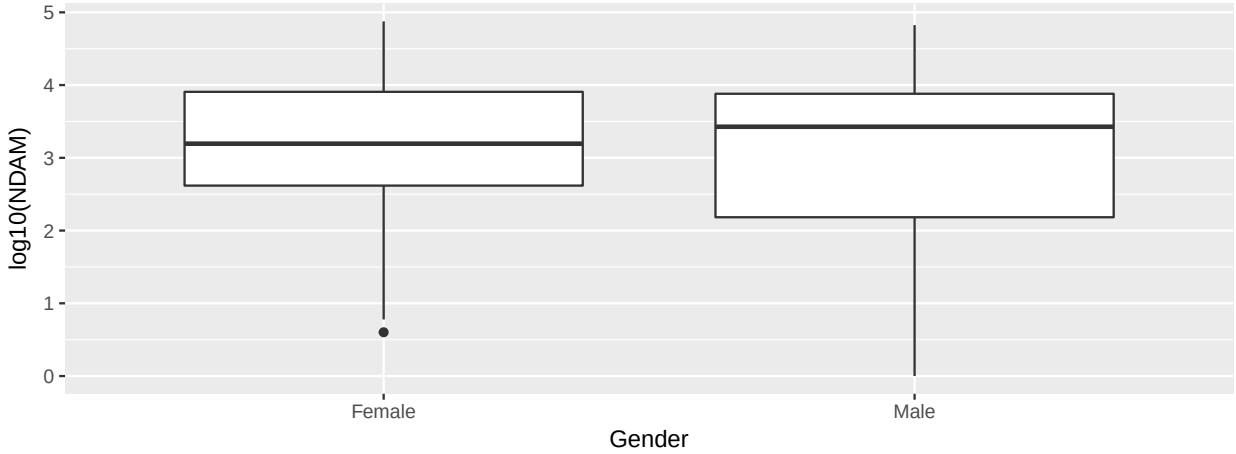
Hurricanes and Himmicanes

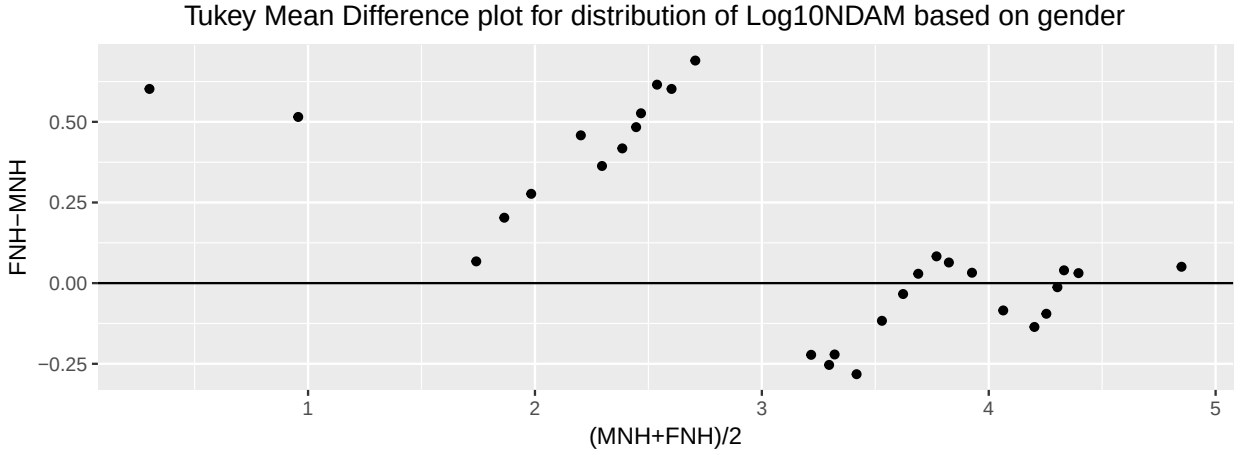
Varun Agarwal, Ajinkya Khamkar, Jivitesh Poojary, Yatin Sharma

Question - Is there a meaningful difference between the distribution of damage caused by hurricanes with female names and the distribution of damage caused by hurricanes with male names?

In order to find if there is some meaningful difference between the distribution of damage caused by hurricanes with female names and the distribution of damage caused by hurricanes with male names, we draw the boxplots of base 10 log of quantitative variable "NDAM" for both genders.

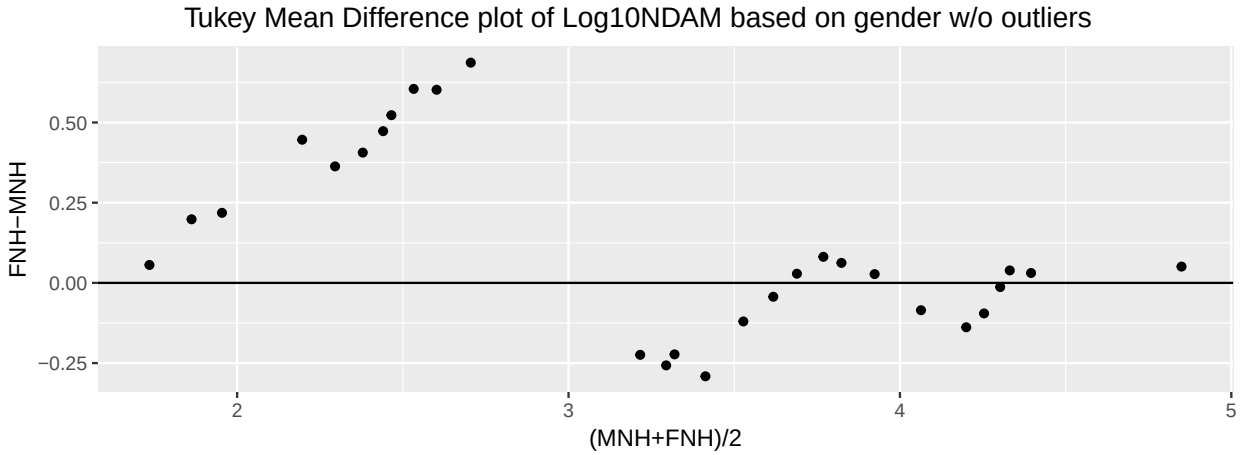


It can be inferred from the above plot that the distribution for males is highly left-skewed, compared to the distribution for females. Also the median for males is higher than the median for females, indicating slightly higher $Log_{10}NDAM$ for males than females. We further try to analyse both the samples by drawing the Tukey mean-difference plot in which on the x-axis goes the mean of the two quantiles(averaging $Log_{10}NDAM$ for Male and Female), while on the y-axis goes the difference between them(Female minus Male.) Going forward, we will refer to female named hurricanes as FNH and Male named hurricanes as MNH.

