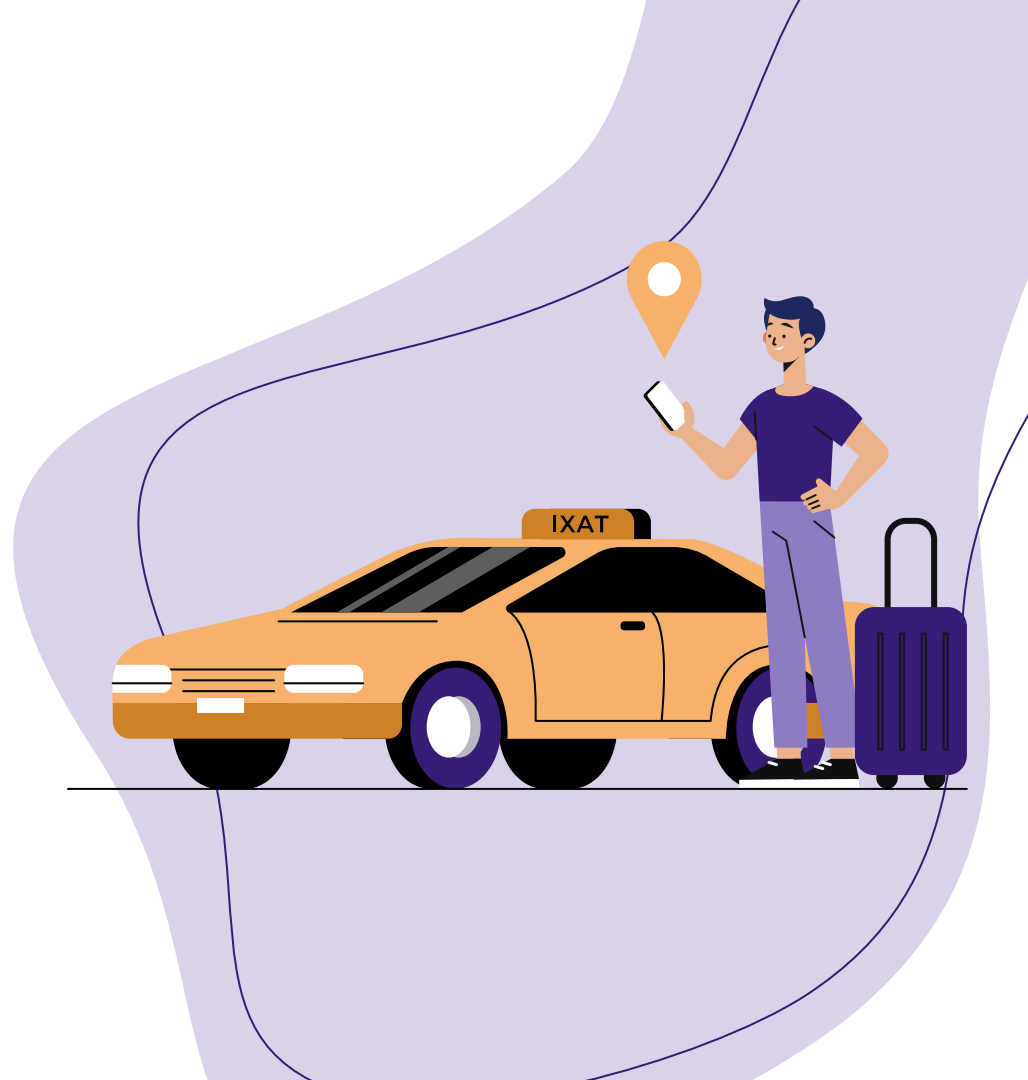


Demand Forecasting for Ride-Sharing Trips using Time Series Analysis

By: Ramadhoni Nasri



Latar Belakang Masalah

Periode Mei hingga September 2019

Total Driver

142.987

Total Passenger

1.175.267

Jumlah driver sangat jauh lebih sedikit daripada jumlah penumpang. Kondisi ini bisa berpotensi menyulitkan perusahaan untuk memenuhi lonjakan permintaan transportasi penumpang pada hari dan jam sibuk dengan kapasitas armada yang dimiliki.



Proposed Solution

Forecasting Trip Daily and Hourly

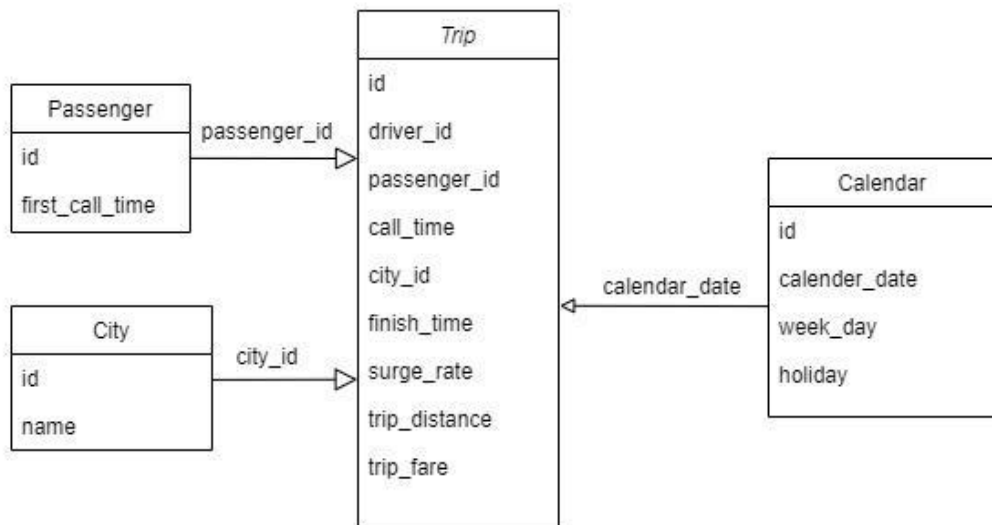
Dengan melakukan analisis data historis dan menggunakan teknik-teknik prediksi yang tepat, perusahaan dapat mengidentifikasi pola-pola dalam permintaan penumpang dan mengantisipasi lonjakan permintaan yang akan terjadi.



