



Getting Started

- What is Segment?
- How Segment Works
- Getting Started Guide
- A Basic Segment Installation
- Planning a Full Installation
- A Full Segment Installation
- Sending Data to Destinations
- Testing and Debugging
- What's Next
- Use Cases

Guides

Connections

Unify

Engage

Privacy

Protocols

Segment App

API

Partners

Glossary

Config API

Help

Keep in mind that Segment's libraries all generate `messageIds` for each event payload, with the exception of the Segment HTTP API, which assigns each event a unique `messageId` when the message is ingested. You can override these default generated IDs and manually assign a `messageId` if necessary. The `messageId` field is limited to 100 characters.

Warehouse deduplication

Duplicate events that are more than 24 hours apart from one another deduplicate in the Warehouse. Segment deduplicates messages going into a Warehouse ([including Profiles Sync data](#)) based on the `messageId`, which is the `id` column in a Segment Warehouse.

Data Lake deduplication

To ensure clean data in your Data Lake, Segment removes duplicate events at the time your Data Lake ingests data. The Data Lake deduplication process dedupes the data the Data Lake syncs within the last 7 days with Segment deduping the data based on the `messageId`.

Need support?

Questions? Problems? Need more info? Contact Segment Support for assistance!

[Visit our Support page](#)

Help improve these docs!

 [Edit this page](#)

 [Request docs change](#)

Was this page helpful?

 [Yes](#)

 [No](#)

Get started with Segment

Segment is the easiest way to integrate your websites & mobile apps data to over 300 analytics and growth tools.

Your work e-mail

[Request Demo](#)

or

[Create free account](#)

© 2025 Segment.io, Inc.

[Privacy](#)

[Terms](#)

[Website Data Collection Preferences](#)

