CREATE TABLE `550DB`.`Providers` (

`FPN` VARCHAR(20) NOT NULL,

`ProviderName` VARCHAR(50) NULL,

`LegalBusinessName` VARCHAR(70) NULL,

`CertifiedBeds` INT NULL,

`AveResidentsPerDay` FLOAT NULL,

`ProviderType` VARCHAR(25) NULL,

`OwnershipType` VARCHAR(40) NULL,

`ProviderResidesInHospital` VARCHAR(5) NULL,

`ContinuingCareRetirementCommunity` VARCHAR(5) NULL,

`AbuseIcon` VARCHAR(5) NULL,

`MostRecentHealthInspectionMoreThan2YearsAgo` VARCHAR(5) NULL,

`WithResidentAndFamilyCouncil` VARCHAR(10) NULL,

`AutoSprinklerSysInAllReqAreas` VARCHAR(3) NULL,

`NumberOfAllBeds` INT NULL,

`TotalNumberofOccupiedBeds` INT NULL,

PRIMARY KEY (`FPN`),

CHECK (ProviderType IN ('Medicare', 'Medicaid', 'Medicare and Medicaid', '')),

CHECK (OwnershipType IN ('For profit - Corporation', 'For profit - Individual', 'For profit - Limited Liability company', 'For profit - Partnership', 'Government - City', 'Government - City/County', 'Government - County', 'Government - Federal', 'Government - Hospital district', 'Government - State', 'Non profit - Church related', 'Non profit - Corporation', 'Non profit - Other', '')),

CHECK (ProviderResidesInHospital IN ('TRUE', 'FALSE', '')),

CHECK (ContinuingCareRetirementCommunity IN ('TRUE', 'FALSE', '')),

CHECK (AbuseIcon IN ('TRUE', 'FALSE', '')),

CHECK (MostRecentHealthInspectionMoreThan2YearsAgo IN ('TRUE', 'FALSE', '')),

CHECK (WithResidentAndFamilyCouncil IN ('Both', 'Family', 'None', 'Resident', '')),

CHECK (AutoSprinklerSysInAllReqAreas IN ('Yes', 'Partial', '')));

USE 550DB;

LOAD DATA LOCAL INFILE 'C:/Users/rjsas/Desktop/CSVs/Providers.csv'

INTO TABLE Providers

FIELDS TERMINATED BY ','

OPTIONALLY ENCLOSED BY '"'

LINES TERMINATED BY '\n'

IGNORE 1 ROWS;

CREATE TABLE `550DB`.`Locations` (

`FPN` VARCHAR(20) NOT NULL,

`Address` varchar(55),

`City` varchar(25),

`State` varchar(2),

`Zip` int,

`SSACounty` int,

`CountyName` varchar(25),

`Longitude` float,

`Latitude` float,

`Phone` int,

PRIMARY KEY (`FPN`),

FOREIGN KEY (`FPN`) REFERENCES `Providers` (`FPN`));

USE 550DB;

LOAD DATA LOCAL INFILE 'C:/Users/rjsas/Desktop/CSVs/Locations.csv'

INTO TABLE Locations

FIELDS TERMINATED BY ','

OPTIONALLY ENCLOSED BY '"'

LINES TERMINATED BY '\n'

IGNORE 1 ROWS;

CREATE TABLE `550DB`.`CMSData` (

`FPN` VARCHAR(20) NOT NULL,

`OverallRating` int,

`HealthInspectionRating` int,

`QMRating` int,

`LongStayQMRating` int,

`ShortStayQMRating` int,

`StaffingRating` int,

`RNStaffingRating` int,

`NursingAide\_ReportedHoursPerResidentPerDay` float,

`LPN\_ReportedHoursPerResidentPerDay` float,

`RN\_ReportedHoursPerResidentPerDay` float,

`LicensedStaffing\_ReportedHoursPerResidentPerDay` float,

`TotalNurse\_ReportedHoursPerResidentPerDay` float,

`PT\_ReportedHoursPerResidentPerDay` float,

`NursingAide\_AdjustedStaffingHoursPerResidentPerDay` float,

`LPN\_AdjustedStaffingHoursPerResidentPerDay` float,

`RN\_AdjustedStaffingHoursPerResidentPerDay` float,

`TotalNurse\_AdjustedStaffingHoursPerResidentPerDay` float,

`TotalWeightedHealthSurveyScore` float,

`NumReportedIncidents` int,

`NumSubstantiatedComplaints` int,

`NumFines` int,

`AmountFines` int,

`NumPaymentDenials` int,

`NumPenalties` int,

PRIMARY KEY (`FPN`),

FOREIGN KEY (`FPN`) REFERENCES `Providers` (`FPN`),

CHECK (`OverallRating` IN (1, 2, 3, 4, 5, '')),

CHECK (`HealthInspectionRating` IN (1, 2, 3, 4, 5, '')),

CHECK (`QMRating` IN (1, 2, 3, 4, 5, '')),

CHECK (`LongStayQMRating` IN (1, 2, 3, 4, 5, '')),

CHECK (`ShortStayQMRating` IN (1, 2, 3, 4, 5, '')),

CHECK (`StaffingRating` IN (1, 2, 3, 4, 5, '')),

CHECK (`RNStaffingRating` IN (1, 2, 3, 4, 5, '')));

USE 550DB;

LOAD DATA LOCAL INFILE 'C:/Users/rjsas/Desktop/CSVs/CMSData.csv'

INTO TABLE CMSData

FIELDS TERMINATED BY ','

OPTIONALLY ENCLOSED BY '"'

LINES TERMINATED BY '\n'

IGNORE 1 ROWS;

CREATE TABLE `550DB`.`COVIDData` (

`FPN` VARCHAR(20) NOT NULL,

`SubmittedData` varchar(1),

`PassedQACheck` varchar(1),

`ResidentsWeeklyAdmissions` int,

`ResidentsTotalAdmissions` int,

`ResidentsWeeklyConfirmed` int,

`ResidentsTotalConfirmed` int,

`ResidentsWeeklySuspected` int,

`ResidentsTotalSuspected` int,

`ResidentsWeeklyAllDeath` int,

`ResidentsTotalAllDeath` int,

`ResidentsWeeklyCovidDeaths` int,

`ResidentsTotalCovidDeaths` int,

`ResidentsAbleToTestAllWithinWeek` varchar(1),

`ResidentsAveTestingTurnaroundTime` varchar(10),

`StaffAbleToTestAllWithinWeek` varchar(1),

`StaffAveTestingTurnaroundTime` varchar(10),

`InHousePOCTMachine` varchar(1),

`ResidentsPOCTTestedSinceLastReport` int,

`StaffPOCTTestedSinceLastReport` int,

`EnoughPOCTSuppliesToTestAllStaff` varchar(1),

`StaffWeeklyConfirmed` int,

`StaffTotalConfirmed` int,

`StaffWeeklySuspected` int,

`StaffTotalSuspected` int,

`StaffWeeklyCovidDeaths` int,

`StaffTotalCovidDeaths` int,

`ShortageOfNursingStaff` varchar(1),

`ShortageOfClinicalStaff` varchar(1),

`ShortageOfAides` varchar(1),

`ShortageOfOtherStaff` varchar(1),

`OneWeekSupplyN95Masks` varchar(1),

`OneWeekSupplySurgicalMasks` varchar(1),

`OneWeekSupplyEyeProtection` varchar(1),

`OneWeekSupplyGowns` varchar(1),

`OneWeekSupplyGloves` varchar(1),

`OneWeekSupplySanitizer` varchar(1),

`AnyCurrentSupplyN95Masks` varchar(1),

`AnyCurrentSupplySurgicalMasks` varchar(1),

`AnyCurrentSupplyEyeProtection` varchar(1),

`AnyCurrentSupplyGowns` varchar(1),

`AnyCurrentSupplyGloves` varchar(1),

`AnyCurrentSupplySanitizer` varchar(1),

`VentilatorDependentUnit` varchar(1),

`NumVentilatorsInFacility` int,

`NumVentilatorsInUse` int,

`AnyCurrentVentilatorSupplies` varchar(1),

`OneWeekVentilatorSupplies` varchar(1),

`TotalResidentConfirmedCasesPer1000` float,

`TotalResidentCovidDeathsPer1000` float,

`TotalResidentCovidDeathsPercentageOfConfirmedCases` float,

`ThreeOrMoreConfirmedCasesThisWeek` varchar(1),

`InitialConfirmedCaseThisWeek` varchar(1),

PRIMARY KEY (`FPN`),

FOREIGN KEY (`FPN`) REFERENCES `Providers` (`FPN`),

CHECK (`SubmittedData` IN ('Y', 'N', '')),

CHECK (`PassedQACheck` IN ('Y', 'N', '')),

CHECK (`ResidentsAbleToTestAllWithinWeek` IN ('Y', 'N', '')),

CHECK (`ResidentsAveTestingTurnaroundTime` IN ('<1 DAY', '1-2 DAYS', '3-7 DAYS', '>7 DAYS', 'N/A', '')),

CHECK (`StaffAbleToTestAllWithinWeek` IN ('Y', 'N', '')),

CHECK (`StaffAveTestingTurnaroundTime` IN ('<1 DAY', '1-2 DAYS', '3-7 DAYS', '>7 DAYS', 'N/A', '')),

CHECK (`InHousePOCTMachine` IN ('Y', 'N', '')),

CHECK (`EnoughPOCTSuppliesToTestAllStaff` IN ('Y', 'N', '')),

CHECK (`ShortageOfNursingStaff` IN ('Y', 'N', '')),

CHECK (`ShortageOfClinicalStaff` IN ('Y', 'N', '')),

CHECK (`ShortageOfAides` IN ('Y', 'N', '')),

CHECK (`ShortageOfOtherStaff` IN ('Y', 'N', '')),

CHECK (`OneWeekSupplyN95Masks` IN ('Y', 'N', '')),

CHECK (`OneWeekSupplySurgicalMasks` IN ('Y', 'N', '')),

CHECK (`OneWeekSupplyEyeProtection` IN ('Y', 'N', '')),

CHECK (`OneWeekSupplyGowns` IN ('Y', 'N', '')),

CHECK (`OneWeekSupplyGloves` IN ('Y', 'N', '')),

CHECK (`OneWeekSupplySanitizer` IN ('Y', 'N', '')),

CHECK (`AnyCurrentSupplyN95Masks` IN ('Y', 'N', '')),

CHECK (`AnyCurrentSupplySurgicalMasks` IN ('Y', 'N', '')),

CHECK (`AnyCurrentSupplyEyeProtection` IN ('Y', 'N', '')),

CHECK (`AnyCurrentSupplyGowns` IN ('Y', 'N', '')),

CHECK (`AnyCurrentSupplyGloves` IN ('Y', 'N', '')),

CHECK (`AnyCurrentSupplySanitizer` IN ('Y', 'N', '')),

CHECK (`VentilatorDependentUnit` IN ('Y', 'N', '')),

CHECK (`AnyCurrentVentilatorSupplies` IN ('Y', 'N', '')),

CHECK (`OneWeekVentilatorSupplies` IN ('Y', 'N', '')),

CHECK (`ThreeOrMoreConfirmedCasesThisWeek` IN ('Y', 'N', '')),

CHECK (`InitialConfirmedCaseThisWeek` IN ('Y', 'N', '')));

USE 550DB;

LOAD DATA LOCAL INFILE 'C:/Users/rjsas/Desktop/CSVs/COVIDData.csv'

INTO TABLE COVIDData

FIELDS TERMINATED BY ','

OPTIONALLY ENCLOSED BY '"'

LINES TERMINATED BY '\n'

IGNORE 1 ROWS;

USE 550DB;

LOAD DATA LOCAL INFILE 'C:/Users/rjsas/Desktop/CSVs/**noData**.csv'

INTO TABLE COVIDData

FIELDS TERMINATED BY ','

OPTIONALLY ENCLOSED BY '"'

LINES TERMINATED BY '\n'

IGNORE 1 ROWS;

**OPTIMIZATION**

USE 550DB;

CREATE TABLE PassedQADistances AS (WITH MinDist AS (

WITH Pairs AS (

WITH NoData AS (

SELECT P.ProviderName, P.FPN, L.State, L.Longitude, L.Latitude

FROM Providers P JOIN COVIDData C ON P.FPN = C.FPN JOIN Locations L on P.FPN = L.FPN

WHERE C.PassedQACheck = "N" AND L.Longitude <> 0.0 AND L.Latitude <> 0.0

)

SELECT N.ProviderName as NoReport, N.FPN as NoFPN, Y.ProviderName as YesReport, N.State as NoState, Y.State as YesState, 12742 \* SIN(SQRT(0.5 - COS((N.Latitude - Y.Latitude) \* PI() / 180) / 2 + (COS(N.Latitude \* PI() / 180) \* COS(Y.Latitude \* PI() / 180) \* (1-COS((N.Longitude - Y.Longitude)\* PI()/180))/2))) as Distance

FROM NoData N, (

SELECT P.ProviderName, L.State, L.Longitude, L.Latitude

FROM Providers P JOIN COVIDData C ON P.FPN = C.FPN JOIN Locations L on P.FPN = L.FPN

WHERE C.PassedQACheck = "Y" AND L.Longitude <> 0.0 AND L.Latitude <> 0.0

) AS Y

)

SELECT P.NoFPN, P.NoReport, MIN(P.Distance) as Distance

FROM Pairs P

GROUP BY P.NoFPN, P.NoReport

)

