The aim of the sub problem on the project is to find the correlation between the population of different age groups in counties and the number of pharmacies present.

Table

Description automatically generated with medium confidence

In the last row of the matrix, we can observe the correlation of different age groups with the pharmacy count. From this we can infer that all the different age groups have a positive correlation with the highest among age 10-14. From the trends in the correlation matrix, we can observe that the correlation among population of age 65 and above is lower than correlation among population of age 65 and below, with the highest being 0.841467 (age 10-14).

Text

Description automatically generated

We can further observe the trends from the scatterplots of different age groups vs pharmacy counts:

Chart, scatter chart

Description automatically generated Chart, scatter chart

Description automatically generated

Chart, scatter chart

Description automatically generated Chart, scatter chart

Description automatically generated

Chart, scatter chart

Description automatically generated Chart, scatter chart

Description automatically generated

Chart, scatter chart

Description automatically generated Chart, scatter chart

Description automatically generated

Chart, scatter chart

Description automatically generated Chart, scatter chart

Description automatically generated

Chart, scatter chart

Description automatically generated Chart, scatter chart

Description automatically generated

We can see that in the last graph- MED\_AGE vs Pharmacy count, as the median age increases the pharmacy counts slightly decreases. Counties with older population (65 and above) have lesser pharmacies than counties with younger population based on the given data. The given data had lots of missing values and outliers thus we cannot make any clear inferences.