



Analisis Dataset Bike Sharing *(2011 - 2012)*

Syifa Gusriani Rohman

December 2024



Latar Belakang



Proyek ini merupakan salah satu tugas pelatihan dicoding berjudul “Belajar Analisis Data dengan Python” menggunakan dataset publik. Meskipun ini merupakan proyek simulasi, pendekatan analisis dirancang menyerupai studi kasus bisnis nyata untuk melatih kemampuan eksplorasi data dan pengambilan keputusan berbasis data.



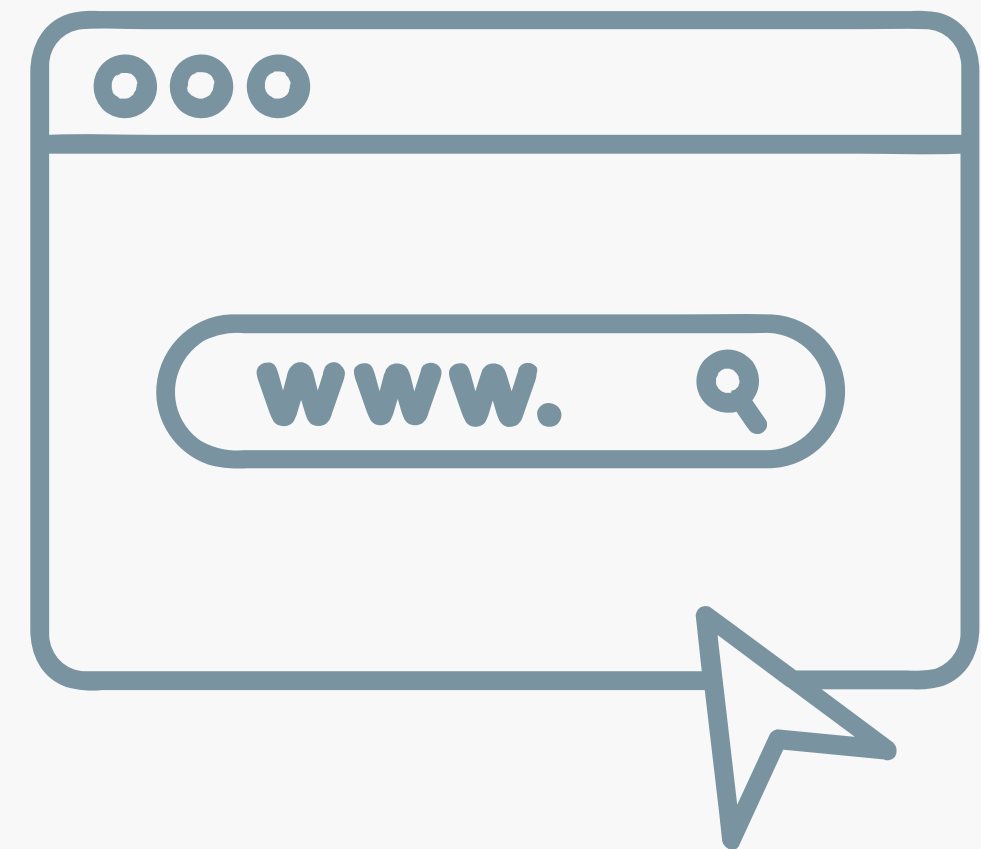
Tujuan Proyek

- Melatih kemampuan dalam mengolah dan mengeksplorasi data real-world
- Mengasah keterampilan dalam menjawab pertanyaan bisnis melalui pendekatan eksplorasi data
- Mendapatkan insight dengan beberapa contoh pendekatan, seperti mengidentifikasi pola penggunaan berdasarkan beberapa faktor (cuaca, waktu dan musim), mengetahui perbedaan perilaku pengguna kasual dan registered (terdaftar), dan lain sebagainya.

Dataset yang Digunakan



- Dataset : Bike Sharing Dataset
- Periode : 2011 - 2012
- Jumlah Data :
 - bs_day : 731 baris
 - bs_hour : 17379 baris
- Sumber : Kaggle ([Link](#))



Dataset yang Digunakan (penjelasan kolom)

Kolom	Penjelasan	Kolom	Penjelasan	Kolom	Penjelasan
season	musim (1:semi, 2:panas, 3:gugur, 4:dingin)	holiday	hari libur, sumber: Data	instant	indeks
		workingday	1:bukan akhir pekan/ libur. 0: selain 1	dteday	tanggal
yr	tahun (0:2011, 1:2012)	weathersit	1: Cerah sd 4:Hujan deras	temp	suhu aktual
mnth	bulan			atemp	suhu terasa
hr	jam	registered	pengguna terdaftar	hum	kelembapan
weekday	hari dalam seminggu	casual	pengguna tidak terdaftar	windspeed	kecepatan angin
				cnt	total penyewaan

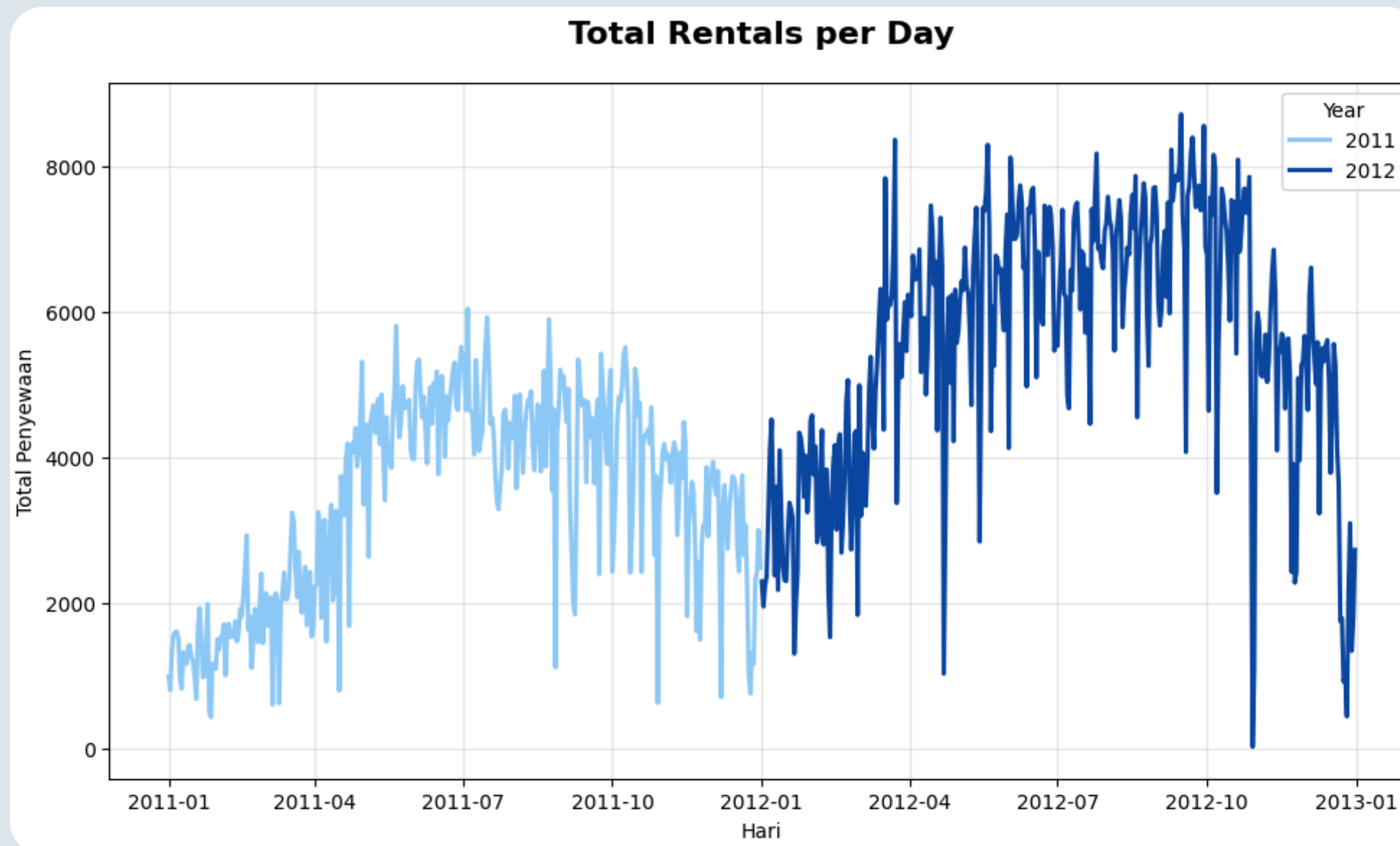
Pertanyaan Bisnis



1. Bagaimana pengaruh suhu, kelembaban , dan kecepatan angin memengaruhi jumlah penyewaan sepeda?
2. Apakah terdapat perbedaan jumlah penyewaan sepeda berdasarkan musim (semi, panas, gugur, dingin)?
3. Bagaimana perbedaan pola penggunaan antara pengguna casual dan registered?
4. Apakah distribusi penyewaan sepeda merata sepanjang tahun atau ada bulan-bulan tertentu dengan permintaan rendah?
5. Kapan waktu paling sering dan jarang terjadi transaksi penyewaan sepeda?

