

# I PENDAHULUAN

## 1.1 Latar Belakang

Segala aspek kehidupan manusia dipermudah dengan berkembangnya internet yang pesat saat ini, salah satu aspeknya yaitu aspek sistem jual beli. Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) pada tahun 2018 menyelenggarakan survei jumlah pengguna internet di Indonesia. Hasilnya 171 juta orang atau 64,8% dari total penduduk Indonesia menggunakan internet (APJII 2019). Sistem jual beli di Indonesia yang awalnya bersifat konvensional beralih ke sistem jual beli digital, perubahan ini dipicu karena tingginya jumlah pengguna internet di Indonesia. Salah satu inovasi dalam perkembangan sistem perdagangan digital adalah *marketplace*. Toko-toko digital yang terbentuk akibat perkembangan sistem perdagangan digital menawarkan berbagai barang kebutuhan konsumen, salah satunya adalah buku anak. Toko-toko digital ini diwadahi oleh *Marketplace*. Orang tua yang tidak memiliki banyak waktu untuk membeli buku bagi anaknya terbantu dengan adanya *marketplace*. Selain itu *marketplace* juga memudahkan orang tua untuk mencari toko yang menjual buku anak yang lengkap dalam waktu singkat.

Banyaknya toko digital di *marketplace* membuat calon konsumen memiliki banyak pilihan untuk berbelanja, namun tidak semua toko digital menyediakan produk buku anak secara lengkap. Hal ini membuat konsumen memiliki permasalahan baru yaitu memilih toko digital yang menyediakan produk buku anak secara lengkap. Oleh sebab itu perlu dilakukan identifikasi toko-toko digital yang menyediakan produk buku anak secara lengkap dengan aspek jenis dan tema buku anak. Salah satu teknik statistika yang relevan untuk diterapkan pada hal ini adalah analisis gerombol. Penggerombolan toko digital di *marketplace* diperlukan untuk mempermudah konsumen memilih toko digital yang menyediakan buku anak yang ingin dibeli. Penggerombolan juga bermanfaat bagi toko digital agar lebih berinovasi dalam menyediakan jenis dan tema buku anak secara lengkap.

Analisis gerombol merupakan metode peubah ganda yang memiliki tujuan utama untuk mengelompokkan amatan-amatan berdasarkan kemiripan karakteristik yang ada (Mattjik dan Sumertajaya 2011). Selain untuk data numerik, analisis gerombol dapat diterapkan pada data kategorik misalnya data biner. Data biner adalah data kategorik yang hanya memiliki dua kategori yang umumnya diberi inisial 0 dan 1. Penggerombolan untuk data biner berbeda dengan penggerombolan data numerik dalam hal ukuran kedekatan (Kariyam 2010). Mengukur kedekatan antara dua objek dapat dilakukan dengan menghitung frekuensi data dalam tabel kontingensi pada nilai yang sama dan nilai yang tidak sama, ukuran kedekatan pada penggerombolan data biner disebut ukuran kesamaan. Terdapat berbagai macam ukuran kesamaan yang digunakan pada penggerombolan data biner. Lima di antara ukuran kesamaan tersebut adalah *simple matching*, Jaccard, Sokal-Sneath 2, Roger-Tanimoto, dan Dice.

## 1.2 Tujuan Penelitian

Tujuan dari penelitian ini adalah mengidentifikasi penggerombolan karakteristik toko-toko yang ada disuatu *marketplace* berdasarkan kelengkapan jenis dan tema buku anak menggunakan analisis gerombol data biner.

@Hak cipta milik IPBUniversity

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
  - b. Pengutipan tidak merugikan kepentingan yang wajar IPBUniversity.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPBUniversity.

## II TINJAUAN PUSTAKA

### 2.1 Marketplace

*Online marketplace* merupakan tempat untuk melakukan pemasaran barang atau jasa melalui media internet. *Online Marketplace* berasal dari dua kata, yaitu *online market* dan *place*. *Online market* mempunyai arti semua usaha yang dilakukan untuk memasarkan suatu barang atau jasa melalui media internet. *Place* mempunyai arti tempat. Maka (Sadgotra *et al.* 2013).

*Marketplace* memiliki fungsi seperti pasar konvensional pada umumnya yaitu tempat bertemunya antara penjual dan pembeli. Perbedaan *marketplace* dengan pasar konvensional berada pada interaksi yang terjadi antara penjual dan pembeli, apabila di *marketplace* interaksi yang terjadi tidak dibatasi oleh ruang maupun waktu, sedangkan di pasar konvensional interaksi yang terjadi terbatas oleh ruang maupun waktu. *Marketplace* tercipta karena perkembangan yang begitu pesat pada teknologi internet dan secara tidak langsung membentuk sebuah pasar dan arena perdagangan (Indrajit 2002).

### 2.2 Analisis Gerombol

Analisis gerombol merupakan metode peubah ganda yang memiliki tujuan utama untuk mengelompokkan amatan-amatan berdasarkan kemiripan karakteristik yang ada (Mattjik dan Sumertajaya 2011). Analisis gerombol memiliki dua metode, yaitu metode penggerombolan berhirarki dan metode penggerombolan tak berhirarki (Anderberg 1973). Proses penggerombolan berhirarki dilakukan secara bertahap dan penggerombolan yang terbentuk terjadi secara alamiah (Mattjik dan Sumertajaya 2011). Amatan-amatan yang terbentuk ke dalam gerombol memiliki karakteristik yang mirip sehingga keragaman yang dihasilkan kecil, sedangkan amatan-amatan antar gerombol yang terbentuk memiliki karakteristik yang tidak mirip sehingga keragaman yang dihasilkan tinggi. Untuk menentukan seberapa mirip suatu amatan dengan amatan lain adalah dengan cara menghitung ukuran kemiripan atau ketidak miripan antar amatan. Pasangan amatan yang mempunyai jarak yang lebih pendek akan lebih mirip satu sama lain dan berbeda dengan amatan yang mempunyai jarak yang jauh. Sehingga analisis gerombol menghasilkan suatu kelompok yang homogen dalam satu kelompok dan heterogen antar kelompoknya (Mattjik dan Sumertajaya 2011). Penggerombolan dapat digunakan sebagai pelapisan data dalam penggolongan tipe objek (Siswadi dan Suharjo 1999).

