**Clustering Analysis Results & Findings**

**Overview**

Using the Elbow Effect, the optimal number of clusters for the data is k = 4, as the WCCS has a drastically lower number at k = 4 compared to adding k clusters will produce diminishing outputs. To visualize this customer segmentation, PCA method was applied to get the clusters and it also confirmed overlapping customer segments. The data can then be used to segment customers with similar – yet individuals attributes for tailored marketing and retention strategies. Most common behaviours observed The clusters identified give you insights on customer behaviour thus helping your drive data driven decisions for customer engagement targeted.

**Finding Optimal k with Elbow Method**

The Elbow method is a widely used method in choosing the best number of cluster (k) in K-means clustering This graph plots Sum of Squared Errors (WCCS) for different values of clusters (k) to visually inspect where the optimal number of clusters lies.

A graph of a number of clusters

AI-generated content may be incorrect.

**Explanation:**

* **Sum of Squared Errors ( WCCS ),** also called Inertia, measures how tightly packed the clusters are. It calculates the sum of the squares of the distance of each point from the centroid of the cluster it belongs to. K-Means aims to minimize this sum which gives rise to tight, well-separated clusters.
* The number of clusters (k) is shown through the X-axis. In our example, we looked at values of k between 1 and 10.
* The graph for each k value, [Image: Y-axis is WCCS]

**Analysis:**

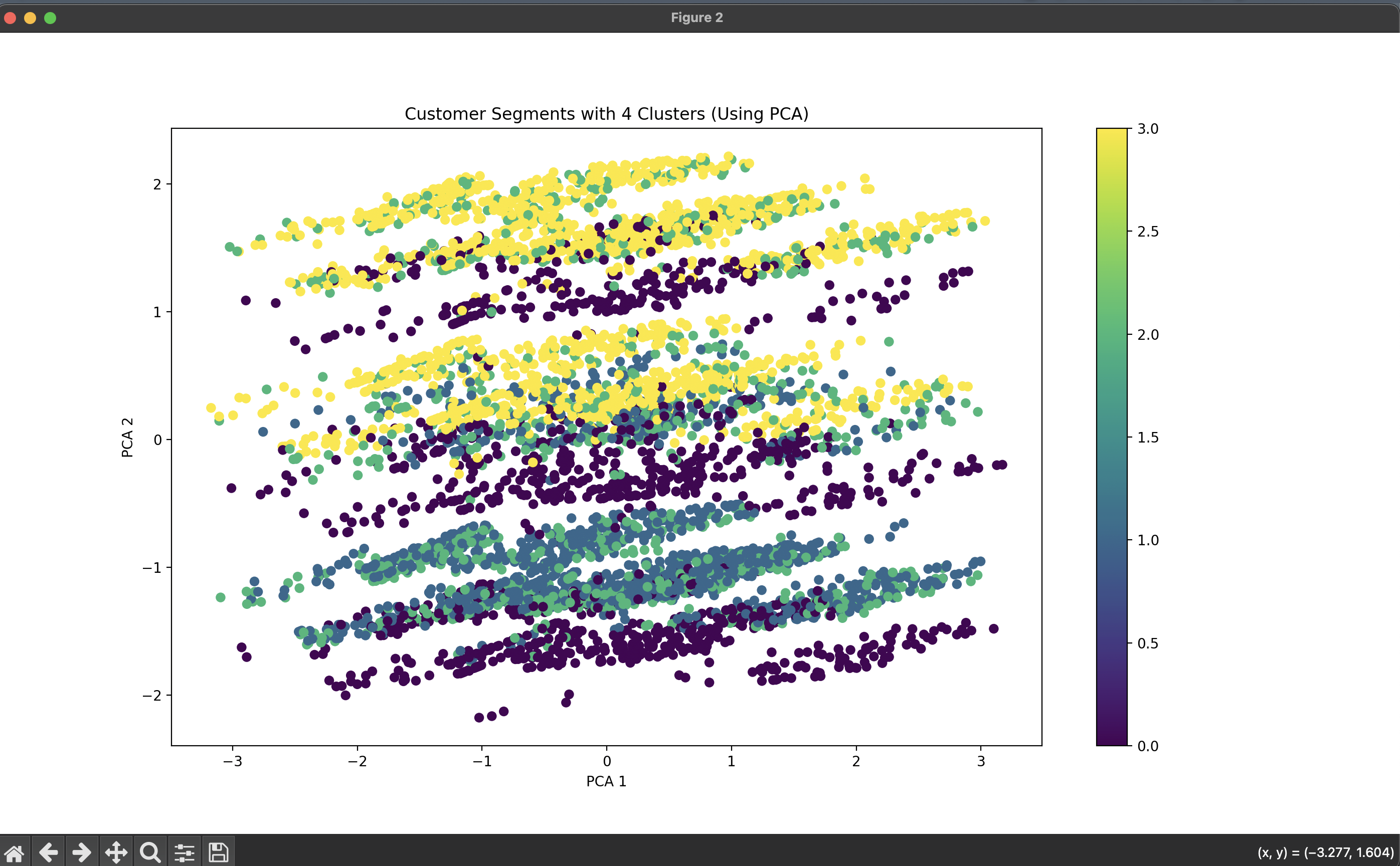
* k = 1: The maximum WCCS is 1 a very tight cluster, because the data points are imposed into one cluster.
* As k grows, the WCCS shrinks since data is clustered in more groups and points are closer to their centroid.
* We see the elbow at  k = 4. This is the point where the WCCS curve "bends" or "levels off," signifying diminishing returns in reducing WCCS with additional clusters. Beyond this point, the addition of further clusters yields little gain in reducing the WCCS, indicating that the data is adequately divided into 4 clusters**.**

