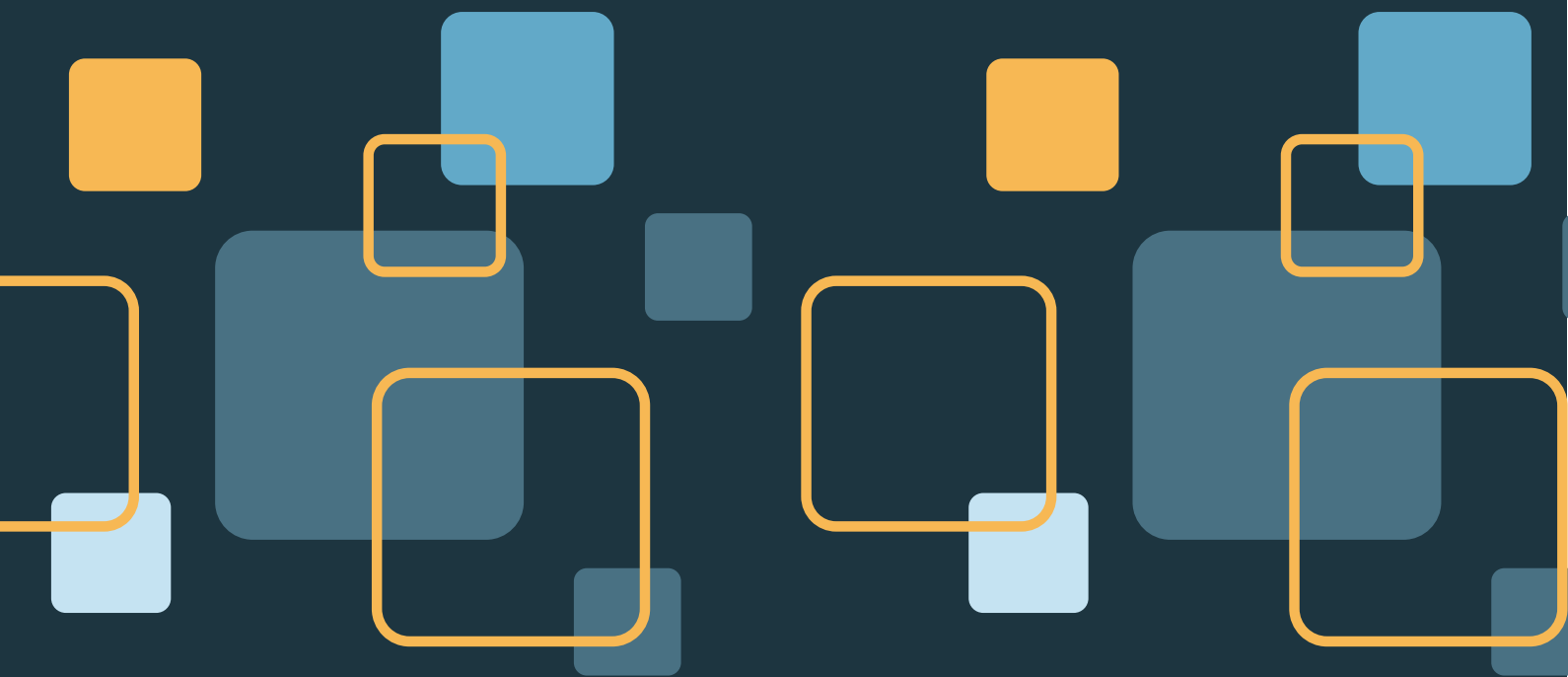


# PANDUAN PENGUNAAN

APLIKASI IMPUTASI DATA KATEGORIK  
PKL D-IV 63 POLITEKNIK STATISTIKA STIS



# DAFTAR ISI

## 1 Daftar Isi

### Pendahuluan

- A. PKL D-IV 63 Politeknik Statistika STIS
- B. Data Hilang
- C. Mekanisme Data Hilang
- D. *Little's Test of Missingness*
- E. KNN & *Hot Deck*
- F. Aplikasi Imputasi Data Kategorik

## 8 Penerapan Aplikasi

- A. Persiapan & Unggah Data
- B. Periksa Data
- C. Imputasi

## 11 Penutup

# PENDAHULUAN

- **PKL D-IV 63 POLITEKNIK  
STATISTIKA STIS**

Praktik Kerja Lapangan (PKL) adalah kegiatan akademik di Politeknik Statistika STIS yang melibatkan kolaborasi antara dosen sebagai pembimbing dan mahasiswa. Kegiatan ini bertujuan untuk menerapkan teori perkuliahan dan meningkatkan keterampilan mahasiswa, serta mendukung tugas dan tanggung jawab di Badan Pusat Statistik (BPS).

