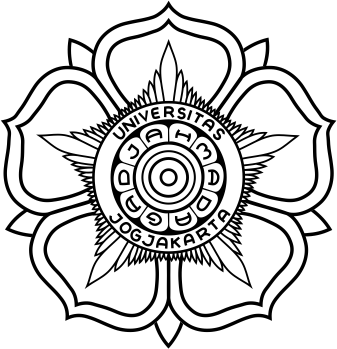
**PERBANDINGAN METODE FORECASTING PADA DATA PANEN DI PROVINSI YOGYAKARTA DAN JAWA BARAT**

****

Disusun oleh :

**Sudiro 16/399873/TK/44887**

**Thea Kirana 16/399921/TK/44935**

**Vicko Pranowo 16/394959/TK/44251**

**DEPARTEMEN TEKNIK ELEKTRO DAN TEKNOLOGI INFORMASI**

**FAKULTAS TEKNIK**

**UNIVERSITAS GADJAH MADA**

**YOGYAKARTA**

**2019**

# Pendahuluan

Tugas yang dikerjakan adalah memprediksi data hasil panen dan menampilkan visualisasi dari data tersebut pada sebuah *dashboard* atau aplikasi web. Metode prediksi yang digunakan adalah regresi linear, regresi polinomial, dan LSTM. Visualisasi data dilakukan menggunakan *platform* aplikasi web Dash by Plotly.

# Metode dan Hasil

* 1. **Data**

Data yang digunakan pada percobaan merupakan data hasil pertanian tomat dan jagung darei berbagai negara di dunia yang diperoleh dari web <http://ourworldindata.org.> Namun pada implementasinya pada percobaan ini, dianggap berasal dari data panen setiap kota di Provinsi Jogjakarta dan Jawa Barat

* 1. **Prediksi**
     1. **Regresi**

Regresi linear atau *Linear Regression* adalah salah satu teknik dari *Machine Learning* untuk menemukan hubungan antara satu atau lebih antara fitur (variabel independen) dengan sebuah variabel target yang kontinyu (variabel dependen). Model regresi linear dapat dinyatakan dalam persamaan berikut,

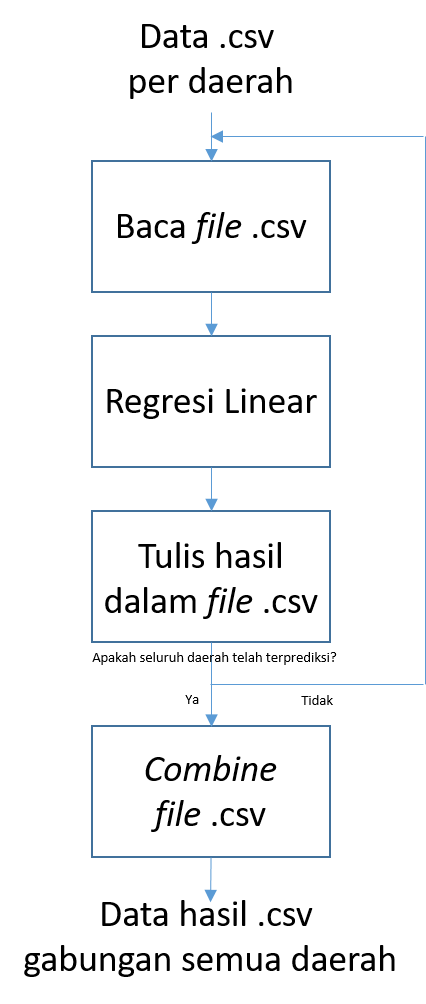
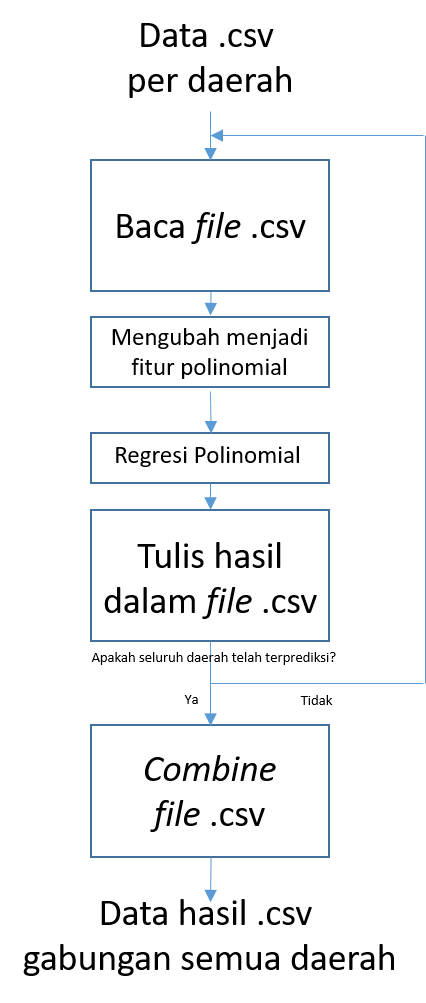
Dimana *Y* adalah variabel target yang diprediksi, adalah nilai bias, adalah parameter model, dan adalah nilai fitur. Model dilatih dengan sebuah kumpulan data untuk menemukan parameter yang tepat sehingga model yang sangat sesuai untuk memprediksi data tersebut. Model berupa garis yang terdiri dari nilai-nilai target yang terprediksi dari kumpulan data latih. Model yang baik memiliki kesalahan atau *error* yang rendah antara nilai terprediksi dengan data latih. Pengukuran yang digunakan untuk mengevaluasi kinerja model adalah *Root Mean Squared Error* (RMSE) dan *Coefficient of Determination* (nilai R2). RMSE adalah akar kuadrat dari rerata jumlah kuadrat *error*. RMSE dapat dinyatakan sebagai berikut.

