

**Analisa Korelasi Antara Emisi Gas Buang Terhadap
Pertumbuhan GDP Indonesia**



**UNIVERSITAS
TEKNOLOGI
SUMBAWA**

Dosen Pengampuh :

Herfandi, M.Kom

Disusun oleh :

- 1. Miftahul Khoir (20.01.013.082)**
- 2. Arif Annursida (20.01.013.045)**
- 3. M.Dimas Sakti M.(20.01.013.037)**
- 4. Aditya Fa'athir B.(20.01.013.035)**
- 5. Bayu Septian (20.01.013.040)**

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS REKAYASA SISTEM

UNIVERSITAS TEKNOLOGI SUMBAWA

2022

ABSTRAK

Dalam kehidupan manusia pasti hidup saling berhubungan satu sama lain entah itu hubungan keluarga,teman ataupun rekan kerja,contoh lainnya suatu perubahan variabel yang di ikuti variabel lainya karena memiliki hubungan atau korelasi antara variabel.

Dan untuk kesempatan kali ini kami akan menganalisa hubungan antara emisi gas buangterhadap pertumbuhan GDP Indonesia menggunakan data mining korelasi.

KATA PENGANTAR

Dengan memanjatkan puji syukur kehadiran Allah swt, karena atas berkat dan rahmat- Nya lah penyusun dapat menyelesaikan makalah yang berjudul **“Analisa Korelasi Antara Emisi Gas Buang Terhadap Pertumbuhan GDP Indonesia”** dengan tepat waktu. Tujuan utama penyusunan makalah ini adalah selain untuk memenuhi salah satu tugas mata kuliah Kecerdasan Buatan.

Dalam penyusunan makalah ini penyusun mengucapkan terima kasih kepada Dosen Pengampu mata kuliah „Kecerdasan Buatan“ Bapak Herfandi. Kedua orang tua kami yang senantiasa memberikan dukungan dan nasihatnya, serta sahabat-sahabat kami tercinta yang selalu memberikan dukungan serta semangatnya dalam penyusunan makalah ini.

Meskipun telah berusaha dengan segenap kemampuan, namun penyusun menyadari bahwa makalah ini masih belum sempurna. Oleh karena itu, segala tegur sapa, kritik, serta saran yang diberikan pembaca akan penyusun terima dengan kelapangan hati guna perbaikan pada masa yang akan datang. Akhir kata, penyusun berharap semoga makalah ini dapat memberikan manfaat bagi pembaca.

Sumbawa, 3 Januari 2022

Penulis

DAFTAR ISI

ABSTRAK.....	ii
DAFTAR ISI.....	iii
BAB I.....	1
PENDAHULUAN	1
2.1. Latar Belakang	1
2.2. Tujuan Penulisan.....	1
2.3. Rumusan Masalah.....	1
2.4. Metodologi Penyusunan.....	1
2.5. Manfaat Penulisan.....	2
BAB II LANDASAN TEORI.....	3
2.1. Data Mining	3
2.2. Algoritma Apriori.....	11
2.3. Analisis Pola Frekuensi.....	12
2.4. CROSS-INDUSTRY STANDARD PROCESS FOR DATA MINING (CRISP-DM) 12	
A. BUSINESS UNDERSTANDING.....	13
BAB III	14
METODE PENELITIAN.....	14
3.1. Metode Penelitian	14
3.2. Pengumpulan Data	14
Tabel 3.1 dataset	14
BAB IV	15
HASIL DAN PEMBAHASAN.....	15
4.1. Tahapan Analisa Korelasi Antara Emisi Gas Buang Terhadap Pertumbuhan GDP Indonesia.....	15
Gambar 4.1 Data Set Kaggle.com.....	15
Gambar 4.2 Data Set bersih	16
Gambar 4.3 Hasil Data Modeling	16
5.1 Kesimpulan	18
5.2 Saran.....	18
REFERENSI	19

BAB I

PENDAHULUAN

2.1. Latar Belakang

Dalam kehidupan tidak ada manusia yang bisa hidup sendiri, pasti tergantung atau berhubungan dengan yang lain. Baik itu berhubungan dengan sesama manusia, maupun dengan alam sekitar. Misalnya: Kalau kita ingin hidup sehat banyak faktor yang berkaitan/ berpengaruh, antara lain: lingkungan rumah, jam istirahat, jam kerja, cuaca dll. Konsep pemikiran tentang hubungan adalah untuk menjawab pertanyaan tentang apakah kemunculan suatu gejala akan diikuti oleh gejala-gejala lain, atau lebih spesifik apakah perubahan suatu variabel akan diikuti oleh perubahan variabel lain. Perubahan suatu variabel diikuti oleh perubahan variabel lain menandakan adanya hubungan (korelasi) antarvariabel.

2.2. Tujuan Penulisan

Adapun tujuan penulisannya adalah :

1. Menjelaskan konsep korelasi
2. Menjelaskan macam-macam korelasi dan perbedaannya
3. Menjelaskan teknik-teknik korelasi
4. Menjelaskan langkah - langkah teknik analisis korelasi
5. Menjelaskan rancangan teknik analisis korelasi

2.3. Rumusan Masalah

- Bagaimana menjelaskan konsep korelasi ?
- Bagaimana menjelaskan macam-macam korelasi ?
- Apakah perbedaannya dari setiap macam-macam korelasi ?
- Bagaimana menjelaskan teknik-teknik korelasi?
- Bagaimana menjelaskan langkah - langkah teknik analisis korelasi?
- Bagaimana menjelaskan rancangan teknik analisis korelasi?

2.4. Metodologi Penyusunan

Adapun metode yang digunakan dalam penyusunan makalah ini adalah studi pustaka.

Yakni dengan mengumpulkan sumber-sumber, baik dari buku ataupun internet tentang faktor- faktor yang mempengaruhi proses dan hasil belajar, yang kemudian kami gabungkan menjadi satu dalam satu makalah.

2.5. Manfaat Penulisan

Adapun setelah disusunnya makalah ini, kami berharap dapat bermanfaat bagi pembacasebagaimana yang kami jadikan tujuan. Yakni memberikan informasi dan pengetahuan tentang Analisis Korelasi, mengetahui beberapa masalah terkait dengan penggolongan dan macam-macamnya, serta terpenuhinya tugas mandiri mata kuliah Statistika Dasar.

BAB II

LANDASAN TEORI

2.1. Data Mining

Data mining telah menarik banyak perhatian dalam dunia sistem informasi dan dalam masyarakat secara keseluruhan dalam beberapa tahun terakhir, karena ketersediaan luas dalam jumlah besar data dan kebutuhan segera untuk mengubah data tersebut menjadi informasi yang berguna dan pengetahuan(Witten, 2012). Informasi dan pengetahuan yang diperoleh dapat digunakan untuk aplikasi mulai dari pasar analisis, deteksi penipuan, dan retensi pelanggan, untuk pengendalian produksi dan ilmu pengetahuan eksplorasi(Han&Kamber, 2007). Adanya ketersediaan data yang melimpah, kebutuhan akan informasi atau pengetahuan sebagai sarana pendukung dalam pengambilan keputusan baik bagi individu, organisasi, perusahaan dan pemerintahan.

Banyaknya data, ditambah dengan kebutuhan untuk alat analisis data yang kuat, telah digambarkan sebagai kaya data tapi miskin informasi. Jumlah data yang tumbuh secara cepat, dikumpulkan dan disimpan dalam repositori data yang besar dan banyak, telah jauh melampaui kemampuan manusia untuk memahami data-data tersebut tanpa mampu mengelolah data tersebut. Akibatnya, data yang dikumpulkan dalam repositori data yang besar menjadi "kuburan data" (Han&Kamber, 2007). Hal ini melatarbelakangi lahirnya suatu cabang ilmu pengetahuan baru yaitu *data mining*. *Data mining* adalah untuk mengekstrasikan atau "menambang" pengetahuan dari kumpulan banyak data(Han&Kamber, 2007).

Data mining adalah teknik yang merupakan gabungan metode-metode analisis data secara berkesinambungan dengan algoritma-algoritma untuk memproses data berukuran besar. Data mining merupakan proses menemukan informasi atau pola yang penting dalam

basis data berukuran besar dan merupakan kegiatan untuk menemukan informasi atau pengetahuan yang berguna secara otomatis dari data yang jumlahnya besar (Kusrini, 2009). *Data mining*, sering juga disebut *knowledge discovery in database (KDD)*, adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan pola keteraturan, pola hubungan dalam set data berukuran besar. Keluaran dari data mining ini dapat dijadikan untuk memperbaiki pengambilan keputusan di masa depan. Dalam *data mining* data disimpan secara elektronik dan diolah secara otomatis, atau setidaknya disimpan dalam komputer. Data mining adalah tentang menyelesaikan masalah dengan menganalisa data yang telah ada dalam database (Kusrini, 2009).

Siklus hidup proyek *data mining* menurut *Cross-Industry Standard Proses for Data Mining*(CRISP-DM)yang dikembangkan tahun 1996 terbagi dalam 6 fase (Kusrini, 2009). Berikut gambar dari Siklus hidup proyek *data mining*

1. Fase Pemahaman Bisnis (*Business Understanding Phase*)
 - a. Penentuan tujuan proyek dan kebutuhan secara detail dalam lingkup bisnis atau unit penelitian secara keseluruhan.
 - b. Menerjemahkan tujuan dan batasan menjadi formula dari permasalahan *data mining*.
 - c. Menyiapkan strategi awal untuk mencapai tujuan.
2. Fase pemahaman data (*Data Understanding Phase*)
 - a. Mengumpulkan data.
 - b. Menggunakan analisis penyelidikan data untuk mengenali lebih lanjut data dan pencarian pengetahuan awal.
 - c. Mengevaluasi kualitas data.
 - d. Jika diinginkan, pilih sebagian kecil group data yang mungkin mengandung pola dari permasalahan.
3. Fase pengolahan data (*Data Preparation Phase*)

Siapkan data awal, kumpulkan data yang akan digunakan untuk keseluruhan fase berikutnya. Fase ini merupakan pekerjaan berat yang perlu dilaksanakan secara intensif.

 - a. Pilih kasus dan variabel yang ingin dianalisis dan yang sesuai analisis yang akan dilakukan.
 - b. Lakukan perubahan pada beberapa variabel jika dibutuhkan.
 - c. Siapkan data awal sehingga siap untuk perangkat pemodelan.
4. Fase Pemodelan (*Modelling Phase*)
 - a. Pilih dan aplikasikan teknik pemodelan yang sesuai.
 - b. Perlu diperhatikan bahwa beberapa teknik mungkin untuk digunakan pada permasalahan *data mining* yang sama.

- c. Jika diperlukan, proses dapat kembali ke fase pengolahan data untuk menjadikan data ke dalam bentuk yang sesuai dengan spesifikasi kebutuhan teknik *data mining* tertentu.

5. Fase Evaluasi
 - a. Mengevaluasi satu atau lebih model yang digunakan dalam fase permodelan untuk mendapatkan kualitas dan efektivitas sebelum disebarkan untuk digunakan.
 - b. Menetapkan apakah terdapat model yang memenuhi tujuan pada fase awal.
 - c. Menentukan apakah terdapat permasalahan penting dari bisnis atau penelitian yang tidak tertangani dengan baik.
 - d. Mengambil keputusan yang berkaitan dengan penggunaan hasil dari *datamining*.
6. Fase Penyebaran
 - a. Menggunakan model yang dihasilkan. Terbentuknya model tidak menandakan telah terselesaikannya proyek.
 - b. Contoh sederhana penyebaran: pembuatan laporan.
 - c. Contoh kompleks penyebaran: penerapan proses *data mining* secara paralel pada departemen lain.

Menurut Larose dalam bukunya yang berjudul "*Discovering Knowledge in Data: An Introduction to Data Mining*", *datamining* dibagi menjadi beberapa kelompok berdasarkan tugas/pekerjaan yang dapat dilakukan (Larose, 2005), yaitu :

1. Deskripsi

Terkadang peneliti dan analisis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data. Deskripsi dari pola kecenderungan sering memberikan kemungkinan penjelasan untuk suatu pola atau kecenderungan.

2. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih ke arah numerik daripada ke arah kategori. Model dibangun menggunakan baris data (*record*) lengkap yang

menyediakan nilai dari variabel target sebagai nilai prediksi. Selanjutnya, pada peninjauan berikutnya estimasi nilai dari variabel target dibuat berdasarkan nilai variabel prediksi.

3. Prediksi

Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang. Beberapa metode dan teknik yang digunakan dalam klasifikasi dan estimasi dapat pula digunakan (untuk keadaan yang tepat) untuk prediksi.

4. Klasifikasi

Dalam klasifikasi, terdapat target variabel kategori. Sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah.

5. Pengklasteran (*Clusterring*)

Pengklasteran merupakan pengelompokan *record*, pengamatan, atau memperhatikan dan membentuk kelas obyek-obyek yang memiliki kemiripan. Klaster adalah kumpulan *record* yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan *record* dalam klaster yang lain. Berbeda dengan klasifikasi, pada pengklasteran tidak ada variabel target. Pengklasteran tidak melakukan klasifikasi, mengestimasi, atau memprediksi nilai dari variabel target, akan tetapi, algoritma pengklasteran mencoba untuk melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang memiliki kemiripan (homogen), yang mana kemiripan *record* dalam satu kelompok akan bernilai maksimal, sedangkan kemiripan dengan *record* dalam kelompok lain akan bernilai minimal.

6. Asosiasi

Tugas asosiasi dalam *data mining* adalah untuk menemukan atribut yang muncul dalam satu waktu. Salah satu implementasi dari asosiasi adalah *market basket analysis* atau analisis keranjang belanja, sebagaimana yang akan dibahas dalam penelitian ini.

Dalam buku “Data Mining – Meramalkan Bisnis Perusahaan, Feri Sulianta & Dominikus Juju, 2010” menyatakan bahwa secara terperinci, ada empat tahap yang dilalui dalam *Data Mining* antara lain :

1. Tahap pertama:

Precise statement of the problem (mendefinisikan permasalahan yang ingin diketahui). Misalnya ingin mengetahui apakah seorang customer berpotensi memiliki kredit macet, atau

mengidentifikasi seorang customer apakah akan pindah ke kompetitor bisnis kita, dan lain sebagainya. Setelah menemukan pertanyaan bisnis yang perlu dijawab oleh data mining, selanjutnya tentukan tipe tugas dasar untuk menjawab pertanyaan bisnis tersebut. Tugas dasar yang menjadi dasar algoritma data mining adalah klasifikasi, regresi, segmentasi, asosiasi dan korelasi.

2. Tahap kedua:

Initial Exploration (Mempersiapkan data yang menjadi sumber untuk *data mining* termasuk data “*cleaning*” untuk dipelajari polanya). Setelah menentukan definisi masalah, langkah berikutnya adalah mencari data yang mendukung definisi masalah. Menentukan porsi data yang digunakan men-training data mining berdasarkan

algoritma data mining yang telah dibuat. Setelah persiapan data selesai dilakukan, langkah berikutnya adalah memberikan sebagian data kedalam algoritma data mining.

3. Tahap Ketiga:

Model building and validation. Validasi apakah data mining memberikan prediksi yang akurat. Setelah *training* data selesai dilakukan, data mining tersebut perlu di-“uji” atau di-validasi keakuratannya terhadap data testing.

4. Tahap Ke-empat:

Deployment. Tahap ini memilih aplikasi yang tepat terhadap *data mining* untuk membuat prediksi

2.2. Algoritma Apriori

Algoritma apriori atau sering disebut juga dengan analisis asosiasi (*association rule mining*) adalah teknik data mining untuk menemukan aturan asosiasi antara suatu kombinasi item. (Han & Kamber, 2006) Misalkan $I = \{i_1, i_2, i_3, \dots, i_n\}$ merupakan himpunan items dan D merupakan himpunan database transaksi yang setiap transaksi T merupakan himpunan item yang terkandung dalam transaksi T . Setiap transaksi diasosiasikan dengan *identifier* yang disebut TID. Aturan yang memenuhi *minimum support* (\min_sup) dan *minimum confidence* (\min_conf) disebut kandidat rule yang kuat. Berdasarkan konvensi, nilai *support* dan nilai *confidence* adalah nilai antara 0% sampai 100% sebanding dengan 0 sampai 1,0 (Linof, 2004).

Himpunan item disebut dengan itemset. Itemset yang mengandung k items disebut k - itemsets. Sebagai contoh himpunan {kopi, gula} merupakan 2-itemset. Kecenderungan kemunculan itemset dalam sejumlah transaksi disebut *frequency*, *support count* atau *count itemset*..

2.3. Analisis Pola Frekuensi

Tahap ini mencari kombinasi *item* yang memenuhi syarat minimum dari nilai *support* dalam basis data. Nilai *support* sebuah *item* diperoleh dengan menggunakan rumus berikut:

$$\text{support}(A) = \frac{\text{jumlah transaksi mengandung } A}{\text{total transaksi}}$$

Sementara, nilai *support* dari 2 *item* diperoleh dengan menggunakan rumus

$$\text{support}(A, B) = \frac{P(A \cap B)}{\sum \text{transaksi mengandung } A \text{ dan } B} \\ \text{support}(A, B) = \frac{\sum \text{transaksi mengandung } A \text{ dan } B}{\sum \text{transaksi}}$$

Frequent itemset menunjukkan *itemset* yang memiliki frekuensi kemunculan lebih dari nilai minimum yang ditentukan (). Misalkan = 2, maka semua *itemsets* yang frekuensi kemunculannya lebih dari atau sama dengan 2 kali disebut *frequent*. Himpunan dari *frequent k-itemset* dilambangkan dengan F_k .

Setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiasi yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* aturan asosiatif A B. Nilai *confidence* dari aturan A B diperoleh dengan rumus berikut :

$$\text{confidence} = P(B|A) \\ = \frac{\sum \text{transaksi mengandung } A \text{ dan } B}{\sum \text{transaksi mengandung } A}$$

2.4. CROSS-INDUSTRY STANDARD PROCESS FOR DATA MINING (CRISP-DM)

Cross-Industry Standard Process for Data Mining atau CRISP-DM adalah salah satu model proses datamining (datamining framework) yang awalnya (1996) dibangun oleh 5 perusahaan yaitu Integral Solutions Ltd (ISL), Teradata, Daimler AG, NCR Corporation dan OHRA. Framework ini kemudian dikembangkan oleh ratusan organisasi dan perusahaan di Eropa untuk dijadikan methodology standard non-proprietary bagi data mining. Adapun tahapan CRISP-DM dibagi menjadi enam yaitu Business Understanding, Data Understanding, Data Preparation, Modeling,

Evaluation, dan Deployment. Berikut ini penjelasan singkat mengenai tahapan CRISP-DM:

A. BUSINESS UNDERSTANDING

Ini adalah tahap pertama dalam CRISP-DM dan termasuk bagian yang cukup vital. Pada tahap ini membutuhkan pengetahuan dari objek bisnis, bagaimana membangun atau mendapatkan data, dan bagaimana untuk mencocokkan tujuan pemodelan untuk tujuan bisnis sehingga model terbaik dapat dibangun.

B. DATA UNDERSTANDING

Secara garis besar untuk memeriksa data, sehingga dapat mengidentifikasi masalah dalam data. Tahap ini memberikan fondasi analitik untuk sebuah penelitian dengan membuat ringkasaan (summary) dan mengidentifikasi potensi masalah dalam data. Tahap ini juga harus dilakukan secara cermat dan tidak terburu-buru, seperti pada visualisasi data, yang terkadang insight-nya sangat sulit didapat jika dihubungkan dengan summary data nya. Jika ada masalah pada tahap ini yang belum terjawab, maka akan mengganggu pada tahap modeling.

C. DATA PREPARATION

Secara garis besar untuk memperbaiki masalah dalam data, kemudian membuat variabel derived. Tahap ini jelas membutuhkan pemikiran yang cukup matang dan usaha yang cukup tinggi untuk memastikan data tepat untuk algoritma yang digunakan. Bukan berarti saat Data Preparation pertama kali dimana masalah-masalah pada data sudah diselesaikan, data sudah dapat digunakan hingga tahap terakhir.

D. MODELING

Secara garis besar untuk membuat model prediktif atau deskriptif. Pada tahap ini dilakukan metode statistika dan Machine Learning untuk penentuan terhadap teknik data mining, alat bantu data mining, dan algoritma data mining yang akan diterapkan. Lalu selanjutnya adalah melakukan penerapan teknik dan algoritma data mining tersebut kepada data dengan bantuan alat bantu.

BAB III

METODE PENELITIAN

3.1. Metode Penelitian

Pada penelitian ini, data yang digunakan adalah data emisi gas buang . Data tersebut akan diolah menggunakan metode algoritma Apiori . Berikut tahapan dalam melakukan penelitian data mining.

1. Pengumpulan data Tahapan ini menerangkan tentang darimana sumber data dalam penelitian inididapatkan dan menemukan informasi yang bisa digunakan untuk penelitian.
2. Pengolahan data awal Tahapan ini menerangkan tentang tahap awal dalam data mining. Pengolahanawal data meliputi proses input data keformat yang dibutuhkan, penggabungan data dan training data.
3. Metode yang diusulkan Tahapan ini dijelaskan pemilihan dan penggunaan metode algoritma Apioripada penelitian.
4. Pengujian dan Validasi hasil Tahapan ini menjelaskan tentang pengujian, hasil pengujian akan di validasi dan kemudian di evaluasi. Penjelasan mengenai hal ini akan di paparkan pada BAB IV.

3.2. Pengumpulan Data

3.2.1. Sumber Data

Data yang digunakan dalam penelitian ini adalah data dengan total 225 data . Setaip baris menandakan bahwa data berada dalam satu transaksi yang sama.

Tabel 3.1 dataset

1	CountryCode,ShortName,TableName,LongName,Alpha2Code,CurrencyUnit,SpecialNotes,Region,IncomeGroup,Wb2Code,NationalAccountsBaseYear,NationalAccountsReferenceYear,SnaPriceValuation,LendingCate
2	AFG,Afghanistan,Afghanistan,Islamic State of Afghanistan,AF,Afghan afghani,Fiscal year end: March 20
3	ALB,Albania,Albania,Republic of Albania,AL,Albanian lek,,Europe & Central Asia,Upper middle income,AL,Original chained constant price data are rescaled,,1996,Value added at basic prices (VAB),IBRD,,Country uses
4	DZA,Algeria,Algeria,People's Democratic Republic of Algeria,DZ,Algerian dinar,,Middle East & North Africa,Upper middle income,DZ,1980,,Value added at basic prices (VAB),IBRD,,Country uses the 1968 System of Na
5	ASM,American Samoa,American Samoa,American Samoa,AS,U.S. dollar,,East Asia & Pacific,Upper middle income,AS,,,,,,Country uses the 1968 System of National Accounts methodology,,2011 (household consumpti
6	ADO,Andorra,Andorra,Principality of Andorra,AD,Euro,,Europe & Central Asia,High income: nonOECD,AD,2000,,Value added at basic prices (VAB),,,Country uses the 1968 System of National Accounts methodology,,...
7	AGO,Angola,Angola,People's Republic of Angola,AO,Angolan kwanza,"April 2013 database update: Based on IMF data, national accounts data were revised for 2000 onward
8	ATG,Antigua and Barbuda,Antigua and Barbuda,AG,East Caribbean dollar,"April 2012 database update: Based on official government statistics, national accounts data were revised for 2000 onward
9	ARB,Arab World,Arab World,Arab World,JA,,Arab World aggregate. Arab World is composed of members of the League of Arab States,,,1A,,,,,,,,,,,,,
10	ARG,Argentina,Argentina,Argentine Republic,AR,Argentine peso,The base year has changed to 2004,,Latin America & Caribbean,High income: nonOECD,AR,2004,,Value added at basic prices (VAB),IBRD,,Country uses
11	ARM,Armenia,Armenia,Republic of Armenia,AM,Armenian dram,,Europe & Central Asia,Lower middle income,AM,Original chained constant price data are rescaled,,1996,Value added at basic prices (VAB),IBRD,Cou
12	ABW,Aruba,Aruba,AW,Aruban florin,SNA data for 2000-2011 are updated from official government statistics
13	AUS,Australia,Australia,Commonwealth of Australia,AU,Australian dollar,Fiscal year end: June 30
14	AUT,Austria,Austria,Republic of Austria,AT,Euro,"A simple multiplier is used to convert the national currencies of EMU members to euros. The following irrevocable euro conversion rate was adopted by the EU Coun
15	AZE,Azerbaijan,Azerbaijan,Republic of Azerbaijan,AZ,New Azeri manat,April 2012 database update: National accounts historical expenditure series in constant prices were revised in line with State Statistical Commi
16	BHR,Bahrain,Bahrain,Kingdom of Bahrain,BH,Bahraini dinar,Based on official government statistics
17	BGD,Bangladesh,Bangladesh,People's Republic of Bangladesh,BD,Bangladeshi taka,Fiscal year end: June 30
18	BRB,Barbados,Barbados,Barbados,BB,Barbados dollar,,Latin America & Caribbean,High income: nonOECD,BB,1974,,Value added at producer prices (VAP),,,Country uses the 1993 System of National Accounts methodo
19	BLR,Belarus,Belarus,Republic of Belarus,BY,Belarusian rubel,,Europe & Central Asia,Upper middle income,BY,Original chained constant price data are rescaled,,2000,Value added at basic prices (VAB),IBRD,,Country u
20	BEL,Belgium,Belgium,Kingdom of Belgium,BE,Euro,"A simple multiplier is used to convert the national currencies of EMU members to euros. The following irrevocable euro conversion rate was adopted by the EU Cc
21	BLZ,Belize,Belize,Belize,BZ,Belize dollar,,Latin America & Caribbean,Upper middle income,BZ,2000,,Value added at basic prices (VAB),IBRD,,Country uses the 1993 System of National Accounts methodology,,2011,"I
22	BEN,Benin,Benin,Republic of Benin,BJ,West African CFA franc,,Sub-Saharan Africa,Low income,BJ,2007,,Value added at basic prices (VAB),IDA,HIPC,Country uses the 1968 System of National Accounts methodology,,...
23	BMU,Bermuda,Bermuda,The Bermudas,BM,Bermuda dollar,,North America,High income: nonOECD,BM,2006,,Value added at basic prices (VAB),,,Country uses the 1993 System of National Accounts methodology,,20
24	BTN,Bhutan,Bhutan,Kingdom of Bhutan,BT,Bhutanese ngultrum,April 2013 database update: Data were updated using the government of Bhutan macroeconomic framework,,South Asia,,Lower middle income,BT,200
25	BOL,Bolivia,Bolivia,Plurinational State of Bolivia,BO,Bolivian Boliviano,,Latin America & Caribbean,Lower middle income,BO,1990,,Value added at basic prices (VAB),Blend,HIPC,Country uses the 1968 System of Nati
26	BIH,Bosnia and Herzegovina,Bosnia and Herzegovina,BA,Bosnia and Herzegovina convertible mark,Based on official government statistics for chain linked series
27	BWA,Botswana,Botswana,Republic of Botswana,BW,Botswana pula,"Fiscal year end: March 31

