We began our investigation by utilizing the p values from investigation A as a quantification of fear. A change in the p value can indicate a change in the bat’s behavior towards the rats. This allowed us to take a step in understanding the change in their behavior. Now that we had a measure of fear, we wanted to know how the p value is changing across the months of the year. We ran the two-sample test for dataset 2 segmented for each of the months. This gave us a p value for each month for rat minutes and rat arrival number. We assimilated this information into the following plots.

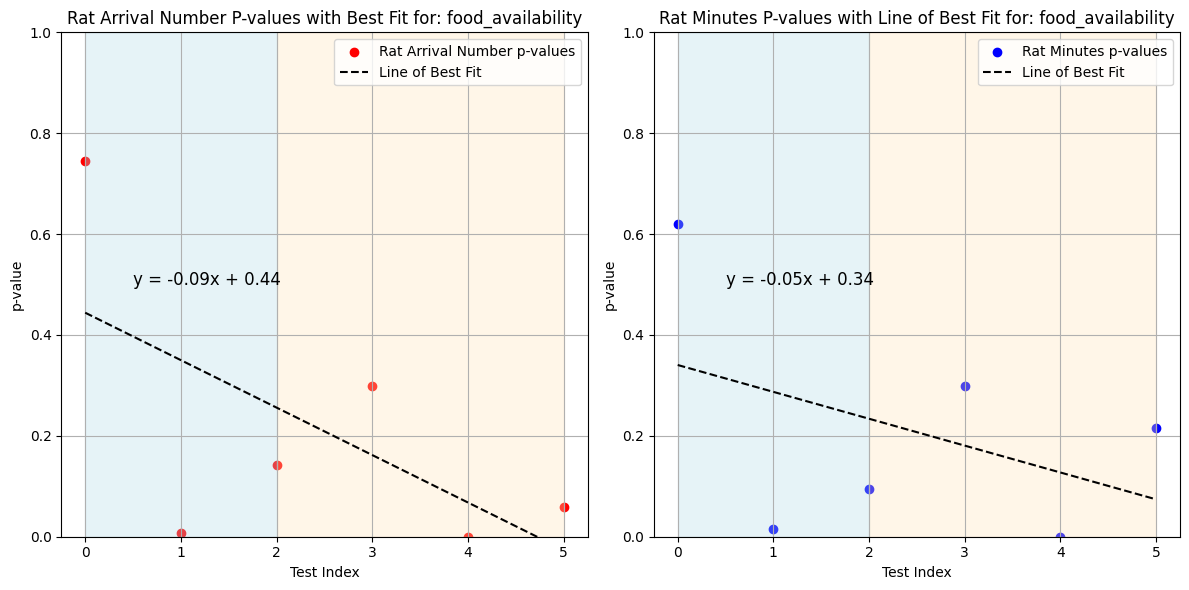


Figure 1: P values plotted for ‘Rat Arrival Number’ and ‘Rat Minutes’ for the test column: ‘food\_availability’

