

Cleaning Data Part II: Takeaways

by Dataquest Labs, Inc. - All rights reserved © 2022

Syntax

- Use **Sort & Filter, Filter** to verify that your data is clean.
- Use **CTRL + 1** to bring up **Format Cells**.
- Use **Format Cells** to standardize the format of your numbers, text, and dates.
- Use **Category: Number** to set the number of decimal places, and be mindful of using currency symbols.
- Use **MEDIAN()** to help avoid the effect of outliers.
- Use **Insert, Recommended Charts** to quickly visualize your data to verify it's correct.
- Use **All Charts** if necessary to manually insert a chart if one isn't recommended.

Concepts

- Computers take things literally — your data needs to be consistent and clean for proper analysis.
- Headers also matter. Make sure your headers are meaningful and consistent.
- Outliers are accurate data, but they often distort an analysis and tell the wrong story.
- Distorted data is also called skewed data.
- Remove outliers or take the median to avoid the effects of skewed data.

Resources

- [MEDIAN function](#)
- [Using Recommended Charts](#)

Takeaways by Dataquest Labs, Inc. - All rights reserved © 2022