AN ANALYSIS OF THE FACTORS INFLUENCING JOURNEY TIME VARIATION IN THE CORK PUBLIC BIKE SYSTEM

Cameron McBain

MAI Student

Trinity College Dublin

Brian Caulfield

Associate Professor

Trinity College Dublin

Abstract
The Cork Public Bike System was opened in December 2014. This system consists of a total of 330 bikes across 31 stations in Cork, with the scheme serving all the main trip attractors in the city. By using datasets of all trips taken in the scheme during 2015 and 2016, a number of new spatial, temporal and weather variables were assigned to all 560,000 trips that have been used for the analysis in this research, in addition to a journey time that has been predicted by Google Maps for each individual route combination. These journey times from Google Maps were then compared with the actual journey times recorded in the dataset to create a new variable called Extra Travel Time.

Trips to and from stations that had a lower number of shops within walking distance were also found to be likely predictors of the quickest trips. The stations with above average public transport links also are associated with quicker trips suggesting that users of public bikes will take quick, direct journeys if they are trip-chaining with other transportation modes such as bus or rail.

Introduction

Cycling in Ireland, like many other countries, has seen increases over the past decade. The numbers cycling in the major cities in Ireland (Dublin, Cork and Galway) are all on the increase (Caulfield, 2014). This increase can be attributed to a number of factors. In 2009, a number of policies were introduced in Ireland to encourage this growth in cycling. A tax-free loan bicycle purchase scheme called "Cycle to Work Scheme". This was introduced to enable the public to purchase bikes with an up to 50% discount and some research has shown this scheme to have been one of the main factors behind the increase in cycling in Ireland (Caulfield and Leahy, 2011).

Cork is a small city in the South-West of Ireland and is the second largest city in Ireland in terms of population with approximately 125,657 people living in the city (CSO, 2017). Cork currently has no light rail system in place and is only served by one heavy rail station, Cork Kent which is located on the North-East corner of the city with both commuter services to Mallow, Midleton and Cobh and InterCity services to Dublin and Tralee in operation. There are also a number of local bus services and there is also a bus station on the quays in the city centre. The introduction of the bike scheme in Cork is seen as a policy intervention to increase cycling in the city in order for Cork to have a broader range of transportation services, as both the Smart Travel (Department of Transport, 2009a) and National Cycling Policy Framework (Department of Transport, 2009b) have outlined the introduction of public bike systems to Irish urban areas as key future transport developments in Ireland. The two of these documents have set a target of 10% of nationwide trips to be by bicycle by 2020, and the public bike system in Cork (along with the other systems in Galway and Limerick) has been implemented to assist in reaching these bold targets.

As it has been two years since the scheme opened, there has been some research into different aspects of how the scheme has been performed. Caulfield et al (2017) examined the usage patterns in a smaller sized city (Cork), in order to see if smaller cities can derive benefits from public bike systems and provide a prominent role in these cities. The research discovered that in Cork habitual trip patterns are found among regular uses and average trip times are mostly short in duration. It was also established that longer trips were more likely to take place in better conditions along with other insights into the dynamics of a smaller than usual public bike system. This paper used data from the first calendar year (2015) of operation while the research conducted in this paper has continued building upon the knowledge of the Cork system with the supplementation of the 2016 data. The primary research objectives were then defined and are as follows:

- How much influence spatial, temporal and weather variables have on the variation in journey times of all the different route combinations;
- How the journey times of different trips compare with the journey times predicted by Google Maps; and
- How the number of trips in the scheme varies with time and in different weather conditions.

These research objectives were established to contribute to the existing literature of public bike systems and to assist policy makers and system operators in their planning processes of public bike system design and improvement, and could possibly argue the case for GPS tracking in the next generation of systems.

Other studies have shown that by using GPS tracking in a bike sharing scheme can result in more in-depth analysis and rich data on how these schemes are used (Wergin and Buehler, 2017; Broach et al, 2012; Hood et al, 2011). The cost involved in fitting bikes with GPS and the data protection issues concerned with this tracking may render wide scale tracking unfeasible. The approach presented in this paper tries to bridge the gap between rental times and travel times by comparing these using Google Maps travel times. The findings are a first step and further research in other cities would be interesting to compare against the findings from Cork.

