**TUGAS KELOMPOK**

**PROGRAM STUDI INDEPENDEN**

**ORBIT FUTURE ACADEMY**

**Identitas Kelompok**

Kelompok : 10

Nama Anggota : Halomoan Filipus Simarmata (Jupyter XXI)

Diana Eka Riyani (Jupyter XXI)

Nyayu Chika Marselina (Jupyter XXI)

Sukma Imelda (Cordoba)

Athiya Shinta Wulandari (Cordoba)

Coach : Ipin Sugiyarto

Program : Foundations of AI and Life Skills for Gen-Z

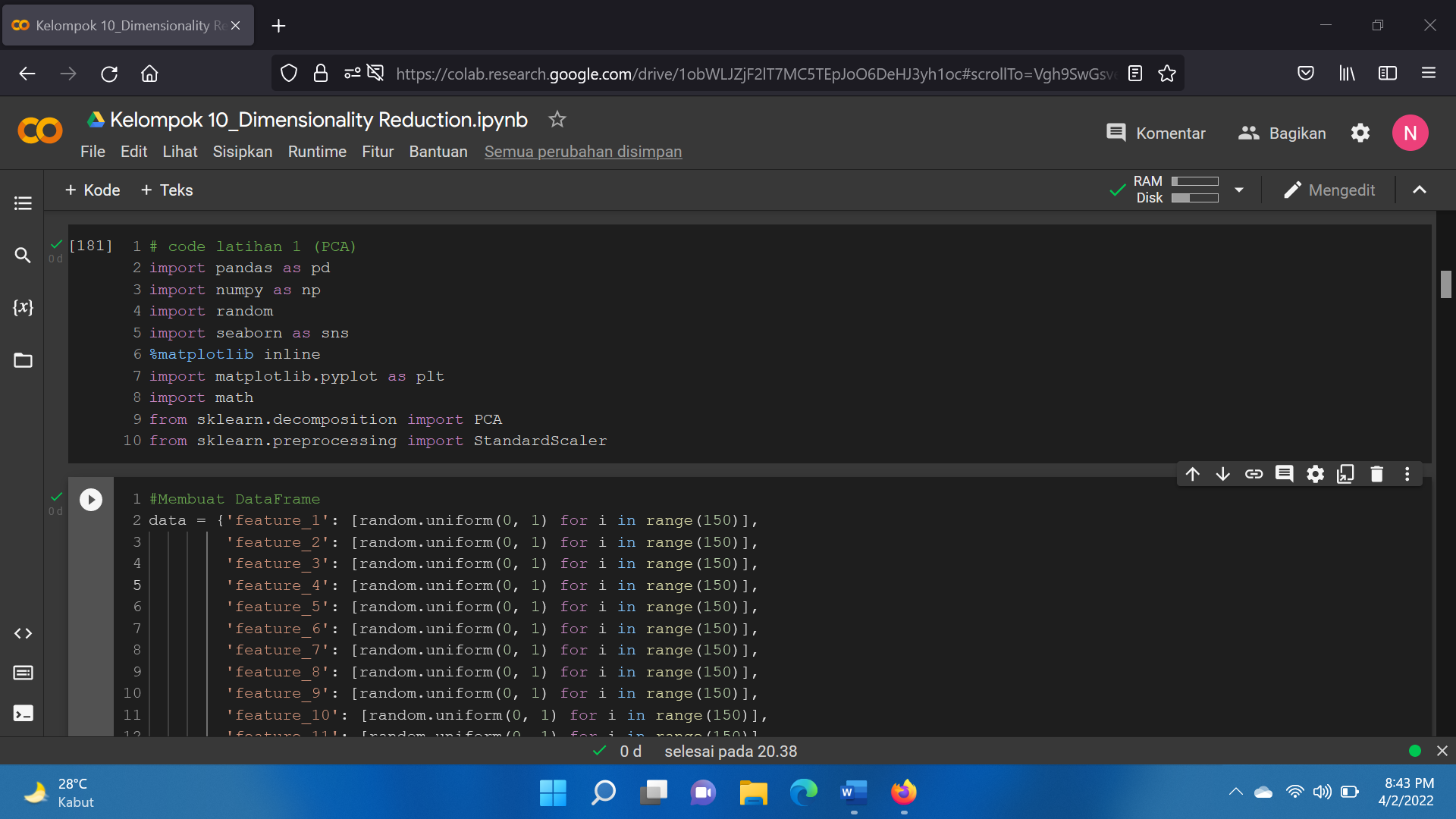
Hari, Tanggal : Kamis, 31 Maret 2022

**LATIHAN 1:** Membuat model data reduction dengan menggunakan PCA (Principle Component Analysis), dengan ketentuan sebagai berikut:

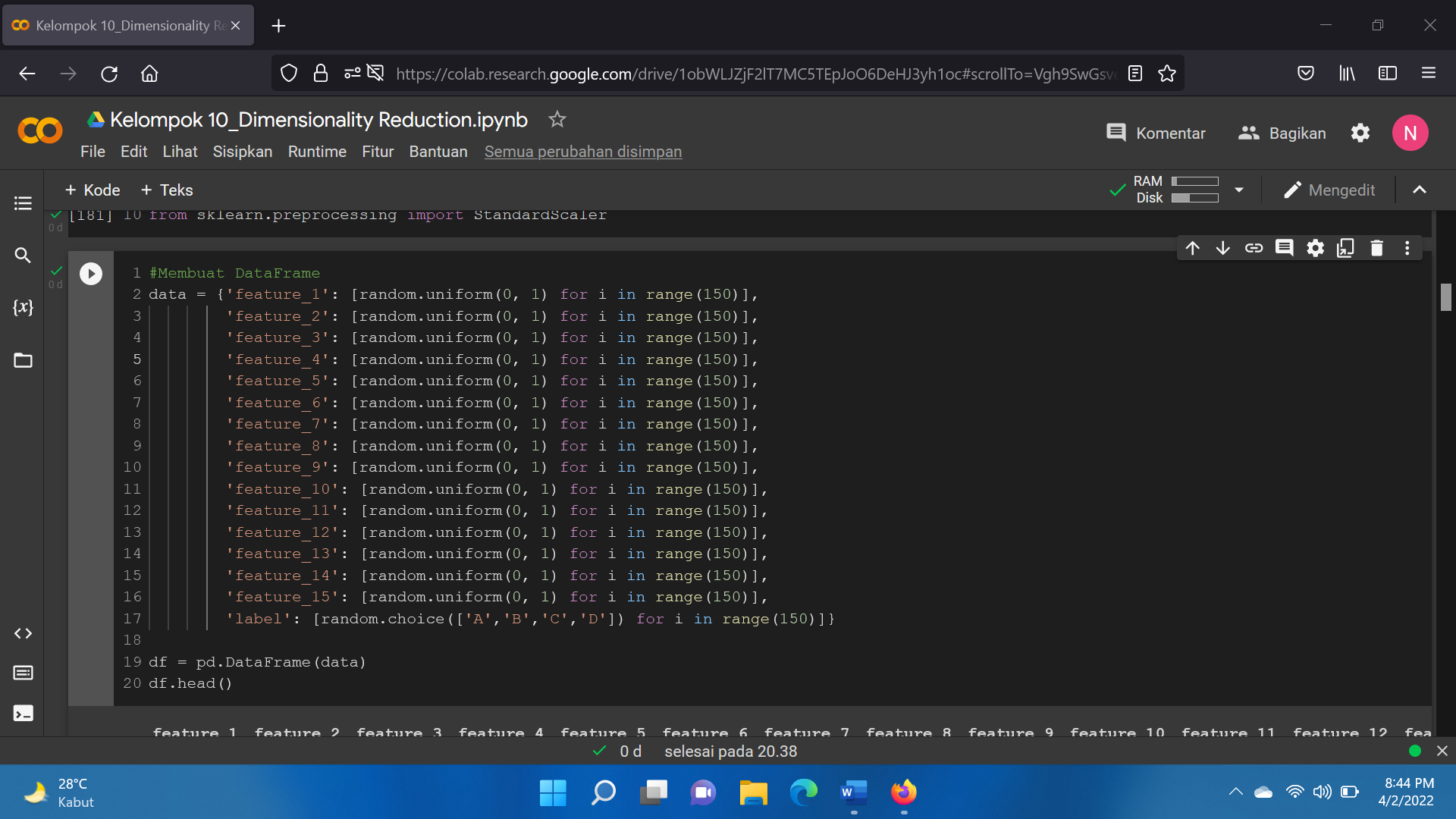
1. Buatlah DataFrame dengan jumlah fiture 15
2. Generate DataFrame dengan bilangan float secara random, antara 0 dan 1, serta buat n\_component berjumlah 2.
3. Membuat Label A, B, C, dan D (masing-masing jumlah 150 data), kemudian lanjutkan proses reduksi data model PCA sampai dengan visualisasi data dalam bentuk plot.

**Penyelesaian:**

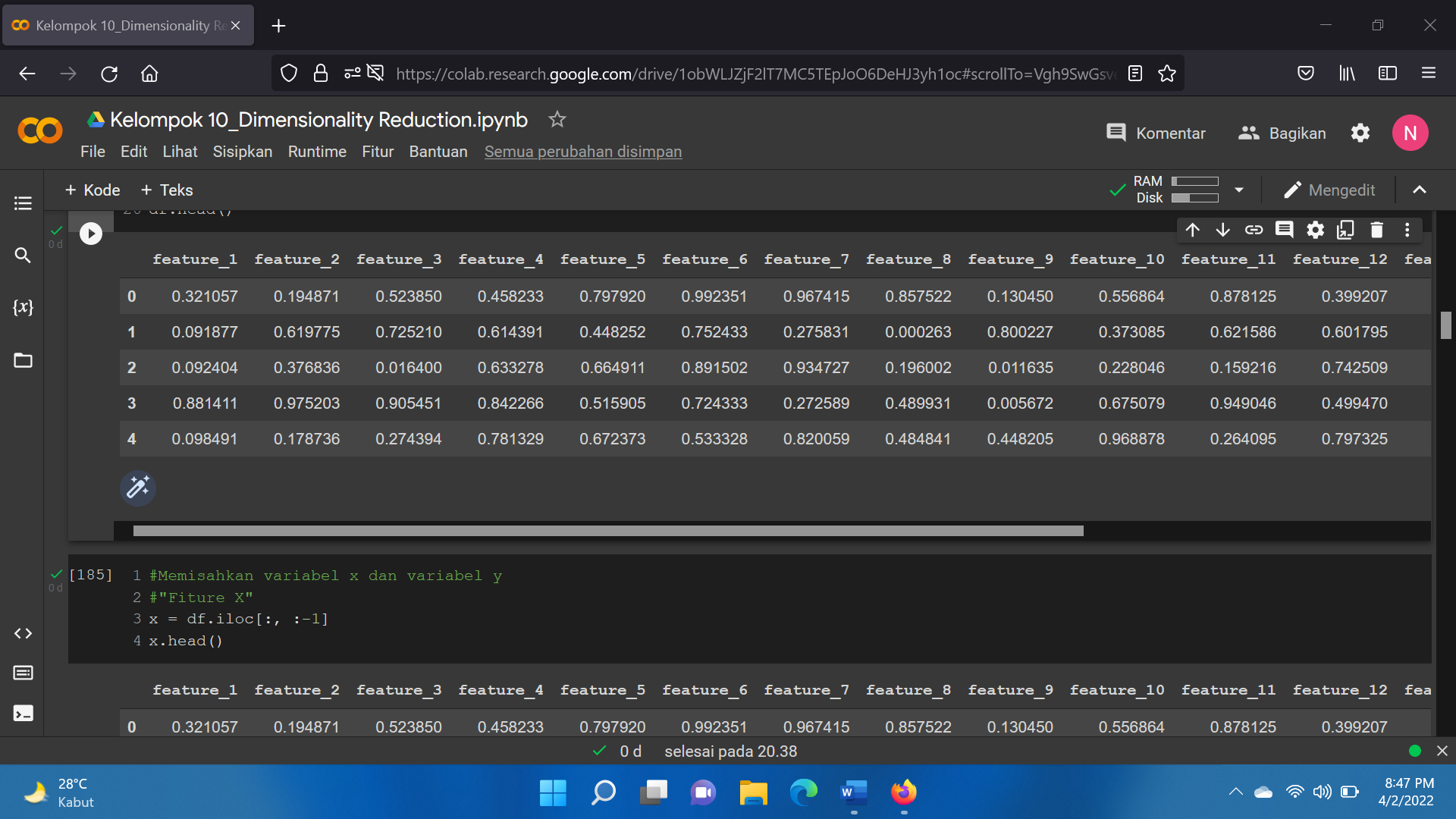
1. Import library yang dibutuhkan terlebih dahulu



1. Membuat DataFrame di mana berisikan 15 feature yang masing-masingnya memiliki jumlah data sebanyak 150 data serta terdiri dari label A, B, C, dan D.

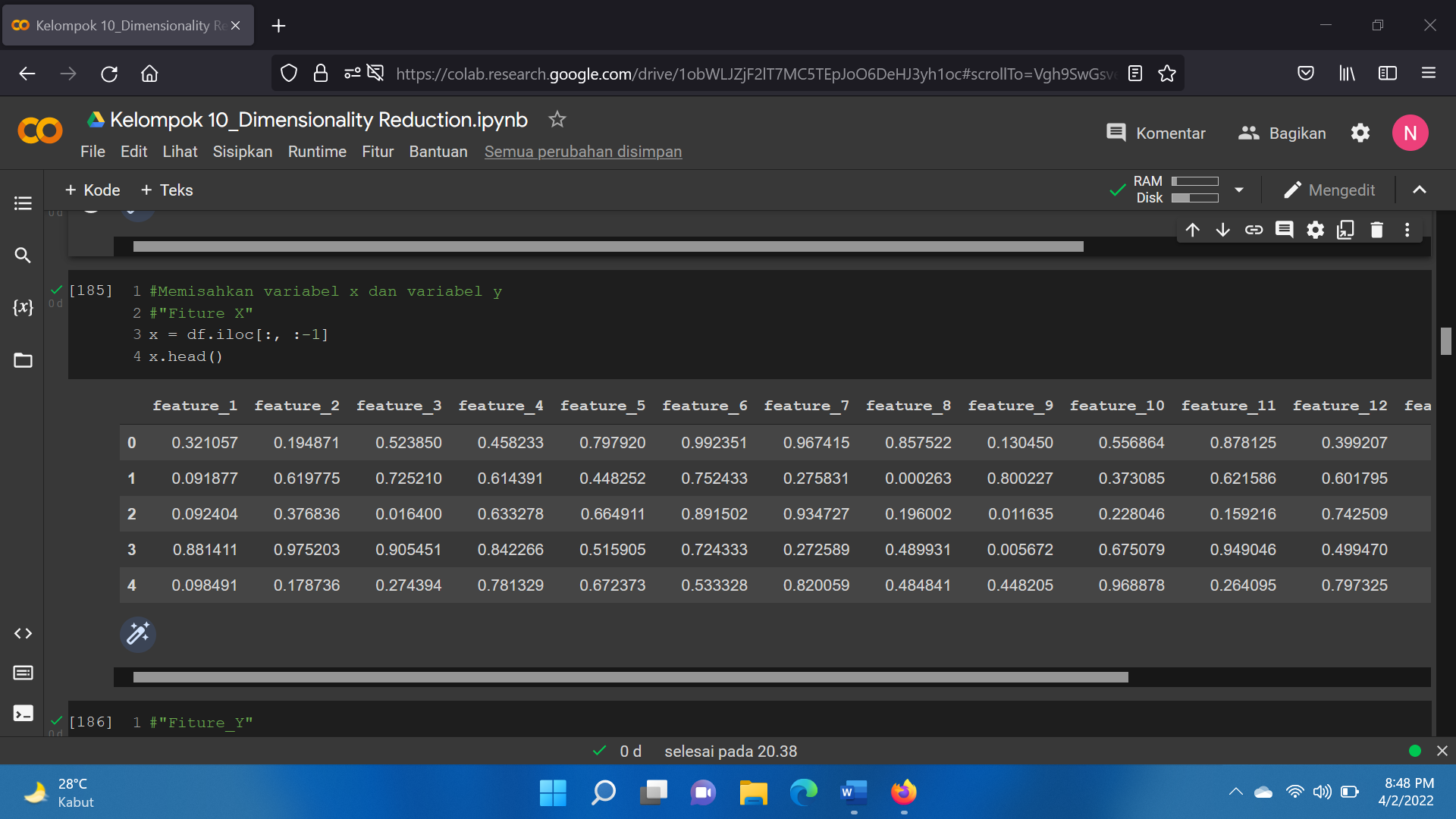


Hasil running coding di atas:

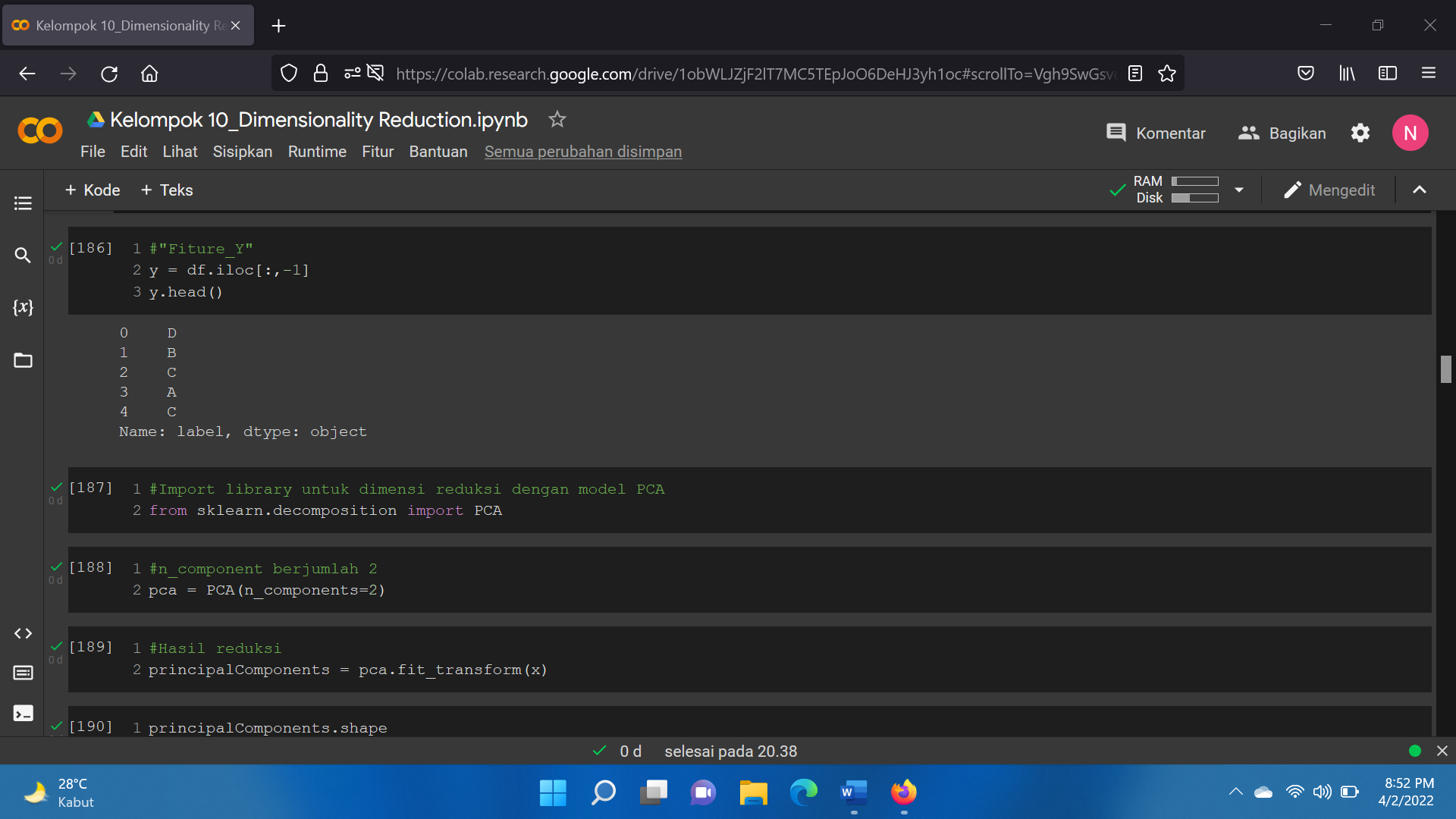


Karena menggunakan df.head(), maka hanya menampilkan 5 baris pertama dari DataFrame.

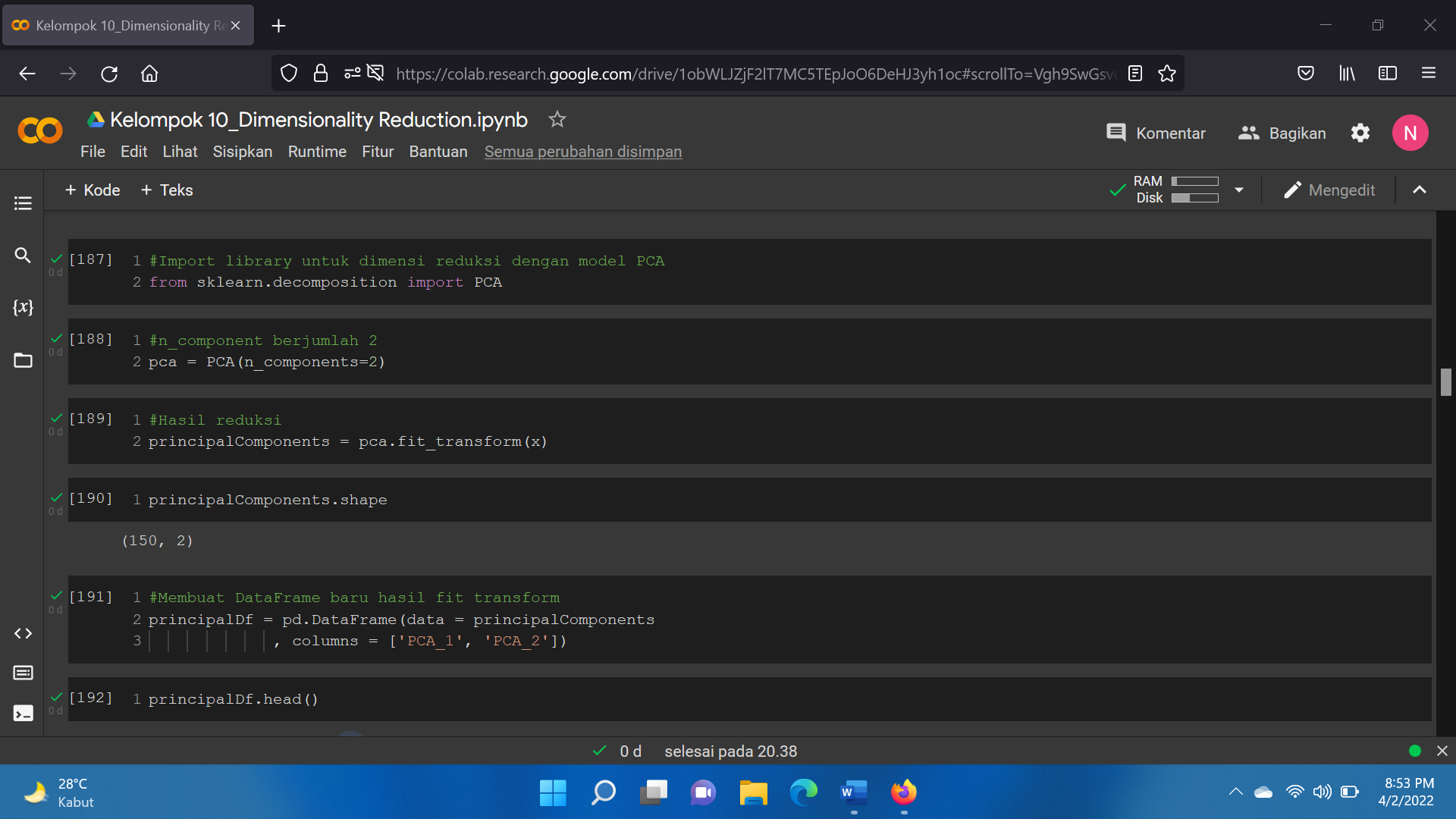
1. Memisahkan variabel x dari DataFrame yang dibentuk. Pada bagian ini, menampilkan 5 baris data pertama dari DataFrame yang dibuat dengan variabel x yang telah dipisahkan.



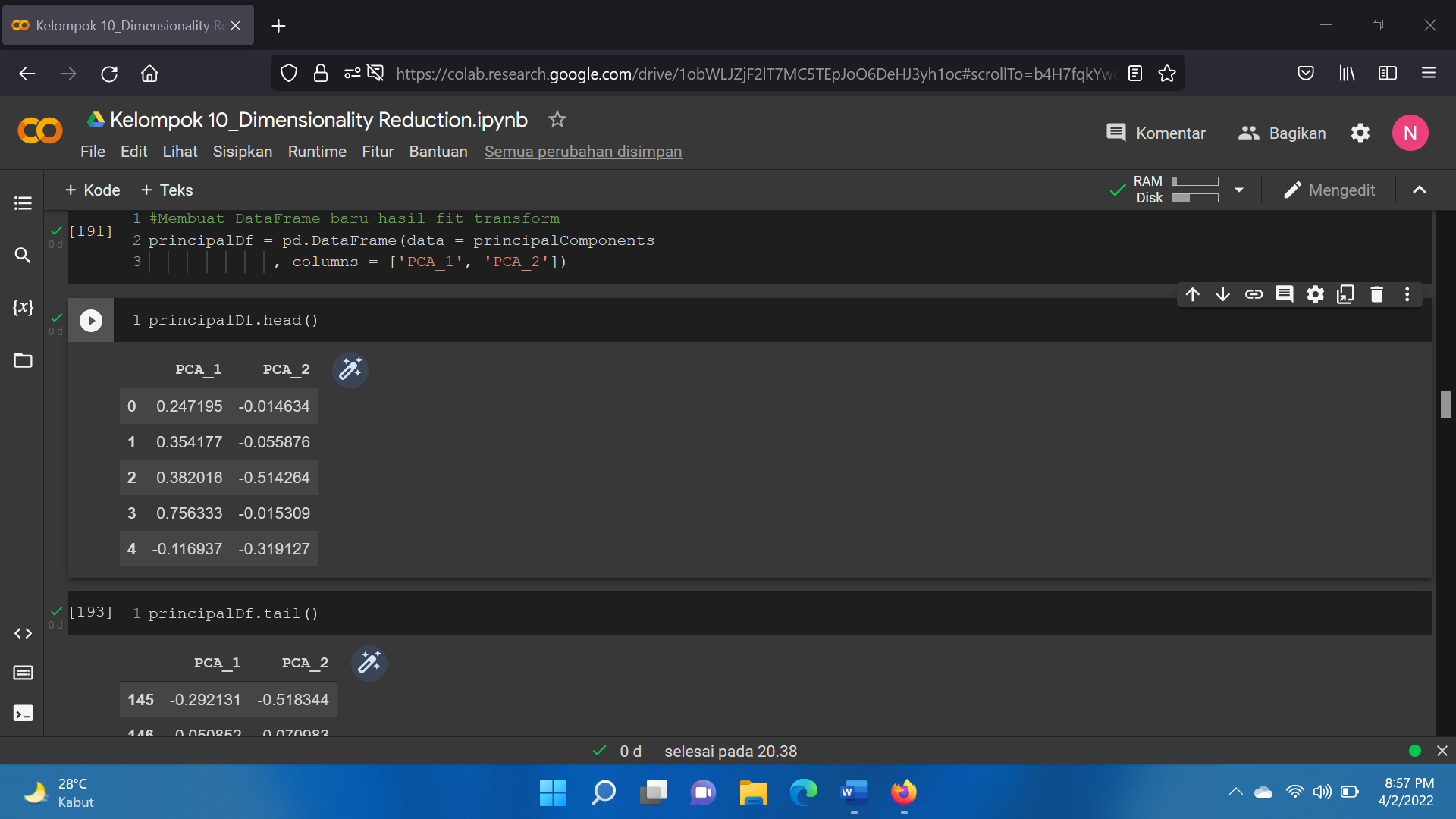
1. Menampilkan 5 baris data pertama dari DataFrame dengan variabel y yang telah dipisahkan.



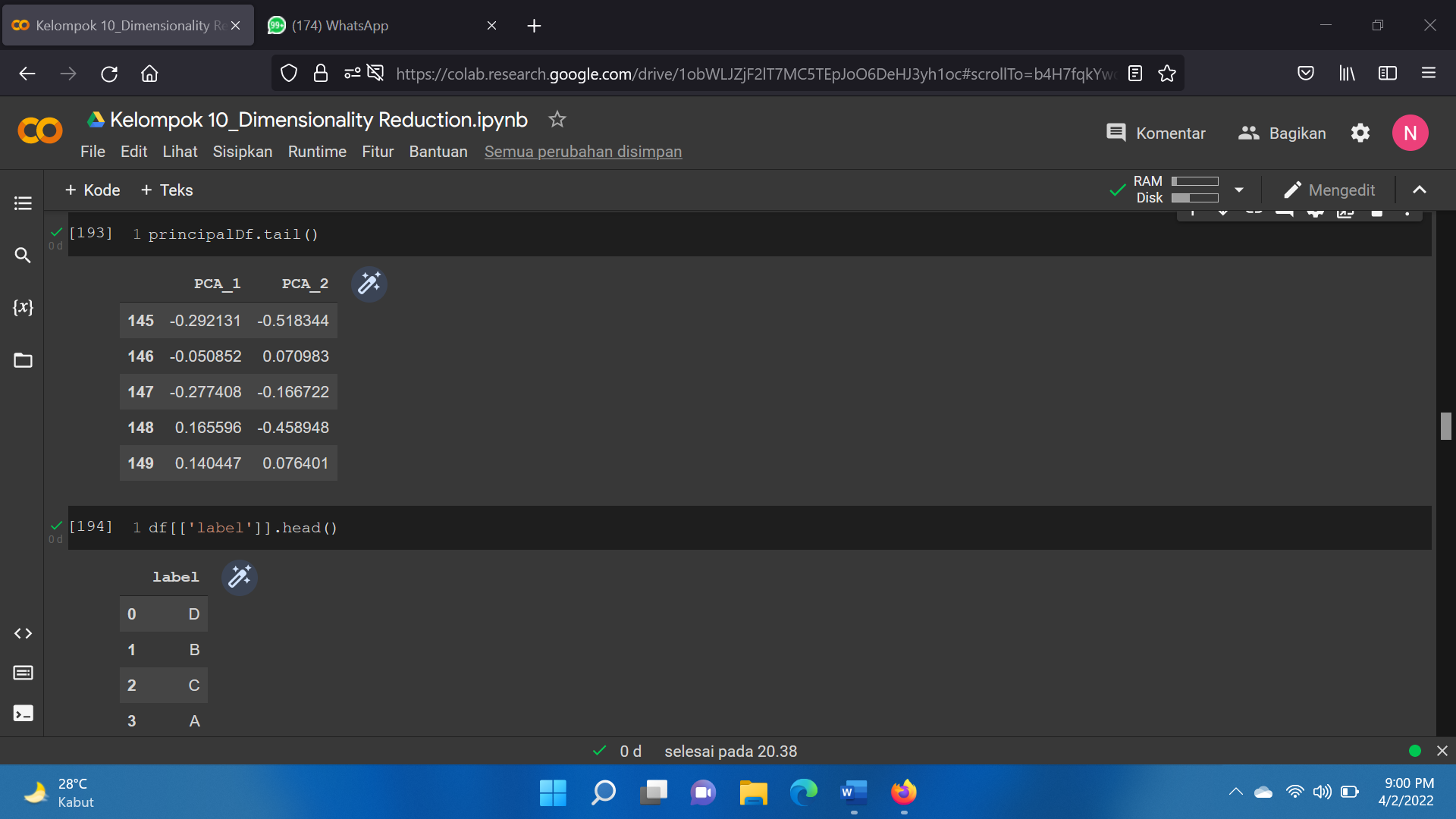
1. Mengimport library yang akan digunakan untuk mereduksi dimensi dengan model PCA. Kemudian membentuk n\_components berjumlah 2 yang disimpan di variabel pca. Lalu, melakukan reduksi yang mana hasilnya disimpan di dalam variabel principalComponents. Lalu mencetak ukuran dari hasil reduksi menggunakan principal.Components.shape. Setelah itu, membuat DataFrame yang baru (hasil dari fit transform) yang disimpan di dalam variabel principalDf. Serta menampilkan 5 baris pertama dari DataFrame baru hasil fit transform.