Analisis - 1

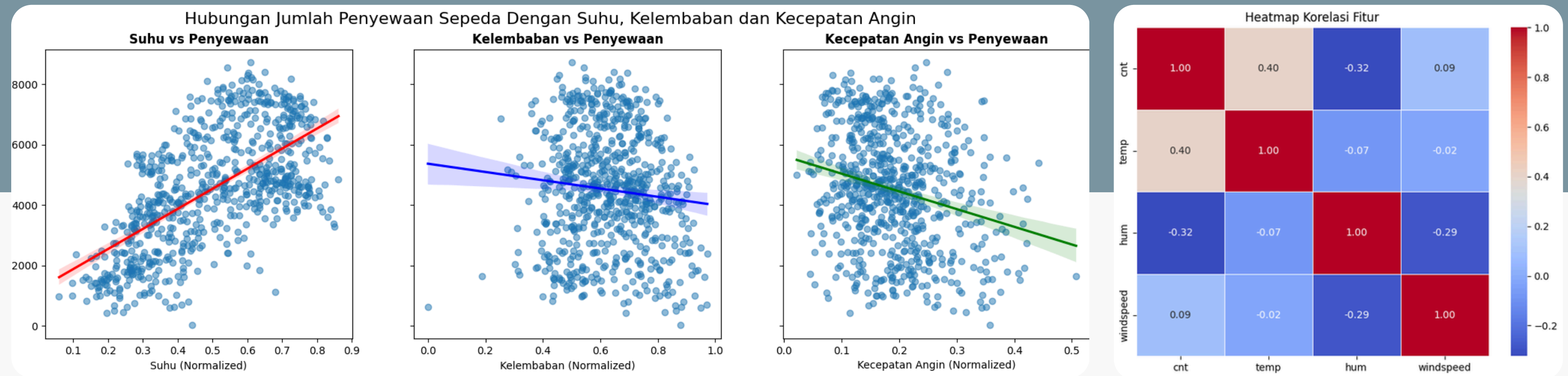
Berdasarkan hasil analisis dataset, **total penyewaan sepeda dari tahun 2011 hingga 2012 mencapai 3.292.679 penyewaan**. Rinciannya, pada **tahun 2011 terdapat 1.243.103 penyewaan**, sedangkan pada **tahun 2012 jumlahnya meningkat menjadi 2.049.576 penyewaan**. Dari data tersebut, terlihat adanya kenaikan sebesar **64,88%** dari tahun 2011 ke tahun 2012.



Grafik di samping menunjukkan total penyewaan yang terjadi selama dua tahun tersebut. Terlihat jelas bahwa jumlah penyewaan pada tahun 2012 mengalami peningkatan dibandingkan tahun sebelumnya.

Analisis - 2 (Pertanyaan 1)

Bagaimana pengaruh suhu, kelembaban, dan kecepatan angin memengaruhi jumlah penyewaan sepeda?



- **Suhu:** Scatter plot menunjukkan pola menaik ke kanan, menandakan **jumlah penyewaan sepeda meningkat seiring suhu yang lebih hangat**. Hal ini diperkuat dengan korelasi cenderung menengah (0.4).

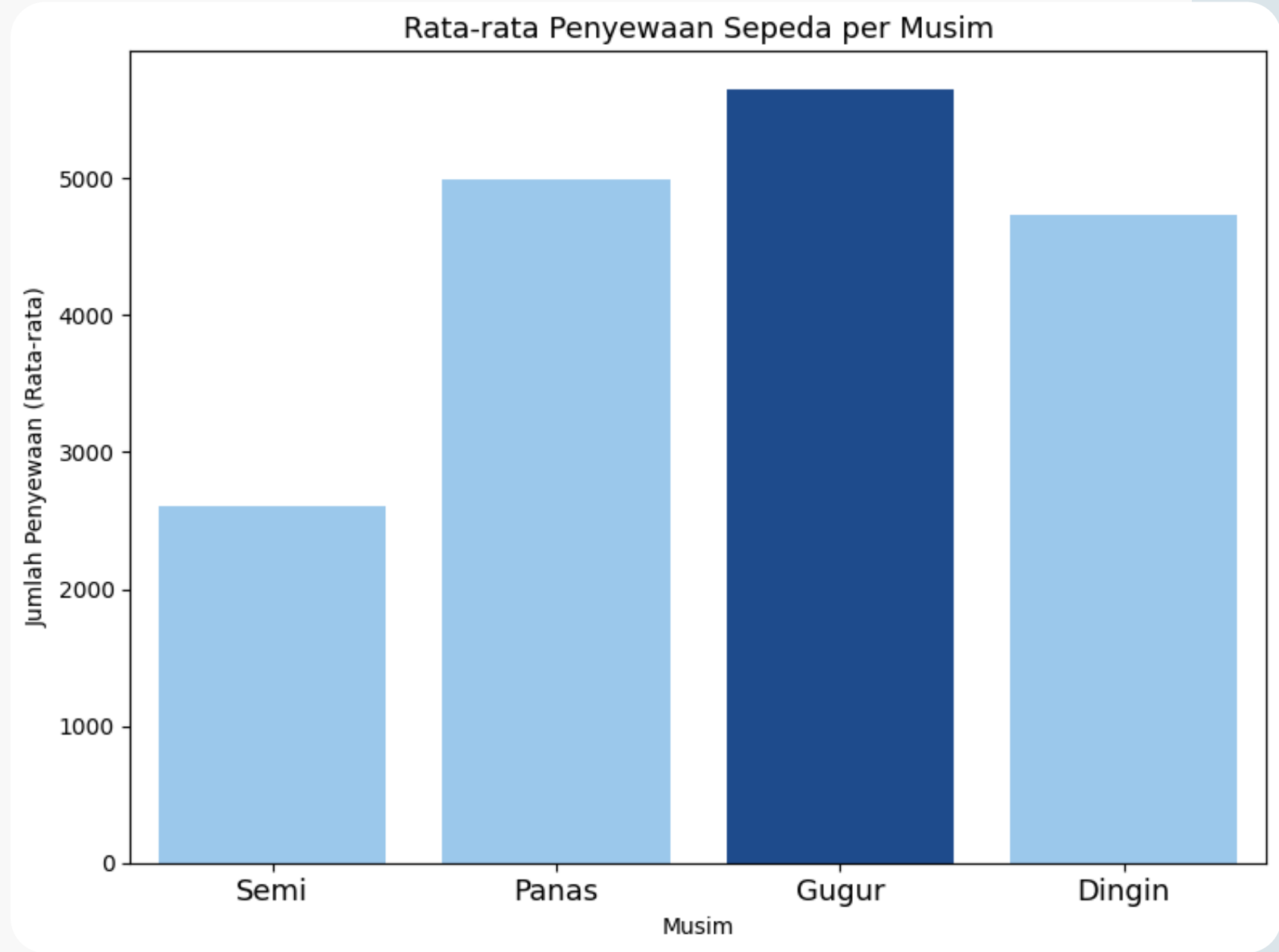
Analisis - 2 (Pertanyaan 1)

Bagaimana pengaruh suhu, kelembaban, dan kecepatan angin memengaruhi jumlah penyewaan sepeda?

- **Kelembaban:** Pola terlihat acak, menunjukkan **kelembaban tidak terlalu berpengaruh**, meskipun ada **sedikit kecenderungan penyewaan meningkat pada kelembaban tinggi**. Tapi pada tabel korelasi, **cenderung berkebalikan (- 0.32) walaupun kecil**.
- **Kecepatan Angin:** Pola juga acak, menunjukkan **kecepatan angin kurang berpengaruh**, namun penyewaan **cenderung lebih banyak saat angin berkecepatan rendah**. Ini juga menunjukkan **kecenderungan naik berdasarkan korelasi, tapi sangat kecil (0.09)**.

Analisis - 3 (Pertanyaan 2)

Apakah terdapat perbedaan jumlah penyewaan sepeda berdasarkan musim (semi, panas, gugur, dingin)?

