**Problem Statement**

The Daily show data analysis

Dataset Description:

YEAR – The year the episode aired.

GoogleKnowlege\_Occupation -Their occupation or office, according to Google’s Knowledge Graph. On the other hand, if they are not in there, how Stewart introduced them on the program.

Show – Air date of the episode. Not unique, as some shows had more than one guest

Group – A larger group designation for the occupation. For instance, U.S senators, U.S presidents, and former presidents are all under “politicians” Raw\_Guest\_List – The person or list of people who appeared on the show, according to Wikipedia.

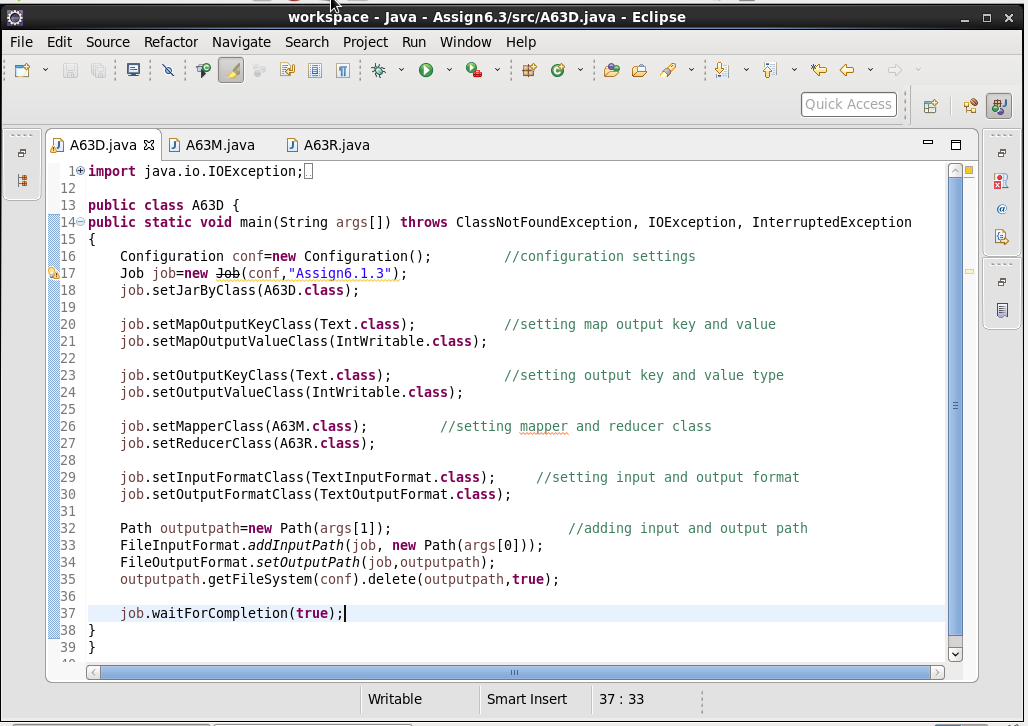
The GoogleKnowlege\_Occupation only refers to one of them in a given row.

**Input File-**

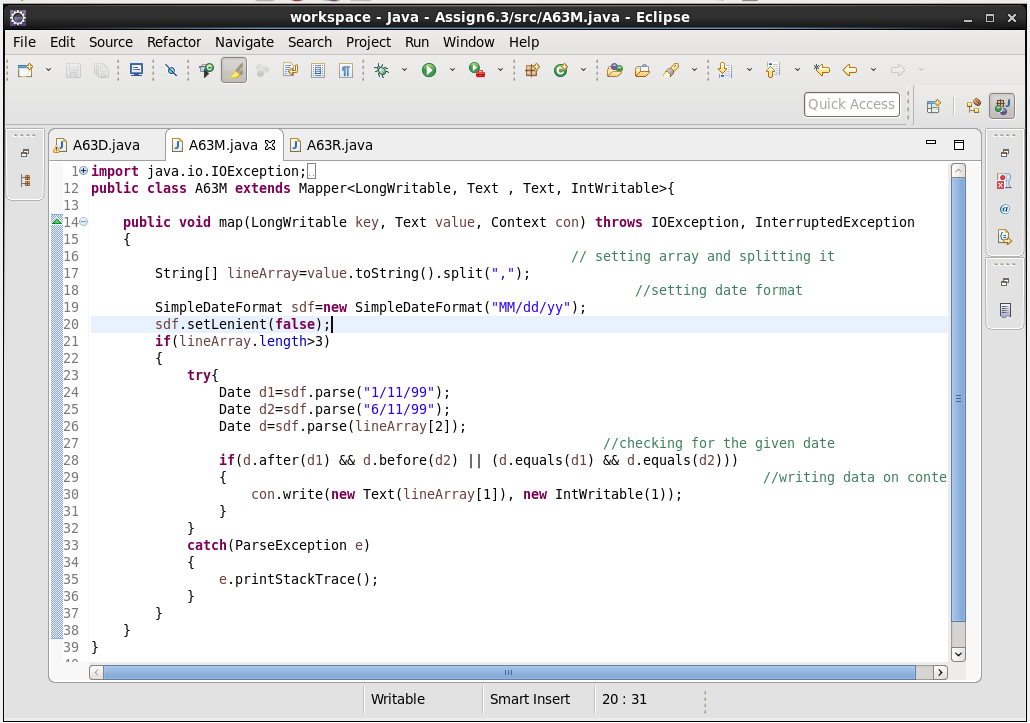
****

**Problem statement 1: Find the number of GoogleKnowlege\_Occupation people who were guests in the show, in a particular time period. Take the time period as 1/11/99 to 6/11/99 i.e., 6 months.**

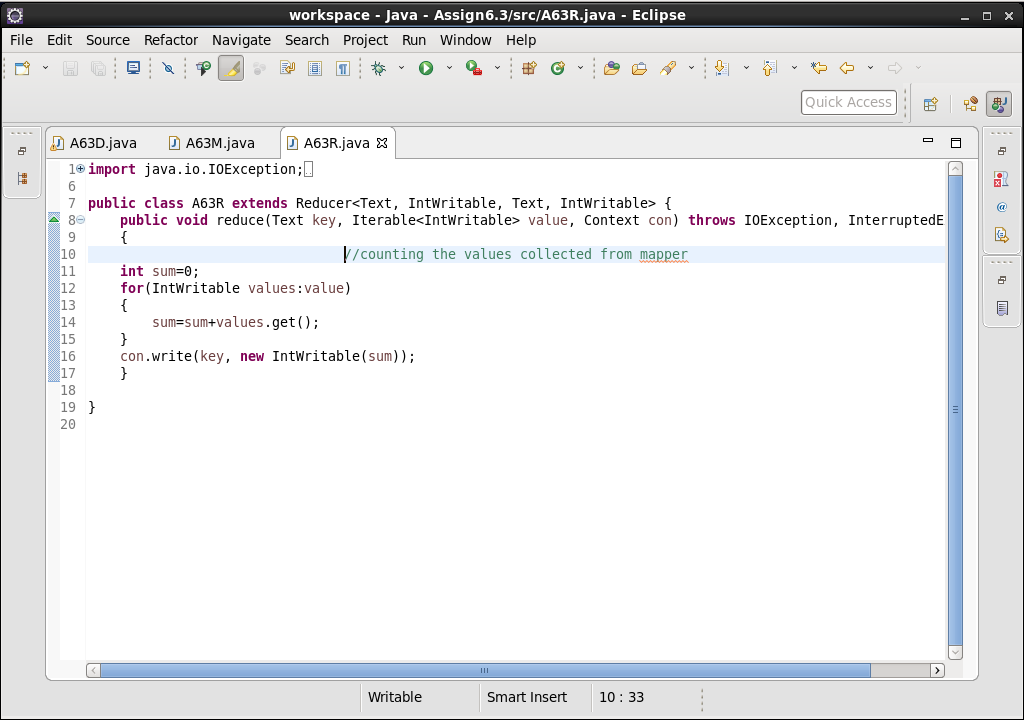
**Driver Code-**

****

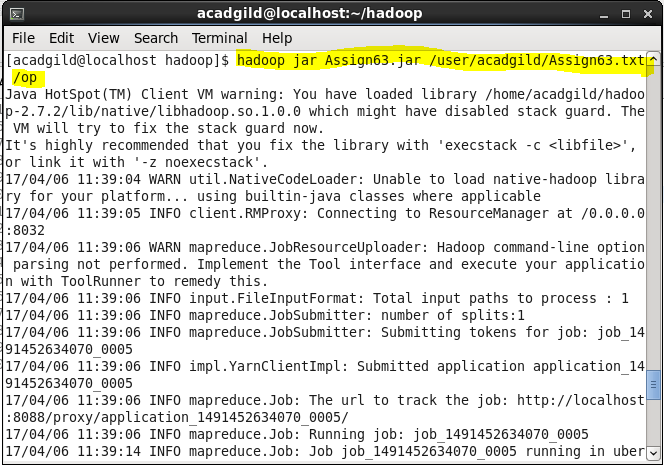
**Mapper Code-**

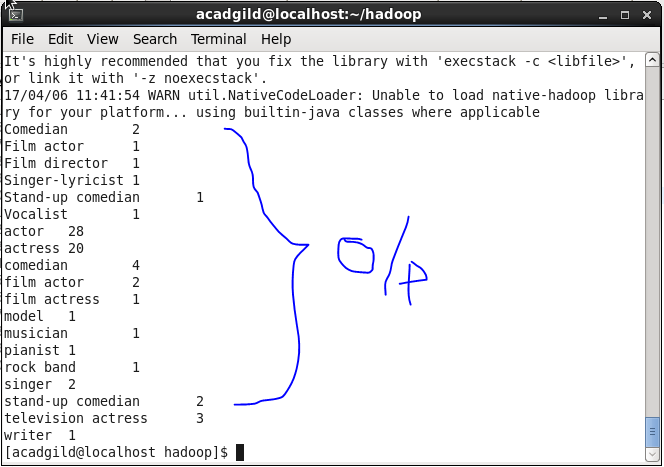
****

**Reducer Code-**

****

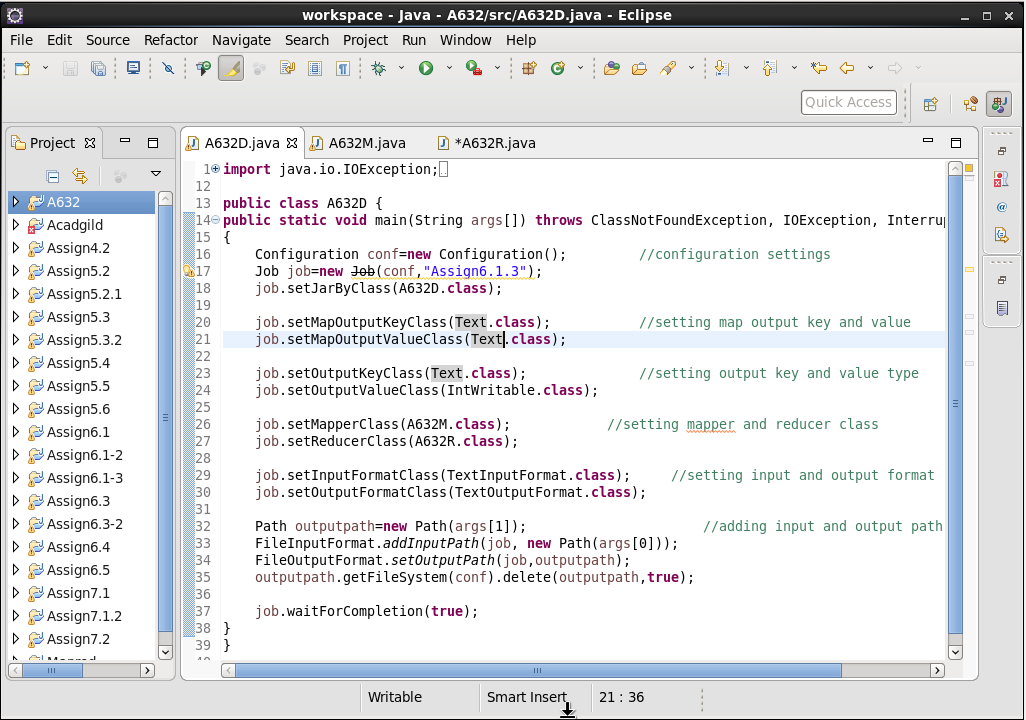
