Gravity Model

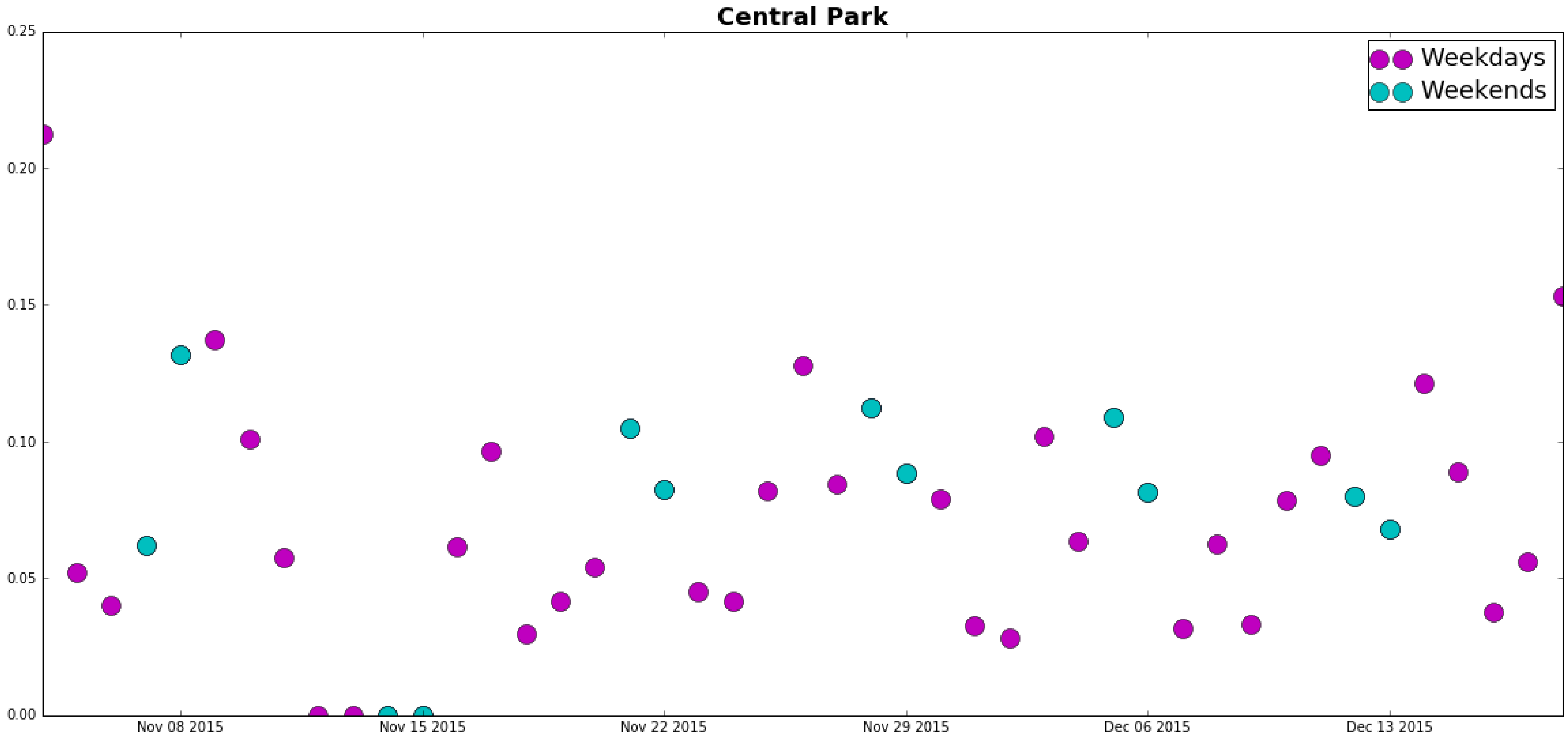
Investigate W^{in}

W_b^{in} trend and normalization

- First we check a few landmarks and places with high level of tweets activities, observe their coefficients, i.e. W_b^{in} change during weekdays and weekends
- Second we select a place (here we simply choose the one with highest volume—10007, city hall) as our reference point and set its W_b^{in} equal to 1, then re-run the OLS to find "new and normalized" W_b^{in} for each zip code b
- In the meantime we'll check if the sum of W_b^{in} changes over time or remain stable (first un-normalized case, then normalized one)

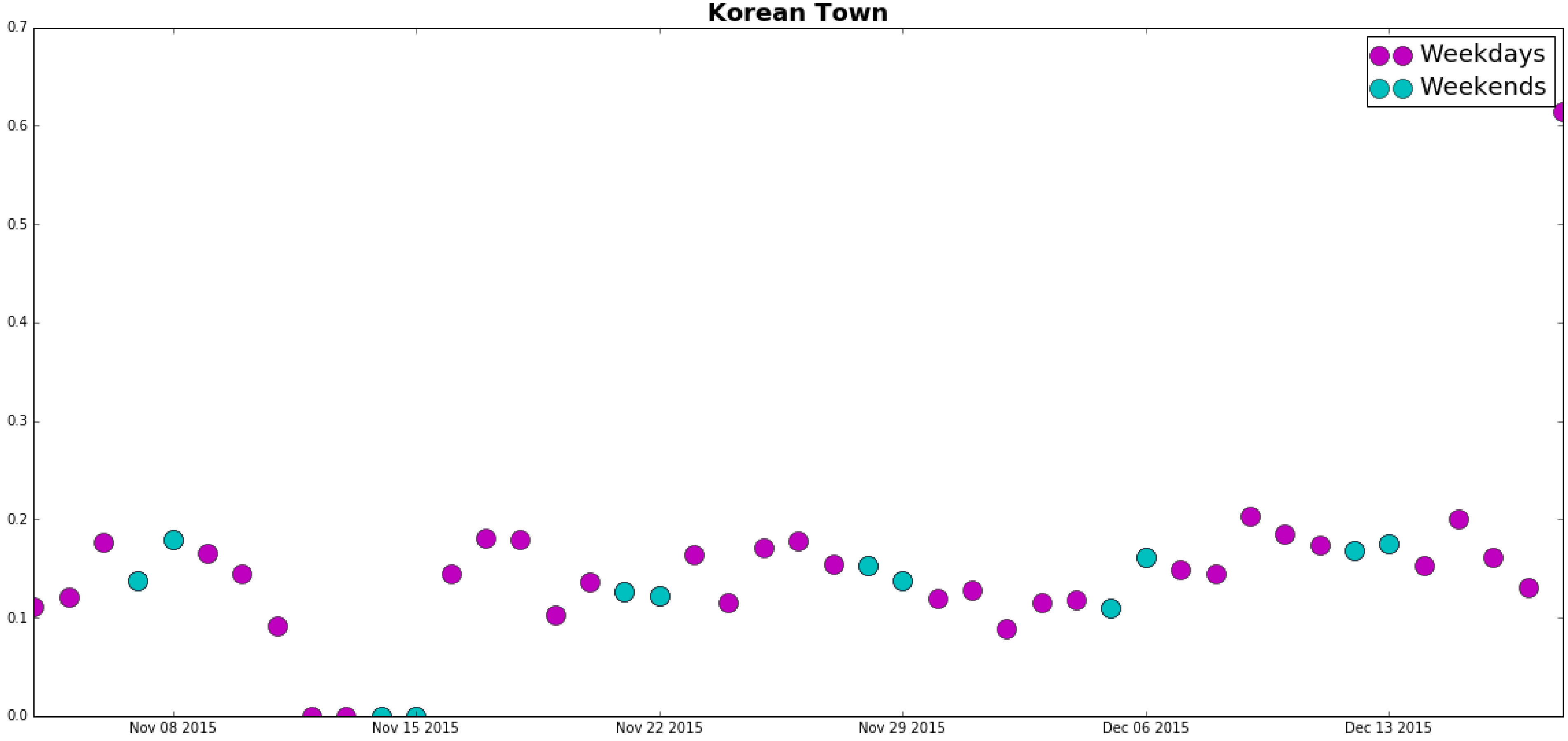
83, Central Park (un-normalized value)

