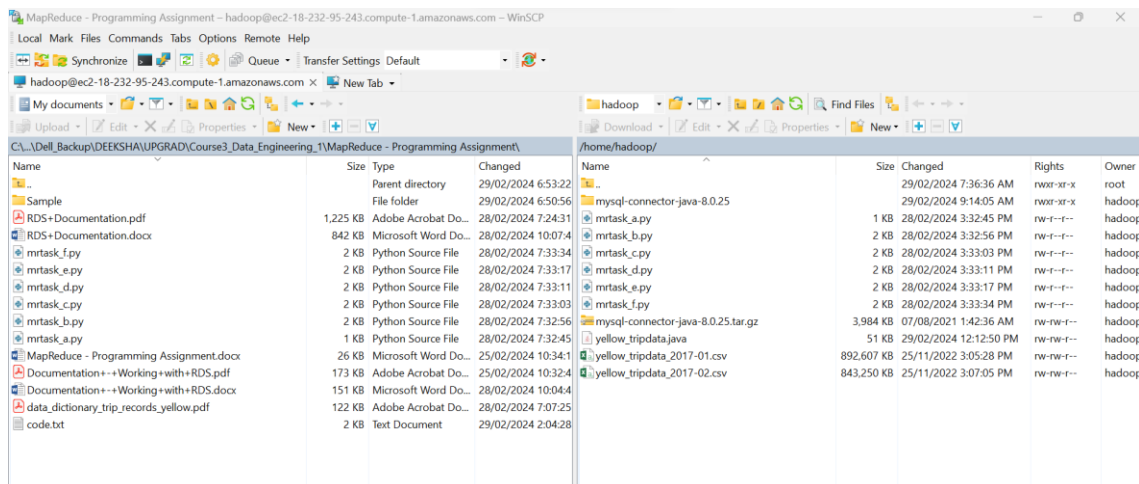
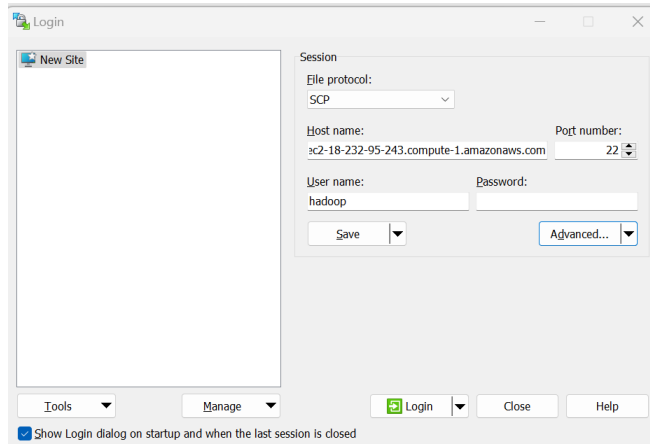


## MapReduce Tasks:

### Task 4. Using MapReduce to perform data analysis on files downloaded to the EMR instance

- Developed 6 MapReduce Jobs with Python Scripts for each question number
- Used WinSCP tool to transfer these scripts to Hadoop



Executed command to give read, write, and execute permissions on all the transferred files  
chmod 770 filename

```
[hadoop@ip-172-31-47-204 ~]$ chmod 770 mrtask_a.py
[hadoop@ip-172-31-47-204 ~]$ chmod 770 mrtask_b.py
[hadoop@ip-172-31-47-204 ~]$ chmod 770 mrtask_c.py
[hadoop@ip-172-31-47-204 ~]$ chmod 770 mrtask_d.py
[hadoop@ip-172-31-47-204 ~]$ chmod 770 mrtask_e.py
[hadoop@ip-172-31-47-204 ~]$ chmod 770 mrtask_f.py
[hadoop@ip-172-31-47-204 ~]$ chmod 770 batch_ingest.py
[hadoop@ip-172-31-47-204 ~]$ ls -lrt
total 3611176
-rw-rw-r-- 1 hadoop hadoop 4079310 Aug  7  2021 mysql-connector-java-8.0.25.tar.gz
-rw-rw-r-- 1 hadoop hadoop 914029540 Nov 25  2022 yellow_tripdata_2017-01.csv
-rw-rw-r-- 1 hadoop hadoop 863487050 Nov 25  2022 yellow_tripdata_2017-02.csv
-rw-rw-r-- 1 hadoop hadoop 969809025 Nov 25  2022 yellow_tripdata_2017-03.csv
-rw-rw-r-- 1 hadoop hadoop 946349441 Nov 25  2022 yellow_tripdata_2017-04.csv
-rwxrwx--- 1 hadoop hadoop 1858 Feb 28 15:32 batch_ingest.py
-rwxrwx--- 1 hadoop hadoop 1022 Feb 28 15:32 mrtask_a.py
-rwxrwx--- 1 hadoop hadoop 1038 Feb 28 15:32 mrtask_b.py
-rwxrwx--- 1 hadoop hadoop 1389 Feb 28 15:33 mrtask_c.py
-rwxrwx--- 1 hadoop hadoop 1834 Feb 28 15:33 mrtask_d.py
-rwxrwx--- 1 hadoop hadoop 1360 Feb 28 15:33 mrtask_e.py
-rwxrwx--- 1 hadoop hadoop 1584 Feb 28 15:33 mrtask_f.py
drwxr-xr-x 3 hadoop hadoop 149 Feb 29 09:14 mysql-connector-java-8.0.25
-rw-rw-r-- 1 hadoop hadoop 51862 Feb 29 12:12 yellow_tripdata.java
```

