**Sistem Rekomendasi Lagu Teratas – Global Berdasarkan Properti Trek menggunakan Metode Content Based Filtering**

|  |  |
| --- | --- |
| Rolly Maulana Awangga | Naomi C.H Tampubolon’ |
| *Jurusan Teknik Informatika* | *Jurusan Teknik Informatika* |
| *Politeknik Pos Indonesia* | *Politeknik Pos Indonesia* |
| Bandung, Indonesia | Bandung, Indonesia |
| awangga@poltekpos.ac.id | naomicindi.hermina@gmail.com |

Spotify saat ini merupakan platform musik, podcast dan video yang banyak digemari banyak penggunanya. Spotify juga memberikan kemudahan akses musik. Platform musik spotify sangat banyak digemari penggunanya, terutama kalangan young adults yaitu berusia 18 hingga 25 tahun.(Wahyu Dewatara & Monik Agustin, 2019) Spotify menyediakan banyak fitur seperti yaitu sekarang memutar, cari, koleksi, dengarkan offline, podcast dan acara serta banyak lagi. Properti trek di Spotify termasuk genre, nama artis, dan banyak lainnya. (Fitriani, n.d.) Berdasarkan properti trek, bagaimana cara merekomendasikan lagu berdasarkan properti trek, dan bagaimana cara memastikan bahwa lagu berdasarkan properti trek disukai. Masalah ini akan menjadi sumber dalam penelitian. *Dataset* yang digunakan dalam penelitian ini adalah 50 Lagu Teratas – Global (*50 Top Songs - Global*) yang diambil dari Spotify playlist chart yang kemudian diconvert ke dalam dalam *file* dengan *format* *cvs* dan kemudian nantinya akan dikelola atau diproses. Penelitian ini bertujuan untuk rekomendasikan lagu berdasarkan trek properti yang ada dan rekomendasikan berdasarkan trek properti disukai. Metode Content-Based Filtering digunakan dalam penelitian ini. *Content-Based Filtering* berfungsi merekomendasikan berdasarkan kemiripan konten seperti *genres* yang merupakan kemiripan jarak *cosinusnya* dan terdapat 5 lagu teratas yaitu *STAYING ALIVE (feat. Drake & Lil Baby), Bam Bam (feat. Ed Sheeran), About Damn Time, Bad Decisions (with BTS & Snoop Dogg)*, dan *Ferrari* yang direkomendasikan menggunakan metode *Content-Based Filtering*.

Kata Kunci: Sistem rekomendasi, musik, spotify, content-based filtering.

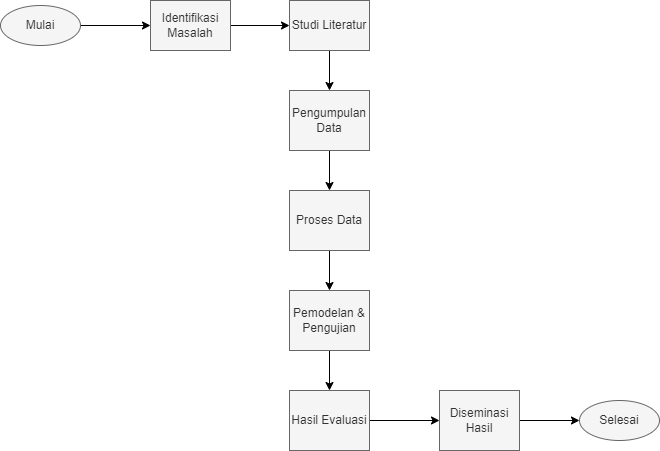
# Pendahuluan

Musik terdiri dari beberapa unsur, seperti melodi, harmoni, ritme, dan timbre. Musik adalah fenomena *intuitif* yang dapat dikembangkan, ditingkatkan, dan disebarluaskan dengan berbagai cara. Musik adalah fenomena unik yang dihasilkan oleh berbagai alat musik. Musik memang menjadi salah satu hobi yang paling digemari, hal tersebut terbukti dengan manfaat yang bisa didapatkan dengan mendengarkan musik. Musik memiliki kemampuan untuk mengurangi stres. Manfaat mendengarkan musik adalah kita bisa mendengarkannya sambil melakukan hal lain. Musik adalah komponen yang tak terpisahkan dari kehidupan dan teknologi. Perkembangan yang semakin berkembang pesat juga mempengaruhi dunia industri.(Muslimin M, n.d.) Dimana pada jaman era digital saat ini tidak perlu repot dalam mengengarkan musik dimanapun dan kapanpun, dikarenakan kecanggihan teknologi tersebut. *Spotify* adalah layanan *streaming* musik terpopuler di Indonesia, telah mencapai 200 juta pengguna perbulan.(Noviani et al., 2020) Data yang digunakan pada penelitian ini adalah data lagu yang diambil dari *playlist* *chart* *spotify* 50 Lagu Teratas – Global (*50 Top Songs - Global*).(Fitriani, n.d.) *Spotify* memiliki properti lagu yang berfungsi untuk mengetahui lebih mendalam mengenai lagu atau sebuah musik. Ada beberapa fitur atau properti lagu pada *spotify* yaitu, sebagai berikut: *track name, album name, artist name, duration, popularity, genres, danceability, energi, key, loudness, speechiness, acousticness, instrumentaness, liveness, tempo* dan *valence*. Berdasarkan properti trek tersebut kita dibutuhkan sitem yang dapat merekomendasikan berdasarkan properti trek yang ada. Maka dari itu bagaimana caranya untuk merekomendasikan sebuah lagu berdasarkan properti trek yang ada. Dari permasalahan tersebut dibutuhkan sebuah sistem yang dapat merekomendasikan lagu berdasarkan properti trek yang ada. Properti trek ‘*genres’* merupakan properti trek yang digunakan dalam rekomendasi sistem.

Dalam penelitian ini sistem rekomendasi akan dianalisa dan dibangun dengan menggunakan metode *content-based filtering* karena dengan algoritma ini berdasarkan *similarity* konten sehingga rekomendasi yang dihasilkan mempunyai kualitas yang baik. Visualisasi hasil rekomendasi tersebut akan diimplementasikan dengan *framework* *Jupyter Notebook*. Target luaran yang ingin dicapai pada penelitian ini adalah jurnal nasional terakreditas.

# Metodologi Penelitian

Metode penelitian yang digunakan dalam penelitian ini akan diwakilkan oleh flowchart pada Gambar 1 sebagai berikut ini :



Gambar 1. Diagram Alur Metodologi Penelitian

## Identifikasi Masalah

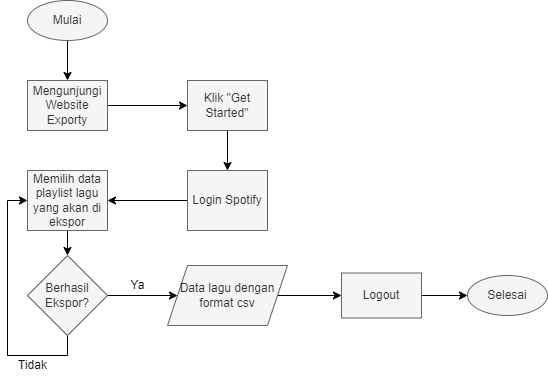
Identifikasi masalah merupakan langkah pertama dalam tahapan metode penelitian yang berfungsi untuk mengidentifikasi masalah. Pada tahap ini peneliti melakukan penelitian dengan cara memnukan masalah (*problem*) yang ada, kemudian melakukan indentifikasi sumber permasalahan (*root* *cause*) sehingga menciptakan permasalahan (*problem* *statement*) yang menjelaskan permasalahan yang sudah diidentifikasi. Adapun Permasalahan yang terdapat pada penelitian ini adalah bagaimana cara merekomendasikan lagu berdasarkan properti trek.

## Studi Literatur

Setelah dilakukan tahap indentifikasi masalah langkah selanjutnya dalam tahapan metode penelitian adalah studi literatur. Pada tahap ini peneliti melakukan pencarian informasi yang relevan dengan masalah yang akan diteliti dan mengkaji beberapa teori dasar yang relevan dengan permsalaha yang akan diteliti kemudian memperdalam pengetahuan peneliti tentang permasalahan yang akan diteliti dan mengkaji hasil-hasil penelitian terdahulu yang berkaitan dengan penelitian yang akan diteliti dan mendapatkan informasi tentang aspek-aspek mana dari suatu permasalahan yang sudah pernah diteliti untuk menghindari agar tidak terjadi *plagiarisme* penelitian.

