

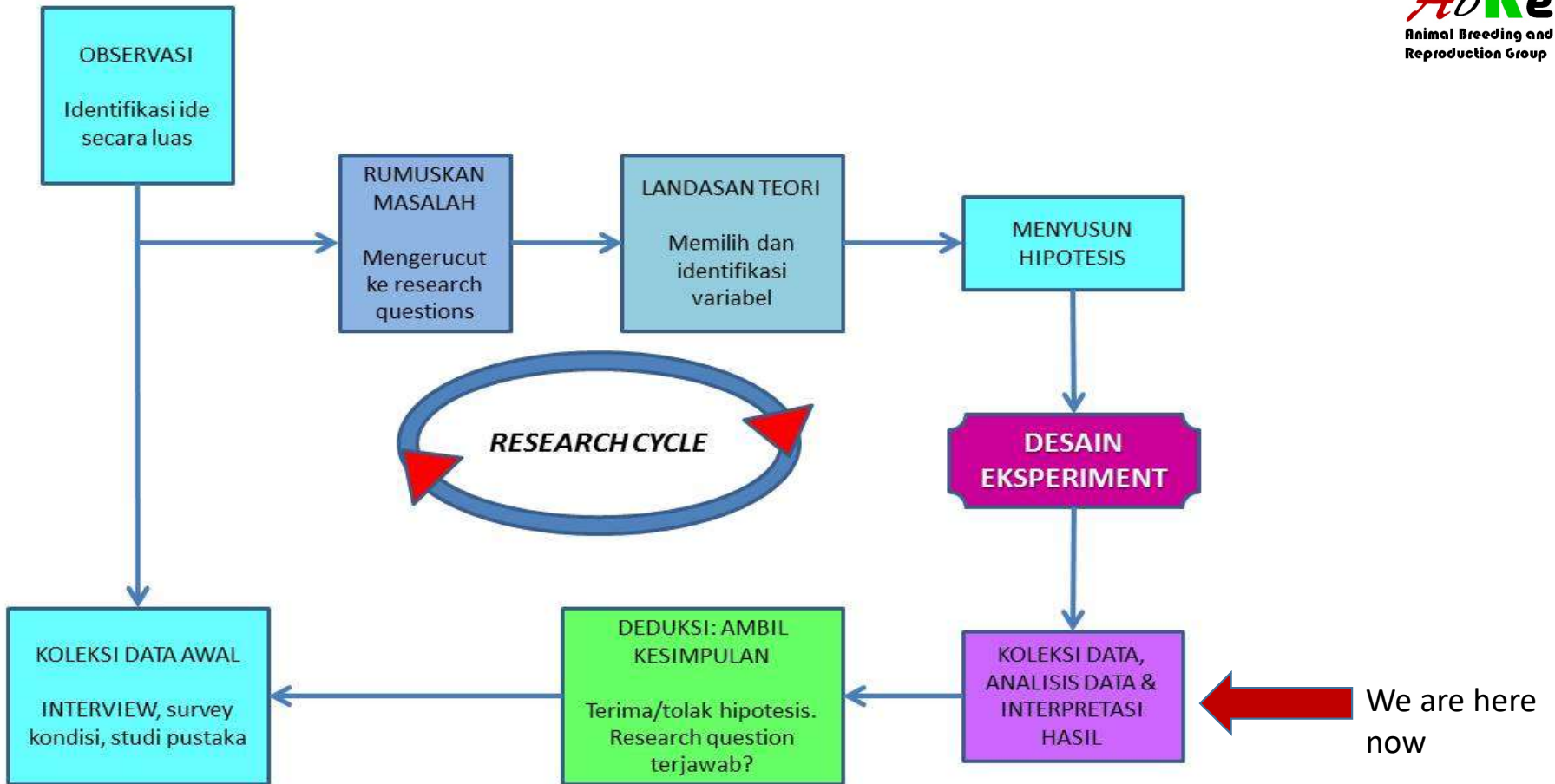


Rancob 4A: Metode analisis data I

Nuzul

AbRe
Animal Breeding and
Reproduction Group





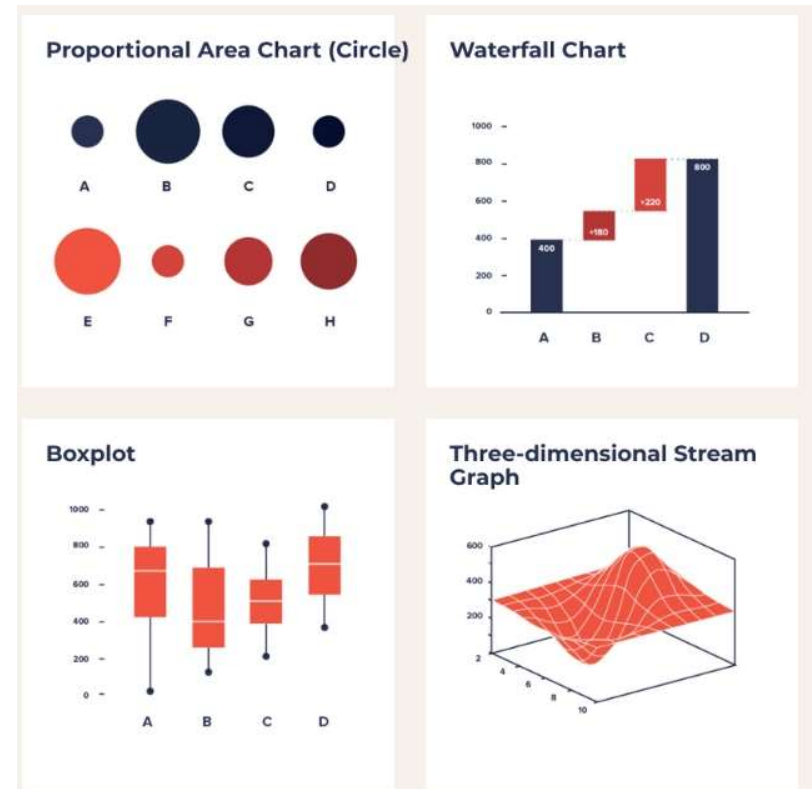


- Setelah topik atau masalah diformulasikan lewat research questions, kemudian peneliti menyusun landasan teori.
- Berikutnya hipotesis akan dibangun berdasarkan landasan teori.
- Untuk menguji benar atau tidaknya hipotesis diperlukan uji hipotesis; dan untuk melakukan uji hipotesis, diperlukan data yang dihasilkan dari eksperimen
- Ekperimen perlu didesain dengan baik agar sesuai dengan tujuan penelitian dan efisien, disinilah fungsi rancob.
- Setelah dirancang, eksperimen dilaksanakan dan data dikumpulkan.
- Data yang terkumpul akan dianalisis untuk menguji hipotesis yang telah dibuat di depan. Metode statistic untuk analisis data diaplikasikan di tahap ini

Analysis 1: Exploratory data analysis (EDA)

Memahami EDA

- EDA merupakan metode yang digunakan untuk merangkum karakteristik utama dari data dan memberikan gambaran akan sebuah data; cara yang sering digunakan adalah visualisasi berupa grafik atau table.
- Tidak ada kesimpulan spesifik yang diambil lewat metode analisis ini
- Tujuan utama EDA adalah untuk melihat trend data, ada atau tidaknya outlier/pencilan, distribusi data dll



EDA 1: Mean ± SD

Mean adalah rerata dari beberapa angka: merupakan hasil perhitungan nilai sentral dari seperangkat angka.