Set up the environment for MR jobs execution:

### Installed the mrjob library

```
[root@ip-172-31-47-204 ~]# pip install mrjob
WARNING: Running pip install with root privileges is generally not a good idea. Try `pip3 install --user` instead.
Collecting mrjob
  Downloading https://files.pythonhosted.org/packages/8e/58/fc28ab743aba16e90736ad4e29694bd2adaf7b879376ff149306d50c4e90/mrjob-0.7.4-py2.py3-none-any.whl (439kB)
    100% |#####| 440kB 789kB/s
Requirement already satisfied: PyYAML>=3.10 in /usr/local/lib64/python3.7/site-packages (from mrjob)
Installing collected packages: mrjob
Successfully installed mrjob-0.7.4
```

For Questions->a to f, used yellow\_tripdata\_2017-03.csv dataset

**Q.a) Which vendors have the most trips, and what is the total revenue generated by that vendor?**

**Ans. Vendor 2 had the most trips with total revenue of 91682368.32**

```
root@ip-172-31-47-204:~
[root@ip-172-31-47-204 ~]#
[root@ip-172-31-47-204 ~]# python mrtask_a.py yellow_tripdata_2017-03.csv > mrtask_a_2017_03_output.txt
No configs found; falling back on auto-configuration
No configs specified for inline runner
Creating temp directory /root/mrtask_a.root.20240229.181746.734059
Running step 1 of 2...
Running step 2 of 2...
job output is in /root/mrtask_a.root.20240229.181746.734059/output
Streaming final output from /root/mrtask_a.root.20240229.181746.734059/output...
Removing temp directory /root/mrtask_a.root.20240229.181746.734059...
[root@ip-172-31-47-204 ~]#

[root@ip-172-31-47-204 ~]# cat mrtask_a_2017_03_output.txt
"2"      91682368.32536966
[root@ip-172-31-47-204 ~]#
```

**Q.b) Which pickup location generates the most revenue?**

**Ans. Pickup location that generated the most revenue is "132"**

```
root@ip-172-31-47-204:~
[root@ip-172-31-47-204 ~]#
[root@ip-172-31-47-204 ~]#
[root@ip-172-31-47-204 ~]# python mrtask_b.py yellow_tripdata_2017-03.csv > mrtask_b_2017_03_output.txt
No configs found; falling back on auto-configuration
No configs specified for inline runner
Creating temp directory /root/mrtask_b.root.20240229.181605.049750
Running step 1 of 2...
Running step 2 of 2...
job output is in /root/mrtask_b.root.20240229.181605.049750/output
Streaming final output from /root/mrtask_b.root.20240229.181605.049750/output...
Removing temp directory /root/mrtask_b.root.20240229.181605.049750...
[root@ip-172-31-47-204 ~]#

[root@ip-172-31-47-204 ~]# cat mrtask_b_2017_03_output.txt
"132"    13307409.48001407
[root@ip-172-31-47-204 ~]#
```

**Q.c) What are the different payment types used by customers and their count? The final**

results should be in a sorted format.

Ans. The different payment types used by customers and their count are as follows:

Payment Type	Count
"1"	6994699
"2"	3231928
"3"	53815
"4"	14999

```
root@ip-172-31-47-204:~  
[root@ip-172-31-47-204 ~]#  
[root@ip-172-31-47-204 ~]# python mrtask_c.py yellow_tripdata_2017-03.csv > mrtask_c_2017_03_output.txt  
No configs found; falling back on auto-configuration  
No configs specified for inline runner  
Creating temp directory /root/mrtask_c.root.20240229.182011.744624  
Running step 1 of 3...  
Running step 2 of 3...  
Running step 3 of 3...  
job output is in /root/mrtask_c.root.20240229.182011.744624/output  
Streaming final output from /root/mrtask_c.root.20240229.182011.744624/output...  
Removing temp directory /root/mrtask_c.root.20240229.182011.744624...  
[root@ip-172-31-47-204 ~]#  
[root@ip-172-31-47-204 ~]# cat mrtask_c_2017_03_output.txt  
"1"      6994699  
"2"      3231928  
"3"      53815  
"4"      14999
```

Q.d) What is the average trip time for different pickup locations?

Ans. The average trip time for different pickup locations are attached in the file



"1"	6.105276981852914
"10"	56.11884170550665
"100"	15.604376259723034
"101"	12.203763440860214
"102"	42.82314814814816
"105"	18.590277777777778
"106"	15.196617436874702
"107"	14.326443236908414
"108"	18.1587962962963
"109"	203.13809523809525
"11"	10.917836257309942
"111"	15.513333333333332
"112"	14.553648561025495
"113"	15.095411481968693
"114"	16.23875211075014
"115"	11.043333333333335
"116"	15.671463848776575
"117"	9.8
"118"	2.986111111111111
"119"	13.447759103641456

