



Scaling and distributions

Feature scaling

- Feature scaling is the process of adjusting values measured on different scales to a common scale.
- The goal of feature scaling is to:
 - prevent features with larger values from dominating the model
 - make the features more comparable
 - return better estimates if the models are based on distance calculations (like KNNs, k-means, and PCA)
 - speed up the convergence of gradient descent in neural networks and SVMs,



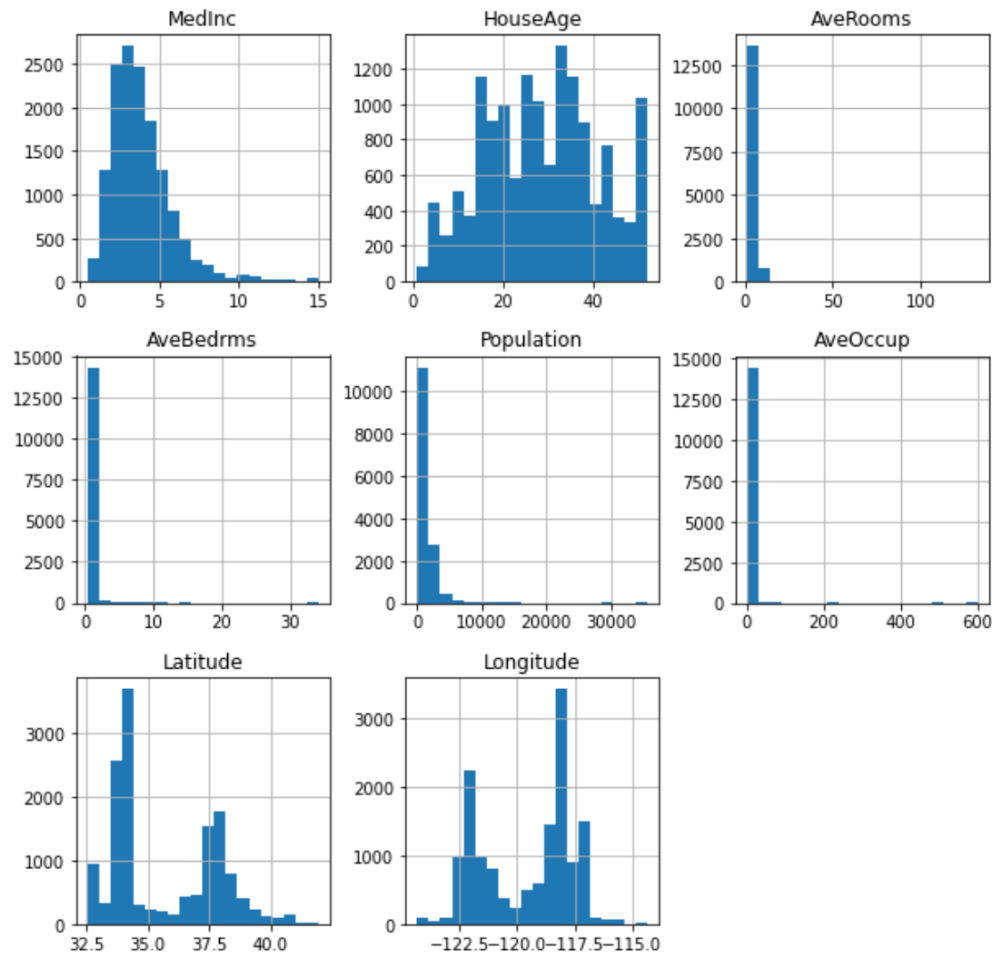
Feature scaling

Feature scaling **does not change the shape** of the distribution.