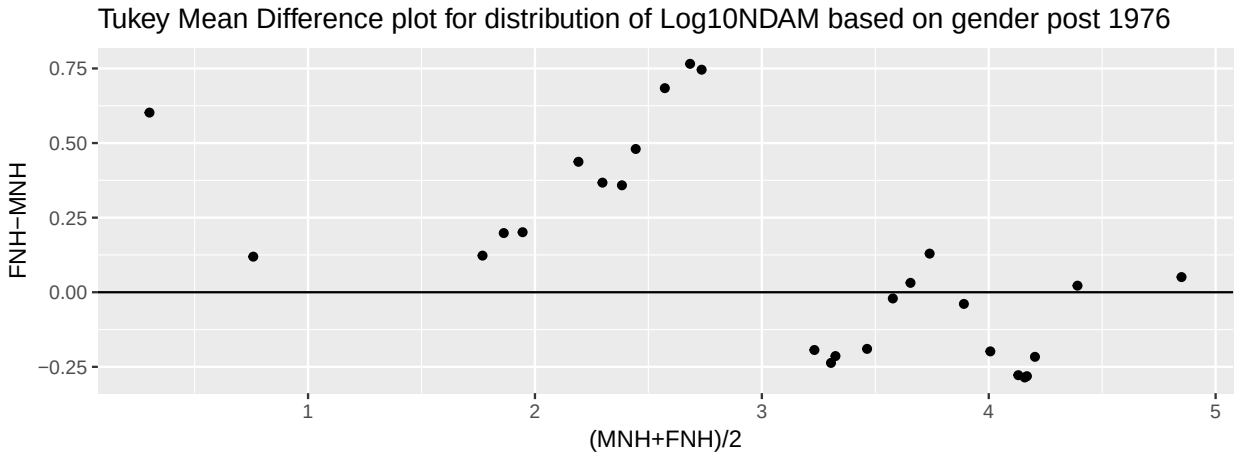


For $(MNH + FNH)/2 < 3$, we see a distinct pattern, indicating that the relationship is neither additive nor multiplicative, but something more complex. Generally, the quantile values for FNH remain higher than MNH. For $(MNH + FNH)/2 > 3$, we see a fluctuating distribution which seems considerably complex to

say anything conclusive. This pattern could be attributed to the natural naming bias in the data or due to the effect of outliers. We will check if the patterns still exist under different modelling conditions
1) Removal of outliers
2) Considering a subset of the data for hurricanes after 1976, since we notice a bias in the naming of hurricanes before this period.



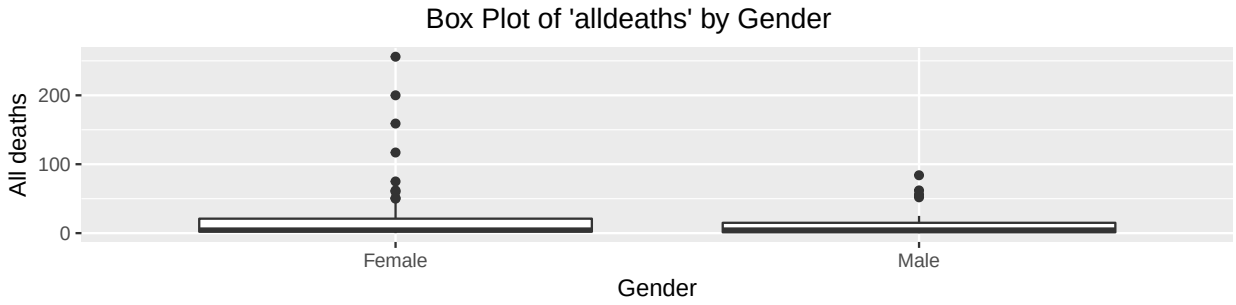
The pattern above looks similar to the first plot, thus we can infer that the left sided outliers do not really have a significant effect on the 2 distinct patterns found in the plot. We now repeat the analysis for the data from 1978 onwards.



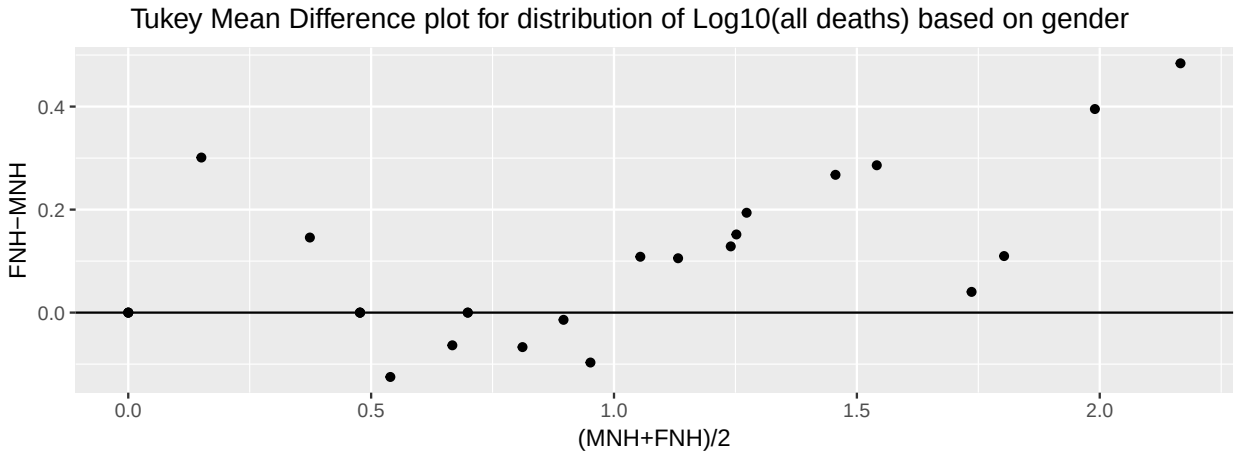
Again, the plot above looks pretty similar to the original plot suggesting that the naming bias in pre 1976 is not responsible for the two distinct patterns in the plot.

