

Penerapan Fungsi Mid Dan Find pada Pembersihan Data Alamat

Aristejo¹, Sigit Wijanarko², Subhiyanto³ Aji Santoso⁴

Abstract—Interacting with data is a common occurrence. However, it should be realized that the data received often requires refinement to make it easier to read and search. The data refinement process is called as data cleansing. For instance, inconsistency of address writing often occurs, such as street names and city names mixed in the same field. Hence, it is difficult to read and be processed further. Excel provides the MID and FIND functions. If these two functions are combined, they will help us separate the city name from the street name and also placing the city name in its own column.

Intisari—Interaksi dengan data adalah hal yang lumrah terjadi. Namun, kita juga harus menyadari bahwa data yang diterima seringkali masih memerlukan perapihan agar lebih mudah dibaca dan dicari. Proses perapihan data ini sering disebut sebagai data cleansing. Pada data alamat, seringkali terjadi penulisan yang tidak seragam, seperti nama jalan dan nama kota yang bercampur dalam satu field yang sama, sehingga akan menyulitkan untuk dibaca dan diolah lebih lanjut. Excel menyediakan fungsi MID dan fungsi FIND. Jika kedua fungsi ini digabungkan, maka akan membantu kita dalam memisahkan nama kota dari nama jalan, sekaligus menempatkan nama kota ke dalam satu kolom tersendiri.

Kata Kunci— data cleansing; excel; mid; find.

I. PENDAHULUAN

Seringkali kita mendapatkan data alamat yang pengisiannya tidak konsisten, terutama saat data alamat hanya diisi dalam satu kolom. Sebagai contoh, berikut adalah hasil pengisian data alamat yang berasal dari satu kolom:

Tabel 1. Contoh Data Alamat

Alamat
Jl Dukuh 1 no 34 Jakarta Selatan
Jl Radio Raya Kelurahan Petukangan Jakarta Selatan
Perumahan Bintaro Regency Pondok Aren Tangerang Selatan
Jl DR Semeru I no 41 Grogol Jakarta Barat
Jl DR Muwardi Raya no 12 Jakarta Barat
Jl Tanjung Duren Selatan no 113 Jakarta Barat

Ketika kita perlu memisahkan nama kota, permasalahan yang dihadapi adalah inkonsistensi jumlah kata pada setiap baris. Inkonsistensi ini mengharuskan kita menggabungkan beberapa fungsi dalam Excel ketika perlu memisahkan nama kota ke dalam kolom tersendiri [1]. Jika data tersebut sedikit, tentu saja

mudah dilakukan. Tetapi bagaimana jika data tersebut melebihi 500 baris? Kita perlu menggunakan fungsi-fungsi di Excel yang dapat membantu kita bekerja lebih efisien namun tetap efektif. Ada beragam fungsi yang dapat digunakan di Excel untuk menyelesaikan masalah tersebut. Dalam tulisan kali ini, kita akan menggunakan fungsi MID dan FIND, dengan melakukan kombinasi dari keduanya [2]. Sebelum melangkah lebih jauh, mari kita pahami terlebih dahulu apa itu fungsi MID dan fungsi FIND, beserta sintaksis dan penggunaannya.

II. KAJIAN LITERATUR

A. Data Cleansing

Data cleansing adalah proses identifikasi, koreksi, dan penghapusan atau perubahan data yang tidak akurat, tidak lengkap, tidak relevan, atau tidak terkonsistensi dalam sebuah dataset [3]. Tujuan dari data cleansing adalah untuk meningkatkan kualitas data dengan memperbaiki kesalahan, menormalisasi format, dan menghilangkan duplikat sehingga data menjadi lebih bermanfaat dan dapat diandalkan untuk analisis atau penggunaan lainnya. Proses ini melibatkan berbagai teknik, termasuk pemisahan data yang tidak terstruktur, normalisasi format, validasi data, penghapusan duplikat, dan penyatuan data yang berasal dari berbagai sumber. Data cleansing merupakan langkah penting dalam siklus hidup data dan menjadi prasyarat untuk analisis data yang akurat dan efektif.