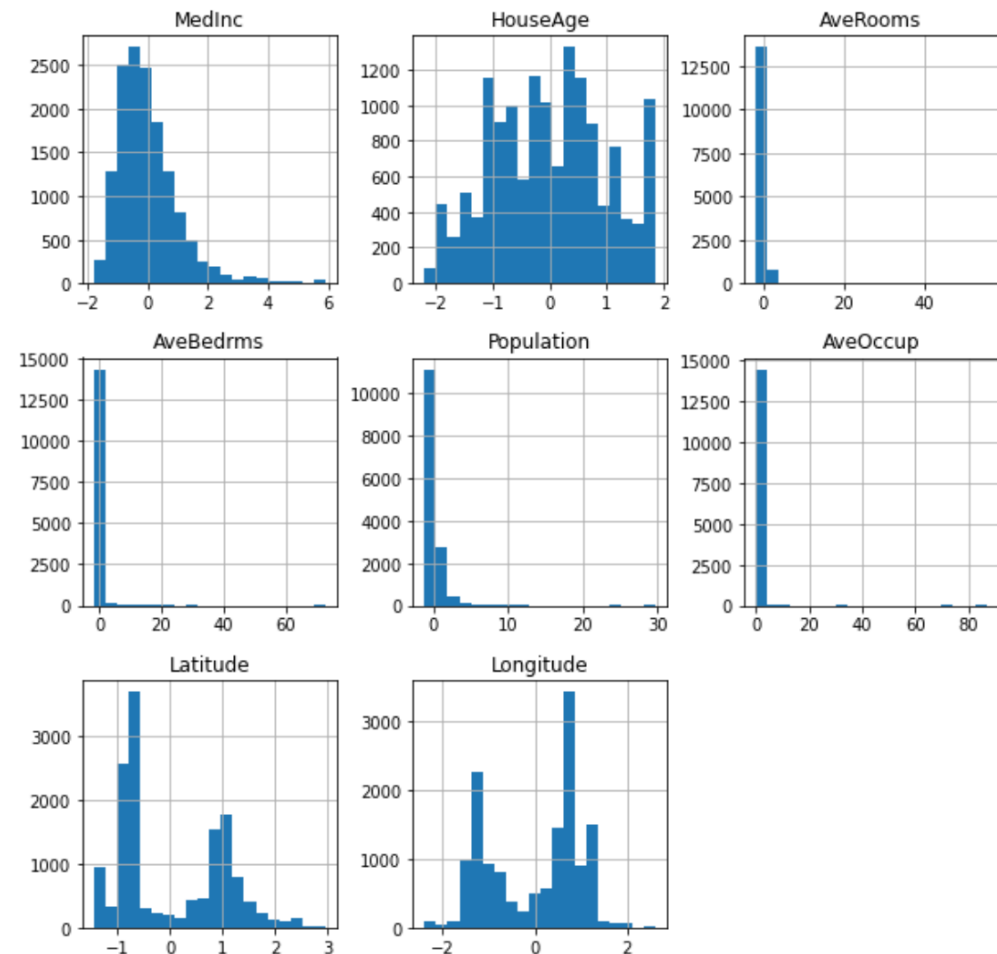
(apart from scaling to the norm).



