Importing Libraries

```
[ ] import numpy as np # linear algebra
    import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)
    from sklearn import preprocessing
    from sklearn.model_selection import train_test_split
    import matplotlib.pyplot as plt
    import seaborn as sns
```

Importing Dataset

DESCRIPTION

> The dataset provided contains movie reviews given by Amazon customers. Reviews were given between May 1996 and July 2014.

Data Dictionary

> UserID – 4848 customers who provided a rating for each movie Movie 1 to Movie 206 – 206 movies for which ratings are provided by 4848 distinct users

---

Data Considerations

- All the users have not watched all the movies and therefore, all movies are not rated. These missing values are represented by NA.
- Ratings are on a scale of -1 to 10 where -1 is the least rating and 10 is the best.

```
[ ] amazon_df = pd.read_csv("Amazon - Movies and TV Ratings.csv")
```

```
[ ] amazon_df_no_id=amazon_df.drop(["user_id"],axis=1)
```

Properties Of Dataset

```
[ ] amazon_df.shape
```

```
(4848, 207)
```

```
[ ] pd.options.mode.chained_assignment = None  # default='warn'
    pd.options.display.max_columns = 999
    pd.options.display.max_rows = 5000
    pd.set_option('display.max_columns', None)
```

```
[ ] amazon_df.head()
```

user_id	Movie1	Movie2	Movie3	Movie4	Movie5	Movie6	Movie7	Movie8	Movie9	Movie10	Movie11	Movie12	Movie13	Movie
DBKS7OM2IR	5.0	5.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	N
QC2PC1VTGP	NaN	NaN	2.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	N
6WPMP9UKX	NaN	NaN	NaN	5.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	N
Y68KEPQ5ZD	NaN	NaN	NaN	5.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	N
WROP5KTTW	NaN	NaN	NaN	NaN	5.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	N

```
[ ] amazon_df.describe()
```

	Movie1	Movie2	Movie3	Movie4	Movie5	Movie6	Movie7	Movie8	Movie9	Movie10	Movie11	Movie12	Movie13	Movie14
count	1.0	1.0	1.0	2.0	29.000000	1.0	1.0	1.0	1.0	1.0	2.0	5.0	1.0	1.0
mean	5.0	5.0	2.0	5.0	4.103448	4.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	4.0
std	NaN	NaN	NaN	0.0	1.496301	NaN	NaN	NaN	NaN	NaN	0.0	0.0	NaN	NaN
min	5.0	5.0	2.0	5.0	1.000000	4.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	4.0
25%	5.0	5.0	2.0	5.0	4.000000	4.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	4.0
50%	5.0	5.0	2.0	5.0	5.000000	4.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	4.0
75%	5.0	5.0	2.0	5.0	5.000000	4.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	4.0
max	5.0	5.0	2.0	5.0	5.000000	4.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	4.0

```
[ ] amazon_df.info(verbose=True,null_counts=True)
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4848 entries, 0 to 4847
Data columns (total 207 columns):
 #   Column    Non-Null Count  Dtype
---  ------    --------------  -----
 0   user_id   4848 non-null   object
 1   Movie1    1 non-null      float64
 2   Movie2    1 non-null      float64
 3   Movie3    1 non-null      float64
 4   Movie4    2 non-null      float64
 5   Movie5    29 non-null     float64
 6   Movie6    1 non-null      float64
 7   Movie7    1 non-null      float64
 8   Movie8    1 non-null      float64
 9   Movie9    1 non-null      float64
 10  Movie10   1 non-null      float64
 11  Movie11   2 non-null      float64
 12  Movie12   5 non-null      float64
 13  Movie13   1 non-null      float64
 14  Movie14   1 non-null      float64
 15  Movie15   1 non-null      float64
 16  Movie16   320 non-null    float64
 17  Movie17   1 non-null      float64
 18  Movie18   1 non-null      float64
 19  Movie19   2 non-null      float64
 20  Movie20   1 non-null      float64
 21  Movie21   1 non-null      float64
 22  Movie22   2 non-null      float64
 23  Movie23   3 non-null      float64
 24  Movie24   5 non-null      float64
 25  Movie25   1 non-null      float64
 26  Movie26   3 non-null      float64
 27  Movie27   1 non-null      float64
 28  Movie28   3 non-null      float64
 29  Movie29   243 non-null    float64
 30  Movie30   2 non-null      float64
 31  Movie31   2 non-null      float64
 32  Movie32   2 non-null      float64
 33  Movie33   1 non-null      float64
 34  Movie34   1 non-null      float64
 35  Movie35   1 non-null      float64
 36  Movie36   1 non-null      float64
 37  Movie37   1 non-null      float64
 38  Movie38   1 non-null      float64
 39  Movie39   4 non-null      float64
 40  Movie40   3 non-null      float64
 41  Movie41   1 non-null      float64
 42  Movie42   1 non-null      float64
 43  Movie43   7 non-null      float64
 44  Movie44   2 non-null      float64
 45  Movie45   1 non-null      float64
 46  Movie46   1 non-null      float64
 47  Movie47   1 non-null      float64
 48  Movie48   1 non-null      float64
 49  Movie49   1 non-null      float64
 50  Movie50   1 non-null      float64
 51  Movie51   2 non-null      float64
 52  Movie52   17 non-null     float64
 53  Movie53   3 non-null      float64
 54  Movie54   1 non-null      float64
 55  Movie55   1 non-null      float64
 56  Movie56   1 non-null      float64
 57  Movie57   1 non-null      float64
 58  Movie58   1 non-null      float64
 59  Movie59   1 non-null      float64
 60  Movie60   1 non-null      float64
 61  Movie61   1 non-null      float64
 62  Movie62   4 non-null      float64
 63  Movie63   1 non-null      float64
 64  Movie64   1 non-null      float64
 65  Movie65   1 non-null      float64
 66  Movie66   1 non-null      float64
 67  Movie67   1 non-null      float64
 68  Movie68   1 non-null      float64
 69  Movie69   1 non-null      float64
 70  Movie70   2 non-null      float64
 71  Movie71   1 non-null      float64
 72  Movie72   1 non-null      float64
 73  Movie73   1 non-null      float64
 74  Movie74   1 non-null      float64
 75  Movie75   1 non-null      float64
 76  Movie76   2 non-null      float64
 77  Movie77   1 non-null      float64
 78  Movie78   1 non-null      float64
 79  Movie79   2 non-null      float64
 80  Movie80   1 non-null      float64
 81  Movie81   12 non-null     float64
 82  Movie82   2 non-null      float64
 83  Movie83   1 non-null      float64
 84  Movie84   1 non-null      float64
 85  Movie85   3 non-null      float64
 86  Movie86   21 non-null     float64
 87  Movie87   1 non-null      float64
 88  Movie88   1 non-null      float64
 89  Movie89   83 non-null     float64
 90  Movie90   18 non-null     float64