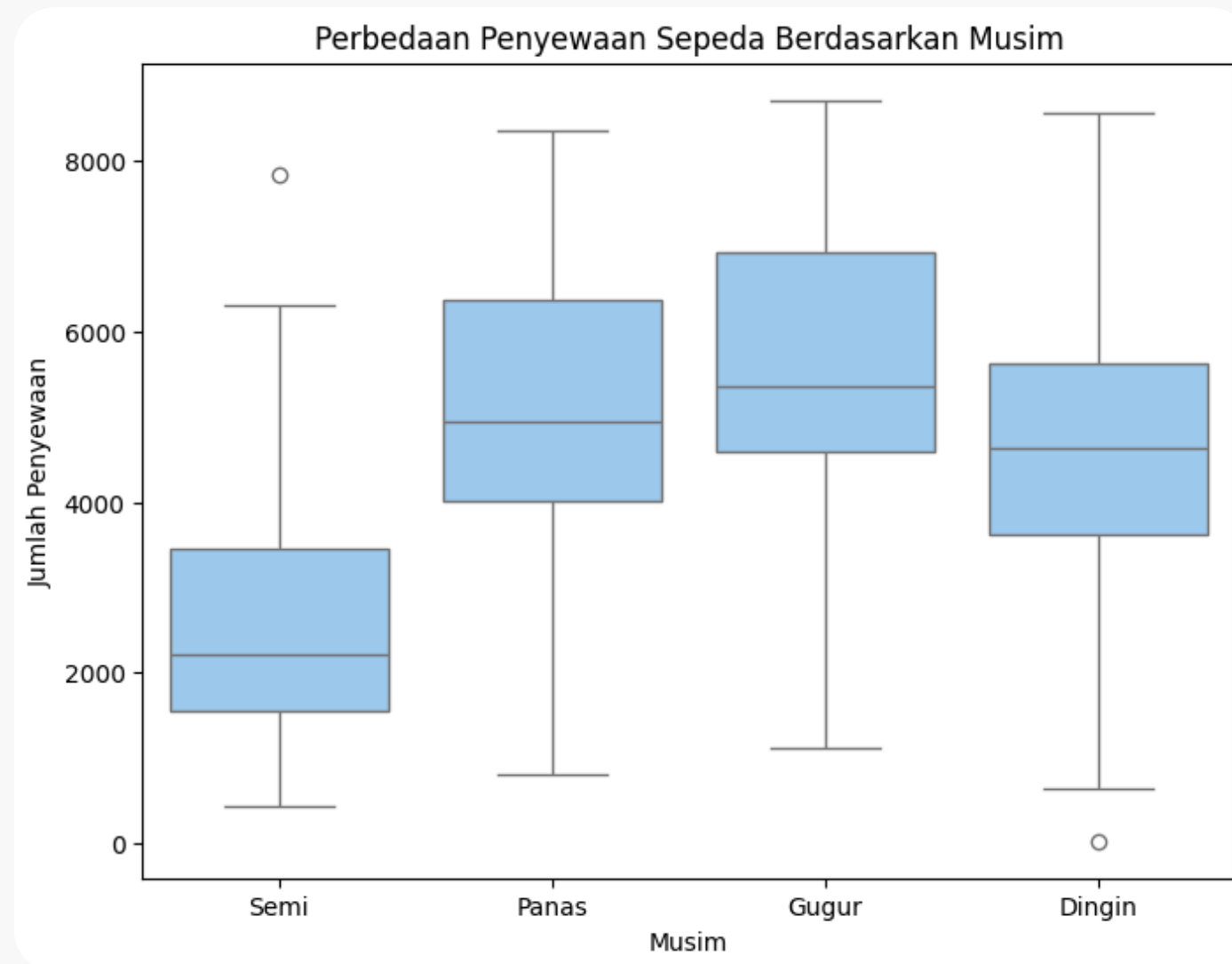


Tabel rata-rata suhu per-musim

Musim	Semi	Panas	Gugur	Dingin
Suhu (°C)	12.21	22.32	28.96	17.34

Berdasarkan grafik disamping, **musim gugur menjadi musim dengan penyewaan terbanyak**. Mengacu pada jawaban nomor 2, bahwa **lebih banyak penyewaan terjadi pada suhu yang lebih hangat**, disini juga terlihat bahwa **rata - rata suhu di musim gugur adalah yang terhangat dibandingkan musim lain**.

Analisis - 3 (Pertanyaan 2) revisi nnggu



Penjelasan Grafik:

- Garis dalam kotak: jumlah penyewaan yang umum.
- Garis miring ke atas/bawah: sebaran tidak merata.
- Kotak condong ke bawah: banyak hari sepi.
- Titik di luar: hari tidak biasa (sangat ramai/sepi).

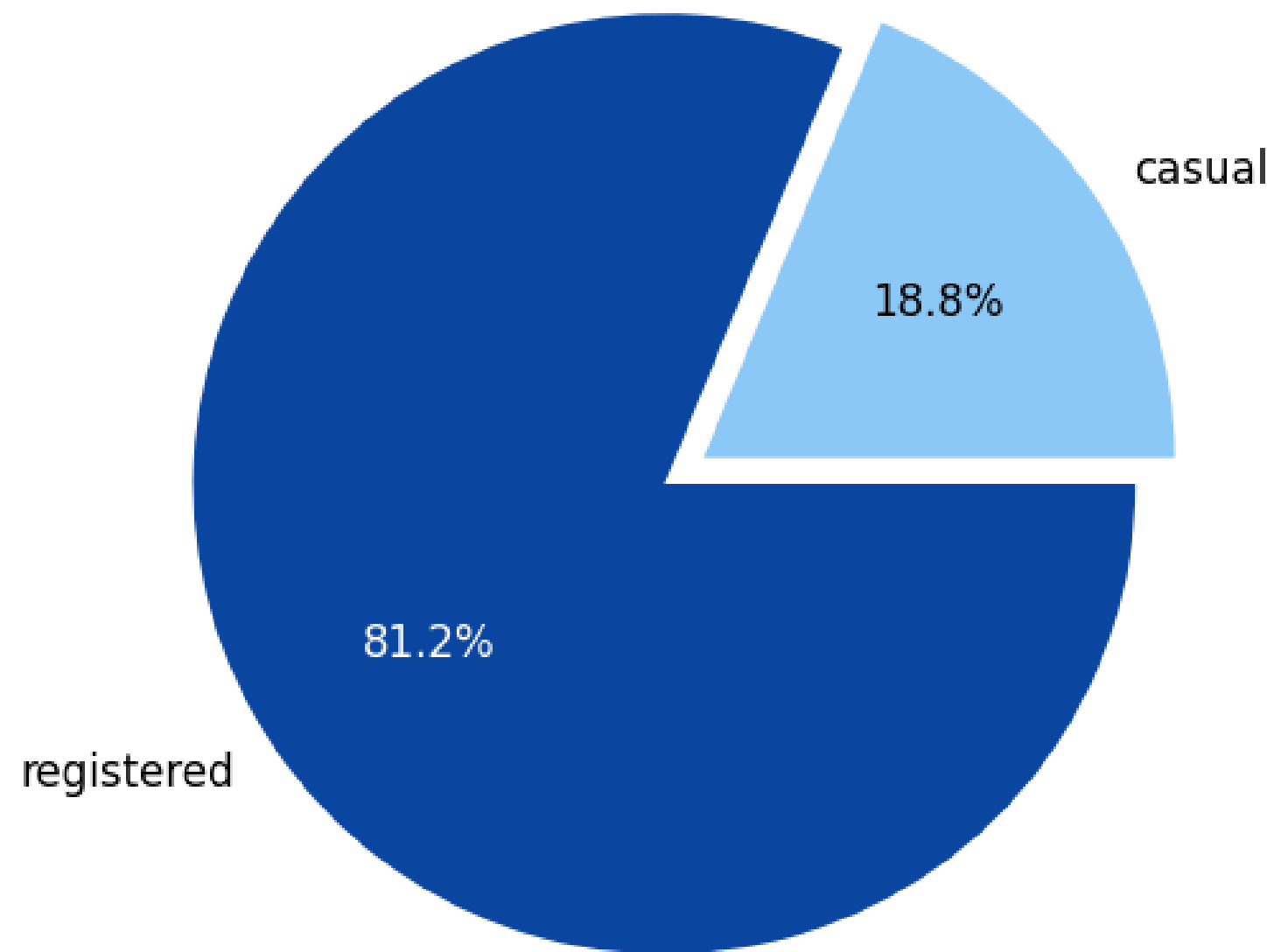
Grafik ini memperlihatkan pola penyewaan sepeda di tiap musim.

- Musim Semi: Sebagian besar hari memiliki jumlah penyewaan yang rendah, tapi ada beberapa hari dengan lonjakan sangat tinggi. Ini menunjukkan penggunaan sepeda masih tidak menentu di musim ini. Rata-rata penyewaan: 2604.13.
- Musim Panas: Jumlah penyewaan biasanya tinggi, tapi tetap bervariasi antar harinya. Banyak orang menyewa sepeda, tapi tidak selalu konsisten setiap hari. Rata-rata penyewaan: 4992.33.
- Musim Gugur: Ini musim dengan jumlah penyewaan tertinggi secara umum. Meskipun jumlahnya besar, selisih antar hari masih cukup terasa, tidak sepenuhnya rata. Rata-rata penyewaan: 5644.30.
- Dingin: Jumlah penyewaan cenderung menurun. Beberapa hari sangat sepi, tapi ada juga yang cukup ramai. Jadi, meskipun rata-rata lebih rendah, selisih antar harinya tetap ada. Rata-rata penyewaan: 4728.16.

Analisis - (Pertanyaan 3)

Bagaimana perbedaan pola penggunaan antara pengguna casual dan registered?