3.3. Data preparation

Pada laporan ini merupakan dataset yang kami teliti sudah sudah merupakan data yang telah di bersikan Sehingga yang di sajikan sudah lengkap dan tidak memiliki data yang missing

BAB IV

HASIL DAN PEMBAHASAN

4.1. Tahapan Analisa Korelasi Antara Emisi Gas Buang Terhadap Pertumbuhan GDP Indonesia

4.1.1. BUSINESS UNDERSTANDING

Adapun latar belakang pembuatan laporan ini adalah untuk mengendalikan dengan cara Analisa Korelasi Antara Emisi Gas Buang Terhadap Pertumbuhan GDP Indonesia

4.1.2. DATA UNDERSTANDING

Data yang akan digunakan ini bersumber dari Kaggle.com. .
Adapaun bentuk awal dari data kotornya sebagai berikut.

	CountryCode,ShortName,TableName,LongName,Alpha2Code,CurrencyUnit,SpecialNotes,Region,IncomeGroup,Wb2Code,NationalAccountsBaseYear,NationalAccountsReferenceYear,SnaPriceValuation,LendingCate
1	AFG,Afghanistan,Afghanistan,Islamic State of Afghanistan,AF,Afghan afghani,Fiscal year end: March 20
2	ALB,Albania,Albania,Republic of Albania,AL,Albanian lek,,Europe & Central Asia,Upper middle income,AL,Original chained constant price data are rescaled.,1996,Value added at basic prices (VAB),IBRD,,Country uses the 1968 System of National Accounts methodology,...
3	DZA,Algeria,Algeria,People's Democratic Republic of Algeria,DZ,Algerian dinar,,Middle East & North Africa,Upper middle income,DZ,1980,,Value added at basic prices (VAB),IBRD,,Country uses the 1968 System of National Accounts methodology,...
4	ASM,American Samoa,American Samoa,AS,U.S. dollar,,East Asia & Pacific,Upper middle income,AS,,,,Country uses the 1968 System of National Accounts methodology,2011,(household consumption expenditure in constant prices of 2007),...
5	ADO,Andorra,Andorra,Principality of Andorra,AD,Euro,,Europe & Central Asia,High income: nonOECD,AD,2000,,Value added at basic prices (VAB),Country uses the 1968 System of National Accounts methodology,...
6	AGO,Angola,Angola,People's Republic of Angola,AO,Angolan kwanza,"April 2013 database update: Based on IMF data, national accounts data were revised for 2000 onward
7	ATG,Antigua and Barbuda,Antigua and Barbuda,AG,East Caribbean dollar,"April 2012 database update: Based on official government statistics, national accounts data were revised for 2000 onward
8	ARB,Arab World,Arab World,1A,,Arab World aggregate. Arab World is composed of members of the League of Arab States.,1A,,Value added at basic prices (VAB),IBRD,,Country uses the 1968 System of National Accounts methodology,...
9	ARG,Argentina,Argentina,Argentine Republic,AR,Argentine peso,The base year has changed to 2004.,Latin America & Caribbean,High income: nonOECD,AR,2004,,Value added at basic prices (VAB),IBRD,,Country uses the 1968 System of National Accounts methodology,...
10	ARM,Armenia,Armenia,Republic of Armenia,AM,Armenian dram,,Europe & Central Asia,Lower middle income,AM,Original chained constant price data are rescaled.,1996,Value added at basic prices (VAB),IBRD,,Country uses the 1968 System of National Accounts methodology,...
11	ABW,Aruba,Aruba,Aruba,AW,Aruban florin,SNA data for 2000-2011 are updated from official government statistics
12	AUS,Australia,Australia,Commonwealth of Australia,AU,Australian dollar,Fiscal year end: June 30
13	AUT,Austria,Austria,Republic of Austria,AT,Euro,"A simple multiplier is used to convert the national currencies of EMU members to euros. The following irrevocable euro conversion rate was adopted by the EU Council: 166,376.036 Austrian schillings = 1 euro.
14	AZE,Azerbaijan,Azerbaijan,Republic of Azerbaijan,AZ,New Azeri manat,April 2012 database update: National accounts historical expenditure series in constant prices were revised in line with State Statistical Committee data
15	BHR,Bahrain,Bahrain,Kingdom of Bahrain,BH,Bahraini dinar,Based on official government statistics
16	BGD,Bangladesh,Bangladesh,People's Republic of Bangladesh,BD,Bangladeshi taka,Fiscal year ends: June 30
17	BRB,Barbados,Barbados,Barbados,BB,Barbados dollar,,Latin America & Caribbean,High income: nonOECD,BB,1974,,Value added at producer prices (VAP),Country uses the 1993 System of National Accounts methodology,...
18	BLR,Belarus,Belarus,Republic of Belarus,BY,Belarusian rubel,,Europe & Central Asia,Upper middle income,BY,Original chained constant price data are rescaled.,2000,Value added at basic prices (VAB),IBRD,,Country uses the 1968 System of National Accounts methodology,...
19	BEL,Belgium,Belgium,Kingdom of Belgium,BE,Euro,"A simple multiplier is used to convert the national currencies of EMU members to euros. The following irrevocable euro conversion rate was adopted by the EU Council: 166,376.036 Belgian francs = 1 euro.
20	BLZ,Belize,Belize,Belize,BZ,Belize dollar,,Latin America & Caribbean,Upper middle income,BZ,2000,,Value added at basic prices (VAB),IBRD,,Country uses the 1993 System of National Accounts methodology,2011,(household consumption expenditure in constant prices of 2007),...
21	BDN,Benin,Benin,Republic of Benin,BI,West African CFA franc,,Sub-Saharan Africa,Low income,BI,2007,,Value added at basic prices (VAB),IDA,IFC,Country uses the 1968 System of National Accounts methodology,...
22	BMU,Bermuda,Bermuda,The Bermuda Islands,BM,Bermuda dollar,,North America,High income: nonOECD,BM,2006,,Value added at basic prices (VAB),Country uses the 1993 System of National Accounts methodology,2006
23	BTN,Bhutan,Bhutan,Kingdom of Bhutan,BT,Bhutanese ngultrum,April 2013 database update: Data were updated using the government of Bhutan macroeconomic framework.,South Asia,Lower middle income,BT,2006
24	BOL,Bolivia,Bolivia,Plurinational State of Bolivia,BO,Bolivian Boliviano,,Latin America & Caribbean,Lower middle income,BO,1990,,Value added at basic prices (VAB),Blend,IFC,Country uses the 1968 System of National Accounts methodology,...
25	BHN,Bosnia and Herzegovina,Bosnia and Herzegovina,BA,Bosnia and Herzegovina convertible mark,Based on official government statistics for chain linked series
26	BWA,Botswana,Botswana,Republic of Botswana,BW,Botswana pula,"Fiscal year end: March 31

Gambar 4.1 Data Set Kaggle.com

4.1.3 DATA PREPARATION

Pada tahap ini, datanya sudah menjadi data bersih yang di masukan di rapitminer. Adapun bentuk datanya sebagai berikut :

File Edit Process View Connections Settings Extensions Help

Views Design Results Turbo Prep Auto Model Deployments

Result History **ExampleSet (Read CSV)**

Open in Turbo Prep Auto Model

Filter (247 / 247 examples): all

Row No.	CountryCode	ShortName	TableName	LongName	Alpha2Code	CurrencyUnit	SpecialNotes	Regi
1	AFG	Afghanistan	Afghanistan	Islamic State	AF	Alghan afghani	Fiscal year e	Sout
2	ALB	Albania	Albania	Republic of AL	AL	Albanian lek	?	Euro
3	DZA	Algeria	Algeria	People's De	DZ	Algerian dinar	?	Midd
4	ASM	American Sa	American Sa	American Sa	AS	U.S. dollar	?	East
5	ADO	Andorra	Andorra	Principality of	AD	Euro	?	Euro
6	AGO	Angola	Angola	People's Rep.	AO	Angolan kwa	April 2013 dat	Sub
7	ATG	Antigua and	Antigua and	Antigua and	AG	East Caribbe	April 2012 dat	Latr
8	ARB	Arab World	Arab World	Arab World	1A	?	Arab World a	?
9	ARG	Argentina	Argentina	Argentine Re	AR	Argentine peso	The base yea	Latr
10	ARM	Armenia	Armenia	Republic of A	AM	Armenian dra	?	Euro
11	ABW	Aruba	Aruba	Aruba	AW	Aruban florin	SHA data for	Latr
12	AUS	Australia	Australia	Commonwea	AU	Australian dol	Fiscal year e	East
13	AUT	Austria	Austria	Republic of A	AT	Euro	A simple mult	Euro
14	AZE	Azerbaijan	Azerbaijan	Republic of A	AZ	New Azari ma	April 2012 dat	Euro

ExampleSet (247 examples, 0 special attributes, 31 regular attributes)

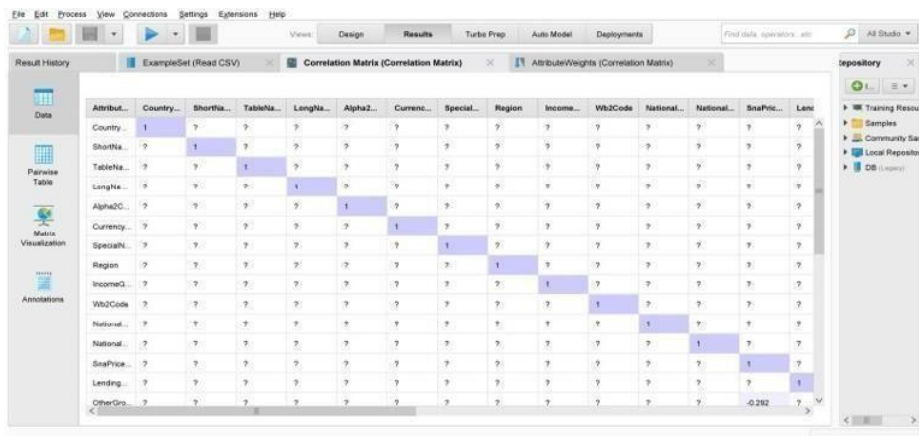
Gambar 4.2 Data Set bersih

4.1.4 MODELLING

Pada tahap modeling ini, kita akan mengatur model dari data di aplikasi rapidminer dan kita bisa melihat hasil grafiknya. Adapun bentuknya sebagai berikut :



Gambar 4.3 Hasil Data Modeling



Gambar 4.3 Hasil Grafik Data

BAB V

PENUTUP

5.1 Kesimpulan

Setelah membaca pembahasan diatas kami menyimpulkan bahwa korelasi sebagai salahsatu metode *data mining* berguna untuk menjelajahi data baru yang dihasilkan dari korelasi – korelasi pada pembahasan diatas yang dapat dijadikan pengetahuan baru.

5.2 Saran

Pada pembahasan diatas dapat dilihat bahwa dataset yang digunakan bersih dan tidak memerlukan proses *cleaning*, namun jumlah data yang digunakan dalam pembuatan laporanini masih kurang
Berdasarkan kesimpulan diatas, untuk memperoleh suatu data baru yang kuat, diperlukan penambahan data pada dataset yang digunakan saat ini

REFERENSI

Dawson, C. W. (2009). *Projects in Computing and Information System A Student's Guide*.

England: Addison-Wesley.

Gunadi - Penerapan Metode Data Mining Market Basket Analysis terhadap Penjualan Bukudengan Algoritma Apriori – 2012

Han, J., & Kamber, M. , 2007,*Data Mining Concepts and Technique*. Morgan Kaufmannpublisher

Julsam dkk, Penambangan Data Format TextExcel dengan Software Tanagra, 2009. Kothari, C. R. (2004). *Research Methology methodes and Technique*. India: New Age

Interntional.

Kusrini, dan Luthfi, E. T. (2007). *Algoritma Data*

mining. Yogyakarta: Andi.Kusrini, dan Luthfi, E. T.

(2009). *Algoritma Data mining*. Yogyakarta: Andi.

Kurniawati - Pemetaan Pola Hubungan Program Studi dengan Algoritma Apriori – 2014.Larose, D. T. (2005). *Discovering Knowledge in Data*. Canada: Wiley Interscience.

Nurchahyo - Penerapan Data Mining dengan Algoritma Apriori untuk Mendukung StrategiPromosi Pendidikan

Vercellis, C. (2009). *Business Intelligence : Data Mining and Optimization for DecisionMaking*. John Wiley dan Sons, Ltd

Witten, H. I., Eibe, F., dan Hall, A. M. (2011). *Data Mining Machine Learning Tools andTechiques*. Burlington: Morgan Kaufmann Publisher.

<https://www.kaggle.com/ardisragen/indonesia-gdp-cases>.