**PYTHON PROJECT DOCUMENTATION**

In this project we have a given a dataset and we have to find some queries to find results for the dataset. This data set is related about the movies and their year of releases, duration of movie. In this project we have to open the text file using file handling. The entire approach is done using file handling method. We solve all queries related to the data set by using file handling method. Then we have to converted the file into the list by using split option, with that we print the results for the queries.

The queries are:

**Find the number of movies released between 1950 and 1960**

f = open(**"movie.txt"**,**"r"**)  
count = 0  
**for** line **in** f:  
 a = line.split(**","**)  
 **if** (int (a [2])>1950 **and** int (a [2]) <1960):  
 count = count +1  
print (**"Number of movies released between 1950 and 1960:"**,count)  
f.close()

**OUTPUT**

Number of movies released between 1950 and 1960: 414

**Find the number of movies having rating more than 4.**

f = open(**"movie.txt"**,**"r"**)  
count = 0  
**for** line **in** f:  
 b = line.split(**","**)  
 **if** b[3]!=**'' and** float(b[3]) > 4.0:  
 count = count+1  
print(**"Total number of movie rating greater than 4: "**,count)  
f.close()

**OUTPUT:**

Total number of movie rating greater than 4: 897

**Find the movies whose rating are between 3 and 4**

f = open(**"movie.txt"**,**"r"**)  
count = 0  
**for** line **in** f:  
 a= line.split(**","**)  
 **if**((a[3])>**'3.0' and** (a[3])<**'4.0'**):  
 count = count +1  
print(**"Number of movies having rating between 3 and 4: "**,count)  
f.close()

**OUTPUT:**

Number of movies having rating between 3 and 4: 7161

**Find the number of movies with duration more than 2 hours (7200 second)**

f = open(**"movie.txt"**,**"r"**)  
count = 0  
**for** line **in** f:  
 line = line.strip()  
 lis = line.split(**","**)  
 **if** lis[4]!=**'' and** int(lis[4])>7200:  
 count = count+1  
print(**"Number of movies with duration more than 2 hours: "**,count)  
f.close()

**OUTPUT:**

Number of movies with more than 2 hours: 641

**Find the list of years and number of movies released each year**

f = open(**"movie.txt"**,**"r"**)  
count = 0  
d =[]  
**for** line **in** f:  
 lis = line.split(**","**)  
 d.append(lis[2])  
c = list(set(d))  
**for** i **in** range(len(c)):  
 a = d.count(c[i])  
 print(**"The year"**,c[i],**"The movies are"**,a)  
f.close()

**OUTPUT:**

**The year 1945 The movies are 9  
The year 1941 The movies are 7  
The year 2004 The movies are 1381  
The year 1994 The movies are 517  
The year 1948 The movies are 13  
The year 1972 The movies are 166  
The year 1991 The movies are 364  
The year 1943 The movies are 7  
The year 1937 The movies are 4  
The year 1962 The movies are 124  
The year 1961 The movies are 119  
The year 1993 The movies are 564  
The year 1935 The movies are 11  
The year 1997 The movies are 788  
The year 1921 The movies are 2  
The year 1940 The movies are 9  
The year 1938 The movies are 5  
The year 2001 The movies are 1173  
The year 1977 The movies are 136  
The year 1982 The movies are 153  
The year 1949 The movies are 9  
The year 1990 The movies are 470  
The year 1986 The movies are 287  
The year 1929 The movies are 5  
The year 1920 The movies are 6  
The year 1926 The movies are 2  
The year 1980 The movies are 107  
The year 1944 The movies are 10  
The year 2010 The movies are 5107  
The year 1953 The movies are 17  
The year 1971 The movies are 131  
The year 1925 The movies are 5  
The year 1989 The movies are 421  
The year 1966 The movies are 103  
The year 1947 The movies are 9  
The year 1927 The movies are 4  
The year 1914 The movies are 20  
The year 1954 The movies are 17  
The year 2002 The movies are 1117  
The year 1919 The movies are 3  
The year 1934 The movies are 8  
The year 1936 The movies are 7  
The year 1915 The movies are 1  
The year 2000 The movies are 902  
The year 1998 The movies are 843  
The year 1996 The movies are 688  
The year 1924 The movies are 5  
The year 1939 The movies are 6  
The year 1968 The movies are 173  
The year 1952 The movies are 15  
The year 2013 The movies are 981  
The year 1964 The movies are 126  
The year 1979 The movies are 140  
The year 1985 The movies are 334  
The year 2007 The movies are 2892  
The year 2006 The movies are 2416  
The year 1918 The movies are 1  
The year 1973 The movies are 168  
The year 1988 The movies are 334  
The year 1933 The movies are 7  
The year 1960 The movies are 123  
The year 1928 The movies are 2  
The year 2008 The movies are 3358  
The year 1922 The movies are 2  
The year 1916 The movies are 1  
The year 1930 The movies are 5  
The year 1965 The movies are 104  
The year 1958 The movies are 73  
The year 1969 The movies are 124  
The year 2005 The movies are 1937  
The year 1967 The movies are 279  
The year 1951 The movies are 33  
The year 1984 The movies are 303  
The year 1987 The movies are 280  
The year 1959 The movies are 87  
The year 1923 The movies are 4  
The year 1950 The movies are 10  
The year 1983 The movies are 270  
The year 1931 The movies are 3  
The year 1946 The movies are 6  
The year 1963 The movies are 88  
The year 1995 The movies are 592  
The year 1970 The movies are 141  
The year 2014 The movies are 1  
The year 2009 The movies are 4451  
The year 1942 The movies are 3  
The year 2003 The movies are 1399  
The year 1992 The movies are 479  
The year 1955 The movies are 14  
The year 1956 The movies are 60  
The year 1932 The movies are 4  
The year 1981 The movies are 112  
The year 1974 The movies are 178  
The year 1975 The movies are 134  
The year 1999 The movies are 1181  
The year 1913 The movies are 3  
The year 1976 The movies are 118  
The year 2012 The movies are 4339  
The year 1978 The movies are 231  
The year 2011 The movies are 5511  
The year 1957 The movies are 98**

**Find total number of movies in dataset**

f = open(**"movie.txt"**,**"r"**)  
count = 0  
a = f.read()  
lis = a.split(**"\n"**)  
**for** i **in** lis:  
 count = count+1  
print(**"Total number of movies in dataset: "**,count-1)  
f.close()

**OUTPUT:**

Total number of movies in dataset: 49590