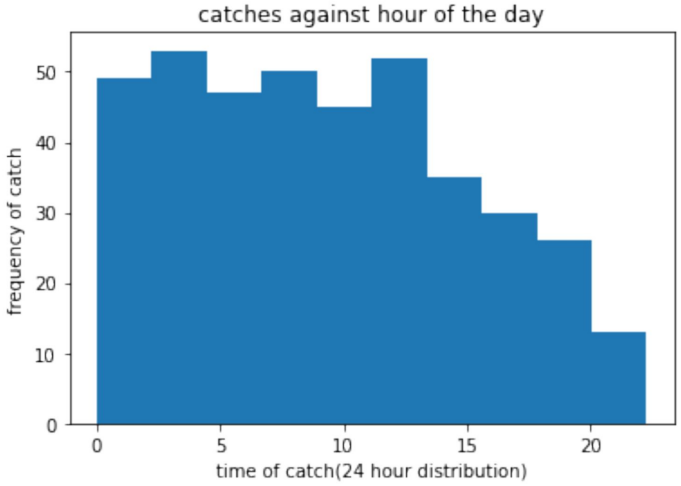
CW1 Statistics

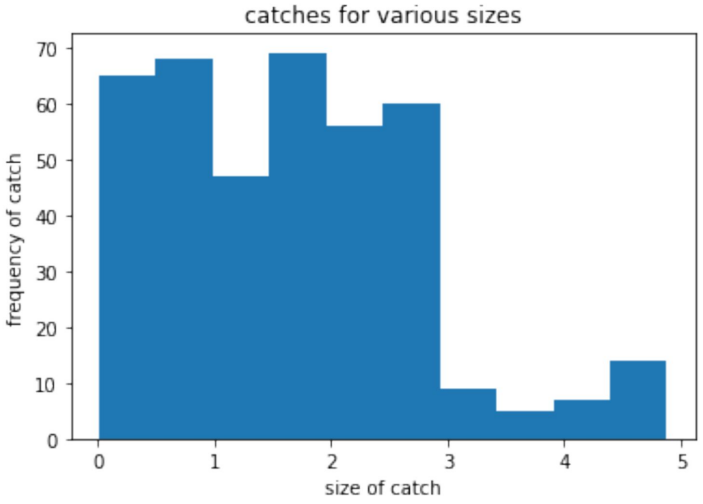
Task 1

Plot for mapping time of catch in the day against its frequency



- **Median** timing : **9.02** hour of the day
- **Mean** timing : **9.37** hour of the day
- Mean is to the right of the median, and hence we have a **right skewed distribution** as clearly visible in the plot.
- **standard deviation** : **5.7892**
- Mean range for **95% confidence interval** (+- 2 standard deviations) : **0 to 20.9488**

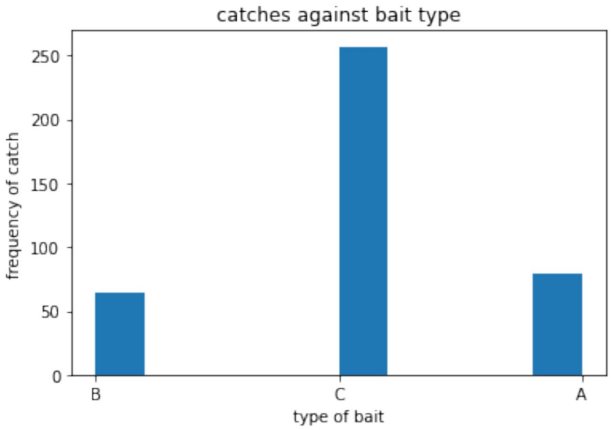
Plot for mapping frequency for various catch sizes



- **Median** catch size : **1.62 kgs**

- **Mean** catch size : **1.67 kgs**
- Mean is to the right of the median, and hence we have a **right skewed distribution** as clearly visible in the plot.
- **standard deviation** : **1.10677**
- Mean range for **95% confidence interval** (+- 2 standard deviations) : **0 to 3.8809**

Plot for mapping frequency of catches against each bait

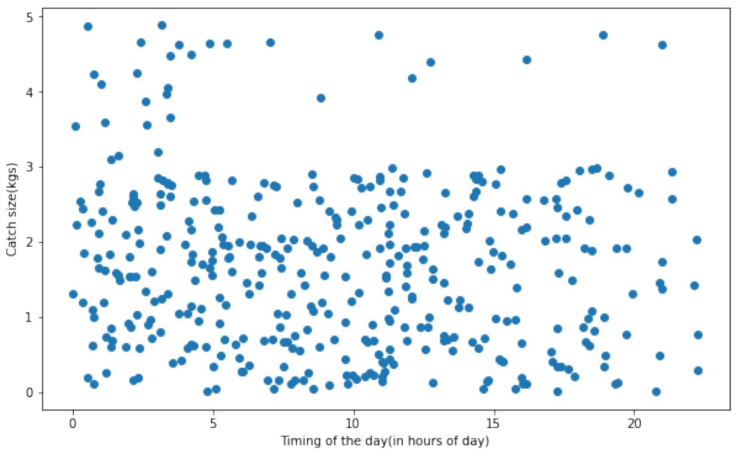


Findings from task 1

1. Early hours of the day yield more fish,and it steadily starts decreasing in the afternoon, from around 1 PM.
2. Most of the catch weighs less than 3 kgs, very few weigh well over it.
3. Bait type C has the most catches associated against it.

Task 2

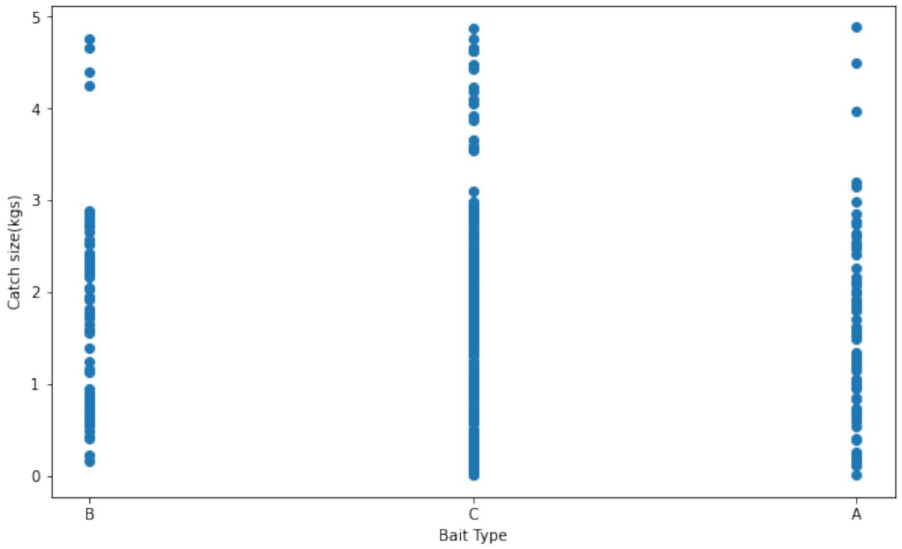
Plot for catch size against time of the day



Spearman's correlation: -0.108

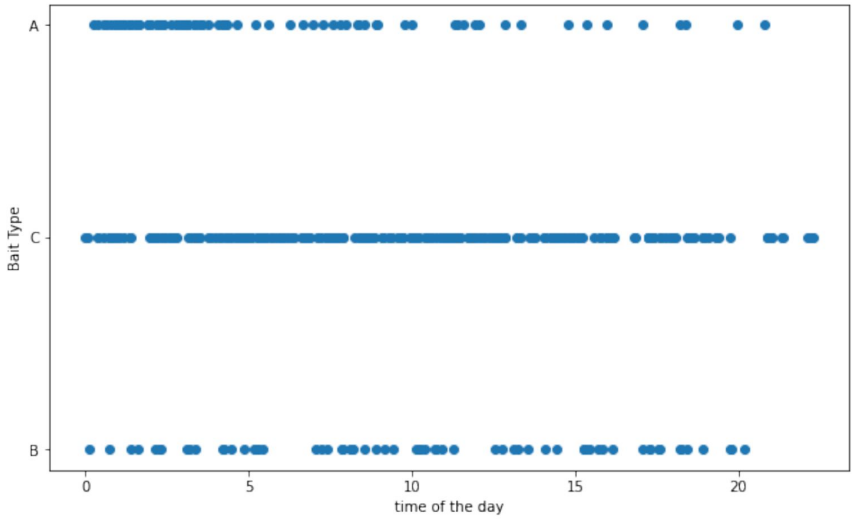
As we can see from the plot above, there is a correlation between the catch size and time of the day, and as its negative, it implies that earlier timings yield larger catch sizes

Plot for catch size against bait type



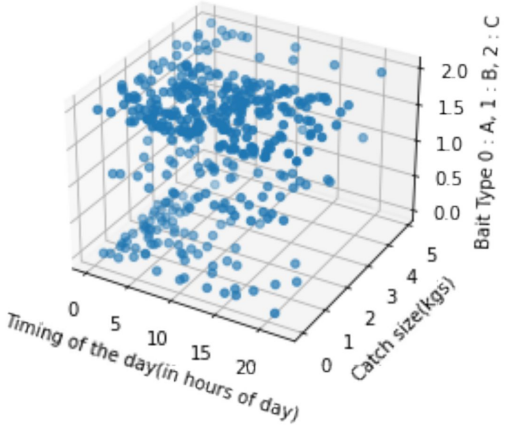
As we can see above, Bait type C is associated with the largest number of catches, and is the most used type of bait. We cannot surely conclude if C is more effective, and hence is used more, or if it has large catches by the sheer frequency of its use as compared to the rest.

Plot for bait type against time of the day



From the above plot, we can infer that bait type A is used more frequently in the early hours of the day, and bait type C is used less frequently in the latter end of the day. Bait type B seems to be used much less frequently, and more distributed, as compared to the other 2.
At 15:00, or 3 PM, bait type C seems to be the more frequently used bait type.

3D scatter plot depicting all the 3 axis, and their relation



Conclusion

1. **Best time** to go fishing is from **2:30AM to 5:00AM**
2. Bait associated with **most catches** : **Bait C**
3. **Best bait** type at 3PM : **Bait C**

Name : Sharang Deepak Gupta
Student ID : 32196946
Code URL : https://github.com/sharang1996/fishing_fleet