"12"	27.171749165193486
"120"	13.396666666666665
"121"	13.307905982905982
"122"	57.585964912280694
"123"	12.791750841750844
"124"	16.08425287356322
"125"	16.281976174199638
"126"	16.042989417989418
"127"	15.29647435897436
"128"	14.925757575757574
"129"	13.91558404110307
"13"	19.149716258261186
"130"	38.636641382216595
"131"	10.526371308016879
"132"	44.534667046862694
"133"	16.76554997208264
"134"	18.376689381033508
"135"	14.40541922290389
"136"	12.584346846846849
"137"	14.060476623848091
"138"	37.442771056676094
"139"	23.40952380952381
"14"	15.749354243542435
"140"	14.278674909240369
"141"	12.391378608082539
"142"	13.910121859205608
"143"	13.206644272409084
"144"	16.929347185768396
"145"	12.501812450748616
"146"	15.172756058431132
"147"	16.578253968253968
"148"	16.584931917192876
"149"	10.220402298850576
"15"	11.010185185185186
"150"	16.546376811594204
"151"	13.512899896800828
"152"	15.19777807305047
"153"	12.657235142118862
"154"	33.008333333333334
"155"	22.139869281045748
"156"	4.492857142857142
"157"	23.1468121968122
"158"	16.683626771528726
"159"	16.156171229900043
"16"	12.13918918918919
"160"	22.163055555555555
"161"	15.558085142830453
"162"	15.217090260480004
"163"	15.765110147709361
"164"	15.66443342276047

"165"	20.56750841750842
"166"	14.858510043132219
"167"	12.437793427230048
"168"	11.897469269703542
"169"	13.474629629629629
"17"	12.939188672024498
"170"	15.023953246154356
"171"	13.095128205128207
"172"	13.606666666666666
"173"	20.23913043478261
"174"	17.088115942028985
"175"	17.32816091954023
"176"	35.36333333333333
"177"	19.516613418530355
"178"	6.0441091954022985
"179"	14.375353943276236
"18"	12.14846153846154
"180"	23.321614583333336
"181"	15.907067007578375
"182"	14.6375
"183"	12.124166666666666
"184"	25.396666666666665
"185"	31.498222222222218
"186"	17.204260372929703
"187"	17.416666666666668
"188"	15.727673796791448
"189"	14.931175059952032
"19"	10.288888888888888
"190"	15.164267676767677
"191"	14.482564102564103
"192"	28.38516819571865
"193"	17.529686693062487
"194"	29.225446998722866
"195"	19.768767990788714
"196"	19.87514619883041
"197"	12.270531400966183
"198"	13.256589147286821
"199"	19.166666666666668
"2"	31.054761904761904
"20"	14.162921348314606
"200"	11.093781094527364
"201"	12.390625
"202"	17.537518037518037
"203"	19.530797101449274
"204"	7.261666666666668
"205"	53.41709401709401
"206"	16.101851851851855
"207"	6.2888795045045045
"208"	14.488125
"209"	18.707017543859653

"21"	18.813793103448276
"210"	14.301785714285712
"211"	17.012496046948232
"212"	12.277672955974841
"213"	16.26485042735043
"214"	8.5
"215"	47.343349106203995
"216"	27.39974437627812
"217"	19.36658950617284
"218"	25.640579710144927
"219"	43.529047619047624
"22"	29.752331002331
"220"	12.537727272727274
"221"	25.062820512820515
"222"	35.830303030303035
"223"	14.028142544485684
"224"	13.60093953704937
"225"	12.91793006636039
"226"	16.998068871844033
"227"	14.764018691588785
"228"	13.832922173274593
"229"	13.568517030272616
"23"	16.120833333333334
"230"	16.891338432887245
"231"	16.9992376857944
"232"	16.809078430444455
"233"	15.595958464012563
"234"	15.265466642469073
"235"	13.28462837837838
"236"	12.944847684221873
"237"	12.370872689480343
"238"	12.96824867660725
"239"	13.144993406845405
"24"	13.869919911805713
"240"	12.571527777777776
"241"	12.262056737588651
"242"	12.809098228663444
"243"	15.543206444254356
"244"	16.231835236764983
"245"	0.9166666666666666
"246"	16.333486238532096
"247"	16.497478220999543
"248"	18.274906367041197
"249"	15.072414367953192
"25"	15.079508640077417
"250"	16.01778606965174
"251"	12.501051051051054
"252"	18.974590163934423
"253"	15.904444444444444
"254"	41.800200803212846

"255"	15.129783635210382
"256"	16.5599060956413
"257"	20.53894691035683
"258"	16.89525386313466
"259"	33.590740740740735
"26"	11.181123388581952
"260"	15.042695201798104
"261"	21.87678272474335
"262"	13.009731723108985
"263"	12.210742434885523
"264"	14.92916963514813
"265"	11.486002886002883
"27"	3.591666666666667
"28"	27.806325632563258
"29"	34.497692307692304
"3"	60.41428571428572
"30"	20.933333333333334
"31"	17.613821138211385
"32"	15.874666666666666
"33"	18.17328661767219
"34"	14.2520944741533
"35"	22.245520833333337
"36"	16.046817625458992
"37"	16.217501017501014
"38"	28.897058823529417
"39"	18.684934086629003
"4"	14.74015566218378
"40"	15.519469776249565
"41"	13.175797583826423
"42"	13.393007571579
"43"	15.305562425593424
"44"	10.887500000000001
"45"	18.507105733555647
"46"	21.533333333333333
"47"	20.015294117647063
"48"	15.062742407526823
"49"	13.62706245461147
"5"	35.0625
"50"	15.72983890425194
"51"	24.8797619047619
"52"	17.87568097347385
"53"	18.165151515151514
"54"	16.60330303030303
"55"	14.587458745874589
"56"	27.260372608257807
"57"	18.451724137931034
"58"	5.412962962962962
"59"	2.9166666666666665
"6"	9.276829268292685
"60"	12.933928571428572

