

Pemanfaatan Metode *Clustering* untuk melihat pola penjualan dan perilaku pembelian konsumen, pada penjualan tiket pesawat PT. Garuda Indonesia, Cabang Batam

Rivort Pormes^{#1}, Daniel H. F. Manongga^{*2}

[#] Magister Sistem Informasi, Universitas Kristen Satya Wacana

Jl. Diponegoro 52-60 Salatiga, 50711-Indonesia

¹rivopormes@gmail.com

^{*} Magister Sistem Informasi, Universitas Kristen Satya Wacana

Jl. Diponegoro 52-60 Salatiga, 50711-Indonesia

²dmanongga@yahoo.co.uk

Abstract — Analysis of *Data mining* on data sales aims to process data with limited attributes to obtain information that can be generated from processing of such data. The utilization of Clustering Method is expected to answer the needs of processing data in two years to figure out the pattern of most purchases and distance of purchases up to the day of departure on PT Garuda Indonesia, Batam Branch. *Data mining* processing aims to process sales data into informations or knowledges which plays an important role as one of the solutions in determining the future strategy. The use of the K-means algorithm aims to explore algorithm processing capabilities with limited attributes. The purpose of using this method is to obtain the largest purchasing frequency information each month as well as consumer purchasing patterns as reference in establishing business recommendations as well as acquiring knowledge which will be useful for PT Garuda Indonesia, Batam Branch. By considering the existence of business competition of other airline companies, the result is expected to become a feedback for the company in the future.

Keywords— *Clustering, K-means, Data mining.*

I. PENDAHULUAN

Bisnis penerbangan di Indonesia belakangan ini mengalami pertumbuhan yang cukup signifikan, terlihat dengan banyaknya orang yang lebih cenderung memilih transportasi udara untuk berpergian dan menghubungkan mereka ke kota-kota yang ada di Indonesia, karna Indonesia merupakan negara maritim, yang sudah sepantasnya untuk menghubungkan satu kota atau daerah dengan daerah lain pilihan yang paling tepat adalah transportasi udara. Dengan transportasi udara waktu yang digunakan untuk menempuh satu daerah atau kota dapat ditempuh dengan waktu yang singkat. Hal ini menjadi peluang bisnis yang cukup besar bagi perusahaan-perusahaan penerbangan di Indonesia, Salah satunya PT. Garuda Indonesia.

Melihat kondisi ini, masyarakat sebagai *customer* (pengguna jasa) dihadapkan beberapa pilihan. Masyarakat dihadapkan pada pilihan apakah menggunakan jasa penerbangan Garuda Indonesia yang menyediakan jasa *full service*, atau menggunakan jasa penerbangan lain yang menyediakan harga terjangkau namun terkesan minim akan pelayanan yang sempurna.

Kondisi seperti ini terbukti dari proses pengambilan keputusan konsumen yang dijelaskan oleh Schiffman dan Kanuk mendefinisikan suatu keputusan sebagai pemilihan suatu tindakan dari dua atau lebih pilihan alternatif.[1]

Hawkins dan Mothersbaugh menjelaskan dalam teori umum mengenai perilaku konsumen bahwa dalam melakukan keputusan pembelian, perilaku seseorang dipengaruhi oleh faktor eksternal seperti budaya, sub-budaya, demografi, status sosial, kelompok, keluarga dan aktivitas pemasaran, faktor internal, yaitu persepsi, pembelajaran, ingatan, motivasi, *personality*, emosi dan sikap serta *self concept* dan *lifestyle*. Ketiga faktor ini akan menjadi kebutuhan dan keinginan yang akan mempengaruhi proses keputusan pembelian, dimana pada gilirannya keputusan akan memberikan pengalaman dan perolehan [2].

Engel et al. menambahkan bahwa proses keputusan pembelian ditentukan oleh tiga hal pokok, yaitu informasi, proses informasi, dan faktor-faktor yang menentukan proses keputusan. Faktor yang menentukan keputusan pembelian pada konsumen terdiri dari pengaruh lingkungan, perbedaan individu, dan proses psikologis konsumen. Hal lain yang perlu diperhatikan adalah sejauh mana kekuatan pasar (*market share*) dari Garuda dibandingkan dengan pesaing usaha lain. [3]

Ketersediaan data yang melimpah, kebutuhan akan informasi sebagai pendukung pengambilan keputusan untuk membuat solusi bisnis, dan dukungan infrastruktur di bidang teknologi informasi merupakan cikal-bakal dari lahirnya teknologi *Data mining*. *Data mining* adalah proses yang

menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai *database* besar [4].

Data mining dimaksudkan untuk memberikan solusi bagi para pengambil keputusan di dunia bisnis. Dalam menghadapi persaingan bisnis kedepan dan peningkatan pendapatan perusahaan. Pihak terkait dalam perusahaan dituntut untuk dapat mengambil keputusan yang tepat dalam menentukan strategi bisnis yang akan dijalkannya.

Dalam dunia bisnis yang semakin dinamis dan penuh persaingan, para pelakunya harus senantiasa memikirkan cara-cara untuk terus *survive* dan jika mungkin mengembangkan skala bisnis mereka. Salah satu cara adalah dengan melakukan analisis data perusahaan.

Data mining adalah suatu metode pengolahan data untuk menemukan pola yang tersembunyi dari data tersebut. Hasil dari pengolahan data dengan metode *Data mining* ini dapat digunakan untuk mengambil keputusan di masa depan. *Data mining* ini juga dikenal dengan istilah *pattern recognition*. *Data mining* merupakan metode pengolahan data berskala besar oleh karena itu *Data mining* ini memiliki peranan penting dalam bidang industri, keuangan, cuaca, ilmu dan teknologi. Secara umum kajian *Data mining* membahas metode-metode seperti, *clustering*, *klasifikasi*, *regresi*, *seleksi variable*, dan *market basket analisis*. [5]

Salah satu metode yang digunakan untuk proses pengolahan *Data mining* adalah metode pengelompokan (*Clustering*). Pada dasarnya *clustering* merupakan suatu metode untuk mencari dan mengelompokkan data yang memiliki kemiripan karakteristik (*similarity*) antara satu data dengan data yang lain. *Clustering* merupakan salah satu metode *Data mining* yang bersifat tanpa arahan (*unsupervised*), maksudnya metode ini diterapkan tanpa adanya latihan (*training*) dan tanpa ada guru (*teacher*) serta tidak memerlukan target output.

Pengolahan *Data mining* bertujuan mengolah data penjualan menjadi informasi atau *knowledge* yang berperan penting sebagai salah solusi dalam menentukan strategi bisnis PT. Garuda Indonesia, cabang Batam. Penggunaan algoritma *K-mean* bertujuan untuk memudahkan pengolahan data dengan atribut-atribut yang terbatas.

