

Structured Transformation ETH

Transforms EtherScan ETH Enriched data

Input:

- Data Content: EtherScan Enriched ETH Data
- Data Type: Parquet
- Data Destination: Processed Layer

Output:

- Data Content: ETH Balance Data
- Data Type: Parquet
- Data Destination: Structured Layer

** Notebook is simply for reference and only takes in a sample collection

```
from pyspark.sql.types import FloatType
from pyspark.sql.functions import format_number
from Azure_configs import processed_data_path, structured_data_path
import datetime
```

```
today=datetime.date.today().strftime('%m-%d-%y')
```

```
def Structured_transformation_eth():

    EScan_ETH_DF=spark.read.parquet(f'{processed_data_path}/{today}/ETH_Balance/NFT=**/')

    structured_ETH_DF=EScan_ETH_DF.select(
        EScan_ETH_DF['owner_address'],\
        format_number((EScan_ETH_DF['ETH_Balance']).cast(FloatType())/10**18),2).alias('ETH')

    structured_ETH_DF.show(10,truncate=False)
    structured_ETH_DF.write.mode('overwrite').parquet(f'{structured_data_path}/{today}/ETH_Balance/')

    return
```

Structured_transformation_eth()

```
+-----+-----+
|owner_address|ETH|
+-----+-----+
|0x040da2c464933005b6d1ffeccec7fb4025dc9ddb|0.35|
|0x04ae2f0bda04f1405991d91c2e8420d6148369ea|0.21|
|0x06e63138f3241a420829bc125e6cb6bebf88c2c2|0.58|
|0x0a8f4037729acdb854da856431a87b8a264d8c40|0.00|
|0x1251122f1d77fa46e1e576c4fd6dd56ab17812ff|0.00|
|0x1db12c2a7c803567c9fd59a202504edf0e56ca78|119.05|
|0x22eab1c78521596f9d6d73dd878778009c39317c|10.31|
|0x266892ed0d40ea5c37f3e0239537999c13468311|0.00|
|0x2e675eeae4747c248bfddbafaa3a8a2fdddaa44b|0.55|
|0x34669322bdfca9e801ca334e7b0e6d69d1f87137|270.97|
+-----+-----+
only showing top 10 rows
```