Dimana adalah fungsi hipotesis atau nilai prediksi. Sedangkan nilai R2 menjelaskan jumlah varians dari variabel dependen yang dapat dikurangi dengan menggunakan *least square regression*. Nilai R2 dapat dinyatakan sebagai berikut,

Dimana *SSr* adalah jumlah dari kuadrat *error*, dan *SSt* adalah jumlah total dari *error* jika menggunakan rerata dari nilai yang diamati sebagai nilai terprediksi.

Model regresi linear memiliki keterbatasan untuk merepresentasikan data kompleks yang tidak linear. Untuk mengatasi hal tersebut dibutuhkan model yang lebih kompleks. Persamaan dengan tingkat yang lebih tinggi, dapat didapat dengan menambahkan pangkat pada fitur asli sebagai fitur yang baru sehingga persamaan model menjadi,

Sehingga model ini dapat menyesuaikan data yang lebih kompleks dengan kurva kuadratik.

Gambar di atas menjelaskan alur program untuk regresi linear dan regresi polinomial. Data hasil panen per daerah dibaca dari *file* .csv kemudian dilakukan perhitungan regresi beserta evaluasi RMSE dan nilai R2. Hasil prediksi dari regresi kemudian dituliskan kembali ke dalam *file* .csv per daerah. Terakhir seluruh *file* .csv tersebut digabungkan menjadi satu *file* untuk digunakan dalam proses visualisasi. Prediksi yang dihasilkan adalah nilai hasil panen pada tahun 2015 hingga 2020.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Hasil panen | : Maize | | | | | |
| **Daerah** | **Linear Regression** | | **Polynomial Regression (degree 2)** | | **Polynomial Regression (degree 3)** | |
| **RMSE** | **R2** | **RMSE** | **R2** | **RMSE** | **R2** |
| Bandung | 0.155 | 0.0776 | 0.152 | 0.0988 | 0.125 | 0.259 |
| Bantul | 0.574 | 0.7533 | 0.558 | 0.76 | 0.272 | 0.882 |
| Bekasi | 0.575 | 0.027 | 0.573 | 0.0294 | 0.5513 | 0.067 |
| Bogor | 0.1936 | 0.8417 | 0.0428 | 0.9649 | 0.0346 | 0.9716 |
| Ciamis | 0.04497 | 0.7854 | 0.04468 | 0.7868 | 0.0375 | 0.8206 |
| Cianjur | 0.0905 | 0.7265 | 0.08613 | 0.7399 | 0.067 | 0.7975 |
| Cirebon | 0.01663 | 0.9613 | 0.0132 | 0.9692 | 0.0126 | 0.9705 |
| DIY | 0.02916 | 0.615 | 0.0244 | 0.6767 | 0.0229 | 0.6977 |
| Garut | 0.0193 | 0.248 | 0.0137 | 0.4665 | 0.0136 | 0.4699 |
| Gunung Kidul | 1.1697 | 0.44565 | 1.0724 | 0.4917 | 1.0321 | 0.5108 |
| Indramayu | 0.0147 | 0.0649 | 0.0107 | 0.3219 | 0.01029 | 0.3487 |
| Karawang | 0.0461 | 0.7856 | 0.0352 | 0.8362 | 0.0349 | 0.8373 |
