

OBJECTIVE 1:

Load file into spark

```
at SparkSqlCaseStudy4.HospitalDataAnalysis.main(HospitalDataAnalysis.scala)
In Patient Data->>163066
```

DRGDefinition	ProviderId	ProviderName	ProviderStreetAddress	ProviderCity	ProviderState	ProviderZipCode	HospitalReferralRegionDescription	TotalDischarges	AverageCoveredCharges
039 - EXTRACRANIA...	10001	SOUTHEAST ALABAMA...	1108 ROSS CLARK C...	DOTHAN	AL	36301	AL - Dothan	91	329
039 - EXTRACRANIA...	10005	MARSHALL MEDICAL ...	2505 U S HIGHWAY ...	BOAZ	AL	35957	AL - Birmingham	14	151
039 - EXTRACRANIA...	10006	ELIZA COFFEE MEMO...	205 HARENGO STREET	FLORENCE	AL	35631	AL - Birmingham	24	375
039 - EXTRACRANIA...	10011	ST VINCENT'S EAST	50 MEDICAL PARK E...	BIRMINGHAM	AL	35235	AL - Birmingham	25	139
039 - EXTRACRANIA...	10016	SHELBY BAPTIST ME...	1000 FIRST STREET...	ALABASTER	AL	35007	AL - Birmingham	18	316
039 - EXTRACRANIA...	10023	BAPTIST MEDICAL C...	2105 EAST SOUTH B...	MONTGOMERY	AL	36116	AL - Montgomery	67	169
039 - EXTRACRANIA...	10029	EAST ALABAMA MEDI...	2000 PEPPERELL FA...	OPELIKA	AL	36801	AL - Birmingham	51	119
039 - EXTRACRANIA...	10033	UNIVERSITY OF ALA...	619 SOUTH 19TH ST...	BIRMINGHAM	AL	35233	AL - Birmingham	32	358
039 - EXTRACRANIA...	10039	HUNTSVILLE HOSPITAL	101 SIVLEY RD	HUNTSVILLE	AL	35801	AL - Huntsville	135	285
039 - EXTRACRANIA...	10040	GADSDEN REGIONAL ...	1007 GOODYEAR AVENUE	GADSDEN	AL	35903	AL - Birmingham	34	752
039 - EXTRACRANIA...	10046	RIVERVIEW REGIONAL...	600 SOUTH THIRD S...	GADSDEN	AL	35901	AL - Birmingham	14	673
039 - EXTRACRANIA...	10055	FLOWERS HOSPITAL	4370 WEST MAIN ST...	DOTHAN	AL	36305	AL - Dothan	45	396
039 - EXTRACRANIA...	10056	ST VINCENT'S BIRM...	810 ST VINCENT'S ...	BIRMINGHAM	AL	35205	AL - Birmingham	43	228
039 - EXTRACRANIA...	10078	NORTHEAST ALABAMA...	400 EAST 10TH STREET	ANNISTON	AL	36207	AL - Birmingham	21	311
039 - EXTRACRANIA...	10083	SOUTH BALDWIN REG...	1613 NORTH MCKENZ...	FOLEY	AL	36535	AL - Mobile	15	254
039 - EXTRACRANIA...	10085	DECATUR GENERAL H...	1201 7TH STREET SE	DECATUR	AL	35609	AL - Huntsville	27	92
039 - EXTRACRANIA...	10090	PROVIDENCE HOSPITAL	6801 AIRPORT BOUL...	MOBILE	AL	36608	AL - Mobile	27	158
039 - EXTRACRANIA...	10092	D C H REGIONAL ME...	809 UNIVERSITY BO...	TUSCALOOSA	AL	35401	AL - Tuscaloosa	31	197
039 - EXTRACRANIA...	10100	THOMAS HOSPITAL	750 MORPHY AVENUE	FAIRHOPE	AL	36532	AL - Mobile	18	107

only showing top 20 rows

Objective 2:

➤ What is the average amount of **AverageCoveredCharges** per state

```
In Patient Data->>163066
```

ProviderState	avg(AverageCoveredCharges)
TOWANDA	17.0
SAN PABLO	27.2
PO BOX 1727"	null
CUMBERLAND	54.57142857142857
HANCOCK	18.0
PRINCETON	51.1
WATERTOWN	30.571428571428573
EDMONDS	23.571428571428573
MCMINNVILLE	38.0
BOAZ	37.5
BAXLEY	14.0
30002	null
140082	null
150089	null
330024	null
750 MORPHY AVENUE	null
2500 ROCKY MOUNTA...	null
20 YORK ST	null
1000 MAR-WALT DR	null
14000 FIVAY ROAD	null

only showing top 20 rows

➤ find out the **AverageTotalPayments** charges per state



➤ find out the **AverageMedicarePayments** charges per state.



Objective 3:

- Find out the total number of **Discharges** per state and for each disease



- Sort the output in descending order of **totalDischarges**

