**Hadoop-hdfs- scoop—hive**

select \* from yellowtripdata;

1. what is the total number of trips(equal to the number of rows)?

select count(\*) from yellowtripdata;

1. what is the total revenue generated by all the trips?

select sum(total\_amount) from yellowtripdata;

1. What fraction of the total is paid for tolls?

select (sum(tolls\_amount)/sum(total\_amount))\*100 from yellowtripdata;

1. What fraction of it is driver tips?

select (sum(tip\_amount)/sum(total\_amount))\*100 from yellowtripdata;

1. What is the average trip amount?

select sum(total\_amount)/count(\*) from yellowtripdata;

1. What is the average distance of the trips?

select sum(trip\_distance)/count(\*) from yellowtripdata;

1. How many different payment types are used?

Select count(distinct payment\_type) from yellowtripdata;

1. For each payment type display the following details:
2. average fare generated

select payment\_type,avg(fare\_amount) from yellowtripdata group by payment\_type order by payment\_type;

1. average tip

select payment\_type, avg(tip\_amount) from yellowtripdata group by payment\_type order by payment\_type;

1. average tax
2. select payment\_type, avg(mta\_tax) from yellowtripdata group by payment\_type order by payment\_type;

select payment\_type, avg(mta\_tax) from yellowtripdata group by payment\_type;

1. On average, which hour of the day generates the highest revenue?