**Output-**

****

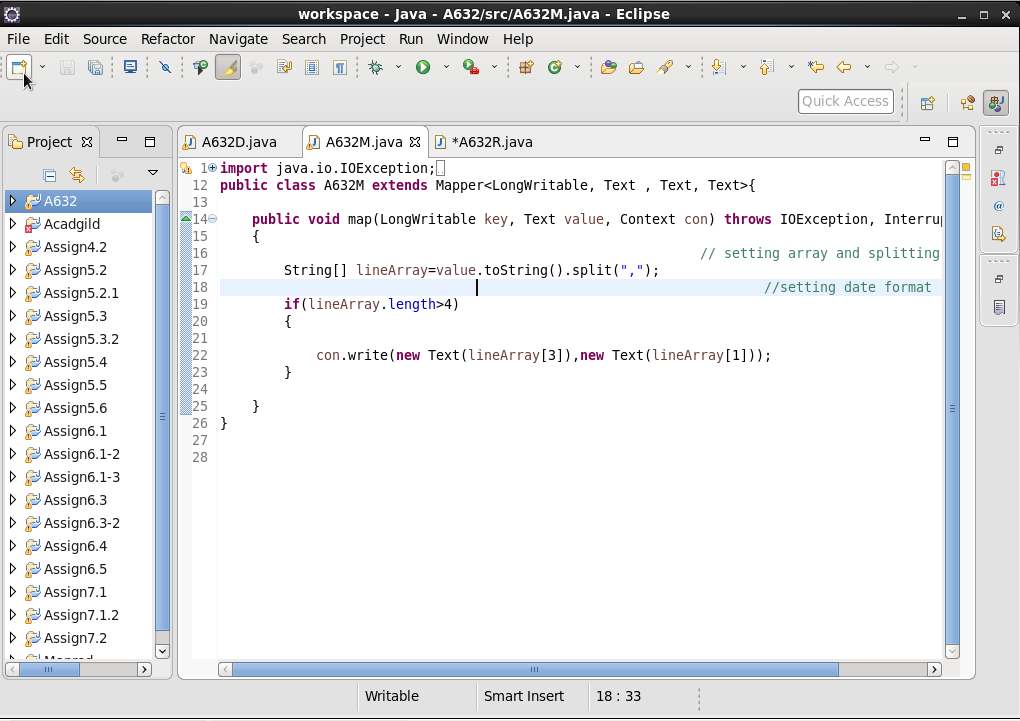
****

**Problem statement 2: Find the number of GoogleKnowledge occupation types in each group, who have been guests on the show.**

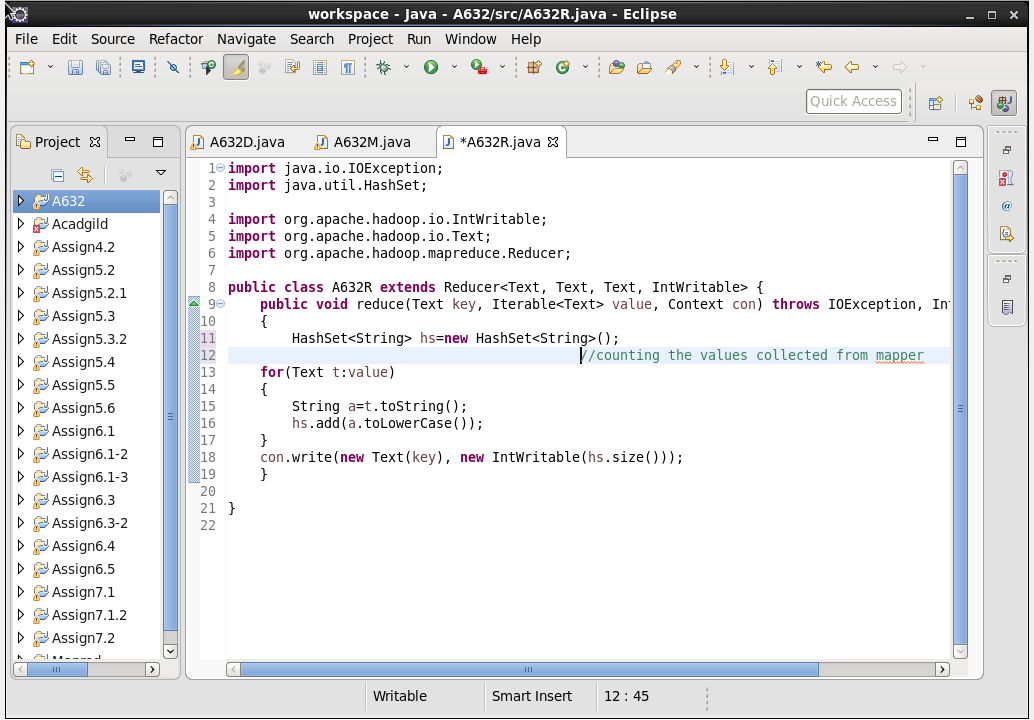
**Driver Code-**

****

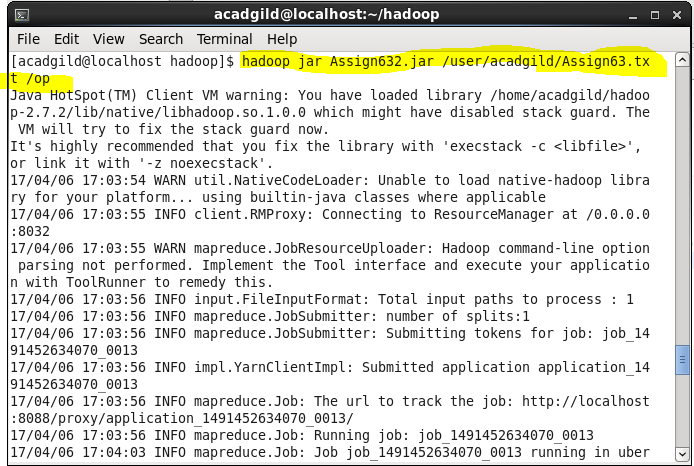
**Mapper Code-**

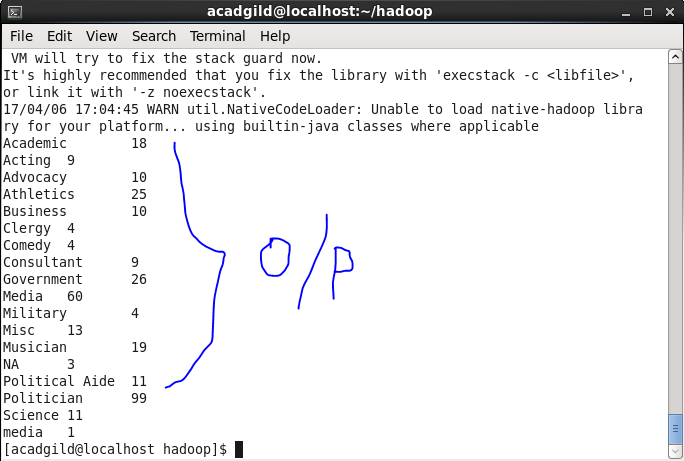
****

**Reducer Code-**

****

**Output File-**

****

****