Dava Virgio Kertawijaya

00000056848

====================================================================

**PROBLEM/QUESTIONS:**

***It is reminded again that the answer given is in accordance with the business scenario, output and the data provided in the IS-429 BDA Theory answer.***

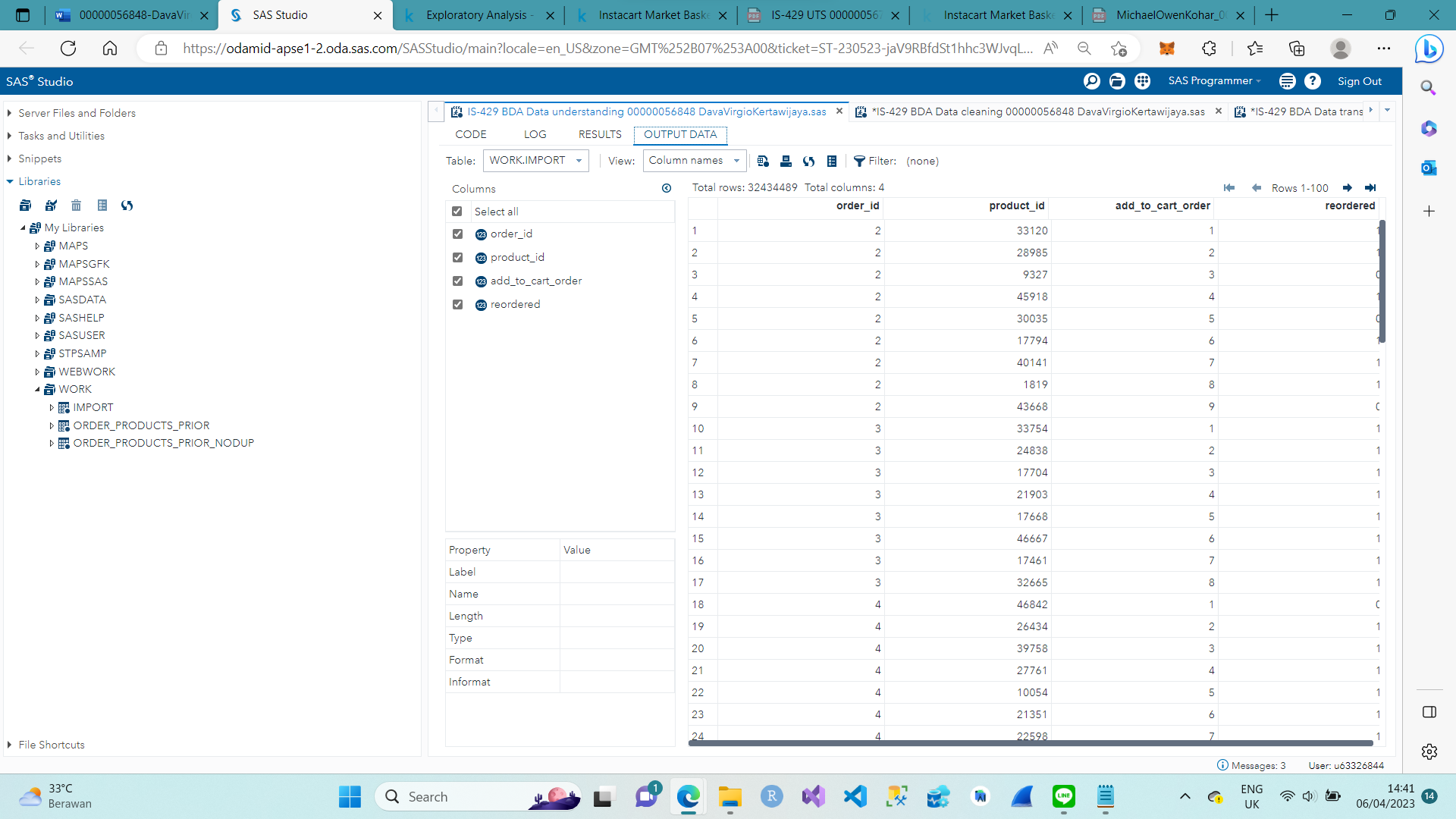
***For further explanation, please pay attention to the instructions for question number 7 above.***

