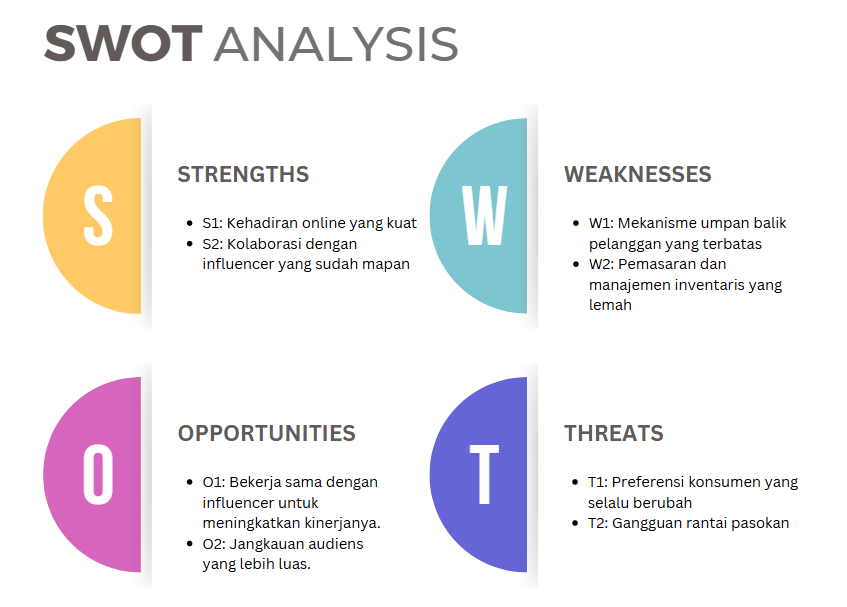
**Dataset :** [**https://www.kaggle.com/datasets/a23bisola/fashion-dataset-uk-us**](https://www.kaggle.com/datasets/a23bisola/fashion-dataset-uk-us)

**Link SAS :** [**https://vfl-071.engage.sas.com/links/resources/report?uri=%2Freports%2Freports%2Ff359b20f-71d3-4e54-9360-c20f70bd9828**](https://vfl-071.engage.sas.com/links/resources/report?uri=%2Freports%2Freports%2Ff359b20f-71d3-4e54-9360-c20f70bd9828)

**QUESTION 1**

**Clothing / StyleMode Fashion Company**

**SWOT :**

****

**Swot i-matrix :**

|  | Opportunities (O) | Threats (T) |
| --- | --- | --- |
| Strengths (S) | **S1O1:** Manfaatkan kehadiran online yang kuat untuk menjangkau audiens yang lebih luas.  **S2O2:** Bekerja sama dengan influencer untuk meningkatkan visibilitas merek. | **S1T1:** Gunakan analitik data untuk beradaptasi dengan cepat terhadap perubahan preferensi konsumen.  **S2T2:** Terapkan strategi manajemen risiko untuk mengurangi gangguan rantai pasokan. |
| Weaknesses (W) | **W1O1:** Tingkatkan mekanisme umpan balik pelanggan untuk lebih memahami preferensi mereka.  **W2O2:** Tingkatkan strategi pemasaran untuk bersaing dengan merek-merek yang sudah mapan. | **W1T1:** Atasi ketergantungan pada mode cepat dengan mempromosikan praktik berkelanjutan.  **W2T2:** Perkuat manajemen inventaris untuk mencegah kelebihan stok atau kekurangan stok. |

Perusahaan ini memiliki kehadiran online yang kuat, yang memungkinkannya menjangkau audiens yang luas dan beragam. Platform digitalnya sudah mapan, sehingga mudah untuk berinteraksi dengan pelanggan dan memberikan pengalaman online yang mulus. Keunggulan ini semakin didukung oleh kolaborasi influencer yang mapan, di mana kemitraan dengan influencer fesyen utama meningkatkan visibilitas perusahaan. Kolaborasi ini tidak hanya mendorong kesadaran merek tetapi juga menumbuhkan kredibilitas, membantu melibatkan dan menarik pelanggan di pasar yang sangat kompetitif. Namun, perusahaan menghadapi beberapa kelemahan. Mekanisme umpan balik pelanggan yang terbatas membuatnya sulit untuk mengumpulkan dan menganalisis wawasan yang berharga tentang preferensi pelanggan. Kurangnya umpan balik langsung dari konsumen ini menghambat perusahaan untuk beradaptasi dengan cepat untuk memenuhi kebutuhan mereka. Selain itu, sistem pemasaran dan manajemen inventaris yang lemah menambah tantangannya. Perusahaan berjuang untuk mengoptimalkan upaya pemasarannya, tertinggal dari para pesaing yang memiliki pesan merek yang lebih kuat. Selain itu, ketidakefisienan dalam manajemen inventaris meningkatkan risiko kelebihan stok atau kehabisan stok, yang dapat menyebabkan hilangnya penjualan dan ketidakpuasan pelanggan.

Terlepas dari tantangan-tantangan ini, perusahaan memiliki peluang yang signifikan untuk berkembang. Dengan memanfaatkan kehadiran online yang kuat, perusahaan dapat menjangkau audiens yang lebih luas dan meningkatkan strategi pemasaran digitalnya untuk menjaring lebih banyak pelanggan. Selain itu, berkolaborasi dengan influencer memberikan peluang untuk terus meningkatkan visibilitas merek dan meningkatkan penjualan. Dengan strategi yang tepat, pemasaran yang digerakkan oleh influencer dapat memperkuat jangkauan perusahaan dan menarik segmen pelanggan baru.

Namun, perusahaan juga harus memperhatikan ancaman eksternal. Perubahan preferensi konsumen dalam industri fesyen dapat terjadi dengan cepat dan tidak terduga. Jika perusahaan tidak dapat dengan cepat beradaptasi dengan tren ini, perusahaan berisiko kehilangan relevansi di pasar. Selain itu, gangguan rantai pasokan juga menimbulkan risiko yang signifikan. Masalah seperti kekurangan stok atau keterlambatan logistik dapat berdampak negatif pada ketersediaan produk, yang menyebabkan ketidakpuasan pelanggan dan hilangnya penjualan. Agar tetap kompetitif, perusahaan harus siap untuk mengelola potensi gangguan ini secara proaktif. Secara keseluruhan, kekuatan perusahaan memberikan fondasi yang kuat untuk pertumbuhan, tetapi mengatasi kelemahan dan mempersiapkan diri untuk menghadapi ancaman eksternal akan sangat penting untuk kesuksesan jangka panjang.

i-Matriks SWOT dapat menghubungkan kekuatan, kelemahan, peluang, dan ancaman tertentu ke dalam strategi yang dapat ditindaklanjuti yang dapat diterapkan oleh perusahaan untuk meningkatkan kinerjanya.

* S1O1: Memanfaatkan kehadiran online yang kuat untuk menjangkau audiens yang lebih luas

Kehadiran online perusahaan yang kuat (S1) dapat digunakan lebih lanjut untuk memperluas jangkauan audiensnya (O1). Dengan berinvestasi pada pemasaran digital yang lebih bertarget dan konten online yang menarik, perusahaan dapat menarik pelanggan baru, menjajaki pasar baru, dan meningkatkan penjualan. Memanfaatkan kekuatan ini selaras dengan peluang untuk memanfaatkan basis pelanggan yang lebih luas melalui platform online.

* S2O2: Berkolaborasi dengan influencer untuk meningkatkan visibilitas merek

Dengan membangun kolaborasi influencer yang sudah ada (S2), perusahaan dapat meningkatkan visibilitas merek (O2). Berkolaborasi dengan influencer populer dapat memperkenalkan merek kepada audiens baru, menciptakan lebih banyak buzz seputar produknya, dan meningkatkan penjualan melalui kampanye pemasaran yang digerakkan oleh influencer. Strategi ini memanfaatkan peluang untuk menumbuhkan kesadaran dan loyalitas merek melalui influencer tepercaya di industri fesyen.

* S1T1: Gunakan analisis data untuk beradaptasi dengan cepat terhadap perubahan preferensi konsumen

Kehadiran online perusahaan yang kuat (S1) menyediakan platform yang sangat baik untuk mengumpulkan data dan wawasan pelanggan. Dengan menggunakan alat analisis data, perusahaan dapat memantau dan merespons dengan cepat terhadap perubahan preferensi konsumen (T1). Hal ini memungkinkan perusahaan untuk tetap kompetitif dalam industri fesyen yang bergerak cepat dengan menyesuaikan penawarannya, menyesuaikan harga, atau menciptakan strategi pemasaran yang ditargetkan berdasarkan data real-time.

* S2T2: Menerapkan strategi manajemen risiko untuk memitigasi gangguan rantai pasokan

Dengan kolaborasi influencer yang telah terjalin (S2), perusahaan dapat mengurangi risiko yang terkait dengan gangguan rantai pasokan (T2). Dengan menjaga hubungan yang kuat dengan para influencer dan pelanggan utama, perusahaan dapat meminimalkan dampak gangguan ini terhadap penjualan. Selain itu, menerapkan strategi manajemen risiko, seperti diversifikasi pemasok atau mengadopsi alat pelacakan inventaris, akan membantu memitigasi risiko dan memastikan ketersediaan produk tetap stabil.

* W1O1: Meningkatkan mekanisme umpan balik pelanggan untuk lebih memahami preferensi

Mekanisme umpan balik pelanggan yang terbatas (W1) menjadi tantangan tersendiri bagi perusahaan, namun ada peluang untuk meningkatkan sistem ini (O1). Dengan membangun saluran yang lebih efektif untuk mengumpulkan umpan balik pelanggan (misalnya, survei, ulasan, atau mendengarkan media sosial), perusahaan dapat memperoleh wawasan yang lebih baik tentang preferensi pelanggan. Hal ini akan memungkinkan perusahaan untuk membuat keputusan berdasarkan data untuk memenuhi permintaan pelanggan dan meningkatkan penawaran produk.

* W2O2: Meningkatkan strategi pemasaran untuk bersaing dengan merek-merek yang sudah mapan

Upaya pemasaran perusahaan yang lemah (W2) dapat diatasi dengan meningkatkan strategi pemasaran (O2) untuk bersaing dengan merek-merek yang sudah mapan. Dengan mengadopsi pendekatan pemasaran yang lebih inovatif, seperti pemasaran influencer, kampanye media sosial, atau pemasaran email, perusahaan dapat memperkuat posisi merek dan menarik pelanggan baru, sehingga dapat bersaing dengan para pemimpin industri.

* W1T1: Mengatasi ketergantungan pada fesyen cepat saji dengan mempromosikan praktik-praktik berkelanjutan

Ketergantungan perusahaan pada fast fashion (W1) merupakan kelemahan potensial di pasar yang semakin menuntut produk berkelanjutan. Perusahaan dapat mengubahnya menjadi peluang dengan mempromosikan praktik-praktik berkelanjutan (T1), seperti bahan yang ramah lingkungan atau kemitraan perdagangan yang adil. Dengan melakukan hal tersebut, perusahaan tidak hanya memitigasi risiko yang terkait dengan perubahan preferensi konsumen terhadap keberlanjutan, tetapi juga memposisikan diri sebagai merek yang bertanggung jawab secara sosial.

