1. Blanket changes to make
    a. Assume that store and forward is N if it's null, also turn into boolean.
        i. Issue: trying to do types both runs an error. - resolved
        ii. Issue: returns the results as a string. - resolved itself
    b. Drop duplicate columns - done
    c. If there are any nulls among the passenger column or such the rows can be dropped. - done
    d. Assure that the columns are all the correct data type. - done
        i. passenger_count to int - done
        ii. RatecodeID to int - done
        iii. Clean store_and_fwd_flag - done
    e. Some of the trip_distance rows are zero and should be dropped. - decided to keep
2. 2009 - Conclusion: Dropped.
    a. Too many irreconcilable issues.
        i. Among others the Rate_Code are missing in their entirety.
3. 2010: - conclusion: dropped, perhaps come back later
    a. Change the names to keep them up to date
        i. Vendor_id needs name change and numeric transformation - done
        ii. Pickup and dropoff to datetime datatype - done
        iii. Payment type transform strings to int - done
    b. Add improvement_surcharge, congestion_surcharge, Airport_fee = 0 - done
    c. Need to transform the longitude and latitude column into their pol's int'
        i. Issue: creates duplicate store_and_fwd_flag column with PULocationID and DOLocationID at the end of the dataframe. - resolved
        ii. Coordinates do not match - unresolvable within time limit.
4. >=2011
    a. Replace the congestion_surcharge and Airport_fee nulls with 0's. - done
5. >=2019: - decided to just leave them be.
    a. Doesn't include the congestion charge and airport fee in the total amount. Can only assume it's purposeful given the order of columns.
    b. Need to decide whether to replace the current totals with the true totals including congestion charge and airport fee.
    c. Might be all of them years?

Post-processing notes:

- August of 2014 cleaned data is too small. Investigation needed.