

# Transforming Numeric Data

You may need to apply two kinds of transformations to numeric data:

- **Normalizing** - transforming numeric data to the same scale as other numeric data.
- **Bucketing** - transforming numeric (usually continuous) data to categorical data.

## Why Normalize Numeric Features?

Normalization is necessary if you have very different values within the same feature (for example, city population). Without normalization, your training could blow up with NaNs if the gradient update is too large.

You might have two different features with widely different ranges (e.g., age and income), causing the gradient descent to "bounce" and slow down convergence. Optimizers like [AdaGrad](http://www.jmlr.org/papers/volume12/duchi11a/duchi11a.pdf) (http://www.jmlr.org/papers/volume12/duchi11a/duchi11a.pdf) and [Adam](https://arxiv.org/abs/1412.6980) (https://arxiv.org/abs/1412.6980) protect against this problem by creating a separate effective learning rate per feature. But optimizers can't save you from a wide range of values within a single feature; in those cases, you must normalize.

ig: When normalizing, ensure that the same normalizations are applied at serving time to avoid skew.

irms:

[numerical data](https://developers.google.com/machine-learning/glossary/#numerical_data) (https://developers.google.com/machine-learning/glossary/#numerical\_data)

[Previous](#)

← [Introduction to Transforming Data](#)

(/machine-learning/data-prep/transform/introduction)

[Next](#)

[Normalization](#) (/machine-learning/data-prep/transform/normalization)



Except as otherwise noted, the content of this page is licensed under the [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/) (https://creativecommons.org/licenses/by/4.0/), and code samples are licensed under the [Apache 2.0 License](https://www.apache.org/licenses/LICENSE-2.0) (https://www.apache.org/licenses/LICENSE-2.0). For details, see the [Google Developers Site](#)

Policies (<https://developers.google.com/site-policies>). Java is a registered trademark of Oracle and/or its affiliates.

Last updated 2019-07-11 UTC.