**UTS IS 529-C**

**Advanced Big Data Analytics**

**“ChizuAutocar”**

**Automotive Business Retail**



**By Christopher Darren - 00000054804**

Saya menamakan perusahaan saya “ChizuAutocar”. ChizuAutocar merupakan sebuah **perusahaan retail** yang bergerak pada **bidang otomotif terutama penjualan mobil** dari berbagai brand yang ada di seluruh dunia. Perusahaan *retail* ini berlokasi di India tepatnya di Mumbai pada tahun 2000. ChizuAutocar menjual berbagai brand mobil yang dipakai di seluruh dunia dimulai dari MercedesBenz, Hyundai, Maruti, Audi, Toyota ,Honda, Ford, Skoda, Volkswagen, Jaguar dan lain lain.

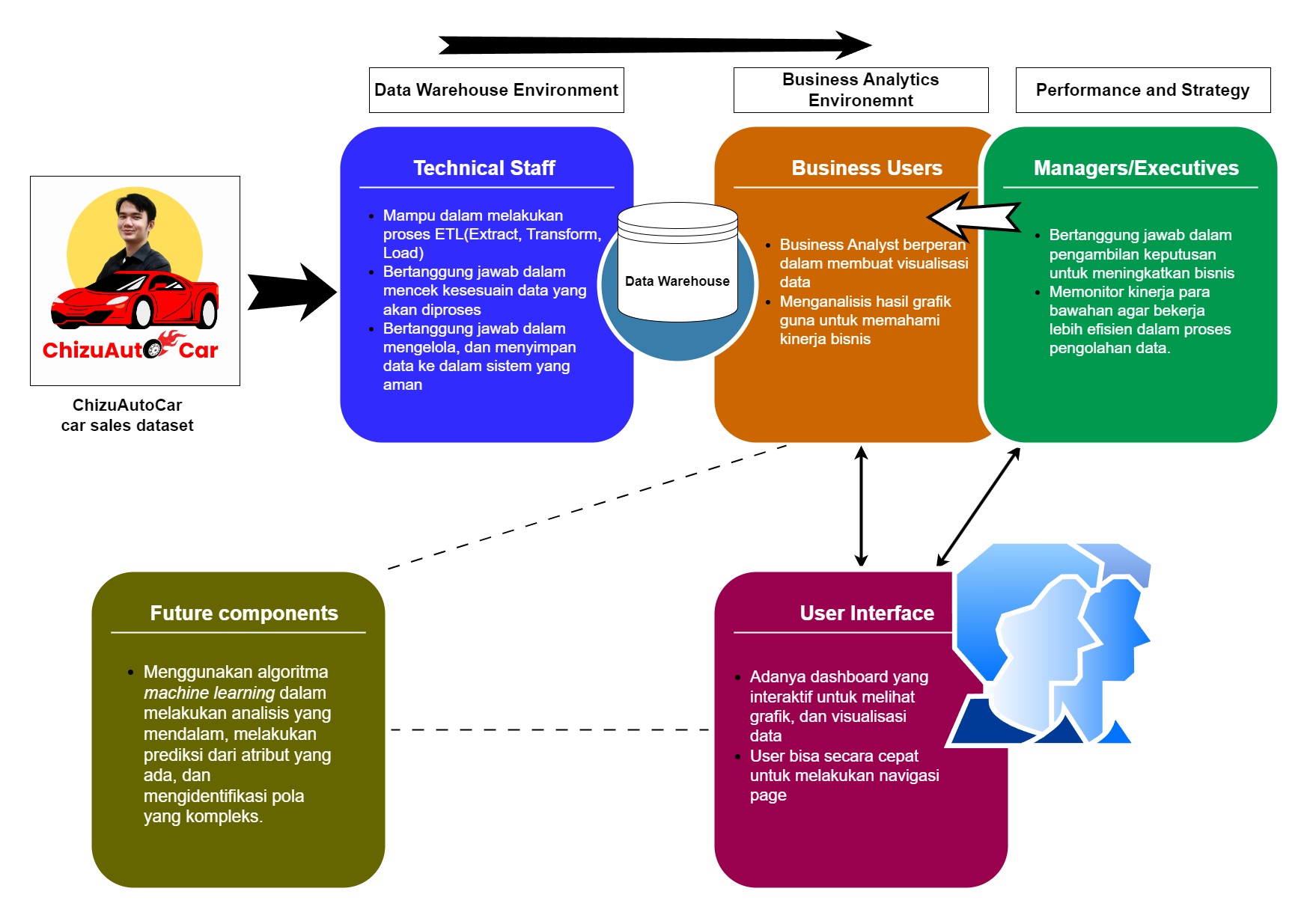
ChizuAutocar merupakan perusahaan *retail* India yang kuat dalam penjualan mobil dari tahun 2005 hingga tahun 2017, namun belakangan ini ChizuAutocar mengalami masalah dimana di tahun 2019 penjualan mereka turun drastis dari berbagai brand yang mereka jual. ChizuAuto car memerlukan solusi agar penjualan mobil meningkat dan mendapatkan keuntungan, maka dari itu Saya beserta tim *business intelligence* ingin mempelajari masalah apa yang terjadi dengan menggunakan *software business* untuk melihat grafik penjualan dan menemukan solusi yang baik untuk meningkatkan penjualan ChizuAutocar. Setelah itu saya juga akan membuat model prediksi menggunakan *software business* inti memprojeksikan kira kira penjualan mobil ChizuAutocar di masa yang akan datang akan terlihat seperti apa.