"61"	16.112351914203767
"62"	13.466998671978754
"63"	20.96495258935084
"64"	17.030952380952378
"65"	16.48728611130251
"66"	17.331985383678443
"67"	10.18055555555554
"68"	15.95617606984292
"69"	14.446176628971461
"7"	13.409697415601121
"70"	25.418271691090954
"71"	17.520917225950786
"72"	26.093424036281178
"73"	33.49
"74"	12.538660697986355
"75"	13.022578643438315
"76"	19.246626984126983
"77"	18.64166666666667
"78"	17.36
"79"	15.416023310583807
"8"	17.4796568627451
"80"	15.720131101813113
"81"	14.933809523809526
"82"	14.795396505376344
"83"	15.461774461028194
"84"	18.5
"85"	19.640858208955223
"86"	4.910256410256411
"87"	20.577113508814527
"88"	21.77548703545406
"89"	19.37323473282443
"9"	74.27303921568627
"90"	14.498818951803642
"91"	22.116152263374484
"92"	23.007697456492636
"93"	36.96678743961352
"94"	11.879012345679012
"95"	17.892524916943522
"96"	14.138690476190476
"97"	15.308801450019969
"98"	11.727536231884056



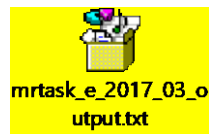
```

root@ip-172-31-47-204:~
[root@ip-172-31-47-204 ~]#
[root@ip-172-31-47-204 ~]# cat mrtask_d_2017_03_output.txt
"1"      6.105276981852914
"10"     56.11884170550665
"100"    15.604376259723034
"101"    12.203763440860214
"102"    42.82314814814816
"105"    18.590277777777778
"106"    15.196617436874702
"107"    14.326443236908414
"108"    18.1587962962963
"109"    203.13809523809525
"11"     10.917836257309942
"111"    15.513333333333332
"112"    14.553648561025495
"113"    15.095411481968693
"114"    16.23875211075014
"115"    11.043333333333335
"116"    15.671463848776575
"117"    9.8
"118"    2.986111111111111

```

Q.e) Calculate the average tips to revenue ratio of the drivers for different pickup locations in sorted format.

Ans. The average tips to revenue ratio of the drivers for different pickup locations are attached in the file



```

"1"      0.1207141821477579
"10"     0.10818785706737385
"100"    0.1036698554215106
"101"    0.11272386860071483
"102"    0.10055306633681427
"105"    0.09171366594360085
"106"    0.10978934961315243
"107"    0.12060767877711343
"108"    0.16459569486203868
"109"    0.2727411419199817
"11"     0.08292554532378368
"111"    0.09254458568885279
"112"    0.10773394359303982
"113"    0.1198356138476097
"114"    0.11911520367510238
"115"    0.06954984925667948
"116"    0.0891850662923476
"117"    0.043243243243243246
"118"    0.06123568538709457

```

"119" 0.07048709650651401  
"12" 0.09293556386304279  
"120" 0.10883550438515564  
"121" 0.11741420610866483  
"122" 0.1387652122291481  
"123" 0.08960279415384258  
"124" 0.08453696528256818  
"125" 0.12518769144404782  
"126" 0.09259487147933362  
"127" 0.08441883718898342  
"128" 0.0996911080691746  
"129" 0.060201630039305176  
"13" 0.12714951929426124  
"130" 0.10414607569293091  
"131" 0.10417221067470275  
"132" 0.10575125324839516  
"133" 0.10062634626777764  
"134" 0.09832702825954157  
"135" 0.08685750264472232  
"136" 0.03613849315545647  
"137" 0.11313062136906334  
"138" 0.1317185832146078  
"139" 0.06289872602766453  
"14" 0.09728965671976669  
"140" 0.11132977364619498  
"141" 0.11224772836685043  
"142" 0.10943407036321535  
"143" 0.11385544735623641  
"144" 0.11534246435160811  
"145" 0.09385839052238519  
"146" 0.08701042916776668  
"147" 0.0407104414538495  
"148" 0.11552592223584679  
"149" 0.07203300930134972  
"15" 0.08857607480435004  
"150" 0.08822992437709598  
"151" 0.10653807861147498  
"152" 0.08032142329328529  
"153" 0.038389942412793886  
"154" 0.07685038639697143  
"155" 0.07278967134703021  
"156" 0.07347670250896057  
"157" 0.10501594923339569

