

## Project Milestone Applied Statistics

The project milestone document will help you to understand about the topics that needs to be covered to complete the project as per the questions.

Question No	Question	Marks	Week	Topic
Part 1 Q1 A	Refer above table and find the joint probability of the people who planned to purchase and actually placed an order.	1 pts	Week 2	Joint Probability, Marginal Probability
Part 1 Q1 B	Refer to the above table and find the joint probability of the people who planned to purchase and actually placed an order, given that people planned to purchase.	1 pts	Week 2	Joint Probability, Marginal Probability
Part 1 Q2 A	Probability that none of the items are defective?	1 pts	Week 2	Probability distribution, Binomial distribution
Part 1 Q2 B	Probability that exactly one of the items is defective?	1 pts	Week 2	Probability distribution, Binomial distribution
Part 1 Q2	Probability that two or fewer of the items are defective?	1 pts	Week 2	Probability distribution, Binomial distribution
Part 1 Q2	Probability that three or more of the items are defective?	1 pts	Week 2	Probability distribution, Binomial distribution
Part 1 Q3	What is Probability that in a given week he will sell some cars?	1 pts	Week 2	Poisson distribution
Part 1 Q3	What is Probability that in a given week he will sell 2 or more but less than 5 cars?	1 pts	Week 2	Poisson distribution
Part 1 Q3	Plot the Poisson distribution function for cumulative probability of cars sold per-week vs number of cars sold per week	1 pts	Week 2	Poisson distribution
Part 1 Q4	What is the probability that all three orders will be recognized correctly?	1 pts	Week 2	Probability distribution, Binomial distribution
Part 1 Q4	What is the probability that none of the three orders will be recognized correctly?	1 pts	Week 2	Probability distribution, Binomial distribution
Part 1 Q4	What is the probability that at least two of the three orders will be recognized correctly?	1 pts	Week 2	Probability distribution, Binomial distribution
Part 1 Q5	Explain 1 real life industry scenario (other than the ones mentioned above) where you can use the concepts learnt in this module of Applied Statistics to get data driven business solution.	3 pts	Week 1 & 2	Mix of both weeks. Scenario based
Part 2 Q1	Read the data set, clean the data and prepare final dataset to be used for analysis.	10 pts	Common Part	Common Part

Part 2 Q2	Perform detailed statistical analysis and EDA using univariate, bi-variate and multivariate EDA techniques to get data driven insights on recommending which teams they can approach which will be a deal win for them. Also, as a data and statistics expert you have to develop a detailed performance report using this data.	10 pts	All Weeks Python Module AS Week 1	EDA and Visualizations
Part 2 Q3	Please include any improvements or suggestions to the association management on quality, quantity, variety, velocity, veracity etc. on the data points collected by the association to perform a better data analysis in future. At-least 1 suggestion for each point.	10 pts	Textual Answer Expected	Textual Answer Expected
Part 3 Q2	Data exploration. A. Check the datatypes of each attribute. B. Check for null values in the attributes.	1 pts	Common Part	Common Part
Part 3 Q3 A	Drop the null values	1 pts	Python, EDA	Python, EDA
Part 3 Q3 C	Plot box plot for funds in million	1 pts	Python, EDA	Python, EDA
Part 3 Q3 D	Check the number of outliers greater than the upper fence	1 pts	Python, EDA	Python, EDA
Part 3 Q3 E	Check frequency of the Operating State features classes	1 pts	Python, EDA	Python, EDA
Part 3 Q4 A	Is there any significant difference between Funds raised by companies that are still operating vs companies that closed down?	1 pts	Week 3	Hypothesis
Part 3 Q4 B	Write the null hypothesis and alternative hypothesis	1 pts	Week 3	Hypothesis
Part 3 Q4 C	Test for significance and conclusion	1 pts	Week 3	Hypothesis
Part 3 Q4 D	Make a copy of the original data frame	1 pts	Python, EDA	Python, EDA
Part 3 Q4 E	Check frequency distribution of Result variables.	1 pts	Python, EDA	Python, EDA
Part 3 Q4 F	Calculate percentage of winners that are still operating and percentage of contestants that are still operating	1 pts	Week 3	Hypothesis
Part 3 Q4 G	Write your hypothesis comparing the proportion of companies that are operating between winners and contestants:	2 pts	Week 3	Hypothesis
Part 3 Q4 H	Test for significance and conclusion	1 pts	Week 3	Hypothesis
Part 3 Q4 I	Select only the Event that has 'disrupt' keyword from 2013 onwards.	1 pts	Python, EDA	Python, EDA