

```
>>> Loading user data
>>> Loading rating data
>>> Loading movie data
```

(i) Average rating per movie:

```
+-----+-----+
|movie_id|      avg_rating|
+-----+-----+
|    474| 4.252577319587629|
|     29|2.6666666666666665|
|     26| 3.452054794520548|
|    964|3.3333333333333335|
|   1677|          3.0|
|     65|3.5391304347826087|
|    191| 4.163043478260869|
|   1224|2.6666666666666665|
|    558|3.6714285714285713|
|   1010|          3.25|
+-----+-----+
```

only showing top 10 rows

(ii) Top 10 movies with highest average rating:

```
+-----+-----+-----+
|movie_id|title                                     |avg_rating|
+-----+-----+-----+
|1293    |Star Kid (1997)                           |5.0       |
|1500    |Santa with Muscles (1996)                 |5.0       |
|1653    |Entertaining Angels: The Dorothy Day Story (1996)|5.0       |
|1122    |They Made Me a Criminal (1939)            |5.0       |
|1536    |Aiqing wansui (1994)                     |5.0       |
|814     |Great Day in Harlem, A (1994)             |5.0       |
|1201    |Marlene Dietrich: Shadow and Light (1996) |5.0       |
|1189    |Prefontaine (1997)                       |5.0       |
|1599    |Someone Else's America (1995)             |5.0       |
|1467    |Saint of Fort Washington, The (1993)      |5.0       |
+-----+-----+-----+
```

only showing top 10 rows

(iii) Users with at least 50 ratings and their favorite genres:

```
25/06/20 09:37:27 ERROR CodeGenerator: failed to compile: org.codehaus.commons.compiler.CompileException: File 'generated.java', Line 1217, Column 28: Redefinition of parameter "agg_expr_21"
org.codehaus.commons.compiler.CompileException: File 'generated.java', Line 1217, Column 28: Redefinition of parameter "agg_expr_21"
```

```
|user_id|genre0_score|      genre1_score|      genre2_score|      genre3_score|
|genre4_score|      genre5_score|      genre6_score|      genre7_score|
|genre8_score|      genre9_score|      genre10_score|      genre11_score|
|genre12_score|      genre13_score|      genre14_score|      genre15_score|
|genre16_score|      genre17_score|      genre18_score|
```

```
+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+
|      26|          0.0|0.32710280373831774| 0.1588785046728972|0.0186915887850467
28|0.037383177570093455| 0.308411214953271| 0.08411214953271028|
0.0|0.38317757009345793|          0.0|0.009345794392523364|0.0280373831
77570093|          0.0|0.056074766355140186|0.18691588785046728| 0.149532
71028037382|0.27102803738317754| 0.07476635514018691|0.009345794392523364|
|      474|          0.0| 0.1651376146788991|0.10397553516819572|0.0091743119266055
05| 0.04892966360856269|0.25382262996941896| 0.07339449541284404|0.0030581039755
35...| 0.5076452599388379|0.012232415902140673| 0.03669724770642202| 0.033639143
73088685| 0.04892966360856269| 0.06422018348623854| 0.2018348623853211| 0.091743
11926605505|0.18654434250764526| 0.10703363914373089|0.024464831804281346|
|      65|          0.0|          0.1375|          0.0875|          0.01
25|          0.0375|          0.3625|          0.0625|
0.025|          0.575|          0.0125|          0.0|
0.0125|          0.0|          0.0625|          0.2875|
0.1125|          0.1125|          0.1|          0.05|
|      541|          0.0|0.22556390977443608| 0.2706766917293233| 0.150375939849624
05| 0.3233082706766917|0.39097744360902253| 0.03007518796992481|0.0075187969924
81203| 0.2556390977443609| 0.03007518796992481|0.007518796992481203|0.0150375939
84962405| 0.12781954887218044| 0.03007518796992481|0.17293233082706766| 0.142857
14285714285|0.08270676691729323| 0.07518796992481203|0.007518796992481203|
|      222|          0.0| 0.2971576227390181|0.16020671834625322| 0.049095607235142
12| 0.1111111111111111| 0.3385012919896641| 0.07493540051679587|0.0025839793281
65375| 0.330749354005168|0.015503875968992248|0.012919896640826873| 0.069767441
86046512| 0.05167958656330749| 0.03359173126614987|0.15762273901808785| 0.13436
69250645595|0.20155038759689922| 0.06976744186046512| 0.020671834625323|
|      270|          0.0|0.21739130434782608|0.08695652173913043|0.0072463768115942
03|0.014492753623188406| 0.2028985507246377|0.043478260869565216|
0.0| 0.4057971014492754|0.007246376811594203|          0.0| 0.173913043
47826086|0.021739130434782608|0.043478260869565216| 0.3115942028985507| 0.108695
65217391304| 0.3188405797101449|0.050724637681159424|0.007246376811594203|
|      293|          0.0| 0.2422680412371134|0.10309278350515463|0.0309278350515463
93| 0.05154639175257732| 0.2860824742268041| 0.09278350515463918|0.0025773195876
28866| 0.4252577319587629|0.010309278350515464|0.020618556701030927| 0.043814432
98969072| 0.05412371134020619| 0.04639175257731959|0.18814432989690721| 0.118556
70103092783|0.24742268041237114| 0.07474226804123711| 0.02577319587628866|
|      938|          0.0|0.32407407407407407| 0.1574074074074074|0.0462962962962962
94| 0.10185185185185185|0.37037037037037035| 0.05555555555555555|
0.0| 0.2962962962962963|0.018518518518518517|0.009259259259259259|0.0370370370
37037035|0.037037037037037035|0.046296296296296294|0.19444444444444445| 0.13888
8888888889| 0.2962962962962963| 0.07407407407407407|0.009259259259259259|
|      243|          0.0|0.07407407407407407|0.04938271604938271|0.0123456790123456
78| 0.07407407407407407| 0.3333333333333333|0.037037037037037035|0.0123456790123
45678| 0.7283950617283951|0.012345679012345678|          0.0|0.0123456790
12345678|          0.0|0.012345679012345678|0.32098765432098764|0.0246913
58024691357|0.04938271604938271| 0.09876543209876543|0.012345679012345678|
|      367|          0.0|0.29310344827586204| 0.1206896551724138|          0
.0|          0.0| 0.1206896551724138| 0.10344827586206896|
0.0|0.20689655172413793|          0.0|0.017241379310344827| 0.53448275
86206896|          0.0| 0.08620689655172414|0.08620689655172414| 0.258620
68965517243| 0.3448275862068966| 0.05172413793103448|          0.0|
+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+
only showing top 10 rows
```

(iv) Users under 20 years old:

user_id	age	gender	occupation	zip
30	7	M	student	55436
36	19	F	student	93117
52	18	F	student	55105
57	16	M	none	84010
67	17	M	student	60402
68	19	M	student	22904
101	15	M	student	05146
110	19	M	student	77840
142	13	M	other	48118
179	15	M	entertainment	20755

only showing top 10 rows

(v) Users aged 30-40 with occupation 'scientist':

user_id	age	gender	occupation	zip
40	38	M	scientist	27514
71	39	M	scientist	98034
74	39	M	scientist	T8H1N
107	39	M	scientist	60466
183	33	M	scientist	27708
272	33	M	scientist	53706
309	40	M	scientist	70802
337	37	M	scientist	10522
430	38	M	scientist	98199
538	31	M	scientist	21010

only showing top 10 rows

[maria_dev@sandbox-hdp ~]\$