

New Bike-Share Route Prediction

How well do machine learning approaches predict new expansion routes for the Chicago bike-share system?

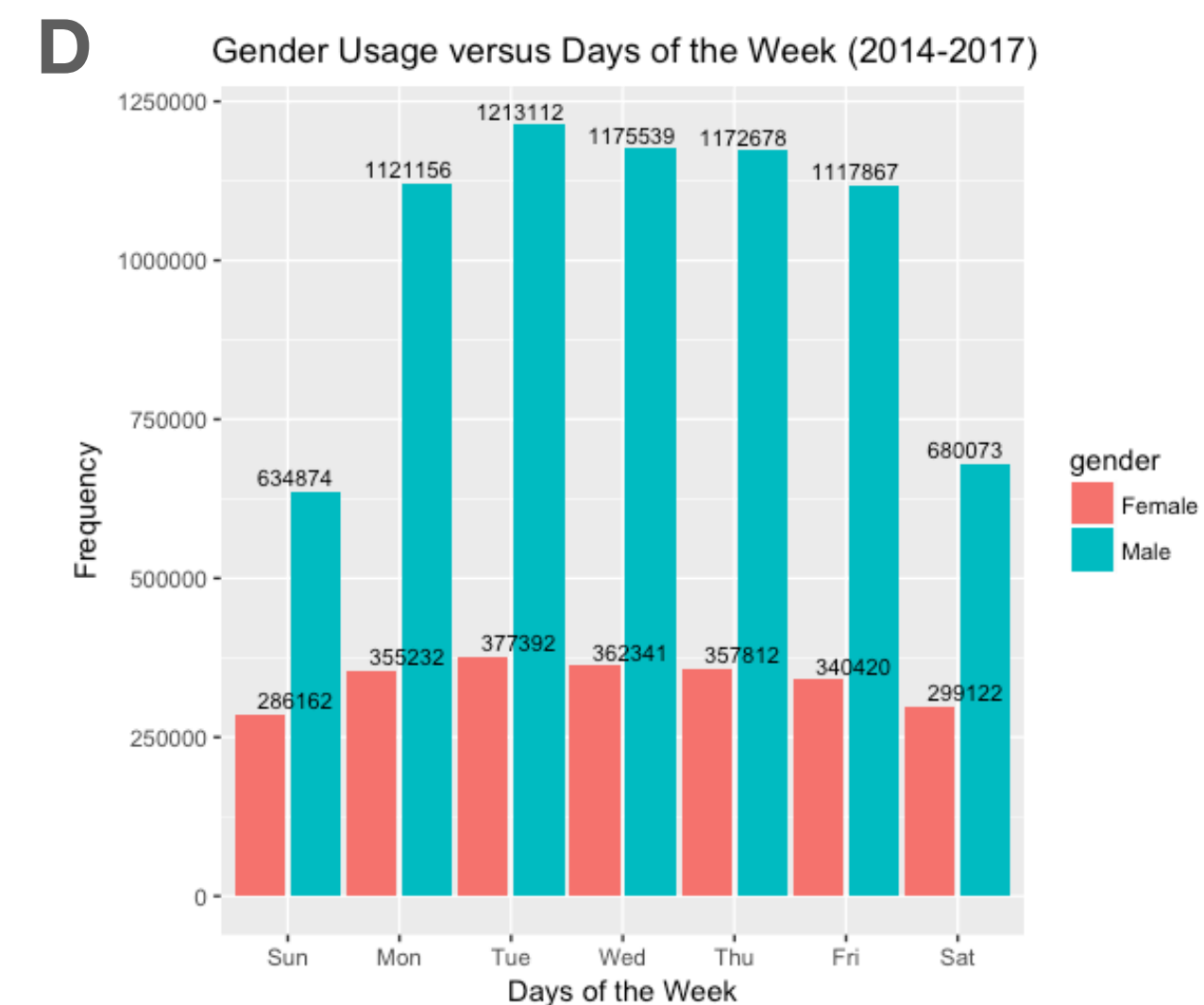
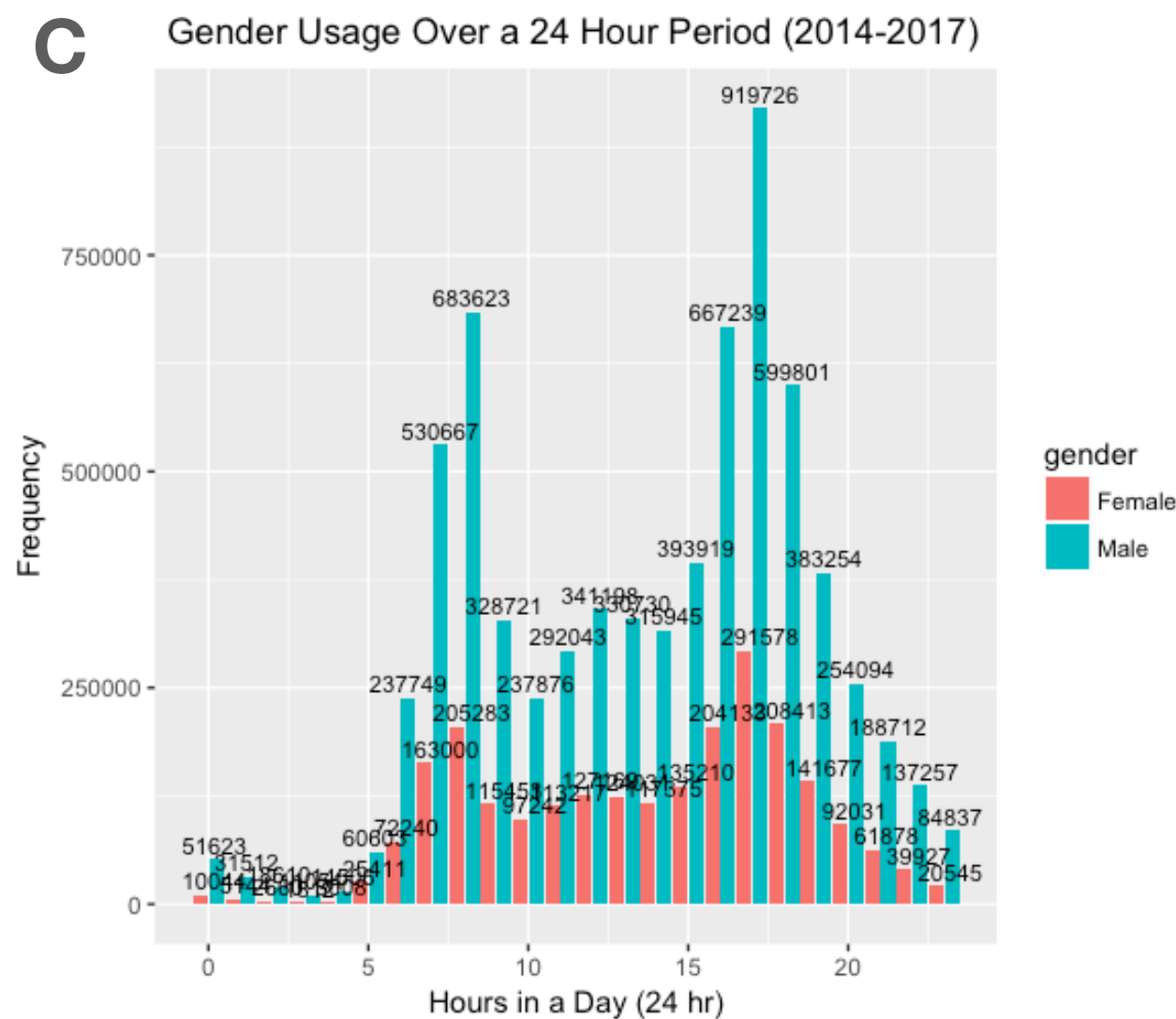
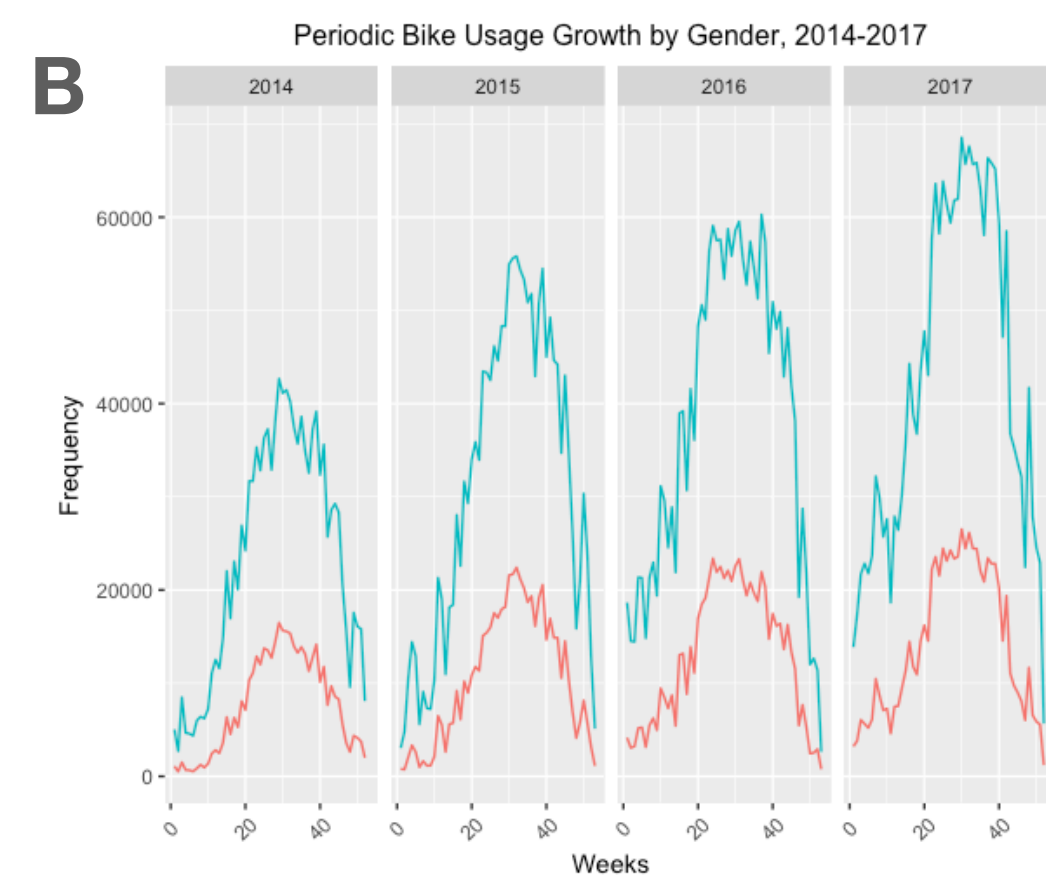
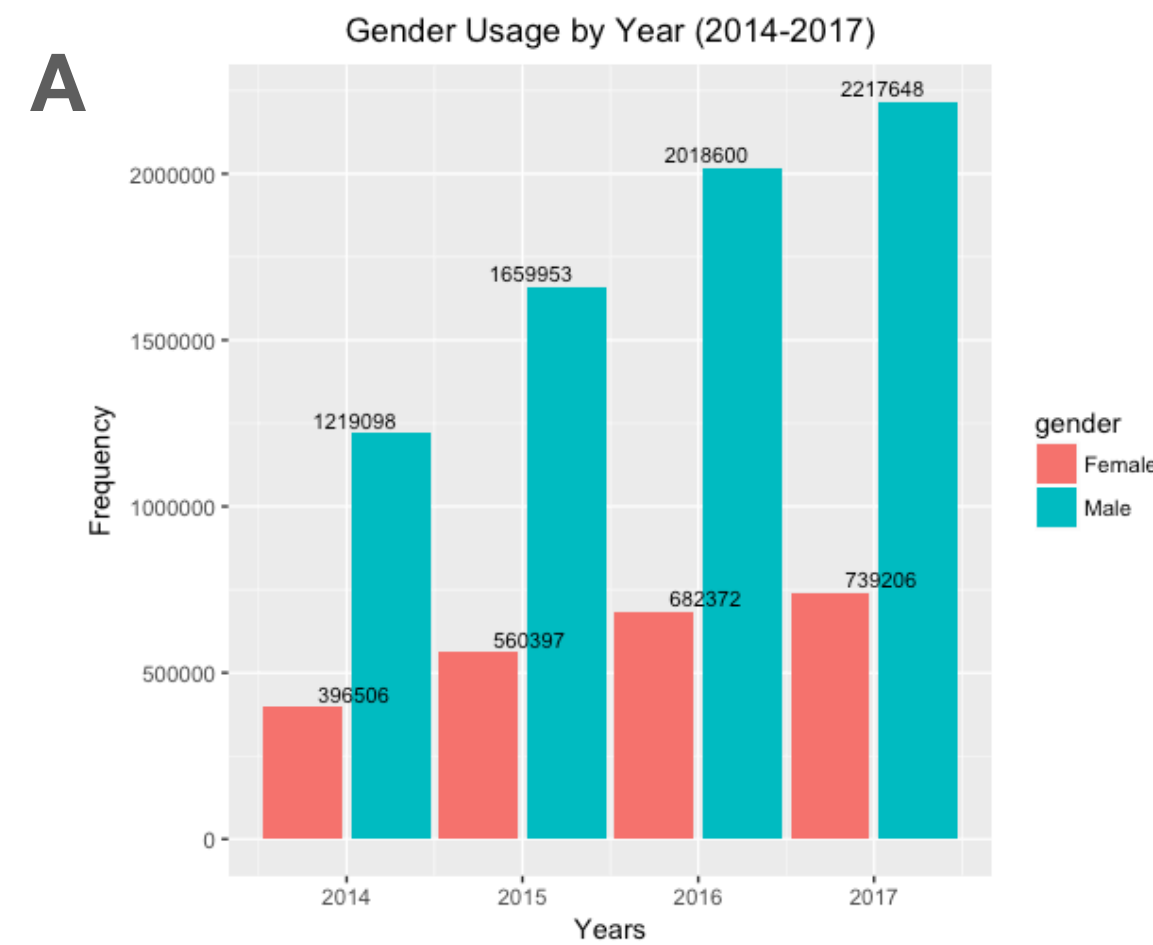
Introduction