"158" 0.12040536538332194  
"159" 0.04321460191532224  
"16" 0.15964166650420136  
"160" 0.11332917205196806  
"161" 0.11910074531590943  
"162" 0.12317702599821087  
"163" 0.11388391450287717  
"164" 0.11356375521257311  
"165" 0.07474841653471122  
"166" 0.11300512805877584  
"167" 0.03467827275684446  
"168" 0.03970975383351151  
"169" 0.0414459260514526  
"17" 0.08139413432067066  
"170" 0.11967256252276448  
"171" 0.08945571641402163  
"172" 0.18060729407061474  
"173" 0.06020800402356662  
"174" 0.05603043490423155  
"175" 0.16555962907051971  
"176" 0.16006448641179177  
"177" 0.059742733123308786  
"178" 0.03763788048031276  
"179" 0.08576712484744223  
"18" 0.02333018387964241  
"180" 0.12101155295619763  
"181" 0.11377482959640933  
"182" 0.060489365909785486  
"183" 0.027458002899292233  
"184" 0.16455893832943014  
"185" 0.11617778281376145  
"186" 0.11058511154798138  
"187" 0.23076923076923078  
"188" 0.07111055908517584  
"189" 0.11508590247613698  
"19" 0.04503657850309511  
"190" 0.10714460915480402  
"191" 0.1963891248937978  
"192" 0.06695344073926153  
"193" 0.04982252686861339  
"194" 0.12353339789282343  
"195" 0.11921275351973504  
"196" 0.08615093221135418

"197" 0.08975374732334052  
"198" 0.09483434515131679  
"199" 0.14926695870154202  
"2" 0.09762134750060372  
"20" 0.0460440256411774  
"200" 0.11615574722990556  
"201" 0.23398594719967258  
"202" 0.09243860905697258  
"203" 0.067581902633449  
"204" 0.027348010502928698  
"205" 0.09300381378536113  
"206" 0.1270032305061126  
"207" 0.05570345080062397  
"208" 0.0682832201745878  
"209" 0.11693289972505848  
"21" 0.050033559889741105  
"210" 0.08695098338938143  
"211" 0.1168006199265725  
"212" 0.0895001994474567  
"213" 0.09725773573506433  
"214" 0.05414483407228269  
"215" 0.11303285020251182  
"216" 0.09869999868671986  
"217" 0.0757981251916151  
"218" 0.05608133485551089  
"219" 0.09630096969726427  
"22" 0.07085825831368699  
"220" 0.07202521484559447  
"221" 0.06275954048073716  
"222" 0.08461812445937007  
"223" 0.09938704365441753  
"224" 0.1155173395730364  
"225" 0.07652485336934048  
"226" 0.08406752730011911  
"227" 0.10962382431369791  
"228" 0.09868491220695112  
"229" 0.11332588933706157  
"23" 0.17110724973453034  
"230" 0.10986270413409384  
"231" 0.12007054766692572  
"232" 0.10693159562361285  
"233" 0.11596157899487383  
"234" 0.12285200245014952

"235" 0.037134449735166664  
"236" 0.1127781738964015  
"237" 0.11068444994275063  
"238" 0.11272397870276392  
"239" 0.11560001435312818  
"24" 0.1063974946892142  
"240" 0.12758717660292465  
"241" 0.05137835975189523  
"242" 0.06999299194073183  
"243" 0.09200401387480461  
"244" 0.09319486267490545  
"245" 0.03149804712107849  
"246" 0.12084865909975609  
"247" 0.06630808026691458  
"248" 0.057240260632887036  
"249" 0.12107964533332406  
"25" 0.10826333581483975  
"250" 0.05085385413718274  
"251" 0.12126772106222404  
"252" 0.14370421057691493  
"253" 0.12804966977102736  
"254" 0.030382997527837872  
"255" 0.11892329702334607  
"256" 0.10775672822983688  
"257" 0.112969964463628  
"258" 0.09056624728682734  
"259" 0.0631120381736458  
"26" 0.05387850647596753  
"260" 0.0697776162218613  
"261" 0.11054257530034575  
"262" 0.11347120517467521  
"263" 0.11377826548459388  
"264" 0.11307165019896352  
"265" 0.11478984484901446  
"27" 0.0  
"28" 0.10388760038789656  
"29" 0.05496613535902854  
"3" 0.021096199673899408  
"30" 0.0993200906545794  
"31" 0.0511827881817308  
"32" 0.03230808739283316  
"33" 0.12194910391730475  
"34" 0.114243218938323

"35"	0.09512543911462057
"36"	0.1045583224356361
"37"	0.09692903345771506
"38"	0.10158848663818103
"39"	0.08943043728568467
"4"	0.10609633227274991
"40"	0.12068726251898797
"41"	0.08679255866679474
"42"	0.06802938558055768
"43"	0.10689328654189433
"44"	0.07384624107041615
"45"	0.10127267526545299
"46"	0.06957819069578192
"47"	0.02703190872023088
"48"	0.10770213232529696
"49"	0.09993093896359685
"5"	0.041104899704044726
"50"	0.11193807472964104
"51"	0.11316964094480096
"52"	0.12949651566118686
"53"	0.07991508514686092
"54"	0.11611870162898161
"55"	0.06966287087572713
"56"	0.09689085739177432
"57"	0.08978281151684783
"58"	0.20677517842210683
"59"	0.0
"6"	0.052264931159250044
"60"	0.031681601623639075
"61"	0.08715004888384095
"62"	0.07443707183475348
"63"	0.11646649621741123
"64"	0.07580341093428032
"65"	0.10533440352393146
"66"	0.12354358405949303
"67"	0.11795703565511574
"68"	0.11484199757426107
"69"	0.041565606728516925
"7"	0.07795579964389203
"70"	0.10419975360897017
"71"	0.13638875572600404
"72"	0.0854619904976244
"73"	0.1347386827278026

