Implementasi *Business Intelligence* dan *Big Data Analytics* Pada Industri Logistik dan Pergudangan Melalui Dataset Perusahaan Ninja Warehouse & Logistic

William

Fakultas Teknik dan Informatika Universitas Multimedia Nusantara,

Tangerang, Indonesia

[william13@student.umn.ac.id](mailto:william13@student.umn.ac.id)

*Abstrak*— Di era industri 4.0, data memiliki peran yang sangat krusial dalam perkembangan teknologi dan implementasinya dalam kehidupan sehari-hari, salah satunya dalam sektor bisnis yang sedang berkembang pesat dalam beberapa tahun kebelakang ini, yaitu dalam bidang logistik dan pergudangan. Dengan berkembangnya dunia *e-commerce* sebagai platform untuk masyarakat melakukan belanja secara online, industri logistik mendapatkan arus volume yang lebih besar, mengingat segala kemudahan yang disediakan untuk melakukan transaksi pembelian secara online melalui layanan website maupun *e-commerce.* Salah satu perusahaan yang bergerak dalam bidang logistik dan pergudangan yaitu perusahaan Ninja Warehouse & logistic yang menyediakan jasa ekspedisi pengiriman produk dengan bebeberapa metode *shipment* seperti lewat jalur darat maupun udara. Sebelum produk-produk customer tersebut dikirimkan, perusahaan Ninja Warehouse & Logistic menyimpan dan mengelola *flow* produk di dalam *warehouse* yang terbagi dalam beberapa block. Dalam keberlangsungan proses binsis perusahaan Ninja Warehouse & Logistic, terdapat beberapa tantangan utama yang di hadapi, seperti keterlambatan proses pengiriman produk sampai ke tujuan (alamat customer). Dengan mengimplementasikan *Business Intelligence* dan *Big Data Analytics,* perusahaan dapat menganalisa data dan mengolah data menjadi informasi yang berguna dapat digunakan menjadi dasar utama suatu perusahaan dalam pengambilan keputusan, dalam kasus ini yaitu di sektor industri logistik dan pergudangan. Dalam penelitian ini, saya menggunakan platform Microsoft Power BI dan SAS Visual Analytics untuk melakukan visualisasi yang dapat digunakan sebagai insight bagi perusahaan Ninja Warehouse & Logistics untuk mengambil keputusan bisnis perusahaan serta perencanaan strategi yang akan diterapkan kedepannya.

Kata Kunci — Big Data Analytics, Business Intelligence, Data Science, e-commerce, Logistic, SAS Studio, Power BI, Warehousing.

# I. LATAR BELAKANG

## A. Latar Belakang

Operasi logistik merupakan komponen krusial di dalam keberlangsungan *e-commerce suply chain* [1]*.* Perkembangan teknologi yang memungkinkan setiap orang untuk melakukan transaksi secara online baik melalui platform *e-commerce* dan platform lainnya seperti website, maupun media sosial mempengaruhi volume transaksi yang terjadi dalam perusahaan logistik. *E-commerce service providers* kerap memfokuskan layanan utama mereka dengan cara melakukan kerja sama dengan perusahan-perusahaan logistik yang terpercaya untuk memperlancar proses bisnis mereka, mengingat waktu pengiriman logistik barang customer memiliki peranan penting dalam menentukan kepuasan pelanggan ketika berbelanja secara *online* .

Dalam upaya untuk memenuhi kebutuhan hidupnya, manusia membutuhkan *daily usage stuff* atau barang-barang bulanan untuk memenuhi kebutuhan sehari-hari mereka. Barang-barang tersebut bisa di dapatkan melalui supermarket maupun toko-toko grosir terdekat di wilayah mereka masing-masing. Kegiatan *purchasing* ini berjalan secara *offline* / datang langsung ke toko fisik selama bertahun-tahun. Seiring dengan berjalannya waktu dan perkembangan teknologi yang pesat, internet menghasilkan inovasi baru yang memungkinkan manusia untuk melakukan transaksi dan mendapatkan kebutuhan mereka masing-masing secara online tanpa datang langsung ke toko fisik. Hal ini dianggap mempermudah kehidupan manusia dan meningkatkan efisiensi tanpa perlu memikirkan jarak yang ditempuh untuk mendapatkan kebutuhan mereka, disinilah tugas logistik yang memiliki peranan penting untuk mengantar produk-produk tersebut ke alamat konsumen [2].

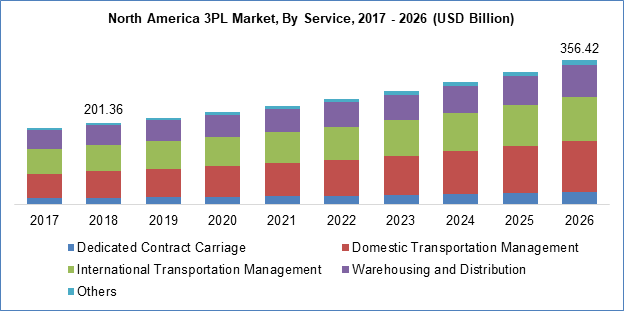


Figure 1 3PL Market size in North Americaperiod 2017-2026 (estimated)

Berdasarkan data yang didapat dari halaman website polaris market research, market value dari perusahaan 3PL (*third-party* *logistics)* mengalami tren kenaikan setiap tahunnya dari periode 2017 sampai dengan sekarang. Perusahaan 3PL sendiri merupakan perusahaan yang bergerak dalam bidang layanan logistik yang mencakup segala sesuatu yang melibatkan pengelolaan satu atau lebih aspek kegiatan *procurement* dan *fullfilment activities.* Hal ini menunjukan bahwa perkembangan indsutri logistik berbanding lurus dengan perkembangan *e-commerce* yang semakin bertambah volume pengguna sertiap tahunnya. Gambar 1 menunjukan data mengenai market value dari perusahaan 3PL dalam beberapa kategori, seperti *Delicated Contract Carriage, International Transportation Management, Domestic Transportation Management,* dan *Warehousing + Distribution.* Jika dilihat dari data tersebut, bisa disimpulkan bahwa mayoritas perusahaan memfokuskan pada bidang *Domestic Transportation Management* [3].



Figure 2 Ninja Warehouse & Logistic

Ninja Warehouse & Logistic merupakan perusahaan berbasis 3PL (*third-party logistics)*yang menyediakan layanan pengiriman ekspedisi dan layanan pengelolaan produk melalui *warehouse* / pergudangan. Ninja Warehouse & Logistic memiliki slogan “*Beyond Speed Delivery!”* yang mengutamakan kepuasan pelanggan dengan mengantar paket logistik customer tepat waktu sesuai dengan visi dan misi perusahaan untuk membantu pelanggan dengan menyediakan jasa logsitik dan pengelolaan pergudangan dengan terjangkau serta membantu pelaku bisnis / UMKM lainnya dalam memasarkan produknya secara online melalui beberapa platform. Layanan yang disediakan oleh perusahaan Ninja Warehouse & Logistic menjamin keselamatan produk dan ketepatan waktu pengriman ke alamat tujuan. Dalam rangka untuk meningkatkan kepuasan pelanggan, setiap pelaku bisnis yang memasarkan secara online harus memperhatikan perusahaan logistik yang dipilih untuk melakukan kerja sama. Ketepatan waktu dan keselamatan produk merupakan faktor utama yang berpengaruh dalam meningkatkan kepuasan pelanggan.

