PDS ASSIGNMENT 3

NAME: Shreyan Reddy Gangwar

ST ID : 12599073

a)The mean and maximum glucose readings are determined using a random sample of 25 observations from a population dataset. The sample statistics are visually compared to the total population, demonstrating the sample's representativeness and capacity to catch extreme values. This visual investigation aids in understanding the heterogeneity between sample and population statistics, offering insights into the sample's representativeness and acting as a useful tool for preliminary data exploration.

A screenshot of a computer

Description automatically generated

Results

A screenshot of a computer

Description automatically generated

Analysis: The sample and population data show that the mean glucose level in the sample is somewhat lower than the population mean, which might be attributable to random fluctuation in small samples. The marginal difference between the maximum of the sample and the maximum of the population shows that the sample's extreme values are near to the population's extreme values. This is consistent with sampling randomness, implying that the sample properly represents the population's center tendency and extreme values. When making conclusions about larger populations, however, recognizing the unpredictability and inherent biases induced by small sample sizes is critical.b) Find the 98th

b)percentile of BMI of your sample and the population and compare the results

A screenshot of a computer

Description automatically generated

Analysis: The distribution of BMI values at the higher percentiles differs little between the sample and the population when the 98th percentile BMI is compared. The lower 98th percentile BMI in the sample (44.68) compared to the population (47.53) shows that the sample had somewhat lower BMI values for persons in this specific percentile. However, the modest discrepancy suggests that the sample and population are quite similar at this percentile. The virtually identical representation in the visualizations of the population's BMI values and the sample's 98th percentile suggests a significant agreement between the two datasets, demonstrating the consistency in capturing the distribution's properties. This alignment confirms the sample's dependability in reflecting the higher percentile trends found in the larger population.

C)

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

The bootstrap samples properly represent the blood pressure distribution and features of the population, resulting in credible estimates for various blood pressure statistics. The stability and closeness of the bootstrap estimates to population values indicate the bootstrap method's resilience in estimating population parameters. This study backs up the value of bootstrap resampling as a tool for statistical inference, allowing for relevant conclusions without requiring considerable data collecting. The use of bootstrap samples to generate estimates closely resembles the real features of the complete population, which contributes to the reliability of statistical findings.

Analysis: The excellent agreement between the bootstrap mean (68.97) and the population mean (69.11) demonstrates the bootstrap method's dependability in capturing the central trend of BloodPressure. The close equivalency of the bootstrap standard deviation (19.25) to the population standard deviation (19.36) indicates that the variability found in BloodPressure readings across the bootstrap samples well encapsulates the entire variability within the population. Furthermore, the congruence of BloodPressure percentiles between bootstrap samples and the population strengthens the integrity of the sample's distribution, confirming its representativeness of the larger population. This convergence in statistical parameters highlights the bootstrap technique's durability in approximating essential BloodPressure features and its application for trustworthy statistical conclusions in circumstances with limited data.