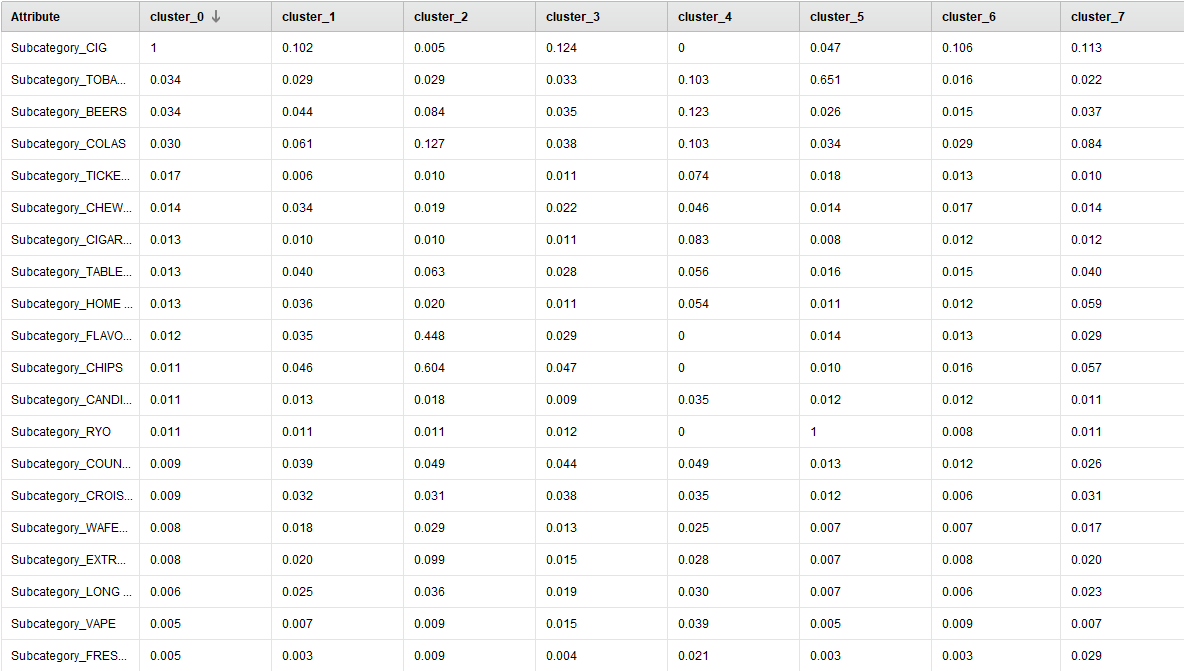
**ΑΡΧΕΙΟ ΑΝΑΛΥΣΗΣ**

Στο αρχείο zip παρατίθενται 11 αρχεία ipynb (τύπου jyputer notebook, που αξιοποιούν την python) τα οποία περιέχουν τις αναλύσεις και τα αποτελέσματα αυτών, που χρησιμοποιήθηκαν για την παρουσίαση αλλά και αποτελέσματα που αποτελούν επιπλέον χρήσιμη γνώση για την εταιρεία. Συγκεκριμένα, το σημειωματάριο ‘**basket\_time\_analysis.ipynb**’ αποκαλύπτει τις περιόδους αιχμής των πωλήσεων και στις επιδόσεις των κατηγοριών προϊόντων που ορίσαμε από το clustering των υποκατηγορίες των προϊόντων του Dataset 2. Το σημειωματάριο ‘**cig\_analysis.ipynb**’ κάνει το ίδιο με προηγούμενο notebook, απλά αυτή τη φορά για 4 πρώτες, ως προς το Revenue που αποφέρουν, κατηγορίες. Το σημειωματάριο ‘**cleaning\_dataset2.ipynb**’ προετοιμάζει τα δεδομένα από το Dataset 2 για δυο κατηγορίες clustering, για Category και Subcategory clustering των προϊόντων και βρίσκει τον καλύτερο αριθμό clusters για κάθε επιλογή. Στη συνέχεια, το clustering πραγματοποιείται στο rapidminer και για τα 2 clusters, όπου επιλέξαμε να εστιάσουμε στο clustering των υποκατηγοριών και αποθηκεύσαμε τα αποτελέσματα του rapidminer στο αρχείο ‘rapidminer\_clust\_data2.csv’. Στιγμιότυπο από τα αποτελέσματα του rapidminer :



Το σημειωματάριο ‘**cluster\_average\_price\_category.ipynb**’ αξιοποιεί τα έτοιμα επεξεργασμένα, μέσω χρήσης excel, δεδομένα από το Dataset 1 και πραγματοποιεί clustering με βάση την τιμή στα μαγαζιά και μετά πραγματοποιεί χρήσιμους στατιστικούς ελέγχους πάνω στα αποτελέσματα του clustering. Το σημειωματάριο ‘**data1analysis.ipynb**’ πραγματοποιεί ανάλυση για την απάντηση των πρώτων δύο ερωτημάτων που παρείχε η Intale σχετικά με το Dataset 1. Το σημειωματάριο ‘**clustering.ipynb**’ πραγματοποιεί επαλήθευση του clustering στις υποκατηγορίες των προϊόντων και παρέχει ανάλυση για την βέλτιστη επιλογή αριθμό clusters. Το σημειωματάριο ‘**revenue\_analysis\_dataset1.ipynb**’ παρέχει ανάλυση του revenue για τα μαγαζιά ως προς πολλές παραμέτρους (ανά τύπο μαγαζιού, ανά περιοχή, ανά περιοχή και ανά τύπο μαγαζιού μαζί). Το σημειωματάριο ‘**specialized\_clustering\_df1.ipynb**’ περιέχει την ενοποίηση της γνώσης από το clustering των τιμών που πραγματοποιήθηκε στο notebook cluster\_average\_price\_category.ipynb και του Dataset 1 , ώστε να ληφθούν χρήσιμα συμπεράσματα πάνω στα basket-missions σχετικά με το Revenue αυτών μέσω του Dataset1 και μετά συνεχίζει με την στατιστική ανάλυση των αποτελεσμάτων της ενοποίησης αυτής αξιοποιώντας στατιστικούς ελέγχους καθώς και . Το σημειωματάριο ‘**specialized\_clustering\_quantity\_df1.ipynb’** πραγματοποιεί την ίδια λειτουργία με το προηγούμενο notebook, αλλά αυτή τη φορά αναλύεται η ποσότητα των προϊόντων. Το σημειωματάριο ‘**time\_series.ipynb’** πραγματοποιεί ανάλυση time series πάνω στα δεδομένα του Dataset 1. Τέλος, το σημειωματάριο ‘**transaction\_analysis\_df2.ipynb’** πραγματοποιεί ανάλυση για την απάντηση των πρώτων δύο ερωτημάτων που παρείχε η Intale σχετικά με το Dataset 2, αλλά και περαιτέρω ανάλυση σχετικά με την ανάλυση των κερδών ανά μέρα και ώρα της εβδομάδας.