"74" 0.06906228366992087  
"75" 0.08762260222707684  
"76" 0.10933075911947004  
"77" 0.08396271580572354  
"78" 0.03167017360983983  
"79" 0.11855989992766253  
"8" 0.11408029797739186  
"80" 0.10761994826311788  
"81" 0.07312019605658907  
"82" 0.06607188316453157  
"83" 0.06176357421006426  
"84" 0.1663871184166387  
"85" 0.06123726655348049  
"86" 0.31385985951687506  
"87" 0.12639901382300125  
"88" 0.11923831790881051  
"89" 0.08679196188505098  
"9" 0.1327844471286185  
"90" 0.1193716516843036  
"91" 0.08170909573057543  
"92" 0.06180773736294442  
"93" 0.11060716927871955  
"94" 0.040732697124377784  
"95" 0.08596793426998711  
"96" 0.15366033684493052  
"97" 0.09782993105540554  
"98" 0.09033586708780211

```
[root@ip-172-31-47-204 ~]# cat mrtask_e_2017_03_output.txt
"1" 0.1207141821477579
"10" 0.10818785706737385
"100" 0.1036698554215106
"101" 0.11272386860071483
"102" 0.1005306633681427
"105" 0.09171366594360085
"106" 0.10978934961315243
"107" 0.12060767877711343
"108" 0.16459569486203868
"109" 0.2727411419199817
"11" 0.08292554532378368
"111" 0.09254458568885279
"112" 0.10773394359303982
"113" 0.1198356138476097
"114" 0.11911520367510238
"115" 0.06954984925667948
"116" 0.0891850662923476
"117" 0.043243243243243246
"118" 0.06123568538709457
"119" 0.07049709650651401
"12" 0.09293556386304279
"120" 0.10883550438515564
"121" 0.11741420610866483
"122" 0.1387652122291481
"123" 0.08960279415384258
"124" 0.08453696528256818
"125" 0.12518769144404782
"126" 0.09259487147933362
"127" 0.08441883718898342
"128" 0.0996911080691746
"129" 0.060201630039305176
"13" 0.12714951929426124
"130" 0.10414607569293091
"131" 0.10417221067470275
"132" 0.10575125224839516
"133" 0.10062634626777764
"134" 0.09832702825954157
"135" 0.08685750264472232
"136" 0.03613849315545647
"137" 0.11313062136906334
```

...

```

"138" 0.1317185832146078
"139" 0.06289872602766453
"14" 0.09728965671976669
"140" 0.11132977364619498
"141" 0.11224772836685043
"142" 0.10943407036321535
"143" 0.11385544735623641
"144" 0.11534246435160811
"145" 0.09385839052238519
"146" 0.08701042916776668
"147" 0.0407104414538495
"148" 0.11552592223584679
"149" 0.07203300930134972
"15" 0.08857607480435004
"150" 0.08822992437709598
"151" 0.10653807861147498
"152" 0.08032142329328529
"153" 0.03838994241279388
"154" 0.07685038639697143
"155" 0.07278967134703021
"156" 0.07347670250896057
"157" 0.10501594923339569
"158" 0.12040536538332194
"159" 0.04321460191532224
"16" 0.15964166650420136
"160" 0.11332917205196806
"161" 0.11910074531590943
"162" 0.12317702599821087
"163" 0.11388391450287717
"164" 0.11356375521257311
"165" 0.07474841653471122
"166" 0.11300512805877584
"167" 0.03467827275684446
"168" 0.03970975383351151
"169" 0.0414459260514526
"17" 0.08139413432067066
"170" 0.11967256252276448
"171" 0.08945571641402163
"172" 0.18060729407061474
"173" 0.06020800402356662
"174" 0.05603043490423155
"175" 0.16555962907051971

```

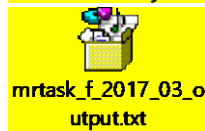
```

"6" 0.052264931159250044
"60" 0.031681601623639075
"61" 0.08715004888384095
"62" 0.07443707183475348
"63" 0.11646649621741123
"64" 0.07580341093428032
"65" 0.10533440352393146
"66" 0.12354358405949303
"67" 0.11795703565511574
"68" 0.11484199757426107
"69" 0.041565606728516925
"7" 0.07795579964389203
"70" 0.10419975360897017
"71" 0.13638875572600404
"72" 0.0854619904976244
"73" 0.1347386827278026
"74" 0.06906228366992087
"75" 0.08762260222707684
"76" 0.10933075911947004
"77" 0.08396271580572354
"78" 0.03167017360983983
"79" 0.11855989992766253
"8" 0.11408029797739186
"80" 0.10761994826311788
"81" 0.07312019605658907
"82" 0.06607188316453157
"83" 0.06176357421006426
"84" 0.1663871184166387
"85" 0.06123726655348049
"86" 0.31385985951687506
"87" 0.12639901382300125
"88" 0.11923831790881051
"89" 0.08679196188505098
"9" 0.1327844471286185
"90" 0.1193716516843036
"91" 0.08170909573057543
"92" 0.06180773736294442
"93" 0.11060716927871955
"94" 0.040732697124377784
"95" 0.08596793426998711
"96" 0.15366033684493052
"97" 0.09782993105540554
"98" 0.09033586708780211

