Sprocket Central Pty Ltd , a medium size bikes & cycling accessories organisation, has approached Tony Smith (Partner) in KPMG’s Lighthouse & Innovation Team. Sprocket Central Pty Ltd  is keen to learn more about KPMG’s expertise in its Analytics, Information & Modelling team. 

Smith discusses KPMG’s expertise in this space (you can read more [**here**](https://home.kpmg/au/en/home/services/advisory/management-consulting/digital/data-analytics-modelling.html)). In particular, he speaks about how the team can effectively analyse the datasets to help Sprocket Central Pty Ltd grow its business.

Primarily, Sprocket Central Pty Ltd needs help with its customer and transactions data. The organisation has a large dataset relating to its customers, but their team is unsure how to effectively analyse it to help optimise its marketing strategy. 

However, in order to support the analysis, you speak to the Associate Director for some ideas and she advised that “*the importance of optimising the quality of customer datasets cannot be underestimated. The better the quality of the dataset, the better chance you will be able to use it drive company growth*.”

The client provided KPMG with 3 datasets:

* Customer Demographic
* Customer Addresses
* Transactions data in the past 3 months

You decide to start the preliminary data exploration and identify ways to improve the quality of Sprocket Central Pty Ltd’s data.

*Hi there – Welcome again to the team! The client has asked our team to assess the quality of their data; as well as make recommendations on ways to clean the underlying data and mitigate these issues.  Can you please take a look at the datasets we’ve received and draft an email to them identifying the data quality issues and how this may impact our analysis going forward?*

*I will send through an example of a typical data quality framework that can be used as a guide. Remember to consider the join keys between the tables too. Thanks again for your help*.”

[Read email below]

*Hi there,*

*As per voicemail, please find the 3 datasets attached from Sprocket Central Pty Ltd:*

* *Customer Demographic*
* *Customer Addresses*
* *Transaction data in the past three months*

*Can you please review the data quality to ensure that it is ready for our analysis in phase two. Remember to take note of any assumptions or issues we need to go back to the client on. As well as recommendations going forward to mitigate current data quality concerns.*

*I’ve also attached a data quality framework as a guideline. Let me know if you have any questions.*

*Thanks for your help.*

*Kind Regards  
Your Manager*

For context, Sprocket Central Pty Ltd is a long-standing KPMG client who specialises in high-quality bikes and accessible cycling accessories to riders. Their marketing team is looking to boost business by analysing their existing customer dataset to determine customer trends and behaviour.

we will be talking about identifying which of the 1000 customers Sprocket Central Pty Ltd should target, based on the existing 3 datasets (Customer demographic, customer address and transactions). by Explaining the three phases: Data Exploration; Model Development, and Interpretation.

Using the existing 3 datasets (Customer demographic, customer address and transactions) as a labelled dataset, please recommend which of these 1000 new customers should be targeted to drive the most value for the organisation.

In building this recommendation, we need to start with a PowerPoint presentation that outlines the approach which we will be taking. The client has agreed on a 3-week scope with the following 3 phases as follows - Data Exploration; Model Development and Interpretation.

Prepare a detailed approach for completing the analysis including activities – i.e. understanding the data distributions, feature engineering, data transformations, modeling, results from interpretation, and reporting. This detailed plan needs to be presented to the client to get a sign-off. Please advise what steps you would take.

Please ensure your PowerPoint presentation includes a detailed approach for our strategy behind each of the 3 phases including activities involved in each - i.e. understanding the data distributions, feature engineering, data transformations, modelling, results from interpretation and reporting. This detailed plan needs to be presented to the client to get a sign-off.