DATA SCIENCE FOR DESIGN

CYCLING EDINBURGH

Project Report

by Hafsa Olcay & Sofia Thuru University of Edinburgh, 2017

ABOUT OUR DATA

Main dataset: The data was gathered from 33 counters across the city. The datasets we used are in xsl format and are arranged into a seven days timeframe ranging from 29/01/2017 to 04/02/2017. Each day includes hourly distributions of bike count, bike speed and pedestrian count. However, there are some inconsistencies from dataset to dataset. E.g separate tables for each day vs. merged tables, pedestrian counts vs. no pedestrian counts and different labels for variables.

SQL Bike Counter Database: We have used SQL database for bike counters in Edinburgh to retrieve data from past years. This database also included data about the exact location of counters with latitudinal and longitudinal coordination which we traced on Google Maps (Figure 1-3).

Additional Resources: We used several facilities of Google Maps to contextualise our data analysis. These include terrain view (Figure 1) which demonstrated the levels of height for a better understanding of the physical qualities of site, bicycling view (Figure 2) (We didn’t really do this/ it is not relevant as we did not analyse bike speed?) to trace cycling routes, Lothian bus routes (Figure 3) to see how transportation routes overlap, and street views for additional information on the properties of path when applicable (Figure 4-5).

We also did qualitative research using published material from SPOKES, the city councils (<http://www.edinburgh.gov.uk/info/20135/cycling_projects>) and looked at local cycling blogs for further information on how people cycle/ their opinion about certain routes (<http://citycyclingedinburgh.info/bbpress/>). I think it’s important to highlight qualitative research as it was in this way that we were able to contextualise the quantitative information we gathered.

OUR OBJECTIVE

The objective of our analyses is to provide insights on the ways in which cycling infrastructures encourage and/or discourage cycling habits. We did this by looking closely at our data and finding odd figures. Once we established some patterns within our datasets we begun to contextualise the data through qualitative research. The starting point of our enquiry was questioning why don’t people cycle more in certain areas which lead us to hypothesise that this could be due to safety concerns related to the site properties and urban infrastructure.

1. (I reworded the whole paragraph because some of it did not make logical sense. For instance, “questioning the reason why people don’t cycle more” doesn’t logically lead to improved cycling infrastructures as the sentence seemed to suggest. It was merely the starting point which we built on)
2. (also I highlighted “in certain areas” because I think it’s important to differentiate between “why people don’t cycle more” - general statement/ has a lot of inbuilt bias and “why don’t people cycle more in certain areas” is more closely related to findings in our data.

Our data analysis and visualisation target the experts in Urban Big Data Centre who work on the improvement of urban infrastructure for cycling.

THE PROCESS & ANALYSES

Initially, we analysed the cycling patterns around 33 counters1 by tracing the daily distribution of bike numbers counted for each day. On average, we have found that people cycle more on weekdays rather than on weekends, with the highest rates located on Wednesday and the lowest on Saturday. People cycle the most during peak hours and it is not as common to cycle as it is to walk in the afternoon. There were cases in which a location was popular in one direction but not in the other. The reason for this remains obscure (i.e. Site 40: Mayfield Road S/Bound had 0 counts for each day for one direction and 452 as the average of the week on the other). As it was not clear if this was due to lack of bike use, the counters’ failure or a mistake in the collection of data we decided to exclude it altogether. However, we detected suspicious differences in the two directions of one of the sites, London Road, which had low rates of bike count on one direction. Moreover, it also had a suspiciously high bike count on the 4th of July. Upon additional research, we found that this was due to a special occasion that took place at the time.



Upon gaining initial insights on the cycling habits per location we then compared the popularity of bike use in each site based on their weekly averages. In order to contextualise our data and further explore the correlation between sites and bike use we traced the roads and counters on a map of Edinburgh. The most popular locations were Melville Drive and Wester Coates. We also detected that Morrison Street, despite being situated in between popular locations, had one of the lowest cycling rates.



We made further analysis for Wester Coates and Morrison Street to see if popularity for these two sites is relevant to cycling habits.



(I don’t like the word “habit” - like cycling habit - because it has a connotation of individual behaviour and has a psychological feel to it. I think sticking to bike counts is better? Just a stylistic note)

SITE 7: WESTER COATES

These figures show that people cycle the most during the morning rush hour on Wester Coates which is much higher than the evening rush hour **that has the second highest rates. (this bit doesn’t make sense as we firstly need to clarify the timeframe of the morning rush/ evening rush. E.g. 8-9am?)** Despite the hourly and daily changes in bike counts, the range of speed is between 10 – 20 KpH and does not show significant differences during the week. Therefore, it can be argued that there is not much obstruction interrupting the flow of cycling which might be a factor that makes the location popular.

(what do you mean by obstruction? Car traffic? Actually, I read there can be a lot of car traffic but why would car traffic impact bikers if they use the bus lane?)

SITE 21: MORRISON STREET

The figures for Morrison Street also show the highest rates of cycling to be during the morning rush hour although they are significantly lower compared to Wester Coates. The hourly distribution of bike speed shows a wider range and a more scattered distribution (of what?) on both hourly and daily basis. This is most evident during Mondays and Thursday, when people cycled the most. The distribution of speed might suggest that Morrison Street does not provide smooth cycling conditions for the cyclists, and therefore, is not very favourable. I like this!!!!!! :D

The figures on the left demonstrate a clearer distribution of speed on the least and most popular days in both sites. The highest bike speed was recorded on Morrison Street which is an unpopular street for cyclists in general. This indicates that the physical properties of the site allow for speeding up which should make it favourable for cyclists. (maybe it’s less busy which is why it’s more popular? Also you previously said that speed was scattered thus suggesting bad infrastructure. You contradict yourself?) To see if the street is unpopular in general we compared the pedestrian and bike counts in Wester Coates and Morrison Street. I also like this!

CONCLUSION

DATA VISUALIZATION AND INTERACTIVE PDF

For our interactive data visualization, we have mapped the counters and showed their popularity distribution on the chosen three days based on bike counts. The days are chosen according to the most and least popular days (We???) and Monday is added as it is the turn of the week.

& CRITICAL REFLECTION