Non-zero part	mean	std
weekdays	0.075187	0.042355
weekends	0.092074	0.021786



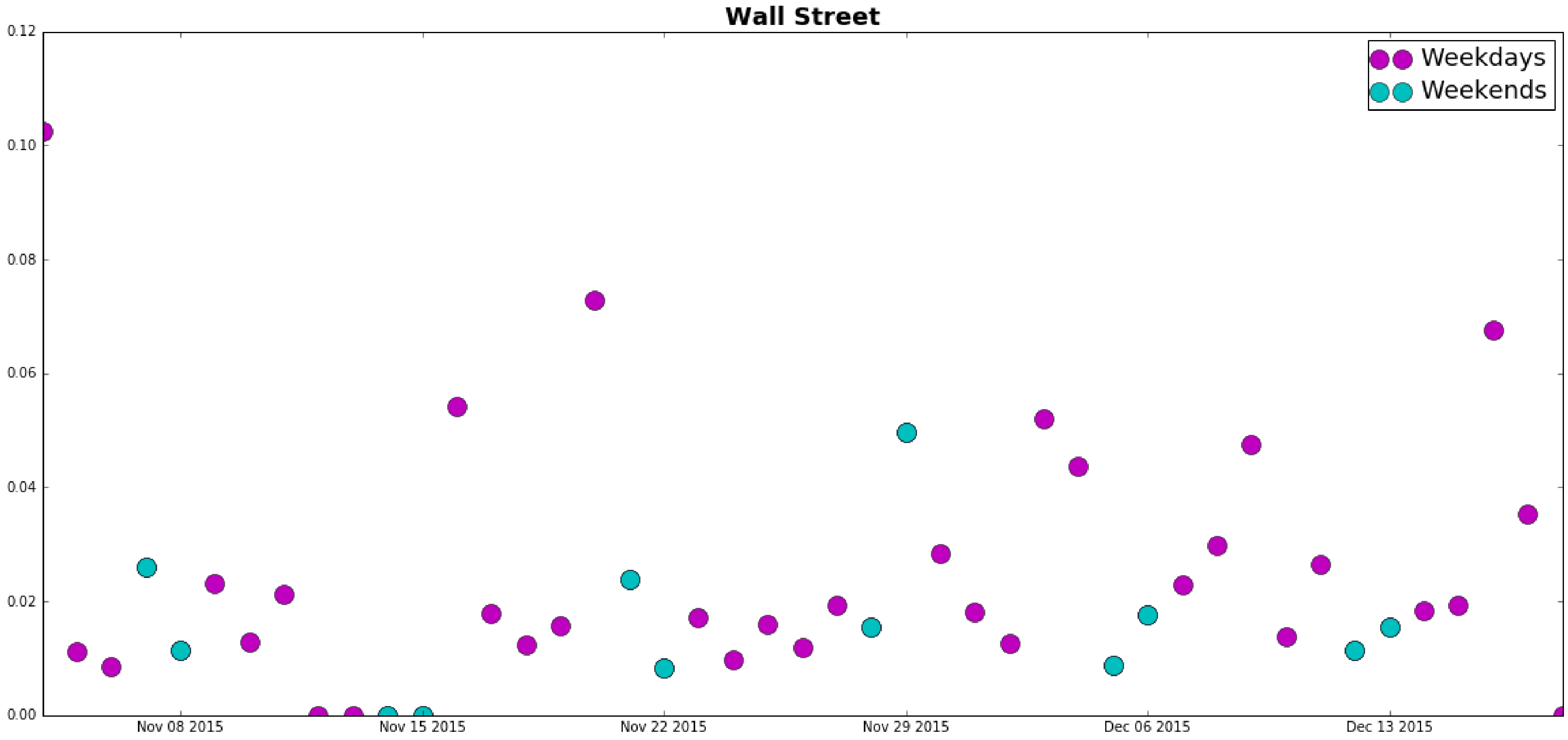
10001, Korean Town (un-normalized value)