| Kulon Progo | 0.1092 | 0.2439 | 0.10523 | 0.2712 | 0.0996 | 0.3096 |
| Kuningan | 0.0629 | 0.4899 | 0.0481 | 0.6096 | 0.04768 | 0.6134 |
| Majalengka | 0.1855 | 0.879 | 0.1273 | 0.9174 | 0.0768 | 0.9502 |
| Purwakarta | 0.0306 | 0.16053 | 0.0306 | 0.1616 | 0.027 | 0.2601 |
| Sleman | 0.0395 | 0.03004 | 0.0151 | 0.6299 | 0.0135 | 0.6686 |
| Subang | 0.625 | 0.5249 | 0.128 | 0.902 | 0.1208 | 0.9081 |
| Sukabumi | 1.4 | 0.101 | 1.376 | 0.1171 | 0.8422 | 0.4596 |
| **Rerata** | **0.283229** | **0.461117** | **0.234555** | **0.56583158** | **0.181125** | **0.62116842** |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Hasil panen | : Tomato | | | | | |
| **Daerah** | **Linear Regression** | | **Polynomial Regression (degree 2)** | | **Polynomial Regression (degree 3)** | |
| **RMSE** | **R2** | **RMSE** | **R2** | **RMSE** | **R2** |
| Bandung | 16.89 | 0.8471 | 8.417 | 0.9238 | 8.398 | 0.924 |
| Bantul | 33.94 | 0.598 | 4.849 | 0.9426 | 4.384 | 0.948 |
| Bekasi | 17.199 | 0.7721 | 8.6688 | 0.88515 | 8.3637 | 0.8891 |
| Bogor | 0.8961 | 0.822 | 0.8802 | 0.826 | 0.7876 | 0.8443 |
| Ciamis | 18.732 | 0.0696 | 17.261 | 0.1427 | 8.8421 | 0.56085 |
| Cianjur | 13.133 | 0.7182 | 7.462 | 0.8398 | 5.1928 | 0.8885 |
| Cirebon | 0.788 | 0.00127 | 0.7835 | 0.008019 | 0.4738 | 0.4001 |
| DIY | 10.252 | 0.8847 | 6.3829 | 0.92824 | 5.4359 | 0.93889 |
| Garut | 0.87984 | 0.1096 | 0.65219 | 0.3399 | 0.419 | 0.5751 |
| Gunung Kidul | 0.1207 | 0.3112 | 0.0498 | 0.716 | 0.02745 | 0.84347 |
| Indramayu | 340.55 | 0.59442 | 280.02 | 0.6665 | 160.546 | 0.80879 |
| Karawang | 13.269 | 0.7408 | 4.784 | 0.9065 | 3.5186 | 0.9312 |
| Kulon Progo | 15.607 | 0.77489 | 10.212 | 0.8527 | 8.9625 | 0.8707 |
| Kuningan | 7.514 | 0.10702 | 2.7504 | 0.673 | 2.0393 | 0.7576 |
| Majalengka | 0.28877 | 0.26633 | 0.16275 | 0.5865 | 0.05817 | 0.85219 |
| Purwakarta | 7.673 | 0.9728 | 6.16027 | 0.978 | 5.828 | 0.9794 |
| Sleman | 1.4638 | 0.7955 | 1.218 | 0.829 | 1.07797 | 0.84941 |
| Subang | 158.931 | 0.6531 | 59.812 | 0.86945 | 38.605 | 0.915 |
| Sukabumi | 0.0856 | 0.487611 | 0.075 | 0.55105 | 0.07396 | 0.5575 |
| **Rerata** | **34.64278** | **0.554013** | **22.13688** | **0.70867942** | **13.84389** | **0.80705789** |

Dari evaluasi RMSE dan R2 yang dilakukan pada dua dataset (Maize dan Tomato), dapat dilihat bahwa semakin tinggi derajat polinomial, semakin sesuai garis prediksi regresi terhadap data.

