

**Dataset**

```
CREATE TABLE IF NOT EXISTS stocks (  
  exch string,  
  symb string,  
  price_open float,  
  price_high float,  
  price_low float,  
  price_close float,  
  volume int,  
  price_adj_close float)  
ROW FORMAT DELIMITED FIELDS TERMINATED BY ',';
```

**HDFS**

**Metadata**

**MySQL**

```
CREATE TABLE IF NOT EXISTS stocks (  
    exch string,  
    symbol string,  
    ymd string,  
    price_open float,  
    price_high float,  
    price_low float,  
    price_close float,  
    volume int,  
    row_idx_close float)  
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
```

## Dataset

1. 在 1990 年，美国有 100 万人口居住在 100 英里或更远的地方。  
 2. 在 1990 年，美国有 100 万人口居住在 100 英里或更远的地方。  
 3. 在 1990 年，美国有 100 万人口居住在 100 英里或更远的地方。  
 4. 在 1990 年，美国有 100 万人口居住在 100 英里或更远的地方。  
 5. 在 1990 年，美国有 100 万人口居住在 100 英里或更远的地方。  
 6. 在 1990 年，美国有 100 万人口居住在 100 英里或更远的地方。  
 7. 在 1990 年，美国有 100 万人口居住在 100 英里或更远的地方。  
 8. 在 1990 年，美国有 100 万人口居住在 100 英里或更远的地方。  
 9. 在 1990 年，美国有 100 万人口居住在 100 英里或更远的地方。  
 10. 在 1990 年，美国有 100 万人口居住在 100 英里或更远的地方。

## Metadata



## MySQL

## HDFS

The diagram illustrates the SerDe process for reading data from HDFS. It shows a sequence of steps:

- HDFS**: The raw data is stored in HDFS. The data shown is a sequence of comma-separated values (CSV) representing stock market data, including symbols, dates, and various price metrics.
- InputFormat**: The data is read from HDFS by the InputFormat.
- SerDe**: The data is converted from its raw format to a structured format (e.g., a table of columns and values).
- ObjectInspector**: The structured data is interpreted into a table of columns and values.

The final output is a table with the following columns and values:

exch	symbol	ymd	price_open	price_high	price_low	price_close	volume	price_adj_close
ABCSE	AGC	2007-05-30	20.10	20.18	20.03	20.15	2830	14.61

NCSTP, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023, 2024, 2025, 2026, 2027, 2028, 2029, 2030, 2031, 2032, 2033, 2034, 2035, 2036, 2037, 2038, 2039, 2040, 2041, 2042, 2043, 2044, 2045, 2046, 2047, 2048, 2049, 2050, 2051, 2052, 2053, 2054, 2055, 2056, 2057, 2058, 2059, 2060, 2061, 2062, 2063, 2064, 2065, 2066, 2067, 2068, 2069, 2070, 2071, 2072, 2073, 2074, 2075, 2076, 2077, 2078, 2079, 2080, 2081, 2082, 2083, 2084, 2085, 2086, 2087, 2088, 2089, 2090, 2091, 2092, 2093, 2094, 2095, 2096, 2097, 2098, 2099, 2100, 2101, 2102, 2103, 2104, 2105, 2106, 2107, 2108, 2109, 2110, 2111, 2112, 2113, 2114, 2115, 2116, 2117, 2118, 2119, 2120, 2121, 2122, 2123, 2124, 2125, 2126, 2127, 2128, 2129, 2130, 2131, 2132, 2133, 2134, 2135, 2136, 2137, 2138, 2139, 2140, 2141, 2142, 2143, 2144, 2145, 2146, 2147, 2148, 2149, 2150, 2151, 2152, 2153, 2154, 2155, 2156, 2157, 2158, 2159, 2160, 2161, 2162, 2163, 2164, 2165, 2166, 2167, 2168, 2169, 2170, 2171, 2172, 2173, 2174, 2175, 2176, 2177, 2178, 2179, 2180, 2181, 2182, 2183, 2184, 2185, 2186, 2187, 2188, 2189, 2190, 2191, 2192, 2193, 2194, 2195, 2196, 2197, 2198, 2199, 2200, 2201, 2202, 2203, 2204, 2205, 2206, 2207, 2208, 2209, 2210, 2211, 2212, 2213, 2214, 2215, 2216, 2217, 2218, 2219, 2220, 2221, 2222, 2223, 2224, 2225, 2226, 2227, 2228, 2229, 2230, 2231, 2232, 2233, 2234, 2235, 2236, 2237, 2238, 2239, 2240, 2241, 2242, 2243, 2244, 2245, 2246, 2247, 2248, 2249, 2250, 2251, 2252, 2253, 2254, 2255, 2256, 2257, 2258, 2259, 2260, 2261, 2262, 2263, 2264, 2265, 2266, 2267, 2268, 2269, 2270, 2271, 2272, 2273, 2274, 2275, 2276, 2277, 2278, 2279, 2280, 2281, 2282, 2283, 2284, 2285, 2286, 2287, 2288, 2289, 2290, 2291, 2292, 2293, 2294, 2295, 2296, 2297, 2298, 2299, 2300, 2301, 2302, 2303, 2304, 2305, 2306, 2307, 2308, 2309, 2310, 2311, 2312, 2313, 2314, 2315, 2316, 2317, 2318, 2319, 2320, 2321, 2322, 2323, 2324, 2325, 2326, 2327, 2328, 2329, 2330, 2331, 2332, 2333, 2334, 2335, 2336, 2337, 2338, 2339, 2340, 2341, 2342, 2343, 2344, 2345, 2346, 2347, 2348, 2349, 2350, 2351, 2352, 2353, 2354, 2355, 2356, 2357, 2358, 2359, 2360, 2361, 2362, 2363, 2364, 2365, 2366, 2367, 2368, 2369, 2370, 2371, 2372, 2373, 2374, 2375, 2376, 2377, 2378, 2379, 2380, 2381, 2382, 2383, 2384, 2385, 2386, 2387, 2388, 2389, 2390, 2391, 2392, 2393, 2394, 2395, 2396, 2397, 2398, 2399, 2400, 2401, 2402, 2403, 2404, 2405, 2406, 2407, 2408, 2409, 2410, 2411, 2412, 2413, 2414, 2415, 2416, 2417, 2418, 2419, 2420, 2421, 2422, 2423, 2424, 2425, 2426, 2427, 2428, 2429, 2430, 2431, 2432, 2433, 2434, 2435, 2436, 2437, 2438, 2439, 2440, 2441, 2442, 2443, 2444, 2445, 2446, 2447, 2448, 2449, 2450, 2451, 2452, 2453, 2454, 2455, 2456, 2457, 2458, 2459, 2460, 2461, 2462, 2463, 2464, 2465, 2466, 2467, 2468, 2469, 2470, 2471, 2472, 2473, 2474, 2475, 2476, 2477, 2478, 2479, 2480, 2481, 2482, 2483, 2484, 2485, 2486, 2487, 2488, 2489, 2490, 2491, 2492, 2493, 2494, 2495, 2496, 2497, 2498, 2499, 2500, 2501, 2502, 2503, 2504, 2505, 2506, 2507, 2508, 2509, 2510, 2511, 2512, 2513, 2514, 2515, 2516, 2517, 2518, 2519, 2520, 2521, 2522, 2523, 2524, 2525, 2526, 2527, 2528, 2529, 2530, 2531, 2532, 2533, 2534, 2535, 2536, 2537, 2538, 2539, 2540, 2541, 2542, 2543, 2544, 2545, 2546, 2547, 2548, 2549, 2550, 2551, 2552, 2553, 2554, 2555, 2556, 2557, 2558, 2559, 2560, 2561, 2562, 2563, 2564, 2565, 2566, 2567, 2568, 2569, 2570, 2571, 2572, 2573, 2574, 2575, 2576, 2577, 2578, 2579, 2580, 2581, 2582, 2583, 2584, 2585, 2586, 2587, 2588, 2589, 2590, 2591, 2592, 2593, 2594, 2595, 2596, 2597, 2598, 2599, 2600, 2601, 2602, 2603, 2604, 2605, 2606, 2607, 2608, 2609, 2610, 2611, 2612, 2613, 2614, 2615, 2616, 2617, 2618, 2619, 2620, 2621, 2622, 2623, 2624, 2625, 2626, 2627, 2628, 2629, 2630, 2631, 2632, 2633, 2634, 2635, 2636, 2637, 2638, 2639, 2640, 2641, 2642, 2643, 2644, 2645, 2646, 2647, 2648, 2649, 2650, 2651, 2652, 2653, 2654, 2655, 2656, 2657, 2658, 2659, 2660, 2661, 2662, 2663, 2664, 2665, 2666, 2667, 2668, 2669, 2670, 2671, 2672, 2673, 2674, 2675, 2676, 2677, 2678, 2679, 2680, 2681, 2682, 2683, 2684, 2685, 2686, 2687, 2688, 2689, 2690, 2691, 2692, 2693, 2

HDFS

InputFormat

ABCSE	AGC	2007-05-30	20.10	20.18	20.03	20.15	2830	14.61
-------	-----	------------	-------	-------	-------	-------	------	-------

SerDe  
DeSerializer  
ObjectInspector

exch	symbol	ymd	price_open	price_high	price_low	price_close	volume	price_adj_close
------	--------	-----	------------	------------	-----------	-------------	--------	-----------------

[illegible]

ABCSE AGC	2007-05-30	20.10	20.18	20.03	20.15	2830	14.61
-----------	------------	-------	-------	-------	-------	------	-------

**SerDe**  
Serializer

ObjectInspector

exch	symbol	ymd	price_open	price_high	price_low	price_close	volume	price_adj_close
------	--------	-----	------------	------------	-----------	-------------	--------	-----------------

OutputFormat

## HDFS

[illegible]

www.hadoopinrealworld.com

## READING DATA

```
ABCSE,B7J,2008-10-28,6.48,6.74,6.22,6.72,44300,5.79
ABCSE,B7J,2008-10-27,6.21,6.78,6.21,6.40,55200,5.51
ABCSE,B7J,2008-10-24,6.39,6.66,6.21,6.40,67400,5.51
ABCSE,B7J,2008-10-23,6.95,6.95,6.50,6.59,59400,5.68
ABCSE,B7J,2008-10-22,6.92,7.17,6.80,6.80,55300,5.86
ABCSE,B7J,2008-10-21,7.20,7.30,7.10,7.10,54400,6.11
ABCSE,B7J,2008-10-20,6.94,7.31,6.94,7.12,45700,6.13
ABCSE,B7J,2008-10-17,6.43,6.93,6.42,6.90,57700,5.94
ABCSE,B7J,2008-10-16,6.61,6.69,6.21,6.53,83200,5.62
ABCSE,B7J,2008-10-15,6.84,6.90,6.36,6.36,78900,5.48
ABCSE,B7J,2008-10-14,7.15,7.32,6.93,6.96,74700,5.99
ABCSE,B7J,2008-10-13,6.00,6.57,6.00,6.57,75700,5.66
ABCSE,B7J,2008-10-10,5.05,5.72,4.79,5.72,158400,4.93
ABCSE,B7J,2008-10-09,6.30,6.41,6.00,6.02,140500,5.18
ABCSE,B7J,2008-10-08,5.60,6.47,5.60,6.28,292000,5.41
ABCSE,B7J,2008-10-07,7.59,7.59,6.66,6.69,89900,5.76
ABCSE,B7J,2008-10-06,7.83,7.90,7.00,7.40,159600,6.37
```

HDFS

InputFormat

ABCSE AGC 2007-05-30 20.10 20.18 20.03 20.15 2830 14.61

SerDe  
DeSerializer

ObjectInspector

exch

symbol

ymd

price\_open

price\_high

price\_low

price\_close

volume

price\_adj\_close

## WRITING DATA

ABCSE AGC 2007-05-30 20.10 20.18 20.03 20.15 2830 14.61

**SerDe**  
Serializer

**ObjectInspector**

↓   ↓   ↓   ↓   ↓   ↓   ↓   ↓  
exch   symbol   ymd   price\_open   price\_high   price\_low   price\_close   volume   price\_adj\_close

**OutputFormat**

ABCSE,B7J,2008-10-28,6.48,6.74,6.22,6.72,44300,5.79  
ABCSE,B7J,2008-10-27,6.21,6.78,6.21,6.40,55200,5.51  
ABCSE,B7J,2008-10-24,6.39,6.66,6.21,6.40,67400,5.51  
ABCSE,B7J,2008-10-23,6.95,6.95,6.50,6.59,59400,5.48  
ABCSE,B7J,2008-10-22,6.92,7.17,6.80,6.80,55300,5.86  
ABCSE,B7J,2008-10-21,7.20,7.30,7.10,7.10,54400,6.11  
ABCSE,B7J,2008-10-20,6.94,7.31,6.94,7.12,45700,6.13  
ABCSE,B7J,2008-10-17,6.43,6.93,6.42,6.90,57700,5.94  
ABCSE,B7J,2008-10-16,6.61,6.69,6.21,6.53,83200,5.62  
ABCSE,B7J,2008-10-15,6.84,6.90,6.36,6.36,78900,5.48  
ABCSE,B7J,2008-10-14,7.15,7.32,6.93,6.96,74700,5.99  
ABCSE,B7J,2008-10-13,6.00,6.57,6.00,6.57,75700,5.66  
ABCSE,B7J,2008-10-10,5.05,5.72,4.79,5.72,158400,4.93  
ABCSE,B7J,2008-10-09,6.30,6.41,6.00,6.02,140500,5.18  
ABCSE,B7J,2008-10-08,5.60,6.47,5.60,6.28,292000,5.81  
ABCSE,B7J,2008-10-07,7.59,7.59,6.66,6.69,89900,5.76  
ABCSE,B7J,2008-10-06,7.83,7.90,7.00,7.40,159600,6.37

**HDFS**

# METASTORE

```
CREATE TABLE IF NOT EXISTS stocks (  
  exch string,  
  symbol string,  
  ymd string,  
  price_open float,  
  price_high float,  
  price_low float,  
  price_close float,  
  volume int,  
  price_adj_close float)  
ROW FORMAT DELIMITED FIELDS TERMINATED BY ',';
```

Metadata



MySQL

Dataset



```
ABCSE,B72,2008-10-25,6.45,6.75,6.22,6.72,44300,5.79  
ABCSE,B72,2008-10-27,6.21,6.78,6.21,6.40,55200,5.51  
ABCSE,B72,2008-10-24,6.39,6.66,6.21,6.40,67400,5.51  
ABCSE,B72,2008-10-23,6.95,6.95,6.50,6.50,59400,5.68  
ABCSE,B72,2008-10-22,6.92,7.17,6.00,6.00,50300,5.86  
ABCSE,B72,2008-10-21,7.29,7.30,7.10,7.10,54400,6.11  
ABCSE,B72,2008-10-20,6.94,7.31,6.94,7.12,45700,6.13  
ABCSE,B72,2008-10-17,6.43,6.93,6.42,6.90,57700,5.94  
ABCSE,B72,2008-10-16,6.61,6.69,6.21,6.53,83200,5.62  
ABCSE,B72,2008-10-15,6.24,6.90,6.16,6.16,78900,5.48  
ABCSE,B72,2008-10-14,7.15,7.32,6.53,6.56,74700,5.99  
ABCSE,B72,2008-10-13,6.59,6.57,6.00,6.57,75700,5.86  
ABCSE,B72,2008-10-10,6.58,6.52,6.19,6.12,158400,5.93  
ABCSE,B72,2008-10-09,6.30,6.41,6.00,6.02,148500,5.18  
ABCSE,B72,2008-10-08,5.60,6.47,5.60,6.28,292000,5.41  
ABCSE,B72,2008-10-07,7.59,7.59,6.66,6.69,89900,5.76  
ABCSE,B72,2008-10-06,7.03,7.90,7.00,7.40,109600,6.37
```

HDFS