B. Fungsi Mid dan Find dalam Excel

Fungsi MID dalam Excel digunakan untuk mengambil sebagian teks dari sebuah string berdasarkan posisi awal dan jumlah karakter yang diinginkan. Fungsi FIND digunakan untuk mencari posisi dari teks tertentu dalam sebuah string [4]. Kombinasi kedua fungsi ini bisa digunakan untuk memisahkan nama kota dari alamat dalam sebuah kolom Excel [5].

1) Fungsi MID mengembalikan jumlah karakter tertentu dari sebuah string teks, dimulai dari posisi yang Anda tentukan, berdasarkan jumlah karakter yang Anda tentukan.

Syntax fungsi MID adalah MID(text, start_num, num_chars). Gunakan tanda sama dengan (=) ketika menuliskan syntanx MID di tempat rumus.

- Teks adalah string teks yang memuat karakter yang ingin Anda ekstrak. Sifatnya mandatory [6].
- Start_num adalah posisi karakter pertama yang ingin Anda ekstrak dalam teks [7]. Karakter pertama dalam teks memiliki start_num 1, dan seterusnya. Sifatnya mandatory.
 - Jika start_num lebih besar dari panjang teks, maka fungsi MID mengembalikan "" (teks kosong).
 - Jika start_num lebih kecil dari panjang teks, tetapi start_num ditambah num_chars melebihi panjang

^{1,2,3,4} STMIK Antar Bangsa, Kawasan Bisnis CBD Ciledug, Jl. HOS Cokroaminoto No.29-35, Karang Tengah, Kec. Ciledug, Kota Tangerang, Banten 15157 (tlp: 0811-9391-441; e-mail: aristejo14@gmail.com, sgtwijanarko23@gmail.com, subhiyanto.bian@gmail.com, ajisantoso074@gmail.com)

teks, maka fungsi MID mengembalikan karakter hingga akhir teks.

- Jika start_num lebih kecil dari 1, maka fungsi MID menampilkan #VALUE! yang merupakan nilai kesalahan.
 - Num_chars menentukan jumlah karakter yang ingin ditampilkan oleh fungsi MID. Sifatnya mandatory. Jika num_chars negatif, maka fungsi MID menampilkan #VALUE! yang merupakan nilai kesalahan.
- 2) Fungsi FIND selalu menghitung setiap karakter, baik byte tunggal atau byte ganda sebagai 1, tak peduli apa pun pengaturan bahasa default-nya. Syntax fungsi FIND adalah FIND(find_text, within_text, [start_num]). Gunakan tanda sama dengan (=) ketika menuliskan syntanx FIND di tempat rumus. Keterangan tentang komponen syntanx FIND adalah sebagai berikut:
- Find_text. Sifatnya mandatory. Teks yang ingin ditemukan.
 - Within_text. Sifatnya mandatory. Teks yang berisi teks yang ingin ditemukan.
 - Start_num. Sifatnya optional. Menentukan karakter yang digunakan untuk memulai pencarian. Karakter pertama dalam within_text adalah angka karakter 1. Jika kita menghapus bagian start_num, maka excel akan mengasumsikan menjadi 1.

Berikut adalah aturan dalam menggunakan fungsi FIND:

- Fungsi FIND bersifat case sensitive (peka terhadap huruf kapital dan huruf kecil). Artinya ketika kita akan mencari sebuah kata maka kita harus perhatikan penggunaan huruf kapital dan huruf kecil pada kata yang kita cari dan kata yang menjadi acuan.
- Fungsi FIND tidak boleh berisi karakter wildcard. Karakter wildcard adalah karakter khusus yang digunakan sebagai pengganti karakter yang tidak diketahui dalam sebuah teks. Penggunaan karakter wildcard sangat berguna untuk menemukan beberapa item dengan data yang sama namun tidak identik.
 - Contoh wildcard adalah ketika kita mencocokkan setiap karakter numerik tunggal diantara angka 1 dan 3 dalam sebuah bilangan maka kita menggunakan tanda #. Pada kasus ini yang dimaksud karakter wildcard adalah 1#3. Begitupun bentuk-bentuk karakter wildcard lainnya, semua tidak diijinkan dalam penggunaan fungsi FIND:
- Jika find_text tidak muncul di within_text, maka FIND akan menampilkan #VALUE! yang merupakan nilai kesalahan.
- Jika start_num kurang dari nol, maka FIND akan menampilkan #VALUE! yang merupakan nilai kesalahan. Jika start_num lebih besar dari panjang

within_text, maka FIND akan menampilkan #VALUE! yang merupakan nilai kesalahan.

III. METODE PENELITIAN

Dalam penelitian ini, digunakan pendekatan metodologi kualitatif. Pendekatan ini bertujuan untuk memahami fenomena yang dialami oleh objek penelitian. Dalam metodologi kualitatif, aspek matematika dan statistik tidak digunakan. Dalam konteks penelitian ini, fungsi MID dan FIND dalam Excel diterapkan untuk pembersihan data alamat disuatu instansi. Tahap awal melibatkan implementasi fungsi MID dan FIND. Selanjutnya, langkah berikutnya adalah data *cleansing*.

IV. HASIL DAN PEMBAHASAN

Hasil dari penelitian ini adalah sebagai berikut.

A. Implementasi Fungsi MID dan FIND

Ketika kita akan mengambil sebuah teks nilai dalam serangkaian teks maka kita dapat menggabungkan rumus MID dan FIND ini. Contoh kasus ketika kita ingin membersihkan data alamat dengan memisahkan nama kota dari teks dan menempatkannya pada kolom terpisah. Data awal seperti terlihat pada gambar 2 berikut:

Tabel 2. Data Alamat Awal

Alamat
Jl Dukuh 1 no 34 Jakarta Selatan
Jl Radio Raya Kelurahan Petukangan Jakarta Selatan
Perumahan Bintaro Regency Pondok Aren Tangerang Selatan
Jl DR Semeru I no 41 Grogol Jakarta Barat
Jl DR Muwardi Raya no 12 Jakarta Barat
Jl Tanjung Duren Selatan no 113 Jakarta Barat

Terlihat disini bahwa teks nama kota masih tercampur dengan teks alamat. Untuk merapihkan data alamat tersebut maka kita akan memisahkan teks nama kota dengan teks alamat seperti terlihat pada tabel 3 berikut:

Tabel 3. Data Alamat Bersih

Alamat	Kota
Jl Dukuh 1 no 34 Jakarta Selatan	Jakarta Selatan
Jl Radio Raya Kelurahan Petukangan Jakarta Selatan	Jakarta Selatan
Perumahan Bintaro Regency Pondok Aren Tangerang Selatan	Tangerang Selatan

Untuk melakukan melakukan hal tersebut maka kita akan menggunakan gabungan fungsi MID dan FIND. Caranya adalah sebagai berikut:

- 1) Tempatkan kursor pada sel sebelah sel alamat pertama. Di contoh ini kita tempatkan kursor di sel B3.

	A	B
1		
2	Alamat	Kota
3	Jl Dukuh 1 no 34 Jakarta Selatan	
4	Jl Radio Raya Kelurahan Petukangan Jakarta Selatan	
5	Perumahan Bintaro Regency Pondok Aren Tangerang Selatan	
6	Jl DR Semeru I no 41 Grogol Jakarta Barat	
7	Jl DR Muwardi Raya no 12 Jakarta Barat	
8	Jl Tanjung Duren Selatan no 113 Jakarta Barat	

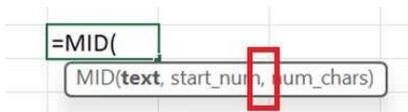
Gbr 4. Posisi Kursor

2) Selanjutnya kita akan mengekstrak nama kota dari kolom Alamat dan kita tampilkan di kolom Kota. Disini kita akan menggunakan fungsi MID dan FIND. Syntax MID adalah MID(text, start_num, num_chars). Karena kita akan menggabungkan dengan fungsi FIND maka bagian start_num pada fungsi MID akan kita ganti dengan fungsi FIND. Sehingga untuk kolom kota baris pertama, penulisan fungsinya menjadi =MID(A3,FIND("Jakarta Selatan",A3,1),20). Maka hasilnya akan menjadi seperti gambar 5 berikut:

	A	B
1		
2	Alamat	Kota
3	Jl Dukuh 1 no 34 Jakarta Selatan	Jakarta Selatan
4	Jl Radio Raya Kelurahan Petukangan Jakarta Selatan	
5	Perumahan Bintaro Regency Pondok Aren Tangerang Selatan	
6	Jl DR Semeru I no 41 Grogol Jakarta Barat	
7	Jl DR Muwardi Raya no 12 Jakarta Barat	
8	Jl Tanjung Duren Selatan no 113 Jakarta Barat	

Gbr 5. Fungsi MID dan FIND

Disini penulis menggunakan koma sebagai pemisah antar komponen syntax. Perlu diketahui bahwa pemisah dapat menggunakan koma atau titik koma tergantung Excel yang digunakan. Jadi pemisahannya disesuaikan saja dengan versi Excel yang anda gunakan. Cara kita mengetahui apakah harus menggunakan koma atau titik koma terlihat pada saat kita mulai menulis fungsi. Misal menulis fungsi MID. Ketika mulai mengetikkan =MID maka akan muncul petunjuk penulisan fungsinya sebagai berikut:



Gbr 6. Pemisah Antar Syntax

Pada petunjuk fungsi, terlihat bahwa pemisah antar komponen syntax menggunakan koma. Jika harus menggunakan titik koma maka akan terlihat pemisahannya menggunakan titik koma.

3) Setelah selesai implementasi fungsi MID dan FIND pada baris pertama dengan hasil seperti pada gambar 5 maka langkah selanjutnya adalah meng-copy fungsi tersebut ke bawah hingga baris terakhir. Ada dua metode untuk mengcopy. Berikut kedua metode tersebut:

- Metode Copy dan Paste.
 - Letakkan kursor pada baris pertama di kolom Kota. Kemudian pada keyboard tekan Ctrl + C. Jika menggunakan mouse maka klik kanan kemudian klik copy.
 - Geser kursor satu baris ke bawah. Masih di kolom Kota.
 - Pada keyboard tekan Ctrl + panah bawah. Blok semua hingga baris terakhir.
 - Tekan Ctrl +V. Jika menggunakan mouse maka klik kanan kemudian klik paste.
- Metode Klik.
 - Letakkan kursor pada baris pertama di kolom Kota.
 - Posisikan mouse ke pojok kanan bawah dari sel yang akan di-copy.
 - Double klik tanda plus tersebut. Excel akan meng-copy fungsi tersebut secara otomatis hingga ke baris terbawah dari data yang ada 4) Setelah dicopy maka hasilnya adalah sebagai berikut:

Tabel 7. Meng-copy

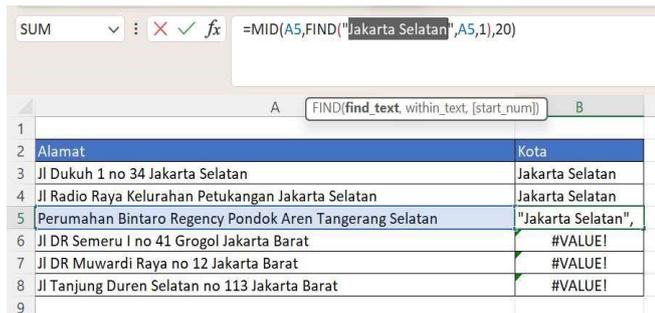
Jl Dukuh 1 no 34 Jakarta Selatan	Jakarta Selatan
Jl Radio Raya Kelurahan Petukangan Jakarta Selatan	Jakarta Selatan
Perumahan Bintaro Regency Pondok Aren Tangerang Selatan	#VALUE!
Jl DR Semeru I no 41 Grogol Jakarta Barat	#VALUE!
Jl DR Muwardi Raya no 12 Jakarta Barat	#VALUE!
Jl Tanjung Duren Selatan no 113 Jakarta Barat	#VALUE!

Karena muncul nilai value pada sel, maka kita harus memperbaiki nilai parameter yang kita masukkan ke dalam fungsi tersebut. Cara memperbaiki nilai parameter, langkah pertama yang dilakukan adalah menempatkan kursor pada sel yang berisi nilai value. Setelah itu perhatikan baris yang berisi syntax fungsi seperti pada gambar berikut:

Gbr 8. Baris Fungsi

Pada baris fungsi terlihat jelas bahwa nilai yang harus dicari (find_text) oleh fungsi FIND adalah Jakarta Selatan sedangkan pada sel A6, A7, dan A8 tidak terdapat nilai Jakarta Selatan, sehingga Excel menampilkan nilai value yang menyatakan terjadi kesalahan nilai parameter. 5) Pada

sel B5, blok nilai Jakarta Selatan kemudian ganti dengan Tangerang Selatan.



Gbr 9. Ganti Nilai Parameter

Setelah diganti kemudian tekan Enter, dan copy fungsi pada sel B5 hingga ke bawah. Maka hasilnya adalah sebagai berikut:

Tabel 10. Nilai Value

Jl Dukuh 1 no 34 Jakarta Selatan	Jakarta Selatan
Jl Radio Raya Kelurahan Petukangan Jakarta Selatan	Jakarta Selatan
Perumahan Bintaro Regency Pondok Aren Tangerang Selatan	Tangerang Selatan
Jl DR Semeru I no 41 Grogol Jakarta Barat	#VALUE!
Jl DR Muwardi Raya no 12 Jakarta Barat	#VALUE!
Jl Tanjung Duren Selatan no 113 Jakarta Barat	#VALUE!

B. Data Cleansing

Data cleansing dapat diterapkan secara bersama-sama dengan menerapkan fungsi MID dan FIND dalam Excel kepada data. Ini berdampak pada peningkatan keterbacaan data.

Masih terlihat adanya nilai Value. Konsep terjadinya hal ini sama seperti penjelasan diatas. Sebagai solusinya maka kita perlu mengulangi langkah tadi dengan mengganti nilai parameter find_text pada fungsi FIND dengan Jakarta Barat, sehingga hasilnya adalah seperti berikut:

Tabel 11. Data Bersih

Jl Dukuh 1 no 34 Jakarta Selatan	Jakarta Selatan
Jl Radio Raya Kelurahan Petukangan Jakarta Selatan	Jakarta Selatan
Perumahan Bintaro Regency Pondok Aren Tangerang Selatan	Tangerang Selatan
Jl DR Semeru I no 41 Grogol Jakarta Barat	Jakarta Barat
Jl DR Muwardi Raya no 12 Jakarta Barat	Jakarta Barat
Jl Tanjung Duren Selatan no 113 Jakarta Barat	Jakarta Barat

V. KESIMPULAN

Berdasarkan implementasi dan perancangan data cleansing yang menerapkan fungsi MID dan FIND dalam Excel Pengguna data diberi kemudahan dalam membaca dan mencari, juga membawa dampak positif dengan menghasilkan data yang lebih terstruktur setelah data cleansing selesai.

1. Dalam melakukan proses pembersihan data dimana kita perlu untuk mengambil nilai terakhir dari suatu teks maka kita dapat menggunakan fungsi MID dan FIND namun kita harus menggabungkan kedua fungsi tersebut.
2. Metode pengambilan nilai terakhir dari suatu teks dengan menggunakan fungsi MID dan FIND yang digabung memiliki kelemahan yaitu kita harus mengganti nilai parameter find_text pada fungsi FIND setiap kita menemukan nilai yang berbeda.

UCAPAN TERIMA KASIH

Penulis bersyukur kepada Tuhan Yang Maha Esa, Pencipta alam semesta, dan ingin mengungkapkan rasa terima kasih kami kepada semua individu dan kelompok yang telah memberikan bantuan, dukungan, dan arahan berharga selama pelaksanaan penelitian ini.

REFERENSI

- [1] Y. Puspita Dewi and A. Pusat Kegiatan Belajar Masyarakat, "PADA PKBM BINA BANGSA LARANGAN SELATAN KOTA TANGERANG," vol. 02, p. 1, 2020.
- [2] F. Marisa, "EDUCATIONAL DATA MINING (KONSEP DAN PENERAPAN)."
- [3] N. Rokhman, A. M. Ningtyas, M. F. Salim, and D. B. Santoso, "Penerapan Sistem Data Cleansing untuk Mencegah dan Menghilangkan Duplikasi Rekam Medis," *Jurnal Pengabdian kepada Masyarakat (Indonesian Journal of Community Engagement)*, vol. 6, no. 4, Dec. 2020, doi: 10.22146/jpkm.51073.
- [4] M. Mujilan, "Pengolahan data keuangan dengan excel: Komputerisasi Akuntansi." Fakultas Bisnis, 2022.
- [5] J. Lambert and C. Frye, *Microsoft Excel step by step : (Office 2021 and Microsoft 365)*.
- [6] A. Prasetyo, *Tip & Trik Menjadi Master PHP*. MediaKita, 2011.
- [7] J. Arifin, *Mengupas Kedahsyatan 340 Fungsi Terapan Microsoft Excel 2016*. Elex Media Komputindo, 2017.



Aristejo, lahir di Jakarta pada tanggal 14 Oktober 1978. Tahun 2001 Lulus dari Jurusan Teknik Sipil di Universitas Trisakti. Tahun 2004 lulus program Pasca Sarjana Ilmu Komputer dengan di Universitas Indonesia. Saat ini aktif mengajar sebagai dosen tetap pada program studi Teknik Informatika di STMIK Antar Bangsa.



Sigit Wijanarko lahir di Sintang pada tanggal 23 Juni 1980. Lulus S1 Program Studi Teknik Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional – Jakarta pada tahun 2004. Lulus Magister Ilmu Komputer Program Pascasarjana Universitas Budi Luhur dengan konsentrasi Rekayasa Komputasi Terapan pada tahun 2019. Saat ini aktif sebagai Dosen tetap di STMIK Antar Bangsa dan praktisi IT di perusahaan swasta.



Subhiyanto, lahir di Brebes pada tanggal 10 Maret 1984. Tahun 2012 Lulus Sarjana Komputer Jurusan Teknik Informatika di STMIK Nusa Mandiri. Tahun 2020 lulus program Pasca Sarjana Ilmu Komputer dengan konsentrasi Rekayasa Komputasi Terapan di Universitas Budi Luhur. Saat ini aktif mengajar sebagai dosen tetap dan sebagai Kepala Biro Teknologi Informasi di STMIK Antar Bangsa



Aji Santoso. Lahir di Tegal pada tanggal 13 Agustus 1993. Lulus dari Program Strata Satu (S1) Jurusan Pendidikan Bahasa Inggris di Universitas Pancasakti Tegal pada tahun 2018. Lulus dari Program Strata Dua (S2) Pascasarjana Pendidikan Bahasa Inggris pada tahun 2021 di Universitas Indraprasta PGRI Jakarta. Saat ini aktif sebagai Dosen Tetap di STMIK Antar Bangsa.