```

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

Ans. The average trip revenue per month - analysing it by hour of the day (day vs night) and the day of the week (weekday vs weekend) are attached in the file



```

[3, 0, 0] 20.40521359874317
[3, 0, 1] 17.276197995863022
[3, 0, 2] 17.601844621054536
[3, 0, 3] 17.25439005270323
[3, 0, 4] 17.319914947221616
[3, 0, 5] 16.53061715387093
[3, 0, 6] 15.555541933661654
[3, 1, 0] 19.68703647305948
[3, 1, 1] 16.41271198977643
[3, 1, 2] 16.925512278579316
[3, 1, 3] 16.61165540540129
[3, 1, 4] 16.160107419186303
[3, 1, 5] 15.832497441155462
[3, 1, 6] 15.364370270396241
[3, 10, 0] 15.792245838246089

```



[3, 10, 1]	15.261493329051431
[3, 10, 2]	15.93810925783485
[3, 10, 3]	16.30214075876192
[3, 10, 4]	16.080729421785076
[3, 10, 5]	13.143838603285518
[3, 10, 6]	13.905262614980442
[3, 11, 0]	16.17422854439143
[3, 11, 1]	15.433695314990839
[3, 11, 2]	16.34197199387861
[3, 11, 3]	16.615542982468714
[3, 11, 4]	16.38474991537654
[3, 11, 5]	13.390678702178493
[3, 11, 6]	13.940995981214432
[3, 12, 0]	16.282810720389826
[3, 12, 1]	15.430570743504797
[3, 12, 2]	16.544916132322516
[3, 12, 3]	17.02373108252729
[3, 12, 4]	16.61745941219783
[3, 12, 5]	13.700263177960306
[3, 12, 6]	14.484117559352311
[3, 13, 0]	16.38637314883534
[3, 13, 1]	15.864056889450703
[3, 13, 2]	17.298944478583774
[3, 13, 3]	17.851531498394657
[3, 13, 4]	17.00853579230275
[3, 13, 5]	14.254259457054172
[3, 13, 6]	15.15122902764985
[3, 14, 0]	16.62469914309496
[3, 14, 1]	16.017420900045288
[3, 14, 2]	17.232419968337283
[3, 14, 3]	18.193252459930395
[3, 14, 4]	17.451055929438706
[3, 14, 5]	14.67993677474162
[3, 14, 6]	16.21387120346392
[3, 15, 0]	16.592225876866596
[3, 15, 1]	15.852790671648188
[3, 15, 2]	17.06561720209403
[3, 15, 3]	18.295222864452015
[3, 15, 4]	17.562402247232335
[3, 15, 5]	14.855918342093966
[3, 15, 6]	16.797449883663646
[3, 16, 0]	18.042703708594157
[3, 16, 1]	17.33124934085766

[3, 16, 2]	18.864639905252556
[3, 16, 3]	19.769173255703436
[3, 16, 4]	18.880767666578635
[3, 16, 5]	14.73631308287822
[3, 16, 6]	16.793010721933754
[3, 17, 0]	16.983800284304177
[3, 17, 1]	16.444131525273395
[3, 17, 2]	17.585606638175037
[3, 17, 3]	18.34339560985185
[3, 17, 4]	17.510178429370153
[3, 17, 5]	14.502202293992317
[3, 17, 6]	16.486961563093033
[3, 18, 0]	16.008802509992595
[3, 18, 1]	17.727856811936284
[3, 18, 2]	16.515468995342214
[3, 18, 3]	16.70180986005624
[3, 18, 4]	16.314710427802392
[3, 18, 5]	13.990148658698722
[3, 18, 6]	15.850448864622033
[3, 19, 0]	15.741789679794332
[3, 19, 1]	15.277445970111039
[3, 19, 2]	16.20599355997634
[3, 19, 3]	16.54377192830582
[3, 19, 4]	15.94514029882594
[3, 19, 5]	14.164873520747108
[3, 19, 6]	15.69852857863186
[3, 2, 0]	18.22057571817039
[3, 2, 1]	16.031631874536583
[3, 2, 2]	17.007477382017846
[3, 2, 3]	15.890491018705909
[3, 2, 4]	15.659892363689025
[3, 2, 5]	15.399850275361139
[3, 2, 6]	15.354934237853096
[3, 20, 0]	15.678553676037785
[3, 20, 1]	15.365148446566707
[3, 20, 2]	16.05710221197255
[3, 20, 3]	16.37797821085915
[3, 20, 4]	15.591262959487404
[3, 20, 5]	14.780932733913335
[3, 20, 6]	16.250329221689736
[3, 21, 0]	16.197433387798725
[3, 21, 1]	16.06818687365222
[3, 21, 2]	16.50227081709119

