Business Case: ACME S.A.

ACME S.A. is a personal care company which manufactures and commercializes different products like shampoo, soap, toothpaste, and others among all the country. This company has different manufacture units which are responsible to produce all the products for the company. However, in last months, some shortages appeared. One hypothesis that explains the cause of shortages is that promotional activities are generating huge variances in demand. Due to promotion effects were not considering in demand planning process, the factories couldn't be prepared to support that additional volume.

Despite this hypothesis is logical, there is some skepticism about it. For that reason, the company has decided to hire you as a Data Scientist to help them to resolve this problem and has a trustable demand planning process.

The leadership team of the company (conformed by managers and directors who don't have a lot of experience in statistics neither in machine learning techniques) needs that you help to answer the following questions:

1) *What is the average impact generated by promotions over the total sales? Is that impact general for all brands and customers?*
2) *Considering the approved promotions for last quarter of 2018, how many tons of products we are going to sell in last quarter (Q4) of 2018?*
3) *For the Q4 2018 forecast, how many tons comes from promotions and how many from "usual" behavior of the series?*

Additionally, the company is working with the British Software Consulting Group to implement an analytical layer on cloud called Databricks. This layer will be the one used for processing data using a language like Python (technically called Pyspark). The leadership is concerned about the costs that the new platform can generate, because the company is going to pay by every second used to process a bite of information. For that reason, they have an extra question for you:

4) *Which strategies are you going to apply to create models, retrain them, and/or debugging them without generating huge costs?*

Instructions:

Please prepare 1 slide to introduce yourself and 1 slide to answer each one of the 4 questions. All slides must be redacted in English, and you must use Python to resolve the case.

Remember you have 30 min to present your resolution

Data:

Attached to this case you will find two excel files with all the data that you require.
Inside those documents, there is a sheet where each variable is detailed.
Please feel free to introduce any assumption that you consider needed.
Also, if in any point you must assume something, please do it!