1. **Question 1: Sub-CLO 2 & 3 Weight (50%)**

Manage you data Using SAS® Studio Software

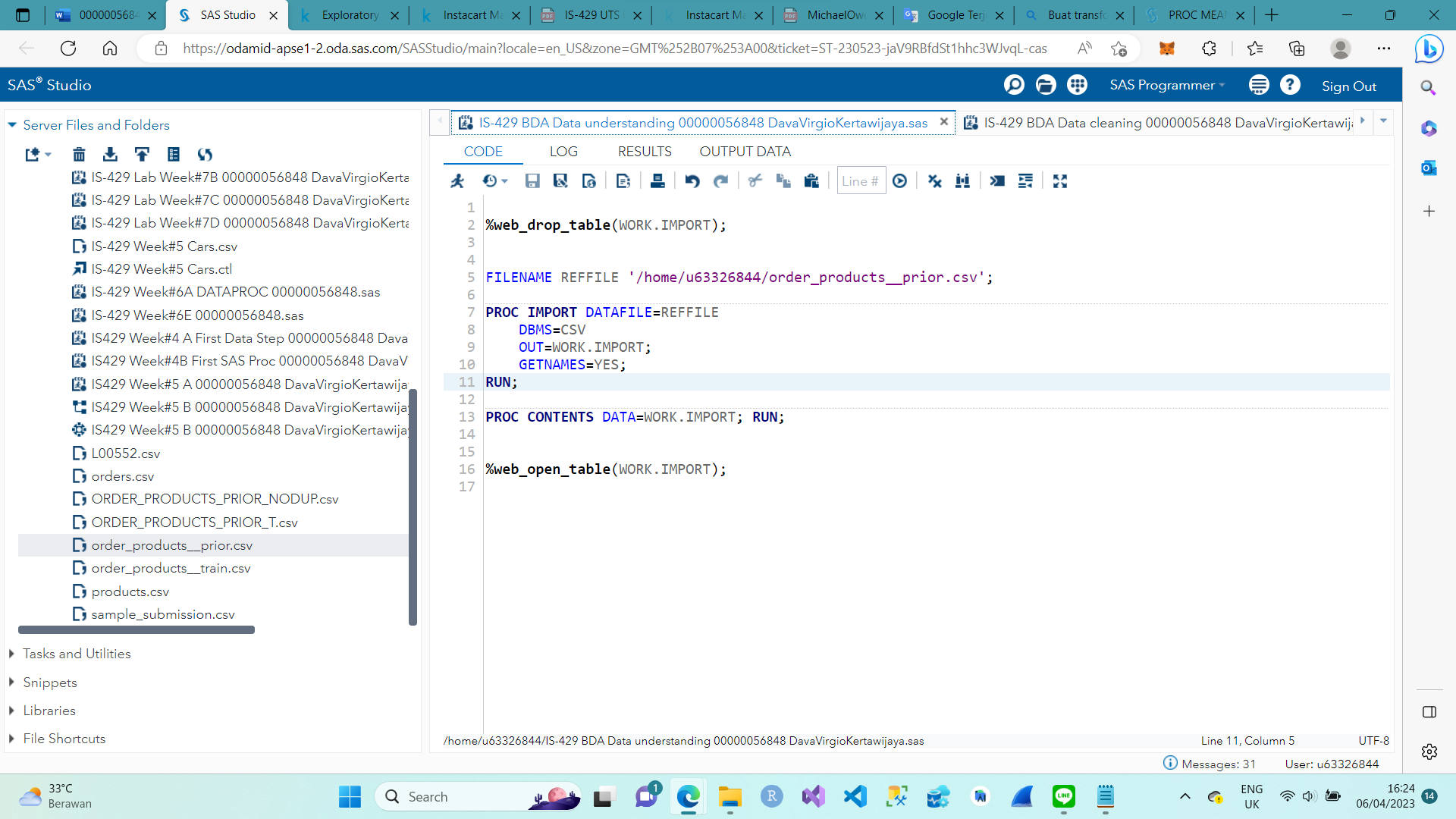
1. Explain what data you store in SAS CAS or GCP BigQuery (If you use GCP Software), provide a **screenshot** of the **library or folder** that stores your data.