Dataset

99 Little Orange

Relation Table:



Description of colom:

- * id: id unik untuk setiap perjalanan
- * driver_id: id pengemudi
- * passenger_id: id penumpang
- * city_id: asal kota perjalanan
- * call_time: waktu pemesanan perjalanan
- * finish_time: waktu perjalanan selesai
- * surge_rate: nilai kenaikan tarif dinamis pada saat peak hours
- * trip_distance: jarak tempuh dalam km
- * trip_fare: biaya/tarif perjalanan
- * first_call_time : waktu/tanggal pemesanan perjalanan pertama dari passenger
- * city: nama kota. Total ada 6 kota pada dataset ini
- * calendar_date: tanggal dalam format MM/DD/YYYY
- * week_day: nama-nama hari dalam seminggu, senin-minggu
- * holiday: apakah tanggal tersebut adalah libur (1) atau tidak (0), total ada 35 hari libur dari periode 2018-01-01 sampai 2020-12-31
- * is_weekend: menentukan apakah akhir pekan (1) atau tidak (0)



Rata-rata jumlah trip per bulan adalah 452.174.6, per hari adalah 15.485, dan per jam adalah 653.



Secara umum, perjalanan memiliki jarak sekitar 2.5 km hingga 8 km. Tarif perjalanan paling banyak berkisar antara 2.6 hingga 7 dolar.



Total Driver 142.987
Total Passenger 1.175.267

**99 Little Orange
Mei-September 2019**



Rata-rata jumlah trip pada akhir pekan lebih tinggi dibandingkan dengan hari kerja.



Kota The Shire memiliki jumlah trip terbanyak. Kota dengan trip tersedikit adalah Kota Gondor.



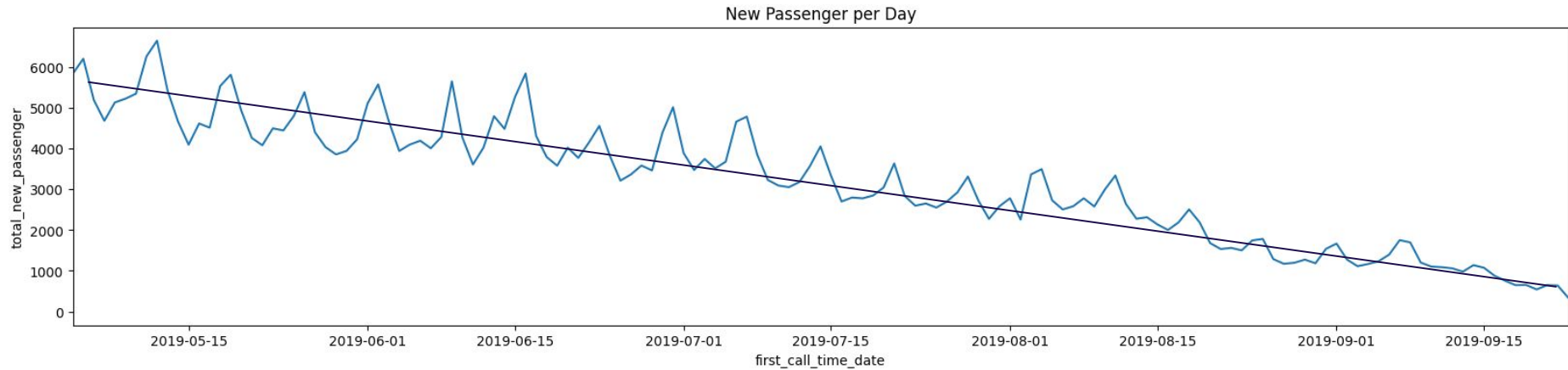
Rata-rata tarif perjalanan lebih mahal pada akhir pekan karena permintaan yang tinggi.



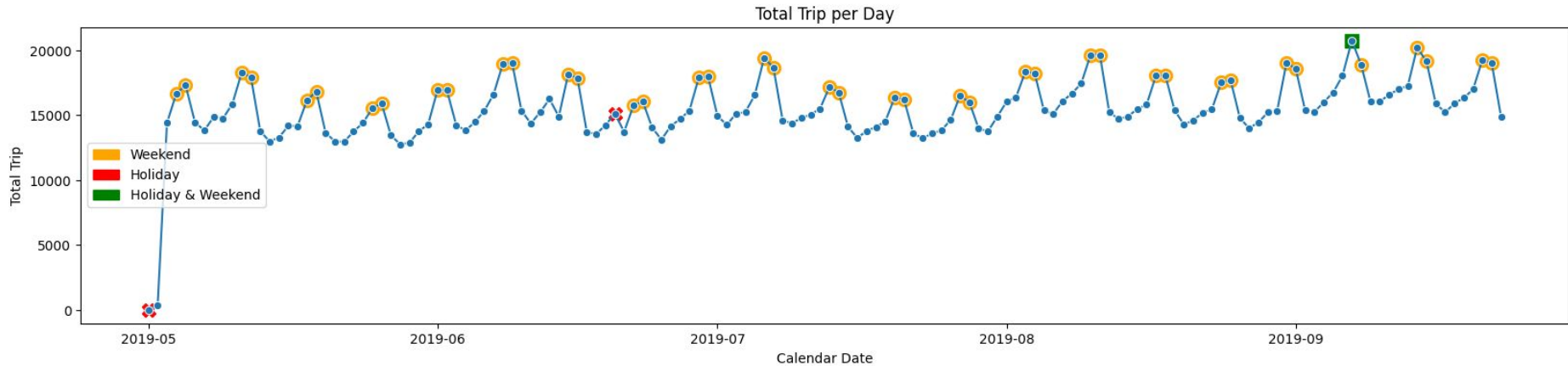
LET'S DIVE DEEPER!

We are going to do Exploratory
Data Analysis from now on.

Pola musiman pada data penumpang baru harian terlihat makin memudar dari waktu ke waktu, serta diiringi tren menurun dalam jangka panjang.



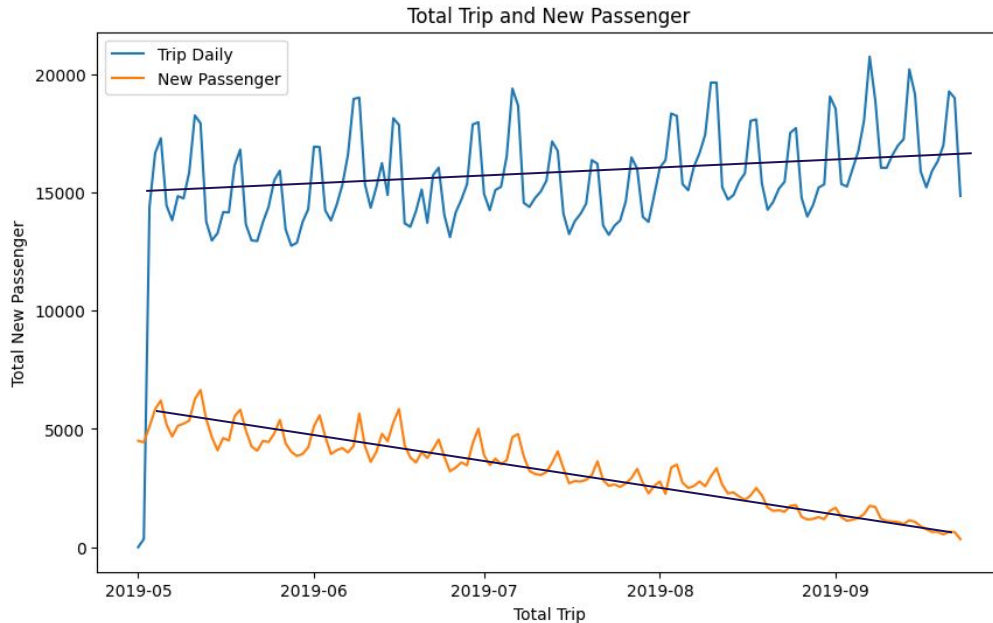
Terdapat pola musiman pada trip harian



Catatan:

- Total trip tertinggi terjadi pada Holiday dan Weekend. (bagian warna hijau).
- Total trip harian tetap stabil tanpa tren naik atau turun yang signifikan. Hal ini menunjukkan adanya konsistensi dalam permintaan layanan ride sharing dari minggu ke minggu.
- Pemodelan deret waktu trip harian akan memperhatikan pola musiman

Penurunan total penumpang baru tidak terlalu mempengaruhi jumlah trip harian



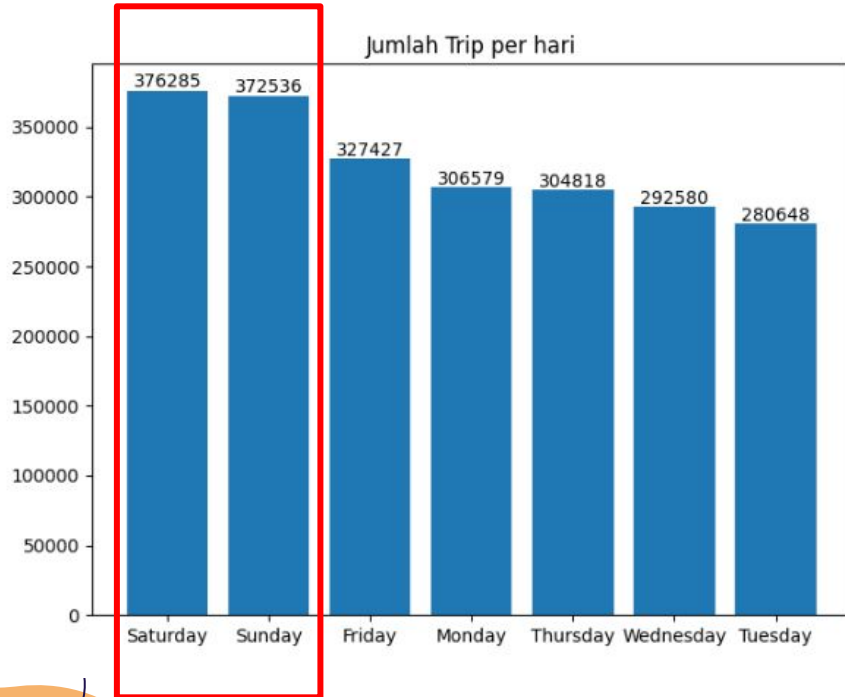
Penyebab:

- Adanya tren penurunan jumlah penumpang baru tanpa mempengaruhi tren total trip menunjukkan bahwa **pasar** mungkin telah mencapai **titik jenuh**.

Rekomendasi:

- Perlunya **perluasan jangkauan** layanan ride sharing ke wilayah baru atau segmentasi pasar yang belum tergarap untuk mengimbangi penurunan jumlah penumpang baru dan meningkatkan pertumbuhan jumlah penumpang baru.
- Selain itu, perusahaan perlu **menjaga reputasi dan citra merek** agar tidak terjadi penurunan jumlah trip.

Jumlah trip paling tinggi terjadi pada hari Sabtu dan Minggu daripada hari-hari lain.