 91  Movie91   128 non-null    float64
 92  Movie92   101 non-null    float64
 93  Movie93   2 non-null      float64
 94  Movie94   3 non-null      float64
 95  Movie95   6 non-null      float64
 96  Movie96   3 non-null      float64
```

```
 97  Movie97   5 non-null      float64
 98  Movie98   1 non-null      float64
 99  Movie99   2 non-null      float64
100 Movie100  1 non-null      float64
101 Movie101  5 non-null      float64
102 Movie102  2 non-null      float64
103 Movie103  272 non-null    float64
104 Movie104  4 non-null      float64
105 Movie105  3 non-null      float64
106 Movie106  1 non-null      float64
107 Movie107  39 non-null     float64
108 Movie108  54 non-null     float64
109 Movie109  13 non-null     float64
110 Movie110  8 non-null      float64
111 Movie111  39 non-null     float64
112 Movie112  2 non-null      float64
113 Movie113  4 non-null      float64
114 Movie114  7 non-null      float64
115 Movie115  1 non-null      float64
116 Movie116  1 non-null      float64
117 Movie117  11 non-null     float64
118 Movie118  5 non-null      float64
119 Movie119  8 non-null      float64
120 Movie120  3 non-null      float64
121 Movie121  4 non-null      float64
122 Movie122  4 non-null      float64
123 Movie123  1 non-null      float64
124 Movie124  4 non-null      float64
125 Movie125  5 non-null      float64
126 Movie126  2 non-null      float64
127 Movie127  2313 non-null   float64
128 Movie128  3 non-null      float64
129 Movie129  3 non-null      float64
130 Movie130  4 non-null      float64
131 Movie131  4 non-null      float64
132 Movie132  3 non-null      float64
133 Movie133  1 non-null      float64
134 Movie134  6 non-null      float64
135 Movie135  1 non-null      float64
136 Movie136  2 non-null      float64
137 Movie137  3 non-null      float64
138 Movie138  13 non-null     float64
139 Movie139  4 non-null      float64
140 Movie140  578 non-null    float64
141 Movie141  7 non-null      float64
142 Movie142  1 non-null      float64
143 Movie143  1 non-null      float64
144 Movie144  1 non-null      float64
145 Movie145  1 non-null      float64
146 Movie146  1 non-null      float64
147 Movie147  1 non-null      float64
148 Movie148  2 non-null      float64
149 Movie149  1 non-null      float64
150 Movie150  3 non-null      float64
151 Movie151  2 non-null      float64
152 Movie152  1 non-null      float64
153 Movie153  1 non-null      float64
154 Movie154  1 non-null      float64
155 Movie155  2 non-null      float64
156 Movie156  1 non-null      float64
157 Movie157  4 non-null      float64
158 Movie158  66 non-null     float64
159 Movie159  2 non-null      float64
160 Movie160  6 non-null      float64
161 Movie161  25 non-null     float64
162 Movie162  15 non-null     float64
163 Movie163  13 non-null     float64
164 Movie164  4 non-null      float64
165 Movie165  1 non-null      float64
166 Movie166  2 non-null      float64
167 Movie167  4 non-null      float64
168 Movie168  4 non-null      float64
169 Movie169  4 non-null      float64
170 Movie170  3 non-null      float64
171 Movie171  1 non-null      float64
172 Movie172  2 non-null      float64
173 Movie173  15 non-null     float64
174 Movie174  4 non-null      float64
175 Movie175  1 non-null      float64
176 Movie176  1 non-null      float64
177 Movie177  1 non-null      float64
178 Movie178  1 non-null      float64
179 Movie179  7 non-null      float64
180 Movie180  1 non-null      float64
181 Movie181  2 non-null      float64
182 Movie182  30 non-null     float64
183 Movie183  1 non-null      float64
184 Movie184  17 non-null     float64
185 Movie185  24 non-null     float64
186 Movie186  9 non-null      float64
187 Movie187  1 non-null      float64
188 Movie188  6 non-null      float64
189 Movie189  5 non-null      float64
190 Movie190  7 non-null      float64
191 Movie191  6 non-null      float64
192 Movie192  10 non-null     float64
193 Movie193  7 non-null      float64
194 Movie194  7 non-null      float64
195 Movie195  1 non-null      float64
196 Movie196  9 non-null      float64
197 Movie197  5 non-null      float64
198 Movie198  2 non-null      float64
199 Movie199  1 non-null      float64
200 Movie200  8 non-null      float64
201 Movie201  3 non-null      float64
202 Movie202  6 non-null      float64
```

```
 202 Movie202  6 non-null      float64
 203 Movie203  1 non-null      float64
 204 Movie204  8 non-null      float64
 205 Movie205  35 non-null     float64
 206 Movie206  13 non-null     float64
dtypes: float64(206), object(1)
memory usage: 7.7+ MB
```

Exploratory Data Analysis Task

Exploratory Data Analysis:

- Which movies have maximum views/ratings?
- What is the average rating for each movie? Define the top 5
- movies with the maximum ratings.
- Define the top 5 movies with the least audience.

Which movies have maximum views/ratings?

```
[ ] #Sum of Count
    Count=[amazon_df_no_id.count(axis=0)]
    count_DF = pd.DataFrame(data=Count)
    count_DF = count_DF.transpose().reset_index()
    count_DF.columns = ['Movies', 'Count_sum']
    ##Which movies have maximum views?
    Top_viewed_movie = count_DF.sort_values(by='Count_sum',ascending=False)
    print(Top_viewed_movie[Top_viewed_movie.Count_sum == Top_viewed_movie.Count_sum.max()])
    ##Top_viewed_movie.head()
```

```
       Movies  Count_sum
126  Movie127       2313
```

```
[ ] #Sum Of Rating
    Sum=[amazon_df_no_id.sum(axis=0)]
    sum_DF = pd.DataFrame(data=Sum)
    sum_DF = sum_DF.transpose().reset_index()
    sum_DF.columns = ['Movies', 'Rating_sum']
    #Which movies have maximum ratings?
    Top_rated_movie = sum_DF.sort_values(by='Rating_sum',ascending=False)
    print(Top_rated_movie[Top_rated_movie.Rating_sum == Top_rated_movie.Rating_sum.max()])
    ##Top_rated_movie.head()
```

```
       Movies  Rating_sum
126  Movie127      9511.0
```

What is the average rating for each movie?

```
[ ] Average_DF = pd.DataFrame([amazon_df_no_id.mean(axis=0)])
    Average_DF = Average_DF.T.reset_index()
    Average_DF.columns = ["Movies","Average"]
    Average_DF
```

	Movies	Average
0	Movie1	5.000000
1	Movie2	5.000000
2	Movie3	2.000000
3	Movie4	5.000000
4	Movie5	4.103448
5	Movie6	4.000000
6	Movie7	5.000000
7	Movie8	5.000000
8	Movie9	5.000000
9	Movie10	5.000000
10	Movie11	5.000000
11	Movie12	5.000000
12	Movie13	5.000000
13	Movie14	4.000000
14	Movie15	5.000000
15	Movie16	4.518750
16	Movie17	3.000000
17	Movie18	5.000000
18	Movie19	3.500000
19	Movie20	3.000000

20	Movie21	5.000000
21	Movie22	5.000000
22	Movie23	5.000000
23	Movie24	4.400000
24	Movie25	5.000000
25	Movie26	3.000000
26	Movie27	5.000000
27	Movie28	3.333333
28	Movie29	4.806584
29	Movie30	4.500000
30	Movie31	5.000000
31	Movie32	4.500000
32	Movie33	5.000000
33	Movie34	5.000000
34	Movie35	5.000000
35	Movie36	5.000000
36	Movie37	5.000000
37	Movie38	5.000000
38	Movie39	5.000000
39	Movie40	5.000000
40	Movie41	5.000000
41	Movie42	5.000000
42	Movie43	4.857143
43	Movie44	5.000000
44	Movie45	1.000000
45	Movie46	5.000000
46	Movie47	5.000000
47	Movie48	5.000000
48	Movie49	5.000000
49	Movie50	5.000000
50	Movie51	4.500000
51	Movie52	3.470588
52	Movie53	2.000000
53	Movie54	5.000000
54	Movie55	5.000000
55	Movie56	5.000000
56	Movie57	5.000000
57	Movie58	1.000000
58	Movie59	2.000000
59	Movie60	1.000000
60	Movie61	5.000000
61	Movie62	3.000000
62	Movie63	5.000000
63	Movie64	3.000000
64	Movie65	5.000000
65	Movie66	5.000000
66	Movie67	1.000000
67	Movie68	5.000000
68	Movie69	1.000000
69	Movie70	5.000000
70	Movie71	4.000000
71	Movie72	5.000000
72	Movie73	2.000000
73	Movie74	5.000000
74	Movie75	5.000000
75	Movie76	5.000000
76	Movie77	5.000000
77	Movie78	5.000000