Standardization

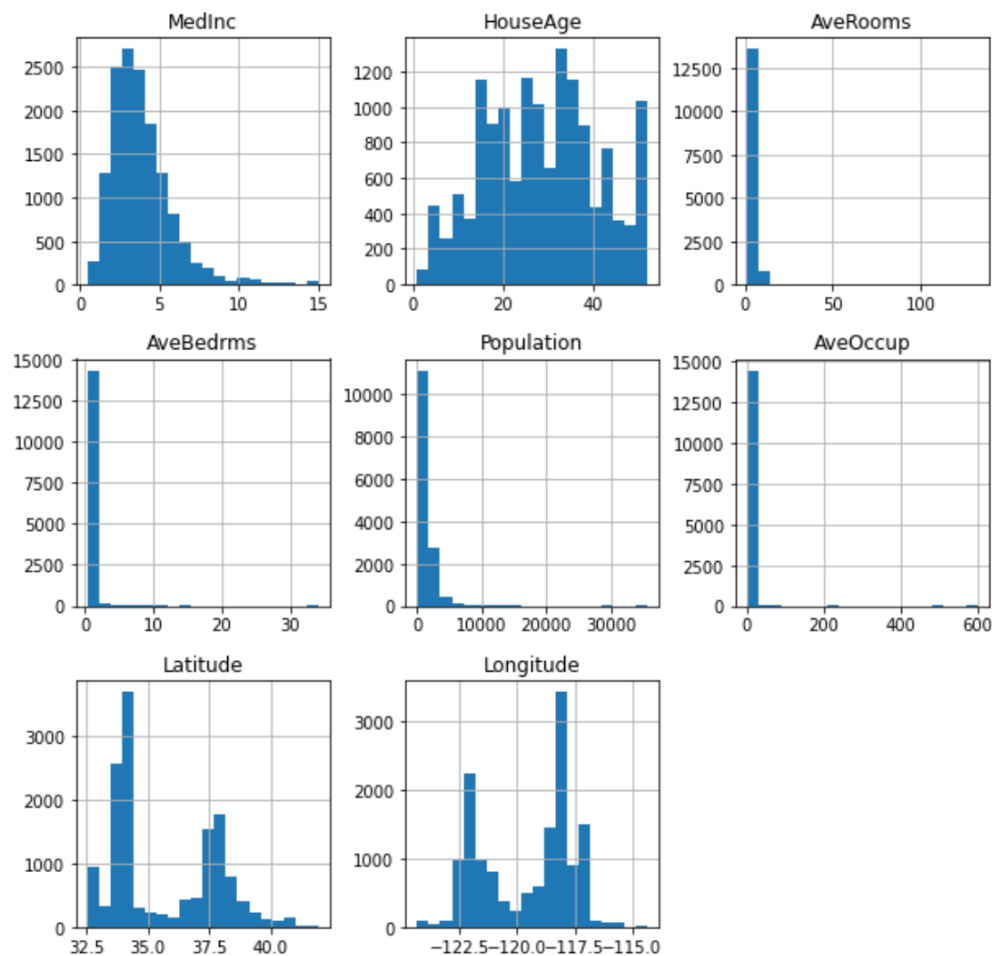


Original scale

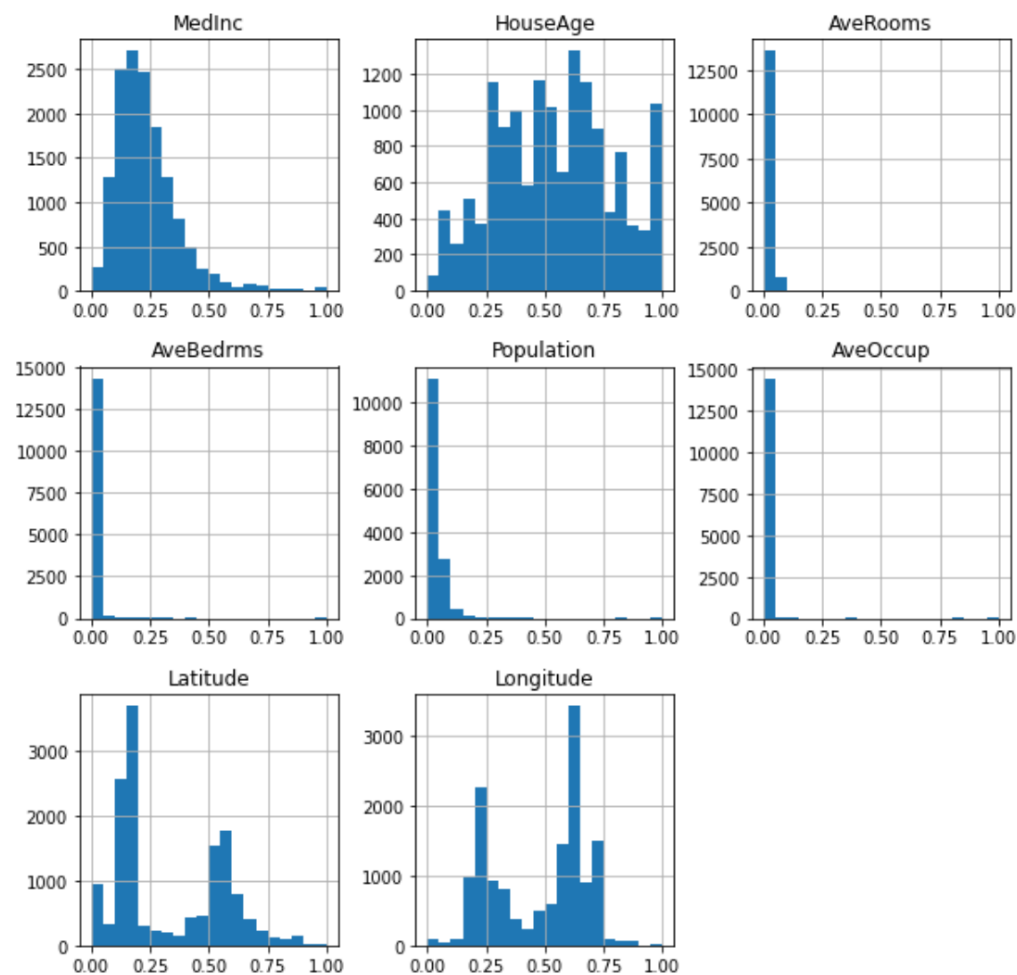


Standardized

MinMax scaling

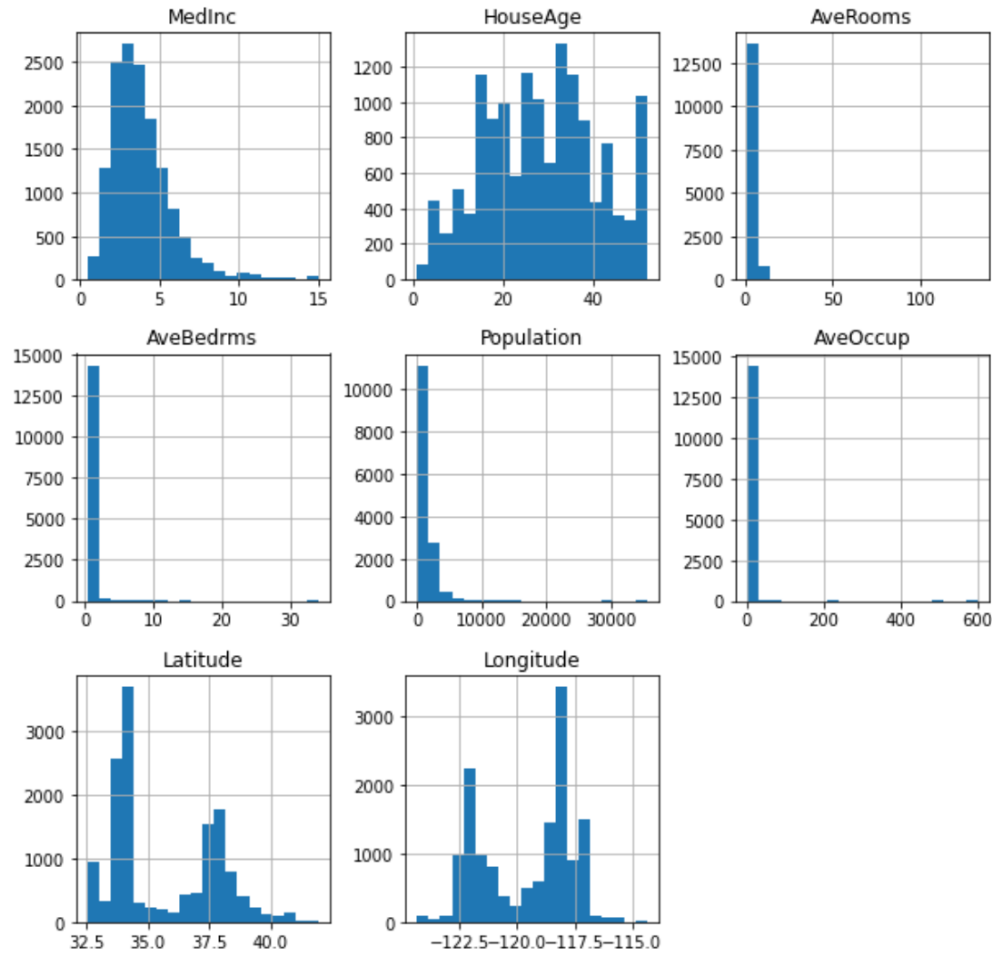