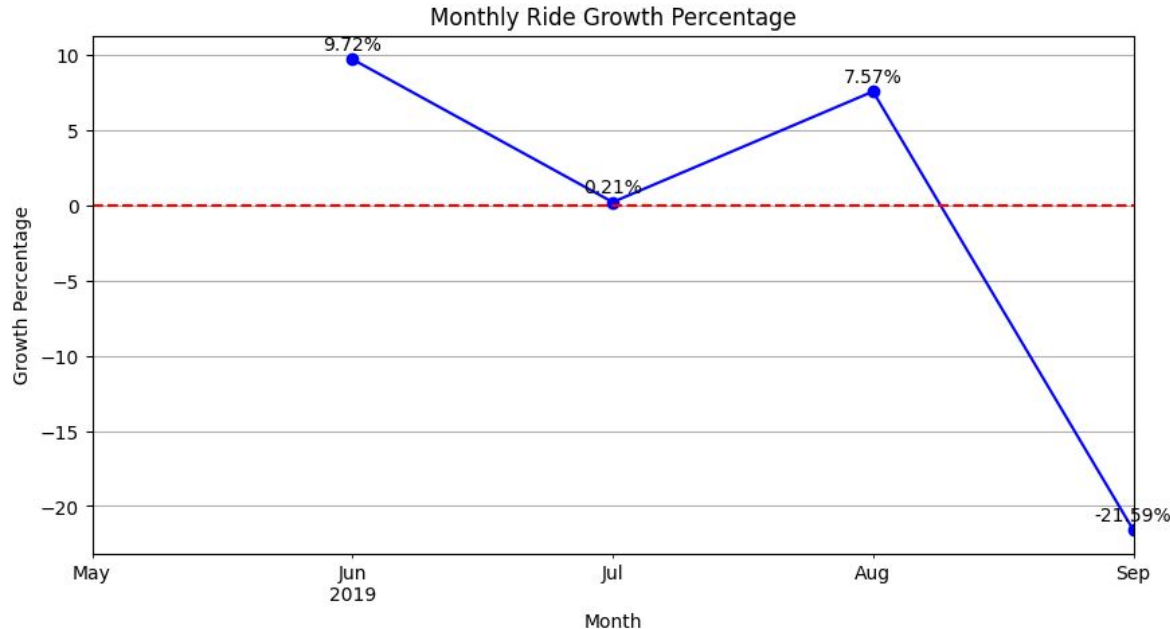
Penyebab:

- Hal ini bisa terjadi aktivitas karena libur & waktu luang lebih banyak. Sehingga banyak orang beraktivitas diluar rumah untuk menikmati akhir pekan. Hal lain yang juga bisa mempengaruhi adalah kemungkinan seringnya ada event khusus pada akhir pekan sehingga banyak orang datang dari luar kota untuk datang ke event tersebut.

Rekomendasi:

- Perlu optimasi suplai driver lebih besar di Sabtu & Minggu untuk mengakomodasi lonjakan permintaan.

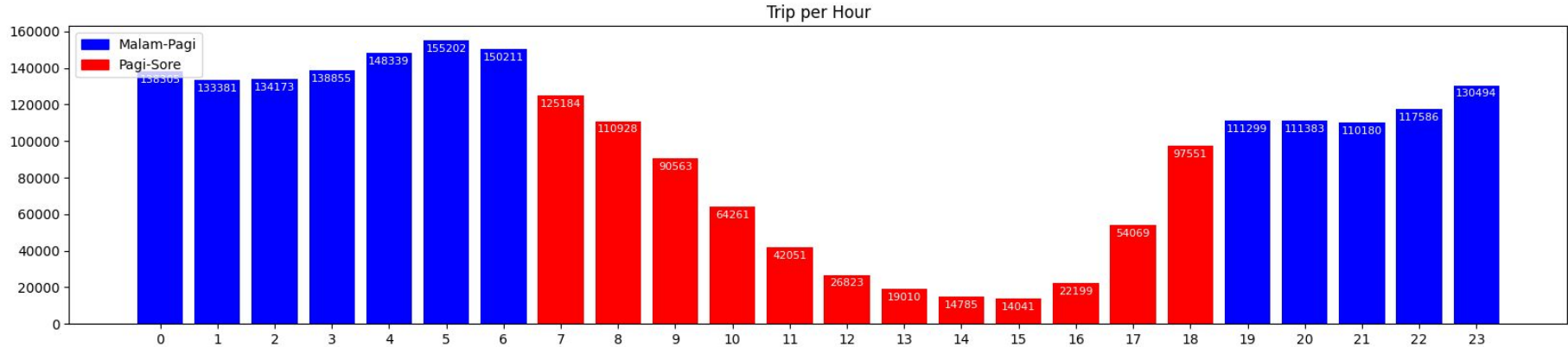
Terjadi pertumbuhan positif jumlah trip pada bulan Juni hingga Agustus. Hal ini menunjukkan peningkatan terjadinya jumlah perjalanan/trip dibandingkan bulan-bulan sebelumnya.



Catatan:

Penurunan signifikan pada bulan September disebabkan karena data ini diambil hanya sampai tanggal 23 September. Walaupun belum lengkap tren pertumbuhan negatif ini cukup mencolok. Bisa jadi akibat faktor musiman akhir tahun atau dampak promosi/insentif bulan lalu yang sudah berakhir.

Jumlah trip lebih banyak terjadi pada malam hingga dini hari dibandingkan dari pagi sampai sore hari.



