**LAPORAN PROYEK DATA MINING**

***Association Rule for Business Strategic and Promo Recommendation in Restaurant ABC***



**Disusun oleh:**

1. 12S17012 – Reza Oktovian Siregar
2. 12S17032 – Angelia Regina Ginting
3. 12S17047 – Christina Clara

**PROGRAM STUDI SARJANA SISTEM INFORMASI**

**FAKULTAS INFORMATIKA DAN TEKNIK ELEKTRO**

**INSTITUT TEKNOLOGI DEL**

**NOVEMBER 2020**

# DAFTAR ISI

[DAFTAR ISI i](#_Toc60682958)

[DAFTAR TABEL ii](#_Toc60682959)

[DAFTAR GAMBAR iii](#_Toc60682960)

[Bab I Business Understanding 1](#_Toc60682961)

[Bab II Data Understanding 5](#_Toc60682962)

[Bab III Tinjauan Pustaka 6](#_Toc60682963)

[**3.1** ***Data Mining*** 6](#_Toc60682964)

[**3.2** **Asosiasi** 7](#_Toc60682965)

[**3.3** **Algoritma Apriori** 8](#_Toc60682966)

[Bab IV *Data Prepration* 9](#_Toc60682967)

[**4.1** ***Data Set*** 9](#_Toc60682968)

[**4.2** ***Select Data*** 9](#_Toc60682969)

[**4.2.1** ***Package* yang digunakan** 10](#_Toc60682970)

[4.3 *Clean Data* 11](#_Toc60682971)

[**4.4** ***Format Data*** 12](#_Toc60682972)

[Bab V *Modelling* 14](#_Toc60682973)

[**5.1** ***Select Modeling Technique*** 14](#_Toc60682974)

[**5.2** ***Build Model*** 16](#_Toc60682975)

[Bab VI *Evaluation* 22](#_Toc60682976)

[**6.1** ***Evaluate Result*** 22](#_Toc60682977)

[**6.2** ***Evaluate Process*** 23](#_Toc60682978)

[Bab VII *Deployment* 27](#_Toc60682979)

[Bab VIII Kesimpulan 29](#_Toc60682980)

[Bab IX Lampiran 30](#_Toc60682981)

[REFERENSI 31](#_Toc60682982)

# DAFTAR TABEL

[**Tabel 1. Project Plan** 2](#_Toc60683023)

[**Tabel 2. Package yang digunakan** 10](#_Toc60683024)

# DAFTAR GAMBAR

[Gambar 1. Tahapan Proses 6](#_Toc60682985)

[Gambar 2 Data Transaksi Rumah Makan 9](#_Toc60682986)

[Gambar 3 Association rule 10](#_Toc60682987)

[Gambar 4 Deskripsi nilai null pada kolom 12](#_Toc60682988)

[Gambar 5 Deskripsi missing value pada data 12](#_Toc60682989)

[Gambar 6 Transformasi nilai quantity 13](#_Toc60682990)

[Gambar 7. Itemsets 15](#_Toc60682991)

[Gambar 8. Antecendent dan Consequent 16](#_Toc60682992)

[Gambar 9. Pemodelan analisis asosiasi 16](#_Toc60682993)

[Gambar 10. Kode program transformasi data 17](#_Toc60682994)

[Gambar 11. Potongan kode frequent\_itemsets 17](#_Toc60682995)

[Gambar 12. Potongan kode association rule 18](#_Toc60682996)

[Gambar 13. Kode program analisis asosiasi 18](#_Toc60682997)

[Gambar 14. Kode program visualisasi bar chart 19](#_Toc60682998)

[Gambar 15. Bar chart pasangan menu 19](#_Toc60682999)

[Gambar 16. Heatmap barang yang kemungkinan dibeli bersamaan 20](#_Toc60683000)

[Gambar 17. Hasil Evaluasi Rule Pertama 22](#_Toc60683001)

[Gambar 18. Hasil Evaluasi Rule Kedua 22](#_Toc60683002)

[Gambar 19. Potongan Kode Pengambilan Data Set 23](#_Toc60683003)

[Gambar 20. Potongan Kode Item Terfavorit 23](#_Toc60683004)

[Gambar 21. Potongan Kode Frozenset 23](#_Toc60683005)

[Gambar 22. Potongan Kode Pembuatan Data Frame 24](#_Toc60683006)

[Gambar 23. Potongan Kode Daftar 5 Item Terfavorit 24](#_Toc60683007)

[Gambar 24. Potongan Kode Frequent Itemsets 24](#_Toc60683008)

[Gambar 25. Potongan Kode Minimum Support 24](#_Toc60683009)

[Gambar 26. Potongan Kode Daftar 10 Itemsets 25](#_Toc60683010)

[Gambar 27. Potongan Kode Rule Pertama 25](#_Toc60683011)

[Gambar 28. Potongan Kode Rule Kedua 26](#_Toc60683012)

[Gambar 29. Visualisasi Heatmap 27](#_Toc60683013)

# Bab I Business Understanding

Rumah makan ABC merupakan usaha bisnis yang bergerak di bidang wisata kuliner. Didalam rumah makan tersebut banyak informasi yang menyediakan berbagai menu pilihan hidangan yang diperuntukkan untuk pelanggan. Bahkan diera teknologi sekarang, banyak pelanggan memanfaatkan teknologi untuk melakukan dan memberikan ulasan terhadap objek hidangan tersebut. Hal ini dapat membantu pelanggan untuk memperoleh informasi yang dibutuhkan terkait hidangan. Namun dengan bertambah banyaknya informasi tersebut dapat membuat pelanggan merasa kebigungan mengambil keputusan terhadap menu hidangan mana yang sesuai dengan keinginan dan selera pelanggan. Dengan ini maka pelanggan dapat mengalami *overload information* atau banyaknya informasi yang diterima oleh pelanggan.Bagi pengelola rumah makan juga didalam perancangan strategi bisnis yang mampu meningkatkan transaksi diantara penjual dan pembeli didalam proses bisnis, maka pengelola membutuhkan pemahaman yang mendalam terkait dengan perilaku konsumen [1].

Permasalahan-permasalahan ini dapat diatasi melalui pengembangan sistem rekomendasi didalam organisasi. Rekomendasi produk dapat dimanfaatkan sebagai aplikasi model yang mampu memahami keadaan dan keinginan pelalggan melalui adanya hasil observasi terhadap data transaksi oleh pelanggan. Melalui Association Rule Mining sebagai pendekatan yang mampu memberikan dukungan keputusan antara produk yang dibeli secara bersama-sama oleh pelanggan [2].

Salah satu teknik untuk membangun sistem rekomendasi yang tepat sasaran tanpa melakukan interaksi secara langsung dengan konsumen adalah dengan memanfaatkan teknik data mining. Teknik data mining yang digunakan didalam penelitian ini yakni menggunakan fungsi *association.* Melalui pengolahan data dengan fungsi *association* dapat mendapatkan informasi terkait dengan menu makanan yang saling berasosiasi maka ketika suatu produk tersebut diminati oleh pelanggan maka produk lain yang memiliki keterkaitan yang kuat akan memliki kemungkinan juga akan diminati.

Melalui pemilahan data transaksi yang digunakan sebagai bahan analisa dikarenakan melalui data transaksi antar penjual dan pembeli dapat menemukan pola-pola asosiasi antar menu hidangan yang dipesan. Metode yang dimanfaatkan dalam mengidentifikasi pola yakni menggunakan *association rule* dengan algoritma apriori. Metode serta algoritma ini, tim peneliti akan menemukan transaksi penjualan dengan *strong association* (keterkaitan yang kuat) antar menu hidangan didalam transaksi yang akan dimanfaatkan oleh sistem rekomendasi sebagai saran menu hidangan yang dapat membantu pelanggan menentukan menu hidangan yang akan dipesan dari rumah makan tersebut. [3]

Oleh karena itu, hasil dari penelitian berupa asosiasi antar menu hidangan, sebagai rekomendasi alternatif dari pemesanan makananan yang didapatkan dari perhitungan minsup dan mincof. Hasil yang diperoleh dari penelitian ini mencakup tujuan yakni :

1. Mengimplementasikan *association rule mining* pada data transaksi penjualan rumah makan abc.
2. Menemukan pola kebiasaan pelanggan yang dapat membantu merekomendasikan produk kepada pelanggan.
3. Membantu meningkatkan penjualan dengan melalui penemuan pola penjualan hidangan yang dapat dimanfaatkan untuk mempromosikan hidangan tertentu dan dapat merancang strategi pemasaran serta promosi produk menjadi lebih tepat sasaran.

Pada penelitian ini menggunakan *dataset* menu hidangan pada rumah makan abc yang diambil dari *kaggle*.

***Project Plan***

Tahap perencanaan dalam mencapai tujuan dari proyek data mining dalam mencapai tujuan dari bisnis “*Association Rule for Business Strategic and Promo Recommendation in Restaurant ABC*” dilampirkan sebagai berikut :

**Tabel 1. Project Plan**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Tahapan | Durasi Waktu | Sumber daya | Kegiatan | Ketergantungan |
| Business Understanding | 3 hari | Semua *analysts* | Tahapan awal yang berfokus pada pemahaman tujuan proyek dan persyratan bisnis, menilai situasi, tujuan data mining , merancang *project plan* yang kemudian dikonversikan menjadi rumusan masalah untuk mencapai tujuan | Pemahaman komsep serta perkembangan penerapan konsep data mining |
| Data Understanding | 3 hari | Semua *analysts* | Tahapan pemahaman data yang dimulai dengan pengumpulan data awal , mendeskripsikan data, mengeksplor data dan mengidentifikasikan masalah kualitas data. | Masalah data dan teknologi |
| Data Prepetation | 2 minggu | Data mining *consultant,* beberapa *database analyst time* | Tahap persiapan data melalui pemilahan data serta transformasi dan pembersihan data untuk pemodelan | Masalah data dan teknologi |
| Modelling | 2 minggu | *Data mining consultant*, beberapa *database analyst time* | Memilih teknik pemodelan dan diterapkan serta parameter model dikalibrasi dengan niali-nilai optimal. | Ketidakmampuan menemukan model yang tepat |
| Evaluation | 1 minggu | Semua *analysts* | Mengevaluasi model dan meninjau langkah-langkah yang telah dilakukan untuk membangun model dengan mempertimbangkan bahwa model yang dihasilkan mencapai tujuan bisnis | Ketidakmampuan untuk menerapkan hasil, terjadi kesalahan pada proses pengerjaan proyek, perkembangan penerapan konsep data mining |
| Deployment | 1 minggu | *Data mining consultant*, beberapa *database analyst time* | Meningkatkan pengetahuan dari data, membuat plan deployment, *Monitoring* dan *Maintance Plan* dan meninjau proyek | Ketidakmampuan untuk menerapkan hasil, perkembangan penerapan konsep data mining |

Didalam proses pengembangan proyek ini akan menggunakan *tools data mining* yang mampu mendukung tahapan proses *Tools* dan teknik *association rule* dengan algoritma Apriori adalah Python. Didalam Python akan melakukan i*mport library,* yang pertama adalah pandas ( Untuk melakukan operasi *dataframe*) dan mlxtend (*library* untuk *machine learning*) dengan mengimport apriori dan *association rules.*

# Bab II Data Understanding

*Data Understanding* adalah fase mengumpulkan data awal dan mempelajari data untuk bisa mengenal data yang akan dipakai. Pada proyek ini menggunakan metode *association rule* dan algoritma apriori di mana *data understanding* dimulai dengan mengidentifikasikan masalah data, mendeteksi *subset* yang menarik dari data untuk membentuk hipotesis awal [4].

1. *Collect Initial Data*

Tahap mengumpulkan data awal dilakukan dari hasil laporan dataset transaksi order makanan dalam bentuk dokumen *spreadsheet* yang berasal dari kaggle (sumber: <https://www.kaggle.com/henslersoftware/19560-indian-takeaway-orders> )[5].

1. *Describe Data*

Dataset Transaksi Restoran yang digunakan terdiri dari beberapa atribut berupa *Order Number, Order Date, Item Name, Quantity, Product Price* dan *Total products* [5].

1. *Explore Data*

Pada tahap ini dilakukan dengan pendekatan EDA dalam memaknai informasi yang terkandung pada data. Penggunaan EDA dalam mengeksplorasi data dapat digunakan untuk mengurangi dimensi data yang terlalu besar. Penggunaan EDA dilakukan untuk mengeksplorasi data dengan menggunakan fungsi statistik, matematik, dan divisualisasikan dalam bentuk grafik. Dengan menggunakan EDA akan didapatkan ringkasan statistik dari data yang dimiliki, antara lain rata-rata, nilai minimum, nilai maksimum, jumlah data, standar deviasi, dan lain-lain [6].

1. *Verify Data Quality*

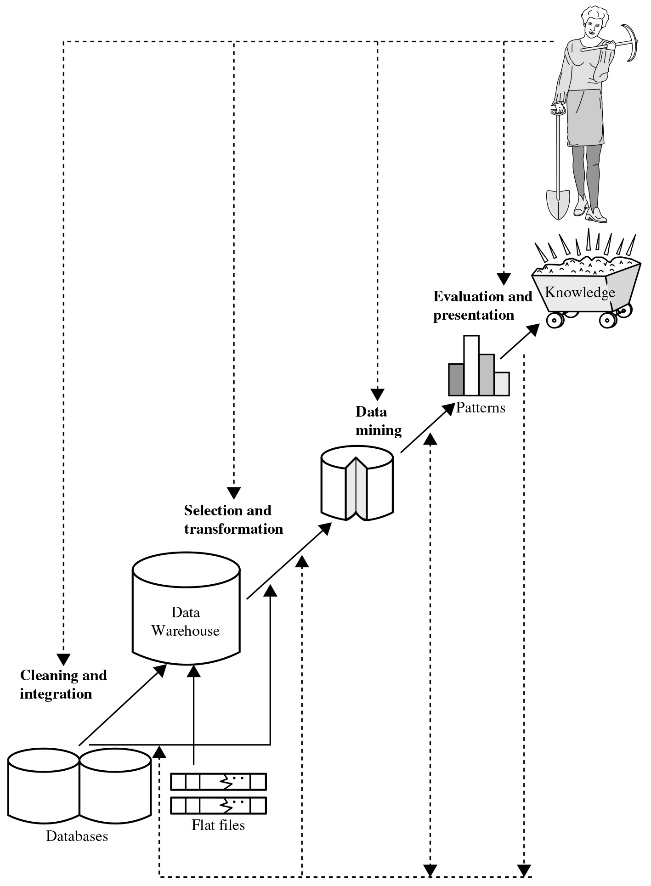
Pada tahap ini dilakukan evaluasi kualitas dan kelengkapan data pada dataset Transaksi Restoran. Evaluasi data ini dilakukan untuk melihat nilai-nilai yang hilang atau *missing value* terutama jika data yang dikumpulkan pada jangka waktu yang lama, memeriksa atribut yang hilang atau kosong (*null*), melihat terdapatnya data rangkap atau data *double*, serta menilai apakah semua nilai masuk akal, pada tahap ini data yang dimiliki tidak ada yang memiliki missing value [5]. Pada *verifiy data quality* daoat dilihat pada subab 4.3.

# Bab III Tinjauan Pustaka

Pada bab ini akan dijelaskan landasan teori terkait proyek yang akan dikerjakan.

## ***Data Mining***

*Data Mining* yang merupakan suatu metode untuk menemukan pengetahuan dalam data yang memiliki ukuran yang cukup besar dengan tujuan untuk menemukan pola atau kecenderuangan yang penting yang keberadaanya biasa tidak disadari. Data mining yang sering disebut juga dengan *knowledge discovery database* (KDD) merupakan suatu kegiatan untuk mengumpulkan, memakai data historis yang berguna untuk menemukan keteraturan, pola atau hubungan pada data set yang berukuran besar. Adapun proses *knowledge discovery database* (KDD) seperti data *cleaning*, data integration, data *selection*, data *transformation*, data *mining*, *pattern evaluation,* *knowledge presentation*. [7]



Gambar 1. Tahapan Proses

Sumber ( <https://hosteko.com/blog/pengertian-fungsi-metode-dan-penerapan-data-mining> )

## **Asosiasi**

*Data mining* dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan. Salah satu tugas dari data mining adalah asosiasi. Asosiasi bertugas untuk menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis asosiasi disebut juga dengan analisis keranjang belanja. Kasus dalam asosiasi salah satu contohnya adalah menentukan barang dalam *supermarket* yang dibeli secara bersamaan dan yang tidak perna dibeli secara bersamaan.

Analisis asosiasi merupakan penemuan dari *association rule* yang dapat menampilkan pola-pola yang sering muncul pada data. Pada analisis asosiasi terdapat *support* dan *confidence* kedua hal ini dapat menunjukan bahwa suatu *rule* dapat dipercaya. *Support* merupakan ukuran dimana seberapa besar tingkat dominasi suatu *item* atau *itemset* pada semua transaksi yang terjadi. Sedangkan *confidence* ukuran yang menampilkan hubungan yang terjadi terhadap dua *item* secara kondisional. Analisis asosiasi bertujuan untuk menemukan aturan-aturan asosiasi yang memperlihatkan kondisi-kondisi nilai atribut yang kemunculannya terjadi secara bersamaan dalam suatu himpunan data. [8]

Metodologi dasar analisis asosiasi terbagi kedalam dua tahap :

1. Analisa pola frekuensi tinggi

Tahap ini mencari kombinasi item yang memenuhi syarat *minimum* dari nilai *support* yang terdapat pada *database.* Nilai *support* sebuah *item* dapat diperoleh dengan rumus

*Support*

Sedangkan nilai dari *support 2 item* diperoleh dengan menggunakan rumus berikut :

1. Pembentukan aturan assosiatif

Setelah semua pola frekuensi tinggi ditemukan selanjutnya aturan asosiatif yang memenuhi *confidence* dengan menghitung *confidence* aturan asosiatif “jika A maka B” nilai *confidence* dari aturan “jika A maka B” akan dicari dan diperoleh rumus seperti berikut [9]

## **Algoritma Apriori**

Algoritma apriori berguna untuk menentukan *frequent itemsets* aturan asosiasi boolean yang diusulkan oleh Agrawal & Srikant pada tahun 1994. Algoritma apriori termasuk jenis aturan asosiasi yang menyatakan asosiasi antara beberapa atribut. Algoritma apriori dibagi menjadi beberapa tahap yang disebut narasi

Algoritma apriori memiliki dua langkah utama dala cara kerja :

1. *Join* (penggabungan)

Setiap *itemset* yang dianggap sebagai kandidat k-itemset. *Frequent itemset* ni yang memenuhi *support* pada proses ini akan digabungkan untuk mendapatkan kandidat *itemset.* Proses ini akan dilakukan secara terus menerus sampai kandidat *itemset* tida ditemukan lagi.

1. *Prune* (Pemangkasan)

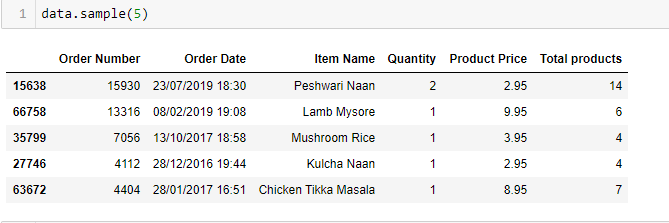
Proses ini basis data akan memindai seluruh *itemset* untuk menemukan apakah *itemset* termasuk *frequent* atau tidak dan *itemset* yang tidak mencukup minimal dari *support* akan dipangkas. [10]

# Bab IV *Data Prepration*

Pada tahapan ini akan dilakukan proses pemilahan serta pengolahan data yang akan dimanfaatkan pada tahapan selanjutnya yakni permodelan dimana model yang telah dirancang akan dapat memberikan hasil maksimal sesuai dengan target yang dibutuhkan [11]. Berikut proses tahapan persiapan data akan yakni seperti deskripsi *data set,* memilih data, membangun data, mengintegrasikan data dan membersihkan data.

## ***Data Set***

Data yang digunakan didalam penelitian mengunakan data transaksi rumah makan abc yang dimana data yang diambil berasal dari kaggle yakni data *Indian takeaway restaurant*. Data yang diperoleh berbentuk .csv (*comma separated values*) yang terdiri dari 74818 baris transaksi serta 7 atribut/ *feature* yang akan dimanfaatkan untuk menemukan pola konsumsi pelanggan.



Gambar 2 Data Transaksi Rumah Makan

## ***Select Data***

Data transaksi yang telah dikumpulkan tidak dapat secara langsung digunakan di dalam proses *association rule mining*, maka sebelumnya data tersebut akan dilakukan pemeriksaan terhadap data yang redundan, data duplikasi, atau bahkan memeriksa data yang tidak lengkap, selain itu diperiksa juga apakah seluruh atribut tersebut akan digunakan semuanya ataukah tidak.

Pada perancangan *association rules* menu pada rumah makan abc, maka tim proyek memutuskan atribut yang digunakan adalah *Order Number, Quantity* serta *Item name* yang akan dikelola didalam membangun *association rule*.



Gambar 3 Association rule

### ***Package* yang digunakan**

Berikut library *python* yang digunakan didalam melakukan *coding* mengembangkan *association rule* [12]:

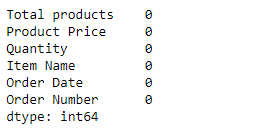
**Tabel 2. Package yang digunakan**

| **Library** | **Deskripsi** |
| --- | --- |
| Pandas | *Library* yang mendukung untuk mengolah data data berbentuk tabular yang tersimpan dalam baris dan kolom. Pandas mendukung pembacaan dan penulisan data dengan media berupa excel *spreadsheet,* CSV, dan SQL yang kemudian akan dijadikan sebagai objek *python* dengan *rows* dan *columns* yang disebut *data frame.* |
| Pandas\_profiling | *Library* yang bersifat *open-source* yang dapat menghasilkan laporan interaktif yang menarik untuk kumpulan data apa pun, hanya dengan satu baris kode. Pembuatan profil panda menghasilkan laporan profil dari panda *DataFrame* |
| os | Modul OS di-*python* menyediakan fungsi untuk membuat dan menghapus direktori (*folder*), mengambil isinya, mengubah dan mengidentifikasi direktori saat ini, dll. Dengan *library* ini dapat melakukan banyak tugas sistem operasi secara otomatis. |
| sys | Modul sys menyediakan fungsi dan variabel yang digunakan untuk memanipulasi bagian berbeda dari lingkungan *runtime python*. Anda akan mempelajari beberapa fitur penting dari modul ini di sini. |
| mlxtend | *Library* untuk *machine learning*, kemudian mlxtend melakukan import fungsi apriori dan association\_rules. Fungsi Apriori untuk mengekstrak *frequent itemsets*  untuk *association rules* |
| matplotlib | *Library* untuk menyajikan visualisasi data |
| numpy | *Library* pada *python* yang berfungsi untuk melakukan operasi vektor dan matriks dengan mengolah *array* dan *array* multidimensi. Biasanya NumPy digunakan untuk kebutuhan dalam menganalisis data. |
| seaborn | *Library visual Python* yang berlandaskan pada matplotlib. Seaborn menyediakan antar-muka tingkat tinggi untuk menangani permasalahan terkait visualisasi data secara statistik agar tampak lebih menarik. |

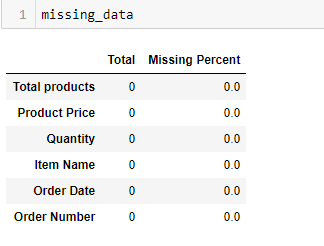
# *Clean Data*

Pada tahapan ini dilakukan untuk memastikan kualitas data yang diperoleh apakah sudah layak data tersebut untuk diproses dengan *data mining*. Tahap data *cleaning* merupakan tahap menghilangkan *noisy* data, melengkapi data dan mengatasi data yang tidak relevan dan tidak konsisten. Aktivitas yang dilakukan didalam proses ini antara lain membersihkan data, menghapus data yang tidak diperlukan serta menyeragamkan isi dari tiap baris data, sehingga data yang akan digunakan siap untuk dimodelkan.

Setelah tim proyek melakukan eksplorasi data ternyata tidak terdapat nilai *null* pada kolom data, isi data tiap barisanya juga sudah sesuai dan tidak terdapat *missing value* pada tiap kolom data. Oleh karena itu, data ini telah dapat dimanfaatkan ke dalam tahapan proses pemodelan.



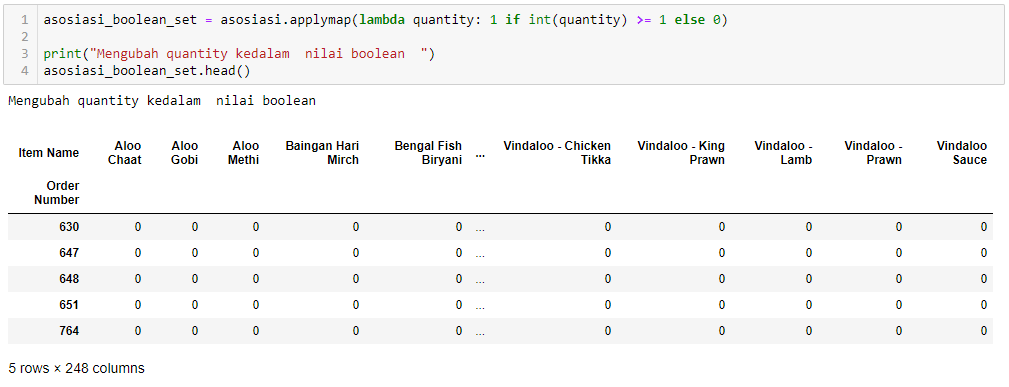
Gambar 4 Deskripsi nilai null pada kolom



Gambar 5 Deskripsi missing value pada data

## ***Format Data***

Bentuk standar adalah adalah bentuk data yang akan diakses oleh algoritma *data mining*. Data dalam penelitian ini dibuat dalam bentuk *binary valued data.* Pengubahan data dalam bentuk *Boolean* yang menujukkan frekuensi setiap produk dalam transaksi.



Gambar 6 Transformasi nilai quantity

# Bab V *Modelling*

Pada bab ini akan dijelaskan mengenai pemodelan yang di implementasikan pada proyek “*Association Rule for Business Strategic and Promo Recommendation in Restaurant ABC*” terdiri dari dua subbab yaitu pemilihan model teknik dan pembangunan model.

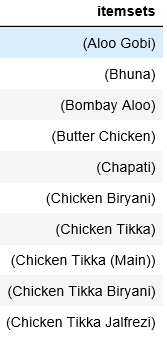
## ***Select Modeling Technique***

Teknik yang digunakan pada proyek ini menggunakan *association rule* dengan algortima apriori. *Association rule* yang dikenal sebagai analisis asosiasi merupakan teknik penambangan data yang bertujuan untuk menemukan aturan asosiatif antara kombinasi *item*. Prinsip penambangan analisis asosiasi dapat dilakukan dengan algoritma apriori. Algoritma apriori berguna untuk menentukan *frequent itemsets*.

Algoritma apriori memiliki tiga komponen utama diantara lain *support, confidence* dan *lift. Support* merupakan ukuran dimana seberapa besar tingkat dominasi suatu *item* atau *itemset* pada semua transaksi yang terjadi., *confidence* ukuran yang menampilkan hubungan yang terjadi terhadap dua *item* secara kondisional, misalnya adalah apabila item A dibeli kemungkinan item B akan dibeli secara bersamaan, dan *lift* merupakan sebuah angka rasio yang menunjukkan berapa banyak berapa banyak kemungkinan menemukan sebuah atribut yang muncul bersamaan dengan atribut lainnya dibandingkan dengan seluruh kejadian adanya atribut yang terpenuhi dengan menggunakan *lift* dapat dilihat tingkat kekuatan aturan atas kejadian acak dari *antecendent* dan *consequent* dilihat dari support masing-masing. Apabila nilai rasio *lift* lebih besar dari 1 berarti semakin besar kemungkinan untuk membeli kedua produk secara bersamaan. Selain itu, pada metode apriori terdapat beberapa definisi yang pelu untuk diketahui seperti :

1. *Itemset*

Kumpulan dari suatu item atau lebih disebut dengan *itemset,* ini merupakan kandidat dari aturan-aturan asosiasi. Contoh dari *itemsets* pada proyek ini dapat dilihat pada gambar dibawah.



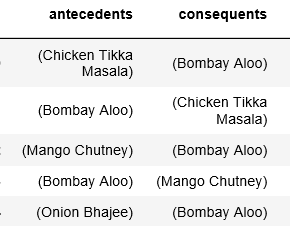
Gambar 7. Itemsets

1. *Minimum Support*

*Minimum Support* merupakan suatu indikator untuk membedakan antara *itemset* yang *frequent* atau tidak. Nilai dari *minimum support* ditetapkan oleh pembuat proyek dengan *minimum support* 3%.

1. *Antecendent* dan *Consequent*

*Antecendent* merupakan kandidat yang menjadi aturan sedangkan *Consequent* adalah kandidat yang menjadi kesimpulan dari aturan tersebut. Contoh dari *antencedent* dan *consequent* dapat dilihat pada gambar dibawah ini dapat dilihat bahwa Jika membeli Chicken Tikka Masala (*antencedent* )maka membeli Bombay Aloo(*consequent)*.

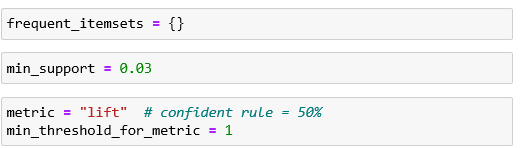


Gambar 8. Antecendent dan Consequent

## ***Build Model***

Implementasi analisis asosiasi menggunakan algoritma apriori dengan dibantu oleh library python MLXtend yang berisi ekstensi dan modul bantuan untuk analisis data dan library pembelajaran mesin lainnya. [13]

Pemodelan analisis asosiasi menggunakan apriori ini menggunakan minimum support 0.03. Metriks yang digunakan untuk mengenerate aturan adalah *support, confidence, lift, leverage, conviction* merupakan kemungkinan dari A dibeli ketika B telah dibeli dengan mempertimbangkan popularitas A jika *lift* > 1 item B kemungkinan besar akan dibeli secara bersamaan dengan A.



Gambar 9. Pemodelan analisis asosiasi

Lakukan transformasi data apabila terdapat jumlah barang lebih atau sama dengan 1 maka akan di set menjadi 1 sedangkan apabila tidak terjadi transaksi akan diset menjadi 0.



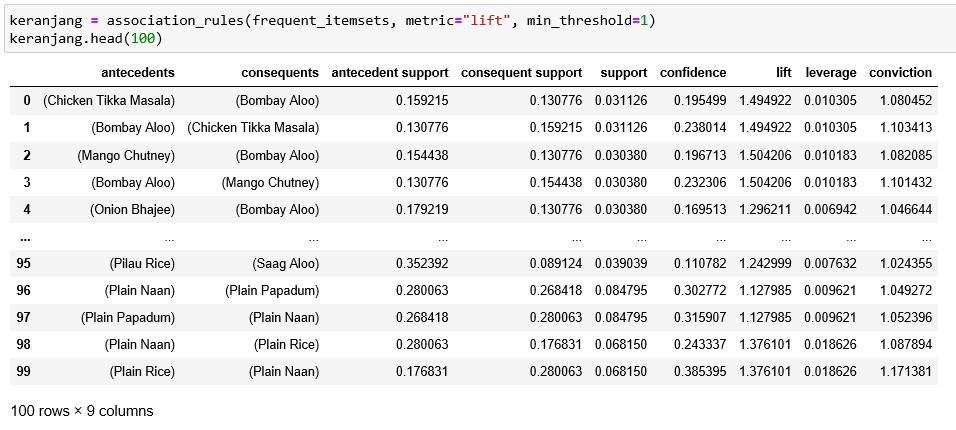
Gambar 10. Kode program transformasi data

Selanjutnya akan dibuat variabel dimana terdiri dari beberapa barang yang sering dibeli dari seluruh transaksi yang sudah di simpan dalam *database* dengan menggunakan perintah apriori dalam proyek ini dibuat dengan minimum support adalah 0.03 atau 3%.



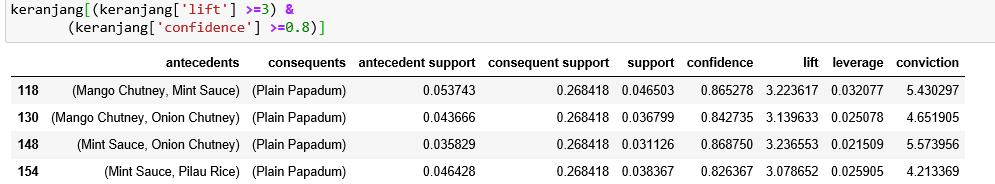
Gambar 11. Potongan kode frequent\_itemsets

Variabel yang memiliki aturan-aturan asosiasi akan dibangun dari setiap barang, variabel keranjang merupakan hasil dari fungsi pencarian asosiasi dimana data yang digunakan berasal dari *frequent\_itemsets* dengan nilai minimum *lift ratio* adalah 1.



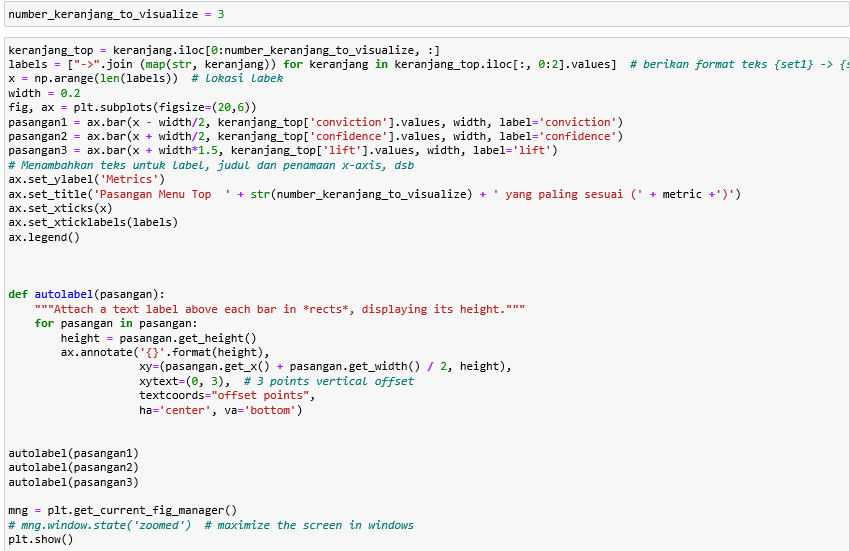
Gambar 12. Potongan kode association rule

Pemodelan juga dapat dilakukan dengan menambahkan ambang batas seperti dilakukan *filter* yang memiliki *lift* >= 3 dan *confidence* >= 0.8



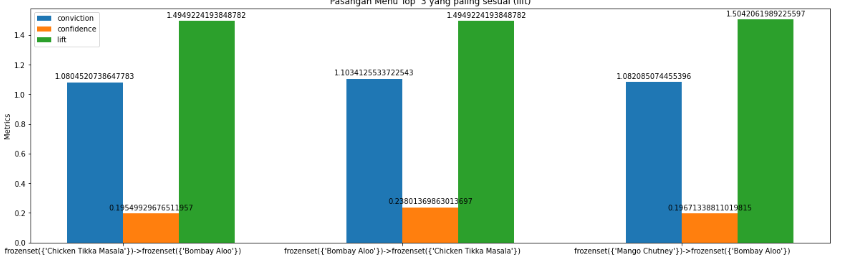
Gambar 13. Kode program analisis asosiasi

Aturan-aturan yang sudah diterapkan sebelumnya akan lebih mudah untuk dipahami oleh perusahaan dengan disajikan dalam bentuk visualisasi. Pada proyek analisis asosiasi akan menggunakan dua buah visualisasi yaitu *heatmap* dan *bar chart*.



Gambar 14. Kode program visualisasi bar chart

Pada visualisasi *bar chart* dapat dilihat bahwa Chicken Tikka Masaka dengan Bombay Aloo memiliki lift lebih besar dari satu yang berarti pada *conviction* yang merupakan kemungkinan dari Chicken Tikka Masaka dibeli ketika Bombay Aloo telah dibeli dengan mempertimbangkan popularitas Chicken Tikka Masaka jika *lift* > 1 item Bombay Aloo kemungkinan besar akan dibeli secara bersamaan dengan Chicken Tikka Masaka .



Gambar 15. Bar chart pasangan menu

Analisis untuk setiap aturan yang ada :

1. {Chicken Tikka Masala}->{Bombay Aloo}

Definisi : jika pelanggan membeli Chicken Tikka Masala maka Bombay Aloo akan dibeli juga.

Dengan *confidence* dari aturan tersebut 0.19, *lift* > 1 yang berarti Bombay Aloo kemungkinan akan dibeli dengan Chicken Tikka Masala.

1. { Bombay Aloo }->{ Chicken Tikka Masala }

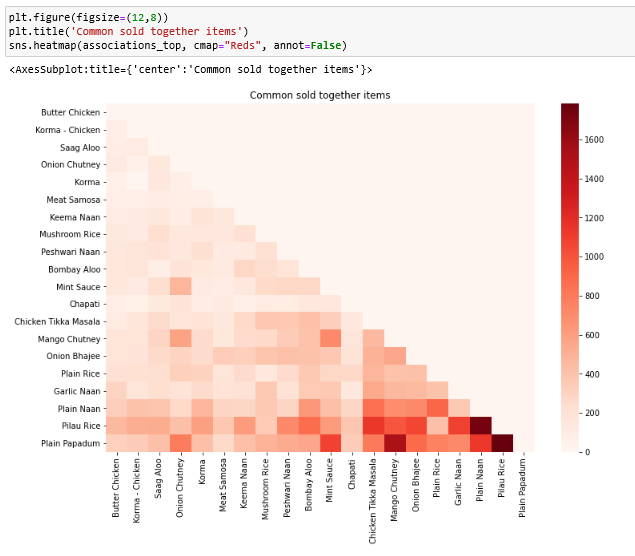
Definisi : jika pelanggan membeli Bombay Aloo maka Chicken Tikka Masala akan dibeli juga.

Dengan *confidence* dari aturan tersebut 0.23, *lift* > 1 yang berarti kemungkinan Chicken Tikka Masala akan dibeli dengan Bombay Aloo. Kemungkinan dari membeli Bombay Aloo maka Chicken Tikka Masala akan dibeli juga lebih besar daripada 3 rule lainnya.

1. {Manggo Chutney }->{Bombay Aloo}

Definisi : jika pelanggan membeli Manggo Chutney maka Bombay Aloo akan dibeli juga.

Dengan *confidence* dari aturan tersebut 0.19, *lift* > 1 yang berarti Bombay Aloo kemungkinan akan dibeli dengan Manggo Chutney.



Gambar 16. Heatmap barang yang kemungkinan dibeli bersamaan

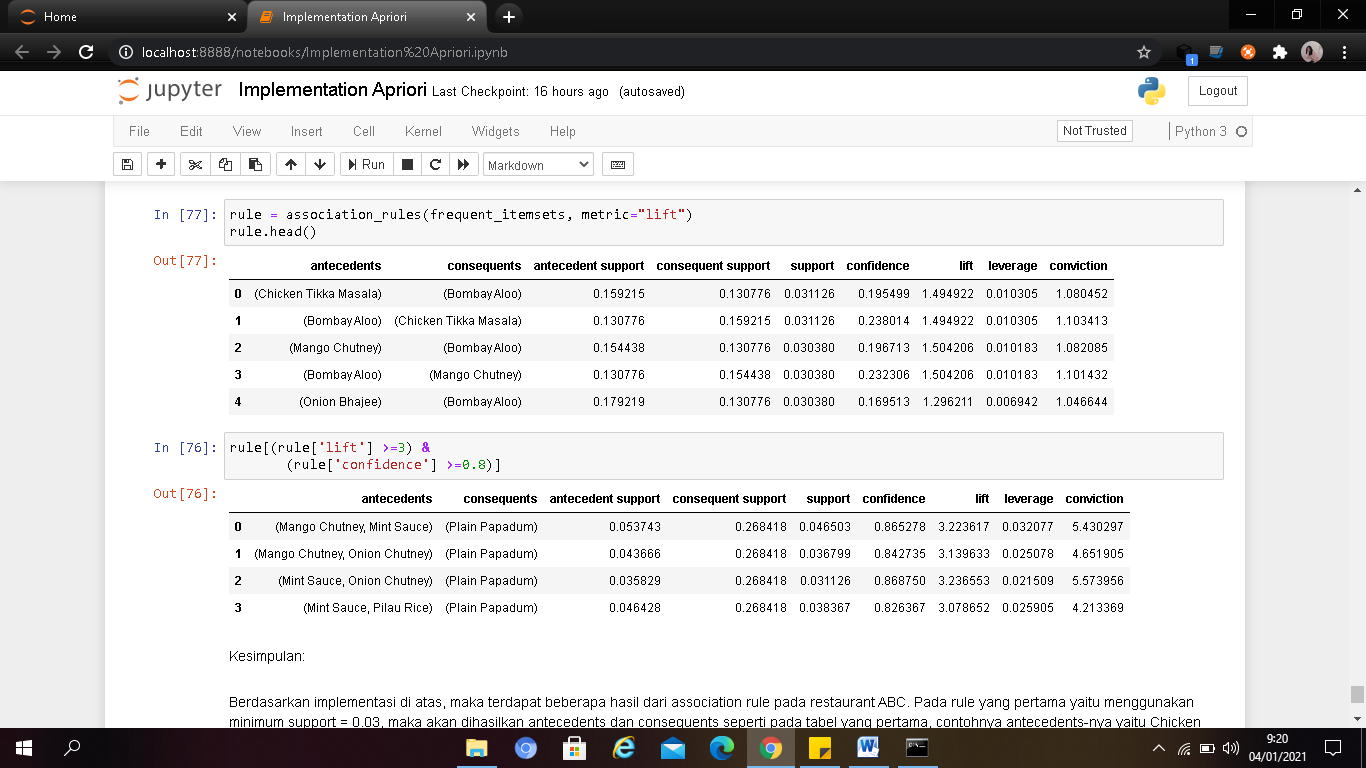
Pada visualisasi *heatmap* dapat dilihat *item* yang akan terjual secara bersamaan yang pada umumnya akan terjadi, seperti Pilau Rice dan Plain Papadum. Pada *heatmap* semakin warna pada kotak merah gelap maka kemungkinan barang tersebut dibeli secara bersamaan semakin tinggi.

# Bab VI *Evaluation*

Pada bab ini membahas mengenai evaluasi atau *evaluation* yang dilakukan terhadap *association rule* dengan menggunakan konsep yang sama seperti tahap *modelling*. Pada tahap evaluasi berisi uraian mengenai *evaluate result* dan *evaluate process*.

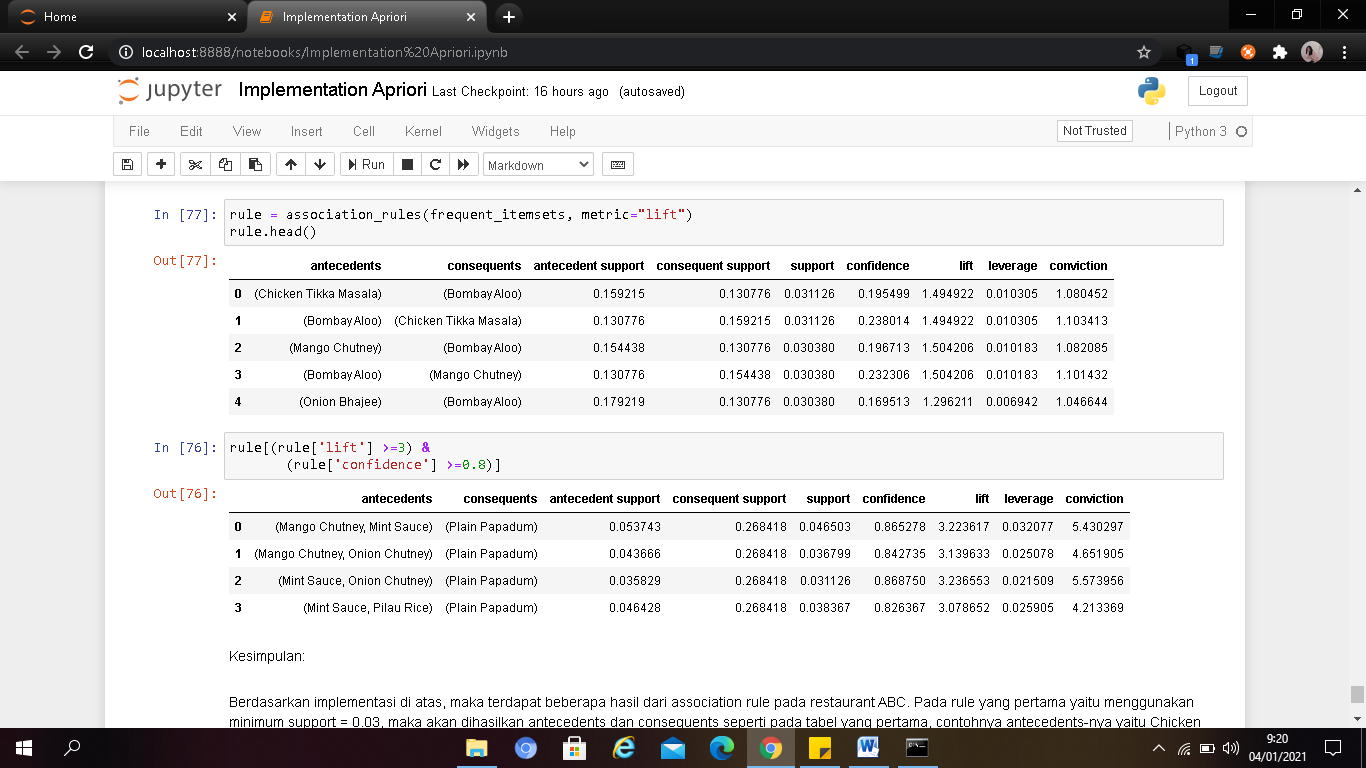
## ***Evaluate Result***

Pada subbab ini akan dijabarkan hasil evaluasi dari *association rule* nama-nama menu pada restoran dengan algoritma apriori. Pengujian dilakukan menggunakan 74818 *record* data tidak termasuk *header*, dimana *rule* atau aturan dilakukan sebanyak dua kali. Pada *rule* yang pertama seperti pada , digunakan *minimum support* sebanyak 0.03 atau 3% dengan menggunakan *metric lift*.



