**SQL Scripts**

* **Created State\_Mapping Table to be consistent with using State Abbreviations vs State Names between Zillow & Realtor.com**

CREATE TABLE state\_mapping (

full\_name TEXT,

abbr TEXT

);

INSERT INTO state\_mapping (full\_name, abbr) VALUES

('Alabama', 'AL'), ('Alaska', 'AK'), ('Arizona', 'AZ'), ('Arkansas', 'AR'),

('California', 'CA'), ('Colorado', 'CO'), ('Connecticut', 'CT'), ('Delaware', 'DE'),

('Florida', 'FL'), ('Georgia', 'GA'), ('Hawaii', 'HI'), ('Idaho', 'ID'),

('Illinois', 'IL'), ('Indiana', 'IN'), ('Iowa', 'IA'), ('Kansas', 'KS'),

('Kentucky', 'KY'), ('Louisiana', 'LA'), ('Maine', 'ME'), ('Maryland', 'MD'),

('Massachusetts', 'MA'), ('Michigan', 'MI'), ('Minnesota', 'MN'), ('Mississippi', 'MS'),

('Missouri', 'MO'), ('Montana', 'MT'), ('Nebraska', 'NE'), ('Nevada', 'NV'),

('New Hampshire', 'NH'), ('New Jersey', 'NJ'), ('New Mexico', 'NM'), ('New York', 'NY'),

('North Carolina', 'NC'), ('North Dakota', 'ND'), ('Ohio', 'OH'), ('Oklahoma', 'OK'),

('Oregon', 'OR'), ('Pennsylvania', 'PA'), ('Rhode Island', 'RI'), ('South Carolina', 'SC'),

('South Dakota', 'SD'), ('Tennessee', 'TN'), ('Texas', 'TX'), ('Utah', 'UT'),

('Vermont', 'VT'), ('Virginia', 'VA'), ('Washington', 'WA'), ('West Virginia', 'WV'),

('Wisconsin', 'WI'), ('Wyoming', 'WY');

* **Added the stateabbr to the cleaned\_housing\_data table (cleaned\_housing\_data\_combined)**

CREATE TABLE cleaned\_housing\_data\_combined AS

SELECT

chd.Date,

COALESCE(sm.abbr, chd.StateName) AS stateabbr,

CASE WHEN TRIM(chd.avg\_listing\_price\_realtor) = '' THEN NULL

ELSE ROUND(CAST(chd.avg\_listing\_price\_realtor AS REAL)) END AS avg\_listing\_price\_realtor,

CASE WHEN TRIM(chd.avg\_active\_listing\_count\_realtor) = '' THEN NULL

ELSE ROUND(CAST(chd.avg\_active\_listing\_count\_realtor AS INTEGER)) END AS avg\_active\_listing\_count\_realtor,

CASE WHEN TRIM(chd.avg\_days\_on\_market\_realtor) = '' THEN NULL

ELSE ROUND(CAST(chd.avg\_days\_on\_market\_realtor AS INTEGER)) END AS avg\_days\_on\_market\_realtor,

CASE WHEN TRIM(chd.avg\_pending\_listing\_count\_realtor) = '' THEN NULL

ELSE ROUND(CAST(chd.avg\_pending\_listing\_count\_realtor AS INTEGER)) END AS avg\_pending\_listing\_count\_realtor,

CASE WHEN TRIM(chd.avg\_listing\_price\_zillow) = '' THEN NULL

ELSE ROUND(CAST(chd.avg\_listing\_price\_zillow AS REAL)) END AS avg\_listing\_price\_zillow,

CASE WHEN TRIM(chd.avg\_active\_listing\_count\_zillow) = '' THEN NULL

ELSE ROUND(CAST(chd.avg\_active\_listing\_count\_zillow AS INTEGER)) END AS avg\_active\_listing\_count\_zillow,

CASE WHEN TRIM(chd.avg\_days\_on\_market\_zillow) = '' THEN NULL

ELSE ROUND(CAST(chd.avg\_days\_on\_market\_zillow AS INTEGER)) END AS avg\_days\_on\_market\_zillow,