Data yang disimpan dan digunakan dalam project kali ini, yaitu data order\_products\_prior.csv dari dataset Instacart yang terdiri dari 7 file csv. Berikut ini merupakan screenshot library yang menyimpan data tersebut.



1. Complete your answer above by **submitting a program** to **store your data**.

Program yang di submit untuk menyimpan data saya dapat dilihat di dalam foldernya. Berikut merupakan kode yang dijalankan untuk menyimpan data ke dalam library.



Saat kode tersebut dijalankan, data order\_products\_prior.csv yang telah diimport dapat dimasukkan ke dalam library untuk diolah lebih lanjut. Outputnya terlihat pada gambar 1. A. di atas.

1. Save the **input data** that you use into a **spreadsheet file format** for later merging into a zip file for your exam answers.

Data csv yang terdapat pada folder ini merupakan data yang telah diolah, bukan raw data. Dikarenakan ukuran file csv raw data yang sangat besar mencapai 681 MB, maka dalam hal ini, saya hanya mencantumkan link dataset raw datanya yang digunakan dari data open source.

Berikut ini link dataset yang digunakan:

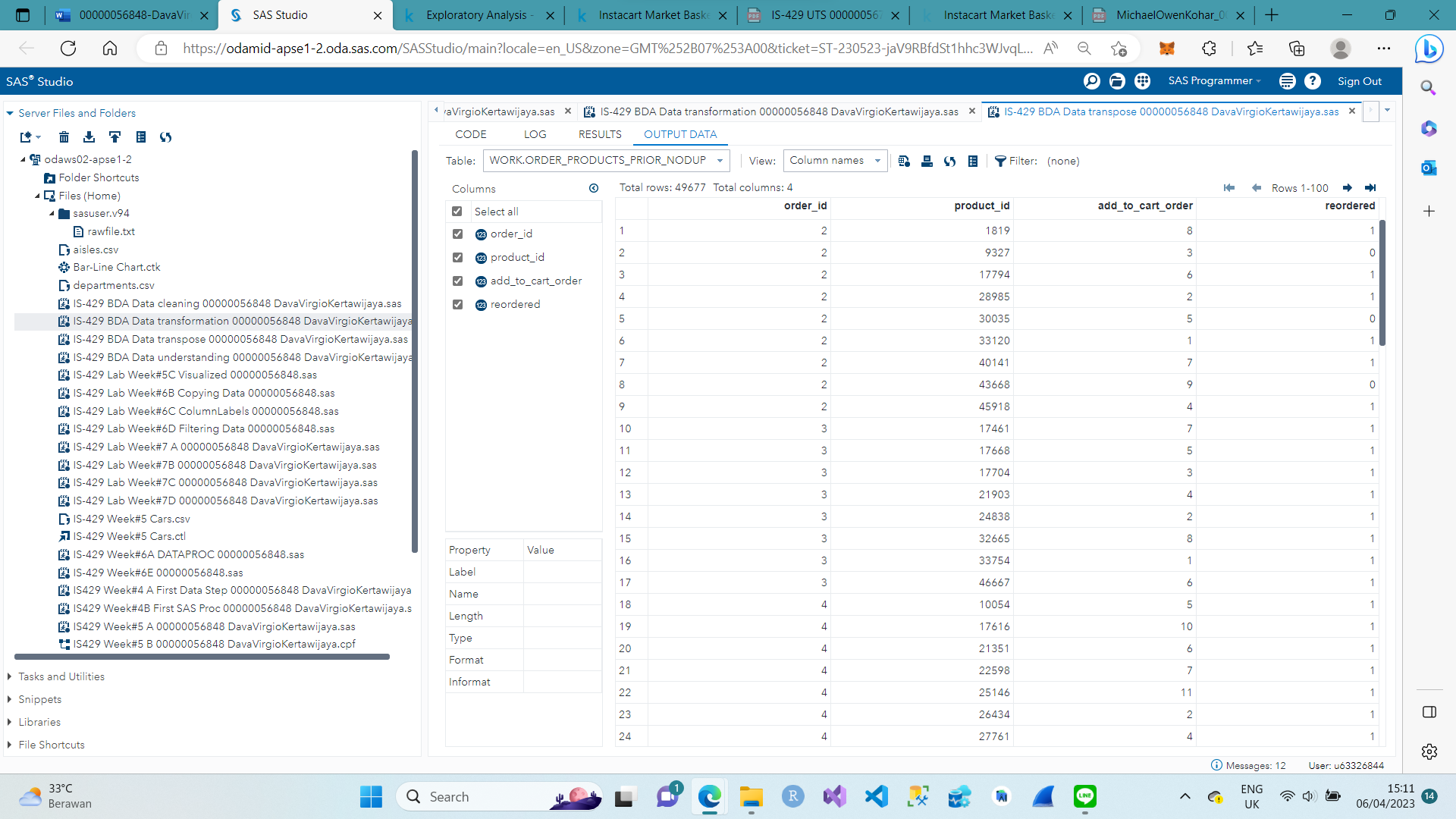
<https://www.kaggle.com/competitions/instacart-market-basket-analysis>

