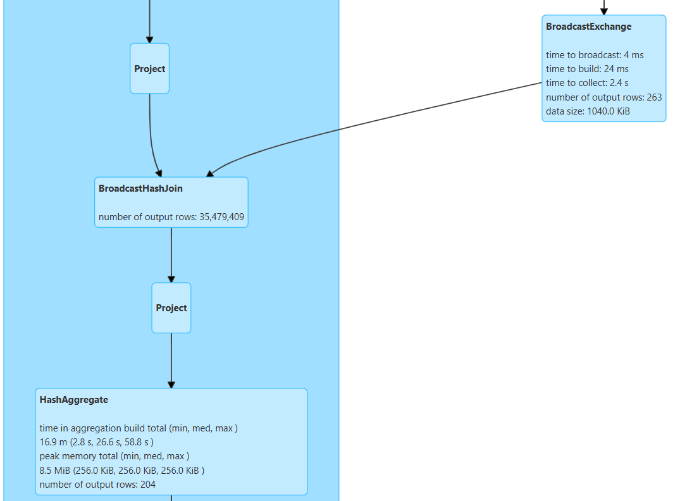
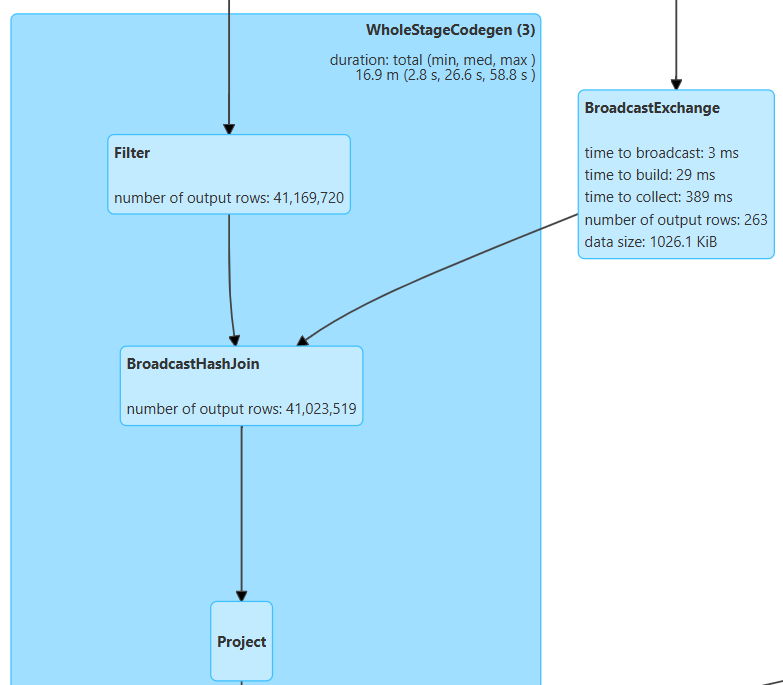
# Στρατηγικές για τα Joins των Q3, Q4,Q5

Για τα παρακάτω Queries , ο Catalyst Optimizer επέλεξε τη στρατηγική BroadcastHashJoin για την εκτέλεση των joins (όπου χρειάστηκε). Και στις δύο περιπτώσεις, το build side ήταν το δεξί dataset (BuildRight), το οποίο είχε μικρό μέγεθος (περίπου 1MB), όπως φαίνεται στο Physical Plan (Statistics(sizeInBytes ≃ 1 ΜB)).

Η επιλογή του BroadcastHashJoin είναι θεωρητικά δικαιολογημένη, καθώς τα μικρού μεγέθους datasets μπορούν να μεταδοθούν αποδοτικά σε όλους τους executors, αποφεύγοντας το κόστος του shuffle που απαιτείται σε στρατηγικές όπως το SortMergeJoin. Συνεπώς, η στρατηγική που επιλέχθηκε είναι σωστή και βελτιστοποιεί τον χρόνο εκτέλεσης.

