A. Which vendors have the most trips, and what is the total revenue generated by that vendor?

```
[hadoop@ip-172-31-20-206 MapReduceAssignment]$ cat yellow_tripdata_2017-01.csv | python 4a_mapper.py | python 4a_reducer.py
Vendor ID: 2
Total Revenue: 5312899.0
[hadoop@ip-172-31-20-206 MapReduceAssignment]$
```

B. Which pickup location generates the most revenue?

```
[hadoop@ip-172-31-20-206 MapReduceAssignment]$ cat yellow_tripdata_2017-01.csv | python 4b_mapper.py | python 4b_reducer.py
Pickup location id:132  Revenue:12660835.90
Pickup location id:138  Revenue:8966861.41
Pickup location id:161  Revenue:5061992.72
Pickup location id:230  Revenue:5019576.38
Pickup location id:162  Revenue:4740598.36
Pickup location id:186  Revenue:4734371.29
Pickup location id:234  Revenue:4396919.37
Pickup location id:237  Revenue:4378238.17
Pickup location id:236  Revenue:4330922.25
Pickup location id:79  Revenue:4226206.24
Pickup location id:170  Revenue:4182349.08
Pickup location id:48  Revenue:4007370.97
Pickup location id:239  Revenue:3737682.25
Pickup location id:163  Revenue:3533381.98
Pickup location id:142  Revenue:3527054.07
Pickup location id:164  Revenue:3200852.44
Pickup location id:107  Revenue:3127727.94
Pickup location id:68  Revenue:3095397.10
Pickup location id:249  Revenue:2850417.24
Pickup location id:141  Revenue:2708975.95
Pickup location id:231  Revenue:2707164.91
Pickup location id:264  Revenue:2690819.60
Pickup location id:100  Revenue:2596307.08
Pickup location id:229  Revenue:2413119.56
Pickup location id:238  Revenue:2359810.38
Pickup location id:90  Revenue:2353495.90
Pickup location id:140  Revenue:2325446.93
Pickup location id:148  Revenue:2224654.37
Pickup location id:113  Revenue:2178950.81
Pickup location id:263  Revenue:2166061.83
Pickup location id:246  Revenue:1925893.55
Pickup location id:114  Revenue:1924494.35
Pickup location id:233  Revenue:1760674.69
Pickup location id:137  Revenue:1661158.07
Pickup location id:13  Revenue:1629016.97
Pickup location id:158  Revenue:1574902.23
Pickup location id:144  Revenue:1573874.41
Pickup location id:262  Revenue:1553326.54
```

```
...Assignment — ssh -i ~/Downloads/vockey.pem hadoop@ec2-3-95-59-164.compute-1.amazonaws.com          - — -zsh                    ... +
Pickup location id:128  Revenue:981.85
Pickup location id:101  Revenue:970.29
Pickup location id:117  Revenue:962.58
Pickup location id:171  Revenue:898.33
Pickup location id:254  Revenue:897.08
Pickup location id:139  Revenue:890.97
Pickup location id:9  Revenue:855.06
Pickup location id:19  Revenue:821.99
Pickup location id:240  Revenue:799.61
Pickup location id:175  Revenue:732.68
Pickup location id:96  Revenue:714.73
Pickup location id:31  Revenue:678.43
Pickup location id:3  Revenue:655.01
Pickup location id:15  Revenue:647.84
Pickup location id:38  Revenue:569.60
Pickup location id:222  Revenue:568.56
Pickup location id:122  Revenue:559.91
Pickup location id:73  Revenue:551.04
Pickup location id:81  Revenue:503.78
Pickup location id:32  Revenue:449.03
Pickup location id:23  Revenue:388.98
Pickup location id:86  Revenue:374.69
Pickup location id:183  Revenue:333.89
Pickup location id:2  Revenue:330.26
Pickup location id:30  Revenue:323.95
Pickup location id:214  Revenue:309.64
Pickup location id:201  Revenue:291.74
Pickup location id:245  Revenue:290.47
Pickup location id:206  Revenue:263.20
Pickup location id:111  Revenue:230.19
Pickup location id:184  Revenue:228.63
Pickup location id:84  Revenue:216.40
Pickup location id:251  Revenue:204.78
Pickup location id:172  Revenue:197.40
Pickup location id:204  Revenue:103.22
Pickup location id:118  Revenue:103.15
Pickup location id:58  Revenue:101.89
Pickup location id:46  Revenue:101.40
Pickup location id:105  Revenue:94.75
Pickup location id:187  Revenue:93.50
Pickup location id:5  Revenue:69.60
Pickup location id:59  Revenue:49.35
Pickup location id:176  Revenue:30.30
Pickup location id:109  Revenue:14.60
Pickup location id:110  Revenue:6.80
[hadoop@ip-172-31-20-206 MapReduceAssignment]$
```

C. What are the different payment types used by customers and their count? The final results should be in a sorted format.

