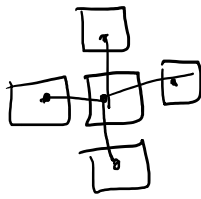
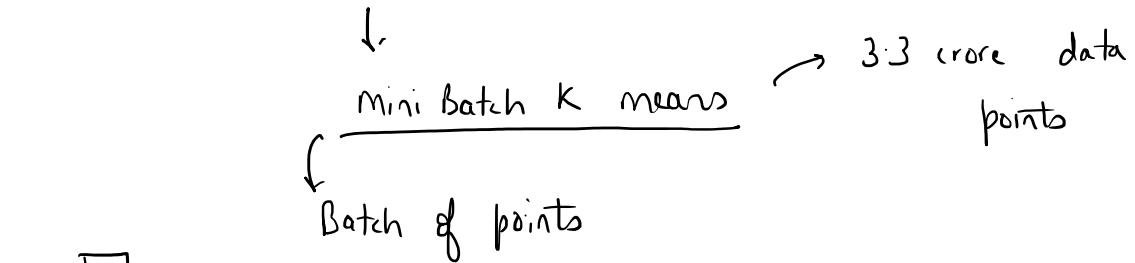


Demand Prediction \rightarrow # of pickups.

Task 1 \rightarrow Breaking our NYC into regions



Inter region \rightarrow 1-1.5 miles.

15 minutes

\downarrow

(30) \rightarrow NYC \rightarrow 30 regions


Task 2 \rightarrow for each region

\downarrow


Time axis \rightarrow Pickup datetime

\downarrow

15 minutes

 \rightarrow X \rightarrow X for each 15 min interval.

total number of pickups.

count \rightarrow  \rightarrow 8:03 pm 8:00 pm - 8:15 pm

\downarrow

Grouping \rightarrow Region.

\downarrow

change the freq (resampling)

\downarrow

count the # of pickups \rightarrow no. of rows

↓
count the # of pickups → no. of rows

Dataset → Each row → each pickup.

index	Region	<u>total pickups</u>	<u>Avg pickups</u>
—	0		—
—			—
—			—
—			—
—	<u>0-1</u>		

EWMA → Avg pickups
Exponentially decaying

$t \leftarrow \text{Max weightage}$
 $t-1 \leftarrow w$
 $t-2 \leftarrow w^2$
 $t-3 \leftarrow w^3$
 $t-4 \leftarrow w^4$

Derived features → lag features → Historical data.
(4)

$t-1, t-2, t-3, t-4$

Day of week → Huge impact on pickup pattern

Experiments → Model selection + HP

↓
(LR) → The best among the group.

8% MAPE

Pipeline → Data, code, model → track

03 March 2025 14:15

