

Dr. Baihaqi Siregar, S.Si., M.T.



PENGANTAR ANALISIS BIG DATA

UNTUK PEMBANGUNAN EKONOMI



PENGANTAR ANALISIS BIG DATA

UNTUK PEMBANGUNAN EKONOMI

Dr. Baihaqi Siregar, S.Si., M.T.



PENGANTAR ANALISIS *BIG DATA* UNTUK PEMBANGUNAN EKONOMI

Penulis:

Dr. Baihaqi Siregar, S.Si., M.T.

Desain Cover:

Andi Juliandi

Sumber Ilustrasi:

www.freepik.com

Tata Letak:

Handarini Rohana

Editor:

Neng Rismawati

ISBN:

978-634-246-844-9

Cetakan Pertama:

Mei, 2026

Hak Cipta Dilindungi Oleh Undang-Undang

by Penerbit Widina Media Utama

Dilarang keras menerjemahkan, memfotokopi, atau memperbanyak sebagian atau seluruh isi buku ini tanpa izin tertulis dari Penerbit.

PENERBIT:

WIDINA MEDIA UTAMA

Komplek Puri Melia Asri Blok C3 No. 17 Desa Bojong Emas
Kec. Solokan Jeruk Kabupaten Bandung, Provinsi Jawa Barat

Anggota IKAPI No. 519/JBA/2025

Website: www.penerbitwidina.com

Instagram: [@penerbitwidina](https://www.instagram.com/penerbitwidina)

Telepon (022) 87355370

KATA PENGANTAR

Puji syukur ke hadirat Allah subhanahu wa ta'ala, karena atas rahmat dan karunia-Nya, buku ajar ***Pengantar Analisis Big Data untuk Pembangunan Ekonomi*** ini dapat diselesaikan dan hadir di hadapan para pembaca.

Ilmu ekonomi sedang mengalami transformasi metodologis terbesar sejak penemuan konsep Produk Domestik Bruto (PDB) pada era 1930-an. Hari ini, denyut nadi perekonomian tidak lagi hanya terekam dalam lembaran kuesioner sensus tahunan yang lambat dan mahal. *Realitas* ekonomi kita kini berdenyut secara *real-time* di dalam jutaan cuitan media sosial, terpotret dalam miliaran piksel citra satelit di malam hari, dan mengalir melalui jejak GPS serta transaksi dompet digital masyarakat.

Sayangnya, ketika mendengar istilah "*Big Data*" atau "*Machine Learning*", banyak mahasiswa ilmu ekonomi sosial merasa terintimidasi. Ada miskonsepsi bahwa ranah ini adalah wilayah eksklusif milik mahasiswa Ilmu Komputer atau *Data Science*. Buku ini ditulis secara khusus untuk mendobrak batasan tersebut.

Buku ajar ini disusun untuk Anda—para mahasiswa sarjana Ekonomi Pembangunan. Anda tidak akan dituntut untuk menjadi seorang *programmer* atau *software engineer* dalam semalam. Sebaliknya, buku ini akan melatih Anda menjadi seorang **Penerjemah dan Arsitek Kebijakan**. Algoritma komputer yang paling canggih sekalipun tidak tahu masalah sosial apa yang harus dipecahkan. Komputer tidak mengerti apa itu kemiskinan struktural, eksternalitas negatif, atau jebakan kelas menengah (*middle-income trap*). Mereka membutuhkan *Anda*—seorang ekonom yang memahami teori kausalitas—untuk mengajukan pertanyaan yang tepat dan memastikan bahwa data tersebut digunakan secara etis untuk hajat hidup orang banyak.

Melalui 17 bab yang terstruktur secara komprehensif, buku ini akan memandu Anda dari pemahaman filosofis dasar hingga aplikasi teknis di dunia nyata. Anda akan belajar bagaimana:

- Mengekstraksi jutaan teks *lowongan* kerja menjadi indikator kesenjangan keterampilan (*skills gap*).

- Mengubah intensitas cahaya dari citra satelit menjadi peta kemiskinan resolusi tinggi.
- Memanfaatkan jejak digital seluler untuk memperluas inklusi keuangan bagi UMKM.
- Meramalkan kondisi makroekonomi saat krisis (*nowcasting*).
- Mengkritisi bias algoritma yang dapat mendiskriminasi kelompok rentan.

Saya menyadari bahwa buku ini masih jauh dari kata sempurna. Teknologi *Big Data* dan Kecerdasan Buatan (AI) berkembang dengan kecepatan eksponensial. Oleh karena itu, kritik, saran, dan masukan dari para pembaca, akademisi, maupun praktisi kebijakan publik akan sangat saya hargai demi penyempurnaan edisi-edisi berikutnya.

Akhir kata, data berskala masif hanyalah sekumpulan angka nol dan satu yang bisu, hingga seorang ekonom memberinya makna. Saya berharap buku ini tidak hanya menjadi literatur wajib di bangku kuliah, tetapi juga menjadi katalis yang menginspirasi Anda untuk menggunakan data sebagai instrumen utama dalam merancang kebijakan publik yang lebih presisi, berkeadilan, dan berdampak nyata bagi pembangunan bangsa.

Selamat membaca, selamat mengeksplorasi, dan mari kita bangun masa depan ekonomi berbasis data!