Ukuran jarak dibutuhkan untuk setiap pasang amatan yang akan dikelompokkan (Manly dan Alberto 2016). Mengukur kemiripan dan ketidak miripan antar amatan bisa dilakukan dengan jarak Euclid. Jarak euclid antara dua amatan pada  $n$ -peubah,  $\mathbf{u}' = [u_1, u_2, \dots, u_n]$  dan  $\mathbf{v}' = [v_1, v_2, \dots, v_n]$  dapat dihitung dengan menggunakan persamaan sebagai berikut (Johnson dan Wichern 2002):

$$d(\mathbf{u}, \mathbf{v}) = \sqrt{(\mathbf{u} - \mathbf{v})'(\mathbf{u} - \mathbf{v})}$$

Pada penggerombolan hirarki terdapat dua proses untuk membentuk gerombol yaitu pemisahan (*divisive*) dan penggabungan (*agglomerative*). Proses

penggabungan (*agglomerative*) diawali dengan semua objek dianggap sebagai gerombol. Setiap amatan kemudian akan bergabung dengan amatan lain berdasarkan kemiripannya. Proses penggabungan berakhir ketika semua amatan membentuk satu gerombol. Terbalik dari proses penggabungan, proses pemisahan dimulai dengan semua amatan dianggap satu gerombol. Semua amatan secara bertahap akan berpisah dan membentuk gerombol-gerombol baru sampai satu gerombol hanya terdiri atas satu amatan. Metode hirarki memiliki beberapa cara dalam proses penggabungan amatan diantaranya adalah pautan tunggal (*Single Linkage*), pautan lengkap (*Complete Linkage*), pautan rata-rata (*Average Linkage*), *Centroid*, *Median*, dan *Ward* (Johnson dan Wichern 2002). Pautan Lengkap dipilih karena subjektif penulis dan tidak ada ketentuan khusus dalam memilih metode perbaikan jarak. Pautan lengkap menggunakan jarak terjauh antara dua amatan  $i$  dan  $j$  pada gerombol  $k$  yang berbeda dalam proses penggabungannya.

$$d_{((ij),k)} = \max \{ d_{(ik)}, d_{(jk)} \}$$

Hasil akhir dari metode penggerombolan berhirarki berbentuk diagram pohon atau dendogram. Jumlah gerombol yang terbentuk ditentukan dari pemisahan gerombol dengan cara pemotongan dendogram pada selisih jarak penggabungan yang terbesar (Everitt *et al.* 2011). Hasil penggerombolan tergantung pada amatan yang digunakan, peubah yang diamati, ukuran kesamaan yang digunakan serta metode penggerombolan yang digunakan (Rencher dan Christensen 2011).

### 2.3 Analisis Gerombol pada Data Biner

Penggerombolan pada data biner dapat dilakukan dengan metode penggerombolan berhirarki. Perbedaannya dengan data numerik terletak pada ukuran kedekatan yang digunakan. Ukuran kedekatan yang digunakan jika seluruh peubah bertipe biner adalah ukuran kesamaan (*similarity measure*) (Kariyam 2010). Menghitung ukuran kesamaan dapat dilakukan dengan *Operational Taxonomic Units* (OTU), langkah awal dimulai dengan membuat tabel kontingensi berukuran  $2 \times 2$  antara vektor  $i$  dan  $j$  yang bertipe biner. Vektor  $i$  dan  $j$  mempunyai ukuran  $1 \times p$ , dengan  $p$  merupakan banyaknya peubah yang diamati. Anggota ke- $n$  dari vektor  $i$  adalah  $i_n$ , akan bernilai 1 apabila peubah ke- $n$  tersedia pada amatan I dan  $i_n$  akan bernilai 0 apabila peubah ke- $n$  tidak tersedia pada amatan I. Anggota ke- $n$  dari vektor  $j$  adalah  $j_n$ , akan bernilai 1 apabila peubah ke- $n$  tersedia pada amatan II dan  $j_n$  akan bernilai 0 apabila peubah ke- $n$  tidak tersedia pada amatan II (Choi *et al.* 2010). Misalkan dua buah amatan I dan amatan II diamati pada  $p$  peubah yang bertipe biner, maka tabel kontingensi ukuran  $2 \times 2$  disajikan sebagai berikut:

Tabel 1 Tabel kontingensi data biner pada dua objek

$i \backslash j$	1 (tersedia)	0 (tidak tersedia)	total
1 (tersedia)	$a$	$c$	$a + c$
0 (tidak tersedia)	$b$	$d$	$b + d$
total	$a + b$	$c + d$	$p = a + b + c + d$

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
    - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
    - b. Pengutipan tidak merugikan kepentingan yang wajar IPBUniversity.
  2. Dilarang mengumumkannya dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPBUniversity.

Keterangan :

- $a$  : banyaknya peubah yang tersedia pada kedua amatan
- $b$  : banyaknya peubah yang tersedia pada amatan I saja
- $c$  : banyaknya peubah yang tersedia pada amatan II saja
- $d$  : banyaknya peubah yang tidak muncul atau tidak tersedia pada kedua amatan

Nilai  $p$  akan selalu sama dengan nilai  $a+b+c+d$ . Nilai  $a$  dan  $d$  disebut nilai sepadan sedangkan nilai  $b$  dan  $c$  disebut nilai tidak sepadan (Choi *et al.* 2010). Terdapat berbagai ukuran kesamaan untuk peubah biner yang diperoleh dari Tabel 1. Adapun lima ukuran kesamaan yang digunakan antara vektor  $i$  dan  $j$  ( $s_{ij}$ ) adalah sebagai berikut:

### 1. Simple Matching

Sokal dan Michener (1958) mengemukakan bahwa ukuran untuk menghitung kesamaan antara dua objek secara sederhana dapat dilakukan dengan menghitung semua nilai sepadan (*positive matches* dan *negative matches*), lalu membaginya dengan jumlah kemungkinan tertinggi dari nilai yang sama. Hal tersebut menjadi dasar dalam menghitung ukuran kesamaan *simple matching* yang dirumuskan sebagai berikut (Timm 2002):

$$S_{ij} = \frac{a + d}{p}$$

### 2. Rogers-Tanimoto

Perhitungan ukuran kesamaan pada ukuran Rogers-Tanimoto dilakukan dengan memberi bobot ganda pada nilai-nilai tidak sepadan. Oleh karena itu, semakin besar nilai-nilai tidak sepadannya maka dua buah objek semakin tidak mirip. Ukuran kesamaan Rogers-Tanimoto dirumuskan sebagai berikut (Timm 2002):

$$S_{ij} = \frac{a + d}{a + d + 2(b + c)}$$

### 3. Jaccard

Jaccard *similarity* biasa digunakan untuk membandingkan objek-objek dan menghitung nilai kemiripan dari dua buah objek (Komang 2017). Jika ketiadaan ciri-ciri/karakteristik pada kedua objek ( $d$ ) dianggap tidak memberikan informasi apapun, maka nilai  $d$  tidak seharusnya dimasukkan ke dalam perhitungan ukuran kesamaan (Gower 1985). Hal tersebut yang menjadi dasar dari ukuran kesamaan Jaccard untuk tidak memasukkan nilai  $d$  seperti yang dirumuskan sebagai berikut (Timm 2002):

$$S_{ij} = \frac{a}{a + b + c}$$

## 4. Dice

Dice (1945) berpendapat bahwa dua objek dapat digabungkan atau dikatakan mirip hanya karna kesamaan ciri-ciri yang ada pada keduanya bukan pada ciri-ciri yang tidak ada. Oleh karena itu, pada ukuran kesamaan Dice, nilai sepadan yang negatif ( $d$ ) tidak terlalu dipentingkan. Ukuran Dice dirumuskan sebagai berikut (Timm 2002):

$$s_{ij} = \frac{2a}{2a + b + c}$$

## Sokal-Sneath 2

Ukuran Sokal-Sneath 2 memberikan bobot yang sama pada objek yang tidak sepadan. Struktur rumus ukuran Sokal-Sneath 2 merupakan variasi dari ukuran Jaccard dengan bobot ganda diberikan pada nilai-nilai tidak sepadan. Namun nilai  $d$  tidak dimasukkan dalam perhitungan. Ukuran kesamaan Sokal-Sneath 2 dirumuskan sebagai berikut (Timm 2002):

$$s_{ij} = \frac{a}{a + 2(b + c)}$$

Ukuran *simple matching*, Dice, Jaccard, Rogers-Tanimoto, dan Sokal-Sneath 2 memiliki batasan nilai dari 0 sampai 1. Semakin besar nilai kesamaannya maka kedua objek dikatakan semakin mirip (Podani 2000). Ukuran-ukuran kesamaan tersebut hanya sesuai untuk data biner dan kurang sesuai untuk data ordinal yang memiliki lebih dari dua kategori (Kariyam 2010). Dengan menggunakan ukuran-ukuran kesamaan tersebut, variabel biner dapat digerombolkan dengan menggunakan metode gerombol berhirarki. Namun, ukuran kesamaan ( $s_{ij}$ ) tersebut harus diubah menjadi ukuran ketaksamaan ( $d_{ij}$ ) dengan rumus sebagai berikut (Timm 2002):

$$d_{ij} = 1 - s_{ij}$$

## 2.4 Evaluasi Hasil Gerombol

Secara umum, ada dua hal penting yang harus diperhatikan dalam mengevaluasi hasil penggerombolan, yaitu banyaknya gerombol yang terbentuk dan keanggotaan gerombol. Hasil penggerombolan pada data kategorik dapat dievaluasi menggunakan rasio keragaman. Rasio keragaman merupakan nilai rasio antara simpangan baku dalam gerombol (SSW) dengan simpangan baku antargerombol (SSB). Hasil penggerombolan dikatakan baik apabila nilai rasio keragamannya kecil. Nilai rasio keragaman yang kecil menunjukkan keragaman dalam gerombol rendah dan keragaman antar gerombol tinggi. (Light dan Margolin 1971).

Nilai rasio keragaman untuk data kategorik dirumuskan sebagai berikut (Light dan Margolin 1971):

$$SSW = \left[ \frac{1}{n-G} \left( \frac{n}{2} - \frac{1}{2} \sum_{g=1}^G \frac{1}{n_g} \sum_{k=1}^K n_{kg}^2 \right) \right]^{1/2}$$

$$SSB = \left[ \frac{1}{G-1} \left( \frac{1}{2} \left( \sum_{g=1}^G \frac{1}{n_g} \sum_{k=1}^K n_{kg}^2 \right) \right) - \frac{1}{2n} \sum_{k=1}^K n_k^2 \right]^{1/2}$$

Rasio keragaman =  $SSW/SSB$

Keterangan :

$n$  : total keseluruhan amatan

$G$  : banyaknya gerombol yang terbentuk

$K$  : banyaknya kategori

$n_k$  : banyaknya amatan kategori ke- $k$

$n_g$  : banyaknya amatan pada gerombol ke- $g$

$n_{kg}$  : banyaknya amatan kategori ke- $k$  gerombol ke- $g$

@Hak cipta milik IPBUniversity

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
  - b. Pengutipan tidak merugikan kepentingan yang wajar IPBUniversity.
2. Dilarang mengumumkannya dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPBUniversity.

IPBUniversity





### III METODE

#### 3.1 Data

Data yang digunakan dalam penelitian ini merupakan data ketersediaan jenis-jenis buku anak dan tema-tema buku anak yang berasal dari toko-toko digital di suatu *marketplace*. Pengambilan data dilakukan secara manual pada tanggal 7 Januari hingga 20 Januari 2020. Pemilihan toko digital dilakukan dengan *non probability sampling*, tidak ada kriteria khusus toko digital yang dipilih. Data yang digunakan dalam penelitian ini bertipe biner dengan 2 kategori, yaitu kategori 0 menggambarkan ketidakterediaan dan kategori 1 menggambarkan ketersediaan buku anak pada toko *online*.

Toko *online* buku anak yang digunakan dalam penelitian ini berjumlah 200 toko. Data memiliki 19 peubah jenis buku anak dan 13 peubah tema buku anak dan 1 peubah buku lainnya. Peubah buku lainnya mencakup jenis buku anak *magic torch book*, *maze book* dan tema buku anak kemandirian. Daftar peubah dalam penelitian ini ditampilkan pada Tabel 2 dan Tabel 3.

Tabel 2 Daftar peubah jenis buku anak

Peubah	Keterangan	Tipe	Kategori
X1	<i>Paper book</i>	Biner	1=ada, 0=tidakada
X2	<i>Board book</i>	Biner	1=ada, 0=tidakada
X3	<i>Cloth book</i>	Biner	1=ada, 0=tidakada
X4	<i>Bath book</i>	Biner	1=ada, 0=tidakada
X5	<i>Sound book</i>	Biner	1=ada, 0=tidakada
X6	<i>Pop up book</i>	Biner	1=ada, 0=tidakada
X7	<i>Touch &amp; feel book</i>	Biner	1=ada, 0=tidakada
X8	<i>Lift the flap book</i>	Biner	1=ada, 0=tidakada
X9	<i>Slide out book</i>	Biner	1=ada, 0=tidakada
X10	<i>Puppet book</i>	Biner	1=ada, 0=tidakada
X11	<i>Jigsaw book</i>	Biner	1=ada, 0=tidakada
X12	<i>Magnet book</i>	Biner	1=ada, 0=tidakada
X13	<i>Wipe clean activity book</i>	Biner	1=ada, 0=tidakada
X14	<i>Sparkle book</i>	Biner	1=ada, 0=tidakada
X15	<i>Peek a boo book</i>	Biner	1=ada, 0=tidakada
X16	<i>Sticker book</i>	Biner	1=ada, 0=tidakada
X17	<i>Ligth beam book</i>	Biner	1=ada, 0=tidakada
X18	<i>Atlas book</i>	Biner	1=ada, 0=tidakada
X19	<i>Augmented Reality book</i>	Biner	1=ada, 0=tidakada

Tabel 3 Daftar peubah tema buku anak

Peubah	Keterangan	Tipe	Kategori
X20	Buku cerita	Biner	1=ada, 0=tidakada
X21	Buku mengenal huruf	Biner	1=ada, 0=tidakada

@Hak Cipta Dilindungi Undang-undang

IPB University

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :  
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah  
 b. Pengutipan tidak merugikan kepentingan yang wajar IPBUniversity.  
 2. Dilarang mengumumkannya dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPBUniversity.



Tabel 3 Daftar peubah tema buku anak (*lanjutan*)

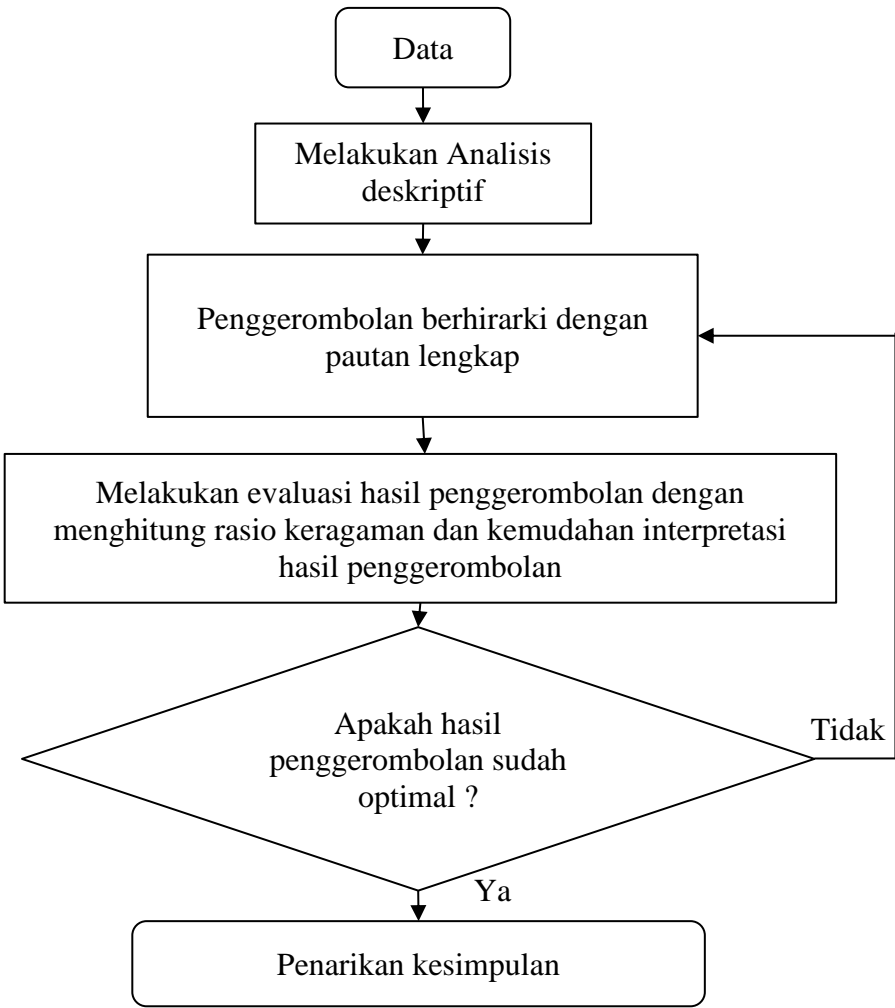
Peubah	Keterangan	Tipe	Kategori
X22	Buku mengenal angka	Biner	1=ada, 0=tidakada
X23	Buku tentang agama Islam	Biner	1=ada, 0=tidakada
X24	Buku tentang agama Kristen	Biner	1=ada, 0=tidakada
X25	Buku tentang agama Katolik	Biner	1=ada, 0=tidakada
X26	Buku tentang keseharian	Biner	1=ada, 0=tidakada
X27	Buku pembentukan karakter	Biner	1=ada, 0=tidakada
X28	Buku mengenal warna	Biner	1=ada, 0=tidakada
X29	Buku mengenal binatang	Biner	1=ada, 0=tidakada
X30	Buku mengenal tumbuhan	Biner	1=ada, 0=tidakada
X31	Buku mengenal benda	Biner	1=ada, 0=tidakada
X32	Buku mengenal kendaraan	Biner	1=ada, 0=tidakada
X33	Lainnya	Biner	1=ada, 0=tidakada