Pada penelitian ini akan dilakukan analisis untuk melihat pola pembelian terbanyak dari tiap-tiap bulan, jarak hari antara pembelian sampai dengan waktu penerbangan, serta kecenderungan pemilihan hari penerbangan, berdasarkan data-data penjualan tiket penerbangan dari kurung tahun 2014-2015 pada PT. Garuda Indonesia, Cabang Batam. Penggunaan algoritma *K-mean* untuk melihat apakah data-data dengan *atribut* terbatas dapat menghasilkan informasi yang dapat dipakai sebagai acuan untuk kepentingan bisnis PT. Garuda Indonesia, Cabang Batam.

Dengan analisis data penjualan ini diharapkan dapat memberikan pengetahuan (*knowledge*) dan *feedback*, sehingga PT. Garuda Indonesia, dapat mengantisipasi persaingan bisnis serta tetap mempertahankan segmen pasar

yang dimiliki saat ini, dengan melihat semakin banyaknya persaingan dari maskapai lain.

II. PENELITIAN TERKAIT

Penelitian sebelumnya oleh Primawati, Verdian, dan Nurcahyo dalam penelitiannya *Classification of Sales Agent for Cement Distribution using K-means Clustering* menjelaskan Agen adalah salah satu aset yang sangat penting bagi para distributor. Pengetahuan yang lebih baik dari agen dan perilaku mereka dibutuhkan, khususnya untuk mendukung keputusan yang berkaitan dengan strategi bisnis perusahaan dan untuk mengelola hubungan baik dengan para distributor. Pengetahuan semacam ini bisa diperoleh dengan mengklasifikasikan agen berdasarkan perilaku mereka melalui Data riwayat, seperti data transaksi jual beli. Salah satu pendekatan yang dapat dilakukan adalah pendekatan segmentasi yang bisa dilakukan dengan cara membagi agen menjadi beberapa segmen. Dalam tulisan ini, teknik *Data mining K-means*, metode clustering dieksplorasi mengklasifikasikan agen penjualan. Dengan mengimplementasikan *K-means*, pengetahuan tentang agen terbaik dapat diperoleh bersama agen yang ada setidaknya kontribusinya kepada distributor. [6]

Penelitian selanjutnya oleh Kashwan dan Velu dalam penelitiannya *Customer Segmentation Using Clustering and Data mining Techniques* menjelaskan teknik clustering merupakan langkah yang sangat penting dalam proses *Data mining*. Teknik ini adalah prosedur multivariat yang sangat sesuai untuk pengaplikasian segmentasi pada peramalan dan perencanaan penelitian di pasar. Laporan penelitian ini merupakan laporan komprehensif dari teknik *K-means* clustering dan Perangkat SPSS untuk membangun sistem real time dan online bagi super market tertentu untuk memprediksi penjualan di berbagai siklus tahunan. [7]

Penelitian selanjutnya oleh Schaidnagel, Abele, Laux, dan Petrov dalam penelitiannya *Sales Prediction with Parametrized Analysis* menjelaskan, ketika prediksi angka penjualan, tidak hanya riwayat penjualan tetapi juga harga di masa mendatang, produk yang akan berpengaruh pada kuantitas penjualan. Solusi yang diajukan menggunakan model univariat namun mengambil harga sebuah produk sebagai parameter yang mempengaruhi sistem prediksi tersebut. Pengaruh harga yang dihitung berdasarkan pada data historis penjualan menggunakan analisis korelasi dan rentang harga disesuaikan untuk pengidentifikasian produk dengan riwayat yang sebanding. [8]

Penelitian lain dari Subiakto dan Sukarno dengan judul *Financial Performance Analysis of PT. Garuda Indonesia, Tbk. in Comparison with Other Regional Airlines*, penelitian ini menganalisis kinerja keuangan industri penerbangan Indonesia, yang dikhususkan pada PT. Garuda Indonesia, Tbk. Penelitian ini membandingkan kinerja maskapai lainnya di Asia Pasifik untuk mendapatkan perbandingan yang lebih jelas dan gambaran dari industri penerbangan Asia Pasifik. Rasio keuangan yang digunakan dalam penelitian ini adalah rasio likuiditas, rasio

profitabilitas, dan rasio solvabilitas untuk mengetahui kinerja keuangan setiap tahun. Data yang dikumpulkan dari prospektus perusahaan, laporan tahunan, dan sumber-sumber lain untuk membantu peneliti untuk menganalisis kinerja keuangan semua maskapai. Hasil analisis ini akan diberikan kepada PT. Garuda Indonesia, Tbk. untuk meningkatkan kinerja keuangan perusahaan dan dapat digunakan sebagai referensi untuk maskapai lain untuk mempelajari kinerja keuangan maskapai lainnya.[9]

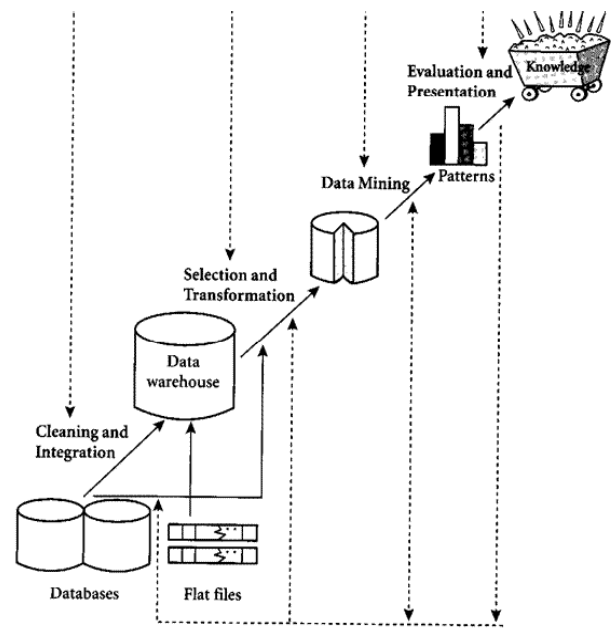
Penelitian dari Siregar dengan judul Implementasi *Data mining* Pada Penjualan Tiket Pesawat Menggunakan Algoritma Apriori (Studi Kasus: Jumbo Travel Medan), dalam penelitian tersebut dikembangkan aplikasi analisis *association* untuk mengekstraksi dan menginterpretasi pola kecenderungan penjualan tiket pesawat yang sering dijual secara bersamaan dari data transaksi menggunakan algoritma *apriori*. Algoritma *apriori* ini akan membentuk *frequent itemset* sebanyak yang telah ditentukan sebelumnya berdasarkan dua parameter, support dan confidence, untuk menemukan aturan asosiasi antara suatu kombinasi item. Proses yang dilakukan diawali dengan persiapan data melalui *preprocessing* data kemudian ditransformasi kedalam bentuk yang dapat diolah pada proses selanjutnya yaitu *join* dan *purge* hingga pembentukan *association rules*. [10]

Pada penelitian ini akan dianalisis data penjualan tiket penerbangan menggunakan metode *Clustering* yang bertujuan melihat pola penjualan, serta pola perilaku pembelian konsumen dalam pembelian tiket penerbangan pada PT. Garuda Indonesia, Cabang Batam. Tujuan analisis ini adalah untuk memperoleh informasi perilaku konsumen untuk menyusun rekomendasi strategi serta *knowledges* yang berguna untuk mempertahankan segmen pasar PT. Garuda Indonesiadengan mempertimbangkan adanya rival bisnis dari perusahaan maskapai lain dan feedback sebagai bahan masukan kedepannya.

III. TINJAUAN PUSTAKA

A. DATA MINING

Data mining (DM) adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual. Patut diingat bahwa kata mining sendiri berarti usaha untuk mendapatkan sedikit barang berharga dari sejumlah besar material dasar. Karena itu DM sebenarnya memiliki akar yang panjang dari bidang ilmu seperti kecerdasan buatan (artificial intelligent), machine learning, statistik dan basis data. Beberapa teknik yang sering disebut-sebut dalam literatur DM antara lain : *clustering*, *classification*, *association rule mining*, *neural network*, dan *genetic algorithm*. [11]



Gambar 1 Tahap-tahap *Data mining*

Pada gambar 1 *Data mining* dapat dibagi menjadi beberapa tahap. Tahap-tahap tersebut bersifat *interaktif* di mana pemakai terlibat langsung atau dengan perantara *knowledge base*. Tahap-tahap ini diilustrasikan di Gambar 1:

1. Pembersihan data (untuk membuang data yang tidak konsisten dan *noise*)
2. Integrasi data (penggabungan data dari beberapa sumber)
3. Transformasi data (data diubah menjadi bentuk yang sesuai untuk di-*mining*)
4. Aplikasi teknik DM
5. Evaluasi pola yang ditemukan (untuk menemukan yang menarik/bernilai)
6. Presentasi pengetahuan (dengan teknik visualisasi). [11]

B. CLUSTERING

Clustering adalah proses pengelompokan benda serupa ke dalam kelompok yang berbeda, atau lebih tepatnya partisi dari sebuah data set kedalam subset, sehingga data dalam setiap subset memiliki arti yang bermanfaat. Sebuah *cluster* terdiri dari kumpulan benda-benda yang mirip antara satu dengan yang lainnya dan berbeda dengan benda yang terdapat pada *cluster* lainnya. Algoritma *clustering* terdiri dari dua bagian yaitu secara hirarkis dan secara partitional. Algoritma hirarkis menemukan *cluster* secara berurutan dimana *cluster* ditetapkan sebelumnya, sedangkan algoritma partitional menentukan semua kelompok pada waktu tertentu. [12]

C. K-means

Algoritma *K-means* merupakan salah satu algoritma dengan *partitional*, karena *K-means* didasarkan pada penentuan jumlah awal kelompok dengan mendefinisikan nilai *centroid* awalnya. Algoritma *K-means* menggunakan proses secara berulang-ulang untuk mendapatkan basis data *cluster*. Dibutuhkan jumlah *cluster* awal yang diinginkan sebagai masukan dan menghasilkan jumlah *cluster* akhir sebagai output.

Jika algoritma diperlukan untuk menghasilkan *cluster* K maka akan ada K awal dan K akhir. Metode *K-means* akan memilih pola K sebagai titik awal *centroid* secara acak. Jumlah iterasi untuk mencapai *cluster centroid* akan dipengaruhi oleh calon *cluster centroid* awal secara random dimana jika posisi *centroid* baru tidak berubah. Data yang memiliki jarak pendek atau terdekat dengan *centroid* akan membentuk sebuah *cluster*. [12]

Algoritma *K-means*

1. Tentukan k sebagai jumlah cluster yang akan dibentuk
2. Tentukan k Centroid (titik pusat cluster) awal secara random/acak.

$$v = \frac{\sum_{i=1}^n x_i}{n} \quad ; i = 1, 2, 3, \dots, n \quad (1)$$

Rumus *K-mean*

Dimana;

v : centroid pada cluster

xi : objek ke-i

n : banyaknya objek/jumlah objek yang menjadi anggota cluster

3. Hitung jarak setiap objek ke masing-masing *centroid* dari masing-masing *cluster*. Untuk menghitung jarak antara objek dengan *centroid* dapat menggunakan *Euclidian Distance*.

$$d(x, y) = \|x - y\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad ; i = 1, 2, 3, \dots, n \quad (2)$$

Rumus *Euclidian Distance*

Dimana;

xi : objek x ke-i

yi : daya y ke-i

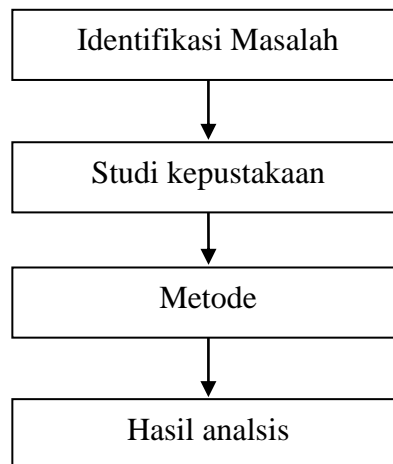
n : banyaknya objek

4. Alokasikan masing-masing objek ke dalam *centroid* yang paling dekat
5. Lakukan iterasi, kemudian tentukan posisi *centroid* baru dengan menggunakan persamaan (1)

6. Ulangi langkah 3 jika posisi *centroid* baru tidak sama. [13]

IV. METODE PENELITIAN

Pada metode penelitian ini diuraikan langkah-langkah sistematis dan terarah yang akan dijadikan acuan sebagai kerangka penelitian. Tahapan penelitian ini dibagi menjadi Empat bagian yaitu, (1) Identifikasi masalah, (2) studi kepustakaan, (3) metode, dan (4) hasil analisis.



Gambar 2 Tahapan penelitian

Pada gambar 2 penelitian ini akan menggunakan metode *Clustering* dengan algoritma *K-means*. Metode ini diharapkan dapat memberikan hasil analisis yang akurat berdasarkan pada data penjualan Tiket Garuda Indonesia, Cabang Batam.

