* Data Prep
  1. Fetch all csvs
  2. Add headers
  3. Add ids col
  4. Add label
  5. Consolidate into single df
  6. Check missing values
  7. Make all dtypes numerical
  8. Again check missing values
* EDA

1. Separate continuous and categorical features by unique count
2. Analyse distribution of continuous variables by box plots and histograms
3. There are some outliers, **have to remove those**
4. Analyze distribution of categorical variables by hist and frequency ratio. **Have to remove where the ratio is below 5% or 10%**