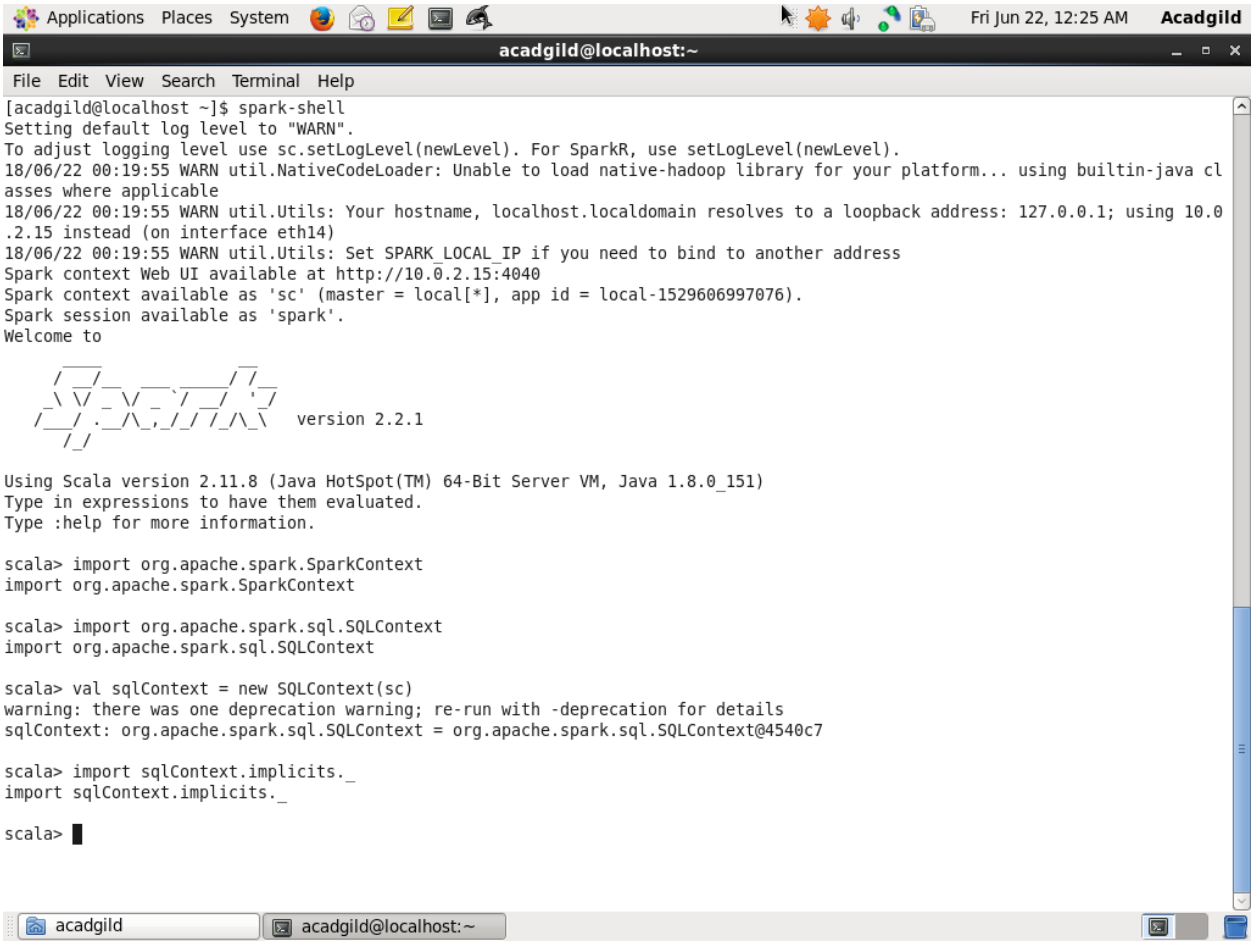
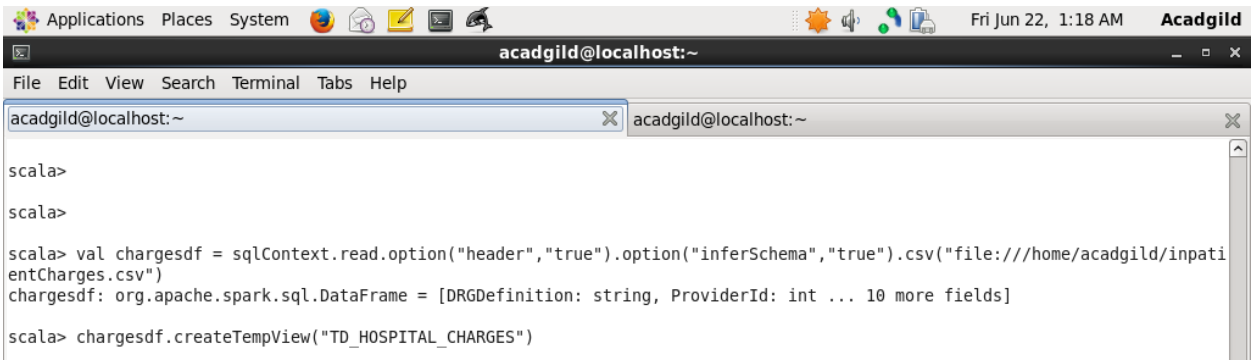