Review of other studies

Public bike systems around the world have attracted plenty of attention in the literature over the last number of years, as they are a sustainable mode of public transportation which can easily be integrated in the existing transportation network in their respective cities. After deciding to conduct this research on the Cork public bike system, the first task was to do a comprehensive search for all the literature on public bike systems. The terms "public bikes" and "bike sharing" were searched and a total of 63 journal articles from recent years were downloaded and reviewed. These relevant journal articles were categorised into: Public Bike System Design; Public Bike System Maintenance; Public Bike Usage; User Surveys; and Impacts of Public Bike Systems. As the topic of this research falls under Public Bike Usage, the literature review in this article will primarily focus on this category. The existing articles that are grouped into the category of Public Bike Usage can also be subcategorised into User Demographics, Typology of Users, Usage and Stations and Travel Patterns.

User Demographics

The demographics of public bike system users is an interesting area of research, and has been investigated a few times in the recent years. Morabia and Costanza (2012) examines the question of public bike systems being an unegalitarian idea by considering the example of the London system where over 100,000 people subscribed to the scheme within the first seven months. The findings from examining the usage data suggested that public bikes is not socially homogenous as there were lower usage levels by people from socially deprived areas and women in general. Ogilive and Goodman (2012) also studied the user demographics of the London system and came to the same conclusions.

The demographics of users of Australian public bike systems (Melbourne and Brisbane) were analysed by Fishman et al (2015) from a user survey on these two schemes. As the Australian systems were found to have lower usage rates than the systems in Europe, North America and China, the authors created a logistic regression model in order to reveal the factors that influenced public bike membership. In terms of demographics that were revealed by the model, the two main membership predictors were those aged 18-34 who had bike stations within 250m of their workplace, and those who were on the higher end of the income scale which agrees with the two previous studies.

Analysis from the Chinese city of Suzhou also produces similar demographic results (Karki & Tao, 2016). Their focus of analysis was on the accessibility and convenience of that particular scheme to those who are more socially deprived, i.e. have a lower income and lower education. The results showed that there is a lower level of usage by the "socially deprived" and also by women. It was concluded that the results warrant making the system more convenient and accessible to the socially deprived. Overall this conclusion can be applied to most public bike systems around the world.

Typology of Users, Usage and Stations

In addition to demographics, users can also be classified into different types as people use public bikes for different purposes such as commuting, shopping, recreation etc. Vogel et al (2014) for example characterised user mobilities and used cluster analysis to produce a user typology by using data mining techniques. There were four main groups of users which specified in this study which were: "Users of heart"; Assiduous users; Multi-modal users; and Sporadic Users. These user types were classified after extensive analysis of travel patterns at an individual user level. Rather than actual user types, Bordagaray et al (2016) classified each trip into a usage type. These usage types were classified into: Round trips; Rental time reset; Bike substitution; Perfectly Symmetrical mobility trips; and Non-perfectly Symmetrical mobility trips.

Jiménez et al (2016) characterised bike stations by creating a new ratio called the Turnover Station ratio, which is based on the station capacity and how many times it is used each day. This new ratio in addition to the cumulative trips and number of available bikes was then used by the authors to classify the bike stations into three different types: Generative stations; Balanced stations; and Attractive Stations. In this literature review these three journal articles were the only three found that purely focussed on typology, meaning that the area of public bike user typology could be taken further in research in the near future.

Travel Patterns

Travel patterns in public bike systems tend to vary in both time and space and are usually related as spatio-temporal patterns. Corcoran et al (2014) analysed the impact weather and calendar events have on these spatio-temporal patterns in public bike usage by analysing data from the Brisbane CityCycle system. The weather variables that were analysed were temperature, rainfall and wind speed and the calendar events that were analysed were weekends, public holidays and school holidays, and these variables were used to create a model that can predict travel patterns within the system. Faghih-Imani & Eluru (2016) have also analysed the New York CitiBike system in order to examine the impacts of spatio-temporal interactions.

O'Brien et al (2014) used data mining techniques on a wide range of public bike systems around the world. They have used the data to create a large database of each of the systems that analyses the variation in usage and infers the likely demographics which can be used to classify each system based on the spatial variations. They propose a wide range of applications which can be used from their new data source that includes rebalancing approaches, identifying areas for new bike stations or cycle lanes and better targeting for new members. Demand estimation can also be determined from data mining which is demonstrated by (Médard de Chardon & Caruso, 2015). Oliveira et al (2016) in their methodology used a historical dataset of the New York CitiBike system to visually analyse the spatio-temporal travel patterns of the system. By mapping a pixel-orientated timeline the authors identified several patterns in the spatial and temporal domain, which could then also be applied to other systems in other cities. These visualisations can aid in analysing the travel patterns in large public bike systems, specifically for public bike system administrators, rebalancing researchers and big data analysts.

