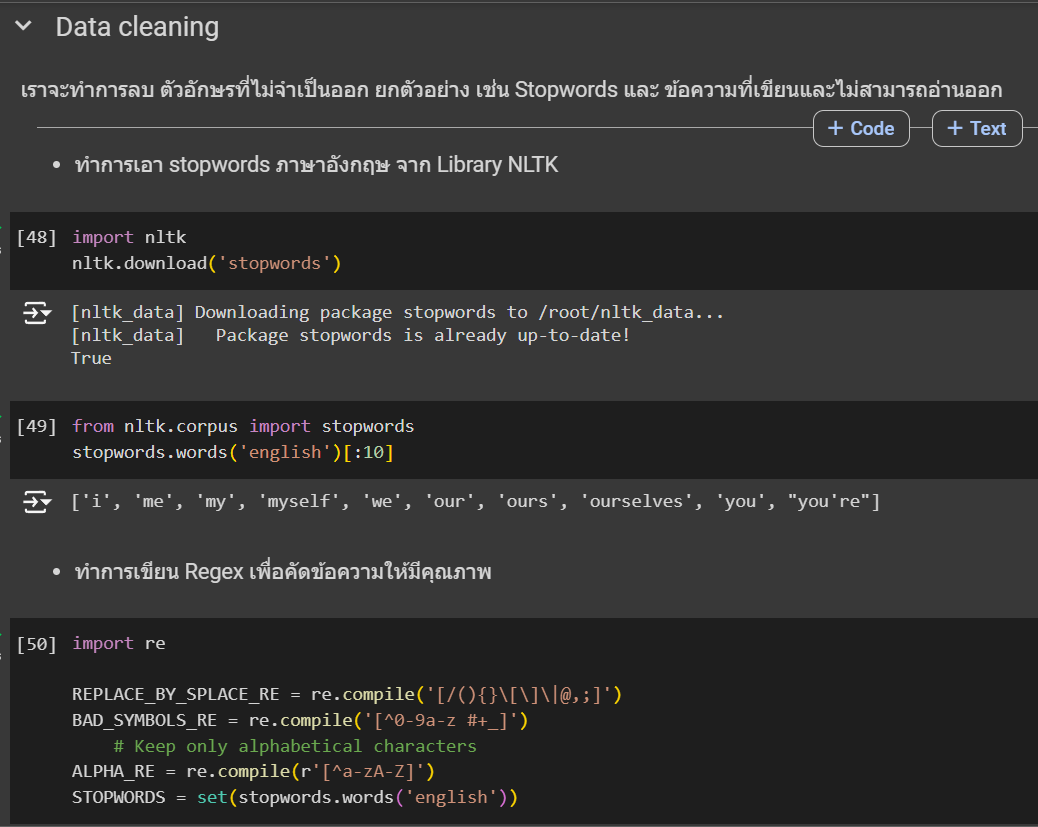
****

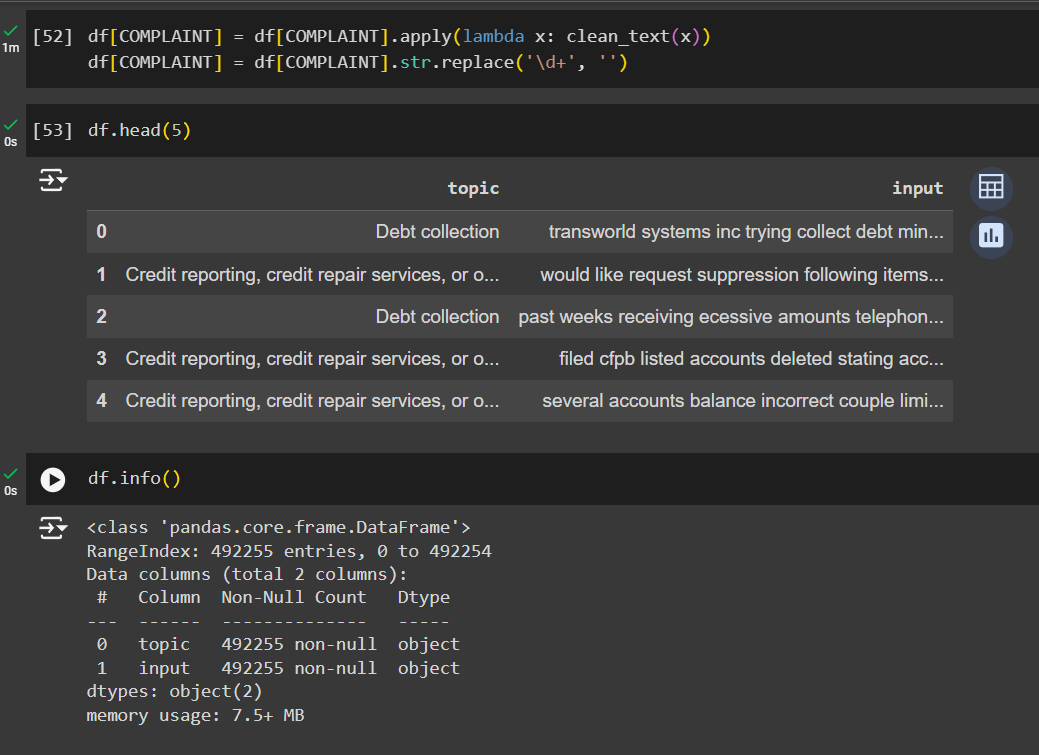
****

* Import Data มาเป็น Dataframe

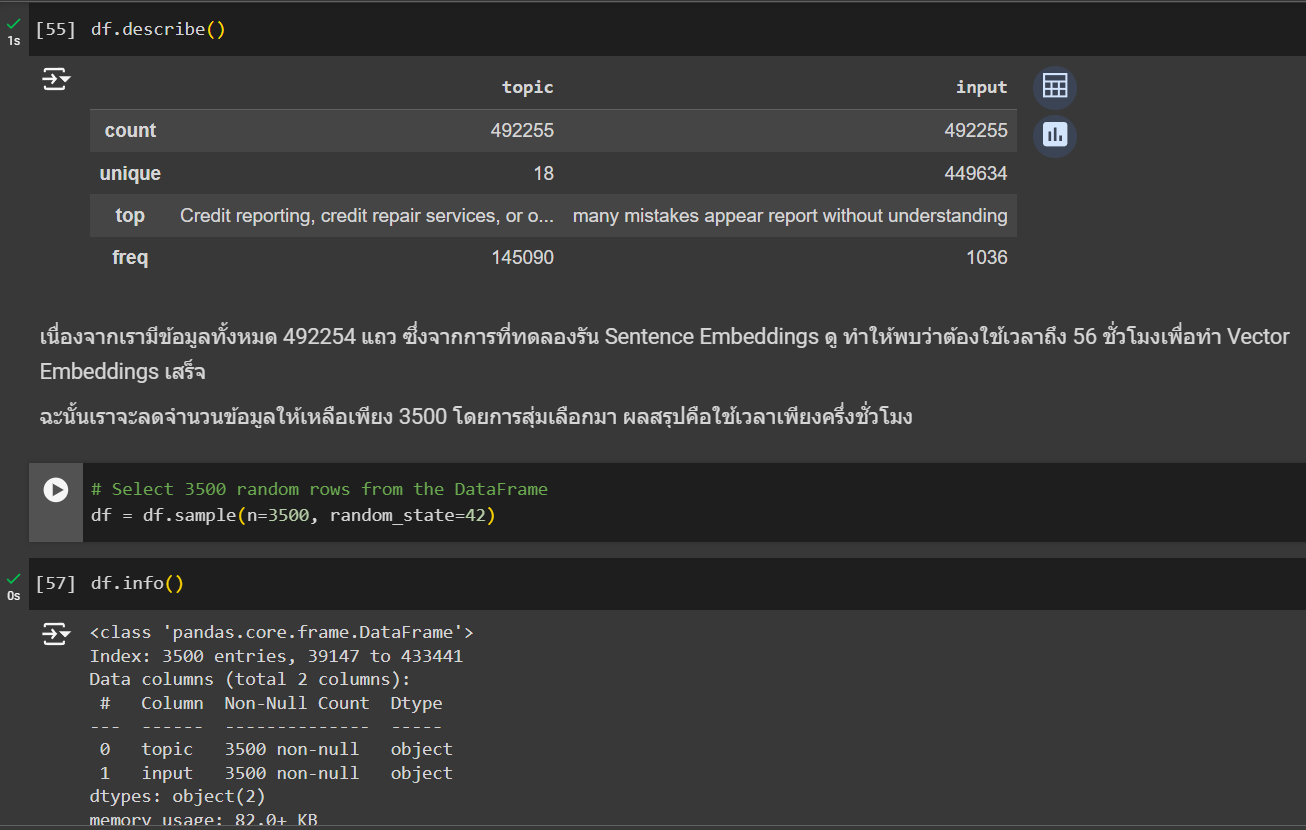
