Dear KPMG Team,

This is Arundhati Das from KPMG Data Analytics (Virtual Internship) team. Thank you for providing us with the three datasets from Sprocket Central Pty Ltd. The below table highlights the summary statistics from the three datasets received. Please let me know if the figures are not aligned with your understanding.

The following are the details of analysis done on the dataset:

|  |  |  |  |
| --- | --- | --- | --- |
| **Table Name** | **Table Records** | | **Table Analysis** |
|  | **Before Data Cleaning** | **After Data Cleaning** |  |
| Transaction Data | 20000 rows & 13 columns  (1542 blank cells) | 19445 rows & 14 columns  (0 blank cell) | * Total profit: $10,930,284 (app.) * ‘Solex’ is the most purchased brand name * The most and least sold product line is ‘Standard’ and ‘Mountain’ respectively |
| New Customer List | 1000 rows & 18 columns  (152 cells) | 878 rows & 18 columns  (0 blank cell) | * Most new customers are from the New South Wales, Australia * Most customers own cars |
| Customer Address | 3999 rows & 6 columns  (0 blank cell) | 3999 rows & 6 columns  (0 blank cell) | * Most customers are from New Sales Wales (NSW) * Most customers have post code between 2000 to 2190 |

# Kind regards,

# Arundhati Das