**Why k = 4 is Optimal:**

* Well, elbow level goes at k=4, this tells us that there are at least this many clusters after which it does not make sense to add a new cluster as our variance inside a cluster is decreasing but at the same time it is also making our model with too many clusters which overfits the model itself.
* After the elbow point (k = 5 or greater) curve goes down adaptively, downward curvature area is down, the WCCS value continues to decrease  but the degree of decrease is reduced, the corresponding clustering solution may reach unnecessary complex.

**Clusters per Customer Segment (Using PCA)**

This second graph is a visualization of K-Means clustering results. It applies Principal Component Analysis (PCA), a dimensionality reduction technique, to reduce the high dimensional data (with many features) to two dimensions to facilitate interpretation and visualization.



**A Brief Breakdown:**

* PCA identifies principal components — directions (or axes) in the data that explain maximum variance and allows us to reduce the dimensions in dataset. Since original dataset may have many features, it allows you to visualize the data on a two-dimensional plane.
* Each dot on the scatter plot represents one customer from the dataset.
* The colors represent the 4 clusters identified by K-Means, with each color corresponding to a different cluster label. For instance, yellow points could indicate one group of customers, purple another, etc.
* X-axis and Y-axis depict first two Principal components (PCA 1 and PCA 2) which covers maximum variance in the data. These few features are combinations of the original features but allow us to visualize the complete dataset in two-dimensions.

**Cluster Separation:**

According to the elbow method, the data points get classified into 4 clusters. Segments

**Color-Coded**: each of the clusters is color-coded (purple, blue, green and yellow). The clear separation between some of these clusters (e.g. purple vs green) means that the K-Means algorithm could separate segments of customers.

**Cluster Spread**: Some clusters are tighter (purple), others (like yellow) are more dispersed. Here, we conclude that some customer sets may have more similar attributes, whilst others vary to greater extents.

**Overlap**: The clusters overlap, which is expected as PCA reduces the dimensions of the data, leading to some information loss But specifically, the K-Means algorithm is indeed working in a higher dimensional space, which makes separation between clusters more clear.

**Insights:**

**Customer Segments**: 4 clusters will signify different customer base of the model (e.g. Customer Demographics, Account + Service usage patterns as per data used in model). One cluster, for example, may describe long-term customers who churn less than average, while another may be an average customer that churns frequently. Understanding which features are most significant in determining these clusters will help you understand the differentiation between segments.

**Business Application**: After you identify the characteristics of each cluster, marketing strategies, service offers, and customer retention strategies can be customized for each segmentation. For instance: Customers that fall in the yellow cluster (which may be a high-risk to churn) can be targeted for retention campaigns. For example, consumers in the purple cluster (maybe representing loyal consumers) can be offered loyalty rewards.

**Connecting Both Graphs:**

**Elbow method & optimal clustering**: From the elbow graph, it is evident that 4 clusters is the optimal choice for this dataset, where there is a good balance between simplicity and accuracy.

**PCA Visualization**: The PCA visual confirms that our clustering solution segments the customers into 4 distinct clusters, with varying degrees of variance and overlap. So, it indicates K-Means clustering is a success in segmentation of the customers based on the patterns present in the data.

**Potential Next Steps:**

**Feature Importance:** Determine which features (e.g. Monthly Charges, Tenure, etc.) are responsible for separating the clusters. Such as, when you study the behavioral patterns of customers in each segment, this will allow you to better understand their features.

**Cluster Profiling:** Profile or describe each cluster based on the key differentiating characteristics. Example:

* Segment 1: “Big and Loyal Customers”
* Segment 2: “Value-Short Term Customers”

**Business Strategy:** Use the groups of customers to tailor business strategy in your marketing, product design, and retention strategies.