Sample Mean	Population Mean
$\bar{x} = \frac{\sum X}{n}$	$\mu = \frac{\sum X}{N}$

where $\sum X$ is sum of all data values

N is number of data items in population

n is number of data items in sample

- Standar deviasi adalah ukuran keragaman atau pola sebaran suatu data set.
- SD adalah akar dari varians atau ragam
- SD yang rendah menunjukkan bahwa nilai2 dari data berada sangat dekat dengan rerata dan sebaliknya

Formula

$$\sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{N}}$$

σ = population standard deviation

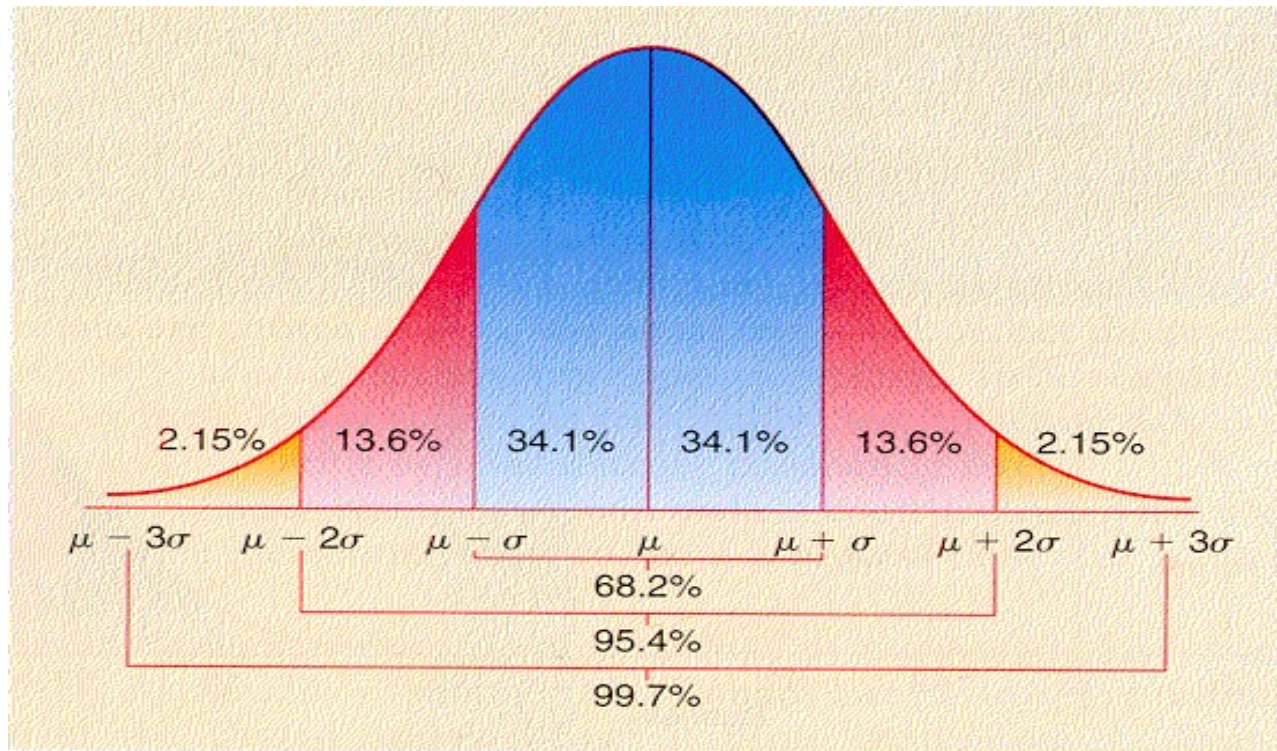
N = the size of the population

x_i = each value from the population

μ = the population mean

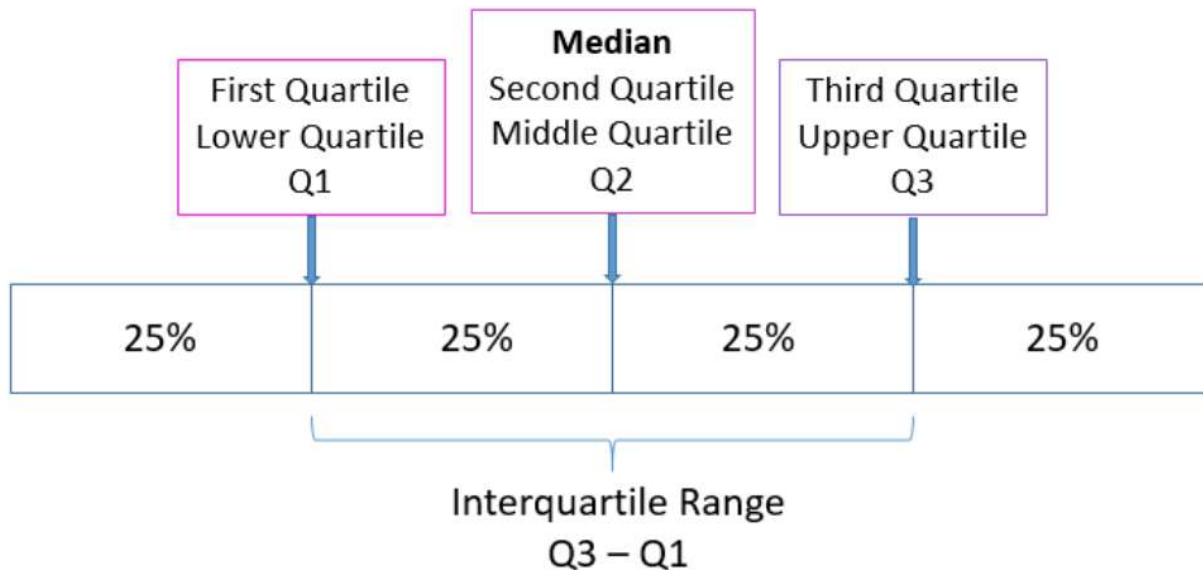
***Suplemen untuk contoh penghitungan akan diberikan dalam catatan tersendiri**