**Sumber dataset**: <https://www.kaggle.com/datasets/adityaa09/used-car-sales-dataset>

**1. Question 1: Sub-CLO-2, Weight (25%).**

**Fill in the blue text in the diagram below that corresponds to the field and company you work for**

**(You will create a new business intelligence diagram that matches your answers)**

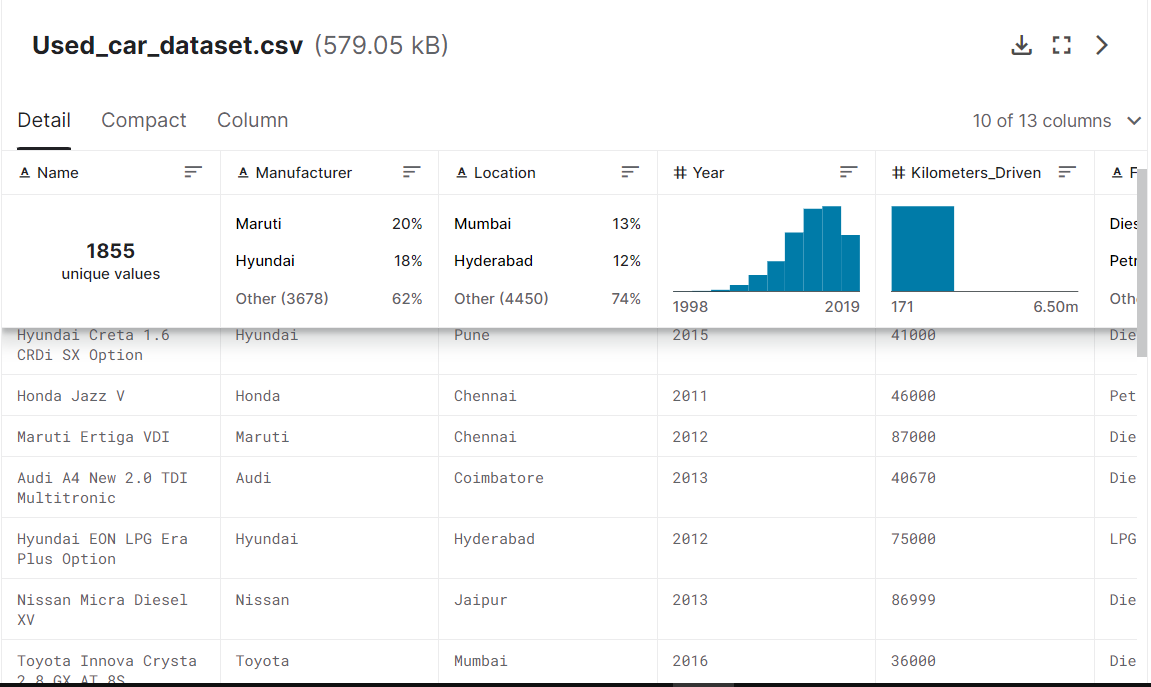


Gambar 1. diagram *business intelligence* ChizuAutocar

**2. Question 2: Sub-CLO-3, Weight (25%)**

**Completely explain each of your answers on the business intelligence diagram you made.**

* ***List of Data Source***



Gambar 2. data set untuk ChizuAutocar

Saya mengambil datasetnya dari situs kaggle.com, dan datanya tersebut adalah data retail penjualan mobil di beberapa daerah pada negara India. Pada dataset ini terdiri dari (13 kolom dan 5976 baris/record data). Dataset ini juga terdiri dari beberapa tipe data seperti Numerikal, Kategorikal, YEAR, dan FLOAT. Berikut penjelasan lebih lengkapnya dari setiap column yang ada.

* + Name : Model mobil yang tertera misal, dari Make “Ford”, model “Focus”, jika digabungkan maka Ford Focus pabrikan yang menciptakan sebuah mobil;
  + Manufacture : pabrikan yang menciptakan sebuah mobil;
  + Location : yakni lokasi dimana mobil tersebut dibeli;
  + Year : tahun mobil tersebut dibeli;
  + Kilometers\_Driven : jarak mobil yang sudah ditempuh;
  + Fuel\_Type : jenis bahan bakar mobil yang digunakan;
  + Transmission : Transmisi kendaraan yang digunakan;
  + Owner\_Type : yaitu kepemilikan mobil;
  + Engine\_CC : yakni volume ruang dalam silinder mesin kendaraan;
  + Power : satuan kekuatan mobil yang bisa dihasilkan oleh mesin dalam berupa HP, maupun BHP, ataupun NM;
  + Seats : jumlah tempat duduk yang tersedia pada kendaraan tersebut;
  + Mileage : jarak mobil yang sudah ditempuh menggunakan sistem imperial dengan satuan *mil*;
  + Price : harga mobil dalam satuan India (Rupee).
* ***Technical Staff***

Seperti pada diagram yang sudah ada di atas, pada bagian *technical staff* ini melibatkan staff bagian data untuk melakukan proses ETL pada dataset ChizuAutocar yang sudah didapat dari internet. Singkatnya dalam proses ETL ini, *extract* yakni mengekstrak data dari berbagai sumber baik dari data internal perusahaan maupun internet. Dalam kasus ini peneliti sudah dapat data dari internet untuk data Used\_car\_dataset dengan format .csv. Proses selanjutnya yakni *Transform* dimana data yang sudah didapat perlu dicek apakah terdapat data kosong, data duplikasi, dan lain lain yang harus dibersihkan dengan tujuan data tersebut konsisten dan layak digunakan oleh para analyst. Proses terakhir dari ETL yakni *load, load* berarti data yang sudah diekstrak dan ditransform akan dimasukkan ke dalam sistem atau *software* yang mendukung untuk membaca data.

* ***Business users***

Seperti pada diagram yang sudah ada di atas, di bagian *business users* bertugas untuk membuat visualisasi visualisasi yang relevan sesuai dataset ChizuAutocar yang tersedia dalam sistem. Visualisasi ini bisa berupa pie chart, bar chart, dan lain lain yang mendukung dalam proses bisnis ChizuAutocar. Tidak hanya itu peran business user juga penting dalam membantu menganalisis grafik yang sudah sudah dibuat guna untuk memahami kinerja bisnis dari ChizuAutocar.

* ***Managers executives***

*Managers* ataupun *executives* pada ChizuAutocar berperan penting dalam proses bisnis dikarenakan mereka berperan dalam pengambilan keputusan akhir dari hasil analitika data yang sudah diproses. Dengan demikian mereka bisa merencanakan strategi bisnis ChizuAutocar agar menjadi perusahaan yang lebih efisien. Terutama untuk *executives* yang bertugas untuk mengarahkan visi dan misi sesuai harapan bisnis kejalan yang benar.

* ***User Interface***

*User Interface* atau desain antarmuka dalam konteks *big data analytics* merupakan tampilan dimana pengguna bisa berinteraksi dengan sistem *big data* itu sendiri. Dari *user interface* ChizuAutocar ini diharapkan pengguna bisa dengan mudah untuk mengnavigasi halaman dari hasil analitikas data yang sudah diproses oleh *data analyst* dan *business intelligence.* Pada *user interface* ChizuAutocar ini juga nantinya akan ada sebuah dashboard supaya bisa melihat hasil visualisasi secara keseluruhan dan bisa melihat perubahan secara langsung ketika user melakukan filter data pada *dashboard* analitika data. *Dashboard*  penting bagi pengguna agar bisa berpindah pindah halaman visualisasi secara cepat dan mudah. Maka dari itu harapan *dashboard* dari ChizuAutocar adalah memudahkan penggunaan dalam mengakses informasi.