Gambar 17. Hasil Evaluasi Rule Pertama

Pada *rule* yang kedua seperti pada , *metric* yang digunakan berupa *metric lift* dan *metric confidence* dimana *metric lift* yang digunakan lebih besar sama dengan 3 dan *metric confidence* yang digunakan lebih besar sama dengan 0.8.



Gambar 18. Hasil Evaluasi Rule Kedua

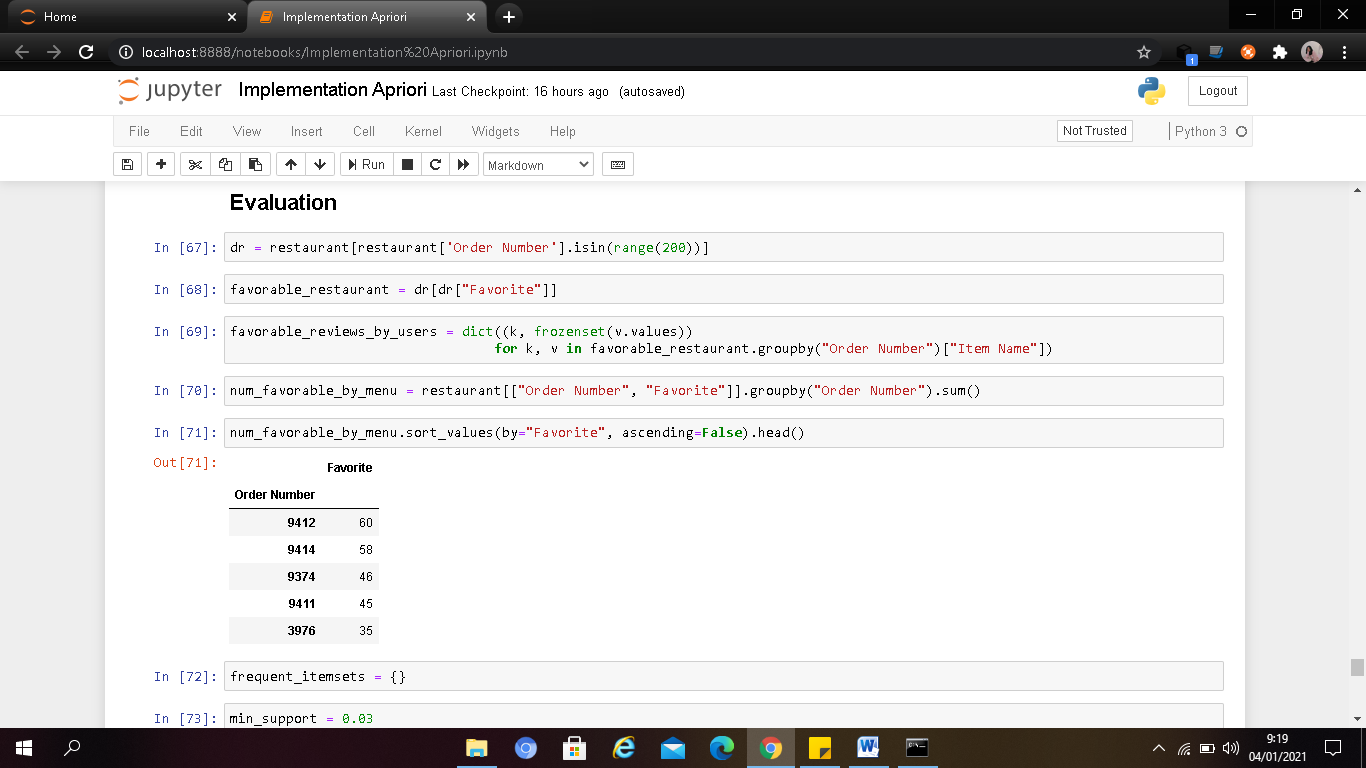
Berdasarkan hasil evaluasi dari kedua *rule* di atas, maka dapat disimpulkan bahwa *antecedents* dan *consequents* yang didapatkan juga berbeda mengikuti *rule* yang ditetapkan. Pada *rule* yang pertama, *item* yang berada pada *antecedents* dan *consequents* tersebut hanya berisi masing-masing satu *item*. Contohnya, apabila *user* memesan *item* Chicken Tikka Masala, maka secara bersamaan *item* Bombay Aloo juga akan dipesan oleh *user*.

Sedangkan pada *rule* yang kedua, *item* yang berada pada *antecedents* dan *consequents* tersebut sudah berbeda dari *rule* yang pertama. *Antecedents* yang berada pada *rule* yang kedua terdiri dari dua *item*, namun *consequents*-nya tetap berisi satu *item*. Contohnya, apabila *user* memesan *item* Mango Chutney dan Mint Sauce, maka secara bersamaan *item* Plain Papadum juga akan dipesan oleh *user*.

## ***Evaluate Process***

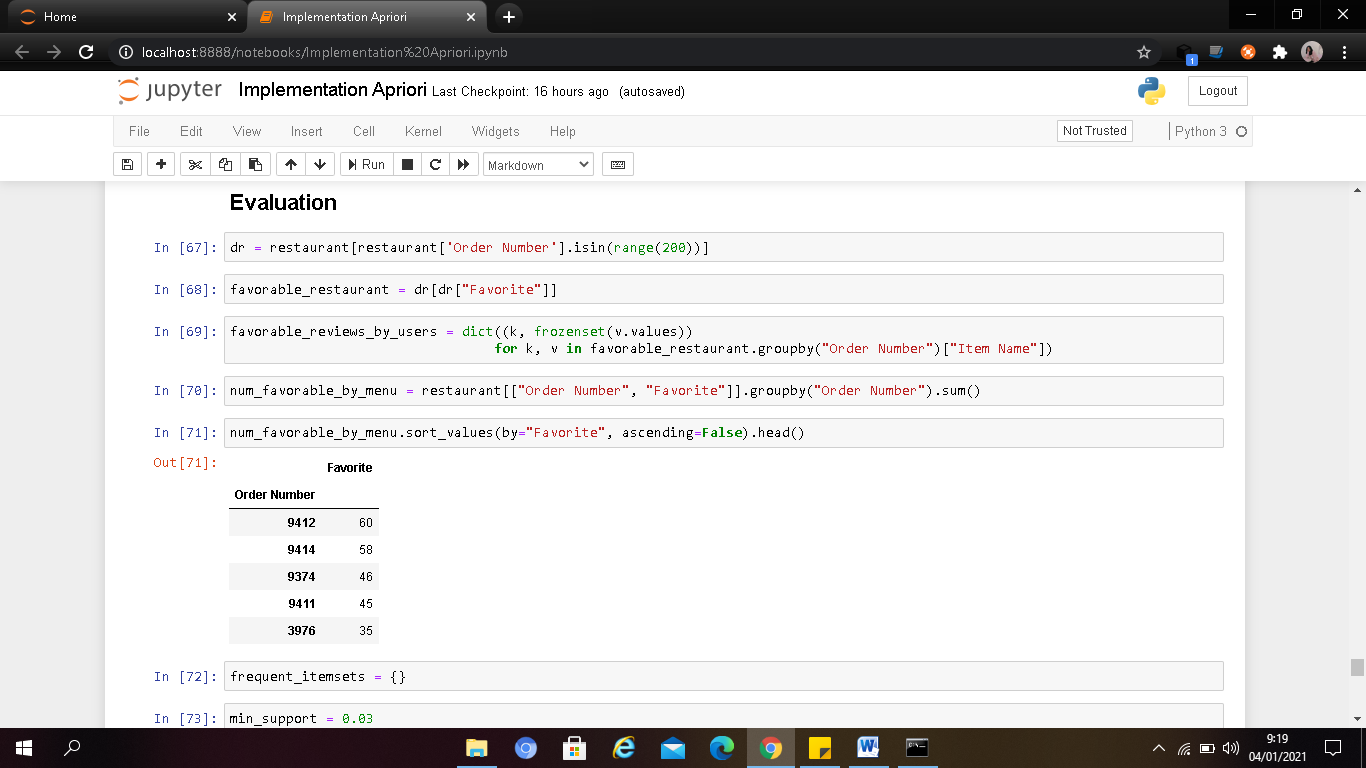
Pada subbab ini akan dijabarkan mengenai proses-proses yang dilakukan terhadap evaluasi yang sudah didapatkan. Secara garis besar, langkah-langkah yang dilakukan pada proses evaluasi hampir sama dengan langkah-langkah yang dilakukan pada proses *data preprocessing* dan *modelling*.

Pada langkah pertama, dilakukan pengambilan *data set* dengan mengurangi ukuran *data set* yang akan dicari sehingga membuat algoritma apriori dapat berjalan dengan lebih cepat. Pada kasus ini, datayang akan digunakan sebesar 200 data.



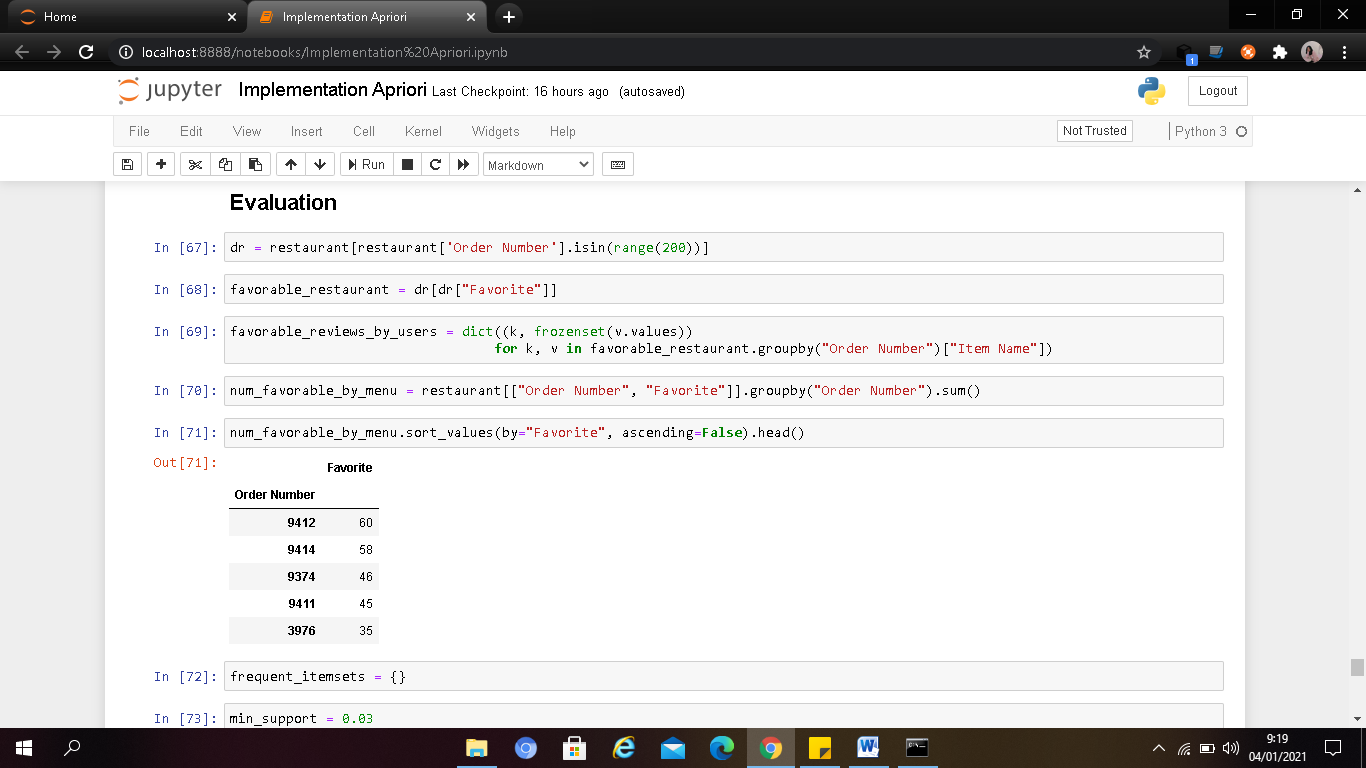
Gambar 19. Potongan Kode Pengambilan Data Set

Kemudian, *data set* tersebut diambil berdasarkan *item* yang paling banyak dipesan atau terfavorit oleh *user*.