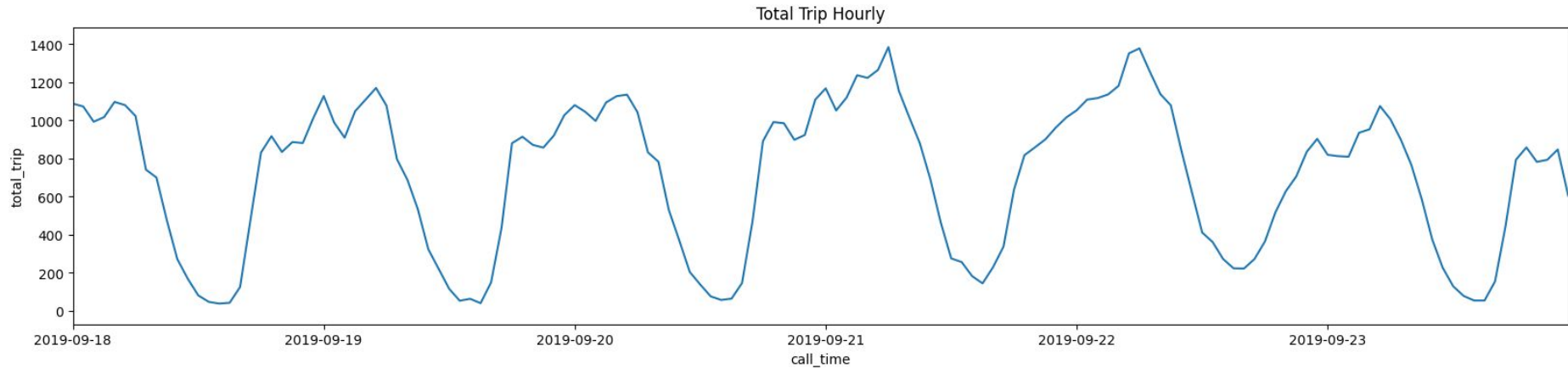
Penyebab:

- Jumlah trip yang banyak dari malam hingga dini hari bisa disebabkan oleh banyak hal seperti kepulangan bekerja, adanya aktivitas malam, serta jam berangkat kerja pada pagi harinya.
- Penurunan trip setelah jam 5 karena sebagian besar orang sudah tiba di tujuan. Trip menurun drastis di siang hari karena orang-orang sedang bekerja dan serta waktu istirahat yang minim akan perpindahan.

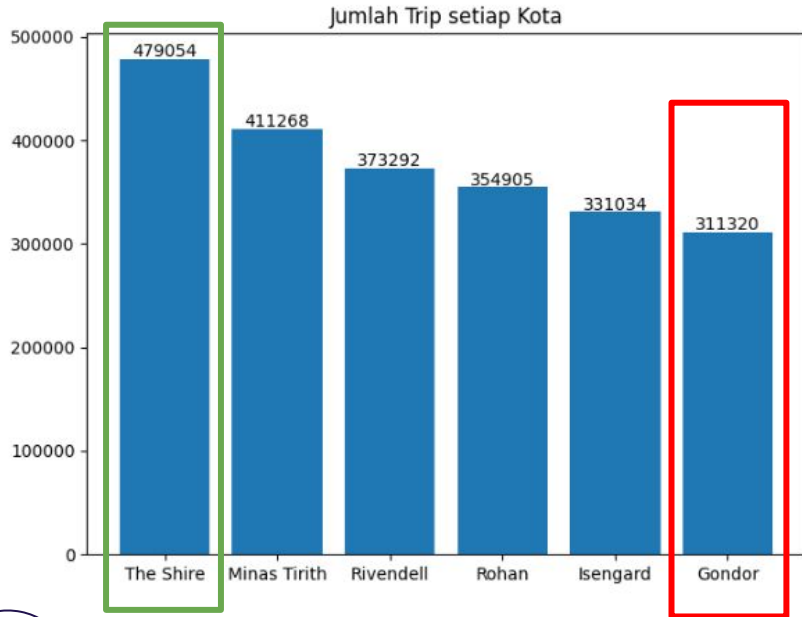
Rekomendasi:

- Dari hal tersebut hal yang bisa dilakukan adalah meningkatkan suplai driver pada jam-jam sibuk tersebut dengan memberikan insentif khusus untuk mendorong lebih banyak driver untuk melayani antar-jemput.

Terdapat pola musiman pada trip hourly, sehingga pemodelan deret waktu untuk total trip hourly akan memperhatikan pola musiman



