Variable Names	Data Quality Issue	Handling Strategy
case_month	None	
res_state	None	
state_fips_code	Contains same data as res_state	Drop Feature
res_county	13 rows with spelling error	Replaced with correct spelling 'DONA ANA'
res_county	5.61% null	Imputed with Missing
county_fips_code	Contains same data as res_county	Drop Feature
age_group	0.726% null	Imputed with Missing
sex	2.168% null	Imputed with Missing
race	12.38% null	Imputed with Missing
ethnicity	13.486% null	Imputed with Missing
case_positive_specimen_interval	Negative Values	Changed to positive
case_positive_specimen_interval	Outliers	Removed
case_positive_specimen_interval	47.36% null	Do nothing
case_onset_interval	57.426% null	5.98% of the null values where symptom_status is Symptomatic where imputed with the mode
case_onset_interval	Negative values	Changed to positive
case_onset_interval	Outliers	Removed
process	90.8% Missing	Drop Feature
exposure_yn	86.14% Missing	Drop Feature
current_status	17.28% of cases with not null case_positive_specimen_interval were Probable case	Changed to Laboratory-confirmed case
symptom_status	2.32% of cases with not null case_onset_interval were Missing	Replaced with Symptomatic
icu_yn	78.14% Missing	Do nothing
death_yn	None	
underlying_conditions_yn	91.77% null	Imputed with missing