Remove Duplicates node documentation

# Remove Duplicates node#

Use the Remove Duplicates node to identify and delete items that are:

• identical across all fields or a subset of fields in a single execution

• identical to or surpassed by items seen in previous executions

This is helpful in situations where you can end up with duplicate data, such as a user creating multiple accounts, or a customer submitting the same order multiple times. When working with large datasets it becomes more difficult to spot and remove these items.

By comparing against data from previous executions, the Remove Duplicates node can delete items seen in earlier executions. It can also ensure that new items have a later date or a higher value than previous values.

Major changes in 1.64.0

The n8n team overhauled this node in n8n 1.64.0. This document reflects the latest version of the node. If you're using an older version of n8n, you can find the previous version of this document here.

## Operation modes#

The remove duplication node works differently depending on the value of the operation parameter:

• Remove Items Repeated Within Current Input: Identify and remove duplicate items in the current input across all fields or a subset of fields.

• Remove Items Processed in Previous Executions: Compare items in the current input to items from previous executions and remove duplicates.

• Clear Deduplication History: Wipe the memory of items from previous executions.

### Remove Items Repeated Within Current Input#

When you set the "Operations" field to Remove Items Repeated Within Current Input, the Remove Duplicate node identifies and removes duplicate items in the current input. It can do this across all fields, or within a subset of fields.

When using the Remove Items Repeated Within Current Input operation, the following parameter is available:

• Compare: Select which fields of the input data n8n should compare to check if they're the same. The following options are available:  
All Fields: Compares all fields of the input data.  
All Fields Except: Enter which input data fields n8n should exclude from the comparison. You can provide multiple values separated by commas.  
Selected Fields: Enter which input data fields n8n should include in the comparison. You can provide multiple values separated by commas.

• All Fields: Compares all fields of the input data.

• All Fields Except: Enter which input data fields n8n should exclude from the comparison. You can provide multiple values separated by commas.

• Selected Fields: Enter which input data fields n8n should include in the comparison. You can provide multiple values separated by commas.

If you choose All Fields Except or Selected Fields as your compare type, you can add these options:

• Disable Dot Notation: Set whether to use dot notation to reference child fields in the format parent.child (turned off) or not (turn on).

parent.child

• Remove Other Fields: Set whether to remove any fields that aren't used in the comparison (turned on) or not (turned off).

### Remove Items Processed in Previous Executions#

When you set the "Operation" field to Remove Items Processed in Previous Executions, the Remove Duplicate node compares items in the current input to items from previous executions.

When using the Remove Items Processed in Previous Executions operation, the following parameters are available:

• Keep Items Where: Select how n8n decides which items to keep. The following options are available:  
  
Value Is New: n8n removes items if their value matches items from earlier executions.  
Value Is Higher than Any Previous Value: n8n removes items if the current value isn't higher than previous values.  
Value Is a Date Later than Any Previous Date: n8n removes date items if the current date isn't later than previous dates.

Keep Items Where: Select how n8n decides which items to keep. The following options are available:

• Value Is New: n8n removes items if their value matches items from earlier executions.

• Value Is Higher than Any Previous Value: n8n removes items if the current value isn't higher than previous values.

• Value Is a Date Later than Any Previous Date: n8n removes date items if the current date isn't later than previous dates.

• Value to Dedupe On: The input field or fields to compare. The option you select for the Keep Items Where parameter determines the exact format you need:  
  
When using Value Is New, this must be an input field or combination of fields with a unique ID.  
When using Value Is Higher than Any Previous Value, this must be an input field or combination of fields that has an incremental value.  
When using Value Is a Date Later than Any Previous Date, this must be an input field that has a date value in ISO format.

Value to Dedupe On: The input field or fields to compare. The option you select for the Keep Items Where parameter determines the exact format you need:

• When using Value Is New, this must be an input field or combination of fields with a unique ID.

• When using Value Is Higher than Any Previous Value, this must be an input field or combination of fields that has an incremental value.

• When using Value Is a Date Later than Any Previous Date, this must be an input field that has a date value in ISO format.

When using the Remove Items Processed in Previous Executions operation, the following option is available:

• Scope: Sets how n8n stores and uses the deduplication data for comparisons. The following options are available:  
Node: (default) Stores the data for this node independently from other Remove Duplicates instances in the workflow. When you use this scope, you can clear the duplication history for this node instance without affecting other nodes.  
Workflow: Stores the duplication data at the workflow level. This shares duplication data with any other Remove Duplicate nodes set to use "workflow" scope. n8n will still manage the duplication data for other Remove Duplicate nodes set to "node" scope independently.

• Node: (default) Stores the data for this node independently from other Remove Duplicates instances in the workflow. When you use this scope, you can clear the duplication history for this node instance without affecting other nodes.

• Workflow: Stores the duplication data at the workflow level. This shares duplication data with any other Remove Duplicate nodes set to use "workflow" scope. n8n will still manage the duplication data for other Remove Duplicate nodes set to "node" scope independently.

When you select Value Is New as your Keep Items Where choice, this option is also available:

• History Size: The number of items for n8n to store to track duplicates across executions. The value of the Scope option determines whether this history size is specific to this individual Remove Duplicate node instance or shared with other instances in the workflow. By default, n8n stores 10,000 items.

### Clear Deduplication History#

When you set the "Operation" field to Clear Deduplication History, the Remove Duplicates node manages and clears the stored items from previous executions. This operation doesn't affect any items in the current input. Instead, it manages the database of items that the "Remove Items Processed in Previous Executions" operation uses.

When using the Clear Deduplication History operation, the following parameter is available:

• Mode: How you want to manage the key / value items stored in the database. The following option is available:  
Clean Database: Deletes all duplication data stored in the database. This resets the duplication database to its original state.

• Clean Database: Deletes all duplication data stored in the database. This resets the duplication database to its original state.

When using the Clear Deduplication History operation, the following option is available:

• Scope: Sets the scope n8n uses when managing the duplication database.  
Node: (default) Manages the duplication database specific to this Remove Duplicates node instance.  
Workflow: Manages the duplication database shared by all Remove Duplicate node instances that use workflow scope.

• Node: (default) Manages the duplication database specific to this Remove Duplicates node instance.

• Workflow: Manages the duplication database shared by all Remove Duplicate node instances that use workflow scope.

## Templates and examples#

For templates using the Remove Duplicates node and examples of how to use it, refer to Templates and examples.

## Related resources#

Learn more about data structure and data flow in n8n workflows.