Exercise:

Please look at the attached data (MockDataSet1.xlsx). This shows an illustrative sample of quotation requests received by an insurer over one month, for life insurance products.

The insurer profit is the NBC, if the Sold flag is "Y" and nothing otherwise.

1. What can you find what is of interest in the data?
2. What are the main drivers of profit and why?
3. Can you build a model to predict the probability of sale?
4. Can you segment this dataset into specific customer groups?
5. How would you build an automated pipeline to ensure data quality as, over time, new customers request new quotations and purchase life insurance products from the insurer?

Ahead of the interview, please send a presentation outlining your findings, and answers to all questions.

Also, please provide all code and analysis that you wrote/developed to answer the questions. Ideally, code should be written in Python.

Thank you, and good luck!

Version 2.0

Jan-25