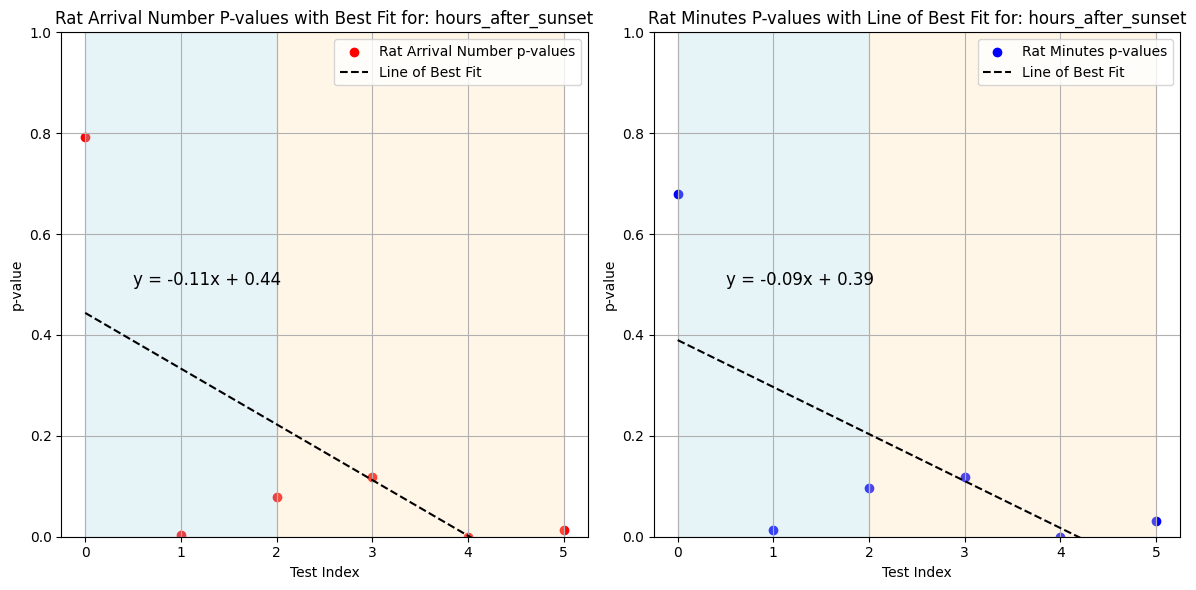


Figure 2: P values plotted for ‘Rat Arrival Number’ and ‘Rat Minutes’ for the test column: ‘hours\_after\_sunset’

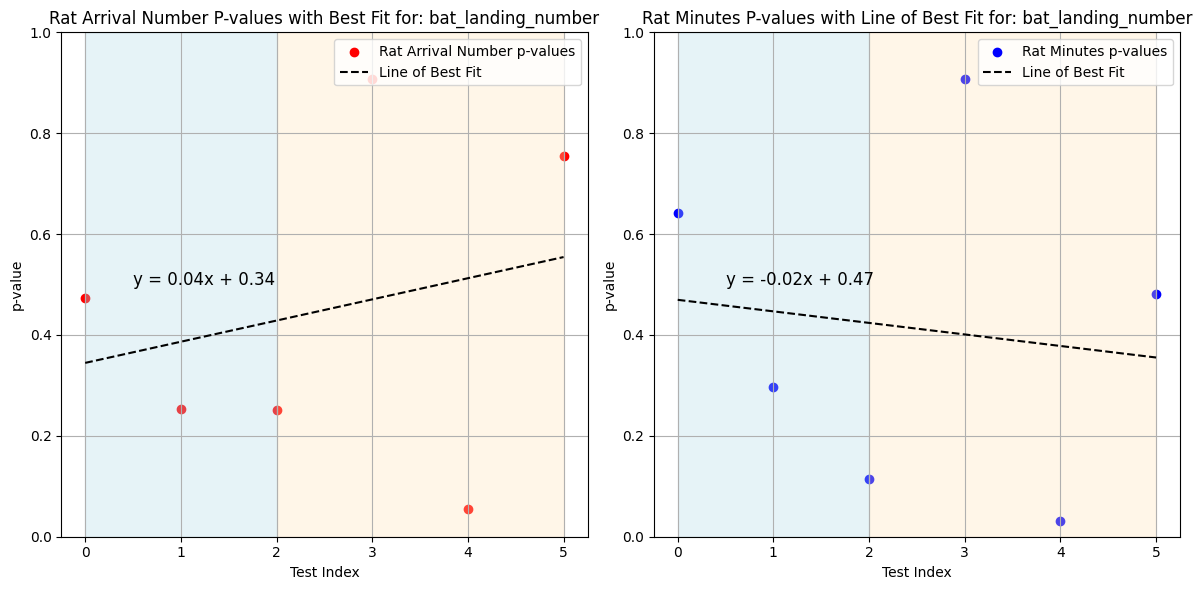


Figure 3: P values plotted for ‘Rat Arrival Number’ and ‘Rat Minutes’ for the test column: ‘bat\_landing\_number’

From figures 1 to 3 above we can see that the p value which we have deemed as a measure of the fear of the bats is changing throughout the months. We have plotted a linear regression line and labelled its equation to show the trend of the data. In all the plots except for the p value plot for rat arrival number for the test column, ‘bat\_landing\_number’, in figure 3, the linear regression line has a negative coefficient. Indicating a general trend of decline for the p value and which indicates a trend of fear increasing across the months.