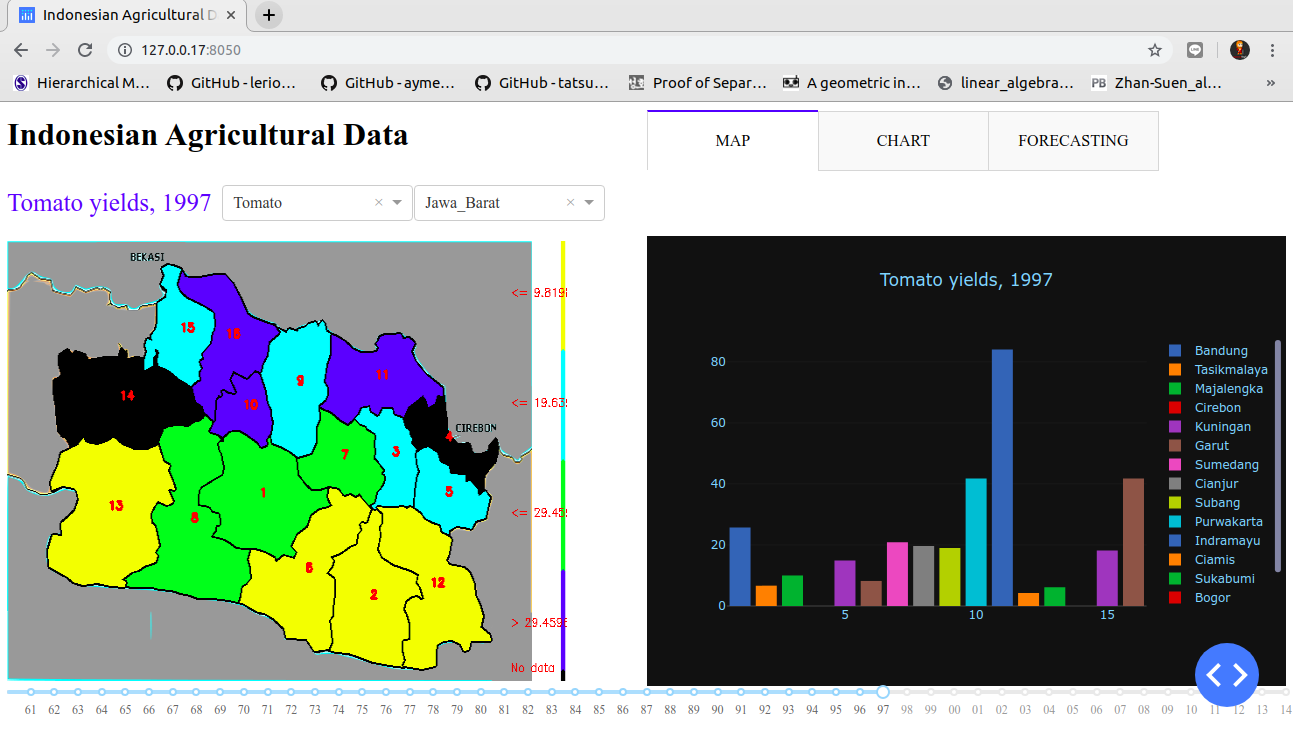
* + 1. **LSTM**

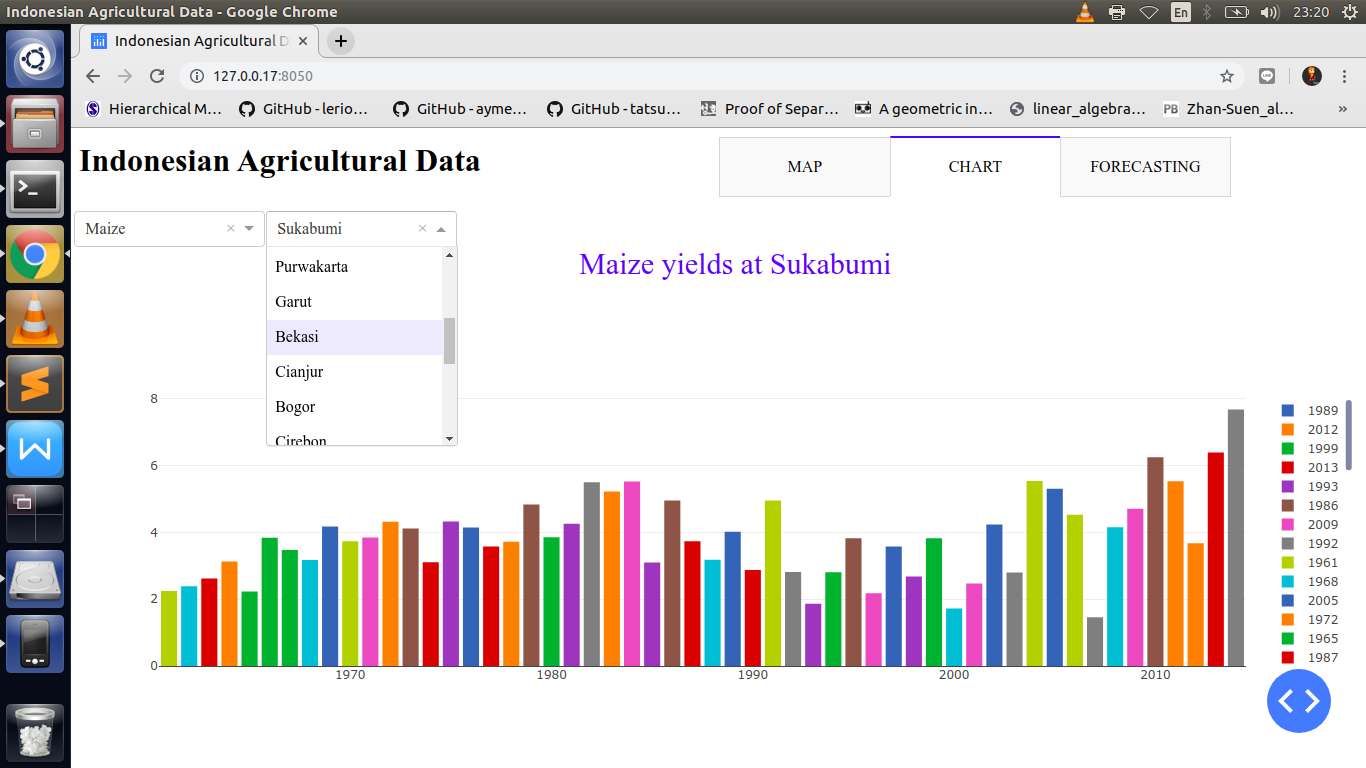
Penjelasan tentang metode LSTM, beserta hasilnya (Vicko)