### 3.1 Prosedur Analisis Data

Tabel 4 Tahapan Prosedur Analisis Data

Tahap	Prosedur
1	Melakukan analisis deskriptif untuk melihat sebaran dan karakteristik data buku anak.
2	Melakukan analisis gerombol berhirarki dengan pautan lengkap dengan lima ukuran kesamaan ( $s_{ij}$ ) untuk data buku anak di <i>marketplace</i> X. Ukuran kesamaan tersebut adalah <i>simple matching</i> , Dice, Jaccard, Rogers-Tanimoto, Sokal-Sneath 2.
3	Melakukan evaluasi hasil penggerombolan untuk mendapatkan penggerombolan optimal dengan melihat nilai rasio keragaman terkecil dan kemudahan interpretasi hasil penggerombolan akhir. Jika hasil penggerombolan sudah optimal dengan nilai rasio keragaman yang terkecil dan kemudahan interpretasi hasil gerombol yang terbentuk maka lanjut ke tahap 4, jika hasil penggerombolan tidak optimal, maka kembali ke tahap 2 dengan menggunakan ukuran kesamaan yang lain.
4	Penarikan kesimpulan dengan mendeskripsikan hasil penggerombolan akhir untuk melihat karakteristik setiap gerombol toko buku anak.

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
    - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
    - b. Pengutipan tidak merugikan kepentingan yang wajar IPBUniversity.
  2. Dilarang mengumumkkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPBUniversity.

