Data Assessment Process

Original Data site: <https://www.kaggle.com/zynicide/wine-reviews>

Updated data Was provided at: <https://www.kaggle.com/zynicide/wine-reviews/discussion/83970>

Downloaded update file from: <https://drive.google.com/file/d/1CU_HASG-vPw0KANv6_8BRKCM94ardLjv/view>

Number of observations: 141,617

Dups: Apparently removed

Added “id” as the name for the 1st column number

Sliced data down:

US Observations: 58821

California: 🡺 39072

All Descriptions Populated: X

Designations missing: 18789

Points populated: X

Blanks Prices removed: -186 🡺 38,886

Taster names blank: 15313 🡺 23,573

Title (wine name) blank: 0

Variety blank: 0

Winery Blank: 0

Vintage Blank: 1060