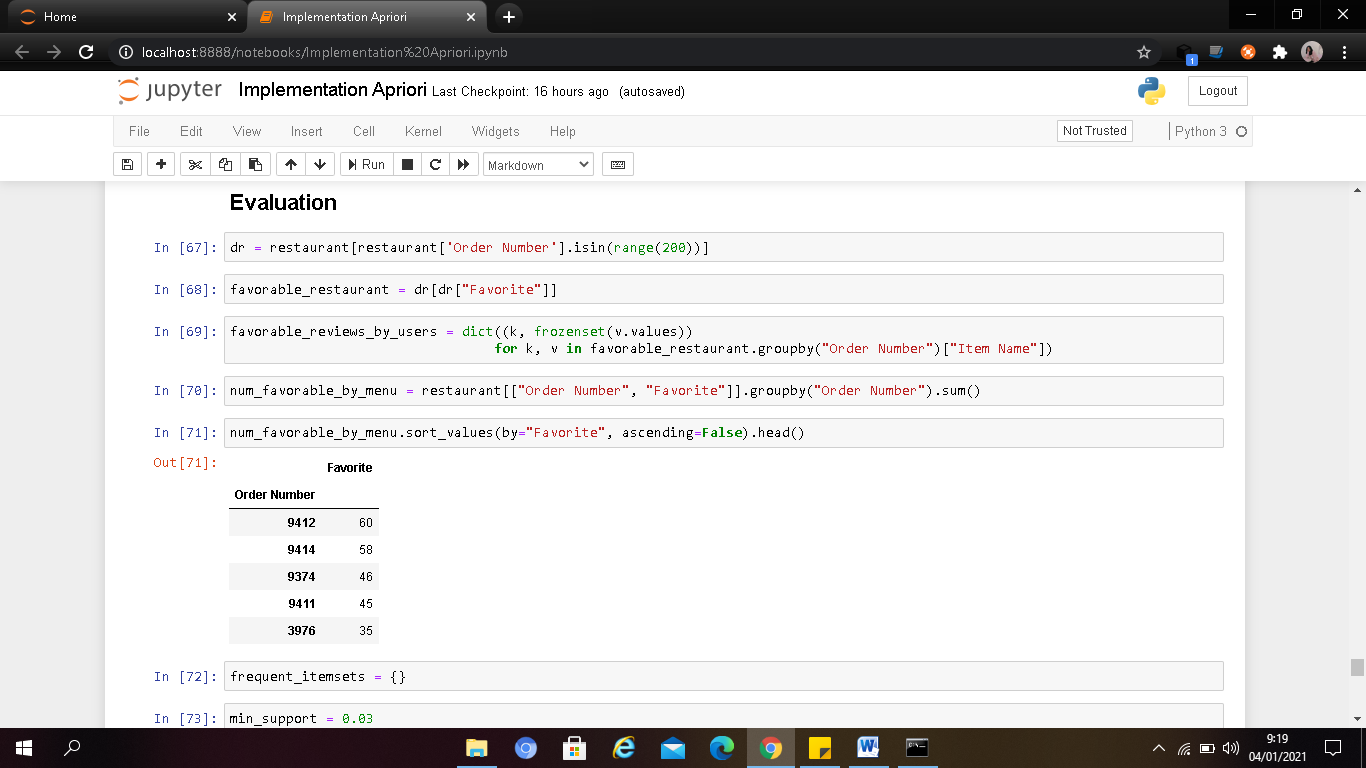
Gambar 20. Potongan Kode Item Terfavorit

Lalu, ulasan terkait *item* terfavorit akan dihitung dengan mengelompokkan *data set* tersebut berdasarkan *Order Number* dan *Item Name*.Dalam hal ini, nilai akan disimpan sebagai *frozenset* untuk memeriksa suatu *item* merupakan *item terfavorit* dengan lebih cepat.



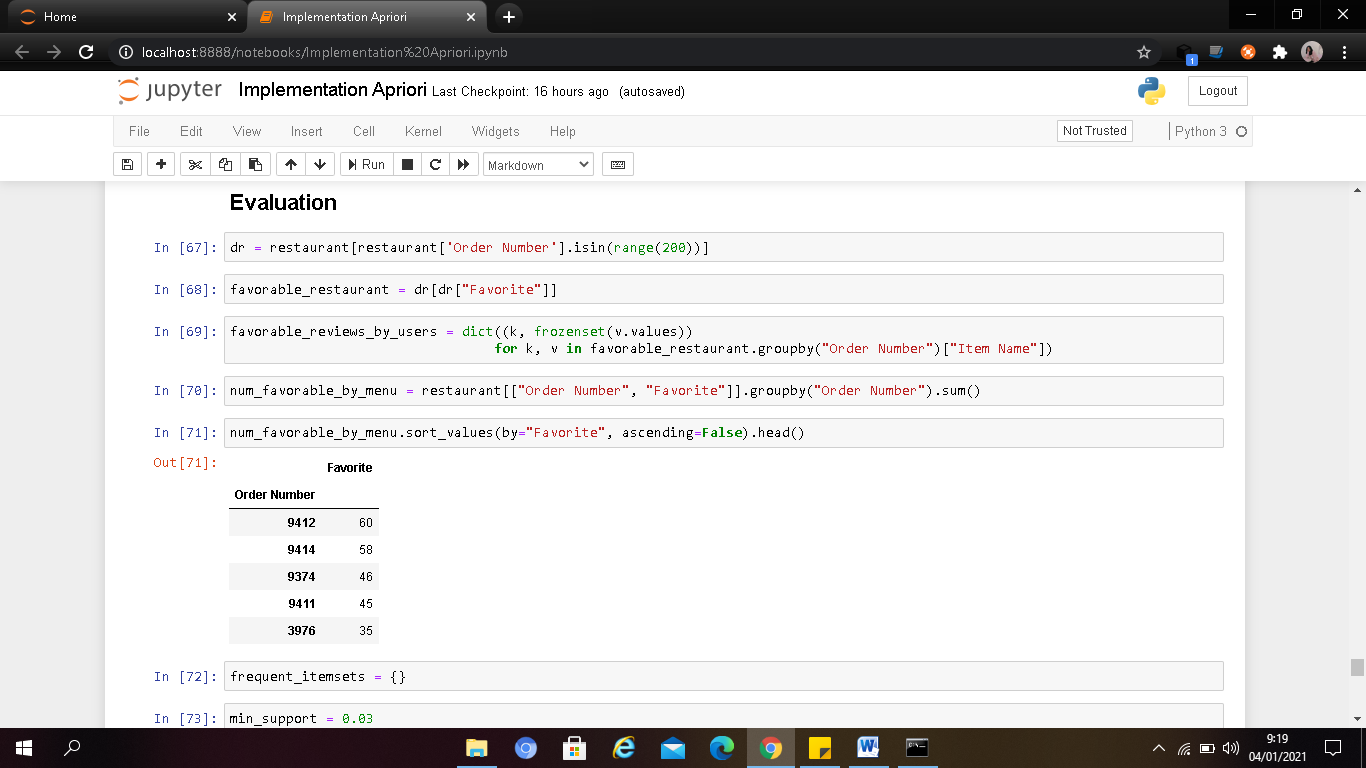
Gambar 21. Potongan Kode Frozenset

Selanjutnya, dilakukan pembuatan *data frame* untuk memberitahukan seberapa sering *item* telah dipesan oleh *user*.



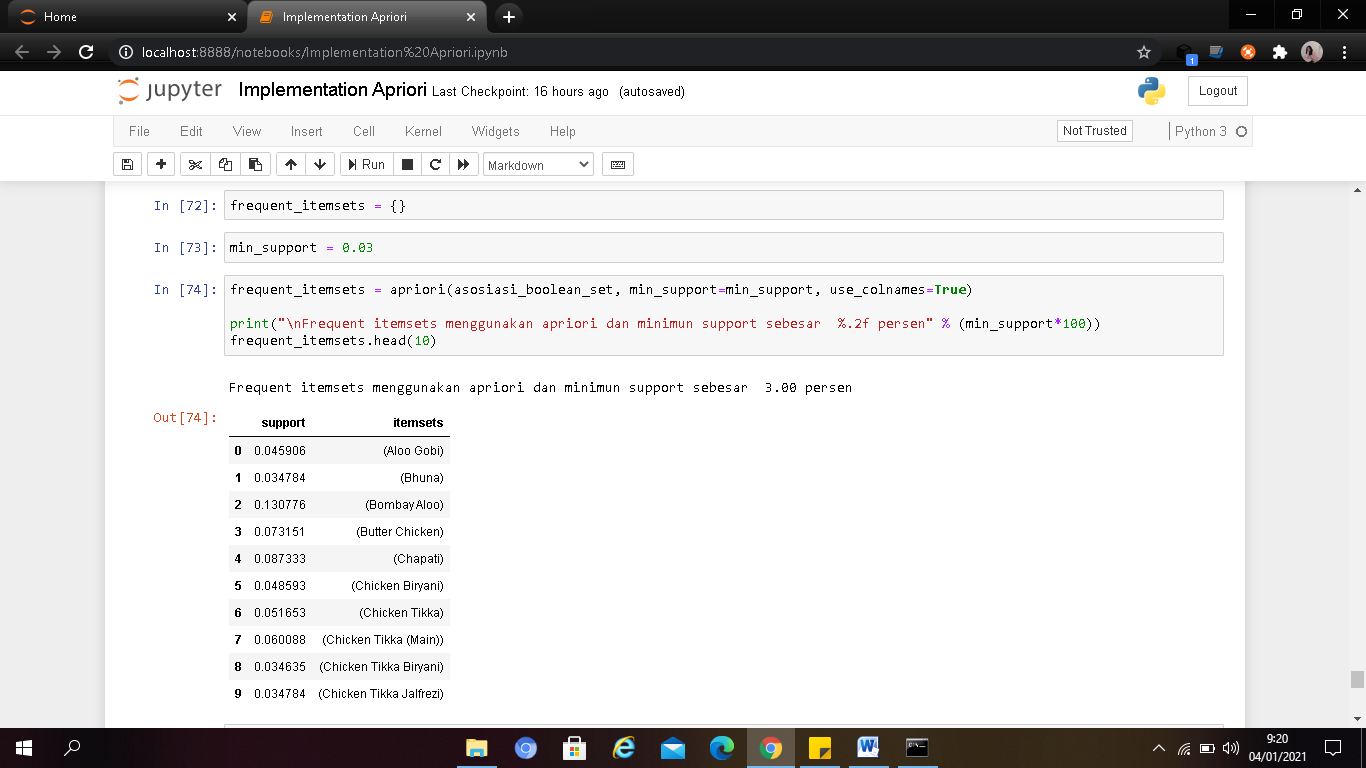
Gambar 22. Potongan Kode Pembuatan Data Frame

Pada potongan kode dibawah, dapat dilihat lima *item* terfavorit atau *item* yang paling banyak dipesan.



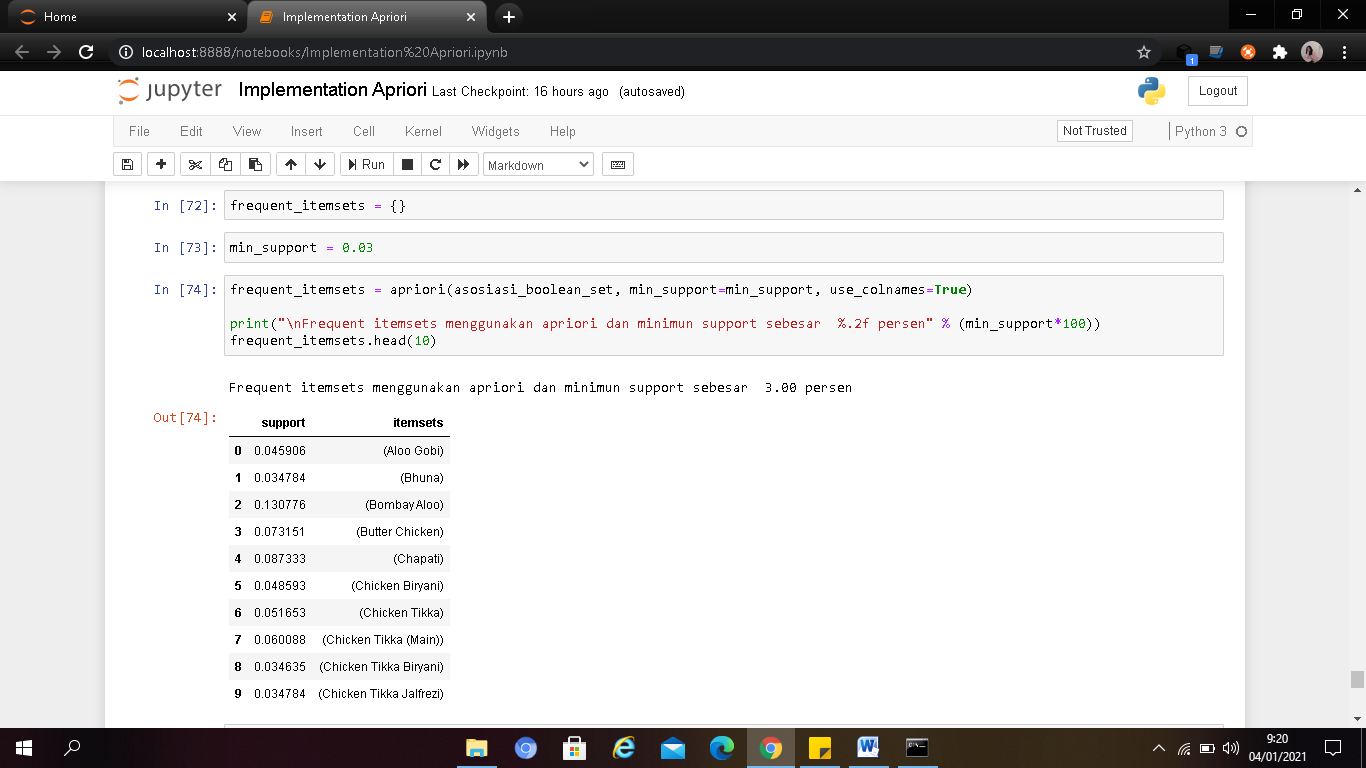
Gambar 23. Potongan Kode Daftar 5 Item Terfavorit

Pada langkah selanjutnya, *item* yang sering ditemukan dalam *dictionary* dapat disimpan apabila *key item* tersebut berupa panjang dari *itemset*. Hal ini memungkinkan *user* untuk mengakses *itemset* dengan panjang tertentu dengan lebih cepat.



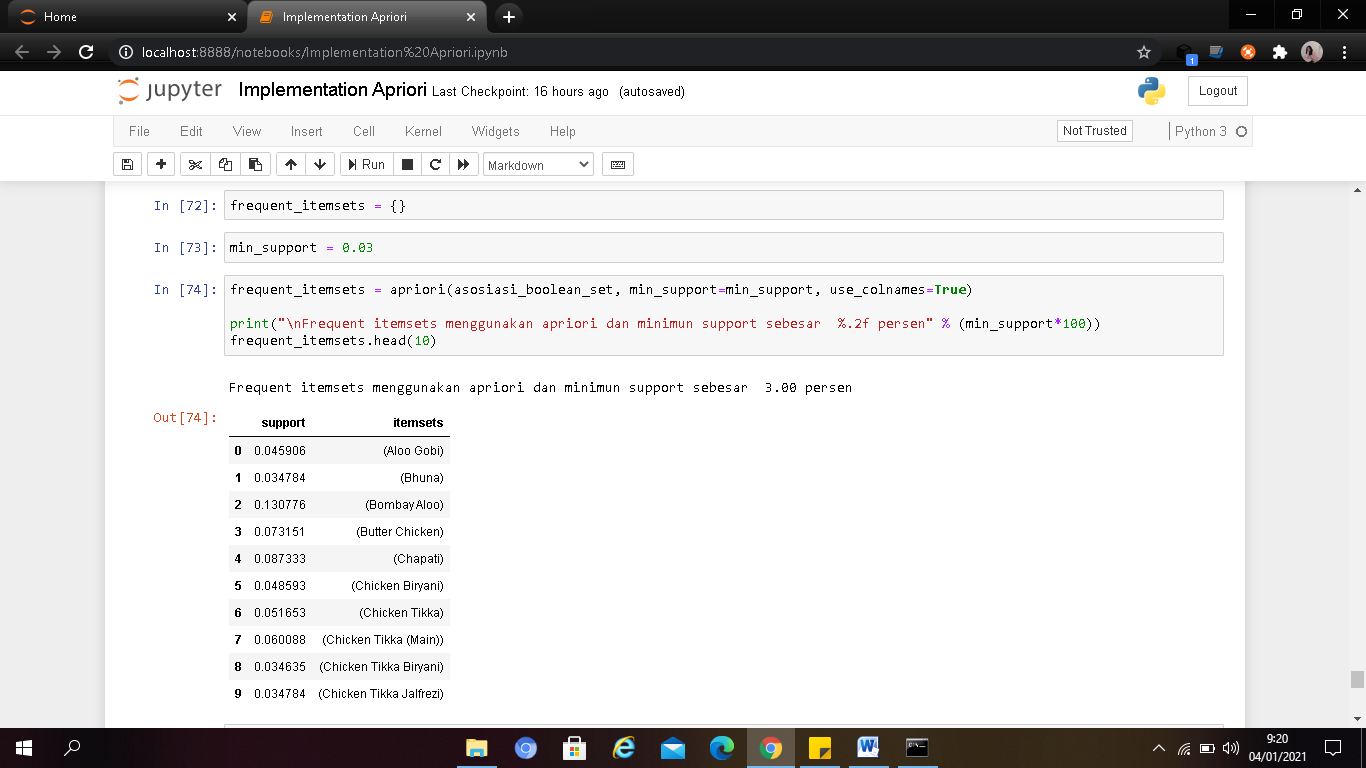
Gambar 24. Potongan Kode Frequent Itemsets

Selanjutnya dilakukan penentuan *minimum support* untuk mempertimbangkan frekuen pada suatu *item*. Dalam hal ini, *minimum support* yang digunakan sebesar 0.03 atau 3%.



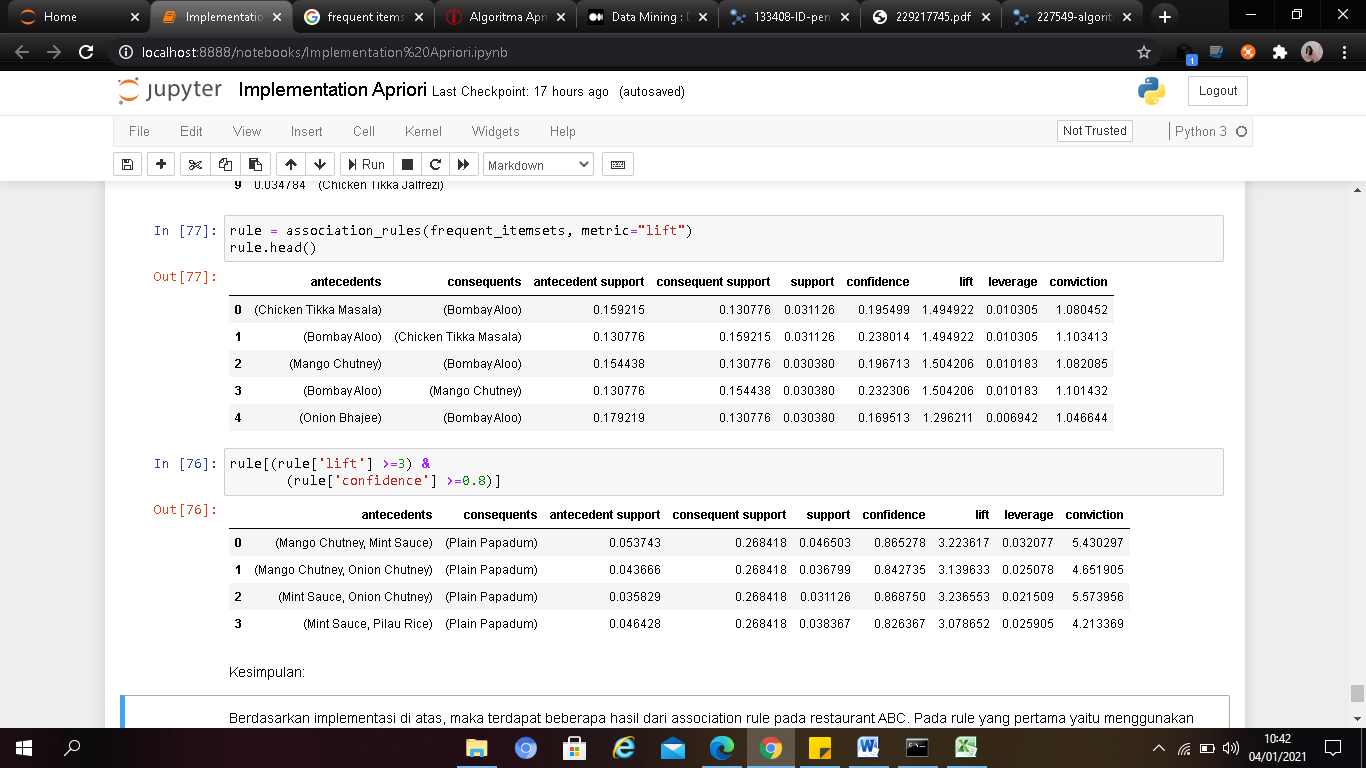
Gambar 25. Potongan Kode Minimum Support

Pada potongan kode dibawah, dapat dilihat 10 *itemsets* yang muncul berdasarkan *minimum support* yang sudah ditetapkan sebelumnya.



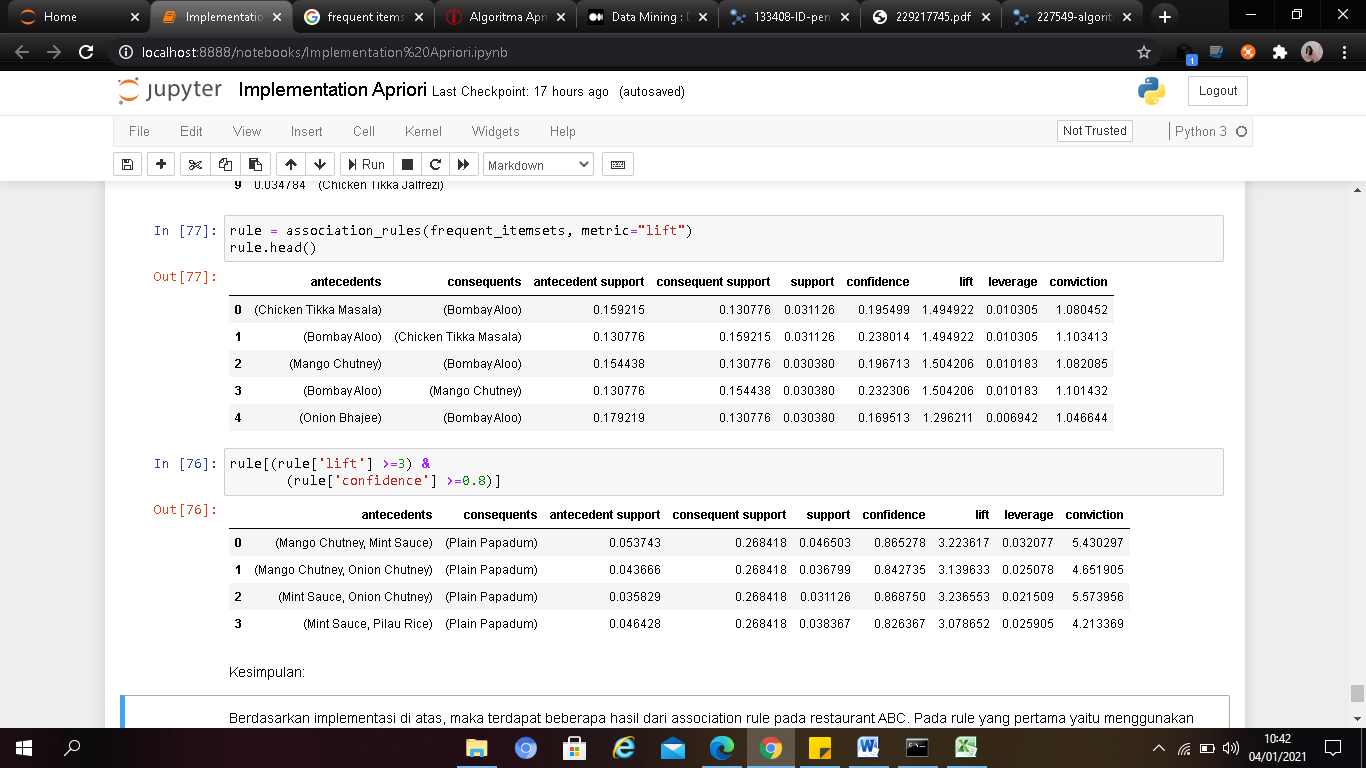
Gambar 26. Potongan Kode Daftar 10 Itemsets

Setelah dilakukan penentuan *minimum support*, maka dapat dibentuk aturan asosiasi (*association rule*) berdasarkan *frequent itemsets* yang sudah dijabarkan sebelumnya. Pada hal ini, *rule* yang ditetapkan terdiri dari dua *rule*. Pada *rule* yang pertama menggunakan *metric lift* dengan *minimum support* sebesar 0.03 atau 3%. Penjelasan mengenai hasil evaluasi dari *association rule* yang pertama ini sudah dijabarkan pada subbab .



Gambar 27. Potongan Kode Rule Pertama

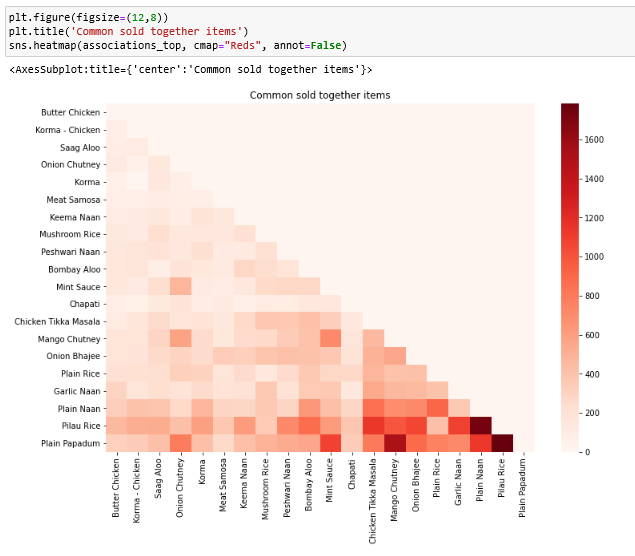
Pada *rule* yang kedua menggunakan *metric lift* berupa lebih besar sama dengan 3 dan *metric confidence* berupa lebih besar sama dengan 0.8 sehingga *antecedents* dan *consequents* yang dihasilkan pada *rule* yang kedua relatif berbeda dengan *antecedents* dan *consequents* yang dihasilkan pada *rule* yang pertama. Penjelasan mengenai hasil evaluasi dari *association rule* yang kedua ini sudah dijabarkan pada subbab .



Gambar 28. Potongan Kode Rule Kedua

# Bab VII *Deployment*

Pada bab ini membahas mengenai tahapan pengembangan sistem yang dapat dilakukan untuk rencana pengembangan selanjutnya. Pada proyek yang telah dibangun sudah memiliki visualisasi seperti *heatmap* dan *bar chart* yang sudah dijelaskan pada bab 3. Dimana visualisasi tersebut membantu restoran untuk menentukan kemungkinan suatu *item* akan dibeli bersamaan dengan *item* lainnya.



Gambar 29. Visualisasi Heatmap

Adapun beberapa pengembangan yang dapat dilakukan, yaitu :

1. Pengembangan *interface* untuk memudahkan komunikasi antara pengguna dengan sistem dengan penambahan fungsi-fungsi pada *interface* sesuai dengan keinginan pengguna. Seperti menyediakan *interface* untuk memasukkan nilai *support, confidence*, maksimum *item* dan *lift* yang akan digunakan.
2. Peningkatan kinerja sistem.
3. Dapat dilakukan pemeliharaan sistem atau *maintenance* untuk dilakukan perbaikan-perbaikan terhadap sistem yang sedang berjalan.
4. Penerapan pada sistem *database* dengan menggunakan *server* yang lebih besar sehingga proses *generate data* akan lebih cepat dan hasil analisis yang didapat akan lebih akurat.
5. Pembuatan dokumentasi berupa laporan dari program yang telah dibuat ataupun kelebihan dan kekurangan yang terdapat pada program sebagai bentuk pengembangan sistem selanjutnya.
6. Diharapkan untuk penelitian selanjutnya dapat menerapkan algoritma pembangkitan *frequent itemset* yang berbeda, seperti FP-Growth.

# Bab VIII Kesimpulan

Adapun kesimpulan yang telah diperoleh olaeh tim proyek setelah melakukan penelitian ini yakni :

1. Metode *Association Rule* dengan algoritma Apriori ini dapat dimanfaatkan untuk membantu pihak pengelola rumah makan memperoleh serta mengetahui pola beli konsumen dan item menu hidangan yang sering dibeli oleh konsumen.
2. Metode *Association Rule* dengan Algoritma Apriori ini dapat meningkatkan penjualan barang pada rumah makan abc melalui analisis menu mana saja yang sering dibeli bersamaan sehingga pengelola rumah makan dapat menyusun rancangan strategi bisnis maupun melakukan promosi terhadap menu yang kemungkinan dipesan bersamaan. Informasi yang saling terkait dengan pelaksanaan promosi dapat tersedia dengan cepat, sehingga pihak pengelola dapat dengan cepat melakukan pengambilan keputusan terhadap bisnis.
3. Penggunaan metode *Association Rule* mempunyai kelemahan yakni membutuhkan jumlah data yang harus mencapai ribuan bahkan jutaan transaksi untuk memperoleh hasil yang maksimal. Sehingga untuk suatu usaha dengan transaksi kurang banyak akan mendapatkan hasil yang diperoleh kurang maksimal.

# Bab IX Lampiran

**Poster**



# REFERENSI

|  |  |
| --- | --- |
| [1] | R. Putra, “PENERAPAN ASSOCIATION RULE DENGAN ALGORITMA A-PRIORI UNTUK SISTEM REKOMENDASI PADA PERUSAHAAN E-COMMERCE X,” UNIVERSITAS KATOLIK PARAHYANGAN, Bandung, 2018. |
| [2] | W. A. Triyanto, “ASSOCIATION RULE MINING UNTUK PENENTUAN REKOMENDASI PROMOSI PRODUK,” Universitas Muria Kudus, 2014. |
| [3] | R. A. H. A. M. Nugroho Wand, “Pengembangan Sistem Rekomendasi Penelusuran Buku Dengan Penggalian Association Rule Menggunakan Algoritma Apriori (Studi Kasus Badan Perpustakaan Dan Kearsipan Provinsi Jawa Timur),” Institut Teknologi Sepuluh Nopember (ITS), Surabaya, 2012. |
| [4] | T. Kristanto and R. Arief, "Open Access Journal of Information Systems (OAJIS)," 02 December 2013. [Online]. Available: http://is.its.ac.id/pubs/oajis/index.php/home/detail/503/Analisa-Data-Mining-Metode-Fuzzy-untuk-Customer-Relationship-Management-pada-Perusahaan-Tour-Travel. [Accessed 16 November 2020]. |
| [5] | Priati, "Penerapan Data Mining Pada Data Transaksi Superstore Untuk Mengetahui Kemungkinan Pelanggan Membeli Product Category Dan Product Container Secara Bersamaan Dengan Teknik Asosiasi Menggunakan Algoritma Apriori," *Jurnal Ilmu Komputer & Teknologi Informasi,* vol. 1, no. 2, pp. 11-18, 2016. |
| [6] | E. D. Wahyuni, A. A. Arifiyanti and M. Kustyani, "Exploratory Data Analysis dalam Konteks Klasifikasi Data Mining," in *Prosiding Nasional Rekayasa Teknologi Industri dan Informasi XIV Tahun 2019 (ReTII)*, Jawa Timur, 2019. |
| [7] | D. Prabowo dan F. Ramdani, “PENERAPAN ALGORITMA APRIORI UNTUK REKOMENDASI BUKU PADA AMIKOM RESOURCE CENTER,” *InformationSystem Journal (INFOS),* vol. III, no. 1, p. 9, 2020. |
| [8] | S. F. Rodiansyah, “Algoritma Apriori untuk Analisis Keranjang Belanja pada Data Transaksi Penjualan,” *Infotech Journal,* vol. I, no. 2, p. 37, 2016. |
| [9] | E. Buulolo, “IMPLEMENTASI ALGORITMA APRIORI PADA SISTEM PERSEDIAAN OBAT (STUDI KASUS : APOTIK RUMAH SAKIT ESTOMIHI MEDAN),” *Pelita Informastika Budi Darma,* vol. IV, no. 1, p. 74, 2013. |
| [10] | D. Prabowo dan F. Ramdani, “PENERAPAN ALGORITMA APRIORI UNTUK REKOMENDASI BUKU PADA AMIKOM RESOURCE CENTER,” *Informastion System Journal (INFOS),* vol. III, no. 1, p. 10, 2020. |
| [11] | R. S. W. Indra Purnama, “IMPLEMENTASI DATA MINING MENGGUNAKAN CRISP-DM PADA SISTEM INFORMASI EKSEKUTIF DINAS KELAUTAN DAN PERIKANAN PROVINSI JAWA TENGAH,” Universitas Diponegoro, Jawa Tengah. |
| [12] | b. W. McKinney, Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython, United States of America: O'REILLY, 2018. |
| [13] | V. Leonardo, L. W. Santoso dan A. N. Tjondrowiguno, “Sistem Rekomendasi Item Pada Game Dota 2 dengan Multilayer Perceptron Neural Network,” p. 3, 2019. |