Methods

4.1 Data Collection by the National Transport Authority

The original datasets that are used in this research have been provided by the National Transport Authority (NTA) of Ireland. The datasets were provided in two separate files, one set for 2015, which contained approximately 290,000 trip records, and the other for 2016, which contained approximately 309,000 trips. Each of these datasets contained numerous variables for every single trip recorded within the Cork system in its first two full calendar years of operation such as the origin and destination stations, the time and date of when the bikes were taken out and returned, the distance between stations, whether the customer is an annual subscriber or using a temporary pass, a unique customer ID number and a bike ID number. The data provided for this research was anonymised so therefore no personal information on the name, age, gender, etc. of the users was provided in order to be in accordance with Irish data protection laws. The data also indicates if a refund was offered for the respective trip or if there was a fault such as the bike being returned immediately.

These two datasets were then merged together to create a master dataset which was then filtered prior to the creation of new variables and evaluation of the data. All temporary pass trips were filtered out as they were considered to be trips taken by visitors to the city and would not reflect the patterns of native users which is what has been examined in this research. Next the trips that were deemed to be faulty by the operator and trips which were refunded were also filtered out. This resulted in the removal of approximately 4,500 temporary pass trips and 35,000 faults or refunded trips.

4.2 Calculation of Spatial and Temporal Variables

A significant number of new spatial and temporal variables were calculated and attributed to each trip, and are listed below:

- **Trip Duration:** The time difference between when the bike was taken out from the origin station and when it is returned to the destination station.
- **Month:** The month of the year the trip took place.
- **Season:** Whether the trip took place in Winter (December - February), Spring (March - May), Summer (June - August) or Autumn (September - November).
- **Day:** The day of the week which the trip took place.
- **Weekend:** Whether the trip took place on either a weekday (Monday - Friday) or Weekend (Saturday - Sunday).
- **Hour:** The time the trip took place rounded to the closest hour.
- **Time of Day:** Whether the trip took place during the AM Peak (7am - 10am) or PM Peak (4pm - 7pm) travel period, Inter-peak (10am - 4pm) travel period or Off Peak (7pm - 7am) period.
- **One-Way Street Start/End:** Whether the street that the station is located on permits travel in one or two directions.
- **Interior/Exterior Start/End:** Stations deemed to be on the outer perimeter of the system were said to be exterior stations, while the remaining stations were said to be interior stations.
- **Cycle Friendly Start/End:** Whether the street that station is located on is deemed to be "Cycle Friendly" according to Google Maps.
- **Shops Start/End:** The number of shops within a 5 minute walk of the station.
- **Restaurants Start/End:** The number of restaurants within a 5 minute walk of the station.
- **Public Transport Links Above or Below Average Start/End:** Whether the station has above or below public transport links in terms of number of bus stops within a 5 minute walking distance
- **Total Origin-Destination Pair Trips:** The total number of trips that took place over the two-year period for that respective Origin-Destination pair.

- **Total User Trips:** The total number of trips that the user for that respective trip took over the two-year period.

Google Maps Journey Times

As the primary goal of this research is to examine the factors that influence journey time variation in the Cork system, the journey times obtained from the Google Maps journey planner perform an essential role in creating the dependent variable used for the MNL Regression modelling. As there are 31 bike stations located around Cork city this meant that there are a total of 961 different route combinations (in a 31 x 31 Origin-Destination matrix) that could be taken on a public bike at any given time. By entering the coordinates for the origin and destination stations for each of these 961 routes into the Google Maps journey planner (Google Developers, 2016), an expected journey time is returned for driving, public transport, cycling and walking. Every single cycling journey time was recorded for all of the routes which were then attributed as a new variable to all of the trips in the dataset, based on their respective origin and destination combination.

By then subtracting the expected Google Maps journey time from each actual observed trip duration, a new variable called Extra Travel Time was created which essentially was the difference between the two journey times, where a positive value indicated a trip that took longer than what Google Maps suggests and a negative value represented a trip that was quicker than the suggested Google Maps journey time. These values were segmented into quartiles based on how fast or slow these trips were compared to a predicted journey time. Table I demonstrates the four quartiles of journey time variation, which were used as the dependent variable for the principal MNL regression model.

Table I Journey Time Variation

Journey Time Variation				
		Frequency	Percent	Cumulative Percent
Valid	More than 46 seconds quicker	139667	25.0	25.0
	Between 46 seconds quicker and 35 seconds slower	139470	24.9	49.9
	Between 35 seconds and 2 minutes 26 seconds slower	140020	25.0	74.9
	More than 2 minutes 26 seconds slower	140296	25.1	100.0
	Total	559453	100.0	

Results and discussion

Station Busyness

The first set of results presented in this chapter are the descriptive statistics of the Cork system which provide an overview of operation in its first two full calendar years of operation. Fig. 1 and Fig. 2 illustrate how the bike stations in the system differ in terms of station busyness, i.e. which stations have the highest frequency of trips as an origin station or a destination station. The busiest stations that trips tend to start and end from are on the outskirts of the city, notably at Kent Station and the three stations that surround University College Cork, which are two major trip attractors in the city. The stations also near St. Patrick's Street and the quays were found to be busy due to the high density of restaurants and retail outlets in these areas. Some of the stations associated with less frequent usage are generally located in areas in the city that have a lower population density or in close proximity to each other, for example Clontarf Street and Lapp's Quay stations are only located 60m from each other.

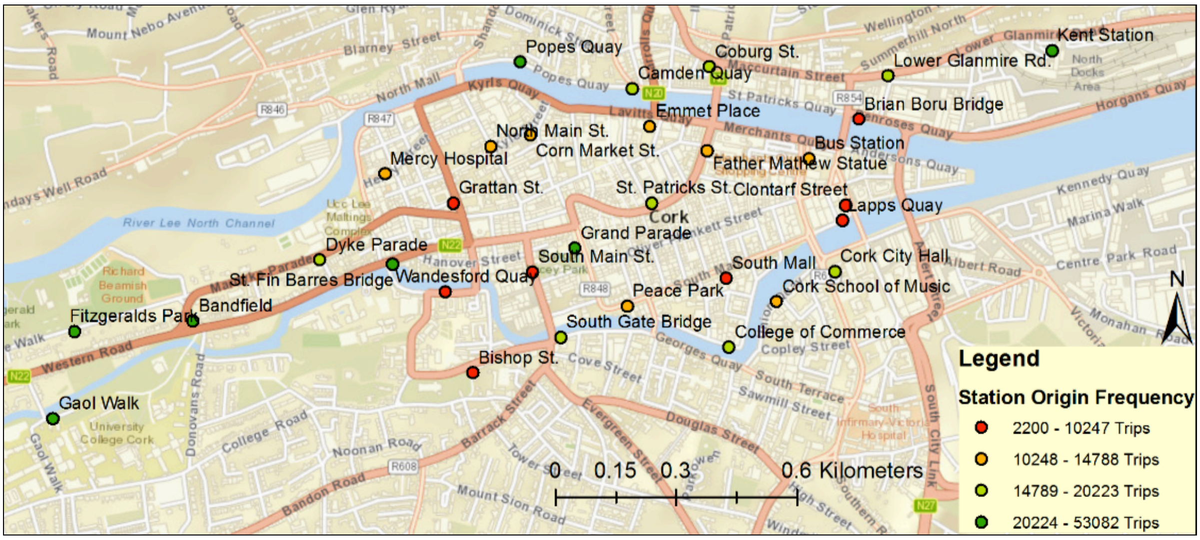

Fig. 1 - Total No. of Trips Originating from each Station (Total number of trips in 2015, 2016)