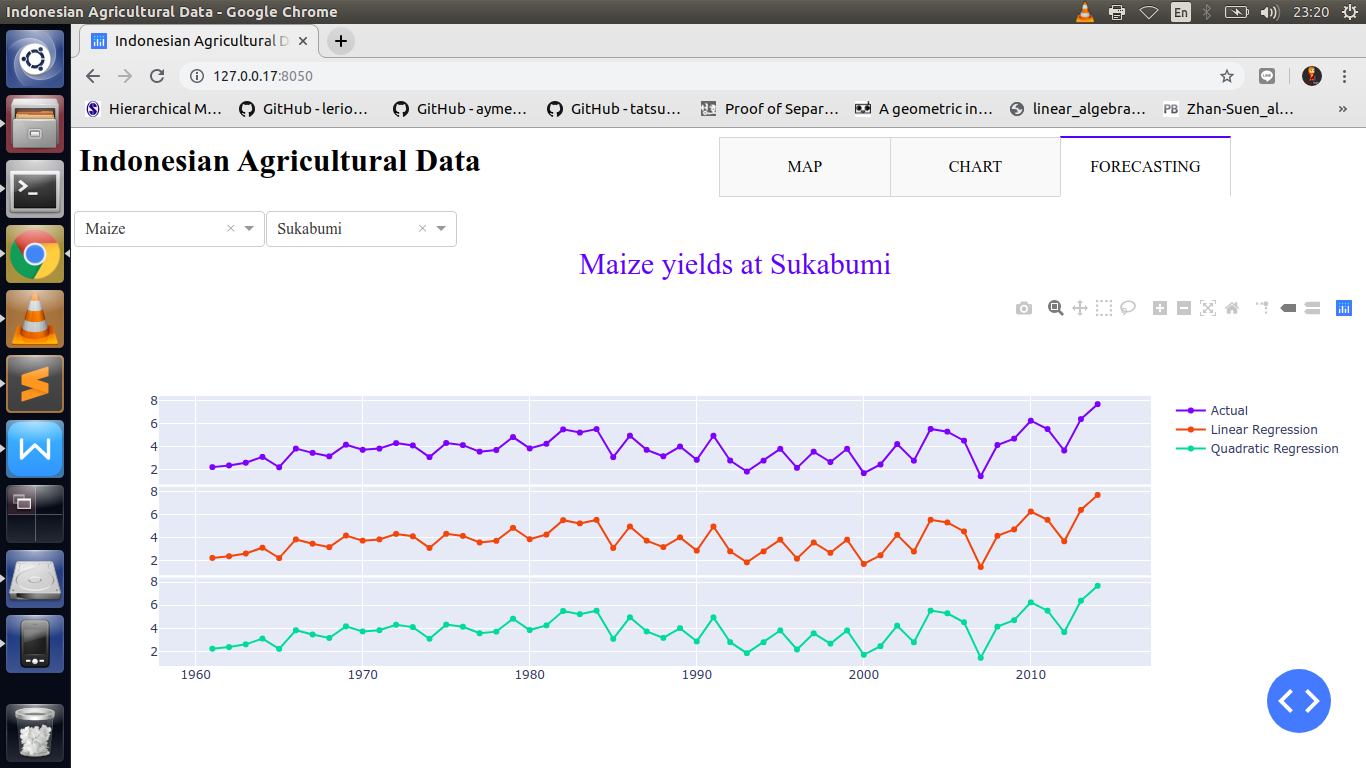
* 1. **Visualisasi Data**

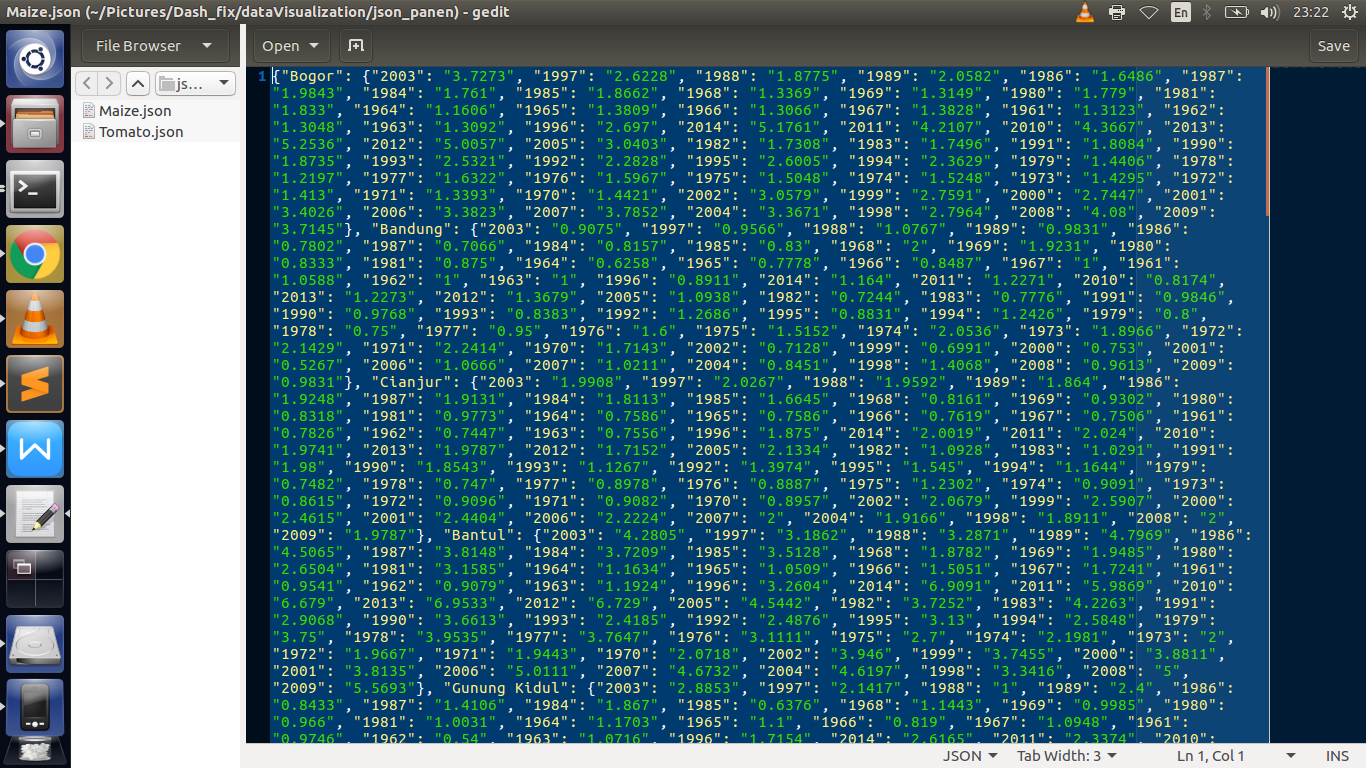
Bagian visualisasi data bertugas untuk menyajikan data angka - angka hasil perhitungan agar secara mudah diamati oleh kalangan manapun berdasarkan pola dari data tersebut.