Industri logistik dan pergudangan di Indonesia sendiri mengalami perkembangan yang cukup pesat seperti yang terjadi dalam negara-negara lain. Hal ini dipengaruhi oleh berkembangnya teknologi dalam bentuk e-commerce dan platform lainnya yang mempermudah masyarakat untuk melakukan transaksi secara online. Dengan contoh kasus dalam *e-commerce* yang ada di Indonesia seperti *Shopee, Tokopedia, Lazada, Bukalapak,* dan platform lainnya. Mayoritas setiap platform *E-commerce* tersebut melakukan kerja sama dengan pihak 3PL (*Thirdt-Party Logistic)* untuk melakukan proses pengiriman produk sampai ke alamat customer. Platform *E-Commerce* melakukan kerja sama baik dengan menberikan diskon-diskon dan promo menarik jika memilih pengiriman dengan perusahan logistik yang bekerja sama, misalkan dengan potongan gratis ongkir serta gratis biaya asuransi produk. Dalam contoh kasus di platform *E-Commerce* yang beredar di Indonesia sendiri, terdapat beberapa *e-commerce* yang menyediakan layanan logsitik mereka sendiri, contohnya yaitu *Shopee* dengan layananan *Shopee Express* mereka. Dalam contoh kasus di platform *E-commerce* Shopee, pelanggan tidak diwajibkan untuk menggunakan layanan logistik mereka saja dan masih bisa mengaplikasikan voucher-voucher diskon potongan ongkir menarik jika memilih layanan logistik lainnya [4].

Tantangan utama yang dihadapi oleh perusahaan Ninja Warehouse & Logistic sendiri memiliki masalah utama yang sama dengan perusahaan logistik lainnya, yaitu masalah keterlambatan dan juga keselamatan produk, khususnya untuk pengiriman produk-produk yang beresiko hancur dalam perjalanan. Jika hal ini diabaikan dan tidak kunjung diperbaiki kedepannya, maka akan berdampak terhadap kepuasan pelanggan ketika menggunakan layanan jasa pengiriman dari Ninja Wrehouse & Logistic dan berpengaruh terhadap penjualan serta omset yang didapat dari perusahaan. *Customer Experience* dalam proses bisnis Ninja Warehouse & Logistic sangatlah penting, mengingat semakin berkembangnya dunia digital dan platform *E-Commerce* yang terus mengalami pertambahan jumlah pengguna setiap tahunnya. Jika terdapat pelanggan yang merasa kurang puas dan dirugikan ketka menggunakan layanan perusahaan, mereka dapat dengan mudah mepersebar review buruk tersebut secara online melalui beberapa platform. Hal ini bisa sangat merugikan perusahaan jika diabaikan.

Dalam rangka untuk meningkatkan kualitas layanan dari Ninja Warehouse & Logistic, perusahaan dapat mengimplementasikan *Business Intelligence* dan *Big Data Analytics* untuk menghasilkan *insight-insight* yang kedepannya bisa digunakan sebagai acuan utama dalam mengambil keputusan bisnis perusahaan. Dalam laporan ini, perusahaan Ninja Warehouse & Logistic akan menggunakan *Business Intelligence Diagram* dan proses-proses visualisasi dan analitik, baik secara *predictive* maupun *deskriptif* untuk mengetahui pola-pola yang terdapat di dalam dataset. *Insight* dan wawasan tersebut kiranya dapat membantu keberlangsungan bisnis perusahaan Ninja Warehouse & Logistic kedepannya sehingga perusahaan dapat berjalan dengan optimal dan mencapai target-target bisnis yang sudah ditentukan sebelumnya.

# II. TINJAUAN TEORITIS & METODE PENELITIAN

## A. Tinjauan Teoritis

1. ***Third-Party Logistics (3PL)***

*Third-Party Logistics* adalah suatu instansi / perusahaan yang menyediakan jasa logistik kepada perusahaan lain untuk mendistribusikan serta menyimpan barang / produk yang akan segera di kirimkan ke alamat tujuan masing-masing. Dalam *e-commerce suply chain,* proses logistik mulai dari pickup sampai dengan delivery memakan lebih dari 60% total *supply-chain cost.*



Figure 3 3PL Main Activities

*Third-Party Logistics* menyediakan layanan dasar dalam industri logistik, yakni transportasi, *warehousing, delivery¸*serta *pick and pack*. *Third-Party Logistics* memiliki peranan penting dalam keberlangsungan bisnis secara online, khususnya untuk para *e-commerce business owner,*yang berhubungan langsung dengan pihak logistik setiap harinya selama proses bisnis berlangsung. Dengan menemukan mitra pelaku 3PL yang tepat, pelaku bisnis secara online bisa mengoptimalkan pelayanan toko mereka serta meningkatkan *Customer Experience,* mengingat pentingnya durasi pengiriman produk yang dilakukan oleh logistik untuk menentukan kepuasan pelanggan dalam melakukan transaksi secara online [5].

1. ***Business Intelligence***

*Business Intelligence* merupakan salah satu prosedur teknikal yang digunakan untuk mengubah data-data perusahaan yang mentah menjadi suatu informasi yang bermanfaat dan *insightful* dengan menggunakan bantuan perangkat lunak [6]. *Business Intelligence* mencakup semua proses, arsitektur dan teknologi yang digunakan untuk mengolah data mentah tersebut. Tools dari *Business Intelligence* memungkinkan penggunanya untuk melakukan data analysis dan membuat laporan dalam bentuk dashboard interaktif yang terdiri dari beberapa objek visual seperti grafik, *charts*, bahkan sampai dengan visualisasi peta sesuai dengan bidang bisnis masing-masing. Beberapa keuntungan dari implementasi *Business Intelligence* ke dalam perusahaan seperti menentukan *Key Performance Indicators* (KPI), membantu visualisasi data dalam bentuk yang interaktif sehingga mudah dipahami. Secara keseluruhan, *Business Intelligence* dapat meningkatkan kualitas operasi bisnis dalam suatu organisasi secara keseluruhan dengan memanfaatkan hasil analisa data yang relevan.



Figure 4 Business Intelligence Critical Aspects

Dengan mengimplementasikan *Business Intelligence* kedalam perusahaan, maka perusahaan dapat memperoleh manfaat yang dihasilkan dari *Business Intelligence*, salah satunya yaitu dapat meningkatkan *revenue* dan *profit* yang dihasilkan oleh perusahaan serta membuat data dan report yang bisa diakses oleh semua pihak dalam perusahaan dan mampu mengoperasikan proses bisnis dengan lebih efisien [7].

Beberapa komponen utama yang ada dalam *Business Intelligence* yakni *Data Warehousing, Data discovery and visualization, data quality management, data integration, self-service BI, Data Governance, dan Data Warehouse Modernization*. Semua komponen yang ada dalam *Business Intelligence* diintegrasikan untuk mengelola data-data perusahaan yang bersifat historis dengan hasil keluaran berupa report dalam dashboard interaktif yang mencakup seluruh informasi-informasi yang dibutuhkan oleh perusahaan dalam mengambil keputusan bisnis [8].

## *Big Data Analytics*

Big Data Analytics merupakan proses yang dilakukan untuk mengekstrak insight-insight yang bermanfaat, seperti korelasi dai pola-pola, trend pasar, dan juga prefrensi pelanggan [9]. Big Data Analytics juga memberikan beberapa manfaat seperti menghasilkan decision making yang akurat serta menghasilkan insight yang dapat dimanfaatkan untuk keberlangsungan bisnis suatu perusahaan. Pada era saat ini, Big Data Analytics menjadi tools yang sangat krusial dalam organisasi perusahaan, mengingat jumlah volume data pada saat ini yang terus bertambah dalam jumlah yang massive setiap tahunnya di setiap bidang indusstri. Dengan memanfaatkan dan menganalisa data tersebut, perusahaan dapat memprediksi tren-tren selanjutnya dengan tujuan untuk membuat keputusan bisnis yang tepat demi keberlangsungan perusahaan serta mengetahui behavior-behavior dari customer berdasarkan data perusahaan tersebut dengan tujuan untuk meningkatkan customer experience [10].