```
[hadoop@ip-172-31-20-206 MapReduceAssignment]$ cat yellow_tripdata_2017-01.csv | python 4c_mapper.py | python 4c_reducer.py
Payment_Type:1 Count=6506189
Payment_Type:2 Count=3144926
Payment_Type:3 Count=46257
Payment_Type:4 Count=13447
Payment_Type:5 Count=1
[hadoop@ip-172-31-20-206 MapReduceAssignment]$
```

D. What is the average trip time for different pickup locations?

```
[hadoop@ip-172-31-20-206 MapReduceAssignment]$ cat yellow_tripdata_2017-01.csv | python 4d_mapper.py | python 4d_reducer.py
The average trip time for pickup location 140 is 13.38 minutes
The average trip time for pickup location 237 is 11.70 minutes
The average trip time for pickup location 41 is 11.64 minutes
The average trip time for pickup location 48 is 13.79 minutes
The average trip time for pickup location 236 is 12.04 minutes
The average trip time for pickup location 238 is 12.33 minutes
The average trip time for pickup location 239 is 12.28 minutes
The average trip time for pickup location 246 is 14.46 minutes
The average trip time for pickup location 114 is 14.79 minutes
The average trip time for pickup location 263 is 11.25 minutes
The average trip time for pickup location 161 is 14.43 minutes
The average trip time for pickup location 163 is 14.68 minutes
The average trip time for pickup location 107 is 13.27 minutes
The average trip time for pickup location 186 is 15.45 minutes
The average trip time for pickup location 262 is 12.15 minutes
The average trip time for pickup location 249 is 13.95 minutes
The average trip time for pickup location 170 is 13.75 minutes
The average trip time for pickup location 211 is 15.02 minutes
The average trip time for pickup location 87 is 18.43 minutes
The average trip time for pickup location 151 is 12.30 minutes
The average trip time for pickup location 142 is 12.89 minutes
The average trip time for pickup location 80 is 13.43 minutes
The average trip time for pickup location 189 is 15.66 minutes
The average trip time for pickup location 234 is 13.82 minutes
The average trip time for pickup location 143 is 12.46 minutes
The average trip time for pickup location 132 is 38.26 minutes
The average trip time for pickup location 229 is 12.72 minutes
The average trip time for pickup location 141 is 11.71 minutes
The average trip time for pickup location 50 is 13.35 minutes
The average trip time for pickup location 90 is 12.97 minutes
The average trip time for pickup location 113 is 13.79 minutes
The average trip time for pickup location 68 is 14.05 minutes
The average trip time for pickup location 145 is 14.65 minutes
The average trip time for pickup location 233 is 14.24 minutes
```

```
The average trip time for pickup location 149 is 39.11 minutes
The average trip time for pickup location 108 is 12.65 minutes
The average trip time for pickup location 210 is 15.02 minutes
The average trip time for pickup location 139 is 14.46 minutes
The average trip time for pickup location 101 is 9.76 minutes
The average trip time for pickup location 73 is 8.09 minutes
The average trip time for pickup location 19 is 8.20 minutes
The average trip time for pickup location 124 is 20.72 minutes
The average trip time for pickup location 201 is 2.81 minutes
The average trip time for pickup location 117 is 33.36 minutes
The average trip time for pickup location 154 is 25.31 minutes
The average trip time for pickup location 222 is 23.93 minutes
The average trip time for pickup location 203 is 30.39 minutes
The average trip time for pickup location 111 is 11.24 minutes
The average trip time for pickup location 191 is 11.09 minutes
The average trip time for pickup location 3 is 10.09 minutes
The average trip time for pickup location 105 is 19.41 minutes
The average trip time for pickup location 38 is 16.90 minutes
The average trip time for pickup location 207 is 8.32 minutes
The average trip time for pickup location 9 is 37.35 minutes
The average trip time for pickup location 6 is 5.71 minutes
The average trip time for pickup location 98 is 13.59 minutes
The average trip time for pickup location 115 is 14.63 minutes
The average trip time for pickup location 86 is 16.18 minutes
The average trip time for pickup location 84 is 0.32 minutes
The average trip time for pickup location 253 is 28.94 minutes
The average trip time for pickup location 251 is 0.68 minutes
The average trip time for pickup location 2 is 31.07 minutes
The average trip time for pickup location 221 is 13.37 minutes
The average trip time for pickup location 187 is 0.25 minutes
The average trip time for pickup location 245 is 1.08 minutes
The average trip time for pickup location 214 is 6.91 minutes
The average trip time for pickup location 172 is 20.09 minutes
The average trip time for pickup location 206 is 16.90 minutes
The average trip time for pickup location 184 is 28.49 minutes
The average trip time for pickup location 5 is 0.37 minutes
The average trip time for pickup location 23 is 12.77 minutes
The average trip time for pickup location 59 is 16.98 minutes
The average trip time for pickup location 110 is 3.18 minutes
The average trip time for pickup location 109 is 3.79 minutes
The average trip time for pickup location 204 is 1.92 minutes
The average trip time for pickup location 176 is 9.70 minutes
The average trip time for pickup location 58 is 5.28 minutes
The average trip time for pickup location 118 is 5.04 minutes
The average trip time for pickup location 30 is 136.47 minutes
[hadoop@ip-172-31-20-206 MapReduceAssignment]$
```

