

Nama: Theophilus Alexander Elvan

NIM : 2702259986

## AOL SC

### **VSCODE Version**

- **Library Yang Digunakan**

Berikut penjelasan mengenai library yang digunakan dalam kode:

1. **NumPy (numpy)**

- Fungsi Utama: Digunakan untuk komputasi numerik yang efisien. Menyediakan struktur array dan berbagai fungsi matematika.
- Contoh Penggunaan: Membuat array `months` untuk mewakili bulan dari 1 hingga 144.

2. **Pandas (pandas)**

- Fungsi Utama: Digunakan untuk manipulasi dan analisis data. Menyediakan struktur data seperti `DataFrame` untuk pengelolaan data tabular.
- Contoh Penggunaan: Membaca file CSV dan mengekstrak kolom data produksi ke dalam `DataFrame`.

3. **Matplotlib (matplotlib.pyplot)**

- Fungsi Utama: Digunakan untuk membuat grafik dan visualisasi data. Menyediakan berbagai fungsi untuk membuat grafik dan diagram.
- Contoh Penggunaan: Membuat grafik yang menampilkan data produksi aktual serta prediksi model eksponensial dan polinomial.

4. **SciPy (scipy.optimize)**

- Fungsi Utama: Digunakan untuk optimisasi dan pemecahan masalah matematika, termasuk fitting kurva dan menemukan akar fungsi.
- Contoh Penggunaan:
  - `curve_fit` digunakan untuk menyesuaikan model eksponensial dan polinomial dengan data.
  - `fsolve` digunakan untuk menemukan bulan di mana produksi melebihi kapasitas gudang yang ditentukan.

- **Alasan penggunaan Library**

1. **NumPy (numpy)**

Alasan Penggunaan:

- Kinerja Tinggi: NumPy menyediakan struktur array multidimensi yang sangat efisien untuk komputasi numerik, lebih cepat dan hemat memori dibandingkan dengan list Python biasa.

- Fungsi Matematis: Menyediakan berbagai fungsi matematis untuk operasi array, seperti fungsi aljabar linear, statistik dasar, dan transformasi Fourier.
- Interoperabilitas: NumPy sering digunakan sebagai dasar untuk library lain seperti SciPy, Pandas, dan Matplotlib, sehingga memungkinkan interoperabilitas yang mulus.

## 2. Pandas (pandas)

Alasan Penggunaan:

- Manipulasi Data yang Mudah: Pandas menyediakan struktur data seperti DataFrame dan Series yang memudahkan pengelolaan, pengelompokan, dan transformasi data tabular.
- Pembacaan dan Penulisan Data: Menyediakan fungsi yang mudah digunakan untuk membaca dari dan menulis ke berbagai format file seperti CSV, Excel, SQL, dan lainnya.
- Operasi Data Tabular: Memungkinkan operasi berbasis label dan indeks, seperti pemfilteran, penggabungan, pengelompokan, dan agregasi.

## 3. Matplotlib (matplotlib.pyplot)

- Alasan Penggunaan:
- Visualisasi Fleksibel: Menyediakan alat yang kuat dan fleksibel untuk membuat berbagai jenis plot dan grafik, dari plot sederhana hingga visualisasi yang lebih kompleks.
- Personalisasi Tinggi: Membolehkan personalisasi plot dengan gaya, anotasi, label, dan format yang sangat kustomisasi.
- Komunitas Besar: Matplotlib adalah library yang sangat populer dengan dokumentasi yang baik dan banyak contoh penggunaan yang bisa diadaptasi.

## 4. SciPy (scipy.optimize)

Alasan Penggunaan:

- Fungsi Optimisasi Lanjutan: SciPy menyediakan berbagai alat untuk optimisasi, pemecahan persamaan, analisis sinyal, statistik, dan lainnya.
- Ekstensif dan Teruji: Berisi banyak fungsi yang telah teruji dan dioptimalkan untuk berbagai aplikasi ilmiah dan teknik.
- Integrasi dengan NumPy: Didesain untuk bekerja dengan baik bersama NumPy, sehingga dapat menangani array dan operasi numerik kompleks secara efisien.