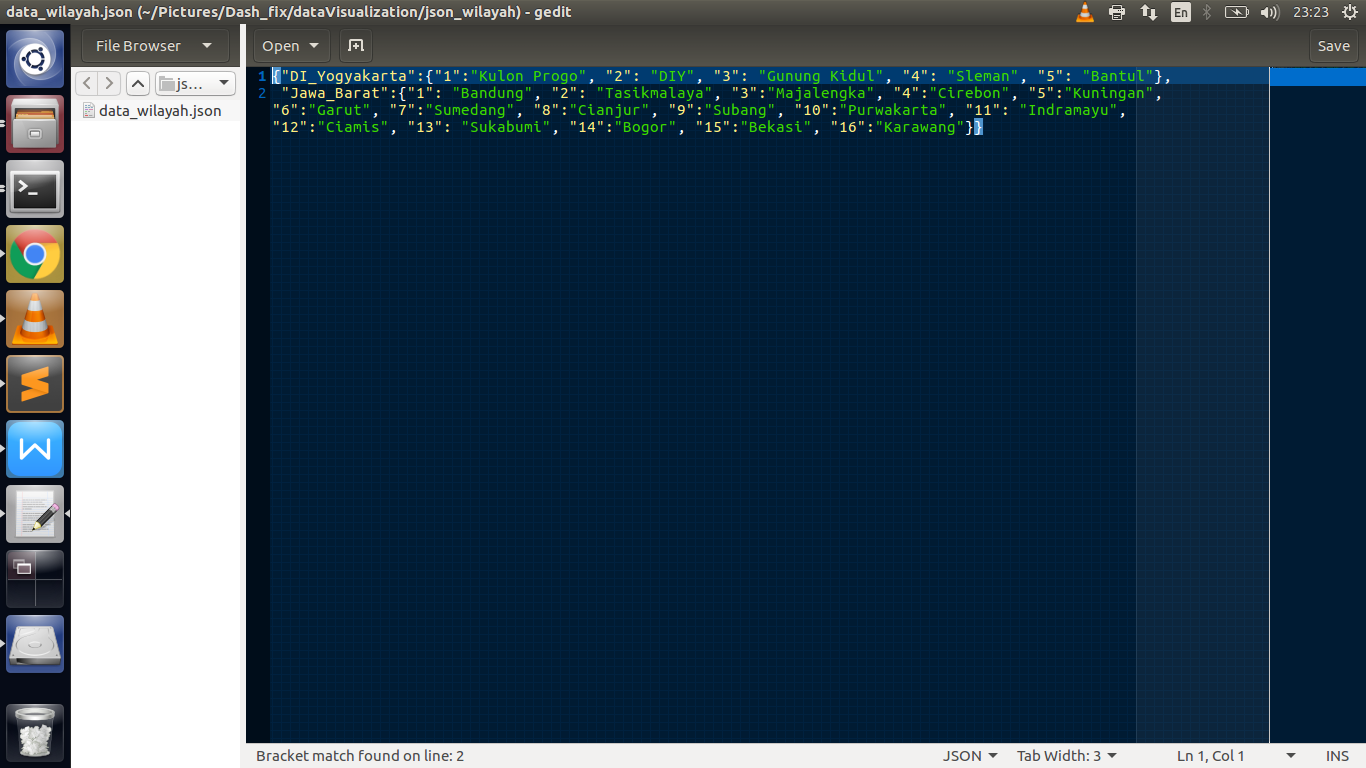
Pada implementasinya terdapat beberapa hal yang disajikan pada proses visualisasi data seperti dapat diamati dari gambar berikut:

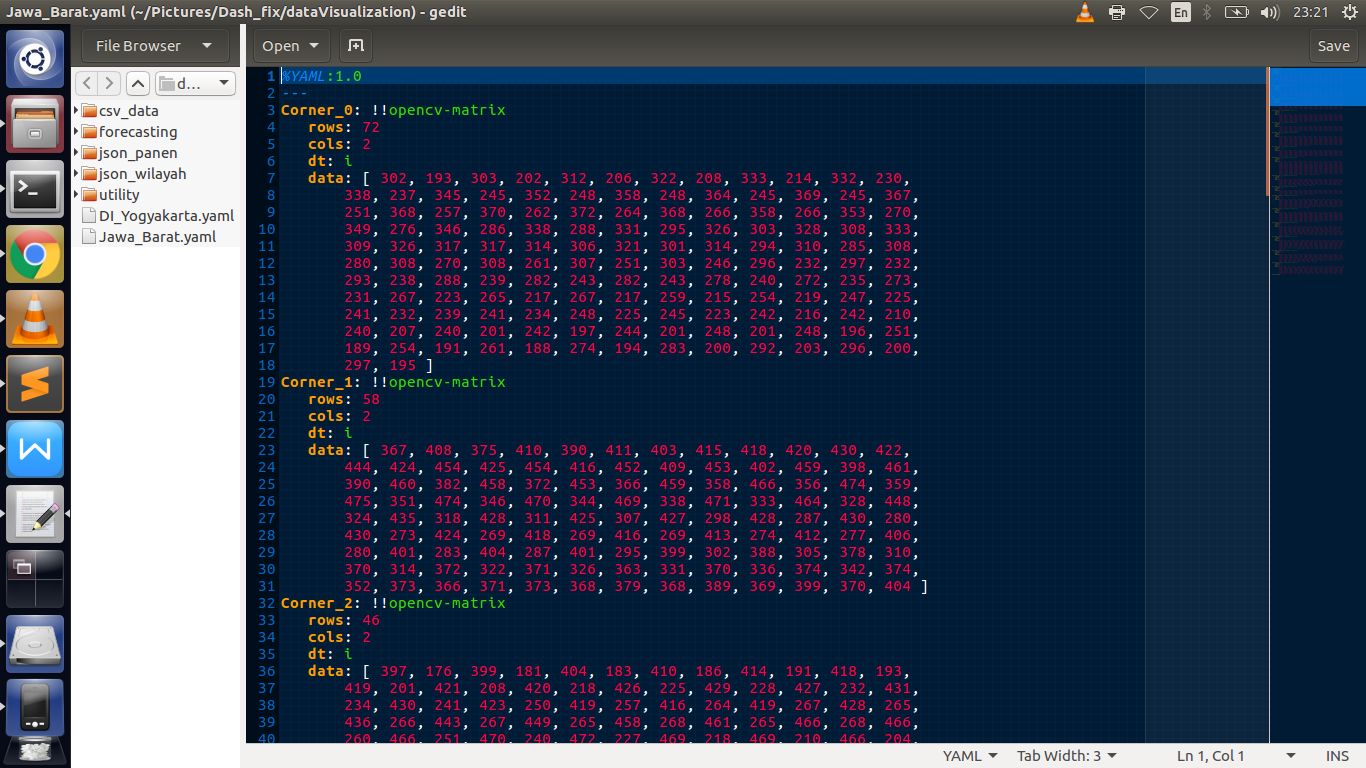




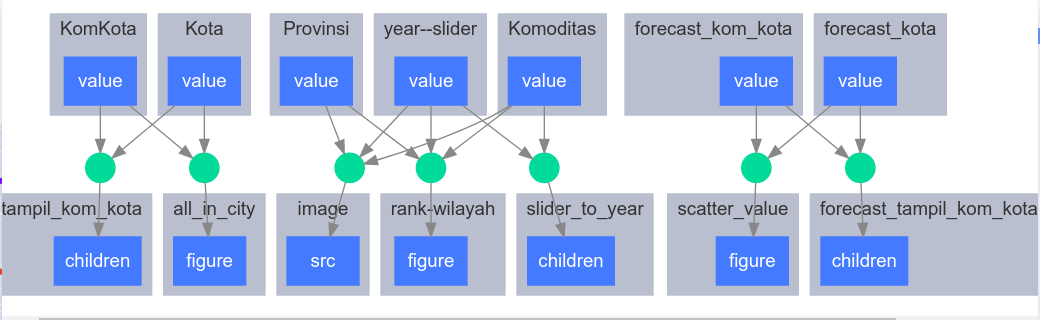












# Kesimpulan

Kesimpulan laporan