

# Use Detre to clean dates

by Zetra Zetra

---

## General metrics

**2,346**

characters

**414**

words

**34**

sentences

**1 min 39 sec**reading  
time**3 min 11 sec**speaking  
time

---

## Writing Issues

No issues found

---

## Plagiarism

This text hasn't been checked for plagiarism

## Unique Words

**46%**

Measures vocabulary diversity by calculating the percentage of words used only once in your document

---

unique words

## Rare Words

**21%**

Measures depth of vocabulary by identifying words that are not among the 5,000 most common English words.

---

rare words

## Word Length

**4.3**

Measures average word length

---

characters per word

## Sentence Length

**12.2**

Measures average sentence length

words per sentence

# Use Detre to clean dates

Messy (or dirty) dates can be a challenge to clean. They come in many different formats, and handling each is not

feasible when you have 100s or 1000s of rows in your data. In the image below, an excel

user wanted to transform the messy dates into the desired format. However, for an average excel user to make

the necessary transformations would be challenging at scale. We understand why the user looked for help from the community because no standard click, click, click would solve the problem.

We built Detre to help someone who faces such a problem with their data. They don't need to have above-average Excel skills or know how to program. They need to upload the data, select the correct data type and be aware of how Detre `sees` such data.

When we look at the dates shown in the image, they have common attributes.

These are:

1. White-space
2. Special characters (\*, etc.)
3. Text
4. Numbers
5. Date Text

The user will provide the guide as to what attributes are present in the given messy date. For example, `8/12/04 (scholarship)` has numbers (i.e. 8,12,04), special characters (i.e. /, (, and ) ) and text (i.e. scholarship) and the task for Detre would be to remove the special characters and text while interpreting the numbers as a date. Fortunately, the ISO 8601 standard, used universally, has symbols to represent days, months and years. The date `8/12/04` is 'd/m/y' in ISO 8601 format. For Detre to clean the messy date data, the user would provide the following guidance: `d*m*y*text*`. In the table below, we show the user's hint for each messy date. The user can change dirty dates with the same pattern at one go as well.

The platform, now, makes certain assumptions related to messy dates. First, there are no spelling mistakes for the months. Second, there are no multiple dates in the field and if there is, then use the first one. Third, all numbers are related to dates. These are quite strong assumptions given how messy dates are in the wild. We are working on fixing this.

As we have shown, Detre can indeed help to clean your messy dates with relative ease. Cleaning data with ease is in line with our mission to `half the time spent on data cleaning`. Are you ready to join?

Then send us your story of messy dates so we can improve our platform for everyone. Or start using the platform today.