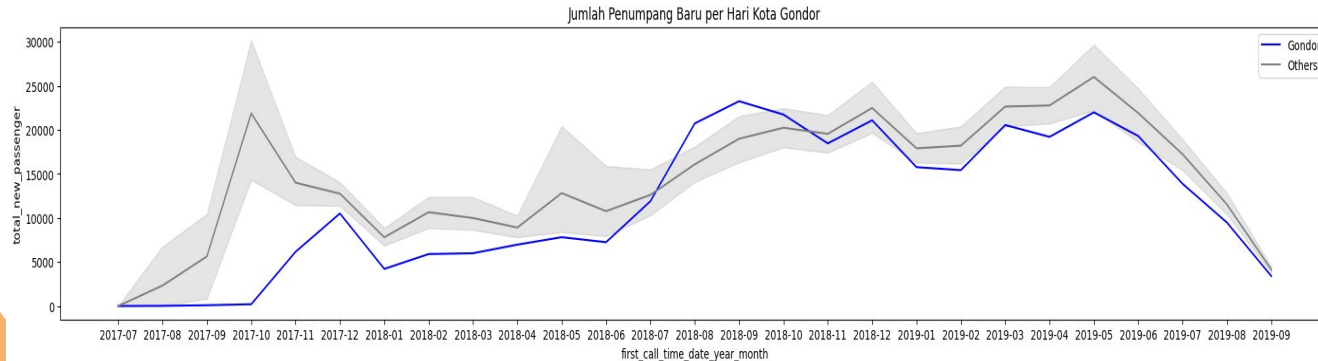
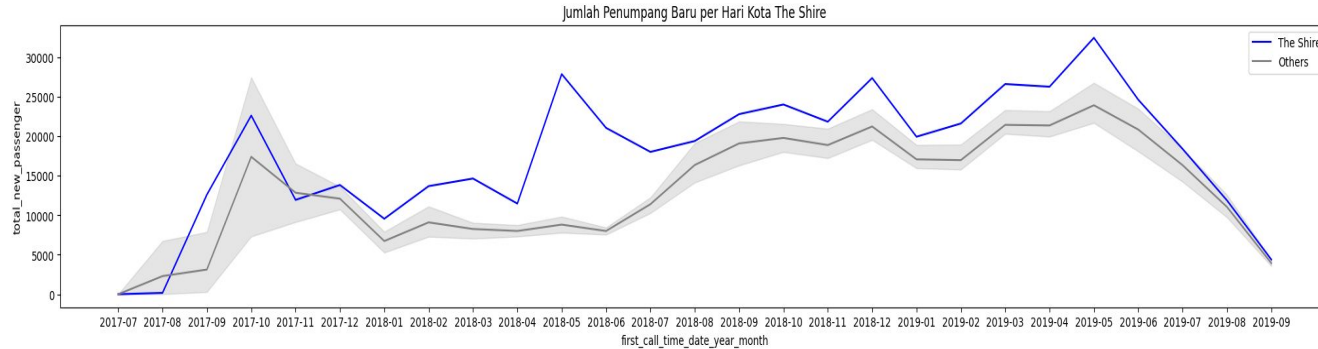
Kota **The Shire** adalah kota dengan jumlah trip terbanyak, sedangkan **Kota Gondor** adalah kota dengan jumlah trip paling sedikit



Rekomendasi:

- Menjaga kualitas layanan di kota The Shire karena basis pengguna terbesar disana
- Perlu memperluas jangkauan layanan trip pada Kota Gondor.

Jumlah penumpang baru harian Kota The Shire hampir selalu diatas kota-kota yang lain. Sedangkan pada Kota Gondor jumlah penumpang barunya hampir selalu dibawah kota-kota lain.

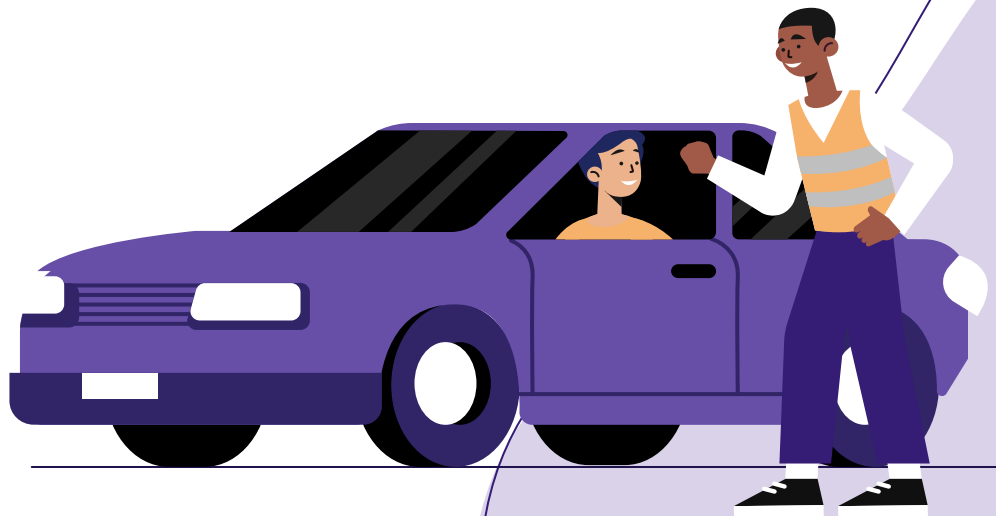


Rekomendasi:

- Perlunya strategi akuisisi penumpang baru yang lebih agresif pada Kota Gondor
- Fokus pada pelayanan pelanggan dan kualitas layanan di The Shire untuk mempertahankan basis pengguna yang besar.

Time Series Modeling

We want to forecast **total trip daily and hourly**



Pre-Processing Steps

1. Aggrasi Data

Karena tidak ada kolom total trip harian dan per jam, maka dilakukan agregasi data berdasarkan tanggal dan waktu (jam).

2. Konversi Indeks ke Deret Waktu

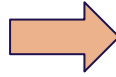
Konversi Indeks ke Deret Waktu adalah proses mengubah indeks sebuah data frame menjadi format waktu untuk analisis data berbasis waktu.

3. Split Data

Membagi data menjadi 80% data train dan 20% data test

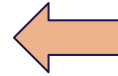
Model Used

1. SARIMA



SARIMA atau Seasonal ARIMA adalah model pada analisis deret waktu (time series) yang secara khusus digunakan untuk menangani pola musiman (seasonal) yang terdapat pada data. Sebagai model lanjutan dari ARIMA, SARIMA mampu menangkap pengaruh musiman seperti faktor cuaca, libur panjang, atau pola aktivitas tahunan/mingguan/harian dalam melakukan peramalan data.

SARIMAX merupakan pengembangan dari model SARIMA dengan menambahkan prediktor eksternal untuk meningkatkan akurasi analisis. Selain menangkap faktor musiman dari data deret waktu, SARIMAX juga memasukkan variabel independen seperti indikator libur nasional, event khusus, atau faktor spesifik lain yang turut mempengaruhi perilaku data dari waktu ke waktu.



2. SARIMAX



Stationarity Check with Dickey-Fuller Test

Hipotesis:

Hipotesis Nol (H_0): Deret waktu tidak stasioner.

Hipotesis Alternatif (H_1): Deret waktu stasioner.