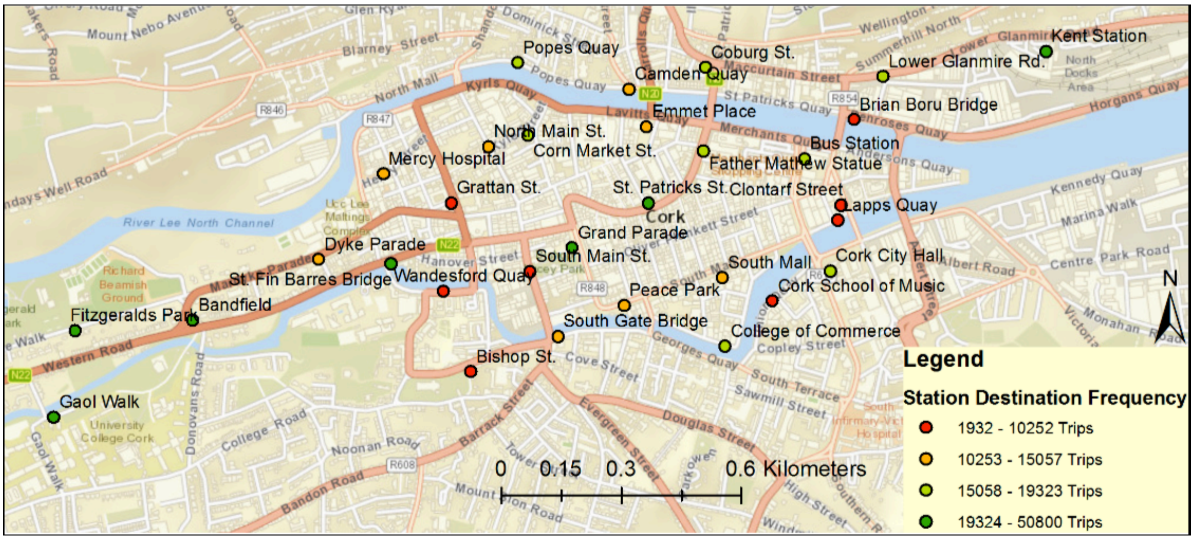

Fig. 2 - Total No. of Trip Destinations per Station (Total number of trips in 2015, 2016)

MNL Regression Models (Journey Time Variation as Dependent Variable)

The MNL Regression model summarised in Table VI has been developed using Journey Time Variation as the dependent variable. After consulting the result tables from the Stepwise Regression results, both temporal and weather variables were removed from the model as these variable types were either omitted during the Stepwise iteration process or were having less than a 0.001 difference in their contribution to the R^2 value of the model. Therefore the spatial variables such as One-Way Street, Interior/Exterior, Cycle Friendly, Shops, Restaurants and Public Transport Links along with Distance, Total Origin-Destination Trips and Total User Trips were the independent variables used for this MNL Regression model.

The model has a Nagelkerke R^2 value of 0.257, suggesting that these independent variables included in the model explains over one quarter of the variance of these different categories of journey time variation. Although this value may seem low, it is noteworthy that the purpose of each trip is not known nor the identity or demographics of each user, meaning that the spatial variables and other variable types included in the model do predict the variance in journey times to a reasonable extent as each of them are deemed to be significant in the full modelling results. In this model the reference category is trips that are more than 2 minutes 26 seconds slower than what Google Maps insinuates as the expected journey time for the quickest safe cycling route between two bike stations. The other three categories that are compared with this reference category are trips that are more than 47 seconds quicker than Google Maps (Category 1), trips that are between 46 seconds quicker and 35 seconds slower (Category 2) and trips that are between 34 seconds and 2 minutes 26 seconds slower (Category 3).

Table II MNL Regression Model

		N	%	Trips more than 47 seconds quicker than google maps	Trips 46 – 35 seconds slower than google maps	Trips 34 - 146 seconds s google maps
	Intercept			-1.057**	-0.439**	-0.071*
One-Way Start	No	173472	31.0	-1.430**	-0.288**	-0.048**
	Yes	385981	69.0	0^b	0^b	0^b
One-Way End	No	172282	30.8	-0.910**	-0.321*	-0.066**
	Yes	387171	69.2	0^b	0^b	0^b
Interior/ Exterior Start	Interior	211332	37.8	-0.332**	-0.352**	-0.163**
	Exterior	348121	62.2	0^b	0^b	0^b
Interior/ Exterior End	Interior	235594	42.1	0.248**	0.119**	-0.055**
	Exterior	323859	57.9	0^b	0^b	0^b
Cycle Friendly Start	No	383914	68.6	1.365**	0.238*	0.024*
	Yes	175539	31.4	0^b	0^b	0^b
Cycle Friendly End	No	395234	70.6	0.494**	0.222*	0.127**
	Yes	164219	29.4	0^b	0^b	0^b
Shops Start	6 Shops or less	276984	49.5	2.202**	0.691**	0.220
	7 - 12 Shops	72590	13.0	1.818*	0.851**	0.108
	13 - 31 Shops	99798	17.8	1.295**	0.504**	0.097
	32 Shops or more	110081	19.7	0^b	0^b	0^b
Shops End	6 Shops or less	260977	46.6	0.996**	0.698*	0.328
	7 - 12 Shops	67250	12.0	0.773**	0.443**	0.207
	13 - 31 Shops	98454	17.6	0.390*	0.170**	0.139
	32 Shops or more	132772	23.7	0^b	0^b	0^b
Restaurants Start	6 Restaurants or less	250900	44.8	-0.507**	-0.232**	-0.138
	7 - 16 Restaurants	102036	18.2	-0.813**	-0.482**	-0.284
	17 - 30 Restaurants	98389	17.6	-0.708**	-0.047**	0.043
	31 Restaurants or more	108128	19.3	0^b	0^b	0^b
Restaurants End	6 Restaurants or less	233070	41.7	-1.261**	-0.532**	-0.162
	7 - 16 Restaurants	99236	17.7	-0.670*	-0.464**	-0.287
	17 - 30 Restaurants	101302	18.1	-1.008**	-0.595*	-0.206
	31 Restaurants or more	125845	22.5	0^b	0^b	0^b

