Ashish Narmen, G1901836L, CI6227-2021-Assignment-2

Source Code and Data Sets

https://github.com/ashishnarmen/ci6227-2021-assignment-2

Data Pre-processing

- Annotate trip data with time segment'
- Group and Count the data according to the time segments
- Merge the in-flow and out-flow across time segments for each station

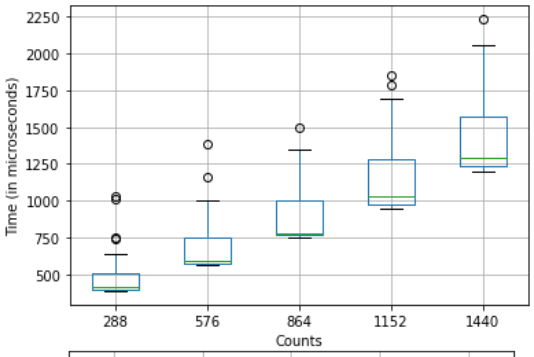
Transactions

- Station ID
- Time of the day (Discretized into Morning, Noon, Afternoon, Evening and Night)
- Flow Count (Incoming Vehicles and Outgoing Vehicles)

Association Rule Mining

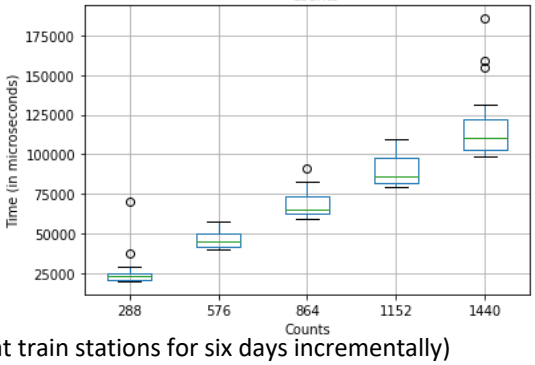
Apriori

Dataset Count	Average Time	Std. Deviation	Standard Error
288	470.960000	123.286465	14.235895
576	677.826667	149.056163	17.211523
864	896.746667	170.277206	19.661918
1152	1142.733333	221.922578	25.625412
1440	1425.973333	252.953845	29.208594



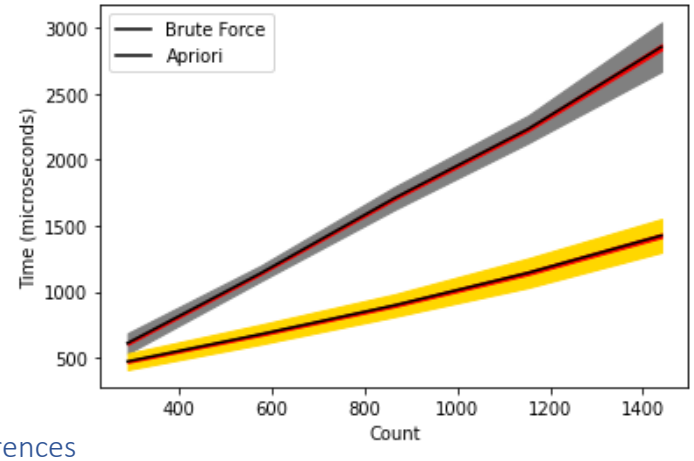
Brute Force

Dataset Count	Average Time	Std. Deviation	Standard Error
288	24379.680000	6114.909695	706.088952
576	45662.093333	4669.166513	539.148909
864	68375.413333	6877.697006	794.168044
1152	89307.706667	8325.061074	961.295250
1440	114251.080000	14778.380611	1706.460405



Performance Comparison

- No. of datasets: 5
- Size of datasets: 288 to 1440 (inflow and outflow data at train stations for six days incrementally)
- Time of the day and flow count in a station
- No. of measurements: 70
- Python Notebook: https://github.com/ashishnarmen/ci6227-2021-assignment-2/blob/main/CI6227_2021_Assignment_2.ipynb



References

Hsu, C. (2018, January 1). *Mining Association Rules on New York City Bike Dataset*. An Explorer of Things. Retrieved October 21, 2021, from https://chih-ling-hsu.github.io/2018/01/01/association-rule-mining

Pisharody, V. K. (n.d.). *GitHub - vinay-k-pisharody/Apriori-Implementation: Brute Force Implementation of Apriori Algorithm* GitHub. Retrieved October 21, 2021, from https://github.com/vinay-k-pisharody/Apriori-Implementation