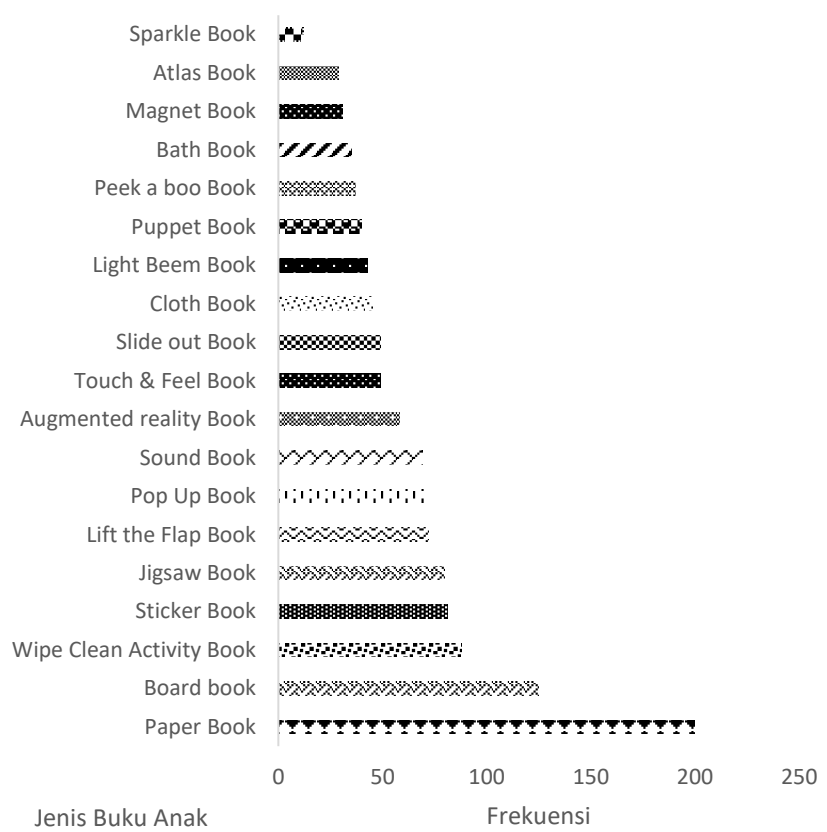


Gambar 1 Diagram alir prosedur analisis data

## IV HASIL DAN PEMBAHASAN

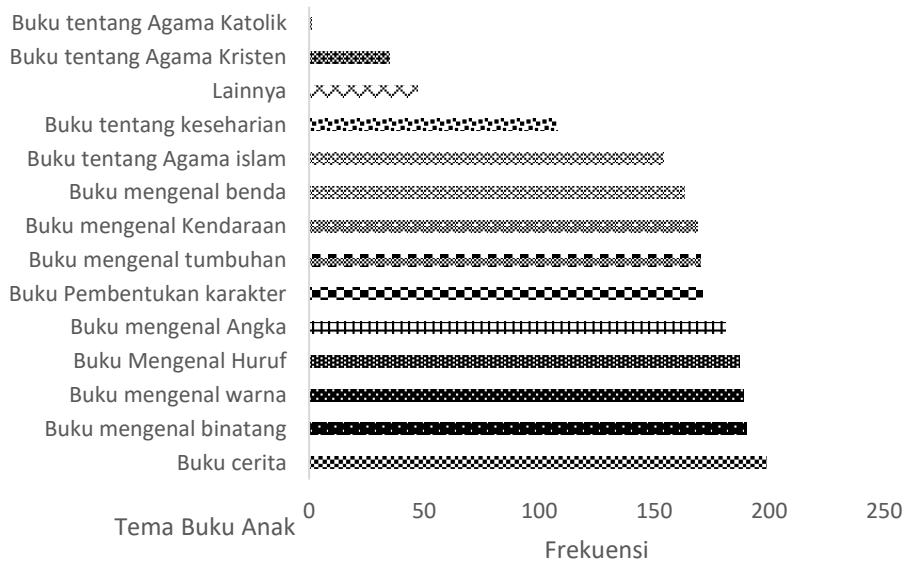
### 4.1 Data

Deskripsi data diperlukan untuk mengetahui sebaran data serta gambaran umum karakteristik dari data buku anak. Data buku anak memiliki 33 peubah biner yang terdiri dari 19 peubah jenis buku anak, 14 peubah tema buku anak. Peubah terkait jenis dan tema buku anak ditampilkan pada Gambar 2 dan Gambar 3. Peubah jenis dan tema buku anak tersebut menunjukkan tersedia atau tidaknya suatu jenis dan tema buku anak pada suatu toko digital.



Gambar 2 Diagram batang peubah jenis buku anak

Hak Cipta Dilindungi Undang-undang  
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :  
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah  
b. Pengutipan tidak merugikan kepentingan yang wajar IPBUniversity.  
2. Dilarang mengumumkkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPBUniversity.



Gambar 3 Diagram batang peubah tema buku anak

Gambar 2 menggambarkan banyaknya jenis buku anak yang tersedia pada toko digital. Pada Gambar 2 terlihat bahwa dari 200 toko digital, *paper book* dan *board book* paling banyak tersedia dengan 200 dan 125 toko digital, sedangkan jenis buku *sparkle book* paling sedikit tersedia dengan 12 toko digital. Pada Gambar 3 dapat dilihat ketersediaan tema buku anak pada setiap toko digital. Tema buku anak buku cerita paling banyak tersedia yaitu sebesar 199 toko digital, sedangkan tema buku anak tentang agama katolik paling sedikit tersedia dengan hanya 1 toko digital. Hal ini menunjukkan mayoritas toko digital menyediakan buku anak dengan jenis buku *paper book*, *board book* dan tema buku anak yang beragam.

## 4.2 Hasil Penggerombolan

### 4.2.1 Penentuan Gerombol

Penggerombolan dilakukan berdasarkan aspek jenis dan tema buku anak. Hasil penggerombolan akan digunakan untuk melihat beda antar karakteristik setiap gerombol yang terbentuk dari toko digital. Setiap gerombol buku anak yang terbentuk akan memberikan sebuah kesimpulan yang berguna untuk toko digital dan konsumen buku anak di *marketplace X* kedepannya.

Tabel 5 Rasio keragaman setiap ukuran kesamaan pada penggerombolan toko buku anak

Ukuran kesamaan	Banyaknya gerombol	Rasio keragaman
Simple matching	3	0,008
Rogers-Tanimoto	3	0,008
Jaccard	5	0,012
Dice	5	0,012
Sokal-Sneath 2	5	0,012

Tabel 5 menunjukkan jumlah gerombol yang terbentuk dan rasio keragaman pada masing-masing ukuran kesamaan dengan pautan lengkap. Terlihat bahwa *simple matching* dan Rogers-Tanimoto memiliki nilai rasio keragaman yang sama. Hal yang sama juga terjadi pada ukuran kesamaan Dice, Jaccard dan Sokal-Sneath 2. Dilihat dari struktur rumusnya, ukuran *simple matching* dan Rogers-Tanimoto memasukkan nilai sepadan yang negatif. Hal yang sama juga berlaku untuk ukuran kesamaan Dice, Jaccard dan Sokal-Sneath 2. Ukuran Dice, Jaccard, dan Sokal-Sneath 2 mengeluarkan nilai sepadan yang negatif. Tetapi, ukuran Dice memberikan bobot ganda pada nilai sepadan dan ukuran Sokal-Sneath memberikan bobot ganda pada nilai tidak sepadan.

Berdasarkan hasil tersebut dapat dikatakan bahwa pemberian bobot ganda pada nilai-nilai sepadan atau tidak sepadan tidak membuat hasil penggerombolan yang terbentuk jadi lebih baik bila digunakan pautan lengkap berdasarkan nilai rasio keragamannya (Finch 2005). Pendekatan untuk mengevaluasi hasil penggerombolan yang optimal adalah dengan melihat nilai rasio keragaman terkecil dan jumlah gerombol yang terbentuk untuk mempermudah hasil interpretasi penggerombolan akhir. Ukuran kesamaan *simple matching* dan Rogers-Tanimoto membentuk 3 gerombol, dendogram yang sama (Lampiran 5 dan 6), dan memiliki nilai rasio keragaman yang lebih kecil daripada ukuran kesamaan lainnya. Hal ini berarti hasil penggerombolan yang terbentuk sudah baik secara rasio keragaman terkecil.

Namun, dari jumlah gerombol yang terbentuk untuk kemudahan interpretasi hasil penggerombolan dengan ukuran kesamaan *simple matching* dan Rogers-Tanimoto kurang memberikan kesimpulan yang bermanfaat dan relevan untuk toko digital dan konsumen buku anak di *marketplace X*. Maka dari itu ukuran kesamaan Jaccard, Dice, dan Sokal-Sneath 2 dipilih karena menghasilkan nilai rasio keragaman yang kecil dan membentuk lima gerombol optimal dengan kemudahan interpretasi hasil penggerombolan serta kesimpulan yang bermanfaat dan relevan bagi para toko digital dan konsumen buku anak di *marketplace X*.