Mean dan SD dapat digunakan untuk kurva distribusi normal



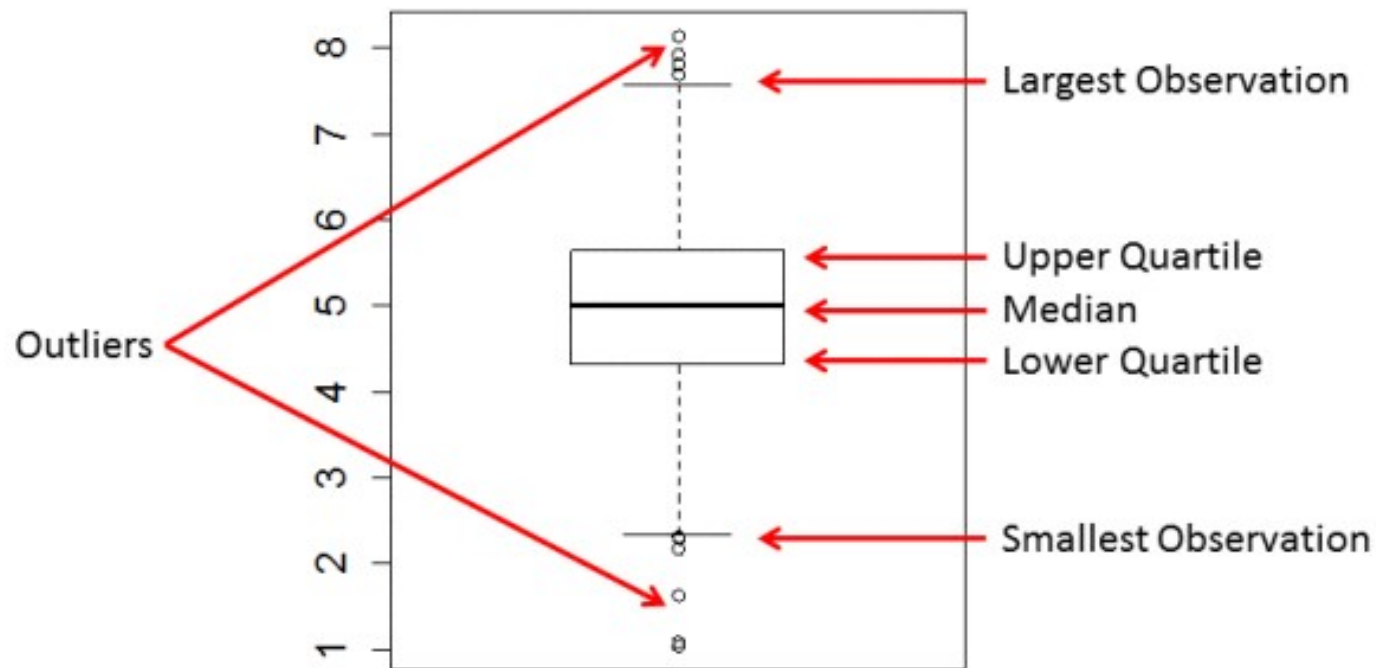
EDA 2: Median dan Quartil

Median and Quartiles



- Median adalah nilai tengah dari sebuah dataset ketika nilai-nilai data tersebut diurutkan dari kecil ke besar.
- Sedangkan kuartil adalah parameter yang dapat membagi data yang sudah diurutkan tadi menjadi empat bagian sama besar

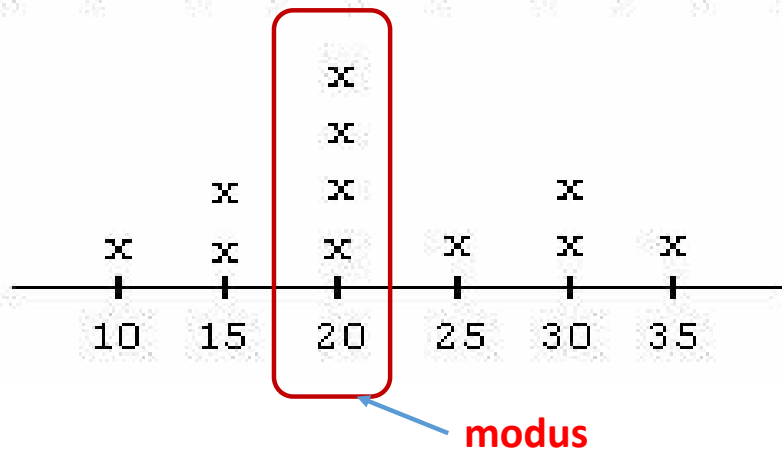
Median dan quartil digunakan sebagai dasar membangun box plot



EDA 3: Modus (Mode)

- Nilai yg paling sering muncul
- Merupakan rerata nominal: Deskriptif statistik tentang rentangan data yang telah diurutkan

10, 20, 15, 20, 25, 30, 35, 20, 20, 30, 15



Contoh:

EDA untuk rangkuman data hasil penelitian

Table 2. Summary statistics of the pre-weaning traits

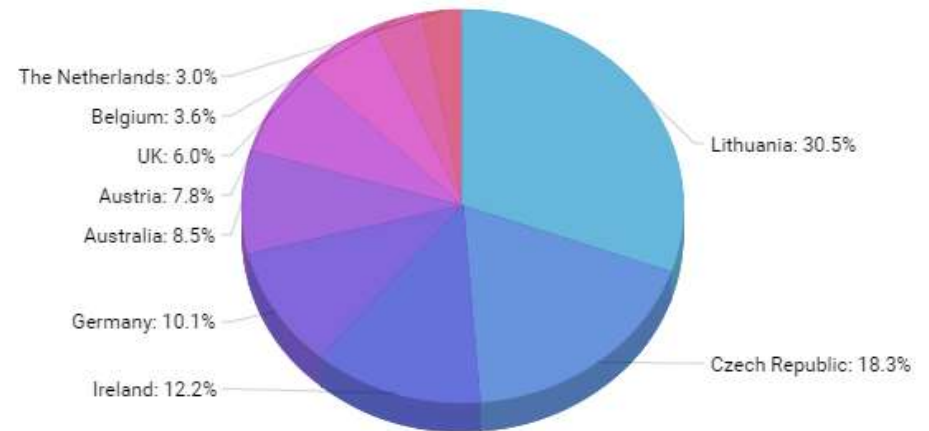
Traits	Breed	N	Mean \pm sd ^{sd} *	Median	Minimum	Maximum
Birth weight (kg)	B×B	461	3.197 \pm 0.605 ^a	3.000	1.500	5.500
	B×J	2164	2.835 \pm 0.580 ^c	3.000	1.500	5.500
	B×(B×J)	1179	3.003 \pm 0.618 ^b	3.000	1.500	5.000
Weaning weight (kg)	B×B	461	15.069 \pm 3.652 ^a	14.933	8.000	27.397
	B×J	2164	13.488 \pm 3.666 ^b	13.151	5.490	26.111
	B×(B×J)	1179	13.518 \pm 3.942 ^b	13.333	5.063	28.000
Average daily gain (kg)	B×B	463	0.154 \pm 0.046 ^a	0.154	0.058	0.297
	B×J	2169	0.138 \pm 0.045 ^b	0.135	0.032	0.287
	B×(B×J)	1183	0.137 \pm 0.049 ^b	0.134	0.016	0.318