V. HASIL DAN PEMBAHASAN

Dalam pengolahan data penjualan penerbangan ini digunakan metode *Clustering* dengan *algoritma K-mean*. Untuk manajemen data yang cukup banyak dilakukan proses pivoting terlebih dahulu pada program Microsoft Excel agar memudahkan dalam proses penjumlahan data sehingga pengolahan dapat dilakukan dengan mudah pada program SPSS, karena data mentah harus dirubah kedalam bentuk numeric agar proses pengolahan dapat dilakukan. Pada pengolahan data penjualan tiket penerbangan dengan jumlah total data sebanyak 210328 selama kurung waktu 2014-2015 dengan menggunakan Program SPSS.

| Name | Type | Width |
|----------|---------|-------|
| Bulan | String | 10 |
| Domes | Numeric | 20 |
| Inter | Numeric | 20 |
| Trans | Numeric | 20 |
| Langsung | Numeric | 20 |

Gambar 3 Atribut data

Pada gambar 3 adalah atribut-atribut yang digunakan dalam pengolahan data selama kurang waktu 2 tahun dari periode 2014-2015.

| Name | Type | Width |
|-----------|---------|-------|
| Bulan | String | 15 |
| Sunday | Numeric | 10 |
| Monday | Numeric | 10 |
| Tuesday | Numeric | 10 |
| Wednesday | Numeric | 10 |
| Thursday | Numeric | 10 |
| Friday | Numeric | 10 |
| Saturday | Numeric | 10 |

Gambar 4 Atribut data hari

Pada gambar 4 adalah atribut hari untuk menentukan rata-rata pemelihan penerbangan berdasarkan hari dalam kurung waktu seminggu pada data selama tahun 2014-2015.

Hasil pengolahan data dijelaskan pada pembahasan di bawah ini.

A. Analisis SPSS Penerbangan Domestik Tahun 2014-2015

Pada analisis *Final Cluster Centers* pada tabel I di bawah, data penerbangan domestik tahun 2014-2015 dikelompokkan berdasarkan angka yang meliputi:

- Angka 1 untuk jumlah penerbangan terbanyak
- Angka 2 untuk jumlah penerbangan yang terkecil, dan
- Angka 3 untuk jumlah penerbangan sedang.

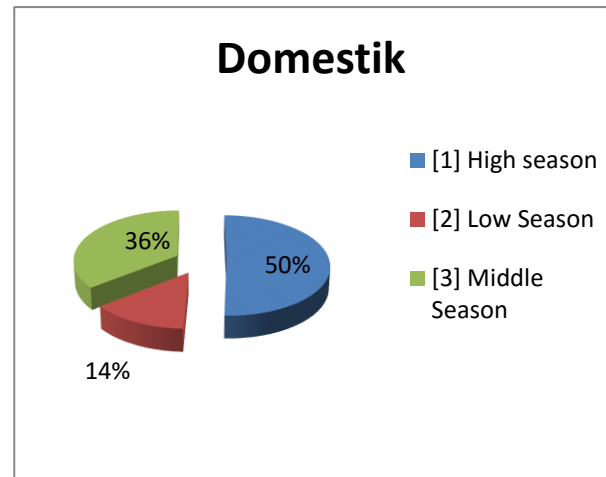
Pada hasil final cluster pengelompokan dilakukan secara otomatis dari program SPSS. Jadi angka-angka tersebut hanya merupakan label untuk pengelompokan dari hasil terkecil sampai yang terbesar.

TABEL I

HASIL CLUSTER PENERBANGAN DOMESTIK

| | Final Cluster Centers | | |
|-------|-----------------------|------|------|
| | Cluster | | |
| | 1 | 2 | 3 |
| Domes | 10889 | 2958 | 7673 |

Penggambaran hasil cluster dari tabel I di atas menggambarkan populasi pengelompokan data yang dapat di dilihat pada diagram dibawah ini.



Gambar 5 Diagram penerbangan domestik

Pada gambar 5 digambarkan populasi yang dibagi menjadi 3 bagian masing-masing memiliki nilai berbeda-beda. Populasi dengan nilai 50% untuk penerbangan dengan kategori *high season* atau pada waktu padat, 36% untuk penerbangan dengan kategori *middle season* atau pada waktu sedang, dan 14% untuk penerbangan dengan kategori paling sedikit atau *low season*.

Dari pengolahan data-data penjualan tiket penerbangan domestik, hasil *cluster* dapat dilihat pada tabel II dibawah ini:

TABEL II

PENERBANGAN DOMESTIK

| Bulan | Domes | QCL_1 |
|-------------|-------|-------|
| January | 11328 | 1 |
| February | 10095 | 1 |
| March | 8570 | 3 |
| April | 7494 | 3 |
| May | 8803 | 3 |
| June | 9707 | 1 |
| July | 11432 | 1 |
| August | 10865 | 1 |
| September | 8122 | 3 |
| October | 10276 | 1 |
| November | 9085 | 3 |
| December | 11537 | 1 |
| January15 | 2958 | 2 |
| February15 | 7502 | 3 |
| March15 | 8933 | 3 |
| April15 | 8369 | 3 |
| May15 | 8900 | 3 |
| June15 | 8678 | 3 |
| July15 | 11870 | 1 |
| August15 | 7050 | 3 |
| September15 | 5949 | 3 |
| October15 | 6297 | 3 |
| November15 | 5338 | 3 |
| December15 | 6006 | 3 |

Tabel II adalah hasil pengolahan pada SPSS, QCL_1 adalah hasil dari proses *cluster* domes yang secara otomatis di kelompokkan berdasarkan kemiripan jumlah data. Angka 1,2, dan 3 menunjukan kelompok populasi pada tiap bulan dalam kurung waktu 2014-2015.

Pada tabel II kelompok 1 adalah bulan-bulan pada data dua tahun selama kurung waktu 2014-2015, dengan jumlah penerbangan sebanyak 10889.

Untuk kelompok 2 adalah bulan-bulan pada data dua tahun selama kurung waktu 2014-2015 dengan jumlah terkecil yang terdapat pada bulan Januari15 dengan jumlah 2958. Sedangkan untuk kelompok 3 adalah bulan-bulan pada data dua tahun selama kurung waktu 2014-2015 dengan jumlah sedang sebanyak 7673.

Dilihat dari tabel II, penerbangan terbanyak yang dikelompokkan dengan angka 1 maka bulan-bulan dengan frekuensi terbanyak berada pada bulan:

- Januari 2014
- Febuari 2014
- Juni 2014
- Juli 2014
- Agustus 2014
- Oktober 2014
- Desember 2014
- Juli 2015

Dengan jumlah data selama 2 tahun untuk penerbangan domestik didapatkan hasil selama 8 bulan terjadi puncak penerbangan terbanyak. Dengan 7 bulan di tahun 2014 dan 1 bulan di tahun 2015.

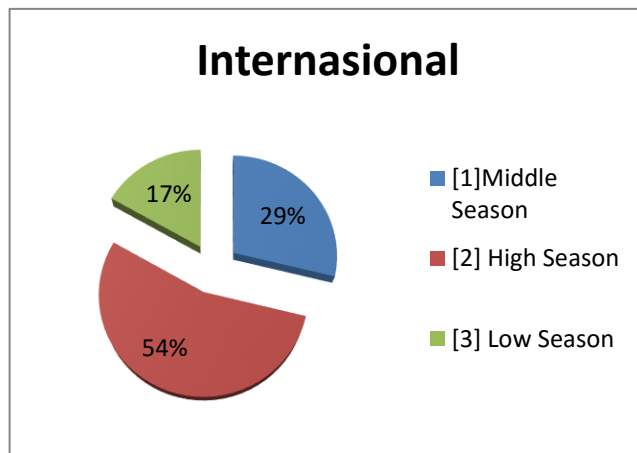
B. Analisis SPSS penerbangan internasional tahun 2014-2015

TABEL III
HASIL CLUSTER PENERBANGAN
INTERNASIONAL

| Final Cluster Centers | | | |
|-----------------------|---------|-----|-----|
| | Cluster | | |
| | 1 | 2 | 3 |
| Inter | 248 | 469 | 145 |

Pada tabel III analisis *Final Cluster Centers* di atas, data penerbangan internasional tahun 2014-2015 dikelompokkan berdasarkan angka yang meliputi:

- Angka 1 untuk jumlah penerbangan sedang
- Angka 2 untuk jumlah penerbangan terbanyak, dan
- Angka 3 untuk jumlah penerbangan terkecil



Gambar 6 Diagram penerbangan internasional

Pada gambar 6 digambarkan populasi yang dibagi menjadi 3 bagian masing-masing memiliki nilai berbeda-beda. Populasi dengan nilai 54% untuk penerbangan dengan kategori high season atau pada waktu padat, 29% untuk penerbangan dengan kategori middle season atau pada waktu sedang, dan 17% untuk penerbangan dengan kategori paling sedikit atau low season.

TABEL IV
PENERBANGAN INTERNASIONAL

| Bulan | Inter | QCL_1 |
|-------------|-------|-------|
| January | 172 | 3 |
| February | 122 | 3 |
| March | 101 | 3 |
| April | 99 | 3 |
| May | 209 | 1 |
| June | 242 | 1 |
| July | 225 | 1 |
| August | 533 | 2 |
| September | 226 | 1 |
| October | 342 | 1 |
| November | 255 | 1 |
| December | 182 | 3 |
| January15 | 177 | 3 |
| February15 | 143 | 3 |
| March15 | 142 | 3 |
| April15 | 145 | 3 |
| May15 | 328 | 1 |
| June15 | 207 | 1 |
| July15 | 230 | 1 |
| August15 | 125 | 3 |
| September15 | 147 | 3 |
| October15 | 404 | 2 |
| November15 | 190 | 3 |
| December15 | 218 | 1 |

Tabel IV adalah hasil pengolahan pada SPSS, QCL_1 adalah hasil dari proses *cluster inter* yang secara otomatis di kelompokkan berdasarkan kemiripan jumlah data. Angka 1,2, dan 3 menunjukkan kelompok populasi pada tiap bulan dalam kurung waktu 2014-2015.

Pada tabel IV kelompok 1 adalah bulan-bulan pada data dua tahun selama kurung waktu 2014-2015 dengan jumlah penerbangan sebanyak 248. Untuk kelompok 2 dengan jumlah penerbangan sebanyak 469. Dan kelompok 3 dengan jumlah penerbangan sebanyak 145.

Pada tabel IV penerbangan terbanyak dikelompokkan dalam angka 2, maka bulan-bulan dengan frekuensi penerbangan terbanyak pada bulan:

- Agustus 2014
- Oktober 2015

Dengan jumlah data selama 2 tahun untuk penerbangan internasional didapat 2 bulan dengan frekuensi terbanyak 1 bulan di tahun 2014 dan 1 bulan di tahun 2015.

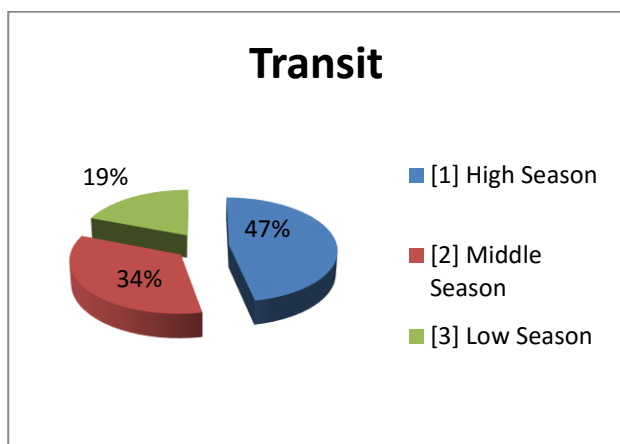
C. Analisis SPSS penerbangan transit tahun 2014-2015

TABEL V
HASIL CLUSTER PENERBANGAN TRANSIT

| Final Cluster Centers | | | |
|-----------------------|---------|------|------|
| | Cluster | | |
| | 1 | 2 | 3 |
| Trans | 3710 | 2686 | 1518 |

Pada tabel V analisis *Final Cluster Centers* di atas, data penerbangan Transit tahun 2014-2015 dikelompokkan berdasarkan angka yang meliputi:

- Angka 1 untuk jumlah penerbangan tinggi
- Angka 2 untuk jumlah penerbangan sedang, dan
- Angka 3 untuk jumlah penerbangan terkecil.



Gambar 7 Diagram penerbangan transit

Pada gambar 7 digambarkan populasi yang dibagi menjadi 3 bagian, masing-masing memiliki nilai yang

berbeda-beda. Populasi dengan nilai 49% untuk penerbangan dengan kategori high season atau pada waktu padat, populasi dengan nilai 34% dengan kategori middle season atau pada waktu sedang, dan populasi dengan nilai 19% dengan kategori low season atau dengan kategori paling sedikit.

TABEL VI
PENERBANGAN TRANSIT

| Bulan | Trans | QCL_1 |
|-------------|-------|-------|
| January | 3683 | 1 |
| February | 2958 | 2 |
| March | 2238 | 2 |
| April | 2177 | 2 |
| May | 2961 | 2 |
| June | 3335 | 1 |
| July | 3974 | 1 |
| August | 3895 | 1 |
| September | 3119 | 2 |
| October | 3838 | 1 |
| November | 3422 | 1 |
| December | 4066 | 1 |
| January15 | 1188 | 3 |
| February15 | 2462 | 2 |
| March15 | 3113 | 2 |
| April15 | 3002 | 2 |
| May15 | 3111 | 2 |
| June15 | 2797 | 2 |
| July15 | 3465 | 1 |
| August15 | 2376 | 2 |
| September15 | 2333 | 2 |
| October15 | 2711 | 2 |
| November15 | 1847 | 3 |
| December15 | 2252 | 2 |

Tabel VI adalah hasil pengolahan pada SPSS, QCL_1 adalah hasil dari proses *cluster trans* yang secara otomatis di kelompokkan berdasarkan kemiripan jumlah data. Angka 1,2, dan 3 menunjukkan kelompok populasi pada tiap bulan dalam kurung waktu 2014-2015.

Pada tabel VI kelompok 1 adalah bulan-bulan pada data dua tahun selama kurung waktu 2014-2015 dengan jumlah penerbangan sebanyak 3710. Untuk kelompok 2 dengan jumlah penerbangan sebanyak 2686. Dan kelompok 2 dengan jumlah penerbangan 1518.

Pada tabel VI penerbangan terbanyak dikelompokkan dengan angka 1, maka bulan-bulan dengan frekuensi terbanyak ada pada bulan:

- Januari 2014
- Juni 2014
- Juli 2014
- Agustus 2014
- Oktober 2014
- November 2014
- Desember 2014
- Juli 2015

Dari jumlah data selama 2 tahun untuk penerbangan transit dengan 8 bulan untuk frekuensi penerbangan terbanyak. Terdiri dari 7 bulan di tahun 2014 dan 1 bulan di tahun 2015.

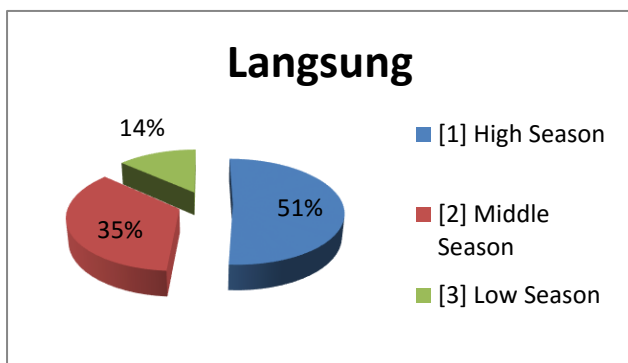
D. Analisis SPSS Penerbangan Langsung Tahun 2014-2015

TABEL VII
HASIL CLUSTER PENERBANGAN LANGSUNG

| | Final Cluster Centers | | |
|----------|-----------------------|------|------|
| | Cluster | | |
| | 1 | 2 | 3 |
| Langsung | 7375 | 5120 | 1947 |

Pada tabel VII analisis Final Cluster Centers di atas, data penerbangan Langsung tahun 2014-2015 dikelompokkan berdasarkan angka yang meliputi:

- Angka 1 untuk jumlah penerbangan terbanyak
- Angka 2 untuk jumlah penerbangan yang sedang, dan
- Angka 3 untuk jumlah penerbangan kecil.



Gambar 8 Diagram penerbangan langsung

Pada gambar 8 di atas digambarkan populasi yang dibagi menjadi 3 bagian, masing-masing memiliki nilai yang berbeda-beda. Populasi dengan nilai 51% dengan kategori high season atau pada waktu padat, 35% dengan kategori middle season atau pada waktu sedang, dan 14% dengan kategori low season atau dengan kategori paling sedikit.

TABEL VIII
PENERBANGAN LANGSUNG

| Bulan | Langsung | QCL_1 |
|-----------|----------|-------|
| January | 7817 | 1 |
| February | 7259 | 1 |
| March | 6433 | 1 |
| April | 5416 | 2 |
| May | 6051 | 2 |
| June | 6614 | 1 |
| July | 7683 | 1 |
| August | 7503 | 1 |
| September | 5229 | 2 |

| | | |
|-------------|------|---|
| October | 6780 | 1 |
| November | 5918 | 2 |
| December | 7653 | 1 |
| January15 | 1947 | 3 |
| February15 | 5183 | 2 |
| March15 | 5962 | 2 |
| April15 | 5512 | 2 |
| May15 | 6117 | 2 |
| June15 | 6088 | 2 |
| July15 | 8635 | 1 |
| August15 | 4799 | 2 |
| September15 | 3763 | 2 |
| October15 | 3990 | 2 |
| November15 | 3681 | 2 |
| December15 | 3972 | 2 |

Tabel VII adalah hasil dari pengolahan pada Program SPSS, QCL_1 adalah hasil dari proses pengolahan menggunakan algoritma cluster langsung yang secara otomatis di kelompokkan berdasarkan kemiripan jumlah data, angka 1,2, dan 3 menunjukkan kelompok populasi pada setiap bulan dalam kurung waktu 2014-2015

Pada tabel VII bulan-bulan dari kelompok 1 pada data-data dalam kurung waktu dua tahun selama 2014-2015 dengan jumlah penerbangan sebanyak 7375. Untuk kelompok 2 jumlah penerbangan sebanyak 5120. Dan yang terakhir merupakan kelompok 3 dengan jumlah sebanyak 1947.

Pada tabel VII penerbangan terbanyak dikelompokkan dengan angka 1, maka bulan-bulan dengan frekuensi terbanyak ada pada bulan

- Januari 2014
- Februari 2014
- Maret 2014
- Juni 2014
- Juli 2014
- Agustus 2014
- Oktober 2014
- Desember 2014
- Juli 2015

Dari data selama 2 tahun untuk penerbangan langsung dengan 9 bulan untuk frekuensi penerbangan terbanyak yang terdiri dari 8 bulan di tahun 2014 dan 1 bulan di tahun 2015.

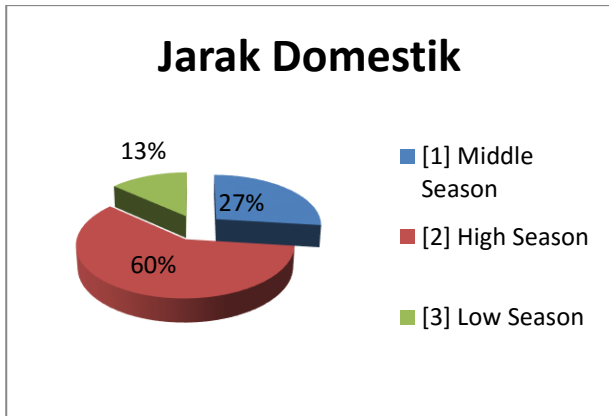
E. Analisis SPSS jarak hari pemesanan penerbangan domestik 2014-2015

TABEL IX
HASIL CLUSTER JARAK PEMESANAN DOMESTIC

| | Final Cluster Centers | | |
|-------|-----------------------|----|---|
| | Cluster | | |
| | 1 | 2 | 3 |
| Domes | 14 | 31 | 7 |

Pada tabel IX analisis Final Cluster Centers di atas data jarak hari pemesanan penerbangan domestik tahun 2014-2015 dikelompokkan berdasarkan angka yang meliputi:

- Angka 1 untuk jumlah penerbangan Sedang
- Angka 2 untuk jumlah penerbangan yang terbanyak, dan
- Angka 3 untuk jumlah penerbangan terkecil.



Gambar 9 Diagram jarak penerbangan domestik

Pada gambar 9 digambarkan populasi yang dibagi menjadi 3 bagian, masing-masing memiliki nilai yang berbeda-beda. Populasi dengan nilai 60% untuk penerbangan dengan kategori high season, populasi dengan nilai 27% untuk penerbangan dengan kategori middle season, dan 13% untuk penerbangan dengan kategori low season.

TABEL X
PENERBANGAN DOMESTIK

| Bulan | Domes | QCL_1 |
|-------------|-------|-------|
| January | 12 | 1 |
| February | 8 | 3 |
| March | 7 | 3 |
| April | 7 | 3 |
| May | 8 | 3 |
| June | 11 | 1 |
| July | 20 | 1 |
| August | 14 | 1 |
| September | 5 | 3 |
| October | 7 | 3 |
| November | 7 | 3 |
| December | 15 | 1 |
| January15 | 14 | 1 |
| February15 | 5 | 3 |
| March15 | 4 | 3 |
| April15 | 5 | 3 |
| May15 | 7 | 3 |
| June15 | 10 | 3 |
| July15 | 31 | 2 |
| August15 | 11 | 1 |
| September15 | 10 | 3 |

| | | |
|------------|----|---|
| October15 | 7 | 3 |
| November15 | 5 | 3 |
| December15 | 14 | 1 |

Tabel X adalah hasil dari pengolahan pada program SPSS, QCL_1 adalah hasil dari proses pengolahan menggunakan algoritma cluster domes, yang secara otomatis di kelompokkan berdasarkan kemiripan jumlah data, angka 1,2, dan 3 menunjukkan kelompok populasi pada data setiap bulan dalam kurung waktu 2014-2015

Pada tabel X bulan-bulan dari kelompok 1 pada data-data dalam kurung waktu dua tahun selama 2014-2015 dengan jumlah penerbangan sebanyak 14, untuk kelompok 2 dengan jumlah penerbangan sebanyak 31, dan kelompok 3 dengan jumlah penerbangan sebanyak 7.

Pada tabel X penerbangan terbanyak dikelompokkan dengan angka 2, maka bulan-bulan dengan frekuensi penerbangan domestik, terbanyak ada pada bulan Juli tahun 2015.

F. Analisis SPSS jarak hari pemesanan penerbangan Internasional 2014-2015

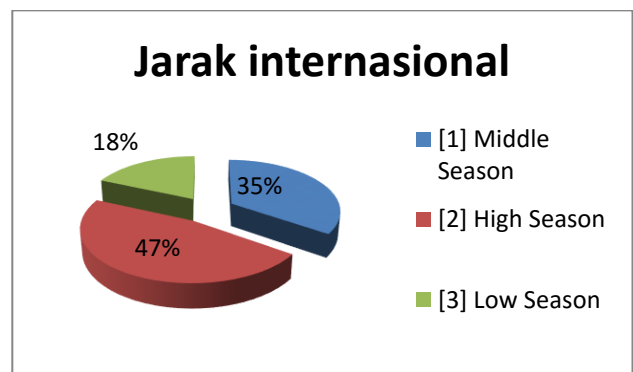
TABEL XI

JARAK HARI PEMESANAN INTERNASIONAL

| | Final Cluster Centers | | |
|-------|-----------------------|----|---|
| | Cluster | | |
| | 1 | 2 | 3 |
| Inter | 17 | 23 | 9 |

Pada tabel XI analisis Final Cluster Centers di atas data jarak hari pemesanan penerbangan domestik tahun 2014-2015 dikelompokkan berdasarkan angka yang meliputi:

- Angka 1 untuk jumlah penerbangan Sedang
- Angka 2 untuk jumlah penerbangan yang terbanyak, dan
- Angka 3 untuk jumlah penerbangan terkecil.



Gambar 10 Diagram jarak penerbangan internasional

Pada gambar 10 diagram di atas digambarkan populasi yang dibagi menjadi 3 bagian, masing-masing memiliki nilai yang berbeda-beda. Populasi dengan nilai 47% untuk

penerbangan dengan kategori high season, populasi dengan nilai 35% untuk penerbangan dengan kategori middle season, dan 18% untuk penerbangan dengan kategori low season.

TABEL XII
PENERBANGAN INTERNASIONAL

| Bulan | Inter | QCL_1 |
|-------------|-------|-------|
| January | 17 | 1 |
| February | 22 | 2 |
| March | 10 | 3 |
| April | 15 | 1 |
| May | 11 | 3 |
| June | 10 | 3 |
| July | 16 | 1 |
| August | 17 | 1 |
| September | 10 | 3 |
| October | 8 | 3 |
| November | 9 | 3 |
| December | 19 | 1 |
| January15 | 26 | 2 |
| February15 | 14 | 1 |
| March15 | 9 | 3 |
| April15 | 7 | 3 |
| May15 | 9 | 3 |
| June15 | 10 | 3 |
| July15 | 21 | 2 |
| August15 | 15 | 1 |
| September15 | 18 | 1 |
| October15 | 18 | 1 |
| November15 | 20 | 1 |
| December15 | 18 | 1 |

Tabel XII adalah hasil dari pengolahan pada program SPSS, QCL_1 adalah hasil dari proses pengolahan menggunakan algoritma *cluster ineter*, yang secara otomatis di kelompokkan berdasarkan kemiripan jumlah data. Angka 1, 2, dan 3, menunjukkan kelompok dari populasi pada data setiap bulan dalam kurung waktu 2014-2015.

Pada tabel XII bulan-bulan dalam kelompok 1 pada data-data dalam kurun waktu dua tahun selama 2014-2015 dengan jumlah penerbangan sebanyak 17, untuk kelompok 2 dengan jumlah penerbangan sebanyak 23, dan kelompok 3 dengan jumlah penerbangan sebanyak 9.

Pada tabel XII penerbangan terbanyak dikelompokkan dengan angka 2, maka bulan-bulan dengan frekuensi terbanyak ada pada bulan:

- Februari 2014
- Januari 2015
- Juli 2015

Dari data 2 tahun pada penerbangan internasional dengan hasil 3 bulan dengan frekuensi penerbangan terbanyak yang terdiri dari 1 bulan di tahun 2014 dan 2 bulan di tahun 2015.

G. Analisis rata-rata jarak hari pemesanan penerbangan Internasional dan Domestik 2014-2015

TABEL XIII

RATA-RATA JARAK HARI PEMESANAN

| Descriptive Statistics | | | | | |
|------------------------|----|---------|---------|-------|----------------|
| | N | Minimum | Maximum | Mean | Std. Deviation |
| Domes | 24 | 4 | 31 | 10.12 | 6.052 |
| Inter | 24 | 7 | 26 | 14.47 | 5.094 |
| Valid N (listwise) | 24 | | | | |

Pada tabel XIII analisis rata-rata pemesanan tiket penerbangan internasional dan domestik di dapatkan hasil bahwa rata-rata pemesanan tiket untuk penerbangan domestik 10 hari sebelum keberangkatan, artinya jarak untuk pemesanan domestik kira-kira 1 minggu 3 hari sebelum keberangkatan para penumpang telah melakukan pembelian tiket.

Sedangkan untuk penerbangan internasional rata-rata pemesanan 14 hari sebelum keberangkatan, artinya 2 minggu sebelum keberangkatan internasional para penumpang telah melakukan pembelian tiket.

