

# PEMETAAN DAN ANALISIS KANDUNGAN NUTRISI PADA MINUMAN STARBUCKSDENGAN METODE K-MEANS

*Mapping and Analysis of Nutritional Content in Starbucks Beverages Using The K-Means Methode*

Mochammad Raffi Radithya<sup>1\*</sup>, Syalma Syahrafina Nurachma<sup>1</sup>, Ratna Dian Saputri<sup>1</sup>

<sup>1</sup>Program Studi Teknik Industri, Telkom University, Bandung  
email: raffiradithya@student.telkomuniversity.ac.id

Received: 30 Maret 2024 | Accepted: 25 April 2024 | published: 31 Mei 2024

## ABSTRAK

Dalam industri minuman, khususnya di Starbucks yang memiliki jangkauan global, keberlanjutan pertumbuhan bergantung pada penyesuaian dengan perkembangan tren dan kebutuhan konsumen, termasuk kebutuhan nutrisi. Penelitian ini bertujuan untuk menganalisis nutrisi dalam beberapa minuman Starbucks menggunakan metode klusterisasi K-Means, dengan fokus pada identifikasi variasi nutrisi di antara minuman tersebut. Metode K-Means digunakan dengan input berdasarkan data sekunder yang mencakup atribut-atribut terkait. Tahapan penelitian meliputi aktivasi library, eksplorasi data melalui boxplot, preprocessing data untuk kolom numerik, tuning model untuk menentukan jumlah kluster, pembangunan model, dan visualisasi hasil. Hasil analisis menunjukkan bahwa dari 242 objek yang dianalisis, terbagi menjadi 7 kluster. Kluster tersebut memiliki variasi kandungan nutrisi yang berbeda-beda, dengan kluster tertentu menunjukkan kandungan nutrisi tertinggi dengan cluster value 1.38, nutrisi cukup tinggi dengan cluster value 0.76, nutrisi tinggi dengan cluster value 0.49, nutrisi sedang dengan cluster value 0.24, nutrisi rendah dengan cluster value -0.21, nutrisi cukup rendah dengan cluster value -0.48, dan nutrisi paling rendah dengan cluster value -0.72. Temuan ini memberikan wawasan tentang variasi nutrisi dalam minuman Starbucks, yang dapat digunakan oleh perusahaan untuk menginformasikan keputusan strategis terkait dengan formulasi produk, pemasaran, dan penargetan konsumen. Selain itu, metode analisis yang digunakan dalam penelitian ini dapat diterapkan pada industri minuman lainnya untuk analisis nutrisi yang serupa.

**Kata Kunci** : industri minuman; starbucks; kandungan nutrisi; k-means; clustering

## ABSTRACT

*In the beverage industry, especially at Starbucks, which has a global reach, sustainable growth depends on adapting to developing trends and consumer needs, including nutritional needs. This study aims to analyze the nutrients in several Starbucks drinks using the K-Means clustering method, focusing on identifying nutritional variations between these drinks. The K-Means method is used with input based on secondary data, which includes related attributes. The research stages include library activation, data exploration via boxplots, data preprocessing for numeric columns, model tuning to determine the number of clusters, model building, and results visualization. The analysis results show that of the 242 objects analyzed, they were divided into 7 clusters. These clusters have different variations in nutritional content, with specific clusters showing the highest nutritional content with a cluster value of 1.38, moderately high nutrition with a cluster value of 0.76, high nutrition with a cluster value of 0.49, medium nutrition with a cluster value of 0.24, low nutrition with a cluster value -0.21, nutrition is relatively low with a cluster value of -0.48, and the lowest nutrition is with a cluster value of -0.72. These findings provide insight into nutritional variations in Starbucks beverages, which the company can use to inform strategic decisions related to product formulation, marketing, and consumer targeting. In addition, the analytical methods used in this research can be applied to other beverage industries for similar nutritional analysis.*

**Keywords**: beverage industry; starbucks; nutritional content; k-means; clustering

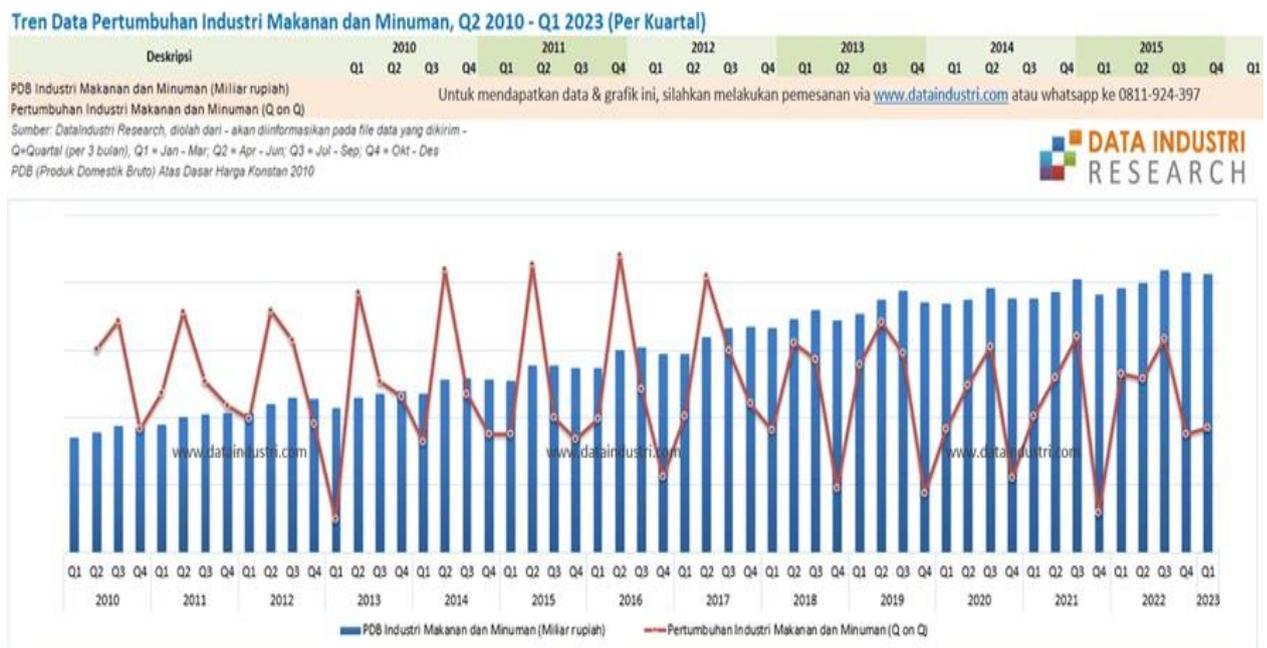
## 1. PENDAHULUAN

Berbagai macam industri semakin berkembang pada perkembangan revolusi saat ini, salah satu industri yang terus berkembang

tersebut adalah industri makanan dan minuman, yang dimana saat terdapat kenaikan pertumbuhan industri makanan dan minuman yang ditunjukkan dengan adanya data grafik di setiap kuartal fiskalnya atau dalam periode yang ditunjukkan dengan adanya data grafik di

setiap kuartal fiskalnya atau dalam periode tiga bulan yang disimbolkan dengan Q1 untuk bulan Januari, Februari, dan Maret. Lalu Q2 untuk bulan April, Mei, dan Juni. Q3 untuk bulan Juli, Agustus, dan September. Serta Q4 untuk bulan Oktober, November, dan Desember. Grafik pada Gambar 1

menunjukkan bahwa pada setiap perpindahan Q2 atau pun Q3 serta perpindahan Q4 pada tahun 2022 menuju Q1. Pada tahun 2023 pun juga mengalami kenaikan terhadap pertumbuhan industri makanan dan minuman yang dibuktikan pada grafik pada gambar 1.



Gambar 1. Grafik Tren Data Pertumbuhan Industri Makanan dan Minuman, Q1 2010 - Q2 2023  
Sumber: Data Industri Research (2011 - 2023)

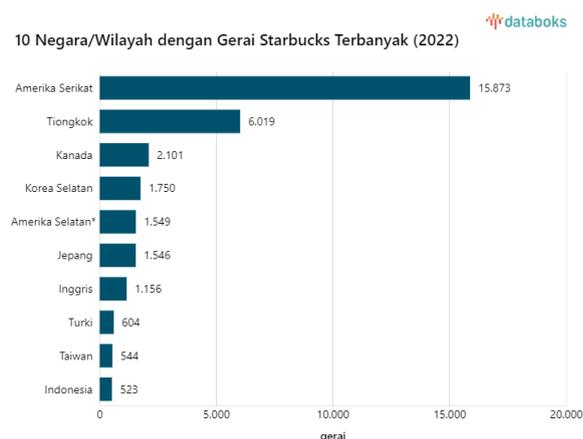
Peningkatan pada tren pertumbuhan industri makanan dan minuman menjadikan industri pada bidang tersebut perlu memperhatikan persaingan usaha yang ada, khususnya di Indonesia. Berdasarkan laporan tahunan Komisi Pengawasan Persaingan Usaha (KPPU) periode 2021 yang dipublikasikan pada situs resminya di April 2022, KPPU melakukan pengukuran Indeks Persaingan Usaha (IPU) dengan sistem skor 1-7 dengan skor 1 adalah yang tingkat persaingan yang rendah dan skor 7 adalah tingkat persaingan yang tinggi. KPPU melaporkan bahwa “Nilai IPU tahun ini (2021) juga semakin mendekati Target Nasional yang tertuang dalam Rencana Pembangunan Jangka Menengah Tahun 2024, yakni Indeks Persepsi Persaingan Usaha 5,0 poin”. Pernyataan tersebut merupakan laporan yang dinyatakan berdasarkan grafik terkait indeks persaingan usaha Indonesia di tahun 2018 sampai 2021 pada Gambar 2.



Gambar 2. Indeks Persaingan Usaha Indonesia  
Sumber: (Databoks, 2022)

Persaingan usaha yang berada di Indonesia tersebut akan membuat industri makanan dan minuman mengadakan berbagai macam rasa serta kandungan didalam makanan dan minuman pun yang menjadi semakin bervariasi untuk dapat memenuhi keinginan, kebutuhan, dan kepuasan konsumen. Salah satu industri minuman adalah Starbucks. Menurut Royani et al. (2023) Starbucks merupakan

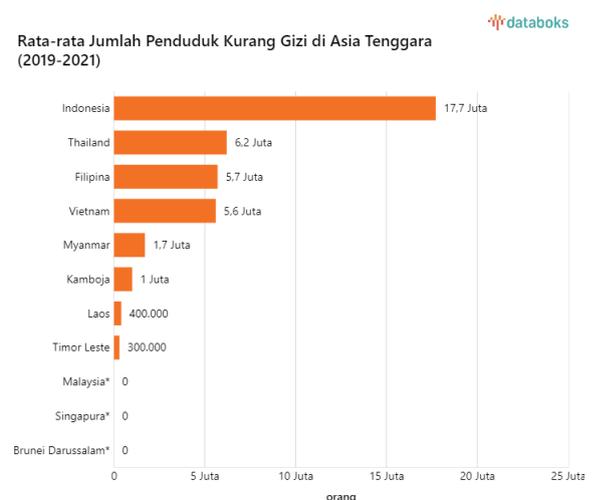
perusahaan kedai kopi terbesar di dunia dengan berbagai minuman panas dan dingin yang dijual dipasaran, serta ditemukan bahwa Starbucks Corporation adalah perusahaan kopi dalam jaringan kedai global yang berasal dari Amerika Serikat yang terpusat pada Seattle, Washington dengan jumlah gerai sebanyak 20.336 di 61 negara. Tidak hanya menjual minuman, Starbucks pun juga menjual makanan dan barang-barang untuk menarik perhatian konsumen. Didukung berdasarkan data Macrotrends pada pendapatan Starbucks di tahun 2017 sampai 2021, Starbucks mencapai pendapatan sebesar US \$29,06 miliar atau sekitar Rp415,6 triliun pada 2021 yang merupakan peningkatan pendapatan 23,6% dibandingkan tahun sebelumnya sebesar US \$23,5 miliar (Databoks, 2022). Data tersebut menunjukkan bahwa pada tahun 2021 perusahaan Starbucks memiliki kenaikan customers sehingga menjadikan pendapatan Starbucks meningkat. Hal ini menjadi alasan yang memperkuat kami untuk menjadi Starbucks sebagai objek penelitian kami. Selain itu, Indonesia adalah negara yang termasuk kedalam sepuluh negara atau wilayah dengan gerai Starbucks terbanyak pada tahun 2022 sebagaimana yang ditunjukkan pada data grafik Gambar 3.



Gambar 3. Data Gerai Starbucks Terbanyak  
Sumber: (Databoks, 2022)

Gambar 3 menunjukkan bahwa konsumen dari Indonesia adalah salah satu dari 10 besar dengan gerai Starbucks terbanyak dari seluruh negara yang berpengaruh terhadap banyaknya gerai. adanya berbagai variasi rasa dan kandungan nutrisi di dalam minuman Starbucks. Namun, dengan berbagai variasi rasa dan kandungan nutrisi yang ada pada minuman

tersebut perlu diperhatikan pemenuhan kebutuhan gizi untuk konsumen. Hal tersebut perlu dipertimbangkan dikarenakan makanan dan minuman yang dibeli oleh konsumen akan sangat berpengaruh terhadap kebutuhan gizi dengan kandungan nutrisi yang didapatkan. Nutrisi merujuk pada zat-zat yang dibutuhkan oleh tubuh untuk fungsi normal, pertumbuhan, perkembangan, dan pemeliharaan kesehatan secara keseluruhan (Vandenbrink *et al.*, 2020). Kandungan nutrisi pada minuman pun perlu diperhatikan dengan baik untuk dapat memenuhi kebutuhan konsumen dengan implikasinya agar dapat membantu konsumen dalam memilih minuman bergizi, terutama negara Indonesia yang menjadi negara pertama dari beberapa negara lainnya dengan rata-rata jumlah penduduknya yang kurang gizi di Asia Tenggara pada tahun 2019 sampai 2021 yang ditunjukkan pada Gambar 4.



Gambar 4. Grafik Rata-rata Jumlah Penduduk Kurang Gizi di Asia Tenggara  
Sumber: (Databoks, 2022)

Berdasarkan sinkronisasi perkembangan industri minuman yang dapat terus meningkat dengan kebutuhan nutrisi untuk konsumen menjadikan rincian situasi dan kondisi menjadi fokus kami dalam penelitian ini memberikan wawasan berharga untuk menganalisis nutrisi yang terdapat pada beberapa minuman Starbucks dalam cluster minuman nya berdasarkan data sekunder atributnya, yaitu kalori, gula, protein, dan kafein menggunakan metode K-Means.

Penggunaan metode K-means digunakan karena memiliki keunggulan dengan merupakan algoritma dalam *clustering* yang sederhana dibandingkan metode lainnya, K-Means pun

memiliki beberapa keunggulan seperti penerapan yang mudah dan relatif cepat, beradaptasi dengan baik dan mudah, serta paling banyak dalam implementasi penugasan data mining (Abbas *et al.*, 2022; Achmad *et al.*, 2023). Diharapkan penelitian ini dapat memberikan informasi untuk membantu customer melakukan pemilihannya terhadap menu-menu yang ada di Starbucks. Selain itu, penelitian ini dapat digunakan oleh perusahaan untuk menginformasikan keputusan strategis terkait dengan formulasi produk, pemasaran, dan penargetan konsumen. Secara keseluruhan, penggunaan metode K-Means dalam penelitian ini tidak hanya memberikan wawasan baru tentang nutrisi dalam minuman Starbucks, tetapi juga memiliki dampak yang potensial dalam meningkatkan kinerja dan daya saing perusahaan (Achmad *et al.*, 2023).

## 2. METODE PENELITIAN

### 2.1 Sample and Data Collection

Penelitian ini merupakan pengumpulan data yang dilakukan dengan menggunakan metode studi literatur terhadap data sekunder terkait menu Starbucks beserta kandungan yang berada di dalamnya. Data menu tersebut sebagai sampel pada bulan Agustus – Desember 2023, yang ditunjukkan pada tabel 1.

Tabel 1. Sampel Pengamatan

No	Beverage_category	Beverage
1	Coffee	Brewed Coffee
2	Coffee	Brewed Coffe
3	Classic Espresso Drinks	Caffee Latte
...		...
...		...
...		...
241	Frappuccino Blended Caramel	Vanilla Beam
242	Frappuccino Blended Caramel	Vanilla Beam (Without Whipped Cream)

Data sebanyak 242 sampel ini memiliki peran penting di dalam penelitian karena sampel tersebut merupakan sampel awal yang akan digunakan untuk mengelompokkan menu dengan

cara *clustering* ke dalam beberapa atribut kandungan menu pada minuman Starbucks seperti kalori, gula, protein, dan kafein.

### 2.2 K-Means Clustering

*Clustering* yang merupakan proses dalam mengelompokkan data-data ke dalam berbagai kelompok serupa sehingga setiap objek yang serupa akan menjadi suatu *cluster* sedangkan setiap objek yang tidak serupa lainnya akan menjadi suatu *cluster* lainnya (Achmad *et al.*, 2023). Setiap *cluster* tersebut berisikan data-data yang semirip mungkin dengan kemiripan yang berjarak sedekat mungkin (Abbas *et al.*, 2021). Deskripsi tersebut menjadikan *clustering* berarti terdapat beberapa pengukuran atau parameter penting dalam mewakili serupa dan tidak serupanya antar cluster terkait (Muflikhah *et al.*, 2018).

Menurut Irwansyah, E. dan Faisal, M. (2015) K-Means adalah metode clustering non-hirarki, yaitu metode dalam mengelompokkan data menjadi satu atau lebih cluster. K-Means adalah algoritma clustering yang menggunakan metode partisi (*partitioning method*) berbasis pada titik pusat (*centroid*).

Berikut merupakan langkah-langkah dalam metode penelitian yang dilakukan untuk mencapai tujuan dalam klasterisasi menu Starbucks, yaitu (Abu Khurma *et al.*, 2021; Achmad *et al.*, 2023):

#### 1. Input Data

Tahap pertama yang dilakukan adalah input data, yaitu dengan mengaktifkan terlebih dahulu *library numpy, pandas, seaborn, dan matplotlib* yang akan digunakan untuk memvisualisasi data.

#### 2. Eksplorasi Data

Tahap kedua, yaitu eksplorasi data yang bertujuan untuk menampilkan dataset, identifikasi tipe data, mendeteksi keberadaan nilai yang hilang (*missing values*), dan menemukan outlier melalui visualisasi dengan menggunakan *boxplot*.

#### 3. Data Preprocessing

Tahap ketiga, yaitu data *preprocessing*. Tahap ini, dilakukan normalisasi atau standarisasi menggunakan nilai *z-score* dengan memanfaatkan fungsi *StandardScaler* untuk atribut – atributnya. Kemudian, hasil normalisasi tersebut disimpan sebagai kerangka data baru yang dinamakan “*Standardized Attributes*”

#### 4. Model *Tuning*

Tahap keempat, yaitu model *tuning*. Terdapat parameter  $k$  yang perlu dipilih sebelumnya dalam algoritma *k-means clustering*. Sebelum mengambil keputusan subjektif tentang nilai  $k$ , dapat dilakukan penyesuaian model atau model *tuning*, contohnya dengan menggunakan metode Elbow (Abu-Jamous & Nandi, 2015). Untuk melakukan analisis nilai  $k$  optimal dengan menggunakan metode Elbow, disusun sebuah program sederhana yang membuat plot antara nilai within-cluster sum of squares (Abu-Jamous & Nandi, 2015).

Proses ini diimplementasikan melalui perulangan menggunakan fungsi *for i in*. Semakin kecil nilai inersia maka semakin baik. Nilai  $k$  optimal dapat diidentifikasi melalui titik balik pada kurva, di mana penurunan inersia menjadi relatif kecil dan kurva cenderung datar. Kemudian, untuk memastikan nilai  $k$  yang optimal dilakukan analisis *Silhouette Coefficient* berdasarkan konfigurasi setiap anggota cluster (Abu-Jamous & Nandi, 2015).. Visualisasi hasilnya dapat dicapai dengan menggunakan fungsi *Silhouette Visualizer*.

#### 5. Model *Building*

Tahap Kelima, yaitu model *building*. Setelah menentukan nilai  $k$  yang dipilih, langkah berikutnya adalah membangun model *k-means clustering* dengan  $k=7$  menggunakan fungsi *Kmeans*, dengan input berupa atribut – atribut yang telah dinormalisasi sebelumnya (Abu-Jamous & Nandi, 2015). *Output* dari proses pengelompokan ini disajikan dalam bentuk dataframe yang diberi nama “hasil”.

#### 6. Model *Visualization*

Tahap keenam, yaitu model *visualization*. Hasil dari pengelompokan dapat dipresentasikan secara visual dalam bentuk plot dua dimensi. Dengan mempertimbangkan jumlah atribut sebanyak “angka”, untuk membuat plot dua dimensi dapat diterapkan teknik reduksi dimensi menggunakan analisis komponen utama (PCA).

#### 2.3 Atribut Penelitian

**Kalori:** Kalori merupakan istilah umum yang seringkali digunakan untuk satuan unit pengukur energi, yang dimana dibutuhkan sebagai sumber energi dalam melakukan aktivitas (Daniswara, 2022).

Umumnya kalori bergantung pada kandungan karbohidrat, protein, lemak, dan lainnya. Asupan kalori yang tidak sesuai untuk tubuh akan memberikan efek lemas sementara dan apabila berlebihan dapat mengakibatkan obesitas (Asih & Widyasiti, 2016).

**Gula:** Gula disebut juga sebagai kandungan monosakarida ( $C_6H_{12}O_6$ ) yang merupakan gula paling sederhana serta terdiri dari molekul tunggal (Handayani, 2023). Gula atau monosakarida juga dapat memiliki kandungan glukosa yang terdiri dari satu gugus aldehyd, yaitu disebut “aldose” serta kandungan fruktosa yang mengandung gugus keton, yaitu disebut sebagai “ketosa”.

**Protein:** Protein adalah proses katabolisme yang berasal dari kelebihan glukosa, sehingga akan disimpan menjadicadangan energi dihati dan di otot rangka dalam bentuk glikogen atau polimer kompleks yang akan diubah menjadi lemak (trigliserida) di sel adiposa (adiposit). Katabolisme protein umumnya terjadi pada hati atau organ lainnya seperti ginjal, usus kecil, otot, hingga pada jaringan adiposa (Luo et al., 2016).

**Kafein:** Kafein merupakan zat khas yang didapatkan dengan diekstrak dari tanaman, kafein berasal dari bahan yang biasanya dapat ditemukan pada butir kopi, daun teh, serta butir kakao (Aina, 2021). Kafein adalah zat psikoaktif dengan efek stimulan dan digunakan di seluruh dunia. Kafein memiliki kandungan sedikit nutrisi, namun mengandung lebih banyak bahan kimia alami seperti karbohidrat, lipid, senyawa nitrogen, vitamin, mineral, alkaloid, dan senyawa fenolik (Spiller dalam Wachmo, 2017)

### 3. HASIL DAN PEMBAHASAN

Tujuan penelitian ini diadakan untuk dapat menyelidiki dan menganalisis kandungan pada menu Starbucks yang dapat membantu konsumen dalam memilih minuman bergizi dengan menggunakan metode *K-Means Clustering*. Identifikasi dalam klasterisasi menu Starbucks menggunakan beberapa atribut, yaitu kalori, gula, protein, dan kafein. Dengan adanya penelitian ini, diharapkan dapat membantu konsumen untuk dapat melakukan pemilihan menu Starbucks yang sesuai dengan kepuasan pelanggan tanpa

melupakan kebutuhan dan pemenuhan gizi yang terkandung. Berikut adalah langkah-langkah yang dilakukan dalam melakukan pengolahan data.

### 3.1 Input Data

Langkah pertama dalam melakukan input data yaitu melakukan aktivasi terlebih dahulu terhadap *library* pandas dan numpy serta seaborn dan matplotlib untuk melakukan visualisasi data terhadap dataset yang akan di input. Gambar 5 merupakan input data dimana pada tahap ini memasukkan dataset yang sudah didapat ke dalam Google Colab.

```
[1] import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline
```



Gambar 5. Tahapan Input Data

### 3.2 Eksplorasi Data

Gambar 6 merupakan tahap eksplorasi data *frame* dimana ini digunakan untuk menampilkan *dataset*, jenis data, mengetahui ada atau tidak adanya *missing values*, serta mengidentifikasi *outlier*, *boxplot*. Lalu, pada Gambar 6 merupakan *output dataframe* berupa *calories*, *sugars (g)*, *protein(g)*, *caffeine (mg)*.

Kemudian, dilakukan untuk membaca data dari file Excel, yaitu 'Kandungan Menu Starbucks.xlsx' dengan menghasilkan *output* berupa '*df*' atau data *frame* yang merupakan struktur data yang sebelumnya di input datanya ke *library pandas* sehingga akan menghasilkan *output* sebagai berikut dengan mencetak data *frame* tersebut dengan menuliskan '*print(df)*'.

```
df=pd.read_excel('Kandungan Menu Starbucks.xlsx')
print(df)
```

	Calories	Sugars (g)	Protein (g)	Caffeine (mg)
0	3	0	0.3	175
1	4	0	0.5	260
2	5	0	1.0	330
3	5	0	1.0	410
4	70	9	6.0	75
...	...	...	...	...
237	320	64	5.0	0
238	170	38	4.0	0
239	200	38	3.0	0
240	180	35	3.0	0
241	240	55	5.0	0

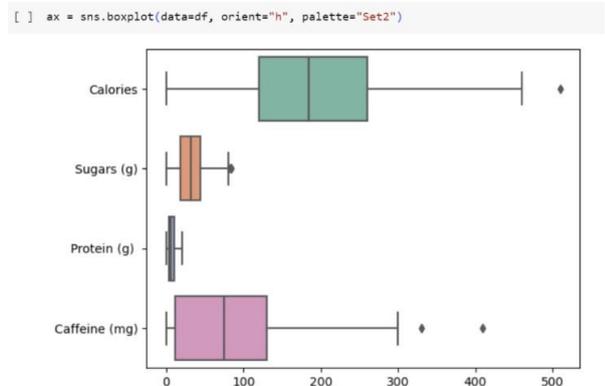
[242 rows x 4 columns]

Gambar 6. Tahapan Eksplorasi Data

Setelah dihasilkan cetakan atau *output* berupa data *frame* dari file 'Kandungan Menu Starbucks.xlsx', selanjutnya adalah melakukan identifikasi *outlier*. Identifikasi *outlier* tersebut akan digunakan melalui *output* berupa *boxplot* yang berasal dari '*import seaborn as sns*' sebelumnya dan membuat *boxplot* dengan '*data=df*' untuk mengidentifikasi data yang berasal dari data *frame*, '*orient="h"*' untuk mengorientasikan *boxplot* secara *horizontal*, dan '*palette="Set2"*' adalah penentuan *output* warna palet yang ditunjukkan pada Gambar 7.

```
[ ] df.dtypes
df.isnull().sum()
```

Calories	0
Sugars (g)	0
Protein (g)	0
Caffeine (mg)	0
dtype:	int64



Gambar 7. Output Eksplorasi Data Awal

Berdasarkan *boxplot* diatas, diketahui bahwa terdapat beberapa *outlier* pada kandungan nutrisi dari beberapa menu Starbucks. *Outlier-outlier* tersebut dapat dilihat di hampir semua kategori kecuali protein. Setelah diteliti, diketahui bahwa *outlier* tersebut berasal dari menu seperti *Signature Espresso Drinks* dengan tambahan *White chocolate mocha (without whipped cream)* dan *2% milk*, *frappuccino blended coffee* dengan tambahan *Java Chip* dan *Venti NonfatMilk* serta beberapa menu lainnya. Hal tersebut juga menandakan bahwa pada menu-menu yang menjadi *outlier* merupakan menu yang memiliki kandungan yang berlebihan seperti kalori, karbohidrat, dan lain-lain jika dibandingkan dengan kandungan menu lainnya.

### 3.3 Data Preprocessing

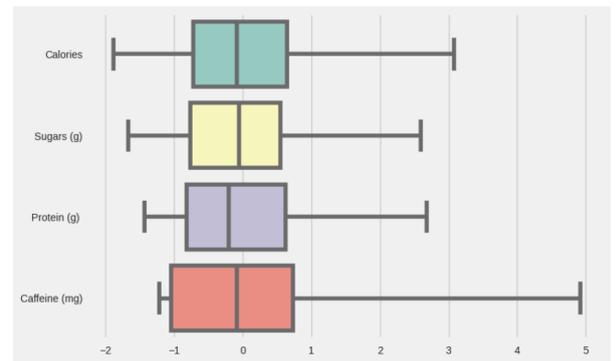
Tahap selanjutnya adalah *preprocessing* yang bertujuan untuk menghilangkan *outlier* dengan menggunakan normalisasi. Gambar 8 dan 9 merupakan tahap dan *output* dilakukan normalisasi atau standarisasi *z-score* dengan menggunakan fungsi *StandardScaler*. Kemudian, hasilnya disimpan sebagai sebuah data *frame* dan dilakukan *print "std\_atr"*. Tahap ini dilakukan untuk penskalaan standar terhadap data numerik ke dalam data *frame* yang dimana *output* yang dihasilkan berupa kolom-kolom berdasarkan '*numeric\_df.columns*'. Lalu, data *frame* tersebut diprint yang berisi data atribut numerik per-kolom-nya.

```
[ ] from sklearn.preprocessing import StandardScaler
scale=StandardScaler()
std_atr=scale.fit_transform(df)
std_atr=pd.DataFrame(std_atr,columns=df.columns)
print(std_atr)
```

	Calories	Sugars (g)	Protein (g)	Caffeine (mg)
0	-1.859434	-1.674141	-1.373732	1.406828
1	-1.849692	-1.674141	-1.332593	2.679127
2	-1.839950	-1.674141	-1.229746	3.726902
3	-1.839950	-1.674141	-1.229746	4.924360
4	-1.206734	-1.217042	-0.201274	-0.089995
...	...	...	...	...
237	1.228713	1.576341	-0.406969	-1.212612
238	-0.232555	0.255833	-0.612663	-1.212612
239	0.059699	0.255833	-0.818357	-1.212612
240	-0.135137	0.103466	-0.818357	-1.212612
241	0.449370	1.119242	-0.406969	-1.212612

[242 rows x 4 columns]

Gambar 8. Tahapan Normalisasi

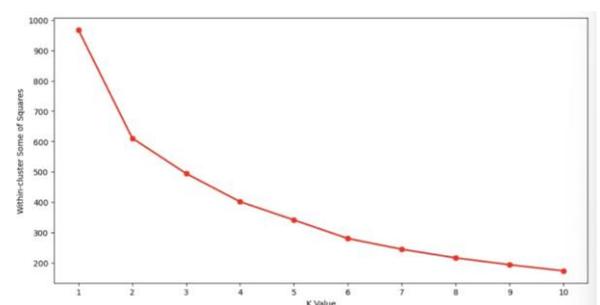


Gambar 9. Boxplot Normalized

### 3.4 Model Tuning

Berdasarkan gambar 10 merupakan tahap model *tuning* dimana untuk melakukan analisis nilai *k* optimal dengan metode *Elbrow*, dapat dibuat program sederhana plot antara nilai *within-cluster some of squares* (*wcss*) atau inerti model untuk setiap nilai *k* yang disetting pada fungsi *k-means*.

Tahap berikut ini digunakan untuk mengevaluasi *cluster* dalam *clustering* dengan melakukan import *K-Means* dengan '*wcss*' yang merupakan singkatan dari *Within-Cluster Sum of Squares*. Tahap ini digunakan '*range (1,11)*' untuk menghasilkan jangkauan grafik dari 1 sampai 10 pada *ouput K-Value* nantinya. Lalu, pada '*plt*' adalah tahap melakukan analisis grafik *cluster* untuk menentukan jumlah *cluster* mana yang akan dipilih. Kemudian, pada gambar 10 merupakan *output* grafik *cluster*.

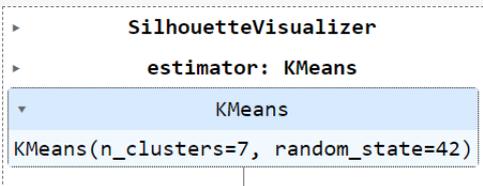


Gambar 10. Output: Grafik Kluster

Berdasarkan *output* yang dihasilkan, jumlah *cluster* yang dipilih adalah tujuh (*k=7*). Dikarenakan pada titik *K-Value* tersebut, adalah titik nilai inerti mengalami

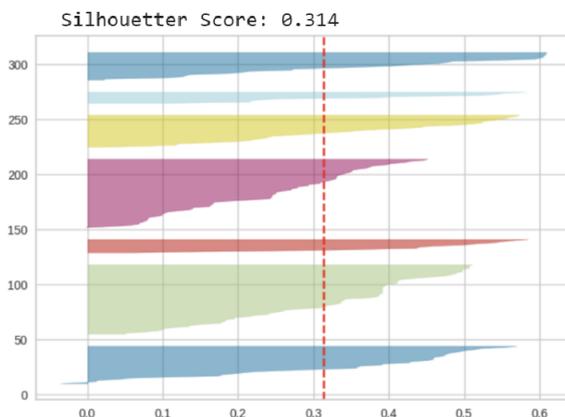
penurunan yang mengecil secara konstan. Setelah itu, untuk meningkatkan keyakinan pada nilai K-optimal yang terpilih, dilakukan analisis berupa *silhouette coefficient* seperti pada gambar 11 dan 12 yang didasari oleh konfigurasi pada objek *cluster* nya yang dimana dilakukan dengan menggunakan fungsi *SilhouetteVisualizer*.

```
[ ] from yellowbrick.cluster import SilhouetteVisualizer
model=KMeans(7,random_state=42)
visualizer=SilhouetteVisualizer(model,colors='yellowbrick')
visualizer.fit(std_atr)
```



Gambar 11. *SilhouetteVisualizer*

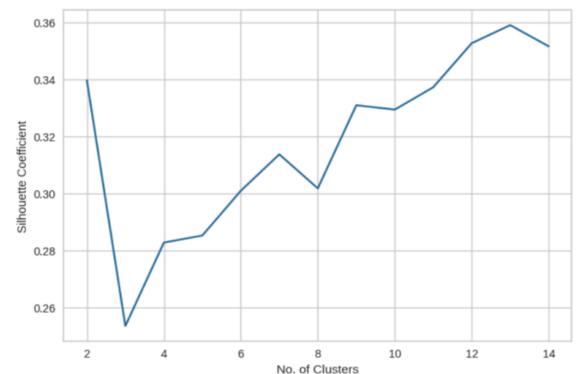
```
[ ] from sklearn.metrics import silhouette_score
km = KMeans(n_clusters=7, random_state=42)
km.fit(std_atr)
score = silhouette_score(std_atr,km.labels_)
print('Silhouetter Score: %.3f' % score)
```



Gambar 12. *Output: Skor Silhouette*

Berdasarkan *output* yang didapatkan, diketahui bahwa skor yang dihasilkan adalah bernilai positif, yaitu k=7 dengan skor *silhouette* adalah 0,314 yang dimana artinya atribut yang dikelompokkan sudah terkonfigurasi dengan sesuai. Gambar 18 merupakan *import Silhouette Score* dan gambar 13 merupakan *output silhouette Coefficient* dan grafik *Silhouette Coefficient*.

```
For n_clusters=2, Silhouette Coefficient = 0.3397083175910184
For n_clusters=3, Silhouette Coefficient = 0.2535085791459062
For n_clusters=4, Silhouette Coefficient = 0.2827697985552126
For n_clusters=5, Silhouette Coefficient = 0.28525605579993657
For n_clusters=6, Silhouette Coefficient = 0.3009618795523781
For n_clusters=7, Silhouette Coefficient = 0.3137459328204555
For n_clusters=8, Silhouette Coefficient = 0.3017979274170174
For n_clusters=9, Silhouette Coefficient = 0.33098118992126646
For n_clusters=10, Silhouette Coefficient = 0.3294464996538695
For n_clusters=11, Silhouette Coefficient = 0.33729768673021643
For n_clusters=12, Silhouette Coefficient = 0.352763233450069
For n_clusters=13, Silhouette Coefficient = 0.3590971525668146
For n_clusters=14, Silhouette Coefficient = 0.35164741913659947
```



Gambar 13. *Silhouette Coefficient*

Berikutnya dilakukan *import* terhadap skor *silhouette* yang telah didapatkan untuk menghasilkan data *n-kcluster* beserta dengan *silhouette coefficient* yang telah tersinkronisasi dengan baik dan didapatkan *output* sebagaimana pada gambar 13.

### 3.5 Model Building

Gambar 23 merupakan tahap model *building* dimana dilakukan pembangunan model *K-Means clustering* dengan menggunakan fungsi *K-Means* dan *input* atribut – atribut yang sudah dinormalisasi. Kemudian, pada gambar 24 didapatkan *output dataframe* model *building* dimana terdapat nilai *inertia* sebesar 245.764 dan terdapat *cluster* untuk masing-masing *centorids* yang terdiri dari *calories*, *sugars (g)*, *protein (g)*, dan *caffein (mg)*. Lalu, pada gambar 14 merupakan *output* hasil *cluster* dimana diperoleh 242 objek yang terbagi ke dalam 7 *cluster*.

```
inertia: 245.76497124656998
cluster_centroids: [[ 1.69144822  1.12876465  1.79138913  0.93907051]
[ 1.29253878  1.69192891 -0.31476094  0.37969765]
[-0.97232207 -0.91468973 -0.567346 -0.45835354]
[ 0.83092357  0.38492081  1.51284476 -0.74797309]
[-1.57249374 -1.36479076 -1.29519413  2.28450975]
[-0.01216701  0.2308546 -0.49059526 -0.59179848]
[-0.1140738 -0.45658275  0.66597723  0.87080327]]
```

Gambar 13. *Data Frame Model Building*

Dikarenakan telah terpilihnya  $k=7$ , maka pada tahap *model building* dilakukan Pembangunan model klasterisasi *7-means* dengan menggunakan fungsi *K-Means* dengan memasukkan atribut-atribut yang sebelumnya sudah dinormalisasikan dan dihasilkanlah *output* berupa hasil pengelompokkan *cluster*, yang dimana pada *output* diatas dari 242 objek yang telah terbagi menjadi tujuh *cluster*.

### 3.6 Model Visualization

Berdasarkan gambar 14 merupakan tahap model *visualization* dimana pada tahap ini

hasil dari *clustering* selanjutnya dapat dilakukan visualisasi ke dalam *plot* dua dimensi seperti pada gambar 15.

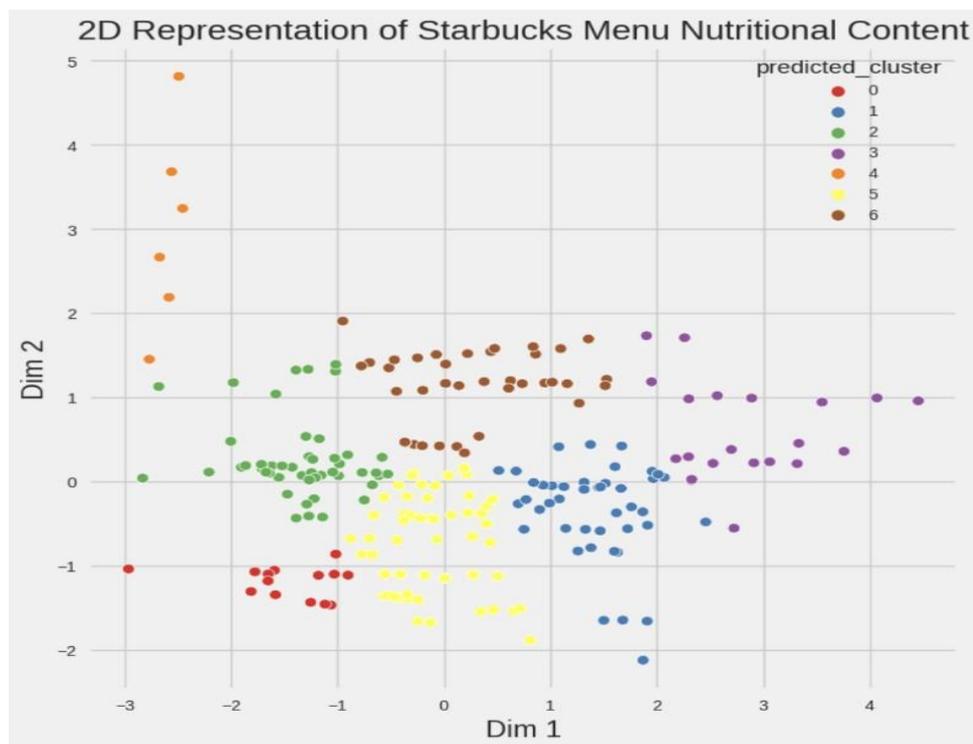
```
[ ] from sklearn.pipeline import Pipeline
from sklearn.preprocessing import LabelEncoder, StandardScaler
from sklearn.decomposition import PCA

[ ] preprocessor=Pipeline([("scaler",StandardScaler()),("pca",PCA(n_components=2,random_state=42)),])
clusterer=Pipeline([("kmeans",KMeans(n_clusters=7,init="k-means++",n_init=50,max_iter=500,random_state=42,)))])
pipe=Pipeline([("preprocessor",preprocessor),("clusterer",clusterer)])
pipe.fit(df)

- Pipeline
- preprocessor: Pipeline
  - StandardScaler
  - PCA
- clusterer: Pipeline
  - KMeans

[ ] preprocessed_data=pipe["preprocessor"].transform(df)
predicted_labels=pipe["clusterer"]["kmeans"].labels_
```

Gambar 14. Tahapan Model *Visualization*



Gambar 15. *Visualization 2D Representation of Starbucks Menu Nutritional Content*

Tahap ini telah dihasilkan *output* keseluruhan berdasarkan *dataset* yang dianalisis dengan hasil dimana terdiri dari tujuh *cluster* atau pengelompokkan nutrisi dari menu Starbucks. *Cluster* pertama (0) terdiri dari 16 menu, *cluster* kedua (1) terdiri dari 29 menu, *cluster* ketiga (2) terdiri dari 64 menu, *cluster* keempat (3) terdiri dari 24 menu, *cluster* kelima (4) terdiri dari 24 menu, *duster* keenam (5) terdiri dari 61 menu, dan

*cluster* ketujuh (6) yang terdiri dari 37 menu dari keseluruhan 242 menu yang diambil berdasarkan pertimbangan atribut-atribut yang berada pada *ouput* data *frame* ditahap kedua.

Berdasarkan tabel 2 diketahui nilai *centroid* untuk setiap atribut pada setiap *cluster*. Nilai *centroid* pada setiap atribut dipengaruhi oleh setiap *cluster* yang terbuat.

Tabel 2. Nilai *Centroid*

Cluster	Centroids			
	Calories	Sugars (g)	Protein (g)	Caffeine (mg)
0	1.691	1.129	1.791	0.939
1	1.293	1.692	-0.315	0.380
2	-0.972	-0.915	-0.567	-0.458
3	0.831	0.385	1.513	-0.748
4	-1.572	-1.365	-1.295	2.285
5	-0.012	0.231	-0.491	-0.592

Berdasarkan pada tabel 2 diketahui bahwa pada *cluster* 0 didapat nilai *centroid* tertinggi pada atribut 'calories' dan 'protein' dengan nilai sebesar

1.691 dan 1.791. Lalu pada *cluster* 1 didapat nilai *centroid* tertinggi pada atribut *Sugars* dengan nilai sebesar 1.692. Sedangkan nilai *centroid* tertinggi pada atribut *caffeine* terdapat pada *cluster* 4 dengan nilai sebesar 2.285.

### 3.7 Evaluation Result

Tahapan selanjutnya adalah mengetahui apa saja menu yang terdapat pada masing-masing *cluster*. *Cluster* yang dibuat didasarkan dari 4 atribut yang sebelumnya sudah ditentukan yaitu kandungan kalori, gula, protein, dan kafein. Keempat atribut tersebut dipilih dikarenakan atribut-atribut tersebut sering dijadikan acuan oleh konsumen untuk memilih minuman yang akan dibeli. Penjelasan lebih lanjut bisa dilihat pada tabel 3.

Tabel 3. *Cluster Value*

Cluster	Cluster Value	Cluster Name
Cluster 0	Paling Tinggi (1.38)	Kandungan Tertinggi
Cluster 1	Cukup Tinggi (0.76)	Kandungan Cukup Tinggi
Cluster 2	Paling Rendah (-0.72)	PKandungan Terendah
Cluster 3	Tinggi (0.49)	Kandungan Tinggi
Cluster 4	Cukup Rendah (-0.48)	Kandungan Cukup Rendah
Cluster 5	Rendah (-0.21)	Kandungan Rendah
Cluster 6	Sedang (0.24)	Kandungan Sedang

Berdasarkan pada tabel 3 bahwa setiap *cluster* memiliki namanya sendiri dimana *cluster* 0 dengan nilai *cluster* paling tinggi sebesar 1.38 disebut *cluster* dengan kandungan tertinggi. Selanjutnya pada *cluster* 1 dengan nilai *cluster* cukup tinggi dengan nilai sebesar 0.76 disebut *cluster* dengan kandungan cukup tinggi. Lalu, pada *cluster* 2, diketahui bahwa nilai *cluster* paling rendah dengan nilai sebesar -0.72 dan disebut *cluster* dengan kandungan terendah. Selanjutnya, pada *cluster* 3 dengan nilai *cluster* tinggi yang bernilai 0.49

memiliki sebutan nama *cluster* dengan kandungan tinggi. Kemudian, pada *cluster* 4 diketahui nilai *cluster* cukup rendah dengan nilai sebesar -0.48 disebut dengan *cluster* dengan kandungan cukup rendah. Sedangkan untuk *cluster* 5 dan *cluster* 6 dengan nilai rendah dan sedang yang bernilai -0.21 dan 0.24 dengan *cluster* yang biasa disebut dengan *cluster* dengan kandungan rendah dan *cluster* dengan kandungan sedang. Hasil dari klasterisasi yang telah dibuat bisa dilihat pada tabel 4.

Tabel 4. Cluster Menu Starbucks Berdasarkan Kandungan

<i>Cluster Name</i>	<i>Menu Starbucks</i>
Kandungan Tertinggi	<i>Classic Espresso Drinks CaffA-Mocha (Without Whipped Cream), Venti Nonfat Milk, Classic Espresso Drinks CaffA-Mocha (Without Whipped Cream) 2 % Milks, Classic Espresso Drinks CaffA-Mocha (Without Whipped Cream) Soymilk</i> ... ... <i>TazoA Tea Drinks TazoA Green Tea Latte 2% Milk</i>
Kandungan Cukup Tinggi	<i>Signature Espresso Drinks Caramel Apple Spice (Without Whipped Cream) Grande, TazoA Tea Drinks TazoA Chai Tea Latte Soymilk, TazoA Tea Drinks TazoA Chai Tea Latte Venti Nonfat Milk</i> ... ... <i>FrappuccinoA Blended CrAme Strawberries &amp; CrAme (Without Whipped Cream) Whole Milk.</i>
Kandungan Terendah	<i>Coffee Brewed Coffee Venti, Classic Espresso Drinks CaffA-Latte Short Nonfat Milk, Classic Espresso Drinks CaffA-Latte 2% Milk,</i> ... ... <i>FrappuccinoA Light Blended Coffee Mocha Venti Nonfat Milk.</i>
Kandungan Tinggi	<i>Signature Espresso Drinks White Chocolate Mocha (Without Whipped Cream) Soymilk, Signature Espresso Drinks White Chocolate Mocha (Without Whipped Cream) Tall Nonfat Milk, Signature Espresso Drinks White Chocolate Mocha (Without Whipped Cream) 2% Milk</i> ... ... <i>Smoothies Strawberry Banana Smoothie 2% Milk</i>
Kandungan Cukup Rendah	<i>Coffee Brewed Coffee Short, Coffee Brewed Coffee Tall, Coffee Brewed Coffee Grande</i> ... ... <i>Shaken Iced Beverages Ice Brewed Coffee (With Milk &amp; Classic Syrup) 2% Milk</i>
Kandungan Rendah	<i>Signature Espresso Drinks Caramel Macchiato Soymilk, Signature Espresso Drinks White Chocolate Mocha (Without Whipped Cream) Short Nonfat Milk, Signature Espresso Drinks White Chocolate Mocha (Without Whipped Cream) Short Nonfat Milk</i> ... ... <i>Frappuccino Blended Vanilla Bean (Without Whipped Cream) Soymilk</i>
Kandungan Terendah	<i>Classic Espresso Drinks CaffA-Latte Tall Nonfat Milk, Classic Espresso Drinks CaffA-Latte Soymilk, Classic Espresso Drinks CaffA-Latte Grande Nonfat Milk</i> ... ... <i>Frappuccino Light Blended Coffee Mocha Grande Nonfat Milk</i>

Identifikasi kandungan nutrisi pada menu Starbucks, konsumen dari tabel 4 dapat disimpulkan bahwa terdapat beberapa rekomendasi yang dapat diberikan untuk meningkatkan penjualan dan meningkatkan kesadaran konsumen dalam membeli menu-menu yang terdapat di Starbucks. Berikut merupakan rekomendasi yang dapat diberikan untuk setiap *clusternya*.

*Cluster* dengan Kandungan Tertinggi (*Cluster 0*): *Classic Espresso Drinks CaffA-Mocha (Without Whipped Cream) Venti Nonfat Milk, Classic Espresso Drinks CaffA-Mocha (Without Whipped Cream) 2 % Milks, Classic Espresso Drinks CaffA-Mocha (Without Whipped Cream) Soymilk, TazoA Tea Drinks TazoA Green Tea Latte 2% Milk*. Tingkatkan promosi produk dengan kandungan nutrisi tinggi, seperti minuman rendah kalori, tinggi protein, dan rendah gula. Perkenalkan menu yang menonjolkan aspek-aspek kesehatan ini. Dengan adanya menu – menu ini dapat membuat *customer* memilih dengan lebih sadar terkait aspek kesehatan dan nutrisi dari produk yang mereka konsumsi.

*Cluster* dengan Kandungan Cukup Tinggi (*Cluster 1*): *Signature Espresso Drinks Caramel Apple Spice (Without Whipped Cream) Grande, TazoA Tea Drinks TazoA Chai Tea Latte Soymilk, TazoA Tea Drinks TazoA Chai Tea Latte Venti Nonfat Milk, FrappuccinoA Blended CrAme Strawberries & CrAme (Without Whipped Cream) Whole Milk*. Lanjutkan mempromosikan produk dengan kandungan yang cukup seimbang, mungkin dengan penekanan pada pengurangan gula atau peningkatan protein.

*Cluster* dengan Kandungan Terendah (*Cluster 2*): *Coffee Brewed Coffee Venti, Classic Espresso Drinks CaffA-Latte Short Nonfat Milk, Classic Espresso Drinks CaffA-Latte 2% Milk, FrappuccinoA Light Blended Coffee Mocha Venti Nonfat Milk*. Lihat produk dalam kategori ini untuk mengetahui apakah mereka memiliki nilai nutrisi yang rendah atau keseimbangan nutrisi yang salah dengan mempertimbangkan inovasi untuk memenuhi kebutuhan konsumen yang lebih sehat.

*Cluster* dengan Kandungan Tinggi (*Cluster 3*): *Signature Espresso Drinks White Chocolate Mocha (Without Whipped Cream) Soymilk, Signature Espresso Drinks White Chocolate Mocha (Without Whipped Cream) Tall Nonfat Milk, Signature Espresso Drinks White Chocolate Mocha (Without Whipped Cream) 2% Milk, Smoothies Strawberry Banana Smoothie 2% Milk*. Lanjutkan mengembangkan produk dengan kandungan tinggi, tetapi pastikan untuk menjaga keseimbangan nutrisi yang sehat seperti protein yang terkendali dengan kalori.

*Cluster* dengan Kandungan Cukup Rendah (*Cluster 4*): *Coffee Brewed Coffee Short, Coffee Brewed Coffee Tall, Coffee Brewed Coffee Grande, Shaken Iced Beverages Ice Brewed Coffee (With Milk & Classic Syrup) 2% Milk*. Untuk meningkatkan nilai nutrisi mereka, evaluasi ulang atau perbaikan produk dari kategori ini. Mengurangi kandungan gula atau meningkatkan protein mungkin diperlukan.

*Cluster* dengan Kandungan Rendah (*Cluster 5*): *Signature Espresso Drinks Caramel Macchiato Soymilk, Signature Espresso Drinks White Chocolate Mocha (Without Whipped Cream) Short Nonfat Milk, Signature Espresso Drinks White Chocolate Mocha (Without Whipped Cream) Short Nonfat Milk, Frappuccino Blended Vanilla Bean (Without Whipped Cream) Soymilk*. Memperhatikan lagi komposisi produk dalam *duster* ini. Buat inovasi atau perubahan yang signifikan untuk menarik minat konsumen dengan menambah nilai nutrisi.

*Cluster* dengan Kandungan Sedang (*Cluster 6*): *Classic Espresso Drinks CaffA-Latte Tall Nonfat Milk, Classic Espresso Drinks CaffA-Latte Soymilk, Classic Espresso Drinks CaffA-Latte Grande Nonfat Milk, Frappuccino Light Blended Coffee Mocha Grande Nonfat Milk*. Fokus pada peningkatan nutrisi atau keseimbangan yang lebih baik dalam produk. Promosikan produk dengan nilai nutrisi seimbang untuk menarik konsumen yang peduli dengan kesehatan.

#### 4. KESIMPULAN

Tujuan untuk menganalisis nutrisi yang terdapat pada beberapa minuman Starbucks digunakan metode klusterisasi dalam penelitian *K-Means* yang didukung oleh atribut-atribut terkait dengan menggunakan input data sekunder. Metode klusterisasi diawali dengan melakukan *data mining* yang merupakan salah satu bagian dari *Knowledge Discovery in Database* (KDD) untuk mengelolah kumpulan data menjadi data yang berguna dan bermanfaat yang berasal dari *database*. Berdasarkan *data mining*, klusterisasi adalah metode yang paling efektif dengan *clustering* atau klusterisasi yang merupakan analisis pengelompokan sebagai alat bantu. *Clustering* tersebut dilakukan untuk pengelompokan ke dalam penggunaan cluster-cluster dengan 6 tahap untuk menghasilkan cluster. Diawali dengan sebelum menginput *database* perlu dilakukan aktivasi *library* yang dimana akan menghasilkan *output dataframe* dari *database* yang telah diinput. Berikutnya adalah melakukan eksplorasi data berupa *output* akhir *boxplot*, dengan tahap berikutnya data *preprocessing* dengan *output* yang dihasilkan berupa data numerik per-kolom-nya. Lalu, *model tuning* untuk menghasilkan *output* penentuan kluster mana yang dipilih sehingga pada tahap berikutnya *model building* dengan menghasilkan 242 objek yang telah terbagi menjadi tujuh kluster. Tahap akhir, adalah *model visualization* yang dimana pada tahap ini telah dihasilkan *output* keseluruhan berdasarkan *dataset* yang dianalisis dengan hasil dimana terdiri dari tujuh *cluster* atau pengelompokan nutrisi dari menu Starbucks. Berdasarkan hasil evaluasi yang didapatkan terdapat 7 *cluster* dimana *cluster* 0 dengan nilai *cluster* paling tinggi sebesar 1.38 disebut *cluster* dengan kandungan tertinggi. Selanjutnya pada *cluster* 1 dengan nilai *cluster* cukup tinggi dengan nilai sebesar 0.76 disebut *cluster* dengan kandungan cukup tinggi. Lalu pada *cluster* 2, diketahui bahwa nilai *cluster* paling rendah dengan nilai

sebesar -0.72 dan disebut *cluster* dengan kandungan terendah. *cluster* 3 dengan nilai *cluster* tinggi yang bernilai 0.49 memiliki sebutan nama *cluster* dengan kandungan tinggi. Lalu, *cluster* 4 diketahui nilai *cluster* cukup rendah dengan nilai sebesar -0.48 disebut dengan *cluster* dengan kandungan cukup rendah. Sedangkan, untuk *cluster* 5 dan *cluster* 6 dengan nilai rendah dan sedang yang bernilai -0.21 dan 0.24 dengan *cluster* yang biasa disebut dengan *cluster* dengan kandungan rendah dan *cluster* dengan kandungan sedang dari keseluruhan 242 menu yang diambil berdasarkan pertimbangan atribut-atribut yang berada pada *output data frame* ditahap kedua. Hasil dari klusterisasi produk-menu Starbucks berdasarkan karakteristik nutrisi seperti kalori, gula, protein, dan kafein mengungkapkan karakteristik unik dari setiap *cluster*. *Cluster* dengan kandungan tertinggi menawarkan produk dengan nilai nutrisi yang tinggi, seperti kalori yang rendah, gula yang rendah, protein yang tinggi, atau kafein yang cukup. Rekomendasi untuk *cluster* ini berpusat pada pengiklanan produk kesehatan yang menonjol. *Cluster* dengan kandungan tinggi, di sisi lain, menunjukkan keseimbangan nutrisi yang baik tetapi masih dapat ditingkatkan, mungkin dengan peningkatan gula atau protein. *Cluster* dengan kandungan terendah, di sisi lain, menunjukkan nilai nutrisi yang rendah atau keseimbangan nutrisi yang kurang sesuai, jadi evaluasi ulang dan inovasi diperlukan untuk memenuhi kebutuhan konsumen yang lebih sehat. *Cluster* lainnya memiliki ciri-ciri yang sebanding dengan rekomendasi yang sesuai untuk meningkatkan nilai nutrisi produk, menyesuaikan komposisi produk, atau menampilkan produk sesuai dengan *preferensi* konsumen yang berbeda. Secara keseluruhan, klusterisasi ini memberikan arahan untuk mengkustomisasi pemasaran, inovasi produk, dan penyesuaian nilai nutrisi untuk meningkatkan minat dan penjualan produk menu Starbucks.

## DAFTAR PUSTAKA

- Abbas, M., El-Zoghbi, A., & Shoukry, A. (2021). *DenMune: Density peak based clustering using mutual nearest neighbors*. *Pattern Recognition*, 109, 107589.
- Abbas, A. H., Abdulsattar, N. F., Mansour, H. S., Mutar, M. H., Habelalmateen, M. I., Audah, L., ... & Mohammed, A. (2022, May). *A New Hybrid Approach Cluster-Heads Election to reduce the number of clusters in VANETs*. In *2022 5th International Conference on Engineering Technology and its Applications (IICETA)* (pp. 436-440). IEEE.
- Abu Khurma, R., & Aljarah, I. (2021). *A review of multiobjective evolutionary algorithms for data clustering problems*. *Evolutionary data clustering: algorithms and applications*, 177-199. DOI: [https://doi.org/10.1007/978-981-33-4191-3\\_8](https://doi.org/10.1007/978-981-33-4191-3_8)
- Abu-Jamous, B., Fa, R., & Nandi, A. K. (2015). *Integrative cluster analysis in bioinformatics*.
- Achmad, F., Abdillah, I. T., & Amani, H. (2023). *Decision-making process for tourism potential segmentation*. *International Journal of Innovation in Enterprise System*, 7(01), 19-30. DOI: <https://doi.org/10.25124/ijies.v7i01.204>
- Achmad, F., Prambudia, Y., & Rumanti, A. A. (2023). *Sustainable Tourism Industry Development: A Collaborative Model of Open Innovation, Stakeholders, and Support System Facilities*. *IEEE Access*, 11, 83343 – 83363. DOI: <https://doi.org/10.1109/ACCESS.2023.33301574>
- Achmad, F., Prambudia, Y., & Rumanti, A. A. (2023). *Improving tourism industry performance through support system facilities and stakeholders: The role of environmental dynamism*. *Sustainability*, 15(5), 4103. DOI: <https://doi.org/10.3390/su15054103>
- Achmad, F., Prambudia, Y., & Rumanti, A. A. (2023, December). *Examining the Feedback Effects of Support System Facilities on Tourism Industry Performance: A Causal Loop Diagram Modeling Approach*. In *2023 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)* (pp. 0768-0772). IEEE.
- Ahdiat, A. (2022). *Grafik Rata-rata Jumlah Penduduk Kurang Gizi di Asia Tenggara (2019- 2021)*. Databoks.
- Alkhairi, P., & Windarto, A. P. (2019, February). *Penerapan K-Means Cluster Pada Daerah Potensi Pertanian Karet Produktif di Sumatera Utara*. In *Seminar Nasional Teknologi Komputer & Sains (SAINTEKS)* (Vol. 1, No. 1).
- Rizaty, Monavia A. (2022). *Pendapatan Starbucks (2017-2021)*. Databoks.
- Handayani, I. A. (2023). *Pengaruh Mengunyah Permen Karet Gula Dan Non Gula Terhadap Ph Saliva pada Mahasiswa Tingkat 1 Jurusan Kesehatan Gigi Poltekkes Tanjungkarang Tahun 2023* (Doctoral dissertation, Poltekkes Kemenkes Tanjungkarang).
- Herviany, M., Delima, S. P., Nurhidayah, T., & Kasini, K. (2021). *Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Daerah Rawan Tanah Longsor Pada Provinsi Jawa Barat: Comparison of K-Means and K-Medoids Algorithms for Grouping Landslide Prone Areas in West Java Province*. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 1(1), 34-40. DOI: <https://doi.org/10.57152/malcom.v1i1.60>
- Irwansyah, E., & Faisal, M. (2015). *Advanced Clustering: Teori dan Aplikasi*. Deepublish.
- Luo, Y., Li, M., Zhu, K. X., Guo, X. N., Peng, W., & Zhou, H. M. (2016). *Heat-induced interaction between egg white protein and wheat gluten*. *Food Chemistry*, 197, 699-708. DOI: <https://doi.org/10.1016/j.foodchem.2015.09.088>

Nurhasanah, S., & Dewi, C. (2019). *Defining the Role of Multinational Corporations: Starbucks and Coffee Culture in Indonesia*. *Jurnal Sentris*, 2(2), 63-77. DOI: <https://doi.org/10.26593/sentris.v2i2.4181.63-77>

Research, DataIndustri. (2023). *Grafik Pertumbuhan Industri Makanan dan Minuman, 2011 – 2023*. DataIndustri Research.

Rizaty, Monavia A. (2021). *Grafik 10 Negara dengan Jumlah Gerai Starbucks Resmi Terbanyak (Kuartal III-2021)*. Databoks.

Royani, I., Maharani, S., Sandi, S. P. H., & Hidayaty, D. E. (2023). *Penerapan Strategi Pelayanan Di Starbucks*. *Profit: Jurnal Manajemen, Bisnis dan Akuntansi*, 2(3), 124- 130. DOI: <https://doi.org/10.58192/profit.v2i3.989>

Santika, Erlina F. (2023). *Grafik 10 Negara/Wilayah dengan Gerai Starbucks Terbanyak (2022)*. Databoks.

Virgo, I., Defit, S., & Yuhandri, Y. (2020). *Klasterisasi Tingkat Kehadiran Dosen Menggunakan Algoritma K-Means Clustering*. *Jurnal Sistim Informasi dan Teknologi*, 23-28. DOI: <https://doi.org/10.37034/jsisfotek.v2i1.17>

Vandenbrink, D., Pauzé, E., & Potvin Kent, M. (2020). *Strategies used by the Canadian food and beverage industry to influence food and nutrition policies*. *international journal of behavioral nutrition and physical activity*, 17, 1-13. DOI: <https://doi.org/10.1186/s12966-019-0900-8>