Question - Are there any meaningful differences between hurricanes with female names and hurricanes with male names?

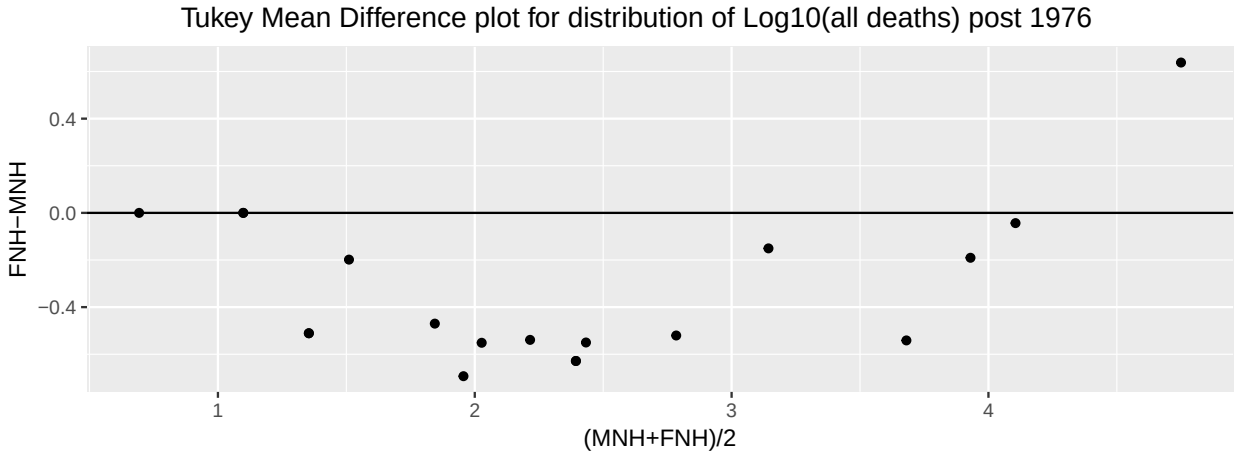
In an attempt to find any other meaningful differences between FNH and MNH, we ty to explore various variables present in the data set. We first try to examine the effect of gender on total number of deaths.



The boxplot above compares the distribution of "alldeaths" variable for FNH and MNH. We see that FNH have a slightly higher variance, and present a good number of high-valued outliers. We further try to analyse both the samples by drawing the Tukey mean-difference plot in which on the x-axis goes the mean of the two quantiles(averaging $Log_{10}alldeaths$ for Male and Female), while on the y-axis goes the difference between them(Female minus Male.)



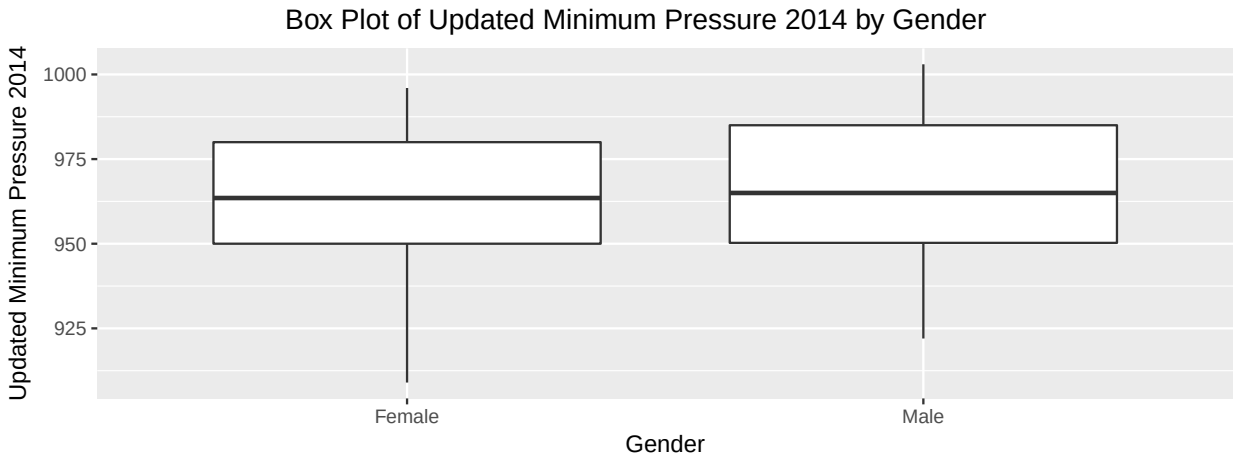
We see that majority of the points are scattered on the positive side of y-axis suggesting that FNH are conducive to higher number of deaths. To confirm, we will further check if the pattern exists under different modelling conditions to come to a conclusion.Since we notice a bias in the naming of hurricanes before this period, we repeat our analysis by considering only the subset of the data from 1976 onwards.



The above graph shows an opposite trend than the one we saw with the originial data. We see that majority of the points lie below the x-axis, indicating that MNH have higher deaths as compared to FNH. The fact that the two graphs are very different shows that different modeling choices lead to completely different

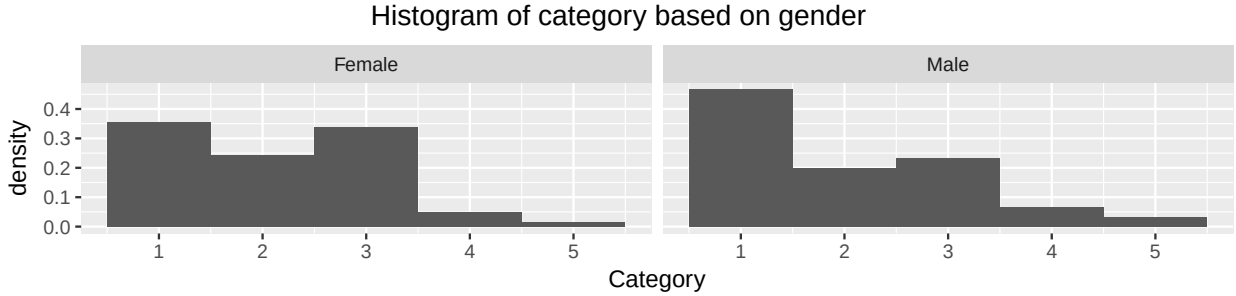
results. Thus we cannot draw any meaningful conclusions about the distribution of deaths, based on gender.

Examining the effect of gender on Updated Minimum Pressure.



The boxplot above compares the distribution of Updated Minimum Pressure variable for FNH and MNH. Both the distributions seem to be left skewed, with the female-named hurricanes having a slightly greater skewness than the male counterpart. We also see that the variance is slightly higher for male-named hurricanes. Both the differences can be attributed to random noise, so we do not observe any significant effect of gender on Updated Minimum Pressure.

Examining the effect of gender on Category.



We compared the density based histograms of hurricane category with respect to gender. MNH have a high probability of Category 1 hurricanes, with subsequent categories having a lower probability. However, in the case of FNH, we observe that they have a high probability of occurrence for Categories 1 to 3, and significantly lower probabilities for categories 4 and 5. However the differences are not significant enough to convey any meaningful difference.