****

* Data Cleaning – Stopwords, Regex

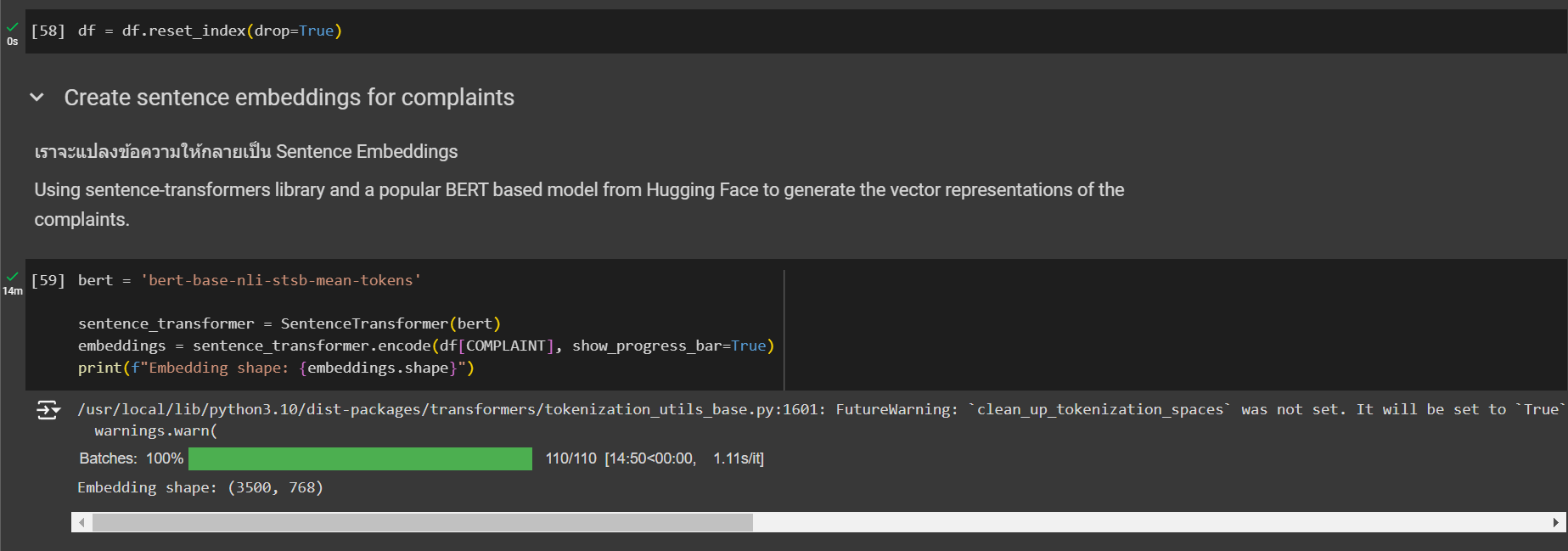
****

****

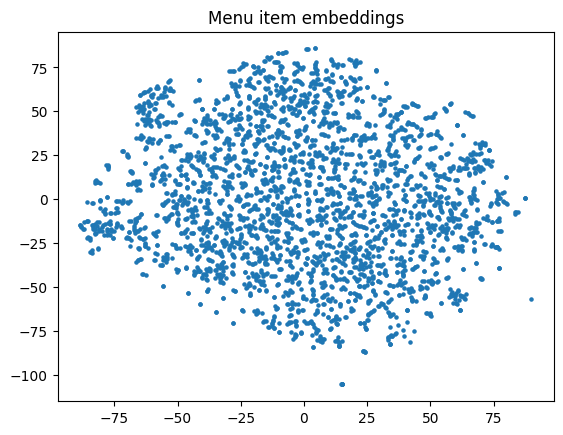
* ทำการลดจำนวนข้อมูลให้เหลือเพียง 3500 (เพราะต้องรอ Vector Embedding 56 ชั่วโมง นานเกินไป)

****

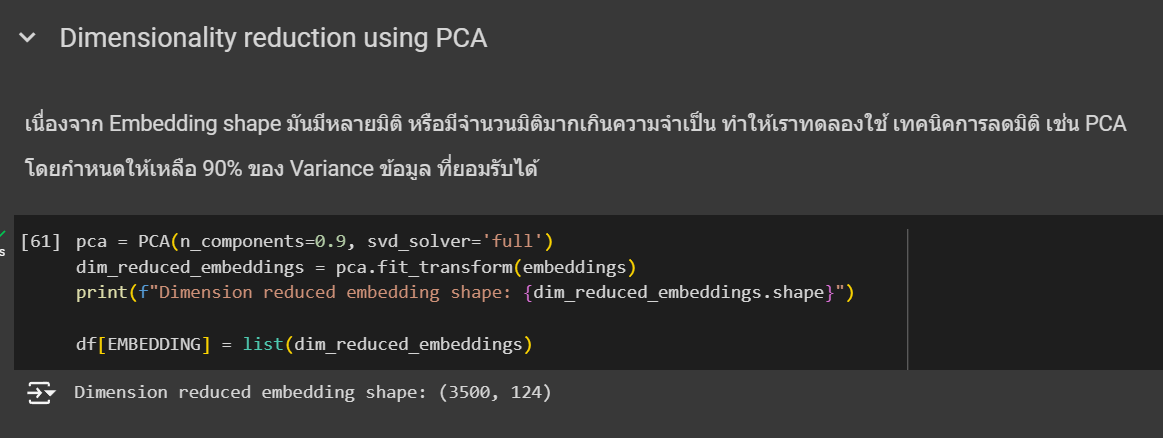
* สร้าง Sentence Embeddings ด้วยโมเดล Sentence Transformer ตระกูล BERT