[3, 21, 3]	16.90277786154788
[3, 21, 4]	15.868536399794348
[3, 21, 5]	15.181319354501266
[3, 21, 6]	16.72184799033714
[3, 22, 0]	16.649822266774542
[3, 22, 1]	16.64261921436027
[3, 22, 2]	16.978409009557595
[3, 22, 3]	17.34474521935207
[3, 22, 4]	16.438219032737116
[3, 22, 5]	15.522444929681589
[3, 22, 6]	18.422694147476886
[3, 23, 0]	17.47287307865713
[3, 23, 1]	17.27000934736513
[3, 23, 2]	17.394545329682096
[3, 23, 3]	17.670885100926263
[3, 23, 4]	16.986778054355376
[3, 23, 5]	16.03912976620376
[3, 23, 6]	19.860621998696352
[3, 3, 0]	18.28883687490132
[3, 3, 1]	16.230832383125353
[3, 3, 2]	17.50601797483577
[3, 3, 3]	16.90531820287067
[3, 3, 4]	16.424534197095852
[3, 3, 5]	15.996357103092873
[3, 3, 6]	15.677216547294654
[3, 4, 0]	21.795900838862302
[3, 4, 1]	18.38125510899298
[3, 4, 2]	19.90678104392431
[3, 4, 3]	20.192967669592427
[3, 4, 4]	20.22441697976039
[3, 4, 5]	17.57629589548885
[3, 4, 6]	17.08128764502286
[3, 5, 0]	21.591044198516926
[3, 5, 1]	18.083058239027036
[3, 5, 2]	19.34133408941352
[3, 5, 3]	20.327765055594767
[3, 5, 4]	21.095153131186052
[3, 5, 5]	20.826286076749316
[3, 5, 6]	20.056067382880748
[3, 6, 0]	16.393886722866988
[3, 6, 1]	14.814524377766153
[3, 6, 2]	15.135161792161815
[3, 6, 3]	15.892669724196614

[3, 6, 4]	16.296061269151785
[3, 6, 5]	19.677382655728223
[3, 6, 6]	20.116308211178346
[3, 7, 0]	14.8068903925165
[3, 7, 1]	14.135772450457887
[3, 7, 2]	14.499646568853017
[3, 7, 3]	14.606266366748153
[3, 7, 4]	14.532465861387703
[3, 7, 5]	16.146780367468107
[3, 7, 6]	17.884824131089324
[3, 8, 0]	15.082638943192709
[3, 8, 1]	14.36155575358702
[3, 8, 2]	14.907604712334534
[3, 8, 3]	15.118939288941126
[3, 8, 4]	15.066199441551346
[3, 8, 5]	13.945318140974235
[3, 8, 6]	15.442457006755673
[3, 9, 0]	15.407340852285504
[3, 9, 1]	14.753088266997892
[3, 9, 2]	15.382085430660746
[3, 9, 3]	15.736632612012666
[3, 9, 4]	15.7315044278208
[3, 9, 5]	13.149422371699014
[3, 9, 6]	14.123488949133877

```

root@ip-172-31-47-204:~
[root@ip-172-31-47-204 ~]#
[root@ip-172-31-47-204 ~]# cat mrtask_f_2017_03_output.txt
[3, 0, 0] 20.40521359874317
[3, 0, 1] 17.276197995863022
[3, 0, 2] 17.601844621054536
[3, 0, 3] 17.25439005270323
[3, 0, 4] 17.319914947221616
[3, 0, 5] 16.53061715387093
[3, 0, 6] 15.555541933661654
[3, 1, 0] 19.68703647305948
[3, 1, 1] 16.41271198977643
[3, 1, 2] 16.925512278579316
[3, 1, 3] 16.61165540540129
[3, 1, 4] 16.160107419186303
[3, 1, 5] 15.832497441155462
[3, 1, 6] 15.364370270396241
[3, 10, 0] 15.792245838246089
[3, 10, 1] 15.261493329051431
[3, 10, 2] 15.93810925783485
[3, 10, 3] 16.30214075876192
[3, 10, 4] 16.080729421785076

```

The various MR job scripts and output files generated:

```
-rwxr-x--- 1 root root      1584 Feb 29 17:17 mrtask_f.py
-rwxr-x--- 1 root root      1360 Feb 29 17:17 mrtask_e.py
-rwxr-x--- 1 root root      1834 Feb 29 17:17 mrtask_d.py
-rwxr-x--- 1 root root      1389 Feb 29 17:17 mrtask_c.py
-rwxr-x--- 1 root root      1038 Feb 29 17:17 mrtask_b.py
-rw-r--r-- 1 root root 969809025 Feb 29 17:18 yellow_tripdata_2017-03.csv
-rwxr-x--- 1 root root      1023 Feb 29 18:08 mrtask_a.py
-rw-r--r-- 1 root root        24 Feb 29 18:20 mrtask_b_2017_03_output.txt
-rw-r--r-- 1 root root        22 Feb 29 18:22 mrtask_a_2017_03_output.txt
-rw-r--r-- 1 root root        44 Feb 29 18:24 mrtask_c_2017_03_output.txt
-rw-r--r-- 1 root root      6628 Feb 29 18:32 mrtask_e_2017_03_output.txt
drwxr-xr-x 4 root root        32 Feb 29 18:38 mrtask_f.root.20240229.183825.942161
-rw-r--r-- 1 root root     4916 Feb 29 18:44 mrtask_f_2017_03_output.txt
-rw-r--r-- 1 root root         0 Feb 29 18:45 mrtask_f_2017_03_output_new.txt
drwxr-xr-x 4 root root        32 Feb 29 18:45 mrtask_f.root.20240229.184558.491327
-rw-r--r-- 1 root root     6230 Feb 29 18:48 mrtask_d_2017_03_output.txt
[root@ip-172-31-47-204 ~]#
```

-----X-----