* W2T2: Memperkuat manajemen inventaris untuk mencegah kelebihan stok atau kehabisan stok

Kelemahan dalam manajemen persediaan (W2) membuat perusahaan terekspos pada risiko kelebihan persediaan atau kehabisan persediaan (T2), yang dapat menyebabkan hilangnya penjualan atau sumber daya yang terbuang. Memperkuat manajemen inventarisnya melalui peramalan yang lebih baik, perencanaan permintaan, dan sistem kontrol inventaris akan membantu perusahaan mengelola tingkat stok secara efektif, memastikan perusahaan dapat memenuhi permintaan pelanggan tanpa kelebihan persediaan.

**Rencana Strategis :**

1. Memperluas Kehadiran Online: Tingkatkan strategi pemasaran digital untuk menarik audiens yang lebih luas dan berinteraksi dengan konsumen melalui media sosial dan kemitraan dengan influencer.

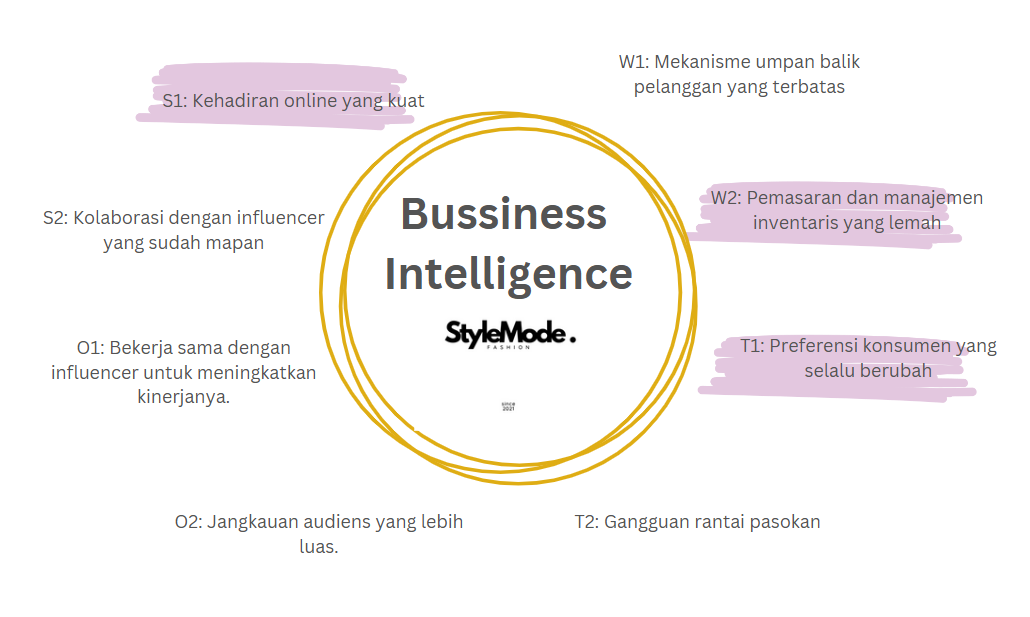
2. Menggunakan Bahan Ramah Lingkungan: Transisi ke bahan dan praktik yang berkelanjutan untuk memenuhi permintaan konsumen akan fesyen yang bertanggung jawab terhadap lingkungan.

3. Menerapkan Analisis Pelanggan: Gunakan analisis data untuk mengumpulkan wawasan tentang preferensi dan perilaku pelanggan, sehingga memungkinkan manajemen inventaris dan perkiraan tren yang lebih baik.

4. Diversifikasi Pemasok: Bangun rantai pasokan yang lebih tangguh dengan mendiversifikasi pemasok dan membuat rencana kontingensi untuk mengatasi potensi gangguan.

5. Meningkatkan Mekanisme Umpan Balik: Kembangkan saluran yang kuat untuk umpan balik pelanggan untuk terus meningkatkan penawaran produk dan layanan pelanggan. Rencana strategis ini bertujuan untuk memposisikan StyleMode Fashion sebagai pemimpin dalam industri fast fashion dengan memanfaatkan kekuatan dan peluang sekaligus mengatasi kelemahan dan ancaman secara efektif.

**Penggunaan Intelijen Bisnis pada StyleMode**



Pembuatan dasbor intelijen bisnis dari StyleMode ini akan berfokus pada poin-poin yang digaris bawahi, akan tetapi, poin-poin yang lain akan tetap dibahas.

* Kehadiran Online yang Kuat (S1):

StyleMode memiliki platform online yang solid, memungkinkannya menjangkau konsumen di Inggris dan Amerika Serikat. Kehadiran digital ini memudahkan pelanggan untuk mengakses dan membeli produk mereka dengan cepat. Di tengah meningkatnya belanja online, ini adalah kekuatan besar yang bisa mendukung pertumbuhan pasar dan mempertahankan pangsa pasar.

* Kolaborasi dengan Influencer yang Sudah Mapan (S2):

Kerja sama dengan influencer terkenal di industri fashion memungkinkan StyleMode untuk meningkatkan visibilitas dan menarik segmen pelanggan yang lebih muda dan sadar tren. Kolaborasi ini memperkuat kredibilitas merek di mata konsumen yang lebih digital-savvy dan menciptakan buzz positif di media sosial, yang pada gilirannya dapat meningkatkan penjualan.

* Mekanisme Umpan Balik Pelanggan yang Terbatas (W1):

Meskipun StyleMode fokus pada kepuasan pelanggan, terdapat kelemahan dalam mekanisme pengumpulan dan analisis umpan balik. Umpan balik yang kurang optimal dapat membuat perusahaan kurang peka terhadap perubahan kebutuhan konsumen dan memperlambat respon terhadap tren mode baru. Memperbaiki saluran komunikasi pelanggan dan mempercepat analisis data ini adalah langkah penting untuk tetap kompetitif.

* Pemasaran dan Manajemen Inventaris yang Lemah (W2):

Ada masalah dalam manajemen inventaris dan strategi pemasaran yang kurang efektif. Hal ini membuat StyleMode rentan terhadap fluktuasi dalam permintaan dan persediaan produk yang tidak tepat. Inventaris yang tidak optimal dapat mengakibatkan kelebihan stok atau kekurangan stok, yang berpengaruh pada biaya dan pengalaman pelanggan.

* Bekerja Sama dengan Influencer untuk Meningkatkan Kinerjanya (O1):

Dengan kolaborasi yang sudah berjalan dengan influencer mapan, ada peluang besar bagi StyleMode untuk memperluas kemitraan ini. Meningkatkan kolaborasi dengan influencer dari niche berbeda atau memperluas kolaborasi ke platform digital lain dapat membuka lebih banyak peluang pertumbuhan, menarik audiens baru, dan memperluas cakupan pasar.

* Jangkauan Audiens yang Lebih Luas (O2):

Dengan kehadiran online yang sudah kuat, StyleMode memiliki peluang untuk memperluas jangkauannya lebih jauh ke pasar internasional lainnya. Selain itu, kampanye yang lebih ditargetkan dan penggunaan data yang lebih baik bisa meningkatkan konversi di segmen pasar yang belum terjangkau, terutama di kalangan milenial dan Gen Z yang mendominasi industri fashion.

* Preferensi Konsumen yang Selalu Berubah (T1):

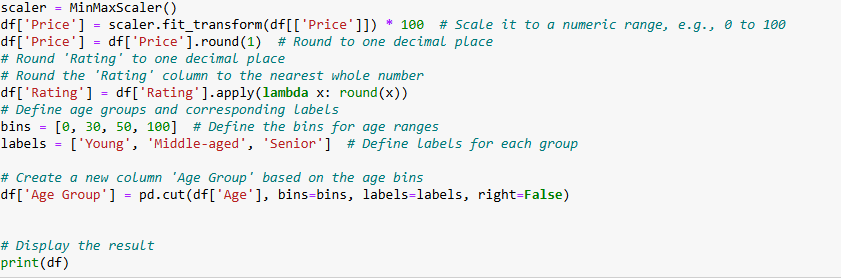
Perubahan tren fashion yang cepat adalah tantangan besar bagi StyleMode. Mode cepat membutuhkan responsivitas yang tinggi dalam merespons preferensi konsumen yang terus berkembang. Jika StyleMode tidak dapat mengikuti perubahan tren ini dengan cepat, mereka berisiko kehilangan pangsa pasar ke kompetitor yang lebih adaptif.

* Gangguan Rantai Pasokan (T2):

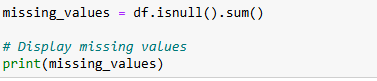
Ketidakpastian dalam rantai pasokan, baik itu disebabkan oleh pandemi, keterbatasan sumber daya, atau isu global lainnya, bisa mengganggu kemampuan StyleMode untuk mempertahankan stok yang cukup. Gangguan ini berpotensi menyebabkan keterlambatan pengiriman, biaya tambahan, atau kekurangan stok yang merusak reputasi merek di mata konsumen.

**QUESTION 2**

**PREPROCESSING**

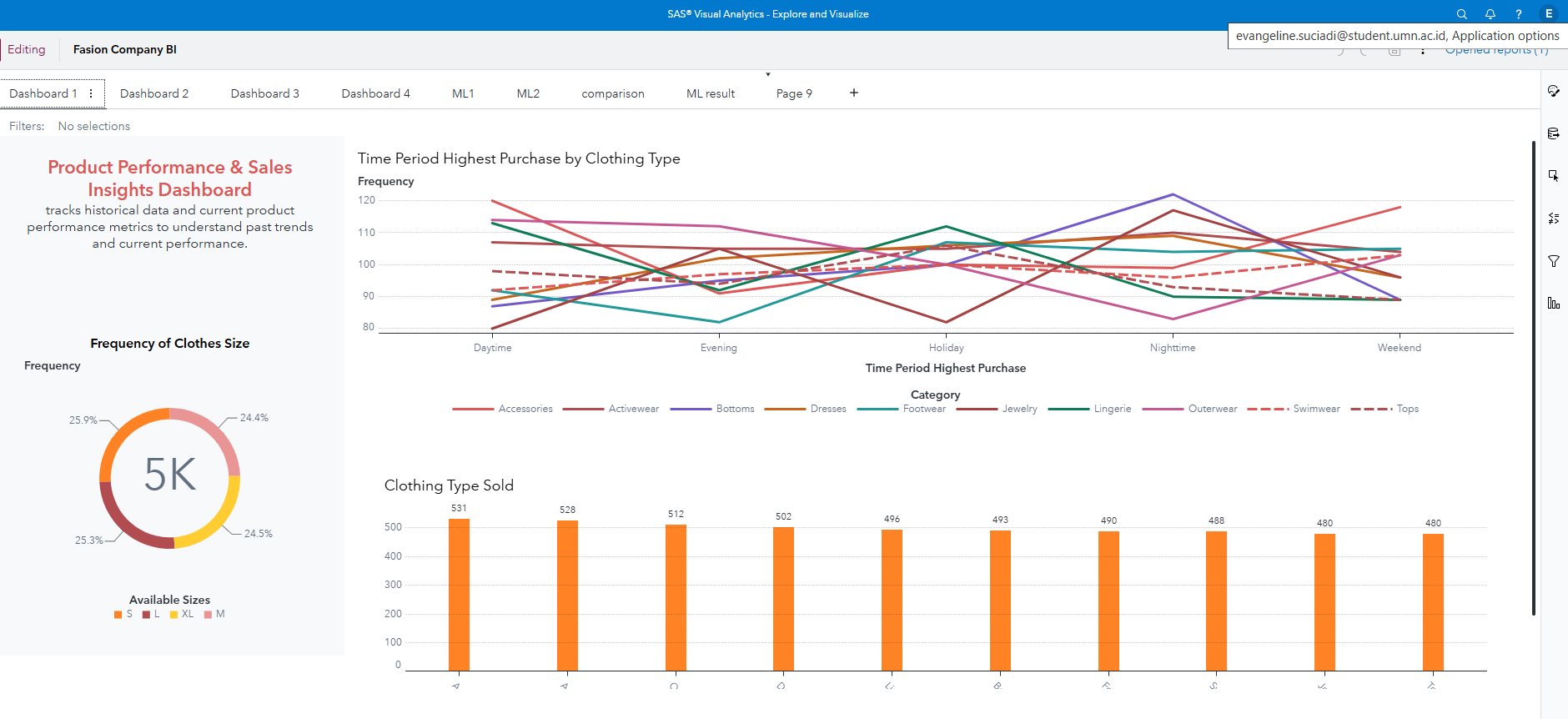
****

Pembulatan price, rating agar lebih rapi dilakukan, column baru pengelompokan umur juga dibuat agar analisa lebih mudah.

****

Pengecekan missing values juga dilakukan , dipastikan tidak ada data yang missing value

**i) Product Performance & Sales Insights Dashboard**



Menganalisis Hasil:

Dasbor Wawasan Kinerja & Penjualan Produk memberikan gambaran umum yang komprehensif tentang penjualan produk dan metrik kinerja, memungkinkan pemahaman yang lebih baik tentang tren historis dan perilaku pasar saat ini. Dasbor menyoroti wawasan utama di berbagai kategori produk, periode waktu pembelian, dan preferensi ukuran pakaian. **Bagan Donat** menunjukkan Ukuran Pakaian yang terdistribusi secara merata

Pada bagian Jenis Pakaian yang Terjual, **diagram batang** menggambarkan bahwa Pakaian Olahraga adalah kategori produk yang paling populer, mencapai 10,62% dari total penjualan (531 unit dari 5.000). Diikuti oleh Aksesoris, yang menyumbang 10,56% (528 unit), dan Pakaian Luar, dengan 10,24% (512 unit). Ketiga kategori ini mendominasi distribusi penjualan, didorong oleh faktor-faktor seperti Jumlah Ulasan, Usia, dan Riwayat Pembelian. Kategori lainnya, termasuk Gaun, Lingerie, Bawahan, Alas Kaki, Pakaian Renang, Perhiasan, dan Atasan, semuanya memiliki kinerja yang sama, masing-masing terjual antara 480 hingga 502 unit. Bagian Frekuensi Ukuran Pakaian memberikan wawasan tentang preferensi ukuran di antara para pelanggan. Ukuran S (kecil) adalah yang paling umum, mewakili 25,86% dari total pembelian, yang setara dengan 1.300 unit. Ukuran L dan XL menyusul di belakangnya, masing-masing sebesar 25,28% dan 24,46%. Grafik ini menunjukkan bahwa ukuran yang lebih kecil sedikit lebih populer di kalangan konsumen, dengan faktor-faktor yang mempengaruhi seperti Usia, Jumlah Ulasan, dan Harga yang membentuk preferensi ukuran ini.

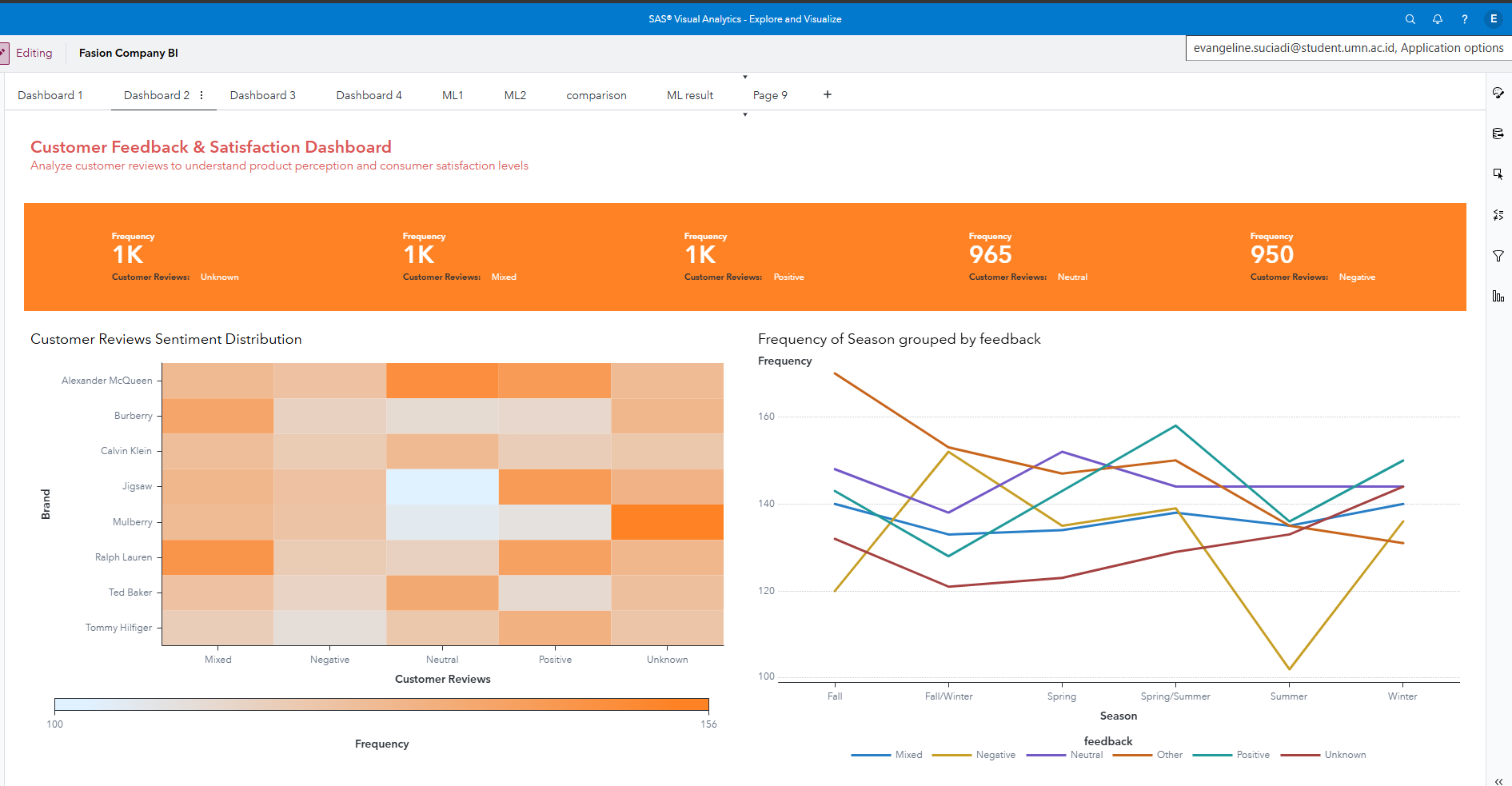
Terakhir, **grafik garis** Periode Waktu Pembelian Tertinggi menunjukkan bahwa Malam Hari adalah waktu puncak pembelian, menyumbang 20,46% dari penjualan, diikuti oleh periode Liburan sebesar 20,36% dan pembelian Akhir Pekan sebesar 19,84%. Tren ini menunjukkan bahwa pelanggan cenderung berbelanja lebih banyak pada jam-jam di luar jam kerja, terutama pada malam hari dan selama hari libur. Faktor yang paling berpengaruh terhadap perilaku ini adalah Harga, Jumlah Ulasan, dan Atribut Gaya. Grafik tersebut juga menampilkan tren spesifik kategori, seperti Bawahan yang mengalami lonjakan penjualan pada malam hari, sementara Alas Kaki lebih populer selama liburan.

Singkatnya, dasbor menyoroti faktor-faktor penting yang mendorong penjualan produk. Ukuran yang lebih kecil, terutama Ukuran S, memiliki permintaan yang lebih tinggi, dengan puncak penjualan pada malam hari dan periode liburan. Kategori seperti pakaian olahraga dan aksesori memimpin dalam hal volume yang terjual, yang menunjukkan minat konsumen yang kuat pada segmen ini.

Dasbor ini berhubung dengan beberapa SWOT :

* S1: Kehadiran online yang kuat – Dasbor menunjukkan bahwa kategori seperti Pakaian Olahraga dan Aksesori sangat populer, yang dapat menunjukkan dampak positif dari kehadiran online yang kuat dalam meningkatkan penjualan.
* W2: Pemasaran dan manajemen inventaris yang lemah – Beberapa kategori produk, seperti Gaun dan Pakaian Renang, memiliki penjualan yang lebih rendah. Ini bisa mengindikasikan kelemahan dalam pemasaran atau manajemen inventaris di kategori tersebut.
* O2: Jangkauan audiens yang lebih luas – Dasbor menunjukkan variasi preferensi ukuran pakaian, dengan ukuran S sebagai yang paling umum, tetapi ukuran L dan XL juga signifikan. Ini menciptakan peluang untuk memperluas jangkauan pasar dengan memenuhi permintaan berbagai ukuran.
* T2: Gangguan rantai pasokan – Informasi tentang penjualan puncak pada periode Malam Hari dan Liburan mengindikasikan bahwa gangguan rantai pasokan dapat menjadi ancaman, terutama selama masa penjualan yang sibuk. Jika rantai pasokan terganggu pada saat permintaan tinggi, hal ini dapat berdampak negatif pada ketersediaan produk dan kepuasan pelanggan.

**ii) Customer Feedback and Satisfaction Dashboard**



Menganalisis Hasil:

Berdasarkan **kartu KPI**, Tidak Diketahui adalah nilai yang paling umum yang mewakili 20,80% (1K dari 5.000) Ulasan Pelanggan. Nilai umum lainnya adalah Campuran (sebesar 20,72%) dan Positif (sebesar 20,18%). Tiga faktor yang paling terkait adalah Usia, Atribut Gaya, dan Jumlah Ulasan.

**Peta panas** ini memvisualisasikan distribusi sentimen pelanggan di berbagai merek pakaian mewah, memberikan wawasan tentang bagaimana pelanggan memandang setiap merek berdasarkan ulasan mereka. Salah satu pengamatan yang mencolok adalah dominasi kategori sentimen “Tidak Diketahui”, yang menyumbang 20,80% dari semua ulasan pelanggan. Hal ini menunjukkan bahwa sebagian besar umpan balik tidak memiliki klasifikasi sentimen yang jelas, mungkin karena data yang tidak lengkap atau ambigu. Merek-merek seperti Tommy Hilfiger dan Mulberry memiliki konsentrasi ulasan “Tidak Diketahui” yang lebih tinggi, dengan Mulberry yang menonjol dalam kategori ini.

Selain “Tidak Diketahui”, **peta panas** menunjukkan bahwa ulasan “Campuran” (20,72%) dan “Positif” (20,18%) juga cukup umum. Khususnya, Ralph Lauren dan Alexander McQueen memiliki distribusi yang seimbang antara sentimen campuran dan positif. Merek-merek ini menunjukkan persepsi pelanggan yang bervariasi, tetapi umumnya condong ke arah kesan yang baik. Ralph Lauren, khususnya, memiliki konsentrasi ulasan campuran yang menonjol, yang dapat mengindikasikan tingkat ambivalensi atau opini pelanggan yang terbagi tentang penawaran merek tersebut Berfokus pada merek itu sendiri, Alexander McQueen memimpin dalam umpan balik pelanggan, menyumbang 13,42% dari total ulasan, diikuti oleh Ralph Lauren sebesar 13,00%. Bersama-sama, kedua merek ini menghasilkan lebih dari seperempat ulasan, menjadikan mereka pemain kunci di pasar. Kedua merek tersebut memiliki ulasan yang tersebar di berbagai kategori sentimen, meskipun mereka menunjukkan frekuensi yang lebih tinggi di segmen “Campuran” dan “Positif”, yang mencerminkan keberadaan mereka yang kuat dan basis pelanggan yang beragam, tren spesifik merek juga muncul dari peta panas ini. Jigsaw menonjol dengan konsentrasi ulasan netral yang tinggi, yang mengindikasikan bahwa banyak pelanggan yang mungkin merasa acuh tak acuh terhadap merek tersebut. Di sisi lain, Mulberry dan Calvin Klein menunjukkan tingkat sentimen negatif yang lebih tinggi, yang menandakan potensi ketidakpuasan atau ekspektasi yang tidak terpenuhi dari pelanggan. Wawasan ini dapat membantu merek-merek ini untuk fokus meningkatkan kepuasan pelanggan.

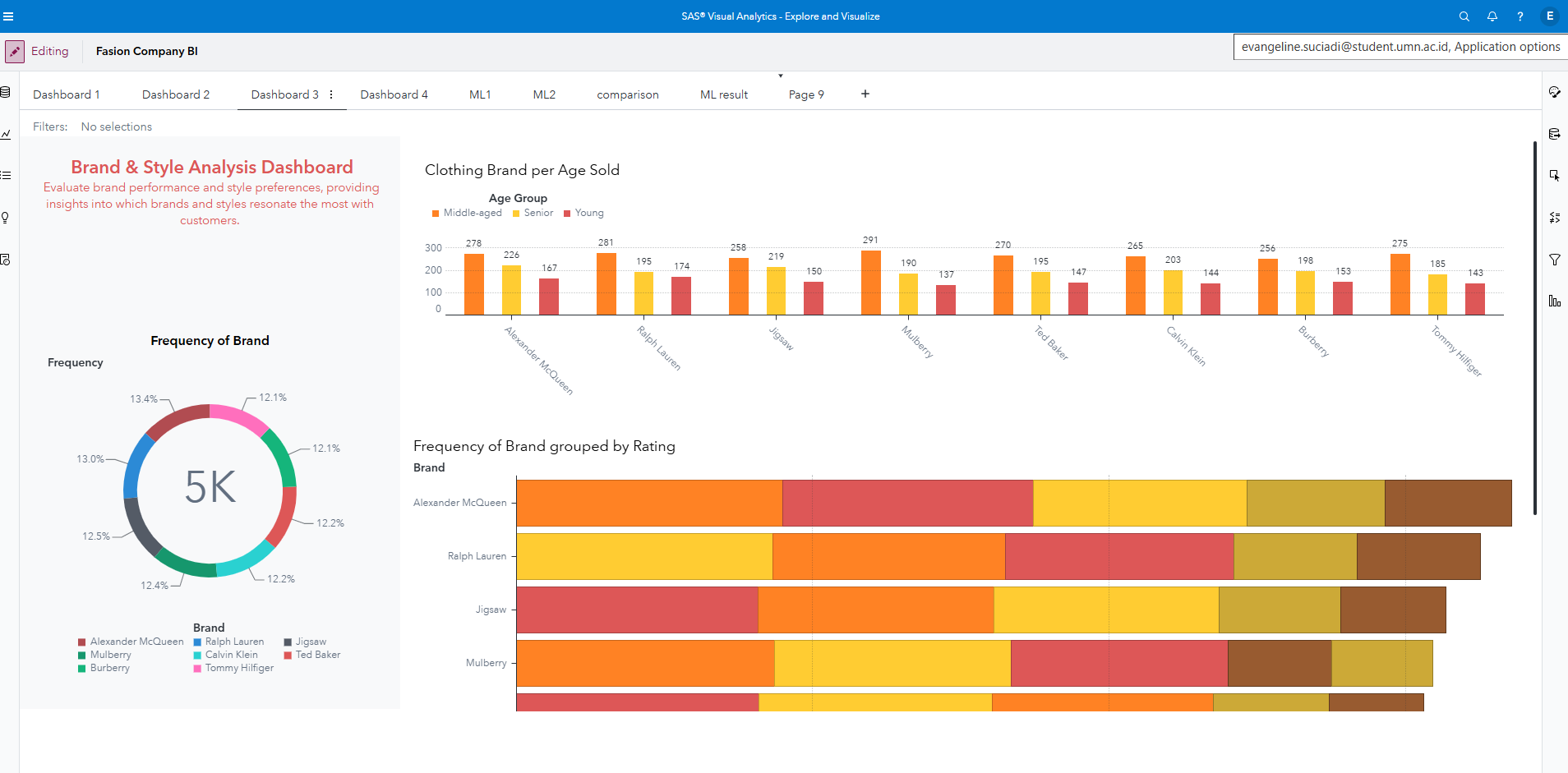
Terakhir, distribusi sentimen dalam **peta panas** ini tampaknya dipengaruhi oleh beberapa faktor, termasuk Usia, Atribut Gaya, dan Jumlah Ulasan. Demografi usia dan preferensi gaya dapat membentuk bagaimana pelanggan memandang merek yang berbeda. Selain itu, Fashion Influencer dan Harga juga dapat berperan dalam visibilitas dan volume ulasan yang tinggi untuk merek-merek seperti Alexander McQueen dan Ralph Lauren. Wawasan ini dapat memandu keputusan strategis, seperti kampanye pemasaran yang ditargetkan dan upaya pemosisian merek.

Grafik garis menggambarkan frekuensi umpan balik pelanggan di berbagai musim dan kategori sentimen yang berbeda, memberikan wawasan tentang bagaimana umpan balik bervariasi tergantung pada waktu dalam setahun, Pertama, Musim Semi/Musim Panas muncul sebagai musim yang paling umum untuk umpan balik pelanggan, mewakili 17,16% dari semua ulasan, diikuti oleh Musim Gugur sebesar 17,06% dan Musim Dingin sebesar 16,90%. Hal ini menunjukkan bahwa musim-musim ini menghasilkan volume ulasan tertinggi. Sebaliknya, Musim Panas jauh lebih jarang terjadi, hanya mewakili 15,70% dari umpan balik

Dasbor ini berhubung dengan beberapa SWOT :

* S1: Kehadiran online yang kuat – Merek seperti Alexander McQueen dan Ralph Lauren menerima banyak ulasan pelanggan, yang menunjukkan keberadaan online yang signifikan dan keterlibatan pelanggan yang tinggi.
* W1: Mekanisme umpan balik pelanggan yang terbatas – Dengan dominasi sentimen “Tidak Diketahui” yang mencapai 20,80%, ini menunjukkan bahwa sebagian besar umpan balik tidak jelas atau tidak terklasifikasi. Ini bisa menandakan kelemahan dalam mekanisme pengumpulan dan analisis umpan balik pelanggan.
* T1: Preferensi konsumen yang selalu berubah – Merek-merek seperti Mulberry dan Calvin Klein memiliki sentimen negatif yang lebih tinggi, menunjukkan risiko ketidakpuasan pelanggan dan kemungkinan perubahan preferensi konsumen. Jika tren ini berlanjut, merek-merek ini dapat kehilangan pelanggan.

**iii) Brand & Style Analysis Dashboard**



Menganalisis Hasil :

Dasbor memberikan wawasan berharga tentang hubungan antara merek pakaian, kelompok usia, dan peringkat pelanggan. Setiap bagan menyoroti pola spesifik yang membantu untuk memahami segmentasi pasar dan kepuasan pelanggan di berbagai merek pakaian.

**Bagan donat** menunjukkan bahwa Alexander McQueen adalah nilai yang paling umum yang mewakili 13,42% (671 dari 5.000) Merek. Nilai umum lainnya adalah Ralph Lauren (sebesar 13,00%) dan Jigsaw (sebesar 12,54%). Tiga faktor yang paling terkait adalah Harga, Usia, dan Fashion Influencer.

Diagram batang yang **dikelompokkan**, yang berfokus pada merek pakaian yang dijual berdasarkan kelompok usia, individu paruh baya mendominasi penjualan sebagian besar merek, membuat porsi terbesar di semua merek yang ditampilkan. Lansia juga mewakili pangsa penjualan yang signifikan, terutama untuk merek-merek seperti Ralph Lauren dan Calvin Klein. Di sisi lain, konsumen muda merupakan bagian terkecil dari penjualan, dengan Ralph Lauren dan Tommy Hilfiger menunjukkan distribusi yang relatif seimbang dibandingkan dengan merek lainnya. Penjualan tertinggi untuk individu paruh baya terlihat pada Alexander McQueen dan Mulberry, sementara Ralph Lauren dan Calvin Klein berkinerja baik di kalangan manula. Untuk pembeli yang lebih muda, Tommy Hilfiger dan Ralph Lauren adalah merek dengan daya tarik terbesar.

**Diagram batang bertumpuk**, yang menunjukkan frekuensi merek yang dikelompokkan berdasarkan peringkat, menunjukkan konsentrasi peringkat di rentang tengah, dengan 2, 3, dan 4 sebagai peringkat yang paling umum. Peringkat 2 memiliki frekuensi tertinggi, yang menunjukkan bahwa banyak pelanggan yang sedikit tidak puas dengan pengalaman merek mereka. Namun, peringkat 4, yang menandakan kepuasan umum, juga cukup sering muncul. Peringkat 3, yang mewakili netralitas, sedikit lebih jarang ditemukan namun tetap menonjol. Peringkat ekstrem (1 dan 5) lebih jarang terjadi, menunjukkan bahwa ketidakpuasan ekstrem dan kepuasan ekstrem jarang terjadi pada merek-merek ini. Pola spesifik merek menunjukkan bahwa Alexander McQueen memiliki jumlah peringkat netral hingga positif yang lebih tinggi, sementara Ralph Lauren dan Calvin Klein memiliki distribusi yang lebih luas, termasuk lebih banyak peringkat negatif. Namun, Burberry dan Tommy Hilfiger tampaknya memiliki umpan balik yang lebih positif secara keseluruhan, dengan jumlah peringkat 4 dan 5 yang signifikan.

Berdasarkan rincian tambahan yang diberikan, Alexander McQueen menonjol sebagai merek yang paling umum, mewakili lebih dari 13% dari total penjualan, diikuti oleh Ralph Lauren dan Jigsaw. Bersama-sama, merek-merek ini menyumbang lebih dari seperempat dari seluruh penjualan. Dalam hal kelompok usia, individu paruh baya merupakan mayoritas pembeli dengan persentase lebih dari 43%, diikuti oleh manula dengan persentase lebih dari 32%, dan pembeli muda sekitar 24%. Sedangkan untuk penilaian, skor yang paling umum adalah 2, yang menunjukkan tingkat ketidakpuasan pelanggan, meskipun diikuti oleh jumlah 4 yang tinggi, yang menunjukkan bahwa pendapat terpolarisasi. Peringkat ekstrem (1 dan 5) hanya berjumlah kurang dari 25% dari total, menunjukkan bahwa pelanggan pada umumnya tidak terlalu tidak puas atau sangat puas.

Secara keseluruhan, dasbor ini menyoroti bahwa pelanggan paruh baya membentuk pasar inti untuk sebagian besar merek pakaian ini, sementara kepuasan merek terbagi, dengan sebagian besar pelanggan merasa agak tidak puas atau cukup puas, tetapi lebih sedikit yang memberikan umpan balik ekstrem

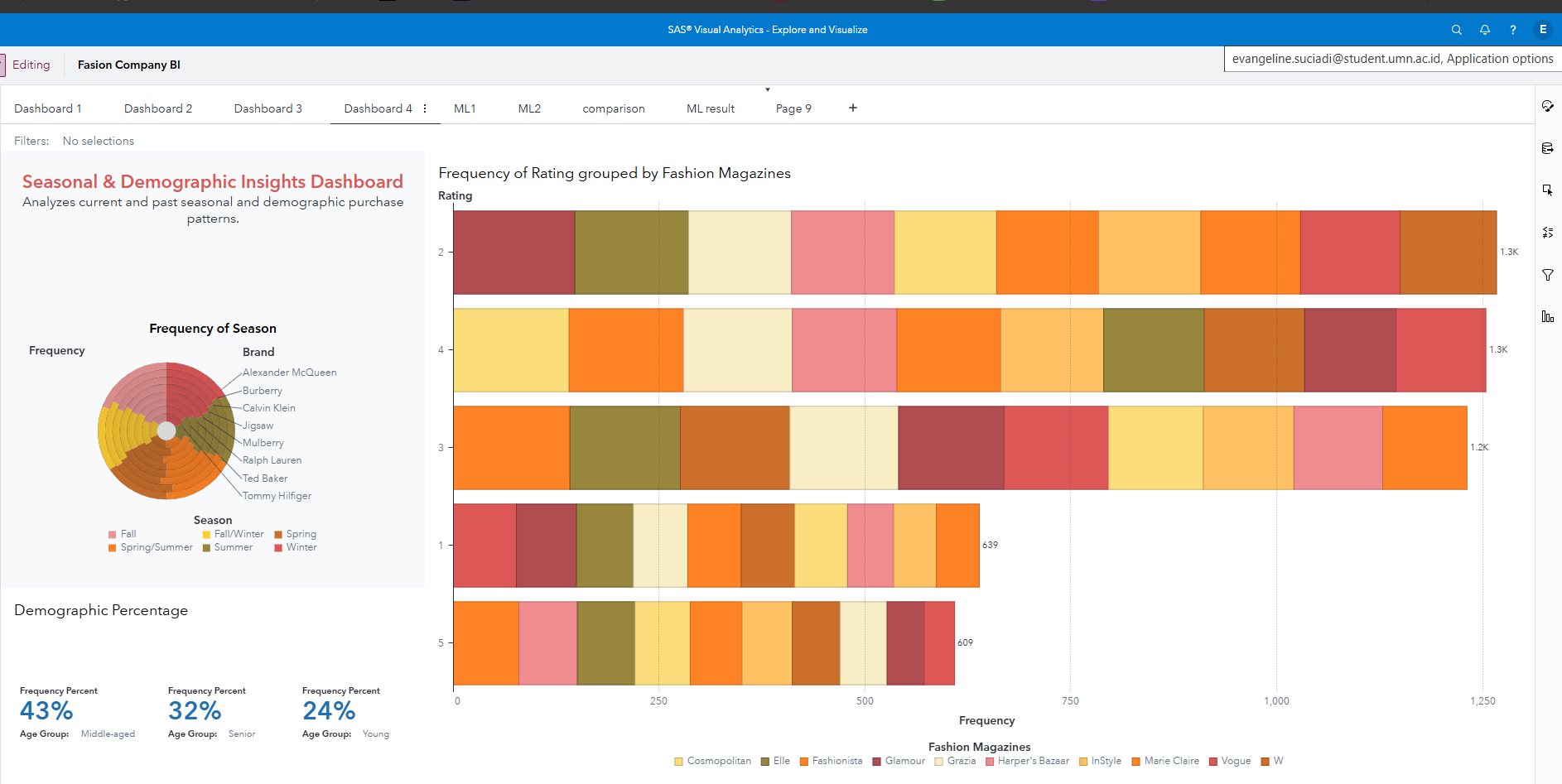
Dasbor ini terhubung dengan beberapa SWOT :

* W2: Pemasaran dan manajemen inventaris yang lemah – Merek seperti Ralph Lauren dan Calvin Klein memiliki distribusi peringkat yang luas, dengan banyak ulasan negatif atau netral, menunjukkan tantangan dalam strategi pemasaran dan potensi masalah kepuasan pelanggan.
* O2: Jangkauan audiens yang lebih luas – Dashboard menunjukkan bahwa merek seperti Tommy Hilfiger dan Ralph Lauren memiliki daya tarik yang lebih besar di kalangan konsumen muda, yang bisa dimanfaatkan lebih lanjut untuk memperluas basis pelanggan di segmen ini.

### 

* T1: Preferensi konsumen yang selalu berubah – Ketidakpuasan yang muncul dari peringkat 2 yang tinggi menunjukkan adanya ancaman terkait perubahan preferensi konsumen, terutama di kalangan pelanggan middle-age, yang mendominasi pembelian.

**iv) Seasonal & Demographic Insights**



Analisa Hasil

**Bagan radar** memvisualisasikan frekuensi musim di berbagai merek. Musim yang paling sering terjadi untuk pembelian adalah Musim Semi/Musim Panas, yang mewakili 17,16% penjualan. Di belakangnya ada Musim Gugur dengan 17,06% dan Musim Dingin dengan 16,90%. Ketiga musim ini menunjukkan distribusi yang cukup merata dalam hal pembelian merek. Sebaliknya, Musim Panas adalah yang paling jarang terjadi, dengan 15,70% pembelian, yang mengindikasikan lebih sedikit aktivitas belanja di musim panas dibandingkan dengan musim-musim lainnya. Tiga faktor teratas yang mempengaruhi preferensi musiman adalah usia, atribut gaya, dan harga. Alexander Mcqueen adalah merek yang paling banyak dibeli di setiap musim, diikuti oleh Burberry.

**Kartu KPI** dengan rincian persentase demografis menunjukkan bahwa konsumen paruh baya mendominasi pembelian, dengan persentase 43,48% dari total penjualan. Lansia mengikuti dengan 32,22%, sementara konsumen muda mewakili kelompok terkecil dengan 24,30%. Konsumen paruh baya merupakan segmen terbesar di sebagian besar merek dan musim. Tiga faktor yang paling erat kaitannya dengan rincian demografis ini adalah majalah fesyen, jumlah ulasan, dan harga. Hal ini menyiratkan bahwa pelanggan paruh baya dan senior lebih dipengaruhi oleh ulasan dan poin harga saat memilih merek pakaian, dibandingkan dengan konsumen yang lebih muda.

**Diagram batang bertumpuk** di sebelah kanan menunjukkan frekuensi peringkat di berbagai majalah mode. Peringkat yang paling umum adalah 2, mewakili 25,34% dari semua peringkat, diikuti oleh peringkat 4 (25,08%) dan peringkat 3 (24,62%). Bersama-sama, peringkat kelas menengah ini (2, 3, dan 4) merupakan mayoritas umpan balik pelanggan. Sebaliknya, peringkat ekstrem seperti 1 dan 5 jauh lebih jarang, dan hanya mencakup 24,96% dari total peringkat. Hal ini menunjukkan bahwa sebagian besar pelanggan memiliki pengalaman yang cukup positif atau cukup negatif dengan merek tempat mereka berbelanja, dengan lebih sedikit pelanggan yang melaporkan kepuasan atau ketidakpuasan yang ekstrem. Faktor yang paling signifikan yang terkait dengan peringkat ini adalah harga, kategori produk, dan jumlah ulasan.

Dalam hal majalah fesyen, Elle adalah yang paling berpengaruh, mewakili 10,66% dari semua penyebutan, diikuti oleh Grazia sebesar 10,24% dan Cosmopolitan sebesar 10,22%. Ketiga majalah ini memiliki pengaruh yang kuat terhadap preferensi konsumen, dengan Elle yang sangat menonjol. Sebaliknya, majalah seperti Harper's Bazaar, InStyle, dan Vogue jauh lebih jarang, dan secara bersama-sama mencapai 28,60% dari penyebutan majalah fesyen. Hal ini menunjukkan bahwa meskipun majalah-majalah tersebut masih berperan dalam memengaruhi pembelian fesyen, mereka tidak memiliki bobot sebanyak tiga majalah teratas.

Secara keseluruhan, dasbor menunjukkan bahwa preferensi musiman tersebar secara merata di sebagian besar musim, dengan konsumen paruh baya membentuk kelompok pelanggan terbesar. Peringkat pelanggan terkonsentrasi di sekitar pengalaman yang moderat, dan meskipun Elle, Grazia, dan Cosmopolitan memainkan peran utama dalam memengaruhi pembelian fesyen, majalah lain seperti Harper's Bazaar, InStyle, dan Vogue memiliki pengaruh yang lebih kecil terhadap keputusan konsumen. Hubungan antara majalah fesyen dan merek, ditambah dengan penilaian pelanggan, menyoroti pentingnya harga, ulasan, dan atribut gaya dalam membentuk perilaku pembelian di seluruh musim dan demografi.

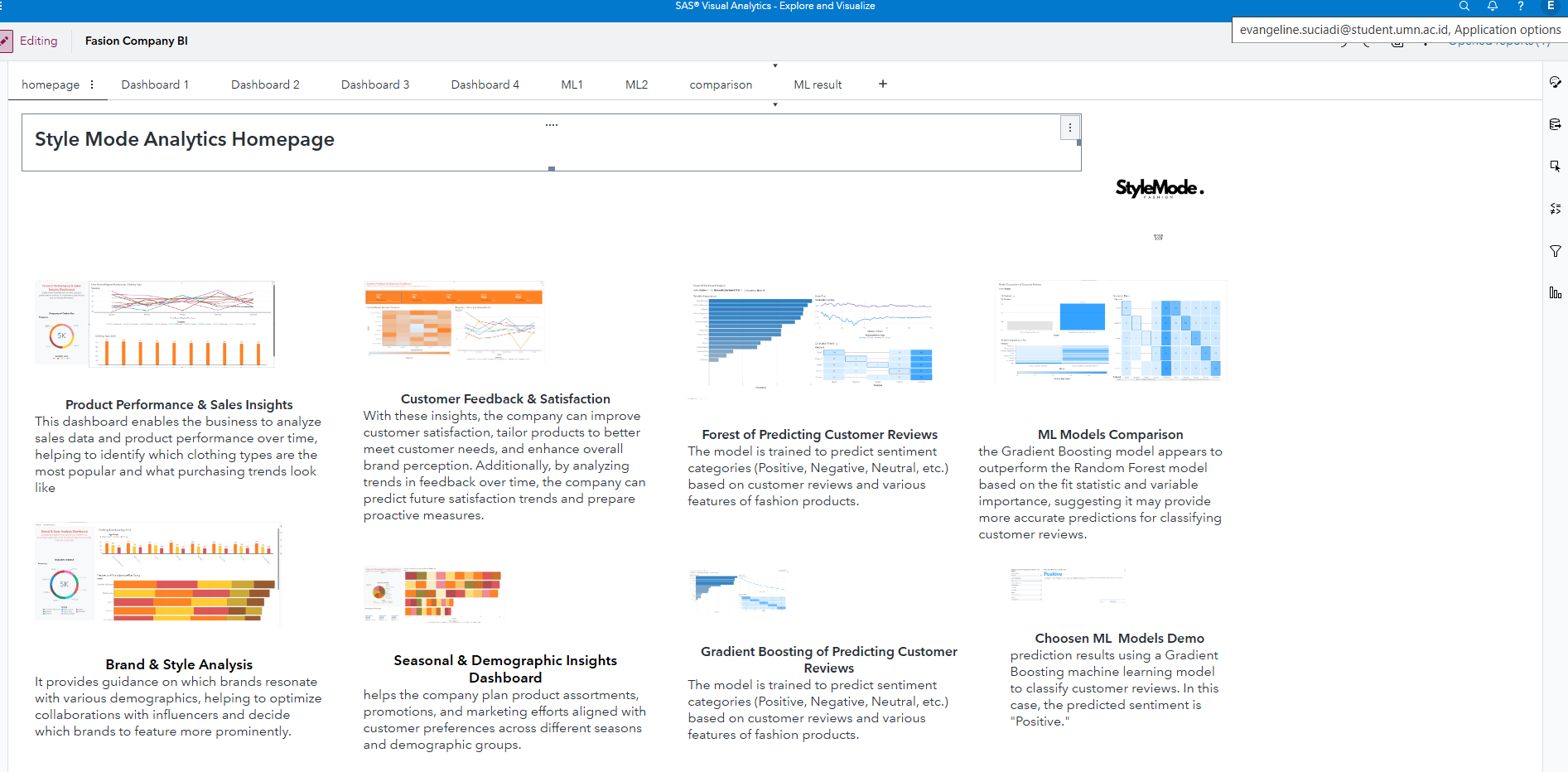
Beberapa SWOT yang dipakai pada dasbor ini :

### 

* S2: Kolaborasi dengan influencer yang sudah mapan – Majalah fesyen seperti Elle, Grazia, dan Cosmopolitan yang mempengaruhi pembelian di seluruh musim dapat dianggap sebagai influencer industri. Merek yang mendapatkan eksposur dari media ini menunjukkan potensi kekuatan dalam kolaborasi dan visibilitas di media yang berpengaruh.
* W1: Mekanisme umpan balik pelanggan yang terbatas – Distribusi peringkat yang berkisar di angka moderat (peringkat 2, 3, dan 4) menunjukkan bahwa banyak pelanggan memiliki pengalaman yang tidak terlalu memuaskan, tapi juga tidak sepenuhnya buruk. Hal ini menunjukkan perlunya perbaikan dalam menangkap dan menindaklanjuti umpan balik pelanggan untuk meningkatkan kepuasan mereka.
* O1: Bekerja sama dengan influencer untuk meningkatkan kinerja – Dengan majalah fesyen seperti Elle dan Grazia yang memiliki pengaruh kuat terhadap keputusan pembelian, ada peluang untuk lebih mengoptimalkan kolaborasi dengan influencer di sektor ini untuk menarik lebih banyak pembeli di segmen yang berbeda.
* T2: Gangguan rantai pasokan – Perubahan preferensi musiman yang menunjukkan penurunan pembelian di musim panas dan variasi dalam pembelian berdasarkan musim lainnya bisa menjadi tantangan bagi manajemen inventaris. Kegagalan dalam merespons permintaan musiman yang fluktuatif dapat menyebabkan kelebihan stok atau kekurangan stok di musim-musim tertentu.

**QUESTION 3**

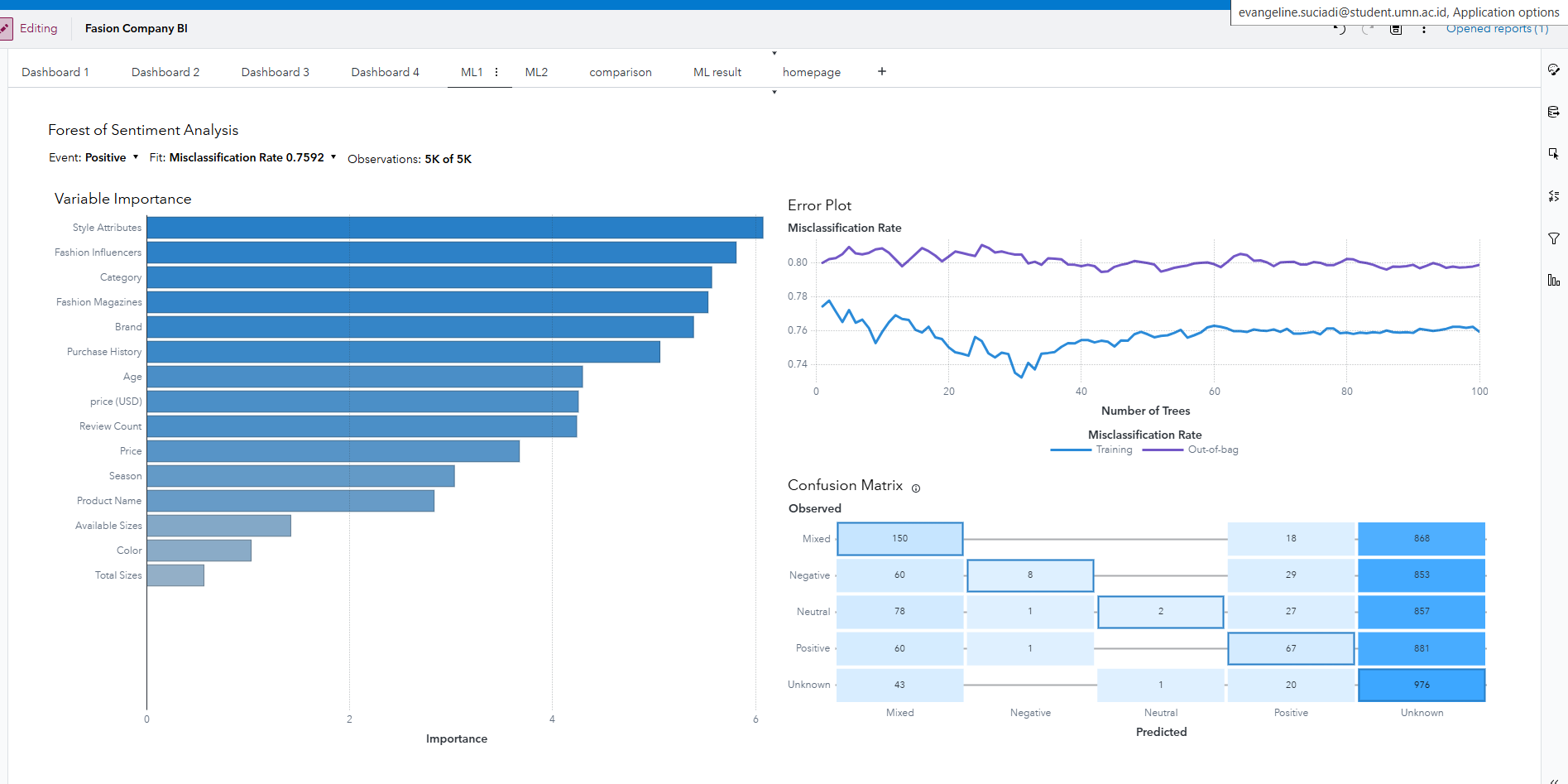
**Homepage**



Tujuan dari beranda “Style Mode Analytics” adalah untuk menyediakan seperangkat dasbor dan wawasan yang komprehensif untuk membantu bisnis dalam industri mode membuat keputusan berdasarkan data. Situs ini berfokus pada analisis kinerja produk dan tren penjualan, umpan balik dan kepuasan pelanggan, resonansi merek dengan demografi yang berbeda, dan preferensi produk musiman. Selain itu, laman ini menyoroti penggunaan model pembelajaran mesin, seperti Gradient Boosting dan Random Forest, untuk memprediksi sentimen pelanggan dari ulasan. Dengan membandingkan kinerja model-model ini, dasbor membantu mengidentifikasi pendekatan yang paling akurat untuk mengklasifikasikan umpan balik pelanggan, sehingga bisnis dapat mengoptimalkan produk, strategi pemasaran, dan kepuasan pelanggan secara keseluruhan.

**QUESTION 4**

**i) Forest Machine Learning Model**



Tujuan model ini adalah untuk memprediksi sentimen pelanggan berdasarkan atribut dan ulasan produk. Model ini menggunakan algoritme hutan acak, yang merupakan metode ansambel yang terdiri dari banyak pohon keputusan. Prediksi akhir ditentukan dengan menggabungkan suara dari pohon-pohon ini, memastikan prediksi yang lebih kuat dibandingkan dengan menggunakan pohon keputusan tunggal. Dataset terdiri dari 5.000 pengamatan, yang sesuai dengan 5.000 ulasan pelanggan.

Tingkat kesalahan klasifikasi adalah 0,7592, yang berarti bahwa sekitar 76% dari prediksi yang dibuat oleh model salah. Tingkat kesalahan yang relatif tinggi ini menunjukkan bahwa model ini berjuang untuk mengklasifikasikan sentimen pelanggan secara akurat dan menyarankan ruang untuk perbaikan, baik melalui penyempurnaan model atau rekayasa fitur.

Model ini memberikan peringkat pentingnya berbagai fitur (atau prediktor) yang digunakan untuk mengklasifikasikan sentimen. Atribut Gaya adalah fitur yang paling signifikan dalam memprediksi sentimen, yang menunjukkan bahwa bagaimana gaya suatu produk dirasakan memainkan peran penting dalam umpan balik pelanggan. Fitur penting lainnya termasuk Fashion Influencer dan Majalah Mode, yang mencerminkan dampak dari opini eksternal terhadap sentimen pelanggan. Kategori dan Merek juga memainkan peran penting, yang menunjukkan bahwa beberapa kategori atau merek membangkitkan respons emosional yang lebih kuat dari pelanggan. Fitur tambahan, seperti Riwayat Pembelian, Usia, Harga, dan Jumlah Ulasan, berkontribusi pada tingkat yang berbeda-beda, tetapi kurang berpengaruh dibandingkan atribut yang berhubungan dengan gaya.

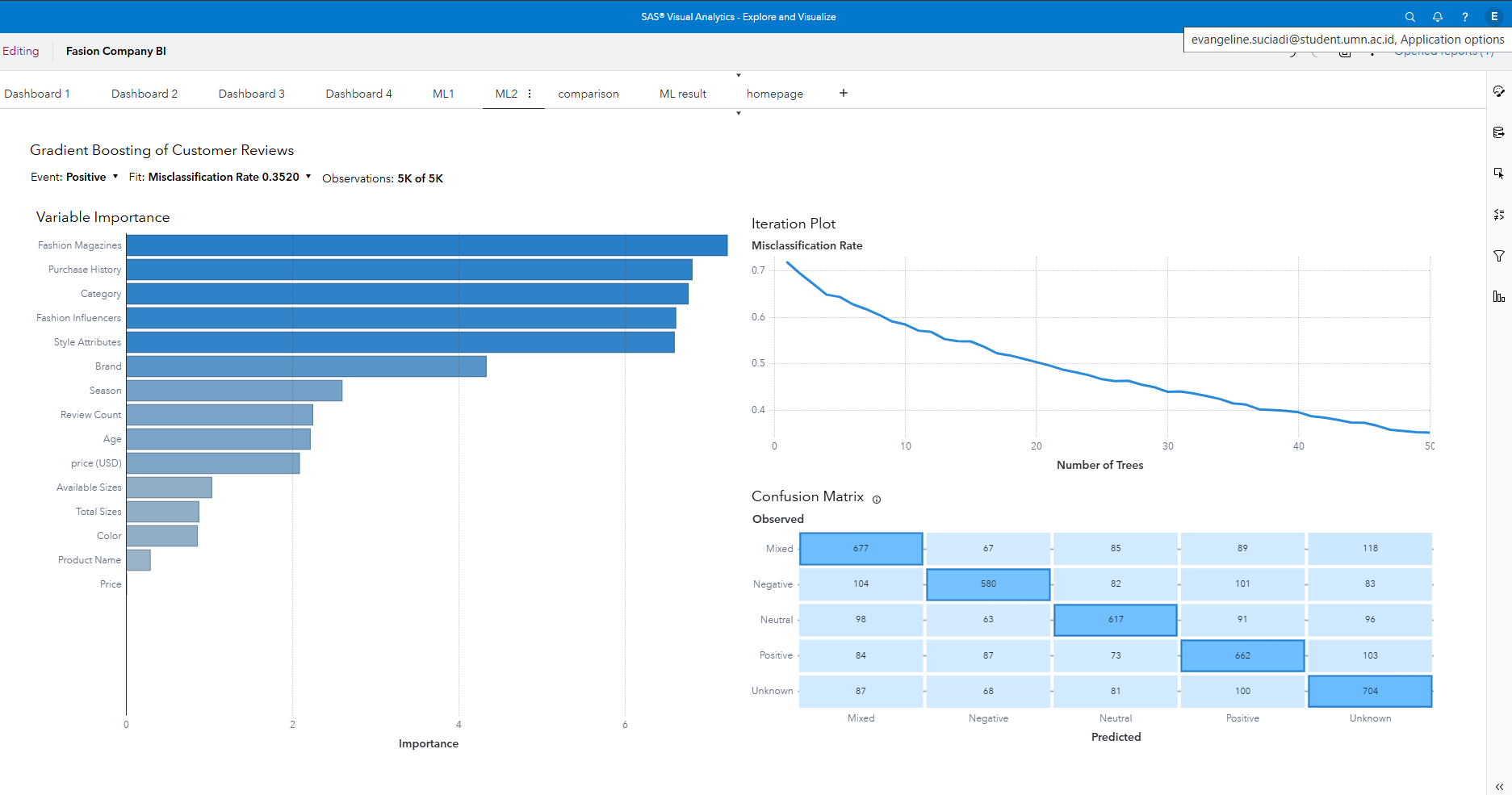
Plot kesalahan memvisualisasikan bagaimana kinerja model berkembang seiring dengan bertambahnya jumlah pohon keputusan dalam hutan acak. Training Error menurun seiring dengan bertambahnya jumlah pohon, dan menjadi stabil setelah sekitar 100 pohon, yang mengindikasikan bahwa model tersebut cocok dengan data pelatihan. Namun, Out-of-Bag (OOB) Error, yang merupakan ukuran kinerja generalisasi model, tetap lebih tinggi daripada training error. Hal ini menunjukkan adanya overfitting, yang berarti model mungkin terlalu selaras dengan data pelatihan dan dapat mengalami kesulitan saat membuat prediksi pada data baru yang tidak terlihat.

Confusion Matrix menunjukkan seberapa baik kinerja model di berbagai kategori sentimen: Campuran, Negatif, Netral, Positif, dan Tidak Diketahui. Matriks tersebut menunjukkan bahwa model kesulitan untuk mengklasifikasikan Sentimen Campuran secara akurat, dengan banyak ulasan yang salah diklasifikasikan sebagai Tidak Diketahui. Untuk ulasan Positif, model berkinerja cukup baik, tetapi model salah mengklasifikasikan sejumlah besar ulasan Netral dan Negatif, yang mengindikasikan kesulitan dalam membedakan perbedaan halus dalam sentimen pelanggan. Kelas Tidak Diketahui memiliki tingkat kesalahan klasifikasi yang tinggi, menyiratkan bahwa model ini tidak pasti atau tidak dapat dengan yakin menetapkan sentimen pada banyak ulasan.

Berbagai macam fitur, termasuk karakteristik produk (Musim, Atribut Gaya, Harga, Merek), atribut yang berhubungan dengan pelanggan (Usia, Riwayat Pembelian), dan pengaruh eksternal (Majalah Mode, Fashion Influencer), digunakan dalam model. Prediktor-prediktor ini menangkap berbagai dimensi sentimen dan perilaku pelanggan, tetapi beberapa di antaranya tampaknya memiliki kekuatan prediksi yang lebih besar daripada yang lain, seperti yang disorot dalam peringkat kepentingan variabel.

Model random forest mengidentifikasi Atribut Gaya, Fashion Influencer, dan Kategori sebagai pendorong sentimen pelanggan yang paling signifikan. Namun, dengan tingkat kesalahan klasifikasi lebih dari 75%, terdapat area yang jelas untuk perbaikan, terutama dalam menangani sentimen Campuran dan Tidak Diketahui. Meningkatkan model melalui rekayasa fitur, data tambahan, atau menyetel parameter random forest dapat membantu mengurangi tingkat kesalahan dan meningkatkan kemampuan model untuk memprediksi sentimen pelanggan secara akurat.

**ii) Gradient Boosting Machine Learning Model**



Model Gradient Boosting bertujuan untuk memprediksi sentimen pelanggan (misalnya, Positif, Negatif, Netral) dari ulasan produk dan fitur terkait. Berbeda dengan Random Forest, Gradient Boosting membangun pohon secara berurutan, dengan setiap pohon memperbaiki kesalahan dari pohon sebelumnya, sehingga menghasilkan model yang lebih kuat dari kumpulan model lemah. Dataset terdiri dari 5.000 observasi (ulasan pelanggan), dan targetnya adalah memprediksi sentimen positif, seperti yang ditunjukkan dalam header.

Tingkat kesalahan klasifikasi dari model ini adalah 0,3520, yang merupakan peningkatan signifikan dibandingkan dengan model Random Forest yang memiliki tingkat kesalahan sebesar 0,7592. Tingkat kesalahan klasifikasi yang lebih rendah menunjukkan bahwa model Gradient Boosting ini lebih akurat dalam mengklasifikasikan sentimen ulasan pelanggan. Sifat iteratif dari model ini memungkinkannya untuk terus mengurangi kesalahan seiring dengan penambahan pohon.

Pada model ini, Majalah Fashion adalah fitur paling penting untuk memprediksi sentimen, menunjukkan bahwa opini eksternal dari majalah berperan penting dalam membentuk sentimen pelanggan. Riwayat Pembelian juga sangat penting, menunjukkan bahwa perilaku pembelian masa lalu memengaruhi sentimen di masa depan, kemungkinan karena pelanggan yang setia atau pembeli berulang mungkin memiliki pendapat yang lebih kuat. Fitur lain seperti Kategori dan Influencer Fashion juga memiliki peran penting, menyoroti bahwa kategori produk tertentu atau dukungan influencer secara signifikan mempengaruhi perasaan pelanggan terhadap suatu produk. Fitur lainnya seperti Atribut Gaya, Merek, dan Musim juga cukup berpengaruh, menunjukkan bahwa karakteristik produk dan identitas merek berkontribusi pada prediksi sentimen. Variabel tambahan seperti Jumlah Ulasan, Usia, dan Harga juga dipertimbangkan, meskipun pengaruhnya lebih kecil dalam proses pengambilan keputusan model.

Plot Iterasi menunjukkan bagaimana kinerja model meningkat seiring dengan penambahan lebih banyak pohon. Awalnya, tingkat kesalahan klasifikasi tinggi (mendekati 0,7), tetapi secara bertahap menurun seiring dengan penambahan pohon, dan stabil di sekitar 50 pohon dengan tingkat kesalahan sekitar 0,35. Penurunan kesalahan yang bertahap ini mencerminkan kekuatan dari Gradient Boosting dalam meningkatkan performa dengan setiap iterasi.

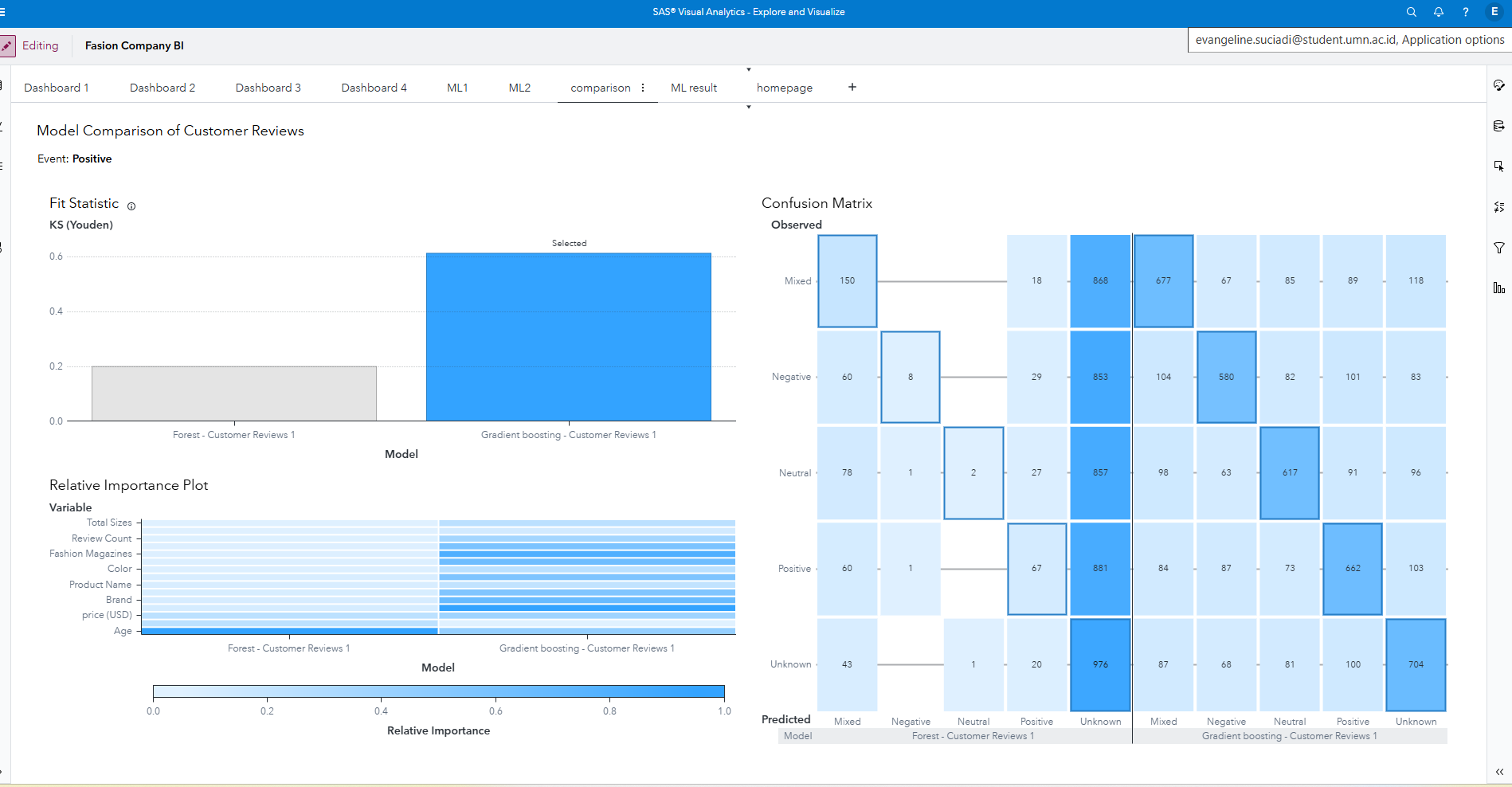
Matriks Kebingungan menggambarkan kinerja model Gradient Boosting di berbagai kelas sentimen (Mixed, Negative, Neutral, Positive, Unknown). Model ini berkinerja baik dalam mengklasifikasikan sentimen Netral dan Positif, seperti yang ditunjukkan oleh angka yang lebih tinggi di sepanjang diagonal untuk kategori tersebut (masing-masing 617 dan 662). Namun, masih terdapat beberapa kesalahan klasifikasi, seperti ulasan Negatif yang salah diklasifikasikan sebagai Netral atau Positif.

Untuk sentimen Mixed dan Unknown, model juga menunjukkan performa yang cukup baik, tetapi terdapat beberapa kesalahan klasifikasi, terutama dalam kategori Unknown, di mana sejumlah besar ulasan diklasifikasikan secara salah sebagai Mixed atau Negatif.

Model ini menggunakan set fitur yang sama dengan model Random Forest, termasuk karakteristik terkait produk (Musim, Atribut Gaya, Merek, Warna) dan variabel terkait pelanggan (Usia, Riwayat Pembelian), serta pengaruh eksternal seperti Majalah Fashion dan Influencer Fashion. Namun, pentingnya relatif dari fitur-fitur ini berbeda dari model sebelumnya, di mana Majalah Fashion dan Riwayat Pembelian memainkan peran lebih menonjol dalam model Gradient Boosting ini.

Model Gradient Boosting menunjukkan kinerja keseluruhan yang jauh lebih baik, dengan tingkat kesalahan klasifikasi sebesar 0,3520 dibandingkan dengan 0,7592 pada Random Forest. Majalah Fashion dan Riwayat Pembelian muncul sebagai fitur paling berpengaruh dalam memprediksi sentimen pelanggan, sementara Atribut Gaya dan Merek juga secara signifikan berkontribusi terhadap prediksi model. Plot Iterasi menunjukkan kemampuan model untuk secara bertahap mengurangi kesalahan seiring dengan penambahan lebih banyak pohon. Meskipun model ini berkinerja baik di sebagian besar kategori sentimen, masih ada beberapa tantangan dalam hal kesalahan klasifikasi, terutama di kategori Unknown. Secara keseluruhan, model ini memberikan pendekatan yang lebih akurat dan efisien untuk klasifikasi sentimen dibandingkan dengan model Random Forest.

**iii) Model Comparison**

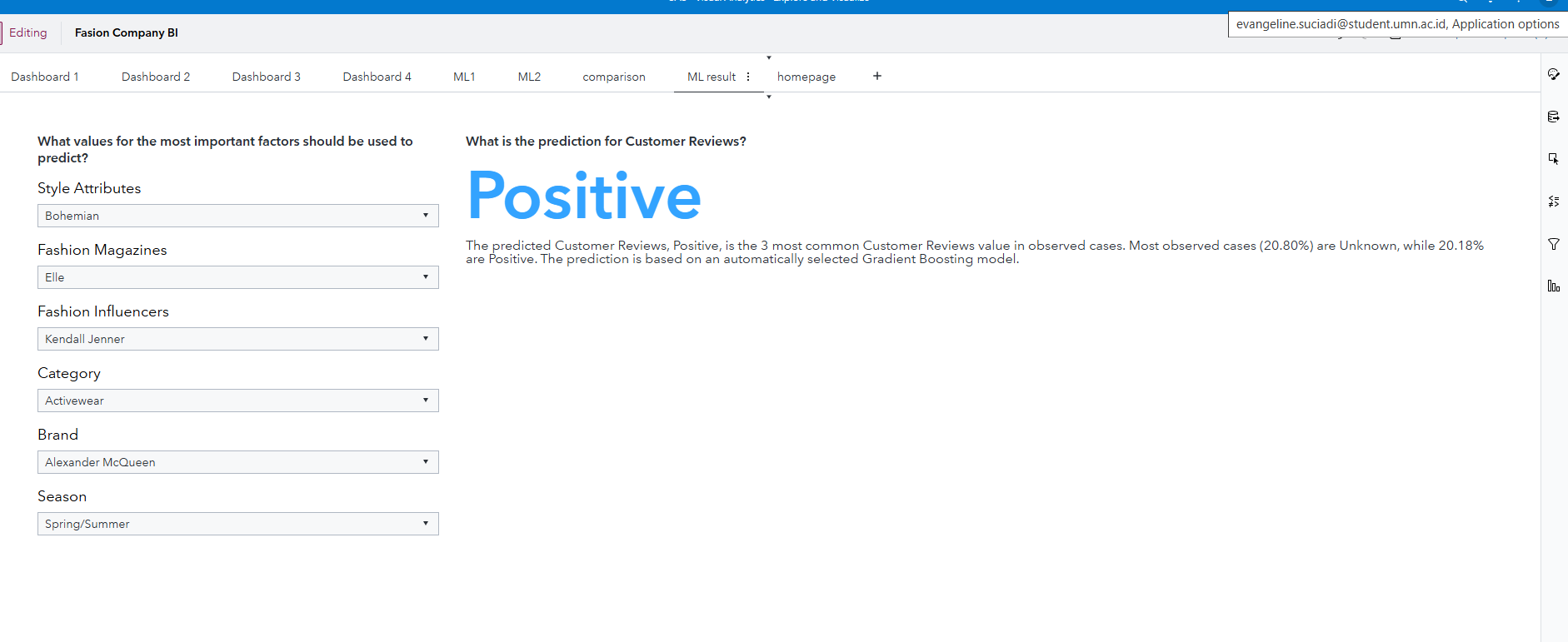


Perbandingan antara model Random Forest dan Gradient Boosting dalam menganalisis ulasan pelanggan menunjukkan perbedaan kinerja yang signifikan. Statistik kesesuaian (fit statistic), yang didasarkan pada indeks Youden yang diterapkan pada uji Kolmogorov-Smirnov (KS), menunjukkan bahwa model Gradient Boosting memiliki nilai KS yang lebih tinggi (mendekati 0,6), yang menunjukkan pemisahan kelas yang lebih baik dibandingkan dengan model Random Forest yang memiliki nilai KS lebih rendah (sekitar 0,2).

Matriks kebingungan lebih lanjut menggambarkan kinerja kedua model dalam berbagai kategori sentimen—Mixed, Negative, Neutral, Positive, dan Unknown—dengan baris yang mewakili kelas aktual dan kolom mewakili kelas yang diprediksi. Sel pada matriks dengan warna lebih gelap menunjukkan jumlah prediksi yang lebih tinggi, dan model Gradient Boosting umumnya menunjukkan kinerja yang lebih baik dengan prediksi yang lebih akurat di sepanjang diagonal, yang menunjukkan klasifikasi yang lebih baik di sebagian besar kategori.

Plot Kepentingan Relatif (Relative Importance Plot) menyoroti pentingnya berbagai fitur seperti Total Ukuran, Jumlah Ulasan, Majalah Fashion, Warna, Nama Produk, Merek, Harga (USD), dan Usia. Model Random Forest memberikan tingkat kepentingan yang relatif sama pada fitur-fitur ini, sedangkan model Gradient Boosting menekankan satu atau dua variabel kunci, yang menunjukkan proses seleksi fitur yang lebih halus. Secara keseluruhan, model Gradient Boosting menunjukkan kinerja yang lebih unggul berdasarkan statistik kesesuaian dan distribusi kepentingan variabel, menjadikannya pilihan yang lebih akurat untuk memprediksi sentimen ulasan pelanggan.

**iv) Prediction Demonstration using Gradient Boosting**



Dari algoritma Machine Learning yang dipilih, kami menciptakan solusi menggunakan model Gradient Boosting untuk memprediksi sentimen ulasan pelanggan, dengan hasil klasifikasi sebagai "Positif." Prediksi ini dipengaruhi oleh berbagai faktor: Bohemian dipilih sebagai atribut gaya, Elle sebagai majalah fashion terkait, Kendall Jenner sebagai influencer fashion, Activewear sebagai kategori produk, Alexander McQueen sebagai merek, dan Spring/Summer sebagai musim. Model ini menunjukkan bahwa "Positif" adalah salah satu dari tiga sentimen yang paling umum diamati dalam dataset, dengan 20,80% kasus diklasifikasikan sebagai "Unknown" dan 20,18% sebagai "Positif." Prediksi ini bergantung pada kemampuan model untuk secara otomatis mempertimbangkan bobot dari faktor-faktor ini dan mendeteksi pola yang terkait dengan umpan balik positif, menggunakan atribut yang diberikan untuk meningkatkan akurasi klasifikasi.