## • Akurasi Pada Program

Evaluasi keakuratan program ini melibatkan beberapa aspek, termasuk pemilihan model, proses fitting, visualisasi, dan pemecahan masalah terkait keputusan bisnis. Berikut adalah penjelasan mengenai keakuratan dari program tersebut:

### 1. Pemilihan Model

#### ➤ Model Eksponensial:

- Keakuratan: Model ini cocok untuk data yang menunjukkan pertumbuhan eksponensial (pertumbuhan cepat dari waktu ke waktu).

- Keterbatasan: Jika data tidak menunjukkan pola eksponensial (misalnya, data lebih linier atau memiliki pola berbeda), model ini mungkin kurang akurat. Dalam hal ini, ada kemungkinan overfitting atau underfitting tergantung pada data yang sebenarnya.

➤ Model Polinomial (Derajat 3):

- Keakuratan: Model ini mampu menangkap berbagai pola yang lebih kompleks karena memiliki lebih banyak derajat kebebasan. Ini memungkinkan model untuk menyesuaikan kurva lebih baik pada data yang memiliki pola perubahan yang lebih kompleks dari waktu ke waktu.
- Keterbatasan: Model polinomial derajat tinggi bisa overfit jika jumlah parameter terlalu banyak dibandingkan data. Ini berarti model mungkin mengikuti fluktuasi kecil pada data yang sebenarnya hanya kebisingan.

## 2. Proses Fitting

➤ Curve Fitting (curve fit):

- Keakuratan: Proses ini mencari parameter optimal yang meminimalkan selisih antara model dan data aktual. Jika fitting dilakukan dengan baik, model akan mengikuti data aktual dengan cukup baik.
- Keterbatasan: Hasil fitting sangat tergantung pada kualitas data dan asumsi awal dari model. Tanpa parameter awal yang baik atau dengan data yang sangat berisik, hasil fitting bisa kurang akurat.

## 3. Evaluasi dan Visualisasi Model

➤ Evaluasi Model:

- Program menilai keakuratan dengan membandingkan prediksi model dengan data aktual.
- **Keakuratan:** Jika model mengikuti pola data dengan baik, maka prediksi yang dihasilkan akan lebih akurat.
- Keterbatasan: Tidak ada metrik kuantitatif seperti Mean Squared Error (MSE) atau R-squared yang ditampilkan untuk memberikan ukuran keakuratan yang lebih formal. Evaluasi visual saja tidak selalu cukup untuk menilai keakuratan.

➤ Visualisasi:

- Keakuratan: Memvisualisasikan data aktual dan model membantu dalam memahami sejauh mana model mengikuti data aktual.
- Keterbatasan: Visualisasi dapat menipu jika skala atau rentang data tidak ditampilkan dengan benar. Mungkin sulit untuk menilai keakuratan secara kuantitatif hanya dari grafik.

#### 4. Pemecahan Masalah

##### ➤ Menentukan Waktu untuk Gudang Baru (fsolve):

- Keakuratan: Fungsi fsolve digunakan untuk menemukan ketika produksi diproyeksikan akan melebihi kapasitas tertentu. Ini bergantung pada akurasi model polinomial yang digunakan.
- Keterbatasan: Jika model yang digunakan tidak akurat, proyeksi waktu untuk melebihi kapasitas juga akan tidak akurat. Solusi ini juga sensitif terhadap nilai-nilai parameter yang mungkin tidak mencerminkan data aktual sepenuhnya.

#### 5. Peningkatan Keakuratan

##### ➤ Untuk meningkatkan keakuratan program, langkah-langkah berikut dapat diambil:

1. Evaluasi Alternatif Model: Uji model lain, seperti regresi linear, non-linear, atau model machine learning untuk pola yang lebih kompleks.
2. Penggunaan Metrik Evaluasi: Tambahkan metrik seperti MSE, RMSE, atau R-squared untuk mengukur keakuratan model secara kuantitatif.
3. Validasi Model: Gunakan teknik validasi silang untuk menghindari overfitting dan mendapatkan perkiraan keakuratan yang lebih baik.
4. Pengolahan Data: Periksa dan bersihkan data untuk mengurangi kebisingan dan outlier yang bisa mempengaruhi keakuratan model.
5. Eksplorasi Parameter Fitting: Uji berbagai parameter awal untuk fitting dan tentukan yang terbaik melalui eksperimen atau optimasi.
6. Visualisasi Lanjutan: Gunakan plot residu atau diagnosis lain untuk memeriksa kecocokan model lebih lanjut.