Another visual cue from the

Due to there being multiple columns in dataset one, we needed to first find an appropriate variable that describes the behavior of the rats. To find this variable, we needed to refer to the one and two samples' tests that were conducted for the purposes of assessment 2.

Table 1: One sample test table from Darwin Group 25 Assessment 2 Presentation

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Column** | **Type of test** | **µ** | **H₀** | **Hₐ** | **Result** |
| ‘Risk’ from Dataset\_1 | Binom test(greater) | 50% | 50% chance of the bat showing risk avoidance/taking behavior | Greater than 50% chance of the bat showing risk avoidance (risk = 0) | Null hypothesis was accepted |
| ‘Reward’ from Dataset\_1 | Binom test(greater) | 50% | 50% chance of the bat’s behavior being rewarding or not. | Greater than 50% chance of the behavior not being rewarding | Null hypothesis was accepted |
| ‘Habit’ from Dataset\_1 split into fear and not fear | Binom test(greater) | 50% | 50% chance of the bat’s behavior showing fear vs not fear | Greater than 50% chance of the behavior showing fear of the rat | Null hypothesis was accepted |

As you can see in Table 1 above, we conducted various one sample tests. This included columns such as habit, risk and reward. We deemed the null hypothesis to be that if the bats are exhibiting behavior of fear, then certain attributes will have a higher than 50% chance of occurring than others. As an example, if the bats are exhibiting fear, there should be a significantly higher than 50% chance of the risk column showing 1 rather than 0. A similar modus operandi can be applied to all the other aforementioned columns.

Table 2: Two sample test table from Darwin Group 25 Assessment 2 Presentation

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Column** | **Split basis** | **H₀** | **Hₐ** | **Result** |
| 'bat\_landing\_number' | rat\_minutes | μ1 = μ2 | μ1≠ μ2 | Null hypothesis was Rejected |
| 'food\_availability' | rat\_minutes | μ1 = μ2 | μ1≠ μ2 | Null hypothesis was Rejected |
| 'hours\_after\_sunset' | rat\_minutes | μ1 = μ2 | μ1≠ μ2 | Null hypothesis was Rejected |
| 'bat\_landing\_number' | rat\_arrival\_number | μ1 = μ2 | μ1≠ μ2 | Null hypothesis was Rejected |
| 'food\_availability' | rat\_arrival\_number | μ1 = μ2 | μ1≠ μ2 | Null hypothesis was Rejected |
| 'hours\_after\_sunset' | rat\_arrival\_number | μ1 = μ2 | μ1≠ μ2 | Null hypothesis was Rejected |

Now let’s analyze the two-sample tests shown above in Table 2. For the two sample tests, we first had to split the data into two samples. We chose to split based on the columns ‘rat\_minutes,’ and ‘rat\_arrival\_number.’ If the columns exhibited a non-zero value, we assumed that to illustrate rat presence and if the columns exhibited a zero value, we associated that with rat absence. After we split the datasets into these two samples, we then conducted the two-sample hypothesis test on the test columns: 'food\_availability', ‘hours\_after\_sunset,’ and 'bat\_landing\_number'. When conducting these two-sample tests, we were trying to ascertain whether the bats are exhibiting fear. We were trying to measure this by looking at the p value from these tests. In all cases we found it to be below 0.05, rejecting our null hypothesis and suggesting that the bats are experiencing fear, which was our conclusion of Investigation A.