Pada PKL D-IV angkatan 63, mahasiswa melakukan Survei Dampak Paparan TIK dan Pola Asuh Orang Tua terhadap Perilaku Generasi Z di Provinsi Bali, dengan dua modul: orang tua dan Generasi Z. Modul budaya pada survei orang tua memasukkan pilihan jawaban “tidak ingat”, yang tidak bisa dianalisis secara mendalam.

# PENDAHULUAN

- **DATA HILANG (MISSING VALUE)**

Data hilang merupakan nilai yang tidak tersimpan pada suatu variabel atau observasi, dan sering terjadi dalam penelitian maupun pengumpulan data besar. Penyebabnya beragam, mulai dari desain survei yang membuat pertanyaan tidak berlaku (*non-applicable*), responden memilih tidak menjawab karena alasan privasi, hingga faktor teknis seperti kerusakan alat, kesalahan pencatatan, atau *input* manual yang kurang teliti.

Munculnya pilihan jawaban seperti “tidak tahu” atau “tidak ingat” dalam kuesioner dan diperlakukan sebagai data hilang karena tidak bisa dianalisis lebih lanjut. Opsi ini terkadang diperlukan untuk menjaga kualitas data meskipun penggunaan respons tersebut masih menjadi pembahasan. Para peneliti seperti Manisera & Zuccolotto (2014) mendukung opsi ini karena memberikan responden ruang untuk tidak memberi jawaban yang tak pasti. Karena tidak memiliki nilai, respons ini biasanya diperlakukan sebagai *missing value*.

# PENDAHULUAN

- **MEKANISME DATA HILANG**

Memahami mekanisme data hilang adalah langkah pertama penanganan, karena mekanisme menunjukkan hubungan antara nilai yang ada dan yang hilang serta menentukan metode analisis yang tepat.

1) MCAR (Missing Completely at Random)

Nilai hilang terjadi secara sepenuhnya acak, tanpa bergantung pada data lainnya. Contoh sederhana: nilai hilang karena alat pengukur tiba-tiba mati tanpa sebab yang berkaitan dengan data.

2) MAR (Missing at Random)

Nilai hilang bergantung pada data yang ada, tetapi bukan nilai itu sendiri. Misalnya, jika responden perempuan cenderung tidak mengisi pertanyaan berat badan, maka kehilangan data berkaitan dengan variabel gender yang sudah terekam.

# PENDAHULUAN

- **LITTLE'S TEST OF MISSINGNESS**

Jika data hilang diduga bersifat MCAR, maka dapat digunakan *Little's Test of Missingness* untuk mengujinya. Uji ini, diperkenalkan oleh **Little** (1988), menggunakan pendekatan *chi-square* ( $\chi^2$ ) untuk membandingkan rata-rata variabel yang diamati pada berbagai pola data hilang dengan rata-rata populasi yang diestimasi melalui algoritma *Expectation-Maximization* (EM).

Jika hasil uji menunjukkan nilai *p-value*  $> 0.05$ , maka data dianggap memenuhi asumsi MCAR. Sebaliknya, jika *p-value*  $\leq 0.05$ , hipotesis MCAR ditolak, yang berarti data hilang mengikuti mekanisme lain.

# PENDAHULUAN

- **K-NEAREST NEIGHBOR (KNN)**

Salah satu metode yang baik menangani MCAR adalah KNN *Imputation*, metode populer mengisi nilai yang hilang baik numerik maupun kategorik dengan memanfaatkan kemiripan antar pengamatan. Untuk setiap data yang kosong, KNN mencari ‘k’ tetangga terdekat (berdasarkan jarak seperti Euclidean atau Hamming), lalu mengganti nilai hilang dengan rata-rata/modus dari tetangga tersebut.

