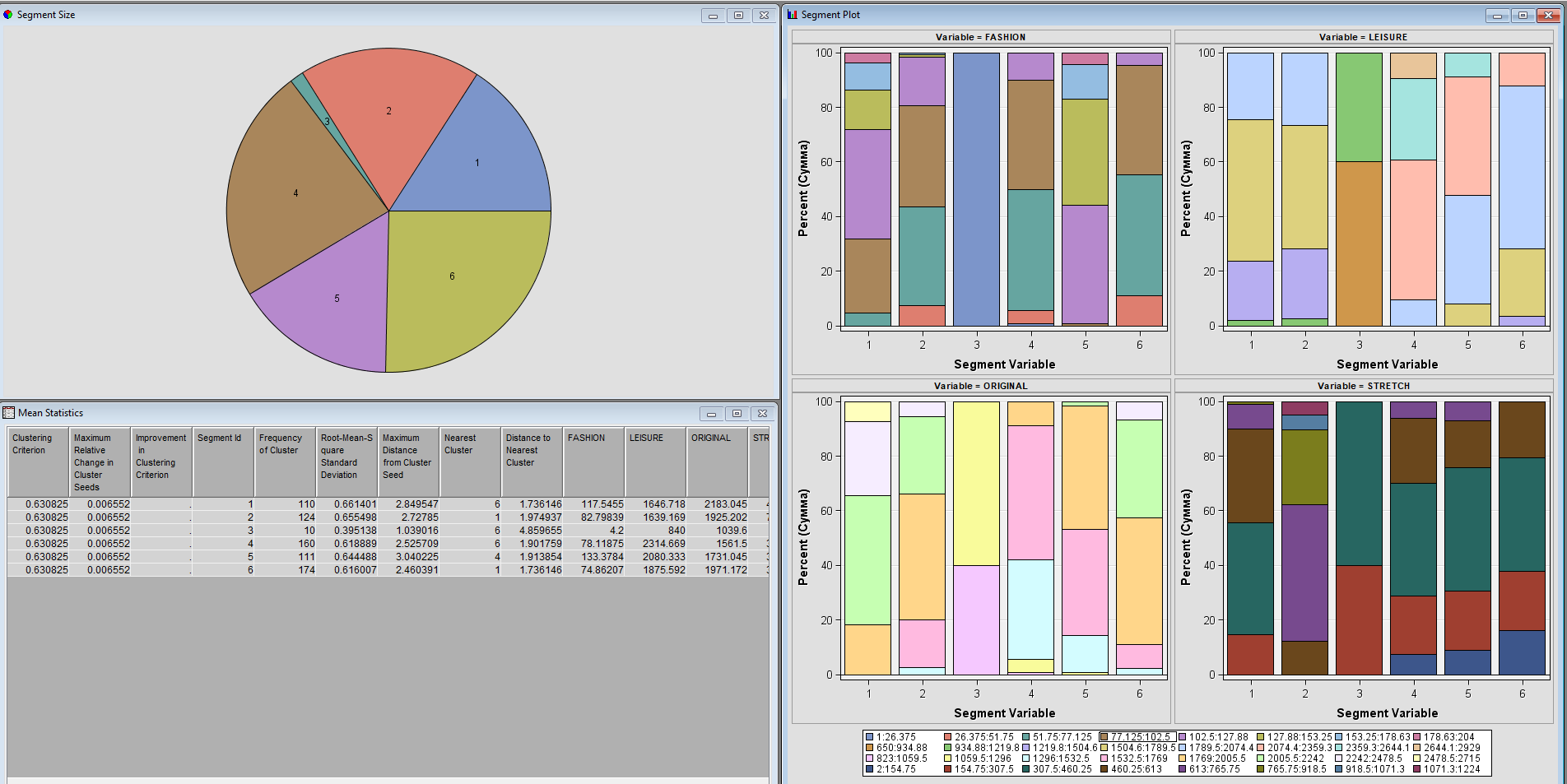
**SAS**

Результаты кластеризации в SAS.

Столбец SALESTOT, содержащий сумму FASHION + LEISURE + STRETCH + ORIGINAL, я удалил. Была проведена нормализация данных.

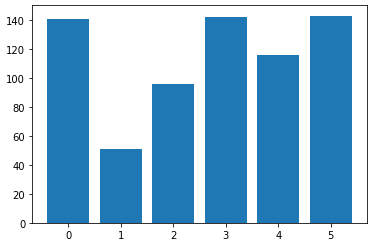
Кластеры в основном получились правдоподобными. При этом кластер 3 состоит всего из 10 элементов (всего датасет содержит 689 строк). Это – недостаток модели, так как кластер, состоящий из 10 элементов, нерепрезентативен.



**Python**

Ссылка на Jupyter Notebook:

<https://drive.google.com/file/d/1GmWzL60SytuUYHShowIUWaaTykWV18KY/view?usp=sharing>

Данные я отмасштабировал с помощью StandardScaler, для кластеризации использовал KMeans.

Распределение по кластерам:

2 159

4 144

1 143

3 105

0 87

5 51

Количество элементов в кластерах получилось в целом более правдоподобным, чем в SAS.