****

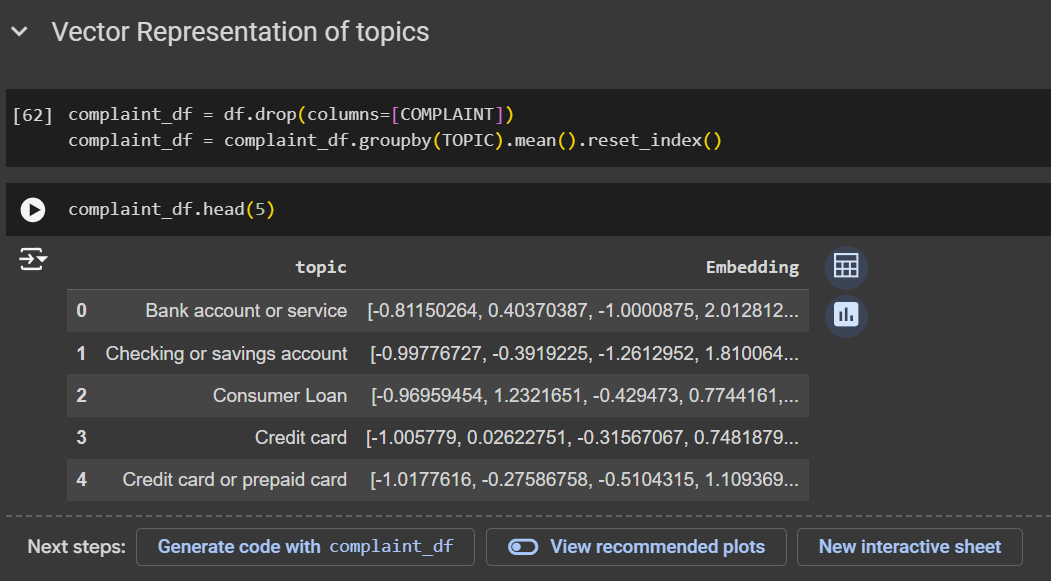
****



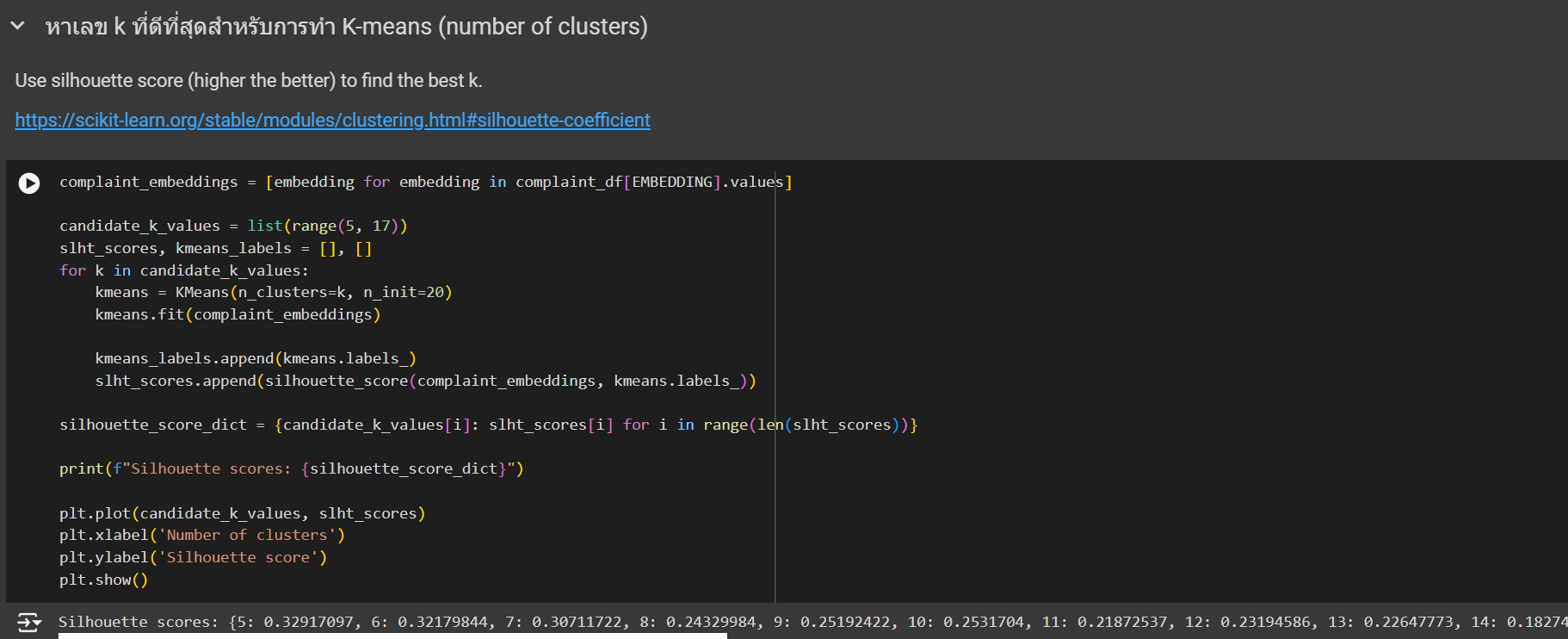
* ลดขนาดของมิติด้วยเทคนิค PCA