- **HOT-DECK**

Salah satu metode yang baik menangani MAR adalah *Hot-Deck Imputation*, menggantikan nilai yang hilang dengan data dari observasi lain (donor) yang serupa dalam *dataset*. Proses ini dimulai dengan membentuk kelompok donor—pengamatan lengkap yang relevan—kemudian mencocokkan setiap data penerima yang hilang (recipient) berdasarkan kemiripan karakteristik, dan akhirnya menggantikan nilai yang hilang menggunakan nilai dari donor yang paling cocok.

# PENDAHULUAN

## • APLIKASI IMPUTASI DATA KATEGORIK

Aplikasi dirancang membantu pengguna menangani data dengan nilai hilang khususnya pada PKL D-IV 63 Politeknik Statistika STIS modul orang tua blok kebudayaan. Pengguna dengan mudah mengunggah data, memeriksa data hilang, serta menjalankan proses imputasi secara otomatis menggunakan metode yang sesuai.

Salah satu kelebihan aplikasi ini adalah integrasi uji *Little's* MCAR yang secara otomatis mengevaluasi apakah data yang hilang bersifat MCAR. Berdasarkan hasil uji ini, aplikasi menentukan metode imputasi yang paling sesuai. Menggunakan KNN jika data dianggap MCAR, atau Hot-Deck jika data tidak MCAR secara otomatis. Pengguna diberikan visualisasi melalui grafik batang dan tabel, serta dapat mengunduh hasil imputasi dalam format Excel. Nilai-nilai terimputasi ditandai huruf tebal.



# PENERAPAN APLIKASI

## PERSIAPAN

1. Unduh data *dummy* PKL D-IV 63 Politeknik Statistika STIS [ [dummy](#) ]
2. Akses aplikasi melalui tautan berikut  
<https://sereia.shinyapps.io/imputate/>

## UNGGAH DATA

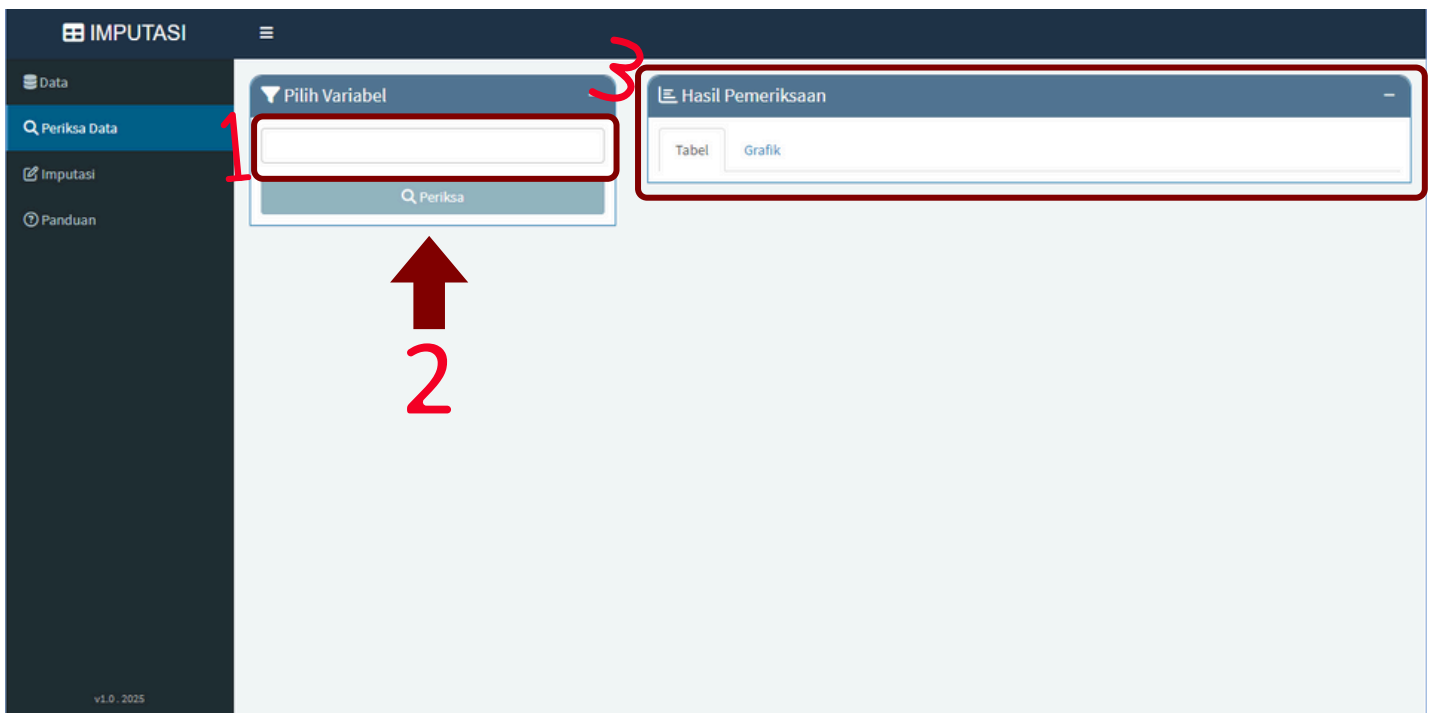
The screenshot shows the IMPUTASI application interface. On the left, the 'Unggah Data' (Upload Data) panel is visible, showing a file named 'Data - Sheet1.csv' selected. A red box highlights the 'Muat Data' (Load Data) button, with a red number '2' next to it. On the right, the 'Pratinjau Data' (Data Preview) panel is visible, showing a table of data. A red box highlights the first row of data, with a red number '3' next to it. The table has columns labeled m2b401 through m2b408 and rows numbered 1 to 10. The data values are as follows:

	m2b401	m2b402	m2b403	m2b404	m2b405	m2b406	m2b407	m2b408
1	1	3	3	2	5	6	1	
2	3	1	1	3	3	2	1	
3	2	2	3	1	6	5	5	
4	2	3	2	1	4	6	4	
5	1	2	3	6	6	6	5	
6	3	3	1	5	5	2	2	
7	3	1	1	3	5	1	3	
8	2	1	2	6	2	4	2	
9	2	2	2	4	5	2	1	
10	3	1	2	4	1	1	1	

1. Klik 'Pilih File' lalu pilih dokumen yang ingin diunggah pada *file manager*.
2. Klik 'Muat Data' dan tunggu hingga data muncul pada Pratinjau Data
3. Di bawah *header* merupakan *filter* yang dapat membantu menyaring data.

# PENERAPAN APLIKASI

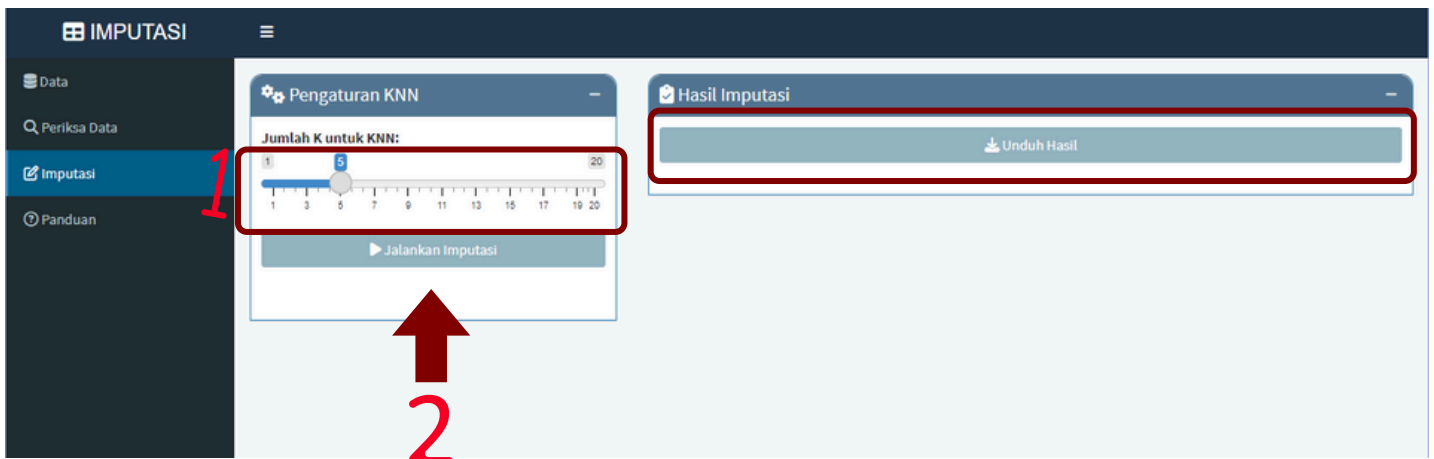
## PERIKSA DATA



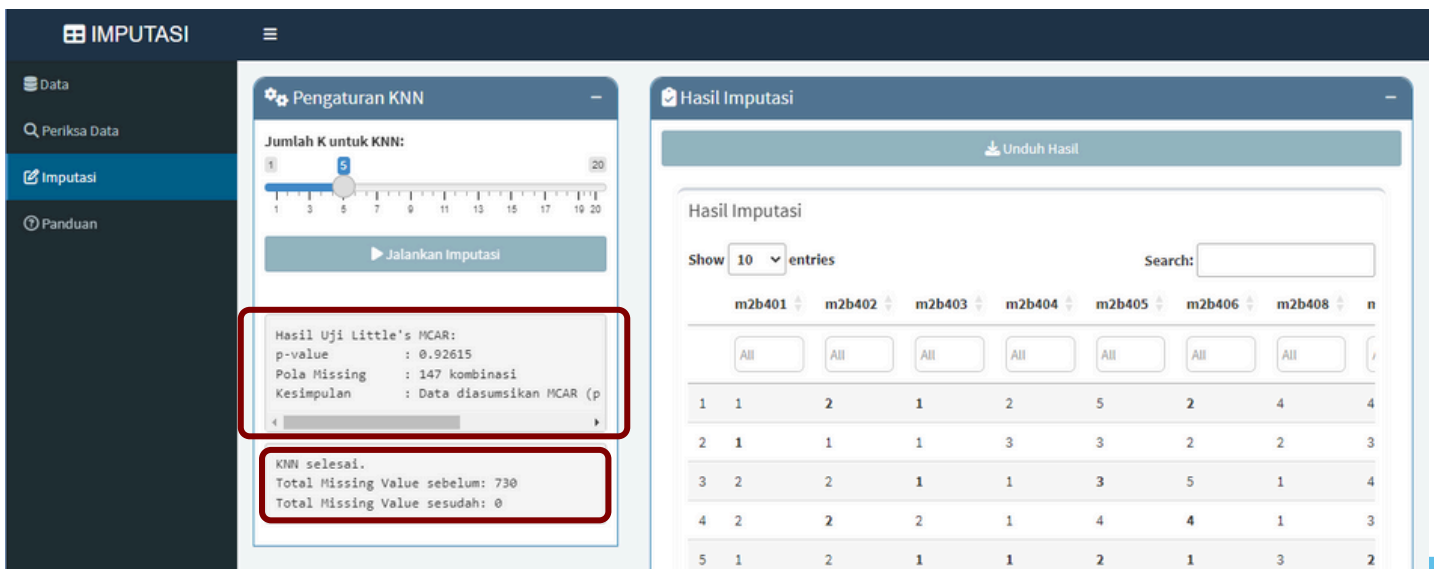
1. Klik boks pada Pilih Variabel lalu pilih variabel mana yang akan diperiksa mengandung variabel 'tidak ingat' dan akan dilakukan imputasi.
2. Klik 'Periksa' dan tunggu hingga pemeriksaan data selesai.
3. Data akan muncul pada Hasil Pemeriksaan, klik tab grafik jika ingin melihat hasil pemeriksaan dalam bentuk grafik begitu juga tabel.

# PENERAPAN APLIKASI

## IMPUTASI



1. Tentukan 'k' yang akan digunakan jika data memenuhi asumsi MCAR.
2. Klik 'Jalankan Imputasi' dan tunggu hingga proses selesai.
3. Klik 'Unduh Hasil' untuk mengunduh hasil imputasi dalam bentuk excel.
4. Berikut merupakan contoh implementasi imputasi



# PENUTUP

Panduan ini diharapkan telah memandu secara jelas dan sistematis, mulai dari persiapan, pengoperasian fitur-fitur, hingga pengunduhan hasil imputasi. Apabila terdapat pertanyaan atau kendala selama penggunaan, silakan menuju ke bagian Panduan pada aplikasi atau hubungi pengembang melalui kontak berikut🚀



+6281335742721



222111864@stis.ac.id

**TERIMA  
KASIH**