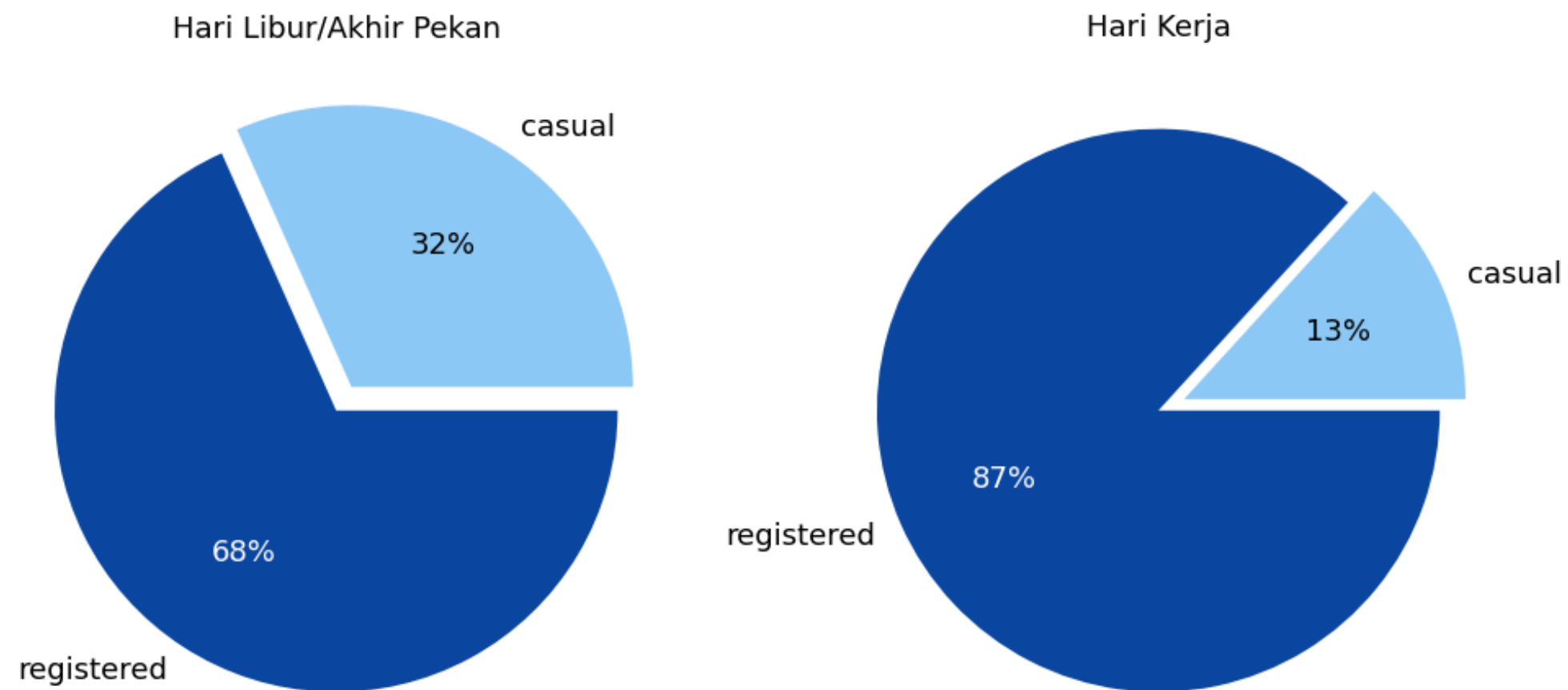
Presentase Total Pengguna Casual dan Registered



Berdasarkan pie chart disamping, terlihat bahwa **presentase pelanggan terdaftar (registered) jauh lebih banyak dibandingkan dengan pelanggan casual.**

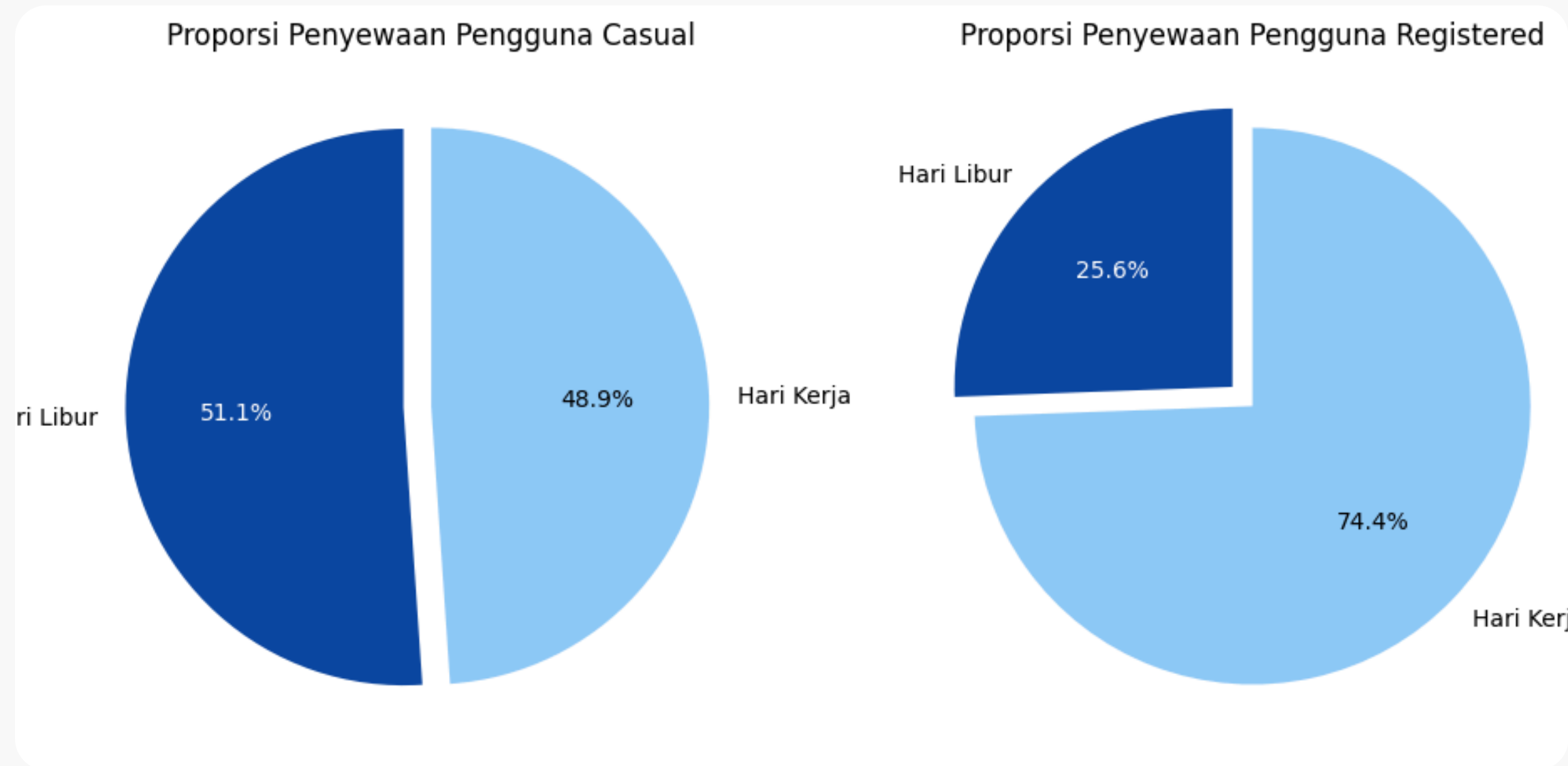
Analisis - (Pertanyaan 3)

Presentase Pengguna Casual dan Registered



Berdasarkan pie chart disamping, total pengguna casual sedikit lebih banyak di hari libur dibandingkan pada hari kerja, namun tetap pengguna registered yang mendominasi.