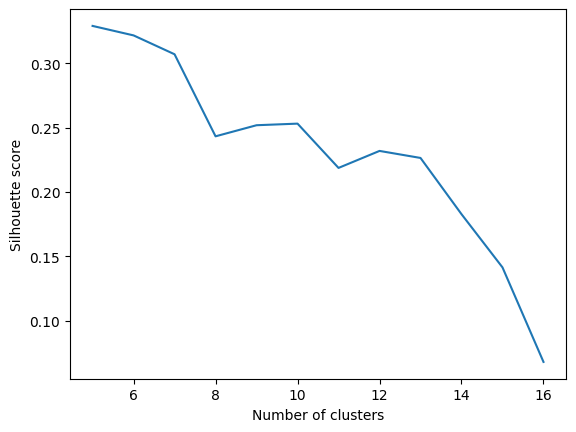
****

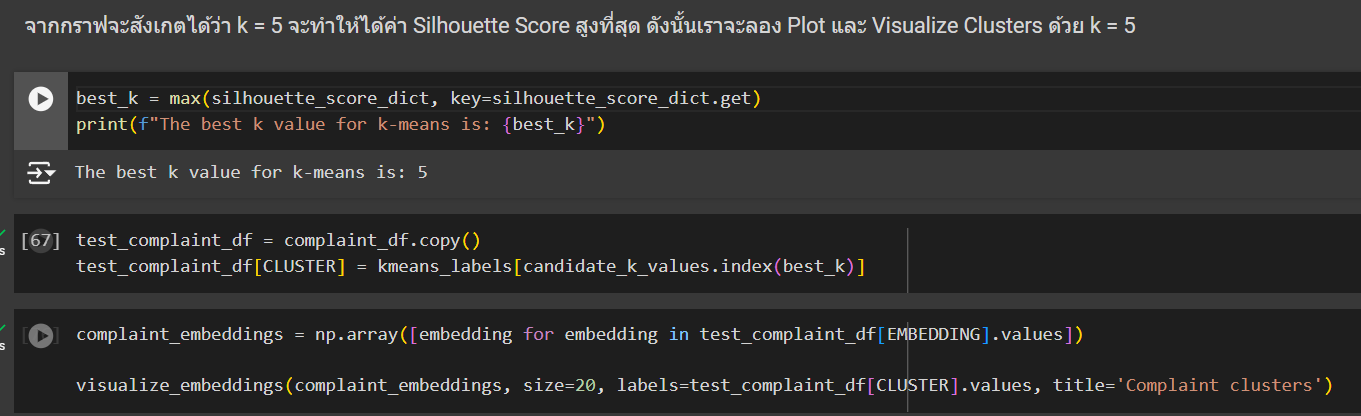
* Vector Representation ของ Embedding

