**Preparing data for Athena: *Python Scripts are available in the Python folder.***

The data sets ‘Eligible Pricing Combinations’, ‘Pool Option Data’ and all loan results data sets has columns names with spaces or commas.

Athena doesn’t support column names with spaces and commas. So, I have written python code in jupyter notebook to read json file, rename column name and written it to json files.

A picture containing screenshot

Description automatically generated

**Uploading data in S3:**

Uploaded to S3, so than I can directly access data from Athena.

A screenshot of a social media post

Description automatically generated

**Created tables in Athena: *Scripts are available in SQL folder.***

**eligible\_pricing\_combinations**:

*SELECT \* FROM eligible\_pricing\_combinations limit 10;*

*A screenshot of a computer

Description automatically generated*

Number of records: 45913 (Data loaded completely)

A screenshot of a cell phone

Description automatically generated

**loan\_data**:

*SELECT \* FROM loan\_data limit 10;*

A screenshot of a cell phone

Description automatically generated

Number of records: 3549 (Data loaded completely)

A screenshot of a cell phone

Description automatically generated

**pool\_option\_data**:

*SELECT \* FROM pool\_option\_data limit 10;*

**A screenshot of a cell phone

Description automatically generated**

Number of records: 458 (Data loaded completely)

**A screenshot of a cell phone

Description automatically generated**

**baseline**:

A screenshot of a cell phone

Description automatically generated

Number of records: 3549 (Data loaded completely)

A screenshot of a cell phone

Description automatically generated

**opt1\_loan\_data**:

A screenshot of a cell phone

Description automatically generated

Number of records: 309 (Data loaded completely)

A screenshot of a cell phone

Description automatically generated

**opt1\_loan\_results\_data:**

A screenshot of a cell phone

Description automatically generated

Number of records: 699 (Data loaded completely)

A screenshot of a cell phone

Description automatically generated

**opt2\_loan\_data**

A screenshot of a cell phone

Description automatically generated

Number of records: 3556 (Data loaded completely)

A screenshot of a cell phone

Description automatically generated

**opt2\_loan\_results\_data**:

A screenshot of a cell phone

Description automatically generated

Number of records: 44070 (Data loaded completely)

A screenshot of a cell phone

Description automatically generated

**opt3\_loan\_data**:

A screenshot of a cell phone

Description automatically generated

Number of records: 3556 (Data loaded completely)

A screenshot of a cell phone

Description automatically generated

**opt3\_loan\_results\_data**:

A screenshot of a cell phone

Description automatically generated

Number of records: 44070 (Data loaded completely)

A screenshot of a cell phone

Description automatically generated