INFERENTIAL STATISTICS

This section presents the results of inferential statistics methods applied on two hypothesis tests namely:

1.) Relationship between fatality and activity
2.) Relationship between fatality and gender

Relationship between fatality and activity

This test was performed to test whether there is a relationship between fatality and activity in other words to see if the activity influences the fatality. To do this we only took the top 10 activities since there are many other activities that are big strings and they also don't make sense.Since both the variables are categorical variables we do the chi-square test to check for dependencies.We made a contingency table using the two variables

Is_Fatal	N	Y	All
Activity_new			
Bathing	75	61	136
Bodyboarding	95	24	119
Boogie boarding	46	2	48
Diving	260	86	346
Fishing	309	63	372
Floating	32	6	38
Kayaking	41	3	44
Playing	20	1	21
Snorkeling	64	8	72
Spearfishing	270	44	314
Standing	122	49	171
Surfing	894	53	947
Swimming	541	346	887
Wading	197	33	230
All	2966	779	3745

1.There is significant relationship between fatality and activity
2.P-value obtained was 4.14 x 10^-69

Relationship between fatality and gender

Here again we perform chi-square test for finding dependencies between the two variables

Is_Fatal	N	Y	All
Sex			
F	362	77	439
M	2600	702	3302
All	2962	779	3741

1.There is no significant relationship between gender and fatality

2.P-Value obtained here was 0.516