Public Transport Links Start	Below Average	333049	59.5	-0.747**	-0.225**	-0.043
	Above Average	226404	40.5	0[b]	0[b]	0[b]
Public Transport Links End	Below Average	328226	58.7	-0.373**	-0.100**	0.083
	Above Average	231227	41.3	0[b]	0[b]	0[b]
Distance	718 m or less	84730	15.1	1.013**	1.493*	0.709
	719 - 1087 m	118428	21.2	1.242*	1.302**	0.814
	1088 - 1596 m	149259	26.7	0.548**	0.872**	0.571
	1597 m or more	207036	37.0	0[b]	0[b]	0[b]
Total Origin-Destination Pair Trips	0 - 500	130698	23.4	-0.517**	-0.469**	-0.476
	501 - 1000	129145	23.1	-0.340**	-0.298**	-0.449
	1001 - 2000	159745	28.6	-0.042*	-0.178*	-0.293
	Over 2000	139865	25.0	0[b]	0[b]	0[b]
Total User Trips	7 Trips or less	7334	1.3	-2.374**	-1.710**	-1.059
	8 - 26 Trips	31721	5.7	-1.392**	-0.973**	-0.571
	27 - 81 Trips	104679	18.7	-0.799**	-0.548**	-0.289
	82 Trips or more	415719	74.3	0[b]	0[b]	0[b]
N						559,453
-2 log likelihood at convergence						213319.603
Log Likelihood ratio test						2.660
Nagelkerke R^2						0.257
Chi-squared statistic						154209.6
Degrees of Freedom						87

a. The reference category is: More than 2 minutes 26 seconds slower.
b. This parameter is set to zero because it is redundant.
*. This has a significance p-value > 0.01.
**. This has a significance p-value > 0.05.

From examining Table II in closer detail, each individual independent variable can be inspected and evaluated for their contribution in predicting the journey time variation among the trips. One-Way starts and ends to each trip were associated with the quicker trips indicating that cyclists may be more conservative with their cycling speeds when there is two-directional traffic. To a lesser extent trips that commenced at one of the exterior stations, and trips that ceased at interior stations were also linked with the more direct, quicker trips. Remarkably stations that were on roads considered to be "Cycle Friendly" by Google Maps were related to slower trips which may contradict the evidence from the existing literature. However the argument could be proposed that "Cycle Friendly" streets are travelled at slower speeds because there isn't the pressure on cyclists on public bikes to sustain cycling speeds similar to adjacent traffic. There may be other explanations for this result, but as the database does not have a track on the route on which the cyclist takes.

Conclusions

As seen from the MNL Regression modelling results, the variation in the journey times of different trips depends on a number of different spatial variables, such as whether the origin and destination bike stations are on one-way streets, are deemed to be "Cycle Friendly" or are an interior or exterior station in Cork city. The number of shops, restaurants and public transport links also influence journey time variation, and each of these variables can be examined for their strength as predictors in the results tables. Another interesting conclusion that can be made is that temporal and weather variables have little to no impact on the variation in journey times.

It would be interesting to see this journey planning tool being used in further research in order to assess its accuracy when measuring against actual trip journey times, but based on this research alone the journey times that were recorded from Google are precise in their predictions.

