**Statistics and Probability Analysis using Excel**

**This is an individual assignment.**

Using Microsoft Excel, create a spreadsheet with a separate tab for each dataset appropriately labeled (dataset1, dataset2, etc.) For best calculation results, enter the data provided below in column format. Each data set is a complete sample. Calculations should use the appropriate Excel formula. Hand or manual calculations will not be given credit.

Within each dataset tab, imbed a textbox and provide a paragraph description of your analysis of the data and the results of each problem. What does the data mean, what is the significance of the statistical analysis, what can be learned from the data/calculations etc.

Upload one single Excel file and name your file yourlastnameLab1.xls

**DATASET1 (data provided in Lab)**

Determine the mean, median and appropriate standard deviation for the selected sample of test scores in Dataset table 1 from last year’s FPST 4333 final exam. Explain your results

**DATASET2 (data provided in Lab)**

In order to acquire the data necessary for further safety analyses, the system safety team has requested that a series of test be run on a brass bushing to be run without lubrication until it seizes (failure). The times to seizure for the sample set are provided in Dataset table 2. Develop a set of statistics for the test and an appropriate histogram graph with no more than 7 groupings (bins) that describe and efficiently communicate the data. Explain your results.

**DATASET3 (data provided in Lab)**

The hour meter reading for all engine failures, which have been reported to operations and maintenance the readings are provided in dataset table 3. Calculate the average failure time and a standard deviation. Develop a Histogram chart with 5-7 categories to show the data graphically. Based on the graph, recommend a time for overhaul to prevent 95% of engine failures in these “critical” systems.

**DATASET4**

Flip a single coin 51 times and record the value. Calculate the percentage of the time each value was shown. How did this compare with expected values?

**DATASET5 (data provided in Lab)**

For the data provided in DATASET Table 5, develop a scatter plot and determine a regression slope equation that approximates the relationship.

|  |  |
| --- | --- |
| **Protection Factor** | **Multiplier** |
| Elimination | 0.1 |
| Substitution | 0.4 |
| Engineering - Multiple | 0.6 |
| Engineering - Single | 0.7 |
| Adm Control  Warning  PPE | 0.8 |

**DATASET6**

Think of at least ten hazards that you have observed or identified in your recent past. These items could be conditions that you observed on Lab Tours, Internships, or just simply observations in daily life. They can be personal safety hazards, process safety hazards, fire hazards, environmental hazards etc.

Create an Excel Table and calculate the risk of each item and build a scatter plot depicting the ISO risk curve.

Calculate the Initial Risk Mishap Index (IMRI) and Final Risk Mishap Index FRMI before and after countermeasures have been implemented See Chapter Ch6, Section 6.6 in your book.

Risk = Probability x Severity x Protection Factor

Use the following columns for your excel calculation sheet.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Hazard Description | Probability | Severity | Existing Countermeasures | PF | IRMI | Proposed Countermeasures  Corrective Actions | PF | FRMI |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |

**DATASET7 (data provided in Lab)**

A manufacturer claims that their Steel Cable has a breaking strength of 800 lbs. A process reliability engineer tests 30 cables and accumulates data the provided in dataset table 7.

Using a level of significance (α) of 0.05, determine the average breaking strength and, perform a z- test to determine if the manufacture’s claims are supported. Use the appropriate method and statistical terms for hypothesis testing.

**DATASET8 (data provided in Lab)**

An operations engineer claims that the average concentration of Hydrogen Sulfide in the refinery process units are less than the OSHA 10-minute exposure limit of 50 ppm. You sample 15 process units and gather data presented in dataset table 8.

Determine the mean, and standard deviation. With a significance level (α) of 0.10, is there enough evidence to support or refute the operations engineers claim?