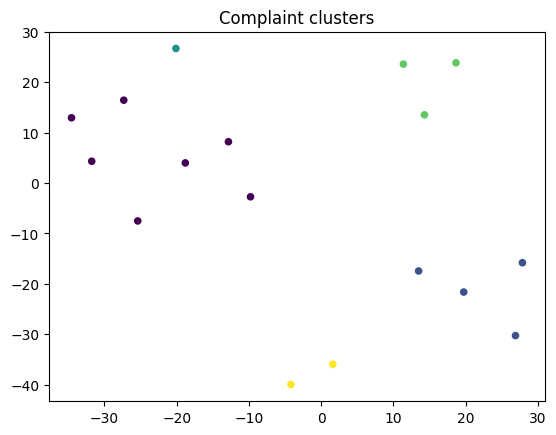
****

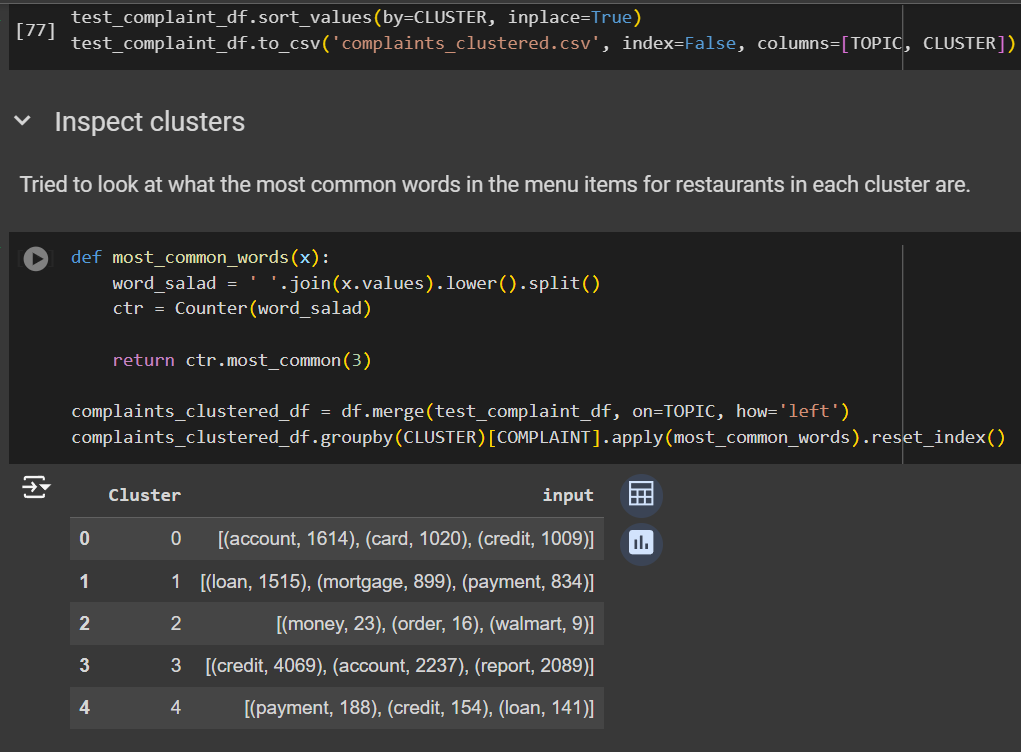
* หาค่า k ที่ดีที่สุด จาก Silhouette Score