CASE WHEN TRIM(chd.avg\_pending\_listing\_count\_zillow) = '' THEN NULL

ELSE ROUND(CAST(chd.avg\_pending\_listing\_count\_zillow AS INTEGER)) END AS avg\_pending\_listing\_count\_zillow

FROM cleaned\_housing\_data chd

LEFT JOIN state\_mapping sm

ON chd.StateName = sm.full\_name OR chd.StateName = sm.abbr;

* **Created cleaned\_housing\_data\_ final: Renamed columns & updated blank rows with “NULL”**

CREATE TABLE cleaned\_housing\_data\_final AS

SELECT

Date AS "Date",

stateabbr AS "State Abbr",

CASE WHEN avg\_listing\_price\_realtor = 0 THEN NULL ELSE avg\_listing\_price\_realtor END AS "Avg Listing Price Realtor -",

CASE WHEN avg\_active\_listing\_count\_realtor = 0 THEN NULL ELSE avg\_active\_listing\_count\_realtor END AS "Avg Active Inventory Realtor -",

CASE WHEN avg\_days\_on\_market\_realtor = 0 THEN NULL ELSE avg\_days\_on\_market\_realtor END AS "Avg Days on Market Realtor -",

CASE WHEN avg\_pending\_listing\_count\_realtor = 0 THEN NULL ELSE avg\_pending\_listing\_count\_realtor END AS "Avg Pending Inventory Realtor -",

CASE WHEN avg\_listing\_price\_zillow = 0 THEN NULL ELSE avg\_listing\_price\_zillow END AS "Avg Listing Price Zillow -",

CASE WHEN avg\_active\_listing\_count\_zillow = 0 THEN NULL ELSE avg\_active\_listing\_count\_zillow END AS "Avg Active Inventory Zillow -",

CASE WHEN avg\_days\_on\_market\_zillow = 0 THEN NULL ELSE avg\_days\_on\_market\_zillow END AS "Avg Days on Market Zillow -",

CASE WHEN avg\_pending\_listing\_count\_zillow = 0 THEN NULL ELSE avg\_pending\_listing\_count\_zillow END AS "Avg Pending Inventory Zillow -"

FROM cleaned\_housing\_data\_combined;

* **Created a new table to export to excel (cleaned\_housing\_data\_export)**

CREATE TABLE cleaned\_housing\_data\_export AS

SELECT

Date AS "Date",

"State Abbr" AS "State Abbr",

COALESCE(CAST(ROUND(NULLIF([Avg Listing Price Realtor -], 0), 0) AS TEXT), 'NULL') AS "Avg Listing Price Realtor -",

COALESCE(CAST(ROUND(NULLIF([Avg Active Inventory Realtor -], 0), 0) AS TEXT), 'NULL') AS "Avg Active Inventory Realtor -",

COALESCE(CAST(ROUND(NULLIF([Avg Days on Market Realtor -], 0), 0) AS TEXT), 'NULL') AS "Avg Days on Market Realtor -",

COALESCE(CAST(ROUND(NULLIF([Avg Pending Inventory Realtor -], 0), 0) AS TEXT), 'NULL') AS "Avg Pending Inventory Realtor -",

COALESCE(CAST(ROUND(NULLIF([Avg Listing Price Zillow -], 0), 0) AS TEXT), 'NULL') AS "Avg Listing Price Zillow -",

COALESCE(CAST(ROUND(NULLIF([Avg Active Inventory Zillow -], 0), 0) AS TEXT), 'NULL') AS "Avg Active Inventory Zillow -",

COALESCE(CAST(ROUND(NULLIF([Avg Days on Market Zillow -], 0), 0) AS TEXT), 'NULL') AS "Avg Days on Market Zillow -",

COALESCE(CAST(ROUND(NULLIF([Avg Pending Inventory Zillow -], 0), 0) AS TEXT), 'NULL') AS "Avg Pending Inventory Zillow -"

FROM cleaned\_housing\_data\_final;