Original scale

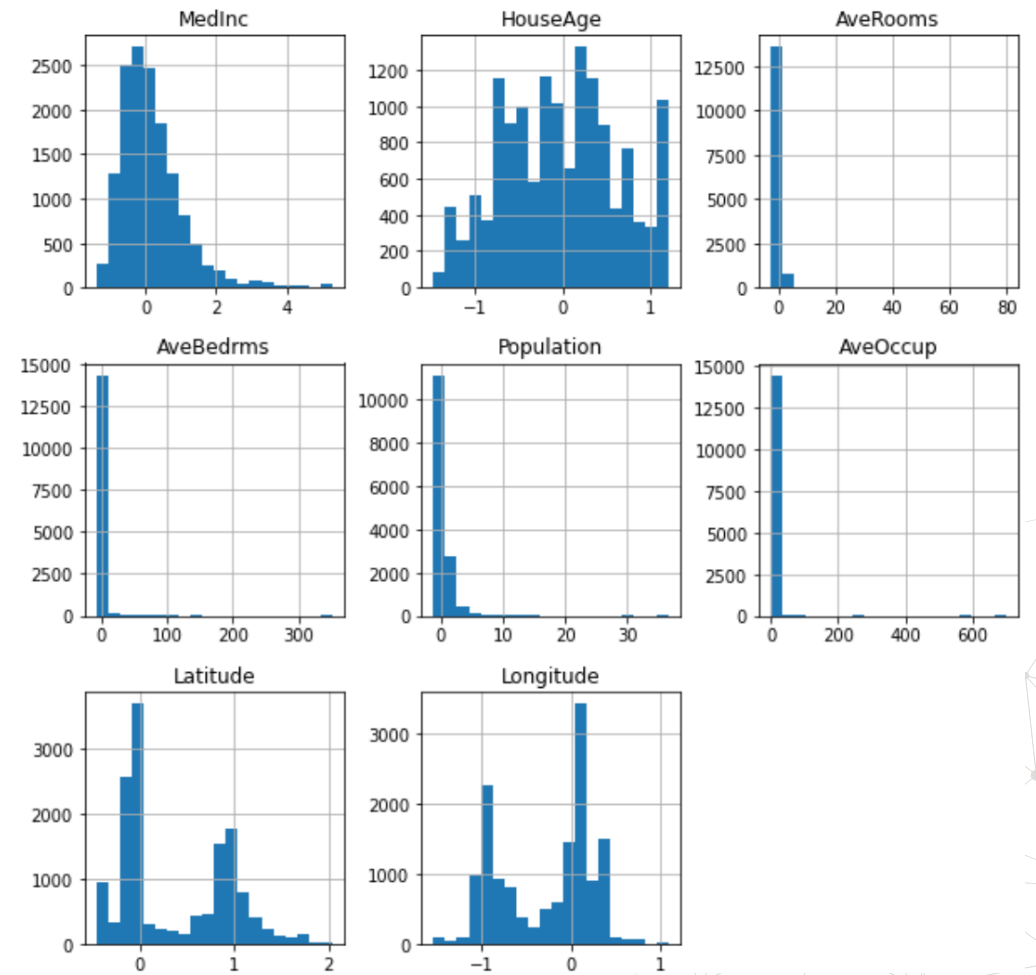


MinMaxScaled

Robust scaling

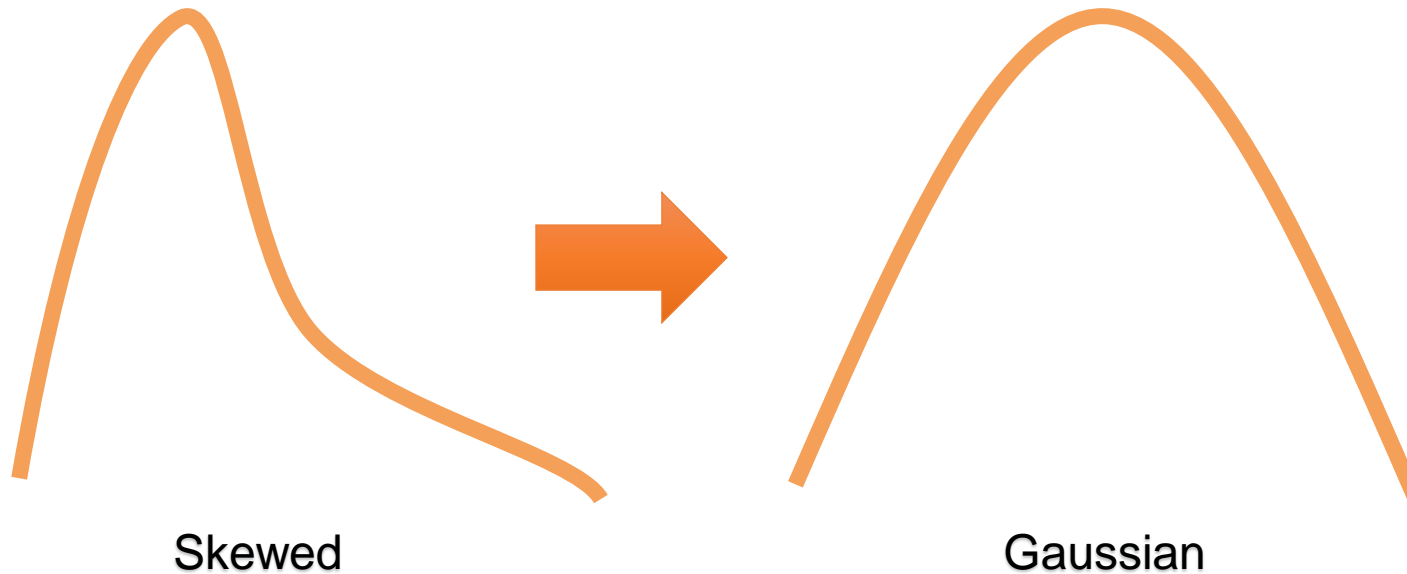


Original scale



Robust scaled

• To change the distribution:



Variable transformation

- Logarithmic
- Reciprocal
- Square-root
- Arcsin
- Power
- Box-Cox
- Yeo-Johnson

Accompanying Jupyter Notebook



Feature scaling with Scikit-learn.

THANK YOU

www.trainindata.com