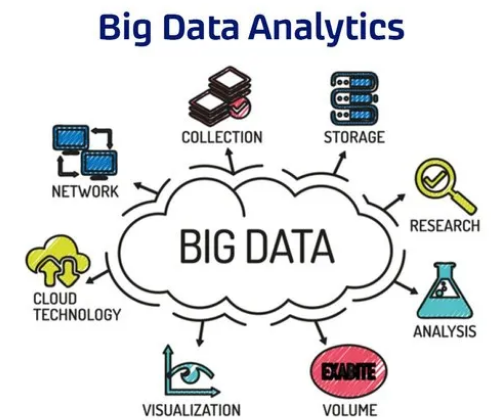


Figure 5 Big Data Analytics Component

Dalam perusahaan Ninja Warehouse & Logistics sendiri mennggunakan implementasi descriptive analytics dan predictive analytics. Berikut merupakan penjelasan dari kedua tipe analitik yang digunakan dalam perusahaan Ninja Warehouse & Logistics:

* ***Descriptive Analytics***

*Descriptive Analytics* merupakan salah satu jenis business analytics yang sifatnya mendeskripsikan analisa dari data historik dan menggambarkan situasi terkini dari perusahaan tersebut yang di implementasikan dalam bentuk laporan bisnis, dashboards, scorecards, dan data warehousing [11]. Jenis analitik ini ditujukan untuk mendefinisikan masalah-masalah yang timbul dalam bisnis perusahaan berdasarkan report data perusahaan pada periode waktu sebelumnnya. Descriptive Analytics penting untuk diimplementasikan karena dengan melakukan jenis analitik ini, pelaku bisnis dapat mempelajari past behaviors dari customer untuk menghasilkan outcome yang lebih baik kedepannya. Dalam contoh kasus perusahaan Ninja Warehouse & Logistic, Descriptive Analytics dapat membantu perusahaan untuk memahami permasalahan apa yang sedang terjadi dalam perusahaan dan dapat menganalisa behavior dari customer untuk mengetahui aspek apa saja yang bisa di perbaiki .

* ***Predictive Analytics***

***Predictive Analytics***merupakan salah satu jenis dari business alaytics yang bersifat “prediktif” untuk memprediksi hal apa saja yang akan terjadi dan berpengaruh bagi perusahaan di masa mendatang*.* Tidak hanya memprediksi apa yang akan terjadi, predictive analytics juga menganalisa faktor-faktor penyebab utama / pemicu hal tersebut terjadi sehingga memungkinkan perusahaan untuk lebih siap mengantisipasi hal tersebut. Predictive Analytics diimplementasikan dalam bentuk pembuatan model prediktif dengan menggunakan algoritma dari machine learning seperti decision tree, linear regression, neural network model, dan gradient boosting. Dalam contoh kasus perusahaan Ninja Warehouse & Logistic, hasil output dari Predictive Analytics berupa model prediktif dari algoritma yang diimplementasikan dengan tingkat akurasi yang berbeda-beda[12]*.*

## B. Metode Penelitian

Metode penelitian yang digunakan dalam artikel jurnal ini yaitu metode penelitian kuantitatif yang terdiri dari angka dan hasil akurasi dari pengetesan model dengan tujuan untuk menentukan hubungan antara satu variabel (independen) dan variabel lainnya (dependen). Penelitian ini juga menggunakan *Business Intelligence Framework* sebagai kerangka utama dalam melakukan proses implementasi ke dalam perusahaan Ninja Warehouse & Logistic.

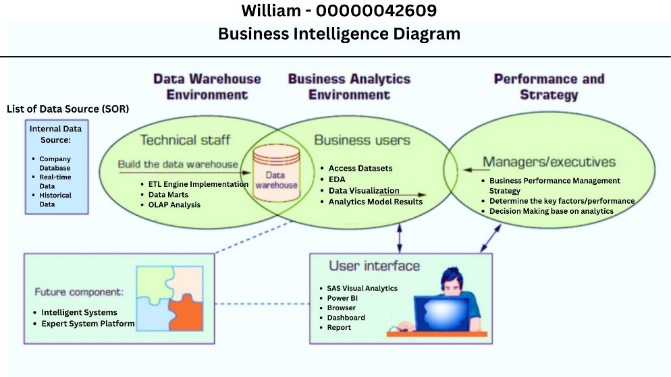


Figure 6 Business Intelligence Framework Diagram

Beberapa bagian utama yang terdapat dalam *Business Intelligence Framework Diagram* yaitu *List of Data Source (*SOR), *Data Warehouse Environment, Buiness Analytics Environment, Performance and Strategy, Future Component,* dan User Interface. Berikut merupakan penjelasan dari masing-masing bagian:

Penelitian kali ini menggunakan dataset yang didapat dari salah satu website penyedia dataset yang bisa diakses secara gratis, yakni melalui website kaggle.com. penulis sudah memastikan dataset yang digunakan dalam penelitian kali ini tidak memiliki data NULL dan variabel yang terdapat dalam dataset memiliki hubungan yang relevan dengan topik yang dibahas.

* ***List of Data Source* (SOR)**

Descriptive Analytics merupakan salah satu jenis business analytics yang sifatnya mendeskripsikan List of Data Source (SOR) merupakan bagian pertama dari Business Intelligence Diagram yang menjelaskan mengenai sumber data yang beragam sesuai dengan industri dan requirements yang dibutuhkan oleh setiap perusahaan. Proses pengumpulan data yang dibutuhkan bisa didapatkan dari sumber data yang beragam seperti CRM, ERP sampai dengan database internal perusahaan. Dalam contoh kasus perusahaan Ninja Warehoue & Logistic, List of data source yang digunakan memanfaatkan data internal perusahaan yang didapat dari database perusahaan yang mencakup data-data baik bersifat real-time seperti data mengenai history purchase customer, harga barang yang dibeli customer, sampai dengan metode pengiriman yang dipilih customer (reguler/instant).

* ***Data Warehouse Environment***

Data Warehouse Environment merupakan bagian dari komponen business intelligence diagram yang menjelaskan mengani teknik / skill yang dibutuhkan untuk membangun arsitektur data warehouse yang baik. Proses Data Warehouse Environment berkaitan dengan seberapa kompleks data yang diperoleh pada bagian List of Data Source. beberapa teknik / abilities yang diperlukan dalam Data Warehouse Environment seperti *ETL Implementation* yang satu proses integrasi data yang melakukan ekstraksi data sebelum di proses ke Business Intelligence Data Warehouse. Sesuai dengan namanya, proses ETL dibagi menjadi 3 proses utama yakni EXTRACT untuk menarik dataset yang akan digunakan dari sumber external (melakukan ekspor data) sebelum menuju ke proses selanjutnya, yakni TRANSFORM.

Teknik lainnya yang digunakan dalam Data Warehouse Environment yaitu Data Marts yang merupakan bagian dari data warehouse yang di desain untuk digunakan dalam 1 department di suatu perusahaan secara spesifik karena data marts mengambil beberapa source data saja jika dibandingkan dalam data warehouse (dalam cakupan yg lebih kecil dan lebih fleksibel). Dalam contoh kasus perusahaan Ninja Warehouse & Logistic, Data marts bisa digunakan untuk memudahkan masing-masing department perusahaan mengakses data yang dibutuhkan sesuai dengan bidang masing-masing.

Teknik terakhir yang digunakan dalam Data Warehouse Enviroment yaitu Olap Analysis (On-Line Analytical Processing) yang merupakan salah satu abilities yang dibutuhkan dalam proses pembangunan data warehouse. Proses OLAP mengimplementasikan multidimensional analysis datawarehouse dengan tujuan untuk mendapatkan insight mengenai dataset yang di analisa. Dalam contoh kasus perusahaan Ninja Warehouse & Logistic, OLAP Analysis memungkinkan perusahaan untuk mengorganisir data dalam bentuk multi dimensional model sehingga data-data dapat mudah dimengerti oleh user.

* ***Business Analytics Environment***

Business Analytics Environment merupakan bagian dari Business Inteligence Diagram yang mencakup langkah apa saja yang harus dilakukan dalam proses analytics di sebuah perusahaan. Berikut merupakan komponen-komponen dalam Business Analytics Environment:

1. Access Datasets

Access Datasets mencakup proses mengakses datasets yang di miliki oleh perusahaan sebelum melakukan proses analytics. Datasets yang digunakan sesuai dengan yang sudah ditentukan berdasarkan pemilahan dalam proses List of Data Souce (SOR). Dalam contoh kasus perusahaan Ninja Warehouse & Logistic, dataset yang akan digunakan berasal dari database internal perusahaan yang mencakup detail history purchasing yang dilakukan oleh customer.