* ***Future components***

*Future components* dalam konteks *big data analytics* adalah sebuah komponen kedepannya yang bisa mendukung kegiatan penelitian dalam sebuah analitika data. ChizuAuto car pasti kedepannya bisa menghadapi permasalahan dan jika tidak mempunyai *tools* yang mendukung untuk menyelesaikan permasahan maka ChizuAutocar akan kewalahan dalam pengelolaan analitika data selanjutnya. Dalam bisnis ini akan menggunakan algoritma *machine learning* dalam melakukan analisa yang lebih mendalam , memungkinkan untuk melakukan prediksi data penjualan, serta melihat pola data yang kompleks menggunakan *machine learning.*

**3. Question 3: Sub-CLO-5, Weight (25%)**

**a) Big Data Analytics that you implement in the company where you work is what type of analytics?**

Pada perusahaan ChizuAutocar, tipe *big data analytics* yang mungkin akan saya gunakan adalah *Descriptive Analysis* dan *Predictive Analysis*. Peneliti akan uraikan lebih lanjut mengenai 2 tipe analitik ini dalam bisnis ChizuAutocar.

* *Descriptive Analysis*

Secara singkat *descriptive analysis* adalah sebuah analisa dimana menjelaskan kejadian di masa lampau. Dalam data ChizuAutocar akan ada visualisasi data dimulai dari penjualan secara keseluruhan, distribusi penjualan, serta model mobil yang paling banyak dibeli dan *dashboard* untuk menjelaskan data yang sudah ada sebelumnya agar bisa lebih dimengerti oleh *business analyst*.

* *Diagnostic Analysis*

Secara singkat *diagnostic analysis* adalah dengan melihat indikator apa saja yang membuat suatu value meningkat atau menurun pada suatu kejadian. Nantinya akan ada visualisasi indikator yang jelas variable apa yang paling berpengaruh dalam penjualan mobil pada perusahaan ChizuAutocar.

* *Predictive Analysis*

Secara singkat *predictive analysis* adalah sebuah analisa dimana bertujuan untuk memprediksi di masa depan berdasarkan data historis ataupun tren. Hal ini mungkin bisa dicapai menggunakan *machine learning.* Pada perusahaan ChizuAutocar akan membuat visualisasi prediksi singkat menggunakan bantuan *Microsoft PowerBI* supaya bisa tahu penjualan kedepannya akan naik atau turun.

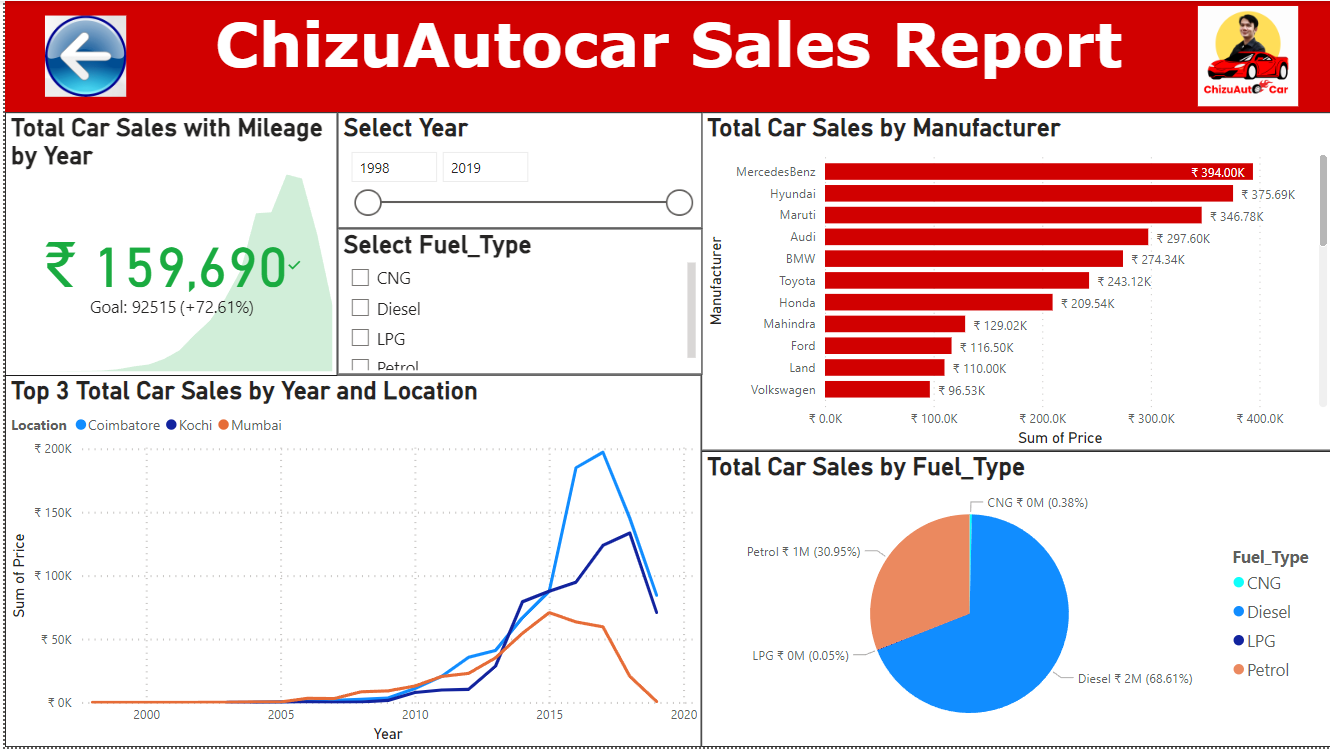
**b) Illustrate what kind of output you produce from the analytical process you do.**

Berdasarkan penjelasan pada point a) berikut hasil output berdasarkan 2 tipe analitik yang saya gunakan untuk bisnis ChizuAutocar.

* *Descriptive Analysis*

Hasil dari descriptive analysis pada ChizuAutocar menghasilkan 3 dashboard yang mempunyai tujuan yang berbeda beda. Berikut 3 dashboard ChizuAuto car dari segi *descriptive analysis*.