78	Movie79	5.000000
79	Movie80	4.000000
80	Movie81	4.416667
81	Movie82	5.000000
82	Movie83	3.000000
83	Movie84	5.000000
84	Movie85	5.000000
85	Movie86	4.666667
86	Movie87	5.000000
87	Movie88	4.000000
88	Movie89	4.578313
89	Movie90	1.833333
90	Movie91	4.578125
91	Movie92	4.772277
92	Movie93	5.000000
93	Movie94	4.333333
94	Movie95	3.333333
95	Movie96	5.000000
96	Movie97	4.800000
97	Movie98	5.000000
98	Movie99	4.000000
99	Movie100	4.000000
100	Movie101	5.000000
101	Movie102	4.000000
102	Movie103	4.562500
103	Movie104	4.500000
104	Movie105	5.000000
105	Movie106	5.000000
106	Movie107	4.000000
107	Movie108	4.666667
108	Movie109	4.384615
109	Movie110	4.875000
110	Movie111	4.641026
111	Movie112	5.000000
112	Movie113	3.750000
113	Movie114	4.428571
114	Movie115	4.000000
115	Movie116	5.000000
116	Movie117	4.727273
117	Movie118	5.000000
118	Movie119	4.375000
119	Movie120	5.000000
120	Movie121	4.250000
121	Movie122	5.000000
122	Movie123	5.000000
123	Movie124	4.750000
124	Movie125	4.800000
125	Movie126	4.500000
126	Movie127	4.111976
127	Movie128	5.000000
128	Movie129	4.000000
129	Movie130	4.500000
130	Movie131	5.000000
131	Movie132	5.000000
132	Movie133	5.000000
133	Movie134	4.000000
134	Movie135	5.000000
135	Movie136	5.000000

136	Movie137	4.000000
137	Movie138	4.076923
138	Movie139	5.000000
139	Movie140	4.833910
140	Movie141	4.142857
141	Movie142	5.000000
142	Movie143	5.000000
143	Movie144	1.000000
144	Movie145	5.000000
145	Movie146	4.000000
146	Movie147	5.000000
147	Movie148	5.000000
148	Movie149	5.000000
149	Movie150	5.000000
150	Movie151	4.500000
151	Movie152	5.000000
152	Movie153	5.000000
153	Movie154	1.000000
154	Movie155	4.000000
155	Movie156	4.000000
156	Movie157	5.000000
157	Movie158	4.818182
158	Movie159	3.000000
159	Movie160	4.666667
160	Movie161	4.600000
161	Movie162	4.866667
162	Movie163	4.461538
163	Movie164	5.000000
164	Movie165	5.000000
165	Movie166	4.000000
166	Movie167	5.000000
167	Movie168	5.000000
168	Movie169	5.000000
169	Movie170	4.333333
170	Movie171	2.000000
171	Movie172	4.500000
172	Movie173	4.733333
173	Movie174	4.750000
174	Movie175	5.000000
175	Movie176	5.000000
176	Movie177	5.000000
177	Movie178	5.000000
178	Movie179	4.714286
179	Movie180	5.000000
180	Movie181	5.000000
181	Movie182	4.733333
182	Movie183	5.000000
183	Movie184	4.823529
184	Movie185	4.791667
185	Movie186	5.000000
186	Movie187	5.000000
187	Movie188	5.000000
188	Movie189	5.000000
189	Movie190	4.714286
190	Movie191	5.000000
191	Movie192	4.500000
192	Movie193	4.571429

193	Movie194	4.714286
194	Movie195	4.000000
195	Movie196	4.888889
196	Movie197	3.800000
197	Movie198	5.000000
198	Movie199	5.000000
199	Movie200	4.625000
200	Movie201	4.333333
201	Movie202	4.333333
202	Movie203	3.000000
203	Movie204	4.375000
204	Movie205	4.628571
205	Movie206	4.923077

Define the top 5 movies with the maximum ratings.

```
[ ] Top_rated_movie.head()
```

	Movies	Rating_sum
126	Movie127	9511.0
139	Movie140	2794.0
15	Movie16	1446.0
102	Movie103	1241.0
28	Movie29	1168.0

Define the top 5 movies with the least audience.

```
[ ] Top_viewed_movie.head()
```

	Movies	Count_sum
126	Movie127	2313
139	Movie140	578
15	Movie16	320
102	Movie103	272
28	Movie29	243

Machine Leaening Model

Recommendation Model:

> Some of the movies hadn't been watched and therefore, are not rated by the users. Netflix would like to take this as an opportunity and build a machine learning recommendation algorithm which provides the ratings for each of the users.

- Divide the data into training and test data
- Build a recommendation model on training data
- Make predictions on the test data

Divide the data into training and test data

```
[ ] X_train, X_test = train_test_split(amazon_df_no_id, test_size = 0.3, random_state = 0)
```

Build a recommendation model on training data

```
[ ]
```

Make predictions on the test data

```
[ ]
```