1. EDA (Exploratory Data Analysis)

Proses EDA sendiri mencakup proses Business Understanding untuk menentukan tujuan dan visi-misi dari bisnis yang dilakukan, proses Data Understanding sebagai tahap pengenalan dan pemahaman dataset yang akan digunakan, proses Data Preparation, Data Modelling, Evaluation, sampai dengan tahap Deployment.

1. Data Visualization

Proses data visualization ditujukan untuk memvisualisasikan data dalam bentuk grafik interaktif agar mudah di mengerti oleh user sehingga dapat menyampaikan informasi mengenai dataset yang di analisa dengan jelas. Dalam perusahaan Ninja Warehouse & Logistic, proses data visualization dilakukan dalam bentuk dashboard interaktif yang memuat berbagai visualisasi dari variabel dataset.

1. Analytics Model Results

Analytics Model Results mencakup proses analisa untuk mendapatkan informasi bisnis yang ditujukan dalam pengambilan keputusan perusahaan dengan membuat model prediksi dengan bantuan algoritma tertentu seperti Decision Tree, Linear Regression, Logistic Regression, Gradient Boosting, dan Neural Network Model.

* ***Performance and Strategy***

*Performance and Strategy* mencakup evaluasi performa kinerja perusahaan dan melakukan perancangan strategy yang akan dilakukan oleh perusahaan kedepannya. Proses ini dilakukan oleh petinggi perusahaan maupun manager yang memiliki wewenang. Bagian performance and strategy terdiri dari komponen Business process management (BPM) untuk memonitoring peforma kinerja perusahaan, menentukan key factors/key performance (KPI) untuk mengukur faktor-faktor utama yang memiliki dampak krusial terhadap kinerja perusahaan, dan komponen terakhir yaitu melakukan Decision Making based on Analytics Report untuk memenuhi target yang ditentukan dalam perusahaan.

* ***Future Component***

Future Component merupakan bagian dari Business Inteligence Diagram yang mencakup sistem yang akan disediakan oleh perusahaan untuk digunakan pada masa sekarang dan masa mendatang. Pada contoh kasus dalam perusahaan Ninja Warehouse & Logistic, perusahaan menyediakan Intelligent Systems yang merujuk pada software-software yang digunakan oleh perusahaan untuk melakukan keputusan analitik. Selain itu, terdapat Expert System sebagai platform analitik lanjutan untuk menunjang peforma perusahaan. Sebagai contoh dalam perusahaan Ninja Warehouse & Logistic, Future Component digunakan ketika akan melakukan tipe tipe analytics dalam big data seperti descriptive analytics dan predictive analytics.

* ***User Interface***

User Interface merupakan bagian dari Business Intelligence Diagram yang menjelaskan mengenai fitur/function apa yang akan ditampilkan melalui bantuan dari berbagai platform. Dalam contoh kasus perusahaan Ninja Warehouse & Logistic menggunakan SAS Studio yang diakses melalui browser dari device PC untuk melakukan implementasi EDA dan pembuatan model prediktif (neural network model, linear regression model, decision tree model, dll). Selain menggunakan SAS Studio, Ninja Warehouse & Logistic juga menggunakan platform PowerBI untuk menampilkan analisa deskriptif dalam bentuk dashboard interaktif.

## C. Objek Penelitian

Objek Penelitian yang digunakan yakni menggunakan dataset dari perusahaan Ninja Logistic & Warehouse yang berisikan mengenai informasi-informasi dan riwayat seputar transaksi dan deskripsi dari produk yang melewati proses logistik, baik dari jenis perjalanan yang ditempuh (darat / laut), berat barang, produk, sampai dengan data mengenai *gender* dari customer yang menggunakan jasa logistik dari Ninja Warehouse & Logistic

## D. Pemahaman Data

Proses *data understanding* merupakan proses pengenalan dan pemahaman dataset perusahaan yang akan digunakan dalam proses data science. Proses utama dalam data understanding secara umum dibagi menjadi 3 proses utama, yakni pendeskripsian data seperti menghitung jumlah rows dan berapa banyak variabel yang terdapat di dalam dataset yang akan digunakan.

Didalam dataset perusahaan Ninja Warehouse & Analytics sendiri, terdapat 12 kolom utama dan terdiri dari 11.000 baris Terdapat 16 variabel utama yakni yang akan di analisa, yakni ID, Warehouse, Mode of Shipment, Customer Care Calls, Customer Rating, Cost of The Product, Prior Purchase, Product Importance, Gender, Discount Offered, Weight in Grams, dan Reached on Time yang merupakan acuan utama untuk menentukan apakah pesanan pelanggan sampai dengan tepat waktu ke alamat masing-masing Berikut merupakan tampilan secara keseluruhan dataset:

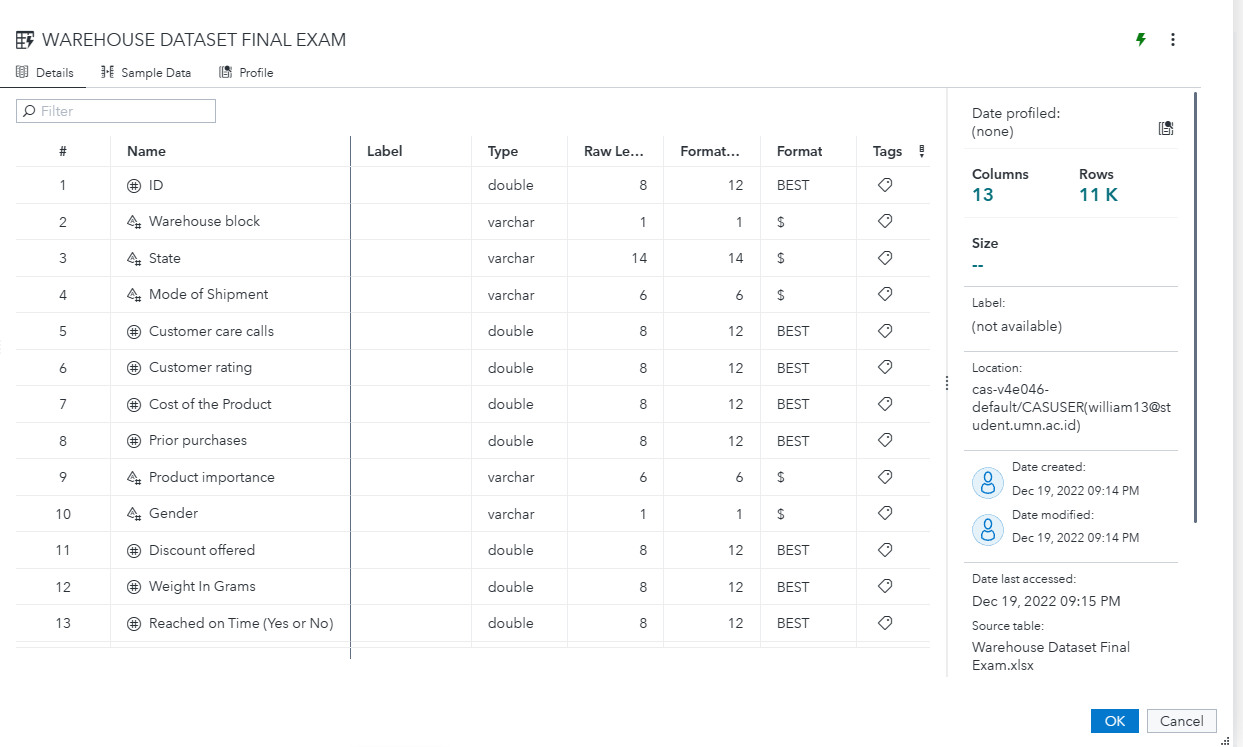


Figure 7 Dataset yang digunakan

# III. PEMBAHASAN

## Pengumpulan Data

Penelitian kali ini menggunakan dataset sekunder yang didapat dari salah satu website penyedia dataset yang bisa diakses secara gratis, yakni melalui website kaggle.com. Penulis sudah memastikan dataset yang digunakan dalam penelitian kali ini tidak memiliki data NULL dan variabel yang terdapat dalam dataset memiliki hubungan yang relevan dengan topik yang dibahas.

