The following are the Requirement for Inferential Statistics:

**Business Details:** We conduct Beneficiary satisfaction survey, these surveys are done periodical and sample data is collected. Currently Descriptive statistics are applied on survey data.

**Requirements:**

I as Business Intelligence consultant want to introduce inferential statistics to find various population parameters and other probabilities using sample data.

Deadline: 30th May 2020

**Implementation plan:**  We will implement project in incremental approach i.e. we will complete Req No 1 and then move to Req No 2..and so on.

|  |  |  |
| --- | --- | --- |
| Req  no | Requirement Details | Comments |
| 1 | Currently Descriptive statistics are applied on Sample data. I want to introduce Inferential statistics and proof that The Mean of Descriptive Statistics should be avoided and the “The Mean of Sample Means” should be used based CLT Theorem. Based on CLT theorem that the “The mean of sample Means” should be used as **Population Mean**    **Step-1:Data**: Take 300+ sample data of citizen age  **Step-2:Apply Descriptive Statistic:** The Mean of 300+ sample ages is around 41.xx  **Step-3:CLT**: Create n samples from 300+ sample data of citizen age, find “The mean of sample Means”, the result of this exercise should be different Step-2  I did above exercises and the results of Step-2 and Step-3 are equal. |  |

|  |  |
| --- | --- |
| 2 | I want to use sample data i.e.300+ sample data of citizen response to Survey Questions, create n samples with n>30 and generate Histogram using “Sampling Distribution of Sample Means”  Below link has proof using Exponential distribution, but I need to proof using Discrete data i.e. survey Answers by Citizens  <https://www.youtube.com/watch?v=Pujol1yC1_A> |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 3 | I want to create Hypothesis Test for 300+ sample data of citizen response to Survey Questions. Survey Answers are rated from 1 to 5 with below Satisfaction levels   |  |  | | --- | --- | | satisfaction | satisfaction\_text | | 1 | Dissatisfied at all | | 2 | Somewhat dissatisfied | | 3 | Neutral | | 4 | Somewhat satisfied | | 5 | Completely satisfied |   **Null Hypothesis:**   1. **Is the Population satisfaction level is Equal to “Somewhat Satisfied” .**   **Alternate Hypothesis:**   1. **Population satisfaction level is NOT Equal to “Somewhat Satisfied” .**   **Technical:**   1. **μ0: μ =4** 2. **μ1: μ # 4** 3. **α = 0.05 ( 5% Significance Level)** |

|  |  |
| --- | --- |
| 4 | I need Help from Expert Statistician for following as Online meeting or through emails   1. May ask for help for How to solve specific problem or how it should be done ? 2. May ask questions on Statistics? 3. May ask to validate my finding after applying specific statistical approach? 4. May ask any other question on Inferential Statistic applications on sample survey data |