Case Study: Data Prep and Cluster Analysis

Company XYZ wants to use advanced data analytics to cluster its customer base around their usage. The analysis will allow XYZ's CEM team to make more informed business decision in the future based on the usage patterns of their client base. XYZ has provided you with large data samples with over 500 usage variables and survey answers.

XYZ has sought your assistance in developing, testing, and performing the data analytics models. As part of the work, you will complete the initial data discovery and preparation, use advanced statistical modelling to answer the business questions at hand and report the findings in a format easy to communicate across different business units.

Task: Create usage-based segmentation

1. Create at least 8 clusters.

2. Only use input variables related to how much people use each service, in what way they use it, etc.

   - Do not use variables on demographics, product holding, customer satisfaction, etc. <u>Only usage</u>

   - Use a combination of usage variables – in general, use enough variables (as you see fit) from each of the following **4 dimensions** – internet usage, fixed phone usage, mobile usage (data, minutes, etc), television usage (viewing of different channels, video on demand, digital and linear tv, etc.), over-the-top usage (streaming/p2p usage).

3. **Describe each segment** – how are the segments different beyond their usage – NPS, demographics, revenue, perception KPIs, ecosystem.

   - Find a way to present of all this in a business way – short, crisp and non-technical language.

There is no one correct solution! It's all about the process.