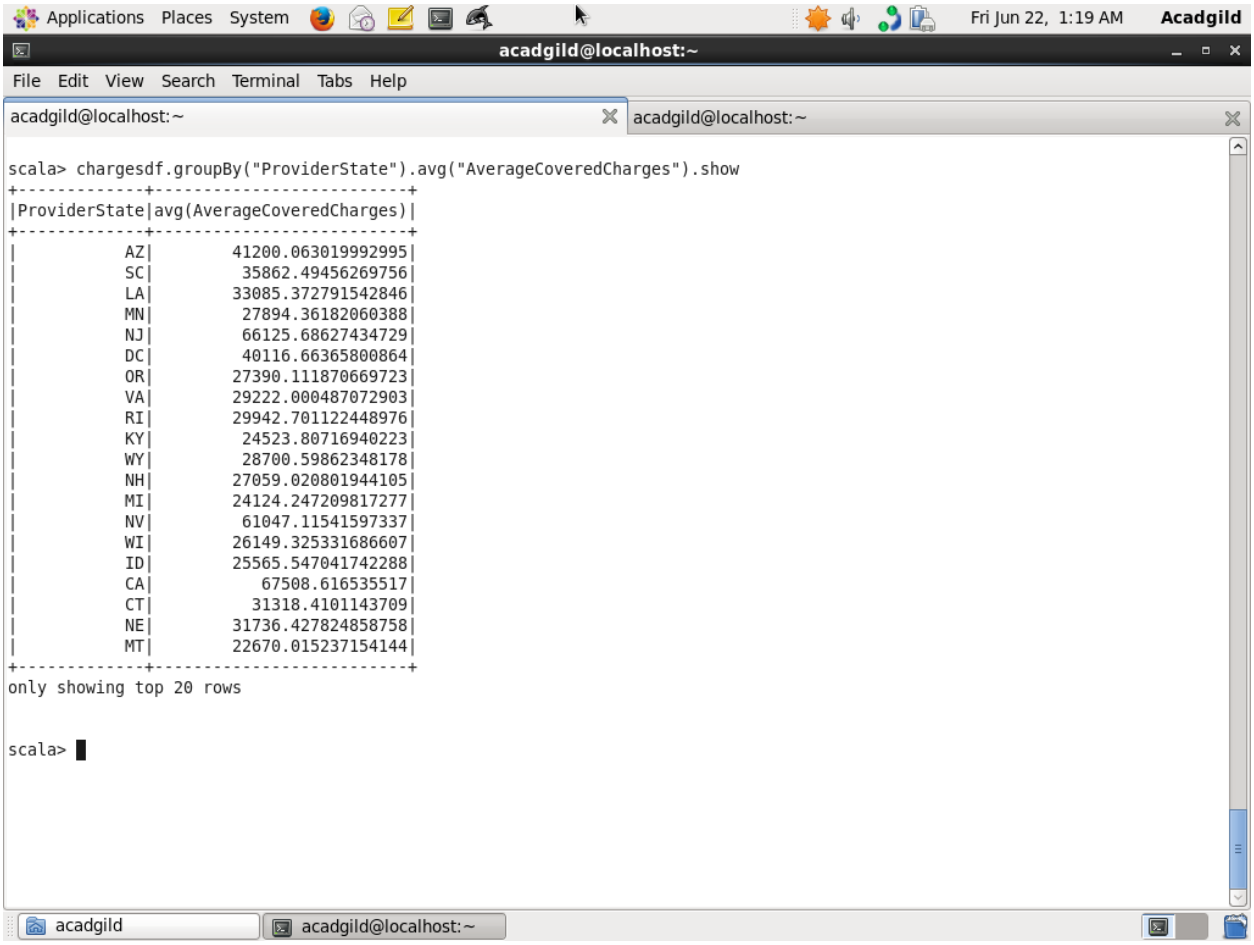
Creating SQL context and importing implicits:



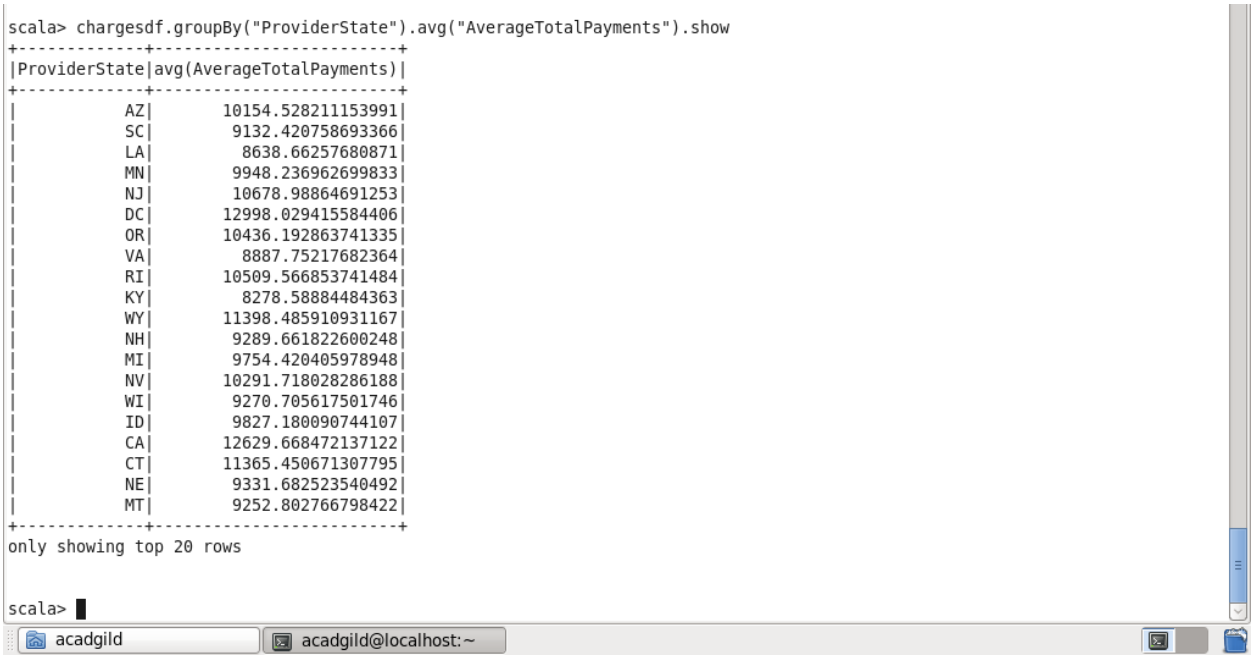
Loading csv file:



Average amount of AverageCoveredCharges per state :



AverageTotalPayments charges per state:



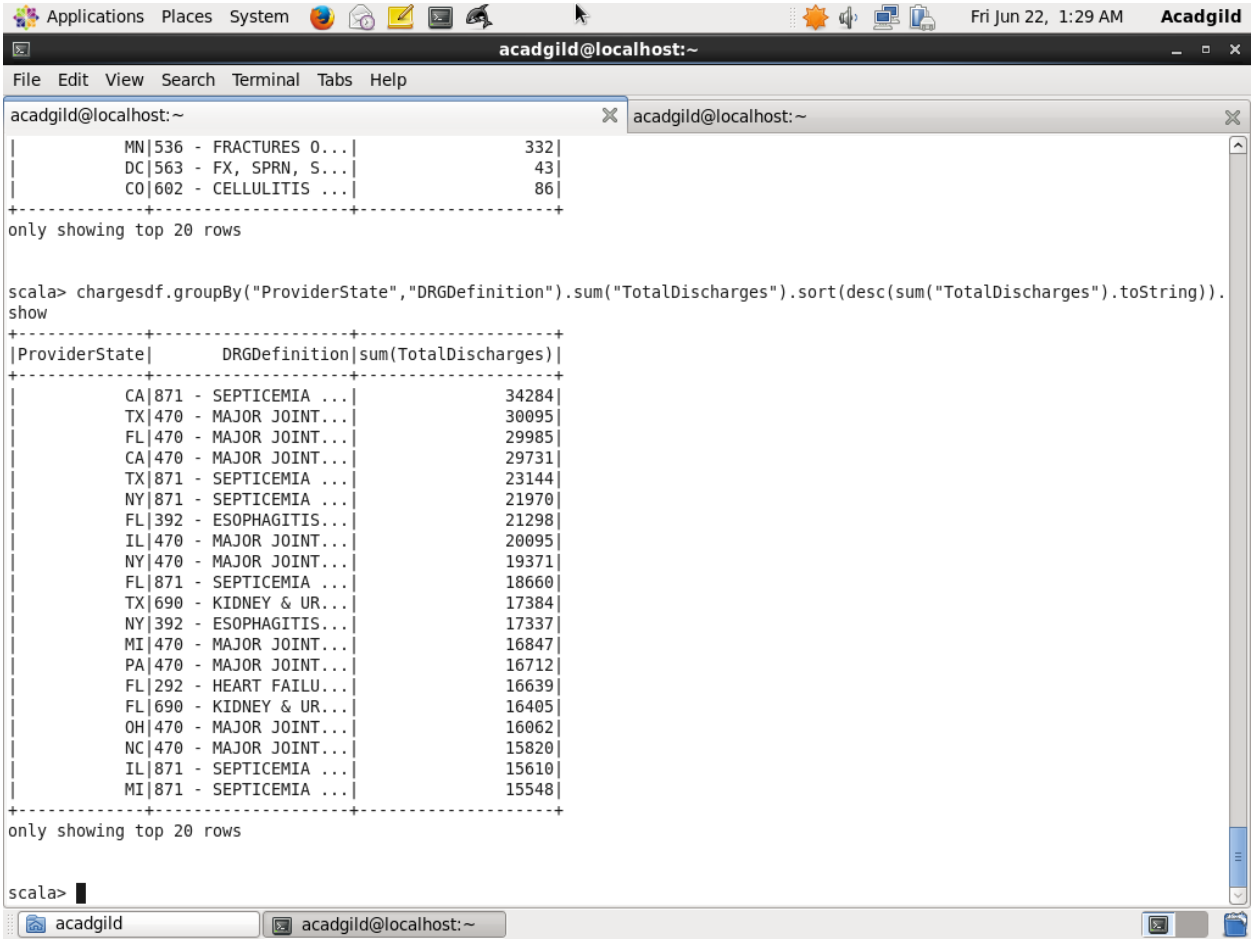
Average Medicare payments per state

```
only showing top 20 rows


scala> chargesdf.groupBy("ProviderState").avg("AverageMedicarePayments").show
+-------------+---------------------------+
|ProviderState|avg(AverageMedicarePayments)|
+-------------+---------------------------+
|           AZ|          8825.717239565045|
|           SC|           7876.33152441167|
|           LA|          7387.704625041281|
|           MN|          8619.214982238007|
|           NJ|          9586.940055946912|
|           DC|         11811.967705627709|
|           OR|          9035.259961508847|
|           VA|          7538.847006001846|
|           RI|          9317.939115646255|
|           KY|          7185.227810467647|
|           WY|          9539.392024291496|
|           NH|          8124.506852976913|
|           MI|          8662.157756043543|
|           NV|          8747.602828618963|
|           WI|          8002.597911079731|
|           ID|          8461.977513611617|
|           CA|         11494.381677893474|
|           CT|         10104.592943809059|
|           NE|         7992.6272504707995|
|           MT|          7981.088063241104|
+-------------+---------------------------+
only showing top 20 rows


scala>
```

acadgild | acadgild@localhost:~

total number of Discharges per state and for each disease :

```
scala> chargesdf.groupBy("ProviderState","DRGDefinition").sum("TotalDischarges").show
18/06/22 01:25:14 WARN executor.Executor: Managed memory leak detected; size = 17039360 bytes, TID = 509
+-------------+--------------------+--------------------+
|ProviderState|       DRGDefinition|sum(TotalDischarges)|
+-------------+--------------------+--------------------+
|           KY|065 - INTRACRANIA...|                1937|
|           NY|101 - SEIZURES W/...|                4503|
|           IN|149 - DYSEQUILIBRIUM|                 700|
|           IA|178 - RESPIRATORY...|                 540|
|           WI|202 - BRONCHITIS ...|                 338|
|           MO|208 - RESPIRATORY...|                1840|
|           WI|251 - PERC CARDIO...|                 417|
|           AR|281 - ACUTE MYOCA...|                 413|
|           AZ|292 - HEART FAILU...|                2643|
|           NY|292 - HEART FAILU...|               13289|
|           NV|293 - HEART FAILU...|                 519|
|           SD|303 - ATHEROSCLER...|                  53|
|           TN|305 - HYPERTENSIO...|                 730|
|           ME|308 - CARDIAC ARR...|                 312|
|           NV|372 - MAJOR GASTR...|                 126|
|           WA|392 - ESOPHAGITIS...|                3148|
|           WI|439 - DISORDERS O...|                 215|
|           MN|536 - FRACTURES O...|                 332|
|           DC|563 - FX, SPRN, S...|                  43|
|           CO|602 - CELLULITIS ...|                  86|
+-------------+--------------------+--------------------+
only showing top 20 rows


scala>
```

acadgild | acadgild@localhost:~

Sort the output in descending order of totalDischarges :