GROWTH GUIDE DATA ANALYTICS TASK

DATASET - The dataset provides the attribute of women and their incidence of diabetes. Data - https://drive.google.com/drive/folders/1200VL9GIzB8DL0AJrHyrbAdnaDDi6d1k?usp=sharing

Methods to apply on dataset-

1. Check if all the features are numeric.
2. Given dataset contains many (zero) values, replace 0 with NaN value.
3. Replace NaN with mean , median, mode.
4. Get the correlation matrix of each column,
5. Get the descriptive statistics of each column.
6. Inferential Statistics
    a. **Normal Deviate Z Test**
    (Population mean and std is known in this case)
    #Get the mean and std value of the population from descriptive stats table.
    I am providing a sample dataset, check whether this sample is drawn from population data using the Z test.
    Hypothesis
    H0 - Samples are likely drawn from the same distributions.
    H1 - Samples are likely drawn from different distributions.
    Using the Z test, check whether the null hypothesis is going to accept or reject.

    b. **One Sample T- Test.**
    (Population mean is known but std is unknown)
    #Get the mean value of population from descriptive stats table and find the estimated std of population.
    I am providing a sample dataset, check whether this sample is drawn from population data using the one paired t-test.
    Hypothesis
    H0 - Samples are likely drawn from the same distributions.
    H1 - Samples are likely drawn from different distributions.
    Using the One paired t-test, check whether the null hypothesis is going to accept or reject.