**Statistics**

**Question Paper- 1**

**Answer all the Questions (8 marks)**

1. Explain Variance, Standard Deviation with example.
2. Explain Bayes Theorem.
3. Answer in true/false
4. If a fair coin is tossed many times and the last eight tosses are all heads, then the chance that the next toss will be heads is somewhat less than 50%
5. Drawing a face card and drawing an ace from a full deck of playing cards are mutually exclusive events.
6. N( mean= 1500, standard deviation = 300)

Shannon is a randomly selected SAT taker, and nothing is known about Shannon’s SAT Aptitude. What is the probability Shannon scores at least 1630 on her SATs?

**Part-B ( 3 marks each)**

1. For the given quantitative samples, compute the following. [3 Marks]

34,67,40,72,37,33,42,62,49,32,52,40,31,19,68,55,57,54,37,32,

54,38,20,50,56,48,35,52,29,56,68,65,45,44,54,39,29,56,43,42.

1. First Quartile
2. Third Quartile
3. Inter Quartile Range (IQR)
4. Inner fence
5. Outer fence
6. In the sample of 100 students from the 2013 YRBSS survey, the average number of days per week that students lifted weights was 2.78 days with a standard deviation of 2.56 days (coincidentally the same as days active). Compute a 95% confidence interval for the average for all students from the survey.

Assume that conditions for normal model are met.

1. A health drink manufacturing company claims their average fill volume of their bottle is 235ml across their factories with a standard deviation of 5ml.
2. What is the probability of randomly picked bottle measuring less than 230ml?
3. What is the probability of randomly picked bottle measuring more than 235ml?
4. Give an inference and comments based on the results of a & b.
5. Play the piano. Georgianna claims that in a small city renowned for its music school, the average child takes at least 5 years of piano lessons. We have a random sample of 20 children from the city, with a mean of 4.6 years of piano lessons and standard deviation of 2.2 years
6. Evaluate Georgianna’s claim using hypothesis test
7. Construct a 95 % confidence interval for the number of years students in this city take piano lessons, and interpret it in the context of data.
8. Do your results from the hypothesis test and the confidence interval agree? Explain your reasoning.
9. Air quality measurements were collected in a random sample of 25 country capitals in 2013, and then again in the same cities in 2014. We would like to use these data to compare average air quality between the two years.
10. Should we use a one-sided test or a two-sided test? Explain your reasoning.
11. Should we use a paired or non-paired test? Explain your reasoning.
12. Should we use t-test or z-test ? Explain your reasoning.

**Part – C**

1. Load the data (recruitment\_data.csv) file and perform the following, [7 Marks]
2. Delete all missing value rows and report the reduced data frame
3. Check the variable ‘sales\_quota\_pct’, for its normality, justify with suitable plots
4. Check whether the sales\_quota\_pct vary significantly with respect to the mode of requirement (Referral candidates Vs Applied Online)
5. Check the proportion of Attrition vary significantly or not, with respect to all mode of recruitment