## Pengumpulan Data

Pengumpulan data adalah tahap setelah studi literatur. Tujuan pengumpulan data adalah untuk memperoleh data yang akan digunakan dalam penelitian. Peneliti melakukan pengumpulan studi literatur yang relevansi dan sesuai dengan apa yang dibutuhkan untuk menunjang penelitian untuk dapat melakukan teknik pengumpulan data studi literatur. Maka dari itu data yang digunakan pada penelitian ini adalah data yang di *convert* dari *website* [*https://exportify.net/#playlists*](https://exportify.net/#playlists) kedalam bentuk *format* *csv*. W*ebsite* tersebut berfungsi untuk mengekspor daftar putar *spotify* menggunakan *Web API*. Alur dari pengumpulan data dapat dilihat dari *flowchart* berikut:



Penjelasan *flowchart,* dimulai dengan cara mengunjungi *website* *Exportify* kemudian akan muncul tampilan awal dari *website*. Klik “*Get Started*”, kemudian akan dialihkan ke halaman *login* dan melakukan proses *login*. Setelah masuk *login* selanjutnya akan di tampilkan daftar *spotify* yang sudah ditambahkan kedalam koleksi *playlist* pada spotify. Kemudian *klik* *export* lalu *playlist* yang berisi lagu yang ingin diambil datanya akan terekspor menjadi data yang berbetuk *file csv*. kemudian *logout* setelah selesai melakukan *export* data.

## Proses Data

Setelah dilakukan pengumpulan data langkah selanjutnya dalam tahap metodologi penelitian adalah proses data. Pada tahap proses data atau pengolahan data dilakukan data *cleaning*. *Cleaning* data adalah prosedur untuk memastikan keakuratan, konsistensi, dan kegunaan data dalam kumpulan data. Prosedurnya adalah mendeteksi kesalahan atau kerusakan pada data dan kemudian memperbaiki atau menghapusnya jika perlu.

## Pemodelan dan Pengujian

Tahapan selanjutnya yaitu Pemodelan dan Pengujian. Tahap ini ialah tahap menggunakan metode rekomendasi dalam pemodelan dan pengujian data. *Content-Based Filtering* merupakan metode, model atau algoritma yang akan digunakan pada penelitian ini. Pemodelan ada tahap yang dilakukan sebelum pengujian. Dalam pemodelan dipertimbangkan dulu model yang akan dipakai dalam mengelola data. Kemudian langkah selanjutnya setelah pemodelan adalah langkah Pengujian menggunakan model *Content-Based Filtering*.

## Hasil Evaluasi

Tahap hasil evaluasi adalah tahap performansi dari pemodelan dan pengujian. Hasil pengujian data akan diimplementasi pada tahap hasil evaluasi.

## Diseminasi Hasil

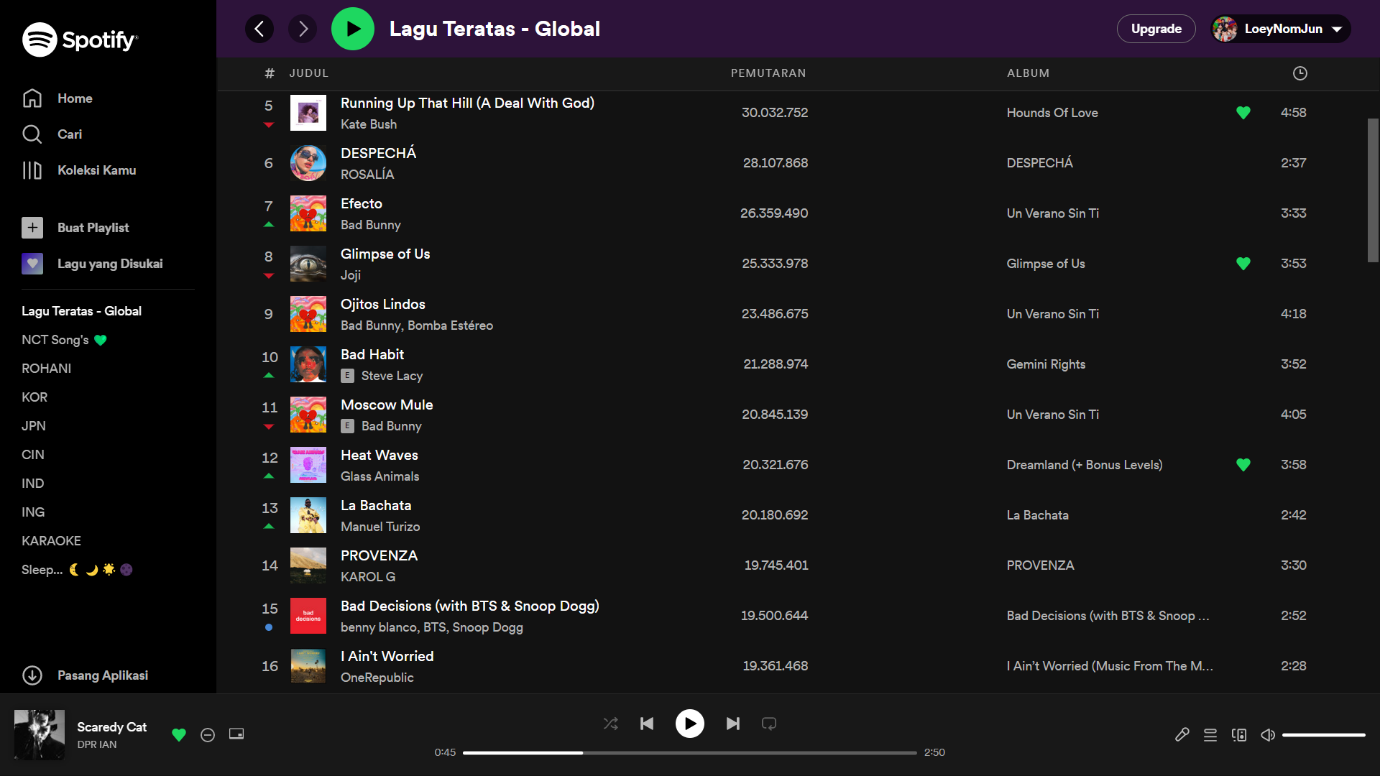
Tahap deseminasi hasil adalah tahap untuk artikel yang diterbitkan dalam laporan penelitian dan *draft* jurnal.