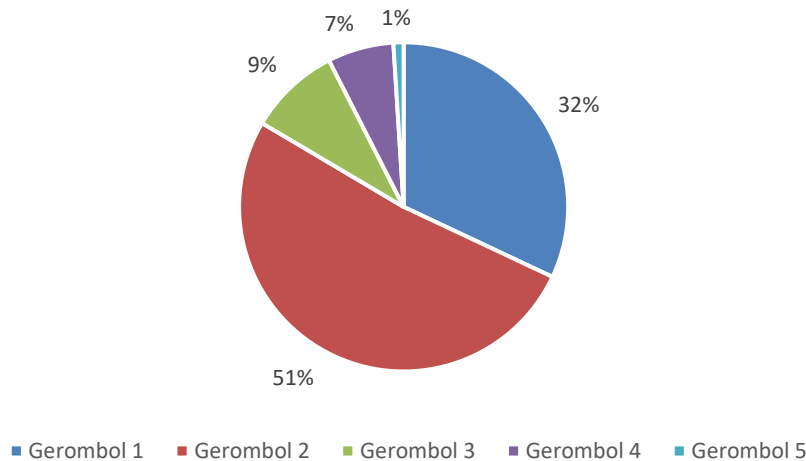
Penggerombolan dengan Jaccard, Dice, dan Sokal-Sneath 2 memberikan hasil penggerombolan yang sama, hal ini disebabkan dendogram yang terbentuk tidak berbeda (Lampiran 2-4), sehingga dapat disimpulkan bahwa penggerombolan menggunakan pautan lengkap dengan ukuran Jaccard, Dice, dan Sokal-Sneath 2 akan menghasilkan anggota yang sama pada tiap

gerombolnya. Hasil penggerombolan dengan ketiga ukuran kesamaan tersebut menghasilkan jumlah gerombol dan anggota pada setiap gerombol yang sama.

Dengan demikian, memilih salah satu saja dari ketiga ukuran kesamaan tersebut dapat dilakukan apabila digunakan pautan lengkap. Setelah mendapatkan hasil penggerombolan, selanjutnya akan ditarik kesimpulan karakteristik setiap gerombol yang terbentuk.

#### 4.2.2 Deskripsi Karakteristik Hasil Gerombol

Banyaknya gerombol yang dihasilkan oleh penggerombolan data biner dengan pautan lengkap dan ukuran kesamaan terpilih adalah lima gerombol. Sebaran hasil penggerombolan dapat dilihat pada Gambar 4. Berdasarkan gambar tersebut dapat dilihat bahwa gerombol kedua memiliki anggota terbanyak dengan 103 toko buku anak dengan persentase 51%, sedangkan gerombol kelima menjadi gerombol dengan anggota paling sedikit yaitu 2 toko buku anak dengan persentase 1%.



Gambar 4 *Pie chart* persentase anggota pada setiap gerombol dengan pautan lengkap

Banyaknya toko buku anak yang menyediakan jenis dan tema buku anak pada setiap gerombol yang terbentuk disajikan pada Tabel 5. Berdasarkan Tabel 5 dapat diketahui bahwa gerombol pertama memiliki anggota 64 toko buku anak dengan karakteristik toko buku anak yang menyediakan jenis buku anak *paper book*, *board book*, *jigsaw book*, dan tema buku anak buku cerita, buku mengenal huruf dan angka, buku mengenai agama islam, buku tentang keseharian dan pembentukan karakter, buku mengenal warna, binatang, tumbuhan, benda, dan kendaraan. Gerombol kedua memiliki anggota 103 toko buku anak dengan karakteristik toko buku anak yang menyediakan semua jenis buku anak dan tema buku anak, kecuali buku anak dengan tema agama katolik. Gerombol ketiga memiliki anggota 18 toko buku anak dengan karakteristik toko buku anak yang hanya menyediakan jenis buku *paper book* dan tema buku anak buku cerita, buku mengenal huruf dan angka, buku mengenai agama islam, buku pembentukan karakter, buku mengenal warna dan binatang. Gerombol keempat memiliki anggota yaitu 13 toko buku anak dengan karakteristik toko

buku anak yang hanya menyediakan jenis buku *paper book* dan tema buku anak buku cerita, buku mengenai agama islam, dan buku membangun karakter. Gerombol kelima memiliki anggota 2 toko buku anak dengan karakteristik toko buku anak yang menyediakan jenis buku *paper book*, *board book*, *augmented reality book* dan tema buku anak buku cerita, buku mengenai agama islam, dan buku mengenai agama kristen.

Tabel 6 Karakteristik hasil penggerombolan toko buku anak

Gerombol	Anggota Gerombol	Tema Buku	Jenis Buku
1	64	Buku cerita, Buku mengenal huruf, angka, Buku pembentukan karakter, Buku mengenal warna, binatang, tumbuhan, benda dan kendaraan	<i>Paper book, board book, dan jigsaw book</i>
2	103	Buku cerita, Buku mengenal huruf, angka, buku mengenai agama islam, buku keseharian, buku pembentukan karakter, buku mengenal warna, binatang, tumbuhan, benda kendaraan	<i>Paper book, board book, sound book, pop up book, lift the flap book, wipe and clean book, sticker book</i>
3	18	Buku cerita, buku mengenal huruf dan angka, buku mengenai agama islam, buku mengenal warna dan binatang	<i>Paper book</i>
4	13	Buku cerita, buku mengenai agama islam, buku pembentukan karakter	<i>Paper book</i>
5	2	Buku cerita, Buku mengenai agama islam, Buku mengenai agama Kristen	<i>Paper book, board book, augmented reality book</i>





## V SIMPULAN

### 5.1 SIMPULAN

Penggerombolan data toko buku anak pada suatu *marketplace* menggunakan penggerombolan biner pautan lengkap dengan ukuran kesamaan Jaccard, Sokal-Sneath2, dan Dice menghasilkan rasio keragaman 0,012 dan lima gerombol optimal. Gerombol kedua memiliki anggota gerombol toko buku anak terbanyak dibandingkan gerombol lainnya dengan 103 anggota toko buku anak. Gerombol kelima memiliki anggota gerombol toko buku anak paling sedikit dibandingkan gerombol lainnya dengan 2 anggota toko buku anak.

