

SOAL

1. Buatlah function Uji Hipotesis rata-rata (t-test) yang bisa mengakomodasi 3 jenis pengujian, uji lower tail, uji upper tail, dan uji two tail
2. Menggunakan function yang telah dibuat pada nomor sebelumnya, selesaikan soal dibawah ini:
Seorang perenang mencatat waktu rata-rata 16,43 detik untuk berenang dengan jarak 25 meter, dengan standar deviasi 0,8 detik. Pelatuhnya kemudian menyuruhnya berenang menggunakan kacamata berenang untuk 15 kali dan dari hasil 15 kali berenang, waktu rata-rata perenang tersebut adalah 16 detik. Pelatuhnya berasumsi bahwa perenang tersebut akan berenang lebih cepat menggunakan kacamata renang. Uji hipotesis pelatih tersebut dengan nilai
3. Buatlah function untuk mengecek apakah angka yang dimasukan kedalam function adalah angka genap atau ganjil
4. Buatlah function untuk mengecek apakah angka yang dimasukan kedalam function tersebut adalah angka prima atau bukan
5. Buatlah function factorial yang akan menghasilkan nilai factorial dari angka yang dimasukan kedalam function
6. *Data Scaling* adalah konsep yang sering digunakan di dalam pre-processing pada Machine Learning untuk meningkatkan performa Machine Learning dan meminimalkan waktu training. Salah satu metode scaling adalah menggunakan metode **Mean-Max Scaling** dengan formula $x_{new} = \frac{x - \min(x)}{\max(x) - \min(x)}$. Dengan menggunakan data **iris** pada R, lakukan **Mean-Max Scaling** pada variabel **Sepal.Length**, kemudian lakukan pengecekan apakah data hasil scaling berdistribusi normal menggunakan simple QQ-Plot atau density plot! Interpretasikan hasilnya!
7. Dari data **iris** buatlah boxplot untuk membandingkan **Sepal.Length** terhadap **Species** nya! Menurut anda, apakah terdapat perbedaan antara **Sepal.Length** antar **Species** dan bagaimana hasil visualisasi tersebut terhadap **Species** nya?
8. IPM (Indeks Pembangunan Manusia) merupakan salah satu indikator dalam mengukur kualitas SDM suatu daerah. IPM dihitung dari 3 dimensi, yakni dimensi kesehatan, pendidikan, dan keuangan. Menggunakan dataset **swiss**, jika dianggap bahwa **Education**, **Fertility**, dan **Agriculture** dianggap merepresentasikan ke-3 dimensi tersebut, hitunglah nilai "*IPM*" = $\sqrt[3]{\log(Education) + \log(Fertility) + \log(Agriculture)} * 10$ dan masukkan ke dalam variabel baru di dalam dataframe yang sama! (**Note:** perhitungan IPM di dalam soal ini hanyalah fiktif belaka. Untuk melihat formula IPM yang asli, silakan merujuk kepada BPS atau UNSTATS)
9. Dari nilai "*IPM*" tersebut, bagi data menjadi kota dengan "*IPM*" di atas rata-rata dan dibawah rata-rata. Apa perbedaan statistika deskriptif diantara keduanya?