Kriteria uji: tolak H_0 jika p-value lebih kecil daripada α (0.05)

Stationarity Check

alpha = 0.05

Daily Trip Time Series Check

Results of Dickey-Fuller Test:

Test Statistic	-3.682207
p-value	0.004370

Berhasil menolak H_0 karena **p-value < alpha**.
Deret waktu trip per hari adalah deret waktu yang **stasioner**.

Hourly Trip Time Series Check

Results of Dickey-Fuller Test:

Test Statistic	-6.716846e+00
p-value	3.564151e-09

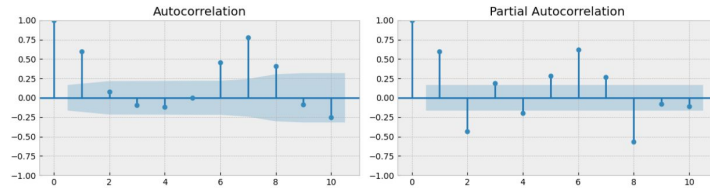
Berhasil menolak H_0 karena **p-value < alpha**.
Deret waktu trip per jam adalah deret waktu yang **stasioner**.

Tidak perlu dilakukan differencing atau transformasi data!

ACF & PACF Check

Daily Trip Time Series Check

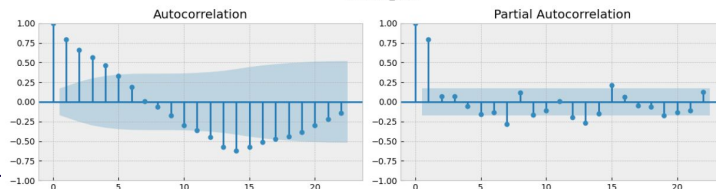
Non Seasonal



Orde Non Seasonal

$p : [0,1,2]$ (Orde AR) ; $d : 0$; $q : [0,1]$ (Orde MA)

Seasonal

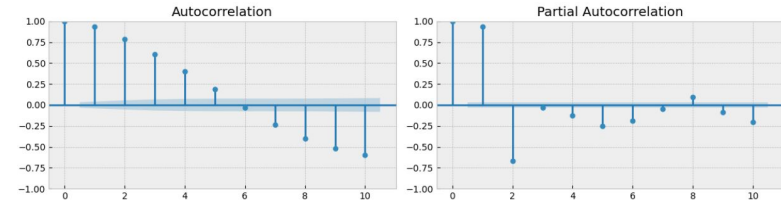


Orde Seasonal

$P = [0,1]$ (Orde AR Seasonal); $D = 0$; $Q = [0,1,2,3]$ (Orde MA Seasonal) ; $s = 7$ (7 hari)

Hourly Trip Time Series Check

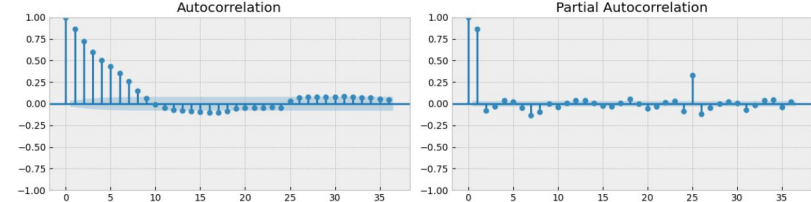
Non Seasonal



Orde Non Seasonal

$p : [0,1,2]$ (Orde AR) ; $d : 0$; $q : [0,1,2]$ (Orde MA)

Seasonal



Orde Seasonal

$P = [0,1,2]$ (Orde AR Seasonal); $D = 0$; $Q = [0,1,2]$ (Orde MA Seasonal) ; $s = 24$ (24 jam)

Tuned & Fit Model

Tuning model dengan konsep Grid Search, yaitu dengan mencoba semua kombinasi orde dari model SARIMA dan SARIMAX

Total kombinasi model untuk deret waktu Daily Trip: 48 kombinasi
Total kombinasi model untuk deret waktu Hourly Trip: 81 kombinasi

Model Selection

Daily Trip Time Series Check

SARIMA best orde model: SARIMA(1,0,0)(1,0,2)7

SARIMAX best orde model: SARIMAX(1,0,0)(0,0,0)7

Evaluasi kedua model:

Model	RMSE Train	RMSE Test	MAPE Train	MAPE Test
SARIMA(1,0,0)(1,0,2)7	1513.54	548.75	0.03	0.02
SARIMAX(1,0,0)(0,0,0)7	1527.46	677.92	0.04	0.03

Model terbaik untuk daily trip adalah model SARIMA(1,0,0)(1,0,2)7

Hourly Trip Time Series Check

SARIMA best orde model: SARIMA(1,0,2)(1,0,2)24

SARIMAX best orde model: SARIMAX(1,0,0)(1,0,2)24

Evaluasi kedua model:

Model	RMSE Train	RMSE Test	MAPE Train	MAPE Test
SARIMA(1,0,2)(1,0,2)24	59.31	58.45	0.093	0.086
SARIMAX(1,0,0)(1,0,2)24	60.56	62.52	0.097	0.092

Model terbaik untuk hourly trip adalah model SARIMA(1,0,2)(1,0,2)24

Interpretasi Evaluasi Model

Daily Trip Model

- MAPE 2% sangat bagus untuk forecast di waktu kedepan, karena kesalahannya sangat kecil.
- RMSE 548 sangat kecil dibanding range data 15-20rb.