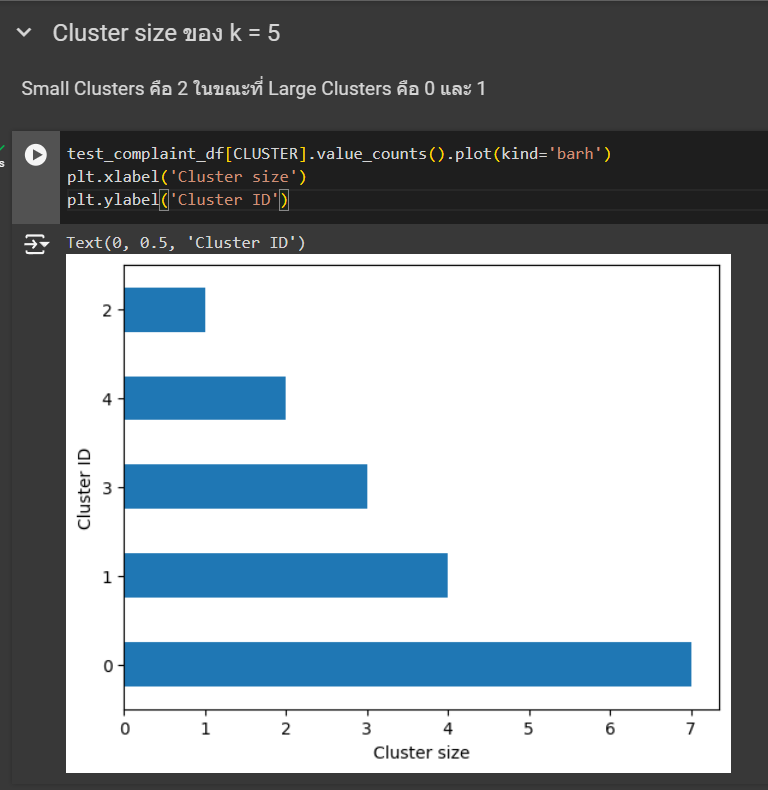
****

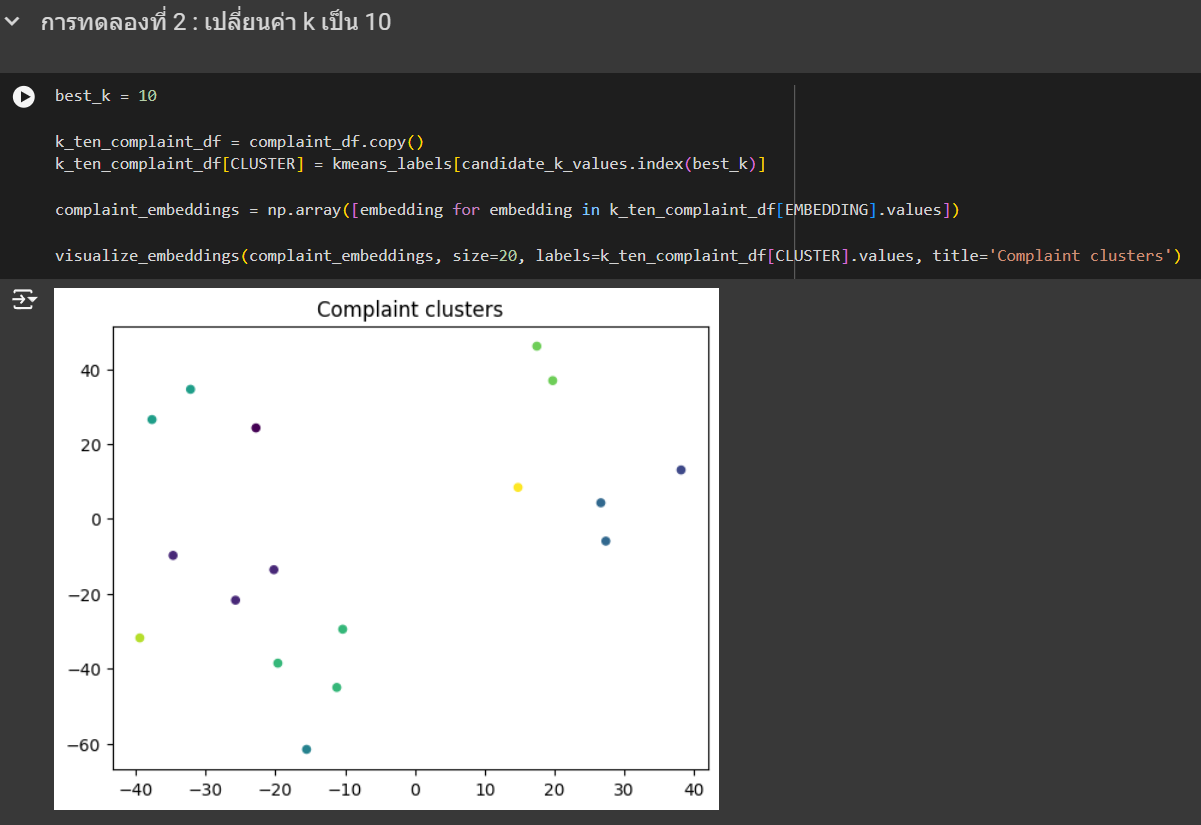


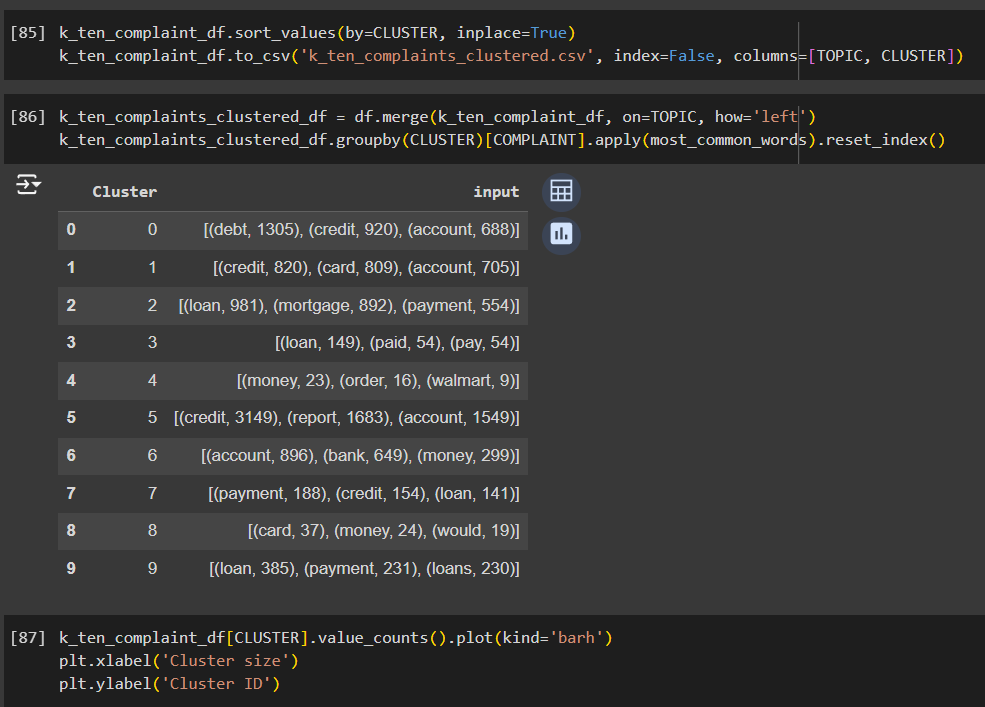