1. **Question 2: Sub-CLO 5 & 7 Weight (50%)**

In order for you to prepare data to be analyzed using analytics, using the features available in SAS Studio GCP BigQuery (If you use GCP Software), you create the following programs:

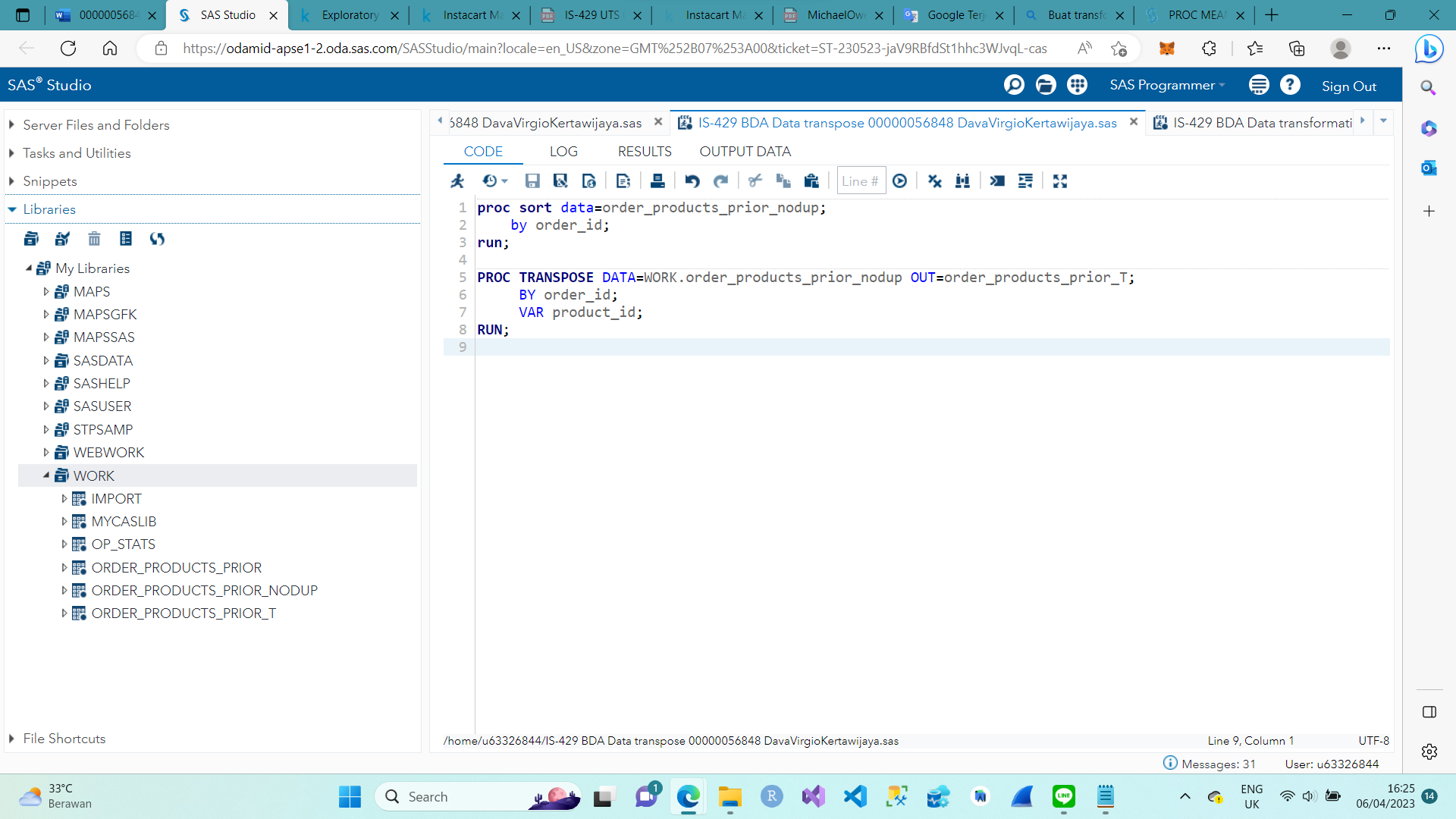
* 1. Make the **process of transposing** the rows and columns in a table, in order to get the data in the proper format for analytics to be applied

Berikut ini merupakan output transposing data yang dilakukan pada data order\_products\_prior\_nodup, yaitu data yang telah dibersihkan dan ditransformasi pada tahap data curation di UTS Theory.



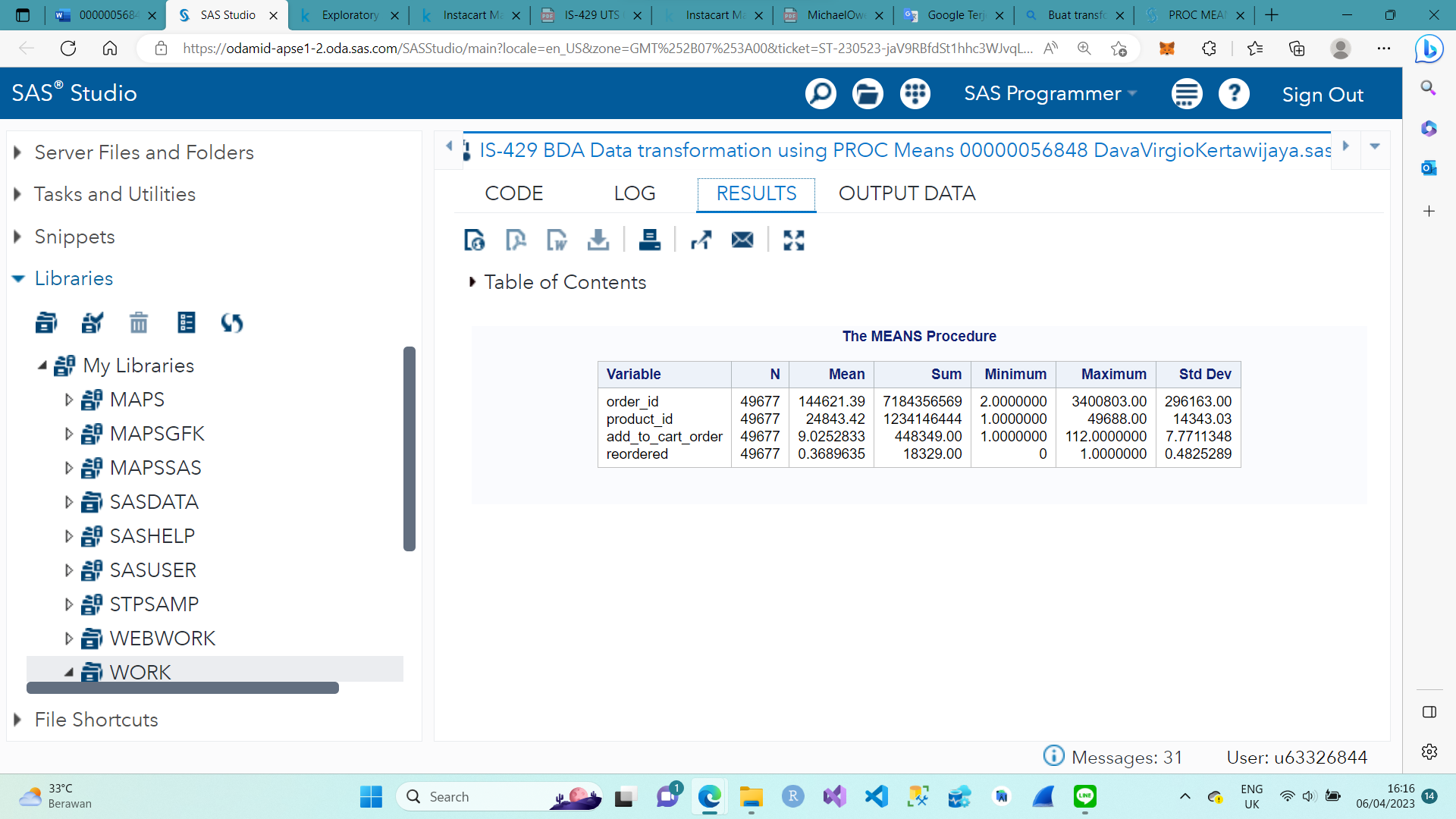
Pada gambar di atas, transpose dilakukan untuk memutar tabel yang awalnya memuat data transaksi belanja di mana setiap baris mewakili satu produk dalam satu transaksi, menjadi tabel di mana setiap baris mewakili satu transaksi dan kolom-kolomnya berisi informasi tentang produk yang dibeli.

Dengan memutar tabel tersebut, pola pembelian produk secara keseluruhan, seperti produk mana yang paling sering dibeli bersamaan dan pada jam berapa pembelian dilakukan dapat dilihat. Selain itu, proses transpose ini juga dapat mempermudah proses analisis dan visualisasi data pada langkah selanjutnya. Berikut ini merupakan kodingannya.

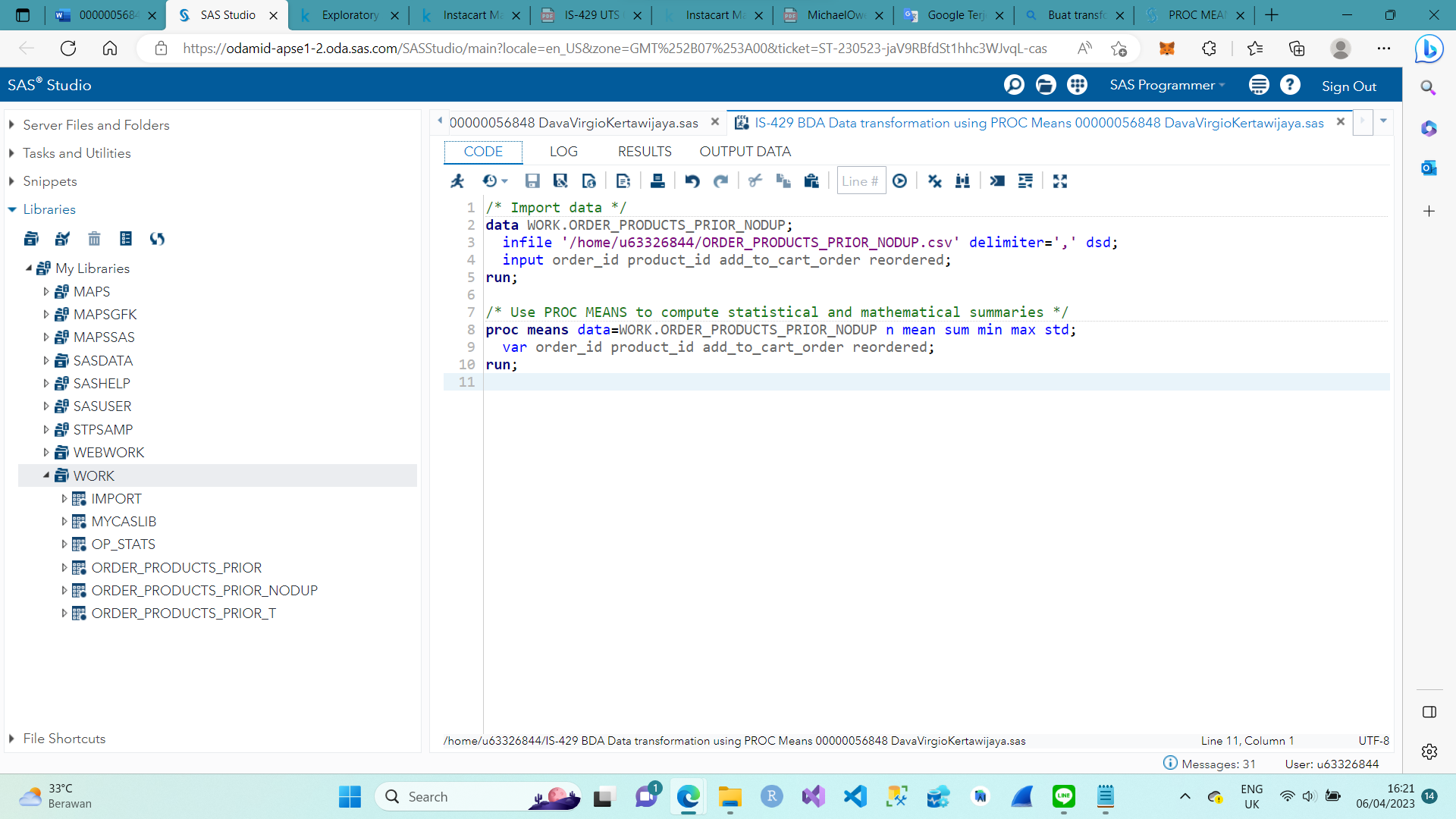


* 1. Make **simple Statistical** and mathematical **data transformations** using SAS **PROC MEANS** (if you use SAS CAS software) in order to enrich the data with statistics related to the existing numeric within the data table you have.

Berikut ini merupakan hasil transformasi data statistik dan matematis sederhana menggunakan SAS PROC MEANS pada data order\_products\_prior\_nodup, yaitu data yang telah di transformasi pada UTS Theory.



Terlihat pada gambar di atas, Means procedure dari data order\_products\_prior\_nodup mencakup jumlah n, mean, sum, minimum, maximum, dan standar deviasi dari data ini. Statistik ini dapat memberikan informasi tentang distribusi data yang sedang dianalisis. Selain itu, informasi ini juga dapat membantu dalam memahami karakteristik data dan membantu dalam pengambilan keputusan. Berikut ini merupakan codingannya.



Collect the **transpose program** and **simple statistics program** mentioned above as part of your examination answer .zip file.