## Google Colab Version

Link notebook colab:

<https://colab.research.google.com/drive/12ql-QinuFU03Mxorg4XkG7r4YSNdCuGM?usp=sharing>

Github Repository:

[Github](#)

### • Library yang digunakan

#### 1. Pandas

- Alasan Penggunaan: Pandas adalah library yang sangat umum digunakan untuk manipulasi dan analisis data dalam Python. Pada kasus ini, Pandas digunakan untuk menyusun data produksi ke dalam bentuk DataFrame, yang memudahkan proses analisis dan pemodelan data.
- Keakuratan: Tingkat keakuratan Pandas sangat tinggi dalam hal manipulasi data struktural, seperti pengurutan, pengelompokan, dan agregasi data. Keandalannya dalam menangani data besar dan kompleks membuatnya sangat cocok untuk analisis awal dan penyajian data.

## 2. Matplotlib

- Alasan Penggunaan: Matplotlib adalah library utama untuk visualisasi data dalam Python. Digunakan untuk membuat grafik dan plot yang membantu dalam memvisualisasikan tren data, dalam hal ini tren produksi tas.
- Keakuratan: Matplotlib memungkinkan visualisasi data yang sangat tepat dan terperinci. Ini memudahkan identifikasi pola dan tren dalam data, serta perbandingan antara data asli dan hasil model.

## 3. Statsmodels (ARIMA)

- Alasan Penggunaan: Statsmodels adalah library Python untuk melakukan analisis statistik dan ekonometri. ARIMA (AutoRegressive Integrated Moving Average) adalah salah satu metode yang ada di dalamnya yang digunakan untuk analisis dan peramalan deret waktu. ARIMA cocok untuk data produksi yang menunjukkan karakteristik musiman dan tren.
- Keakuratan: ARIMA dikenal sangat efektif untuk memodelkan dan meramalkan deret waktu yang memiliki keteraturan atau pola. Keakuratan model ini tergantung pada pemilihan parameter (order) yang tepat. Dalam contoh ini, order (5,1,0) dipilih, yang menunjukkan 5 lag pada bagian autoregressive, differencing sebanyak 1, dan tanpa moving average. Parameter ini dipilih berdasarkan pengujian dan kriteria seperti AIC (Akaike Information Criterion).

## • Evaluasi Keakuratan

### 1. Fit Model Summary:

- AIC/BIC: Ini adalah kriteria informasi yang digunakan untuk mengevaluasi seberapa baik model fit dengan data. AIC (Akaike Information Criterion) dan BIC (Bayesian Information Criterion) lebih rendah menandakan model yang lebih baik.
- Residuals Analysis: Residuals atau sisa adalah perbedaan antara nilai aktual dan nilai yang diprediksi oleh model. Analisis residual dapat menunjukkan apakah model menangkap pola dalam data atau ada pola yang tersisa.

### 2. Forecasting:

- Forecast untuk 12 Bulan Mendatang: Peramalan dilakukan untuk mengetahui tren masa depan. Pada contoh ini, peramalan dilakukan untuk 12 bulan ke depan, memberikan gambaran tentang bagaimana produksi akan berjalan dalam satu tahun ke depan.
- Extended Forecast: Dilakukan untuk jangka panjang (100 bulan), ini berguna untuk melihat kapan produksi mungkin melebihi kapasitas gudang yang ada dan untuk merencanakan tindakan seperti membangun gudang baru.

### 3. Tindakan Berdasarkan Ramalan:

- Kapan Memulai Pembangunan Gudang Baru: Berdasarkan peramalan produksi dan batas kapasitas gudang (25000 unit), kita menghitung kapan produksi akan melebihi kapasitas dan merencanakan kapan harus memulai pembangunan gudang baru, yaitu 13 bulan sebelum kapasitas terlampaui.