



DSI4 Capstone Project Proposal

Blae Quayle

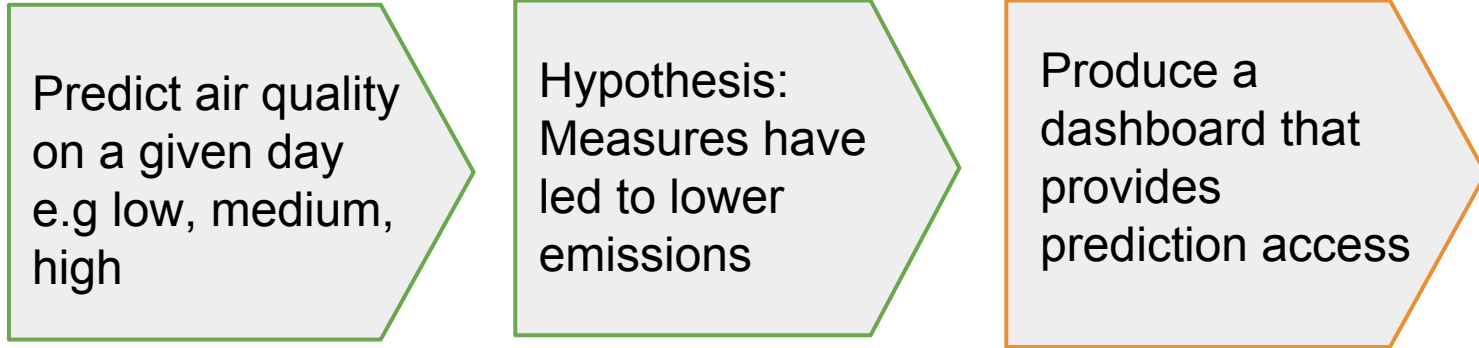
Idea 1 - Putney's Pollution Problem

- NO₂ produced by diesel vehicles has been linked to respiratory and heart problems
- Sites can breach hourly limits of 200 mg of NO₂ per cubic metre of air 18 times per year



- When do you think Putney High Street breached annual limits in 2016?
- Putney High Street London's first Low Emission Bus Zone - March 2017

Goals



Potential audience: Putney residents
Wandsworth Council

Dataset

- Acquire using API from King's College
- Met Office Historical Data



Metrics

- Determine accuracy of prediction, confusion matrix - reduce false negatives
- Allow council to evaluate success of low emission zone
- Potentially identify other areas it could be rolled out

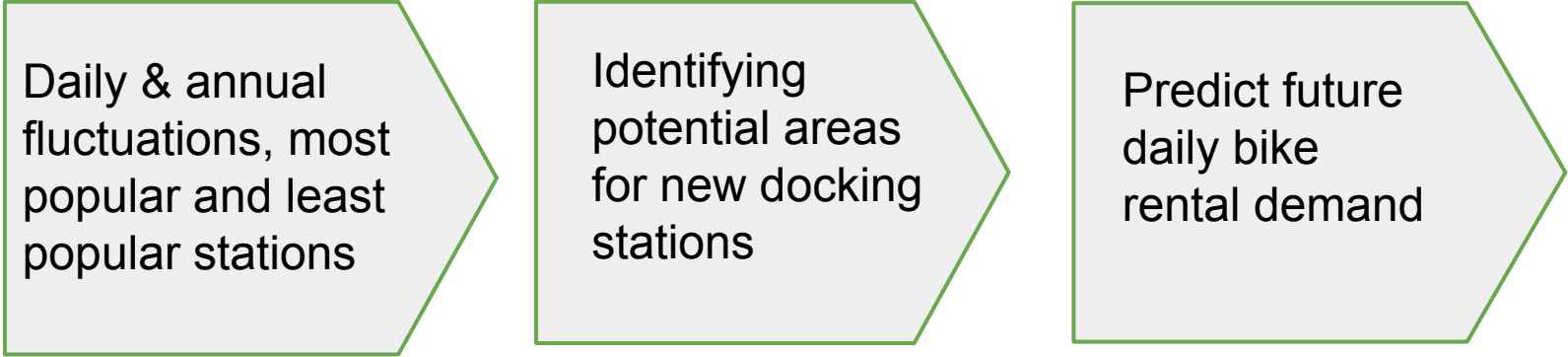
Idea 2 - London Bike Sharing

- Santander cycles started operating in London in July 2010
- Grown to 11,500 bikes and over 750 docking stations
- 8-10 million hires per year



- Lack of bikes and/or docking stations near major railway stations
- Poor coverage in relatively central areas (e.g. SE London) and the suburbs

Goals



Daily & annual fluctuations, most popular and least popular stations

Identifying potential areas for new docking stations

Predict future daily bike rental demand

Potential audience: TFL
Commuters

Dataset



- Monthly and aggregated annual hire data from TFL in .csv format
- Web scraping may to obtain the dock locations
- Live feed of bike availability including the number of docking points

Metrics

- Evaluate residuals on number of hires per day
- RMSE, Mean (or median) absolute error

Questions?