References

Bordagaray, M., Olio, L., Fonzone, A., Ibeas, A. Capturing the conditions that introduce systematic variation in bike-sharing travel behavior using data mining techniques. Transportation Research Part C: Emerging Technologies, 2016. 71: p. 231-248

Broach, J., J. Dill, and J. Gliebe, *Where do cyclists ride? A route choice model developed with revealed preference GPS data.* Transportation Research Part A: Policy and Practice, 2012. **46**(10): p. 1730-1740.

Caulfield, B., O'Mahony, M., Brazil, W., Weldon, P. , Examining usage patterns of a bike-sharing scheme in a medium sized city, *Transportation Research Part A: Policy and Practice*, 100, 2017, p152 - 161

Caulfield, B. Re-cycling a city - examining the growth of cycling in Dublin, *Transportation Research Part A: Policy and Practice*, 61, 2014, p216 – 226

Corcoran, J., Li, T., Rohde, D., Charles-Edwards, E., Mateo-Babiano, D. Spatio-temporal patterns of a Public Bicycle Sharing Program: the effect of weather and calendar events. Journal of Transport Geography, 2014. 41: p. 292-305

CSO. Census of Population 2016 - Profile 3 An Age Profile of Ireland, CSO, 2017

Faghih-Imani, A. and Eluru, N. Determining the role of bicycle sharing system infrastructure installation decision on usage: Case study of montreal BIXI system. Transportation Research Part A: Policy and Practice, 2016. 94: p. 685-698

Fishman, E., Washington, S., Haworth, N., Watson, A. Factors influencing bike share membership: An analysis of Melbourne and Brisbane. Transportation Research Part A: Policy and Practice, 2015. 71: p. 17-30.

Google Developers. Google Maps Distance Matrix API | Google Developers. 2016 [cited 2016 24th November 2016]; Available from: https://developers.google.com/maps/documentation/distance-matrix/.

Hood, J., E. Sall, and B. Charlton, *A GPS-based bicycle route choice model for San Francisco, California.* Transportation letters, 2011. **3**(1): p. 63-75.

IBM Knowledge Center. Pseudo R-Squared Measures. 2017 [cited 2017 6th April]; Available from: https://www.ibm.com/support/knowledgecenter/SSLVMB_24.0.0/spss/tutorials/plum_germcr_rsquare.html

Jiménez, P., Nogal, M., Caulfield, B., Pilla, F., Perceptually important points of mobility patterns to characterise bike sharing systems: The Dublin case. Journal of Transport Geography, 2016. 54: p. 228-239

Karki, T.K. and L. Tao, How accessible and convenient are the public bicycle sharing programs in China? Experiences from Suzhou city. Habitat International, 2016. 53: p. 188-194.

Morabia, A. and M.C. Costanza, Is bike-sharing unegalitarian? Preventive Medicine, 2012. 55(1): p. 1-2.

National Transport Authority. Cork - Coca-Cola Zero® Bikes. 2017 [cited 2017 5th April]; Available from: https://www.bikeshare.ie/cork.html

O'Brien, O., J. Cheshire, and M. Batty, Mining bicycle sharing data for generating insights into sustainable transport systems. Journal of Transport Geography, 2014. 34: p. 262-273

Ogilvie, F. and A. Goodman, Inequalities in usage of a public bicycle sharing scheme: Socio-demographic predictors of uptake and usage of the London (UK) cycle hire scheme. Preventive Medicine, 2012. 55(1): p. 40-45.

Oliveira, G. N., Sotomayor, J.L., Torchelsen, R.P., Silva, C.T., Comba, J.L.D. Visual analysis of bike-sharing systems. Computers & Graphics, 2016. 60: p. 119-129

Statistics Solutions. Conduct and Interpret a Multinomial Logistic Regression - Statistics Solutions. 2017 [cited 2017 6th April]; Available from: http://www.statisticssolutions.com/mlr/

Vogel, M., Hamon, R., Lozenguez, G., Merchez, L., Abry, P., Barnier, J., Borgnat, J.P., Flandrin, P., Mallon, I., Robardet, C. From bicycle sharing system movements to users: a typology of Vélo'v cyclists in Lyon based on large-scale behavioural dataset. Journal of Transport Geography, 2014. 41: p. 280-291.

Wergin, J. and R. Buehler. *Where do Bikeshare Bikes Actually Go? An Analysis of Capital Bikeshare Trips Using GPS Data.* Transportation Research Record (forthcoming)