H. Analisis SPSS rata-rata pemilihan hari keberangkatan penerbangan Tahun 2014

TABEL XIV

PEMILIHAN HARI KEBERANGKATAN

| Descriptive Statistics | | | | | |
|------------------------|----|---------|---------|---------|----------------|
| | N | Minimum | Maximum | Mean | Std. Deviation |
| Sunday | 12 | 1041 | 1945 | 1552.58 | 268.831 |
| Monday | 12 | 950 | 1910 | 1458.33 | 247.023 |
| Tuesday | 12 | 881 | 1646 | 1299.83 | 236.269 |
| Wednesday | 12 | 1018 | 2041 | 1426.75 | 317.476 |
| Thursday | 12 | 1001 | 2099 | 1359.58 | 325.561 |
| Friday | 12 | 976 | 1992 | 1465.42 | 330.833 |
| Saturday | 12 | 961 | 1817 | 1439.33 | 231.567 |
| Valid N (listwise) | 12 | | | | |

Pada tabel XIV menunjukkan rata-rata pemesanan penerbangan berdasarkan hari keberangkatan. Dari hasil analisis data penerbangan di dapatkan hasil rata-rata frekuensi keberangkatan berada pada hari Minggu dengan nilai 1552.58, dari data hasil pengolahan, terlihat bahwa frekuensi penerbangan paling padat rata-rata pada hari Minggu, dan hari keberangkatan dengan frekuensi terkecil berada pada hari Selasa dengan nilai 1299.83.

V. SIMPULAN

Berdasarkan hasil pengolahan data-data penjualan tiket penerbangan PT. Garuda Indonesia, cabang Batam menggunakan SPSS didapatkan kesimpulan sebagai berikut:

1. Metode *Cluster* dengan menggunakan algoritma *K-means* menghasilkan kelompok-kelompok bulan dengan trafik penerbangan terkecil, sedang, dan terbanyak.
2. Dari pengolahan data penjualan PT. Garuda Indonesia, cabang Batam menghasilkan informasi yang cukup berguna dalam melihat pola pembelian tiap bulan, serta jarak rata-rata setiap orang dalam menentukan pembelian dan keberangkatan.
3. Dari hasil pengolahan adanya pola berbeda tiap tahun selama kurung waktu dua tahun, pola dengan pembelian tetap terbanyak pada data penerbangan domestik adalah 8 bulan untuk frekuensi terbanyak dalam kurung waktu 2 tahun jadi jika di bagi berdasarkan 24 bulan berarti dalam 1 tahun hanya 4 bulan dengan frekuensi terbanyak. Untuk 8 bulan dalam 1 tahun hanya diisi oleh penerbangan dengan frekuensi sedang dan kecil. Sedangkan untuk penerbangan internasional hanya 2 bulan dengan frekuensi terbanyak selama kurung waktu 24 bulan.
4. Sedangkan untuk penerbangan transit dan langsung, total penerbangan transit sebanyak 8 bulan dan penerbangan langsung 9 bulan dalam kurung waktu 24 bulan dari data dua tahun. Artinya rasio rata-rata penerbangan transit dan langsung hampir sama berada di frekuensi terbanyak selama 8-9 bulan.
5. Yang berikutnya adalah jarak pemesanan untuk penerbangan. Dari hasil pengolahan data pembelian tiket penerbangan di atas di temukan hasil, rata-rata pembelian tiket penerbangan untuk penerbangan domestik adalah 10 hari sebelum keberangkatan, artinya kurang lebih 1 minggu para penumpang sudah melakukan proses pembelian. Dan untuk penerbangan internasional rata-rata pembelian tiket penerbangan adalah 14 hari sebelum keberangkatan, artinya kurang lebih 2 minggu para penumpang sudah melakukan proses pembelian.
6. Trafik terpadat dalam dalam kurung waktu 7 hari selama seminggu terdapat pada hari Minggu. Dengan hasil rata-rata 1552.58, jadi untuk penerbangan PT. Garuda Cabang Batam, rata-rata *customer* lebih memilih hari Minggu. Dan trafik terkecil terdapat pada hari Selasa dengan hasil rata-rata 1299.83.

DAFTAR PUSTAKA

- [1] Schiffman, Kanuk, *Consumer Behavior*, 10th ed., Prentice Hall; 10 Edition, 2009.
- [2] Hawkins, D. I., & Mothersbaugh D. L., *Consumer Behavior: Building Marketing Strategy*, 12th ed., New York: McGraw-Hill International Edition, 2013.
- [3] Engel, J. F., G. Blackwell, P. W. Miniard, *Perilaku Konsumen, Jilid 1*. Binarupa Aksara, Jakarta, 1994.
- [4] Agrawal, A. & Gupta, H, "Global K-means (GKM) Clustering Algorithm: A Survey." *International Journal of Computer Applications*, LIX(2), pp.20-24, 2013.
- [5] Santosa. B. *Data mining: Teknik pemanfaatan data untuk keperluan Bisnis*, Yogyakarta: Graha Ilmu, 2007.
- [6] Primawati, Verdia, Nurcahyo., "Classification of Sales Agent for Cement Distribution using K-means Clustering," *Journal of Computer Science and Information Technology*, Vol. 1, No 1, pp. 1-9, 2016.
- [7] Kashwan dan Velu, "Customer Segmentation Using Clustering and Data mining Techniques." *International Journal of Computer Theory and Engineering*, Vol. 5, No. 6, December. 2013.
- [8] Schaidnagel, Abele, Laux, Petrov, "Sales Prediction with Parametrized Analysis." *IARIA: The Fifth International Conference on Advances in Databases, Knowledge, and Data Applications*, 2013.
- [9] Subiakto dan Sukarno, "Financial Performance Analysis of PT. Garuda Indonesia, Tbk. in Comparison with Other Regional Airlines." *International Conference on Economics and Business Management*, Juli. 2015.
- [10] Siregar, "Implementasi Data mining Pada Penjualan Tiket Pesawat Menggunakan Algoritma Apriori (Studi Kasus: Jumbo Travel Medan)." *Pelita Informatika Budi Darma, Volume : VII, Nomor: 1*, Juli. 2014.
- [11] Lindawati, "Data mining Dengan Teknik Clustering Dalam Pengklasifikasian Data Mahasiswa Studi Kasus Prediksi Lama Studi Mahasiswa Universitas Bina Nusantara." *Seminar Nasional Informatika*, 2008.
- [12] Merliana, Ernawati, Santoso, "Analisa Penentuan Jumlah Cluster Terbaik Pada Metode K-means Clustering." *Prosiding seminar nasional multi disiplin ilmu & call for papers unisbank*, 2016, ISBN: 978-979-3649-81-8.
- [13] Ediyanto, Mara, Satyahadewi, "Pengklasifikasian Karakteristik Dengan Metode K-means Cluster Analysis." *Buletin Ilmiah Mat. Stat. dan Terapannya*, Volume 02, No. 2, hal 133 – 136, 2013