SELECT \*

FROM `balance-of-trade-us-and-china.US\_Import\_And\_Export\_On\_China.Countries`

WHERE CTYNAME= 'China'

-- checking for null or missing value

SELECT COUNT(\*) AS total\_rows,

       COUNT(EYR) AS non\_null\_rows,

       COUNT(\*) - COUNT(EYR) AS null\_count

FROM `balance-of-trade-us-and-china.US\_Import\_And\_Export\_On\_China.Countries`

--- There were no missing values in all the columns

-- checking for duplicate value

SELECT CTY\_CODE,

       CTYNAME,

      COUNT(\*) AS DUPLICATE\_COUNT

FROM `balance-of-trade-us-and-china.US\_Import\_And\_Export\_On\_China.Countries`

WHERE upper(CTYNAME) IN ('CHINA','VIETNAM','THAILAND','JAPAN','CAMBODIA','SOUTH AFRICA','TAIWAN','EUROPEAN UNION')

GROUP BY CTY\_CODE, CTYNAME

HAVING COUNT (\*) >1

-- checking for the number of countries. There are 259 CTYNAMES

SELECT CTYNAME, COUNT (\*) AS Number\_of\_duplicate

FROM `balance-of-trade-us-and-china.US\_Import\_And\_Export\_On\_China.Countries`

GROUP BY CTYNAME

-- OR

SELECT DISTINCT (CTYNAME)

FROM `balance-of-trade-us-and-china.US\_Import\_And\_Export\_On\_China.Countries`

-- Check for Case Inconsistencies

SELECT Proper (CTYNAME)

FROM `balance-of-trade-us-and-china.US\_Import\_And\_Export\_On\_China.Countries`

-- OR

SELECT

  LOWER(CTYNAME) AS normalized\_name,

  ARRAY\_AGG(DISTINCT CTYNAME) AS actual\_variants,

  COUNT(\*) AS total\_rows

FROM `balance-of-trade-us-and-china.US\_Import\_And\_Export\_On\_China.Countries`

GROUP BY normalized\_name

HAVING ARRAY\_LENGTH(ARRAY\_AGG(DISTINCT CTYNAME)) > 1;

-- I what to get data on these 8 countries ( China, Vietnam, Thailand, Japan, Cambodia, South Africa, Taiwan and the European Union)

-- I am focusing on a 10 year period

CREATE TABLE `balance-of-trade-us-and-china.US\_Trade.US\_Trading\_Imbalance\_With\_Eight\_Countries` AS

SELECT \*

FROM `balance-of-trade-us-and-china.US\_Import\_And\_Export\_On\_China.Countries`

WHERE UPPER(CTYNAME) IN ('CHINA','VIETNAM','THAILAND','JAPAN','CAMBODIA','SOUTH AFRICA','TAIWAN','EUROPEAN UNION')

AND year BETWEEN 2014 AND 2024