Hourly Trip Model

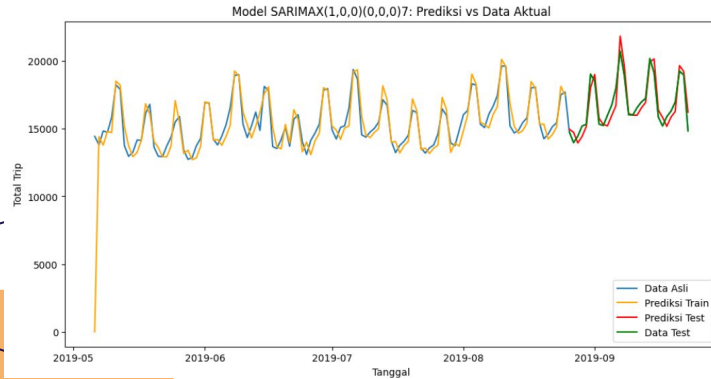
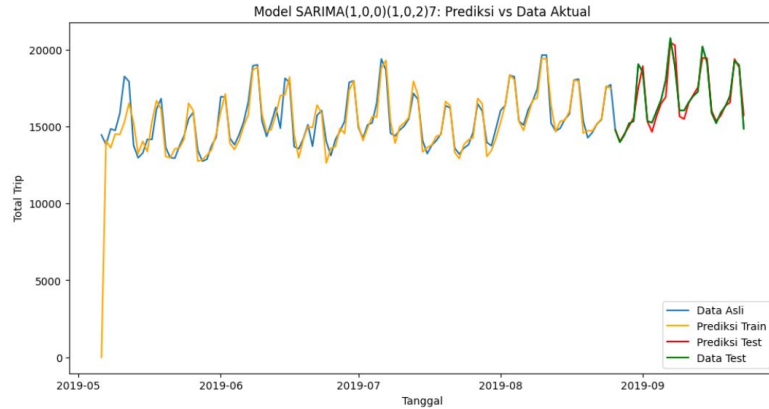
- MAPE 8% masih acceptable untuk peramalan data jam-per jam.
- RMSE 58 juga kecil dibanding range data 200-1400.

Model Daily Trip sangat akurat untuk meramalkan pergerakan harian, sehingga dapat digunakan untuk perencanaan operasional dan pengambilan keputusan jangka pendek.

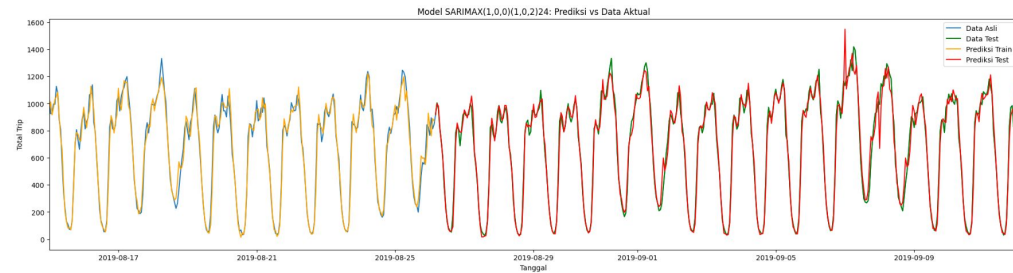
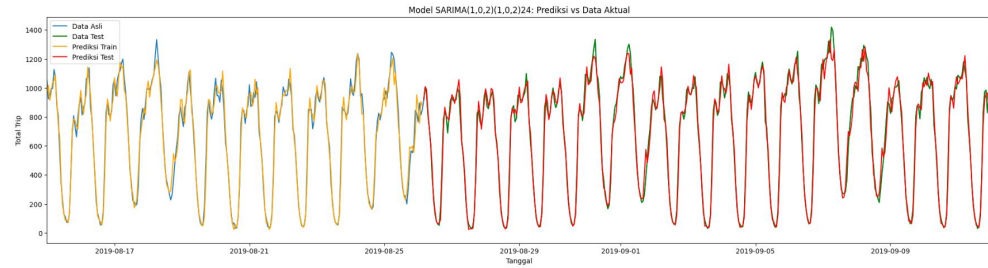
Model Hourly Trip cukup akurat untuk memprediksi pola pergerakan jam per jam, sehingga berguna bagi penjadwalan armada dan alokasi sumber daya.

Model Result

Daily Trip Time Series Check



Hourly Trip Time Series Check



Implementasi Model

- Menggunakan model untuk melakukan forecasting harian dan memberikan notifikasi kepada driver tentang lonjakan permintaan di hari esok.
- Memanfaatkan model untuk memprediksi jam-jam dengan permintaan tinggi dan memberikan notifikasi kepada driver.
- Ketika hasil prediksi menunjukkan lonjakan demand di waktu tertentu, sistem akan otomatis mengaktifkan surge rate untuk periode waktu tersebut guna meningkatkan suplai driver.
- Memantau perbedaan antara permintaan aktual dan hasil prediksi bertujuan untuk mengidentifikasi perubahan pola jumlah trip dengan cepat dan meminimalisir kesalahan prediksi.
- Mengadakan challenge berhadiah kepada driver untuk memenuhi permintaan pada jam-jam tinggi.

Rangkuman Rekomendasi

- Perluasan jangkauan layanan ride-sharing ke wilayah baru atau segmentasi pasar baru. Hal ini disebabkan peningkatan jumlah penumpang baru di area saat ini menunjukkan pasar telah mencapai titik jenuh.
- Diperlukan peningkatan suplai driver pada jam-jam sibuk dan akhir pekan guna menanggapi lonjakan penumpang pada periode tersebut.
- Fokus pada pelayanan pelanggan dan kualitas layanan di The Shire untuk mempertahankan basis pengguna yang besar.
- Pentingnya menerapkan strategi akuisisi penumpang baru yang lebih agresif di Kota Gondor.
- Memaksimalkan strategi penggunaan model forecasting total trip harian dan per jam.

THANKS

My github:

<https://github.com/hubble99>

My email:

ramadhoninasri09@gmail.com

My linkedin:

<https://www.linkedin.com/in/ramadhoni-na-sri-4b514b220/>

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, infographics & images by **Freepik**

Please keep this slide for attribution