****

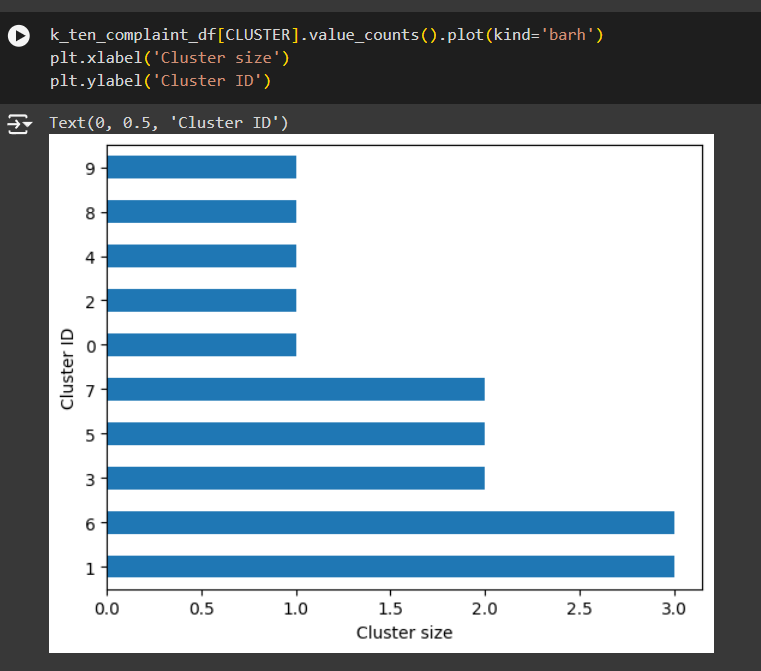


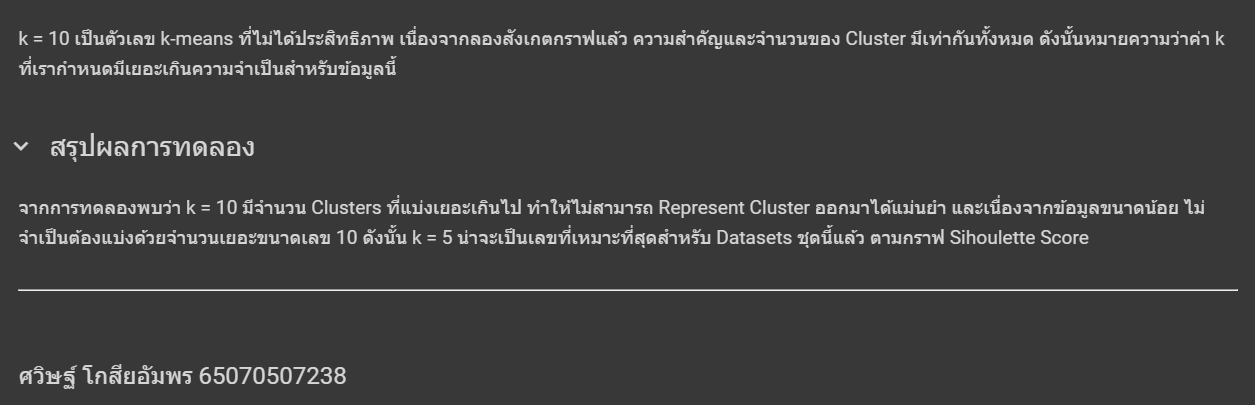
****

****

****

****

****

****