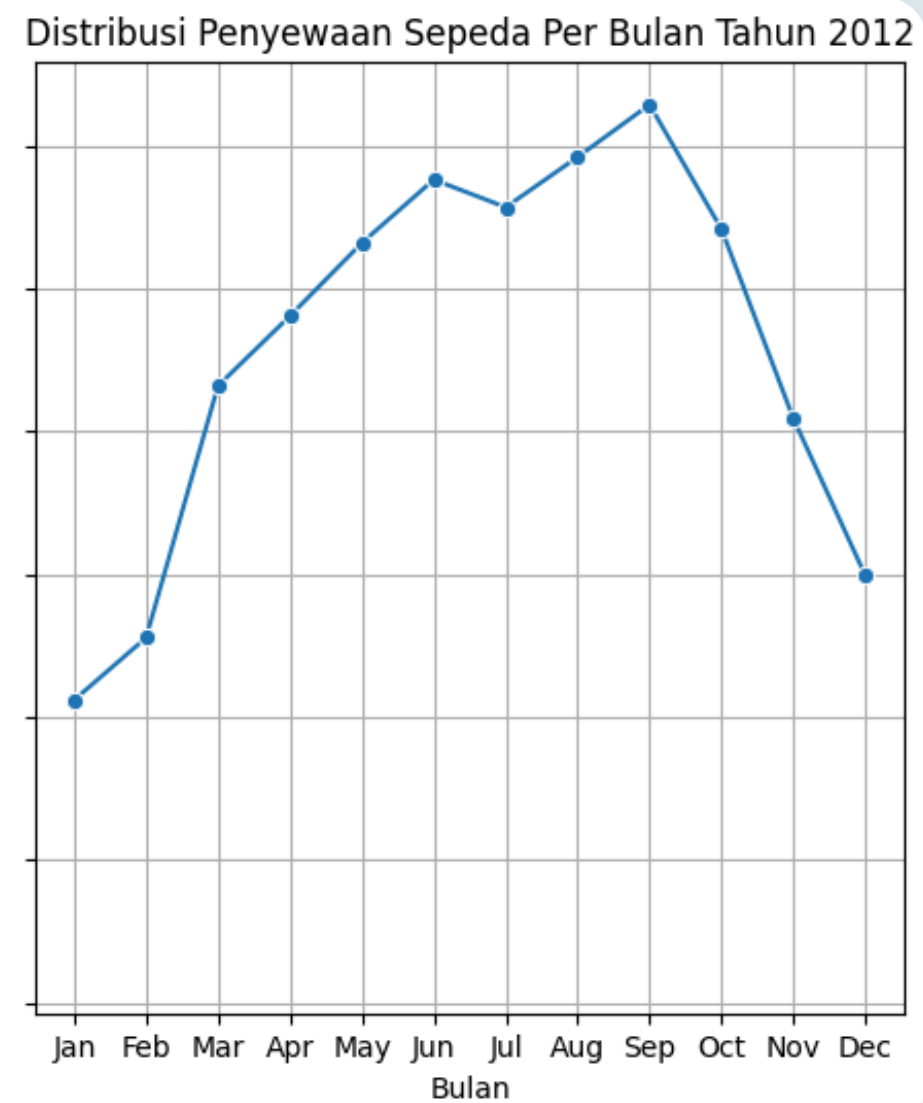
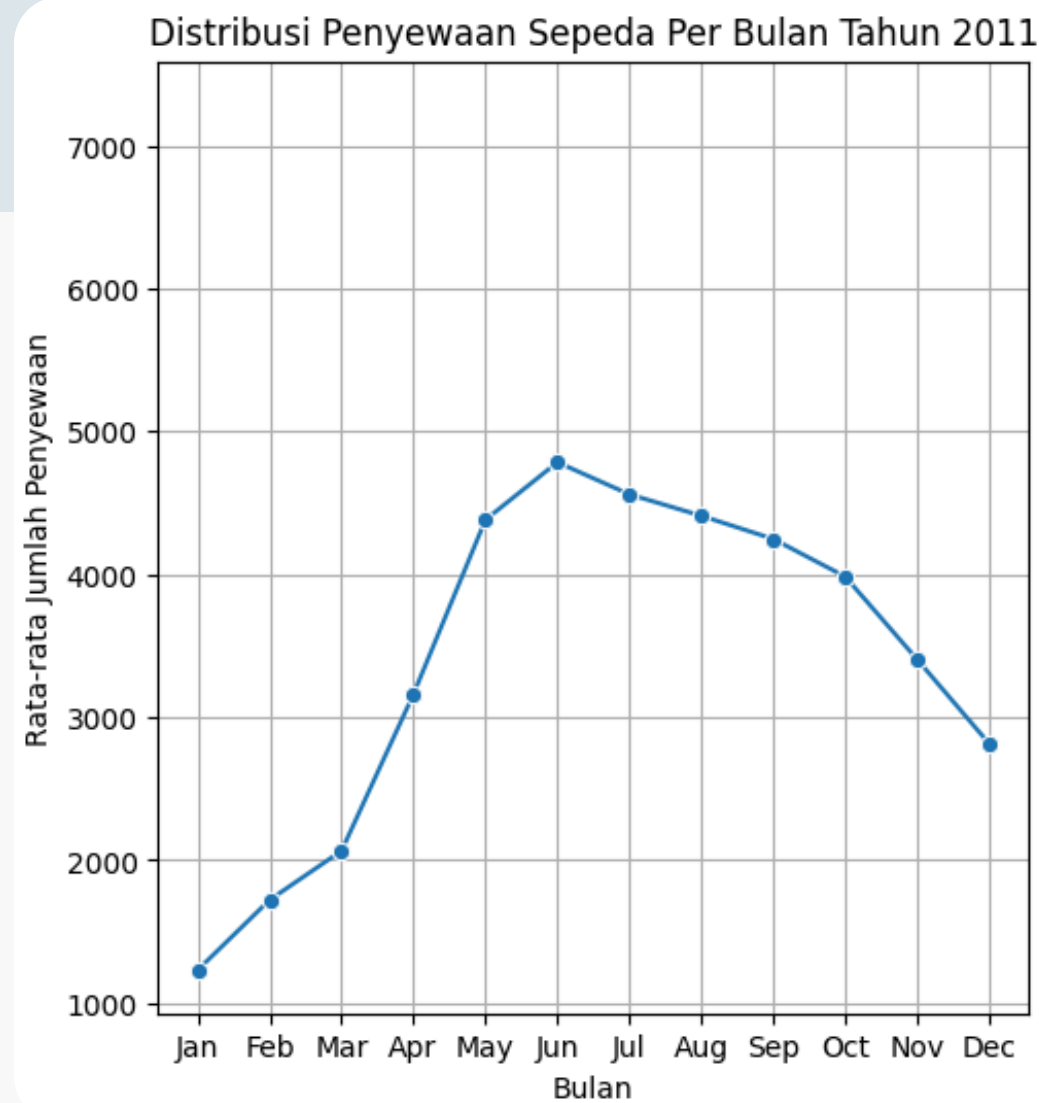
Analisis - (Pertanyaan 3)



Berdasarkan pie chart disamping, pengguna casual lebih banyak melakukan penyewaan sepeda disaat hari libur (baik itu libur nasional maupun akhir pekan), sedangkan pengguna registered jauh lebih banyak melakukan penyewaan sepeda saat hari kerja.

Analisis - (Pertanyaan 4)

Apakah distribusi penyewaan sepeda merata sepanjang tahun atau ada bulan-bulan tertentu dengan permintaan rendah?



Berdasarkan grafik disamping dilihat bahwa terjadi **kenaikan yang cukup signifikan dari tahun pertama ke tahun kedua.**

- Pada tahun pertama, terjadi kenaikan signifikan dalam jumlah penyewaan, terutama di paruh pertama tahun. Meskipun Januari dimulai dengan angka yang relatif rendah, terdapat lonjakan sebesar +39.78% ke Februari, diikuti tren kenaikan yang konsisten hingga Juni, dengan puncaknya di April (+53.07%).

Namun, mulai Juli hingga Desember, tren berbalik menjadi penurunan yang cukup konsisten, dengan penurunan terbesar terjadi pada Desember (-17.29%).

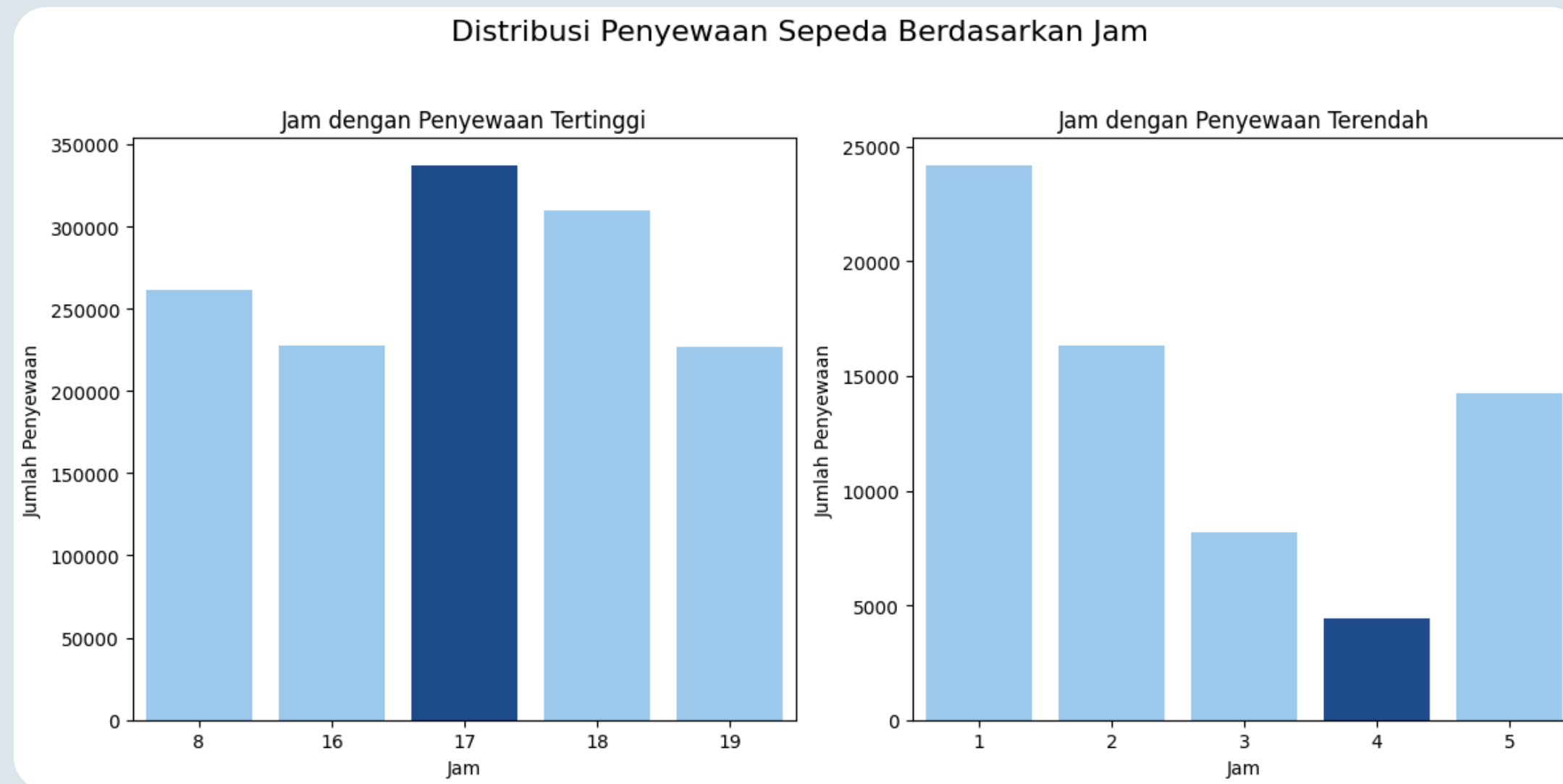
Analisis - (Pertanyaan 4)