E. Calculate the average tips to revenue ratio of the drivers for different locations in sorted format.

```
Pickup Location:43      Average ratio: 0.10
Pickup Location:265     Average ratio: 0.10
Pickup Location:36      Average ratio: 0.10
Pickup Location:132     Average ratio: 0.09
Pickup Location:24      Average ratio: 0.09
Pickup Location:34      Average ratio: 0.09
Pickup Location:100     Average ratio: 0.09
Pickup Location:123     Average ratio: 0.09
Pickup Location:191     Average ratio: 0.09
Pickup Location:108     Average ratio: 0.09
Pickup Location:45      Average ratio: 0.09
Pickup Location:98      Average ratio: 0.09
Pickup Location:190     Average ratio: 0.09
Pickup Location:97      Average ratio: 0.09
Pickup Location:63      Average ratio: 0.09
Pickup Location:210     Average ratio: 0.09
Pickup Location:96      Average ratio: 0.09
Pickup Location:219     Average ratio: 0.09
Pickup Location:257     Average ratio: 0.09
Pickup Location:216     Average ratio: 0.09
Pickup Location:37      Average ratio: 0.09
Pickup Location:180     Average ratio: 0.09
Pickup Location:130     Average ratio: 0.09
Pickup Location:49      Average ratio: 0.09
Pickup Location:124     Average ratio: 0.08
Pickup Location:228     Average ratio: 0.08
Pickup Location:195     Average ratio: 0.08
Pickup Location:128     Average ratio: 0.08
Pickup Location:202     Average ratio: 0.08
Pickup Location:35      Average ratio: 0.08
Pickup Location:120     Average ratio: 0.08
Pickup Location:133     Average ratio: 0.08
Pickup Location:145     Average ratio: 0.08
Pickup Location:75      Average ratio: 0.08
Pickup Location:222     Average ratio: 0.08
Pickup Location:9       Average ratio: 0.08
Pickup Location:61      Average ratio: 0.08
Pickup Location:41      Average ratio: 0.08
Pickup Location:28      Average ratio: 0.08
Pickup Location:223     Average ratio: 0.08
Pickup Location:77      Average ratio: 0.08
Pickup Location:146     Average ratio: 0.08
Pickup Location:12      Average ratio: 0.08
Pickup Location:244     Average ratio: 0.07
Pickup Location:157     Average ratio: 0.07
Pickup Location:116     Average ratio: 0.07
```

```
[hadoop@ip-172-31-20-206 MapReduceAssignment]$ cat yellow_tripdata_2017-01.csv | python 4e_mapper.py | python 4e_reducer.py
Pickup Location:30      Average ratio: 0.23
Pickup Location:176     Average ratio: 0.17
Pickup Location:111     Average ratio: 0.15
Pickup Location:172     Average ratio: 0.15
Pickup Location:117     Average ratio: 0.14
Pickup Location:15      Average ratio: 0.13
Pickup Location:64      Average ratio: 0.13
Pickup Location:138     Average ratio: 0.13
Pickup Location:175     Average ratio: 0.12
Pickup Location:58      Average ratio: 0.12
Pickup Location:125     Average ratio: 0.12
Pickup Location:13      Average ratio: 0.12
Pickup Location:87      Average ratio: 0.12
Pickup Location:252     Average ratio: 0.12
Pickup Location:234     Average ratio: 0.12
Pickup Location:16      Average ratio: 0.12
Pickup Location:249     Average ratio: 0.12
Pickup Location:118     Average ratio: 0.12
Pickup Location:246     Average ratio: 0.11
Pickup Location:150     Average ratio: 0.11
Pickup Location:52      Average ratio: 0.11
Pickup Location:158     Average ratio: 0.11
Pickup Location:221     Average ratio: 0.11
Pickup Location:231     Average ratio: 0.11
Pickup Location:107     Average ratio: 0.11
Pickup Location:90      Average ratio: 0.11
Pickup Location:114     Average ratio: 0.11
Pickup Location:154     Average ratio: 0.11
Pickup Location:113     Average ratio: 0.11
Pickup Location:66      Average ratio: 0.11