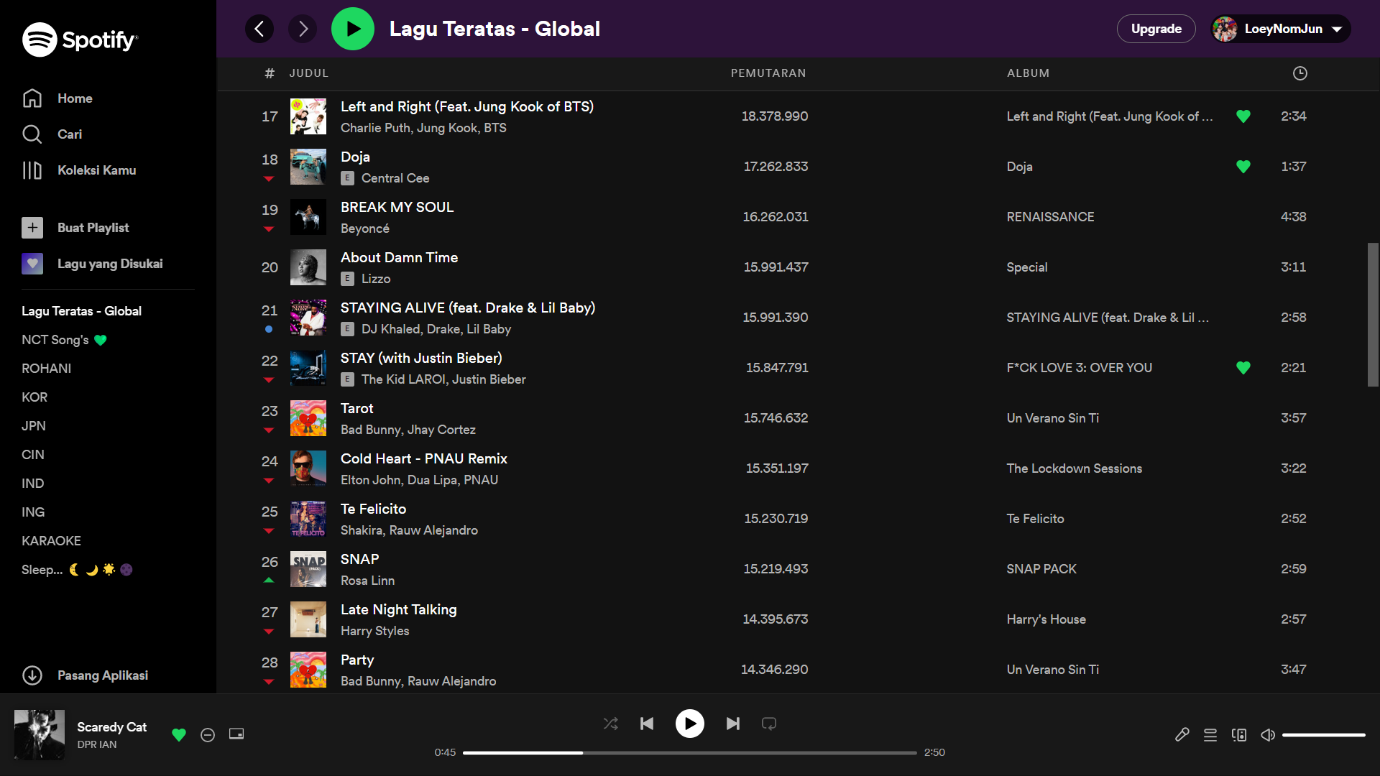
# Pembahasan dan hasil

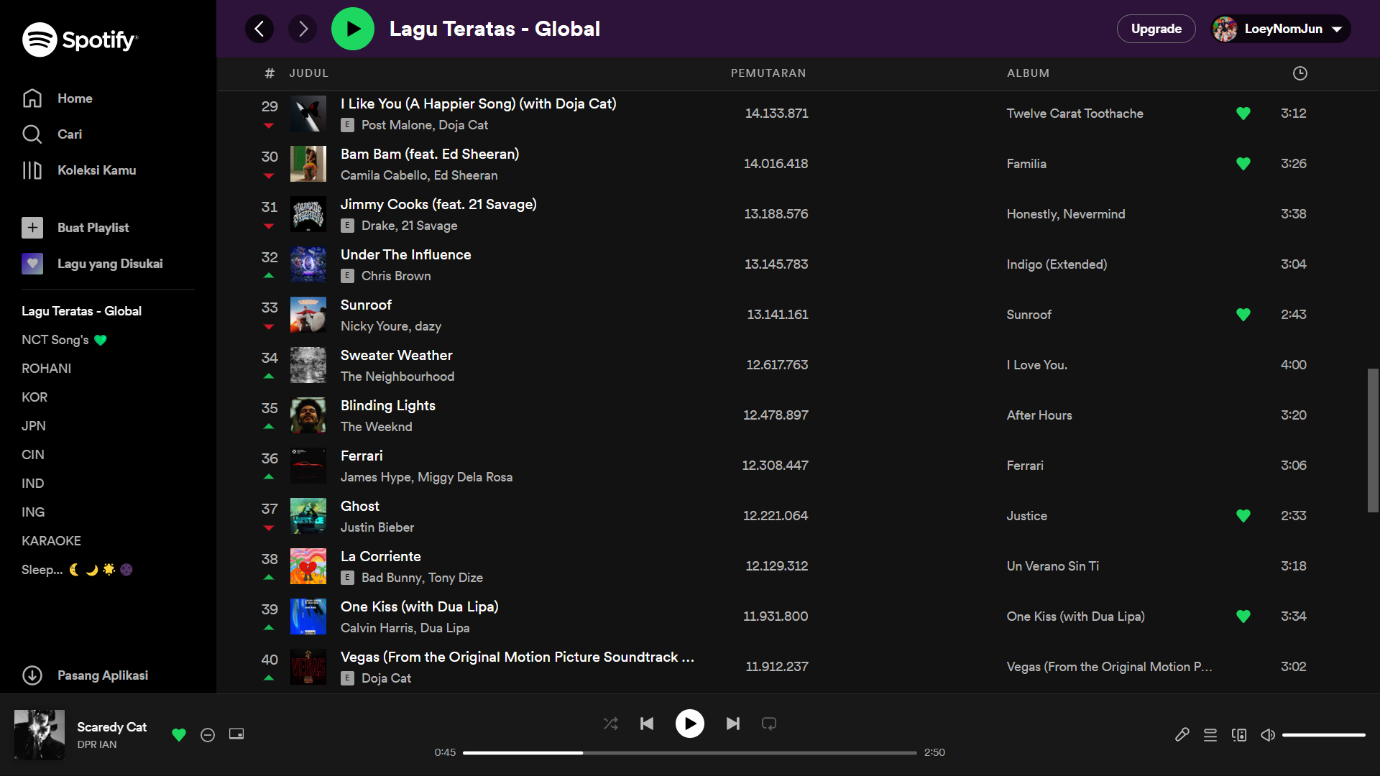
## PEMBAHASAN

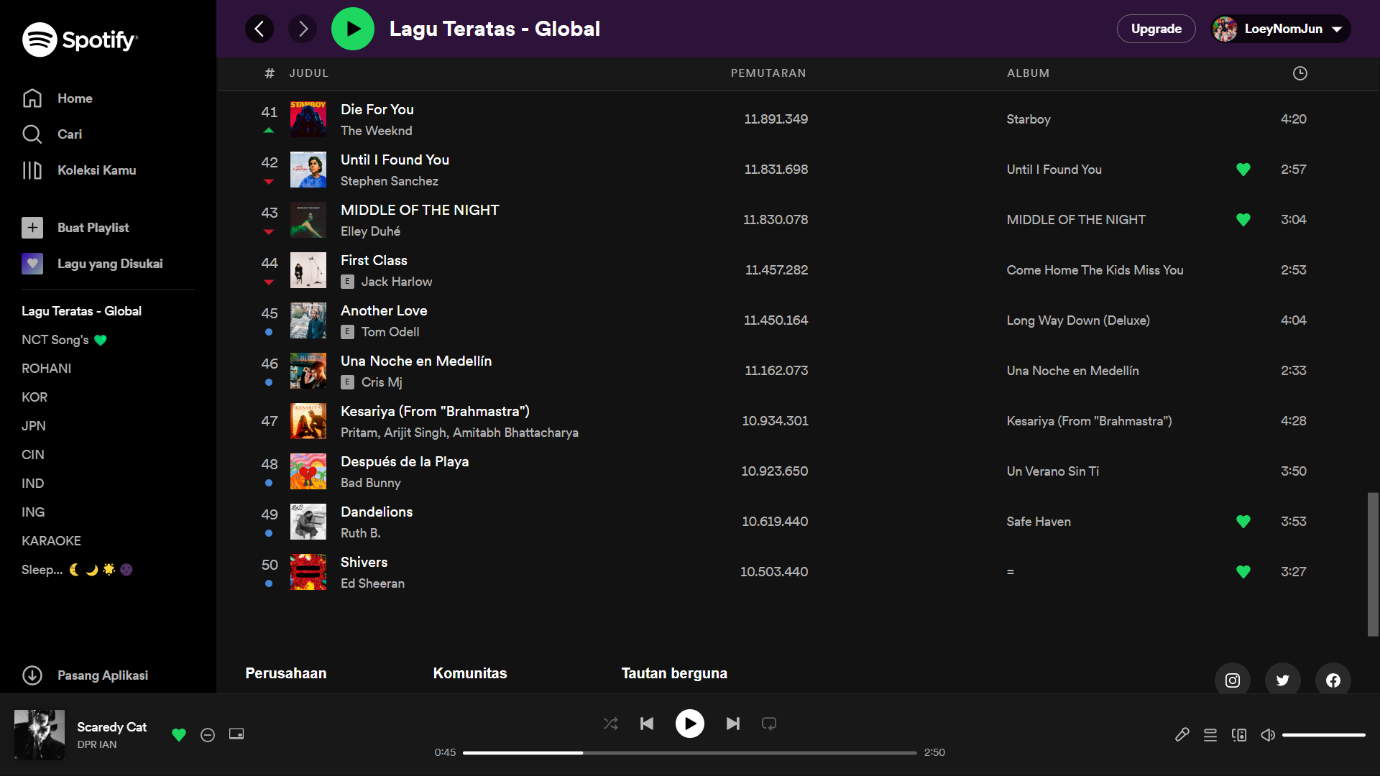
Penelitian ini dilakukan dengan mengambil data selama 1 minggu terakhir dari *playlist* 50 Lagu Teratas - Global (*50 Top Songs - Global*) sebanyak 50 data lagu. Untuk hasil data mentah atau data yang belum dilakukan proses data yaitu data yang ada pada *spotify*, dapat dilihat sebagai berikut:









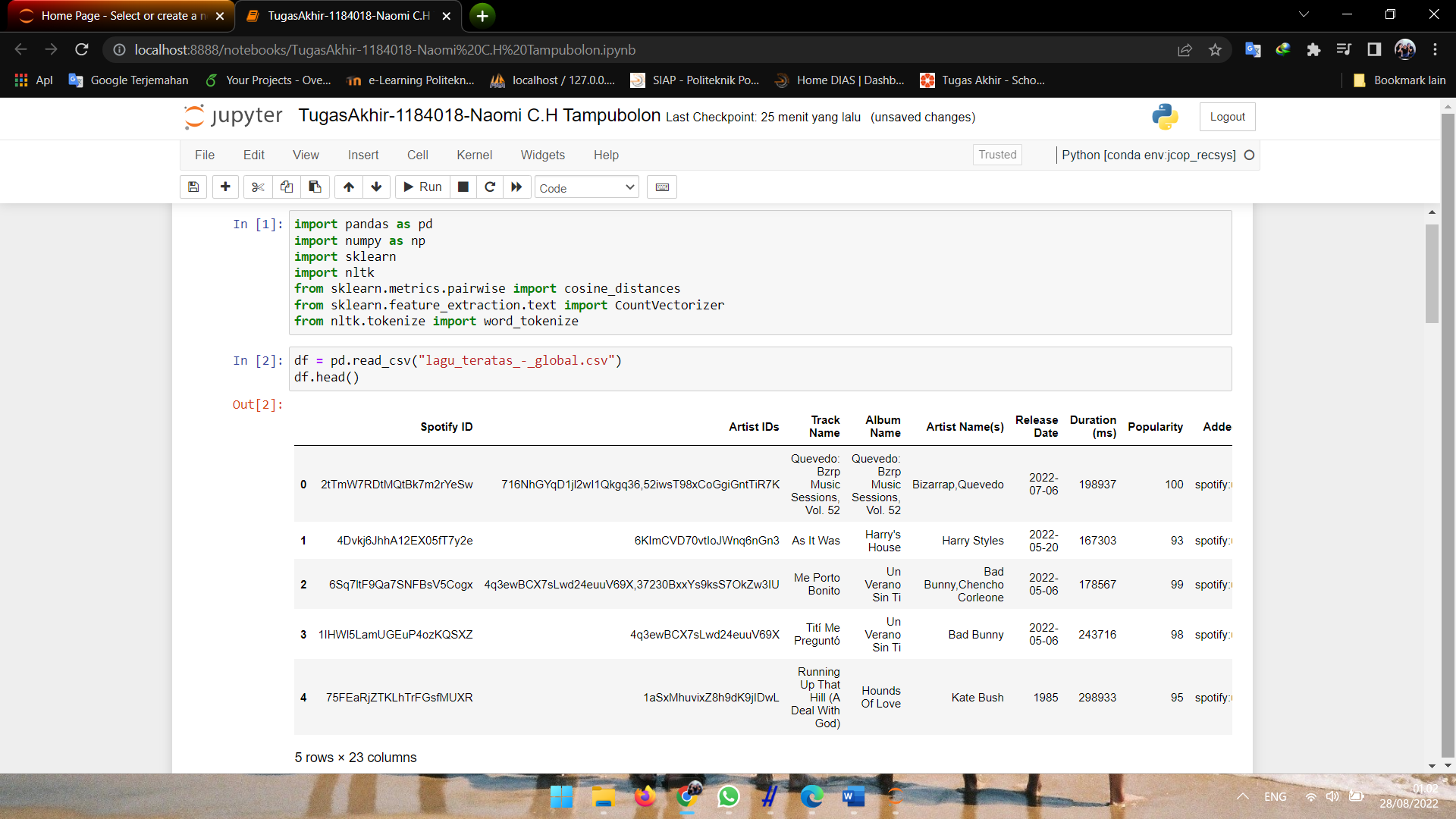


Kemudian data *diconvert* dalam *format* *file* *csv*.

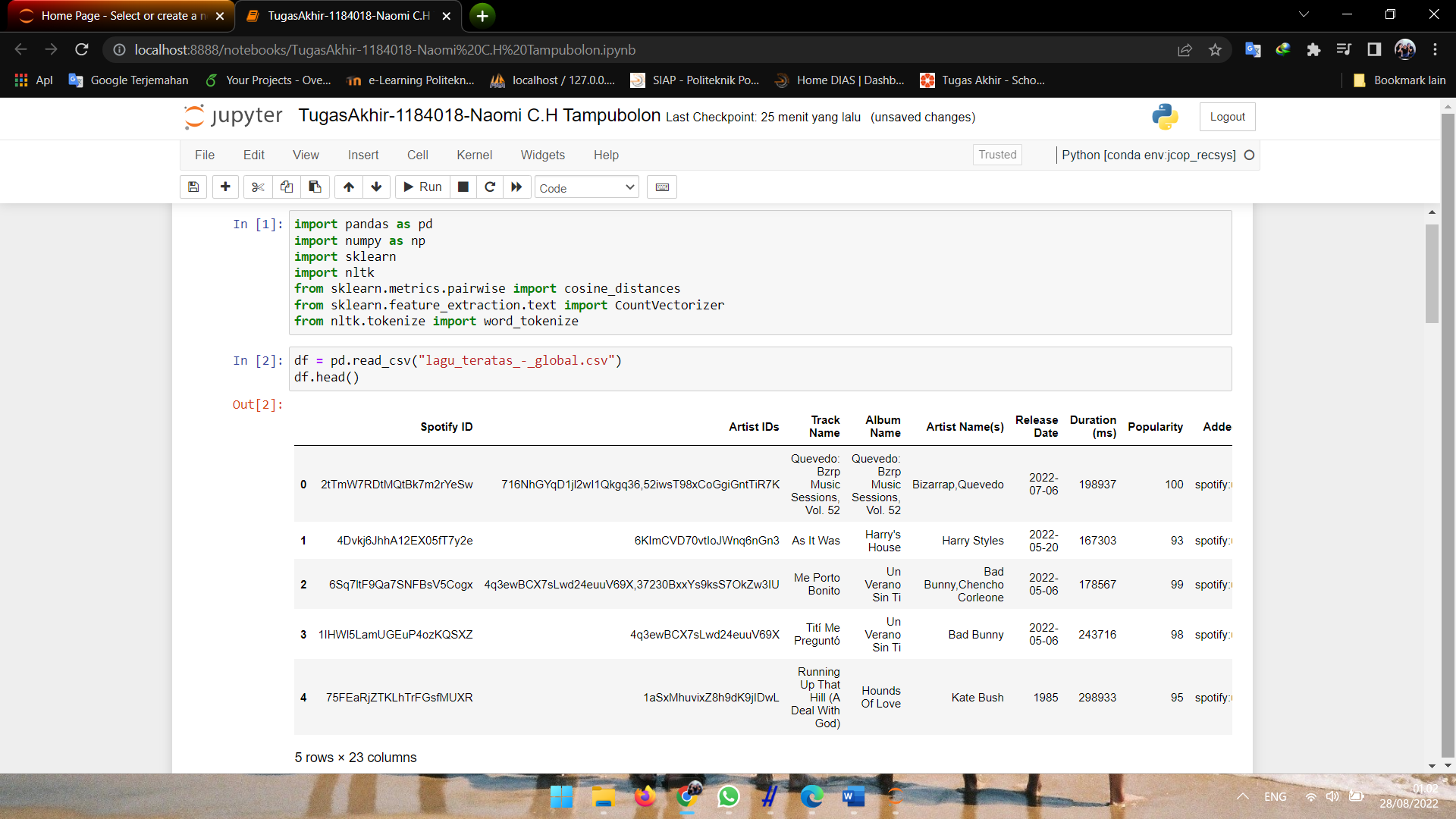


Proses selanjutnya dalam penelitian ini adalah proses pemodelan dan pengujian data. Penelitian ini menggunakan *content-based filtering* dalam pemodelannya. Langkah-langkah dalam pemodelan menggunakan *content-based filtering*.