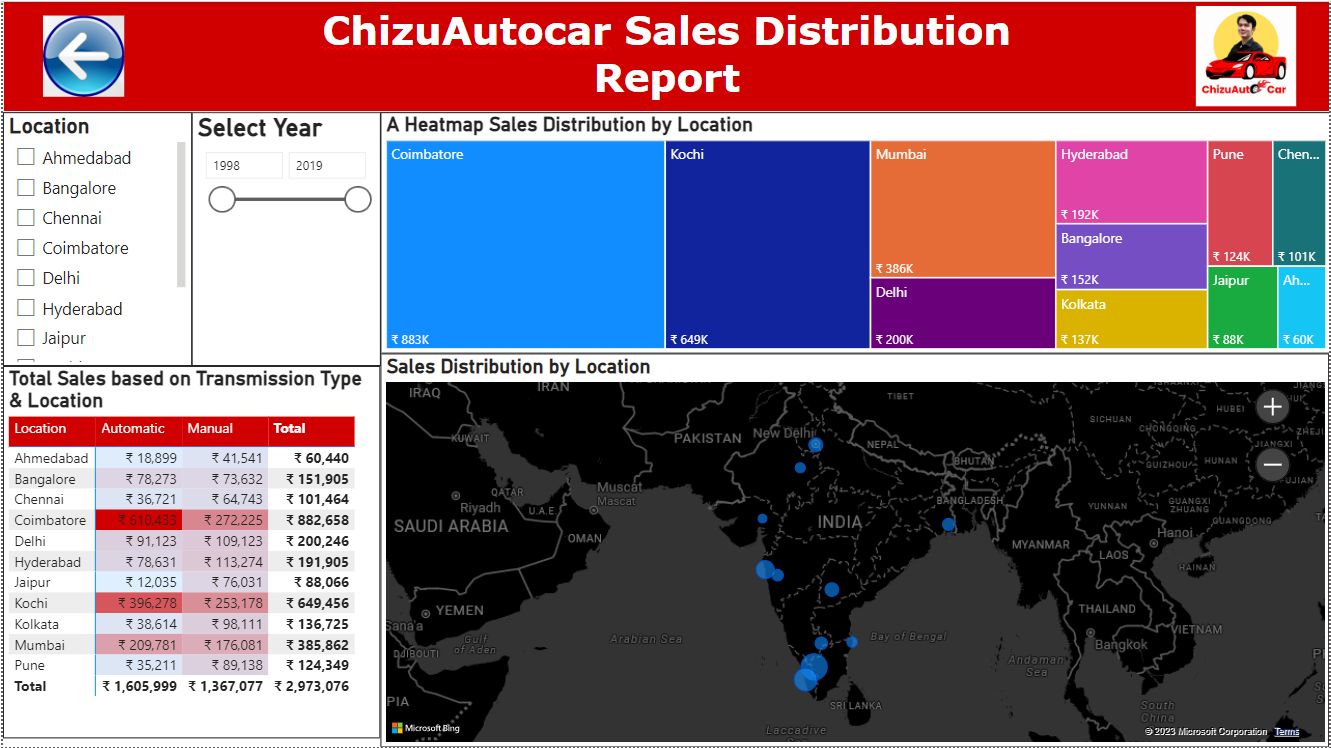
* 1. ChizuAutocar *Sales Report*



Gambar 3. *Dashboard* ChizuAutocar *Sales Report*

*Dashboard* diatas merupakan sebuah dashboard laporan penjualan mobil dari ChizuAutocar. Bisa diketahui bahwa penjualan pertahun berdasarkan *mileage* mobil naik dari tahun ke tahun sebesar 159,690 Rupee. Kemudian dari dashboard ini kita bisa ketahui bahwa keseluruhan hasil penjualan mobil terbanyak adalah dari manufaktur MercedesBenz dengan sebesar 393,955,000 Rupee, diikuti oleh urutan kedua oleh Hyundai, dan urutan ketiga oleh Maruti. Berpindah ke *linechart* di bagian bawah, bisa disimpulkan bahwa ada 3 kota yang dari tahun ke tahun membeli mobil dari ChizuAutocar paling banyak yaitu kota Coimbatore, Koichi, dan Mumbai. Terutama di tahun 2017 dimana penjualan mobil dari berbagai brand terbanyak diraih oleh kota Coimbator , diikuti oleh Koichi dan ketiga Mumbai. Namun setelah tahun 2017 terdapat penurunan yang derastis dari ketiga kota tersebut entah minat membeli mobil sedikit atau tidak laku. Visual terakhir adalah hasil penjualan mobil berdasarkan bahan bakar. Seperti yang diketahui dari tahun ketahun dan dari semua brand yang tercatat dalam penjualan pembeli mobil lebih minat membeli mobil dengan bahan bakar diesel dengan persentase 68,61%, urutan kedua ditempati oleh petrol atau bensin sebesar 30,00%, urutan ketiga ditempati oleh CNG sebesar 0.38%, dan urutan terakhir ditempati oleh LPG sebesar 0.05%.

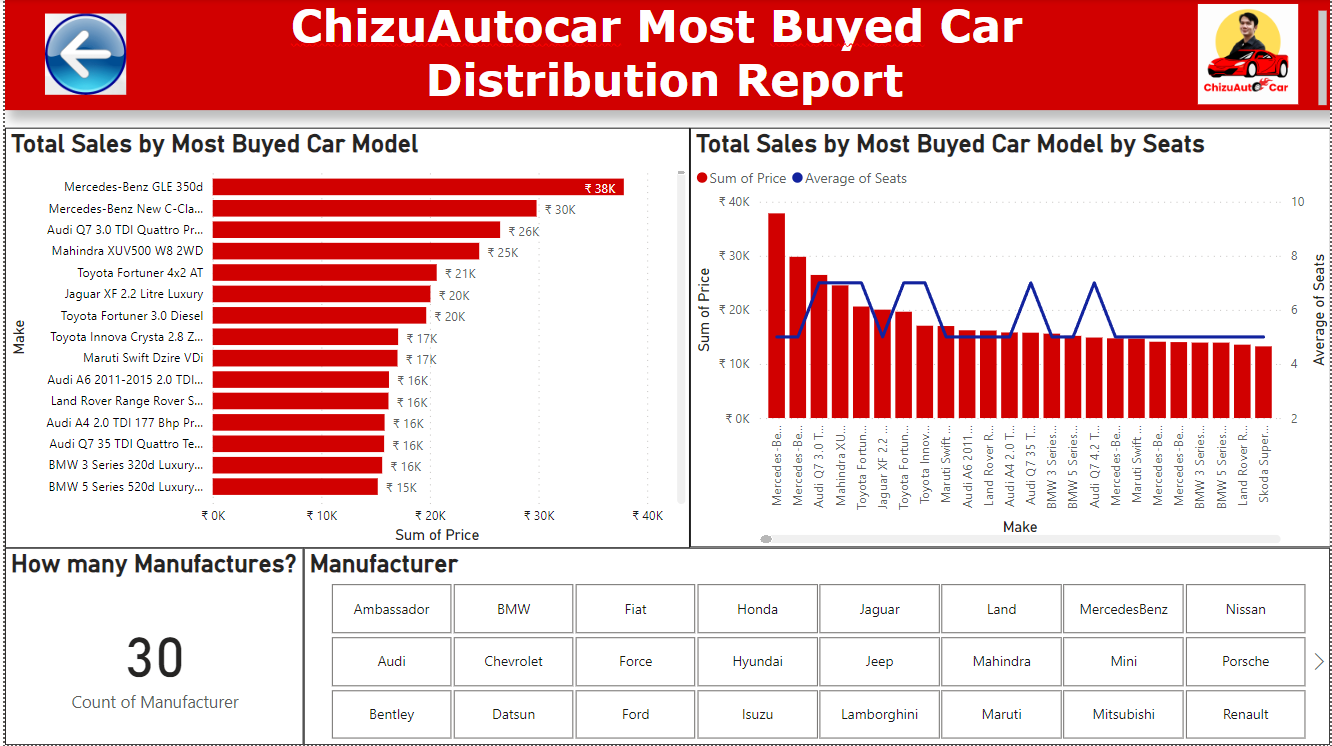
* 1. ChizuAutoCar *Sales* *Distribution Report*



Gambar 4. *Dashboard* ChizuAutocar *Sales Distribution Report*

*Dashboard* diatas merupakan sebuah *dashboard* laporan distribusi penjualan mobil di berbagai lokasi di India. Singkatnya hasil penjualan terbanyak terdapat di lokasi Coimbatore dengan jumlah 883.000 Rupee, diikuti oleh urutan kedua Koichi dengan jumlah 649.000 Rupee, diikuti oleh Mumbai sebesar 386K Rupee begitu pula seterusnya. Hal ini bisa memberikan *insight* bahwa dulu Coimbatore merupakan kota yang membutuhkan banyak mobil dalam menunjang kehidupan orang sehari hari. Berpindah ke table di bagian bawah merupakan table mengenai hasil penjualan mobil di berbagai lokasi India berdasarkan tipe transmisi mobil yang dibeli oleh konsumen. Terlihat dengan jelas bahwa pembeli mobil bertransmisi otomatis yang berlokasi di Coimbatore yang paling banyak jumlahnya dengan angka 610.433 Rupee, diikuti urutan kedua oleh Koichi sebesar 396.278 Rupee dan terakhir Mumbai sebesar 209.781 Rupee. Transmisi manual juga demikian masih diraih oleh tempat yang sama. Melihat hasil *Grand total* dari subtotal masing masing kota ketika digabungkan hasilnya adalah 2,973,076 Rupee.

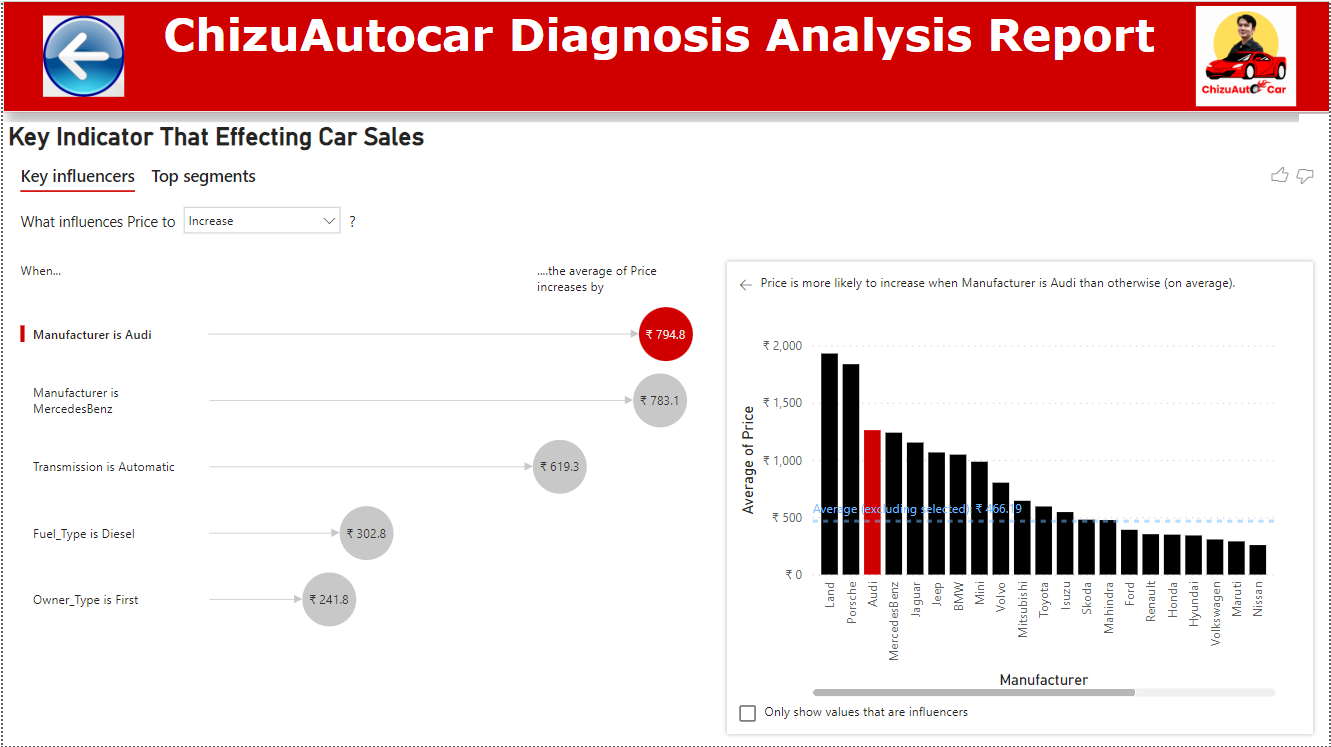
* 1. ChizuAutocar *Most Buyed Car model Distribution Report*



Gambar 5. *Dashboard* ChizuAutocar *Most Buyed Car Distribution* *Report*

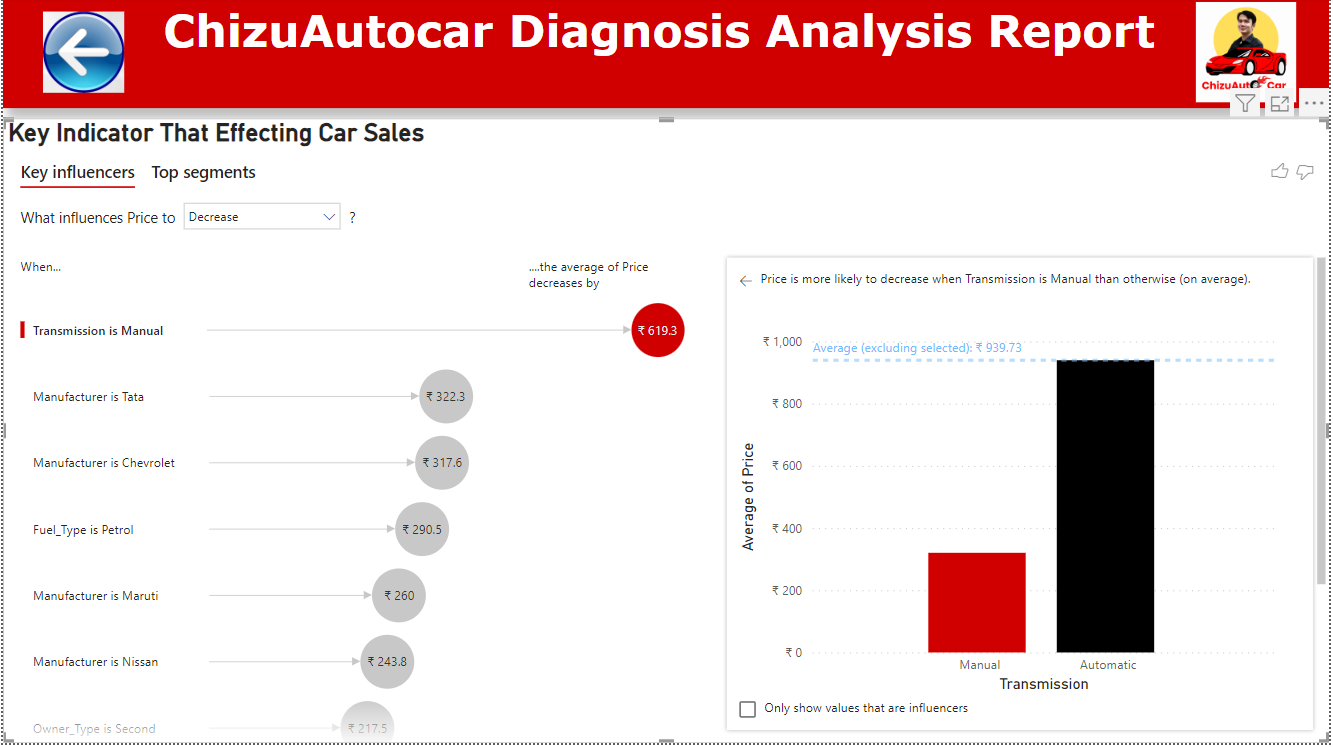
*Dashboard* diatas merupakan sebuah *dashboard* laporan penjualan mengenai distribusi model mobil yang paling banyak dibeli oleh pembeli/*customer*. Dari dashboard dapat memberikan insight bahwa model mobil yang paling banyak dibeli adalah dari brand Mercedes dengan model GLE 350d sebesar 37,853 Rupee, diikuti oleh urutan kedua model New C-Class sebesar 30K Rupee, dan begitu pula seterusnya. Tidak hanya itu data terdahulu ini menggambarkan bahwa masyarakat India suka dengan mobil Mercedes Benz sehingga banyak yang meminat untuk membeli mobil ini, hal ini bisa saja membantu perusahaan ChizuAutocar agar kedepannaya lebih fokus dalam menjual mobil bermerek Mercedes Benz. Kemudian dari data penjualan ChizuAutocar terdapat 30 manufacture dari berbagai belahan dunia sehingga mobil yang dijual variatif dan tidak monoton. Tampilan visualisasi akhir dari dashboard ini adalah total penjualan model mobil yang paling banyak dibeli berdasarkan jumlah tempat duduk. Dari hasil *graph* yang ada , dapat memberikan *insight* bahwa meskipun banyak konsumen yang membeli model GLE 350d, pelanggan lebih memilih untuk membeli mobil dengan jumlah kursi banyak seperti 7-seater pada brand Audi Q7 3.0 T, dan Mahindra. Hal ini disebabkan karena mobil dengan 7-seater lebih bisa muat banyak penumpang dibandingkan oleh 4-5 seater seperti dari brand Mercedes-Benz GLE 350d.

* *Diagnostic Analytics*

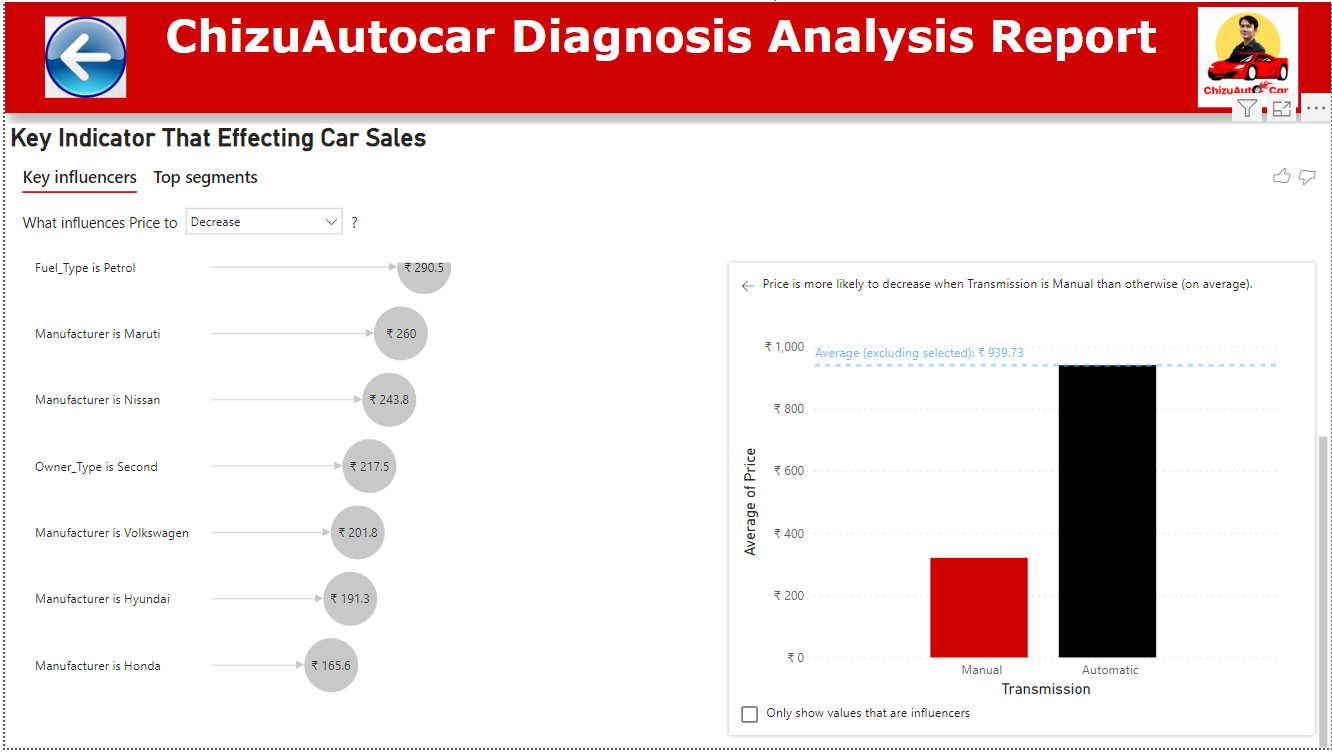


Gambar 6. *Dashboard* mengenai *key indicator* atau variable yang bisa mempengaruhi penjualan mobil ChizuAutocar (pro)

*Dashboard* yang dibuat business user diatas merupakan sebuah dashboard yang berisi tentang *key indicator* apa saja yang bisa mempengaruhi penjualan mobil. Pada visualisasi diatas business user menetapkan 6 variable utama yang bisa mempengaruhi penjualan yaitu Manufacture, Transmission, Owner\_Type dan Fuel\_Type. Visualisasi yang akan dianalysis adalah berdasarkan penjualan *sales(Price)* secara keseluruhan. Secara singkat penjualan mobil akan naik drastis apabila rata rata penjualan mobil tersebut berasal dari manufacture Audi sebesar 794 Rupee diikuti oleh urutan kedua manufactur MercedesBenz sebesar 783.1 Rupee. Sedangkan variable yang paling tidak berpengaruh dalam peningkatan penjualan adalah Owner\_Type dengan segmentasi First sebesar 241.8 Rupee.



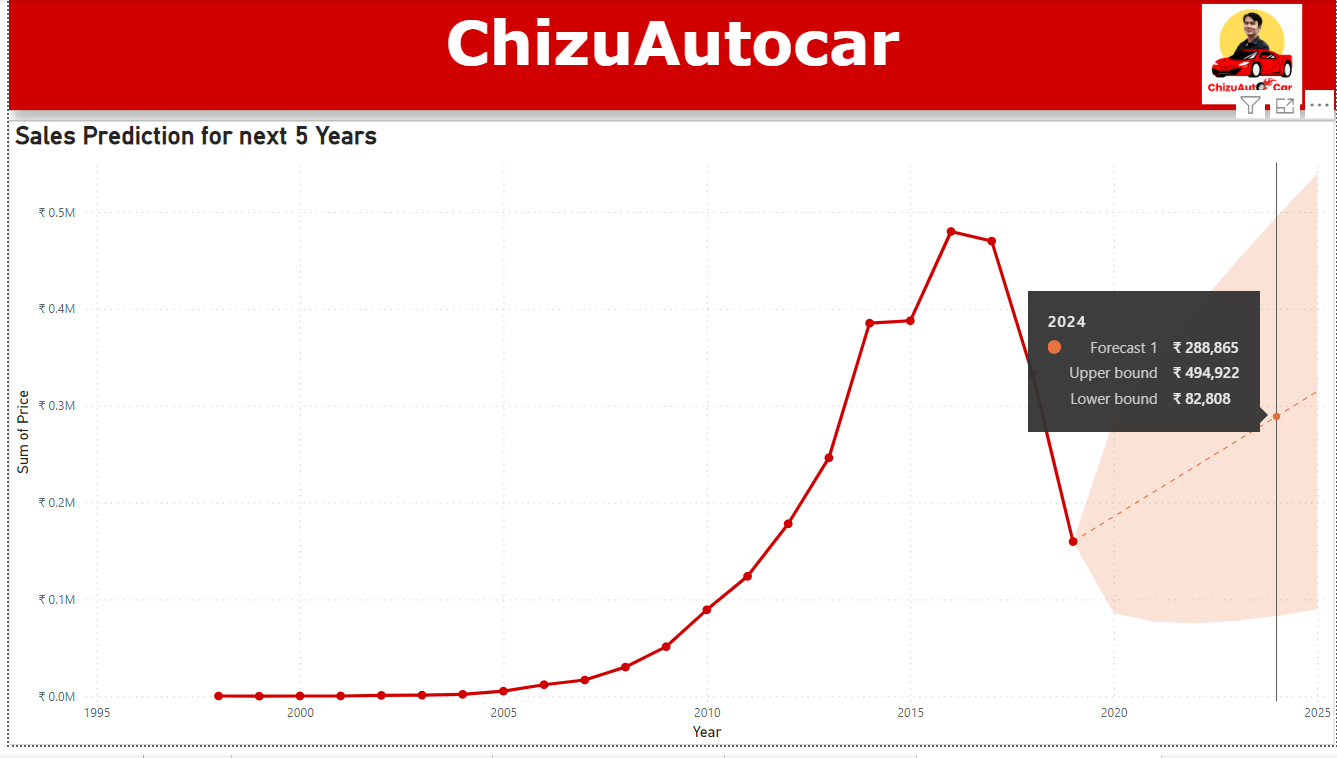
Gambar 7. *Dashboard* mengenai *key indicator* atau variable yang bisa mempengaruhi penjualan mobil ChizuAutocar (kontra)



Gambar 8. *Dashboard* lanjutan mengenai *key indicator* atau variable yang bisa mempengaruhi penjualan mobil ChizuAutocar (kontra)

*Dashboard* yang dibuat business user diatas merupakan sebuah dashboard yang berisi tentang key indicator apa saja yang bisa mempengaruhi penjualan mobil. Pada visualisasi diatas business user menetapkan 6 variable utama yang bisa mempengaruhi penjualan yaitu Manufacture, Transmission, Owner\_Type dan Fuel\_Type. Visualisasi yang akan dianalysis adalah berdasarkan penjualan *sales(Price)* secara keseluruhan. Secara singkat penjualan mobil akan turun apabila rata rata penjualan mobil tersebut berasal dari pembelian mobil yang bertipe transmisi *Manual* sebesar 619.3 Rupee. Diikuti oleh manufaktur Tata yang bisa menyebabkan penurunan penjualan sebesar 322.3 Rupee. Sedangkan variable yang paling tidak berpengaruh dalam penurunan penjualan adalah manufactur Honda sebesar 241.8 Rupee.

* *Predictive analysis*



Gambar 9. Grafik visualisasi *linechart prediction* untuk ChizuAutocar

Grafik visualisasi linechart diatas merupakan sebauh *prediction analysis* yang dibuat oleh business user dalam memprediksi penjualan ChizuAutocar kedepannya. Prediksi dilakukan dari tahun 2019 akhir hingga tahun 2024 dan bisa dilihat bahwa kurva grafiknya cenderung naik.

**4. Question 4: Sub-CLO-7, Weight (25%)**

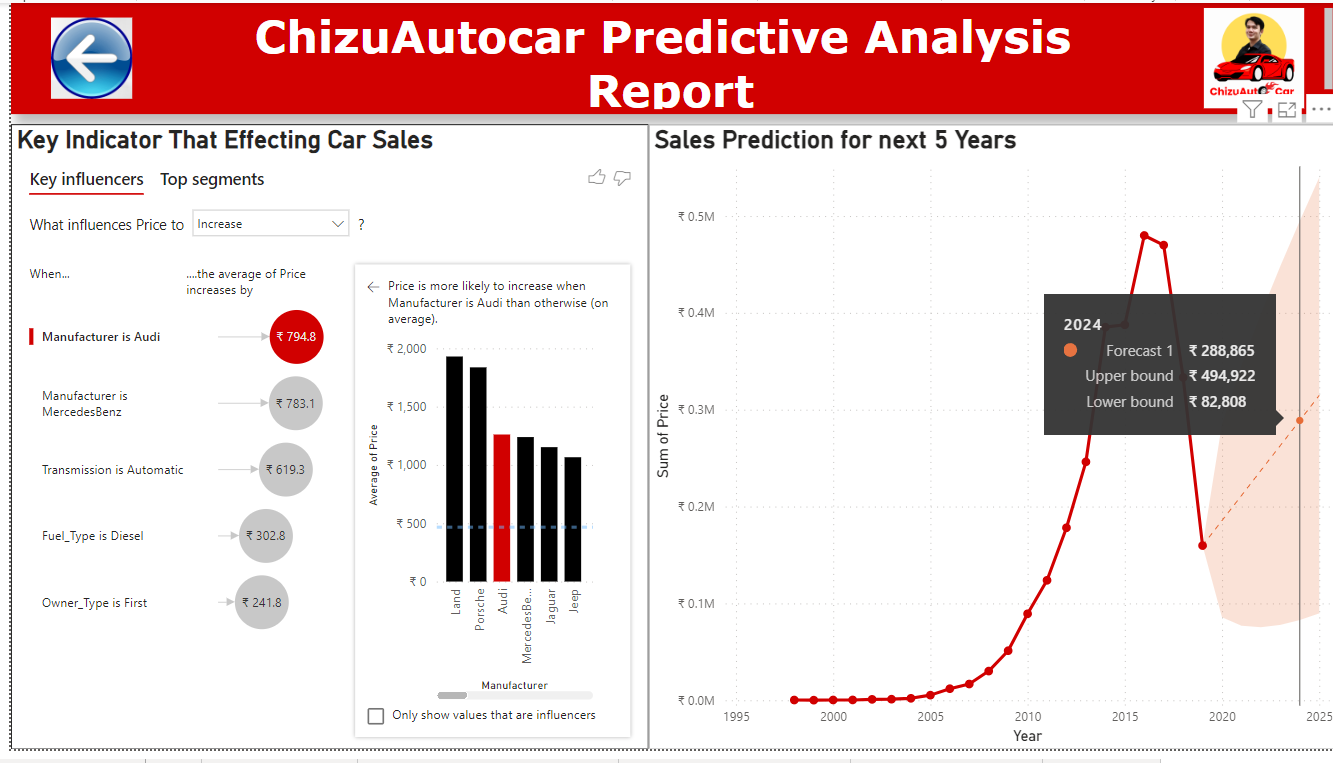
**a) Completing your answer above (#2 question), explain what analytical applications you**

**Implement.**

Dalam perancangan *business intelligence* untuk perusahaan ChizuAutocar saya menggunakan software *Microsoft PowerBI* sebagai tools andalan utama dalam pembuatan visual atau grafis bisnis ChizuAutocar. Saya menggunakan software ini karena kemudahannya dalam membuat sebuah visualisasi yang elegan dan menarik terutama dalam sisi analitik deskriptif. Tidak hanya itu, setiap saya melakukan perubahan data maka akan berubah secara *realtime* dalam artian memudahkan user ataupun pengguna bisnis ketika ingin mengetahui perubahan apa saja ketika sedang di navigasi atau di filter. Contohnya seperti pada gambar 3 dimana dashboard tersebut sangat interaktif apabila diklik, terutama pada bagian filter. Ketika suatu filter akan diklik maka visualnya akan berubah mengikuti perubahan filter yang sudah diaplikasikan, tidak hanya itu *Microsoft PowerBI* juga bisa melakukan filter secara bersamaan seperti memilih lebih dari 1 variable sehingga bisa dikombinasi sekreatif mungkin oleh saya dan para *business user* lainnya. Selain itu, *Microsoft PowerBI* juga bisa membuat analitika data yang kuat dengan cara membuat sebuah diagnosis yang sudah diterapkan pada gambar 6 hingga gambar 8 untuk mengetahui apa variable yang memungkinaan berperan besar dalam sebuah penjualan yang terjadi pada ChizuAutocar. Kemudian langkah terakhir adalah *Microsoft PowerBI* juga bisa melakukan prediksi singkat menggunakan *Linechart* dan *forecasting* sehingga bisa mengetahui kejadian apa yang akan terjadi di masa yang akan datang pada perusahaan ChizuAutocar.

**b) Explain what *statistical methods* or *machine learning algorithms* you use in the analytical process at the company and illustrate the modeling results.**

Dalam pembuatan prediksi analisis pada ChizuAutocar, saya menggunakan *machine learning* *built in* dari *Microsoft Power BI* sendiri. Untuk tipe grafik sudah dijelaskan pada gambar 9 yakni menggunakan linechart sebagai visualisasi pendukung *business*. Berikut *dashboard* predictive analytics dengan menggunakan key indicators yang berpengaruh terhadap penjualan mobil di ChizuAutoocar.



Gambar 10. *Dashboard* Key indicators beserta dengan prediksi penjualan ChizuAutocar dalam 5 tahun kedepan

Seperti penjelasan gambar 6 hingga 8, *dashboard* ini merupakan gabungan antara *key indicator* dan prediksi. Jika *user* bisnis mengarahkan line chart di tahun 2018 maka akan menampilkan key indicator yang ada disebelah kiri. Melengkapi penjelasan yang ada di gambar 9 ini merupakan hasil model prediksi yang mungkin terjadi di masa depan setelah tahun 2019-2024. *Business user* dalam membuat model prediksi ini menetapkan confidence interval sebesar 95% yang merupakan hal standar untuk melakukan statistik karena memberikan keseimbangan yang baik antara tingkat keyakinan yang tinggi dan interval yang cukup ketat. Melihat hasil prediksi di tahun 2024 bisa dilihat bahwa hasil penjualan prediksinya sebesar 288.865 dimana peningkatannya naik drastis dibandingkan tahun tahun sebelumnya seperti di tahun 2020 sebesar 185,525 Rupee. Untuk batas nilai maksimum yang dapat diperoleh pada prediksi di tahun 2024 sebesar 494,922 Rupee sedangkan nilai batas minimum yang diperoleh sebesar 82,808 Rupee.