Gerombol kedua dari penggerombolan biner pautan lengkap dengan ukuran kesamaan Jaccard, Dice, Sokal-Sneath2 menjadi gerombol terbaik, dengan karakteristik toko buku anak yang paling lengkap, menyediakan semua jenis buku anak dan semua tema buku anak, kecuali buku anak dengan tema agama katolik. Gerombol empat menjadi gerombol dengan anggota toko buku anak yang berkarakteristik spesifik menyediakan buku cerita, buku mengenai agama islam, buku pembentukan karakter dan *paper book*. Gerombol lima menjadi gerombol dengan anggota paling sedikit yaitu hanya 2 toko buku anak yang berkarakteristik spesifik menyediakan buku cerita, buku mengenai agama islam, buku mengenai agama kristen, *paper book*, *board book*, dan *augmented reality book*.

@Hak Cipta milik IPB University



## DAFTAR PUSTAKA

- [APJII] Asosiasi Penyelenggara Jasa Internet Indonesia. 2019. *Survei Penetrasi Pengguna Internet Indonesia*. Jakarta: APJII.
- Anderberg MR. 1973. *Cluster Analysis for Application*. New York (NY): Academic Press.
- Choi SS, Tappert CC. 2010. A survey of binary similarity and distance measures. *Journal of Systemics, Cybernetics and Informatics*. 8: 43-48
- Dice LR. 1945. Measures of the amount of ecologic association between species. *J Ecology*. 26(3):297-302.
- Everitt BS, Landau S, Leese M, Stahl D. 2011. *Cluster Analysis*. Ed ke-5. London: John Wiley & Son, Inc.
- Finch H. 2005. Comparison of distance measures in cluster analysis with dichotomous data. *Journal of Data Science*. 3(1): 85-100.
- Gower JC. 1985. Measures of similarity, dissimilarity, and distance. *Encyclopedia of statistical sciences*. 397-405.
- Indrajit RE. 2002. *Konsep & Aplikasi E-Business*. Yogyakarta : Andi.
- Johnson R, Wichern D. 2002. *Multivariate Statistical Analysis*. New Jersey (NJ): Prentice Hall.
- Kariyam. 2010. Kesamaan data biner berdasarkan kategori nilai entropy dan pola struktur. *Jurnal Ilmu Dasar*. 11(2): 177-182.
- Komang R. 2017. Simple query suggestion untuk pencarian artikel menggunakan jaccard similarity. *Jurnal Ilmiah Rekayasa dan Manajemen Sistem Informasi*. 3(1): 30-34.
- Light RJ, Magolin BH. 1971. An analysis of variance for categorical data. *Journal of American Statistical Association*. 335(66): 534-544.
- Manly, Alberto. 2016. *Multivariate statistical methods: A Primer*, 4<sup>th</sup> ed. Chapman dan Hall. New York (NY).
- Mattjik AA, Sumertajaya IM. 2011. *Sidik Peubah Ganda dengan Menggunakan SAS*. Bogor : IPB Press.
- Podani J. 2000. *Introduction to the Exploration of Multivariate Biological Data*. Netherland: Backhuys Publishers.
- Rencher AC, Christensen WF. 2012. *Methods of Multivariate Analysis*. 3<sup>rd</sup> Ed. New York (NY): John Wiley and Sons.
- Sadgotra WY, Saputra EH. 2013. Perancangan *online marketplace* untuk Usaha Kecil Menengah (UKM) di Kabupaten Purworejo. *Jurnal Ilmiah DASI*. 14(4):54-58.
- Siswadi, Suharjo B. 1999. *Analisis Eksplorasi Data Peubah Ganda*. Bogor: Jurusan Matematika FMIPA IPB.
- Sokal RR, Michener CD. 1958. A statistical method for evaluating systematic relationships. *University of Kansas Science Bulletin*. 38(22):1409-1438.
- Timm NH. 2002. *Applied Multivariate Analysis*. New York (NY): Springer.

## RIWAYAT HIDUP

Penulis dilahirkan di Jakarta pada tanggal 25 Maret 1998 dari pasangan Bapak Ade Mahmud dan Ulfah Yahya. Penulis merupakan anak kedua dari dua bersaudara. Penulis menyelesaikan pendidikan dasar di SDIT Raflesia tahun 2009, kemudian melanjutkan Pendidikan Sekolah Menengah Pertama di SMPN 07 Depok dan lulus tahun 2012. Setelah itu penulis melanjutkan studi ke sekolah menengah atas di SMAN 2 Depok dan menyelesaikannya pada tahun 2015. Tahun 2015 penulis diterima sebagai mahasiswa di Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Pertanian Bogor setelah lulus seleksi melalui jalur SNMPTN.

Selama masa perkuliahan, penulis aktif dalam Himpunan keprofesian (Himpro), Himpunan supporter fakultas (Gegana), Himpunan mahasiswa daerah Jakarta (Jco), dan kepanitiaan. Penulis aktif pada Himpunan profesi Gamma Sigma Beta sebagai anggota pada tahun 2017 hingga tahun 2018. Penulis aktif pada himpunan supporter fakultas yang tergabung dalam Gegana 1.8 pada tahun 2016 hingga 2019. Penulis aktif pada himpunan mahasiswa daerah Jakarta pada tahun 2015 hingga 2019. Selain itu, penulis juga aktif di beberapa kepanitiaan seperti Statistika Ria ke-12 tahun 2017 dan Spirit tahun 2018. Pada bulan Juli-Agustus 2018 penulis melaksanakan Praktik Lapang di Badan Meteorologi, Klimatologi dan Geofisika (BMKG).

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
  - b. Pengutipan tidak merugikan kepentingan yang wajar IPBUniversity.
2. Dilarang mengumumkannya dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPBUniversity.



### *@Hak cipta milik IPBUniversity*

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
  - b. Pengutipan tidak merugikan kepentingan yang wajar IPBUniversity.
2. Dilarang mengummumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPBUniversity.



### @Hak cipta milik IPBUniversity

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
  - b. Pengutipan tidak merugikan kepentingan yang wajar IPBUniversity.
2. Dilarang mengummumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPBUniversity.