We have a dataset of travel data analysis Column Name: Column 1: City pair (Combination of from and to): String Column 2: From location: String

Column 3: To Location: String Column

4: Product type: Integer (1=Air, 2=Car, 3 =Air+Car, 4 =Hotel, 5=Air+Hotel, 6=Hotel +Car, 7 =Air+Hotel+Car)

Column 5: Adults traveling: Integer

Column 6: Seniors traveling: Integer

Column 7: Children traveling: Integer

Column 8: Youth traveling: Integer

Column 9: Infant traveling: Integer

Column 10: Date of travel: String

Column 11: Time of travel: String

Column 12: Date of Return: String

Column 13: Time of Return: String

Column 14: Price of booking: Float

Column 15: Hotel name: String

Problem statement 1: Find out how many people has chosen their destination as MIA and MCO

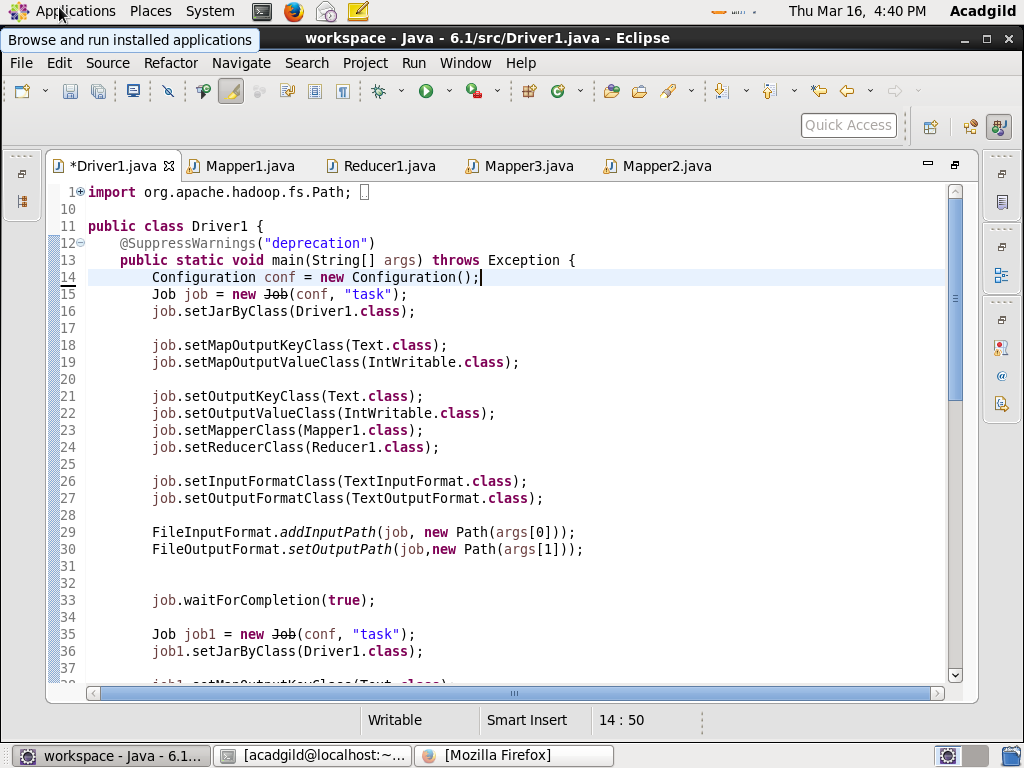
Problem statement 2: Find out the number of people undertaken the trips from the places MIA and HOU

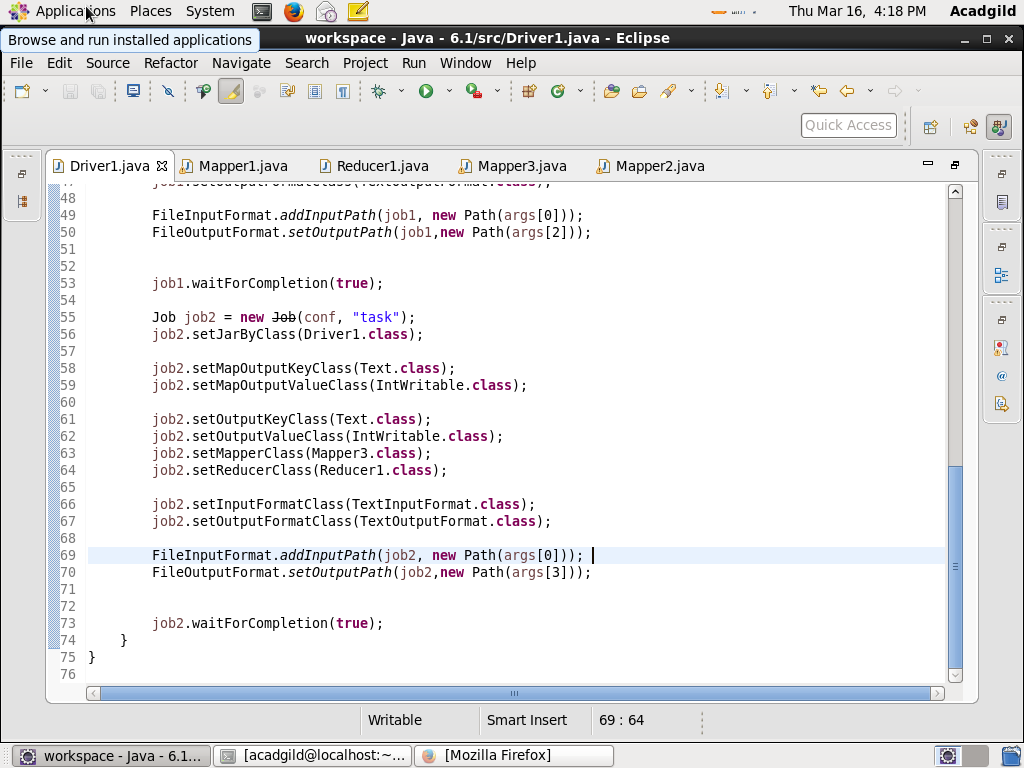
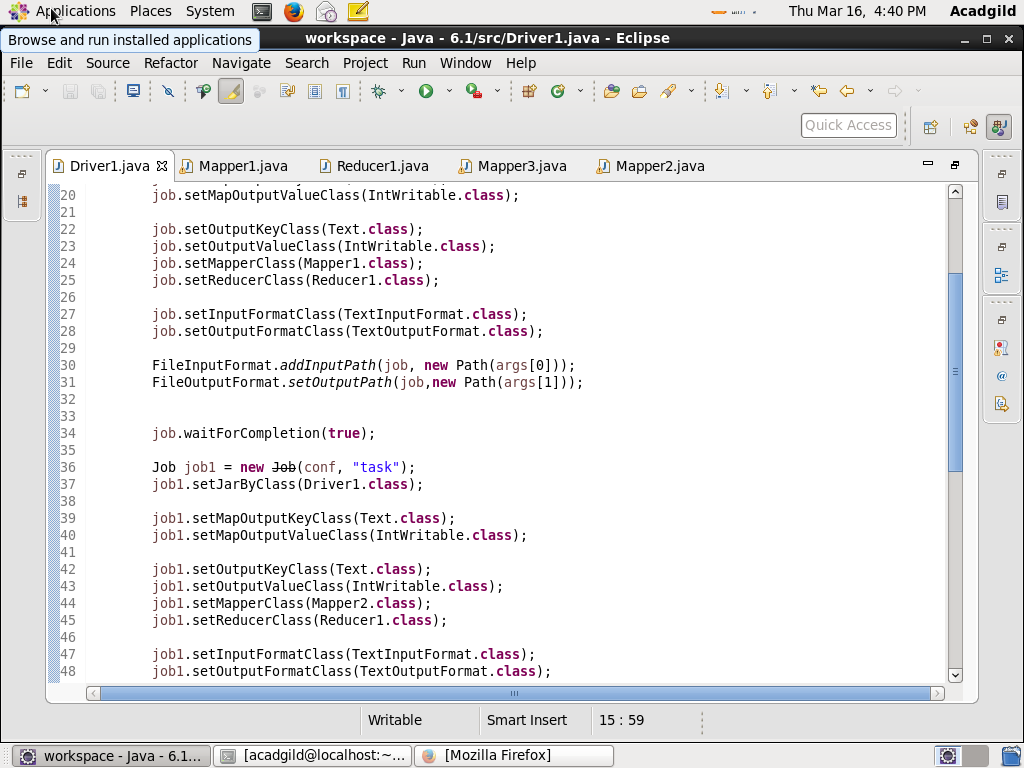
Problem statement 3: Find out how many people has chosen airline mode of travel for the places LAS and LAX

Instead of creating separate project for every task, all 3 tasks are implemented in a single project. So this project has 3 Mapper classes and 1 Reducer class

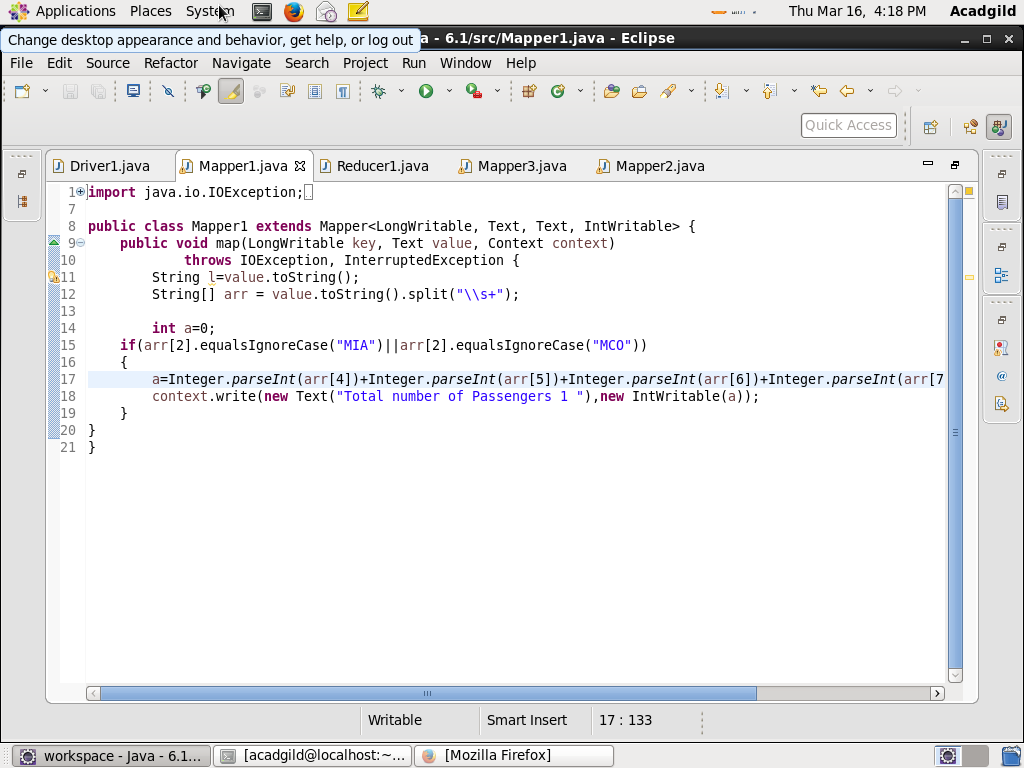
Program:

Driver class

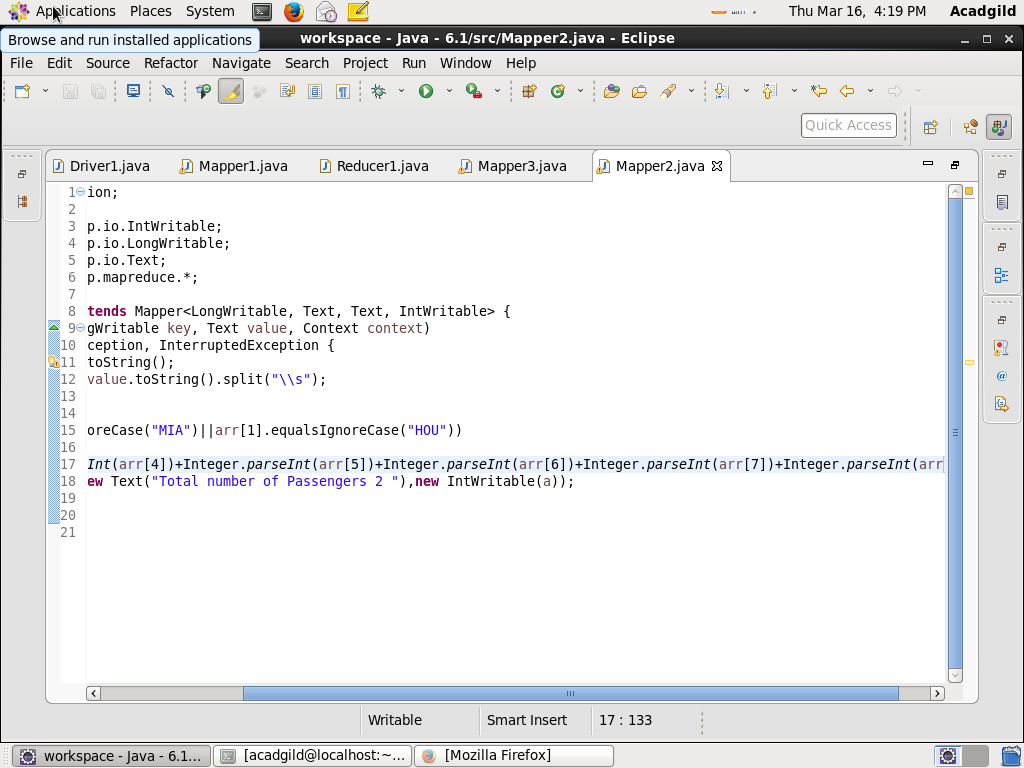




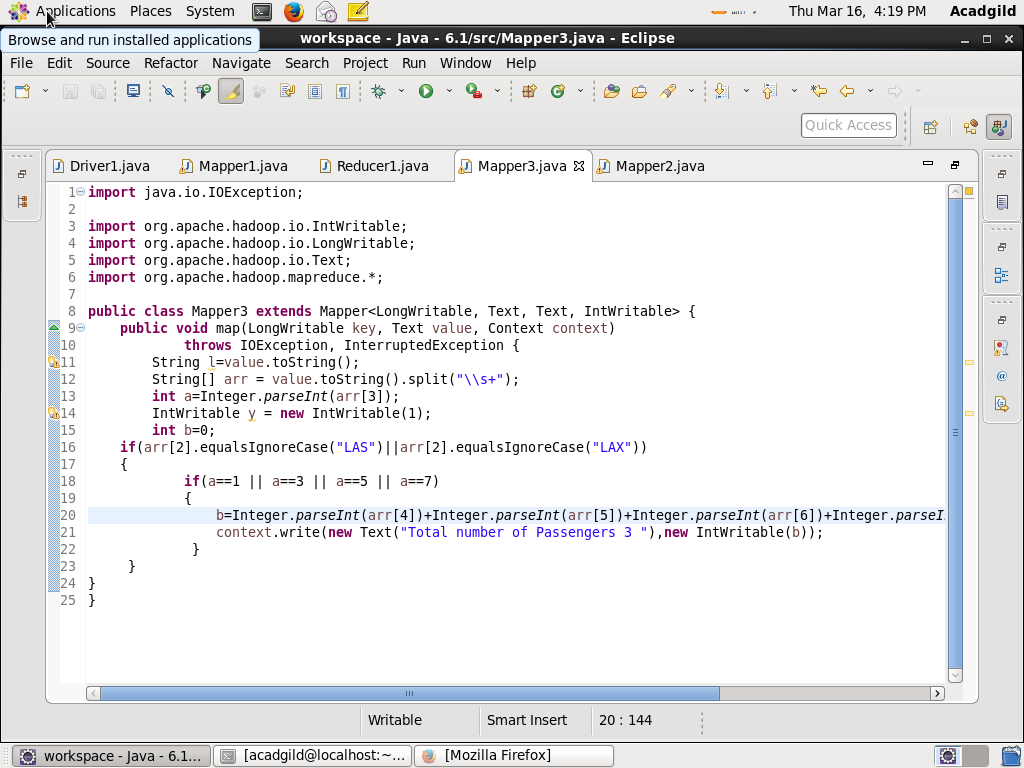
Mapper 1 class



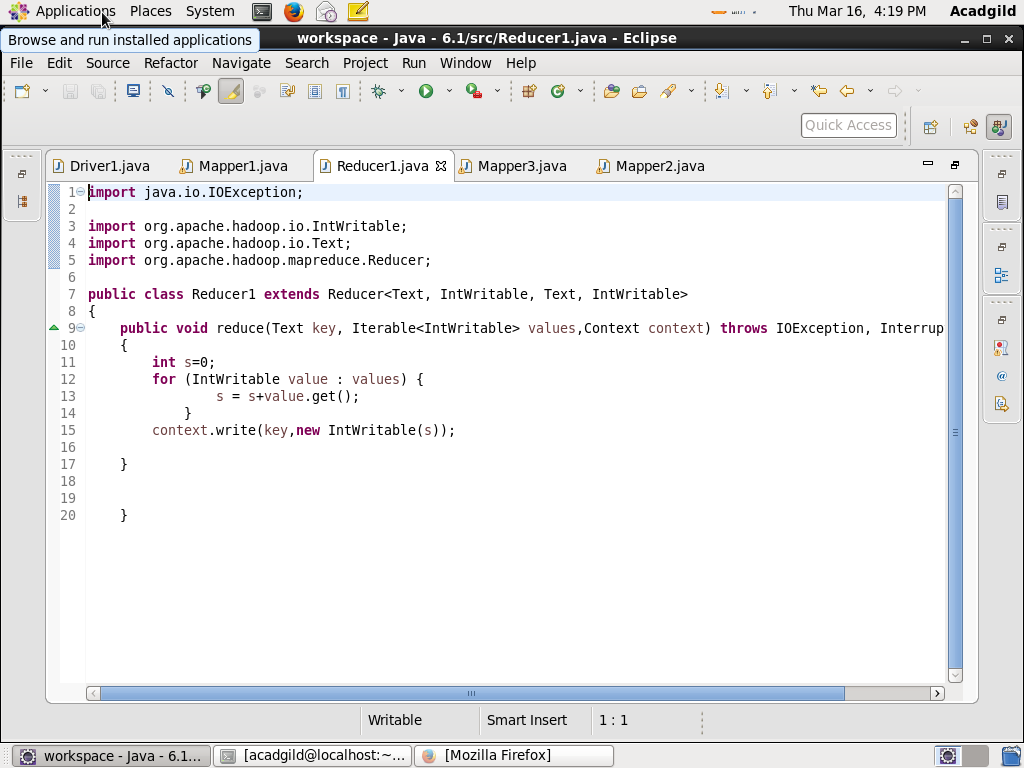
Mapper 2 class



Mapper 3 class:



Reducer class



Output:

