

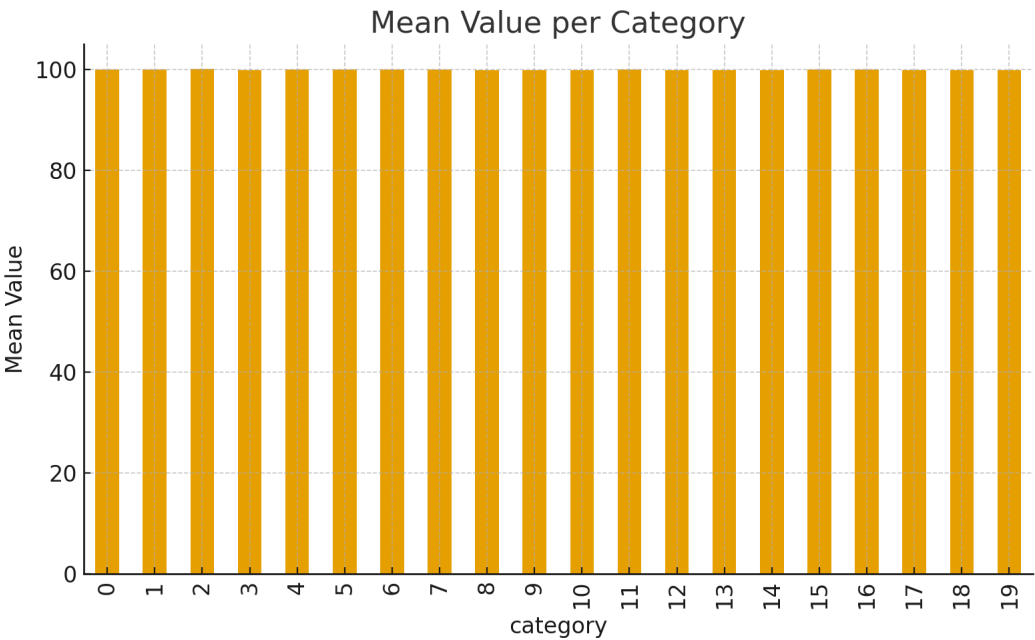
# Scalable Data Analysis Report

This report demonstrates scalable analysis on a synthetic dataset with 2,000,000 rows. The dataset was partitioned and processed incrementally to simulate big-data handling. Three key analyses were performed: 1. Per-category aggregation (count and mean). 2. Daily average value (time-series resampling). 3. Global top-10 records by value.

## Category Aggregation (sample):

category	count	mean
0.0	99820.0	100.0048780594094
1.0	100099.0	100.05037244313873
2.0	99784.0	100.11894541996809
3.0	100061.0	99.87256204216038
4.0	100251.0	100.03812665411144
5.0	99978.0	100.01762654293664
6.0	100450.0	100.05513568709172
7.0	100013.0	100.0551337041116
8.0	99805.0	99.95929089631416
9.0	100482.0	99.93097466750008

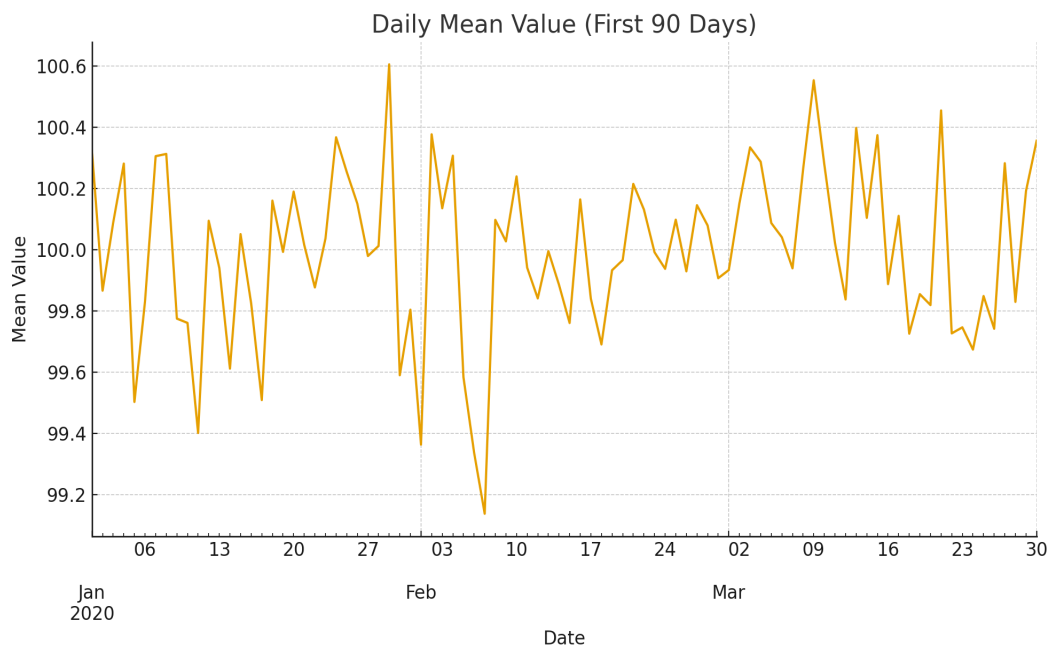
## Category Mean Values Chart:



## Daily Mean (sample of first 10 days):

index	0
2020-01-01 00:00:00	100.32193872018452
2020-01-02 00:00:00	99.86638346450844
2020-01-03 00:00:00	100.08840754387656
2020-01-04 00:00:00	100.28126341775206
2020-01-05 00:00:00	99.50300871412044
2020-01-06 00:00:00	99.832008393459
2020-01-07 00:00:00	100.3054114091454
2020-01-08 00:00:00	100.3128185132251
2020-01-09 00:00:00	99.77510867697012
2020-01-10 00:00:00	99.76094583934837

### Daily Mean Trend (first 90 days):



### Global Top-10 Records by Value:

Unnamed: 0	value	id
0.0	209.70518796425333	1882110.0
1.0	194.9986685938366	336421.0
2.0	192.24086292172737	468632.0
3.0	191.96441014131292	275314.0
4.0	190.953105297307	1185626.0

5.0	190.5926273919291	721536.0
6.0	188.61943529669452	562980.0
7.0	187.22393466164309	1136167.0
8.0	186.87260707155872	1460874.0
9.0	185.90414713836395	605845.0