Medan, Mei 2026

Penulis

DAFTAR ISI

KATA PENGANTAR	iii
DAFTAR ISTILAH.....	v
DAFTAR ISI	x
BAB I PARADIGMA BARU ILMU EKONOMI: MENYAMBUT ERA <i>BIG DATA</i>..	1
A. Mengapa Ilmu Ekonomi Membutuhkan Paradigma Baru?	1
B. Teori Dasar: Membedah Anatomi <i>Big Data</i> (Kerangka 5V)	4
C. Studi Kasus: <i>Big Data</i> dalam Ekonomi Pembangunan.....	6
D. Memahami Alur Kerja (<i>Pipeline</i>) Analitik Ekonomi	9
E. Rangkuman Bab I	11
F. Bahan Diskusi.....	11
G. Latihan Konseptual	12
BAB II LANSKAP DATA BARU: DARI ANGKA MENUJU TEKS, PIKSEL, DAN JEJAK DIGITAL	13
A. Keluar dari Zona Nyaman "Baris dan Kolom"	13
B. Teori Dasar: Tipologi Data di Era <i>Big Data</i>	14
C. Studi Kasus: Mengkuantifikasi yang Kualitatif.....	17
D. Jembatan dari Teks/Piksel Menuju Tabel.....	20
E. Rangkuman Bab II	21
F. Bahan Diskusi.....	22
G. Latihan Eksplorasi	22
BAB III <i>BIG DATA</i> VS. EKONOMETRIKA TRADISIONAL: PREDIKSI VS. KAUSALITAS	23
A. Jebakan "N=All" dan Ilusi Korelasi	23
B. Ekonometrika vs. <i>Data Science</i>	24
C. Teori Dasar: Y vs. β	26
D. Dari <i>Coding</i> ke <i>Learning</i>	28
E. Studi Kasus: Kapan Kita Butuh Prediksi, Kapan Kita Butuh Kausalitas?	32
F. Menggabungkan Dua Dunia (<i>Causal Machine Learning</i>)	36
G. Rangkuman Bab III	38
H. Bahan Diskusi.....	39
I. Latihan Konseptual	39

BAB IV MENGEKSTRAKSI JEJAK DIGITAL: WEB SCRAPING DAN

PENGUNAAN API	41
A. Internet sebagai "BPS" Terbesar di Dunia	41
B. Teori Dasar: Dua Jalur Menuju Gudang Data	42
C. Studi Kasus: Mengekstraksi Data untuk Kebijakan Pembangunan .	44
D. Logika di Balik Layar.....	46
E. Rangkuman Bab IV.....	48
F. Bahan Diskusi.....	48
G. Latihan Konseptual	49

BAB V TEKS SEBAGAI DATA: NATURAL LANGUAGE PROCESSING DALAM

EKONOMI MAKRO.....	51
A. Mengukur yang Tak Terlihat—Ekspektasi dan Sentimen	51
B. Teori Dasar: Apa Itu NLP dan Analisis Sentimen?	53
C. Studi Kasus: NLP dalam Pengambilan Keputusan Publik.....	56
D. Membedah Alur Pra-Pemrosesan Teks	59
E. Rangkuman Bab V.....	61
F. Bahan Diskusi.....	62
G. Latihan Konseptual (Simulasi Lexicon)	62

BAB VI EKONOMI SPASIAL: MEMANFAATKAN CITRA SATELIT DAN

SISTEM INFORMASI GEOGRAFIS	63
A. Memasukkan Kembali "Ruang" ke dalam Ilmu Ekonomi	63
B. Penginderaan Jauh (<i>Remote Sensing</i>)	65
C. Cahaya Malam Hari (<i>Night-Time Lights</i> /NTL).....	66
D. Teori Dasar: Anatomi Data Geospasial.....	67
E. Studi Kasus: Perekonomian dari Luar Angkasa	69
F. Jembatan Menuju Ekonometrika (Statistik Zonal)	72
G. Rangkuman Bab VI.....	74
H. Bahan Diskusi.....	74
I. Latihan Konseptual (Desain Pemetaan)	75

BAB VII PEMETAAN KEMISKINAN DAN KETIMPANGAN DENGAN DATA

NON-KONVENSIONAL.....	77
A. Paradoks Kebutuhan Data Kemiskinan.....	77
B. Teori Dasar: <i>Call Detail Records</i> (CDR) sebagai Proksi Kesejahteraan.....	79
C. Studi Kasus: Memetakan Kemiskinan di Afrika dan Indonesia.....	80

D. <i>Feature Engineering</i> pada Data Telekomunikasi	82
E. Rangkuman Bab VII	84
F. Bahan Diskusi	84
G. Latihan Konseptual (Logika <i>Feature Engineering</i>)	85
BAB VIII MOBILITAS, TRANSPORTASI, DAN PERENCANAAN EKONOMI	
PERKOTAAN	87
A. Kota sebagai Mesin Pertumbuhan dan Jebakan Kemacetan	87
B. Teori Dasar: <i>Opportunity Cost</i> dan Jejak Mobilitas Digital	89
C. Studi Kasus: Mengurai Benang Kusut Perkotaan	91
D. Visualisasi Mobilitas (Matriks OD dan Isokron)	94
E. Rangkuman Bab VIII	96
F. Bahan Diskusi	97
G. Latihan Konseptual (Analisis Aksesibilitas)	97
BAB IX INKLUSI KEUANGAN DAN <i>ALTERNATIVE CREDIT SCORING</i>	99
A. Paradoks "Orang Miskin Tidak Bisa Meminjam Uang"	99
B. Teori Dasar: Asimetri Informasi dan Evolusi <i>Credit Scoring</i>	101
C. Lanskap Baru: Data Alternatif sebagai Pengganti Kolateral	102
D. Studi Kasus: <i>Fintech</i> Memberdayakan UMKM	103
E. Membedah Algoritma Klasifikasi Probabilitas	105
F. Rangkuman Bab IX	106
G. Bahan Diskusi	107
H. Latihan Konseptual (Menjadi Manajer Risiko)	107
BAB X <i>NOWCASTING</i>: PERAMALAN MAKROEKONOMI SECARA	
<i>REAL-TIME</i>	109
A. Mengemudi di Tengah Kabut dengan Kaca Spion	109
B. Teori Dasar: <i>Forecasting</i> vs <i>Nowcasting</i>	111
C. Studi Kasus: Mendeteksi Krisis Saat Itu Juga	112
D. Menjembatani Perbedaan Waktu dengan Model Midas	114
E. Rangkuman Bab X	115
F. Bahan Diskusi	116
G. Latihan Konseptual Daring (Eksplorasi Google Trends)	116