- Transisi dari 2011 ke 2012 dimulai dengan positif, ditandai oleh kenaikan +10.79% dari Desember 2011 ke Januari 2012. Paruh pertama tahun kembali menunjukkan kenaikan stabil, meskipun persentasenya lebih kecil dibanding tahun sebelumnya. Puncak pertumbuhan tercatat di bulan Maret dengan +49.55%. Setelah Juni, penyewaan mulai mengalami fluktuasi: turun sedikit di Juli (-2.86%), lalu naik kembali pada Agustus dan September (+5.35% dan +5.29%). Namun, kuartal terakhir kembali menunjukkan penurunan signifikan, dengan Desember mencatat penurunan terbesar (-21.58%).

Secara keseluruhan, tren penyewaan dari kedua tahun menunjukkan pola musiman: peningkatan pada awal tahun dan penurunan di akhir tahun. Musim gugur meskipun lebih fluktuatif, tetap mencatat angka total yang tinggi—khususnya pada tahun kedua yang memiliki nilai penyewaan cukup besar untuk mengimbangi penurunan di akhir tahun.

Analisis - (Pertanyaan 5)

Kapan waktu paling sering dan jarang terjadi transaksi penyewaan sepeda?



Berdasarkan grafik diatas, **penyewaan tertinggi** itu ada di pukul 17:00, disusul oleh pukul 18. Sementara **penyewaan terendah** itu ada di pukul 4:00, disusul oleh pukul 3:00.

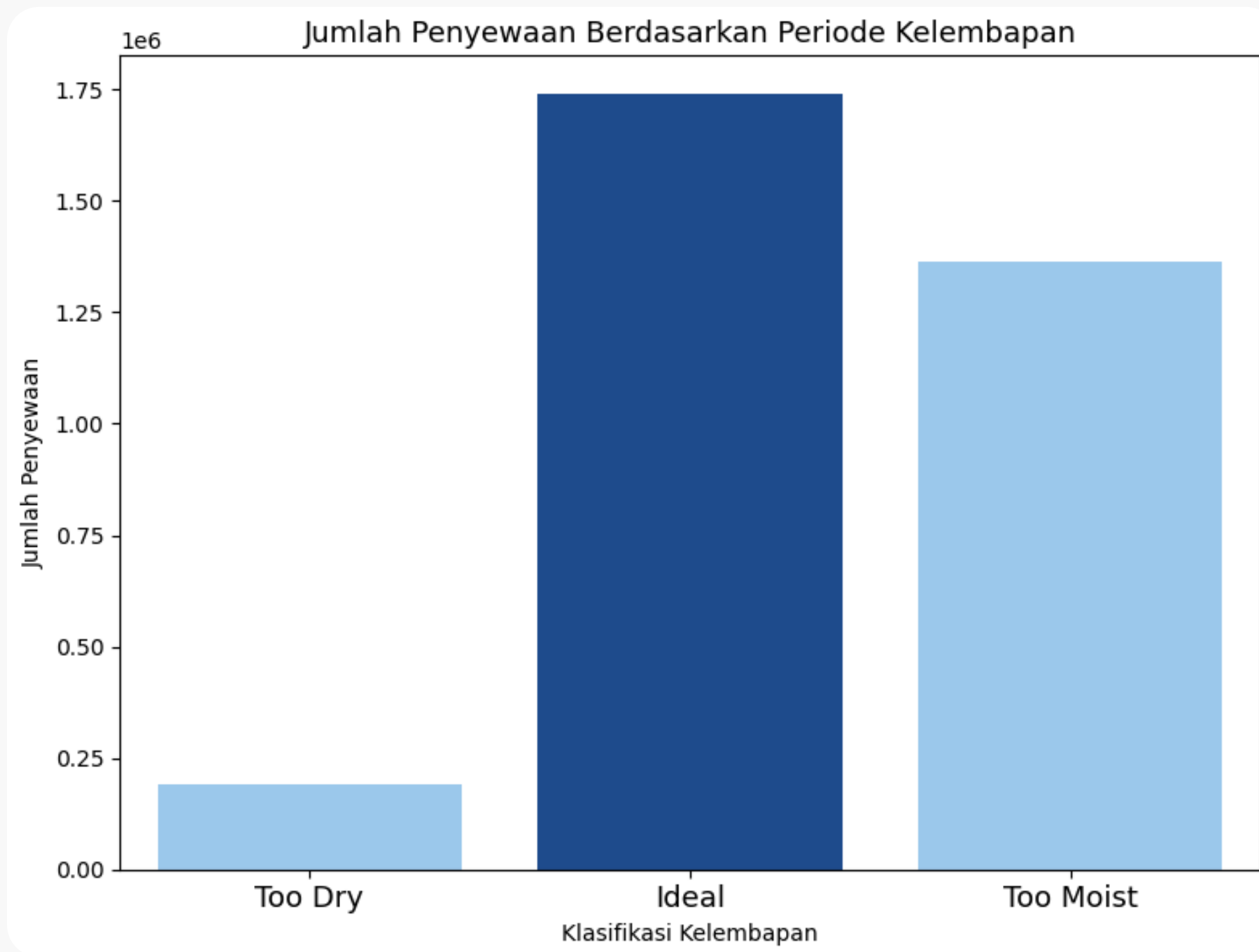
Analisis Lanjutan



Untuk memperkuat analisis, dilakukan analisis lanjutan yakni clustering manual beberapa variabel numerik seperti kelembapan, kecepatan angin, dan suhu.