## Persiapan Data

Tahap persiapan data mencakup proses untuk mengubah dataset menjadi bentuk tabel proporsional yang akan digunakan untuk proses data science selanjutnya, yakni tahap eksplorasi data dan juga data modelling. Saya menggunakan fitur Prepare Data dalam SAS Drive untuk melakukan proses Standardize. Dengan melakukan Standardize dalam tahap persiapan data, kita bisa menghasilkan analisa yang akurat karena menggunakan data dengan format yang sesuai, proporsional dan konsisten sehingga dapat menghasilkan keputusan yang akurat dalam proses bisnis Ninja Warehouse & Logistics. Berikut merupakan hasil akhir dari dataset yang akan digunakan:

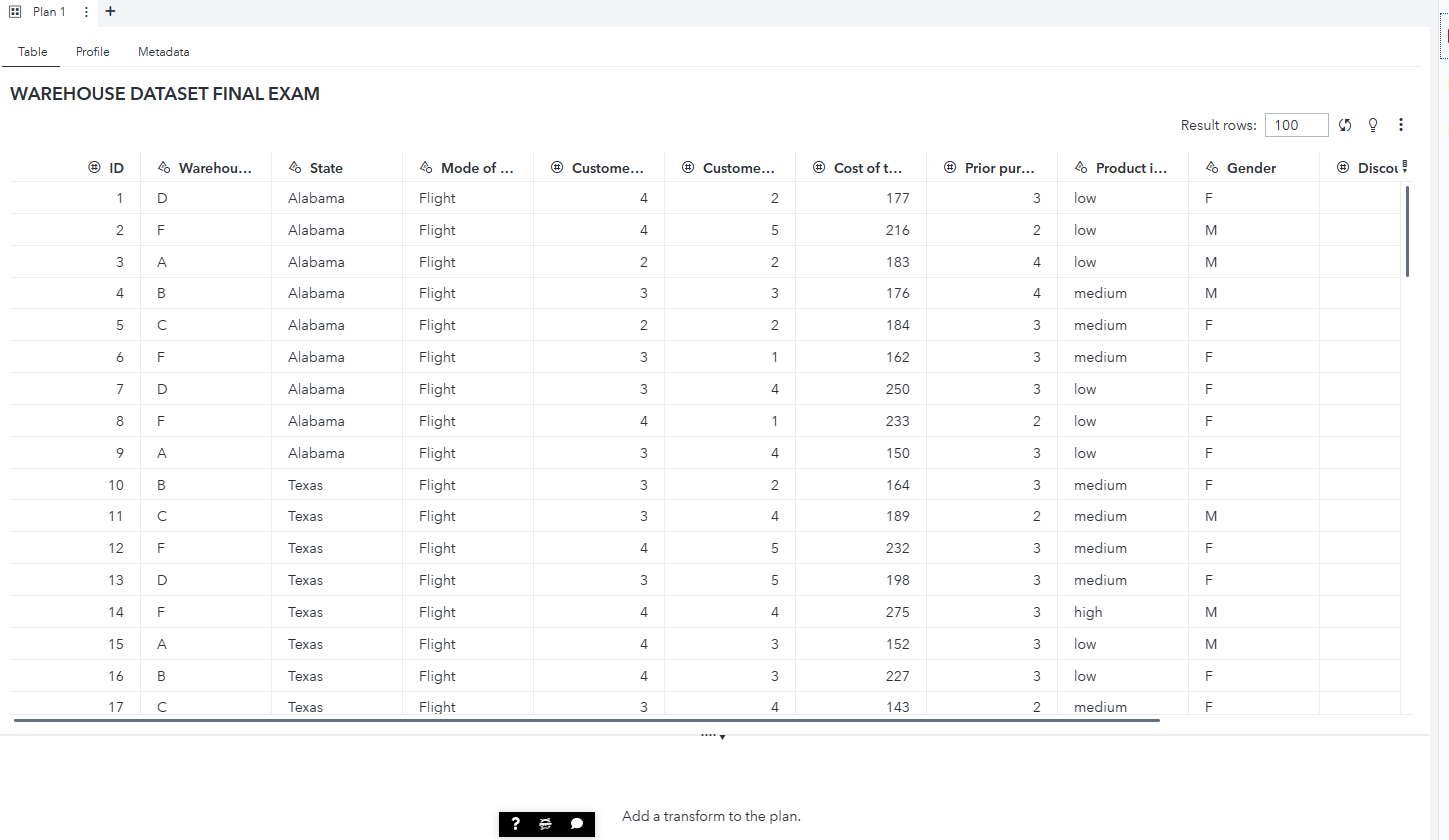


Figure 8 Dataset Setelah melakukan preparasi data

## Eksplorasi Data (Exploratory Data Analysis)

Setelah Tahap preparasi data selesai, langkah selanjutnya dalam mengimplementasikan proses *Data Science* yakni melakukan eksplorasi data. Pada proses ini data scientist menentukan data mana yang perlu dilakukan stukturisasi ulang dan cleansing jika diperlukan. Saya mengimplementasikan *Exploratory Data Analysis* dalam bentuk dashboard interaktif yang memuat grafik-grafik dan chart yang digunakan untuk mengetahui pola-pola dan menganalisa tren tertentu berdasarkan data set perusahaan yang diperoleh. *Exploratory Data Analysis* yang diterapkan dalam penelitian kali ini dibagi menjadi 2 jenis yaitu dalam bentuk *Descriptive Analytics* yang ditampilkan dengan memanfaatkan 2 tools, yaitu Microsoft Power BI dan SAS Visual Analytics.



Figure 9 Ninja Warehouse & Logistic Homepage (Power BI)

Figure 9 menunjukan tampilan homepage Descriptive Analytics perusahaan Ninja Warehouse & Logistic menampilkan kumpulan report-report yang berkaitan langsung dengan kinerja perusahaan Ninja Warehouse & Logistic. Terdapat 4 report utama yaitu Customer Report , Shipping Report, Product Info, dan Purchase Info. Pada tampilan Homepage menggunakan fitur Action dalam power BI untuk pindah ke report yang ingin dituju. Sebagai contoh, ketika user ingin melihat Shipping Report, user bisa memanfaatkan fitur Action dengan memencet tombol CTRL + Klik gambar report yang ingin dituju.

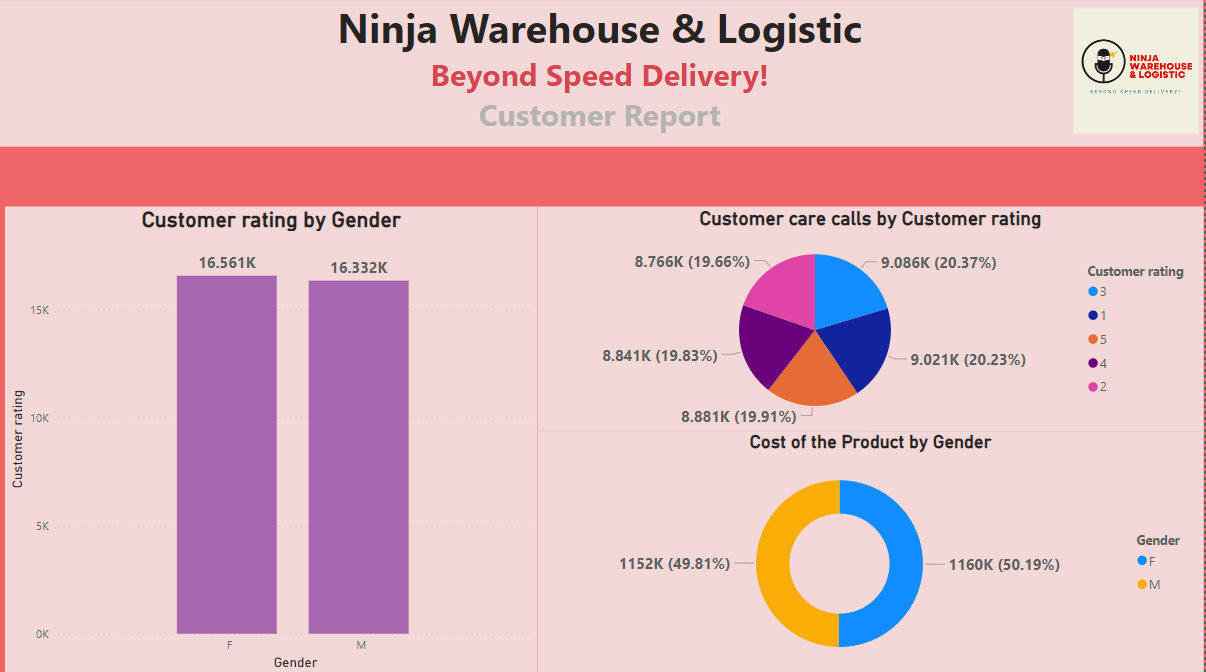


Figure 10 Exploratory Data Analysis (Frequency of Vehicle Year)

Figure 10 menunjukan tampilan 3 visualisasi utama yang menggambarkan behavior utama dari customer. Visualisasi pertama menggunakan Stacked Column Chart Visualization dalam power bi yang menunjukan jumlah nilai Customer Rating by Gender. Dari visualisasi 1 ini didapatkan hasil bahwa gender Female / perempuan memiliki nilai rating sedikit lebih banyak daripada gender male / laki-laki.

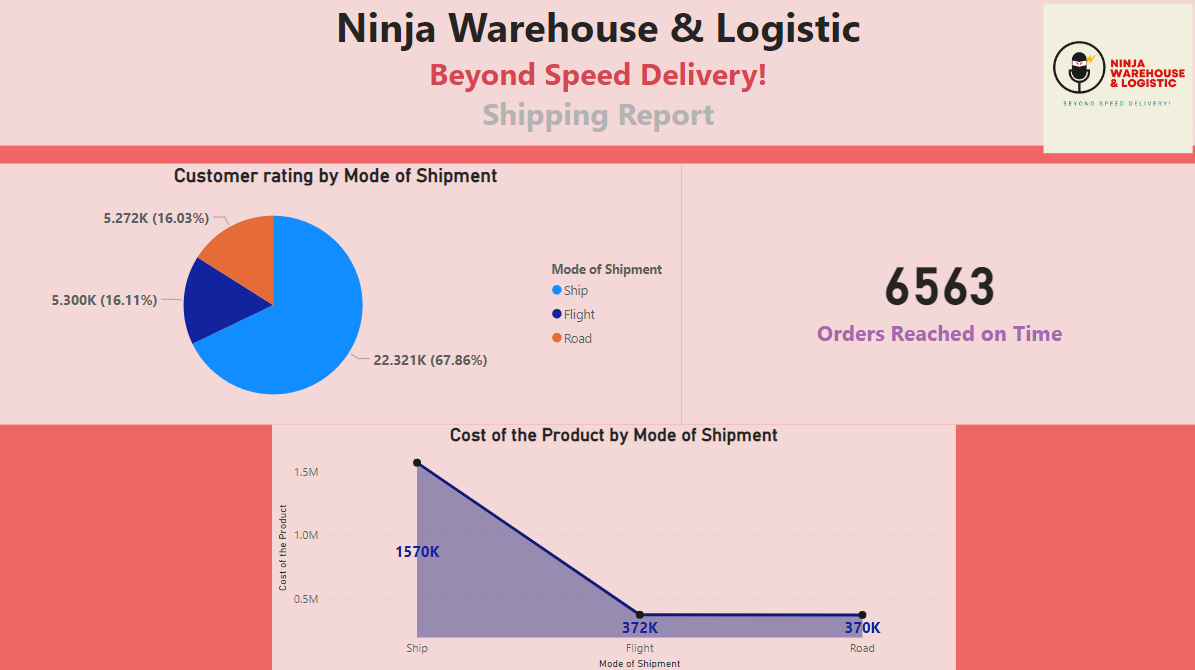


Figure 11 Exploratory Data Analysis (Frequency of Vehicle Type)

Figure 11 visualisasi menggunakan pie chart yang menunjukan frekuensi jumlah panggilan yang dilakukan oleh customer berdasarkan variabel Customer Rating (Customer Care Calls by Customer Rating). Hasil dari visualisasi ke 2 menunjukan bahwa Customer Rating dengan nilai 3 diikuti dengan Customer Rating dengan nilai 5 merupakan kelompok customer yang paling banyak melakukan panggilan sehingga dapat disimpulkan bahwa kelompok Customer yang cenderung “Buruk” tidak menghubungi Customer Care sesering kelompok Customer yang memiliki nilai lebih baik (>1).

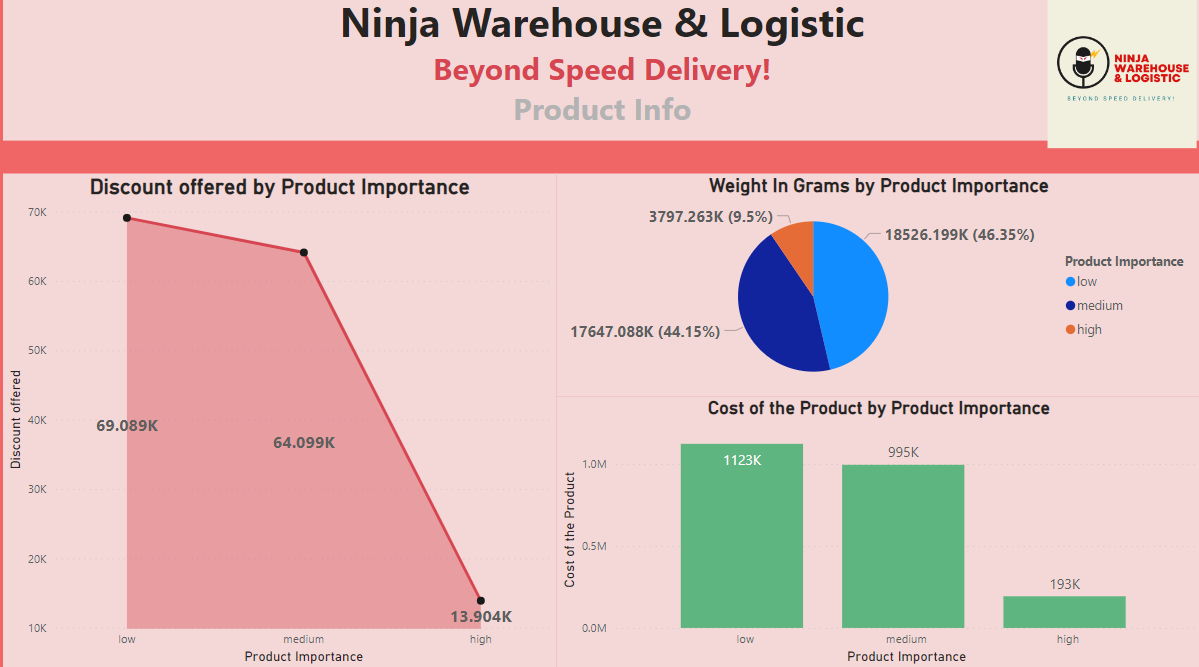


Figure 12 Exploratory Data Analysis (Frequenct of Annual Mileage)

Dalam Figure 12, terdapat 3 visualisasi utama yang menggambarkan laporan dari info produk yang dikirim. Visualisasi pertama menggunakan Stacked Area Chart Visualization dalam power bi yang menunjukan Discount Offered by Product Importance. Dari visualisasi 1 ini didapatkan hasil bahwa Product Importance dengan skala “Low” merupakan tingkatan produk yang paling banyak diberikan diskon oleh perusahaan dengan total diskon mencapai 69,089 Million USD, diikuti dengan tingkatan “Medium

# 

Figure 13 Exploratory Data Analysis (Frequency of Credit Score)

Figure 13 menunjukan bagian report Purchase Report yang terdapat 3 visualisasi utama yang menggambarkan laporan dari info purchasing yang dilakukan oleh customer. Visualisasi pertama menggunakan Line Chart Visualization dalam power bi yang menunjukan Customer Care Calls by Prior Purchases. Prior Purchase sendiri merujuk pada seberapa banyak pembelian sebelumnya yang dilakukan oleh customer. Dari visualisasi 1 ini didapatkan hasil bahwa customer dengan jumlah Prior Puchase sebanyak 3 kali merupakan kelompok customer yang paling banyak menghubungi bantuan customer service. Hal ini bertolak belakang dengan hipotesa awal yang memprediksi bahwa semakin banyak Prior Purchase yang dilakukan oleh customer maka semakin sering customer tersebut menghubungi bantuan Customer Service.

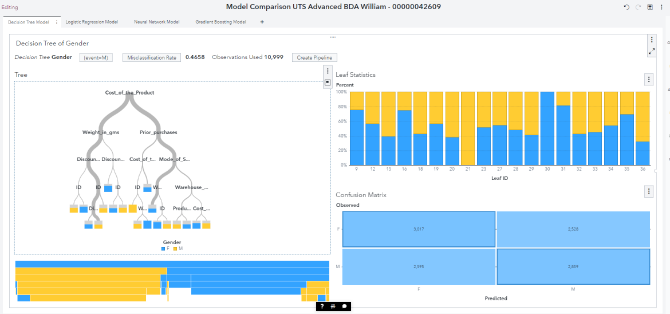


Figure 14 Decision Tree Model

Figure 14 Dari hasil modelling algoritma Decision Tree, didapatkan angka Training Misclassification Rate sebesar 0.4658 yang menunjukan tingkat kesalahan dalam memprediksi subjek sebesar 43,524% sehinnga memiliki akurasi sebesar 46,58%. Pada tampilan model Decision tree, terjadi pembagian pertama yang membagikan variabel Gender menjadi 2 grup, yakni Female dan Male. Berdasarkan hasil dari angka Misclassification Rate, bisa disimpulkan bahwa model decision tree memiliki tingkat akurasi yang kurang baik.

Sebuah gambar berisi meja

Description automatically generated

Figure 15 Logistic Regression Model

Dari hasil modelling algoritma Logistic Regression, didapatkan angka Training Misclassification Rate sebesar 0.4961 yang menunjukan tingkat kesalahan dalam memprediksi subjek sebesar 49,61% sehinnga memiliki akurasi yang lebih kecil dibandingkat dengan angka Misclassification Rate, yakni sebesar 40,39%. Pada tampilan model Logistic regression sama dengan Model Decision Tree dimana terjadi pembagian pertama yang membagikan variabel Gender menjadi 2 grup, yakni Female dan Male. Berdasarkan hasil dari angka Misclassification Rate, bisa disimpulkan bahwa model linear regression dalam dataset yang digunakan memiliki tingkat akurasi yang kurang baik (<50%).



Figure 16 Neural Network Model

Dari hasil modelling algoritma Neural Network Model, didapatkan angka Training Misclassification Rate sebesar 0.4918 yang menunjukan tingkat kesalahan dalam memprediksi subjek sebesar 49,18% sehinnga memiliki akurasi yang lebih kecil dibandingkat dengan angka Misclassification Rate, yakni sebesar 40,82% dengan jumlah observation yang didapat sebanyak 10.999. Pada tampilan model Neural Network sama dengan Model Decision Tree dan Logistic Regression dimana terjadi pembagian pertama yang membagikan variabel Gender menjadi 2 grup, yakni Female dan Male. Berdasarkan hasil dari angka Misclassification Rate, bisa disimpulkan bahwa algoritma Neural Network model dalam dataset yang digunakan memiliki tingkat akurasi yang kurang baik (<50%).

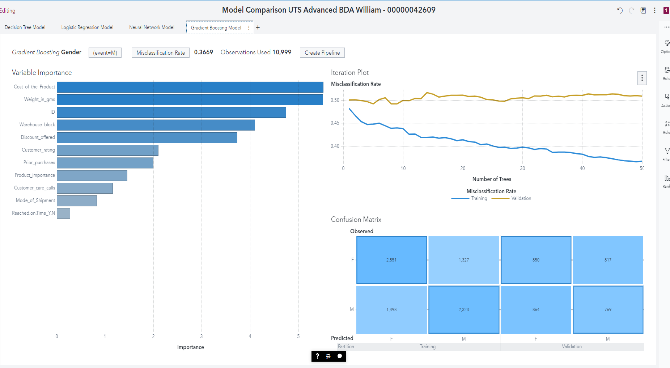


Figure 17 Gradient Boosting Model

Dari hasil modelling algoritma Gradient Boosting Model, didapatkan angka Training Misclassification Rate sebesar 0.3669 yang menunjukan tingkat kesalahan dalam memprediksi subjek sebesar 36,69% sehinnga memiliki akurasi sebesar 63,31% dengan jumlah observation yang didapat sebanyak 10.999. Pada tampilan model Linear regression sama dengan Model Decision Tree dan Linear Regression dimana terjadi pembagian pertama yang membagikan variabel Gender menjadi 2 grup, yakni Female dan Male. Berdasarkan hasil dari angka Misclassification Rate, bisa disimpulkan bahwa algoritma gradient boosting model merupakan algoritma terbaik jika dibandingkan dengan 3 algoritma sebelumnya dengan tingkat akurasi tertinggi (sebesar 63,31%).

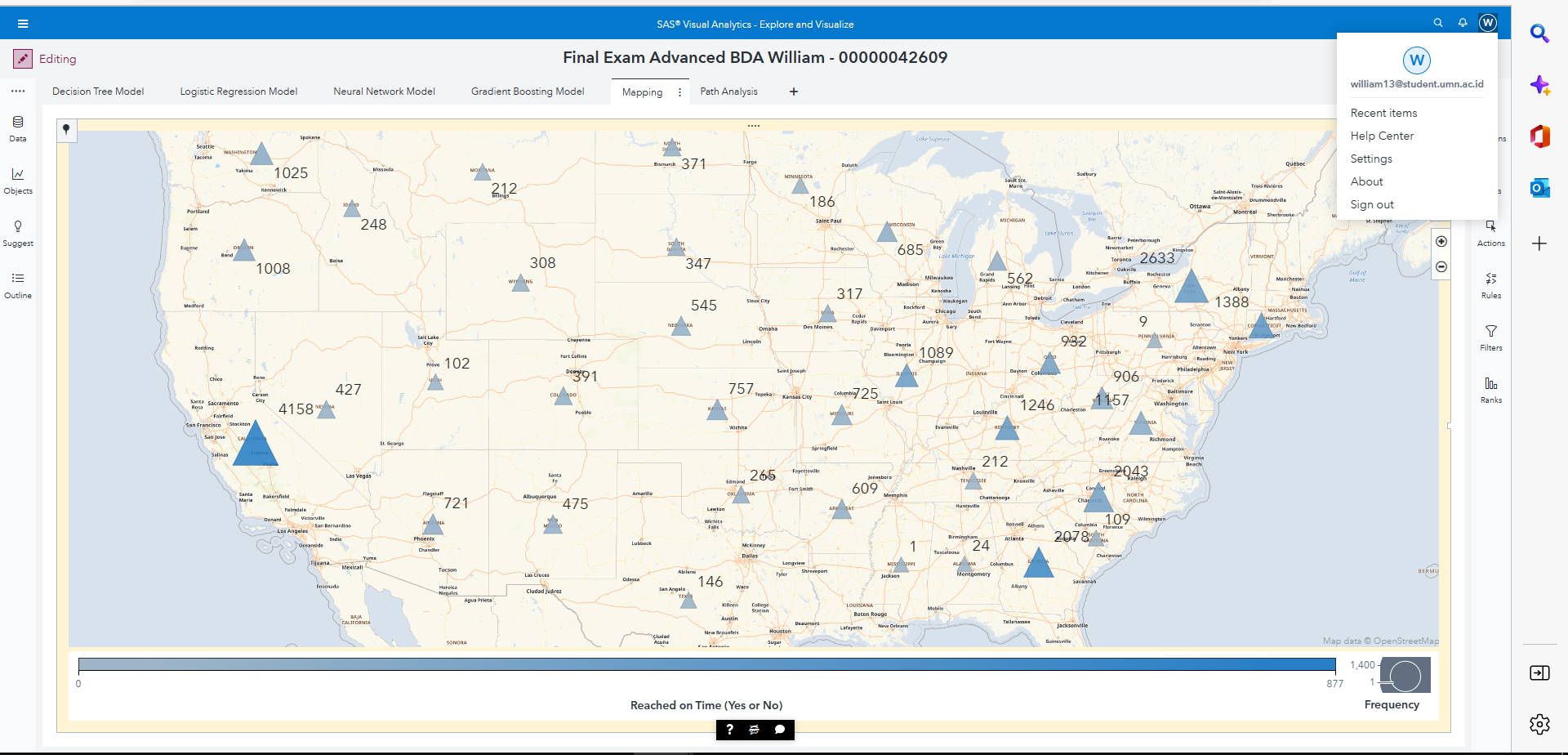


Figure 18 Visualisasi Cartography

Figure 18 menunjukan visualisasi Cartography dengan menggunakan object Geo Coordinate yang disediakan dalam menu SAS Visual Analytics. Dalam visualisasi ini menggunakan colom State yang berisikan negara-negara bagian sesuai dengan dataset dan dijadikan sebagai acuan geografis utama. Sedangkan pada bagian Color menggunakan variabel Reached on Time sebagai indikator utama untuk mengetahui apakah pesanan pelanggan sampai tepat waktu ke tujuan mereka atau tidak. Roles lainnya seperti Data Tip Values menggunakan variable yang sama dengan color, yaitu Reached on Time serta Prior Purchases dan juga Customer Rating. Sedangkan roles data labels menggunakan variabel Customer Rating yang dihitung dari semua total order dari setiap States. Marker Shape yang digunakan dalam visualisasi mapping ini menggunakan bentuk segitiga yang ditandai dengan warna biru. Semakin gelap warnanya, maka hal tersebut menunjukan bahwa skala pesanan sampai dengan tepat waktu semakin tinggi.

Hasil Cartography yang dihasilkan dari figure 18 menunjukan bahwa secara garis besar, *East Coast* atau negara bagian barat timur memiliki skor *Reached on Time* atau sampai dengan tujuan tepat waktu yang lebih baik dibandingkan dengan bagian *West Coast,*meskipun states California merupakan states dengan frekuensi terbaik (bisa dilihat dari label serta marker shape yang berwarna biru tua). Perusahaan Ninja Warehouse & Logistic bisa memfokuskan pada sektor *West Coast* untuk menganalisa faktor-faktor apa saja yang membuat pengiriman terhambat di negara bagian barat sehingga dapat meningkatkan *Customer Satisfaction* yang merata di seluruh negara bagian.

Sebuah gambar berisi peta

Description automatically generated

Figure 19 Interactive Mapping

Figure 19 menunjukan Interactive Mapping dengan menggunakan object Geo Network dengan menggunakan State sebagai source utama dengan target Geographic item 1 yang diambil dari variabel product importance. Sedangkan untuk bagian roles lainnya menggunakan variabel mode of shipment sebagai objek color dalam peta dan menggunakan data tip values dari variabel Reached on Time, Frequency, dan mode of shipment. Object lainnya yang digunakan yaitu list yang menunjukan metode pengiriman yang dibagi menjadi 3 jenis, yaitu flight, road dan ship dan object list (drop-down list) dengan menggunakan product importance sebagai category utama.

# IV. KESIMPULAN

Masalah utama yang dihadapi oleh perusahaan logistik & pergudangan yaitu keterlambatan pesanan serta kerusakan produk yang disebabkan selama perjalanan logistik. Berdasarkan hasil perbandingan modelling menggunakan Decision Tree, Logistic Regression, Gradient Boosting dan Neural Network Model, dapat disimpulkan bahwa Gradient Boosting merupakan model terbaik dengan nilai Misclassification Rate (Event) terendah dibanding dengan nilai model lainnya. Hasil analisa dari keempat model dan visualisasi *mapping (Cartography)* menunjukan bahwa masih terjadi inkonsistensi dalam hal pengantaran pesanan logistik dengan tepat waktu, berpacu pada hasil visualisasi mapping yang menunjukan bahwa *East Cost* memiliki tingkat pengantaran pesanan dengan tepat waktu yang lebih baik jika dibandingkan dengan *West Coast* . Dengan mengimplementasikan *Decsriptive* dan *Predictive Analytics,* diharapkan perusahaan Ninja Warehouse & Logistic dapat menggunakan insight yang didapatkan dari model prediktif sebagai acuan untuk mengambil keputusan bisnis kedepannya serta melakukan perbaikan terhadap sektor-sektor yang memiliki kinerja performa yang belum maksimal.

# UCAPAN TERIMA KASIH

Saya mengucapkan terimakasih sebesar-besarnya kepada Bapak Iwan Prasetiawan, S.Kom., M.M., sebagai dosen pembimbing mata kuliah IS 529 *Advanced Big Data Analytics* kelas B program studi Sistem Informasi Universitas Multimedia Nusantara yang telah membimbing saya selama berlangsungnya kelas 1 semester ini dan membantu mendiskusikan serta memberikan arahan mengenai tugas akhir.

# DAFTAR PUSTAKA

[1] M. I. D. Ranathunga, A. N. Wijayanayake, and D. H. H. Niwunhella, “Simulation-Based Efficiency Assessment of Integrated First-Mile Pickup and Last-Mile Delivery in an E-Commerce Logistics Network,” *Proc. - Int. Res. Conf. Smart Comput. Syst. Eng. SCSE 2022*, no. September, pp. 246–253, 2022, doi: 10.1109/SCSE56529.2022.9905083.

[2] R. Aziz, van J. Hillegersberg, and K. Kumar, “Inter Organizational Relationships Performance in Third Party Logistics: conceptual framework and case study,” *Pioneer. Solut. Supply Chain Manag. A Compr. Insight into Curr. Manag. Approaches*, no. January, pp. 105–126, 2010, [Online]. Available: http://www.narcis.nl/publication/RecordID/oai:doc.utwente.nl:77525

[3] N. Mostafa, A. Galal, and H. Elawady, “A Proposed Approach for Selecting Third Party Logistic Alternatives,” *Egypt. J. Eng. Sci. Technol.*, vol. 35, no. 2, pp. 48–56, 2021, doi: 10.21608/eijest.2021.48088.1023.

[4] B. Y. Wedha, “Enterprise Architecture untuk Industri Truk Logistik di Indonesia,” *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 9, no. 2, pp. 1137–1150, 2022, doi: 10.35957/jatisi.v9i2.1255.

[5] M. K. Hidayad and D. N. Utama, “Third-Party Logistic Selection for Logistic Aggregator Company using Multi-Criteria Decision Making,” *J. Comput. Sci.*, vol. 18, no. 9, pp. 811–820, 2022, doi: 10.3844/jcssp.2022.811.820.

[6] L. Filipova, “Business intelligence systems: modern development trends,” *Sci. J. “Library Sci. Rec. Stud. Informology,”* vol. 002, no. 1, 2022, doi: 10.32461/2409-9805.1.2022.257258.

[7] Z. Sun, “Big Data Analytics Thinking and Big Data Analytics Intelligence,” vol. 5, no. September, pp. 1–11, 2020, doi: 10.13140/RG.2.2.15678.31041.

[8] M. D. Ackermann, J. A. van der Poll, and H. M. van der Poll, “Re-evaluating the Definition of Intelligence in Business Intelligence,” *GATR J. Manag. Mark. Rev.*, vol. 1, no. 1, pp. 33–44, 2016, doi: 10.35609/jmmr.2016.1.1(5).

[9] M. Mach-kr, “applied sciences Conceptual Framework for Implementing Temporal Big Data Analytics in Companies,” 2022.

[10] Z. Sun and Y. Huo, “The Spectrum of Big Data Analytics,” *J. Comput. Inf. Syst.*, vol. 61, no. 2, pp. 154–162, 2021, doi: 10.1080/08874417.2019.1571456.

[11] R. Berman and A. Israeli, “The Value of Descriptive Analytics: Evidence from Online Retailers,” *Mark. Sci.*, no. April, 2022, doi: 10.1287/mksc.2022.1352.

[12] Y. Singh, S. Singh, and D. Kikan, “Literature Review of Predictive Analytics Adoption in Business Decisions,” no. October, 2022.