American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS)
ISSN (Print) 2313-4410, ISSN (Online) 2313-4402

http://asrjetsjournal.org/

# Towards a Better Understanding of Effectiveness of Bike-share Programs: Exploring Factors Affecting Bikes Idle Duration

Boniphace Kutela[a]*, Emmanuel Kidando[b]

[a]*Assistant Lecturer, Department of Civil Engineering Ardhi University, P. O. Box 35176, Dar es salaam, Tanzania*
[b]*Graduate Research Student, Department of Civil and Environmental Engineering FAMU-FSU College of Engineering, 2525 Pottsdamer Street, Tallahassee, FL 32310*
[a]*Email: kutela4@gmail.com*
[b]*Email: ek15f@my.fsu.edu*

**Abstract**

Bike-share program is considered effective and reliable if its stations have bikes and empty docks available at any time of a day. Few studies have considered idle bikes in the system and even lesser have glanced on modeling bikes idle duration (BID) in the bike-share system. This study applied descriptive statistics and log-logistic hazard based model on one year Seattle bike-share ridership data to quantify the BID and determine factors associated with the bikes' idle duration. The findings of the study illustrate that the most and least effective utilized bike were used for 161 hours and 0.19 hours respectively for the entire year. Winter season, especially when raining and snowing was found to increase the likelihood of long BID. On the other end, the bikes located in commercial areas were associated with short BID compared to residential land-use. Moreover, weekend days and evening peak hours (4 p.m. to 6 p.m.) are associated with less likelihood of the BID compared with weekdays and morning peak hours respectively. These findings will facilitate procedures to identify the idle bikes for redistribution strategy and enhancing effective utilization of the bike-share system.

***Keywords:*** Bike-share program; Bike Idle Duration (BID); Hazard based model.

------------------------------------------------------------------------

* Corresponding author.

## 1. Introduction

The bike-share program is a relatively recent initiative to encourage more people making short trips to use bikes for their daily travel. It is one of the effective methods of reducing car dependent and reducing green gas emissions. Bike-share provides users freedom of picking up and return a bicycle at any bike station within the respective scheme's service area. There have been several efforts to make sure that the programs are effective. One of such efforts attempts to make sure that users can find bikes and an empty dock at any time of the day. The absence of the bikes and empty docks at a station is one of the major causes of the users to abandon the bike-share program [1]. Thus, researchers determine ways to balance the system [2]. To make sure that both bikes and empty docks are available in the stations regularly, most of the bike-share programs adopted a bike-rebalancing strategy by which bikes are transported by a vehicle from stations with more bikes to the ones with fewer bikes. However, there are cases where the bikes are available in the system (station) but are unusable [3] or unused. In such situation, the bike-share program becomes less effective and unprofitable. The system creates a hidden cost that most operators have not explored. The longer the bikes remain unused the less the program makes a profit. The operator may incur bike servicing cost or the regular maintenance costs while the bikes have not been used. The maiden review of the Seattle bike-share program (Pronto) data indicated that the effective bike hours utilized per year were only about 2%. For the bike program to be effective, the quantity of the idle time of an individual bike and the factors associated with the bike idle duration must be considered.This study utilized both descriptive and inferential statistics to quantify the bikes idle duration (BID) and determine its association with other independent factors. The Seattle bike-share publicly accessible data that contain trip, weather and station information was used. In addition, Seattle land use data were utilized to assess the association with the BID. To quantify this quantity of individual bike, the difference between the maximum effective bike hour and total bike utilization time was computed. The log-logistic hazard based regression was applied to determine the associated factors. The variables of interest in the model were temporal and land-use factors.

## 2. Previous Studies

Since its establishment in Amsterdam in 1960's, bike-sharing programs have been adopted by over 700 cities around the world [4]. They have been effective in creating a larger cycling population, even in the cities without prior cycling tradition [5]. They have been receiving great attention in academics and practitioners researching at the system level and the station level [6] so they may be a more efficient mode of transportation. Past studies dealt with the wide range of topics regarding design and operations. The design studies focus on locating the stations, the capacity (number of lockers, also known as docking points) of each station and the fleet [7]. On the other end, the operational studies centered their interests on bikes redistribution/balancing [2], since it was the major part of the operational cost [8]. One key strategy to alleviate imbalances was the use of IT-systems of recording data from bike sharing program proposed by [9].

Kaspi and his colleagues [10] described three situations that may cause a bike to remain idle at a dock. The situations involve all three major participants in the bike-share system; the user, the bike, and the station. Per their study, the bike idle situation may occur if (1) no renters have arrived at the station, (2) renters have arrived at the station, but no bicycle can be used and no rent transaction has occurred, or (3) the station is malfunctioning. The last two reasons cannot be traced by the dataset used in this study yet they have a very crucial impact on the user satisfaction if they are not taken care of [3]. Most of the bike-share systems provide to the public on-line aggregated information about each station. Regardless of the distance from the bike station to user location, smartphone users may query the state of each station, thus, obtain in real-time, the number of bicycles and empty locker/docks available. Not only that unusable bike at the station reduces the number of usable lockers, but also, they provide inaccurate information regarding the usability of bicycles. They appear to be available at the station when the user checks the availability online. Parallel to that, the operator incurs unobservable costs by having the idle bikes at the station for a long time. Literature reviews reveal that few studies have focused on the unused bikes at the bike station and particularly on the time the bike remain idle at the station (BID). The time the bikes remain idle in the system and the factors associated with bike idle duration has not been given enough attention by researchers. Therefore, this study quantified the bike idle duration and presented the associated factors.

## 3. Data Description and Processing

This study utilized one-year public accessible bike-share data from Seattle bike-share program (Pronto) collected between October 2014 and October 2015. The bike-share data are publicly accessible and contains information such as trips, stations, and weather. The trip data have information related to trip time (when the trip was initiated and terminated), the bike identification number (which is unique for each bike), the station name and id where the trip was originated and ended, the use type (annual or short time users) and the gender and birth year of the annual members. The time stamp and bike id were the vital information in identifying bikes idle durations (BID). The weather data contained the daily weather information in the service area. Regarding the station data, the station name and id, the total count of the docks at the station and the coordinates are reported. Furthermore, we obtained land use data downloaded from Seattle Department of Construction and Inspections. The Seattle land use data described the land use characteristics of all the locations in Seattle, including those where the bike-share stations are located. These can be categorized into commercial, multi-family/residential (low rise, and high-rise) and mixed land use. The Quantum Geographical Information System (QGIS) software version 2.12.3 facilitated the identification of the major trip generators or attractors located within walkable distance (0.5 miles) from the bike-share station. With consideration of the start time, stop time and bike id variables in the trip data, the BIDs were computed. The BIDs were computed as the difference in the time between the check in of the bike (stop time) to the next check out (start time) of the same bike. In all 142,365-bike usage incidences, the incidences of the idle durations of an individual bike at a time varied from zero minutes to 323 days. This means, there was a situation that a bike that remained idle at the station for 323 days without being used while there were other situations that bikes were checked out as soon as preceding users returned them. To clearly show the difference in these situations, the idle durations were grouped into four categories (Figure 1). With 102,993 incidences, the idle time less than one-day accounts for about 72%, this idle duration can be the situation by which bikes were effectively utilized. The second category by considering the number of incidence (34,026) which accounted for 24% of all incidents was the idle duration between one day and five days. The idle duration between five and ten days accounted for 2.6%, while the longest idle duration category, which is over ten days, explanations 1.1% of all the incidents. Having that long idle duration of the bike does not translate that the bikes were not used, but there were repetitions of usage of the same bikes while some others remained unused for a long period.

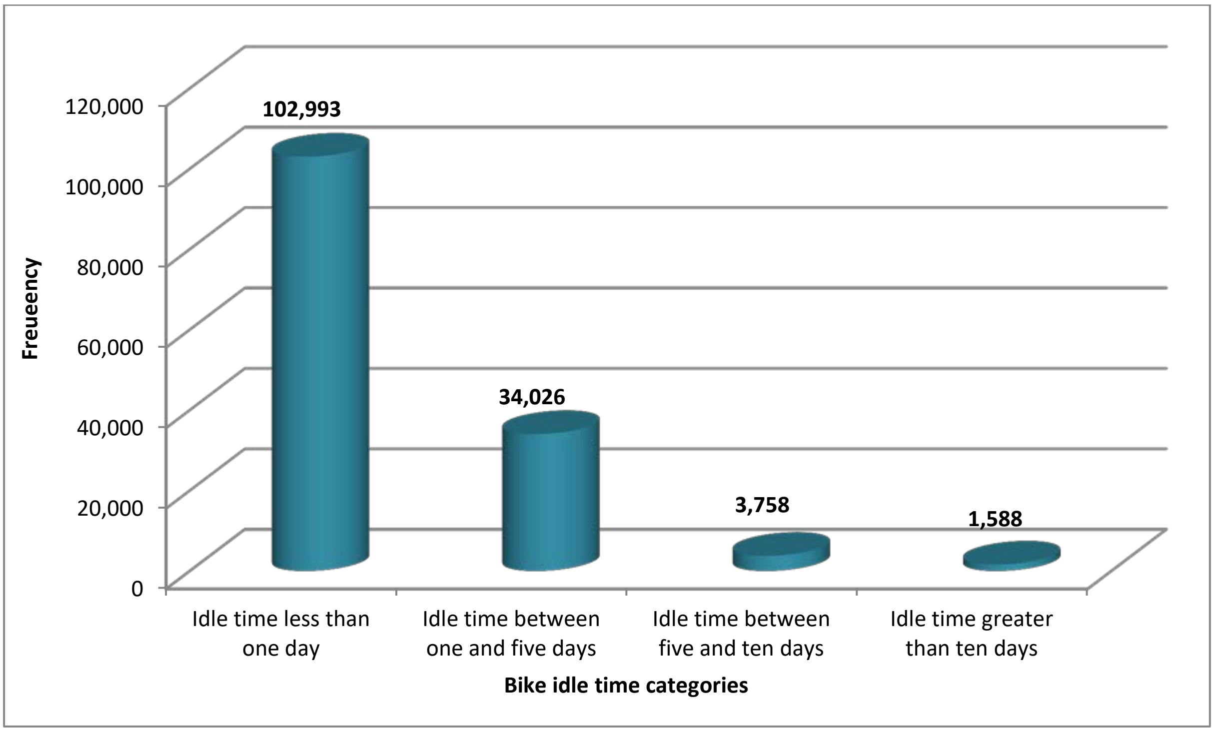


**Figure 1:** Number of observations for each idle time category

### *3.1. Descriptive statistics*

*Aggregating the bikes idle duration*

The aggregated idle duration of the bike (BID) for the entire year was computed as a sum of all idle duration of the same bike for the whole year. Although the data shows that the bike riding activities were done for the whole day, effective bike usage time was less during nighttime. Therefore, we considered only eighteen (18) hours per day as the effective bike usage time in the analysis. The sum of all trip durations for the most effectively used bikes was 161.42 hours, which suggests that the bike remained idle for 6408.58 hours, which equals 268 full days per year. The summation of the trip duration for the least used bikes was 0.19 hours per year which is equivalent to 11.4 minutes, this implies, the bike remained idle for almost for the entire year. The percentage of the utilized and idle bike hour to the total bike hours available per year were computed, Table 1 shows ten least and ten most utilized bikes per year respectively. Considering 18 hours as the effective bike utilization time per day, results in Table 1 indicate that the maximum attainable percentage bike utilization is 2.46%, which means the remaining 97.54% was the idle time.

The summation of the trip duration for the least used bikes was 11.4 minutes, which means the bike remained idle for almost for the entire year. The percentage of the utilized and idle bike hour to the total bike hours available per year were computed, Table 1 shows ten least and ten most utilized bikes per year respectively.

Considering 18 hours as the effective bike utilization time per day, results in Table 1 indicate that the maximum attainable percentage bike utilization is 2.46%, which means the remaining 97.54% was the idle time.

**Table 1:** The Least and Most Utilized Bikes per Year The ten least utilized bikes per year

| SN | Bike ID | Utilization time (hrs.)/year | Maximum bike hours/year | Idle bike hours/year | Percentage bike hour utilized | Percentage idle bike hour |
|---|---|---|---|---|---|---|
| 1 | SEA00001 | 0.19 | 6570 | 6569.81 | 0.00% | 100.00% |
| 2 | SEA00012 | 0.21 | 6570 | 6569.79 | 0.00% | 100.00% |
| 3 | SEA00011 | 1.04 | 6570 | 6568.96 | 0.02% | 99.98% |
| 4 | SEA00378 | 2.00 | 6570 | 6568.00 | 0.03% | 99.97% |
| 5 | SEA00331 | 3.11 | 6570 | 6566.89 | 0.05% | 99.95% |
| 6 | SEA00130 | 5.29 | 6570 | 6564.71 | 0.08% | 99.92% |
| 7 | SEA00225 | 7.36 | 6570 | 6562.64 | 0.11% | 99.89% |
| 8 | SEA00076 | 7.48 | 6570 | 6562.52 | 0.11% | 99.89% |
| 9 | SEA00123 | 8.99 | 6570 | 6561.01 | 0.14% | 99.86% |
| 10 | SEA00050 | 10.52 | 6570 | 6559.48 | 0.16% | 99.84% |
| The ten most utilized bikes per year | | | | | | |
| 1 | SEA00046 | 144.47 | 6570 | 6425.53 | 2.20% | 97.80% |
| 2 | SEA00390 | 144.70 | 6570 | 6425.30 | 2.20% | 97.80% |
| 3 | SEA00142 | 145.98 | 6570 | 6424.02 | 2.22% | 97.78% |
| 4 | SEA00121 | 146.15 | 6570 | 6423.85 | 2.22% | 97.78% |
| 5 | SEA00481 | 146.20 | 6570 | 6423.80 | 2.23% | 97.77% |
| 6 | SEA00029 | 148.83 | 6570 | 6421.17 | 2.27% | 97.73% |
| 7 | SEA00222 | 150.41 | 6570 | 6419.59 | 2.29% | 97.71% |
| 8 | SEA00413 | 153.99 | 6570 | 6416.01 | 2.34% | 97.66% |
| 9 | SEA00218 | 158.32 | 6570 | 6411.68 | 2.41% | 97.59% |
| 10 | SEA00453 | 161.42 | 6570 | 6408.58 | 2.46% | 97.54% |

*Variation of BID by seasons of the year*

The variation of the BID by the seasons of the year was also studied. The aim was to determine the season with high bike usage and the associated bike idle duration regarding the different categories of the idle duration. As it was expected, summer season accounted for most (36.4%) of the trips made in 2014/2015 followed by autumn

(27%), spring (22.2) and the winter season had the least (14.4%) bike utilization. Moreover, the summer season accounted for the most of the idle bikes remained on the docks for less than a day (Figure 2). On the other end, winter season was the leading contributor for the bikes remained idle for five days or more. This might be attributed to the unfavorable weather condition during the winter season.

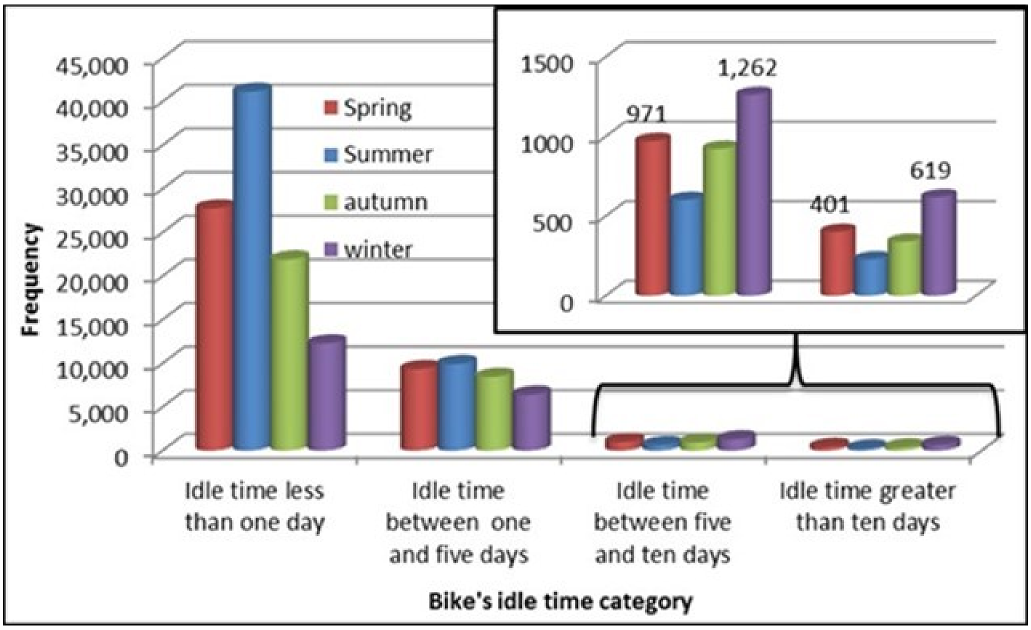


**Figure 2:** Bike idle durations per season of the year.

To explicitly reveal the number of bikes remained idle for every hour of the day, the average number of bikes checked out and returned was compared to the total number of bikes available in the system. The system had 482 bikes stationed in 54 stations.

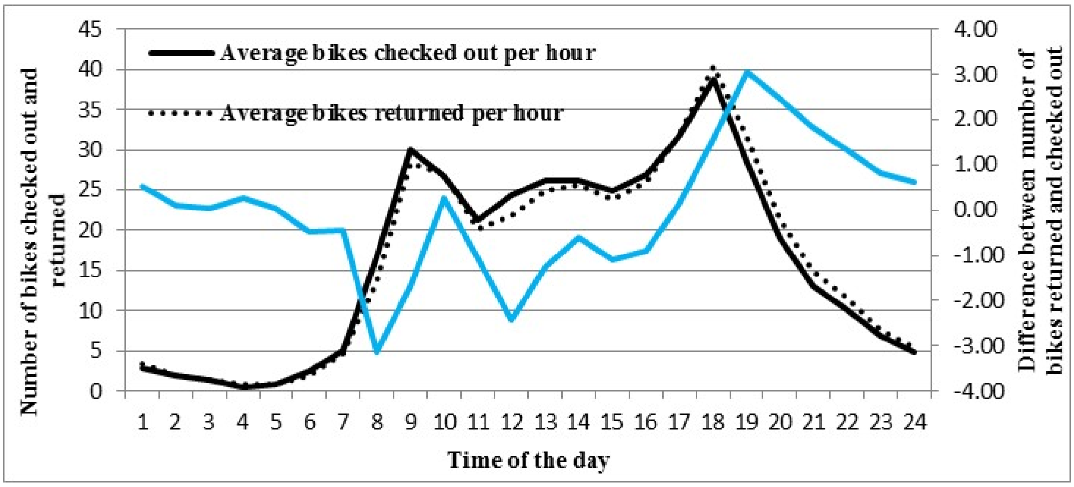


**Figure 3:** Bike utilization per each hour of the day

Figure 3 shows the average bike utilization in terms of the number of checkouts and check-ins (returns) for the 24 hours of the day in the entire year. It can be shown that there were more bike utilization rates during the daytime than in the night times. Among the available 482 bikes in the system, on average, the maximum number of bikes checked out per hour was 40, this occurred between 5 pm and 6 pm. The rest bikes, which were about 440 bikes, remained idle. The trip analysis indicated that peak hours for evening and morning are 9 am and 7 pm respectively.The right-hand side axis of Figure 3 presents the difference between the average number of bikes returned and checked out in the day for the entire year. The positive values imply that more bikes were returned than checked out while the negative values denote the vice versa. It can be observed that from around 4:30 am to 4:15 pm there were more there were more checkouts than returns while from around 4:15 pm to around 4:30 am there more returns than checkouts. At this period (4:15 pm to 4:30 am) the system had more idle bikes. It was also observed that around 4:30 am and 4:15 pm, the system balanced. However, balancing of the system does not imply that the bikes were not idle, but the number of checkouts and those of check-ins was equal.

**4. Bikes Idle Duration (Bid) Modeling Method**

The hazard-based duration (HBD) model was applied in this study. Originated in the medical and industrial engineering field [11], these models have penetrated in transportation engineering, especially in estimating traffic incident duration over the past years. They are based on the survival theory, by which the time until an event of interest occurs is the outcome variable. The bike idle duration aligns with survival model theory and assumptions applied in incident duration estimations. For instance, the survival of the incident on a roadway is the time until it is cleared [12], the same applies for the survival of the idle bike on the dock/station; it is the time until the next checkout is performed. Different hazard-based models have been applied in modeling these time-accelerated events including Cox regression, Proportional hazard Weibull mixtures, log-logistic and others [12-15]. To specify the effects of dependent variables on the hazard function, the proportional hazard (PH) models and accelerated failure time (AFT) models have been applied. The (PH) models rely on the assumption that regression coefficients don’t change with time while the (AFT) models assume the time scale of the survival function is rescaled by the covariates [11]

The hazard function $h(t/_Z)$ and survival function $S(t/_Z)$ can be presented as (*16*);

$$h(t/_Z) = h_o(t)g(\beta, Z) \quad (1)$$

$$S(t/_Z) = S_o(t)g(\beta, Z) \quad (2)$$

$$g(\beta/_Z) = exp(\beta, Z) \quad (3)$$

where $h_o(t)$ represents the baseline hazard function, $S_o(t)$ implies the baseline survival function and $g(\beta, Z)$ indicates the effect of explanatory variable on hazard and survival time. A distribution assumption such as exponential, lognormal, log-logistic, Weibull, and Gompertz are required for the parametric formulation of the baseline hazard function. Almost each of the distribution assumptions has a shortfall. The exponential distribution is constant with time; the Weibull distribution is limited to monotonicity. This challenges are address by applying the log-logistic and lognormal distributions [16]. These models both begin with log linear but are different on the assumption of the error term. Log-logistic error follows the logistic distribution while the lognormal error follows the standard normal distribution [17]. The generalized equation is given as;

$$\ln t_i = X_i\beta + \varepsilon_i \qquad (4)$$

whereby X = vector of covariates, β = vector of estimated coefficients and ε = error term

The log-logistic was found to yield sound result than lognormal by Nam and Mannering study when investigated highway incident duration [18]. Therefore, in this study also log-logistic model was used to evaluate the influence of factors on Bikes Idle Duration (BID). The log-logistic model is given as:

$$f(t) = \frac{\lambda\kappa\left(\lambda t\right)^{\kappa-1}}{\left(1+\left(\lambda t\right)^{\kappa}\right)^{2}};\ t>0; \qquad (5)$$

$$S(t) = \frac{1}{1+\left(\lambda t\right)^{p}};\ t>0 \qquad (6)$$

$$h(t) = \frac{f(t)}{s(t)} = \frac{\lambda\rho\left(\lambda t\right)^{p-1}}{1+\left(\lambda t\right)^{p}};\ t>0 \qquad (7)$$

where, *f(t)* id the distribution function, is the survival function and *h(t)* is the hazard function, λ is a positive scale parameter and *p* is the shape parameter.

**Variable coding and correlation check**

Prior to modeling, it is common to check correlation among variables.

The results of the analysis revealed that the maximum correlation coefficient was about 0.47, which was between precipitation and rain variables, while the minimum was 0.0001 between spring season and residential locations.

Since variables were not highly correlated, all variables were used in the model.

Table 2 shows the summary of the descriptive statistics of the coded variables considered in the model. The average idle duration of the bikes is 27.3 hours whereby the minimum is zero hours and the maximum is 7757 hours.

The descriptive statistics of other variables can be observed from the table 3.

## 5. Model Results and Discussion

The effects of each variable to the bikes idle duration (BID) are as shown in Table 3. A positive sign of a parameter estimate suggests an increase in the BID and a decrease in hazard function associated with an increase or presence of that variable. These paragraphs summarize finding of the model in Table 3.

**Table 2:** Summary of Variables' Descriptive Statistics

| Variable | Type | Observations | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|---|
| **Dependent variable** | | | | | | |
| Bike idle duration (hours) | Continuous | 142364 | 27.321 | 80.965 | 0 | 7757 |
| **Independent variables** | | | | | | |
| **Temporal variables** | | | | | | |
| Fall | Binary (yes 1 no 0) | 142846 | 0.224 | 0.417 | 0 | 1 |
| Winter | Binary (yes 1 no 0) | 142846 | 0.144 | 0.351 | 0 | 1 |
| Summer | Binary (yes 1 no 0) | 142846 | 0.363 | 0.481 | 0 | 1 |
| Spring | Binary (yes 1 no 0) | 142846 | 0.269 | 0.443 | 0 | 1 |
| Weekday | Binary (yes 1 no 0) | 142846 | 0.735 | 0.441 | 0 | 1 |
| Evening peak | Binary (yes 1 no 0) | 142846 | 0.266 | 0.442 | 0 | 1 |
| **Weather condition** | | | | | | |
| Rain | Binary (yes 1 no 0) | 142846 | 0.352 | 0.478 | 0 | 1 |
| Rain and fog | Binary(yes 1 no 0) | 142846 | 0.013 | 0.113 | 0 | 1 |
| Rain and snow | Binary (yes 1 no 0) | 142846 | 0.000 | 0.021 | 0 | 1 |
| Precipitation (in) | Continuous | 142846 | 0.048 | 0.138 | 0 | 2.2 |
| **Spatial variables** | | | | | | |
| Residential land use | Binary (yes 1 no 0) | 142833 | 0.054 | 0.227 | 0 | 1 |
| Mixed land use | Binary (yes 1 no 0) | 142833 | 0.584 | 0.493 | 0 | 1 |
| Commercial land use | Binary (yes 1 no 0) | 142833 | 0.274 | 0.446 | 0 | 1 |
| **Trip attractors or generators within 0.5 miles** | | | | | | |
| Residences | Binary (yes 1 no 0) | 142845 | 0.149 | 0.356 | 0 | 1 |
| Transportation hubs | Binary (yes 1 no 0) | 142833 | 0.067 | 0.250 | 0 | 1 |
| Offices | Binary (yes 1 no 0) | 142833 | 0.492 | 0.500 | 0 | 1 |
| Recreation | Binary (yes 1 no 0) | 142833 | 0.273 | 0.445 | 0 | 1 |

**Table 3:** Log-Logistic Bike Idle Duration Survival Model Results

| **Log-Logistic Model estimates** | | | | | **Marginal effects** | | | |
|---|---|---|---|---|---|---|---|---|
| **Bike idle duration (hrs.)** | **Coeff** | **Percent change (%)** | **Std. Err.** | ***P-value*** | **dy/dx** | **Std. Err.** | **z** | ***P-value*** |
| **Temporal variables** | | | | | | | | |
| Winter | 0.412 | 50.98 | 0.018 | 0.000 | 3.510 | 0.154 | 22.79 | 0.000 |
| Summer | -0.339 | -28.75 | 0.015 | 0.000 | -2.882 | 0.127 | -22.64 | 0.000 |
| Spring | -0.111 | -10.51 | 0.015 | 0.000 | -0.947 | 0.129 | -7.34 | 0.000 |
| Weekday | 0.193 | 21.29 | 0.012 | 0.000 | 1.643 | 0.104 | 15.74 | 0.000 |
| Evening peak | -0.435 | -35.27 | 0.012 | 0.000 | -3.700 | 0.107 | -34.71 | 0.000 |
| **Weather condition** | | | | | | | | |
| Rain | 0.071 | 7.36 | 0.013 | 0.000 | 0.605 | 0.111 | 5.46 | 0.000 |
| Rain and fog | 0.105 | 1107 | 0.046 | 0.022 | 0.891 | 0.39 | 2.29 | 0.022 |
| Rain and snow | 1.363 | 290.8 | 0.237 | 0.000 | 11.608 | 2.023 | 5.74 | 0.000 |
| Precipitation (in) | 0.529 | 69.71 | 0.042 | 0.000 | 4.502 | 0.36 | 12.51 | 0.000 |
| **Land-use variables** | | | | | | | | |
| Mixed land use | -0.332 | -28.3 | 0.016 | 0.000 | -2.830 | 0.142 | -19.97 | 0.000 |
| Commercial land use | -0.439 | -35.5 | 0.018 | 0.000 | -3.733 | 0.159 | -23.43 | 0.000 |
| **Trips generators within 0.5 miles of a bike station** | | | | | | | | |
| Transportation hubs | 0.019 | 0.8 | 0.024 | 0.422 | 0.166 | 0.207 | 0.8 | 0.422 |
| Offices | 0.084 | 1.92 | 0.015 | 0.000 | 0.711 | 0.127 | 5.58 | 0.000 |
| Recreation | -0.018 | -1.9 | 0.017 | 0.291 | -0.154 | 0.146 | -1.06 | 0.291 |
| /ln_gam | 0.119 | | 0.002 | 0.000 | | | | |
| gamma | 1.127 | | 0.002 | | | | | |
| Number of subjects | 141,743 | | | | | | | |
| Number of failures | 141,743 | LR chi$^2$(14) | | 5,541.2 | | | | |
| Time at risk | 3,887,287 | Prob > chi$^2$ | | 0.000 | | | | |
| Log likelihood | -299,731 | | | | | | | |

The result of temporal factors (season of the year, the day of the week and time of a day) in the model indicated all variables are statistically significant at 1% significance level (Table 3). In modeling season, the fall season was the reference category in which the other seasons were compared in the analysis. The regression results revealed that the winter season has the highest impact on BID, suggesting that the BID of an individual bike increases by about 3.5 hours in the winter season compared to the fall season. In contrast, the likelihood of BID in both the summer and spring seasons is lower than fall season. Results indicated a decrease of 2.8 hours and approximately one hour during summer and spring seasons respectively. These findings are consistent with a study by Ma and his colleagues [19] who found there was less bike ridership in winter season while more people rode bikes during summer seasons.Bearing in mind the peak hours, evening peak hours revealed a negative effect on BID as compared to morning peak hours. The model suggests that bikes returned during evening peak hours between 4 and 6 pm were statistically significant less likely to remain idle for a long time without being checked out for the next trip. To be specific, the idle durations of the bikes were more likely to decrease by approximately 3.7 hours during evening peak hour as compared to morning peak hours. Intuitively, we expect to have less idle durations during peak hours since bikes are frequently utilized. This finding supports the descriptive statistics indicated in Figure 3. In addition, regarding the day of the week in our model, weekdays revealed a higher likelihood of BID than weekends. The increase of idle duration estimated by the model about 1.6 hours.

*Weather condition*

Rain, fog, snow and precipitations were the variables considered in the regression model. The variables were interacted in order not only to expose the impact of a single variable but also the combined impact reflecting the reality. It is common to find rain and fog or rain and snow or fog and snow or both simultaneously. The combined effect of rain and snow increases the likelihood of longer idle duration. The results show that the bikes were more likely to remain idle for about 11.6 hours when there were rain and snow compared to clear weather condition. The magnitude of the BID was lower under combined effect of rain and fog compared to the previous combination. It is estimated that the bikes idle duration increases by 0.9 hours during this weather condition as compared to the clear weather. The rain only event was associated with the increase in BID by approximately 0.6 hours. The higher the precipitation amount the higher the likelihood of the longer the BID. Furthermore, results show that for every inch increase in precipitation there is 4.5 hours increase in bikes idle duration. The results resonate with the previous study [20]. Excluding the combined rain and fog variable, which was significant at 5% level, the rest weather condition factors were statistically significant at 1% significance level.

*Land-use factors*

The land-use was defined per the location of the station; a buffer size of 0.5 miles was defined for each station to determine the main generators and attractors of the bike trips. Residential, commercial and mixed (residential and commercial) were the main three land uses while transportation hubs, offices, recreation, and residences were the attractors and generators within 0.5 miles. The bikes located in the stations within commercial and mixed land use locations were 2.8 hours and 3.7 hours less likely to remain idle compared with those at the residential locations. About trips generators and attractors within 0.5 miles of the bike stations, the regression results highlighted that the stations located close to the offices were more likely to remain idle compared to the bikes whose proximity locations were residences. However, the transportation hubs were not statically significant in our model. The results show consistency with findings reported by Bachand-Marleau and his colleagues [21] conducted in Montreal, Canada who utilized online survey data and found that the proximity of docking stations to residential housing increases bike-share trip frequency, thus, decreases the probability of having idle bikes at the station. On the other end, Daddio and Mcdonald [22] results were found contrary to our findings, suggesting that proximity to the metro rail was positively correlated with bike trip generation. In addition, recreation locations such (parks, beaches etc.) were not statistically significant at 5% level in in our study.

## 6. Conclusion and recommendations

This study applied the descriptive and survival model to quantify Bike Idle Duration (BID) and determine the associated factors. It was found that the effective time an individual bike has been utilized in a year range from 11.4 minutes to approximately 6.7 days.

The winter season, rainy weather condition with higher precipitation amount were found to increase BID while during evening peak hour period (4 p.m. to 6 p.m.), and bike in commercial areas was found to decrease the likelihood BID. Comparing with weekend days, weekdays were associated with the increase of the likelihood of the long BIDs. The findings of this study can be used to develop a data-driven decision making regarding the redistribution strategy.

This can be achieved through identifying the idle bikes in the system so they can be transferred and used in other more active stations. To this end, it is recommended that the number of bikes to be reduced from the system during the winter season because they are exposed to unfavorable weather condition while are under-utilized.

## 7. Further study

This study evaluated the impact of temporal, spatial and land use factors on the bike idle duration (BID). Most of these factors are not under human control, thus, it becomes difficult to address them.

Therefore, further research should incorporate the human controllable factors such as the frequency and locations of the bike redistributions, the number of operators, operation modes and others in the model. This will enhance developing countermeasures to improve the efficiency of the bike-share program.

## References

[1] T. Raviv and O. Kolka, "Optimal inventory management of a bike-sharing station," IIE Trans., vol. 45, no. 10, pp. 1077–1093, Oct. 2013.

[2] D. Chemla, F. Meunier, and R. Wolfler Calvo, "Bike sharing systems: Solving the static rebalancing problem," Discret. Optim., vol. 10, no. 2, pp. 120–146, 2013.

[3] M. Kaspi, T. Raviv, and M. Tzur, "Detection of unusable bicycles in bike-sharing systems," 2016.

[4] P. J. Demaio, "Smart Bikes: Public Transportation for the 21st Century," Transp. Q., vol. 57, no. 1, pp. 9–11, 2003.

[5] C. Romero, "SpiCycles in Barcelona," in Chamber of Commerce and Industry of Romania, 2008.

[6] X. Wang, G. Lindsey, J. E. Schoner, and A. Harrison, "Modeling Bike-share Station Activity: Effects of Nearby Businesses and Jobs on Trips to and from Stations," J. Urban Plan. Dev., vol. 142, no. 1, p.

4015001, Mar. 2016.

[7] J.-R. Lin and T.-H. Yang, “Strategic design of public bicycle sharing systems with service level constraints,” Transp. Res. Part E Logist. Transp. Rev., vol. 47, no. 2, pp. 284–294, 2011.

[8] J. Schuijbroek, R. Hampshire, and W.-J. van Hoeve, “Inventory Rebalancing and Vehicle Routing in Bike Sharing Systems,” 2013.

[9] P. Vogel and D. C. Mattfeld, “Strategic and Operational Planning of Bike-Sharing Systems by Data Mining – A Case Study,” Springer Berlin Heidelberg, 2011, pp. 127–141.

[10] M. Kaspi, T. Raviv, and M. Tzur, “Detection of Unusable Bicycles in Bike-Sharing Systems,” 2015.

[11] Kalbfleisch D. John and Ross L. Prentice, The Statistical Analysis of Failure Time Data - John D. Kalbfleisch, Ross L. Prentice - Google Books. New York: John Wiley & Sons, 1980.

[12] D. Chimba, B. Kutela, G. Ogletree, F. Horne, and M. Tugwell, “Impact of Abandoned and Disabled Vehicles on Freeway Incident Duration,” J. Transp. Eng., vol. 140, no. 3, p. 4013013, 2014.

[13] W. Hui, “Proportional Hazard Weibull Mixtures,” Canberra, 1990.

[14] J.-T. Lee and J. Fazio, “Influential Factors in Freeway Crash Response and Clearance Times by Emergency Management Services in Peak Periods,” Traffic Inj. Prev., vol. 6, no. 4, pp. 331–339, Dec. 2005.

[15] S. Hasan, R. Mesa-Arango, and S. Ukkusuri, “A random-parameter hazard-based model to understand household evacuation timing behavior,” Transp. Res. Part C Emerg. Technol., vol. 27, pp. 108–116, 2013.

[16] B. Jones, “Duration models: Parametric models,” Univ. of California, Davis, 2011.

[17] D. Kundu, R. D. Gupta, and A. Manglick, “Discriminating between the log-normal and generalized exponential distributions,” J. Stat. Plan. Inference, vol. 127, no. 1, pp. 213–227, 2005.

[18] D. Nam and F. Mannering, “An exploratory hazard-based analysis of highway incident duration,” Transp. Res. Part A Policy Pract., vol. 34, no. 2, pp. 85–102, 2000.

[19] B. Ting Ma, C. Liu, and S. Erdoğan, “Bicycle Sharing and Transit: Does Capital Bike-share Affect Metrorail Ridership in Washington, D.C.?,” in 83rd Annual Meeting of the Transportation Research Board, 2015.

[20] K. Gebhart and R. B. Noland, “The Impact of Weather Conditions on Capital Bike-share Trips.” 2013.

[21] J. Bachand-Marleau, B. Lee, and A. El-Geneidy, "Better Understanding of Factors Influencing Likelihood of Using Shared Bicycle Systems and Frequency of Use," Transp. Res. Rec. J. Transp. Res. Board, vol. 2314, pp. 66–71, Dec. 2012.

[22] B. David William Daddio and N. Mcdonald, "MAXIMIZING BICYCLE SHARING: AN EMPIRICAL ANALYSIS OF CAPITAL BIKE-SHARE USAGE," University of North Carolina at Chapel Hill, 2012.