

Παρατηρούμε πως στα Parquet αρχεία, έχουμε μια πιο γρήγορη εκτέλεση σε σχέση με τα RDD/API, όπου στο Q2, ενώ τρέξαμε τα ίδια ερωτήματα, με διαφορετικές εντολές, το RDD, έκανε πολύ μεγαλύτερο χρόνο καθώς οι εντολές είναι περισσότερες και πρέπει να επεξεργαστεί περισσότερα δεδομένα για να το εκτελέσει σωστά.