Why is this interesting to me and others?

- Population increase in cities globally, micro-transportation methods arising to move people around more cheaply and efficiently.
- Although bike-sharing concepts have been around in US since 1960s only recently has there been an explosion in bike usage
 - ‘Green’ (less emissions), non-motorized alternative to short city trips
 - Chicago: 8 years ago no bike-sharing, grown to 656 bike stations in network
 - Chicago Divvy (Lyft) bike-share datasets freely available:
 - <https://www.divvybikes.com/system-data>

Where is the Data?

Example EDA: Subscriber Growth and Usage (by gender)



- Approx. 9.5 million entries in 2014-2017 ride-share dataset (clean; 1.2 Gb)
- Additional data available from 2018-2021
- Subscriber/member users have anonymized demographic data (e.g. gender, age)
- Subscriber/member have time-stamped trips logged
- Trip details comprise station location (656 stations by 2017) and coordinates (lat/lon)
- Most use on weekdays (Monday-Friday) during pre- and post work (peak at 8 am & 5 pm)
- Plots generated using R and ggplot2

Can we predict additional expansion?

Where will new stations go?

- 2014 started with 75 stations and 750 bikes (2015 expansion); by end of 2017 there were 656 stations.
- Utilize ML approaches to predict rules of new stations (Python/Keras/Tensorflow?):
 - Use station data from 2014-2016 as a training set, and potentially 2017 as validation set (already available)
 - Acquire additional 2018 or 2019 trip and station data as test set
 - Additional dependencies: (a) socio-economic factors (e.g. zip-code) (b) bike station proximity to existing subway stations
- Who will benefit specifically?
 - Bike-sharing organizations: rationalize and prioritize expansion | optimize logistics of bike inventory/location
 - City planners in municipalities where potential new station may appear (may need new infrastructure e.g. street lighting, retail)
 - End-users benefit: Commuters who enter the city from suburban nodes and would benefit from suburban bike-share stations.