Non-zero part	mean	std
weekdays	0.162118	0.089364
weekends	0.147209	0.023991



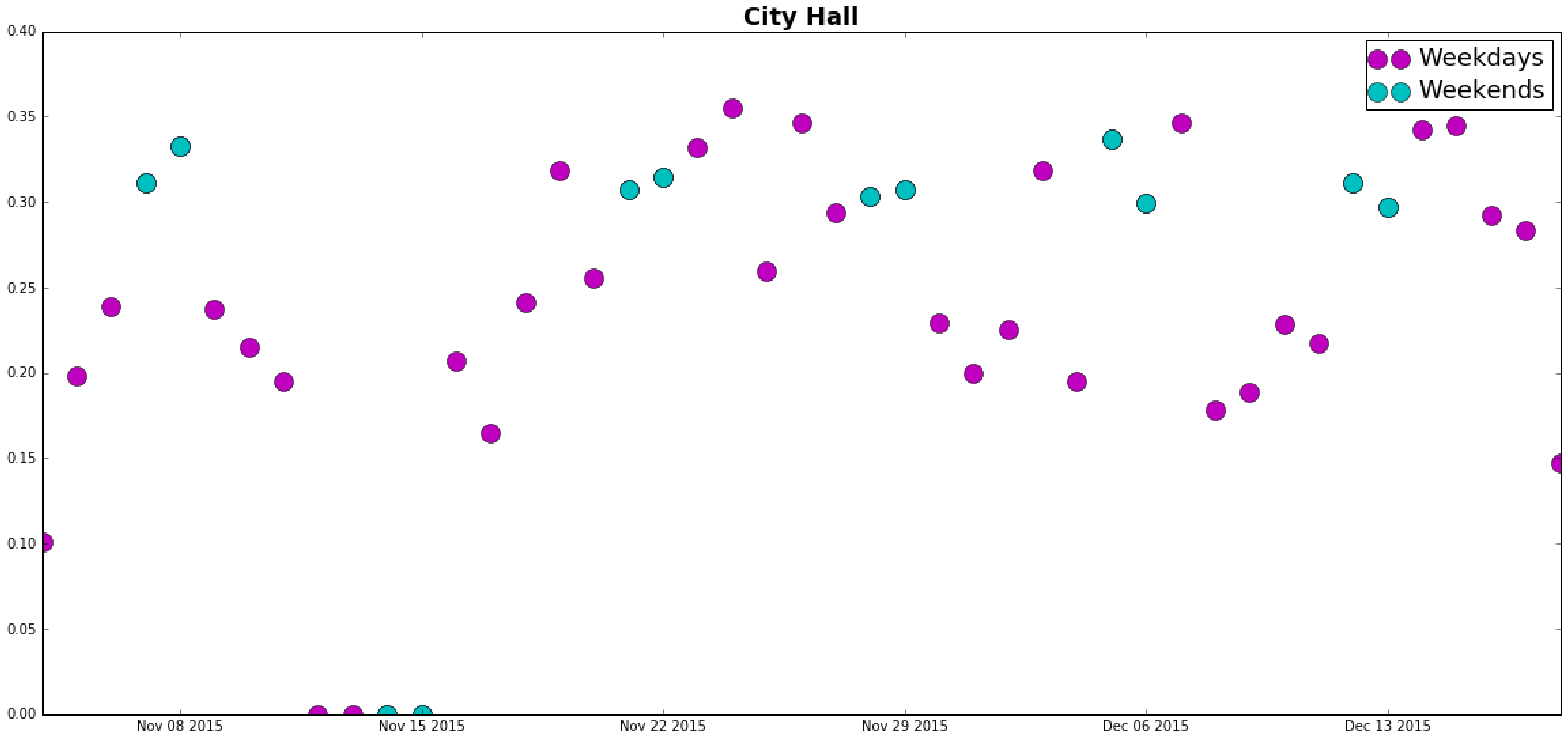
10005, Wall Street (un-normalized value)

Non-zero part	mean	std
weekdays	0.027822	0.022334
weekends	0.018819	0.012331



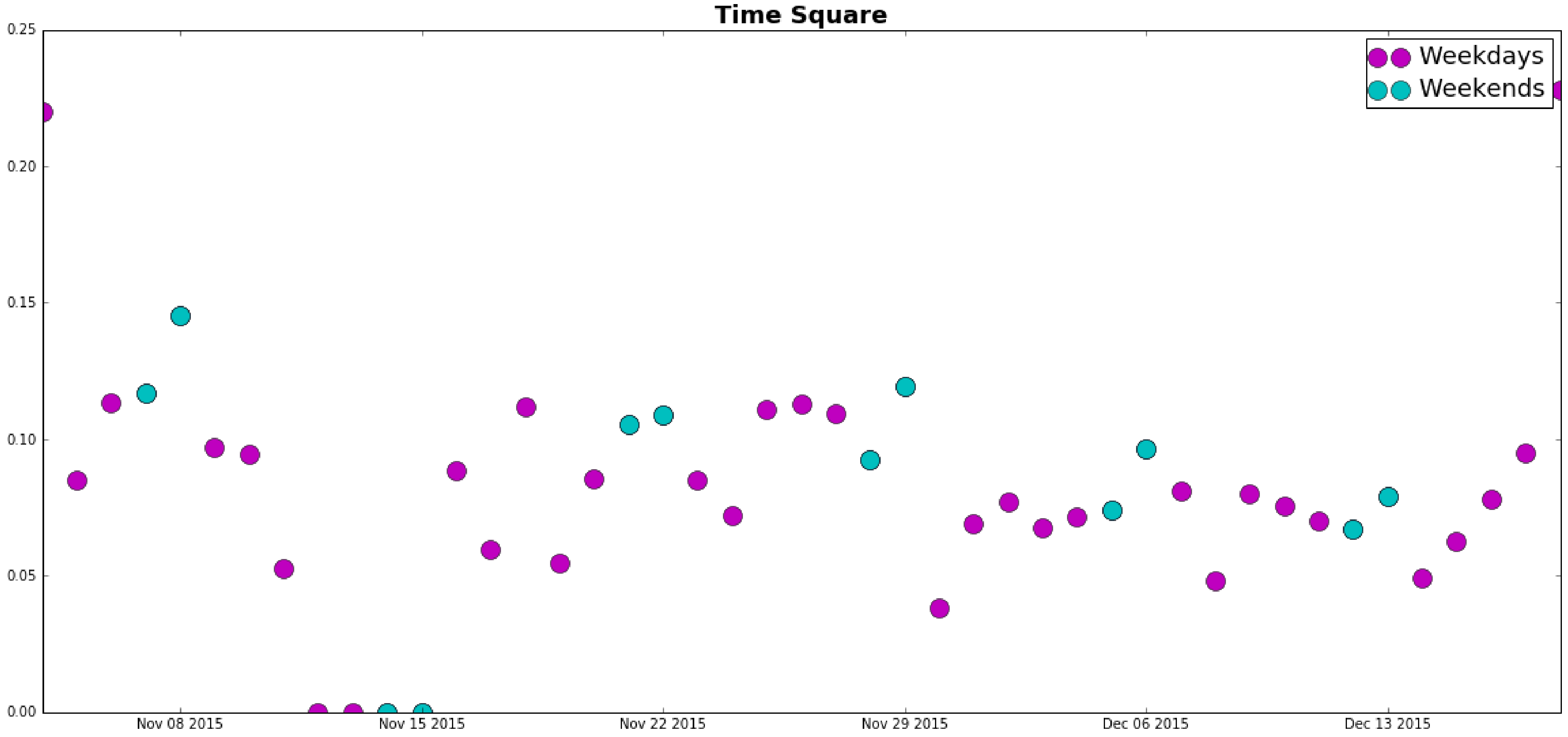
10007, City Hall (un-normalized value)

Non-zero part	mean	std
weekdays	0.248263	0.066790
weekends	0.312098	0.013290



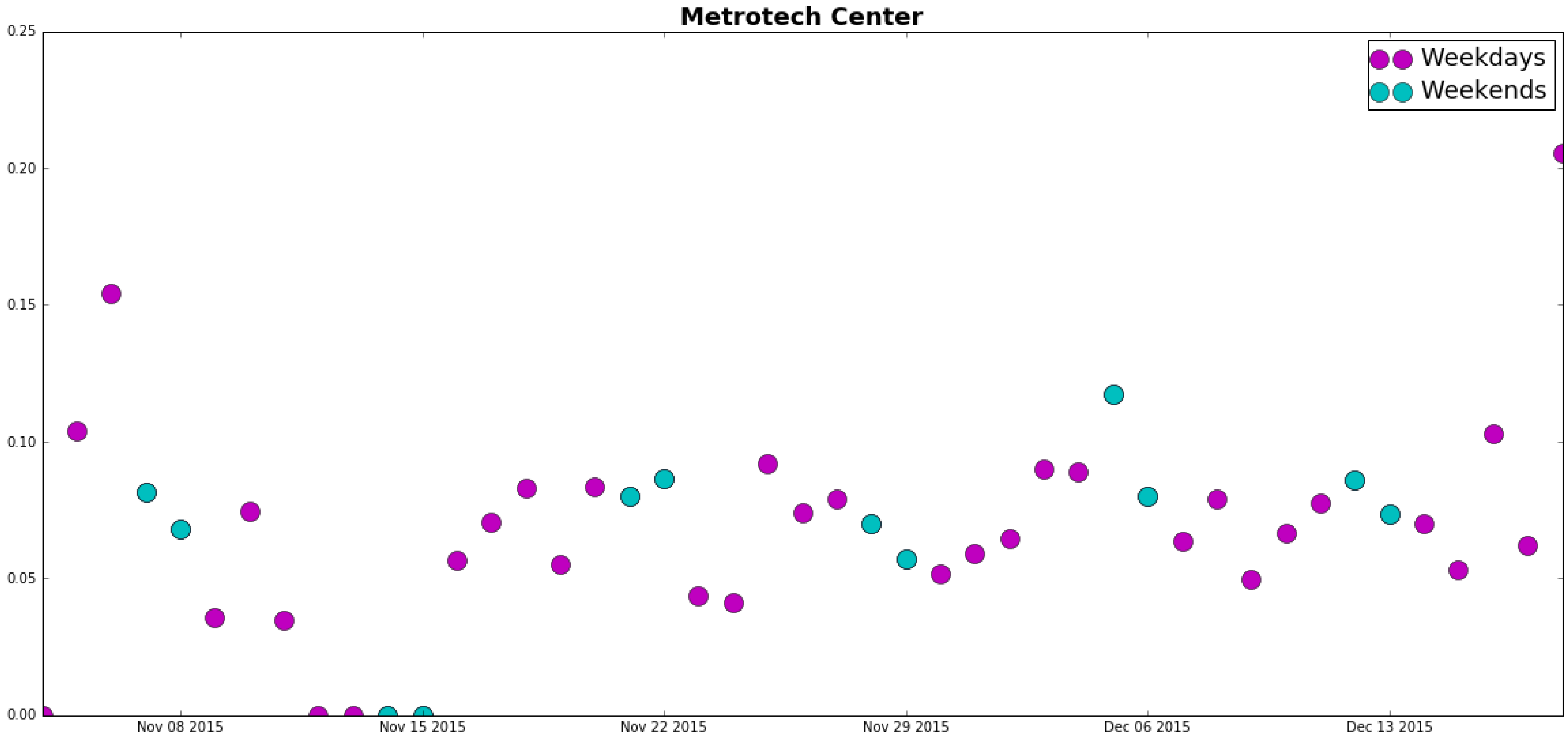
10036, Time Square (un-normalized value)

Non-zero part	mean	std
weekdays	0.088541	0.041426
weekends	0.100421	0.023760



11201, Metrotech Center (un-normalized value)

Non-zero part	mean	std
weekdays	0.073098	0.036409
weekends	0.080092	0.015843



The sum of W_b^{in}

$\sum_b W_b^{in}(t)$ seems to be stable (around 50) if the data volume is high enough (>13000 entries, roughly speaking)