Q3\_DF:



Q3\_DF\_SQL:

A diagram of a project

AI-generated content may be incorrect.

Q3\_PARQ:

A diagram of a project

AI-generated content may be incorrect.

Q3\_SQL\_PARQ:

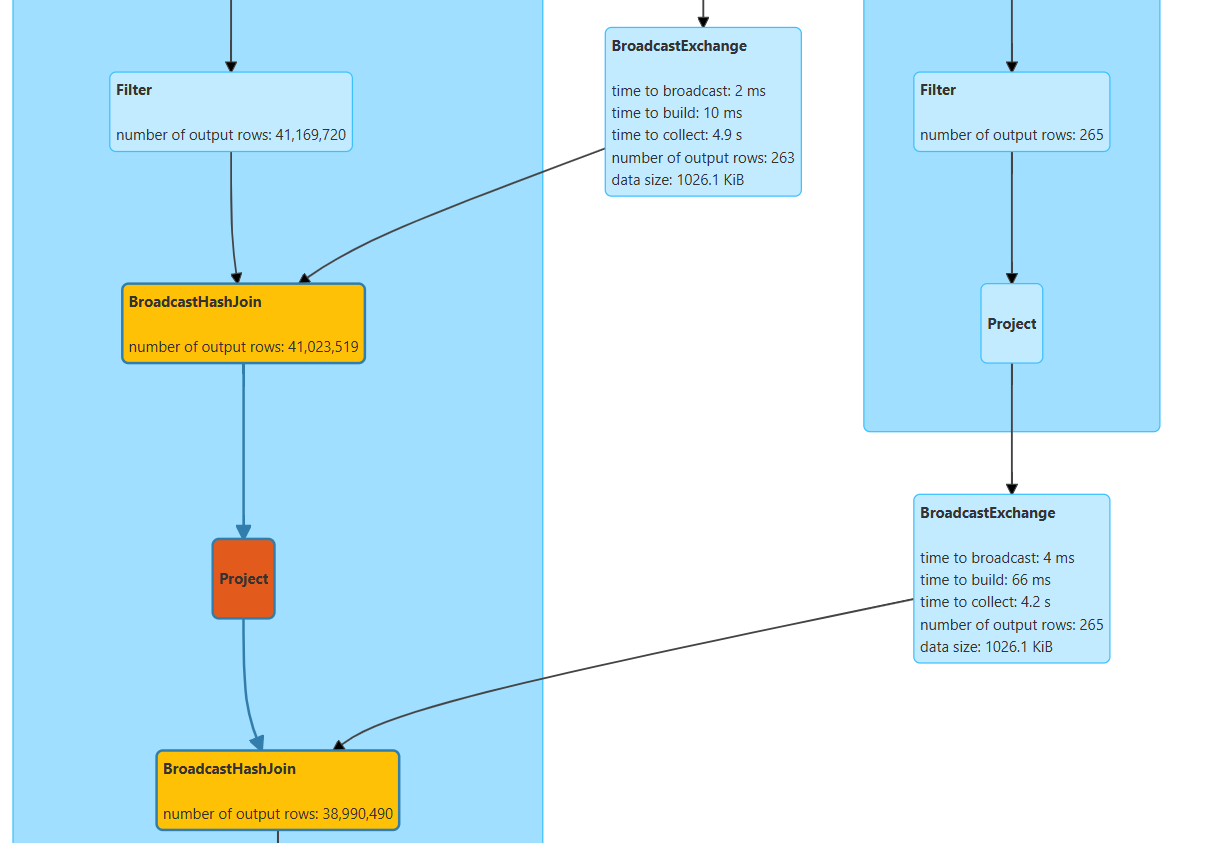
A diagram of a project

AI-generated content may be incorrect.

Q4\_SQL:

Q4\_SQL\_PARQ:

Q5\_PARQ:



Q5\_DF:

A diagram of a project

AI-generated content may be incorrect.