

## Dataset 1: Movie Ratings

### **Raw Data:**

M01,Inception,Christopher Nolan,2010,Sci-Fi,8.8  
M02,Titanic,James Cameron,1997,Romance,7.9  
M03,The Dark Knight,Christopher Nolan,2008,Action,9.0  
M04,Avatar,James Cameron,2009,Sci-Fi,7.8  
M05,Interstellar,Christopher Nolan,2014,Sci-Fi,8.6

### **Tasks:**

1. Add prefix 'MOV-' before ID.
2. Extract decade from Year (e.g., 2010 → 2010s).
3. Classify as 'Hit' if IMDB  $\geq 8.5$  else 'Average'.
4. Count number of movies per Director.
5. Sort by IMDB descending.

### **Expected Output:**

MOV-M03,The Dark Knight,Christopher Nolan,2000s,Action,9.0,Hit,2  
MOV-M01,Inception,Christopher Nolan,2010s,Sci-Fi,8.8,Hit,2  
MOV-M05,Interstellar,Christopher Nolan,2010s,Sci-Fi,8.6,Hit,2  
MOV-M02,Titanic,James Cameron,1990s,Romance,7.9,Average,2  
MOV-M04,Avatar,James Cameron,2000s,Sci-Fi,7.8,Average,2

## Dataset 2: Bus Tickets

### **Raw Data:**

B101,Hyderabad,Bangalore,350,Day,Yes  
B102,Delhi,Jaipur,270,Night,No  
B103,Mumbai,Pune,150,Day,Yes  
B104,Chennai,Coimbatore,500,Night,Yes  
B105,Kolkata,Digha,200,Day,No

### **Tasks:**

1. Add 'BUS-' before BusID.
2. Classify journey as 'Short' if Distance  $< 300$ , else 'Long'.
3. If AC=Yes, fare = Distance\*3; else fare = Distance\*2.
4. Add 10% extra charge for Night type buses.
5. Sort by final fare descending.

### **Expected Output:**

BUS-B104,Chennai,Coimbatore,500,Night,Yes,Long,1650.0  
BUS-B101,Hyderabad,Bangalore,350,Day,Yes,Long,1050  
BUS-B102,Delhi,Jaipur,270,Night,No,Short,594.0  
BUS-B105,Kolkata,Digha,200,Day,No,Short,400  
BUS-B103,Mumbai,Pune,150,Day,Yes,Short,450

## Dataset 3: Student Marks

### **Raw Data:**

S01,Rahul,Maths,78  
S02,Priya,Science,85  
S03,Arjun,English,65  
S04,Sneha,Maths,92  
S05,Vikram,Science,55

### **Tasks:**

1. Add 'STU-' before StuID.
2. Grade students: A if  $\geq 80$ , B if 60–79, C if  $< 60$ .
3. Mark subject category: Science subjects (Maths/Science) vs Non-Science.
4. Calculate Pass/Fail (Pass if  $\geq 40$ ).
5. Sort students by Marks descending.

### **Expected Output:**

STU-S04,Sneha,Maths,92,A,Science,Pass  
STU-S02,Priya,Science,85,A,Science,Pass  
STU-S01,Rahul,Maths,78,B,Science,Pass  
STU-S03,Arjun,English,65,B,Non-Science,Pass  
STU-S05,Vikram,Science,55,C,Science,Pass