*~~sd~~^{sd} = standard deviation; ^{a, b, c} values followed with different superscripts showed significant

difference ($\alpha=0.05$)

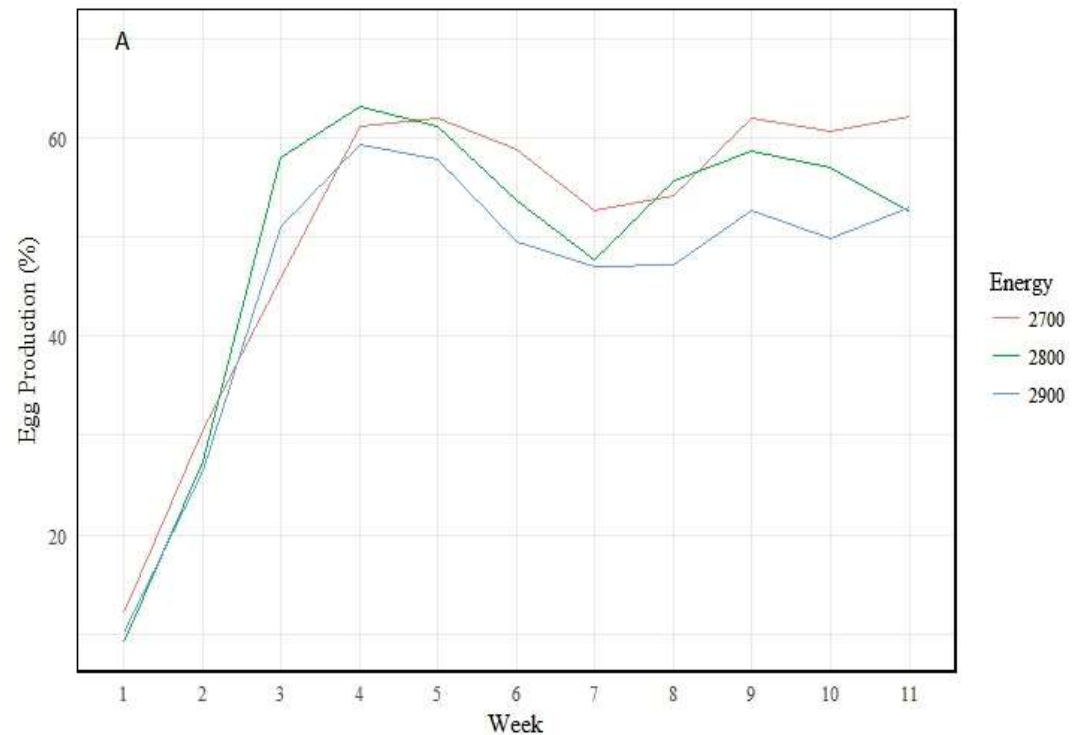
EDA 4: Proporsi

- Proporsi adalah metode pengelompokan dimana tiap kelompok merupakan fraksi atau bagian dari dataset
- Biasanya disajikan dalam bentuk fraksi ataupun prosentase
- Proporsi sebih sering divisualisasikan dalam bentuk grafik



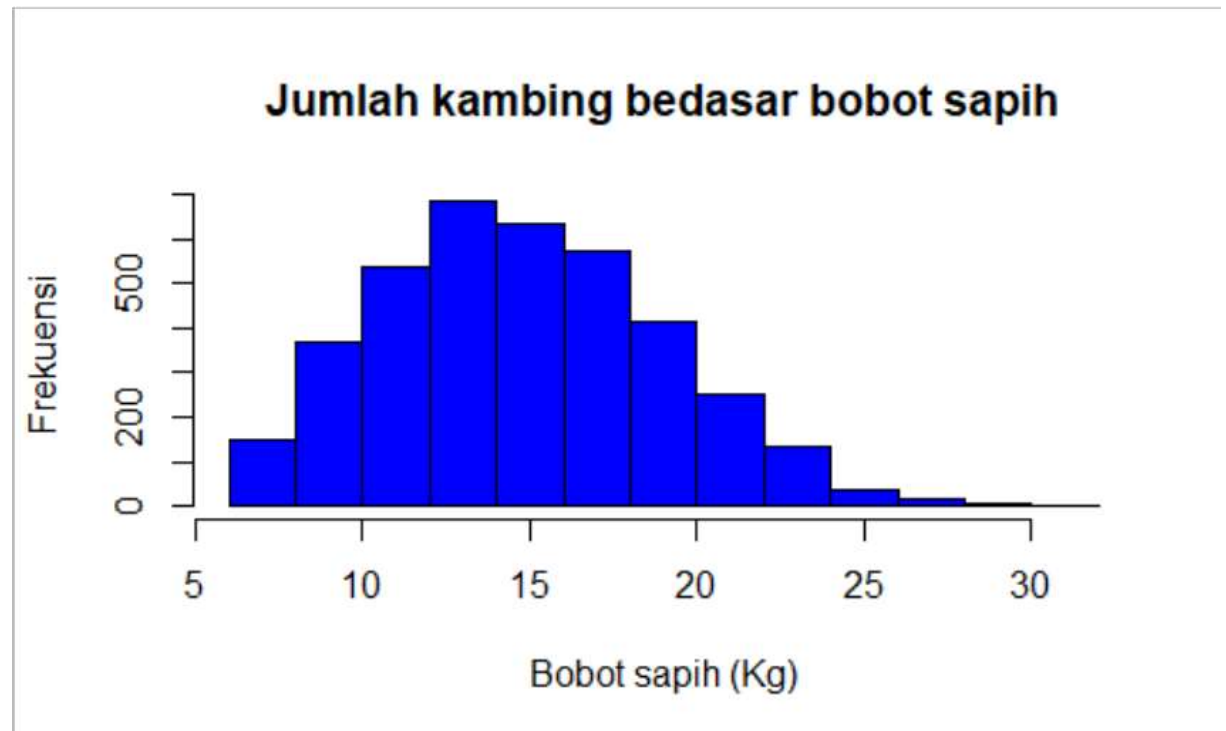
EDA 5: Trend line

- Trend line atau pola menunjukkan pola dari suatu data set
- Biasanya pola berdasarkan waktu
- Contoh: pola produksi telur dan susu



EDA 6: Frekuensi

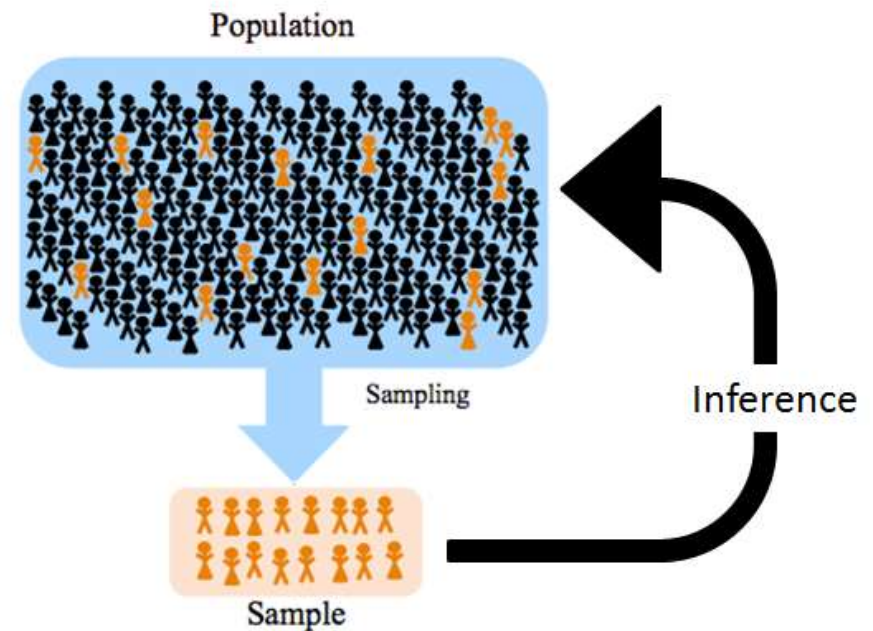
- Menunjukkan jumlah pengamatan dalam suatu range nilai data tertentu



Analisis 2: Uji Hipotesis

Definisi

- Secara statistic, hipotesis adalah sebuah asumsi yang dibuat tentang parameter suatu populasi, yang dalam hal ini adalah populasi eksperimen.
- Asumsi ini bisa jadi benar atau tidak benar
- Uji hipotesis mengacu kepada prosedur statistic formal yang dapat dimanfaatkan untuk menentukan apakah kita akan menerima atau menolak suatu hipotesis
- Karena uji hipotesis dapat digunakan untuk menentukan kesimpulan dari suatu eksperimen, maka sering juga disebut sebagai statistic inferensial



Hipotesis

- Hipotesis ada dua macam:
 - Hipotesis null: dilambangkan dengan H_0 adalah hipotesis yang menyatakan bahwa hasil dari observasi atau eksperimen terjadi murni karena kebetulan (peluang atau probabilitas).
 - Hipotesis alternative dilambangkan sebagai H_A atau H_1 adalah hipotesis yang menyatakan bahwa observasi atau hasil eksperimen dipengaruhi oleh suatu penyebab yang sistematis, yang sering disebut sebagai **factor**.

Contoh: Sebuah penelitian bertujuan untuk mengetahui pengaruh dua jenis ransum dengan modifikasi kandungan protein terhadap bobot badan ayam broiler. Maka perlakuan yang diberikan adalah:

P0 : Ransum basal (**kontrol**)

P1 : Ransum basal + ekstra protein 0.5%

P2 : Ransum basal + ekstra protein 1%

Hipotesis yang dapat dibangun dari topik diatas adalah sbb:

- H_0 : ekstra protein tidak berpengaruh nyata terhadap bobot badan;
- Atau bisa juga dituliskan: $\mu_{P0} = \mu_{P1} = \mu_{P2}$ sedangkan
- H_A : ekstra protein berpengaruh terhadap bobot badan ayam;
- Atau $\mu_{P0} \neq \mu_{P1} \neq \mu_{P2}$

Analisis data yang dipelajari dalam MK rancob ini hanya yang termasuk dalam kategori statistik parametrik

Statistik parametrik adalah suatu cabang statistic dimana data yang akan diolah harus memenuhi asumsi-asumsi berikut:

- Data yang akan dianalisis merupakan sample dari populasi yang memiliki distribusi normal, dengan rerata sebesar μ dan standar deviasi sebesar σ
- Terdapat hubungan yang linier antara respons variable (variable yang diamati = y) dengan factor (variable yang memberikan pengaruh = x)
- Independen: antar titik data tidak memiliki hubungan yang sistematis
- Homoscedaticity: dalam satu factor, varians error antar level adalah setara atau homogen