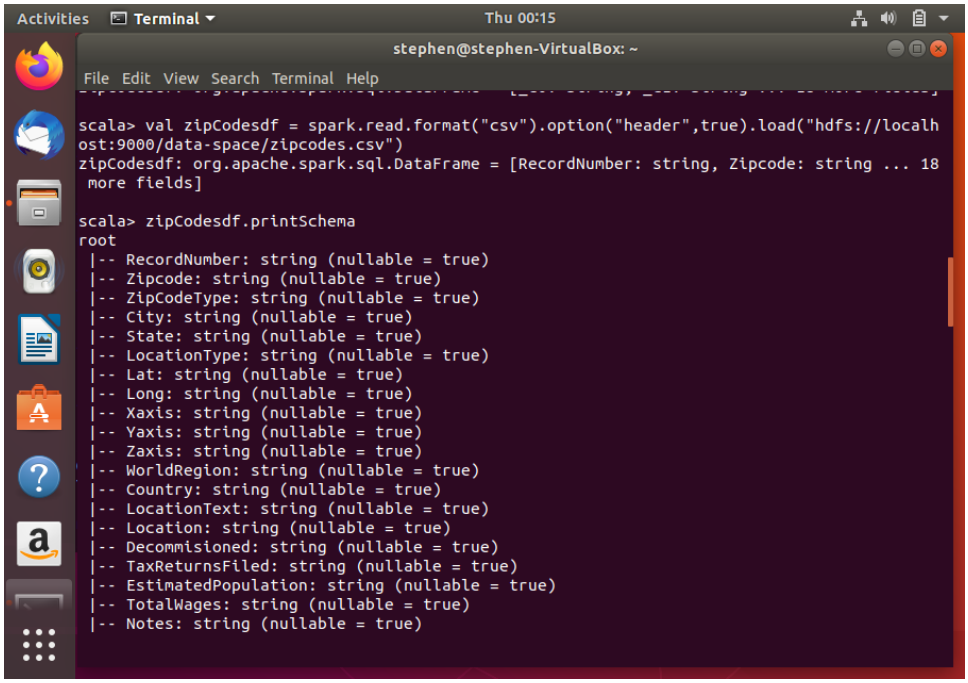
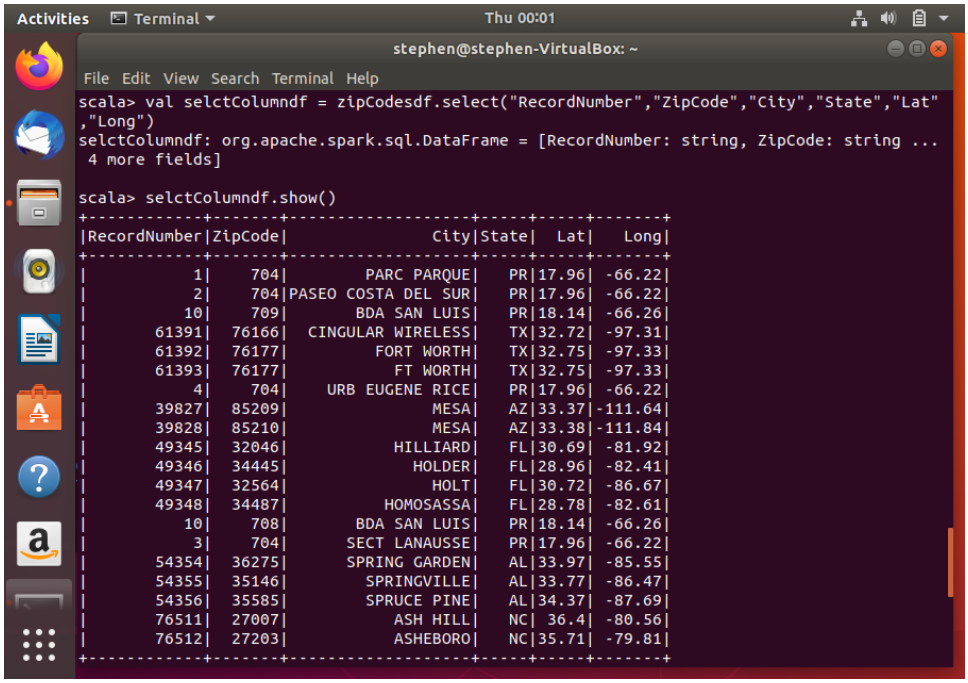
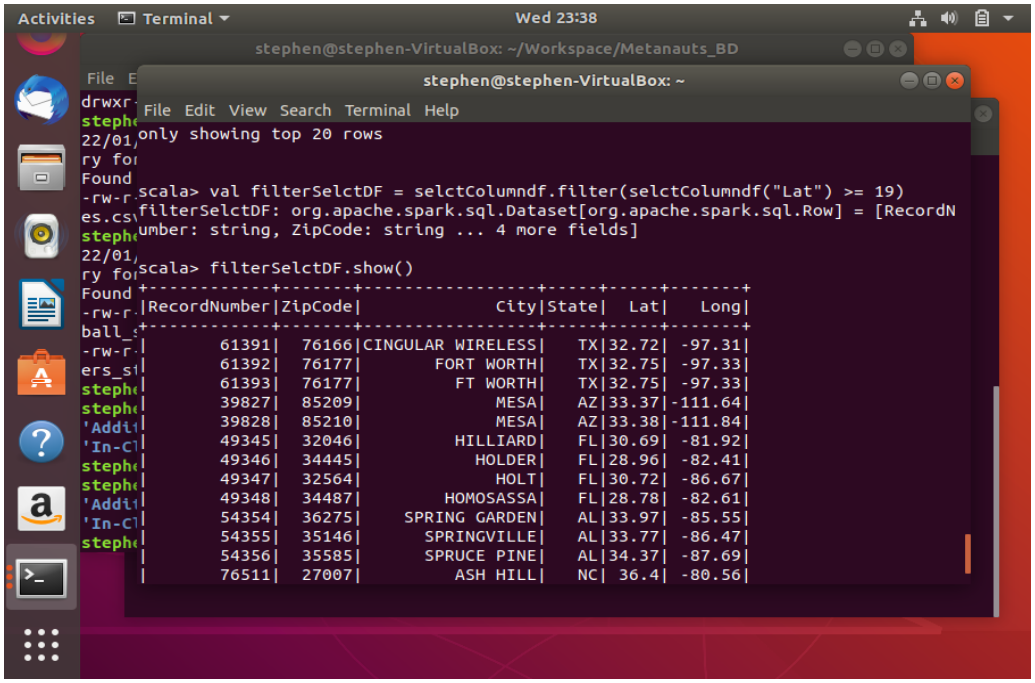
Step 1: Read a file in spark dataframe



Step 2: Filter it on a column



Step 3 & 4: create a dataframe using that filter and print out final dataframe


Activities
Terminal
Wed 23:38
stephen@stephen-VirtualBox: ~/Workspace/Metanauts_BD
stephen@stephen-VirtualBox: ~
File Edit View Search Terminal Help
only showing top 20 rows
scala> val filterSelctDF = selctColumndf.filter(selctColumndf("Lat") >= 19)
filterSelctDF: org.apache.spark.sql.Dataset[org.apache.spark.sql.Row] = [RecordN
umber: string, ZipCode: string ... 4 more fields]
scala> filterSelctDF.show()
+------------+-------+----------------+-----+-----+-------+
|RecordNumber|ZipCode| City|State| Lat| Long|
+------------+-------+----------------+-----+-----+-------+
61391	76166	CINGULAR WIRELESS	TX	32.72	-97.31
61392	76177	FORT WORTH	TX	32.75	-97.33
61393	76177	FT WORTH	TX	32.75	-97.33
39827	85209	MESA	AZ	33.37	-111.64
39828	85210	MESA	AZ	33.38	-111.84
49345	32046	HILLIARD	FL	30.69	-81.92
49346	34445	HOLDER	FL	28.96	-82.41
49347	32564	HOLT	FL	30.72	-86.67
49348	34487	HOMOSASSA	FL	28.78	-82.61
54354	36275	SPRING GARDEN	AL	33.97	-85.55
54355	35146	SPRINGVILLE	AL	33.77	-86.47
54356	35585	SPRUCE PINE	AL	34.37	-87.69
76511	27007	ASH HILL	NC	36.4	-80.56