Partitioning in Hive

Sample Data (sample.csv)

id, name, dept, yoj
1, Girish, SC, 2009
2, Ankit, HR, 2009
3, Ravi, SC, 2010
4, Pankaj, TP, 2010

Create Table to Load Data:

```
>CREATE TABLE dept(id INT, name STRING, dept STRING, yoj INT);
```

Loading Data into Table dept:

```
>LOAD DATA LOCAL INPATH  '<hdfspath>' INTO TABLE dept;
```

Create Partition Table

```
>create table dept_part(id INT,name STRING,dept STRING) PARTITIONED BY(yoj
INT);
```

Load Data into Partition Table

```
>set hive.exec.dynamic.partition.mode=nonstrict;
>INSERT OVERWRITE TABLE dept_part PARTITION(yoj)
SELECT id,name,dept,yoj from  dept;
```

Bucketing in Hive

Create Table to Load Data:

```
>CREATE TABLE dept(id INT, name STRING, dept STRING, yoj INT);
```

Loading Data into Table dept:

```
>LOAD DATA LOCAL INPATH  '<hdfspath>' INTO TABLE dept;
```

Create Table on the basis of Buckets

```
>create table dept_part(id INT,name STRING, dept STRING) PARTITIONED BY(yoj
STRING) clustered by (dept) into 2 buckets;
```

Load Data into Bucketed Tables

```
>set.hive.enforce.bucketing=true;
>INSERT OVERWRITE TABLE dept_part PARTITION(yoj)
SELECT id,name,dept,yoj from dept;
```