BAB XI DINAMIKA PASAR TENAGA KERJA DI ERA DIGITAL.....	117
A. Mengapa Sarjana Menganggur dan Perusahaan Kekurangan Karyawan?	117
B. Teori Dasar: <i>Online Job Postings</i> dan Persamaan Mincer Era Digital	119
C. Studi Kasus: Memetakan Keterampilan dan <i>Gig Economy</i>	120
D. Rangkuman Bab XI.....	124
E. Bahan Diskusi.....	125
F. Latihan Konseptual (Ekstraksi Manual)	125
BAB XII VISUALISASI DATA UNTUK KOMUNIKASI KEBIJAKAN PUBLIK	127
A. Ujian "Lalu Mengapa?"	127
B. Teori Dasar: Dari Eksplorasi Menuju Eksplanasi.....	128
C. Studi Kasus: Saat Grafik Mengubah Kebijakan Negara	130
D. Tiga Langkah Data <i>Storytelling</i>	132
E. Rangkuman Bab XII.....	133
F. Bahan Diskusi.....	134
G. Latihan Konseptual (Mendesain Dasbor)	134
BAB XIII SISI GELAP <i>BIG DATA</i>: ETIKA, PRIVASI, DAN BIAS ALGORITMA .	135
A. Mitos "Mesin yang Objektif"	135
B. <i>The Digital Divide</i>	137
C. Teori Dasar: Anatomi Bias Algoritma dan Paradoks Privasi	137
D. Studi Kasus: Saat Algoritma Menghukum Orang Miskin	139
E. Melawan Bias dan Melindungi Data.....	143
F. Rangkuman Bab XIII.....	145
G. Bahan Diskusi.....	146
H. Latihan Konseptual (Dilema Etika)	146
BAB XIV MENUJU SATU DATA: TATA KELOLA, BIROKRASI, DAN MASA DEPAN EKOSISTEM DATA	147
A. Masalah "Kepulauan Data"	147
B. Interoperabilitas: Menghancurkan Tembok Pemisah	148
C. Satu Data Indonesia (SDI)	149
D. Studi Kasus: Regsosek (Registrasi Sosial Ekonomi)	150
E. Masa Depan: IoT dan AI dalam Kebijakan (<i>Policy 4.0</i>)	151
F. Diskusi Kelas	152

BAB XV DARI WACANA KE EKSEKUSI: PANDUAN PRAKTIS PROYEK

***BIG DATA* EKONOMI153**

- A. Jangan Mulai dari Data, Mulailah dari Masalah 153
- B. Metodologi CRISP-DM untuk Ekonom..... 154
- C. Fase Kritis: Data *Wrangling* 155
- D. Memilih Senjata: *Low-Code* vs. *Pro-Code* 156
- E. Studi Kasus Lokal: Mengawal Harga Cabai 156
- F. Diskusi Kelas 158

BAB XVI MASA DEPAN EKONOMI: *ARTIFICIAL INTELLIGENCE* DAN

TANTANGAN PEMBANGUNAN BARU159

- A. Dari Analisis ke Otomatisasi 159
- B. *Generative AI* untuk Ekonom..... 160
- C. *Digital Twin*: Simulasi Tanpa Risiko 161
- D. Tantangan Baru: *Algorithmic Bureaucracy*..... 162
- E. Studi Kasus: Menghalau "Joki" Bantuan dengan AI & *Machine Learning* 162
- F. Menjadi Ekonom *Hybrid* 164
- G. Diskusi Kelas 164

BAB XVII KAPITA SELEKTA DAN MASA DEPAN EKONOMI

PEMBANGUNAN BERBASIS DATA165

- A. Akhir dari Sebuah Awal..... 165
- B. Teori Dasar: AI Generatif dan Kolaborasi Data (B2G)..... 166
- C. Studi Kasus: *Big Data* sebagai Kompas SDGS 167
- D. Panduan Menyusun Proposal Proyek Akhir (*Capstone Project*).... 169
- E. Masa Depan Profesi: *The Economist-Data Scientist*..... 171
- F. Rangkuman Bab XVII..... 171
- G. Bahan Diskusi..... 172
- H. Ujian Komprehensif/Tugas Akhir..... 172

DAFTAR PUSTAKA173

INDEKS175

PROFIL PENULIS177

BAB I

PARADIGMA BARU ILMU EKONOMI: MENYAMBUT ERA *BIG DATA*

Tujuan Pembelajaran

Setelah menyelesaikan bab ini, mahasiswa diharapkan mampu:

1. Menjelaskan konsep dasar *Big Data* dan membedakannya dari data ekonomi dan demografi tradisional.
2. Memahami karakteristik utama *Big Data* melalui kerangka 5V (*Volume, Velocity, Variety, Veracity, Value*) dalam konteks indikator ekonomi.
3. Memahami pergeseran dari kelangkaan data (*data scarcity*) menuju kelimpahan data (*data abundance*).
4. Menganalisis bagaimana *Big Data* mengubah paradigma perumusan kebijakan publik dan penyelesaian masalah ekonomi pembangunan.
5. Mengidentifikasi kelebihan dan keterbatasan *Big Data* sebagai pelengkap instrumen ekonometrika klasik.

A. MENGAPA ILMU EKONOMI MEMBUTUHKAN PARADIGMA BARU?

Ilmu ekonomi, pada hakikatnya, selalu menjadi ilmu yang digerakkan oleh data (*data-driven science*). Sejak era Simon Kuznets merumuskan konsep Produk Domestik Bruto (PDB) pada tahun 1930-an, para ekonom dan pembuat kebijakan sangat bergantung pada data kuantitatif untuk mengukur kesejahteraan, merancang intervensi pengentasan kemiskinan, dan menjaga stabilitas makroekonomi.

BAB II

LANSKAP DATA BARU: DARI ANGKA MENUJU TEKS, PIKSEL, DAN JEJAK DIGITAL

Tujuan Pembelajaran

Setelah menyelesaikan bab ini, mahasiswa diharapkan mampu:

1. Mengklasifikasikan perbedaan antara data terstruktur, semi-terstruktur, dan tidak terstruktur dalam konteks analisis ekonomi.
2. Memahami konsep dasar transformasi data tidak terstruktur (seperti teks dan citra satelit) menjadi variabel kuantitatif yang siap diolah.
3. Menganalisis contoh penerapan penggunaan teks dan piksel sebagai proksi (*proxy*) untuk indikator pembangunan dan makroekonomi.
4. Mengenal secara konseptual format penyimpanan data modern (seperti JSON) dan bagaimana perangkat lunak seperti R atau *Python* memprosesnya.

A. KELUAR DARI ZONA NYAMAN "BARIS DAN KOLOM"

Sejak semester pertama belajar pengantar ekonomi dan statistika, Anda telah dilatih untuk melihat data dalam satu bentuk baku: tabel. Anda mengunduh data Produk Domestik Bruto (PDB), tingkat inflasi, atau rasio Gini dari situs web Badan Pusat Statistik (BPS) atau Bank Dunia dalam format.xls atau.csv. Data tersebut sangat rapi; setiap baris mewakili satu observasi (misalnya, tahun atau provinsi), dan setiap kolom mewakili satu variabel.

BAB III

BIG DATA VS. EKONOMETRIKA TRADISIONAL: PREDIKSI VS. KAUSALITAS

Tujuan Pembelajaran

Setelah menyelesaikan bab ini, mahasiswa diharapkan mampu:

1. Membedakan secara konseptual antara masalah prediksi (*prediction problem*) dan masalah inferensi kausal (*causal inference problem*) dalam perumusan kebijakan ekonomi.
2. Memahami mengapa volume data yang sangat besar (*Big Data*) tidak secara otomatis menyelesaikan masalah bias variabel yang diabaikan (*omitted variable bias*) atau endogenitas.
3. Menganalisis situasi di mana algoritma *Machine Learning* sangat berguna dan di mana metode ekonometrika tradisional lebih mutlak diperlukan.
4. Mengenal konsep integrasi antara *Machine Learning* dan Ekonometrika untuk evaluasi kebijakan publik.

A. JEBAKAN "N=ALL" DAN ILUSI KORELASI

Dalam dua bab sebelumnya, kita telah melihat bagaimana era digital menyediakan data dengan volume, kecepatan, dan variasi yang belum pernah terjadi sebelumnya. Ada sebuah mitos berbahaya yang sempat beredar di awal kemunculan *Big Data*: gagasan bahwa dengan data yang mencakup seluruh populasi ($N \rightarrow All$), kita tidak lagi membutuhkan teori ekonomi atau pengujian hipotesis. Mitos ini berasumsi bahwa jika kita memiliki data yang cukup besar, korelasi antar variabel sudah cukup untuk menjelaskan segalanya.

BAB IV

MENGEKSTRAKSI JEJAK DIGITAL: WEB SCRAPING DAN PENGGUNAAN API

Tujuan Pembelajaran

Setelah menyelesaikan bab ini, mahasiswa diharapkan mampu:

1. Menjelaskan konsep *Web Scraping* dan *Application Programming Interface* (API) sebagai metode modern dalam pengumpulan data ekonomi.
2. Membedakan kapan seorang ekonom harus menggunakan API resmi dan kapan harus menggunakan teknik *Web Scraping*.
3. Menganalisis penerapan ekstraksi data web dalam menyelesaikan masalah ekonomi pembangunan, seperti pemantauan pasar tenaga kerja dan respons bencana.
4. Memahami prinsip-prinsip etis dan legal dalam mengekstraksi data dari internet, termasuk pemahaman konseptual tentang robots.txt.

A. INTERNET SEBAGAI "BPS" TERBESAR DI DUNIA

Selama puluhan tahun, jika seorang ekonom membutuhkan data mengenai harga sewa rumah di Jakarta atau tren kualifikasi tenaga kerja yang dicari oleh perusahaan manufaktur, mereka hanya memiliki dua pilihan: menunggu laporan tahunan dari Badan Pusat Statistik (BPS) kementerian terkait, atau menyewa *surveyor* untuk turun ke lapangan dan membagikan kuesioner. Kedua opsi tersebut memakan waktu berbulan-bulan dan biaya yang tidak sedikit.

BAB V

TEKS SEBAGAI DATA: *NATURAL LANGUAGE PROCESSING* DALAM EKONOMI MAKRO

Tujuan Pembelajaran

Setelah menyelesaikan bab ini, mahasiswa diharapkan mampu:

1. Memahami konsep dasar *Natural Language Processing* (NLP) dan relevansinya dalam mengekstraksi variabel makroekonomi dari data teks.
2. Menjelaskan mekanisme *Sentiment Analysis* (Analisis Sentimen) dan penerapannya dalam mengukur ekspektasi inflasi dan kepercayaan konsumen.
3. Menganalisis studi kasus penggunaan NLP dalam evaluasi kebijakan publik dan respons krisis di negara berkembang.
4. Menjabarkan tahapan konseptual pra-pemrosesan teks (*Text Preprocessing*) mulai dari teks mentah hingga menjadi matriks data yang siap dianalisis secara statistik.

A. MENGUKUR YANG TAK TERLIHAT—EKSPEKTASI DAN SENTIMEN

Selama ini, kita belajar mengukur ekonomi lewat angka: inflasi 3%, pertumbuhan 5%, pengangguran 7%. Namun, angka-angka ini sering kali merupakan hasil akhir dari sebuah proses panjang yang dimulai dari opini, kepercayaan, dan narasi.

BAB VI

EKONOMI SPASIAL: MEMANFAATKAN CITRA SATELIT DAN SISTEM INFORMASI GEOGRAFIS

Tujuan Pembelajaran

Setelah menyelesaikan bab ini, mahasiswa diharapkan mampu:

1. Menjelaskan pentingnya dimensi "ruang" (spasial) dalam analisis ekonomi pembangunan.
2. Membedakan dua format utama data geospasial: Data Vektor (titik, garis, poligon) dan Data Raster (piksel/citra satelit).
3. Menganalisis bagaimana citra satelit cahaya malam (*night-time lights*) dan citra siang hari (*daytime imagery*) digunakan sebagai proksi indikator kesejahteraan dan PDB.
4. Memahami alur konseptual ekstraksi data spasial menggunakan teknik Statistik Zonal (*Zonal Statistics*) untuk mengubah peta menjadi tabel data yang siap diregresi.

A. MEMASUKKAN KEMBALI "RUANG" KE DALAM ILMU EKONOMI

Sebagian besar model ekonomi makro dan mikro yang Anda pelajari hingga saat ini memiliki satu kelemahan asumsi yang mendasar: mereka mengabaikan geografi. Ketika kita berbicara tentang Produk Domestik Bruto (PDB), inflasi, atau tingkat pengangguran nasional, kita sering kali memperlakukan sebuah negara seolah-olah ia adalah sebuah titik tanpa dimensi ruang (*spaceless world*).

BAB VII

PEMETAAN KEMISKINAN DAN KETIMPANGAN DENGAN DATA *NON-KONVENSIONAL*

Tujuan Pembelajaran

Setelah menyelesaikan bab ini, mahasiswa diharapkan mampu:

1. Menjelaskan keterbatasan survei demografi konvensional dalam menghasilkan peta kemiskinan beresolusi tinggi.
2. Memahami konsep *Call Detail Records* (CDR) dan bagaimana metadata telekomunikasi dapat digunakan sebagai proksi perilaku ekonomi.
3. Menganalisis studi kasus penggabungan data satelit dan data seluler untuk mengestimasi tingkat kemiskinan dan ketimpangan di negara berkembang.
4. Mengidentifikasi teknik *Feature Engineering* (Rekayasa Fitur) dasar yang mengubah log panggilan telepon menjadi variabel indikator kesejahteraan rumah tangga.

A. PARADOKS KEBUTUHAN DATA KEMISKINAN

Tujuan Pembangunan Berkelanjutan (SDGs) nomor satu adalah: *Mengakhiri kemiskinan dalam segala bentuknya di mana pun*. Namun, untuk mengakhiri kemiskinan, pemerintah harus terlebih dahulu menjawab satu pertanyaan fundamental yang sangat sulit: **Di mana sebenarnya orang miskin itu tinggal?** Lihat Gambar 7.1.

BAB VIII

MOBILITAS, TRANSPORTASI, DAN PERENCANAAN EKONOMI PERKOTAAN

Tujuan Pembelajaran

Setelah menyelesaikan bab ini, mahasiswa diharapkan mampu:

1. Menjelaskan konsep ekonomi aglomerasi dan eksternalitas negatif (kemacetan) dalam konteks pembangunan perkotaan.
2. Mengidentifikasi sumber-sumber *Big Data* mobilitas modern, seperti jejak GPS, log aplikasi *ride-sharing*, dan data transaksi kartu pintar (*smart cards*).
3. Menganalisis studi kasus bagaimana data mobilitas digunakan untuk mengevaluasi efisiensi transportasi publik dan fenomena *spatial mismatch* (ketidaksesuaian spasial) di pasar tenaga kerja.
4. Memahami konsep analitis praktis seperti Matriks Asal-Tujuan (*Origin-Destination Matrix*) dan Peta Isokron (*Isochrone Maps*) dalam pengambilan keputusan tata kota.

A. KOTA SEBAGAI MESIN PERTUMBUHAN DAN JEBAKAN KEMACETAN

Dalam ilmu ekonomi pembangunan, terdapat konsensus yang kuat bahwa urbanisasi dan pertumbuhan ekonomi berjalan beriringan. Fenomena ini didorong oleh konsep **Aglomerasi Ekonomi** (*Agglomeration Economies*). Ketika perusahaan, pekerja, dan konsumen berkumpul di satu area geografis yang berdekatan (kota), biaya produksi turun, inovasi meningkat karena pertukaran ide, dan pencocokan (*matching*) antara pencari kerja dan lowongan kerja menjadi jauh lebih efisien. Inilah alasan

BAB IX

INKLUSI KEUANGAN DAN *ALTERNATIVE CREDIT SCORING*

Tujuan Pembelajaran

Setelah menyelesaikan bab ini, mahasiswa diharapkan mampu:

1. Menjelaskan konsep asimetri informasi di pasar kredit dan mengapa bank tradisional gagal melayani masyarakat kelas bawah (*unbanked*).
2. Membedakan antara penilaian kredit tradisional (seperti SLIK OJK/BI *Checking*) dengan *Alternative Credit Scoring* (ACS) yang berbasis jejak digital.
3. Menganalisis bagaimana perusahaan *Financial Technology* (*Fintech*) menggunakan *Big Data* untuk memberdayakan Usaha Mikro, Kecil, dan Menengah (UMKM).
4. Memahami alur kerja algoritma klasifikasi secara konseptual dalam memprediksi probabilitas gagal bayar (*Probability of Default*).

A. PARADOKS "ORANG MISKIN TIDAK BISA MEMINJAM UANG"

Salah satu ironi terbesar dalam ekonomi pembangunan adalah kenyataan bahwa mereka yang paling membutuhkan modal justru adalah mereka yang paling sulit mendapatkannya. Seorang pedagang sayur di pasar tradisional mungkin hanya membutuhkan modal Rp 1.000.000 untuk memperluas dagangannya dan meningkatkan taraf hidup keluarganya. Namun, jika ia datang ke bank konvensional, kemungkinan besar pengajuannya akan ditolak. Mengapa? Lihat Gambar 9.1.

BAB X

NOWCASTING: PERAMALAN MAKROEKONOMI SECARA REAL-TIME

Tujuan Pembelajaran

Setelah menyelesaikan bab ini, mahasiswa diharapkan mampu:

1. Memahami konsep *Nowcasting* dan membedakannya dari *Forecasting* (peramalan masa depan) tradisional.
2. Menjelaskan masalah *publication lag* (jeda publikasi) pada data resmi pemerintah dan dampaknya terhadap respons kebijakan saat krisis.
3. Menganalisis penggunaan data frekuensi tinggi (*high-frequency data*) seperti Google Trends dan transaksi kartu kredit sebagai *leading indicators* (indikator pendahulu).
4. Mengenal secara konseptual model ekonometrika yang mampu menggabungkan data dengan frekuensi yang berbeda, seperti MIDAS (*Mixed-Data Sampling*).

A. MENGEMUDI DI TENGAH KABUT DENGAN KACA SPION

Bayangkan Anda sedang mengemudikan sebuah mobil besar di jalan tol yang dipenuhi kabut tebal. Tiba-tiba, ada jurang di depan Anda. Namun, kaca depan mobil Anda tertutup rapat. Satu-satunya cara Anda untuk menyetir adalah dengan melihat ke kaca spion, yang hanya menunjukkan jalan yang sudah Anda lewati tiga bulan yang lalu.

Dalam ekonomi makro, "mobil besar" itu adalah negara, "jurang" itu adalah krisis ekonomi (seperti pandemi atau resesi global), dan Anda adalah pembuat kebijakan (Menteri Keuangan atau Gubernur Bank Sentral). Kaca spion melambangkan data statistik resmi. Lihat Gambar 10.1.

BAB XI

DINAMIKA PASAR TENAGA KERJA DI ERA DIGITAL

Tujuan Pembelajaran

Setelah menyelesaikan bab ini, mahasiswa diharapkan mampu:

1. Menjelaskan keterbatasan data ketenagakerjaan tradisional dalam menangkap dinamika perubahan keterampilan yang sangat cepat di era digital.
2. Memahami konsep *Skills Mismatch* (Ketidaksesuaian Keterampilan) dan bagaimana *Big Data* dari portal lowongan kerja daring (OJI) digunakan untuk mengukurnya.
3. Menganalisis peran *Gig Economy* (ekonomi pekerja lepas *platform*) sebagai penyerap guncangan ekonomi (*shock absorber*) di negara berkembang menggunakan data *platform*.
4. Memahami alur kerja praktis (*pipeline*) dalam mengekstraksi informasi tak terstruktur dari teks iklan lowongan kerja menjadi basis data intelijen pasar tenaga kerja.

A. MENGAPA SARJANA MENGANGGUR DAN PERUSAHAAN KEKURANGAN KARYAWAN?

Salah satu keluhan paling umum yang didengar oleh pembuat kebijakan di negara berkembang adalah sebuah paradoks: Di satu sisi, tingkat pengangguran terbuka (terutama di kalangan lulusan universitas dan sekolah vokasi) sangat tinggi. Namun di sisi lain, perusahaan-perusahaan manufaktur dan teknologi terus mengeluh bahwa mereka kesulitan mencari tenaga kerja yang memenuhi kualifikasi.

BAB XII

VISUALISASI DATA UNTUK KOMUNIKASI KEBIJAKAN PUBLIK

Tujuan Pembelajaran

Setelah menyelesaikan bab ini, mahasiswa diharapkan mampu:

1. Menjelaskan perbedaan mendasar antara visualisasi eksploratif (untuk peneliti) dan visualisasi eksplanatoris (untuk pembuat kebijakan).
2. Memahami prinsip-prinsip desain visual, seperti rasio data-tinta (*data-ink ratio*) dan beban kognitif (*cognitive load*), dalam menyajikan data ekonomi.
3. Memilih jenis grafik yang tepat berdasarkan pesan analitis yang ingin disampaikan (perbandingan, komposisi, distribusi, atau hubungan spasial).
4. Menganalisis studi kasus bagaimana dasbor interaktif (*dashboards*) dan *data storytelling* mengubah arah kebijakan publik di negara berkembang.

A. UJIAN "LALU MENGAPA?"

Bayangkan Anda baru saja menghabiskan waktu tiga bulan menganalisis jutaan baris data *Call Detail Records* (CDR) dan citra satelit seperti yang kita pelajari di Bab VI dan VII. Anda berhasil membangun model *Machine Learning* yang sangat canggih untuk memprediksi kantong-kantong kemiskinan baru dengan tingkat akurasi 95%.

BAB XIII

SISI GELAP *BIG DATA*: ETIKA, PRIVASI, DAN BIAS ALGORITMA

Tujuan Pembelajaran

Setelah menyelesaikan bab ini, mahasiswa diharapkan mampu:

1. Menjelaskan konsep "Bias Algoritma" (*Algorithmic Bias*) dan bagaimana data masa lalu dapat melanggengkan diskriminasi struktural di masa depan.
2. Menganalisis *trade-off* (tarik-ulur) antara utilitas data (kegunaan untuk kebijakan) dan privasi individu masyarakat.
3. Mengevaluasi studi kasus kegagalan algoritma dalam kebijakan publik, seperti diskriminasi dalam penyaluran bantuan sosial atau penegakan hukum preventif.
4. Memahami konsep praktis audit keadilan algoritma (*Algorithmic Fairness*) dan teknik anonimisasi data dasar untuk melindungi subjek penelitian.

A. MITOS "MESIN YANG OBJEKTIF"

Sepanjang dua belas bab terakhir, kita telah merayakan kekuatan *Big Data* sebagai "minyak baru". Bab ini adalah peringatan bahwa minyak tersebut mudah terbakar. Kita telah melihat bagaimana satelit, log ponsel, dan teks dapat memetakan kemiskinan, menyalurkan kredit bagi yang tidak punya rekening bank, dan mengurai kemacetan. Narasi yang sering dibangun oleh perusahaan teknologi adalah: "*Manusia itu bias, korup, dan penuh prasangka. Komputer hanya melihat angka. Oleh karena itu,*

BAB XIV

MENUJU SATU DATA: TATA KELOLA, BIROKRASI, DAN MASA DEPAN EKOSISTEM DATA

Tujuan Pembelajaran

Setelah mempelajari bab ini, mahasiswa diharapkan mampu:

1. Mengidentifikasi masalah "Silo Data" (*Data Silos*) dalam birokrasi pemerintahan dan dampaknya terhadap pembangunan.
2. Menjelaskan prinsip dasar **Satu Data Indonesia (SDI)** sesuai Perpres No. 39 Tahun 2019.
3. Memahami konsep Interoperabilitas Data: bagaimana membuat sistem yang berbeda bisa saling "berbicara".
4. Menganalisis tren masa depan: *Internet of Things* (IoT) dan peran *Artificial Intelligence* (AI) dalam perumusan kebijakan otomatis.

A. MASALAH "KEPULAUAN DATA"

Indonesia adalah negara kepulauan, dan sayangnya, data pemerintahannya pun berbentuk kepulauan. Lihat Gambar 14.1.

Sering kali, Kementerian Pertanian memiliki data luas sawah yang berbeda dengan Kementerian ATR/BPN, dan berbeda lagi dengan data BPS. Kementerian Sosial punya data orang miskin sendiri, Pemerintah Daerah punya data sendiri. Kondisi ini disebut **Data Silos**.

Akibatnya, pengambilan kebijakan menjadi kacau. Subsidi pupuk mungkin salah sasaran karena data luas sawah yang dipakai tidak akurat. Di sinilah peran **Tata Kelola Data (Data Governance)** menjadi krusial. Ekonom

BAB XV

DARI WACANA KE EKSEKUSI: PANDUAN PRAKTIS PROYEK *BIG DATA* EKONOMI

Tujuan Pembelajaran

Setelah mempelajari bab ini, mahasiswa diharapkan mampu:

1. Memahami siklus hidup proyek data sains (*Data Science Lifecycle*) menggunakan kerangka kerja CRISP-DM.
2. Merumuskan "Pertanyaan Ekonomi" yang tajam sebelum mencari data.
3. Memahami *realitas* proses *Data Wrangling* (pembersihan data) yang memakan waktu 80% dari total proyek.
4. Merancang alur kerja teknis sederhana untuk menyelesaikan masalah ekonomi lokal.

A. JANGAN MULAI DARI DATA, MULAILAH DARI MASALAH

Kesalahan terbesar pemula dalam *Big Data Analytics* adalah **terobsesi dengan *tools* atau data**.

- *Salah*: "Saya punya data Twitter 1 juta baris, enaknya diapain ya?"
- *Benar*: "Saya ingin tahu apakah kenaikan harga beras membuat keresahan sosial. Data apa yang saya butuhkan untuk menjawab itu?"

Sebagai ekonom, nilai tambah Anda bukan pada kemampuan *coding*, tapi pada kemampuan mendefinisikan masalah (*Problem Statement*). Bab ini akan memandu Anda melewati fase-fase kritis dalam sebuah proyek data.

BAB XVI

MASA DEPAN EKONOMI: *ARTIFICIAL INTELLIGENCE* DAN TANTANGAN PEMBANGUNAN BARU

Tujuan Pembelajaran

Setelah mempelajari bab ini, mahasiswa diharapkan mampu:

1. Menjelaskan evolusi dari *Big Data Analytics* menuju *Artificial Intelligence* (AI) dan *Generative AI* dalam konteks ekonomi.
2. Memahami konsep *Digital Twin* (Kembaran Digital) untuk simulasi kebijakan publik tanpa risiko.
3. Menganalisis dampak otomatisasi terhadap pasar tenaga kerja di negara berkembang (studi kasus ketimpangan baru).
4. Merumuskan posisi strategis "Ekonom Masa Depan" yang berkolaborasi dengan AI, bukan digantikan olehnya.

A. DARI ANALISIS KE OTOMATISASI

Jika Bab I hingga XV mengajarkan kita cara "melihat" data (deskriptif) dan "memprediksi" data (prediktif), Bab ini membahas fase selanjutnya: **Preskriptif dan Otomatisasi**.

Di masa depan, ekonom mungkin tidak perlu lagi membersihkan data Excel secara manual. Algoritma AI akan melakukannya secara otomatis, bahkan menuliskan draf awal analisisnya untuk Anda. Tantangan bagi ekonom pembangunan bergeser: bukan lagi soal "bagaimana mengolah data", melainkan "apakah rekomendasi mesin ini adil untuk rakyat?"

BAB XVII

KAPITA SELEKTA DAN MASA DEPAN EKONOMI PEMBANGUNAN BERBASIS DATA

Tujuan Pembelajaran

Setelah menyelesaikan bab ini, mahasiswa diharapkan mampu:

1. Menjelaskan evolusi dari analitik prediktif (*Predictive AI*) menuju Kecerdasan Buatan Generatif (*Generative AI*) dalam konteks perumusan kebijakan publik.
2. Menganalisis peran krusial *Big Data* dalam melacak dan mempercepat pencapaian Tujuan Pembangunan Berkelanjutan (*Sustainable Development Goals/SDGs*).
3. Memahami konsep *Data Collaboratives* (Kolaborasi Data) antara sektor swasta dan pemerintah sebagai model tata kelola masa depan.
4. Menyusun kerangka proposal proyek akhir (*Capstone Project*) yang mengintegrasikan teori ekonomi pembangunan, sumber data *non-konvensional*, dan metodologi analitik yang beretika.

A. AKHIR DARI SEBUAH AWAL

Selamat. Anda telah tiba di bab terakhir dari buku ini. Jika Anda menengok kembali ke Bab I, Anda memulai perjalanan ini dengan melihat data melalui kacamata tradisional: tabel BPS, kuesioner, dan ekonometrika klasik. Seiring berjalannya semester, Anda telah belajar mengekstraksi harga dari *e-commerce* melalui API, menghitung *sentimen* inflasi dari lautan teks, memetakan kemiskinan menggunakan cahaya satelit, dan mengurai kemacetan perkotaan melalui jejak GPS telepon seluler.

DAFTAR PUSTAKA

- Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring economic *policy* uncertainty. *The Quarterly Journal of Economics*, 131(4), 1593–1636. <https://doi.org/10.1093/qje/qjw024>
- Björkegren, D., & Grissen, D. (2018). Behavior revealed in mobile phone usage predicts *credit* repayment. *The World Bank Economic Review*, 32(3), 618–634. <https://doi.org/10.1093/wber/lhx016>
- Blumenstock, J., Cadamuro, G., & On, R. (2015). Predicting poverty and wealth from mobile phone metadata. *Science*, 350(6264), 1073–1076. <https://doi.org/10.1126/science.aac2952>
- Cavallo, A., & Rigobon, R. (2016). The Billion Prices Project: Using *online* prices for measurement and research. *Journal of Economic Perspectives*, 30(2), 151–178. <https://doi.org/10.1257/jep.30.2.151>
- Chetty, R., Friedman, J. N., Hendren, N., Stepner, M., & The Opportunity Insights Team. (2020). *How did COVID-19 and stabilization policies affect spending and employment? A new real-time economic tracker based on private sector data* (NBER Working Paper No. 27431). National Bureau of Economic Research. <https://doi.org/10.3386/w27431>
- Choi, H., & Varian, H. (2012). Predicting the present with Google Trends. *Economic Record*, 88(s1), 2–9. <https://doi.org/10.1111/j.1475-4932.2012.00809.x>
- Ghysels, E., Sinko, A., & Valkanov, R. (2007). MIDAS regressions: Further results and new directions. *Econometric Reviews*, 26(1), 53–90. <https://doi.org/10.1080/07474930600972467>
- Henderson, J. V., Storeygard, A., & Weil, D. N. (2012). Measuring economic growth from outer space. *The American Economic Review*, 102(2), 994–1028. <https://doi.org/10.1257/aer.102.2.994>
- Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., & Ermon, S. (2016). Combining satellite imagery and *machine learning* to predict poverty. *Science*, 353(6301), 790–794. <https://doi.org/10.1126/science.aaf7894>

- Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, 66(1), 35–65. <https://doi.org/10.1111/j.1540-6261.2010.01625.x>
- Mitchell, R. (2018). *Web scraping with Python: Collecting more data from the modern web* (2nd ed.). O'Reilly Media.
- O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.
- Tufte, E. R. (2001). *The visual display of quantitative information* (2nd ed.). Graphics Press.
- UN Global Pulse. (2017). *Integrating big data into the monitoring and evaluation of development programmes*. United Nations.
- Varian, H. R. (2014). Big data: New tricks for econometrics. *Journal of Economic Perspectives*, 28(2), 3–28. <https://doi.org/10.1257/jep.28.2.3>
- World Bank. (2021). *Data for better lives: World Development Report 2021*. World Bank Publications.

PROFIL PENULIS

Dr. Baihaqi Siregar, S.Si., M.T.



Penulis menyelesaikan pendidikan Sarjana di Program Studi Matematika Institut Teknologi Bandung pada tahun 2004, pendidikan Magister di Program Studi Teknik Elektro Universitas Sumatera Utara pada tahun 2011, dan pendidikan Doktor di Program Studi Ilmu Komputer pada tahun 2024. Saat ini penulis berprofesi

sebagai dosen tetap dan Sekretaris Program Studi di Magister *Data Science & Artificial Intelligence* Universitas Sumatera Utara. Saat buku ini ditulis, penulis mendapatkan tugas untuk menjadi pengajar Mata Kuliah Analisis *Big Data* pada Program Studi Sarjana Ekonomi Pembangunan di Fakultas Ekonomi dan Bisnis Universitas Sumatera Utara. Penulis juga merupakan *Founder* dari *IdeaSTEAM/COGNISTEAM* – wadah pembelajaran *Computational Thinking* dan Robotika untuk anak usia sekolah – dan *Naturetech Innovation School*. Penulis dapat dihubungi melalui *platform* [linkedin.com/in/baihaqi-siregar/](https://www.linkedin.com/in/baihaqi-siregar/).

Di era digital, ilmu ekonomi tidak lagi hanya bergantung pada kuesioner kertas dan data kuartalan yang lambat. Denyut nadi perekonomian dan kesejahteraan masyarakat kini terekam secara *real-time* melalui jejak digital yang masif: dari miliaran piksel citra satelit di malam hari, log mobilitas telepon seluler, hingga jutaan keluhan di media sosial. Pertanyaannya: Sudah siapkah para calon ekonom kita membedah lautan data ini?

Buku Pengantar Analisis *Big Data* untuk Pembangunan Ekonomi hadir untuk menjembatani kesenjangan antara ketajaman teori ilmu ekonomi dan kecanggihan teknologi komputasi modern. Ditulis secara khusus untuk mahasiswa tahap sarjana (S1) di program studi Ekonomi Pembangunan, buku ini mendobrak mitos bahwa analitik data besar adalah wilayah eksklusif milik mahasiswa ilmu komputer.

Melalui pendekatan yang sangat bumi dan berfokus pada aplikasi kebijakan publik, pembaca akan diajak menjelajahi revolusi metodologi ekonomi melalui 17 bab yang komprehensif. Buku ini mengupas tuntas bagaimana menterjemahkan data tak terstruktur (teks, gambar, jaringan) menjadi instrumen penyelesaian masalah-masalah pembangunan yang paling mendesak. Pembaca akan mempelajari cara:

- Menggunakan citra satelit dan data spasial untuk memetakan kantong kemiskinan dengan resolusi tinggi.
- Melakukan *nowcasting* (peramalan seketika) indikator makroekonomi seperti inflasi dan PDB menggunakan mesin pencari internet di masa krisis.
- Mengekstraksi jutaan lowongan kerja daring (*web scraping*) untuk mengatasi pengangguran dan *skills mismatch*.
- Mendorong inklusi keuangan bagi UMKM melalui *Alternative Credit Scoring*.

Buku ini tidak mengajarkan Anda sekadar menjadi *programmer*, melainkan melatih Anda menjadi "Arsitek Kebijakan". Oleh karena itu, buku ini juga secara kritis mengupas perbedaan krusial antara prediksi (*Machine Learning*) dan kausalitas (Ekonometrika), serta memperingatkan pembaca tentang bahaya "Sisi Gelap" data berupa pelanggaran privasi dan diskriminasi algoritma.

Dengan perpaduan teori, studi kasus nyata di negara berkembang, dan demonstrasi konseptual, buku ini adalah panduan wajib bagi generasi baru ekonom, pembuat kebijakan, dan analis yang ingin menggunakan data sebagai instrumen untuk menciptakan pembangunan yang presisi, inklusif, dan berkeadilan.

 www.penerbitwidina.com

 @penerbitwidina

 penerbit widina

 penerbitwidina@gmail.com

 widina store

 widina bookstore

Layanan Pembaca & Penerbitan Buku
 0815-7000-699



SCANME

Teknologi & Ekonomi

ISBN 978-634-246-844-9



9 786342 468449