Pickup Location:79      Average ratio: 0.11
Pickup Location:155     Average ratio: 0.11
Pickup Location:162     Average ratio: 0.11
Pickup Location:170     Average ratio: 0.11
Pickup Location:194     Average ratio: 0.11
Pickup Location:33      Average ratio: 0.11
Pickup Location:255     Average ratio: 0.11
Pickup Location:8       Average ratio: 0.11
Pickup Location:211     Average ratio: 0.11
Pickup Location:233     Average ratio: 0.11
Pickup Location:144     Average ratio: 0.11
Pickup Location:54      Average ratio: 0.11
Pickup Location:148     Average ratio: 0.11
Pickup Location:224     Average ratio: 0.11
Pickup Location:88      Average ratio: 0.11
```

```
Pickup Location:73      Average ratio: 0.04
Pickup Location:218     Average ratio: 0.04
Pickup Location:82      Average ratio: 0.04
Pickup Location:86      Average ratio: 0.04
Pickup Location:78      Average ratio: 0.04
Pickup Location:20      Average ratio: 0.04
Pickup Location:126     Average ratio: 0.04
Pickup Location:85      Average ratio: 0.04
Pickup Location:242     Average ratio: 0.04
Pickup Location:60      Average ratio: 0.04
Pickup Location:129     Average ratio: 0.04
Pickup Location:92      Average ratio: 0.04
Pickup Location:165     Average ratio: 0.04
Pickup Location:205     Average ratio: 0.04
Pickup Location:83      Average ratio: 0.04
Pickup Location:46      Average ratio: 0.04
Pickup Location:81      Average ratio: 0.04
Pickup Location:22      Average ratio: 0.04
Pickup Location:26      Average ratio: 0.04
Pickup Location:235     Average ratio: 0.03
Pickup Location:18      Average ratio: 0.03
Pickup Location:119     Average ratio: 0.03
Pickup Location:2       Average ratio: 0.03
Pickup Location:168     Average ratio: 0.03
Pickup Location:173     Average ratio: 0.03
Pickup Location:147     Average ratio: 0.03
Pickup Location:248     Average ratio: 0.03
Pickup Location:169     Average ratio: 0.03
Pickup Location:206     Average ratio: 0.03
Pickup Location:69      Average ratio: 0.03
Pickup Location:241     Average ratio: 0.03
Pickup Location:32      Average ratio: 0.03
Pickup Location:136     Average ratio: 0.03
Pickup Location:159     Average ratio: 0.03
Pickup Location:240     Average ratio: 0.03
Pickup Location:153     Average ratio: 0.02
Pickup Location:47      Average ratio: 0.02
Pickup Location:167     Average ratio: 0.02
Pickup Location:6       Average ratio: 0.02
Pickup Location:212     Average ratio: 0.02
Pickup Location:254     Average ratio: 0.02
Pickup Location:207     Average ratio: 0.02
Pickup Location:38      Average ratio: 0.02
Pickup Location:182     Average ratio: 0.02
Pickup Location:94      Average ratio: 0.02
[hadoop@ip-172-31-20-206 MapReduceAssignment]$
```

F. How does revenue vary over time? Calculate the average trip revenue per month -
    analyzing it by hour of the day (day vs night) and the day of the week (weekday vs
    weekend).

```
[hadoop@ip-172-31-20-206 MapReduceAssignment]$ cat yellow_tripdata_2017-01.csv | python 4f_mapper.py | python 4f_reducer.py
1-2017-night-weekend    Average revenue: $16.33
1-2017-day-weekend      Average revenue: $14.27
1-2017-night-weekday    Average revenue: $15.93
1-2017-day-weekday      Average revenue: $15.55
[hadoop@ip-172-31-20-206 MapReduceAssignment]$
```

```
[hadoop@ip-172-31-20-206 MapReduceAssignment]$ cat yellow_tripdata_2017-01.csv | python 4f_mapper.py | python 4f1_reducer.py
1-2017-day-weekday      Average revenue: $15.55
1-2017-day-weekend      Average revenue: $14.27
1-2017-night-weekday    Average revenue: $15.93
1-2017-night-weekend    Average revenue: $16.33
[hadoop@ip-172-31-20-206 MapReduceAssignment]$
```