



IPB University
— Bogor Indonesia —

Departement of Statistics
Study Program in Statistics and Data Science



SELAMAT DATANG *di* *IPB UNIVERSITY*

Pelopor

Pendidikan Tinggi Statistika dan Sains Data
di Indonesia

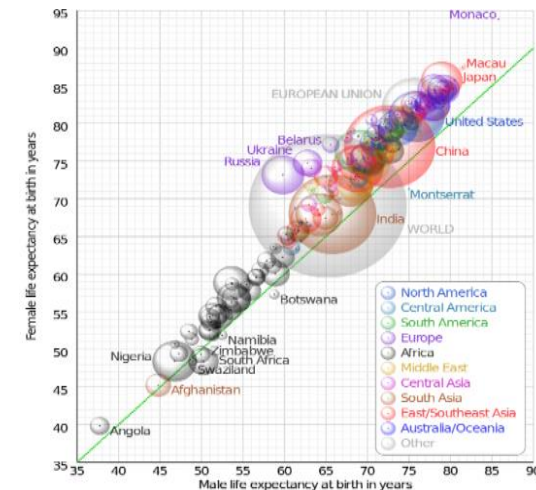
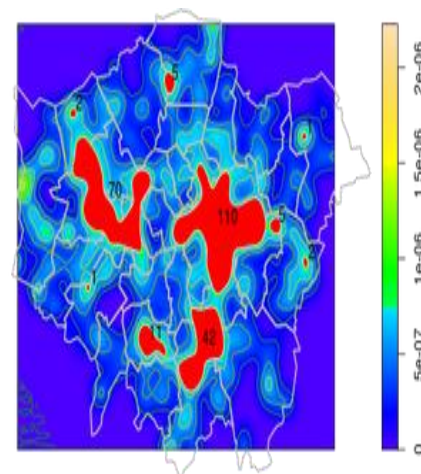


IPB University
— Bogor Indonesia —

Departement of Statistics
Study Program in Statistics and Data Science

STA1211 *Metode Statistika*

Prof. Dr. Ir. Khairil Anwar Notodiputro, MS
Guru Besar Statistika IPB





IPB University
— Bogor Indonesia —

Departement of Statistics

Study Program in Statistics and Data Science

Khairil Anwar Notodiputro

Lulus Sarjana dan Magister Sains dari Institut Pertanian Bogor. Mendapat Ph.D dalam bidang statistika dari *Macquarie University Australia*. **Sekarang sebagai Guru Besar pada Dept. Statistika IPB, Anggota Senat Akademik IPB, dan sebagai Ketua Program Studi Magister dan Doktor Statistika dan Sains Data.**

Di luar Indonesia sebagai *Visiting Professor at Prince of Songkla University and Kasetsart University in Thailand.*

Di IPB pernah menjabat Ketua Jurusan Statistika, Direktur Tingkat Persiapan Bersama, Ketua LSI, dan Dekan Sekolah Pascasarjana. Di luar IPB pernah menjabat Kepala Badan Penelitian dan Pengembangan Kemdikbud, dan Ketua ISI (Ikatan Statistisi Indonesia), Ketua Dewan Pengawas Universitas Hasanuddin.

Aktif melakukan penelitian *Small Area Estimation, Generalized Linear Mixed Models, Time Series Analysis dan Statistical Machine Learning*. Tugas utama mengajar dan membimbing mahasiswa sarjana dan pascasarjana, serta menulis dalam berbagai jurnal ilmiah, seminar nasional dan internasional tetap dilaksanakan. Selain itu pernah menjadi anggota berbagai asosiasi keilmuan internasional seperti ISI (*International Statistical Institute*), ASA (*American Statistical Association*), MSMSSEA (*Muslim Statistician-Mathematician Society in South East Asia*), dan *World Group of Geoinformatics and Digital Government*.



Jobs for Statisticians



The screenshot shows the homepage of 'The World of Statistics' website. At the top is a navigation bar with links: About Us, What is Statistics?, Statistics as a Career, Teacher Resources, Global Supporters, The World of Statistics Posters, and Home. The main heading is 'What Fields Employ Statisticians?'. Below it, a paragraph states: 'One advantage of working in statistics is that you can combine your interest with almost any other field in science, technology, or business, such as:'. This is followed by a list of fields: Agriculture, Animal Population, Astronomy, Biology, Census, Chemistry, Computer Science, Demography, Ecology, Economics, Education, Engineering, Epidemiology, Finance, Forestry, and Genetics. On the right side of the page, there is a search bar, a 'Join THE WORLD OF STATISTICS' button, a video player titled 'Why Statistics is Important to You', a 'Give the gift of significance' button, a 'GET THE WORLD OF STATISTICS NEWSLETTER' button, a 'Statisticians in the News' section with a link to 'Are Men Really Better At Tennis Than Women? Nope—According To Statistics, Women Are Just As Good As Men', a 'Statistical Word of the Week' section with the word 'Hypothesis Testing', a 'Provided by The Institute for Statistics Education' note, and a 'The World of Statistics Poster' section with a world map graphic.

What Fields Employ Statisticians?

One advantage of working in statistics is that you can combine your interest with almost any other field in science, technology, or business, such as:

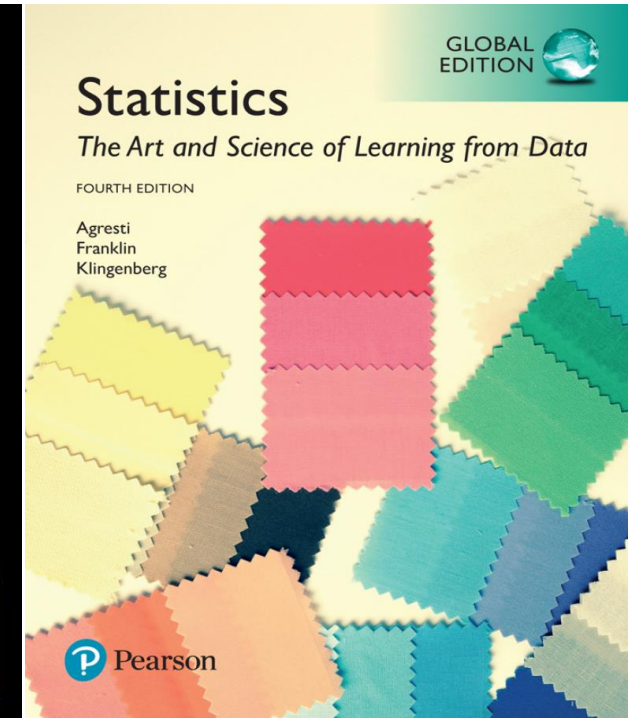
- Agriculture
- Animal Population
- Astronomy
- Biology
- Census
- Chemistry
- Computer Science
- Demography
- Ecology
- Economics
- Education
- Engineering
- Epidemiology
- Finance
- Forestry
- Genetics

What Job Title Might I Have?

Statistician is the most prevalent title in use, but other titles include:

- Business Analyst
- Professor
- Economist
- Software Engineer
- Mathematician
- Risk Analyst
- Quality Analyst
- Investigator
- Environmental Scientist
- Pharmaceutical Engineer
- Researcher
- Data Analyst
- Project Manager
- Manager

Pembelajaran dari Data *(Learning from Data)*



Belajar Statistika?

01

Ungkapan: *"without data, you are just another person with an opinion"*

02

Menggunakan hanya logika bisa salah kalau tidak dikonfirmasi dengan data: *ingat hukum Newton*

1

In 1687, [Sir Isaac Newton](#) published [Principia](#), which hypothesizes the [inverse-square law](#) of universal gravitation. In his own words, "I deduced that the forces which keep the planets in their orbs must be reciprocally as the squares of their distances from the centers about which they revolve....."

2

Newton's theory enjoyed its greatest success when it was used to predict the existence of [Neptune](#) based on motions of [Uranus](#) that could not be accounted by the actions of the other planets.

3

Years later, it was another discrepancy in a planet's orbit that showed Newton's theory to be inaccurate. By the end of the 19th century, it was known that the orbit of [Mercury](#) could not be accounted for entirely under Newtonian gravity, and all searches for another perturbing body (such as a planet orbiting the [Sun](#) even closer than Mercury) have been fruitless.

Belajar Statistika?

03

Semakin sulit kita hidup saat ini kalau tidak memahami data → *perlu belajar statistika*

- Efektivitas vaksin covid-19.
- Pemilihan kepala daerah.
- Perkembangan harga cabai, bawang merah, bawang putih, beras Cianjur.
- Perkembangan indeks saham.
- Komposisi nutrisi dari pangan dalam kemasan.

Semua di atas adalah menyangkut statistik.

04

Masyarakat melek informasi (*modern*) perlu:

- Mengambil sari-pati dari informasi yang ada di dalam tabel, diagram dan grafik.
- Memahami argumen berbasis data, BUKAN asal bunyi (*asbun*).
- Memahami bagaimana cara mengumpulkan data, menyajikan data, menganalisis data dan cara mengambil kesimpulan.

Rencana Materi Kuliah

No.	Pokok Bahasan	Sub Pokok Bahasan	Pertemuan ke
1	Lingkup statistika (<i>Scope of statistics</i>)	<ul style="list-style-type: none">Statistika: Apa dan Mengapa?Statistika Deskriptif dan Inferensia: Metode Deduksi vs InduksiIlustrasi statistika dalam penelitianContoh acak dan penarikan contoh acak sederhana.Kesalahan dalam pengumpulan data	1
2	Konsep dasar peluang (<i>Basic concepts of probability</i>)	<ul style="list-style-type: none">Ruang contoh dan kejadian, operasi-operasi pada kejadianPermutasi dan kombinasiPeluang (aksiomatik dan frekuensi relatif), peluang bersyaratMenghitung peluang suatu kejadianPenarikan contoh acak (pemulihan dan tanpa pemulihan)Kejadian bebas dan Kaidah penggandaanDalil Bayes	2
3	Populasi, contoh, peubah acak dan sebaran peluang peubah acak (<i>population, sample, and probability distributions of random variables</i>)	<ul style="list-style-type: none">Populasi (batasan populasi, populasi terhingga dan tak hingga)Contoh (pengertian contoh, contoh representatif, dan contoh acak)Peubah acak (peubah acak sebagai pemetaan, peubah acak diskrit dan kontinu)Sebaran peubah acak (konsep dan sifat-sifat nilai harapan dan ragam)Beberapa model sebaran peubah acak (Sebaran Binom, Sebaran Poisson, Sebaran Normal)Hampiran Normal terhadap Binom	3-4

Rencana Materi Kuliah (lanjutan)

No	Pokok Bahasan	Sub Pokok Bahasan	Pertemuan ke
4	Sebaran percontohan (<i>sampling distribution</i>)	<ul style="list-style-type: none"> ▪ Penarikan contoh dan Inferensi (pengertian inferensi statistik, statistik dan parameter, dan sebaran percontohan) ▪ Nilai harapan dan ragam ▪ Sebaran percontohan dari populasi Normal (Sebaran t-Student, Khi-Kuadrat) ▪ Rataan contoh terbakukan dan Dalil Limit Pusat 	5-6
5	Pendugaan parameter (<i>Parameter Estimation</i>)	<ul style="list-style-type: none"> ▪ Pengertian penduga takbias linear terbaik (BLUE) ▪ Penduga paramter nilaitengah (μ) dan proporsi (p) ▪ Selang kepercayaan bagi parameter μ, selang kepercayaan bagi parameter p ▪ Penentuan ukuran contoh 	7
6	UTS	Ujian Tengah Semester	
7	Konsep pengujian hipotesis, Pengujian Hipotesis untuk satu populasi (<i>Hypothesis testing for one population</i>)	<ul style="list-style-type: none"> ▪ Prinsip pengujian hipotesis (pengertian, galat jenis I dan galat jenis II). ▪ Pengujian hipotesis rataan populasi dan proporsi (uji z dan atau uji t) ▪ Asumsi pengujian hipotesis ▪ Pengujian hipotesis untuk ragam populasi (uji khi kuadrat) ▪ Hubungan ukuran contoh dan kesensitifan uji 	8-10

Rencana Materi Kuliah (lanjutan)

No.	Pokok Bahasan	Sub Pokok Bahasan	Pertemuan ke
8	Pengujian hipotesis untuk dua populasi (<i>Hypothesis testing for two populations</i>)	<ul style="list-style-type: none"> Selang kepercayaan dan uji hipotesis untuk beda dua nilai tengah populasi (contoh bebas dan berpasangan). Asumsi yang diperlukan 	11
9	Korelasi dan Regresi Linear Sederhana (<i>simple linear regression and correlation</i>)	<ul style="list-style-type: none"> Pengertian dan cara memperoleh korelasi Pengertian regresi linier sederhana Pendugaan koefisien regresi dengan Metode Kuadrat Terkecil Menduga keragaman dari galat Pengujian hipotesis terhadap parameter regresi Ukuran kelayakan model : koefisien determinasi Tabel Analisis Ragam 	12
10	Kecocokan model dan tabel kontingensi (<i>Goodness of fit and contingency table</i>)	<ul style="list-style-type: none"> Uji Khi Kuadrat untuk 'Goodness of Fit' Uji Khi kuadrat untuk uji kebebasan antar dua peubah kategorik 	13
11	Penulisan dan presentasi makalah	<ul style="list-style-type: none"> Mahasiswa diarahkan mencari masalah statistika dalam keseharian yang dapat diselesaikan dalam jangka waktu tertentu dibimbing oleh asisten. Hasilnya ditulis serta dipresentasikan agar tersebar dan mendapat umpan balik. 	14
12	UAS	Ujian Akhir Semester	

Ketentuan Perkuliahan



IPB University
— Bogor Indonesia —

Tugas

01

Mengerjakan soal-soal latihan , kuis, tugas mandiri, pembuatan makalah tentang topik yang terkait, dan mengerjakan tugas-tugas individu maupun kelompok lainnya yang dirasa perlu.

Penilaian

02

Penilaian akhir diberikan di akhir kuliah berdasarkan ujian tengah semester (UTS), ujian akhir semester (UAS), tugas, dan presentasi. Komposisi dari masing-masing adalah 30% UTS, 30% UAS, dan PR (10%), kuis (5%), presentasi (10%), makalah (10%), kehadiran (5%).

Ketentuan Lain

03

- Toleransi waktu keterlambatan: 5 menit
- Berpakaian dan berperilaku sopan sebagaimana ditetapkan dalam aturan IPB
- Mahasiswa wajib mengikuti seluruh kegiatan perkuliahan dan praktikum kecuali berhalangan dengan alasan yang dapat dibenarkan sesuai ketentuan yang berlaku.

Penentuan Huruf Mutu



IPB University
— Bogor Indonesia —

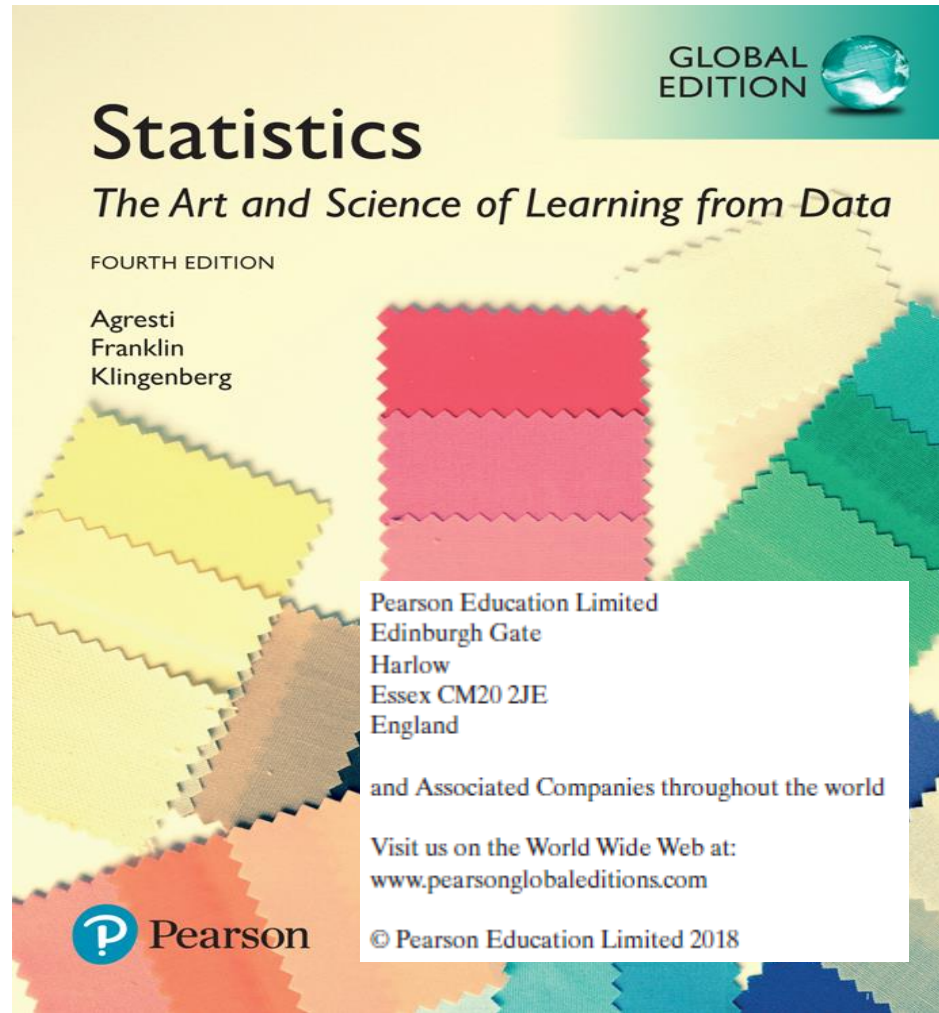
Huruf Mutu: A, AB, B, BC, C, D, dan E

- Penilaian didasarkan pada hasil UTS (30%), UAS (30%) dan PR (10%), kuis (5%), presentasi (10%), makalah (10%), kehadiran (5%).
- Nilai akhir berupa A, AB, B, BC, C, D, dan E
- Nilai A diberikan kepada mahasiswa yang menunjukkan hasil cemerlang dalam UTS, UAS maupun Tugas/responsi/praktikum. Nilai ini kira-kira setara dengan 76 ke atas dalam skala penilaian 0-100.
- Ujian susulan → soal berbeda dengan soal pada ujian utama

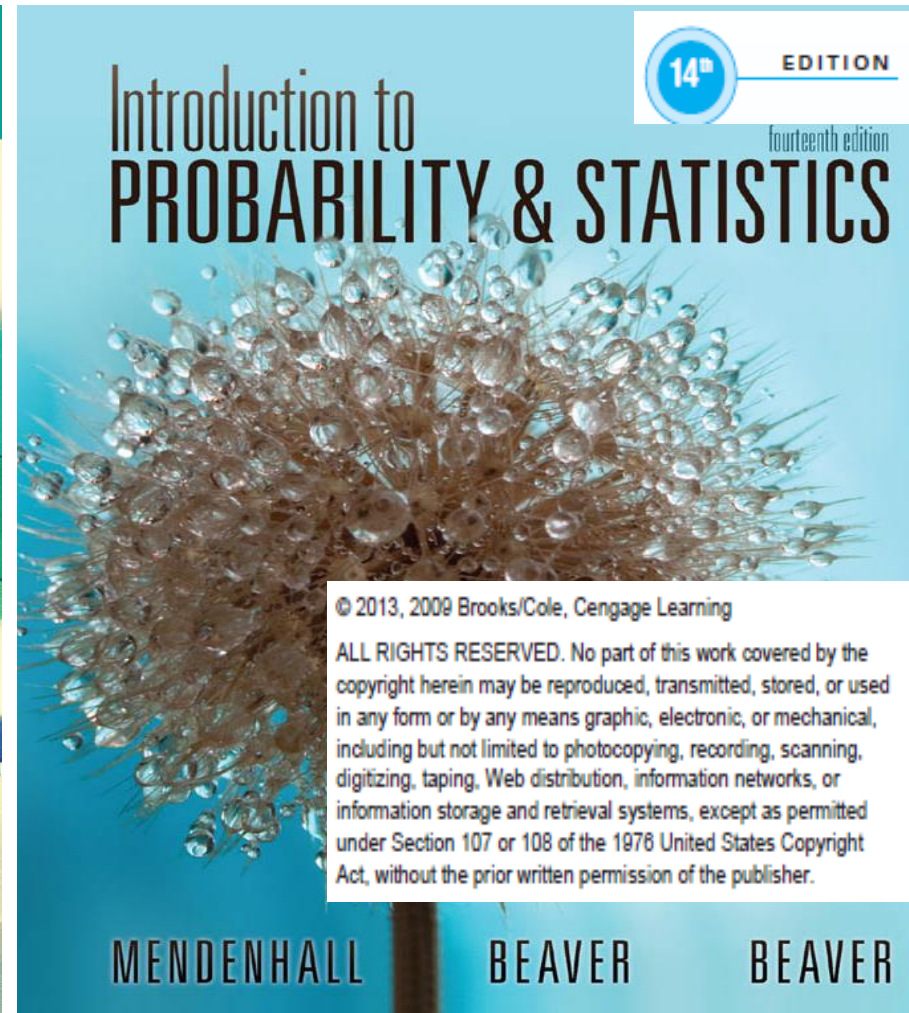
Buku Pegangan



IPB University
— Bogor Indonesia —



Buku Utama



Buku Pelengkap

Rehat dulu...



IPB University
— Bogor Indonesia —

Sesi 1... beres!!!

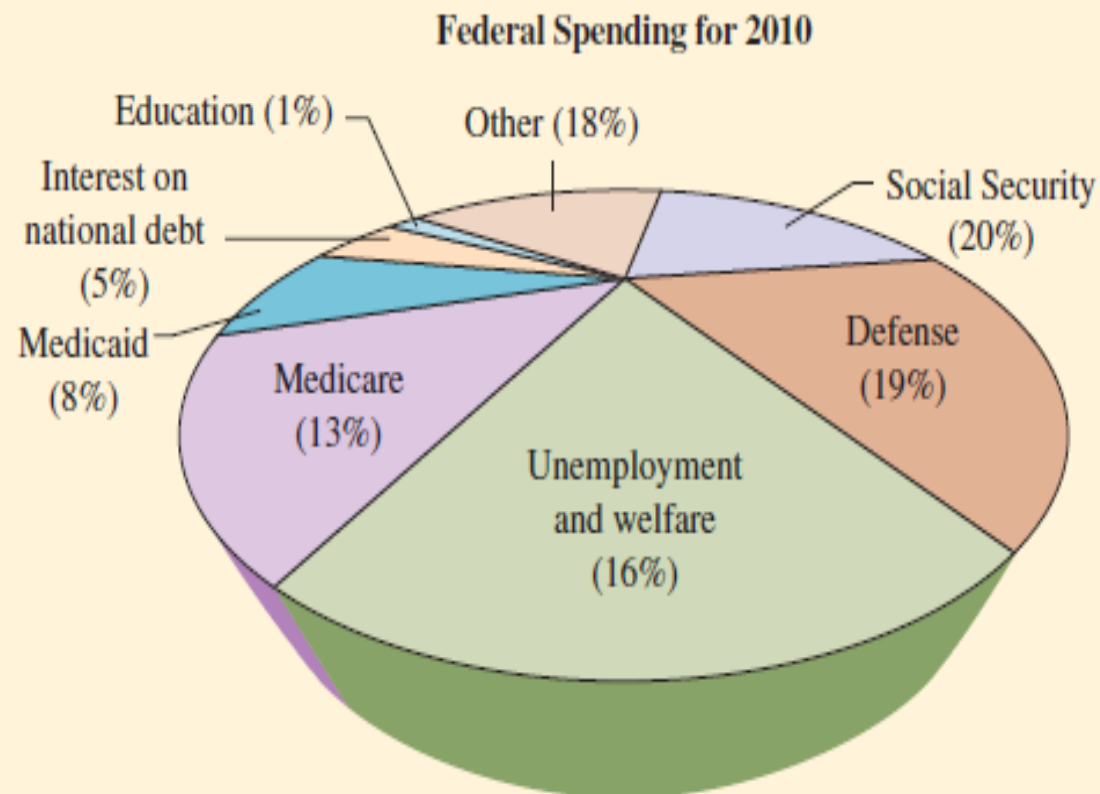


Sesi 2...

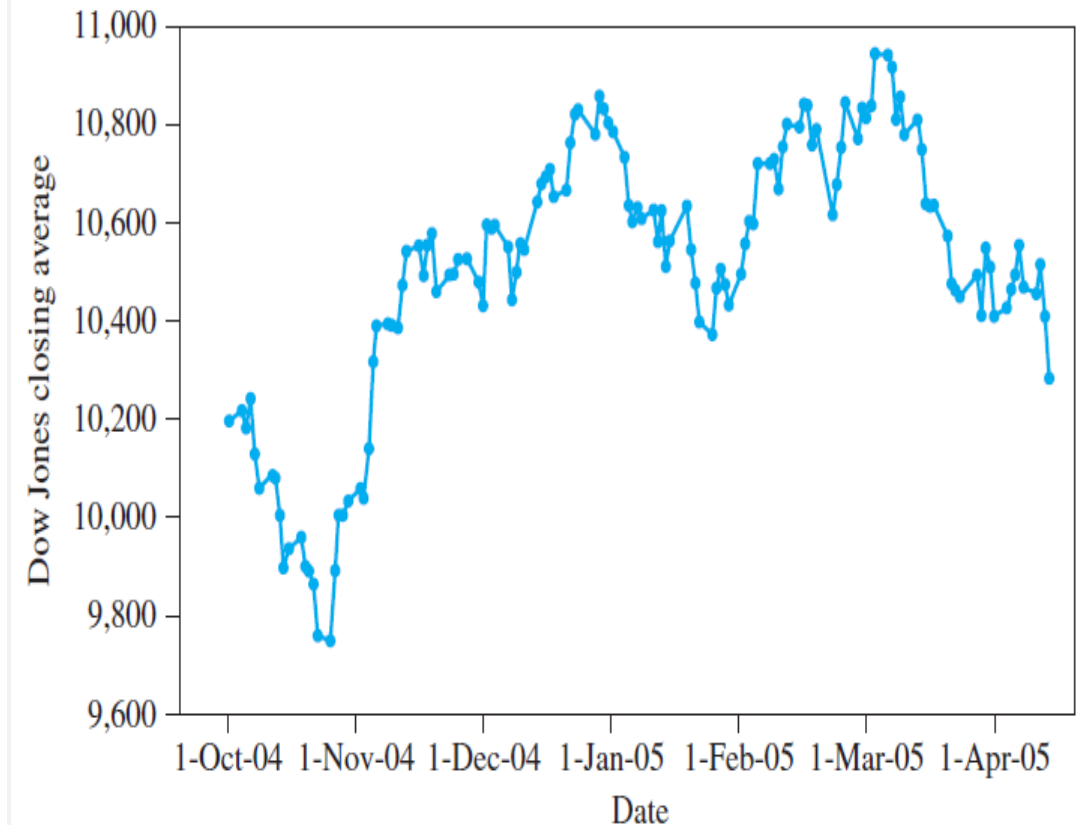


Apa dan Mengapa Statistika?

Pengeluaran Negara 2010



Indeks Saham 2005

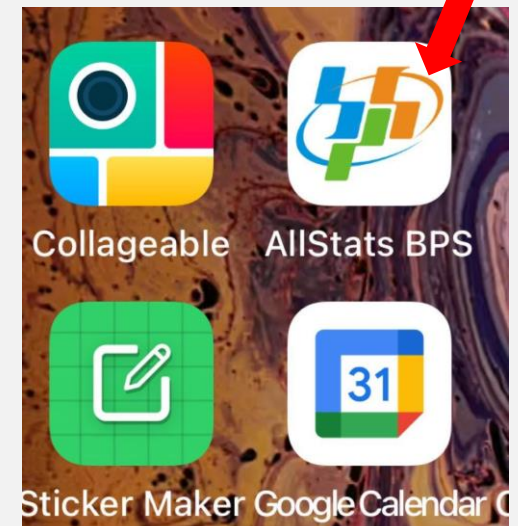


Data dekat dengan kita

Data mengerumuni kita

- Data ada di-mana2 dlm keseharian hidup kita
- Data ada dlm pemberitaan, perdebatan, pemilu, perencanaan dan penelitian
- Statistika menyangkut seluk-beluk data, bahkan jauh melampaui data itu sendiri
- Coba klik tautan ini dan lihat data apa yang ada?
- Coba buka tautan ini: <https://bps.go.id/linkTableDinamis/view/id/895>
- Coba unduh aplikasi “**AllStats BPS**” di HP Anda

Perhatikan data apa saja yang tersedia dalam “**AllStats BPS**” ?



Ilustrasi 1

Ilustrasi ini menunjukkan bagaimana pola pikir statistika dalam menemukan solusi atas suatu permasalahan

- Hipertensi mrpk pembunuh yg takkelihatan di dunia.
- Tahun 1966 di AS → ada kajian efektivitas obat tekanan darah utk hipertensi moderat (130–150 sistolik, 90-100 diastolik)
- Dilibatkan relawan sebanyak 389 pasien dalam pengujian obat tsb:
 - 196 sbg kelompok kontrol (*placebo*) → tdk diberi obat tsb
 - 193 sbg kelompok aktif → diberi obat tsb
- *Questions :*
 - Apakah kelompok aktif lebih baik drpd kelompok kontrol?
 - Apakah kelompok aktif kurang mengalami kerusakan akibat hipertensi?
- Apa saja yang bisa diungkap oleh data yang terkumpul?

Ilustrasi 1

Data dari penelitian ini dapat disajikan menggunakan tabel berikut ini

Pasien	Aktif/Placebo	Tekanan Sistolik	Kejadian hipertensif	Serangan jantung/koroner	Pembengkakan jantung
1	Aktif	125	Tidak	Tidak	Tidak
2	Placebo	140	Ya	Tidak	Ya
3	Aktif	128	Tidak	Ya	Tidak
4	Aktif	124	Ya	Tidak	Tidak
5	Placebo	135	Tidak	Ya	Ya
:	:	:	:	:	:
:	:	:	:	:	:
389	Placebo	150	Ya	Ya	Ya

Ilustrasi 1

Tabel
Jumlah Kasus Berkaitan dengan Jantung dan Tekanan Darah

Uraian		Aktif	Kontrol	
1.	Jumlah kasus dengan tekanan darah ≥ 130	0 (0 %)	24 (12.2 %)	→ konklusif
2.	Jumlah kasus kejadian hipertensif	37	89	→ meyakinkan
3.	Jumlah kasus serangan jantung dan penyakit jantung koroner	35 (18.1%)	38 (19.3%)	→ meyakinkan
4.	Jumlah kasus pembengkakan jantung	12	20	→ perlu pengujian
5.	Jumlah pasien	193	196	

Pentingnya Disain Pengumpulan Data

Andai pengumpulan data sbb:

- a. Tekanan darah sistolik 130-140 → **kel aktif**
- b. Tekanan darah sistolik 141-150 → **kel kontrol**

*Hasilnya akan berbias
menguntungkan produsen obat*

*Bagaimana supaya tidak berbias hasilnya? Maka untuk mengelompokkan ke dalam aktif atau placebo statistisi melakukannya secara acak (**random**).*

Apa yang dimaksud dengan secara acak? Yaitu secara adil, tanpa pilih kasih. Dalam bahasa statistika acak berarti peluang pasien masuk ke dalam kelompok aktif/pasif terukur. Dalam prakteknya bisa dilakukan menggunakan sistem lotere seperti yang kita lihat dalam acara arisan.

Ilustrasi 2

Di AS pernah diperdebatkan apakah markas PBB lebih baik tetap di AS atau dipindahkan ke negara lain? Stasiun TV melakukan jajak pendapat, koran juga melakukan jajak pendapat tapi hasilnya berbeda secara diametral. Mana yg benar?

Pertanyaan: *Apakah Anda setuju jika markas besar PBB tetap ada di Amerika Serikat? (**Ya** setuju/**Tidak** setuju).*

Metode Jajak Pendapat	Ya	Tidak	Jumlah Responden
(1) Responden adalah yg menelpon ke TV (<i>phone-in</i>)	33%	67%	180 000
(2) Responden dipilih acak dari daftar nomor telpon (<i>random</i>)	72%	28%	1 200

Metode (1) atau (2) yang lebih layak dipercaya hasilnya???

telpon masuk (*phone-in*) → **berBIAS** (mengapa?) yaitu ada **bias seleksi** dan **bias respons**

telpon acak (*random poll*) → Biasanya diperkecil karena pemilihan responden dilakukan secara acak

Rehat dulu...



IPB University
— Bogor Indonesia —

Sesi 2... beres!!!



Sesi 3...



Metode Deduksi vs Induksi

Kebenaran dapat diperoleh melalui metode deduksi dan/atau Induksi

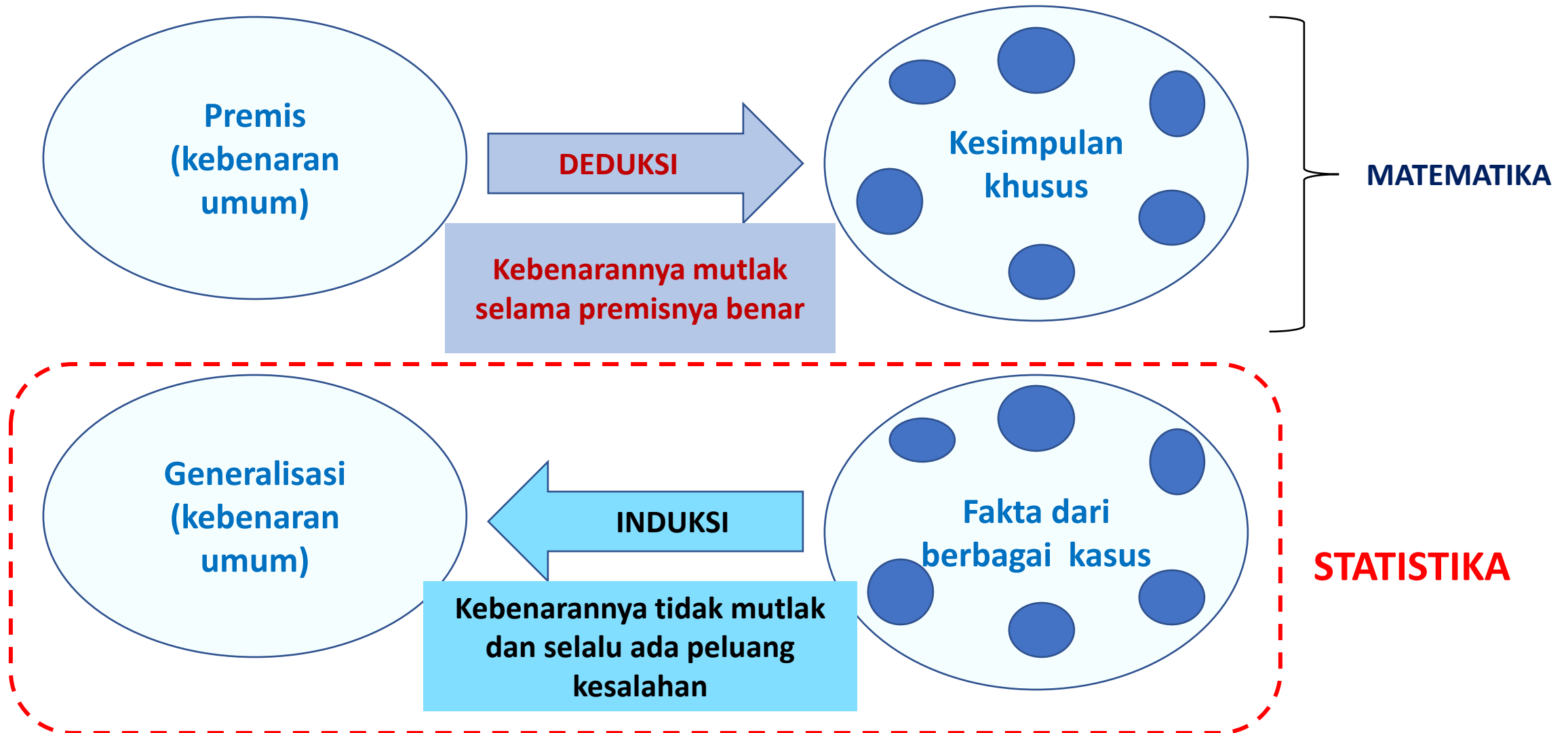
Deduksi

- Dimulai dari premis, yaitu sesuatu yang kita anggap benar (umum). Lalu dari premis ini diturunkanlah kebenaran baru (khusus).
- Metode deduksi yang populer: silogisme, modus tollens dan modus ponens
- Ilustrasi:
Premis: Jika hari hujan maka jalanan basah.
Fakta: Sekarang hari hujan.
Kesimpulan: jalanan basah (kebenaran baru)
- Bagaimana kalau faktanya kita melihat jalanan sedang basah? Apa yang bisa disimpulkan?

Induksi

- Dimulai dari beberapa kasus, artinya pengamatan khusus. Lalu dilakukan generalisasi (umum).
- **Kasus:** Angsa di Aceh berbulu putih, di Semarang juga berbulu putih, di Manado dan Ambon putih.
- **Kesimpulan (umum)** : semua angsa berbulu putih.
- Dalam praktek kita sering ditemui: JIKA p MAKA q ($p \rightarrow q$). Ini deduktif. Pola pikir induktif: apa yang terjadi jika yang diamati adalah q?
- Secara medis seseorang yang kurang tidur wajahnya pucat dan pusing. Jika ada mahasiswa pucat dan pusing. Apakah mahasiswa itu kurang tidur?

Metode Deduksi vs Induksi



Definisi Statistika

Dari ilustrasi ini kita bisa mendefinisikan statistika sebagai ilmu yang mempelajari seluk-beluk tentang data dan statistik.

Statistika adalah ilmu yang **mempelajari seluk beluk data atau statistik**, terutama dalam hal:

- a. Bagaimana cara/teknik mengumpulkan data yang efektif dan efisien;
- b. Bagaimana cara/teknik memodelkan dan menganalisis data sehingga dapat dilakukan deskripsi (*description*), inferensi (*inference*), prediksi (*prediction*) dan bahkan preskripsi (*precription*);
- c. Bagaimana cara/teknik menarik kesimpulan yang sah (*valid*) dan handal (*reliable*).

Statistika Deskriptif dan Inferensi

Secara tradisional statistika dibagi menjadi 2: statistika deskriptif dan statistika inferensi

Statistika Deskriptif

- **Penyajian data** dalam bentuk tabel, grafik atau gambar, serta berbagai ukuran deskriptif: median, rata-rata, simpangan baku, dll.
- **Merupakan upaya agar informasi dasar dari data dapat ditangkap dg jelas.**
- Ingat tadi, dari 193 pasien kel aktif, 19% mengalami kejadian hipertensif → **deskriptif** (hanya fakta)

Statistika Inferensi

- Merupakan proses penarikan kesimpulan populasi (keseluruhan) secara induktif.
- Menggunakan konsep PELUANG untuk menunjukkan ketidakpastian dari kesimpulan.
- Pasien yang mengkonsumsi obat tekanan darah lebih kecil peluangnya utk terkena pembengkakan jantung drpd yang tidak mengkonsumsi obat tersebut → **inferensi** (kita melampaui fakta)

Kesalahan (*error*) dalam Pengumpulan Data

Dalam mengumpulkan data selalu terjadi kesalahan (“error”), ini tidak bisa dihindari, tetapi hanya bisa ditekan sekecil-kecilnya.

Kesalahan percontohan → *sampling error* :

- Kesalahan populasi sasaran
- Kesalahan akibat bias seleksi atau akibat bias respon
- Kesalahan akibat terlalu banyak responden yang tidak bisa menjawab
- Kesalahan akibat alat ukur yang tidak handal

Kesalahan bukan percontohan → *non-sampling error* :

- Kesalahan konsep atau definisi yang ingin diukur
- Kesalahan mencatat, mengentri, atau menganalisis
- Perbedaan kualitas petugas yang mensurvei/mengukur

Kesalahan (*error*) dalam Pengumpulan Data

Potential sources of error

in estimating a population distribution using a sample

Sampling error

Because the sample is not the whole population

Non-sampling error

Poor sampling method

Questionnaire or measurement error

Behavioural effects

“Sampling error” akan semakin kecil jika jumlah responden bertambah banyak.

Reducing Sampling & Non Sampling Errors

Sampling Error



Cause: Small, Un-diverse Sample



Solution: Bigger, More Diverse Sample

Non-Sampling Error



Cause: External Factors

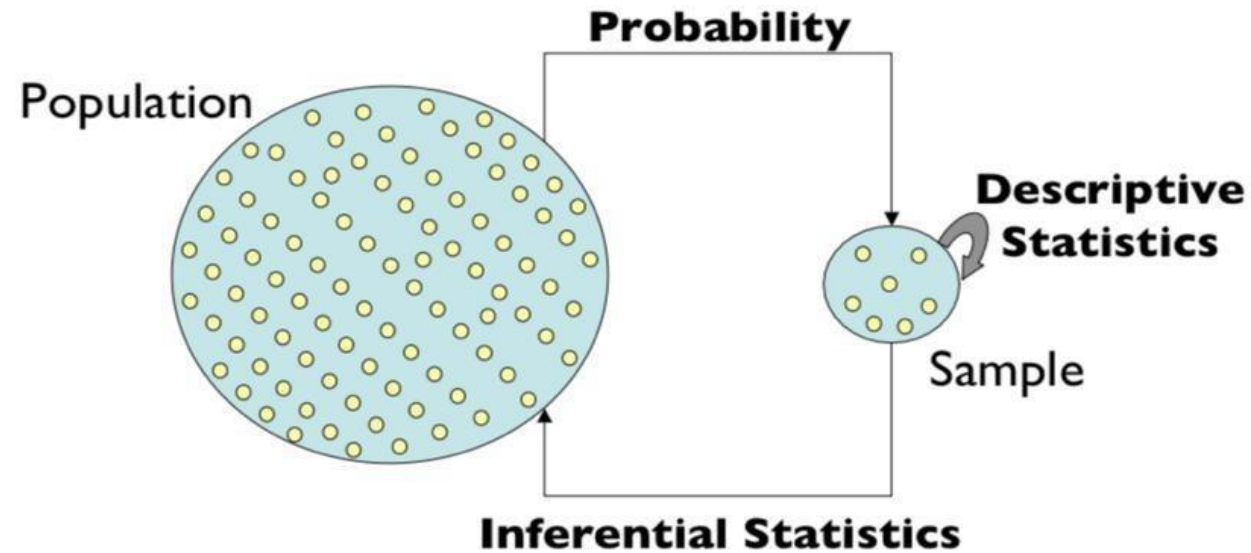


Solution: Study Mechanism Design

“Non-sampling error” akan semakin besar jika jumlah responden bertambah banyak.

Pola Pikir Statistika → Populasi vs Contoh

“Central Dogma” of Statistics



Problem based exercise (sajikan dlm responsi)

**Berikut ini adalah data daerah asal mahasiswa STK 57
(Jabodetabek atau Non-jabodetabek)**

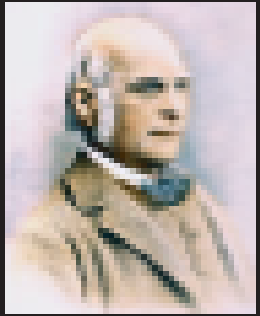
Mhs	Asal	Mhs	Asal	Mhs	Asal	Mhs	Asal	Mhs	Asal	Mhs	Asal	Mhs	Asal	Mhs	Asal
1	N	11	J	21	N	31	N	41	J	51	N	61	N	71	J
2	J	12	N	22	N	32	J	42	N	52	N	62	J	72	J
3	J	13	J	23	N	33	N	43	N	53	J	63	N	73	J
4	N	14	N	24	N	34	J	44	J	54	J	64	N	74	J
5	J	15	N	25	J	35	J	45	N	55	N	65	J	75	N
6	J	16	J	26	N	36	N	46	N	56	J	66	N	76	N
7	N	17	N	27	N	37	N	47	J	57	N	67	N	77	N
8	N	18	J	28	J	38	J	48	N	58	J	68	N	78	J
9	J	19	N	29	J	39	J	49	N	59	N	69	N	79	N
10	J	20	J	30	N	40	J	50	J	60	N	70	J		

Problem based exercise (sajikan dlm responsi)

Lakukan hal berikut ini terhadap data asal mahasiswa STK 57

1. Hitunglah berapa % mhs STK 57 yang dari Jabodetabek!!!
2. Anda sebagai calon statistisi hanya mengamati sebagian saja dari mahasiswa itu; coba ambil secara acak 10 mhs sebagai contoh dan jelaskan secara lengkap bagaimana cara Anda mengambilnya, dan sajikan mhs nomor berapa saja yang terambil.
3. Dari data contoh nomor 2, hitunglah brp % mhs STK 57 yang dari Jabodetabek! Apakah hasilnya sama dengan hasil pada nomor 1? Jelaskan.
4. Ulangi, ambil contoh acak sebanyak 30 mhs. Hitunglah brp % mhs STK 57 yang dari Jabodetabek! Apakah hasilnya sama dengan hasil pada nomor 1 dan nomor 2? Jelaskan!
5. Menurut Anda mana yang lebih bagus untuk mengambil kesimpulan, hasil nomor 2 atau nomor 4.
6. Tulislah laporan singkat (2-3 halaman) yg memuat: Judul, Permasalahan (yang ingin dijawab), Metode (bagaimana cara Anda menjawabnya), Hasil (statistik deskriptif), Kesimpulan.

Tokoh Statistika



Francis Galton

On the Shoulders of... Francis Galton

Heights and peas—what do they have in common?

—Francis Galton (1822–1911)

Francis Galton, like his cousin Charles Darwin, was interested in genetics. In his work on inheritance, he collected data on parents' and children's heights. He found that, on average, the heights of children were not as extreme as those of the parents and tended toward the mean height. There was regression toward the mean. This refers to this average, not to every observation. It does not imply that over the course of many generations, heights will all vary less from the mean height.



Galton also conducted experiments examining the weight of sweet peas. The mean weight of offspring peas was closer than the weight of the parent group to the mean for all peas. Again, Galton observed regression toward the mean.

Galton discovered a way to summarize linear associations numerically. He named this numerical summary the “coefficient of correlation.” In 1896, the British statistician Karl Pearson (who proposed the chi-squared test for contingency tables in 1900) derived the current method for estimating the correlation using sample data.





IPB University
— Bogor Indonesia —

Departement of Statistics

Study Program in Statistics and Data Science

