Performance Improvement

# Data Shuffling

<https://stackoverflow.com/questions/31386590/when-does-shuffling-occur-in-apache-spark/31387957>

# AQE : Adaptive Query Execution

<https://stackoverflow.com/questions/62603531/adaptive-query-execution-in-spark-3>

# Repartition

<https://stackoverflow.com/questions/31610971/spark-repartition-vs-coalesce>