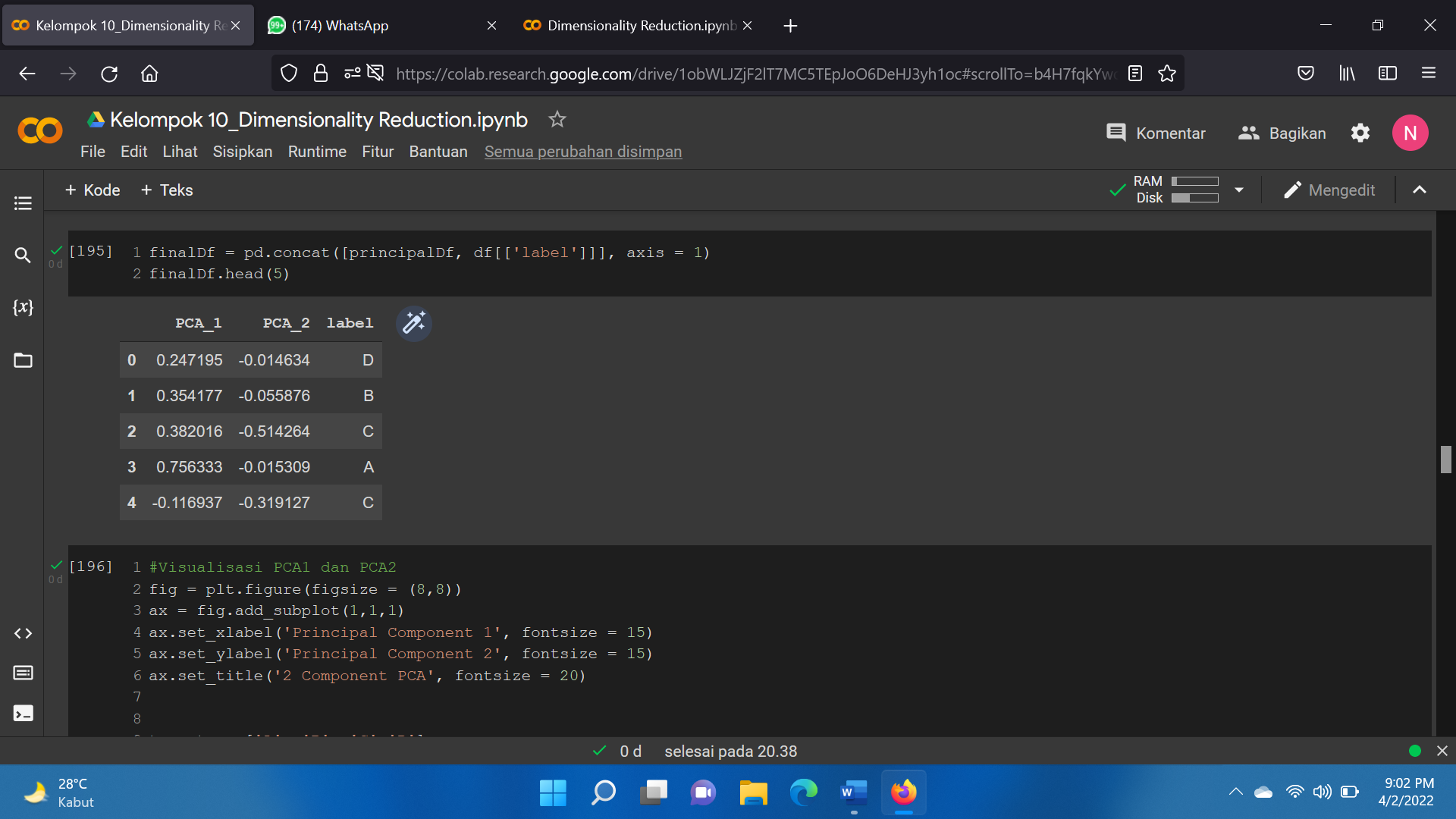
1. Menampilkan 5 baris pertama dan 5 baris terakhir DataFrame baru hasil fit transform. Berikut tampilan 5 baris pertama DataFrame baru:



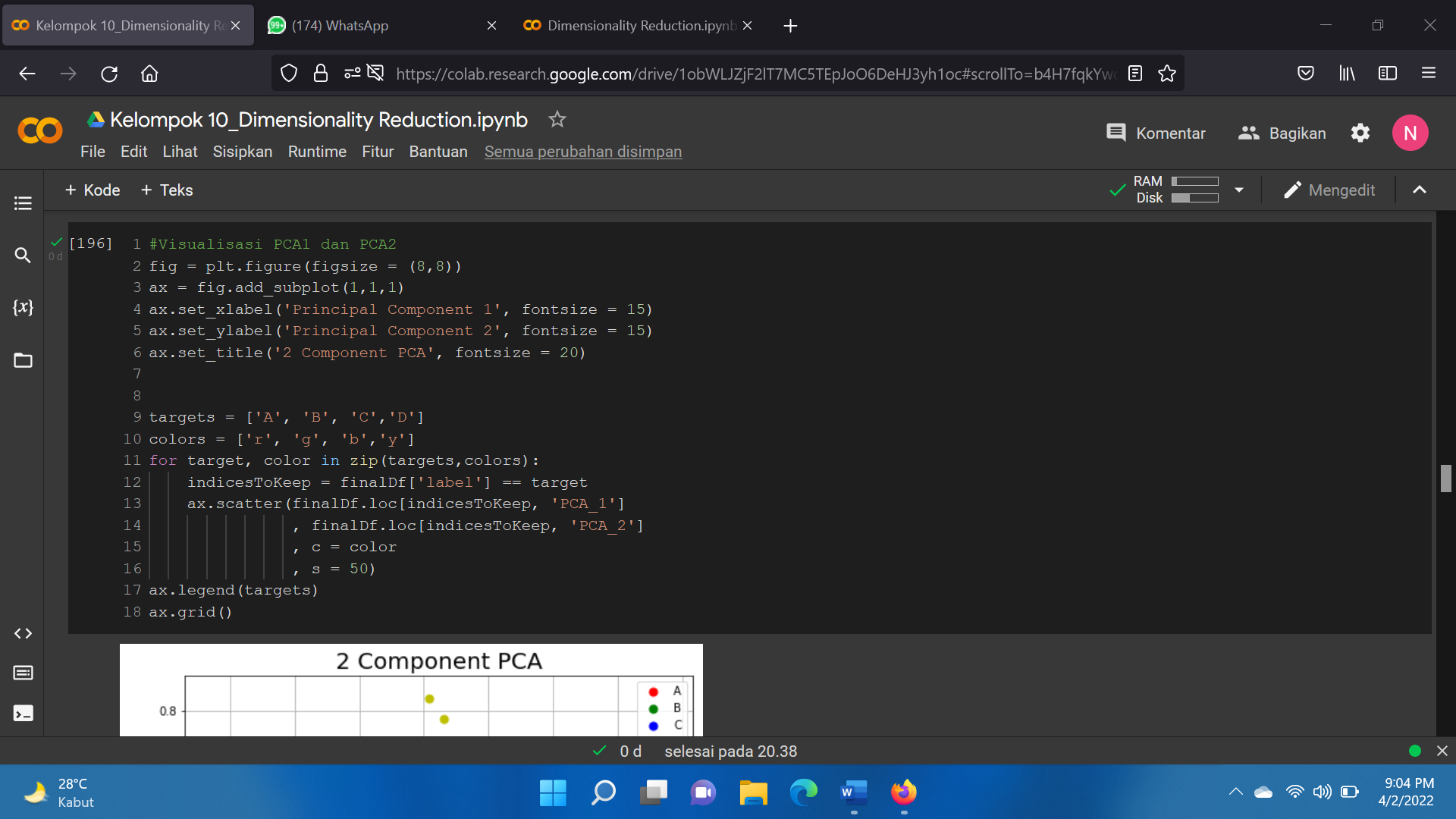
Berikut tampilan 5 baris terakhir DataFrame baru:



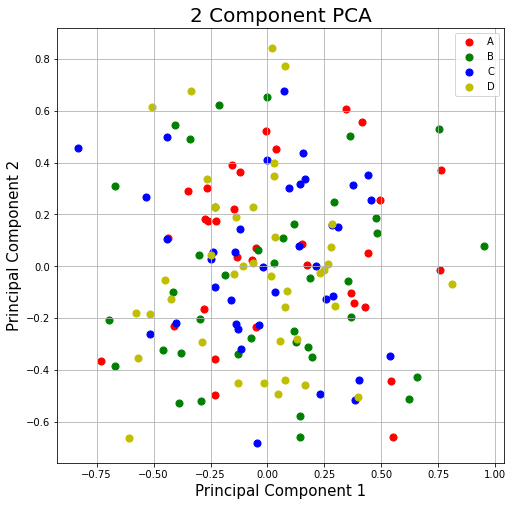
1. Menampilkan 5 baris pertama final DataFrame yang isinya meliputi nilai PCA1, PCA2, dan label.



1. Membuat visualisasi PCA1 dan PCA2



Berikut running code untuk visualisasi PCA1 dan PCA2:



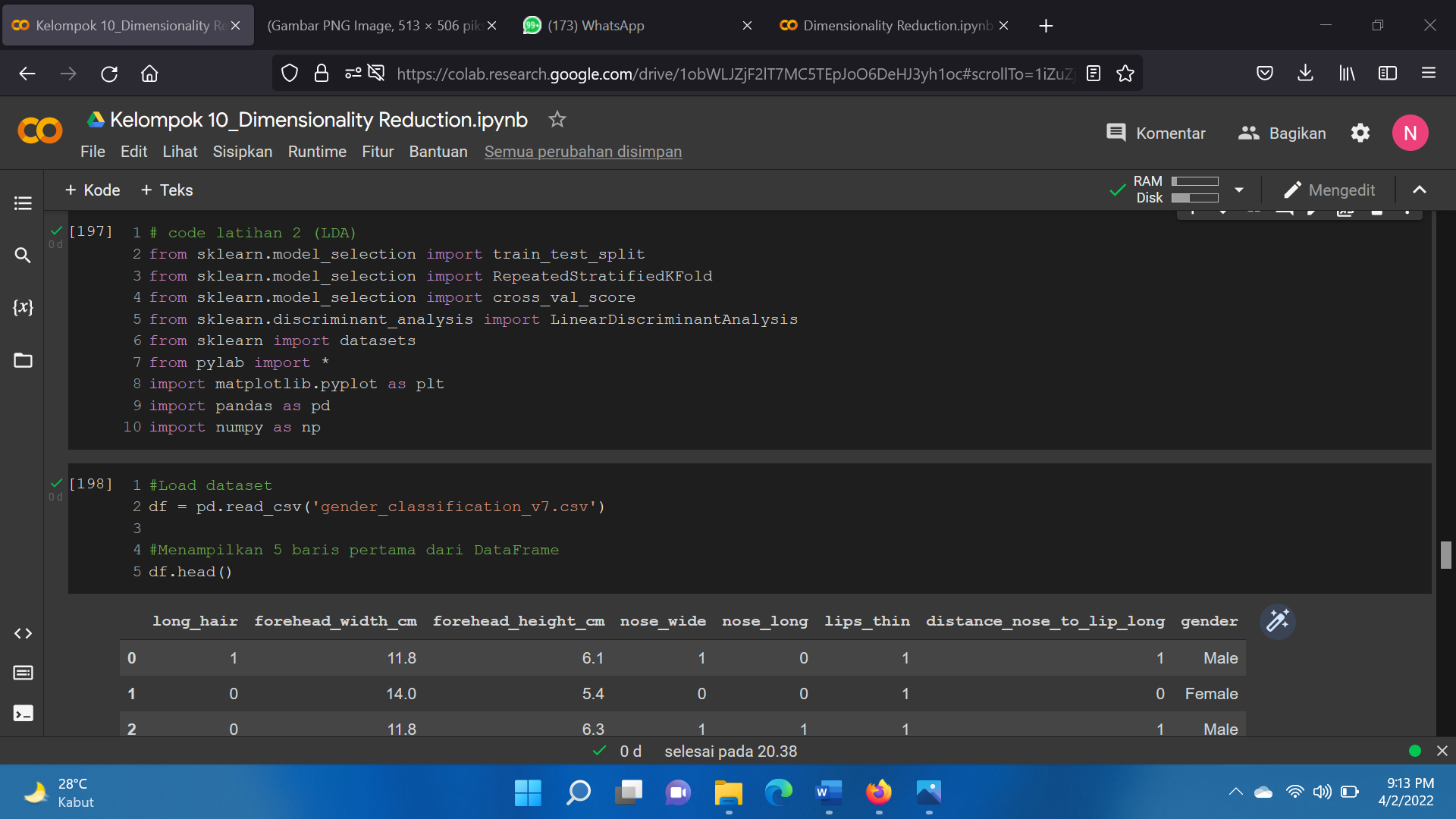
Dari visualisasi di atas, plot berwarna merah untuk label A, plot berwarna hijau untuk label B, plot berwarna biru untuk label C, plot berwarna cream untuk label berwarna D.

**LATIHAN 2:** Membuat model data reduction dengan menggunakan LDA (Linear Discriminant Analysis) dengan ketentuan sebagai berikut:

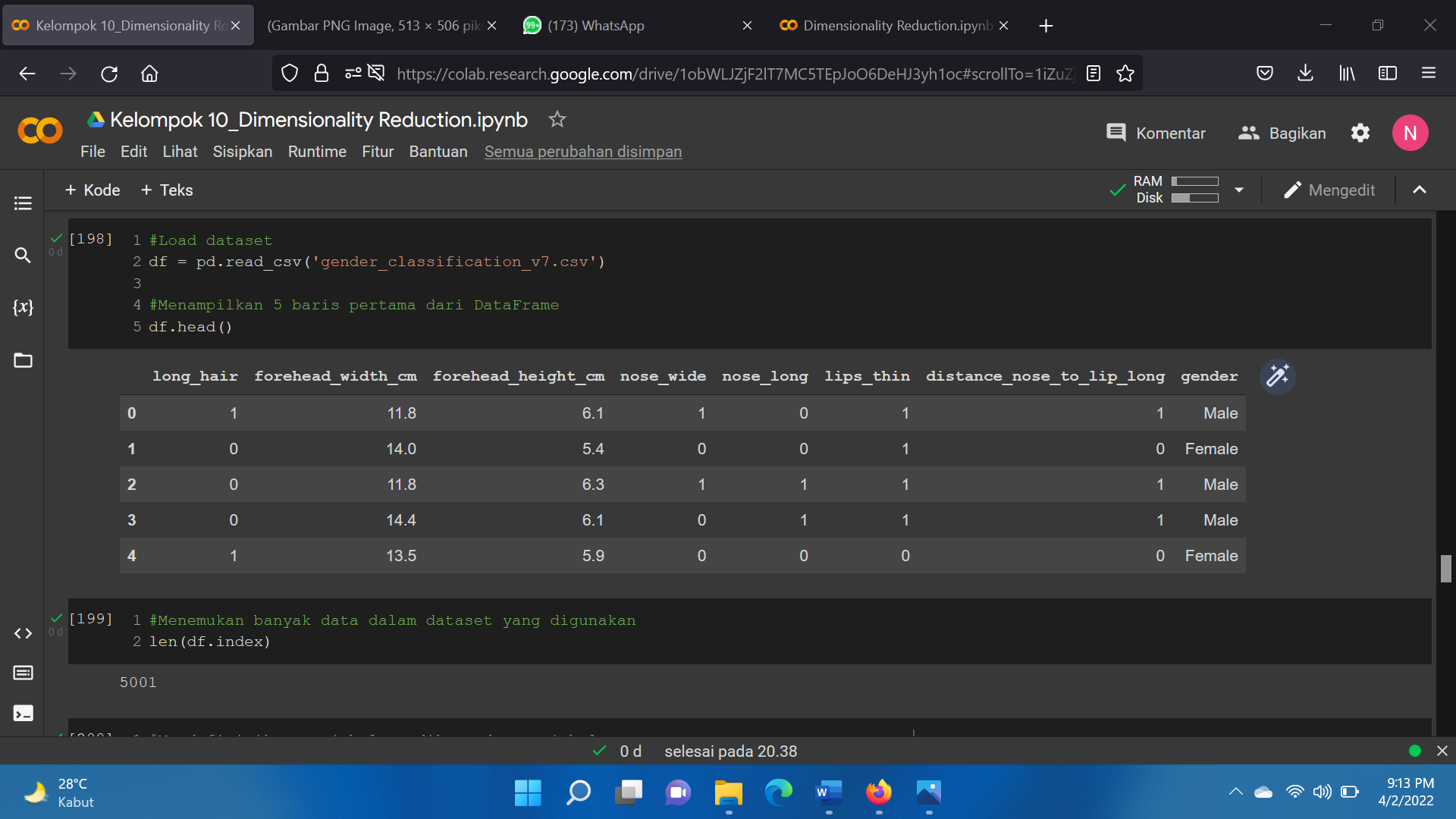
1. Menggunakan dataset yang berbeda dari latihan yang kemarin.
2. Tampilkan DataFrame untuk variabel x dan y (data tabular). Kemudian, lanjutkan proses reduksi data model LDA sampai dengan visualisasi data dalam bentuk plot.

**Penyelesaian:**

1. Mengimport library yang akan digunakan

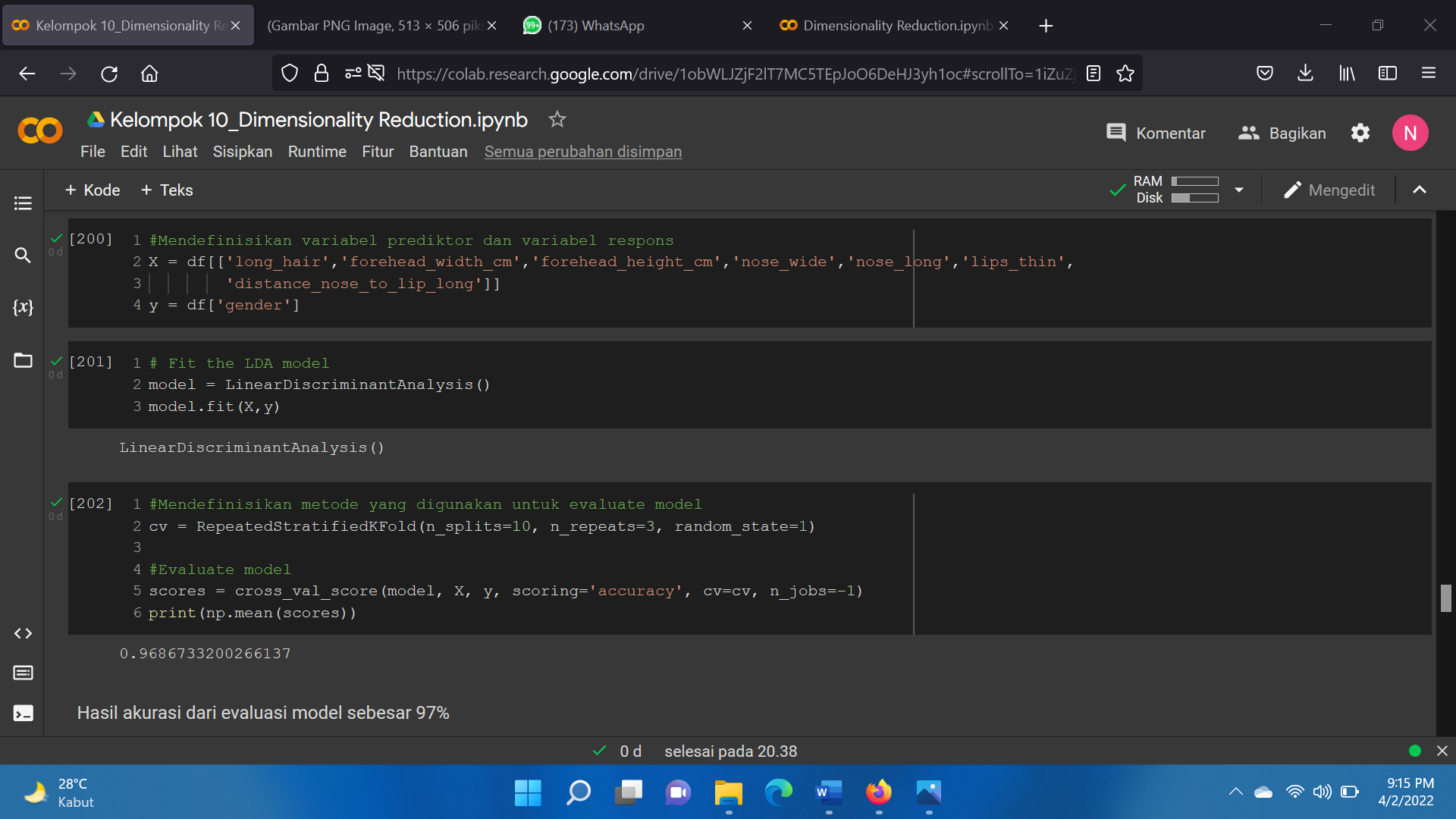


1. Membaca dataset yang akan digunakan ke dalam bentuk DataFrame, kemudian menampilkan 5 baris pertama dari DataFrame. Kelompok 10 menggunakan dataset pengklasifikasian gender. Selain itu, menemukan banyak data dalam dataset yang digunakan.



Dari hasil running code di atas, terdapat 5001 data pada dataset yang digunakan.

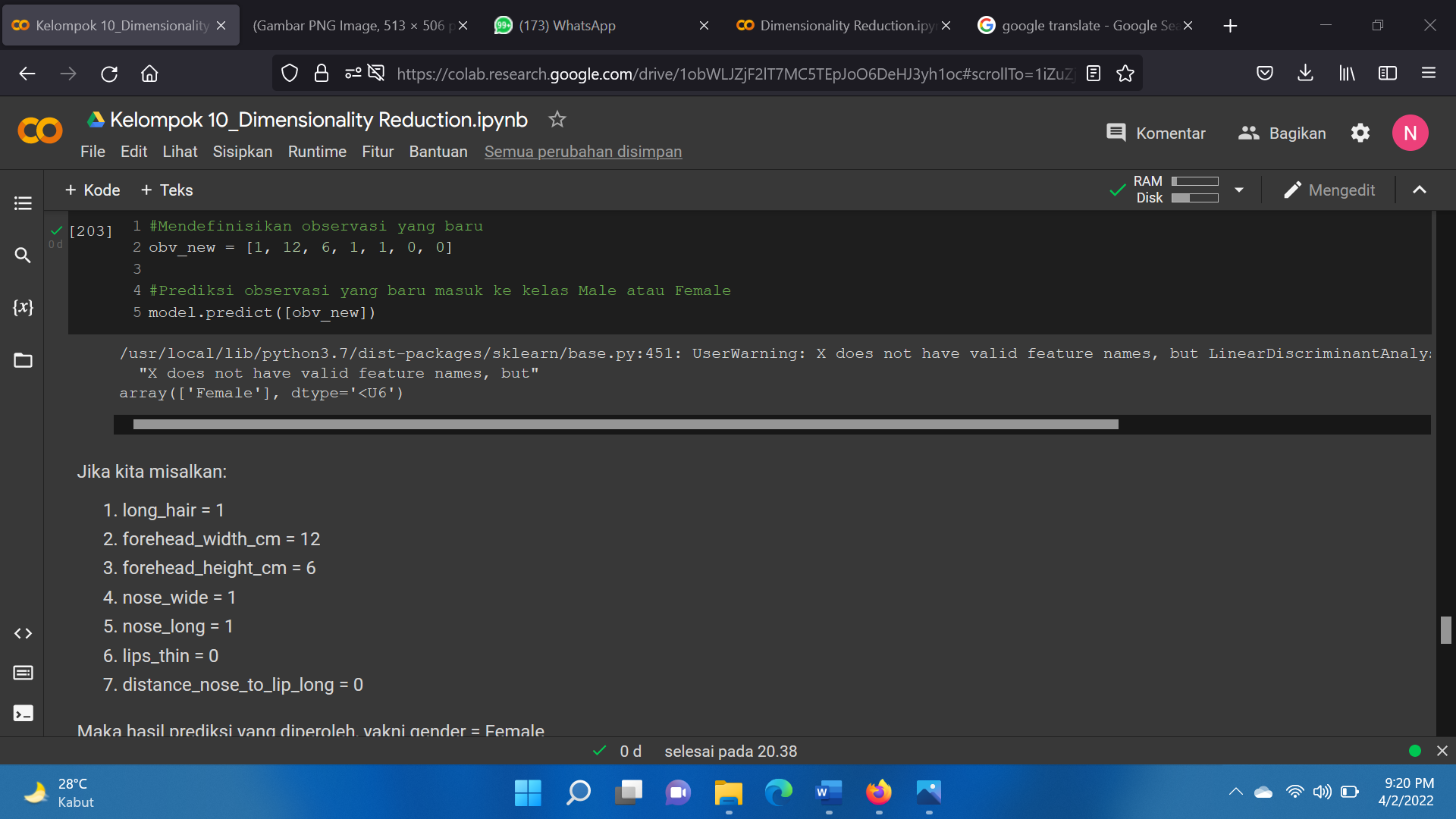
1. Mendefinisikan variabel predictor dan variabel response. Fit model LDA dari variabel X dan y. Kemudian mendefinisikan metode yang digunakan untuk meng-evaluate model lalu dilakukan evaluate model sehingga diperoleh nilai akurasi sebesar 97%.



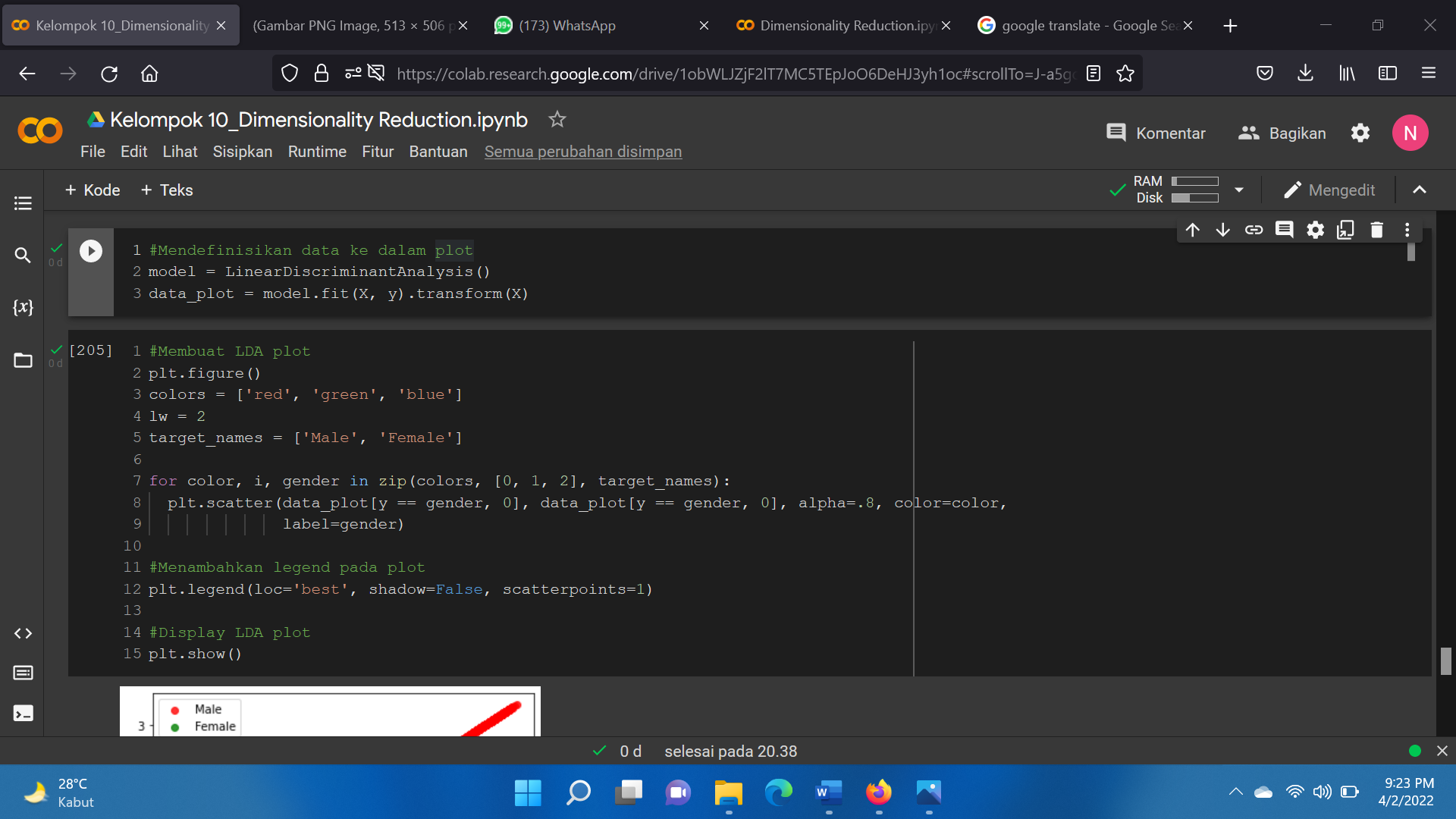
1. Mendefinisikan observasi yang baru lalu melakukan prediksi dari observasi yang baru. Jika dimisalkan:

* long\_hair = 1
* forehead\_width\_cm = 12
* forehead\_height\_cm = 6
* nose\_wide = 1
* nose\_long = 1
* lips\_thin = 0
* distance\_nose\_to\_lip\_long = 0

Diperoleh hasil prediksinya, yakni Female



1. Mendefinisikan data ke dalam plot, lalu membuat visualisasi LDA plot. Lalu, menambahkan legend pada plot dan menampilkan LDA plot yang sudah dibentuk.



Berikut hasil visualisasi LDA plot:

