Table 1 Reviewed Studies on Customer Segmentation

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Article | Business Context | Dataset | Segmentation  Features | Techniques |
| Vucetic et al,2018 | Marketing | Restaurant customer feedback data period January 1, 2016, to December 31, 2016 | Customer name, type of food ordered, food review, gender | K-Means using PL based algorithm |
| Aziz et al, 2019 | Small Medium Enterprise | SME customers are all pulse server operators AR-Pulsabiz Malang Indonesia period January 1, 2018, to June 30,2018 | ID Customer, Length, Recency, Frequency, Monetary | K-Means Clustering and LRFM Model |
| Ye Jingyi, 2021 | E- Commerce | Online Retail Data Set period 12 January 2010 and 12 September 2011 from UK | Invoice No, Stock Code, Description, Quantity, Invoice Date, Unit Price, Customer ID, Country, Total Price | K-Means Clustering |
| Sano et al, 2021 | Marketing | Transaction Supermarket data January 1, 2017, to December 31, 2018 | ID Customer, Product Name, Length, Recency, Frequency, Monetary | AEF, RFM, k-means, Ward method, FCM, and the decision tree |
| Antony et al,2019 | RFM Analysis | Sales data of a pharmacy in Palembang period January 2015 until December 2015 | ID Customer, Product Name, Length, Recency, Frequency, Monetary | K-Means Method and eight indexes of validity to determine the optimal number of clusters namely Elbow Method, Silhouette Index, Calinski-Harabasz Index, Davies-Bouldin Index, Rutkowski Index, Hubert Index, Ball-Hall Index, and Krakowski-Lai Index |
| Puh et al, 2020 | Food Retailing | Questionnaire data consisting of 500 consumers in Croatia in 2020 | Demographic characteristics (Age, Gender, Education, Occupation, Monthly Income in HRK), Product, Frequency, Percentage | Latent Class Model |
| Abdi et al, 2018 | Telecom Company | Customers of a telecom company period January 1,2017 to December 31, 2017 | Socio-demographic attributes (Region, Age, Marital, Address, Income, Education, Employment, Retire, Gender), Behavioral Attributes (Hours of Usage (Longmont, Tollmon, Equipmon, Cardmon, Wiremon), Selected Service (Multiline, Voice, Pager, Internet, Call Id, Call wait, Forward, Confer, Call card, Wireless, Churn)) | K-Means Clustering, Neural network, |
| McLoughlin et al, 2014 | Electricity | Experimental data by installing smart meters to more than 4000 residences in Ireland, period January 1, 2009, to December 31, 2010 | Dwelling type, No. of bedrooms, Age, Social Class, Electronic Type | k-means, k-medoid and Self Organizing Maps (SOM) |
| Li et al, 2012 | Transportation | Historical data from the vehicle sharing platform database at the university detailed data of all customers from November 30, 2015, to November 30, 2017 | User ID, driving mileage, points, discounts and 29 other attributes. The Variable are used User id, current miles, cost, car id | K-Means Method and KLRFMD model |
| Marisa et al, 2019 | Small Medium Enterprise | The sale of electric pulses period January 1, 2016, to December 31, 2017 | Name Customer, Average of transaction/ week, Payment System | K-Means Clustering |
| Chindyana et al ,2021 | Tourism | Google review rating in 2020 | Id Customer, Gender, Place, Review | K-Means Method and DBSCAN Method |
| Zhao et al,2021 | E-commerce | The customer transaction online customer company in the UK period January 1, 2016, to December 31,2017 | Invoice number, quantity, price, address, and zip code. | K-Means Clustering and RFM Model |