1. *Import Library*



1. Menampilkan 5 dataset teratas.

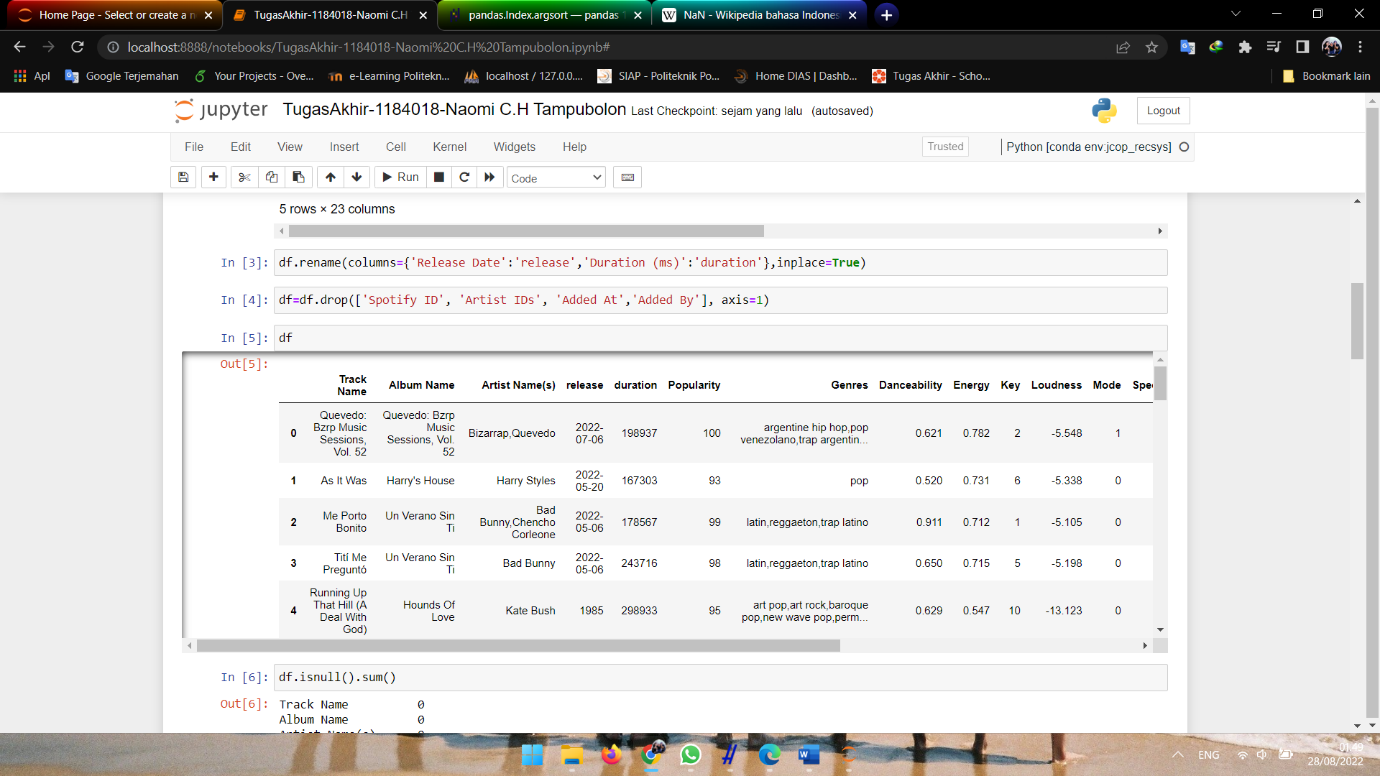


1. *Cleaning* Data

*Cleaning* data adalah prosedur untuk memastikan keakuratan, konsistensi, dan kegunaan data dalam kumpulan data. Prosedurnya adalah mendeteksi kesalahan atau kerusakan pada data dan kemudian memperbaiki atau menghapusnya jika perlu. Berikut adalah proses data yang telah dilakukan *cleaning* data:

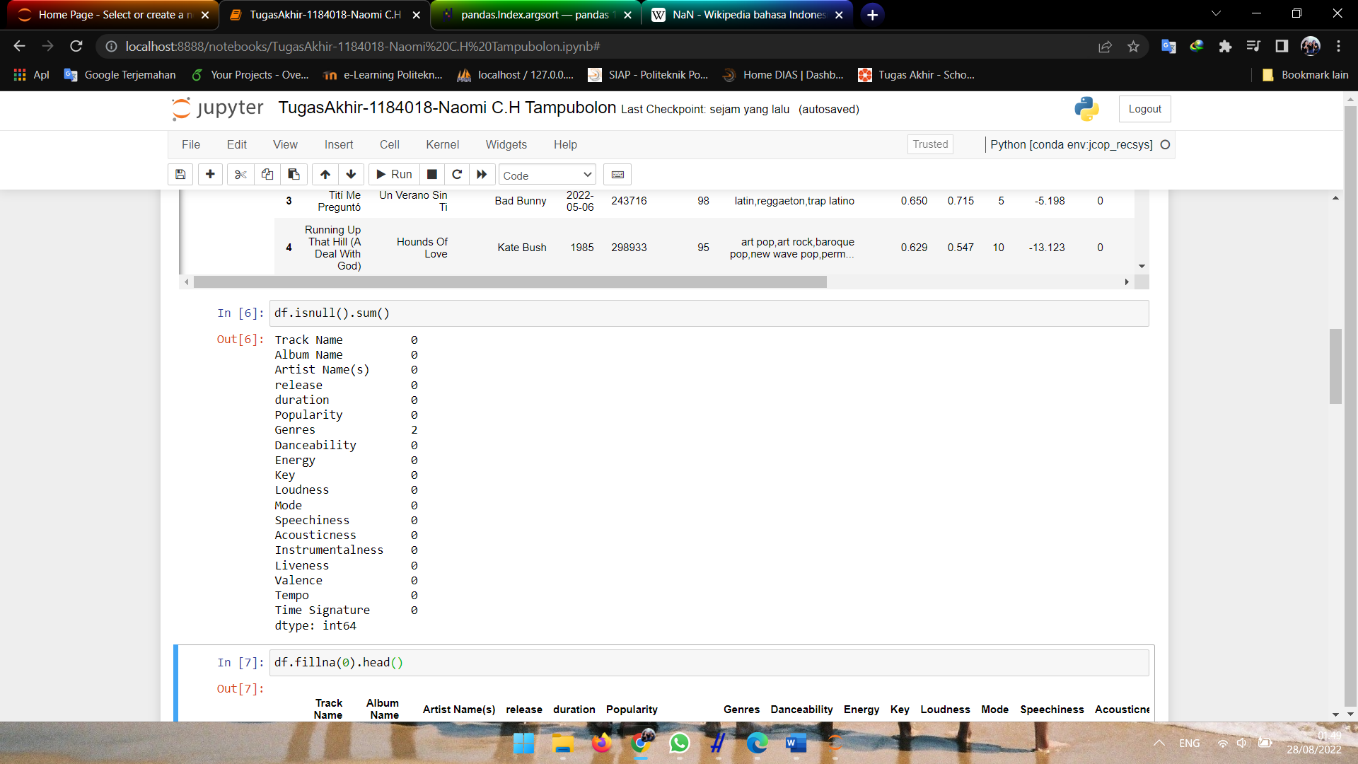
* *Drop* tabel

*Drop* tabel yang befungsi untuk menghilangkan tabel yang tidak digunakan atau diperlukan dalam proses penelitian. Terdapat beberapa data yang tidak diperlukan dalam penelitian ini, dan di lakukan *drop* tabel. Dapat dilihat sebagai berikut:

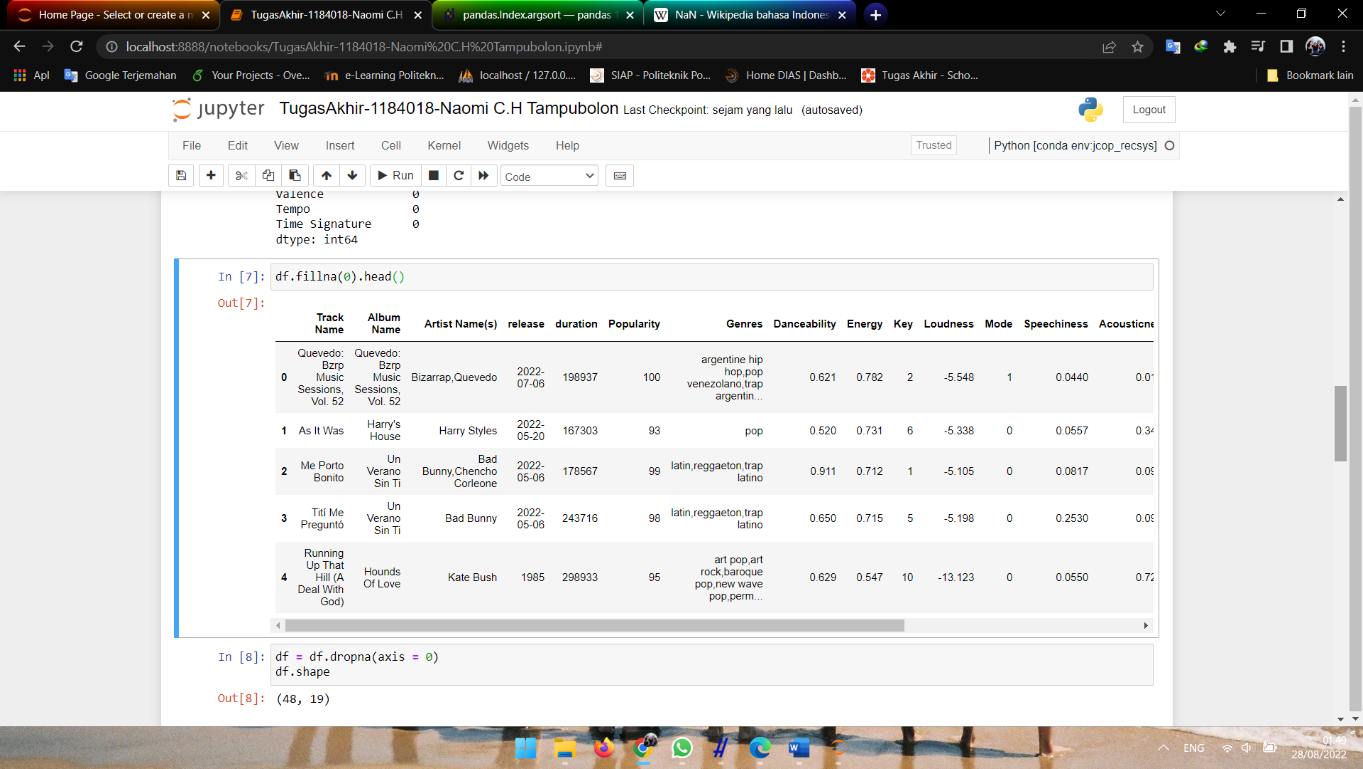


* Cek Data

Mengecek apakah data tersebut memiliki *missing* *value* atau tidak. Setelah dilakukan pengecekan data terdapat 2 data yang memiliki *missing* *value*:*Drop* tabel yang tidak digunakan.



Lalu dilakukan proses mengisi nilai *NA/NaN (Not a Number)* menjadi 0.



Kemudian menghapus atau menghilangkan semua entri dengan nilai *NaN* atau 0 seperti berikut:

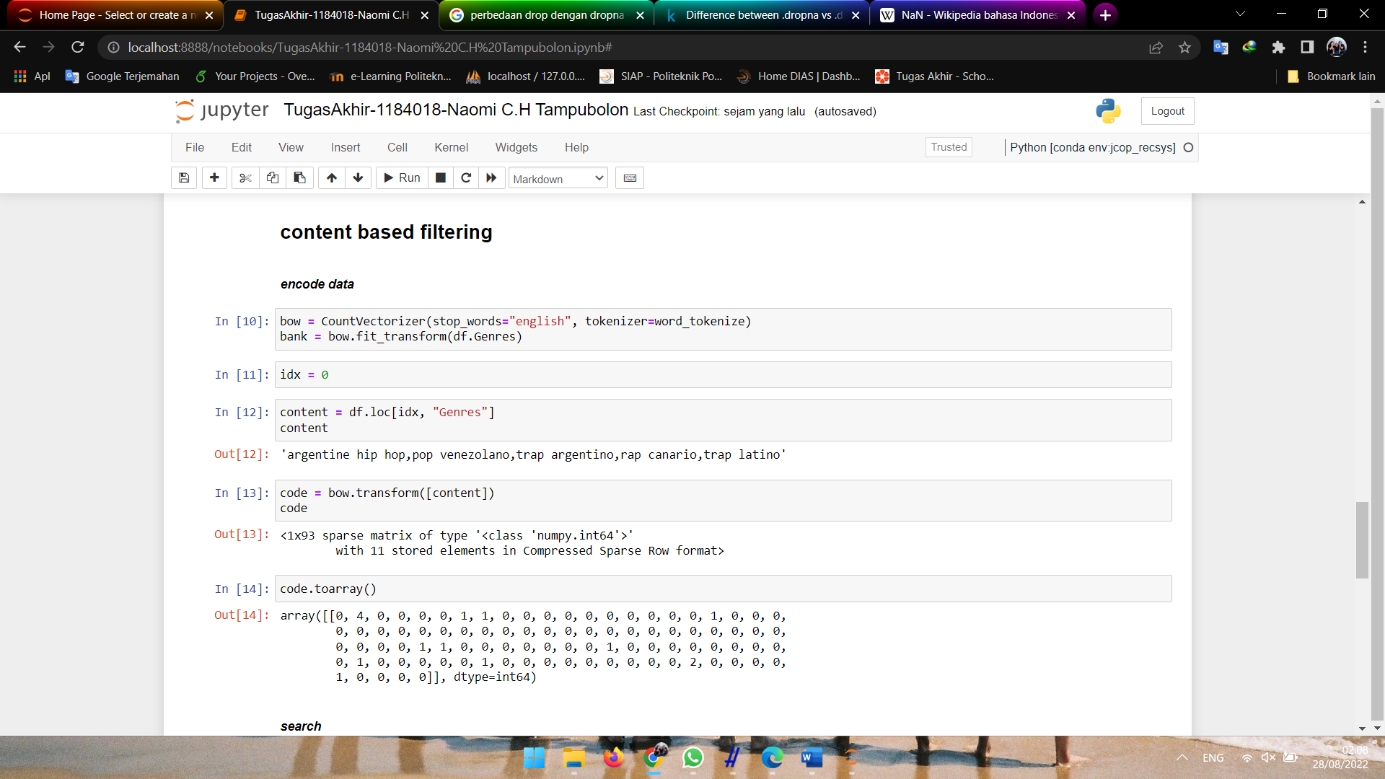


1. Pemodelan sistem rekomendasi berdasarkan konten.

Pada pemodelan menggunakan *content-based filtering* memiliki beberapa tahap dalam menampilkan rekomendasi berdasarkan konten. Dapat dilihat sebagai berikut:

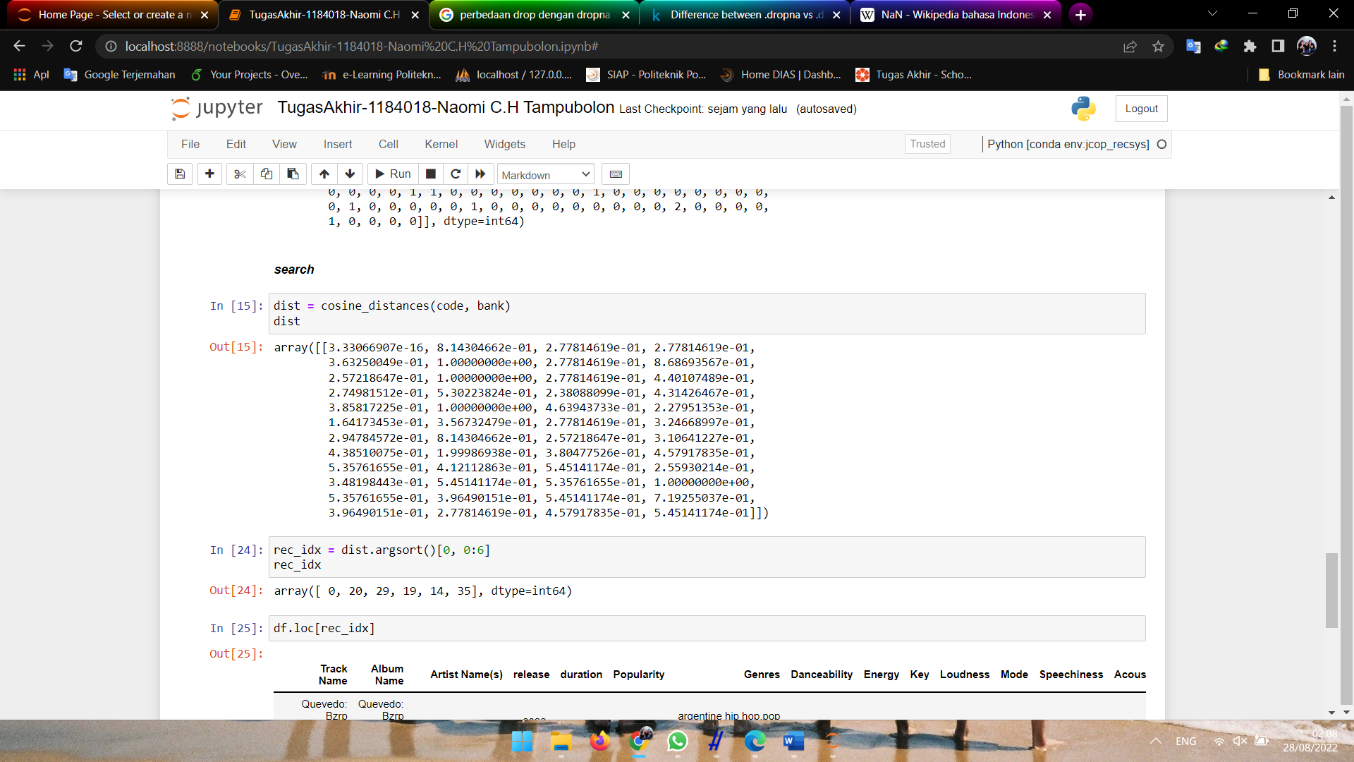
* *Encode Data*

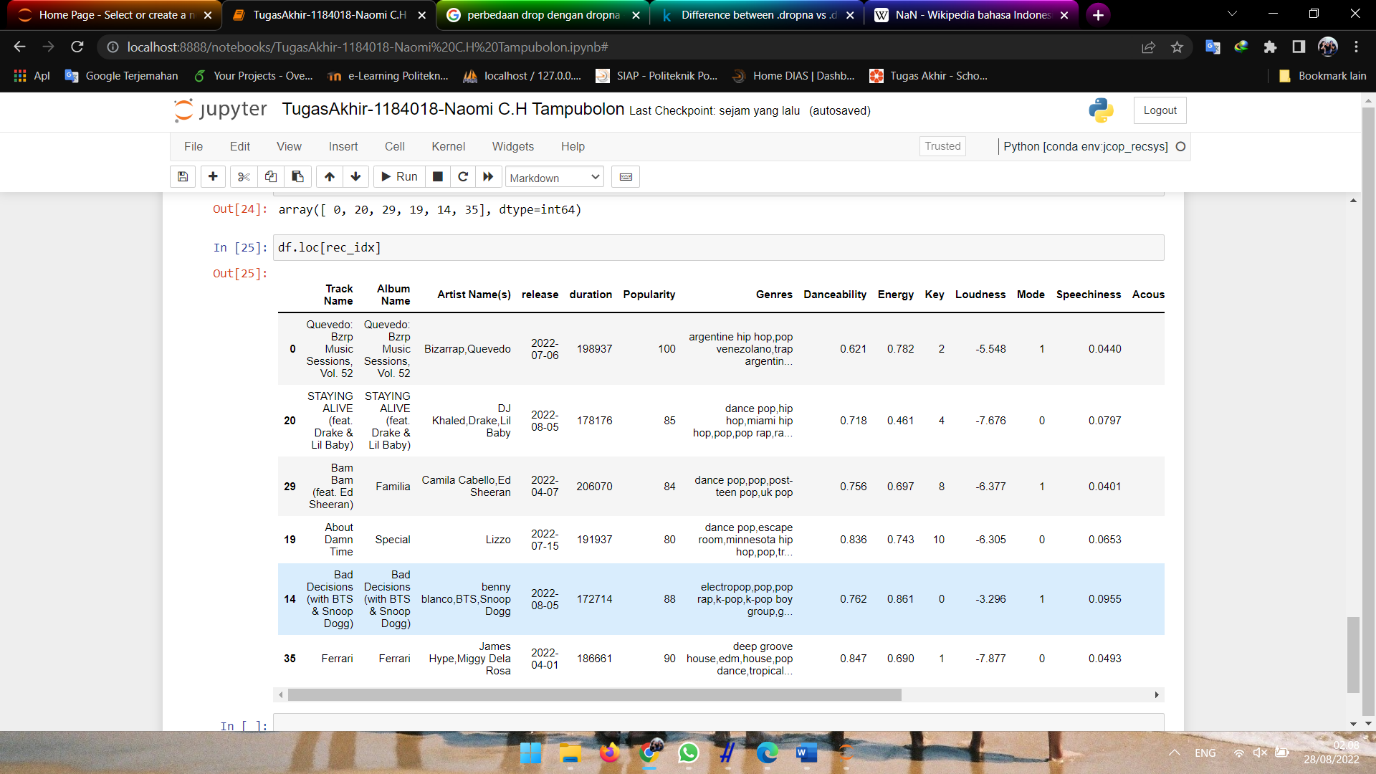
Pada tahap ini *genres* di *encode* dengan cara mengkonversi *teks* menjadi *vector* menggunakan *CountVectorizer* dengan konten yang digunakan dalam merekomendasikan yaitu kemiripan konten index 0 pada properti trek *genres*. Dengan langkah sebagai berikut:



* Mencari *similarity.*

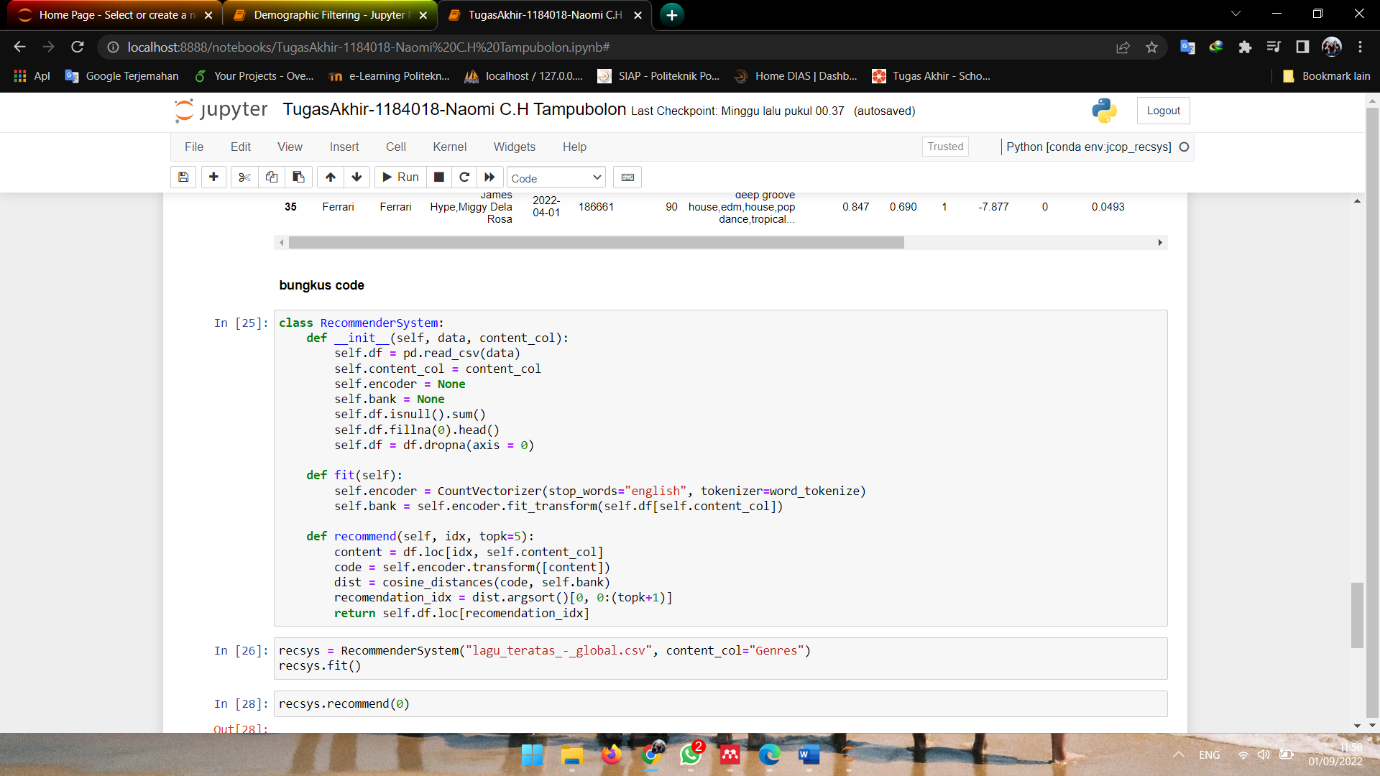
Pada tahap ini kita menggunakan *cosine distance* dalam mencari kesamaan atau kemiripan konten dengan cara menghitung jarak *cosinus* yaitu ukuran kesamaan antara dua vektor berdasarkan sudut kosinus di antara konten. (Usino et al., 2019) Kemudian mengembalikan nilai index bilangan bulat yang akan mengurutkan indeks. Dari hasil mengurutkan tersebut maka didapat hasil rekomendasi berdasarkan *index* 0 yaitu 5 lagu teratas yang mirip dengan konten yang sama dengan *index* 0 dan dapat dilihat sebagai berikut:

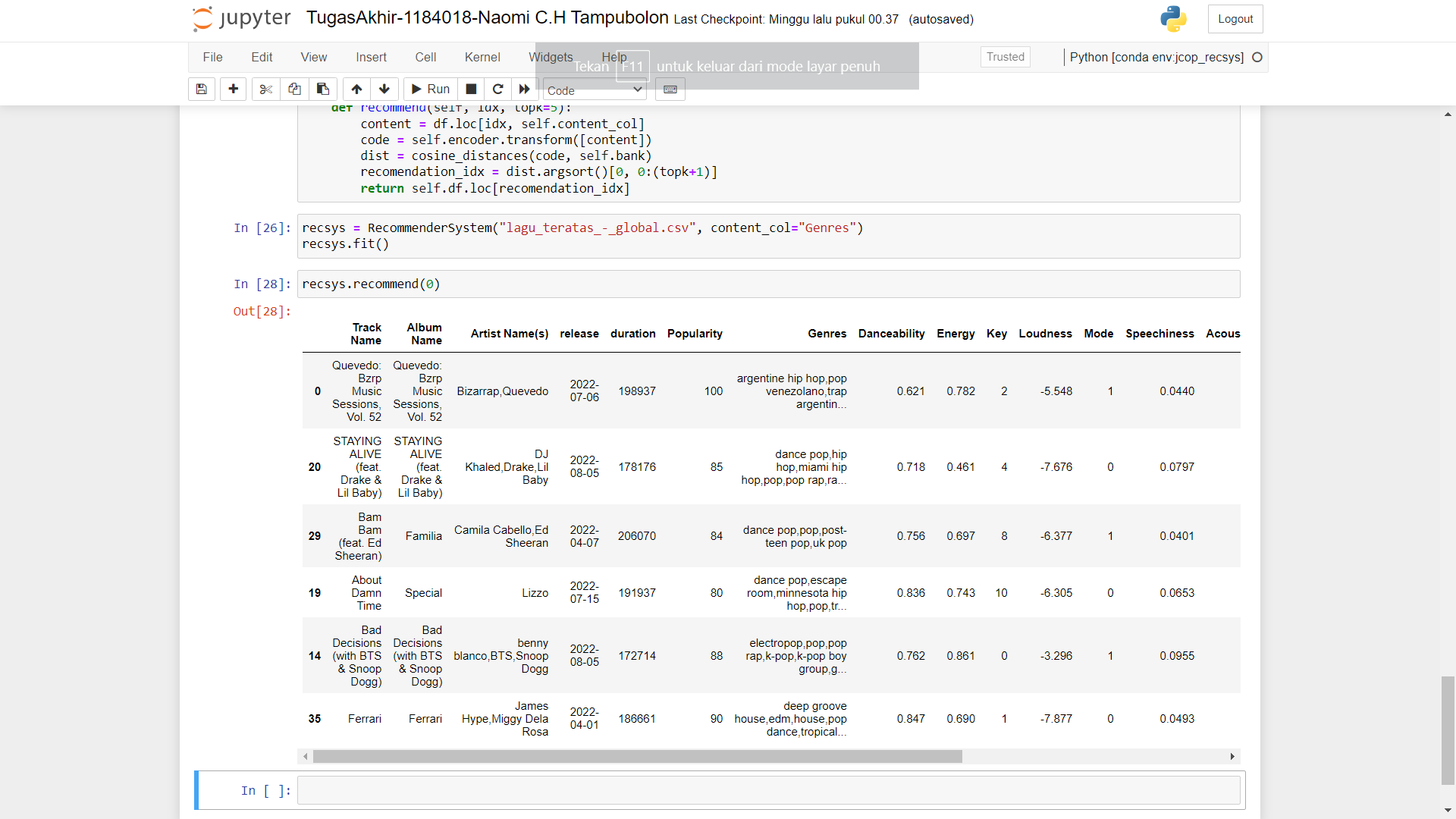




* Pemodelan *Content-Based Filtering.*

Tahap pemodelan *content-based filtering* adalah pemodel sistem rekomendasi berdasarkan *content-based filtering*. Pada pemodelan ini kita dapat menentukan konten atau item yang ingin kita rekomendasikan berdasarkan *index* yang ingin digunakan. standar rekomendati yang digunakan ialah *index* 0, kemudian akan menampilkan rekomendasi lagu berdasarkan *genre* yang mirip dengan *index* 0. Dapat dilihat sebagai berikut:





## HASIL

Berikut tabel lagu yang di rekomendasikan berdasarkan properti trek menggunakan *content-based filtering*.

|  |
| --- |
| *Content-Based Filtering* |
| *Quevedo: Bzrp Music Sessions, Vol. 52* |
| *STAYING ALIVE (feat. Drake & Lil Baby)* |
| *Bam Bam (feat. Ed Sheeran)* |
| *About Damn Time* |
| *Bad Decisions (with BTS & Snoop Dogg)* |
| *Ferrari* |

# Kesimpulan dan Saran

## KESIMPULAN

Berdasarkan hasil proses penelitian diatas, maka dapat diambil kesimpulan pada penelitian ini sebagai berikut:

1. 50 lagu tertas – global (*50 top songs - global*) dianalisa dengan cara melakukan *cleaning* data seperti: penyuntingan (*editing*), perumusan, pengujian data. Sistem rekomendasi digunakan dalam metode. *Content-Based* *Filtering* digunakan sebagai pemodelan dalam merekomendasikan lagu berdasarkan parameter dari properti trek serta menggunakan *Jupyter Notebook* untuk menampilkan hasil rekomendasi lagu berdasarkan properti trek.
2. Berdasarkan properti trek ‘*genre’* terdapat 5 lagu teratas yang direkomendasikan menggunakan metode *Content-Based Filtering* yaitu *STAYING ALIVE (feat. Drake & Lil Baby), Bam Bam (feat. Ed Sheeran), About Damn Time, Bad Decisions (with BTS & Snoop Dogg)*, dan *Ferrari*.

## SARAN

Adapun saran untuk penelitian ini adalah hasil dari penelitian ini dapat menjadi bahan untuk penelitian selanjutnya yang lebih baik lagi adanya.

# References

Fitriani, S. A. (n.d.). *Analisis Klaster Atribut Musik pada Global Top 50 dengan Data Spotify dengan Menggunakan Algoritma K-Means*.

Muslimin M. (n.d.). *Perkembanfan Teknologi dalam Industri Media*.

Noviani, D., Pratiwi, R., Silvianadewi, S., Alexandri, M. B., & Hakim, M. A. (2020). Pengaruh Streaming Musik Terhadap Industri Musik di Indonesia. In *Jurnal Bisnis STRATEGI •* (Vol. 29, Issue 1).

Usino, W., Satria Prabuwono, A., Hamed Allehaibi, K. S., Bramantoro, A., & Amaldi, W. (2019). Document Similarity Detection using K-Means and Cosine Distance. In *IJACSA) International Journal of Advanced Computer Science and Applications* (Vol. 10, Issue 2). www.ijacsa.thesai.org

Wahyu Dewatara, G., & Monik Agustin, S. (2019). Pemasaran Musik pada Era Digital Digitalisasi Industri Musik dalam Industri 4.0 di Indonesia. *WACANA: Jurnal Ilmiah Ilmu Komunikasi*, *18*, 1–10.