SELECT DISTINCT M.NoReport, M.NoFPN, P.YesReport, P.YesFPN, P.YesState, M.Distance

FROM MinDist M JOIN (

WITH NoData AS (SELECT P.ProviderName, L.State, L.Longitude, L.Latitude

FROM Providers P JOIN COVIDData C ON P.FPN = C.FPN JOIN Locations L on P.FPN = L.FPN

WHERE C.PassedQACheck = "N" AND L.Longitude <> 0.0 AND L.Latitude <> 0.0)

SELECT N.ProviderName as NoReport, Y.ProviderName as YesReport, Y.FPN as YesFPN, N.State as NoState, Y.State as YesState, 12742 \* SIN(SQRT(0.5 - COS((N.Latitude - Y.Latitude) \* PI() / 180) / 2 + (COS(N.Latitude \* PI() / 180) \* COS(Y.Latitude \* PI() / 180) \* (1-COS((N.Longitude - Y.Longitude)\* PI()/180))/2))) as Distance

FROM NoData N, (SELECT P.FPN, P.ProviderName, L.State, L.Longitude, L.Latitude

FROM Providers P JOIN COVIDData C ON P.FPN = C.FPN JOIN Locations L on P.FPN = L.FPN

WHERE C.PassedQACheck = "Y" AND L.Longitude <> 0.0 AND L.Latitude <> 0.0) AS Y

)

AS P ON M.Distance = P.Distance

WHERE M.NoReport = P.NoReport

ORDER BY M.NoReport);

USE 550DB;

CREATE TABLE SubmittedDataDistances AS (WITH MinDist AS (

WITH Pairs AS (

WITH NoData AS (

SELECT P.ProviderName, P.FPN, L.State, L.Longitude, L.Latitude

FROM Providers P JOIN COVIDData C ON P.FPN = C.FPN JOIN Locations L on P.FPN = L.FPN

WHERE C.SubmittedData = "N" AND L.Longitude <> 0.0 AND L.Latitude <> 0.0

)

SELECT N.ProviderName as NoReport, N.FPN as NoFPN, Y.ProviderName as YesReport, N.State as NoState, Y.State as YesState, 12742 \* SIN(SQRT(0.5 - COS((N.Latitude - Y.Latitude) \* PI() / 180) / 2 + (COS(N.Latitude \* PI() / 180) \* COS(Y.Latitude \* PI() / 180) \* (1-COS((N.Longitude - Y.Longitude)\* PI()/180))/2))) as Distance

FROM NoData N, (

SELECT P.ProviderName, L.State, L.Longitude, L.Latitude

FROM Providers P JOIN COVIDData C ON P.FPN = C.FPN JOIN Locations L on P.FPN = L.FPN

WHERE C.SubmittedData = "Y" AND L.Longitude <> 0.0 AND L.Latitude <> 0.0

) AS Y

)

SELECT P.NoFPN, P.NoReport, MIN(P.Distance) as Distance

FROM Pairs P

GROUP BY P.NoFPN, P.NoReport

)

SELECT DISTINCT M.NoReport, M.NoFPN, P.YesReport, P.YesFPN, P.YesState, M.Distance

FROM MinDist M JOIN (

WITH NoData AS (SELECT P.ProviderName, L.State, L.Longitude, L.Latitude

FROM Providers P JOIN COVIDData C ON P.FPN = C.FPN JOIN Locations L on P.FPN = L.FPN

WHERE C.SubmittedData = "N" AND L.Longitude <> 0.0 AND L.Latitude <> 0.0)

SELECT N.ProviderName as NoReport, Y.ProviderName as YesReport, Y.FPN as YesFPN, N.State as NoState, Y.State as YesState, 12742 \* SIN(SQRT(0.5 - COS((N.Latitude - Y.Latitude) \* PI() / 180) / 2 + (COS(N.Latitude \* PI() / 180) \* COS(Y.Latitude \* PI() / 180) \* (1-COS((N.Longitude - Y.Longitude)\* PI()/180))/2))) as Distance

FROM NoData N, (SELECT P.FPN, P.ProviderName, L.State, L.Longitude, L.Latitude

FROM Providers P JOIN COVIDData C ON P.FPN = C.FPN JOIN Locations L on P.FPN = L.FPN

WHERE C.SubmittedData = "Y" AND L.Longitude <> 0.0 AND L.Latitude <> 0.0) AS Y

)

AS P ON M.Distance = P.Distance

WHERE M.NoReport = P.NoReport

ORDER BY M.NoReport);