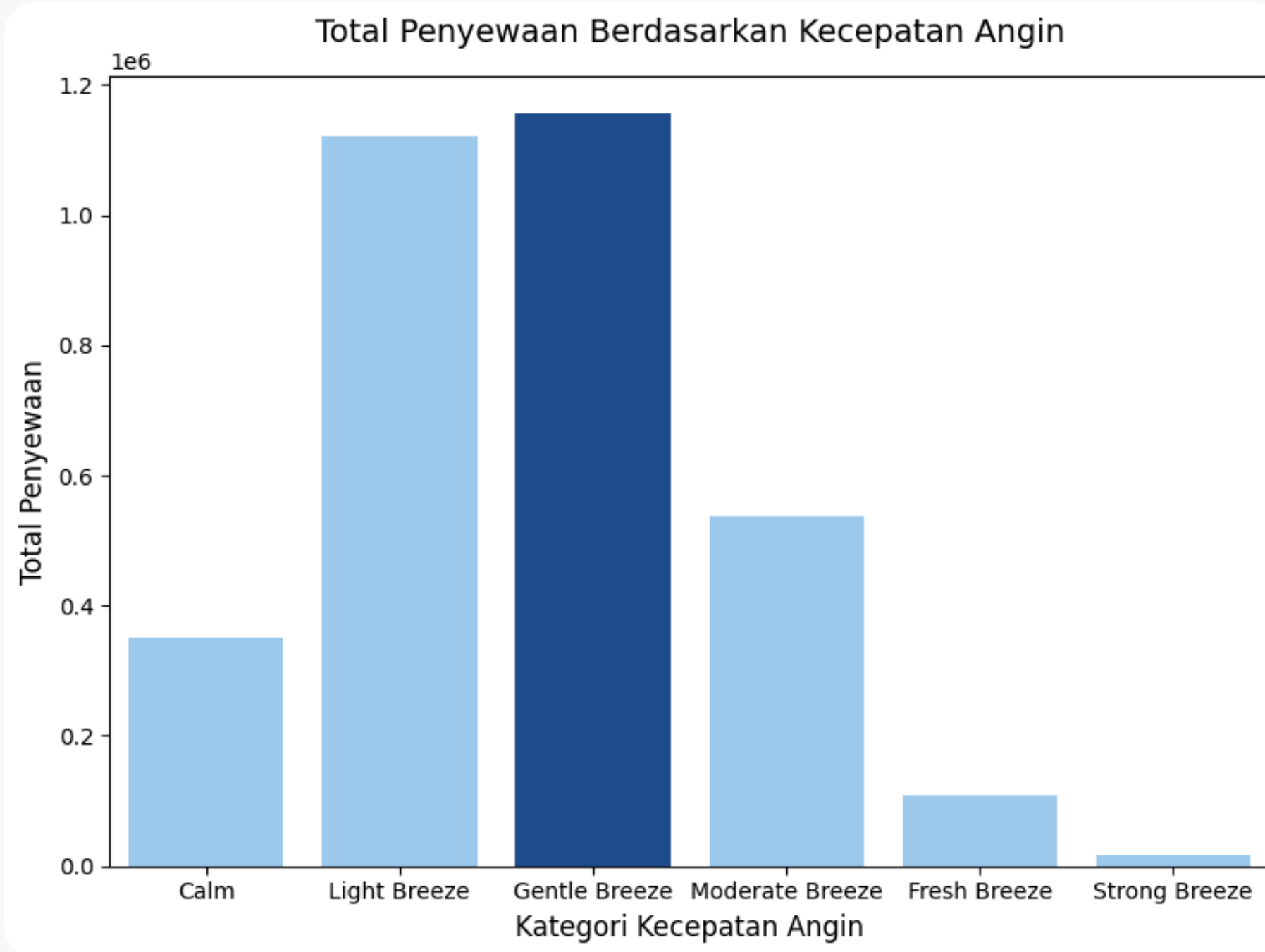


Analisis lanjutan - Kelembapan



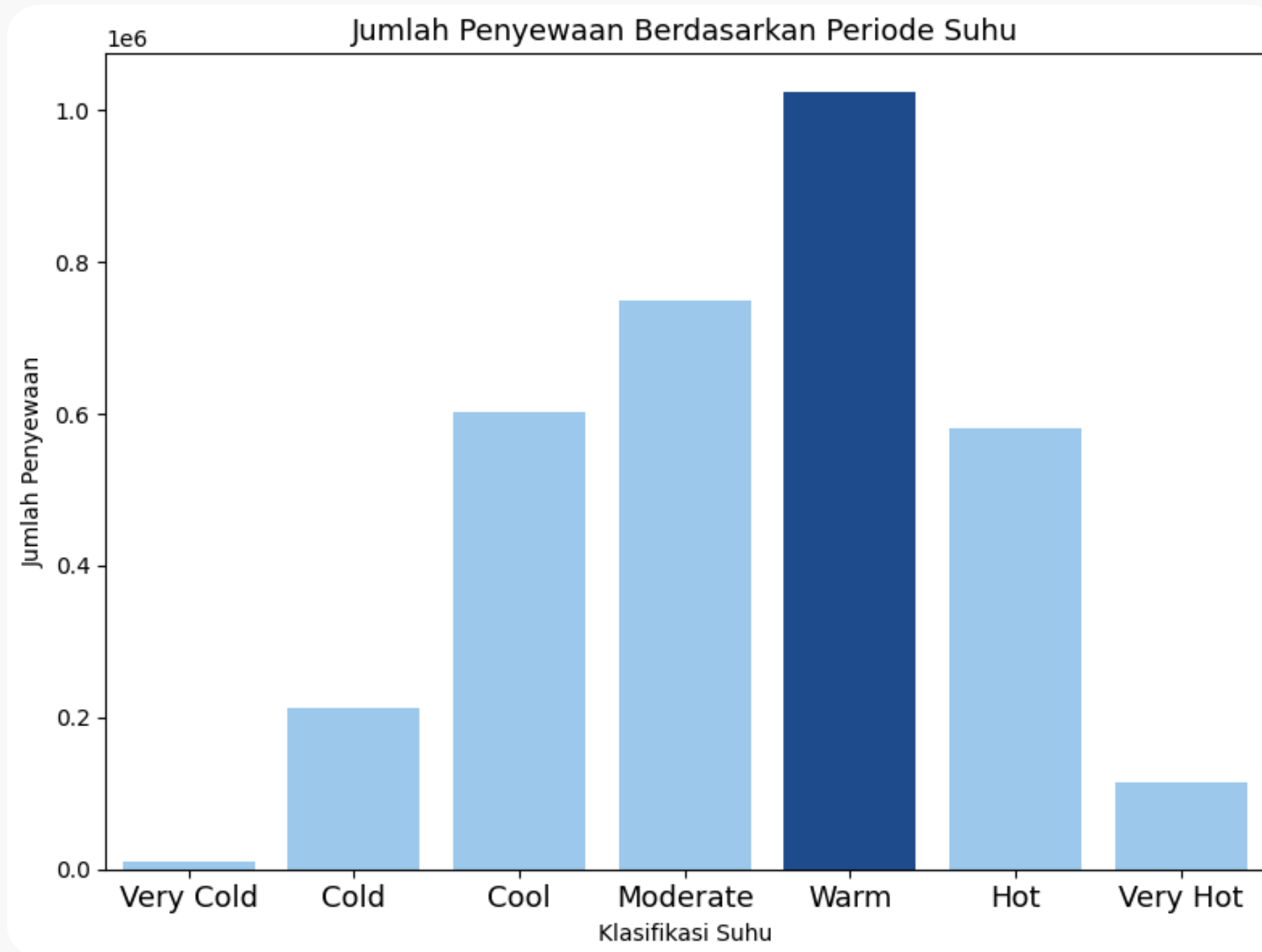
Idealnya, kelembapan berkisar antara 30% - 60%. Diatas 60% sudah dikategorikan terlalu lembap, dan dibawah 30% dikategorikan terlalu kering. Dari hasil grafik disamping, penyewaan banyak dilakukan saat kelembapan ideal.

Analisis lanjutan - Kecepatan angin



Pengelompokan kecepatan angin dilakukan berdasarkan skala beaufort. Dari hasil grafik disamping, penyewaan banyak dilakukan saat kecepatan angin yang pelan (gentle breeze), dengan kecepatan kisaran 8-12 mph. Diposisi kedua, saat kecepatan angin lebih pelan lagi (light breeze) dengan kecepatan kisaran 4-7 mph.

Analisis lanjutan - Suhu



Pengelompokan suhu terdiri dari 7 kelompok. Dari hasil grafik disamping, penyewaan banyak dilakukan saat suhu hangat sekitar 24 - 30°C (penyewaan naik sebesar 33.19%). Diposisi kedua, saat suhu sedang, sekitar 18 - 24°C..

- *Analisis lanjutan - Cuaca dan faktor lain*

Cuaca/Faktor	Suhu (°C)	Kelembapan (%)	Kecepatan Angin (mph)	Jumlah Penyewaan
Cerah/Berawan	20.953	57.375	12.767	2338173
Kabut/Berawan	19.452	69.881	12.058	795952
Salju Ringan/Hujan	18.724	82.763	14.661	158331
Hujan Lebat/Salju	7.653	88.33	13.668	223

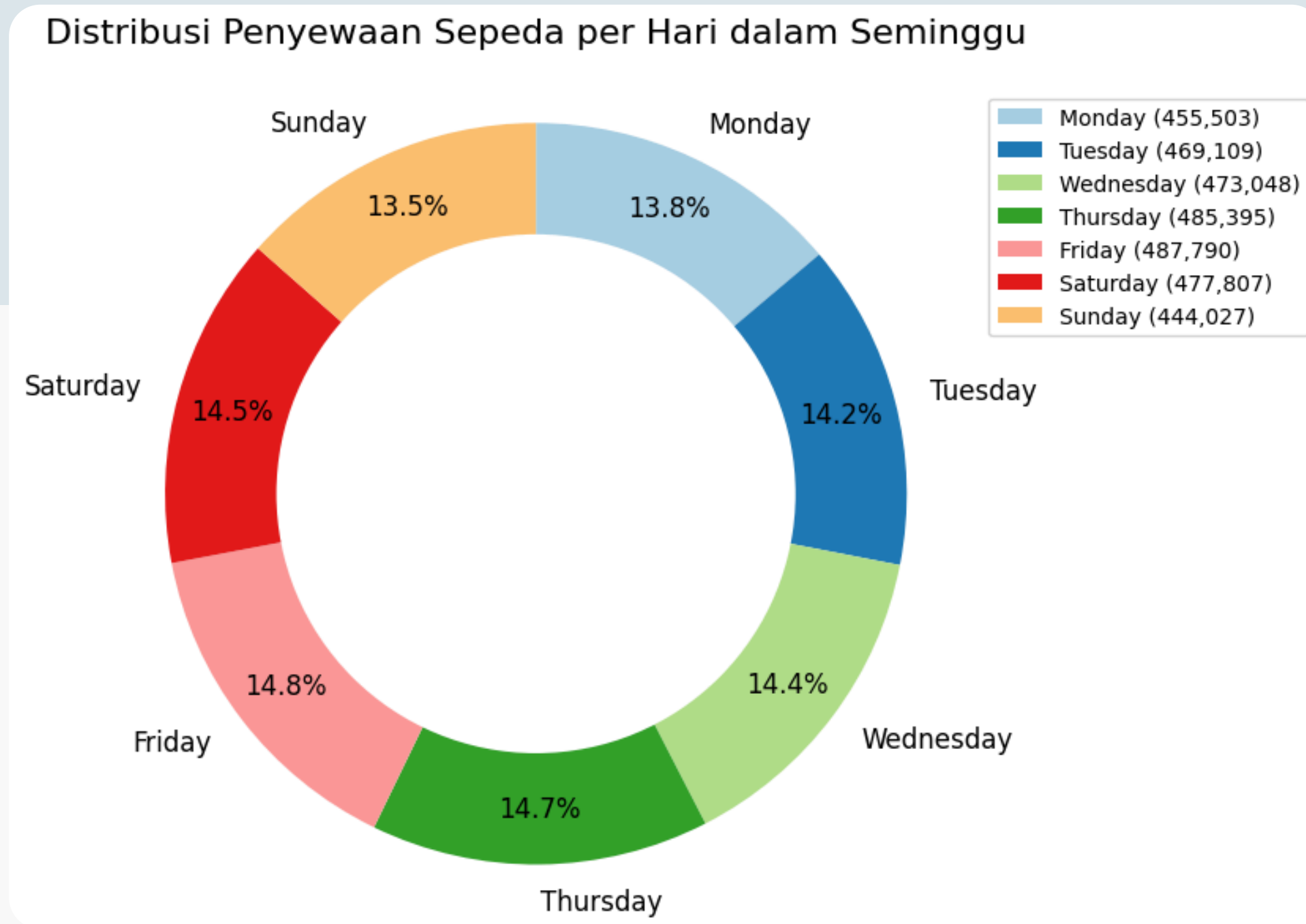
Tabel disini merupakan hubungan cuaca dan faktor lain seperti suhu, kelembapan, dan kecepatan angin. Dari keempat cuaca, cerah berawan menjadi cuaca dengan penyewaan terbanyak.

Dilihat dari tabel perbandingan cuaca dengan suhu, di cuaca tersebut rata - rata suhu sekitar 20.953°C (sedang), Kelembapan di angka 57.375% (ideal) dan Kecepatan angin di angka 12.767 mph (gentle breeze).

Sementara cuaca yang paling dihindari ialah hujan deras atau bersalju, dilihat dari jumlah penyewaan yang paling sedikit.



- *Analisis lanjutan - Faktor hari*



Untuk pengelompokan berdasarkan hari dalam seminggu, hasilnya berkisar diatas angka 400000 penyewaan pada semua hari (perbedaan yang tidak telalu signifikan antar hari) dan hari jumat menjadi hari dengan penyewaan terbanyak, dan memiliki selisih 2395 penyewaan dengan hari kamis (kedua terbanyak) dan selisih 43763 penyewaan dengan hari minggu (penyewaan paling sedikit).

Kesimpulan

Berdasarkan dataset bike sharing tahun 2011 - 2012, terjadi kenaikan pendapatan sebesar 64.88% dari tahun 2011 ke tahun 2012. Secara keseluruhan, suhu menjadi faktor cuaca yang paling memengaruhi jumlah penyewaan sepeda ($r=0.4$), diperkuat saat cuaca hangat ($24-30^{\circ}\text{C}$), penyewaan naik sebesar 33.19%. Sementara kelembapan ($r=-0.32$) bahkan cenderung memiliki pengaruh yang berkebalikan walaupun kecil, dan kecepatan angin ($r=0.09$) hanya memiliki pengaruh kecil. Musim gugur mencatat jumlah penyewaan tertinggi dan paling stabil (rata-rata penyewaan : 5644.30) diikuti musim panas yang juga tinggi meski sering hujan. Musim dingin memiliki aktivitas yang cukup beragam, sementara musim semi jadi yang paling sepi meski kadang muncul lonjakan. Dari sisi pengguna, mayoritas merupakan pengguna registered (81.2%) yang lebih aktif saat hari kerja, sedangkan pengguna casual lebih sering menyewa saat hari libur (51.1%).

Kesimpulan - 2

Dilihat dari tren tahunan, tahun pertama menunjukkan pola kenaikan dari Januari hingga Juni dengan puncak kenaikan di bulan April sebesar 53.07%, lalu menurun secara stabil hingga Desember dengan penurunan mencapai 17.29%. Pada tahun kedua, pola serupa terlihat dengan kenaikan dari Januari ke Juni dan lonjakan tambahan pada Agustus-September, meskipun di paruh kedua tahun terdapat penurunan yang cukup signifikan hingga Desember (21.58%). Kenaikan dari Desember 2011 ke Januari 2012 masih positif sebesar 10.79%, menunjukkan transisi antar tahun yang baik. Meski terjadi penurunan di musim gugur tahun kedua, total penyewaan di musim ini tetap tertinggi secara keseluruhan. Secara waktu, puncak penyewaan terjadi pada sore hari pukul 17:00–18:00, sedangkan paling sepi terjadi dini hari pukul 3:00–4:00.

Pola Pengguna

Pengguna sepeda didominasi oleh pengguna terdaftar (registered), aktif di hari kerja terutama sore hari (sekitar pukul 17:00–18:00), kemungkinan besar untuk keperluan pulang kerja. Pengguna casual lebih sedikit dan lebih aktif di akhir pekan atau hari libur, terutama saat musim gugur dan musim panas. Cuaca ideal untuk penyewaan adalah cerah berawan, suhu sekitar 21°C, kelembapan sedang, dan angin ringan. Penyewaan menurun saat cuaca ekstrem seperti hujan atau salju. Pola ini mengindikasikan bahwa mayoritas pengguna adalah komuter rutin yang sensitif terhadap cuaca dan waktu sibuk.



Catatan: Musim semi dengan penyewaan terendah (lebih rendah dari musim dingin)

Musim semi (Januari–Maret dalam dataset) memiliki jumlah penyewaan lebih rendah dibanding musim dingin, karena beberapa faktor:

- Di awal 2011, penyewaan masih sangat rendah—kemungkinan karena bisnis masih baru merintis atau baru aktif kembali.
- Suhu di musim ini masih cukup rendah karena transisi dari musim dingin.
- Musim semi memiliki banyak libur nasional, sehingga pengguna registered yang biasanya aktif saat kerja, justru tidak menggunakan sepeda.
- Musim dingin mencatat angka lebih tinggi karena mungkin sudah mulai ada basis pengguna tetap. Selain itu, musim semi di awal 2012 menunjukkan peningkatan dibanding musim dingin 2011, mengindikasikan pertumbuhan penyewaan dari waktu ke waktu.

Strategi Bisnis



Data tahun 2011–2012 menunjukkan pola penggunaan sepeda yang cenderung stabil, seperti dominasi pengguna terdaftar, jam sibuk di sore hari, serta tingginya penggunaan saat cuaca cerah dan suhu hangat (sekitar 24-30°C). Meskipun data historis, pola ini kemungkinan masih relevan, terutama di sistem bike-sharing perkotaan modern. Maka, strategi disusun dengan menyesuaikan konteks era 2020-an.

1. Retensi Pengguna Terdaftar (81.2%) – Prioritas Utama

Pengguna registered mendominasi penyewaan (81.2%) dan aktif di hari kerja.

Strategi:

- Tawarkan langgananan fleksibel (harian/mingguan/bulanan) sesuai pola kerja kota.
- Terapkan program loyalitas berbasis poin melalui aplikasi (misal: gratis 1x sewa tiap 10 kali).
- Tambahkan tier subscription (silver-gold-platinum) untuk pengguna aktif.

2. Optimasi Operasional: Cuaca, Musim, dan Jam Puncak

Data menunjukkan penyewaan cukup sensitif terhadap suhu ($r=0.4$), musim, dan jam.



Strategi Bisnis



Strategi:

- Gunakan API prakiraan cuaca untuk memprediksi lonjakan harian. Sediakan sepeda ekstra di area perkantoran saat cuaca cerah & suhu hangat, khususnya pukul 17:00-18:00.
- Sepanjang musim dingin dan transisi musim dingin ke semi (terendah):
 - Jadikan masa maintenance dan uji coba fitur baru (jika memungkinkan).
 - Kurangi armada di lokasi sepi dan rotasi ke area yang tetap aktif.
 - Buat konten ringan seperti "Safe Cycling Tips in Cold Weather".
 - Beri poin loyalty ekstra atau cashback kecil untuk pemakaian rutin (misalnya, 5 hari berturut-turut).
- Transisi ke musim panas (kenaikan besar): Diskon untuk pengguna casual di akhir pekan serta tambahkan stok di lokasi wisata dan taman kota.
- Musim gugur (puncak penyewaan, >5.600 sewa/bulan): Paket langganan + diskon karena bulan ini mulai ajaran baru, berlaku untuk registered users.
- Pasang iklan bertema "Sepeda untuk Musim Terbaik".



Strategi Bisnis



3. Aktivasi Pengguna Casual (51.1% dominan saat weekend)

Targetkan gaya hidup & liburan untuk menarik casual users.

Strategi:

- Kolaborasi dengan tempat wisata (taman kota, museum, food festival): paket tiket + rental sepeda.
- Buat campaign “Explore the City with Bike” di media sosial, dorong UGC (user-generated content) dengan hadiah sewa gratis.
- Gunakan momen libur panjang & cuaca hangat (misal: “Traffic-Free Weekend”).

4. Perluasan Lokasi Strategis

Gunakan heatmap permintaan dan mobilitas pekerja/kampus.

Strategi:

- Tambahkan stasiun di area kampus, transit (stasiun/halte utama), dan distrik wisata.
- Prioritaskan lokasi dengan lonjakan penyewaan di jam sibuk dan hari kerja.



Strategi Bisnis



Catatan:

Strategi ini disusun murni berdasarkan pola yang ditemukan dalam dataset historis Bike Sharing 2011–2012, tanpa mempertimbangkan data eksternal, survei pengguna, atau realitas pasar saat ini. Interpretasi dan penerapannya tetap perlu disesuaikan dengan konteks bisnis dan lingkungan terkini.



Visualisasi dengan Tableau

